

A Guide for New Reactome Curators

Overview of Curation

This set of slides (Appendix C1) provides a high-level overview of the curation process described below.

Quick Guide

This is a list of the most commonly used steps in the annotation process with links to the relevant guides, and is provided as PDF in Appendix C2.

Introduction

The goal of Reactome is to describe the known biochemical details of human biological pathways. A Reactome curator's job is to work with experts in different fields of biology to identify and curate suitable human pathways (see below) by breaking them down into subpathways and reactions and describing them in a format that is compatible with the Reactome data model. The purpose of this guide is to describe each step of that curation process to help the curator fully understand the steps involved. The guide is a useful reference for the experienced curator as it is nearly impossible to remember all of the steps and details of the curation process. Curators are actively encouraged to add to sections of the guide providing useful hints that may be missing for current and future curators. Regular use and addition of material to the guide will facilitate the process of curation and increase annotation consistency. If in doubt about any of these steps...do not hesitate to e-mail the internal list for clarification (and then contribute any helpful info that you get to save someone else the trouble of asking later!)

Here we provide basic guidelines as to how data for relevant pathways should be collected, organized and entered into Reactome via the Curator Tool. The process of converting a biological topic should maintain the referential integrity of the reactions, both newly added as well as those already present in Reactome. It is assumed that you (the reader of the guide) have some familiarity with Reactome, and have read the Reactome papers, including the very important paper describing naming conventions for both events and entities (Jupe et al, 2014).

Our publications are listed on the Publications page of the Reactome website (https://reactome.org/community/publications).

The entire dataset, website, curator, and author tools can be downloaded from the Reactome download page (https://reactome.org/download-data). Installation and configuration instructions are also provided here (https://reactome.org/download-data/reactome-curator-tool).

If you have any problems with the install or tools please contact us at help@reactome.org.

This rest of this guide is divided into the following sections:

1. Important general guidelines for new curators
2. Choosing a topic/pathway for curation

3. Understanding and using the data model
4. Guideline for Naming Entities and Events
5. Using Data Inferred from Other Species
6. Providing evidence for assertions
7. Using the curator tool
8. The annotation process for traditional Reactome pathways
9. Cell lineage path curation guide
10. Disease pathway curation guide
11. Infectious disease and Innate Immune System pathway curation guide
12. Drug curation guide
13. Complexities of disease and drug curation: special cases
14. Modifying deleting and merging instances
15. QA
16. Curation Tools/Automation
17. Preparing a pathway report for internal and external review
18. Reporting bugs
19. Appendices

Important general guidelines for new curators

- Evidence and references:
  - When considering which type of evidence to use when creating events in Reactome, note that you should **NOT** include events based solely on microarray, proteomics or similar high throughput data, unless either 1) an expert author has indicated that they believe the results are correct and represent the current consensus, or 2) The results are a meta-data analysis that combines several independent studies to show the overlap between them. See guidelines describing External confidence criteria here (Appendix C3).
  - In general any time you make a statement in your summation you should include a reference. This is to point the user to the paper you found that proves the statement is true, otherwise the reviewer can't go and check. It also shows that it is an established fact, not your own finding or opinion.
  - Making inferences: If reactions (RLEs) have not been demonstrated experimentally, it is OK to include these as long as somewhere in subsequent events there is some kind of experimental data that suggests your inferred events must logically occur to get you to that point. It's essential that you explain the inference you have made (or others have suggested). You must include citations for the characterized mechanism that allows you to make the inference, i.e. why you believe the events you are describing are similar, despite lack of direct experimental proof in humans. It is best to have an expert's approval of this type of event before

including it. Note that you should NOT create a pathway in non-human species and use that to infer a human pathway. Instead, create individual non-human reactions and use those to infer the human RLEs where evidence for the RLE is missing. See the section below on Creating an inferred event.

- Updating existing content
  - Be extremely careful changing ANY released instance. This should almost never be done unless new information has become available. Be absolutely sure to check any/all referrers of the instance being changed (entity/event) and make sure that those changes are relevant in ALL referrers. Never change the compartment of an existing entity as that will change the compartment everywhere the entity is used.
  - Changing the species of an event or entity, once it has been committed to gk_central, will change its stableIdentifier. Please contact the managing editor if this has happened. Changing a species once an entity/event has been released will require the deletion of the instance and creation of a new one. The old deleted instance must point to the new one via the deletedInstance that is created. See section on deleting instances in this guide.
  - Adding an entity with a new species to an existing set (say a new pathogen entity is added to a set of entities from other pathogens to be used in an innate immune response event) can be a bit complicated. You have to then make sure that the species setting for the set includes the new pathogen species and that all referrers of that set are adjusted accordingly with the appropriate relatedSpecies). You can do this by looking for each referrer of the changed entity (all entities and ReactionLikeEvents) and then use the curatorTool species QA check to make sure there are no conflicts.
  - If you change an existing event, especially if it has been released, you need to change the summation to reflect this.
  - Substantive changes to released material (physical entities and events) are automatically tracked each release by the UpdateTracker script (see Appendix C4 for a description of changes that are caught by UpdateTracker)

- Naming events and entities:
  - Use the published guidelines (Jupe et al, 2014) on naming events and entities in Reactome. Note that, despite the statement in this paper, current practice supports using coordinates for *all post-translational modifications, if known.

- Naming of growing complexes: Generally, for serial binding events the 'new' input is named after the existing complex, this gives a better sense that a complex is growing or changing.
- Generally, descriptive words like ‘complex’ or ‘receptor’ are not included in physical entity names except in cases where the common name uses those terms, ie RNA pol II complex, Mediator complex etc
- Use of the word 'activated': Activated has a special meaning in Reactome and should only be used to indicate a change of conformational state from inactive (this word is not used) to active or activated.
- Refer to the more detailed “Guidelines for naming Entities and Events” section in this guide.

References

1. Jupe S, Jassal B, Williams M, Wu G. A controlled vocabulary for pathway entities and events. Database (Oxford). 2014 Jun 20;2014:bau060. doi: 10.1093/database/bau060. PMID: 24951798; PMCID: PMC4064568.

- Naming people and consortia
  - *initial* is a mandatory attribute on a **Person** record, and refers to the first name initial. If no first name is available, enter a dash “-” as the *initial* AND *firstName* attribute
  - if a consortium is cited as an author in a **LiteratureReference**, the consortium name should be listed as the *surname* attribute, and a dash “-” should be listed as the *initial* attribute
- Use of compartments
  - Be careful to use the correct compartment. For example: Nucleus vs. nucleolus. Also, be as specific as possible. If you are talking about an event occurring inside the nucleus, use nucleoplasm not nucleus (Also use cytosol not cytoplasm).

Choosing a topic/pathway for curation

Focus first on:

1. Areas of biology that you know well (or one in which you have good contacts).
2. Pathways that are well characterized at the molecular level AND for which there is considerable HUMAN biochemical data and proteins that are new to Reactome.

Points to consider:

1. Give lower priority to topics that will require a large amount of inferences from other species.
2. Don't try to tackle huge pathways all at once.
3. Break large topics into *manageable* pieces. Experts are generally unwilling to commit to a project bigger than a dozen rxns and it is easy to get lost in the details and lose focus which is costly in terms of time.
4. Create a logical hierarchy for the events in the pathway that you are curating.
5. Work on several (preferably related) small projects simultaneously. Experts often fail to come through at critical times so it is good to have a few projects at different stages to keep work flowing through the pipeline.
6. Check the Editorial Calendar to ensure that your plans don't overlap with another curator. If in doubt, contact the curator. Discuss plans with Editor in Chief and managing editor.

Understanding and using the data model

Data Model Glossary

A detailed understanding of the Reactome data model is important for accurate and consistent curation. The best way to start learning about the different data classes and how to use them is to browse the Data Model Glossary (Appendix C5). Some classes are used infrequently but you still need to be aware of them. Some of the more important data classes are described in the annotation examples below.

More complex examples are cited in the advance curation section as well as the disease annotation guide.

Key Data Classes

PhysicalEntity

PhysicalEntities can be single entities, such as proteins, small molecules, RNA, DNA, carbohydrates, lipids, or sub-atomic particles. They can also be complexes consisting of a combination of any of the single entities, or polymers synthesized from the single entities. Related entities can be grouped into a set. The use of sets is described below. PhysicalEntities can be inputs, outputs, catalysts, regulators, or requirements in ReactionlikeEvents.

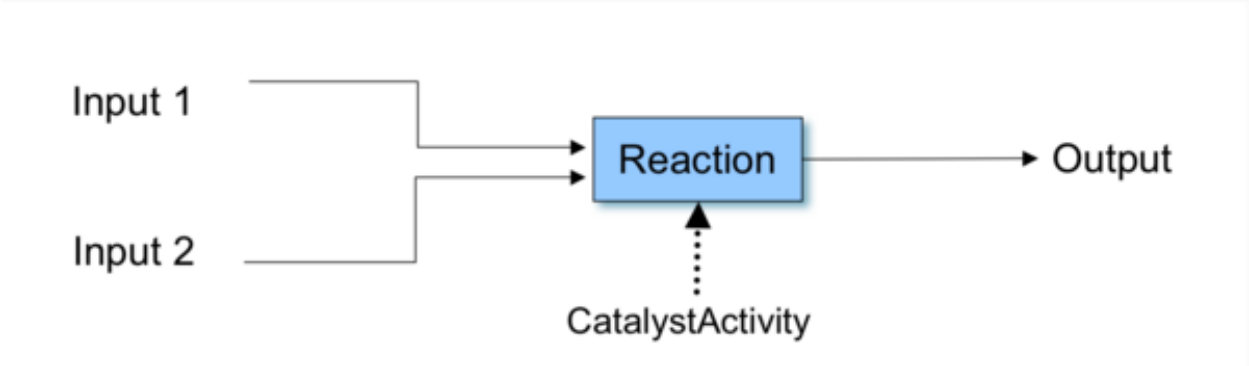
Events

The main concepts in the Reactome data model are Event and PhysicalEntity. At present there are two subclasses of Event: ReactionlikeEvent and Pathway.

ReactionLikeEvent

ReactionLikeEvent (RLE) in Reactome is a single step event in which input entities are converted to output entities by means of structural, chemical or compartmental

changes.



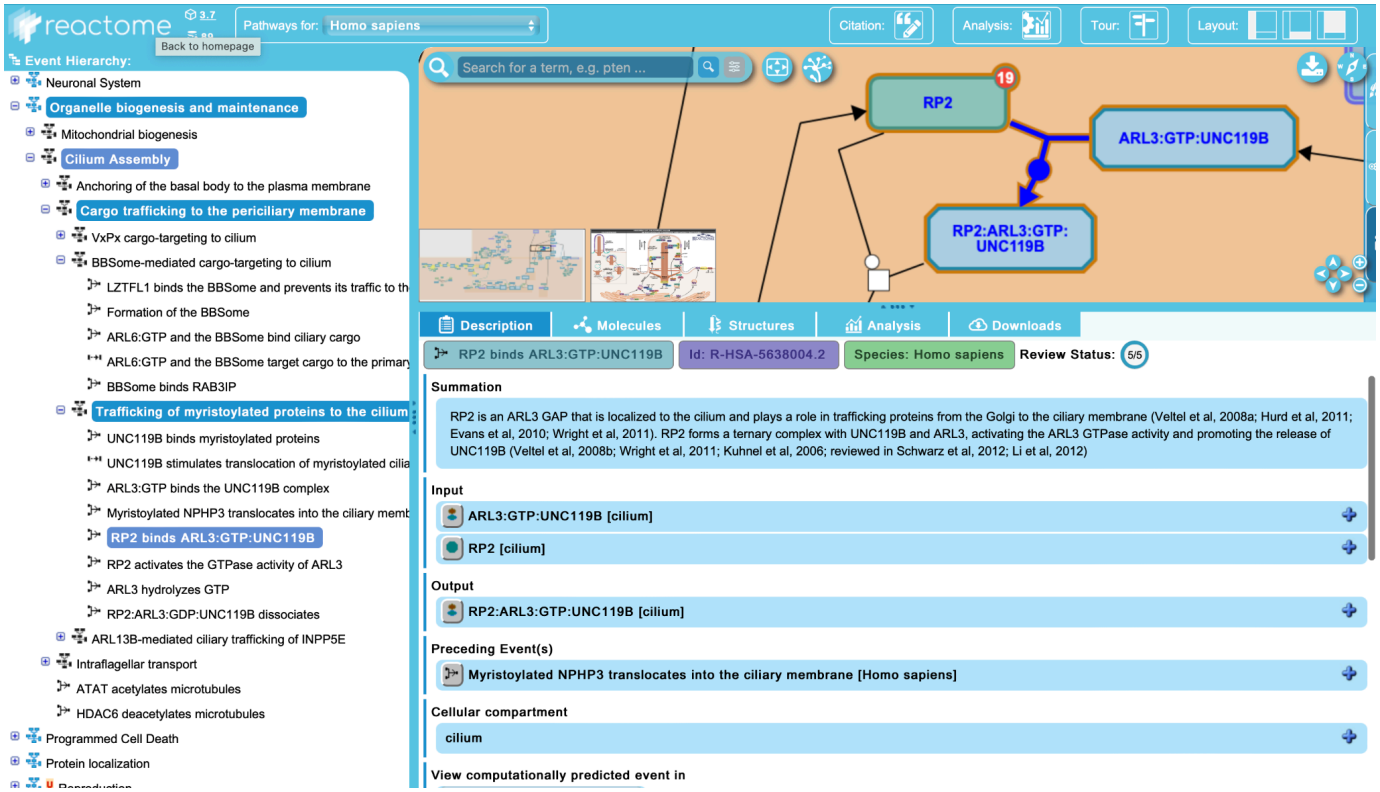
The Reactome data model robustly represents biology. The central concept in Reactome is the reaction, which is used together with pathways, macromolecules, small molecules, complexes, and catalyst activities to represent biological processes. The reaction itself is a single step biological event in which input entities are converted to output entities.
The ReactionLikeEvent subclass has 6 subclasses: Reaction, BlackBoxEvent, Depolymerisation, and Polymerisation, FailedReaction, and CellDevelopmentStep. ReactionLikeEvents may be linked to other events that precede them, regulate or are regulated by them.
Participants of ReactionLikeEvents – inputs, outputs, catalysts, regulators and requirements – are PhysicalEntities.

Here are some common examples of different types of ReactionsLikeEvents. The compartment of an RLE that involves participants in more than one compartment (or movement of participants across multiple compartments), should reflect the compartments of all of the participants (inputs, outputs, catalyst).

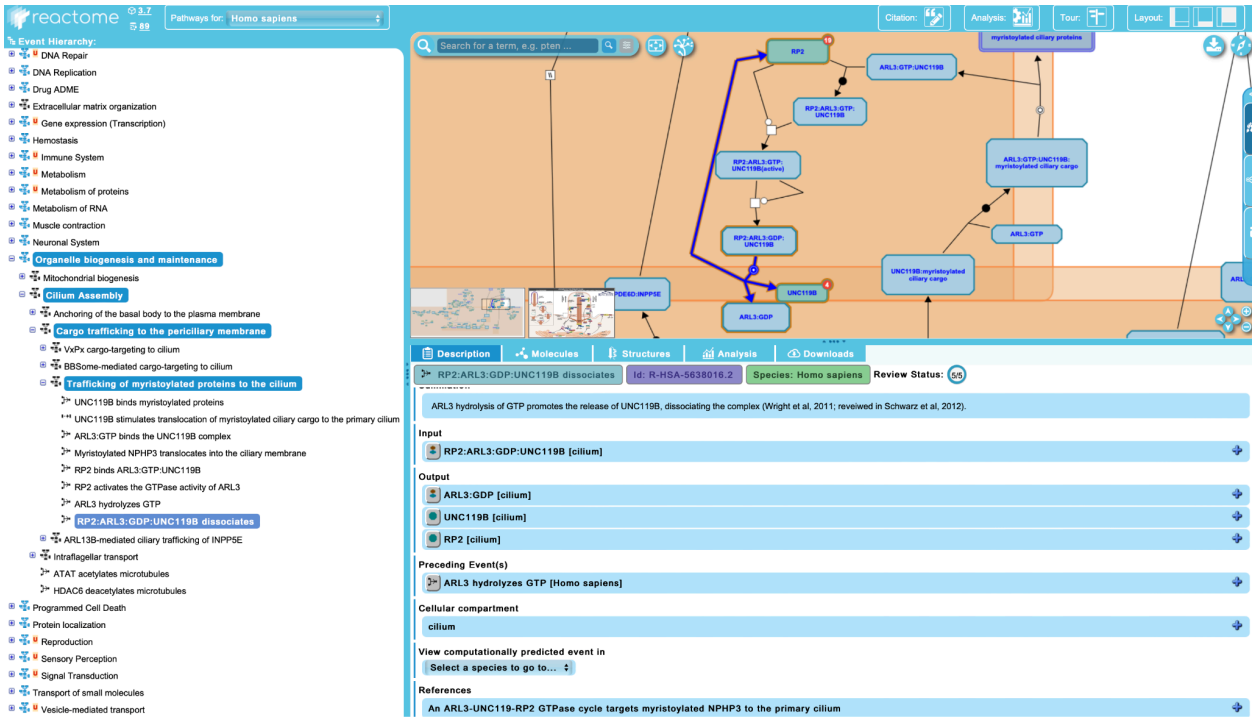
Reaction

A Reaction is an event that converts inputs to outputs in a single step. Any enzymatically-catalyzed event, or transport of physical entity to adjacent cell compartments, or protein-DNA binding or protein-protein complex formation or disassociation, for example, can be annotated as a reaction:

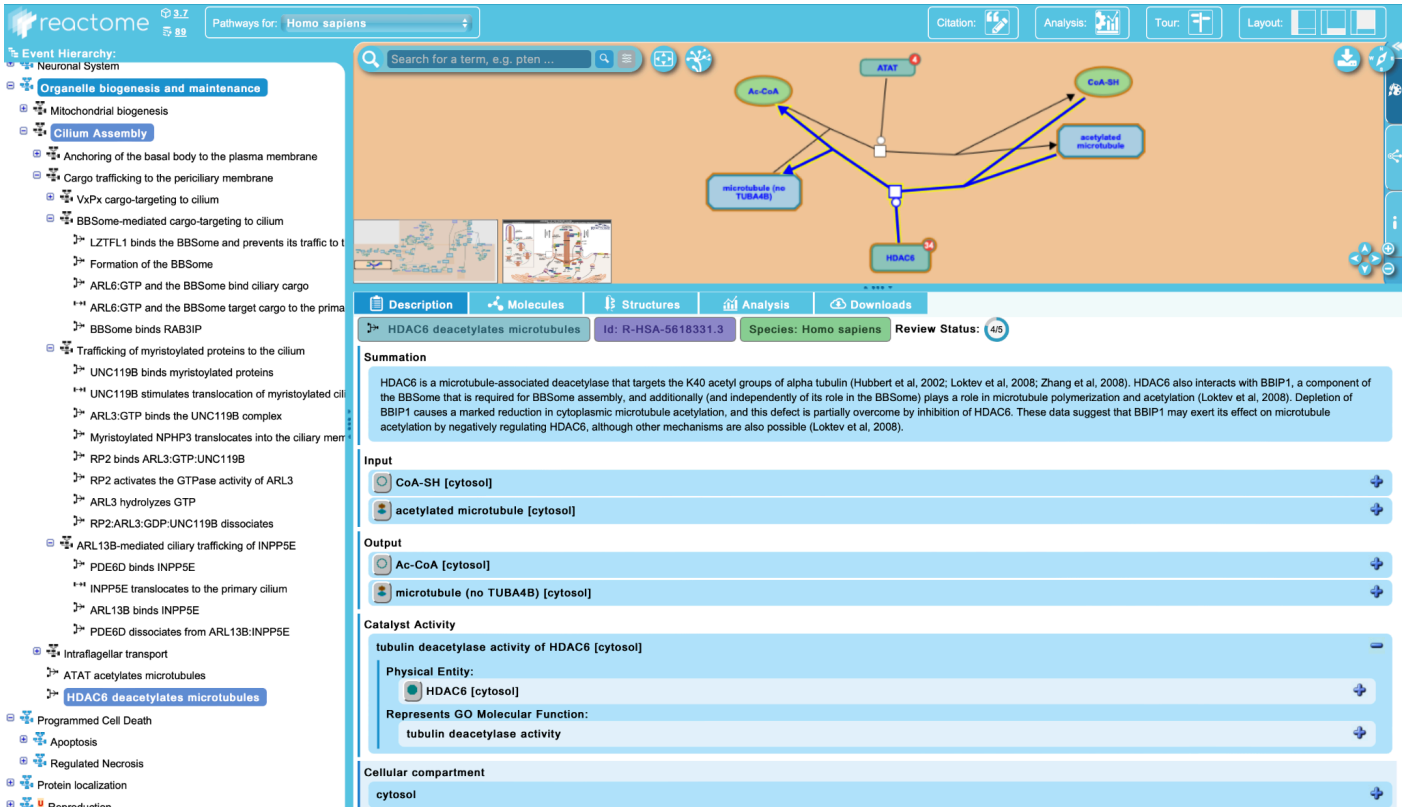
Complex Association:



Complex Dissociation:



Enzyme-catalyzed biochemical reactions:



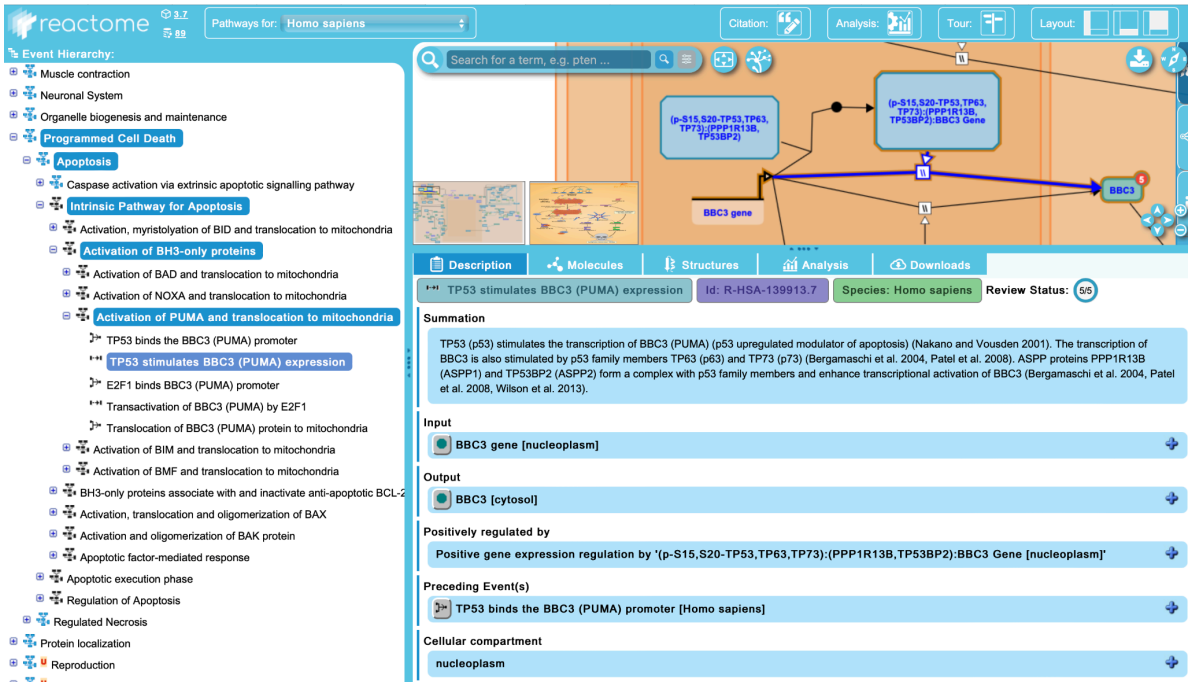
The compartment of a Reaction that involves participants in more than one compartment (or movement of participants across multiple compartments), should reflect the compartments of all of the participants (inputs, outputs, catalyst, regulator).

BlackBoxEvent

A BlackBox event converts inputs to outputs of a multiple step biological process in which intervening steps are not annotated because:

1. We don't know all of the intervening steps.
2. The intervening steps are known but we do not want to curate them for the purposes of the module. Examples are:
    1. The transactivation of a gene leading to production of a protein....we don't want to include all the steps of translation of that protein
    2. The degradation of a protein.
    3. Transport from one cellular organelle to another.

The compartment of a BlackBoxEvent that involves participants in more than one compartment (or movement of participants across multiple compartments), should reflect the compartments of all of the participants (inputs, outputs, catalyst).

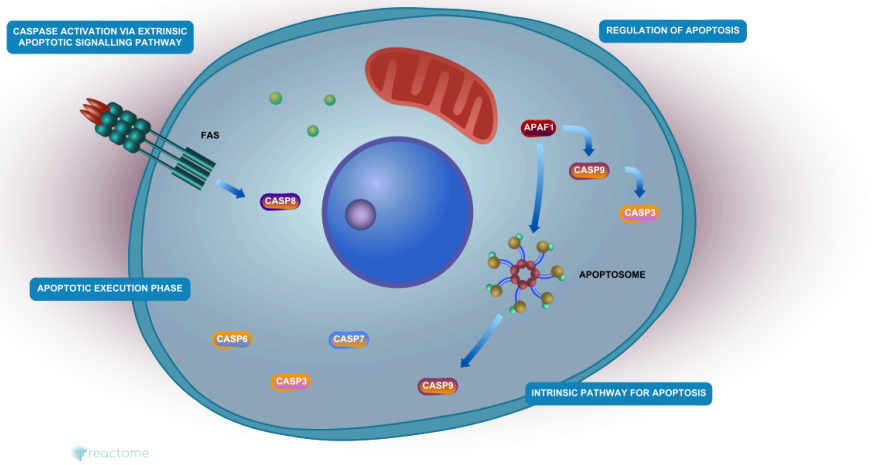


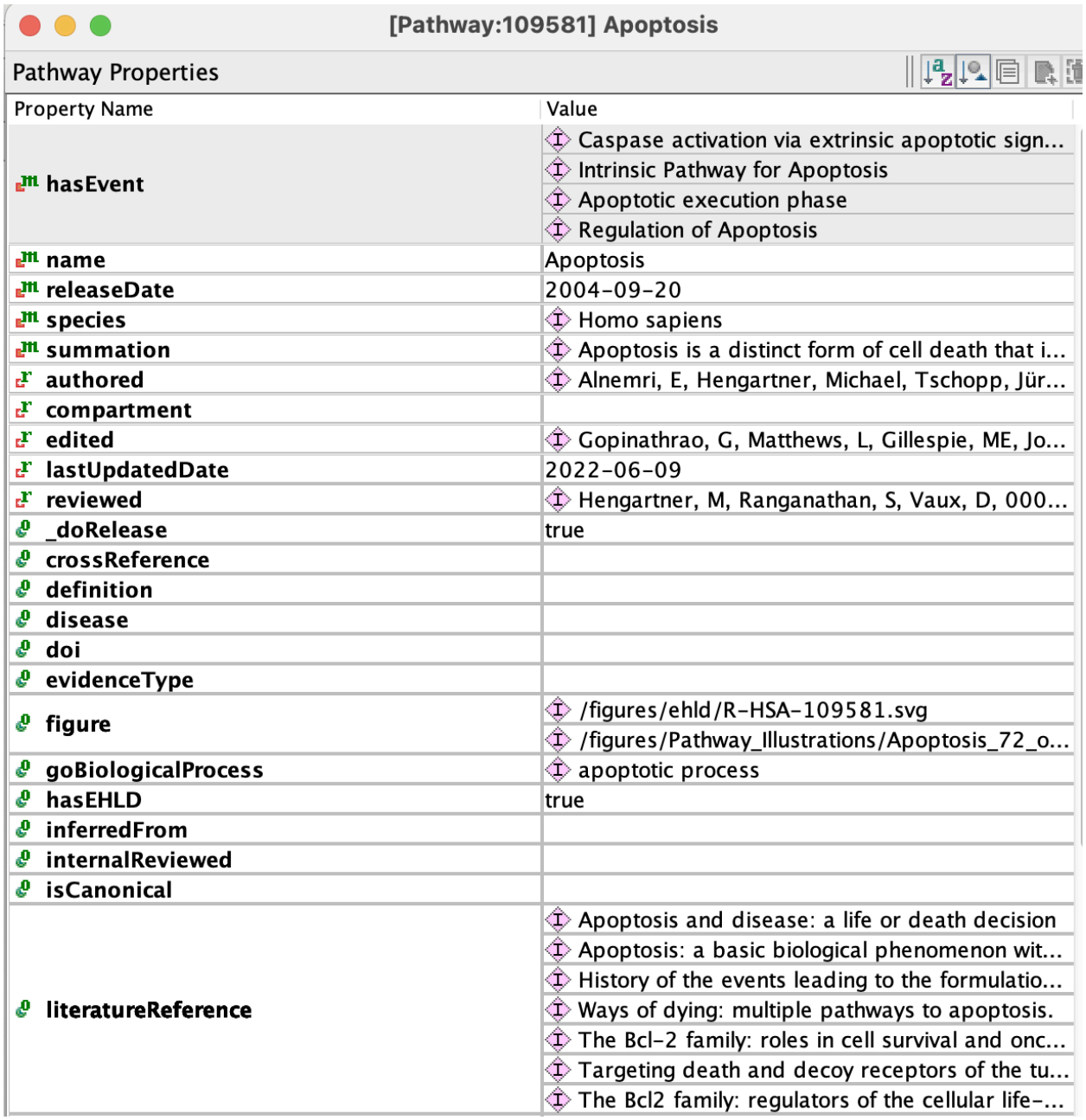
FailedReaction

See Disease section of the curator guide here (Annotating loss-of-function reactions).

Pathway

Pathway in Reactome is a multiple step event which can be broken down to subpathways and/or ReactionLikeEvents. A concrete curated example is the Reactome Apoptosis pathway which consist of the subpathways "Caspase activation via extrinsic apoptotic signalling pathway"", "Intrinsic pathway for Apoptosis", "Apoptotic execution phase" and "Regulation of Apoptosis".

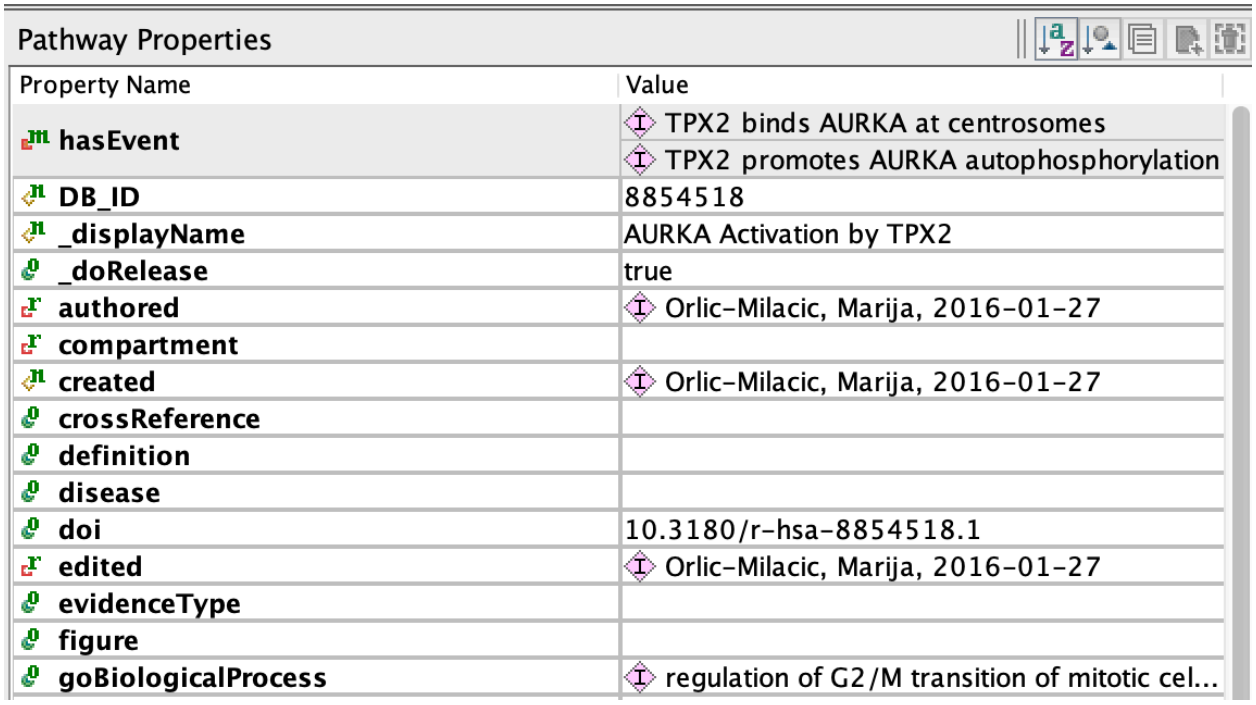




Mandatory, Required, and Optional properties (fields) for data classes

All the classes of data in Reactome have a number of properties (fields of information) that MUST be filled in before they can be 'released', i.e. made visible to the public. They also have required (where relevant) fields, optional fields, and non-editable fields.

When viewing an instance of a class in the curator tool, its properties all have a symbol that indicates the type, either a red box with a letter superimposed, or a yellow diamond with n superimposed. The yellow diamond with n indicates fields that are automatically completed. The red squares indicate, by letter,



- m - Mandatory - you must provide this information.
- r - Required - you must provide this information if it is possible - a good example of this is species, required for defined sets of proteins, but not required for defined sets of small molecules.
- o - Optional - not available for every instance, so cannot be “Mandatory” or 'Required' but in some circumstances may be a procedural requirement, e.g. inferredFrom must be completed for human reactions that are inferred from a model organism.

Guidelines for Naming Entities and Events

The guidelines were first set out in this paper (Jupe et al, 2014) describing “A controlled vocabulary for pathway entities and events”.
Please read the paper to learn how to name EWASs and events.

Here are some guidelines for the different classes of entities:

EWAS

Names are based on the UniProt gene symbol (the gene name), which is typically the same as the HGNC symbol. Post-translational modifications are represented by a prefix, cleavage events or differences from the canonical form are represented by a suffix.

Small Molecules

Please use the short name recommendations for small molecules found here (Appendix C6).

Complexes

These guidelines are an extension of the ideas for a pathway controlled vocabulary set out in this paper (Jupe et al, 2014). You should familiarize yourself with the rules for naming EWASs and simple entities.

In a typical scenario the names of sets and complexes are simply a joined list of the entities they contain, which in the Curator Tool can be copied to the name field from the hasComponent or hasMember fields.

Complexes should be named after the names of their components, separated by the colon symbol. Some examples: GRB2:SOS1 IL3:IL3RA:IL3RB:JAK2 IL4:p-Y-IL4R:JAK2:p-Y-IL2RG:JAK3:p-Y-JAK1:p-Y705-STAT3,p-Y641-STAT6 RXFP4:RXFP4 ligands

For cardinality (number of copies) use 2x, 3x etc. in front of the name. Is it OK to use dimer, tetramer etc. to indicate multimerization but NOT hetero- or homo- or the word complex.

When a complex is extended by the addition of further components, the name of the complex produced should show the 'added' component on the right end of the complex name.

For complexes larger than 4 components, it is acceptable, though less desirable, to use a descriptive name from the literature IF widely used, e.g. high-affinity IL3 receptor.

What about synonyms? You can add as many extra names as you like, listed after the 'correct' name in the name slot.

Authors/reviewers should NOT be allowed to dictate the name used. We have a controlled vocabulary designed to consistently name entities throughout the Reactome database. Authors may not like the name we use but that is usually because they have become used to a different name from the literature. Unfortunately there are often many alternative names in the literature and there is rarely 100% agreement on the best name to use. We can't make everyone happy, but we must pick one name from the

alternatives. The only way we can justify not using names suggested by (highly biased) authors is to be 100% consistent and always use our controlled vocabulary!

Remember - if you diverge from the simple list of colon separated entity names, you **must** check your name is not the same as existing complexes!

Sets

Set names: Small sets (less than 5 members) should be named as a list of the members e.g. IL3,IL3RA,CSF2RB.

Candidate sets should include the candidates within parentheses, e.g. CCR2,(CCR4,CCR6) is a candidate set where CCR2 is a member and CCR4 and CCR6 are candidates.

For large sets (5 or more) it is acceptable, though less desirable, to use a descriptive name from the literature IF widely used, e.g. XXXXXX or a 'function-defining' name that indicates the usage of the set, e.g. 'RXFP4 ligands'. If the large set includes an entire family of proteins, it is OK to use the name of the family, e.g. CCRs (note the s suffix to indicate that there is more than one entity represented), there's no need to add the word family. If the set represents a range of members, use hyphen to indicate the range e.g. CCR2-5. For complicated sets that have diverse members, a descriptive name based on function may be the best option. Try to describe the attribute that is the basis for grouping the entities, their shared property or function, e.g. SLC receptors, VAV2-activated RhoGEFs.

What about sets that have a common name stem, but are not so simple? It's tempting to try and come up with complicated abbreviations around the gene name common prefixes, e.g. for the real example set of genes PLA2G1B PLA2G2A PLA2G2D PLA2G2E PLA2G2F PLA2G5 PLA2G10 PLA2G12A, the name might be PLA2G1B,G2A,G2D,G2E,G2F,G5,G10,G12A but don't - this is too hard to interpret.

Remember - if you diverge from the simple list of comma separated entity names, you **must** check your name is not the same as existing sets!

Sequence variants

Reactome's approach to annotating sequence variants of proteins is described in detail in the section on Disease Annotation, below (Annotating Disease Entities), and is informed by the Human Genome variation society description (http://www.hgvs.org/mutnomen/examplesAA.html) of sequence changes at protein level.

References

1. Jupe S, Jassal B, Williams M, Wu G. A controlled vocabulary for pathway entities and events. Database (Oxford). 2014 Jun 20;2014:bau060. doi: 10.1093/database/bau060. PMID: 24951798; PMCID: PMC4064568.

Using Data Inferred From Other Species

Sometimes there is not sufficient experimental support for a set of reactions in one species, simply because the experiments have not been done in the target species. The InferredFrom slot addresses this problem. If the experts (perhaps also the curator) knows that this reaction really occurs in the target species, the reaction or reactions can be constructed in the experimentally supported species and then the InferredFrom slot can be filled in for the reactions in the target species. In this case the experimental species *becomes* the support for the target species. See a more complete description here under “Creating an inferred event”.

Providing evidence for assertions in Reactome

Principles for assigning literature references for reactions

The references associated with a reaction MUST provide direct experimental evidence for the occurrence of that reaction in the species you are annotating (i.e. human for human Reactome). If there is no direct experimental evidence in humans, but in other species, then you need to create an inferred human reaction as described in the section “Creating an inferred event”, below. See Guidelines on evidence confidence here (Appendix C3).

Other evidence types

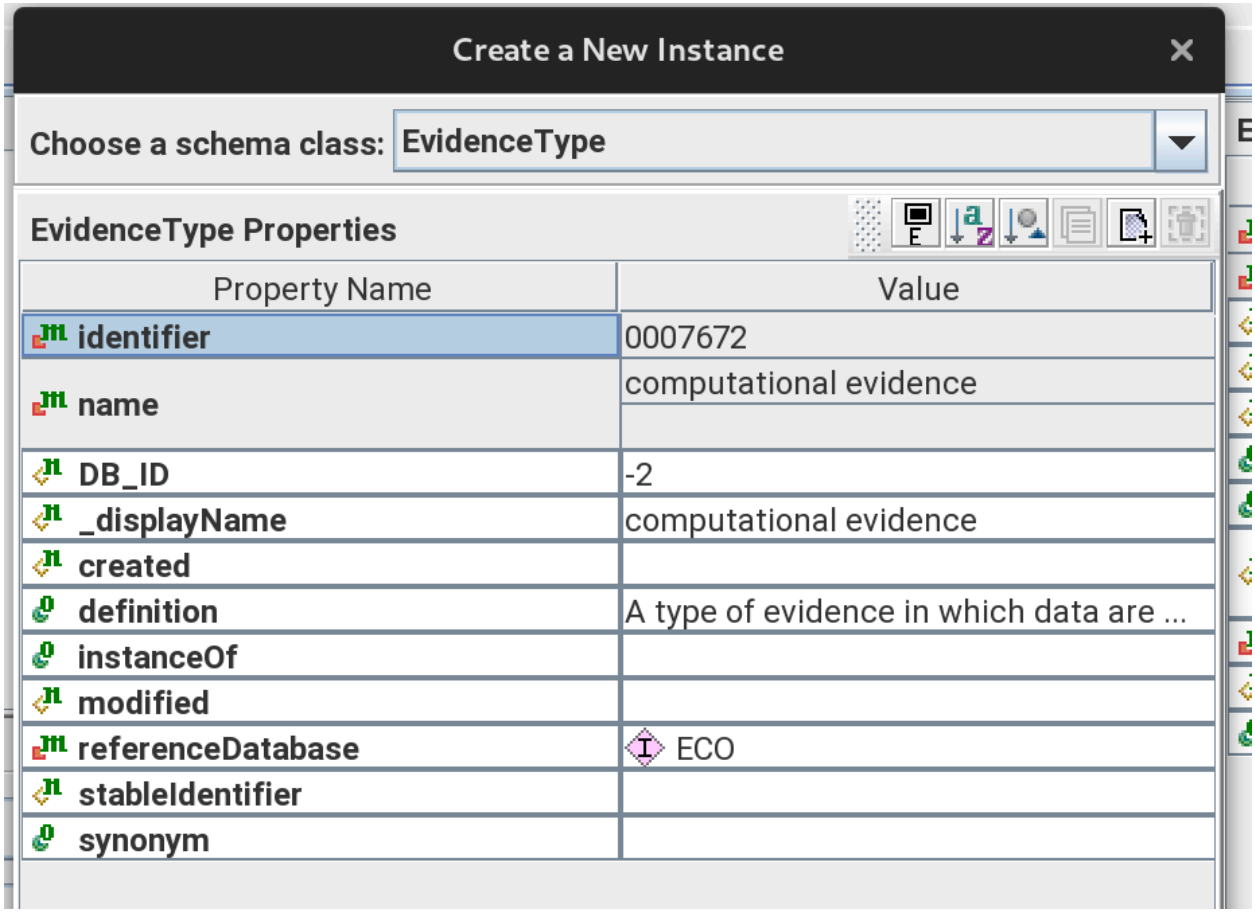
If applying the fundamental “direct evidence” rule leads to a case where a pathway just misses one or few reactions to be complete, it is possible to add missing reactions with less direct evidence. Such reactions MUST have the evidenceType slot filled and the summary MUST contain rationale on why the reaction was added, despite the weaker evidence.

It is also possible to use black box reactions when specific evidence is missing, in the following cases:

- localization reactions where the means of transport is unknown but the localization itself has direct evidence (e.g. microscopy)
- enzymatic reactions where the catalytic activity is unknown but the end product(s) themselves are supported by direct evidence (e.g. mass spectrometry, crystals)

How to assign an evidence type

1. In the curator tool, view your reaction and focus on the **evidenceType** slot
2. Set its value from a list of already existing values in your project, or values existing in the Reactome database, like with other slots
3. If you think a value is missing in Reactome database, create a new one by using a term from the Evidence & Conclusion Ontology (ECO) at https://www.ebi.ac.uk/ols/ontologies/eco
   In particular, the **name, definition, identifier, and referenceDatabase** slots need to be filled (see picture)



Citing biological databases

Generally, there shouldn't be a need to create ReferenceDatabase instances. Contact dev@reactome.org if you think you need a new ReferenceDatabase instance.

- Attributes: AccessURL- template used to form the URL that will be used to link to a particular record in an external database. Generally, this should not need to be touched. Contact the developers if you think any changes are necessary.
- URL- URL for the site which gives "summary information" about the database including a description of what information the db contains. This attribute should not need to be modified manually.

In certain cases, there are external databases that include information that is relevant to an annotated pathway.

One example is rare disease database such as RettBase, which describes Rett syndrome-related mutations in MECP2, FOXG1 and CDKL5: http://mecp2.chw.edu.au/. This database provides information about variants described in the disease pathway: Loss of function of MECP2 in Rett syndrome.
We are not able to cross-reference RettBase for individual MECP2 variants in the way that we cross-reference COSMIC for cancer variants because RettBase is not organized in a way that allows us to do this. However, we can mention RettBase in the pathway summation and cite their most recent publication (https://www.ncbi.nlm.nih.gov/pubmed/28544139) as a literatureReference for the Loss of function of MECP2 in Rett syndrome pathway. In addition, we can also create a second Publication instance of class "URL" and use this to cite the database itself.

NOTE that you should NOT cite disease databases at the reaction level **unless these disease databases capture biochemical/functional information.

Using the curator tool

Downloading, Installing, And Maintaining The Curator Tool

For instructions on downloading and installing the curator tool, see the public website here (https://reactome.org/download-data/reactome-curator-tool)

Windows:

Instructions

- After downloading, double-click install.zip to unzip it to an empty folder. This folder should be named "ReactomeCuratorTool" though any name can be used. Please use the latest version of WinZip if you use this application for unzipping.
- After unzipping, an application called ReactomeCuratorTool.exe should be found in the named folder. Double click this application to launch the Reactome Curator Tool.

Notes

- You need to install Java 11. You can download one from OpenJDK.

Linux:

Instructions

- After downloading open a shell, and cd to the directory where you downloaded the installer.
- At the prompt, type unzip install.zip to unzip the installer.

- Use ./ReactomeCuratorTool.sh to launch the Reactome Curator Tool.

Notes

- You need to install Java 11. You can download one from OpenJDK.

MAC:

Updated March 3, 2024 - Mac users now uses the Linux version of the curator tool. It is just too complicated to hack the native Mac installer. To run the curator tool under Mac, please following the following:

1) Download the installer for linux using a browser. After unzip it (it may be automatically unzipped after download), move it out from your Download folder to other place (e.g. Desktop).

2) Open a Mac OS terminal by double clicking “Terminal” in the Applications -> Utilities folder

3) Enter the following in the terminal to check if you have Java installed:

java --version

- If this fails ("The operation couldn’t be completed.", or something like that), install a Java JRE or JDK from https://adoptopenjdk.net (use Java 11, JDK. Make sure you choose the correct CPU architecture.)

4) At the same terminal after validating Java 11 is installed, issue command:

bash ReactomeCuratorTool.sh.

- The curator tool window should popup as usual.

5) Still another alternative for Mac users. After downloading and unzipping the Linux installer (as in step 1 above), assemble all the files into a single folder and name that folder CuratorTool (or any other unique and easy-to-remember name), and put that folder in a convenient directory on your Mac. (PD leaves it on the desktop). To launch the curator tool, open a terminal window on your Mac and cd to the CuratorTool folder:

```
cd desktop/CuratorTool
```

then, still in the terminal window:

```
bash ReactomeCuratorTool.sh
```

(The terminal window remains open while the curator tool is running.) The CuratorTool should open as usual. After you "quit" the CuratorTool, enter "exit" in the terminal window and quit the terminal window.

Curator tool default settings

All routine curation uses a connection between your local curator tool and the gk_central database. These are the settings for that connection, which you make via the "database info" tab of "curator tool preferences". Other values of these settings can be used for various specialized purposes, but these are outside the scope of regular curation and can be dangerous.
database host: curator.reactome.org
database name: gk_central
database port: 3306
user name: [your user name]
user password: [your password]
Guanming Wu assigns usernames and passwords for new curators; yours will persist unchanged thereafter.
Some institutional firewalls see connecting to port 3306 of a remote server as a security risk and will block your access to it. If this happens, you, Guanming, and your institutional IT people will need to negotiate a workaround, e.g., whitelisting this port at curator.reactome.org for your individual computer.

Curator tool user interface

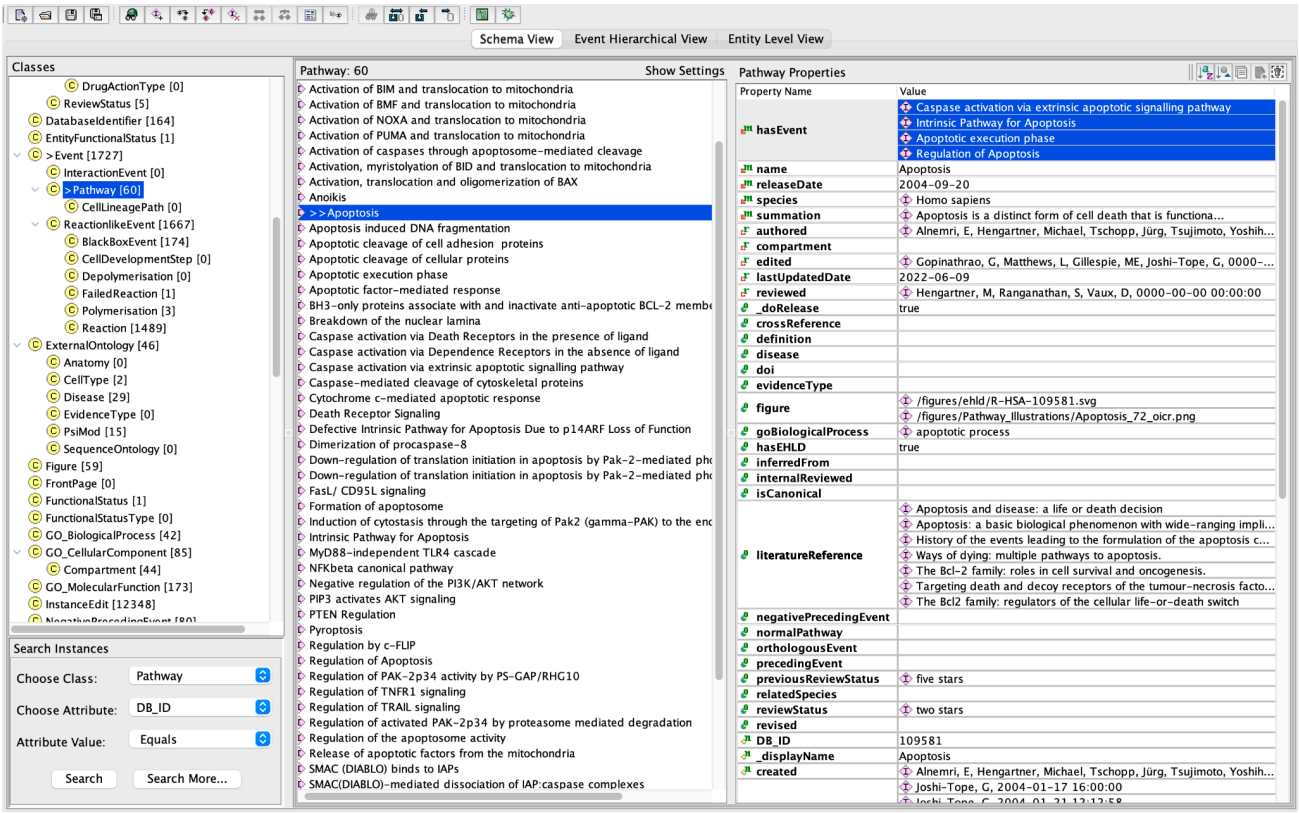
The curator tool has three panes to view the content. These different views can be switched using the tabs in the upper left hand corner of the curator tool window.

The Menu Bar

The menu bar contains a number of shortcut buttons, a scroll over message will tell you the function of each button.

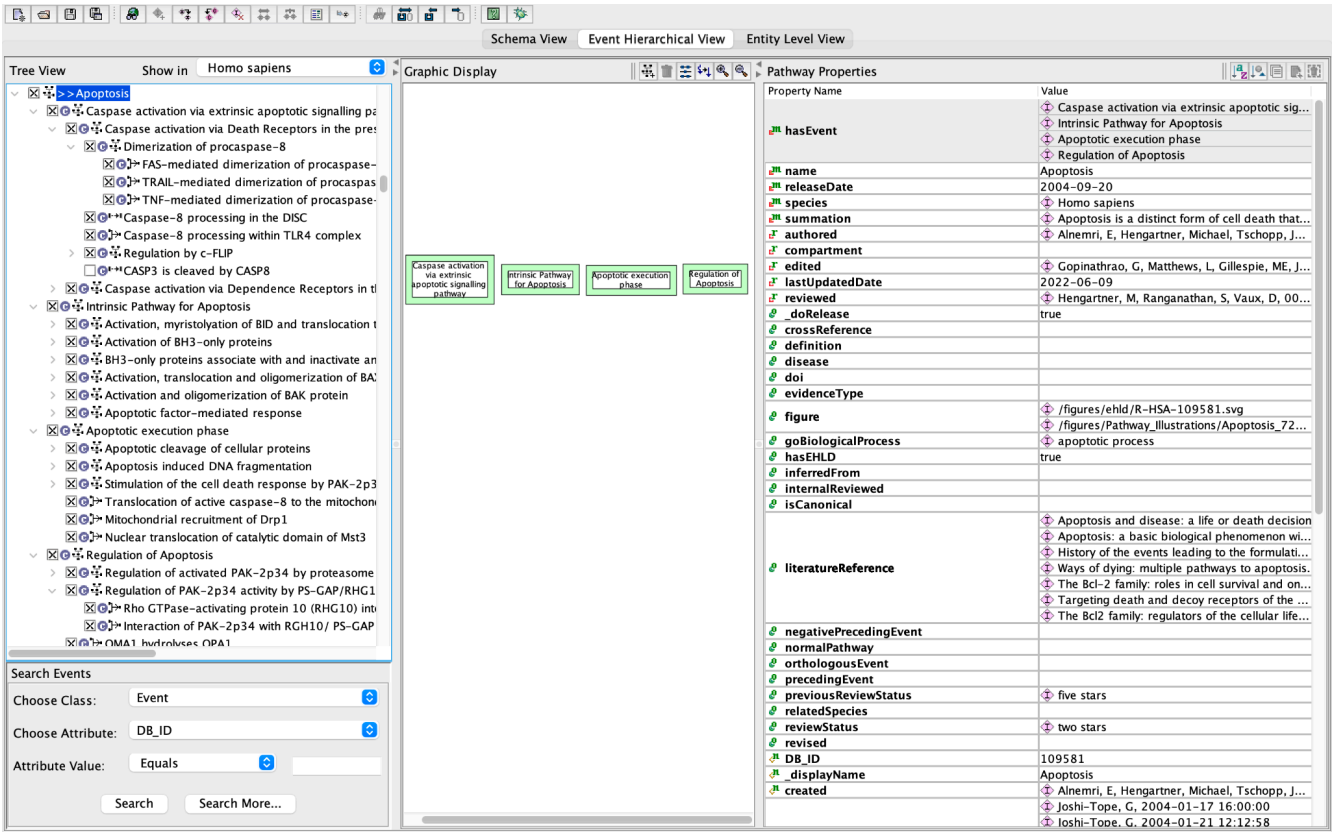
The Schema View

The Schema view provides a hierarchical list of the data classes. In this view you can search, open, edit and create instances of these classes. This panel is used most often for creating new items, such as an EWAS derived from a ReferenceGeneProduct.



The Event View

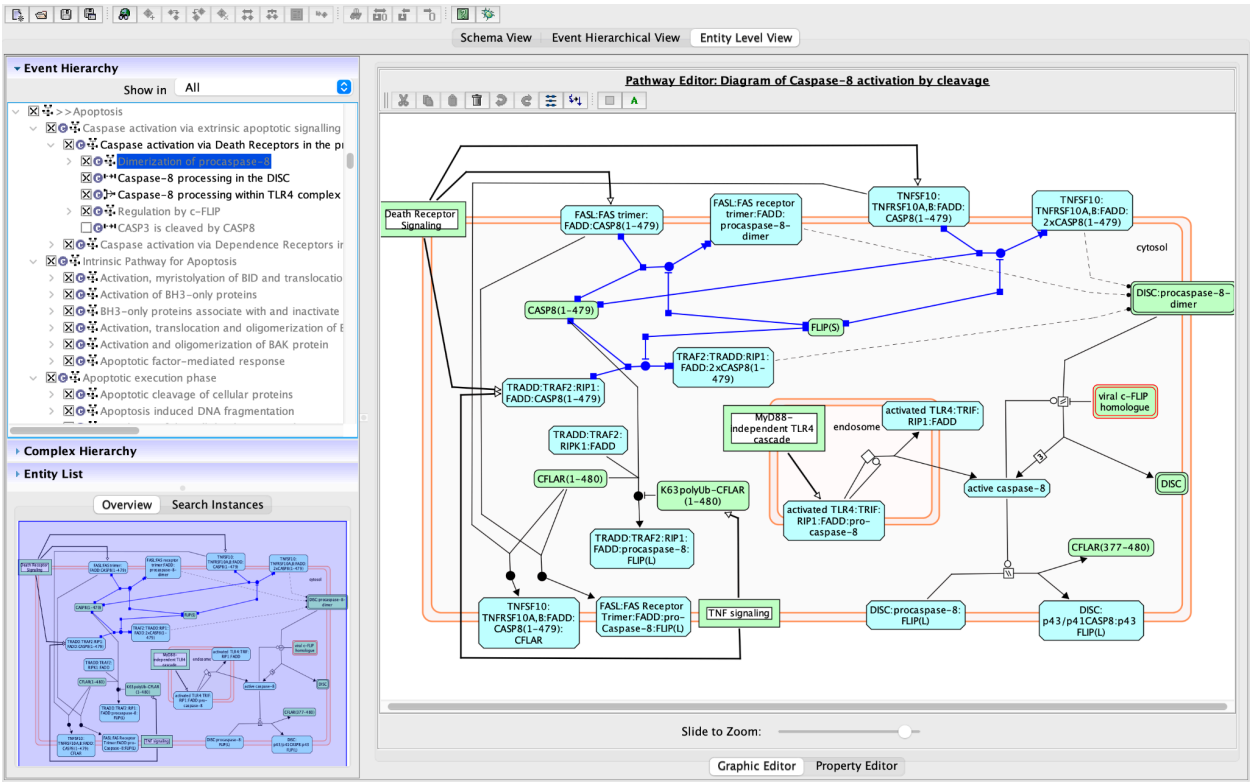
This view is where most curation is done. The event view displays a hierarchical, alphabetical list of the events in the local project on the left, a graphic in the middle that shows the relationship between objects for selected events, and on the right the details of selected events. Unfurling a pathway by clicking on the + symbol to the left of its name in this view reveals all of its component Reaction events. DON'T CONFUSE these symbols for the check box - clicking on an unchecked box marks the event as ready for release onto the public website; clicking on a checked box has the opposite effect. For pathways the check box selects the ENTIRE pathway, doing this by mistake can be time-consuming and tedious to reverse so be warned!



The Entity Level View (ELV)

The ELV is where pathway diagrams are created and subsequently represented. In this view the pathway is layed out graphically showing the inputs, outputs, catalyst where appropriate, and regulatory molecules, each in the correct cellular compartment.

For the usage please see the chapter on “Drawing a pathway diagram”, below.



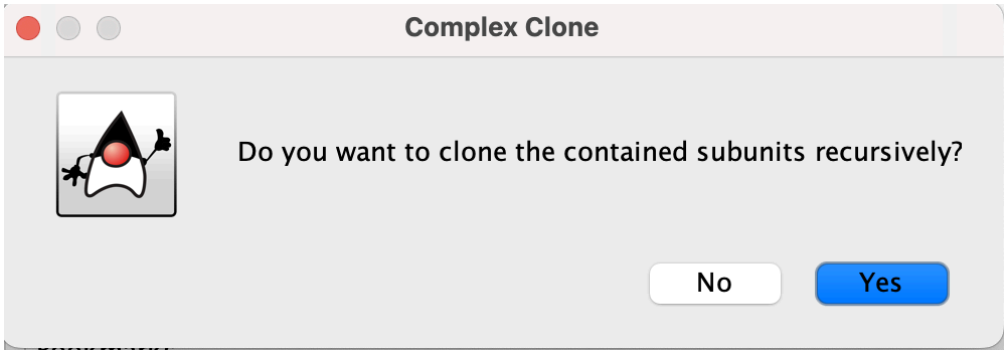
USING THE PRECEDINGEVENT SUGGESTION FEATURE IN THE CURATOR TOOL

See slides describing this process here (Appendix C7).

CLONING TOOL

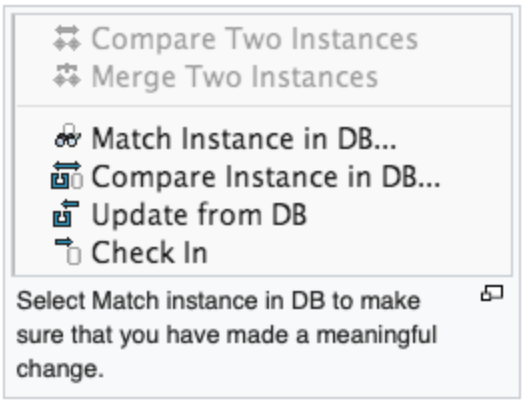
BE CAUTIOUS WHEN USING **THE CLONING TOOL.** Use it carefully and thoughtfully. The use of the cloning tool is restricted to making a copy of something that will be changed. When using the tool you must be absolutely sure that at least one of the defining attributes of that entity is different from the original.

Be particularly careful when cloning a complex or a set and are faced with the option to “clone contained subunits recursively” - the answer to this is almost always “no” since the danger of creating hard-to-track duplicate entities is high.



Once the cloning is done, immediately go back and change the defining attributes that need to be changed.

Once changed use the Select Match instance in DB to make sure that you have made a meaningful change.



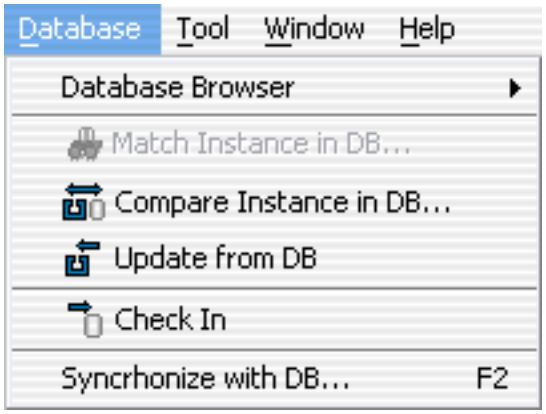
"Match instance in DB" feature to ensure that the instance is really different from the entity in gk_central. Do this by right clicking on the newly created instance.

Synchronizing local projects with the database

Synchronize at least once a day. It is good practice to synchronize your local project with the gk_central database to ensure that your work is not lost should the local copy become corrupted or lost due to hard drive failure. It is useful for other curators who may be able to reuse EWASs or other instances that you have created. It is perfectly acceptable to 'check-in' incomplete work unless it is marked as 'doRelease'.

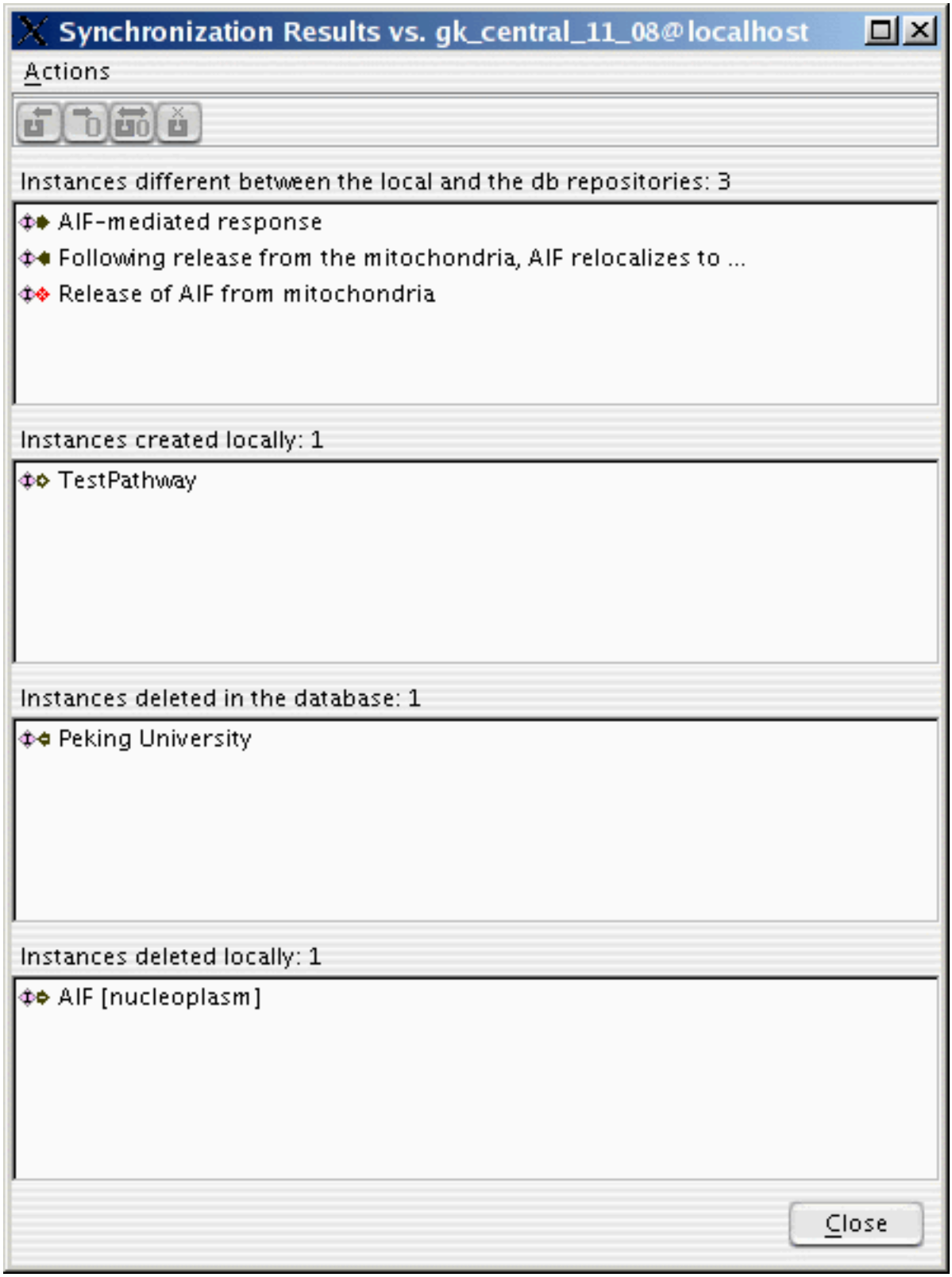
To keep the data in the opened project and the database repository consistent, you need to synchronize the local project with the database periodically. Several actions are available for you to do synchronizing. The following is the "Database" menu for doing

database-related actions:



"Match Instance in DB..." can find a matched instance in the database repository for the selected instance in the local repository. A matched instance has the same defining attributes as the local one. If a matched instance can be found, you can merge the local one to the database one. "Compare Instance in DB..." can compare a selected instance with the corresponding one in the database. Corresponding means the same DB_IDs. So you should not compare one instance checked from database A to an instance in database B because the same DB_ID might be assigned to different instances. "Update from DB" will update a checked out instance from the database. "Check In" is used to check in newly created or modified instances to the database. It is recommended that you should use "Compare Instance in DB..." first before "Update from DB" or "Check In". After a new instance or modified instance is checked into the database, the marker ">" will be removed to indicate that the local and the database copies are the same.

The "Synchronize with DB..." menu is used to check all instances in the selected schema class in the schema view or in the whole opened project if no class is selected. There are four possible inconsistency categories between the two repositories: instances different between the local and the db repositories resulting from instance modification either locally or in the database, instances created locally, instances deleted in the database by you or others, and instances deleted locally by you. You can choose appropriate action for the selected instances in different categories. Please be aware that if you do a multiple selection from different categories, the enabled actions are applied to all selected categories. For example, actions "Update from DB" and "Commit to DB" can be applied to instance "AIF-mediated response" in the first category, but only "Commit to DB" can be applied to instance "TestPathway" in the second category. If you select both "AIF-mediated response" and "TestPathway", only "Commit to DB" is enabled and "Update from DB" is disabled. Double clicking an instance in the first category will popup a comparison dialog of the local and database instances, while double clicking an instance in other categories will show the contents of the clicked instance. To deselect a single instance, hold the control key and click the selected instance.



There are three different cases in the first category, instances different between the

local and the db repositories: instance is modified in the local project but not in the database, instance modified in the database but not in the local project, and instances are modified in both the local and database. The user can commit changes to the database or overwrite changes by updating from database for the first case. The user cannot commit changes for an instance in the second or third case, but can update from database for both cases.

The following is the list of icons used in the synchronization dialog:

Icon	Applicable actions	Note
	Update from DB, Commit to DB, Show Comparison	New changes in local instance only
	Update from DB, Show Comparison	New changes in database instance only
	Update from DB, Show Comparison	New changes in both local and database instances
	Commit to DB	New local instance
	Update from DB, Commit to DB	Local instance is deleted but still exists in the database
	Update from DB, Commit to DB, Clear Record	Database instance is deleted but still exists in the local

What do I do if the curator tool flags instances as being duplicated in gk_central

You will need to evaluate each instance that is flagged as being a duplicate and compare it to the instance that is flagged in gk_central. This is easiest to do before synchronizing your project with the database as described below:
Search the local project for all database instances that contain '-', then highlight each one in turn, right-click to open a pop-up menu, and choose "match instance in DB". Any results returned by that query are already-existing instances that, by the rules of gk_central, are duplicated by the new local instance. Mostly, that's right and you should accept the option offered by the form, to replace the local instance in the local project with the already-existing one from gk_central. That cleanly gets rid of the duplication and preserves all references correctly in the local project and in gk_central. Sometimes (rarely) the form is wrong because, despite identical defining attributes, two instances are genuinely different, e.g., if a person instance has already been created for J(ohn) Doe and you have now created one for J(ane) Doe. In this case, you should refuse the offer made by the form and should make a note to force the new instance into gk_central at synchronize-and-commit time.

What to do if the curator tool reports an exception when committing a project

In the event that an exception is reported by the curator tool during the commit to gk_central, this may be the result of a dropped connection with the database. Try to rerun the synchronization. If any conflicts in instances are revealed (this should be rare), send Guanming the project and explanation of the problem. If, after resynchronizing and attempting to commit again, if you still get an exception, verify that you have a connection to the database. If no connection is available, try the commit again when the connection is reestablished. Contact Guanming if there are any problems.

Problem of commit size

Usually as a curator in a team, you want to make many small commits instead of one big chunk. The reason is to minimize the chance of other curators changing the database entities while you are working on them. Resolving these conflicts is much easier with small project files

Solution: selected rewrite

After checking a project file for correctness the Reactome editor synchronizes the local project with the database. Here conflicts with commits of other curators become visible, and a list of DB_IDs that are in conflict is presented. Instead of rewriting the whole project file, the external curator can follow these steps:

-
- Refresh your view of the gk_central (menu Database-->Database Browser-->...)
- Select the conflicting entities, one by one, in your project file (not in the database)
- Right-click on an instance with the conflict, select "Update from DB"—this resets the instance to the version in the current database
- Re-enter your work on that reset instance
- save the project file with a different name
- Repeat above three steps for each instance with a conflict.
- Commit project.

The Annotation Process for Traditional Reactome Pathways

Note that this section refers to “traditional” Reactome pathway. The annotation of a special class of pathways referred to as Cell Lineage Paths is described in detail in a separate section below.

Create the framework for the pathway(s) you intend to curate before filling in the details. Start by creating a pathway, add to this new or existing reactions in the correct order, complete the summations and literature citations, then identify the EWASs, Complexes, Sets etc. required and complete the details of the reactions consecutively. Cascading signaling processes can involve very complicated Complexes, in these circumstances

the Graphic Display in Entity Hierarchical View is very useful as an overview of the order of events.

PRECURATION

1. Create a basic outline of pathway: Regulation of Apoptosis as an example. Further description of the annotation process will focus on the subpathway highlighted below in yellow.

I. Regulation of Apoptosis

II. Regulation of activated PAK-2p34 by proteasome mediated degradation

IIa Ubiquitination of PAK-2p34

IIb Proteasome mediated degradation of PAK-2p34

III. Regulation of PAK-2p34 activity by PS-GAP/RHG10

IIIa. Rho GTPase-activating protein 10 (RHG10) interacts with caspase-activated PAK-2p34

IIIb. Interaction of PAK-2p34 with RGH10/ PS-GAP results in accumulation of PAK-2p34 in the perinuclear region

IV. Role of DCC in regulating apoptosis

IVa. Caspase cleavage of DCC

IVb. DCC interacts with DIP13alpha

IVc. Caspase-9 binds DCC:DIP13alpha complex

IVd. Activation of caspase-3

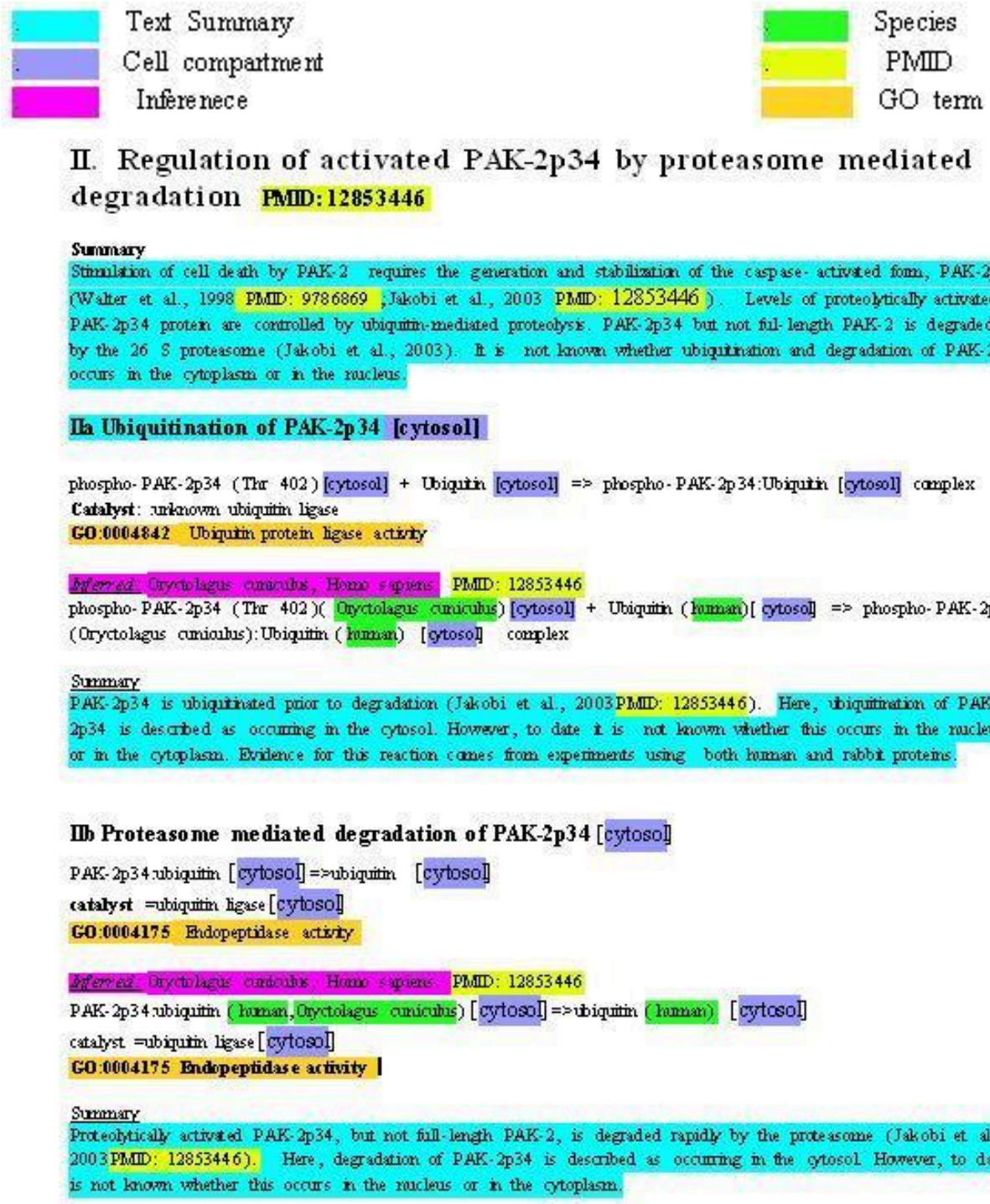
IVe. Caspase cleavage of UNC5B

IVf. DAPK binds UNC5B

IVg Caspase cleavage of UNC5A

IVh UNC5A binds NRAGE

2. Flesh out outline with information including: molecules, compartment, species, text summary, references (PMIDs)



3. Create table of of molecules participating in the pathway

-Each modified form of a protein as a separate entry
-Look up/enter corresponding UniProtID identifiers

	A	B
1	**Molecule**	**Uniprot ID**
2	Graf-2/ Arhgap10	A1A4S6
3	Death-associated protein kinase 3|DAPK3|DAPK3_HUMAN|DAPK3	O43293
4	DCC's fragment with the ADD domain (fragment)	P43146
5	DCC|DCC_HUMAN|Netrin receptor DCC	P43146
6	DCC's inhibitory C-terminal domain (fragment)	P43146
7	Death-associated protein kinase 1|DAPK1|DAPK1_HUMAN|DAPK1	P53355
8	phospho-PAK-2p34 (Thr 402)	Q13177
9	UNC5C	O95185-1
10	UNC5A fragment with death domain	Q6ZN44
11	UNC5A|Netrin receptor UNC5A|UN5A_HUMAN	Q6ZN44
12	UNC5B fragment with death domain	Q8IZJ1
13	UNC5B|Netrin receptor UNC5B|UN5B_HUMAN	Q8IZJ1
14	Death-associated protein kinase 2|DAPK2|DAPK2_HUMAN|DAPK2	Q9UIK4
15	DCC-interacting protein 13-alpha|DP13A_HUMAN|APPL1	Q9UKG1
16	NRAGE|Melanoma-associated antigen D1|MGD1_HUMAN|MAGED1	Q9Y5V3
17	unidentified caspase	unknown
18	Ubiquitin ligase	unknown
19	DAPKs	Candidate set
20	ubiquitin	Defined set

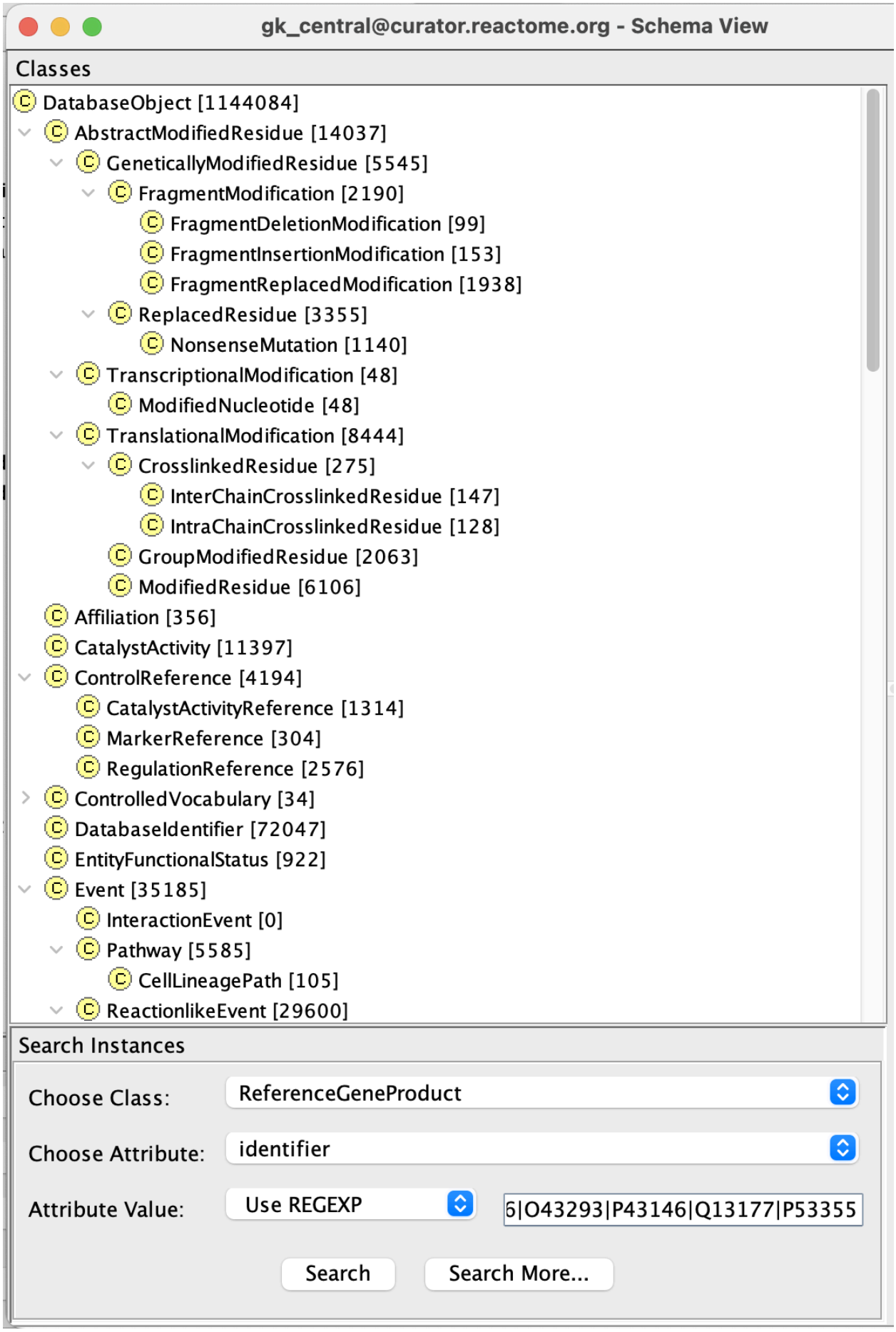
Data entry

Identify existing proteins/molecules

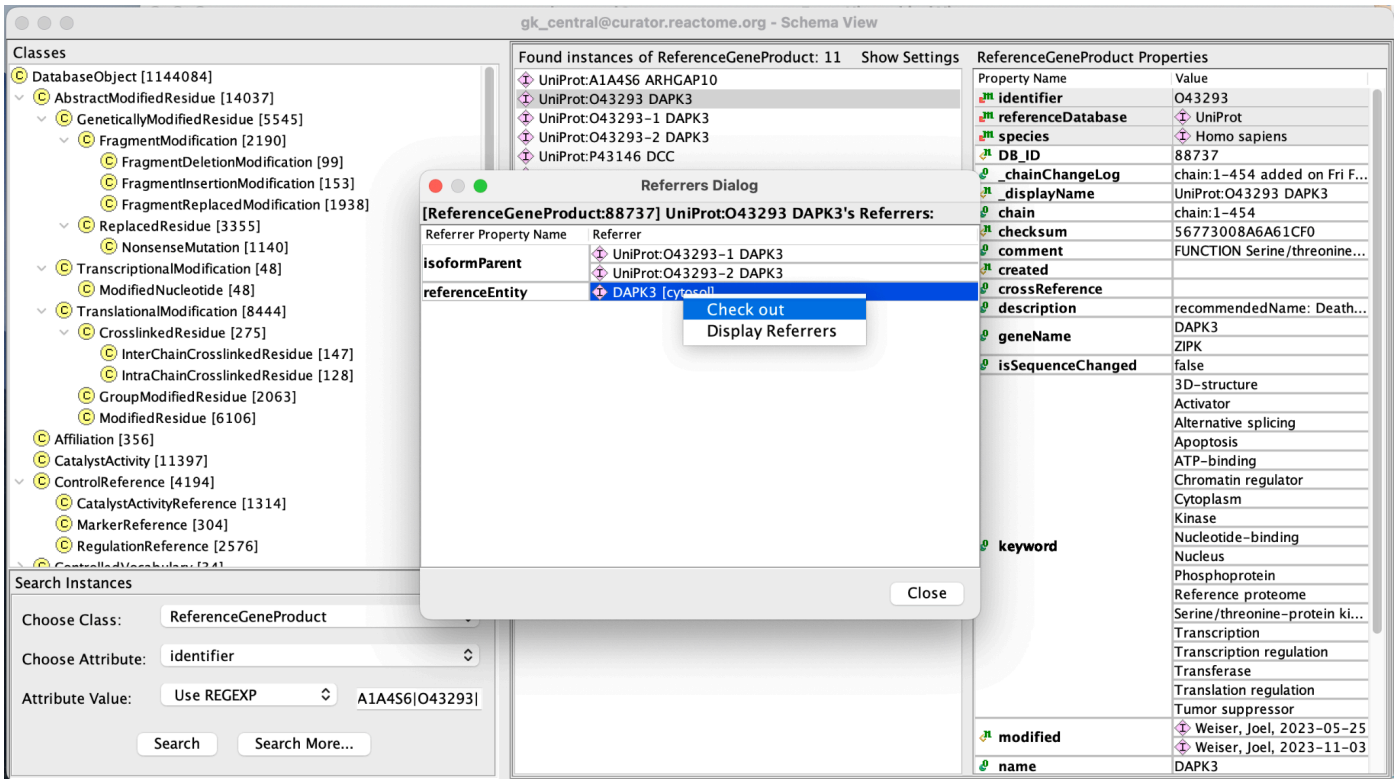
In many cases the proteins on your list will already exist in the database. You should make every effort to reuse existing instances wherever possible to avoid unnecessary and confusing duplications. Use the curator tool to search the ReferenceGeneProduct (RGP) class using Uniprot identifiers as follows:

Searching the database

Choose Class: ReferenceGeneProduct
Choose attribute: identifier
Attribute value: Use REGEXP
Enter your identifier in the search box. You can enter several as a pipe separated list (e.g A1A4S6|O43293|P43146|P43146....).


gk_central@curator.reactome.org - Schema View
Classes
DatabaseObject [1144084]
AbstractModifiedResidue [14037]
GeneticallyModifiedResidue [5545]
FragmentModification [2190]
FragmentDeletionModification [99]
FragmentInsertionModification [153]
FragmentReplacedModification [1938]
ReplacedResidue [3355]
NonsenseMutation [1140]
TranscriptionalModification [48]
ModifiedNucleotide [48]
TranslationalModification [8444]
CrosslinkedResidue [275]
InterChainCrosslinkedResidue [147]
IntraChainCrosslinkedResidue [128]
GroupModifiedResidue [2063]
ModifiedResidue [6106]
Affiliation [356]
CatalystActivity [11397]
ControlReference [4194]
CatalystActivityReference [1314]
MarkerReference [304]
RegulationReference [2576]
ControlledVocabulary [34]
DatabaseIdentifier [72047]
EntityFunctionalStatus [922]
Event [35185]
InteractionEvent [0]
Pathway [5585]
CellLineagePath [105]
ReactionlikeEvent [29600]
Search Instances
Choose Class:
ReferenceGeneProduct
Choose Attribute:
identifier
Attribute Value:
Use REGEXP
6|O43293|P43146|Q13177|P53355
Search
Search More...

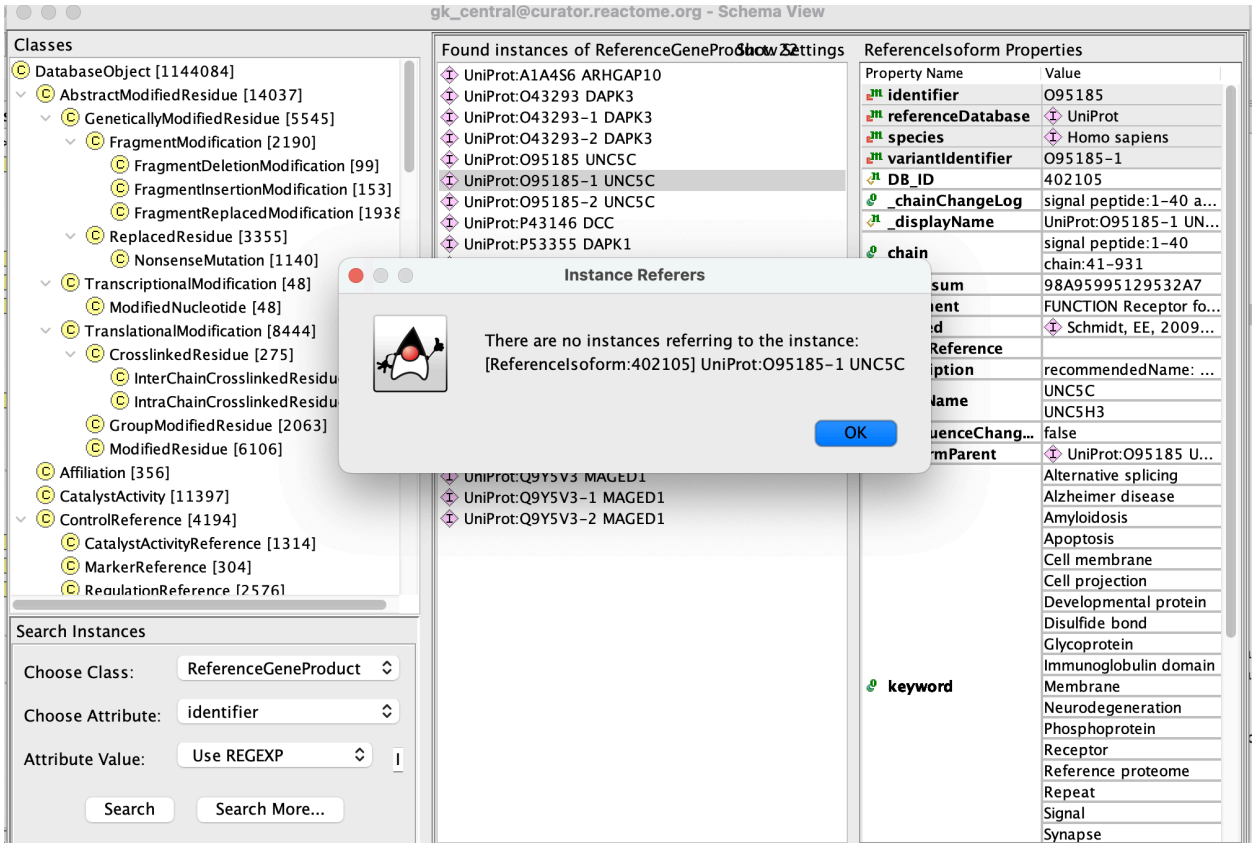
Select the ReferenceGeneProduct (RGP) returned, if a list select them one at a time, right click and opt to "View referrers". The resulting Referrers Dialog box lists Referrers by property name. At the top of the list can be isoforms of the protein, if present in Uniprot. Do not use isoforms unless you are certain that only specific isoforms have the functionality you intend to represent. Isoforms may have their own referrers and you should check this - if someone took the trouble to create an isoform-specific EWAS they probably had good reason to do so. Items listed with the property name referenceEntity are EntitiesWithAccessionedSequence (EWASs), a Reactome identifier for specific forms/locations of a protein. Often there will be more than one EWAS for a single RGP, because post-translationally modified forms of proteins and proteins in different cellular compartments are each represented by a different EWAS. If any of the listed EWASs correspond to your needs, right click and opt to "Check Out" that referrer. If the correct molecular compartment or post-translationally modified form is not present, you can still check out an EWAS and later use the Curator tool to clone it and modify it to your needs.



Create new proteins/molecules

Please see the guidelines on naming entities in this paper (Jupe et al, 2014)! If you search for a RGP that does NOT have referrers in the database you will get this

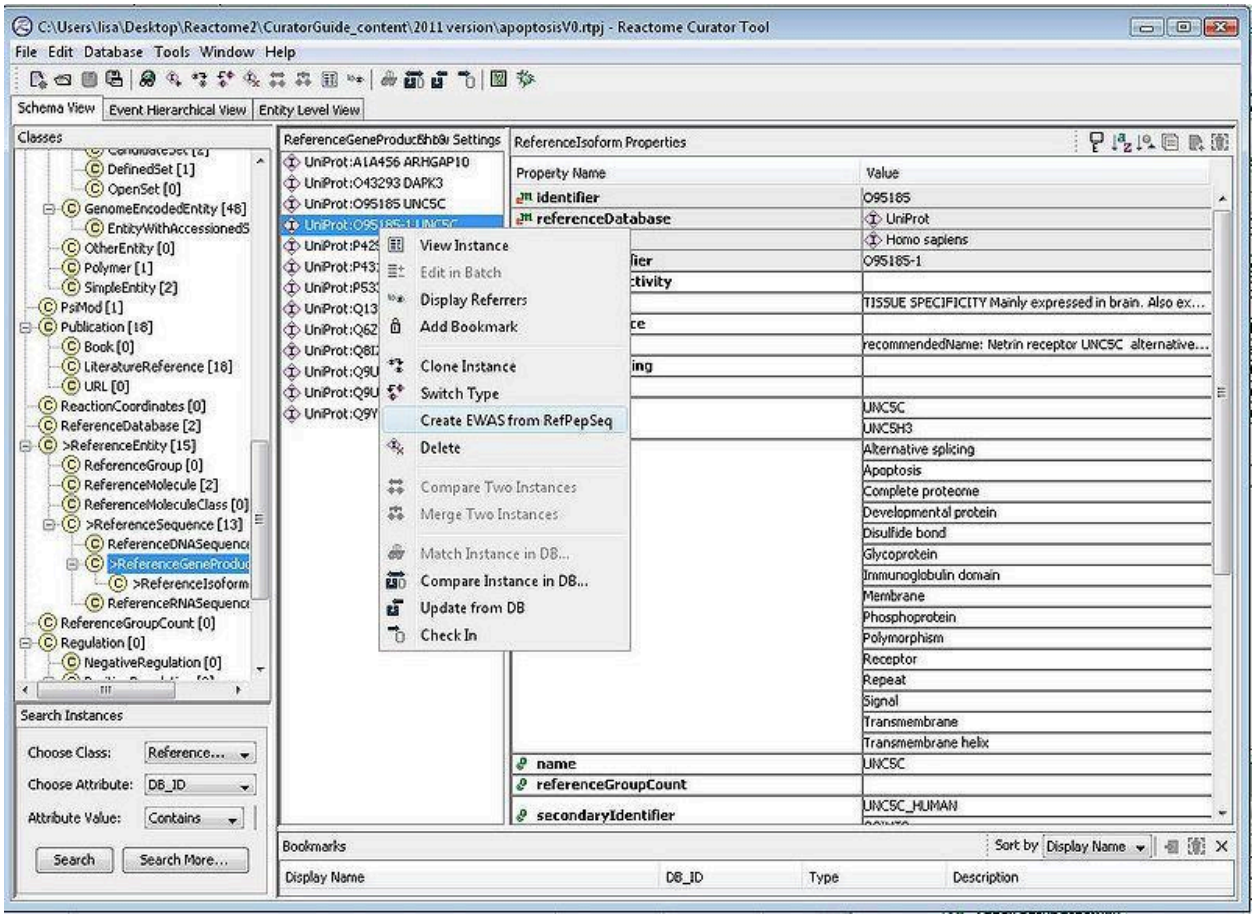
message:



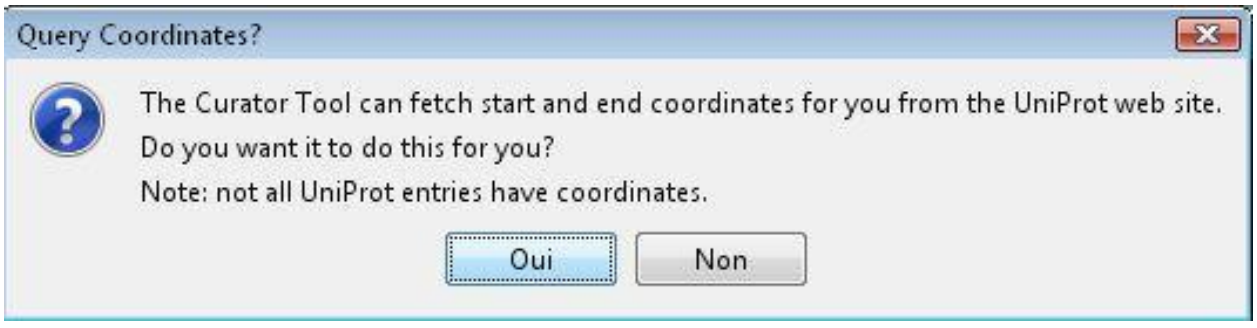
In this case, you will need to create the EWAS.

Creating a new EWAS

To do this, first check out the RGP of interest into your local project. Then, in the curator tool, select RGP in the class list, then scroll or search for the RGP of interest. Right click and opt to create EWAS from RGP.

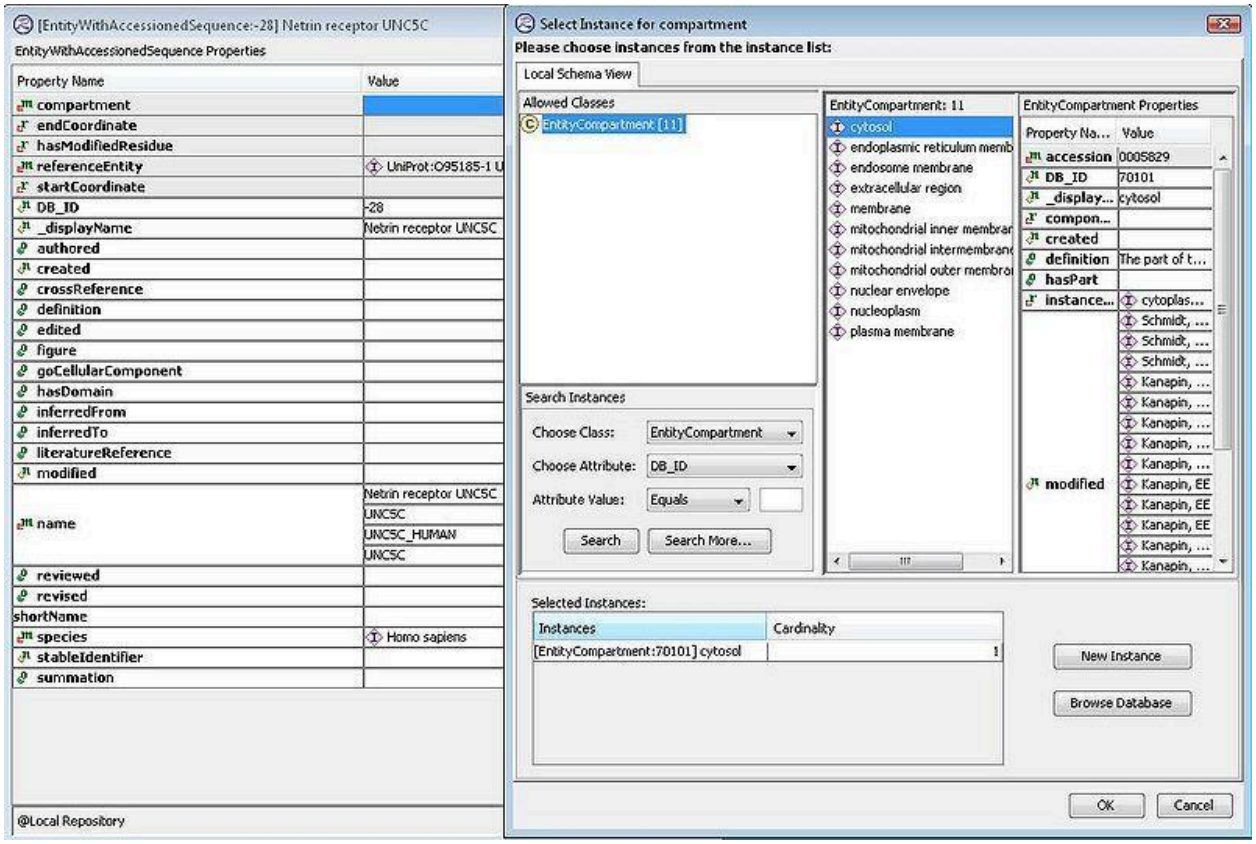


It will ask if you want to accept the end coordinates described by Uniprot. Only say yes if you can confirm them to be accurate. If the end coordinates are not certain, the convention is to represent the start as 1, and end as -1.



In the newly created EWAS, the ReferenceEntity will have a name and species entered by default. You can add an alternative name if this was specified by the Author, do this by right-clicking on the existing name and select Add. You must define the compartment. To do this, right click in the compartment slot and select the correct compartment in your local repository. Select the compartment and hit OK. If you don't

see the desired compartment in your local project , click the "Browse database" button to search in gk_central. ***If you still don't find the necessary compartment, DO NOT create a new one. Contact the managing editor. New compartment creation requires multiple steps as well as changes in the UI which need to be coordinated with developers.***



References

1. Jupe S, Jassal B, Williams M, Wu G. A controlled vocabulary for pathway entities and events. Database (Oxford). 2014 Jun 20;2014:bau060. doi: 10.1093/database/bau060. PMID: 24951798; PMCID: PMC4064568.

Translational modifications (see note below about how to make sure all PTMs register correctly in diagrams):

If the protein you want to represent with an EWAS is **post-translationally modified**, then the hasModifiedResidue slot of the EWAS is to be filled. Right clicking on the hasModifiedResidue slot of the EWAS will allow you to select a modification instance from one of the 3 child subclasses of the 'TranslationalModification' class, namely ModifiedResidue, GroupModifiedResidue and CrosslinkedResidue. You will be

prompted to choose from your local project, browse gk_central or create a new modified residue instance. Almost always you will want to create a new instance, as modified residue instances are specific to the RGP and residue position. If EWAS instances are created from the same RGP but belong to two or more cellular compartments, then these EWASs, e.g., EWAS (cytosol) and EWAS (nucleoplasm), may share modified residue instances when modified at the same residue.

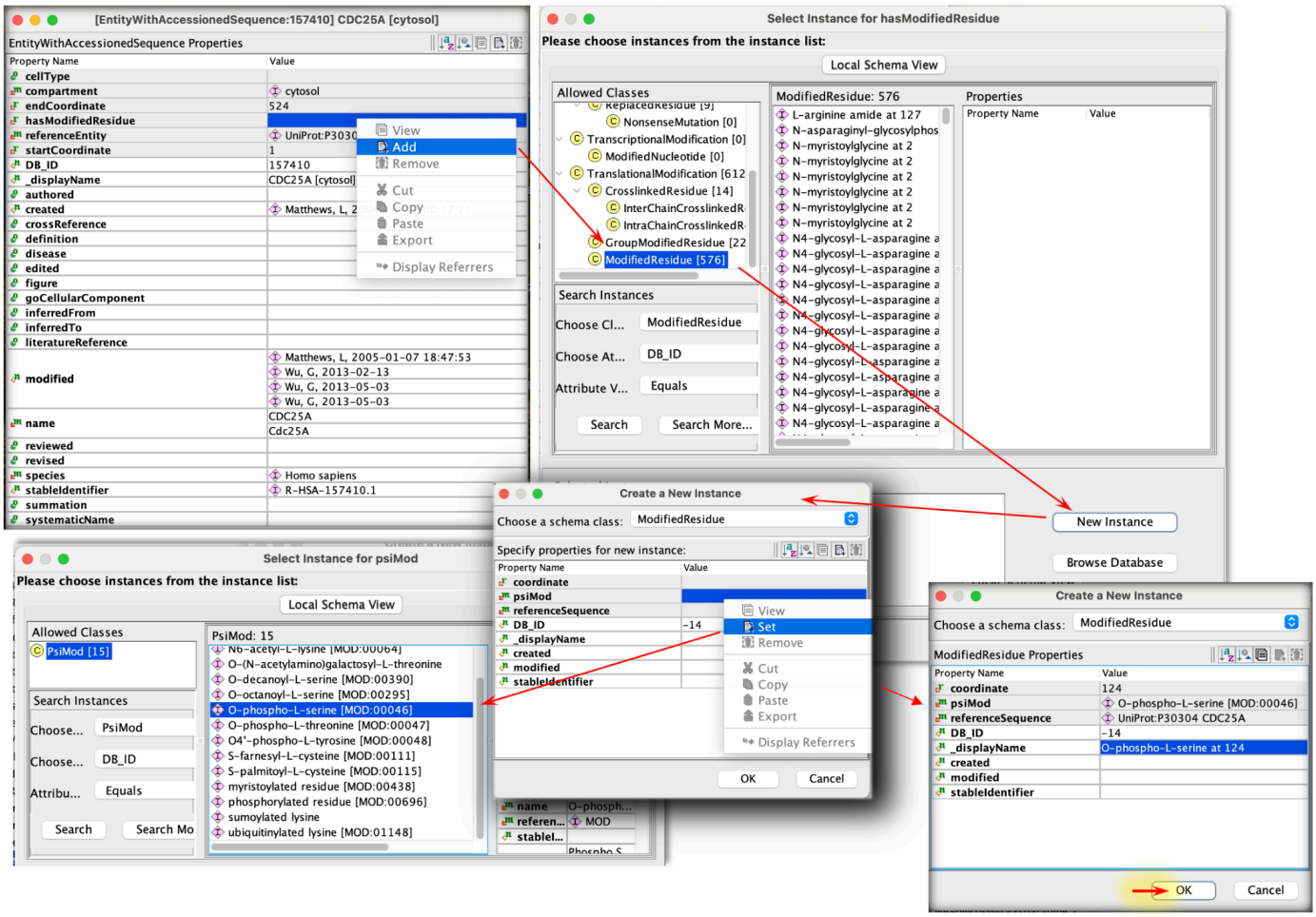
- *ModifiedResidue:*

This instance is added to EWAS to indicate that a chemical group such as phosphate or methyl group is added to an amino acid residue of a protein. Two mandatory slots need filling for a modifiedResidue instance - psiMod and referenceSequence. For example, to indicate that CDC25A protein has a phospho-serine modification at residue 124 you need to right click in the hasModifiedResidue slot of the EWAS, then select ‘Add’ from drop-down list and choose ModifiedResidue as a schema class in the pop-up window. Then, click on the ‘New Instance’ button to open a ‘Create a New Instance’ form. Right click in the PsiMod slot of the modification instance to select O-phospho-L-serine [MOD:00046] as a modification type and hit ok (a modification type can be found within the local repository or by searching the gk_central database).

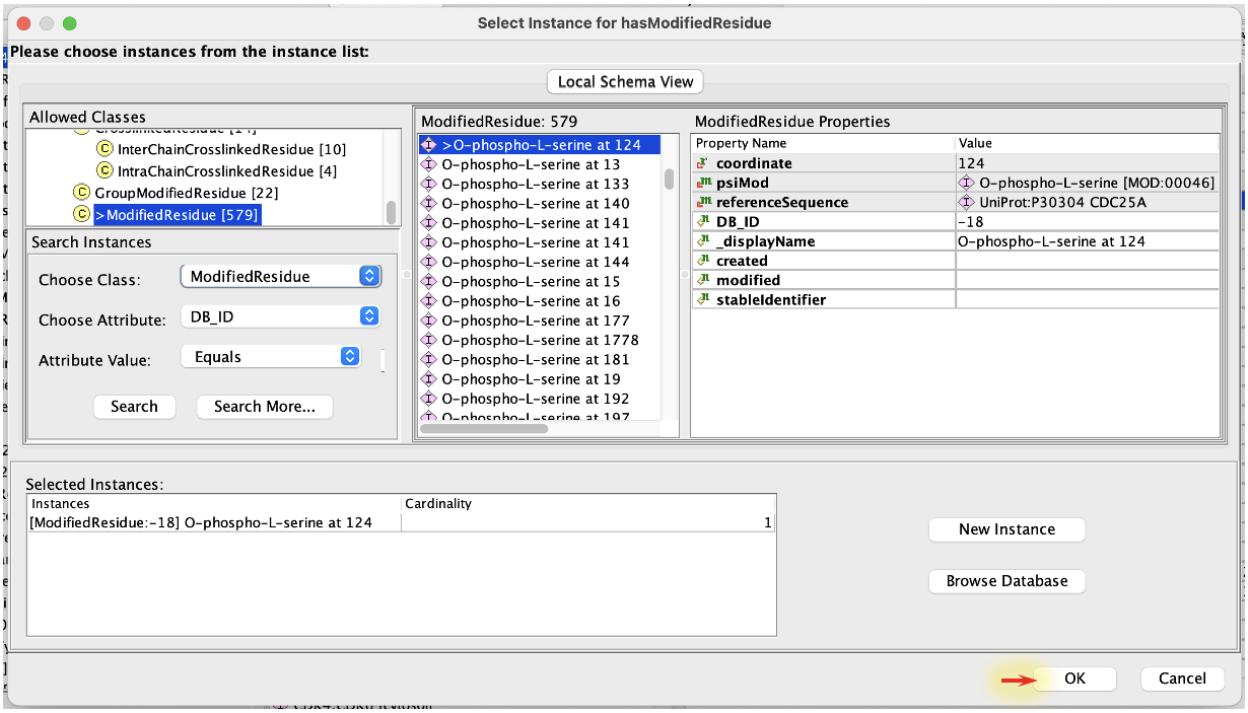
If a needed modification is not already present in gk_central, look up its identifier (a five-digit number) at the PSI-MOD website; copy-paste it into the "create new instance" form in the curator tool, and allow the wizard associated with the form to retrieve all other needed data from PSI-MOD. Once created, such an instance (like all other reference instances in gk_central) should not be modified. If an instance is needed that is not already in PSI-MOD, or if a PSI-MOD instance appears to be incorrect, contact PSI-MOD to resolve the problem.

Similarly, right clicking in the ReferenceSequence slot of the modification instance allows selecting RGP with Uniprot identifier for CDC25A protein. Finally, enter the residue number 124 in the coordinate slot and hit ok.

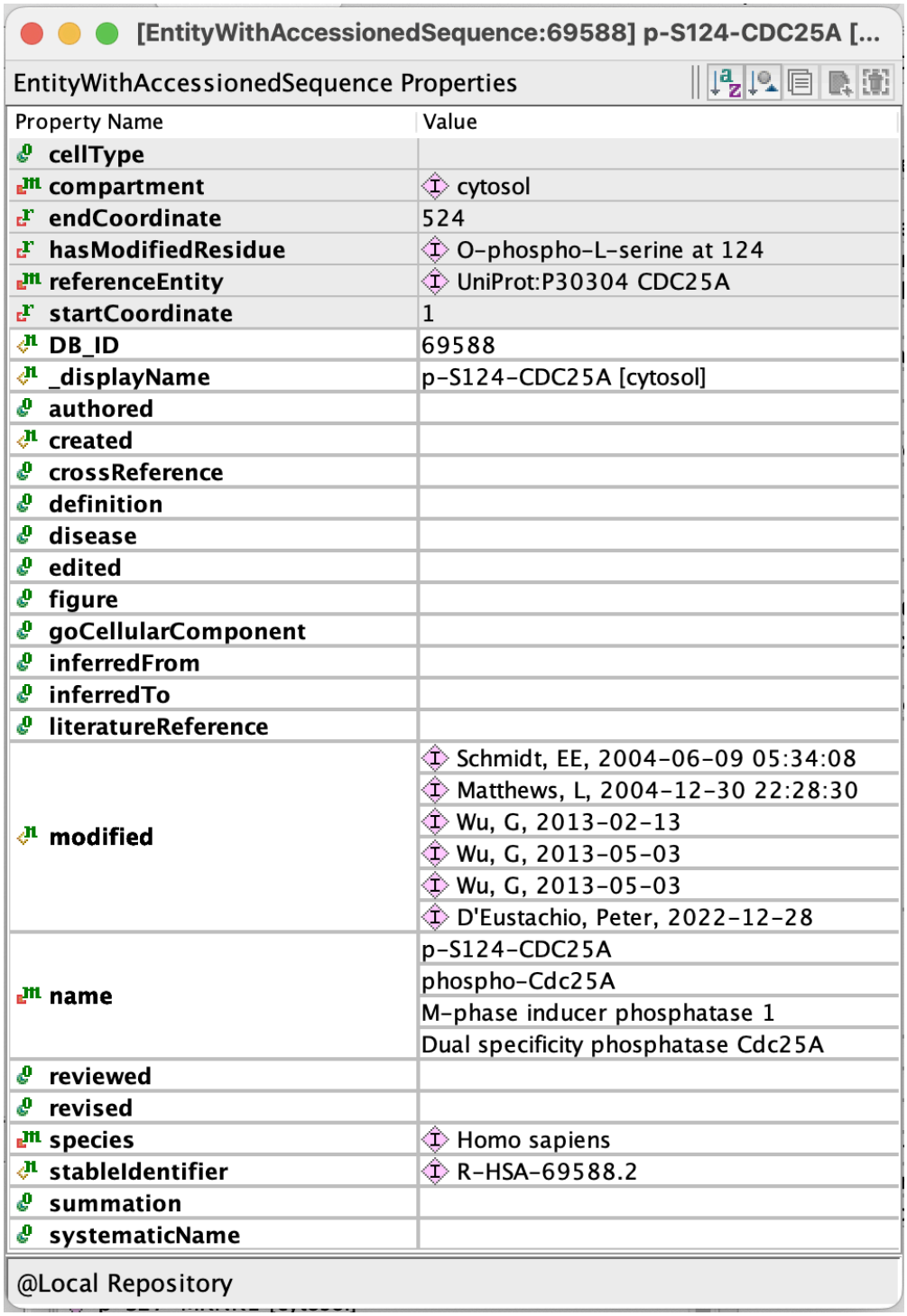
Note that the curator tool does not prevent you from adding a modifiedResidue instance with one ReferenceGeneProduct onto the EWAS record with a different ReferenceGeneProduct, and the confusion that results is substantial!



The modified residue instance can now be applied to the EWAS by clicking ok.

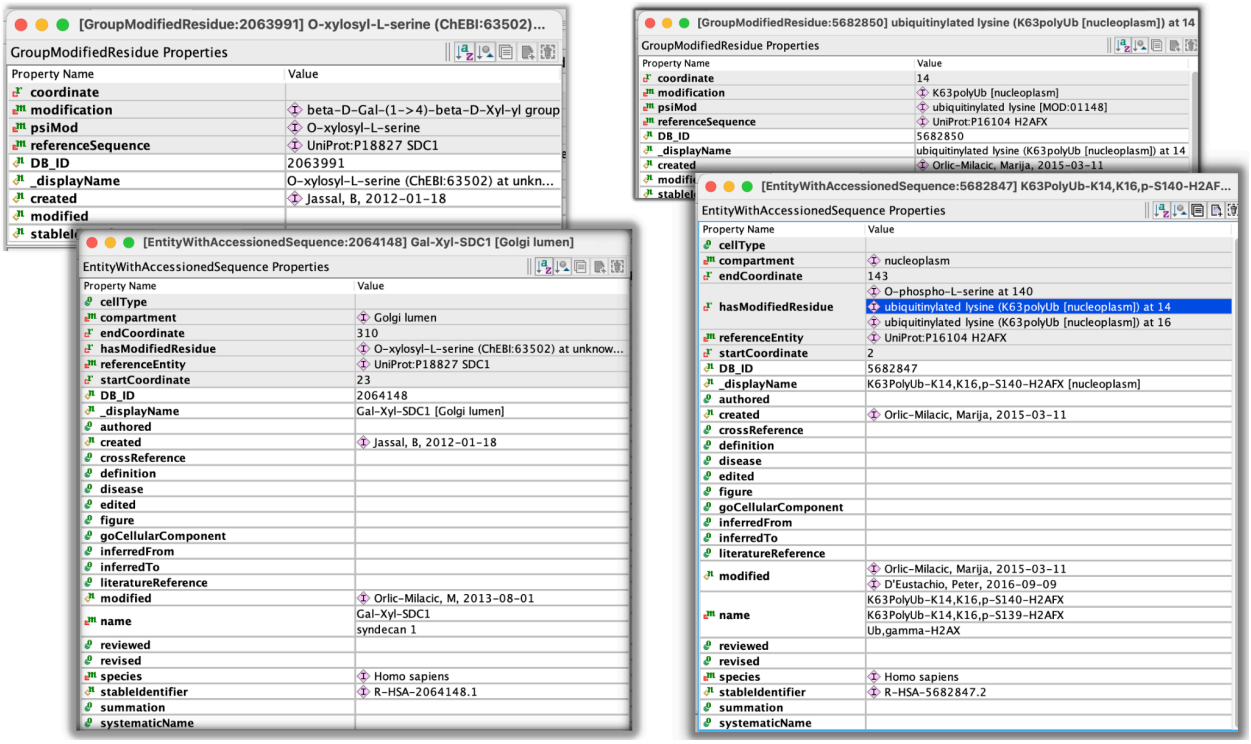


The modified EWAS is shown below. Make sure that the EWAS name is edited to reflect the modification, e.g., p-S124-CDC25A.



- *GroupModifiedResidue (GMR)*

This instance is linked to EWAS instance to indicate that a protein has a modification formed by attachment of two (or more) either peptide or chemical moieties. 3 mandatory slots need filling in a GMR instance - modification, psiMod and referenceSequence. For example, GMR is extensively used to annotate EWASs in protein glycosylation process that involves stepwise addition of sugar moieties, attached to an amino acid residue of a protein. In the synthesis of several GAGs, the 4 sugar moieties, i.e., xylose (Xyl), 2 x galactose (Gal) and glucuronic acid (GlcA), are added in a stepwise fashion. The first reaction is the addition of Xyl to a Ser residue. The product, a protein with a Xyl-modified Ser residue, is represented by EWAS with a modifiedResidue instance where psiMod slot value is O-xylosyl-Ser [MOD:00814]. The second reaction requires Gal to be attached to O-xylosyl-Ser to form the protein product with Gal-Xyl-yl group attached to Ser. Right click in the hasModifiedResidue slot of the EWAS instance and select GroupModifiedResidue class. In the modification instance right click in the modification slot and add 'beta-D-Gal-(1->4)-beta-D-Xyl-yl group' [ChEBI:63502) as a referenceGroup. The psiMod slot still needs to be filled with O-xylosyl-Ser [MOD:00814]. Enter the Uniprot identifier for the ReferenceGeneproduct in the ReferenceSequence slot. GMR may also be linked to an EWAS instance to indicate that a protein is modified by attaching polypeptide chains such as polyUb or polySUMO. Polypeptide chain instances are selected from the Polymer class and added in the modification slot in GMR.

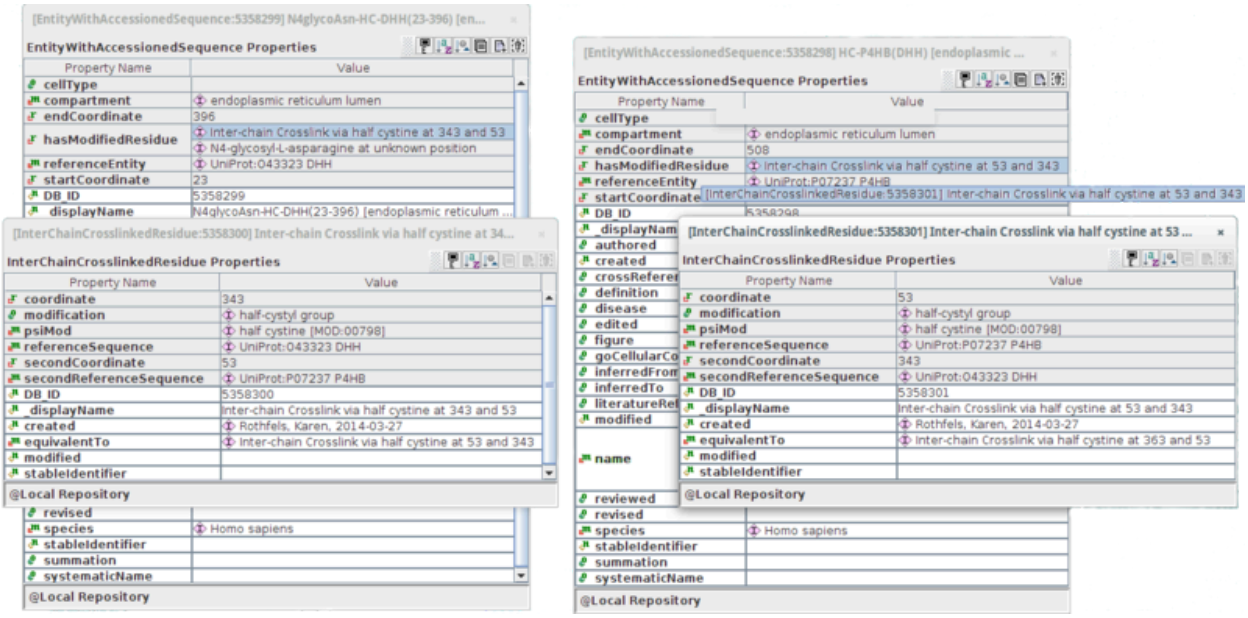


- *CrosslinkedResidue*

This instance is linked to a target EWAS instance to annotate disulfide, thioester, mono-sumoylated lysine, monoubiquitinated lysine or other protein crosslinks, where

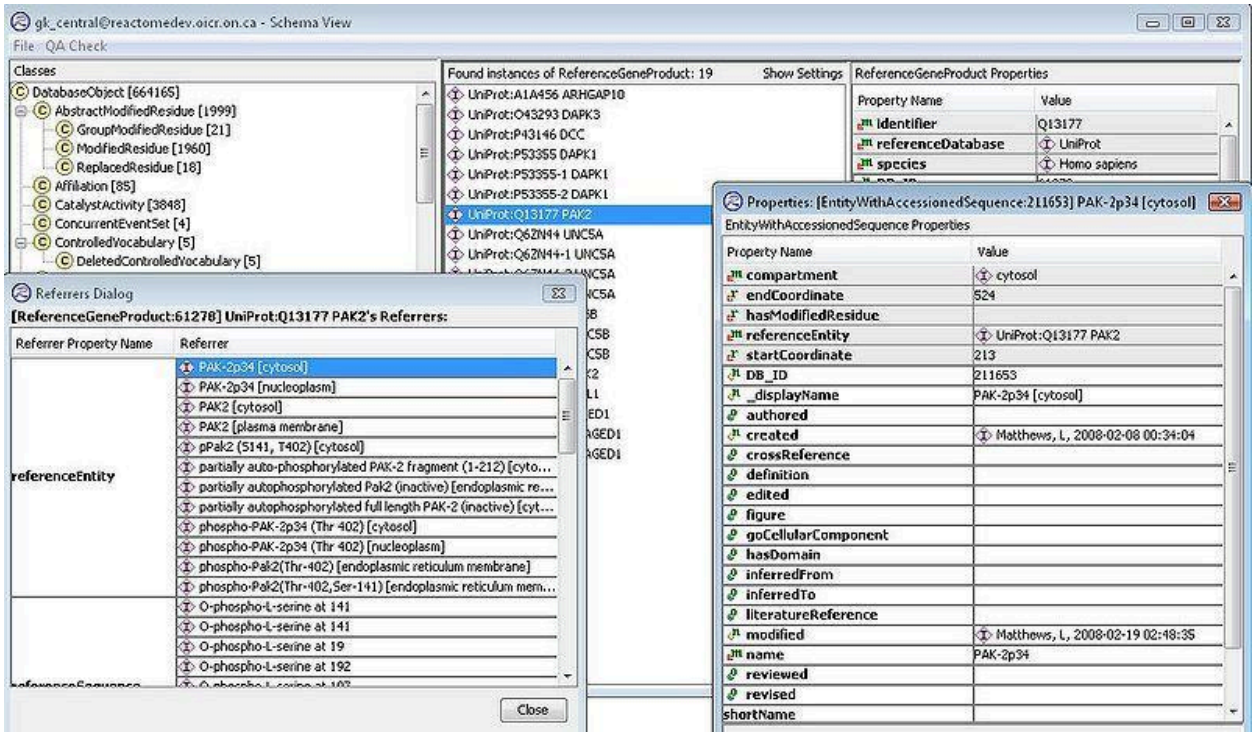
two coordinates describe a modification. The "crosslinkedResidues" subclass has two child classes - IntraChainCrosslinkedResidue and InterChainCrosslinkedResidue.

1. **IntraChainCrosslinkedResidue**
   is a value of hasModifiedResidues slots in EWAS records used to annotate crosslinking within the same amino acid sequence. The attributes and requirements are similar to those of ModifiedResidue, but it also includes a secondCoordinate slot.
2. **InterChainCrosslinkedResidue**
   To indicate that two different peptide chains are crosslinked you create an InterChainCrosslinkedResidue instance for each corresponding EWAS instance of these two proteins. The interchain modification instances are linked by having each other as values in their equivalentTo slots. There are 4 mandatory slots - psiMod, referenceSequence, secondReferenceSequence and equivalentTo. For example, to indicate that disulfide bond occurs between Cys343 of DHH and Cys53 of P4HB, two interChainCrosslinkedResidue instances are needed, one instance for modified DHH, the second instance for modified P4HB protein. Right click in the hasModifiedResidue slot of EWAS of modified DHH protein and select InterChainCrosslinkedResidue class. Then, right click in the referenceSequence slot of the modification instance and enter RGP with UniProt identifier 'UniProt:O43323 DHH'. A modified cysteine residue involved in the disulfide bond can be described as a half cystine (MOD:00798 , DB_ID 448182). Enter half cystine [MOD:00798] as a modification type in the psiMod slot. Enter half-cystyl group as a reference group value in the modification slot (optional). Enter 343 in the coordinate slot to state that DHH protein has a half cystyl group at position 343. Enter 53 in the secondCoordinate slot and RGP with identifier 'UniProt:P07237 P4HB' in the secondReferenceSequence slot to state the position on the second amino acid sequence. Repeat these steps for EWAS of modified P4HB but enter 53 as a coordinate value and select RGP with identifier 'UniProt:P07237 P4HB' in the referenceSequence slot, while secondCoordinate value is 343 and secondReferenceSequence value is RGP with the identifier 'UniProt:O43323 DHH'. Finally, right click in the equivalentTo slot of the first instance (linked to EWAS of modified DHH) and select the other one that is linked to EWAS of modified P4HB. Similarly, enter the first instance as a value in the equivalentTo slot of the second one.
3. In the event that you are creating a crosslink between two monomers of the same protein, create one InterChainCrosslinkedResidue instance and indicate that the referenceSequence and secondReferenceSequence are the same as are the coordinate and secondCooordinate.



Reactome does not aim to describe all disulfide bonds in all proteins. Curators may choose to annotate disulfide bonds at their discretion when, for example, the bonds or bonding process may be involved in a disease, may regulate the protein's activity, or may be required to hold a complex together.

If you are annotating a protein fragment, you can define the start and end coordinates of the fragment as shown below: Don't forget to change the default name to indicate that it is a fragment.



***To be sure that all PTMS get added correctly to the reactions in diagrams** you have to first make sure you have downloaded TranslationalModifications fully in your project. In the schema view, choose that class to get the whole list of these instances. Select all in the middle list and then select “Update Instances” from the popup menu. Then go to the diagram, right click, and select the option “update node features”. Say yes to the question “instances must be fully downloaded, do you want to continue, yes no”.

Creating a Complex

This is done in one of two ways. Either:

Go to the Schema view, select Complex, right-click and select Create Instance. The Create a New Instance dialog box appears. Enter a name in the field for Name. Typically you would also enter the *Compartment, Species and identify the entities (EWASs, sets or complexes) that make up this complex using the field hasComponent. All of these fields are completed by either double-clicking to type, or right clicking to select Add and identify the correct item from the local project.

- *We record two location attributes for every complex. The "core compartment" (the place most closely associated with the function of the complex) will be entered into the compartment slot. This slot will be made single-valued and will be filled manually. All of the "other" compartments of the participating components will be automatically added in the "includedLocation" attribute of the complex at the time of database slicing for release.

Example: Complexes that have component(s) localized to membrane as well as component(s) localized to other adjacent compartments should have the membrane listed in the "Compartment" slot . The other compartments would be assigned automatically.

Creating a Set

Reactome has several types of set - refer to the Glossary and User Guide for definitions.

The most commonly used sets are Defined Sets and Candidate Sets.

Defined Set members should be proven equivalents, i.e. all of them have been demonstrated to perform the function that is described by the event they participate in.

Candidate Sets have two categories of inclusion, members, equivalent to defined set members, and candidates, members that are not proven to be functionally equivalent, but are believed to be equivalent based on phylogeny, domain structure etc.

Creating sets is a similar process for all subtypes, select the appropriate type in the Schema view, right-click and select Create Instance, fill in the Name and Species, right click to add the set members.

- All members of a set must have the same compartment. The only time a set can have multiple compartment attributes is if its members themselves all have the same multiple compartment attributes, e.g., a set of membrane-spanning complexes with components explicitly located [on this side], in the membrane, and [on that side].
- The members or candidates of a set should (in most cases) belong to the same class...e.g all EWASs or all simpleEntities or all Complexes. However, there are exceptions under circumstances in which structurally different entities are really functionally indistinguishable (i.e either a monomer or dimer can function in a particular reaction). Since these exceptions are likely to be fairly rare, a QA check is in place to flag sets that have members that are from different classes. The ids of such exceptional acceptable sets may be forwarded to the editorial release manager to be put on a QA skip list.

Using sets

Representing outputs that include sets: When representing the output of a reaction: Set [A, A', A"] + EWAS B
Use the 'set-to –complex' approach: Set [A, A', A"] + EWAS B → Complex (Set [A, A', A"]:EWAS B), where no individual complex entities are created for the reaction. Entity of the individual complex (for example A:B) will be created if we need to functionally distinguish the potential output complex somewhere else in the diagram. The same dashed lines used to link an EntitySet and its contained members will be used to link the individual complex A:B to Complex (Set [A, A', A"]:EWAS B).

Creating a Pathway

Please see *note* below if you are adding a pathway that will be a new top level pathway. Here is a description of how the outlined mini pathway "Regulation of activated PAK-2p34 by proteasome mediated degradation" is built from its component events in the curator tool. The pathway consists of 2 reactions: "Ubiquitination of PAK-2p34" and "Proteasome mediated degradation of PAK-2p34". For simplicity, the reactions have already been created (see section on creating inferred reactions for an example.)

II. Regulation of activated PAK-2p34 by proteasome mediated degradation PMID: 12853446

Summary

Stimulation of cell death by PAK-2 requires the generation and stabilization of the caspase-activated form, PAK-2p34 (Walter et al., 1998 PMID: 9786869 ,Jakobi et al., 2003 PMID: 12853446). Levels of proteolytically activated PAK-2p34 protein are controlled by ubiquitin-mediated proteolysis. PAK-2p34 but not full-length PAK-2 is degraded by the 26 S proteasome (Jakobi et al., 2003). It is not known whether ubiquitination and degradation of PAK-2p34 occurs in the cytoplasm or in the nucleus.

IIa Ubiquitination of PAK-2p34 [cytosol]

phospho-PAK-2p34 (Thr 402) [cytosol] + Ubiquitin [cytosol] => phospho-PAK-2p34:Ubiquitin [cytosol] complex

Catalyst: unknown ubiquitin ligase

GO:0004842 Ubiquitin protein ligase activity

Inferred: Oryctolagus cuniculus, Homo sapiens PMID: 12853446

phospho-PAK-2p34 (Thr 402)(Oryctolagus cuniculus) [cytosol] + Ubiquitin (human) [cytosol] => phospho-PAK-2p34 (Oryctolagus cuniculus):Ubiquitin (human) [cytosol] complex

Summary

PAK-2p34 is ubiquitinated prior to degradation (Jakobi et al., 2003 PMID: 12853446). Here, ubiquitination of PAK-2p34 is described as occurring in the cytosol. However, to date it is not known whether this occurs in the nucleus or in the cytoplasm. Evidence for this reaction comes from experiments using both human and rabbit proteins.

IIb Proteasome mediated degradation of PAK-2p34 [cytosol]

PAK-2p34:ubiquitin [cytosol] =>ubiquitin [cytosol]

catalyst =ubiquitin ligase [cytosol]

GO:0004175 Endopeptidase activity

Inferred: Oryctolagus cuniculus, Homo sapiens PMID: 12853446

PAK-2p34:ubiquitin (human,Oryctolagus cuniculus) [cytosol] =>ubiquitin (human) [cytosol]

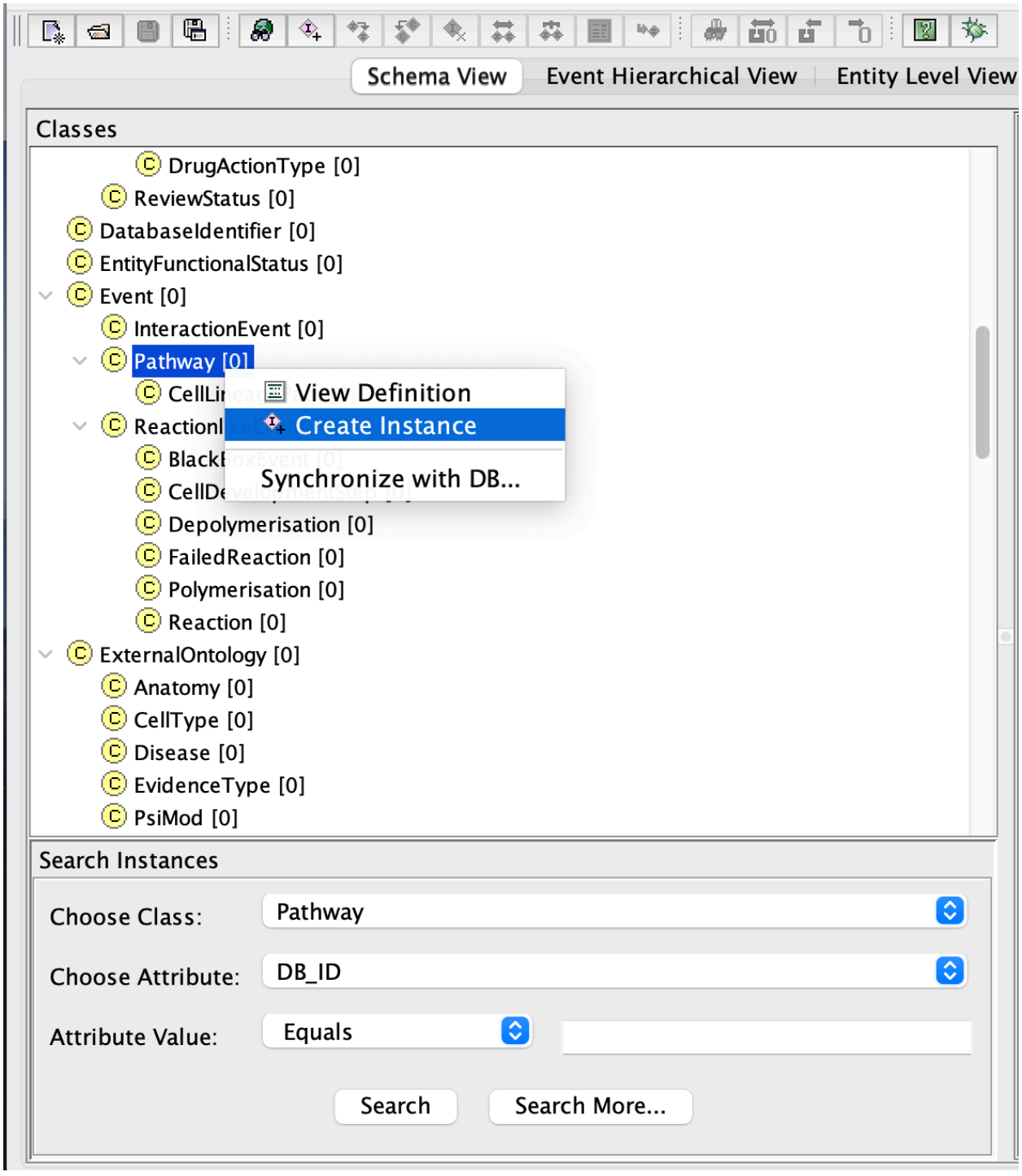
catalyst =ubiquitin ligase [cytosol]

GO:0004175 Endopeptidase activity

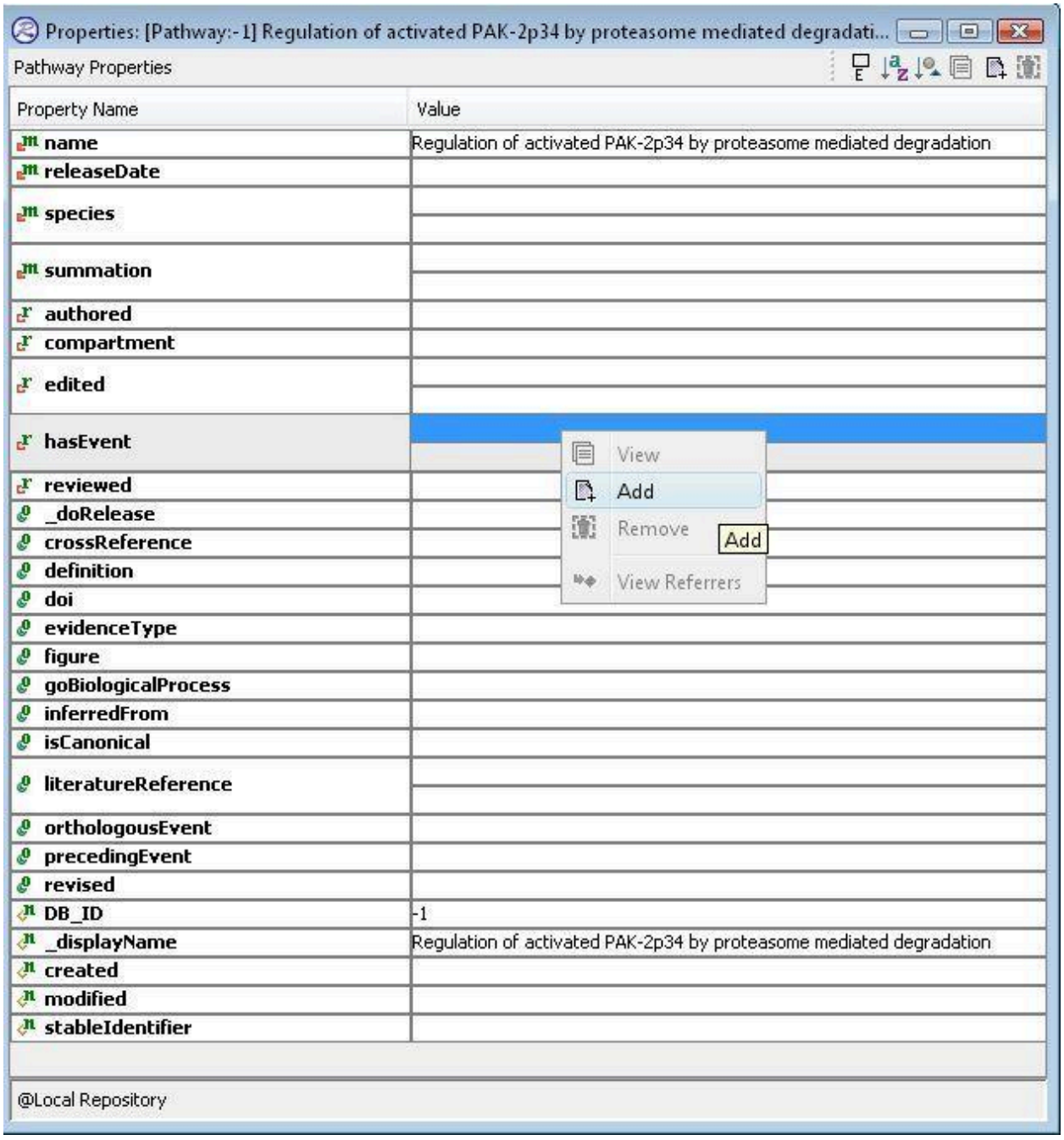
Summary

Proteolytically activated PAK-2p34, but not full-length PAK-2, is degraded rapidly by the proteasome (Jakobi et al., 2003 PMID: 12853446). Here, degradation of PAK-2p34 is described as occurring in the cytosol. However, to date it is not known whether this occurs in the nucleus or in the cytoplasm.

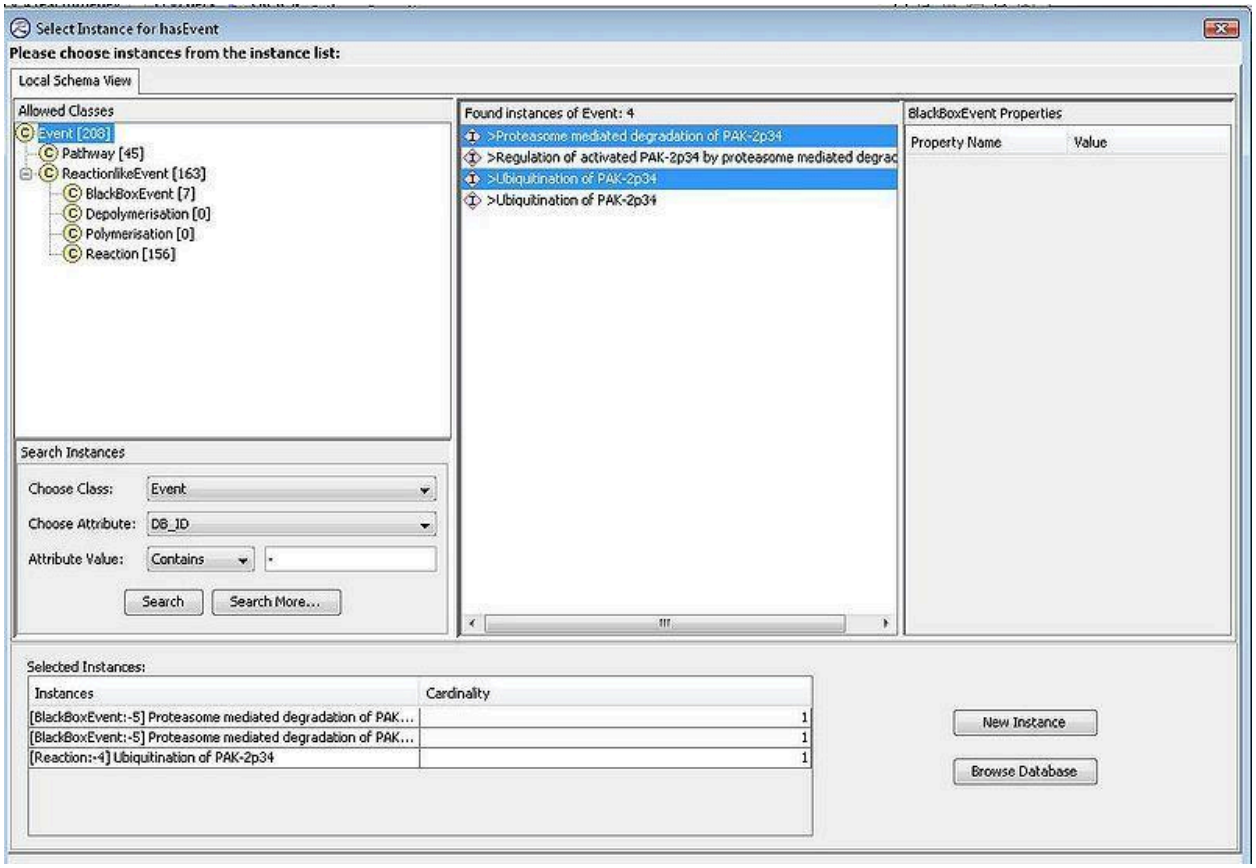
With the pathway class highlighted in the class hierarchy, right click and select Create instance, or use the create instance button in the menu bar at the top of the tool.



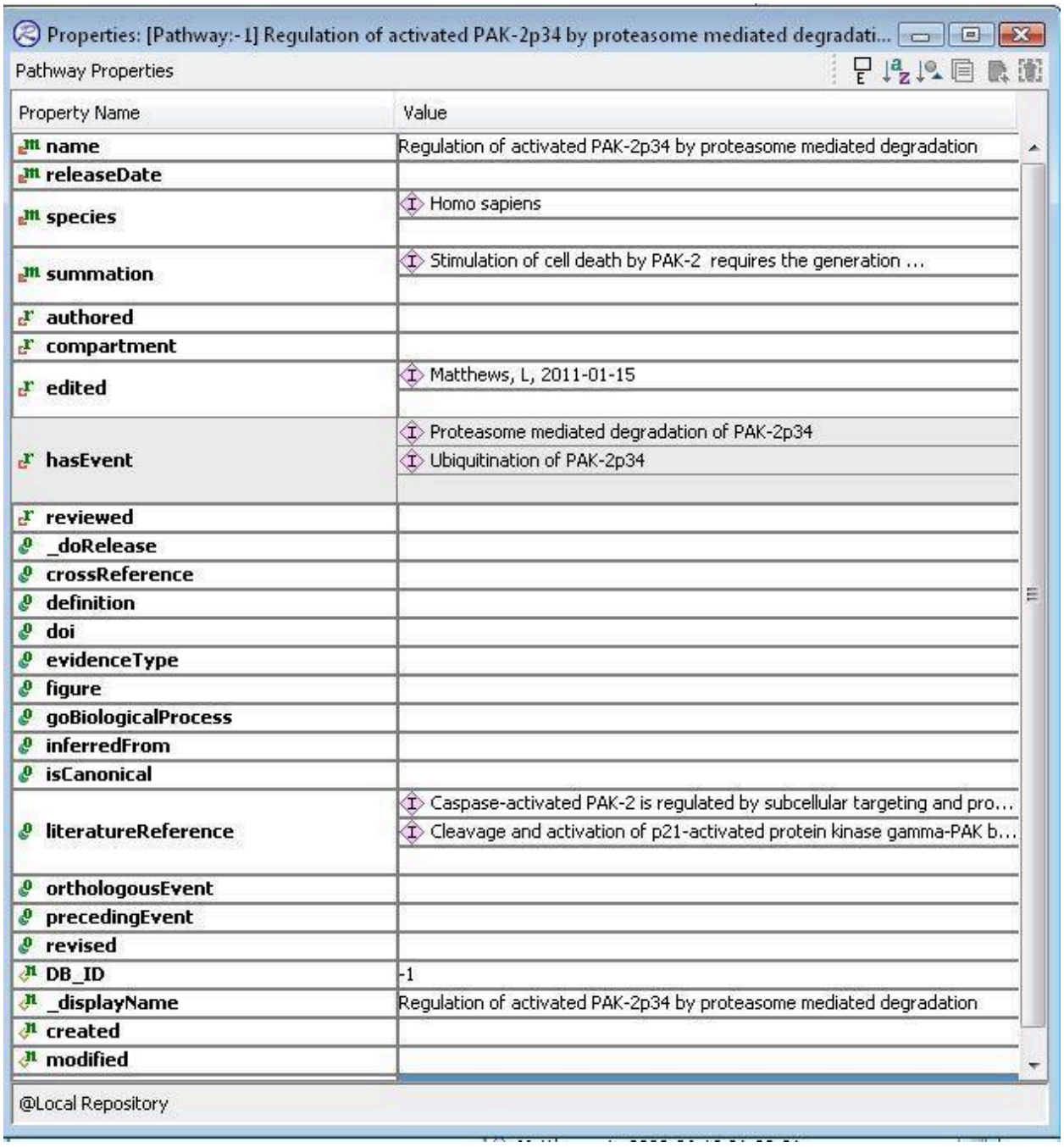
After adding the pathway title, associate reactions as component events of the pathway by right clicking on the hasEvent slot and selecting "Add"



You can then select the events that you need one at a time or, as a shortcut, you can search for your newly created events in your project if they have not yet been submitted to gk_central (they will all have DB_ID attributes that are negative). To do this, search in your project for events with DB_ID containing - . From this list, you can hold the control key and select the events of interest.



The order of the events in a pathway is described through the use of the "precedingEvent" attribute on the reactions that are components of the pathway. If/when no preceding events are specified, the order of the events displayed on the webpage reflects the order in which they are listed as components in the pathway instance that you are creating. Curators should check for precedingEvents that they may have missed within a pathway diagram by using the “Infer reaction relationship” curator tool feature (described in more detail here, Appendix C7). This feature can be accessed by right-clicking in a diagram. Here, the curator is offered a list of preceding events based on shared entities. The curator can then opt to approve a preceding event, in which case that reaction is automatically entered into the appropriate reaction’s precedingEvent attribute. If the curator decides the precedingEvent is not biologically relevant, it can be rejected with a reason for the rejection specified. This “rejection” of a preceding event is recorded in the “negativePrecedingEvent” attribute for the reaction.



Once the component events have been added, the remaining required attributes are added. Identify a GO Biological process (BP) term that corresponds to your pathway, find the corresponding GO_BiologicalProcess instance in gk_central, right click on the goBiologicalProcess slot and select set to add it. If you can't find the term of interest, double check with Peter or Lisa. If no term is available, you can use GOC GitHub Tracker to create a ticket to request a new term as described here. When suggesting a GO term, be sure to look over the guidelines described here (https://github.com/geneontology/go-ontology/blob/master/CONTRIBUTING.md) and

include label (name), definition, references, and position in hierarchy. ***Never*** create a new GO term instance yourself in gk_central!

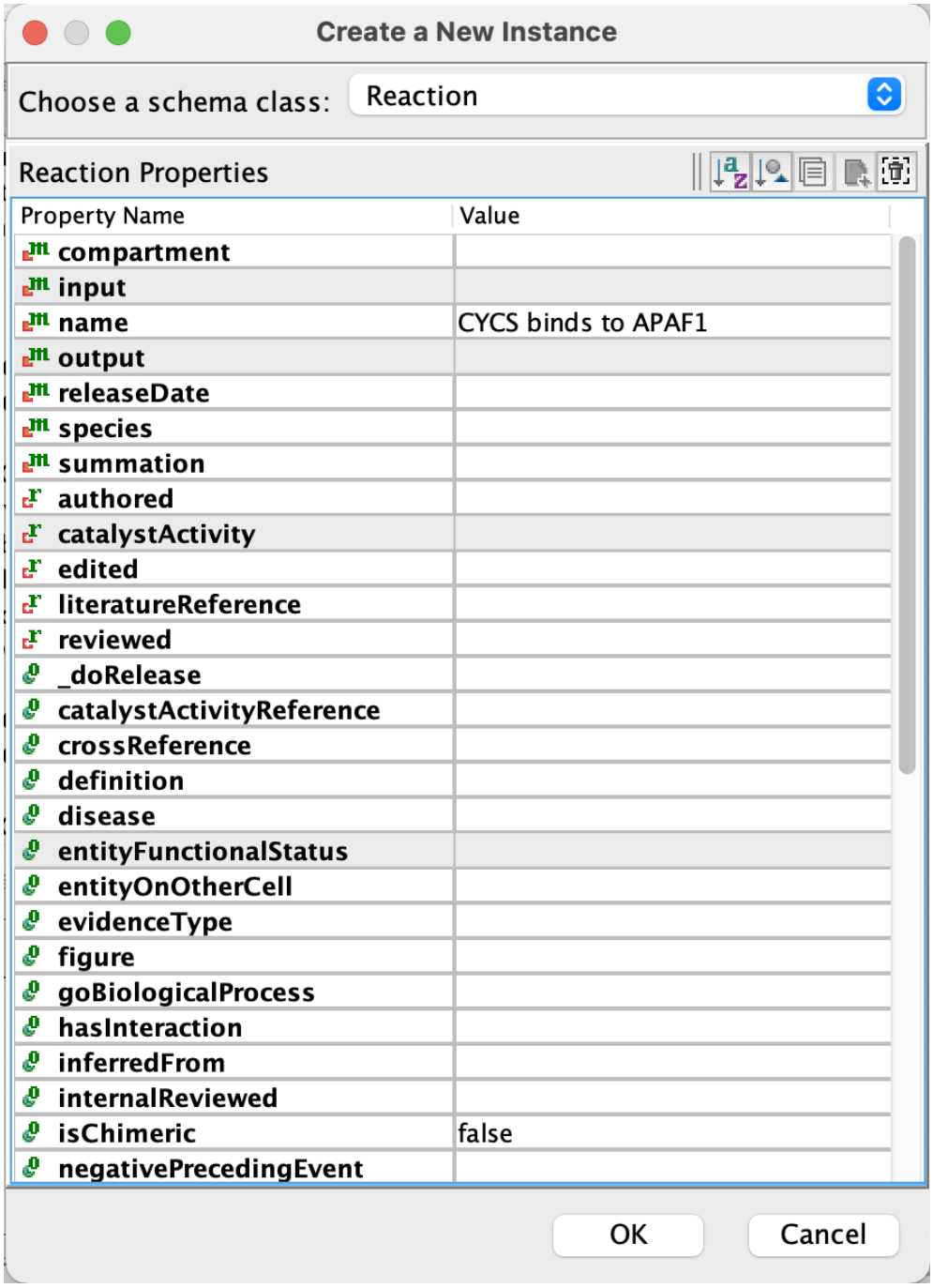
- note: If you are creating a top level pathway(check with Peter if this is appropriate), it must be listed as frontPage item. To do this, check out the frontPage instance from gk_central and add your pathways in the frontPageItem slot. Also please mark this in the editorial calendar as a front page item.

Creating a ReactionLikeEvent

There are subclasses of ReactionlikeEvent: Reaction, BlackBox, Polymerization, Depolymerization, FailedReaction, and CellDevelopmentStep. The latter two will be covered in the Disease and the Cell Lineage Path annotation sections of this guide.

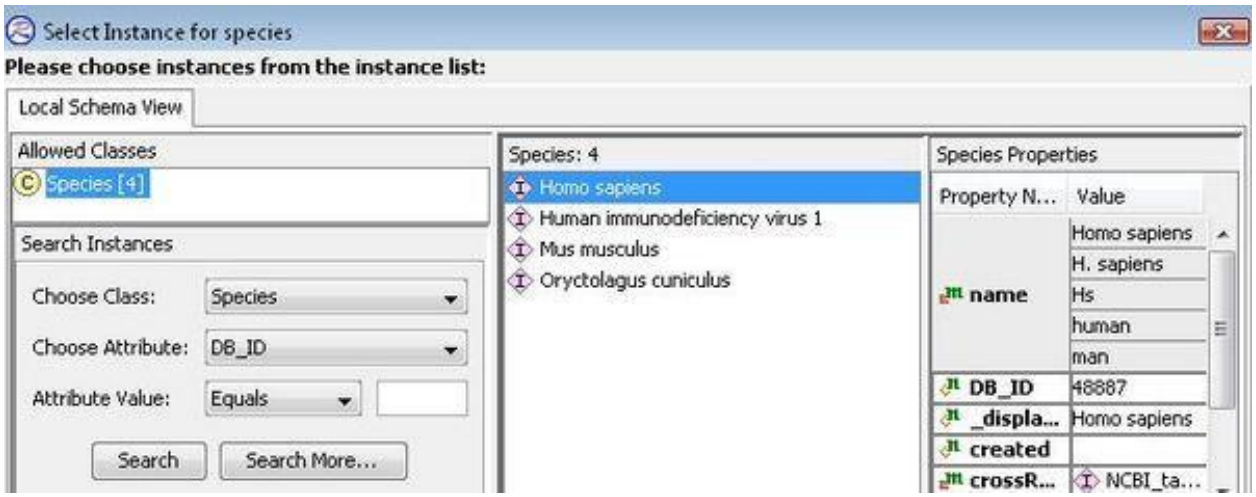
Like all other classes, you can create a new reaction by selecting the Reaction class in the Schema view, right-click and select Create Instance, but it is perhaps better practice and more intuitive to create new reactions inside a pathway. To do this, select a pathway in the Event Hierarchical View, select the hasEvent property name in the details panel on the right, right-click and select Add. This leads to a dialogue for providing details of the reaction.

When creating a reaction, first enter the name:

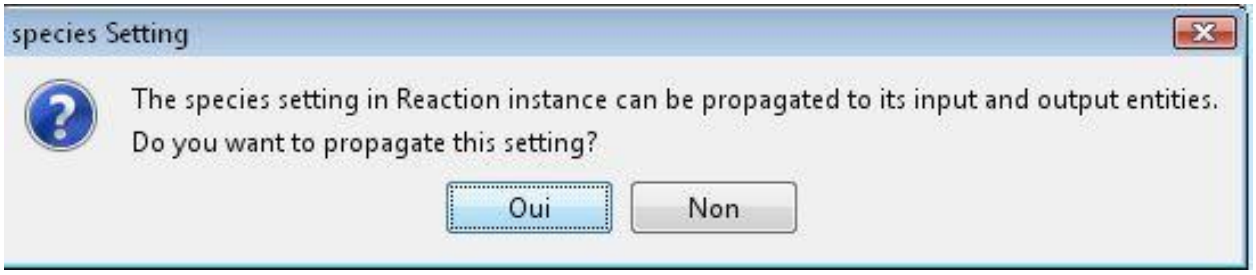


This may be enough detail for the moment, if you are simply creating a placeholder click OK.

To set the species of the reaction right click and select add in the species field.

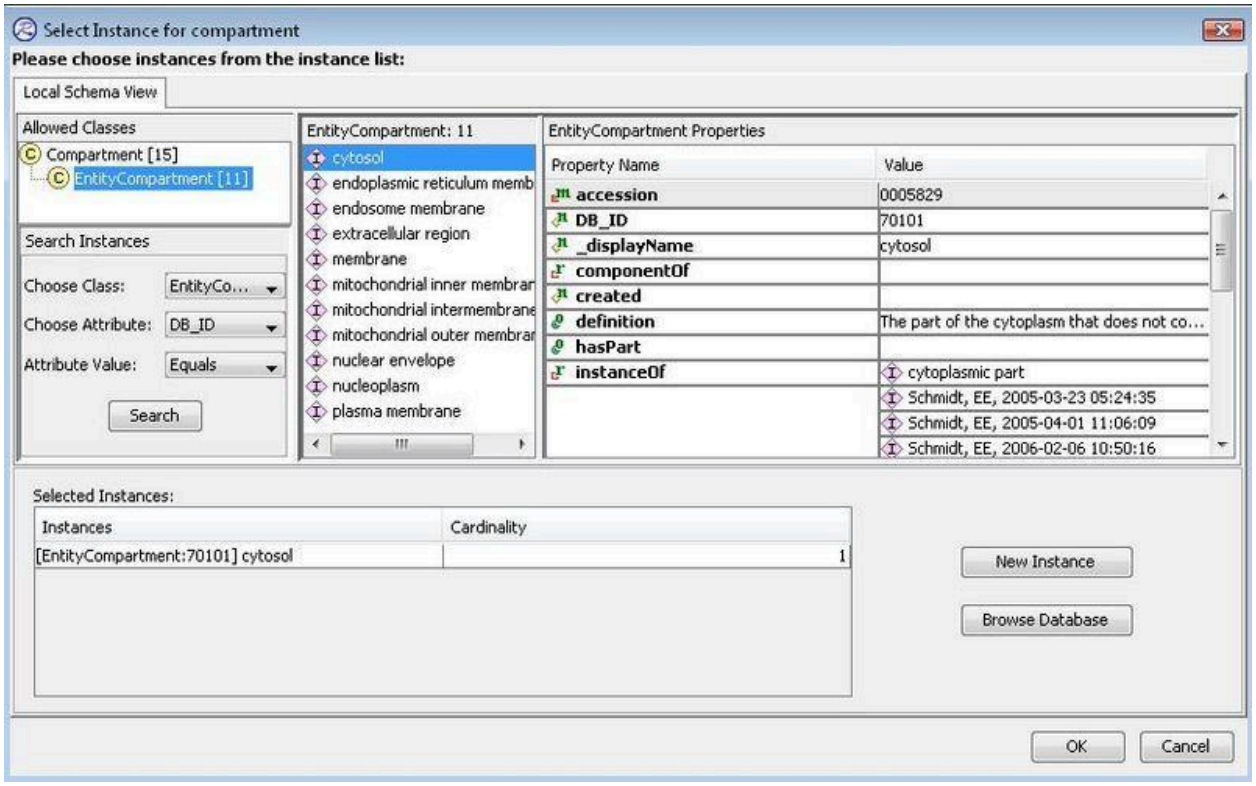


If the species that you happen to be working with is not in your local project you can opt to search for it in the database using the "Browse Database" Button in the dialog box. When you set/change the species or the compartment of a reaction, it will ask if you want to propagate the species/compartment to all of the component molecules as well.



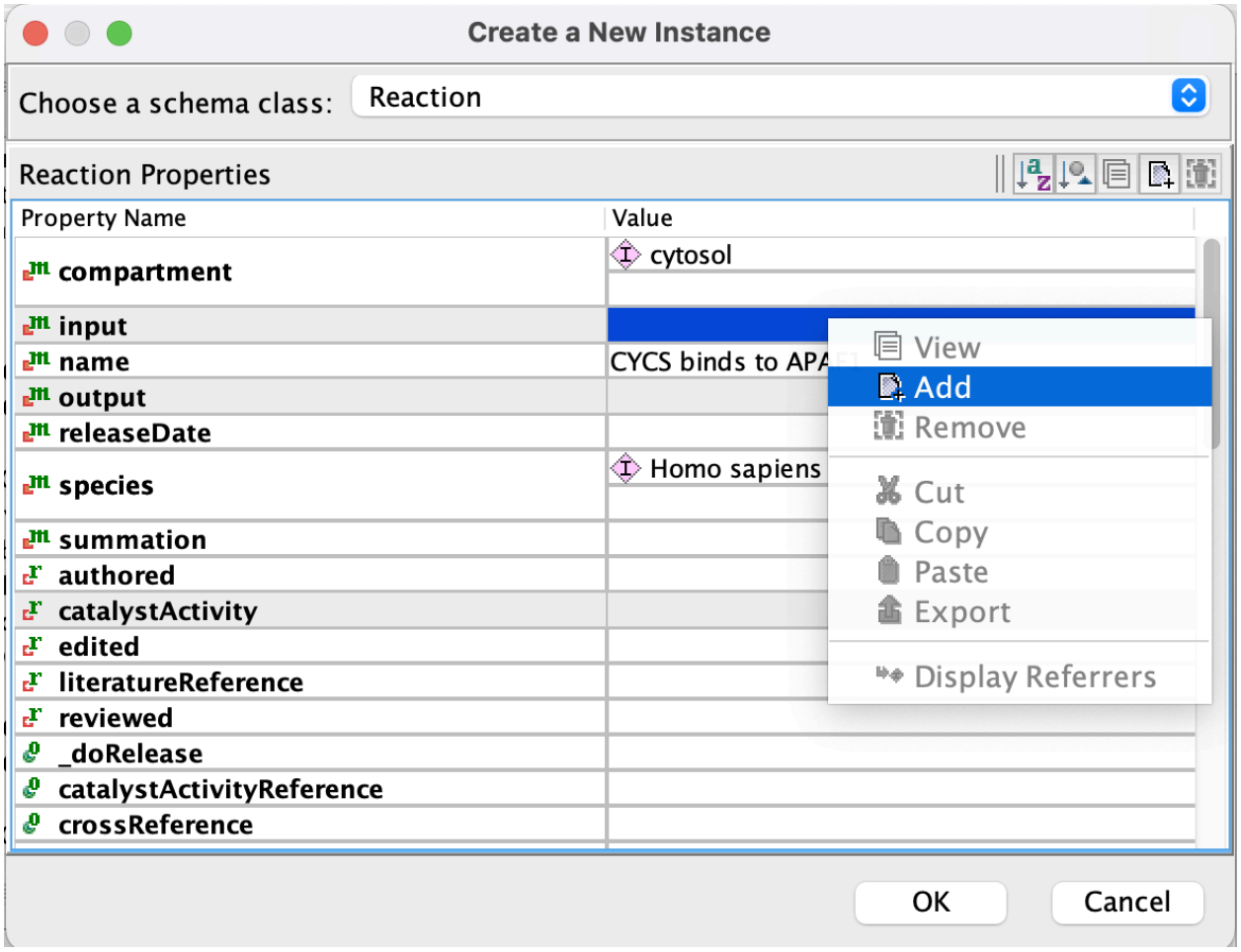
Use caution when selecting Yes. Generally, this is ONLY done when you have just created (but not committed) all of the entities used in the reaction. **ONLY** say yes here if you know that event and contained molecules have no referrers in the database that would be affected. (In other words ALL other reactions or complexes or sets in the db that make use of these now "changed" molecules would be affected).

Next add the compartment to the reaction. Right click in the compartment box and select add.

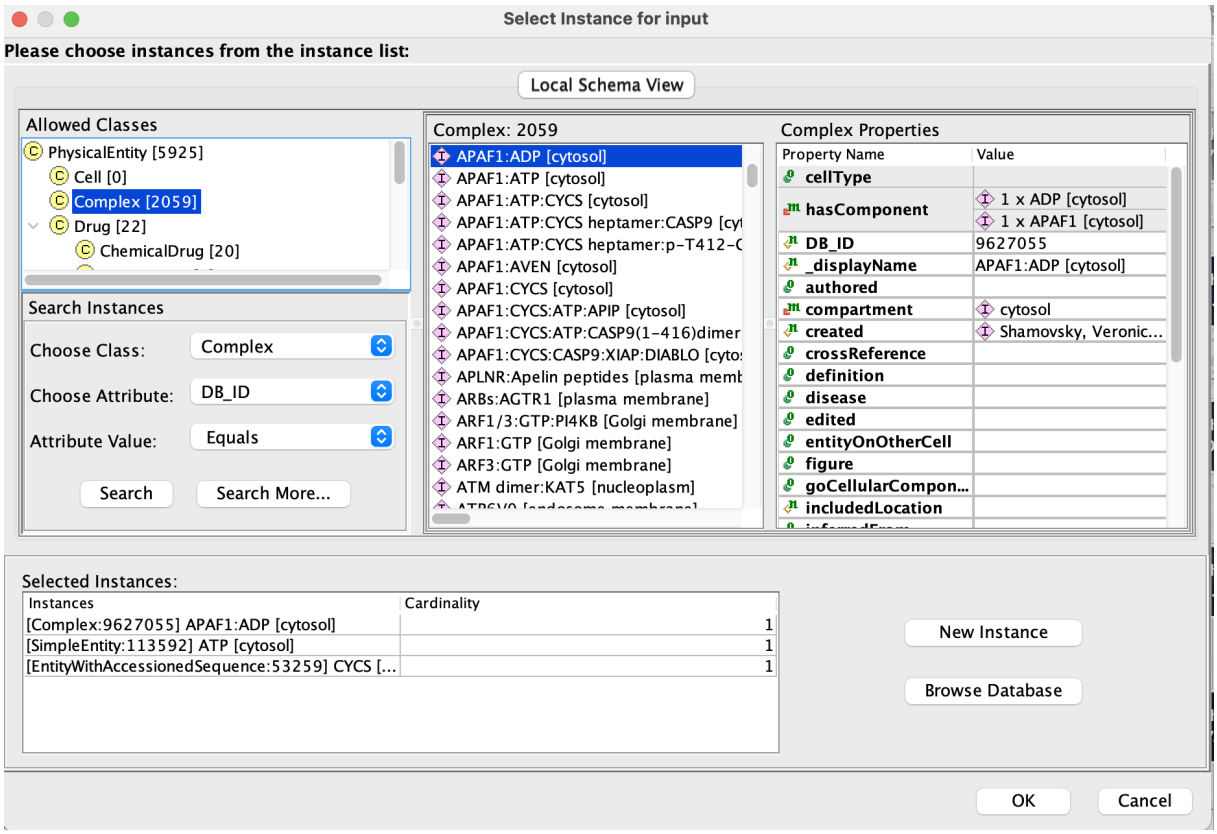


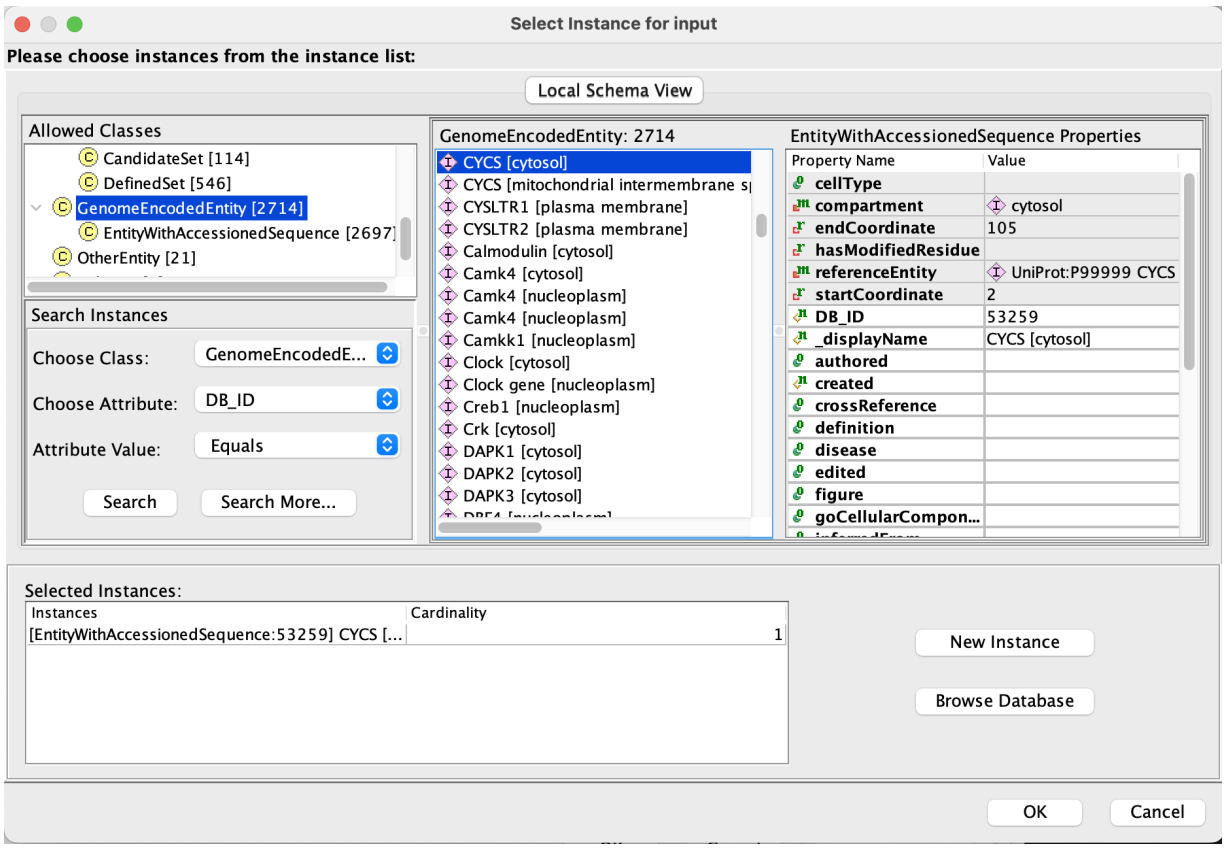
Again if you don't have the compartment you need in your project, you can Browse Database to find the one you need. ***If you still don't find the necessary compartment, DO NOT create a new one. Contact the managing editor. New compartment creation requires multiple steps as well as changes in the UI which need to be coordinated with developers.***

Now add the input and output molecules. Right click in the respective box and select "Add".



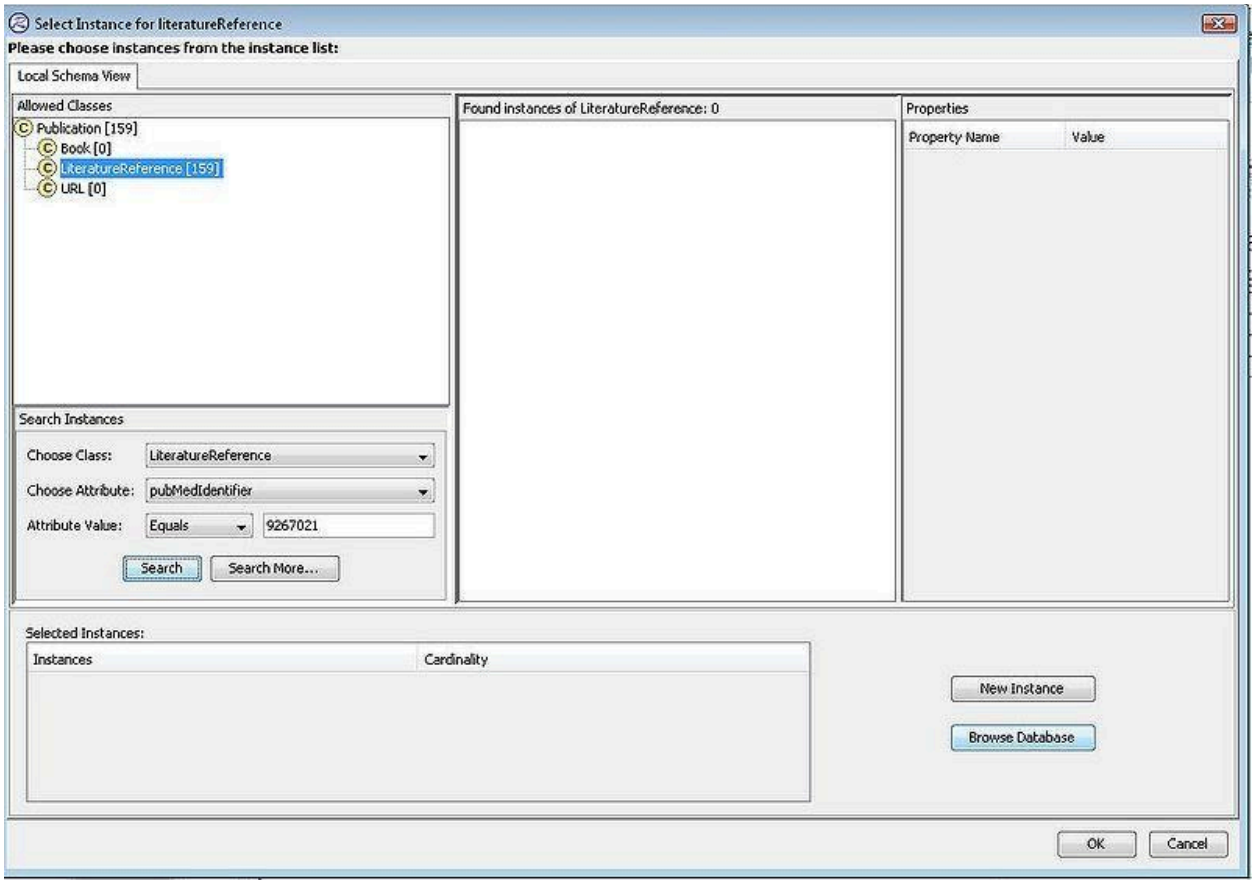
Select the molecule from the appropriate class:

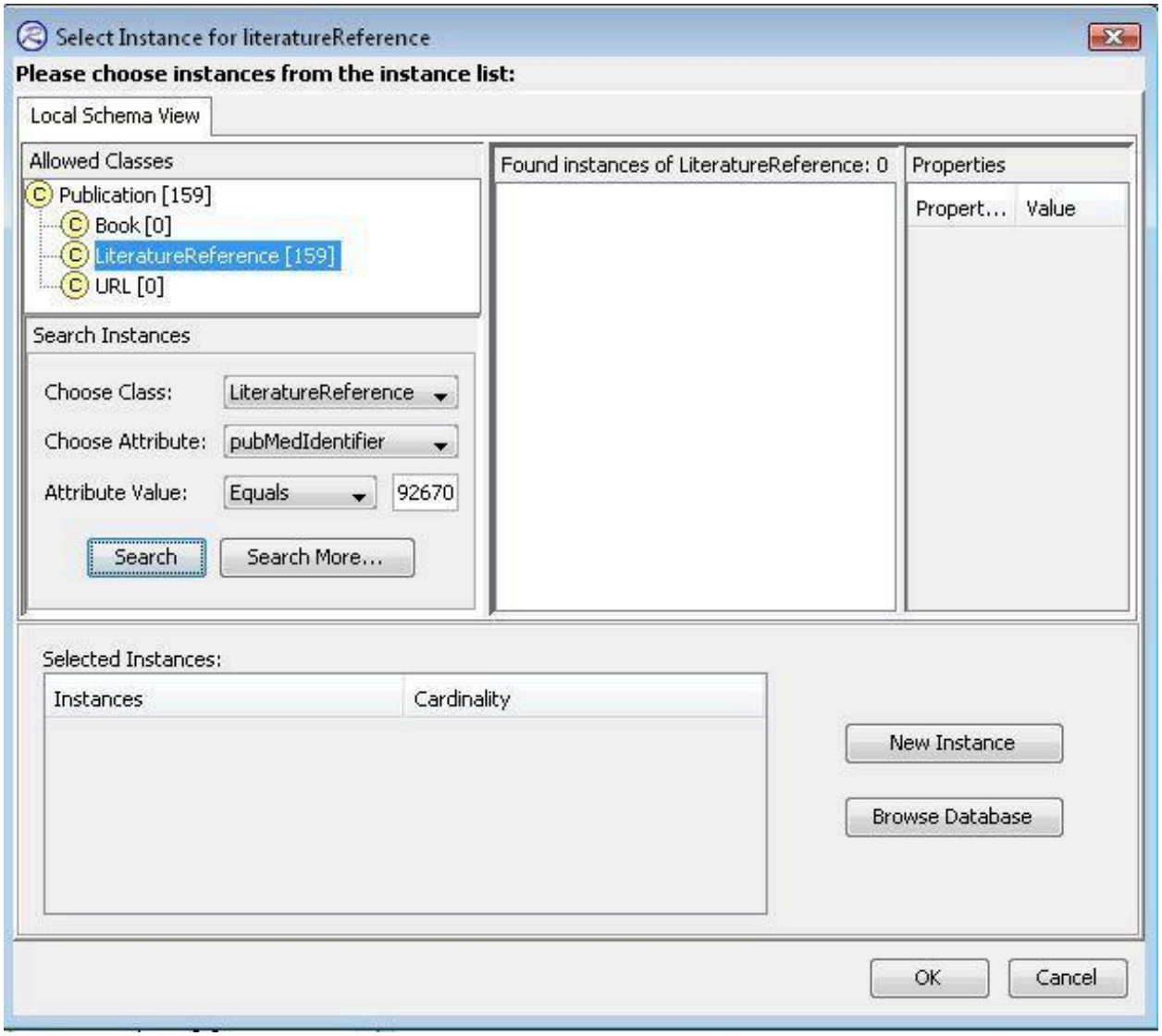

Select Instance for input
Please choose instances from the instance list:
Local Schema View
Allowed Classes
PhysicalEntity [5925]
Cell [0]
Complex [2059]
Drug [22]
ChemicalDrug [20]
Search Instances
Choose Class:
Complex
Choose Attribute:
DB_ID
Attribute Value:
Equals
Search
Search More...
Complex: 2059
APAF1:ADP [cytosol]
APAF1:ATP [cytosol]
APAF1:ATP:CYCS [cytosol]
APAF1:ATP:CYCS heptamer:CASP9 [cyt
APAF1:ATP:CYCS heptamer:p-T412-C
APAF1:AVEN [cytosol]
APAF1:CYCS [cytosol]
APAF1:CYCS:ATP:APIP [cytosol]
APAF1:CYCS:ATP:CASP9(1-416)dimer
APAF1:CYCS:CASP9:XIAP:DIABLO [cytos
APLNR:Apelin peptides [plasma memb
ARBs:AGTR1 [plasma membrane]
ARF1/3:GTP:PI4KB [Golgi membrane]
ARF1:GTP [Golgi membrane]
ARF3:GTP [Golgi membrane]
ATM dimer:KAT5 [nucleoplasm]
Complex Properties
Property Name
Value
cellType
hasComponent
1 x ADP [cytosol]
1 x APAF1 [cytosol]
DB_ID
9627055
_displayName
APAF1:ADP [cytosol]
authored
compartment
cytosol
created
Shamovsky, Veronic...
crossReference
definition
disease
edited
entityOnOtherCell
figure
goCellularCompon...
includedLocation
Selected Instances:
Instances
Cardinality
[Complex:9627055] APAF1:ADP [cytosol]
1
[SimpleEntity:113592] ATP [cytosol]
1
[EntityWithAccessionedSequence:53259] CYCS [...
1
New Instance
Browse Database
OK
Cancel


Select Instance for input
Please choose instances from the instance list:
Local Schema View
Allowed Classes
CandidateSet [114]
DefinedSet [546]
GenomeEncodedEntity [2714]
EntityWithAccessionedSequence [2697]
OtherEntity [21]
Search Instances
Choose Class:
GenomeEncodedE...
Choose Attribute:
DB_ID
Attribute Value:
Equals
Search
Search More...
GenomeEncodedEntity: 2714
CYCS [cytosol]
CYCS [mitochondrial intermembrane sp
CYSLTR1 [plasma membrane]
CYSLTR2 [plasma membrane]
Calmodulin [cytosol]
Camk4 [cytosol]
Camk4 [nucleoplasm]
Camk4 [nucleoplasm]
Camkk1 [nucleoplasm]
Clock [cytosol]
Clock gene [nucleoplasm]
Creb1 [nucleoplasm]
Crk [cytosol]
DAPK1 [cytosol]
DAPK2 [cytosol]
DAPK3 [cytosol]
EntityWithAccessionedSequence Properties
Property Name
Value
cellType
compartment
cytosol
endCoordinate
105
hasModifiedResidue
referenceEntity
UniProt:P99999 CYCS
startCoordinate
2
DB_ID
53259
_displayName
CYCS [cytosol]
authored
created
crossReference
definition
disease
edited
figure
goCellularCompon...
Selected Instances:
Instances
Cardinality
[EntityWithAccessionedSequence:53259] CYCS [...
1
New Instance
Browse Database
OK
Cancel

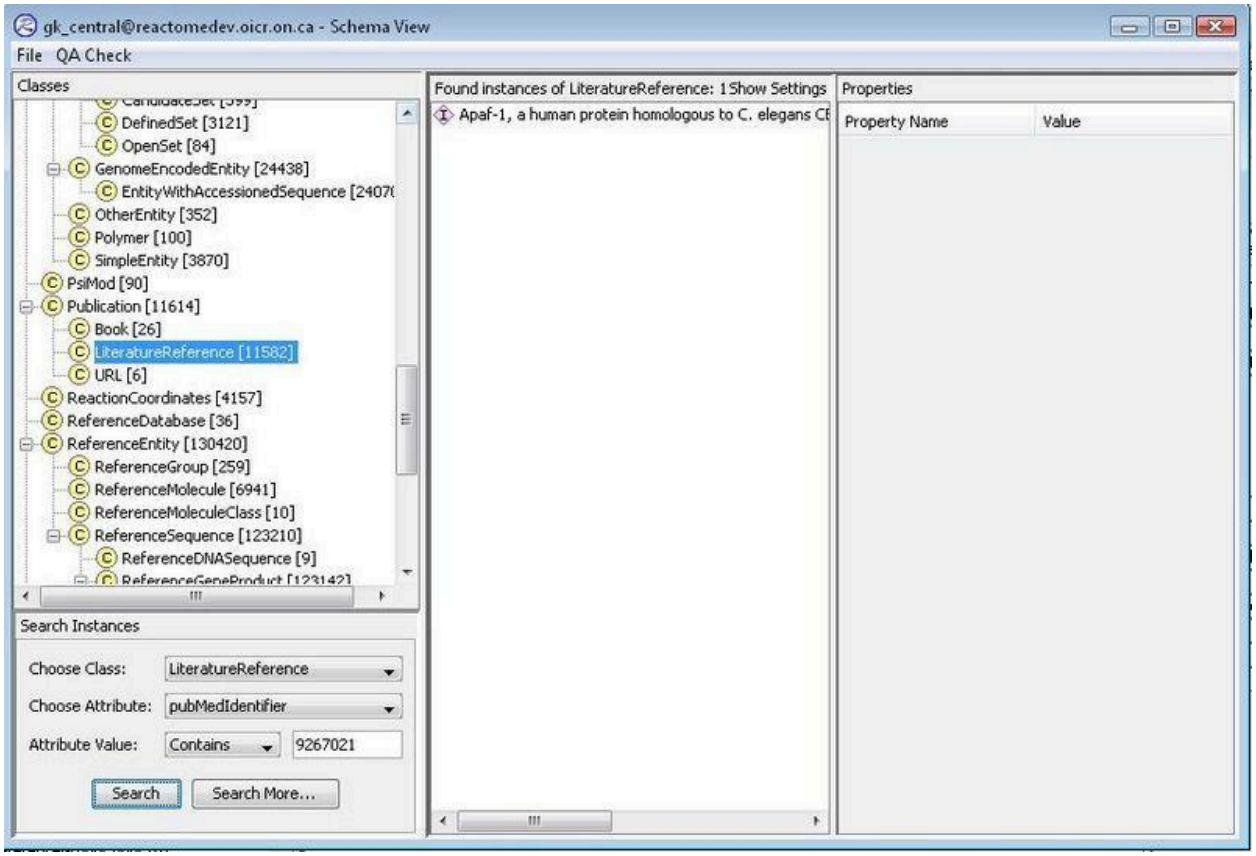
Important: After adding your input and output molecules, it is important to verify that your reaction is balanced (all molecules represented as input are also present as output). See the QA section below for a description of how to do this.

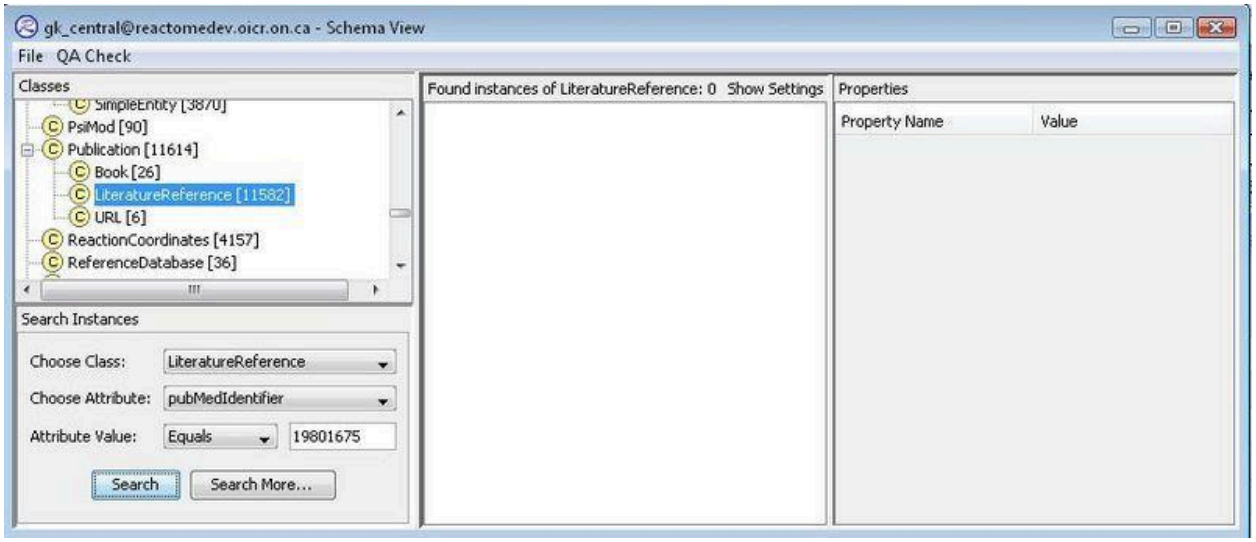
Add the literature reference(s). The references associated with a reaction MUST provide direct experimental evidence for the occurrence of that reaction in the species you are annotating (i.e. human for human Reactome). If there is no direct experimental evidence in human then you need to create an inferred human reaction as described in the section below. Enter the PMID for journal articles, and say yes when prompted to have the details filled in automatically.





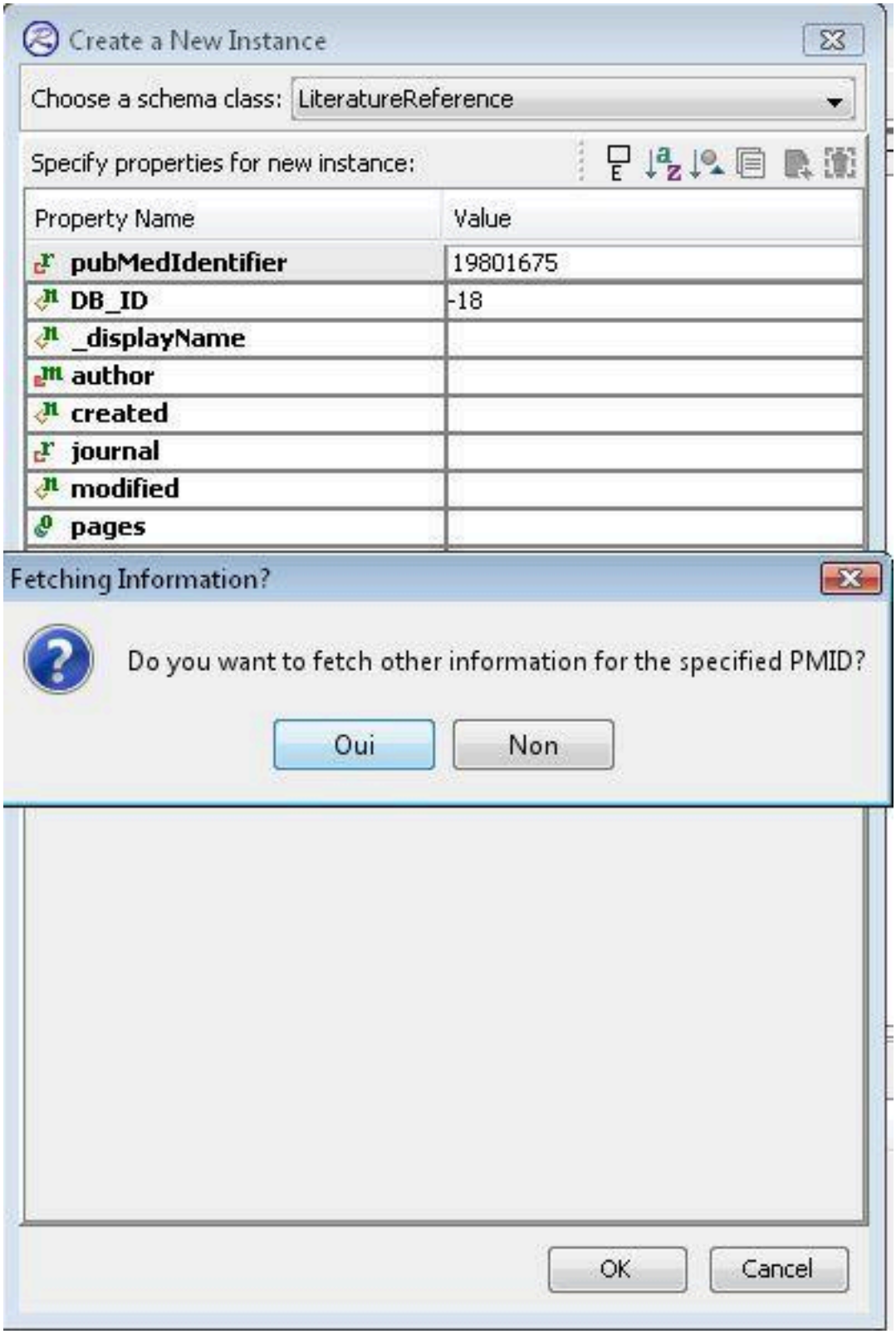
If you don't see the reference in your local repository then opt to Browse Database.



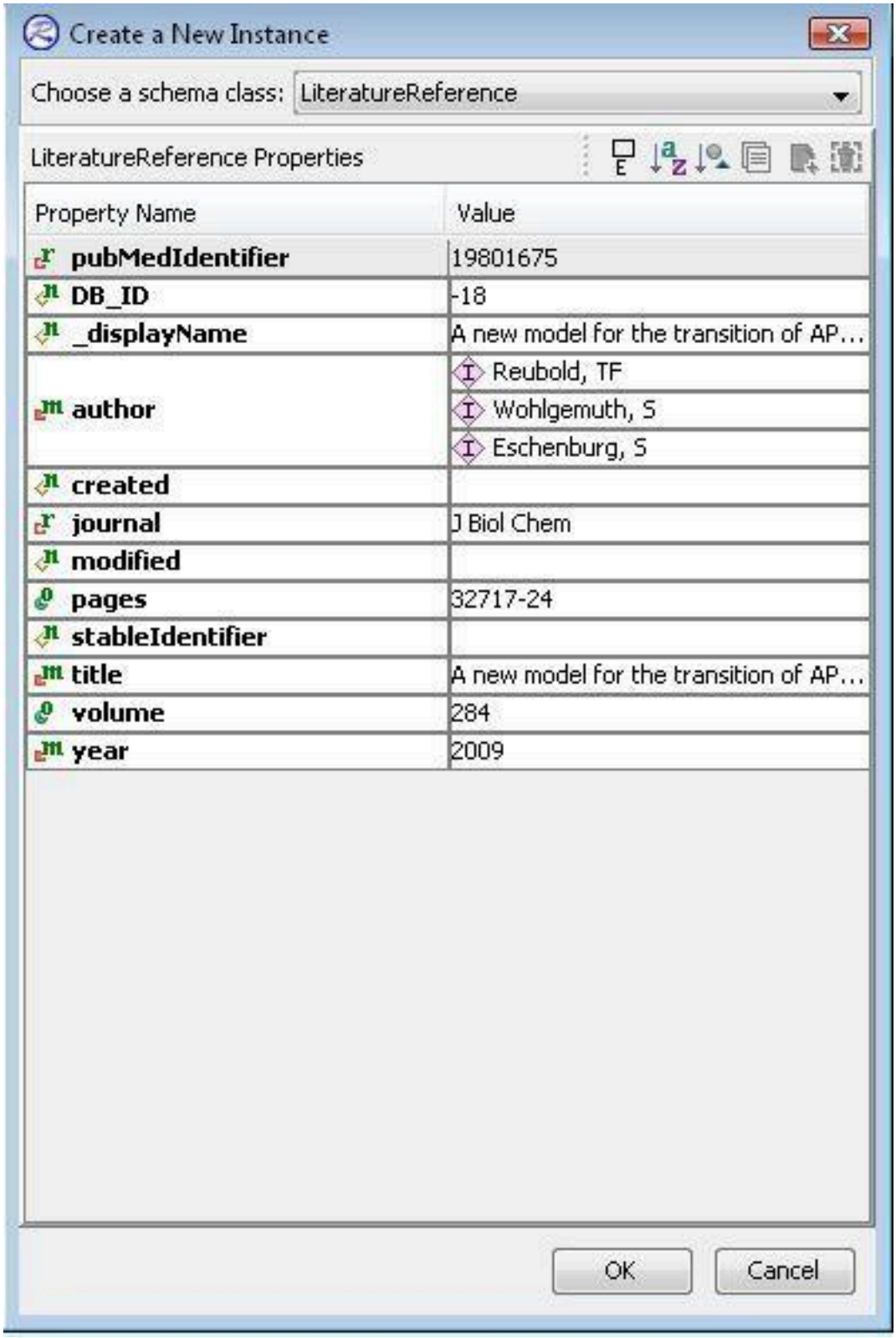


If you can't find the literature reference in the database either, then you need to create a new one:

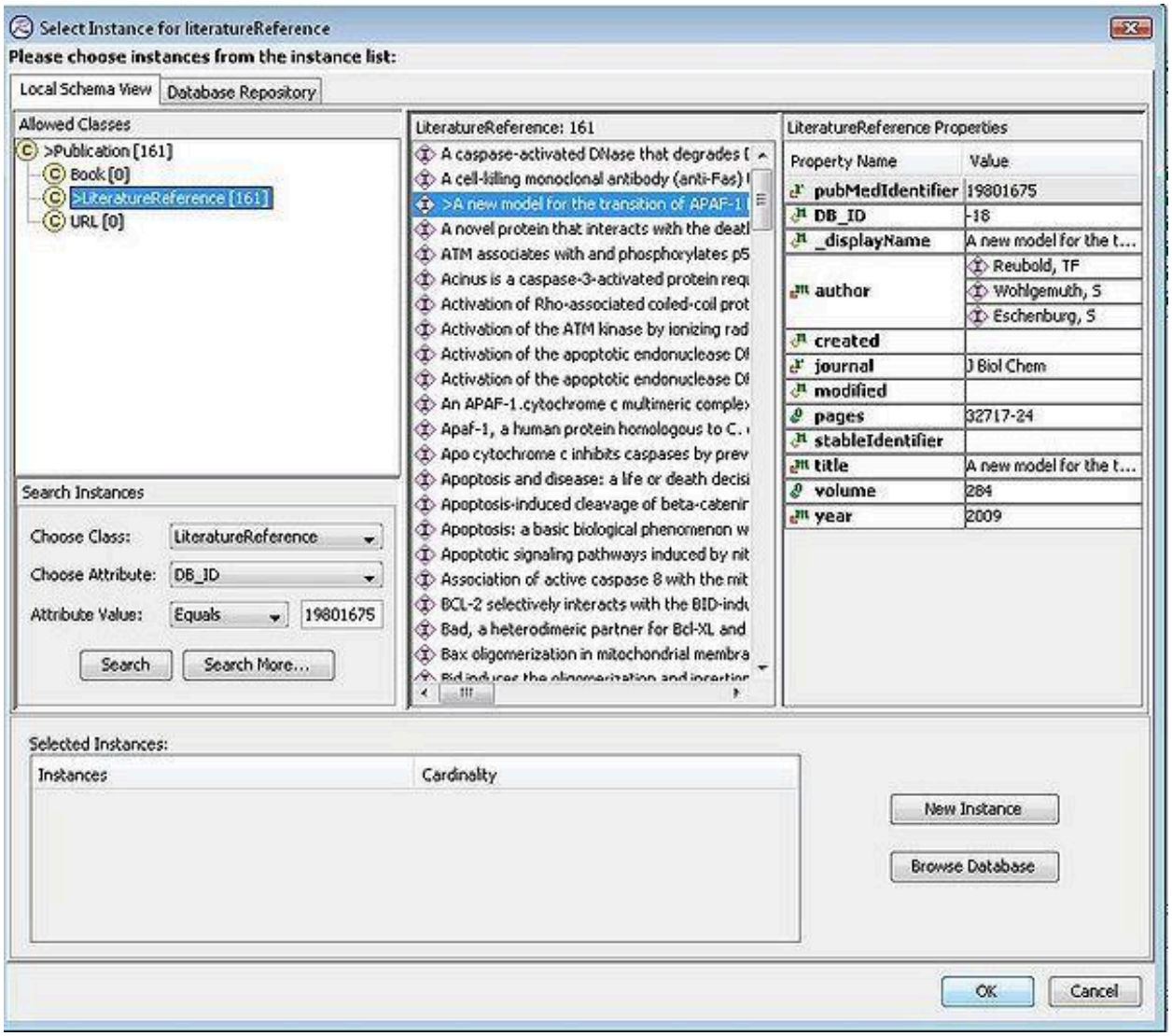
Enter the PMID (number only). Click out of the PMID box. You will be asked if you want to import the PMID record information. Say yes.


Create a New Instance
Choose a schema class: LiteratureReference
Specify properties for new instance:
Property Name
Value
pubMedIdentifier
19801675
DB_ID
-18
_displayName
author
created
journal
modified
pages
Fetching Information?
Do you want to fetch other information for the specified PMID?
Oui
Non
OK
Cancel

You will now see the full record:


Create a New Instance
Choose a schema class: LiteratureReference
LiteratureReference Properties
Property Name
Value
pubMedIdentifier
19801675
DB_ID
-18
_displayName
A new model for the transition of AP...
author
Reubold, TF
Wohlgemuth, S
Eschenburg, S
created
journal
J Biol Chem
modified
pages
32717-24
stableIdentifier
title
A new model for the transition of AP...
volume
284
year
2009
OK
Cancel

Look back in your local repository and you will see the new literature reference. You can now add this as a reference for your reaction.



The process is similar for adding other Publication types (Books and urls) with the exception that all information must be added manually - there is no automatic import.

To make a Publication record for a book, all information must be entered manually: author(s), title, year (mandatory attributes), ISBN, title, chapterTitle, pages (optional) as shown:

[Book:9816605] 9780071624961 Brooks, Geo F Chapter 32. Adenoviruses Jawetz, Melnick & Adel...

Book Properties

Property Name	Value
ISBN	9780071624961
author	Brooks, Geo F
	Carroll, Karen C
	Butel, Janet S
	Morse, Stephen A
	Mietzner, Timothy A
chapterTitle	Chapter 32. Adenoviruses
title	Jawetz, Melnick & Adelberg's Medical Microbiology, Twent...
DB_ID	9816605
_displayName	9780071624961 Brooks, Geo F Chapter 32. Adenoviruse...
chapterAuthors	Butel, Janet S
created	Orlic-Milacic, Marija, 2022-09-12
modified	Orlic-Milacic, Marija, 2022-09-13
pages	423-432
publisher	McGraw Hill
stableIdentifier	
year	2010

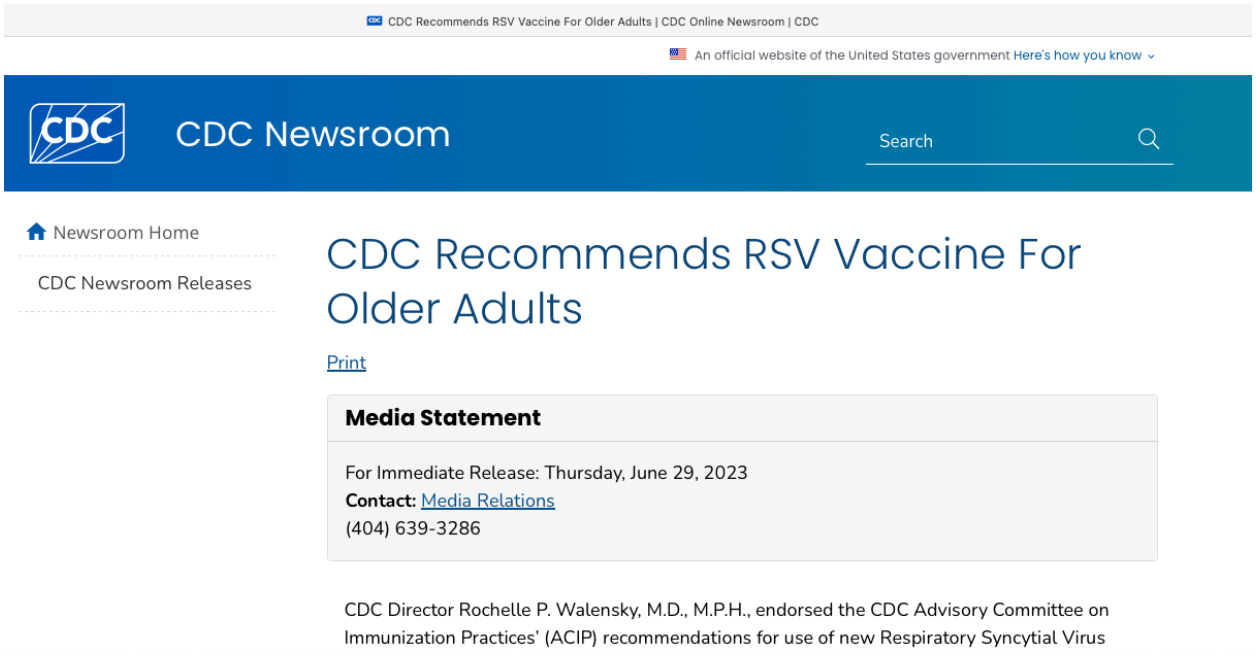
The image below shows an example of a Publication record for a url; again, information must be entered manually.

[URL:9845236] https://www.cdc.gov/media/releases/2023/s0629-rsv.html

URL Properties

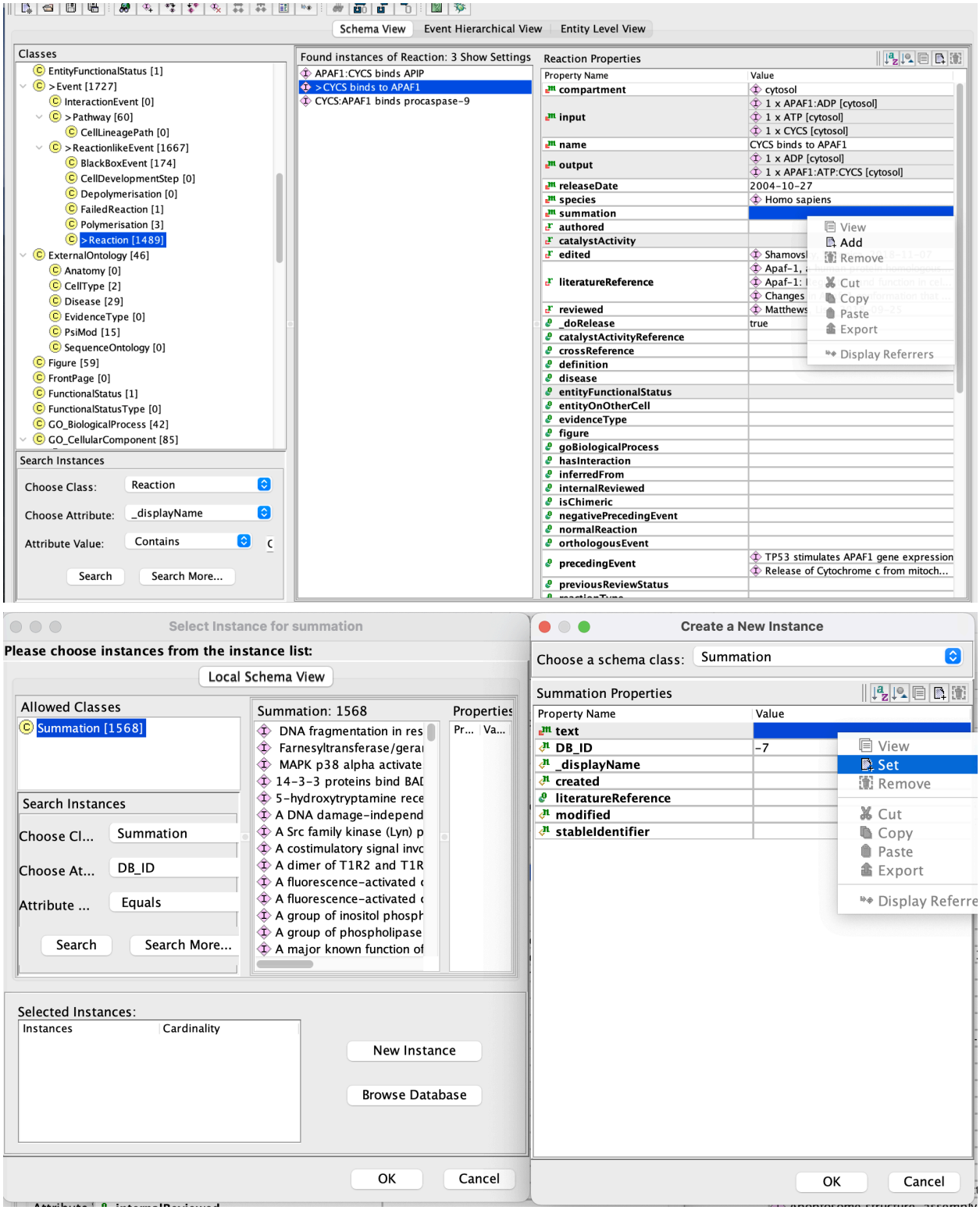
Property Name	Value
uniformResourceLocator	https://www.cdc.gov/media/releases/2023/s0629-rsv.html
DB_ID	9845236
_displayName	https://www.cdc.gov/media/releases/2023/s0629-rsv.html
author	Centers for Disease Control and Prevention USA, CDC
created	Stephan, Ralf, 2023-09-29
modified	
stableIdentifier	
title	CDC Recommends RSV Vaccine For Older Adults

There is some discretionary latitude in what can be entered as the url title and author. The title may be the heading on the page itself, as in the case below:

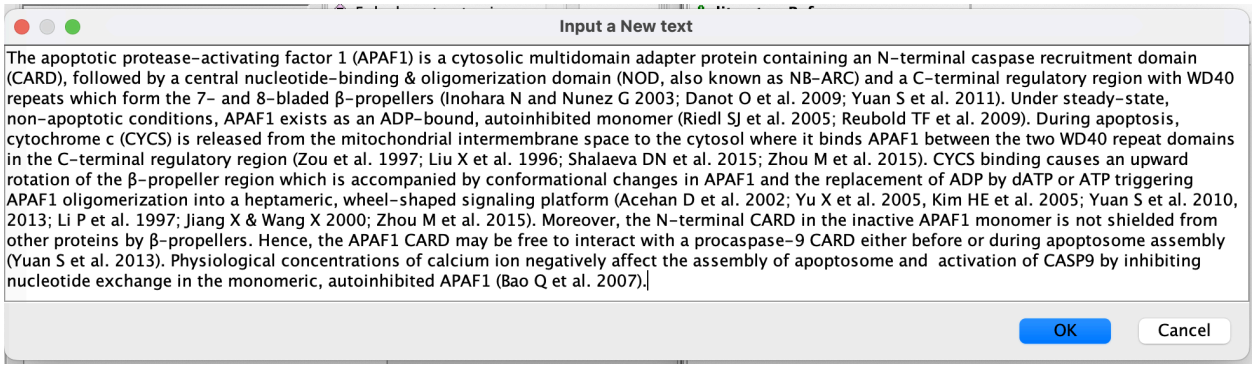


If in doubt, consult gk_central for existing records for similar urls.

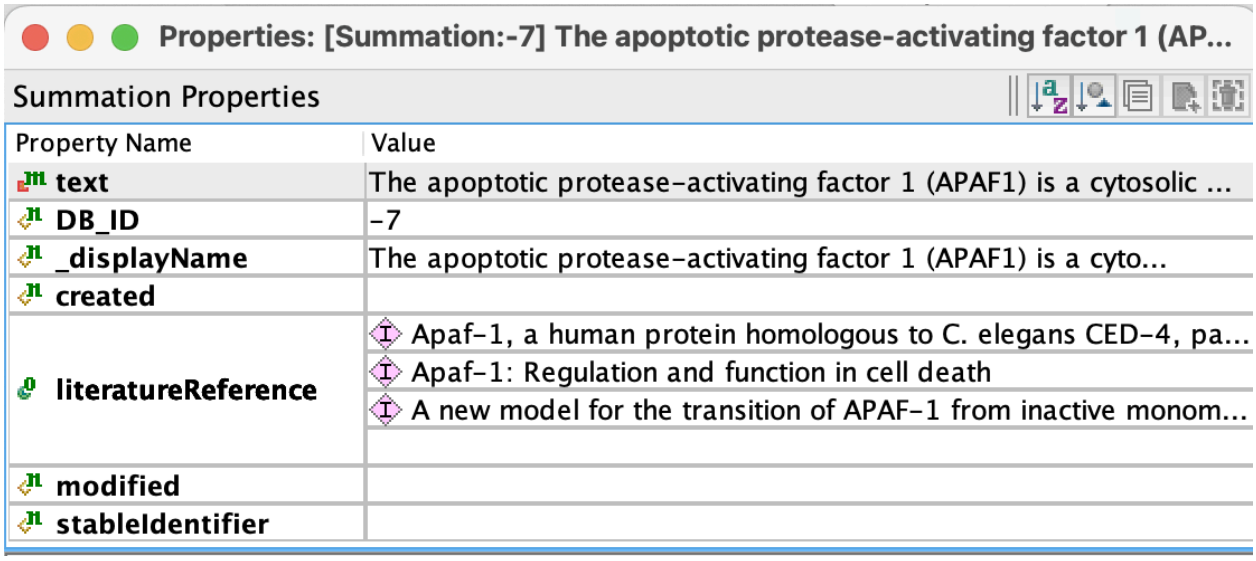
Once you have added your references, you can add a text summary for the reaction in the "summation"slot. Right click and select add.


Schema View
Event Hierarchical View
Entity Level View
Classes
EntityFunctionalStatus [1]
>Event [1727]
InteractionEvent [0]
>Pathway [60]
CellLineagePath [0]
>ReactionlikeEvent [1667]
BlackBoxEvent [174]
CellDevelopmentStep [0]
Depolymerisation [0]
FailedReaction [1]
Polymerisation [3]
>Reaction [1489]
ExternalOntology [46]
Anatomy [0]
CellType [2]
Disease [29]
EvidenceType [0]
PsiMod [15]
SequenceOntology [0]
Figure [59]
FrontPage [0]
FunctionalStatus [1]
FunctionalStatusType [0]
GO_BiologicalProcess [42]
GO_CellularComponent [85]
Search Instances
Choose Class:
Reaction
Choose Attribute:
_displayName
Attribute Value:
Contains
Search
Search More...
Found instances of Reaction: 3 Show Settings
APAF1:CYCS binds APIP
>CYCS binds to APAF1
CYCS:APAF1 binds procaspase-9
Reaction Properties
Property Name
Value
compartment
cytosol
input
1 x APAF1:ADP [cytosol]
1 x ATP [cytosol]
1 x CYCS [cytosol]
name
CYCS binds to APAF1
output
1 x ADP [cytosol]
1 x APAF1:ATP:CYCS [cytosol]
releaseDate
2004-10-27
species
Homo sapiens
summation
authored
catalystActivity
edited
literatureReference
reviewed
_doRelease
true
catalystActivityReference
crossReference
definition
disease
entityFunctionalStatus
entityOnOtherCell
evidenceType
figure
goBiologicalProcess
hasInteraction
inferredFrom
internalReviewed
isChimeric
negativePrecedingEvent
normalReaction
orthologousEvent
precedingEvent
TP53 stimulates APAF1 gene expression
Release of Cytochrome c from mitoch...
previousReviewStatus
View
Add
Remove
Cut
Copy
Paste
Export
Display Referrers
Select Instance for summation
Please choose instances from the instance list:
Local Schema View
Allowed Classes
Summation [1568]
Search Instances
Choose Cl...
Summation
Choose At...
DB_ID
Attribute ...
Equals
Search
Search More...
Summation: 1568
DNA fragmentation in res
Farnesyltransferase/gerai
MAPK p38 alpha activate
14-3-3 proteins bind BAD
5-hydroxytryptamine rece
A DNA damage-independ
A Src family kinase (Lyn) p
A costimulatory signal invc
A dimer of T1R2 and T1R
A fluorescence-activated
A fluorescence-activated
A group of inositol phosph
A group of phospholipase
A major known function of
Properties
Pr...
Va...
Selected Instances:
Instances
Cardinality
New Instance
Browse Database
OK
Cancel
Create a New Instance
Choose a schema class:
Summation
Summation Properties
Property Name
Value
text
DB_ID
-7
_displayName
created
literatureReference
modified
stableIdentifier
View
Set
Remove
Cut
Copy
Paste
Export
OK
Cancel

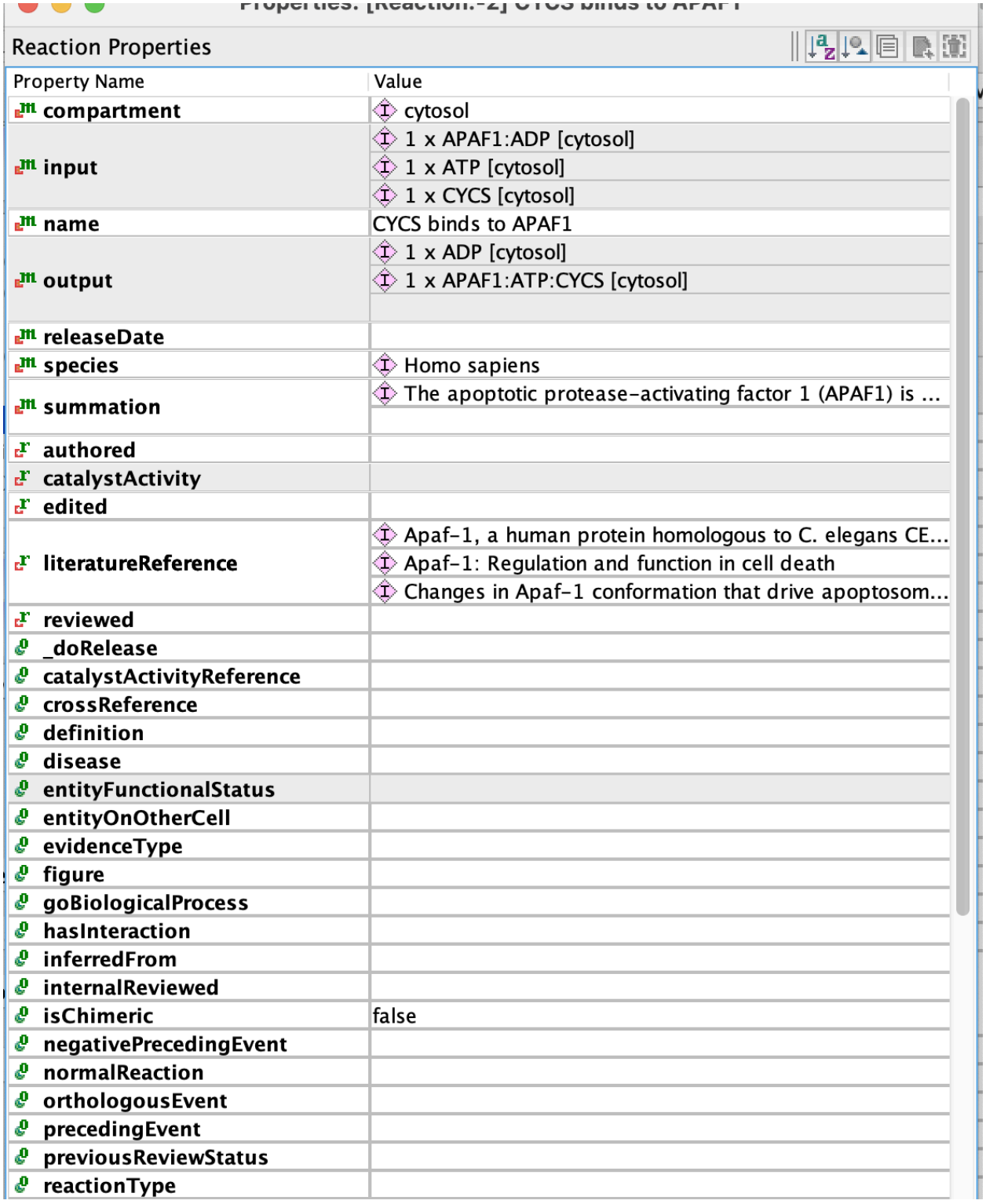
Then input your text and citations.



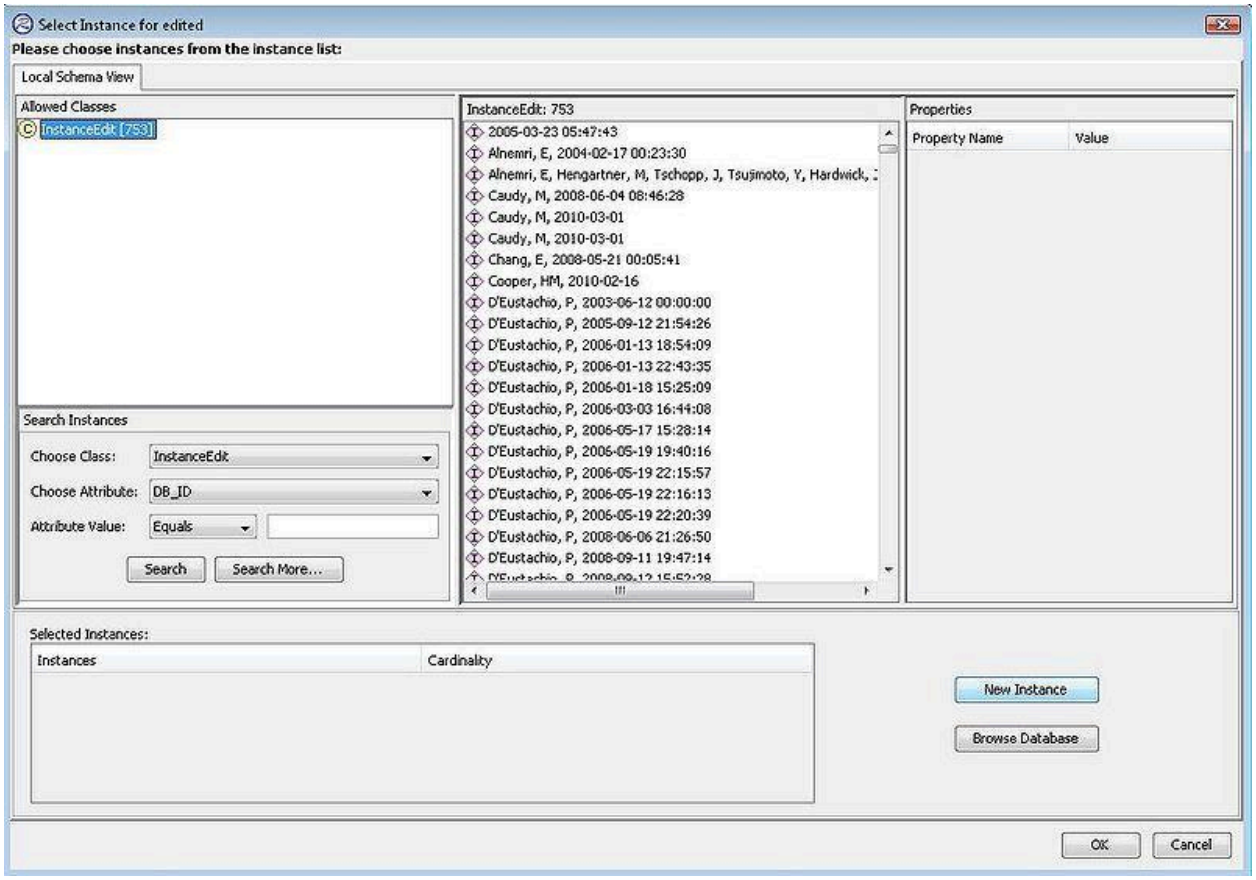
Associate the references associated with the citations in the text summary using the "LiteratureReference" slot for the summation as described above for reactions.



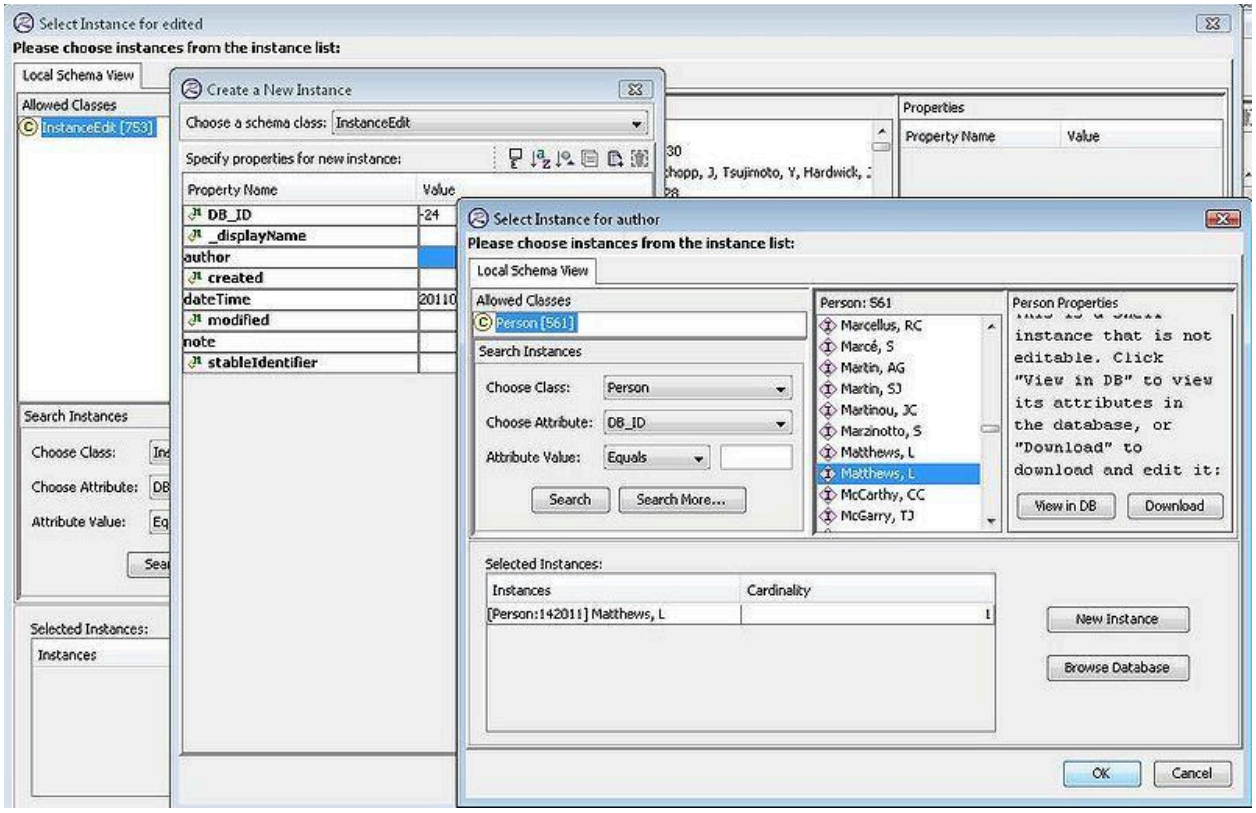
Once the mandatory attributes have been filled, enter the remaining required attributes. Right click on edited to choose an instanceEdit from your project.


Reaction Properties
Property Name
Value
compartment
cytosol
input
1 x APAF1:ADP [cytosol]
1 x ATP [cytosol]
1 x CYCS [cytosol]
name
CYCS binds to APAF1
output
1 x ADP [cytosol]
1 x APAF1:ATP:CYCS [cytosol]
releaseDate
species
Homo sapiens
summation
The apoptotic protease-activating factor 1 (APAF1) is ...
authored
catalystActivity
edited
literatureReference
Apaf-1, a human protein homologous to C. elegans CE...
Apaf-1: Regulation and function in cell death
Changes in Apaf-1 conformation that drive apoptosom...
reviewed
_doRelease
catalystActivityReference
crossReference
definition
disease
entityFunctionalStatus
entityOnOtherCell
evidenceType
figure
goBiologicalProcess
hasInteraction
inferredFrom
internalReviewed
isChimeric
false
negativePrecedingEvent
normalReaction
orthologousEvent
precedingEvent
previousReviewStatus
reactionType

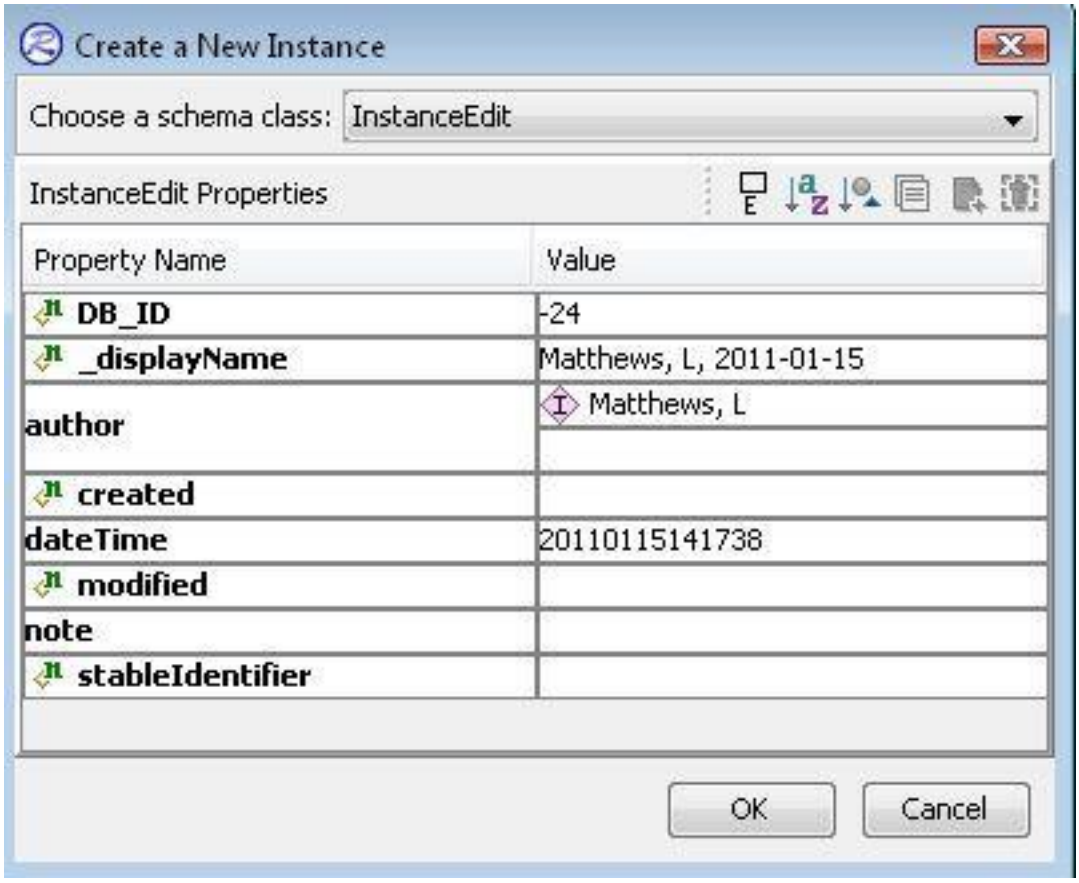
If you haven't created one previously, opt to create a new one by clicking on the New Instance button.



Then right click on author to select a person. If you don't see the one you want locally, browse the database. The edited slot holds the name of the curator that should be credited with creating and editing the event and the date it was edited. This slot is filled with an instanceedit instance that contains this information.



A date time stamp is created automatically for that instanceEdit and clicking "OK" will add this instanceEdit to the edited slot in your reaction.



The authored and reviewed slots hold the instanceEdits describing the author/reviewer and dates of authoring/reviewing respectively. These are created as described above for the "editor" slot. When the reaction is read for release, the do_Release flag should be set to TRUE and the releaseDate slot should be filled with the appropriate release date.

Property Name	Value
compartment	cytosol
input	1 x APAF1:ADP [cytosol]
	1 x ATP [cytosol]
	1 x CYCS [cytosol]
name	CYCS binds to APAF1
output	1 x ADP [cytosol]
	1 x APAF1:ATP:CYCS [cytosol]
releaseDate	
species	Homo sapiens
summation	The apoptotic protease-activating factor 1 (APAF1) is...
authored	
catalystActivity	
edited	Matthews, L, 2011-01-15
literatureReference	Apaf-1, a human protein homologous to C. elegans ...
	Apaf-1: Regulation and function in cell death
	Changes in Apaf-1 conformation that drive apoptoso...
reviewed	
_doRelease	
catalystActivityReference	
crossReference	
definition	
disease	
entityFunctionalStatus	
entityOnOtherCell	
evidenceType	
figure	
goBiologicalProcess	
hasInteraction	
inferredFrom	
internalReviewed	
isChimeric	false
negativePrecedingEvent	
normalReaction	
orthologousEvent	

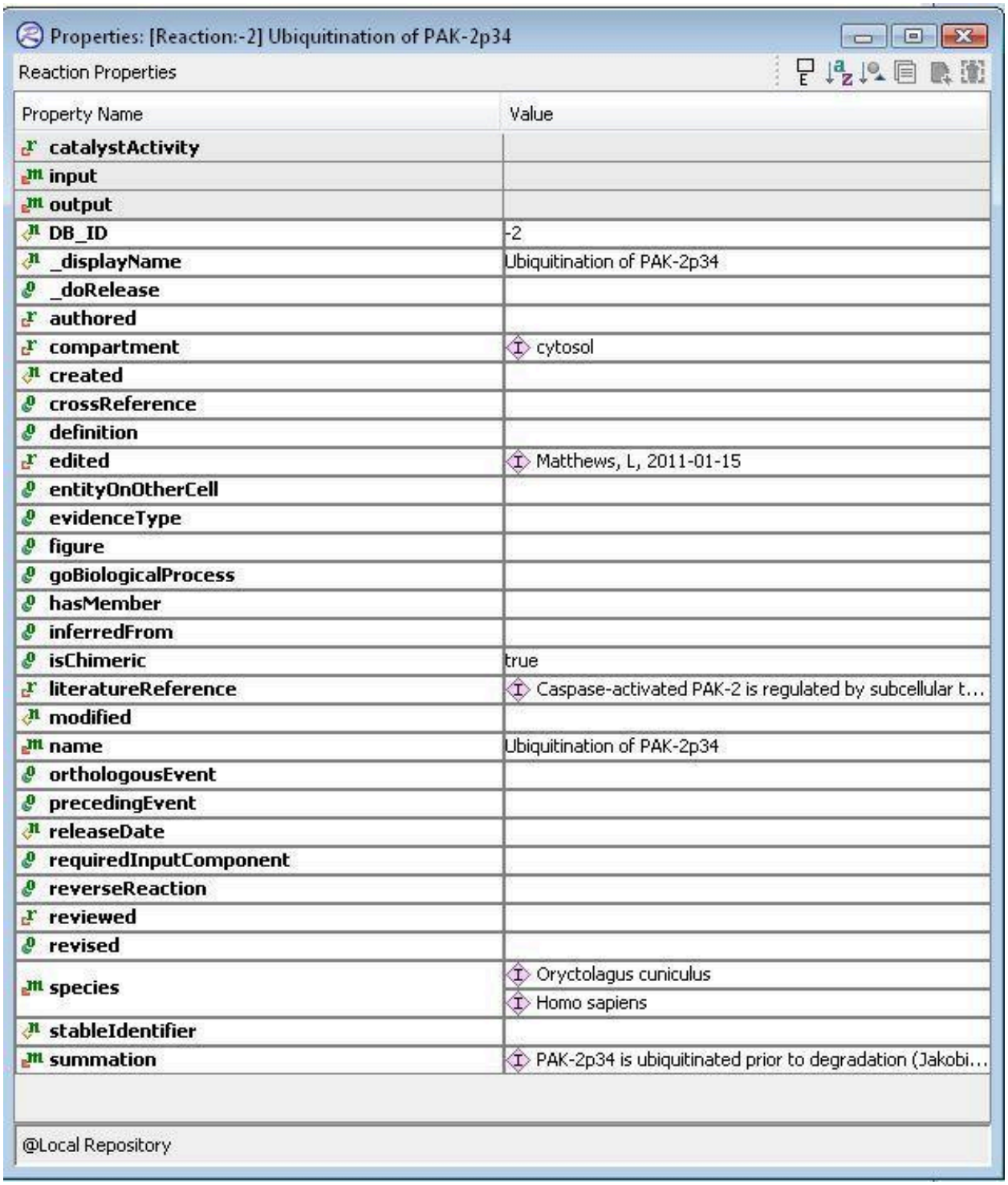
Important: After adding your input and output molecules, it is important to verify that your reaction is balanced (all molecules represented as input are also present as output). See the QA section below for a description of how to do this.

When you can define preceding events for an event it should be done. One precision on that point is that a preceding event for a reaction should always be a reaction/reaction like event and not a pathway and the preceding event for a pathway should be another pathway when relevant and not a reaction or reaction like event.

Creating an inferred event

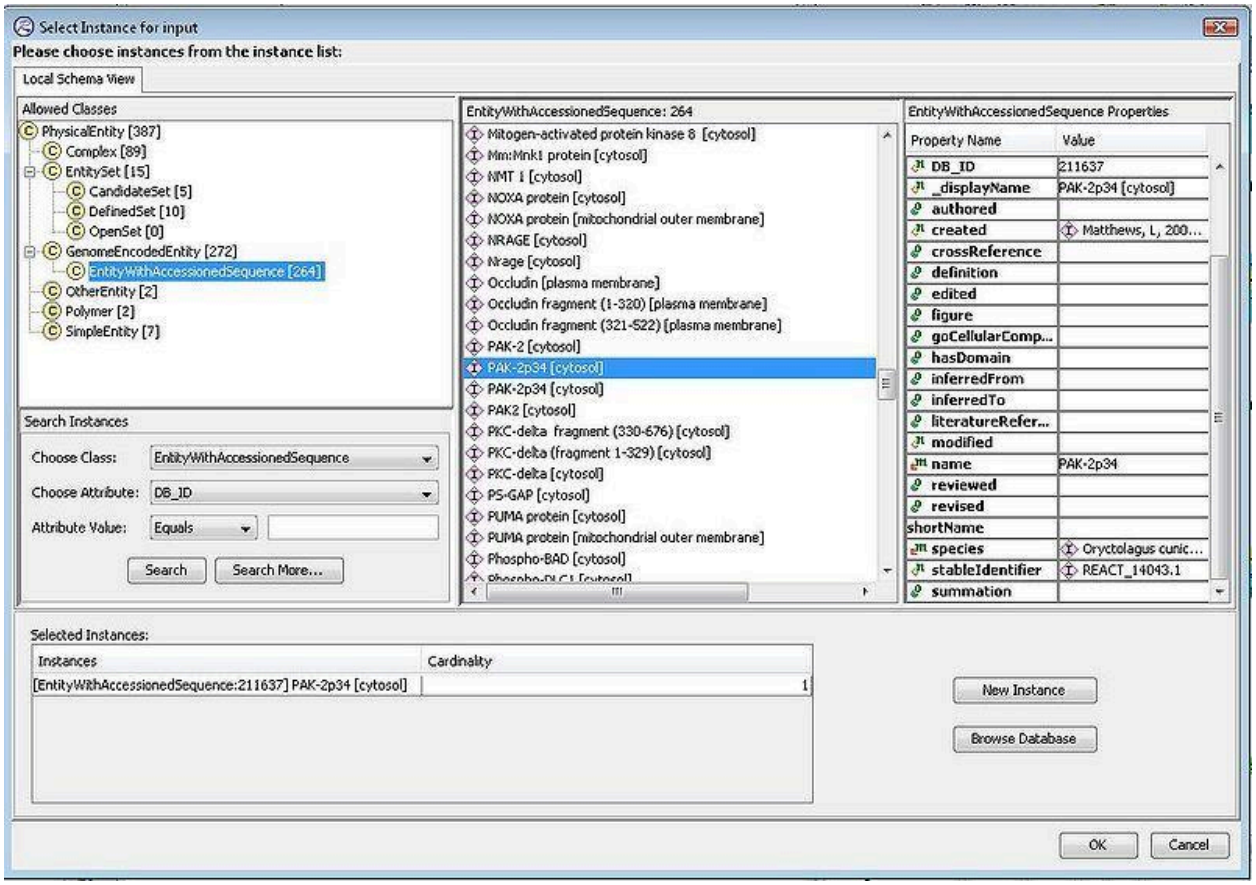
When constructing a human pathway, a curator will come across events that have no direct experimental evidence in humans, but have supporting experimental data from another/other species. If experts in the field believe that the event can in fact occur in humans, the 'other species event' can be used to infer the human event. In the case below a human reaction is inferred from in vitro experimental results using proteins from human and Oryctolagus cuniculus (rabbit).

The reaction "Proteosome mediated degradation of PAK-2p34" is created and the species Homo sapiens and Oryctolagus cuniculus are assigned

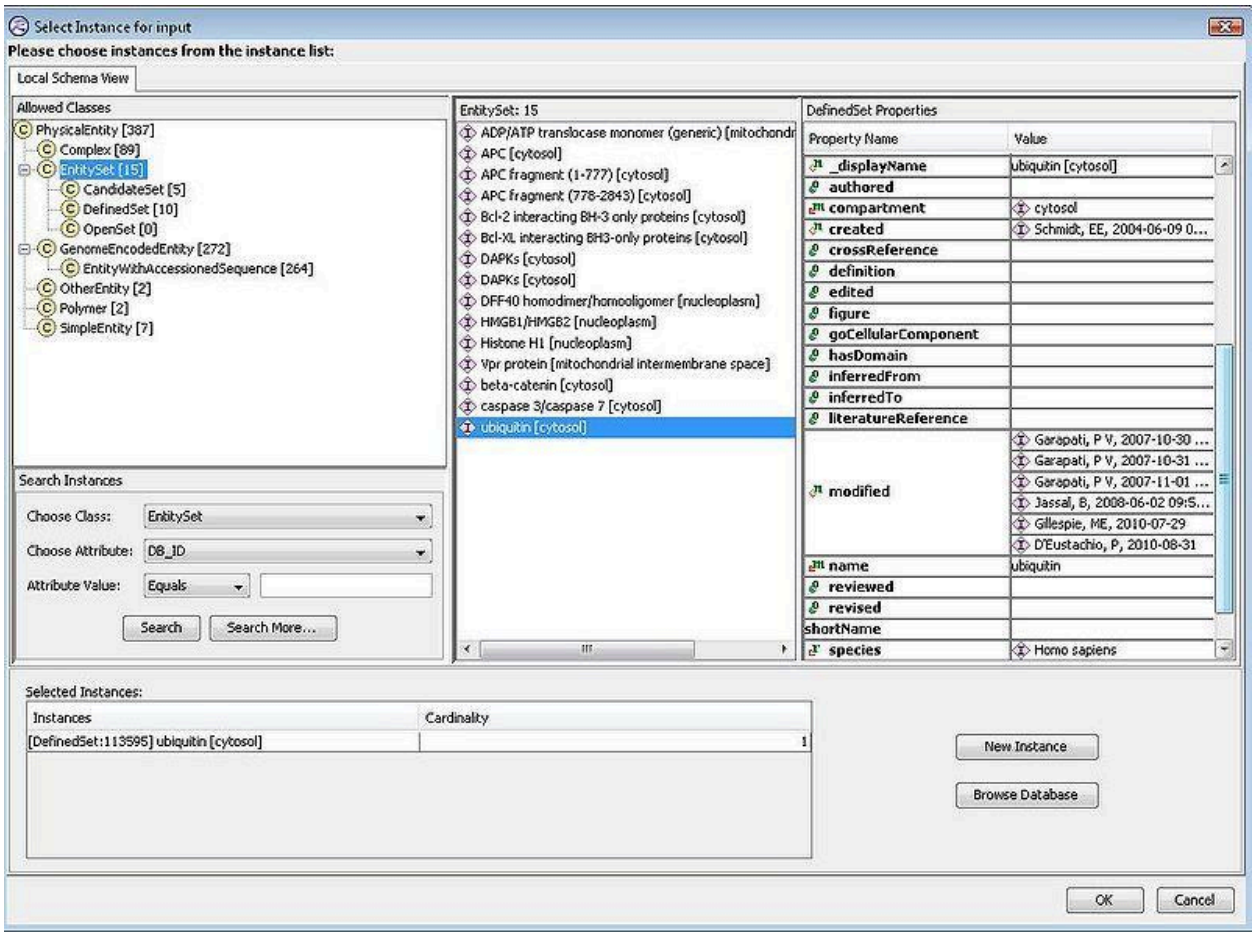

Properties: [Reaction:-2] Ubiquitination of PAK-2p34
Reaction Properties
Property Name
Value
catalystActivity
input
output
DB_ID
-2
_displayName
Ubiquitination of PAK-2p34
_doRelease
authored
compartment
cytosol
created
crossReference
definition
edited
Matthews, L, 2011-01-15
entityOnOtherCell
evidenceType
figure
goBiologicalProcess
hasMember
inferredFrom
isChimeric
true
literatureReference
Caspase-activated PAK-2 is regulated by subcellular t...
modified
name
Ubiquitination of PAK-2p34
orthologousEvent
precedingEvent
releaseDate
requiredInputComponent
reverseReaction
reviewed
revised
species
Oryctolagus cuniculus
Homo sapiens
stableIdentifier
summation
PAK-2p34 is ubiquitinated prior to degradation (Jakobi...
@Local Repository

To avoid having a human and a non-human reaction with identical names, it can be useful to use capitalized forms of object names for the non-human reaction and all-uppercase names for human, e.g. Jak2 and JAK2 for the non-human and human proteins respectively. A text summary and the literature reference providing evidence for this mixed species reaction is added.

The rabbit protein "PAK-2p34" is selected as an input

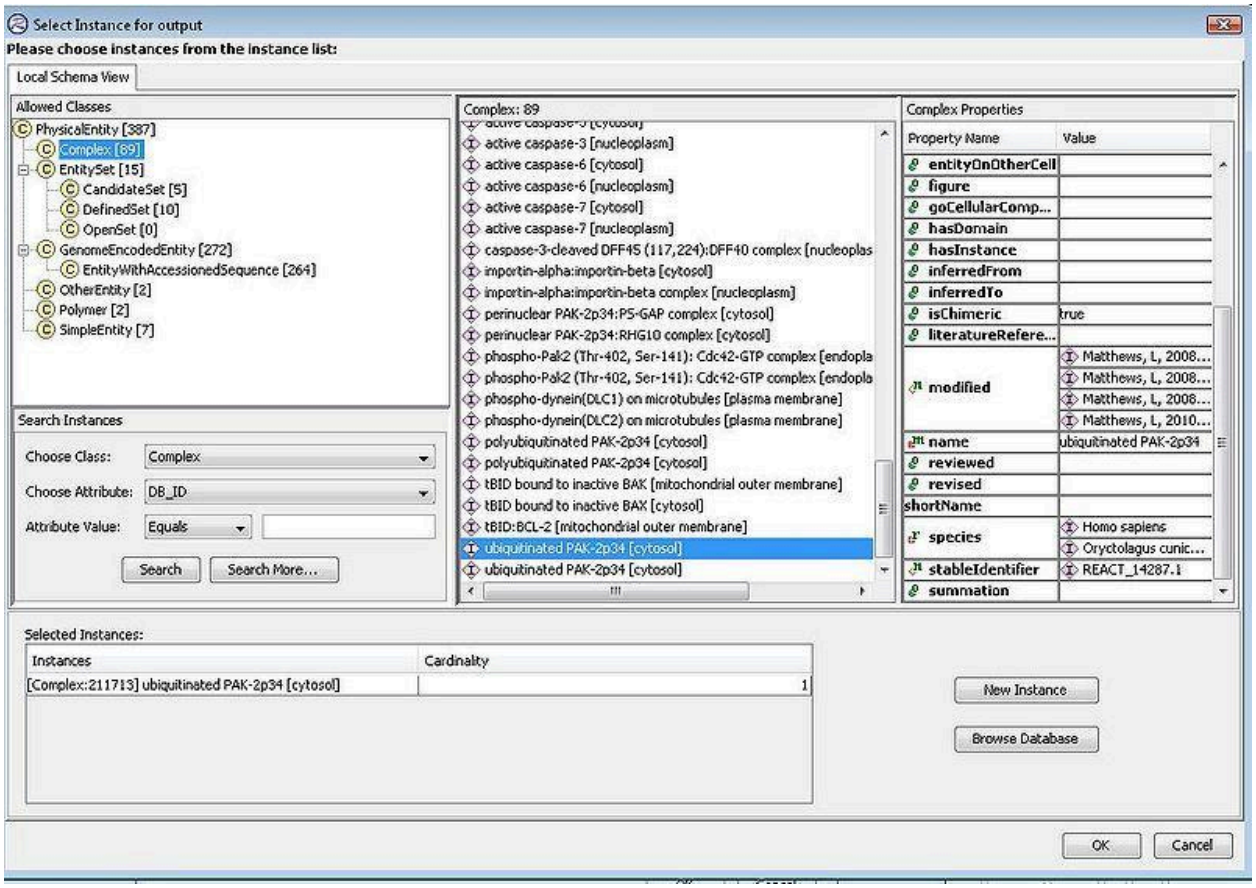


The human set "ubiquitin" is selected as an input...

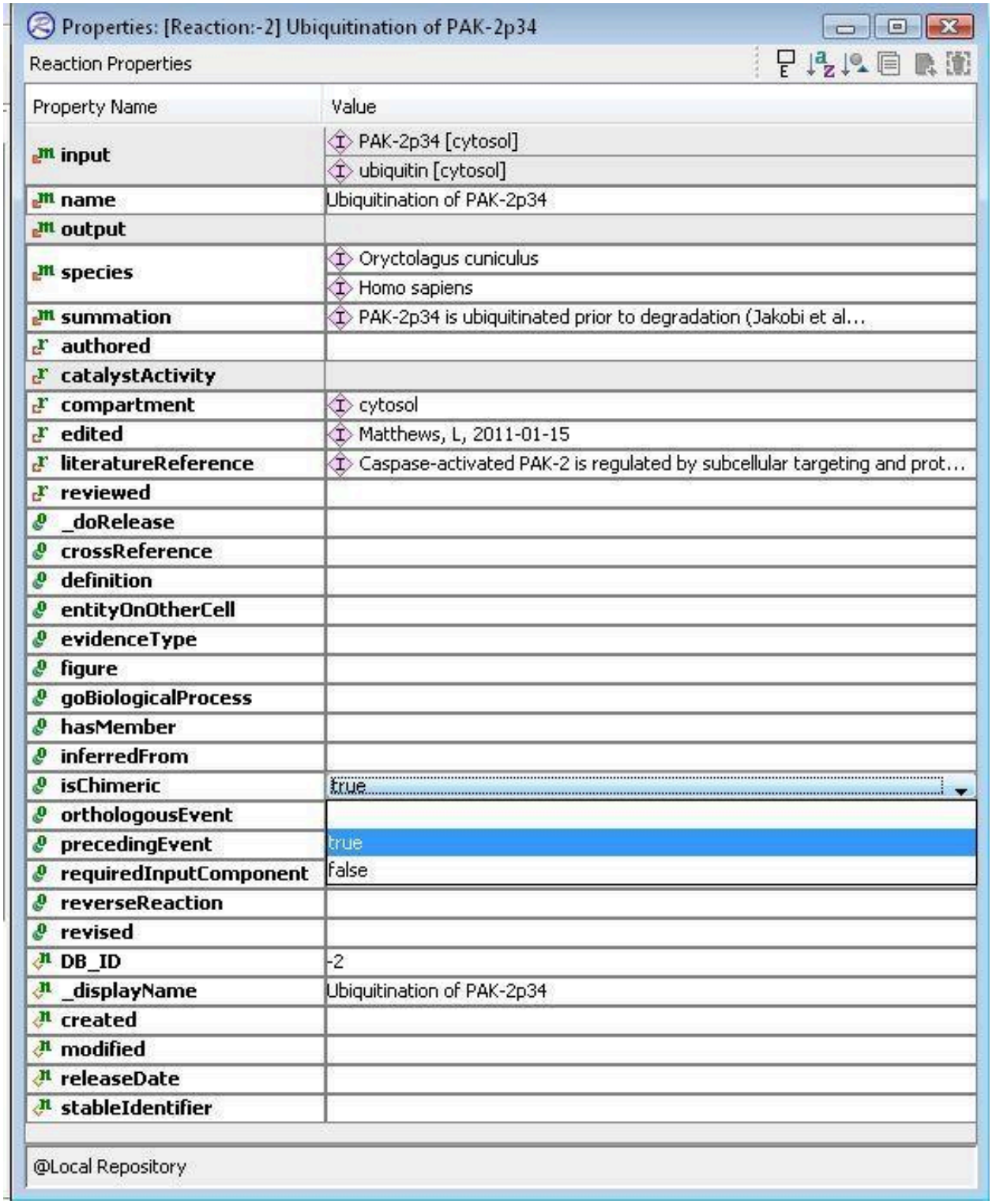


...and the mixed species output complex "PAK-2p34" is selected as the output.

Note that this complex is not natural, it only occurs in vitro. To flag this the complex "isChimeric" attribute is set to True.

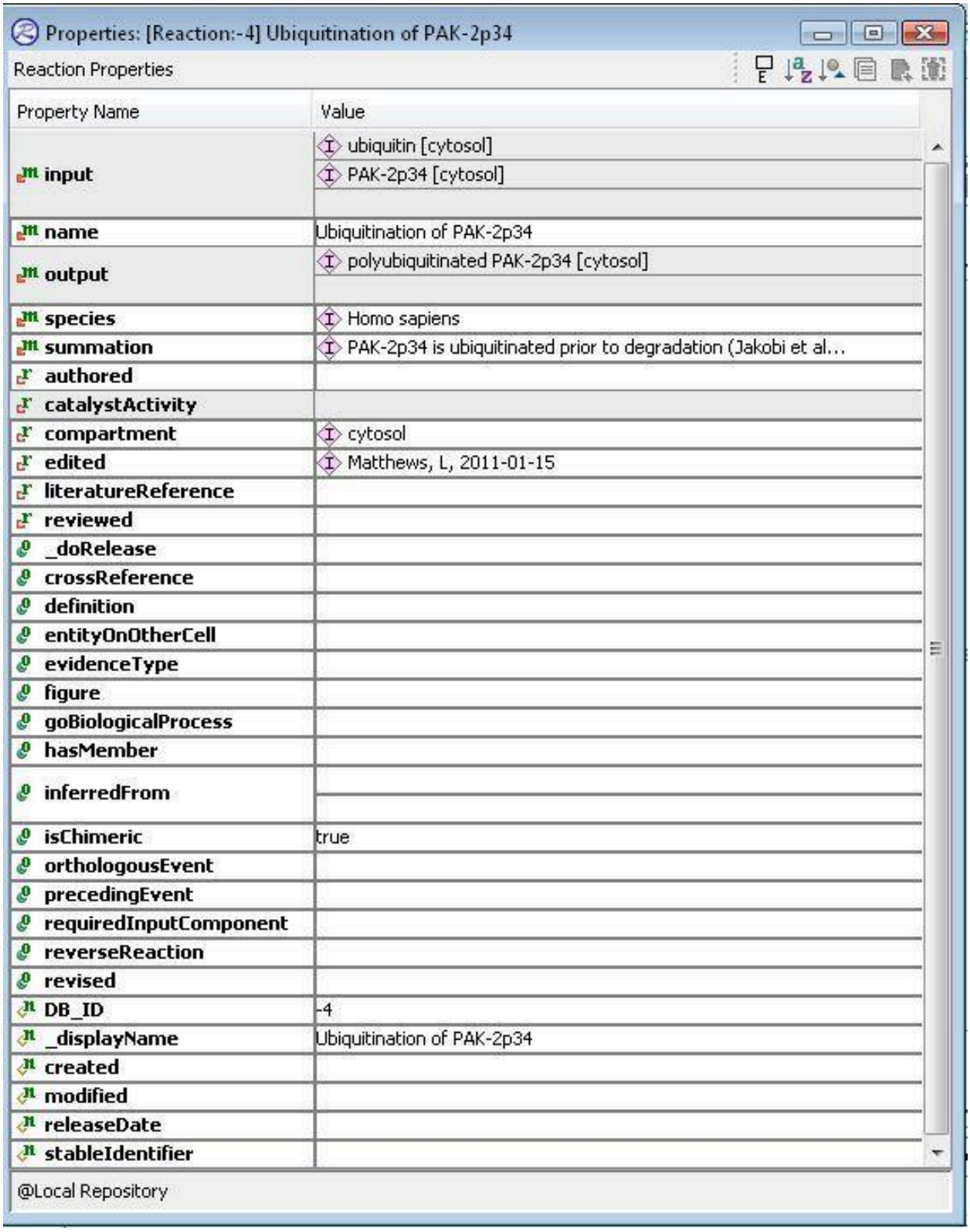


Since the reaction is also multispecies and not a natural event, the isChimera flag is also set to true for the reaction.

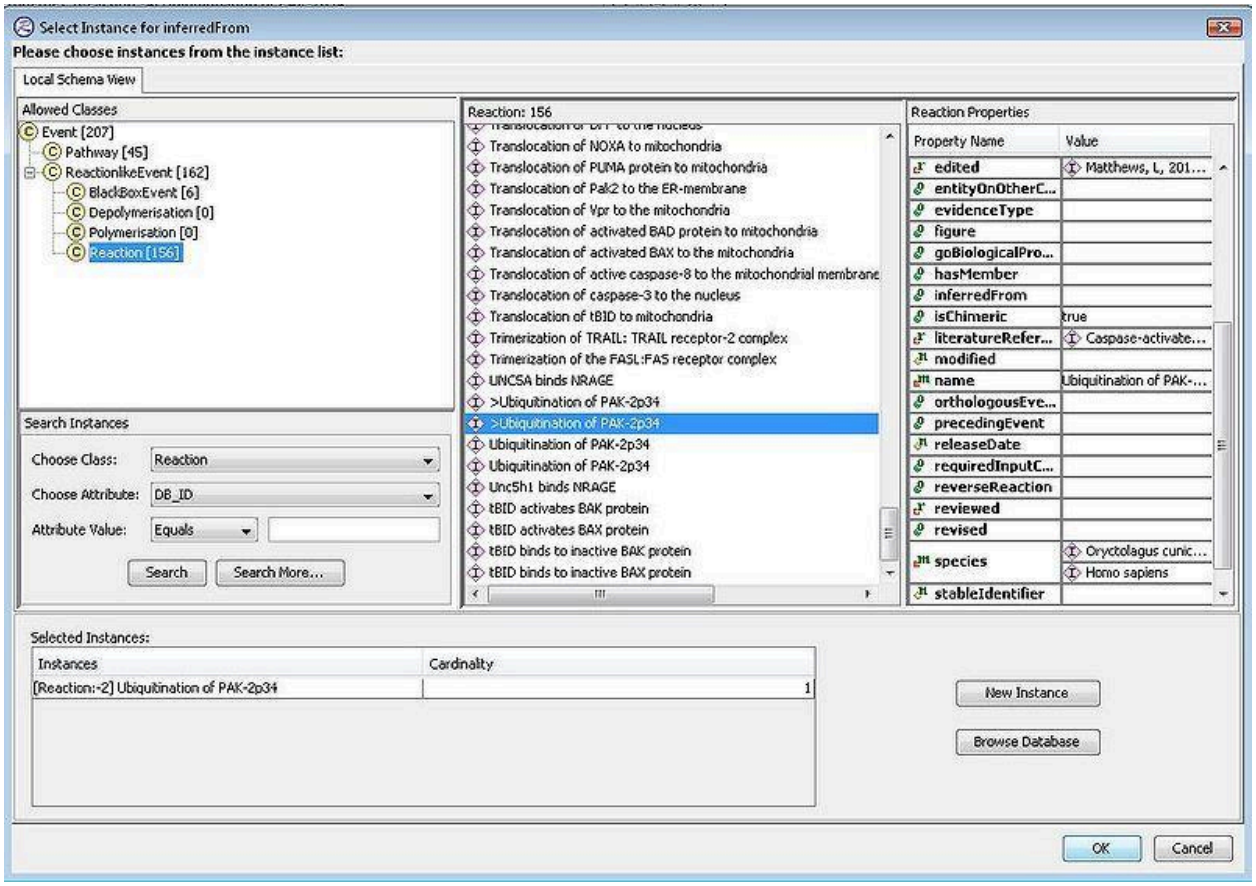


Now that the reaction to be used for inference has been created, create the same reaction for human, using human participating molecules. In the "inferredFrom" field, right click and select "Add" to enter the non-human reaction used for inference. Add

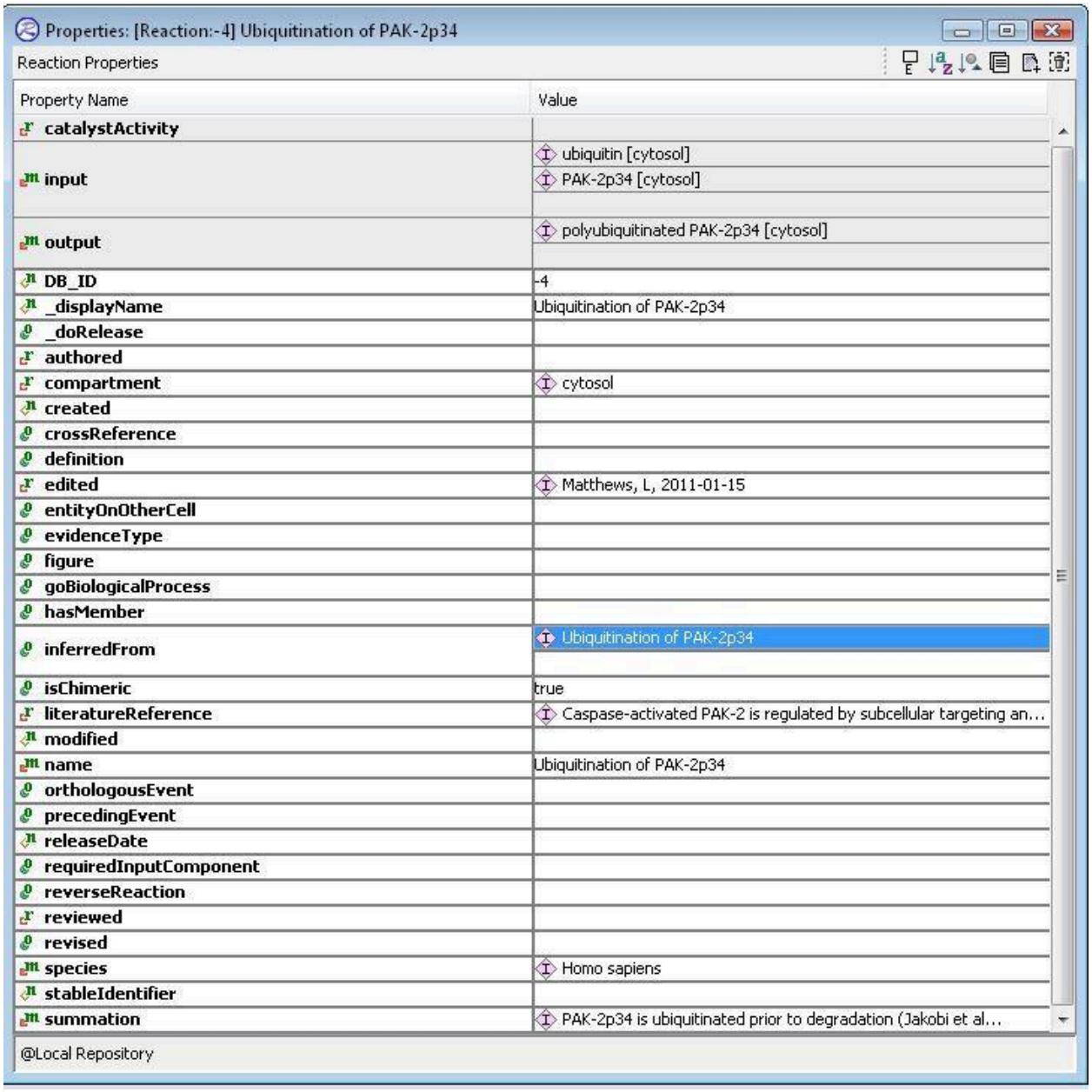
literature references and summation as usual.



Selecting the mixed species reaction created previously...



...now the link to the inferred reaction has been made.



- Note also that if your annotations involve a species that is new to Reactome, please contact Peter D'Eustachio as it will need to have an abbreviation assigned by him.

Creating a Catalyst

Reactions that involve a catalyst should include this information. Within the Reaction Details, the field is called catalystActivity. To complete this you need two things: the physicalEntity or object that is acting as catalyst, and the Activity of that object, defined as a GO molecular function. The physicalEntity will be a molecule or set or complex probably in your local project. The GO molecular function can be identified by consulting

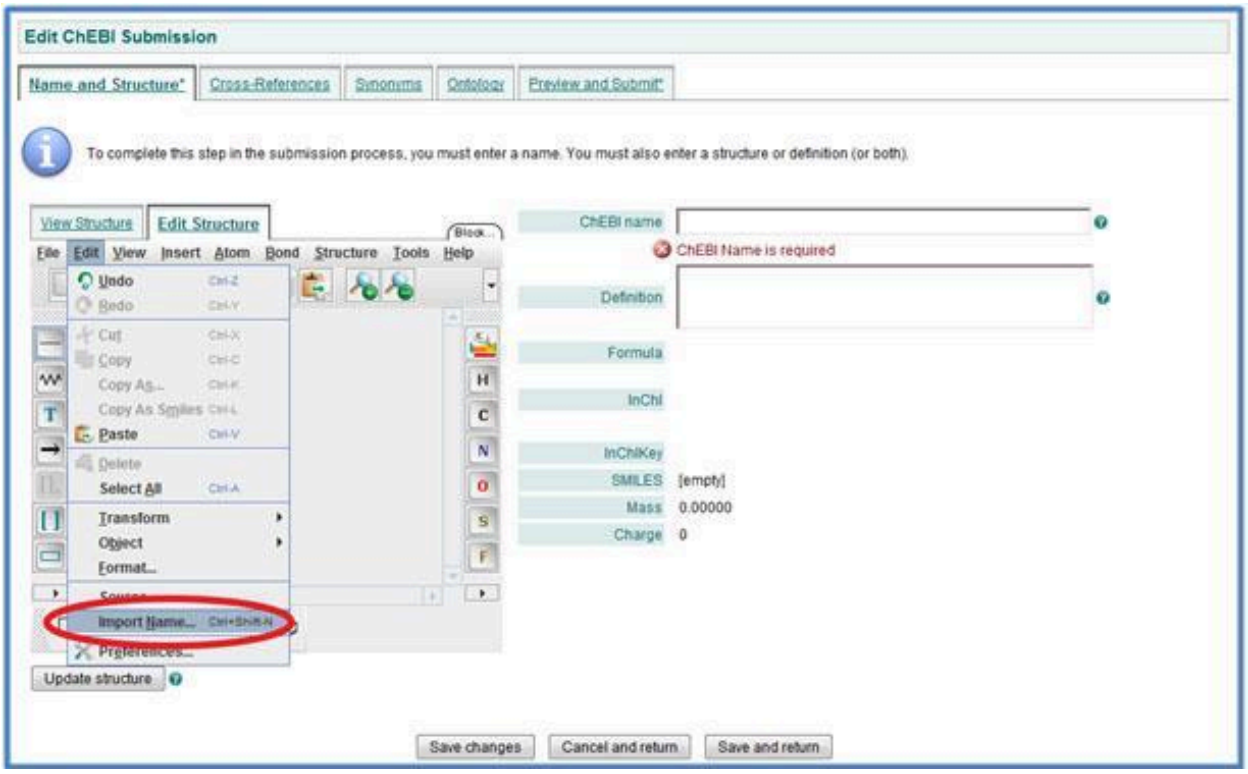
Uniprot, look at the ontologies section for GO Molecular Function. If none of the listed terms seems to be appropriate, either Browse Database for terms in gk_central, or use the OLS website at http://www.ebi.ac.uk/ontology-lookup/ to identify the correct term first. Use the most specific term possible.

Creating a new ChEBI entry

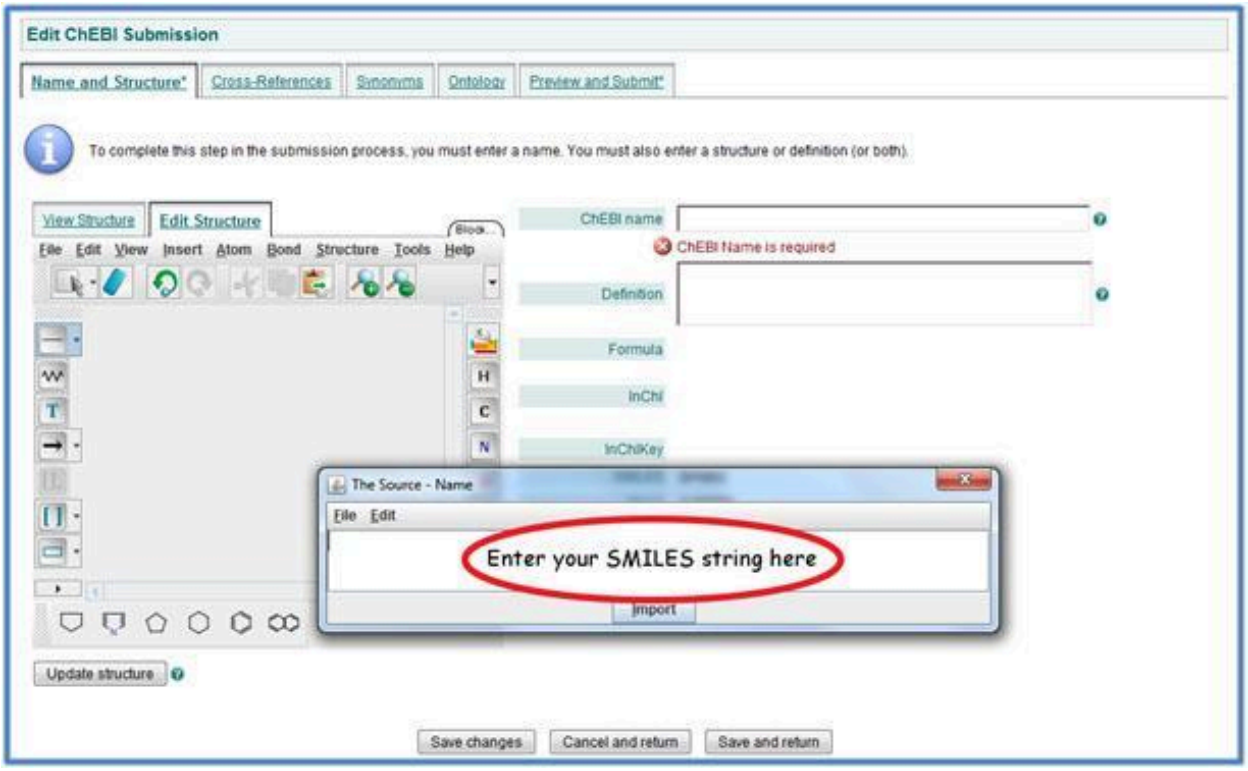
See also Best_Practices:_Using_Chemical_Compound_Databases, Appendix C8 .

Using SMILES strings as input for the ChEBI submission tool

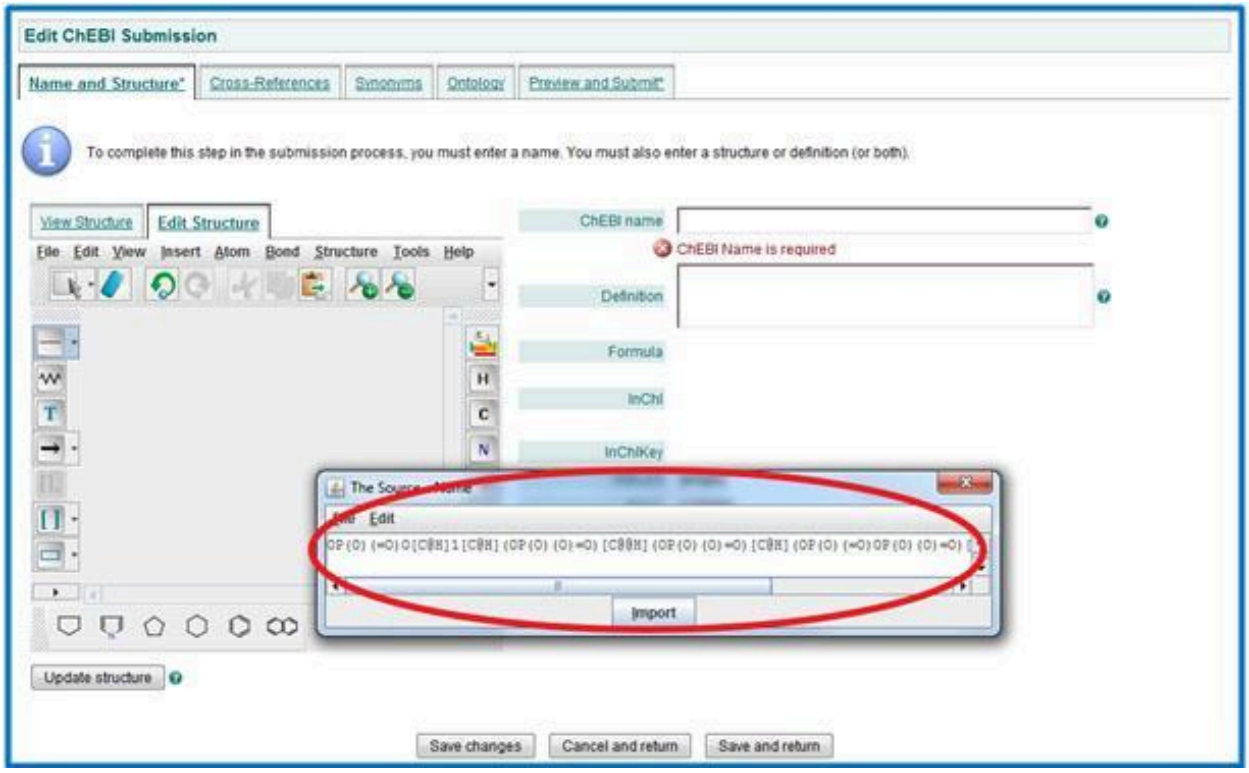
To use the submission tool, you must get a user name and password, and log in. Go here (https://www.ebi.ac.uk/chebi/submission/login) to do that. Once you have logged in, click "create a new submission" from the choices at the bottom of the page. That will cause a new line to appear in the table of "your active submissions" on that page. Click the "edit submission" option on that line to open the actual submission form.
Under the ‘Name And Structure’ section of the Submission tool, select ‘Edit Structure’.
Under the ‘Edit’ Menu, select ‘Import Name’.



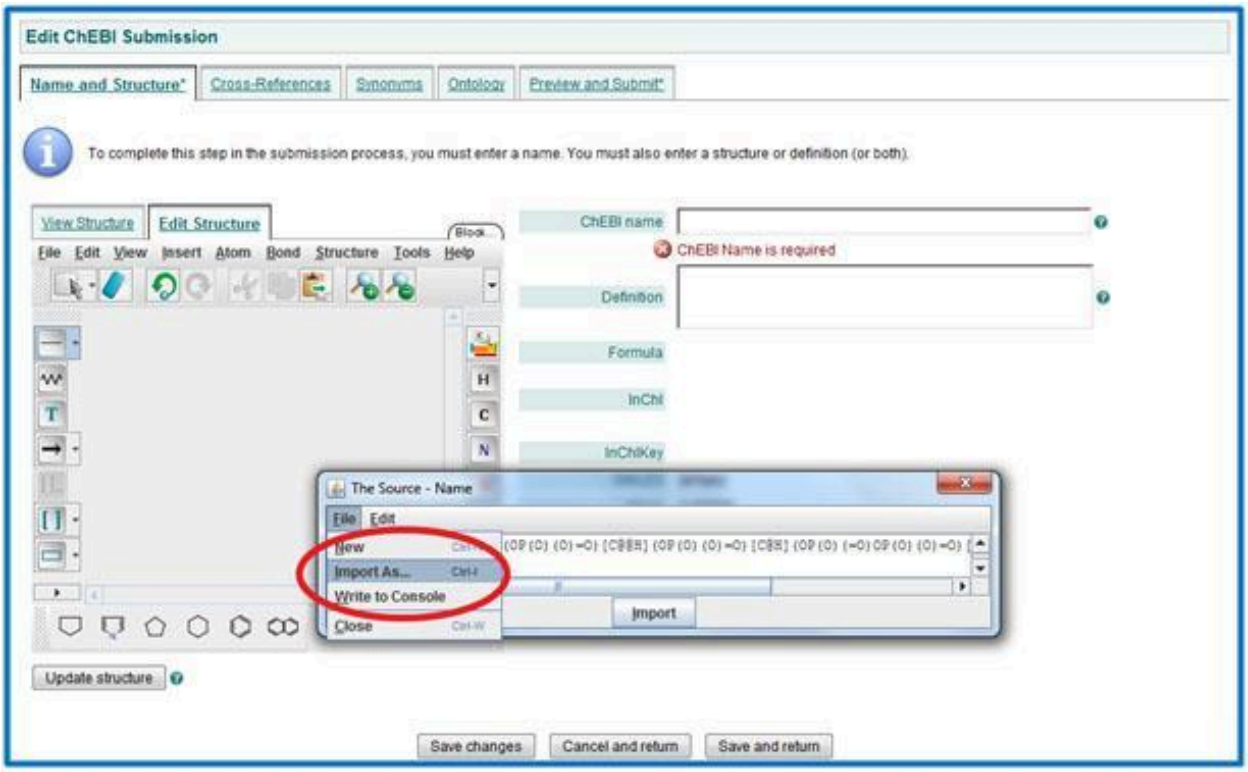
An input box named ‘The Source – Name’ appears, this is the place to paste your SMILES string.



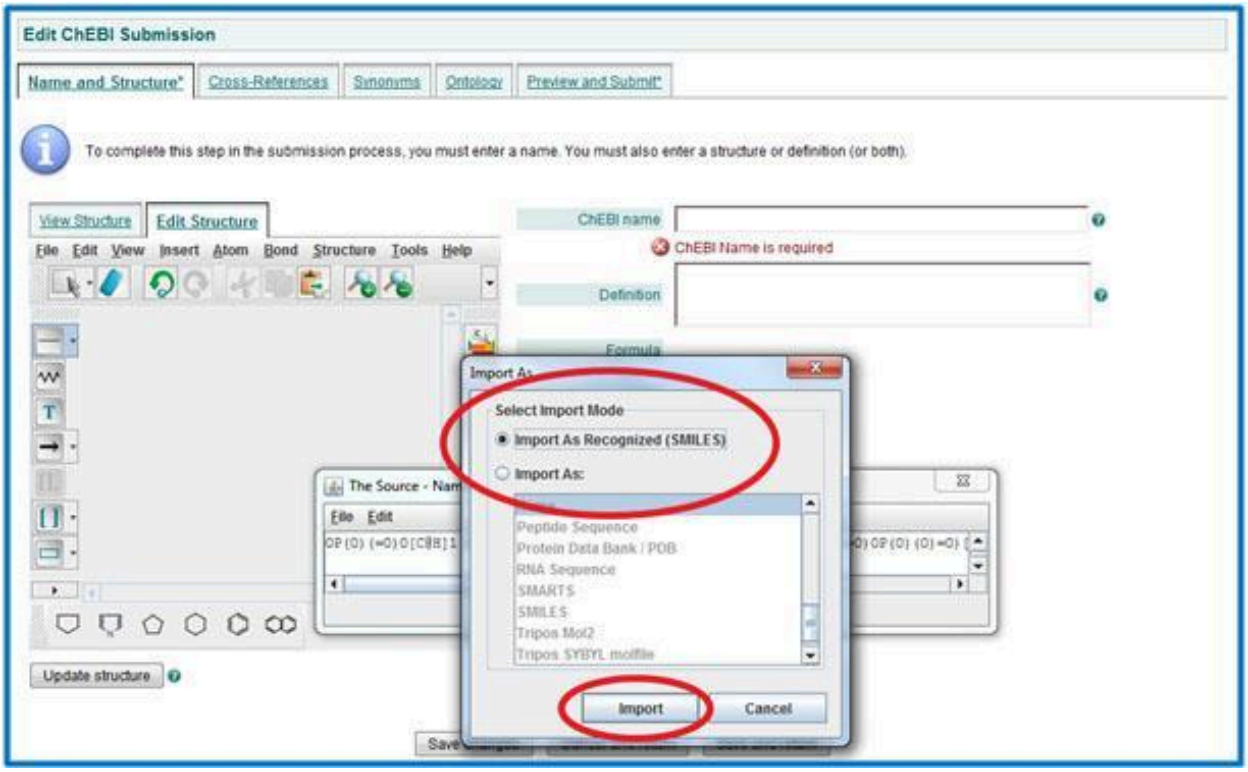
In this example, the following string for 1-PP-IP5 is used:
OP(O)(=O)O[C@H]1[C@H](OP(O)(O)=O)[C@@H](OP(O)(O)=O)[C@H](OP(O)(=O)OP(O)(O)=O)[C@H](OP(O)(O)=O)[C@@H]1OP(O)(O)=O



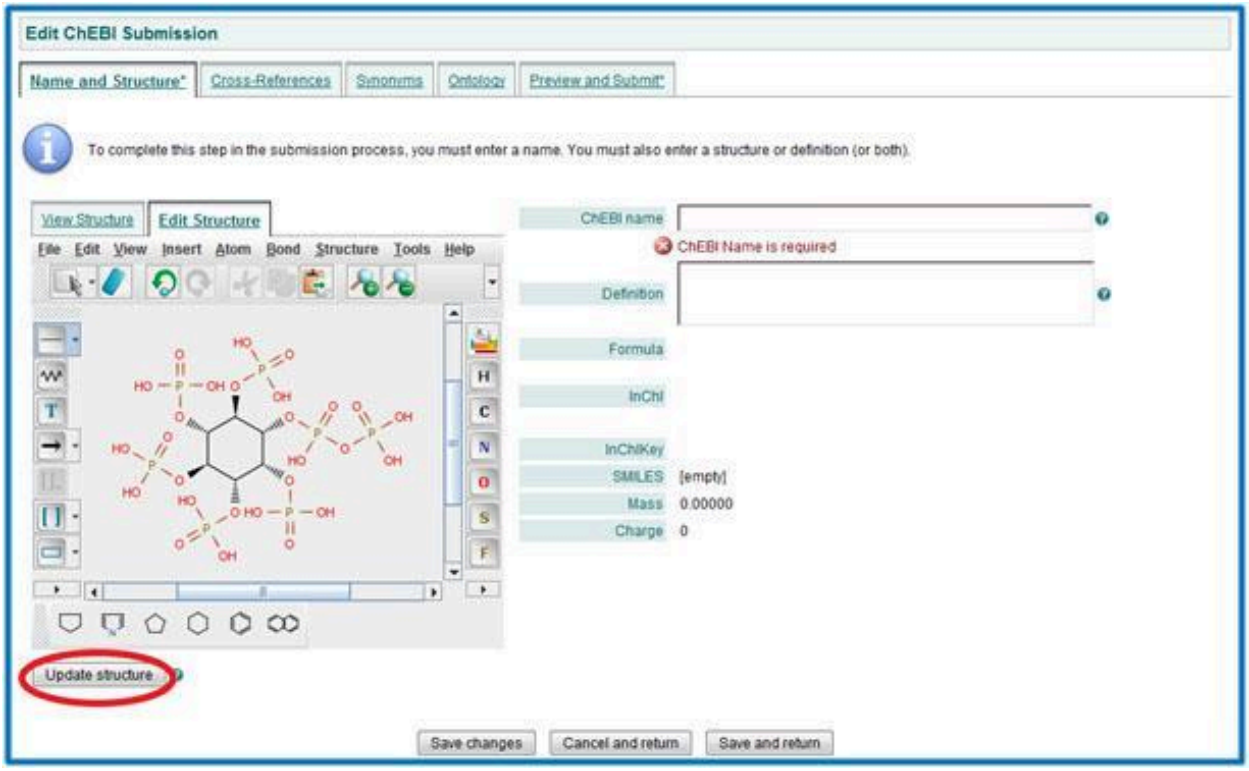
After the SMILES string has been pasted, select the ‘File’ menu in ‘The Source – Name’ input box. Select ‘Import As’. Maybe displace circle or enlarge so as not to hide word “File” in menu



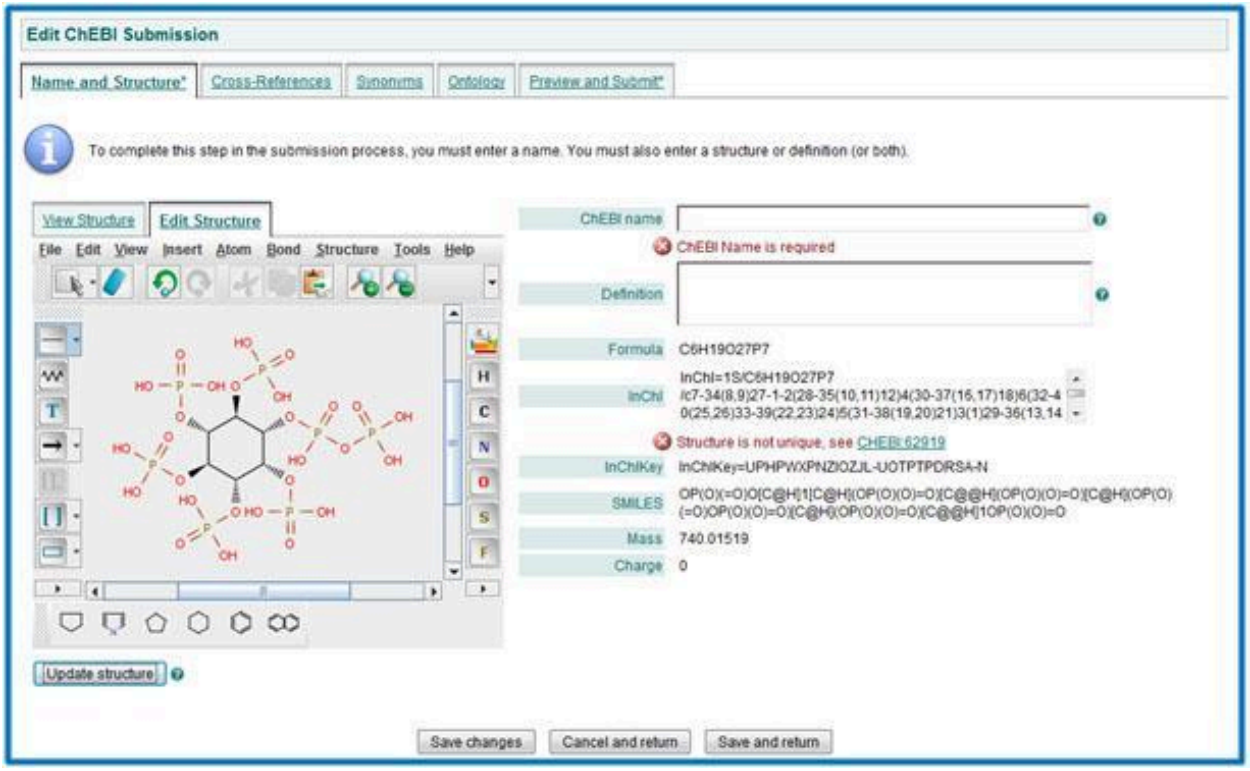
Make sure ‘Import as Recognized (SMILES)’ Import Mode has been selected and click on ‘Import’.



The chemical structure defined by the SMILES string should now be present in the box, ready for editing. Press ‘Update structure’ to obtain details of the structure.



Structure details are now displayed on the right of the page. The structure on the left hand side can now be edited as you wish.



Creating a new Species

If a curator needs to create a new "species" instance for their annotation, please contact the Editor-in-Chief, Peter D'Eustachio as it will need to have an abbreviation assigned by him. the superTaxon slot must* be filled with the "taxon" instance from gk_central **corresponding to the genus of the new species** as specified in the NCBI taxonomy entry for the species. For the bacterial species Bacillus thermoproteolyticus, that genus taxon is Bacillus, so that instance should be imported into the local project and used to fill the superTaxon slot for the new species. If the genus taxon does not exist, then the curator should create it. It too has a superTaxon slot which must be filled with the taxon corresponding to the appropriate family as given in the NCBI taxononmy. This process continues to the root of all life. Note that taxonomists use more levels than the simple kingdom - phylum - class - family - genus - species list. All of the levels shown in NCBI should be instantiated in gk_central for a species.

Annotating the regulation of a process

The following organization of regulation events works well for many kinds of processes, and it fits with our view that all parts of a process should be grouped while respecting GO's view that regulatory events should be distinguishable from the rest of the process.

All about [process] (pathway) --The steps of [process] (pathway)

[process]reaction 1

[process] reaction 2

etc.

--Regulation of [process] (pathway)

[process] regulatory reaction 1

[process regulatory reaction 2

etc.

Regulation here can include reactions that are themselves concrete molecular transformations whose effect is to modulate one of the main process reactions by activating an enzyme, or providing or sequestering an input molecule but can also include airy things like "[this regulatory event] by an unknown molecular mechanism positively or negatively regulates [process reaction #] . Note that regulation instances contain a "regulatedEntity" and a "regulator". Note that we should not be annotating the regulation of a Pathway using regulation instances. We should show the RLE within the pathway that is being regulated. The reason for this is that we want to be as specific as possible about the molecular mechanism of the regulation. Thus pathways should not be listed as the regulatedEntity in Regulation instances.

Annotating Gene Expression:

As far as gene expression regulation is concerned, note that:

- Even if the regulation is at the transcriptional level, show the protein as the output, since we are a protein-oriented database, and we assume that mRNA will get translated into a protein.
- Only show mRNA as the output of the transcription BBE only if this mRNA can be used as an input for another reaction – mainly regulation by microRNAs.
-

A detailed description of how to annotate regulation of gene expression is described below.

NOTE: *When considering which type of evidence to use when creating events in Reactome, note that you should NOT include events based solely on microarray, proteomics or similar high throughput data, unless either 1) an expert author has indicated that they believe the results are correct and represent the current consensus, or 2) The results are a meta-data analysis that combines several independent studies to show the overlap between them.*

Regulation where direct binding of a transcription factor to a regulatory sequence of a target gene has been demonstrated.

At least two of the following three criteria need to be met to create gene expression events and regulatory instances of this type:

1. Presence of specific transcription factor binding elements in regulatory regions of the target gene and/or demonstration of physical association of transcription factor with target gene DNA via chromatin immunoprecipitation (ChIP) or electrophoretic mobility shift assay (EMSA)
2. Evidence that the amount of a transcription factor (altered by overexpression from a recombinant vector and/or downregulation by gene deletion, targeted mutation and/or RNA interference) correlates with the level of target gene expression as measured by luciferase reporter assays and/or quantitative RT-PCR
3. Evidence that disruptive mutations of the transcription factor DNA binding sites in regulatory regions of a target gene affect target gene expression.

For this type of regulation we annotate the binding reaction between transcription factor and target gene. In the following example, the transcription factor is the MTF1 dimer bound to zinc ions (MTF1 dimer:12Zn2+). This transcription factor binds to the CSRP1 gene:

Reaction Properties

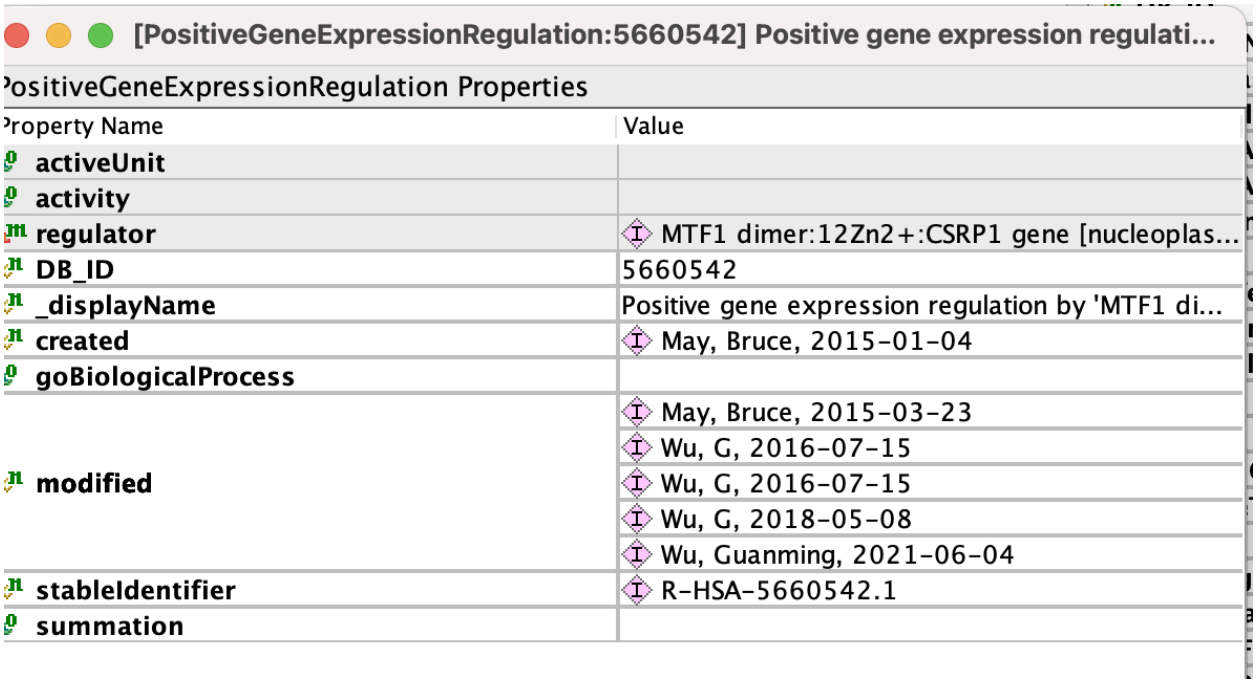
Property Name	Value
catalystActivity	
entityFunctionalStatus	
input	MTF1 dimer:12Zn2+ [nucleoplasm] CSRP1 gene [nucleoplasm]
output	MTF1 dimer:12Zn2+:CSRP1 gene [nucleoplasm]
DB_ID	5660514
_displayName	MTF1 dimer:12Zn2+ binds CSRP1 gene
_doRelease	true
authored	May, Bruce, 2014-12-27
compartment	nucleoplasm
created	May, Bruce, 2015-01-04
crossReference	
definition	
disease	
edited	May, Bruce, 2014-12-27
entityOnOtherCell	
evidenceType	
figure	
goBiologicalProcess	
inferredFrom	Mtf1 dimer:12Zn2+ binds Csrp1 gene
isChimeric	false
literatureReference	
modified	May, Bruce, 2015-03-15 Wu, G, 2016-07-15 May, Bruce, 2017-01-27
name	MTF1 dimer:12Zn2+ binds CSRP1 gene
normalReaction	
orthologousEvent	
precedingEvent	MTF1 dimer:12Zn2+ translocates from cytosol to n...
relatedSpecies	
releaseDate	2017-03-27
releaseStatus	
requiredInputComponent	
reverseReaction	
reviewed	Ford, Dianne, 2017-01-27
revised	
species	Homo sapiens
stableIdentifier	R-HSA-5660514.1
summation	As inferred from the mouse homolog, the promote...
systematicName	

This transcription factor to DNA binding reaction is subsequently defined as a preceding event for the gene expression black box event, where input is a gene and the output is either a protein or an mRNA. In the following example, the input is the CSRP1 gene, and the output is protein CSRP1. Please note the preceding event:

BlackBoxEvent Properties

Property Name	Value
deletedInstanceDB_ID	
entityFunctionalStatus	
input	1 x CSRP1 gene [nucleoplasm]
output	1 x CSRP1 [nucleoplasm]
DB_ID	5660503
_displayName	Expression of CSRP1
_doRelease	true
authored	May, Bruce, 2014-12-27
catalystActivity	
catalystActivityReference	
compartment	nucleoplasm
created	May, Bruce, 2015-01-04
crossReference	
definition	
deletedStableIdentifier	
disease	
edited	May, Bruce, 2014-12-27
entityOnOtherCell	
evidenceType	
figure	
goBiologicalProcess	
hasInteraction	
inferredFrom	Expression of Csrp1
internalReviewed	
isChimeric	false
literatureReference	Human cysteine-rich protein. A memb...
	Zinc sensing by metal-responsive tran...
modified	May, Bruce, 2015-03-15
	Wu, G, 2016-07-15
	May, Bruce, 2017-01-07
	May, Bruce, 2017-01-27
	Wu, G, 2018-04-19
	May, Bruce, 2018-10-31
	Wu, G, 2019-04-04
	Wu, Guanming, 2021-06-04
	Matthews, Lisa, 2023-03-08
name	Expression of CSRP1
negativePrecedingEvent	
normalReaction	
orthologousEvent	
precedingEvent	MTF1 dimer:12Zn2+ binds CSRP1 gene
previousReviewStatus	
reactionType	
regulatedBy	Positive gene expression regulation by...
regulationReference	Positive gene expression regulation by...
relatedSpecies	

Regulation, in cases where regulator is a complex of transcription factor and target gene, is annotated as a **PositiveGeneExpressionRegulation' *or a* NegativeGeneExpressionRegulation**:

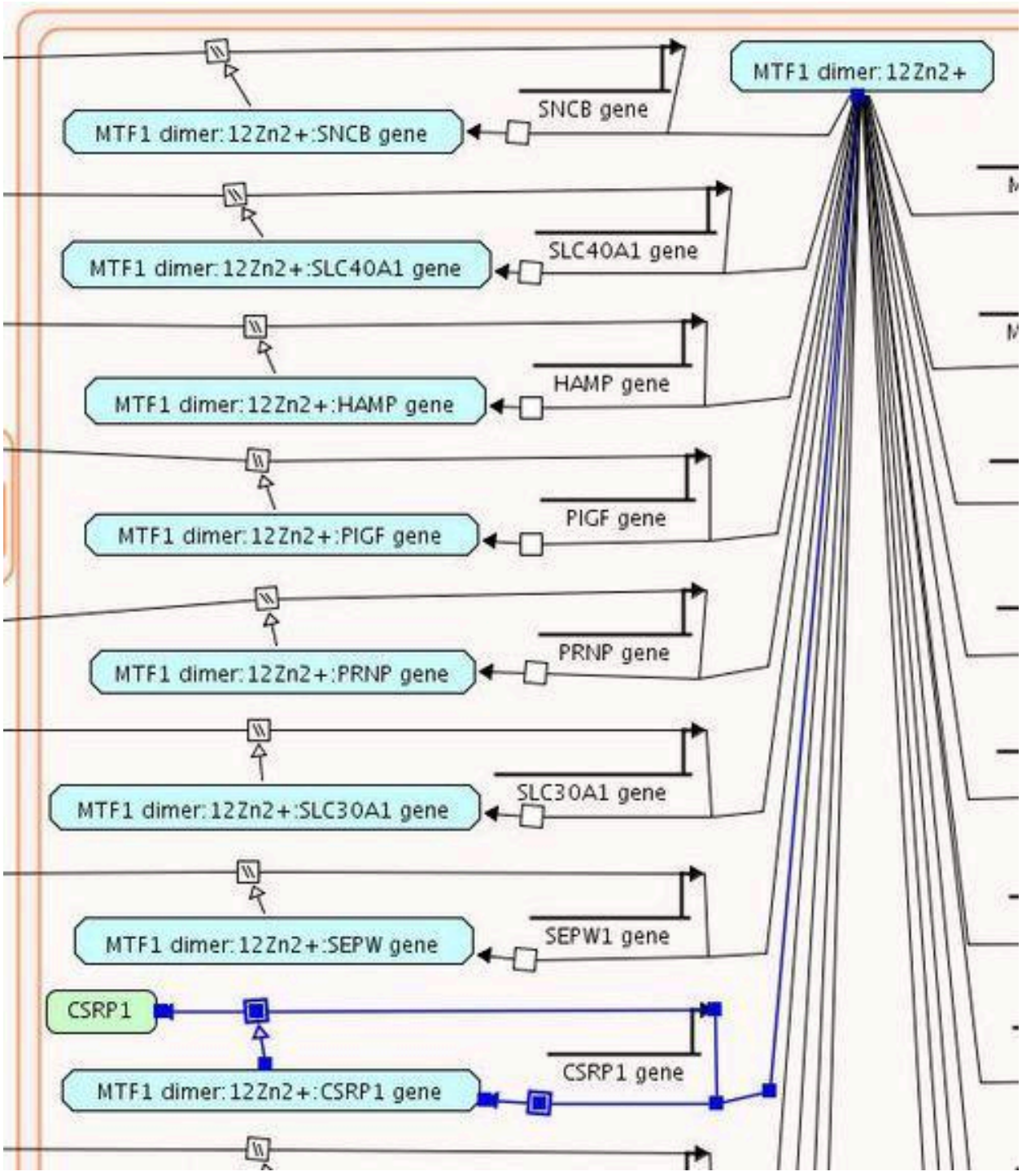


[PositiveGeneExpressionRegulation:5660542] Positive gene expression regulati...

PositiveGeneExpressionRegulation Properties

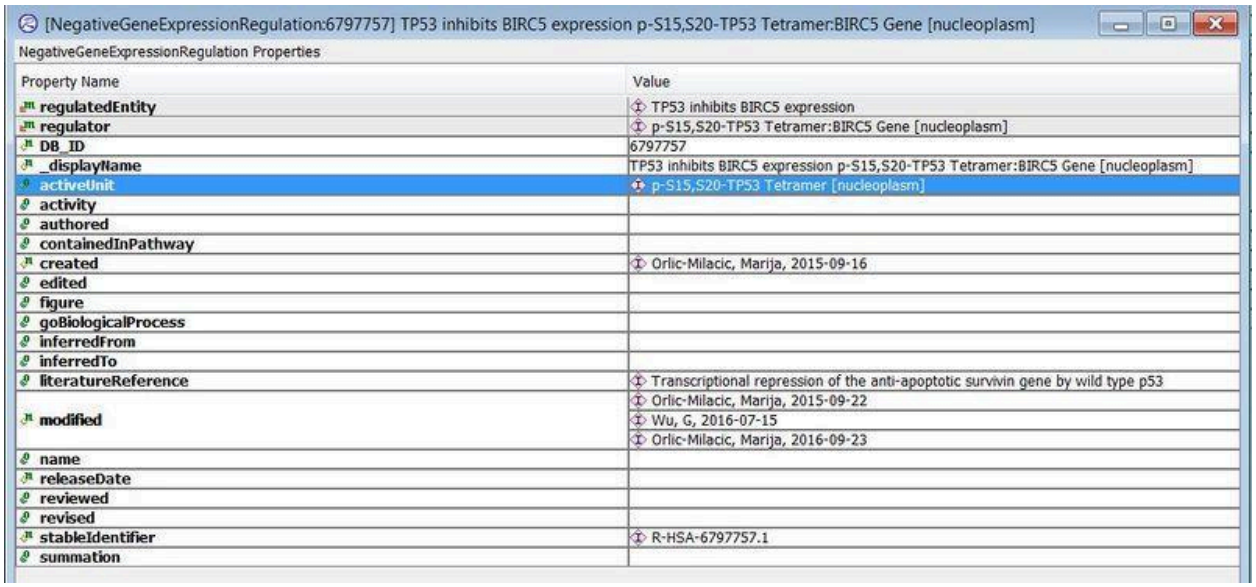
Property Name	Value
activeUnit	
activity	
regulator	MTF1 dimer:12Zn2+:CSRP1 gene [nucleoplas...
DB_ID	5660542
_displayName	Positive gene expression regulation by 'MTF1 di...
created	May, Bruce, 2015-01-04
goBiologicalProcess	
modified	May, Bruce, 2015-03-23
	Wu, G, 2016-07-15
	Wu, G, 2016-07-15
	Wu, G, 2018-05-08
	Wu, Guanming, 2021-06-04
stableIdentifier	R-HSA-5660542.1
summation	

This is how this two-step CSRP1 gene expression regulation (transcription factor binding to DNA, followed by gene expression black box event) looks like in a pathway diagram:



From a computational modelling standpoint, when a regulator is a complex of a transcription factor and its target gene, it is important to define the **activeUnit of the regulation instance**. For example, TP53 negatively regulates BIRC5 gene expression by binding to the BIRC5 gene promoter. The regulator of the NegativeGeneExpressionRegulation instance is the complex of TP53 and BIRC5 gene (p-S15,S20-TP53 Tetramer:BIRC5 gene). Since BIRC5 gene is both and input and a

component of the negative regulator of the gene expression black box event, computational modelling tools may predict that increased copy number of the BIRC5 gene results in a decreased amount of BIRC5 protein. If TP53 is annotated as the activeUnit of the negativeGeneExpressionRegulation instance, this computational problem is prevented:



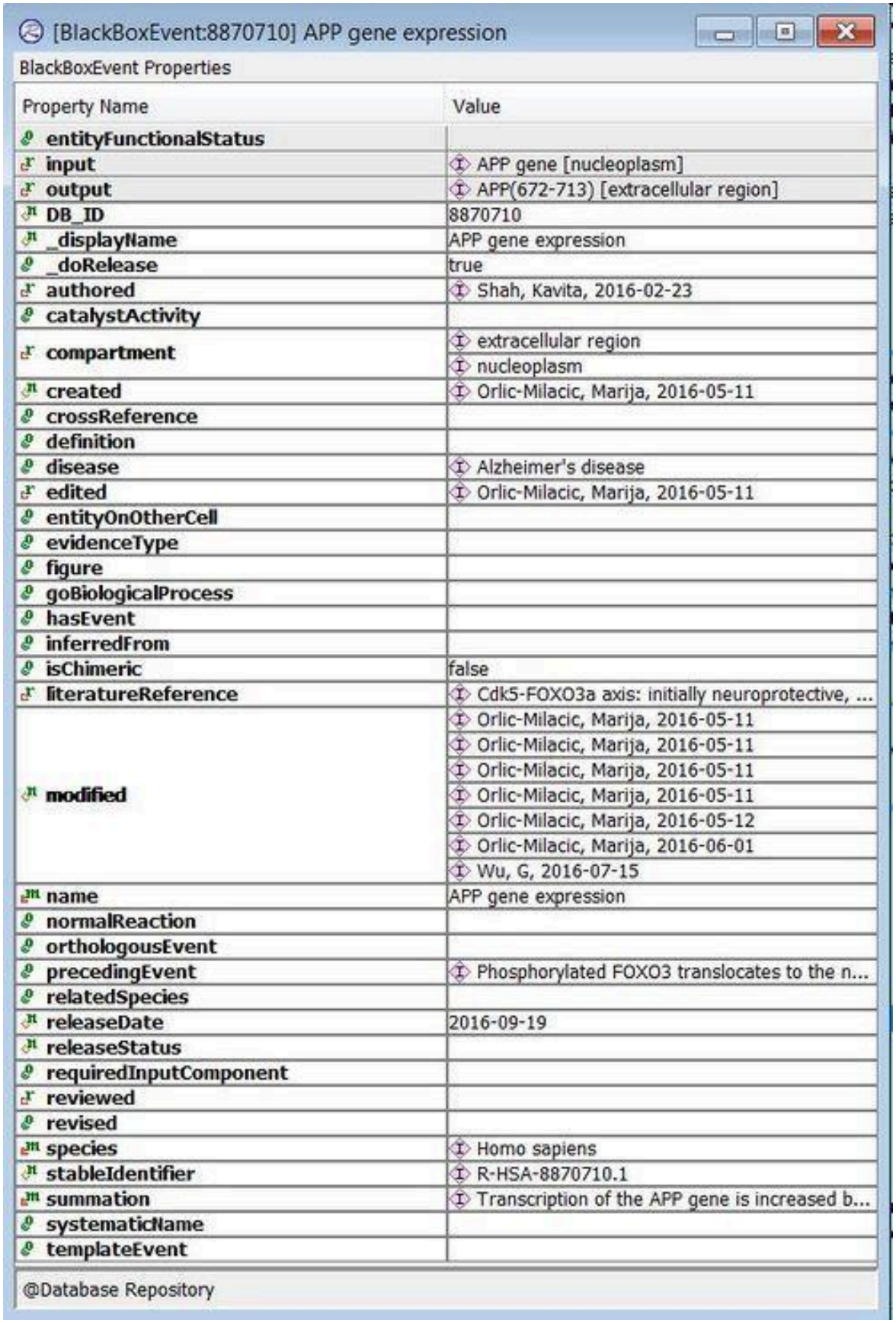
If a translational regulation will be shown, e.g. via noncoding RNAs, then mRNA is annotated as the output of a gene expression black box event. If it is anticipated that a translational regulation will not be made presently but may be added in future releases, then a protein should be annotated as the output of a gene expression black box event that will act as a releasable "placeholder". In a later release, separate transcription and translation black box events will be made and substituted for the "placeholder" black box event. The "placeholder" black box event should be deleted with a deletion instance listing the new transcription and translation black box events as replacements.

Regulation where no direct binding of a regulator to a regulatory DNA sequence has been demonstrated.

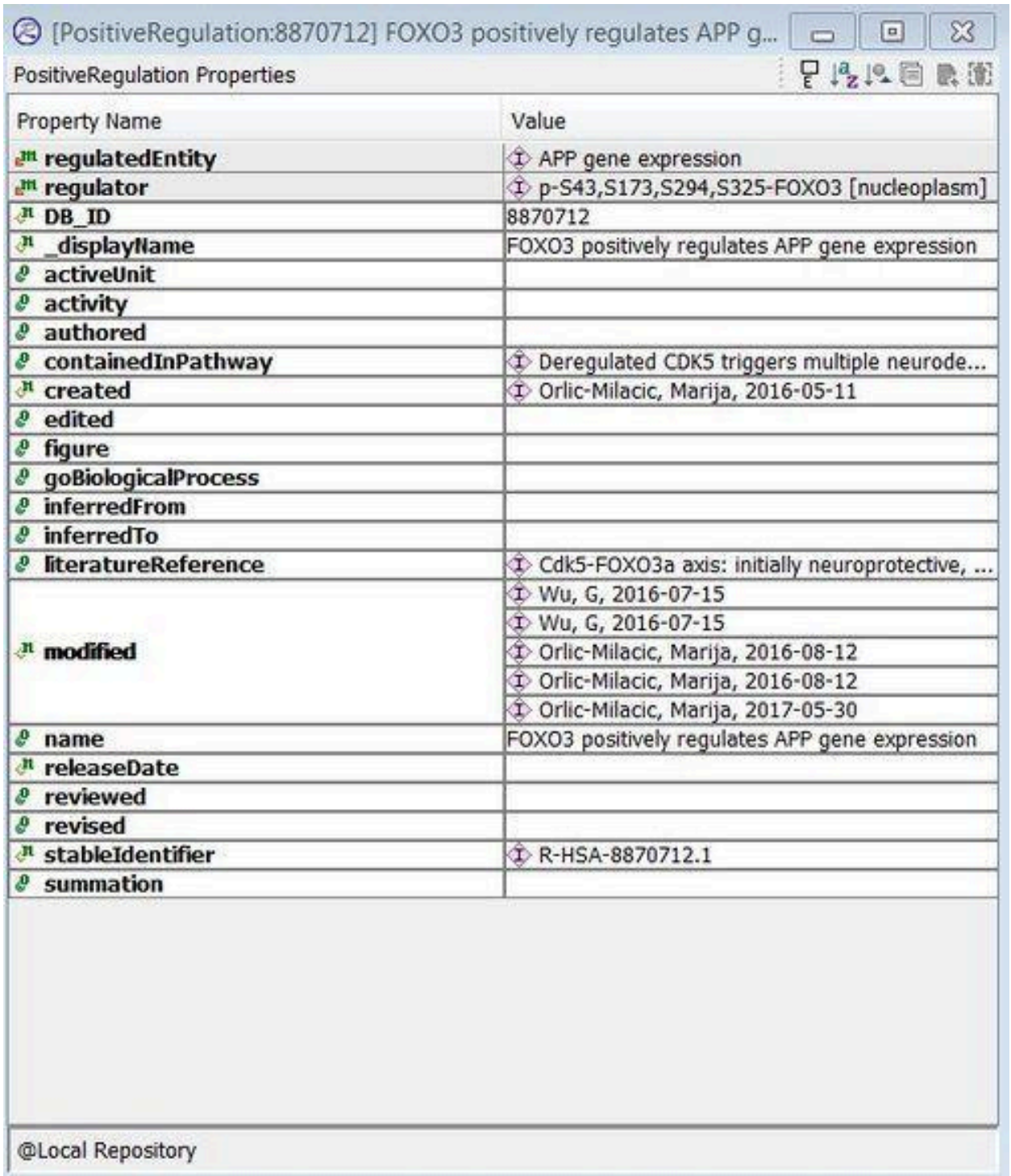
For this type of annotation, one criterion needs to be met:

4. Evidence that the amount and/or the activity of a regulator (altered by overexpression from a recombinant vector and/or downregulation by gene deletion, targeted mutation, RNA interference and/or agonists/antagonists) correlates with the level of target gene expression measured by luciferase reporter assays and/or quantitative RT-PCR.

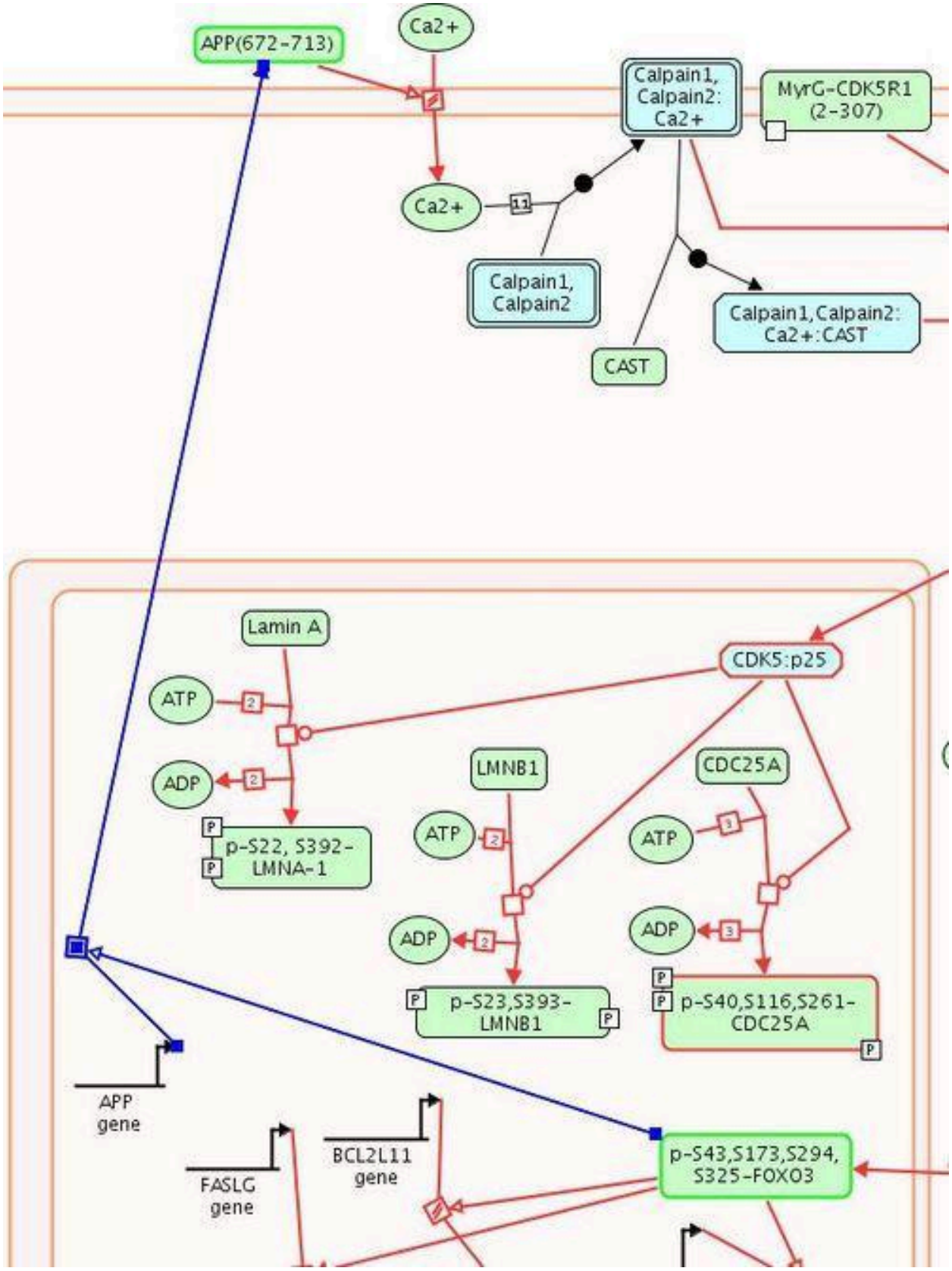
In this case, just a black box gene expression event is created. For example, FOXO3 positively regulates transcription of APP, but no direct binding of FOXO3 to the APP gene locus has been demonstrated:


[BlackBoxEvent:8870710] APP gene expression
BlackBoxEvent Properties
Property Name
Value
entityFunctionalStatus
input
APP gene [nucleoplasm]
output
APP(672-713) [extracellular region]
DB_ID
8870710
_displayName
APP gene expression
_doRelease
true
authored
Shah, Kavita, 2016-02-23
catalystActivity
compartment
extracellular region
nucleoplasm
created
Orlic-Milacic, Marija, 2016-05-11
crossReference
definition
disease
Alzheimer's disease
edited
Orlic-Milacic, Marija, 2016-05-11
entityOnOtherCell
evidenceType
figure
goBiologicalProcess
hasEvent
inferredFrom
isChimeric
false
literatureReference
Cdk5-FOXO3a axis: initially neuroprotective, ...
modified
Orlic-Milacic, Marija, 2016-05-11
Orlic-Milacic, Marija, 2016-05-11
Orlic-Milacic, Marija, 2016-05-11
Orlic-Milacic, Marija, 2016-05-11
Orlic-Milacic, Marija, 2016-05-12
Orlic-Milacic, Marija, 2016-06-01
Wu, G, 2016-07-15
name
APP gene expression
normalReaction
orthologousEvent
precedingEvent
Phosphorylated FOXO3 translocates to the n...
relatedSpecies
releaseDate
2016-09-19
releaseStatus
requiredInputComponent
reviewed
revised
species
Homo sapiens
stableIdentifier
R-HSA-8870710.1
summation
Transcription of the APP gene is increased b...
systematicName
templateEvent
@Database Repository

Regulation instance, when no direct transcription factor binding has been demonstrated, is captured as a general **positiveRegulation** or **negativeRegulation**:



This is how the one-step APP gene expression regulation looks like in a pathway diagram:


APP(672-713)
Ca2+
Calpain1, Calpain2: Ca2+
MyrG-CDK5R1 (2-307)
Ca2+
11
Calpain1, Calpain2
Calpain1,Calpain2: Ca2+:CAST
CAST
Lamin A
CDK5:p25
ATP
2
ADP
2
p-S22, S392-LMNA-1
LMNB1
ATP
2
ADP
2
p-S23,S393-LMNB1
CDC25A
ATP
3
ADP
3
p-S40,S116,S261-CDC25A
APP gene
FASLG gene
BCL2L11 gene
p-S43,S173,S294, S325-FOXO3

Regulation of gene expression at the level of translation

mRNA translation into protein can be regulated by non-coding RNAs, such as microRNAs (miRNAs). We only annotate miRNAs for which direct binding to target mRNAs and regulation of protein expression has been experimentally validated by either:

5. A reporter assay, where the miRNA is combined with a reporter-fused 3'UTR of the mRNA and observed for altered levels of reporter expression;
6. Co-immunoprecipitation of the miRNA:mRNA complex, together with an assay demonstrating that the miRNA alters the levels of either the mRNA (qRT-PCR) or protein (western blot).

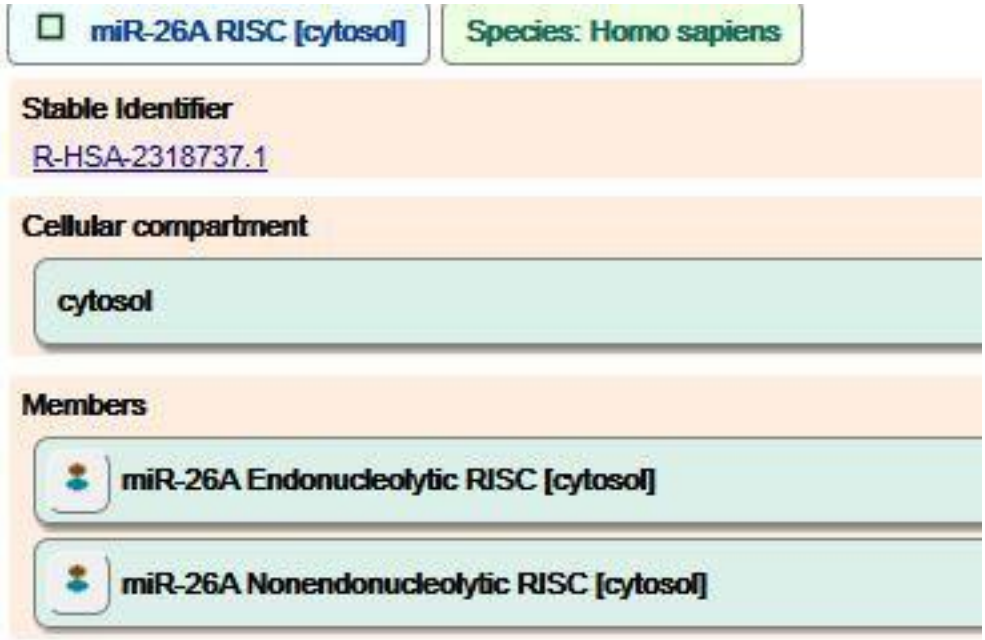
MicroRNAs are always shown to bind to target mRNAs in the context of the RISC complex. Two types of RISC complex exist: nonendonucleolytic RISC (interferes with translation without degrading target mRNA) and endonucleolytic RISC (degrades target mRNA to prevent translation). Composition of different RISCs and their interaction with microRNAs is described in detail in the pathway Gene Silencing by RNA (R-HSA-211000):

- If an miRNA causes reduction of protein levels, but not mRNA levels of a target gene, it is annotated as a component of a nonendonucleolytic RISC.



- If an miRNA causes reduction of both mRNA and protein levels of a target gene, we usually annotate a RISC set where endonucleolytic RISC is a

member and non-endonucleolytic RISC, depending on the experimental evidence, is either a candidate or a member.



If a non-endonucleolytic RISC for an miRNA of interest was definitely experimentally excluded, we would annotate the miRNA as a component of an endonucleolytic RISC only:

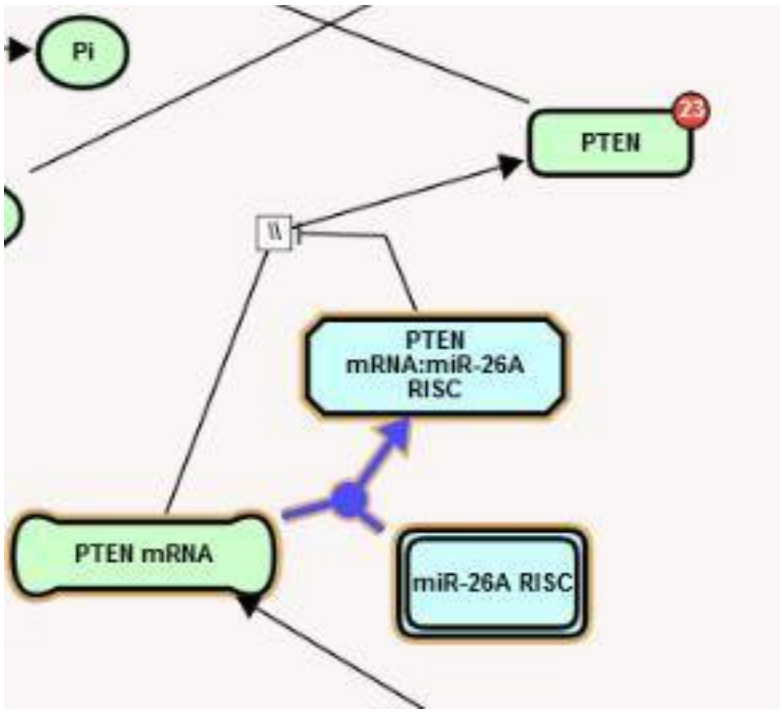


The easiest way to create endonucleolytic RISC and non-endonucleolytic RISC is by cloning an already annotated RISC complex and replacing the miRNA component of the clone by the miRNA of interest:

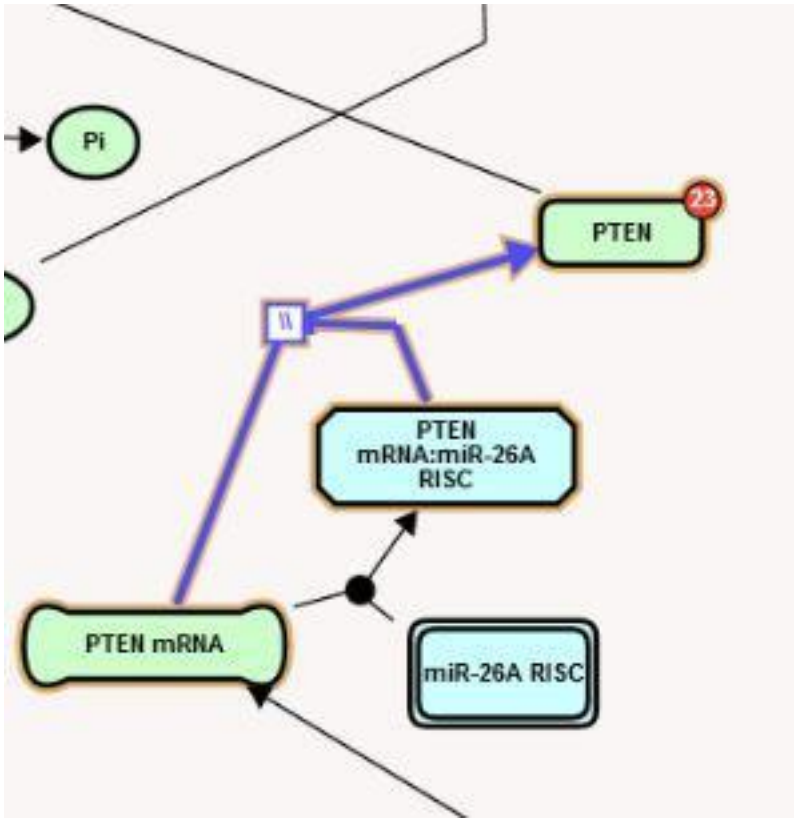
1. As there are over 100 miRNA RISC complexes in gk_central, please make sure that the specific RISC that you are trying to create does not already exist in the central repository.
2. Check out an existing human endonucleolytic and/or non-endonucleolytic RISC to your local project.
3. Note that both endonucleolytic RISC and non-endonucleolytic RISC are nested complexes, in which miRNA is a component of a subcomplex named “minimal endonucleolytic RISC” and “minimal non-endonucleolytic RISC”, respectively. Create separate clones of the minimal endonucleolytic RISC/minimal non-endonucleolytic RISC, and endonucleolytic RISC/non-endonucleolytic RISC, **declining the offer to clone contain subunits recursively.**
4. In the clone(s) of minimal endonucleolytic RISC/minimal non-endonucleolytic RISC, replace the miRNA component with your miRNA of interest. Rename the clone to reflect its new miRNA component.
5. In the clone(s) of endonucleolytic RISC/non-endonucleolytic RISC, replace the minimal endonucleolytic RISC/minimal non-endonucleolytic RISC components with cloned and edited minimal endonucleolytic RISC/minimal non-endonucleolytic RISC. Rename the clone to reflect its new minimal endonucleolytic RISC/minimal non-endonucleolytic RISC component.

Translational regulation of gene expression is annotated in two steps, with step1 defined as a preceding event of step2:

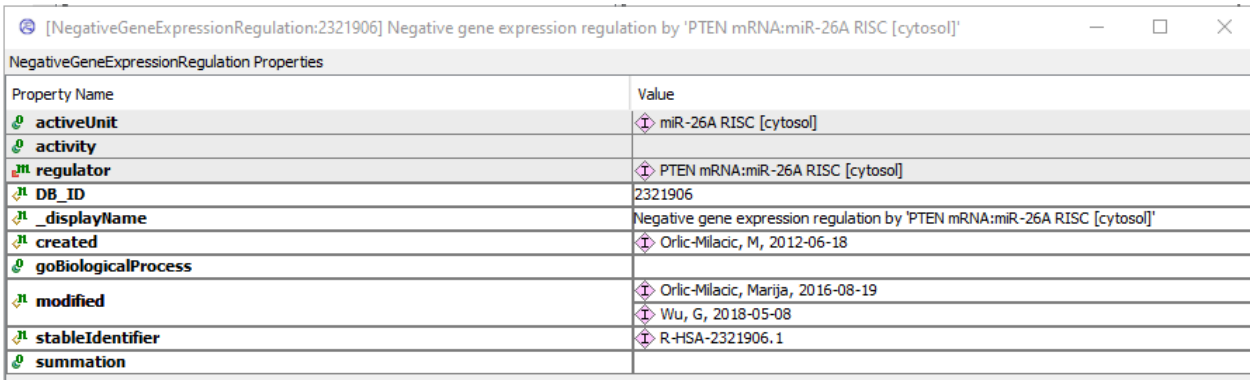
1. Binding of an inhibitor (e.g. miRNA-containing RISC complex) to a target mRNA (**step1**).



2. Gene expression black box event where mRNA is an input and protein is an output (**step2**).



Step2 has regulatedBy attribute populated by a **negativeGeneExpressionRegulation** instance. The output of step1 is an inhibitor:mRNA complex, which is annotated as a negative regulator of step2. Again, from a computational modeling standpoint, when a regulator is a complex of an inhibitor and its target mRNA, it is important to define the **activeUnit of the regulation instance**. For example, miR-26A negatively regulates PTEN mRNA translation by binding to the 3'UTR of PTEN mRNA. The regulator of the NegativeGeneExpressionRegulation instance is the complex of PTEN mRNA and miR-26A RISC (PTEN mRNA:miR-26A RISC). Since PTEN mRNA is both an input and a component of the negative regulator of the translation black box event, computational modeling tools may predict that the increased level of PTEN mRNA results in a decreased amount of PTEN protein. If miR-26A RISC is annotated as the activeUnit of the negativeGeneExpressionRegulation instance, this computational problem is prevented:



Drawing a pathway diagram

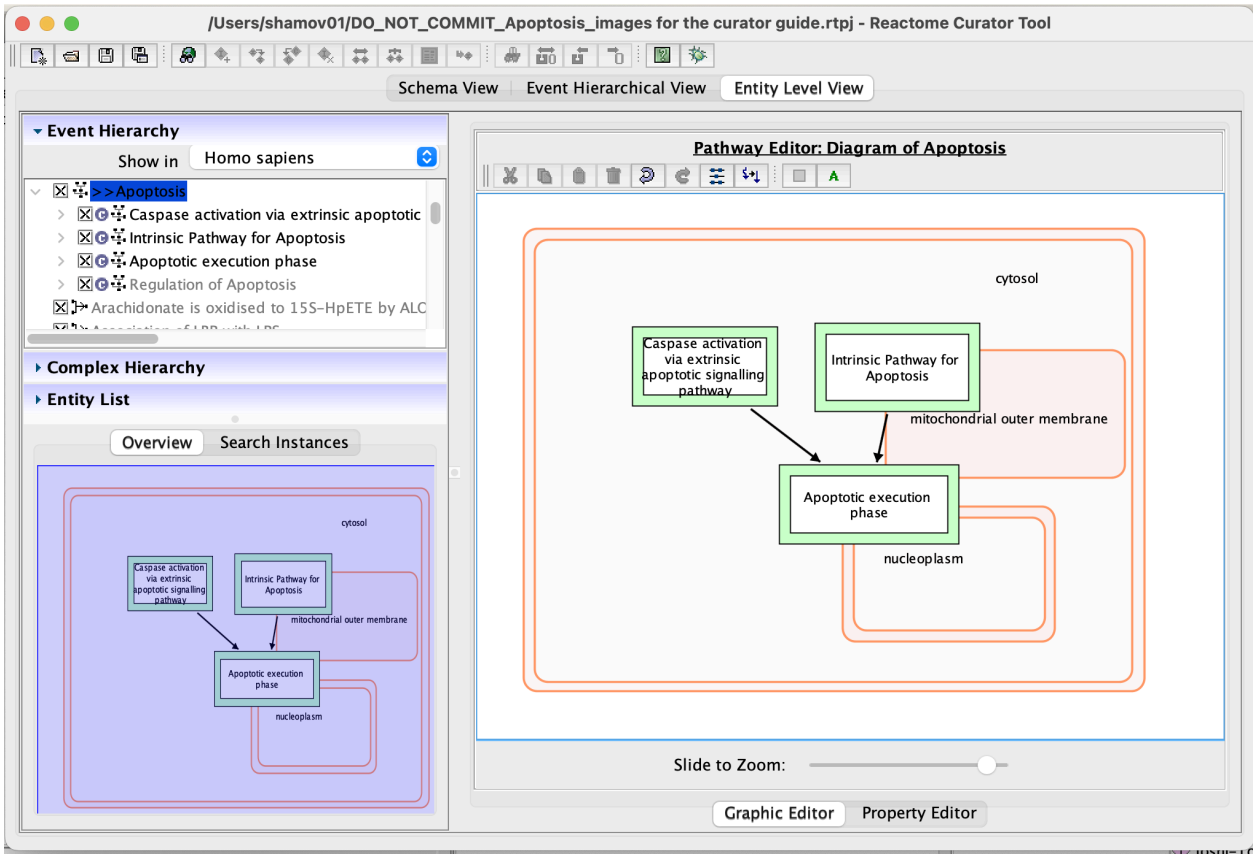
Reactome pathway diagrams are drawn and viewed in the curator tool using the ELV pane of the tool.

Below is an example to show how the superpathway "Regulation of Apoptosis" is diagrammed, followed by important hints about revising existing diagrams.

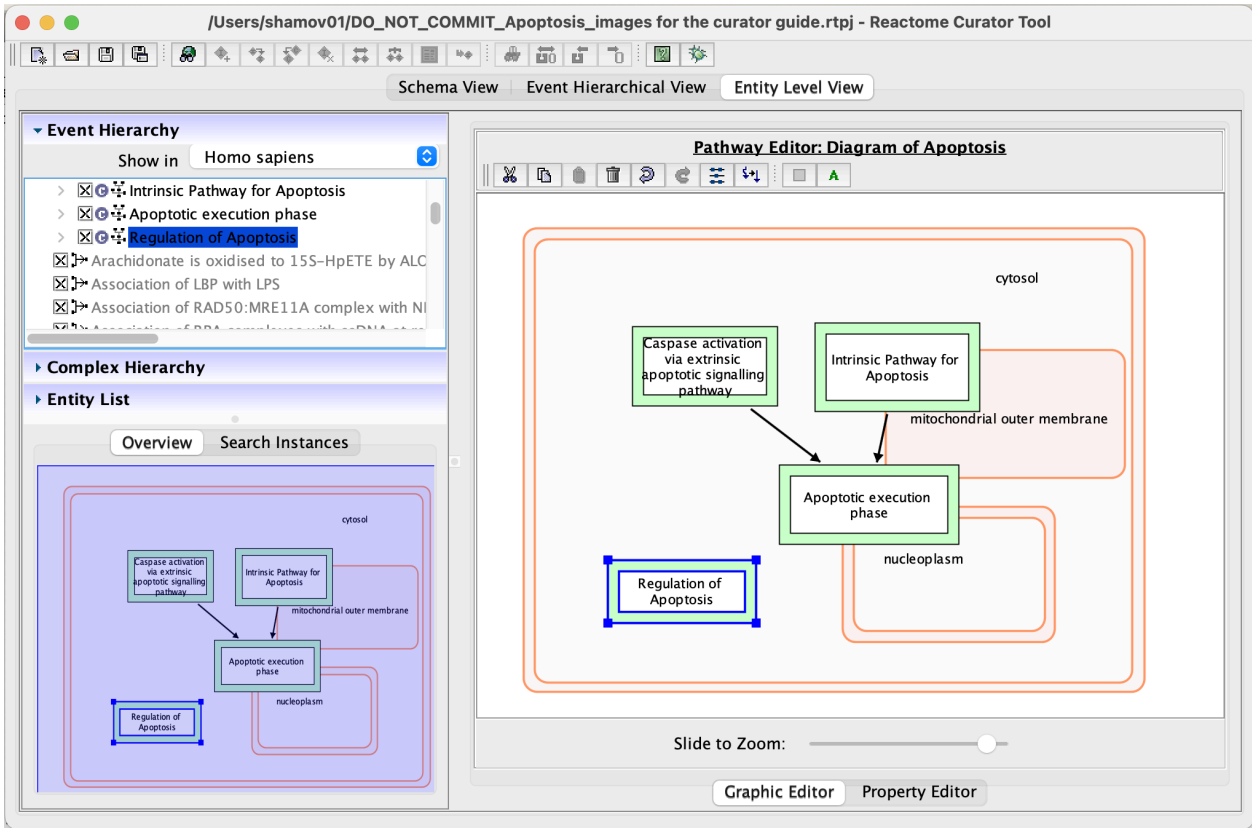
Creating new diagrams

The pathway Regulation of Apoptosis is part of the supercanonical Apoptosis pathway. First, go to gk_central in the Event browser and identify the pathway that you want to diagram. Select this, right click and opt to "build project". In this example, you can tell that Regulation of Apoptosis has been annotated but not yet incorporated in the Apoptosis diagram. You can tell this because the pathway is greyed out in the event hierarchy.

Click on apoptosis and open the diagram.



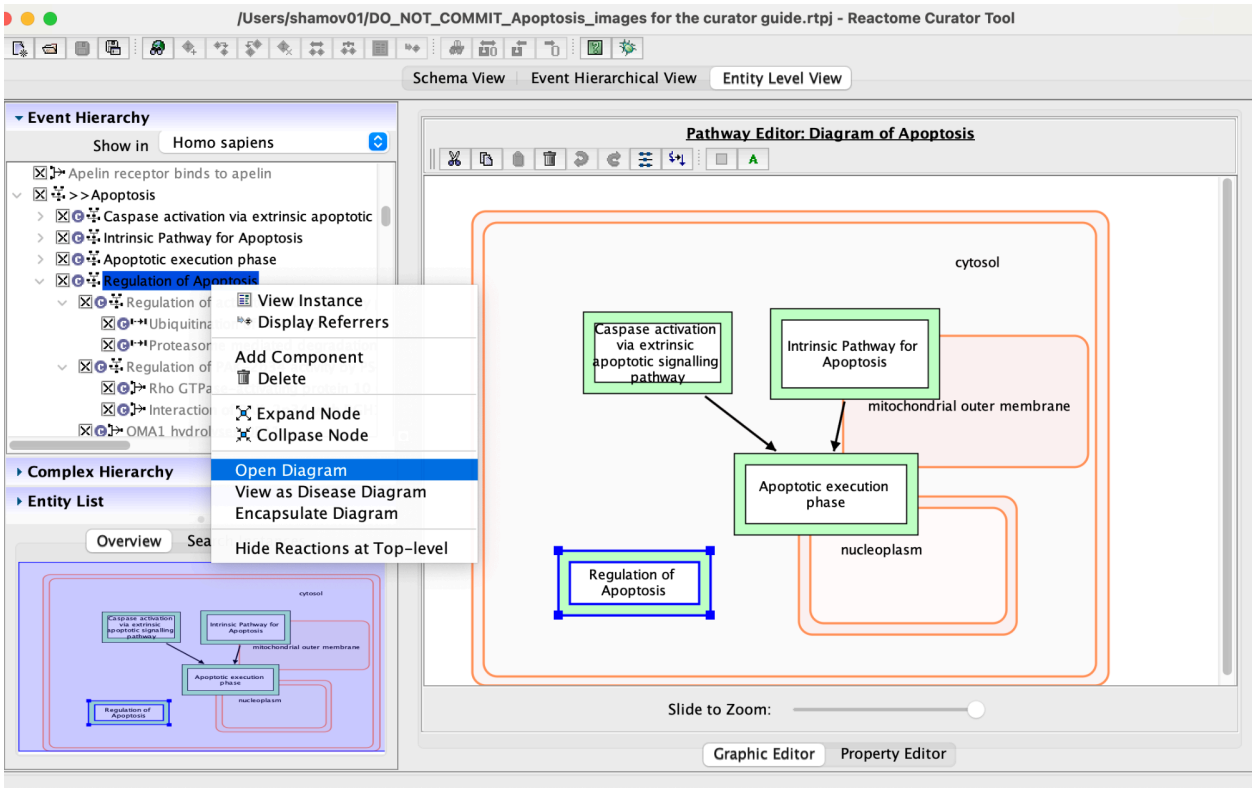
To incorporate the Regulation of Apoptosis pathway into the Apoptosis diagram, simply click on, hold, and drag the pathway from the hierarchy to the diagram.



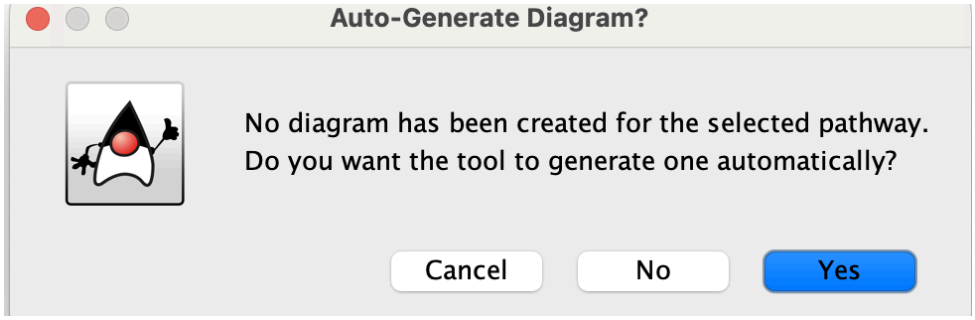
EXAMPLE: CREATE A DIAGRAM FOR REGULATION OF APOPTOSIS

Right click on pathway in the hierarchy or in the diagram and opt to "open diagram"

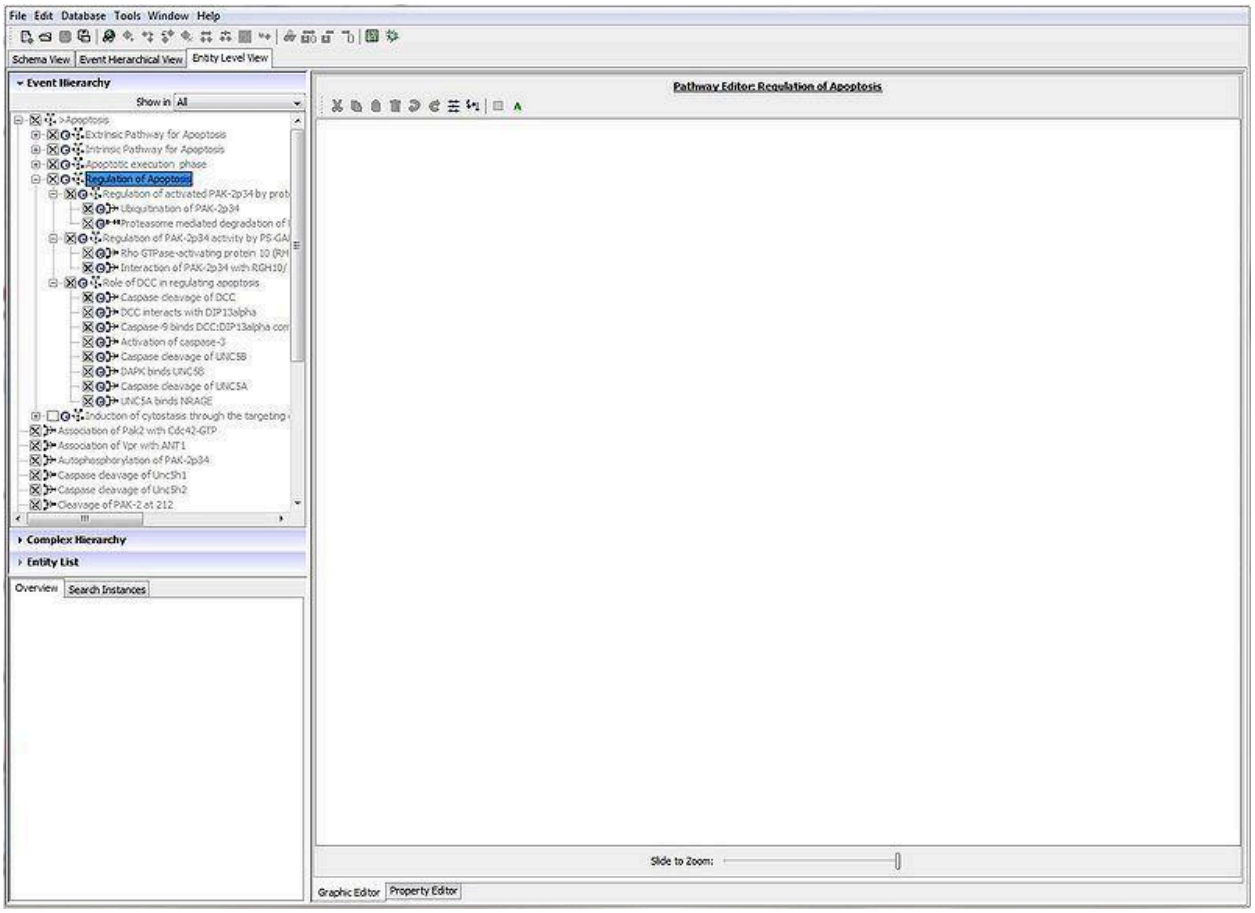
If one has been created it will open.



If it has not, as in the case of Regulation of Apoptosis, you get the below message.

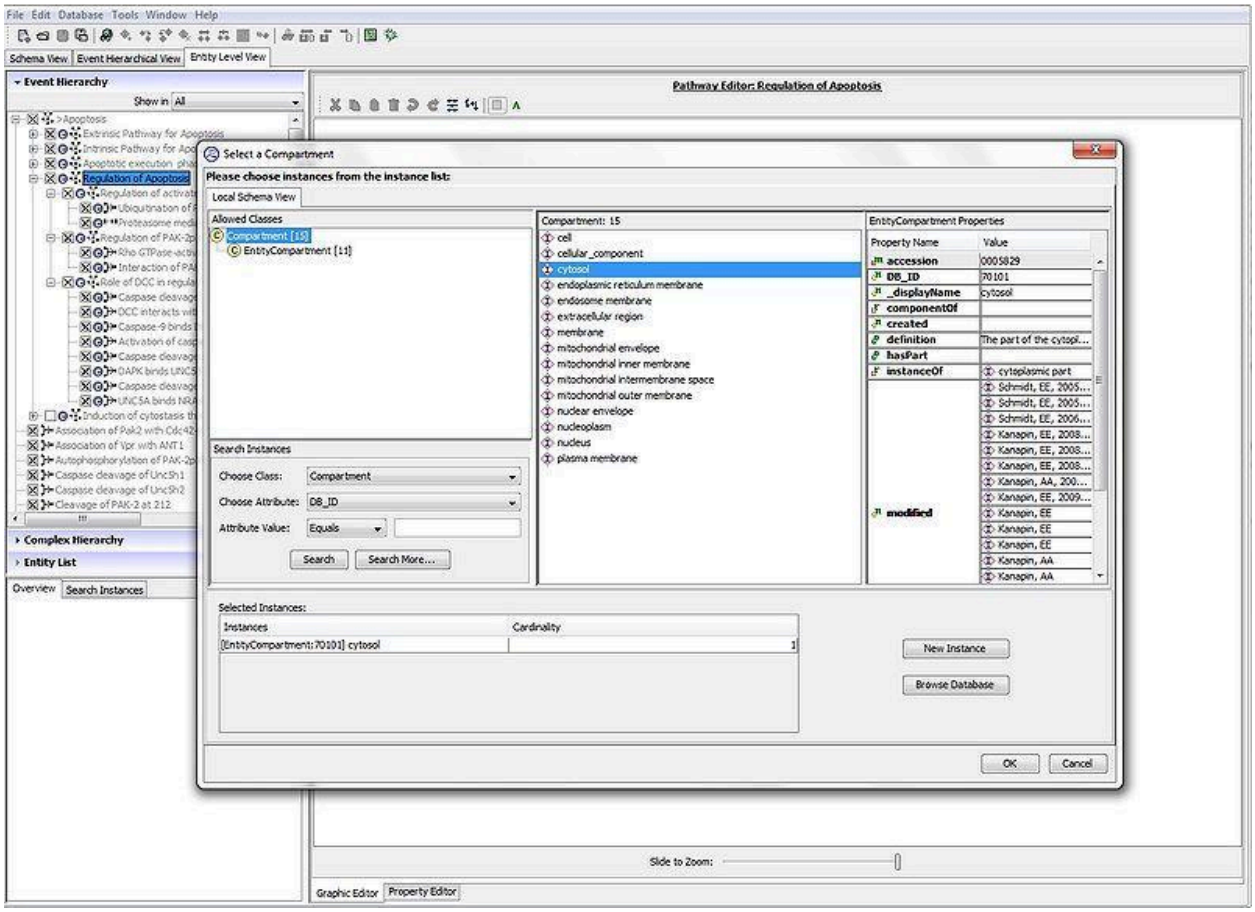


Select "No" and an empty diagram will be opened in the pathway editor pane.

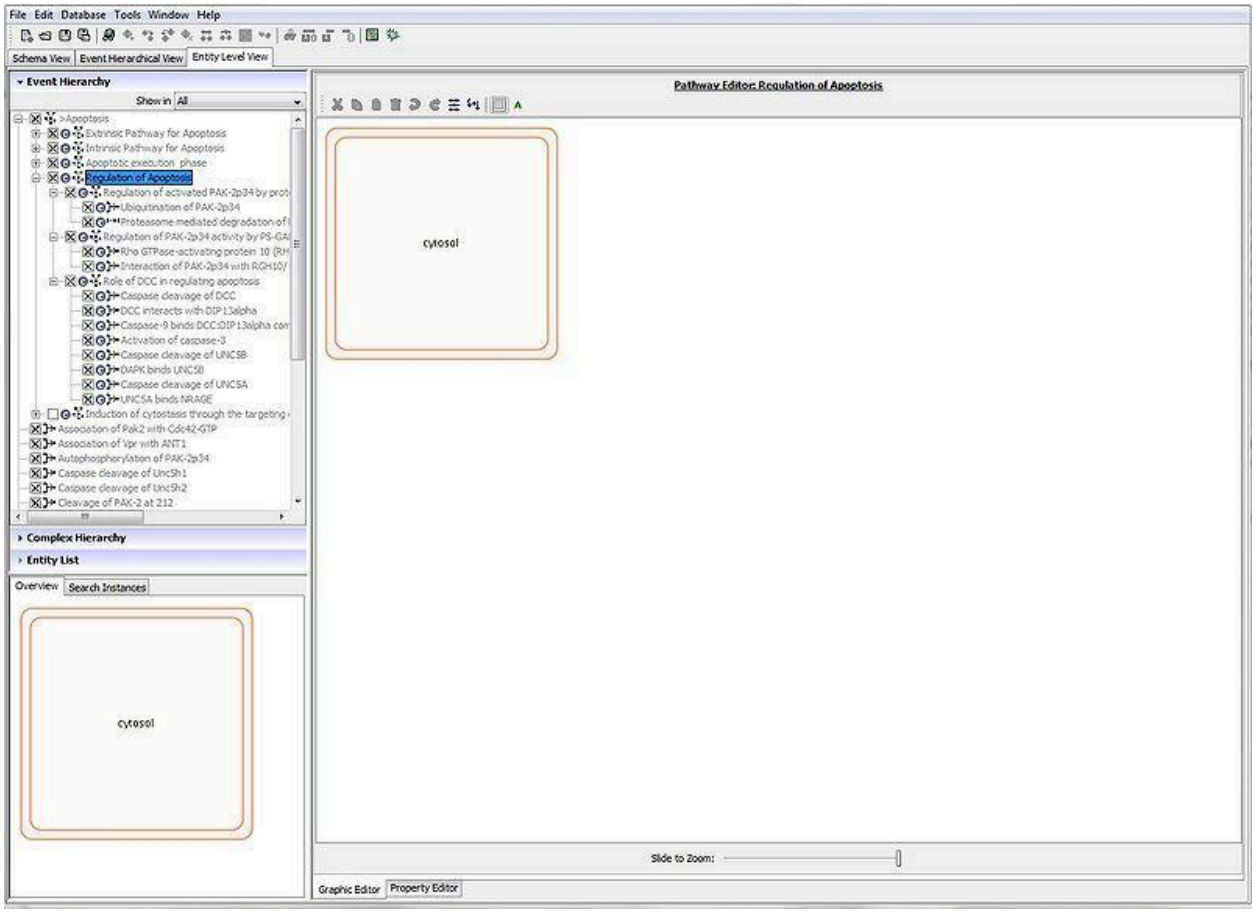


Adding cellular compartments to diagram

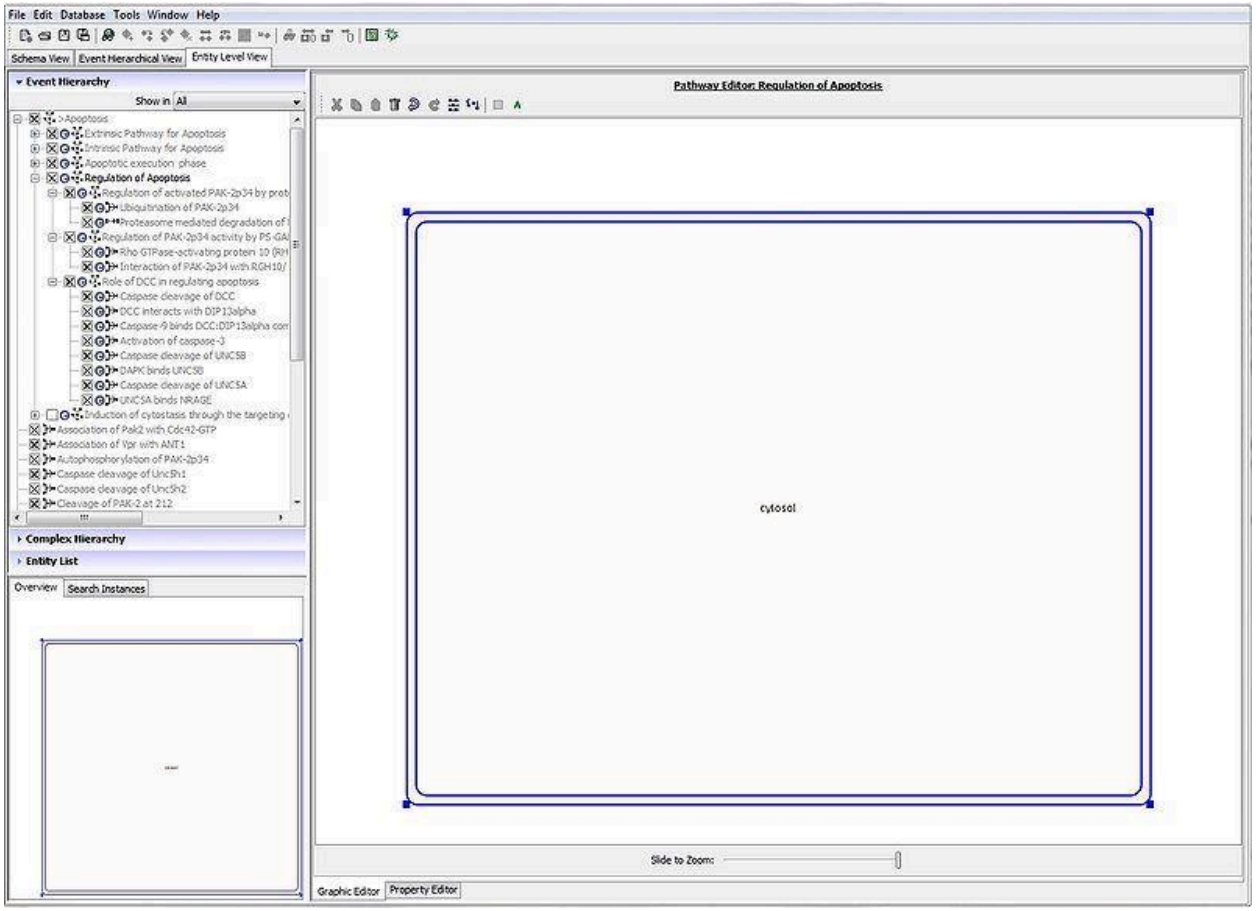
Click on the shaded square in the menu bar for the Pathway editor. It is best to add all the compartments that you will need before you start to lay out reactions. Here cytosol is added.



To reposition the compartment, click on it and drag.

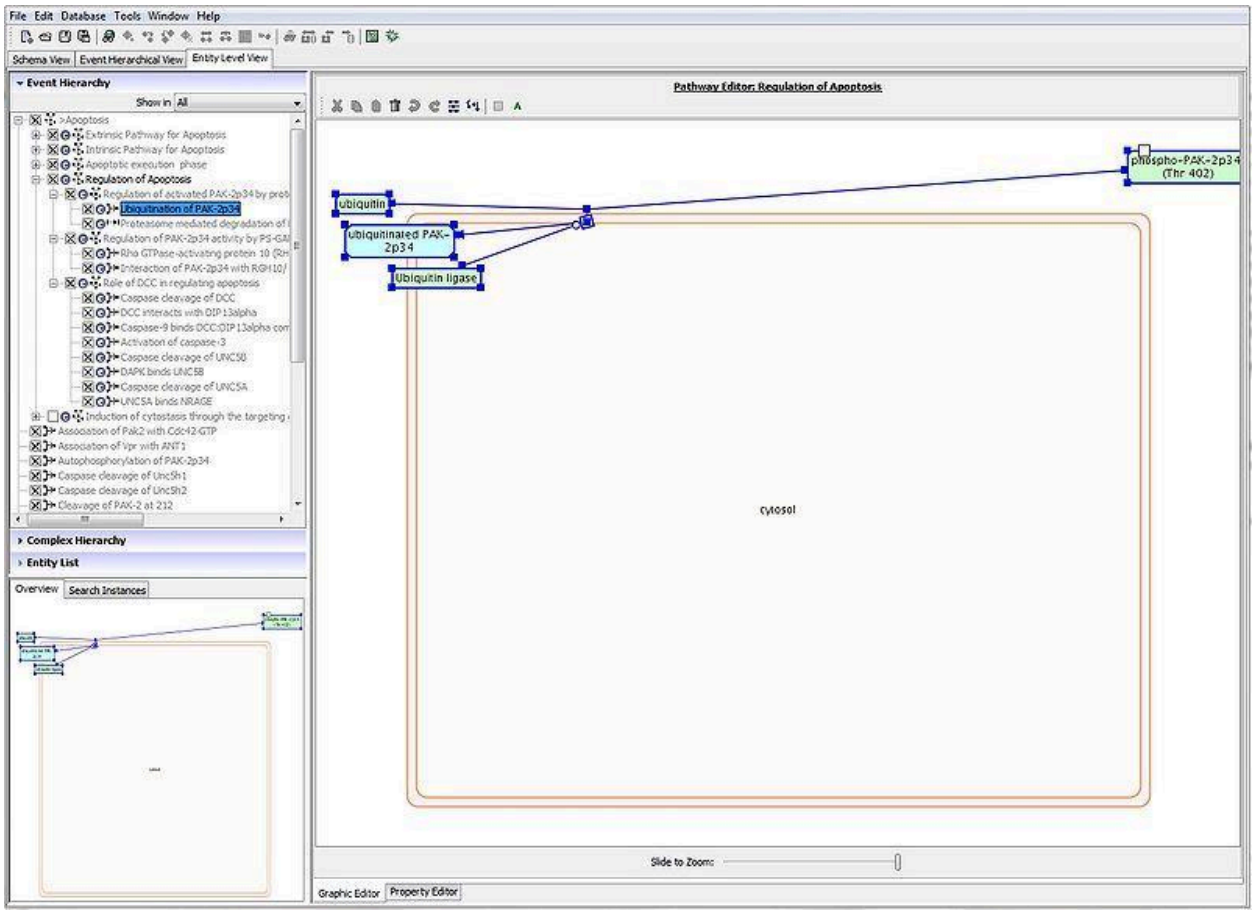


To enlarge the compartment, click on it to select it and then grab the compartment at one of its nodes in the corners. Then drag outward.

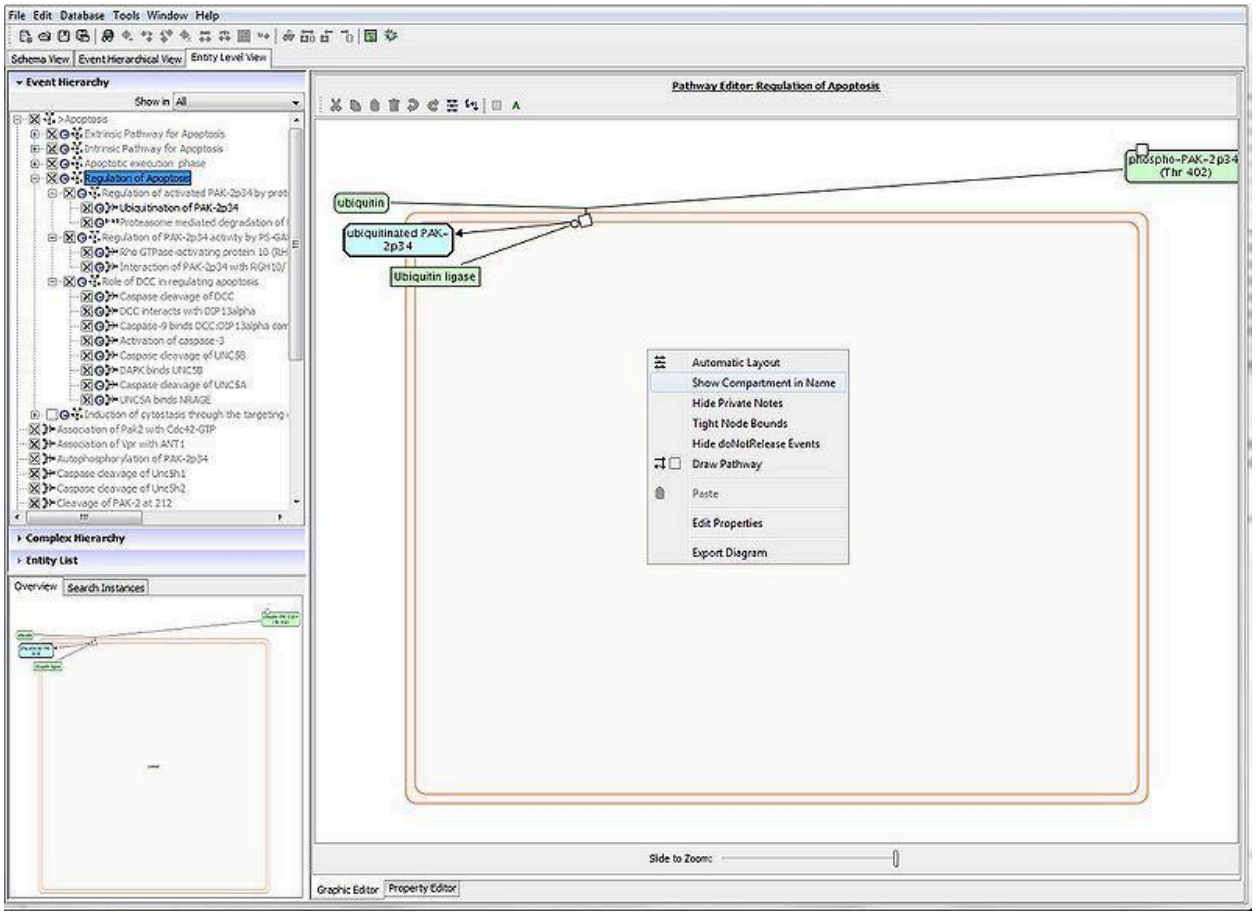


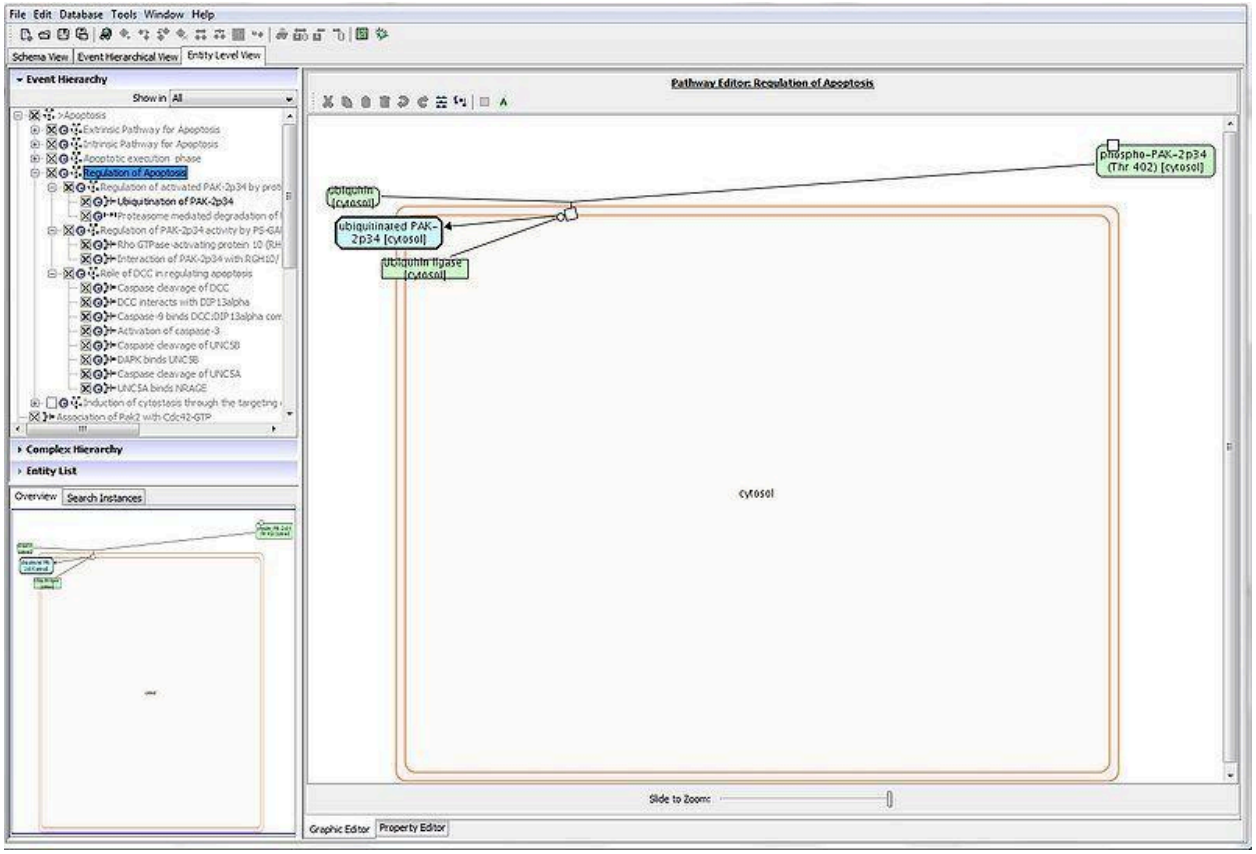
Drawing the pathway

To begin drawing, select a reaction from the event hierarchy and drag it onto the diagram.

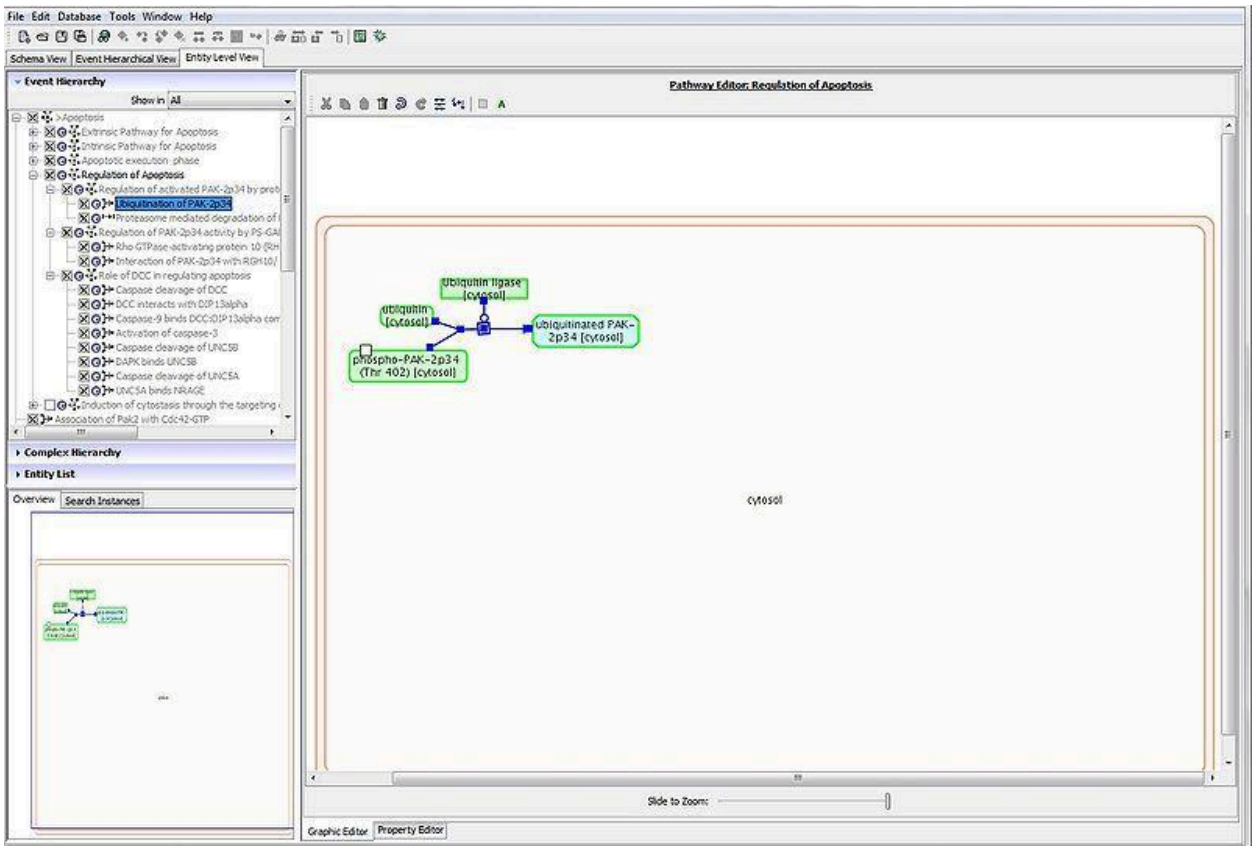


To see the names of the compartments of the reaction participating molecules, right click anywhere on the diagram and opt to show compartment names. This will make it easier to see that the molecules have been positioned in the correct compartment.

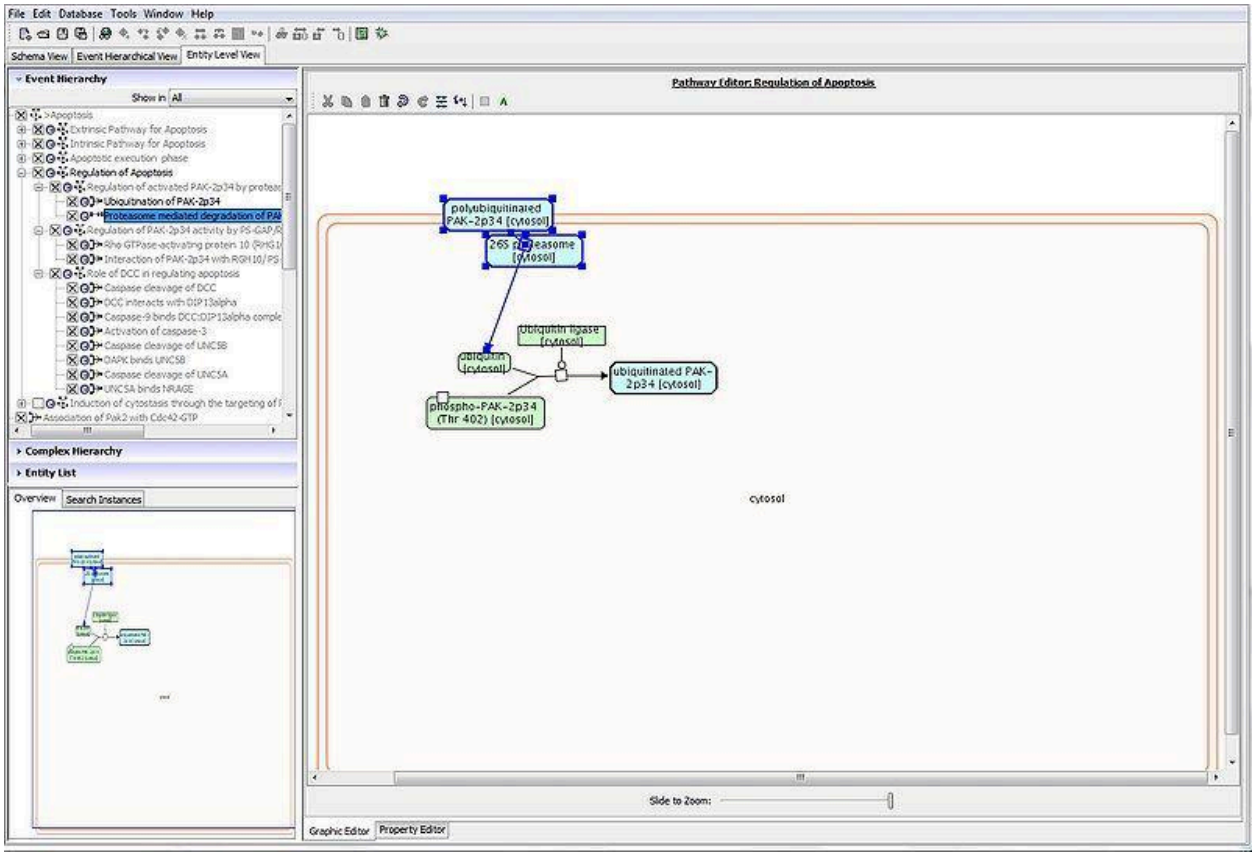

File Edit Database Tools Window Help
Schema View
Event Hierarchical View
Entity Level View
Event Hierarchy
Show in All
Apoptosis
Extrinsic Pathway for Apoptosis
Intrinsic Pathway for Apoptosis
Apoptotic execution phase
Regulation of Apoptosis
Ubiquitination of PAK-2p34
Complex Hierarchy
Entity List
Overview
Search Instances
Pathway Editor: Regulation of Apoptosis
phospho-PAK-2p34 (Thr 402)
ubiquitin
ubiquitinated PAK-2p34
Ubiquitin ligase
Automatic Layout
Show Compartment in Name
Hide Private Notes
Tight Node Bounds
Hide doNotRelease Events
Draw Pathway
Paste
Edit Properties
Export Diagram
Slide to Zoom:
Graphic Editor
Property Editor

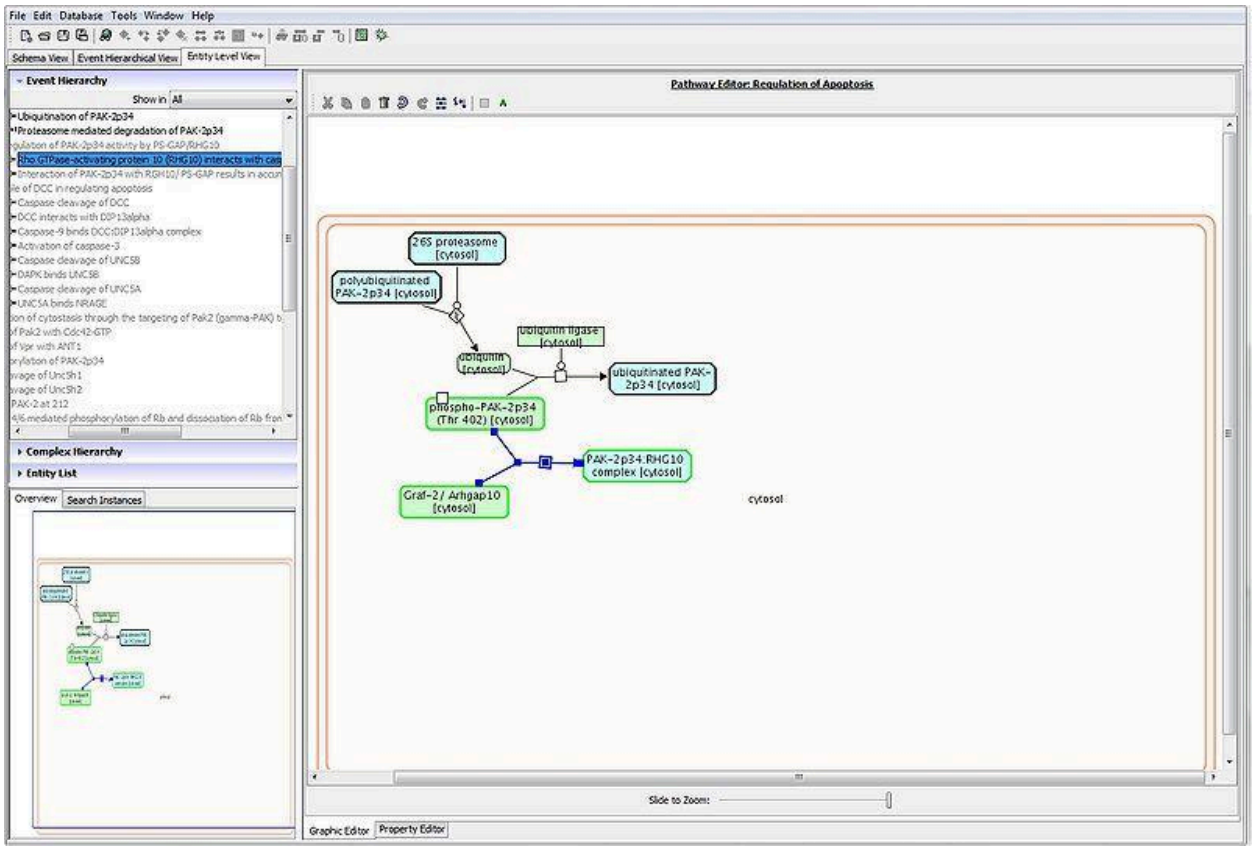


To reposition the reaction you can click and drag different components or you can select the entire reaction by clicking and dragging a selection box over it. Then the reaction and all its component molecules can be moved as a unit.



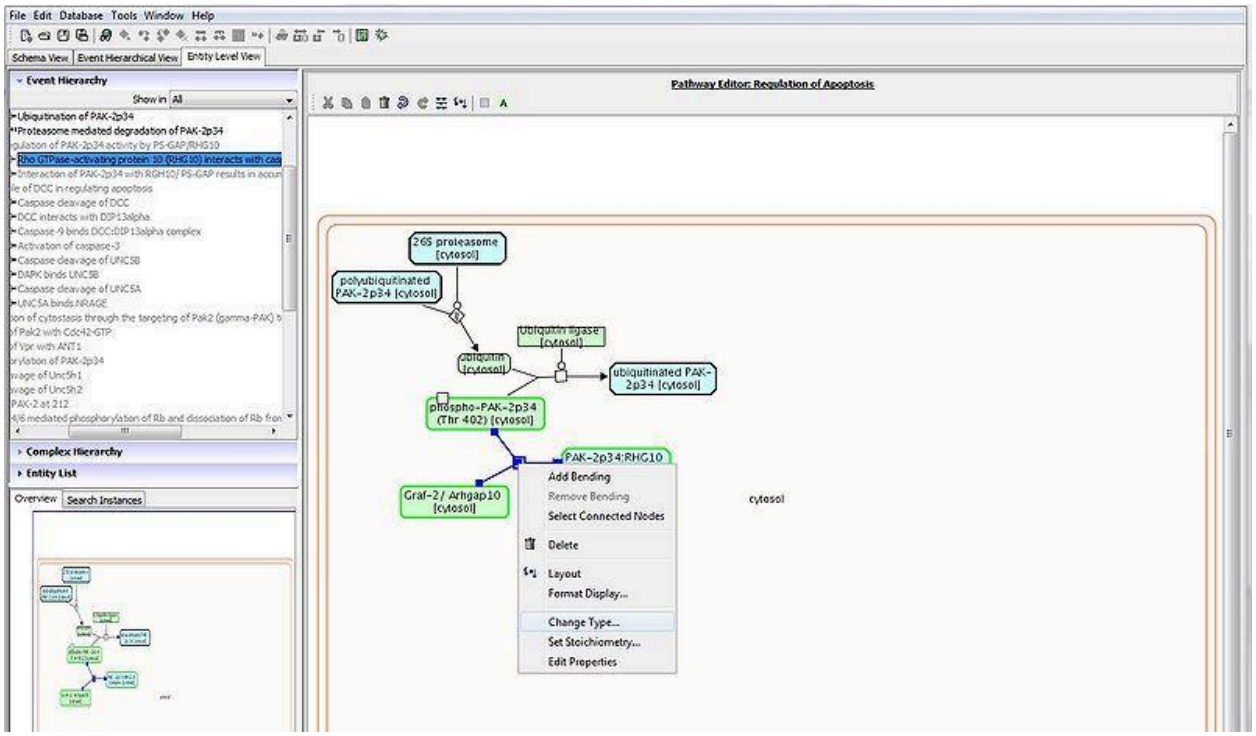
Additional reactions are dragged out and positioned one at a time.


File Edit Database Tools Window Help
Schema View
Event Hierarchical View
Entity Level View
Event Hierarchy
Show in All
Apoptosis
Extrinsic Pathway for Apoptosis
Intrinsic Pathway for Apoptosis
Apoptotic execution phase
Regulation of Apoptosis
Ubiquitination of PAK-2p34
Proteasome mediated degradation of PA
Rho GTPase-activating protein 10
Interaction of PAK-2p34 with RGH10/PS
Role of DCC in regulating apoptosis
Caspase cleavage of DCC
DCC interacts with DIP13alpha
Caspase-9 binds DCC:DIP13alpha comple
Activation of caspase-3
Caspase cleavage of UNC5B
DAPK binds UNC5B
Caspase cleavage of UNC5A
UNC5A binds NRAGE
Association of Pak2 with Cdc42-GTP
Complex Hierarchy
Entity List
Overview
Search Instances
Pathway Editor: Regulation of Apoptosis
polyubiquitinated PAK-2p34 [cytosol]
26S proteasome [cytosol]
Ubiquitin ligase [cytosol]
ubiquitin [cytosol]
ubiquitinated PAK-2p34 [cytosol]
phospho-PAK-2p34 (Thr 402) [cytosol]
cytosol
Slide to Zoom:
Graphic Editor
Property Editor

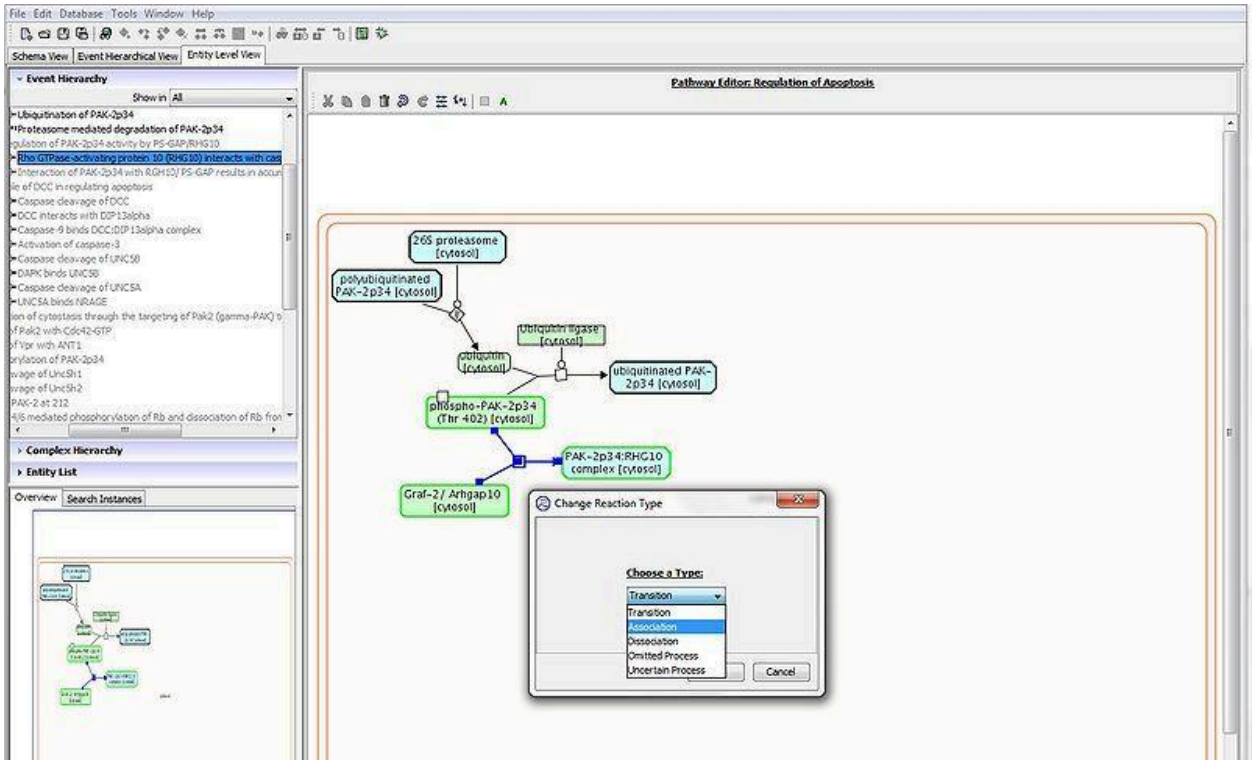


Reaction types

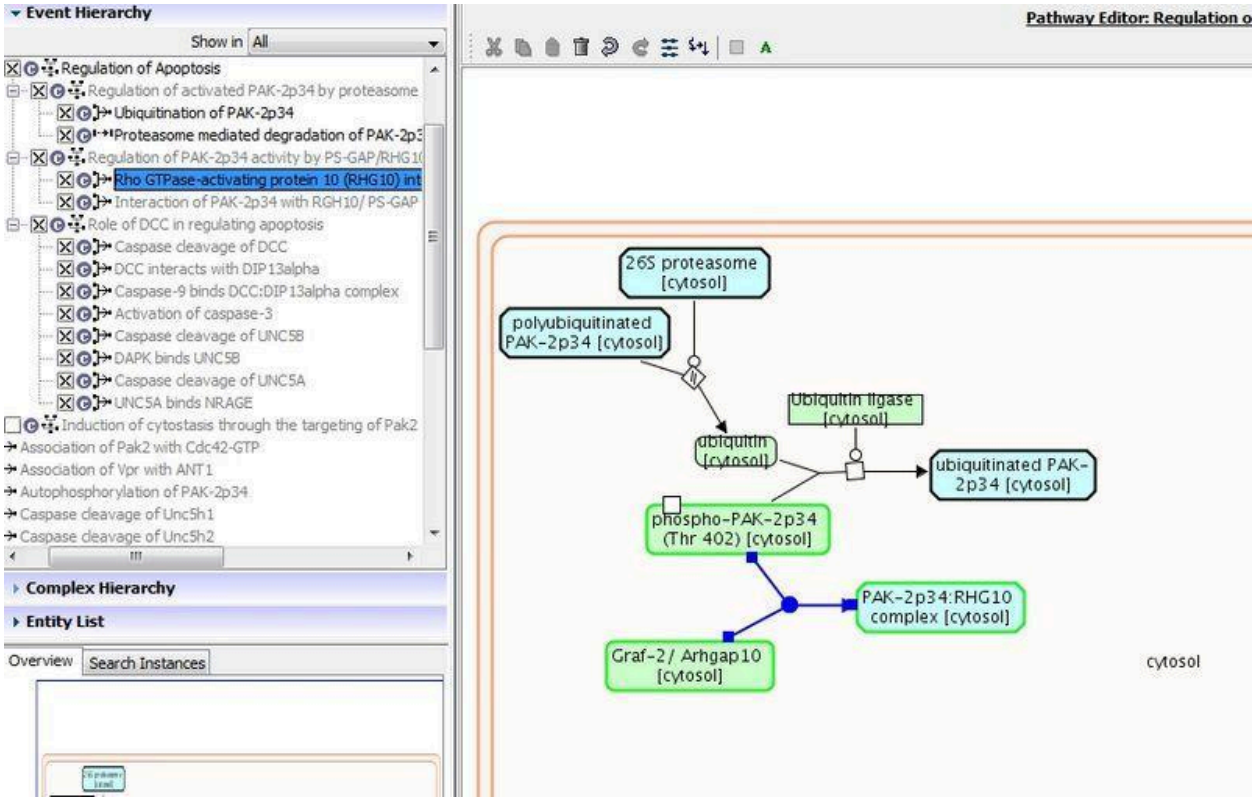
Reactions are of one of 5 types: Transition, Association, Dissociation, Omitted process, and Uncertain process. Transitions involve the molecules changing state, Association is a binding reaction, Dissociation is the Dissociation of a complex. Omitted process, and Uncertain process can be used to describe Black Box Events (BBEs). Omitted process for a BBE describes a process where steps have been deliberately consolidated eg. a gene producing a protein. Uncertain process for a BBE describes a process where the exact mechanism is unknown. To apply a reaction type to a reaction, right click on reaction, select change type .

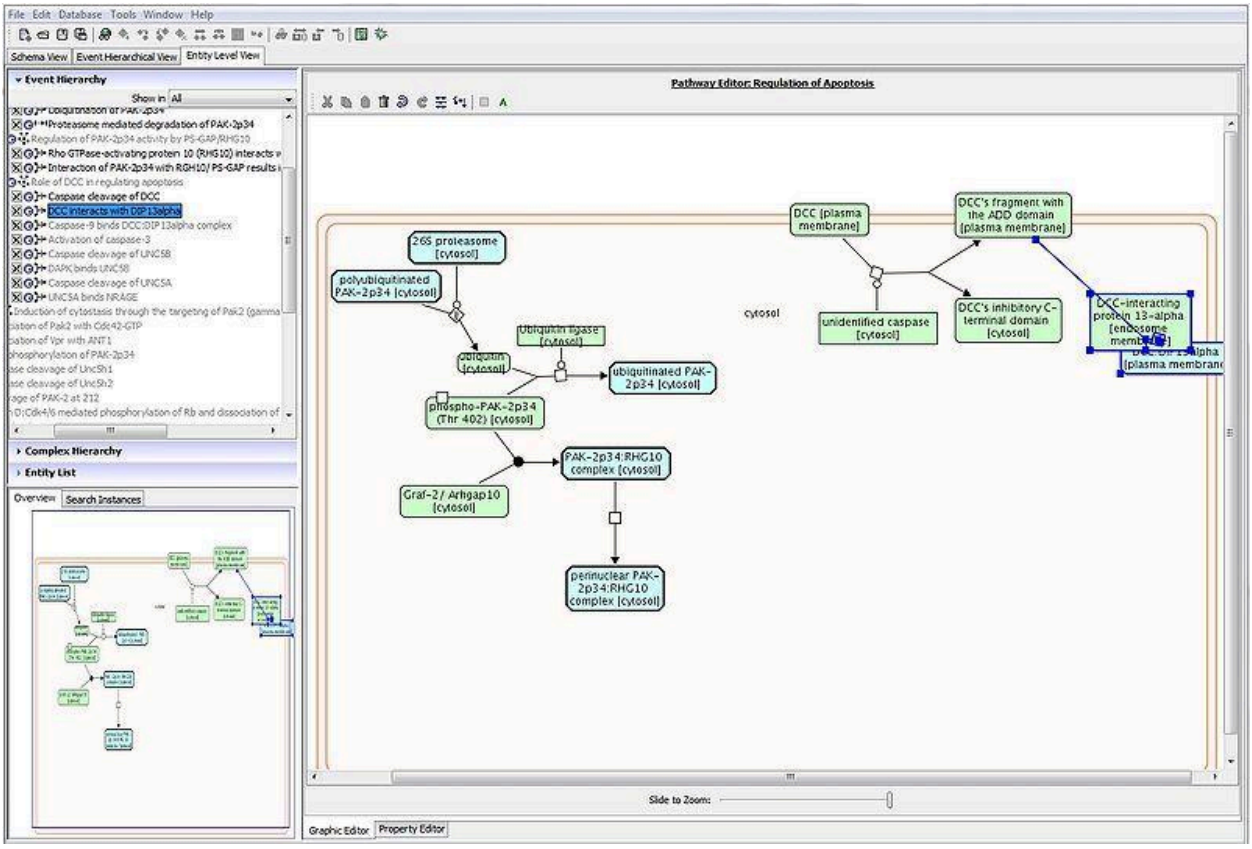


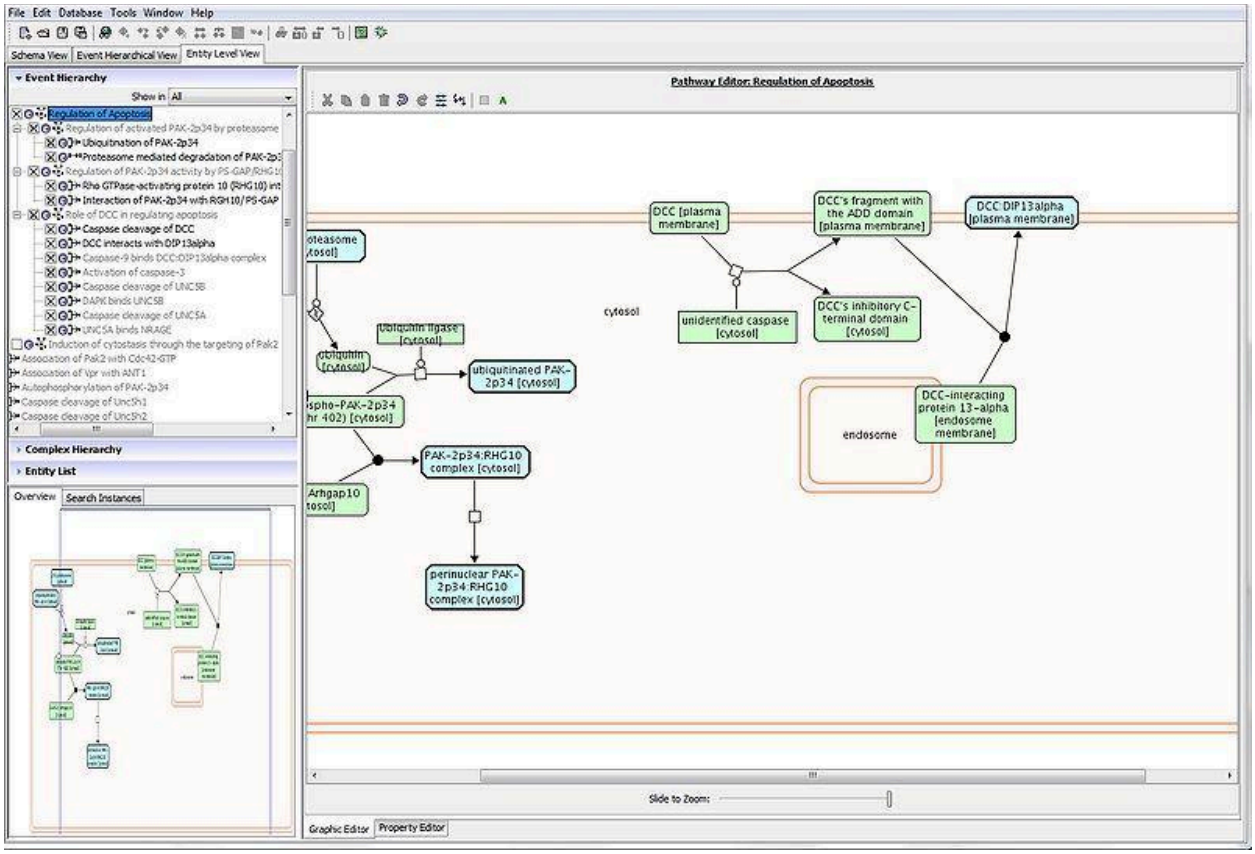
Select type of interest. Here it is an association.

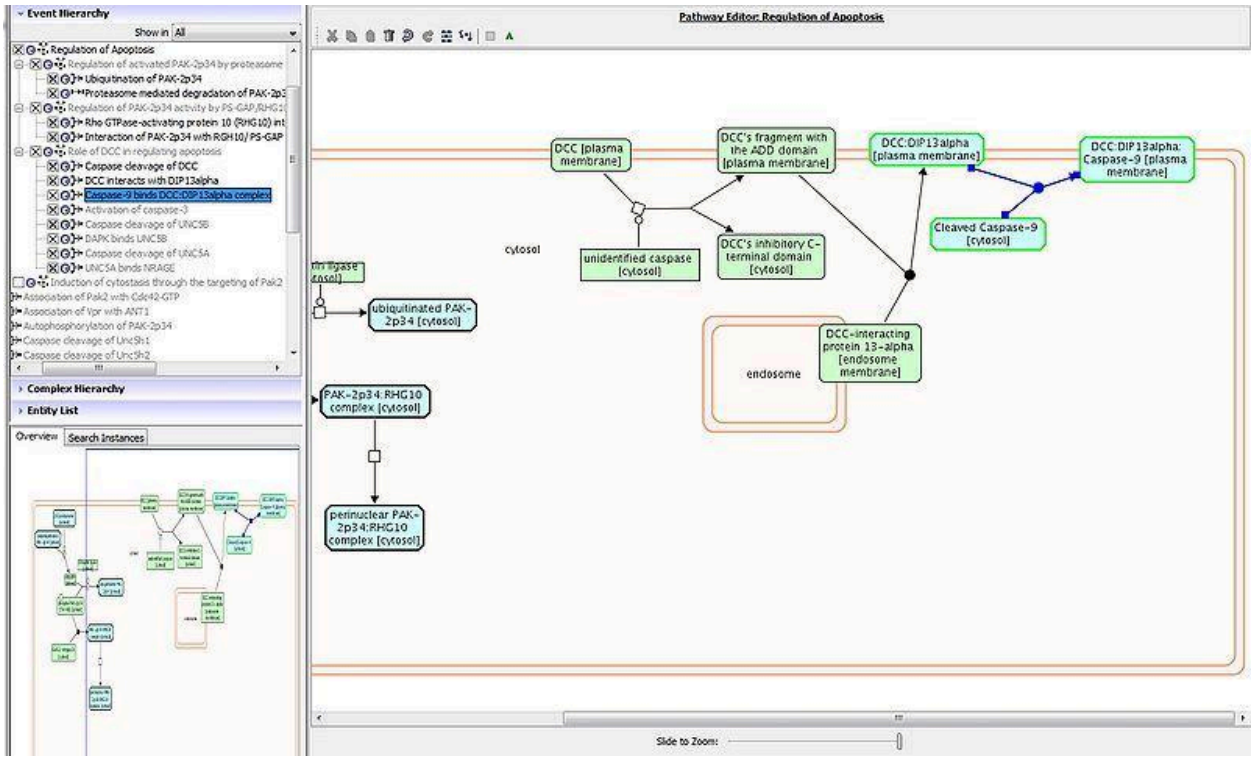


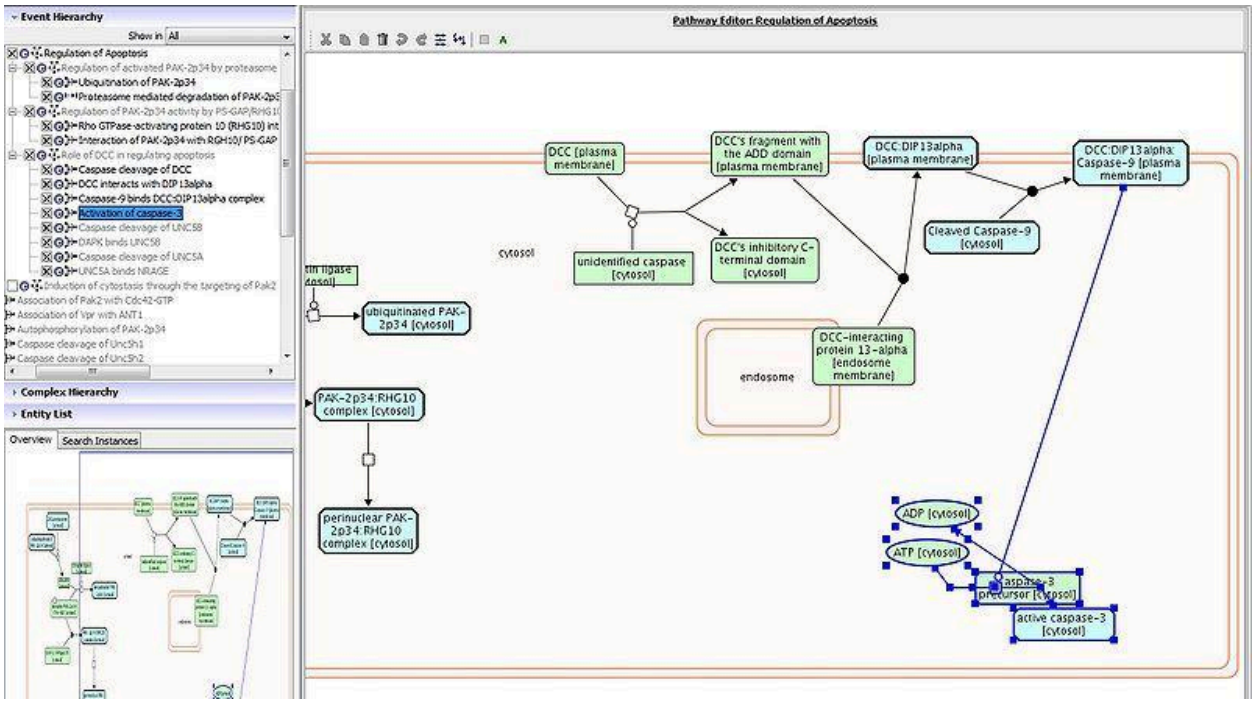
Continue to drag out, and map, and assign reaction type to the remaining reactions.

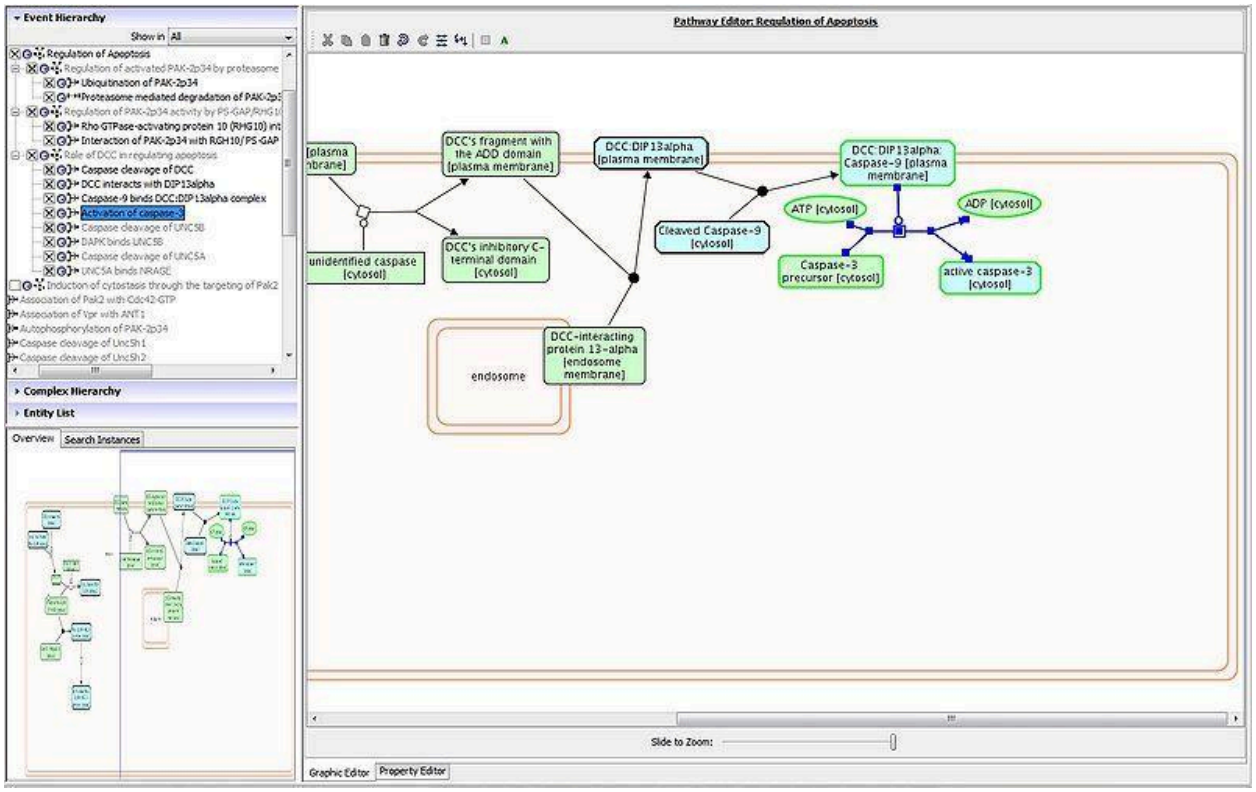


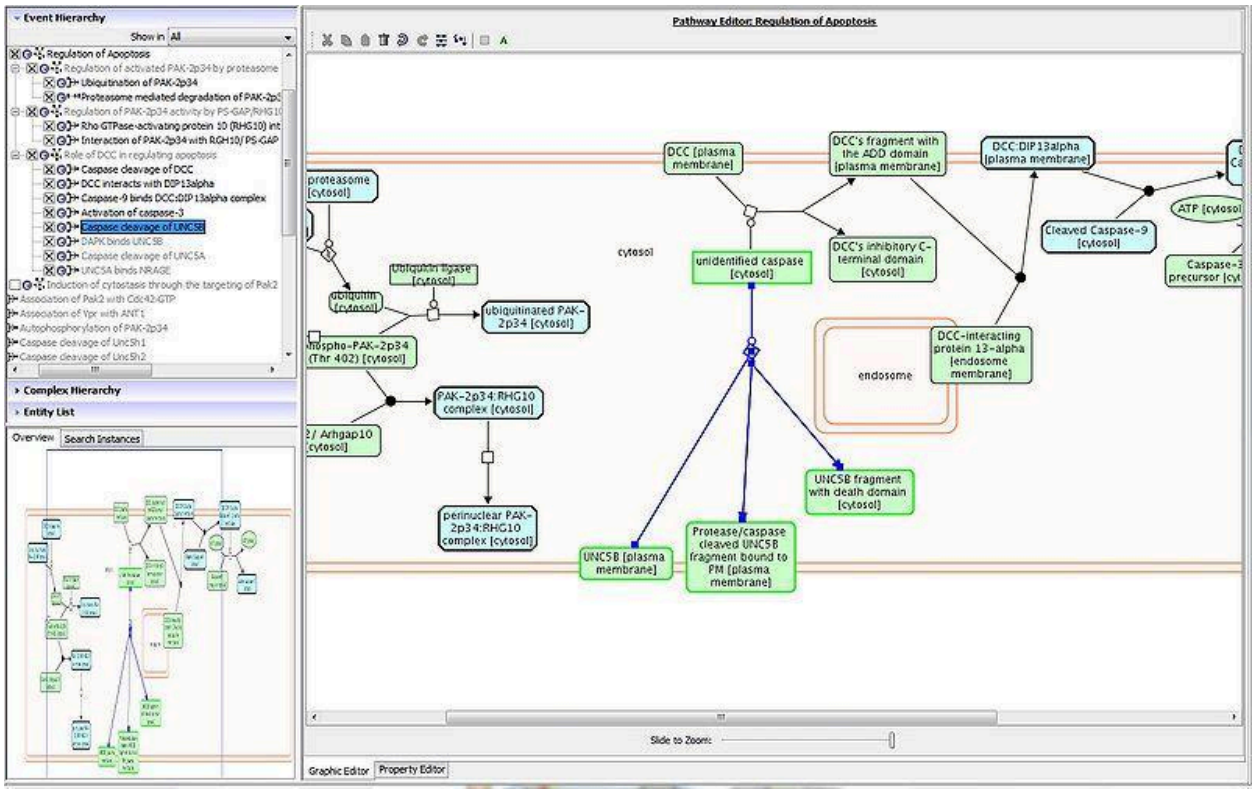

File Edit Database Tools Window Help
Schema View
Event Hierarchical View
Entity Level View
Event Hierarchy
Show in All
Proteasome mediated degradation of PAK-2p34
Regulation of PAK-2p34 activity by PS-GAP/RHG10
Rho GTPase-activating protein 10 (RHG10) interacts
Interacton of PAK-2p34 with RGH10/ PS-GAP results
Role of DCC in regulating apoptosis
Caspase cleavage of DCC
DCC interacts with DIP13alpha
Caspase-9 binds DCC:DIP 13alpha complex
Activation of caspase-3
Caspase cleavage of UNC5B
DAPK binds UNC5B
Caspase cleavage of UNC5A
UNC5A binds NRAGE
Complex Hierarchy
Entity List
Overview
Search Instances
Pathway Editor: Regulation of Apoptosis
26S proteasome [cytosol]
polyubiquitinated PAK-2p34 [cytosol]
Ubiquitin ligase [cytosol]
ubiquitin [cytosol]
ubiquitinated PAK-2p34 [cytosol]
phospho-PAK-2p34 (Thr 402) [cytosol]
Graf-2 / Arhgap10 [cytosol]
PAK-2p34:RHG10 complex [cytosol]
perinuclear PAK-2p34:RHG10 complex [cytosol]
DCC [plasma membrane]
DCC's fragment with the ADD domain [plasma membrane]
cytosol
unidentified caspase [cytosol]
DCC's inhibitory C-terminal domain [cytosol]
DCC-interacting protein 13-alpha [endosome
[plasma membran
Slide to Zoom:
Graphic Editor
Property Editor

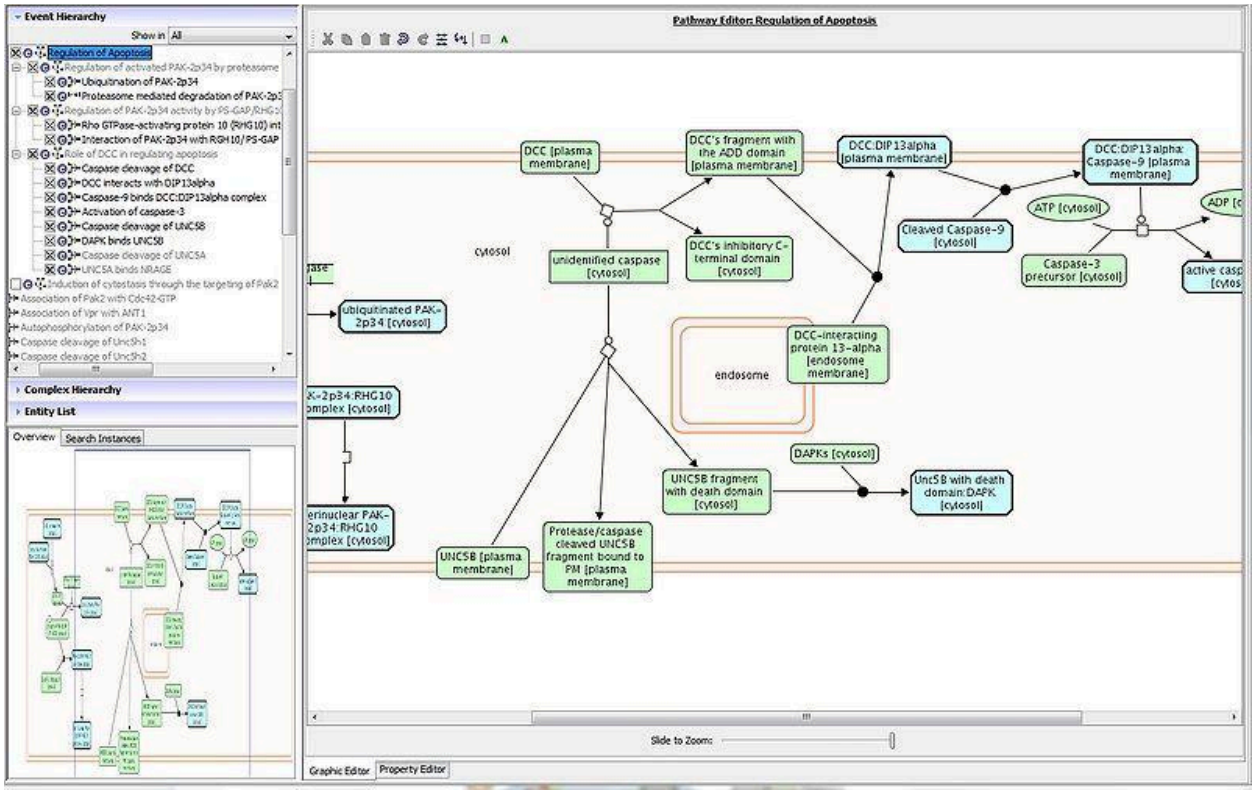

File Edit Database Tools Window Help
Schema View
Event Hierarchical View
Entity Level View
Event Hierarchy
Show in All
Regulation of Apoptosis
Regulation of activated PAK-2p34 by proteasome
Ubiquitination of PAK-2p34
Proteasome mediated degradation of PAK-2p3
Regulation of PAK-2p34 activity by PS-GAP/RHG1
Rho GTPase-activating protein 10 (RHG10) int
Interaction of PAK-2p34 with RGH10/PS-GAP
Role of DCC in regulating apoptosis
Caspase cleavage of DCC
DCC interacts with DIP13alpha
Caspase-9 binds DCC:DIP13alpha complex
Activation of caspase-3
Caspase cleavage of UNC5B
DAPK binds UNC5B
Caspase cleavage of UNC5A
UNC5A binds NRAGE
Induction of cytostasis through the targeting of Pak2
Association of Pak2 with Cdc42:GTP
Association of Vpr with ANT1
Autophosphorylation of PAK-2p34
Caspase cleavage of Unc5h1
Caspase cleavage of Unc5h2
Complex Hierarchy
Entity List
Overview
Search Instances
Pathway Editor: Regulation of Apoptosis
DCC [plasma membrane]
DCC's fragment with the ADD domain [plasma membrane]
DCC DIP13alpha [plasma membrane]
cytosol
unidentified caspase [cytosol]
DCC's inhibitory C-terminal domain [cytosol]
ubiquitinated PAK-2p34 [cytosol]
endosome
DCC-interacting protein 13-alpha [endosome membrane]
PAK-2p34:RHG10 complex [cytosol]
perinuclear PAK-2p34:RHG10 complex [cytosol]
Slide to Zoom:
Graphic Editor
Property Editor

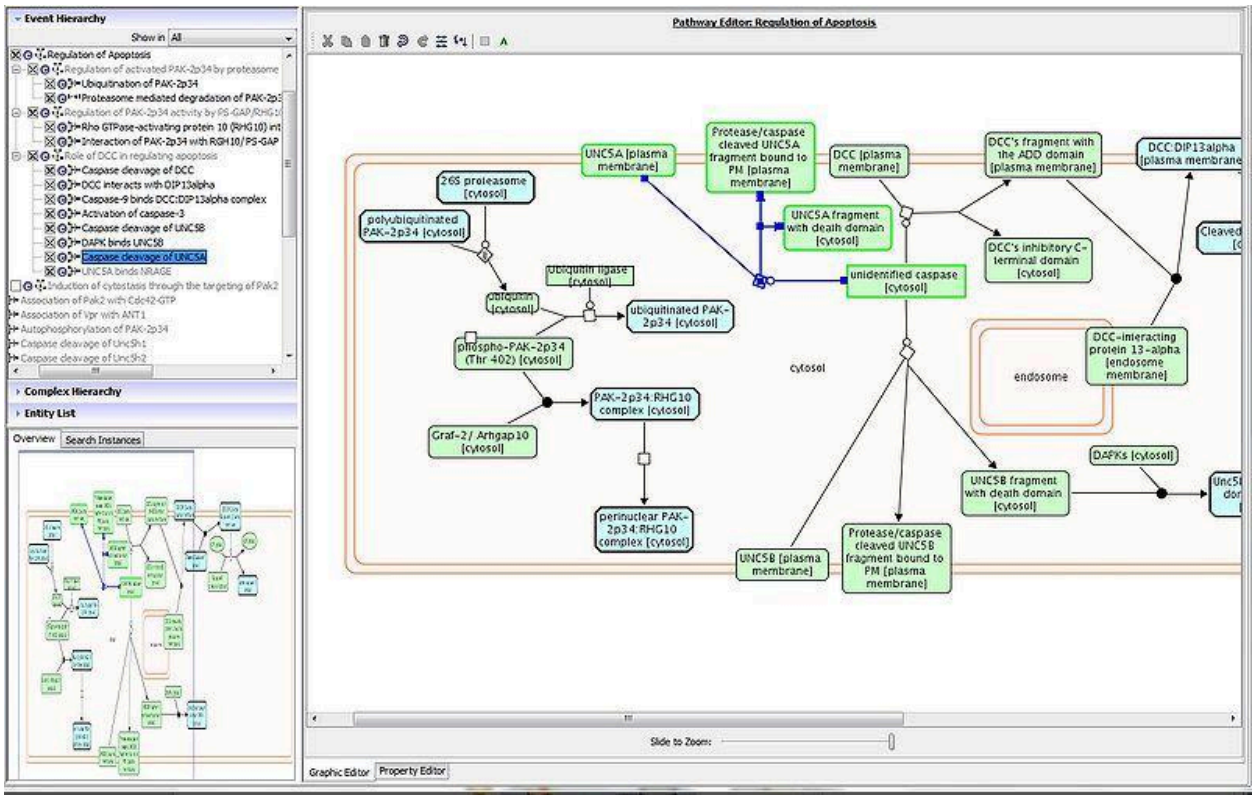

Event Hierarchy
Show in All
Regulation of Apoptosis
Regulation of activated PAK-2p34 by proteasome
Ubiquitination of PAK-2p34
Proteasome mediated degradation of PAK-2p3
Regulation of PAK-2p34 activity by PS-GAP/RHG1
Rho GTPase-activating protein 10 (RHG10) int
Interaction of PAK-2p34 with RGH10/ PS-GAP
Role of DCC in regulating apoptosis
Caspase cleavage of DCC
DCC interacts with DIP13alpha
Caspase-9 binds DCC:DIP13alpha complex
Activation of caspase-3
Caspase cleavage of UNC5B
DAPK binds UNC5B
Caspase cleavage of UNC5A
UNC5A binds NRAGE
Induction of cytostasis through the targeting of Pak2
Association of Pak2 with Cdc42-GTP
Association of Vpr with ANT1
Autophosphorylation of PAK-2p34
Caspase cleavage of UncSh1
Caspase cleavage of UncSh2
Complex Hierarchy
Entity List
Overview
Search Instances
Pathway Editor: Regulation of Apoptosis
DCC [plasma membrane]
DCC's fragment with the ADD domain [plasma membrane]
DCC:DIP13alpha [plasma membrane]
DCC:DIP13alpha: Caspase-9 [plasma membrane]
Cleaved Caspase-9 [cytosol]
cytosol
unidentified caspase [cytosol]
DCC's inhibitory C-terminal domain [cytosol]
ubiquitinated PAK-2p34 [cytosol]
endosome
DCC-interacting protein 13-alpha [endosome membrane]
PAK-2p34:RHG10 complex [cytosol]
perinuclear PAK-2p34:RHG10 complex [cytosol]
Slide to Zoom:

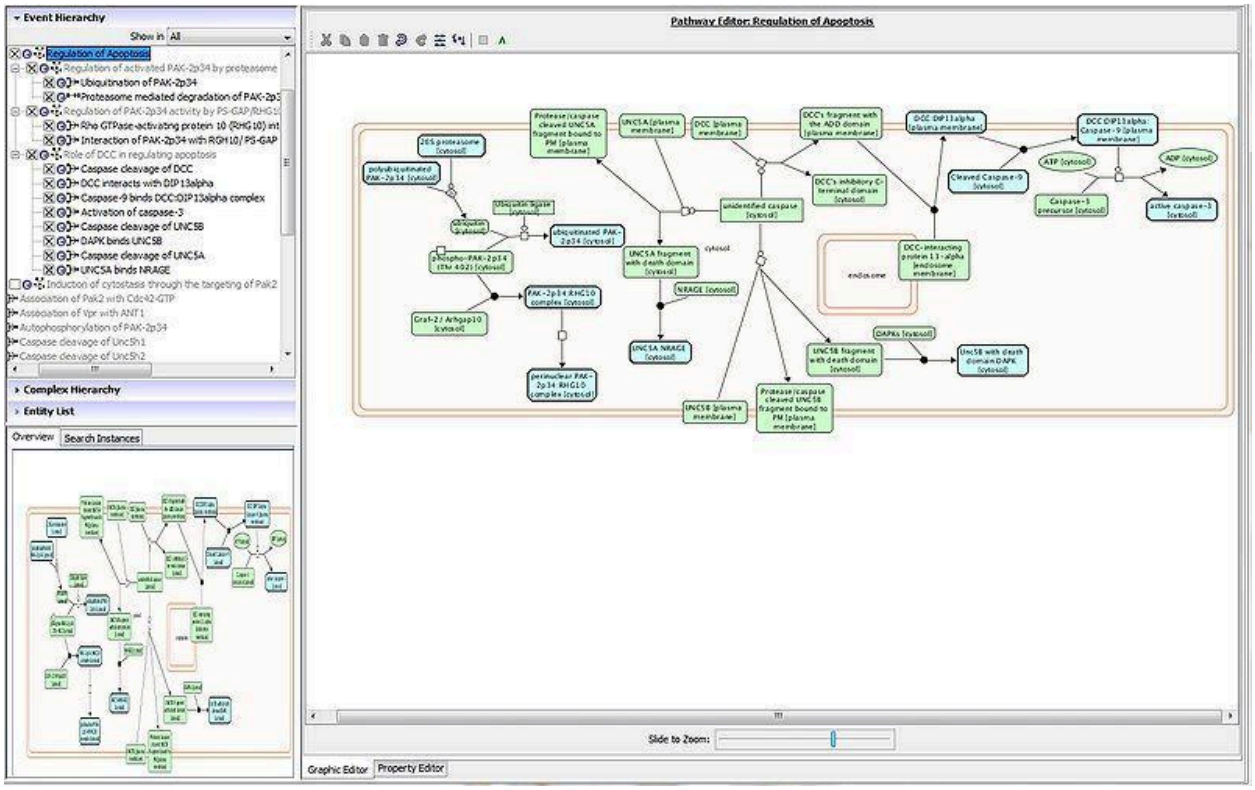

Event Hierarchy
Show in All
Regulation of Apoptosis
Regulation of activated PAK-2p34 by proteasome
Ubiquitination of PAK-2p34
Proteasome mediated degradation of PAK-2p3
Regulation of PAK-2p34 activity by PS-GAP/RHG1
Rho GTPase-activating protein 10 (RHG10) int
Interaction of PAK-2p34 with RGH10/ PS-GAP
Role of DCC in regulating apoptosis
Caspase cleavage of DCC
DCC interacts with DIP13alpha
Caspase-9 binds DCC:DIP13alpha complex
Activation of caspase-3
Caspase cleavage of UNC5B
DAPK binds UNC5B
Caspase cleavage of UNC5A
UNC5A binds NRAGE
Induction of cytostasis through the targeting of Pak2
Association of Pak2 with Cdc42-GTP
Association of Vpr with ANT1
Autophosphorylation of PAK-2p34
Caspase cleavage of UncSh1
Caspase cleavage of UncSh2
Complex Hierarchy
Entity List
Overview
Search Instances
Pathway Editor: Regulation of Apoptosis
DCC [plasma membrane]
DCC's fragment with the ADD domain [plasma membrane]
DCC:DIP13alpha [plasma membrane]
DCC:DIP13alpha: Caspase-9 [plasma membrane]
Cleaved Caspase-9 [cytosol]
cytosol
unidentified caspase [cytosol]
DCC's inhibitory C-terminal domain [cytosol]
ubiquitinated PAK-2p34 [cytosol]
endosome
DCC-interacting protein 13-alpha [endosome membrane]
PAK-2p34:RHG10 complex [cytosol]
perinuclear PAK-2p34:RHG10 complex [cytosol]
ADP [cytosol]
ATP [cytosol]
caspase-3 precursor [cytosol]
active caspase-3 [cytosol]


Event Hierarchy
Show in All
Regulation of Apoptosis
Regulation of activated PAK-2p34 by proteasome
Ubiquitination of PAK-2p34
**Proteasome mediated degradation of PAK-2p3
Regulation of PAK-2p34 activity by PS-GAP/RHG1
Rho GTPase-activating protein 10 (RHG10) int
Interaction of PAK-2p34 with RGH10/PS-GAP
Role of DCC in regulating apoptosis
Caspase cleavage of DCC
DCC interacts with DIP13alpha
Caspase-9 binds DCC:DIP13alpha complex
Activation of caspase-3
Caspase cleavage of UNC5B
DAPK binds UNC5B
Caspase cleavage of UNC5A
UNC5A binds NRAGE
Induction of cytostasis through the targeting of Pak2
Association of Pak2 with Cdc42-GTP
Association of Vpr with ANT1
Autophosphorylation of PAK-2p34
Caspase cleavage of Unc5h1
Caspase cleavage of Unc5h2
Complex Hierarchy
Entity List
Overview
Search Instances
Pathway Editor: Regulation of Apoptosis
DCC's fragment with the ADD domain [plasma membrane]
DCC:DIP13alpha [plasma membrane]
DCC:DIP13alpha: Caspase-9 [plasma membrane]
ATP [cytosol]
ADP [cytosol]
Cleaved Caspase-9 [cytosol]
unidentified caspase [cytosol]
DCC's inhibitory C-terminal domain [cytosol]
Caspase-3 precursor [cytosol]
active caspase-3 [cytosol]
endosome
DCC-interacting protein 13-alpha [endosome membrane]
Slide to Zoom:
Graphic Editor
Property Editor

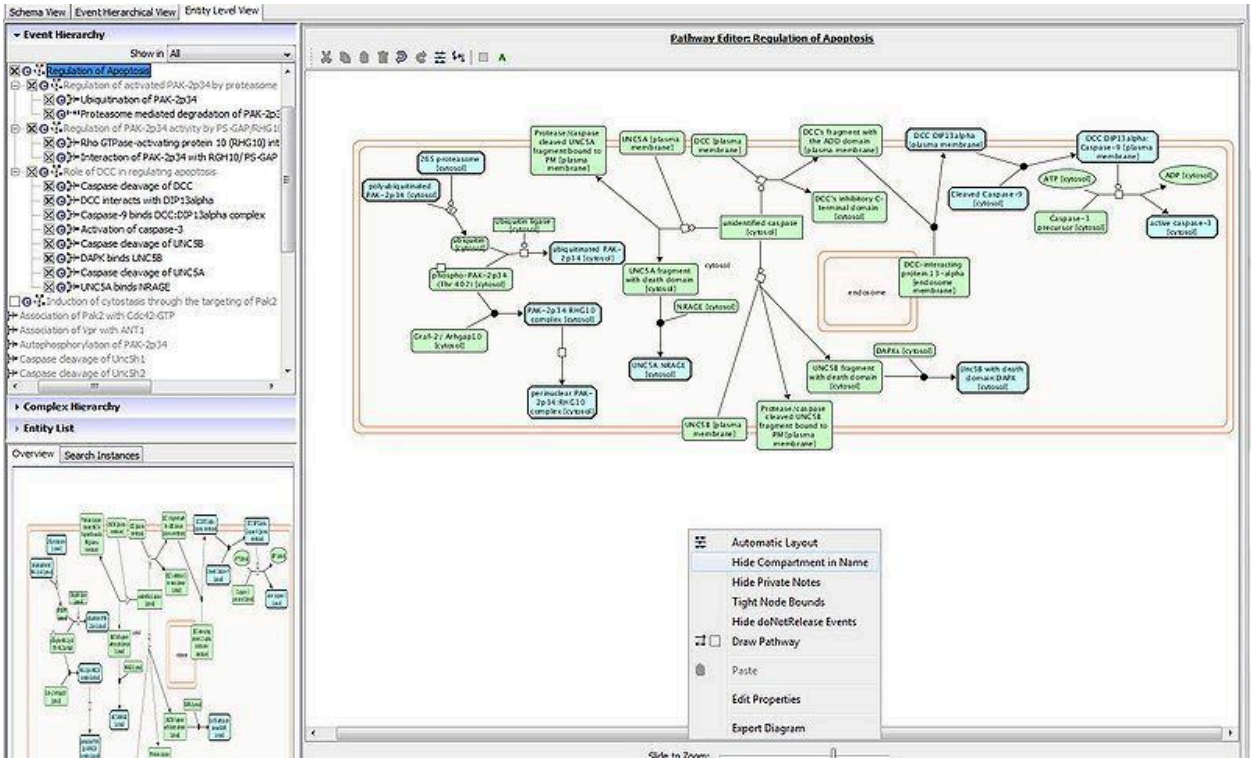

Event Hierarchy
Show in All
Regulation of Apoptosis
Regulation of activated PAK-2p34 by proteasome
Ubiquitination of PAK-2p34
**Proteasome mediated degradation of PAK-2p3
Regulation of PAK-2p34 activity by PS-GAP/RHG1
Rho GTPase-activating protein 10 (RHG10) int
Interaction of PAK-2p34 with RGH10/PS-GAP
Role of DCC in regulating apoptosis
Caspase cleavage of DCC
DCC interacts with DIP13alpha
Caspase-9 binds DCC:DIP13alpha complex
Activation of caspase-3
Caspase cleavage of UNC5B
DAPK binds UNC5B
Caspase cleavage of UNC5A
UNC5A binds NRAGE
Induction of cytostasis through the targeting of Pak2
Association of Pak2 with Cdc42-GTP
Association of Vpr with ANT1
Autophosphorylation of PAK-2p34
Caspase cleavage of Unc5h1
Caspase cleavage of Unc5h2
Complex Hierarchy
Entity List
Overview
Search Instances
Pathway Editor: Regulation of Apoptosis
DCC [plasma membrane]
DCC's fragment with the ADD domain [plasma membrane]
DCC:DIP13alpha [plasma membrane]
ATP [cytosol]
Cleaved Caspase-9 [cytosol]
proteasome [cytosol]
cytosol
unidentified caspase [cytosol]
DCC's inhibitory C-terminal domain [cytosol]
ubiquitinated PAK-2p34 [cytosol]
PAK-2p34:RHG10 complex [cytosol]
endosome
DCC-interacting protein 13-alpha [endosome membrane]
UNC5B fragment with death domain [cytosol]
perinuclear PAK-2p34:RHG10 complex [cytosol]
UNC5B [plasma membrane]
Protease/caspase cleaved UNC5B fragment bound to PM [plasma membrane]
Slide to Zoom:
Graphic Editor
Property Editor


Event Hierarchy
Show in All
Regulation of Apoptosis
Regulation of activated PAK-2p34 by proteasome
Ubiquitination of PAK-2p34
Proteasome mediated degradation of PAK-2p3
Regulation of PAK-2p34 activity by PS-GAP/RHG1
Rho GTPase-activating protein 10 (RHG10) int
Interaction of PAK-2p34 with RGH10/PS-GAP
Role of DCC in regulating apoptosis
Caspase cleavage of DCC
DCC interacts with DIP13alpha
Caspase-9 binds DCC:DIP13alpha complex
Activation of caspase-3
Caspase cleavage of UNC5B
DAPK binds UNC5B
Caspase cleavage of UNC5A
UNC5A binds NRAGE
Induction of cytostasis through the targeting of Pak2
Association of Pak2 with Cdc42-GTP
Association of Vpr with ANT1
Autophosphorylation of PAK-2p34
Caspase cleavage of Unc5h1
Caspase cleavage of Unc5h2
Complex Hierarchy
Entity List
Overview
Search Instances
Pathway Editor: Regulation of Apoptosis
DCC [plasma membrane]
DCC's fragment with the ADD domain [plasma membrane]
DCC:DIP13alpha [plasma membrane]
DCC:DIP13alpha: Caspase-9 [plasma membrane]
ATP [cytosol]
ADP [c
Cleaved Caspase-9 [cytosol]
Caspase-3 precursor [cytosol]
active casp [cytos
cytosol
unidentified caspase [cytosol]
DCC's inhibitory C-terminal domain [cytosol]
ase
ubiquitinated PAK-2p34 [cytosol]
endosome
DCC-interacting protein 13-alpha [endosome membrane]
K-2p34:RHG10 omplex [cytosol]
DAPKs [cytosol]
UNC5B fragment with death domain [cytosol]
Unc5B with death domain DAPK [cytosol]
erinuclear PAK-2p34:RHG10 omplex [cytosol]
UNC5B [plasma membrane]
Protease/caspase cleaved UNC5B fragment bound to PM [plasma membrane]
Slide to Zoom:
Graphic Editor
Property Editor

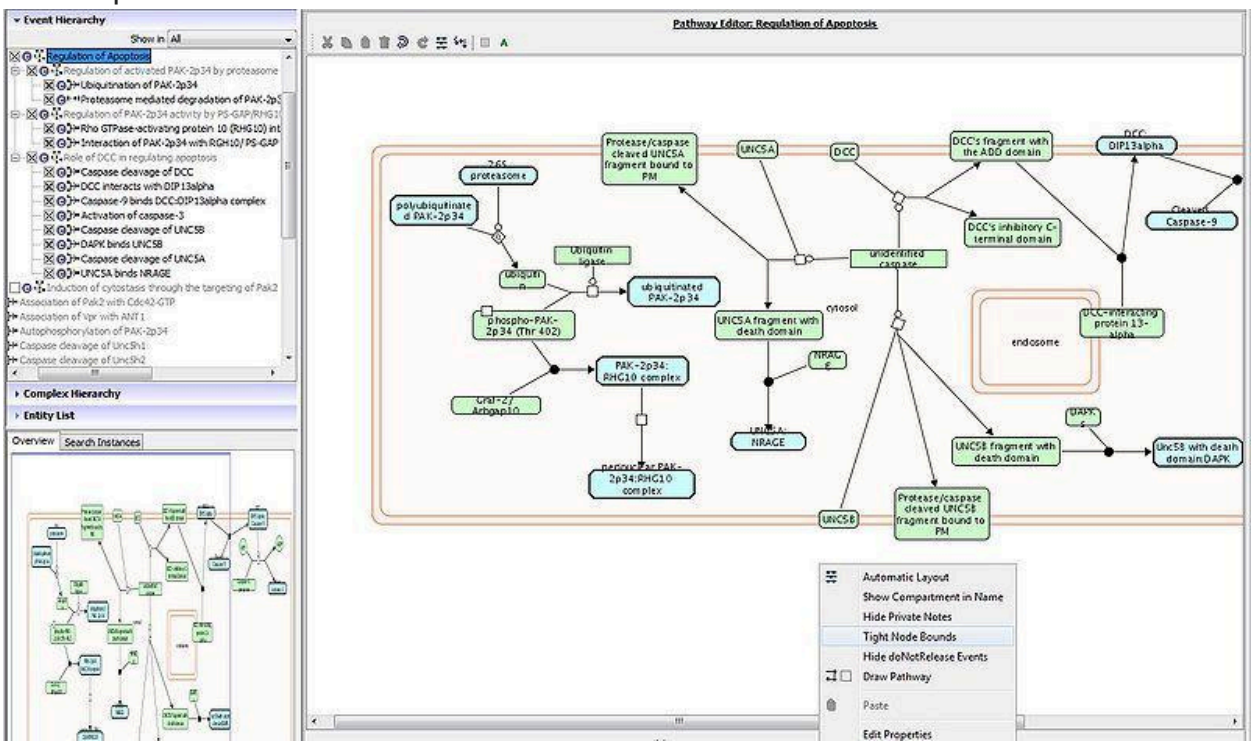

Event Hierarchy
Show in All
Regulation of Apoptosis
Regulation of activated PAK-2p34 by proteasome
Ubiquitination of PAK-2p34
Proteasome mediated degradation of PAK-2p3
Regulation of PAK-2p34 activity by PS-GAP/RHG1
Rho GTPase-activating protein 10 (RHG10) int
Interaction of PAK-2p34 with RGH10/PS-GAP
Role of DCC in regulating apoptosis
Caspase cleavage of DCC
DCC interacts with DIP13alpha
Caspase-9 binds DCC:DIP13alpha complex
Activation of caspase-3
Caspase cleavage of UNC5B
DAPK binds UNC5B
Caspase cleavage of UNC5A
UNC5A binds NRAGE
Induction of cytostasis through the targeting of Pak2
Association of Pak2 with Cdc42-GTP
Association of Vpr with ANT1
Autophosphorylation of PAK-2p34
Caspase cleavage of Unc5h1
Caspase cleavage of Unc5h2
Complex Hierarchy
Entity List
Overview
Search Instances
Pathway Editor: Regulation of Apoptosis
UNC5A [plasma membrane]
Protease/caspase cleaved UNC5A fragment bound to PM [plasma membrane]
DCC [plasma membrane]
DCC's fragment with the ADD domain [plasma membrane]
DCC:DIP13alpha [plasma membrane
26S proteasome [cytosol]
polyubiquitinated PAK-2p34 [cytosol]
UNC5A fragment with death domain [cytosol]
DCC's inhibitory C-terminal domain [cytosol]
Ubiquitin ligase [cytosol]
unidentified caspase [cytosol]
Ubiquitin [cytosol]
ubiquitinated PAK-2p34 [cytosol]
phospho-PAK-2p34 (Thr 402) [cytosol]
cytosol
endosome
DCC-interacting protein 13-alpha [endosome membrane]
PAK-2p34:RHG10 complex [cytosol]
Graf-2/ Arhgap10 [cytosol]
DAPKs [cytosol]
UNC5B fragment with death domain [cytosol]
perinuclear PAK-2p34:RHG10 complex [cytosol]
UNC5B [plasma membrane]
Protease/caspase cleaved UNC5B fragment bound to PM [plasma membrane]
Slide to Zoom:
Graphic Editor
Property Editor



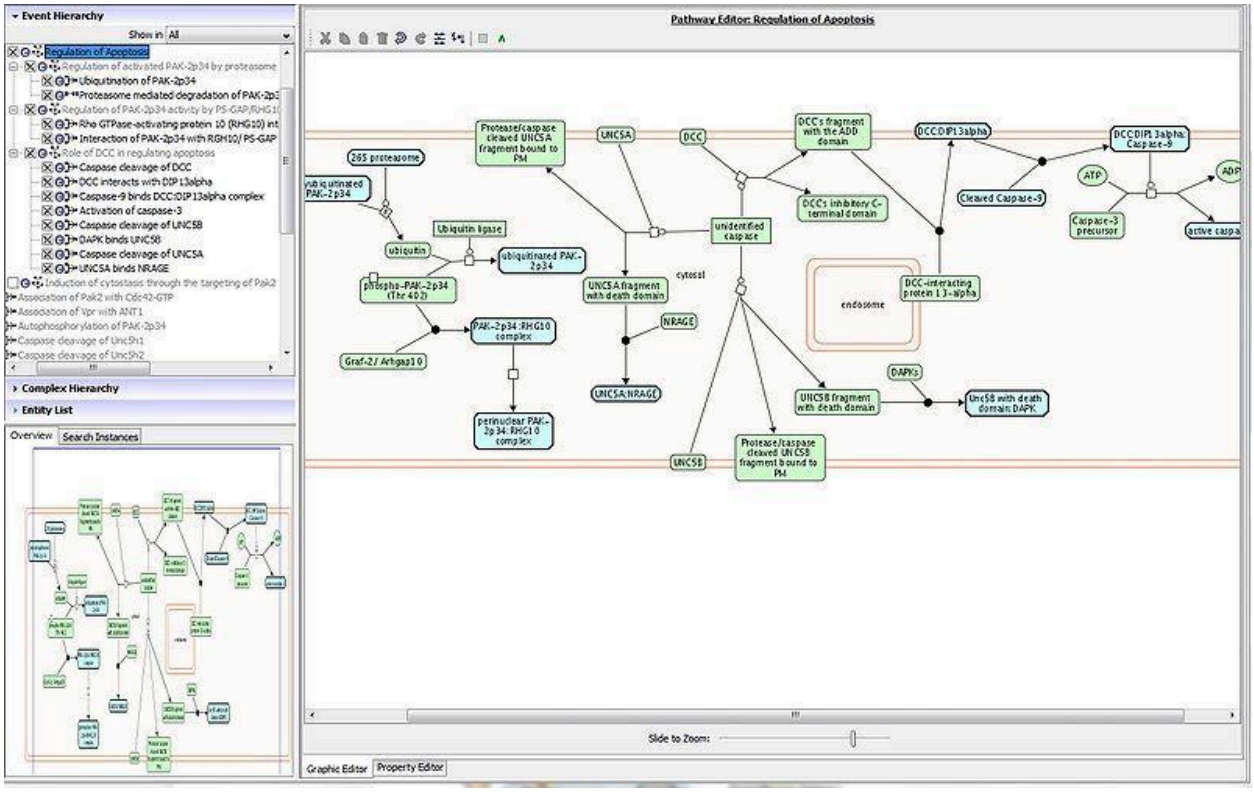
Once all of the reactions have been laid out, the compartment names on molecules can be hidden by right clicking on the diagram anywhere and selecting "Hide compartment in names".



Right click and select "Tight Node Bounds. This will reduce the space left by removing the compartment names.



Here is the completed diagram.



Verify that all entity names are contained in the icons, resizing manually if necessary.

Connecting sets with their members

In some cases sets of entities are used in a pathway, and also the entities themselves. In the diagram you may want to indicate that an entity belongs to a set. To make the connection, click on the set and one of the entities you want to connect to (done by holding shift as you choose both), right-click with both highlighted and choose “Add EntitySet and Member Link”. This will link the set to the chosen entity with a dotted line.

Use of pathways icons as links to related "pathways" in diagrams

If you want to include a link to a pathway that is not actually "part of" the pathway that you are diagramming, you can do this by checking out both the pathway you are diagramming (pathway A) as well as the pathway you'd like to include as a icon (pathway B) using the Event view. Open the diagram of pathway A in the ELV view. Then, drag the icon of pathway B into the ELV. **Note that you must indicate the reaction that links the diagrams via a flowline between the pathway icon and the output of the connecting reaction. Save and commit the changes. Note that the pathway whose icon you would like to include MUST have a pathway diagram associated with it.

To make an association between a reaction in one diagram and another pathway diagram in which the reaction (entity) acts as a preceding event (participant), you can draw a flow line from the entity to the pathway icon. Note that the pathway displayed as an icon in such a case MUST, itself, have an ELV in order to link this way.

To make this connection between events (on a mac) go to preferences under Reactome curator tool on top bar and unclick 'Use ELV as drawing tool'.

Then if you go into a diagram and hover over entity you want to connect, you get the shadow arrows that you can extend to your green box icon, with options for either activate or inhibit.
On a PC I go to the Tools tab, select Options, and then uncheck "Use ELV as a drawing tool".

Drawing a new pathway diagram

If you are creating a diagram for a new top level pathway (check with Peter/Lisa if this is appropriate) ,remember that the pathway itself must be listed as frontPage item in order to see the diagram. To do this, check out the frontPage instance from gk_central and add your pathways in the frontPageItem slot. Also, please mark this in the editorial calendar as a front page item and inform.

Editing existing pathways and corresponding diagrams

Any time a change is made to events in a pathway, the corresponding reactions/RLEs in a diagram must be adjusted accordingly. To start on revisions, be sure to "Build" project into a local CT project. ** Note that if you are revising a "normal pathway" that has a disease counterpart, you must be sure that you start off by "Building" (as opposed to"Checking out") the full disease pathway. Then open the Disease diagram and make any necessary edits there.

Encapsulating a pathway diagram

Encapsulating a sub-pathway

Encapsulate Diagram: this feature is used to refactor a section of a diagram for a sub-pathway contained by a displayed pathway diagram into a new diagram in the Entity Level View. It works as follows using the diagram for "Metabolism of steroid hormones and vitamins A and D" created by Peter and Bijay as an example:

1. Use "Build Project" or "Check Out" to create a local copy of project for Event 196071 (Pathway: Metabolism of steroid hormones and vitamins A and D)

2. In the Entity Level View, "Open Diagram" to open the diagram for the top-level pathway (Metabolism of steroid hormones and vitamins A and D).

3. Select "Vitamin D (calciferol) metabolism" in the "Event Hierarchy" at the left panel. Right click it to get a popup menu called "Encapsulate Diagram".

4. Select this action to create a diagram for this selected sub- pathway. A new pathway node for this sub-pathway will be added to the displayed diagram for pathway "Metabolism of steroid hormones and vitamins A and D". All reactions, compartments and reactions used by this sub-pathway only will be deleted from this displayed diagram.

5. To view the diagram for sub-pathway "Vitamin D (calciferol) metabolism", use "Open Diagram" after selecting this sub-pathway. 6). The user may need some fine adjustments: e.g. resize compartments in both old and new diagrams, add links to the new pathway node.

Encapsulating a connected pathway

Pathway diagrams may not be in a hierarchical relationship but still may have connections via input/output of their reactions, e.g. drug ADME pathways provide the input of drug-target reactions in metabolic pathways. To show a connected pathway B in the diagram of pathway A:

1. Build both projects together from the Database Browser --> Event View by selecting A and Ctrl-Click selecting B, then right-click Build Project

2. Open the Entity Level View, in Event Hierarchy right-click B with Open Diagram, then right-click A with Open Diagram

3. With Diagram A open select pathway B in Event Hierarchy and move it into the diagram. The encapsulated B should appear.

4. Select both the encapsulated B and the entity that is connected with B (e.g. the drug), right-click Create Flow Link.

Guidelines for encapsulating events and reusing events in other diagrams

1. If a pathway (pathway X) that has been drawn out in one pathway diagram is found to be participating in other pathways, consider encapsulating the pathway into its own diagram as described above. In this case, pathway X will be represented as an icon in the pathway diagram(s) in which pathway X participates. **Note that you can ONLY represent a pathway as a pathway icon in a diagram IF that pathway has a pathway diagram of its own.

2. It follows from the above rule, that IF pathway X has a diagram of its own, you cannot draw out pathway X in another pathway diagram. It MUST be represented as a pathway icon. You CAN, however, draw out subevents of pathway X (reactions and subpathways) in another pathway (Y) diagram, as long as the events are represented in

the hierarchy of pathway Y AND as long as the subpathway you wish to represent does not have a diagram of its own. If it does you'd use the pathway icon of the subpathway.

Using the precedingEvent suggestion feature in the curator tool
See slides here (Appendix C7).

Cell Lineage Path Curation Guide

Introduction

Our bodies are built of >30 trillion cells specialized to fulfill diverse roles within our tissues, organs and organ systems. All these cells originate from a single cell, a zygote formed at conception. From zygote to fetus, and throughout childhood, adolescence and adulthood, cells divide and commit to different fates in order for the organism to develop, sustain and regenerate. The series of steps that lead from an undifferentiated progenitor cell, such as a stem cell, to one of its several possible specialized descendants constitutes a cell lineage path. Recent technological advances have allowed researchers to harvest high-throughput omics data from single cells of multicellular organisms and use it to track and manipulate cell fates (Burgess 2018; Saelens et al. 2019). This opens the door to the possibility of deciphering cell lineage paths at single cell resolution, a critical requirement for the advancement of regenerative medicine and cancer medicine.

Researchers who conduct experiments that involve cell lineage tracing through the study of genomic, epigenomic, transcriptomic and proteomic features of single cells currently expend significant effort extracting information from published sources to create short-lived, limited-scope local annotations of biological markers that are needed to interpret and analyze their high-throughput experimental data. This creates the urgent need for an open source, comprehensive, continuously maintained reference repository of cell lineage paths framed on the pre-existing knowledge of tissue morphogenesis, as well as a cell type-specific pathway resource, which can serve as a framework for interpretation and integration of latest experimental findings. Cytomics Reactome, a database of cell lineage paths and a toolset for analysis of single cell omics data, aims to be an integrative systems biology resource that fulfills this need.

Cytomics Reactome will systematically curate published studies and datasets to create a lasting, up-to-date, broad-scope and peer-reviewed online repository of cell

type-specific biological markers, cell lineages, and relevant regulatory networks. The resource will be freely available to the research community, and researchers will be able to take part in the growth and maintenance of this repository by participating in community curation efforts, acting as peer reviewers, and by providing user feedback. To analyze single cell omics data, researchers will be able to download the repository and incorporate it into their own analysis pipelines, or to use online Cytomics Reactome analytic tools to apply cell lineage-oriented or pathway-oriented cell type-tailored analyses to their datasets.

By providing a peer-reviewed, standardized reference of cell lineage paths and milestones, Cytomics Reactome will increase the efficiency and reproducibility of the single cell-based research, and assist in addressing long-standing fundamental questions concerning regulation of progenitor cell division and differentiation by highlighting scientific gaps and acting as a catalyst for hypothesis generation.

Data Model

The Reactome data model provides a robust framework for annotation of biological entities and events (Joshi-Tope et al. 2005) and was easily expanded to accommodate new classes needed for Cytomics Reactome. Reactome web application and pathway visualization tools (Croft et al. 2011) have been extended to enable Cytomics Reactome visual display on the curator website, and we are working on enabling the display and search functions on the live website.

The basic unit of the Reactome data model is a reaction, an event that converts input to output physical entities. Physical entities, which can be simple or complex molecules, or molecular complexes, can also act as reaction catalysts and regulators. Reactions are grouped into causal chains to form pathways (Joshi-Tope et al. 2005).

The Event and PhysicalEntity classes were expanded to accommodate cell lineage path curation. A new Cell class, added as a subclass of the PhysicalEntity class (Figure 1A), has CellType, TissueLayer, Tissue and Organ as single-valued attributes. Multi-valued attributes of the Cell are ProteinMarker, RNAMarker, Proteome, Transcriptome and Epigenome (Figure 1C).

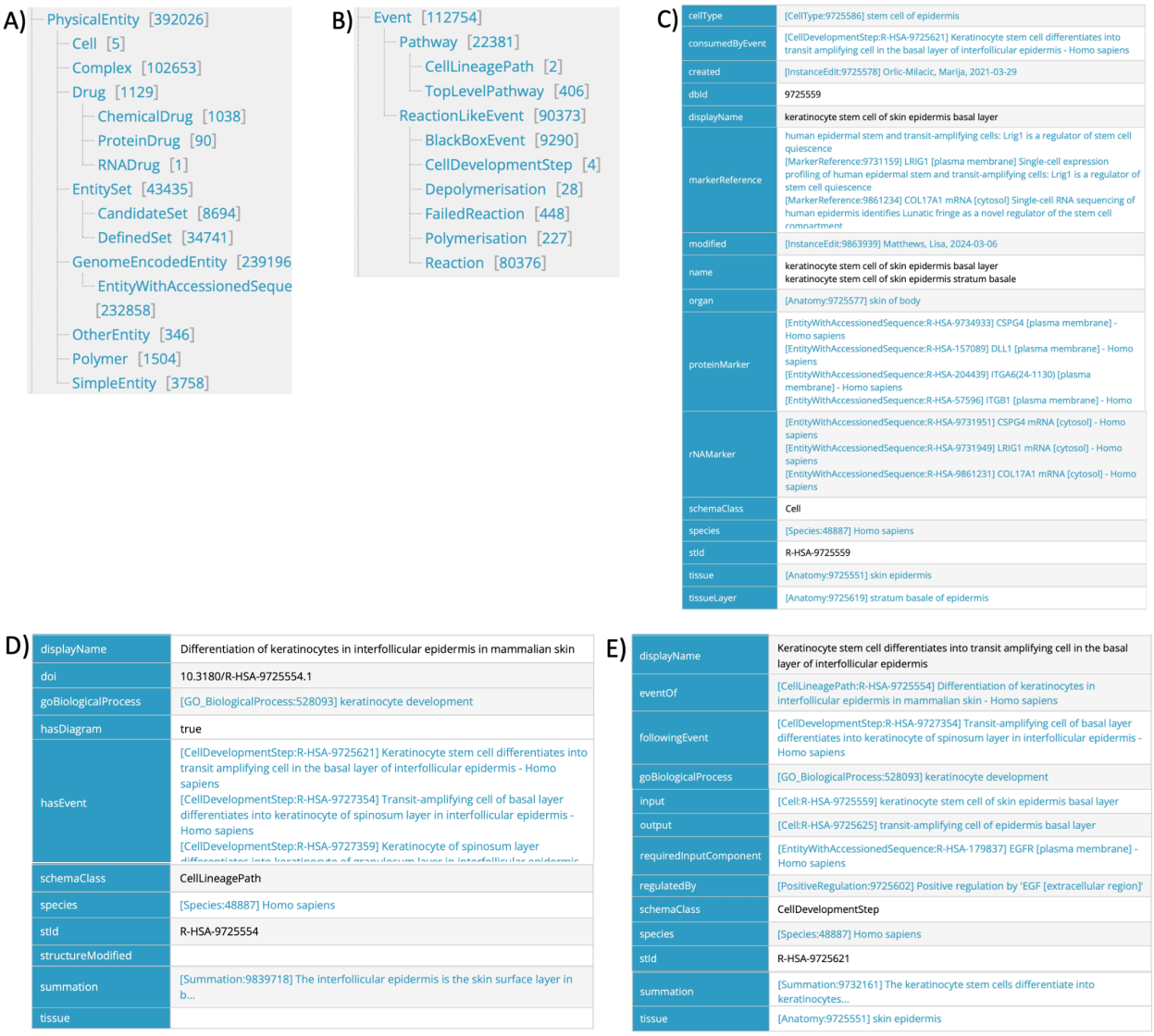


Figure 1. Changes to the Reactome data model to enable Cytomics Reactome annotations. A) Addition of the Cell subclass to the Physical Entity class. B) Addition of Cell Development Step and Cell Lineage Path subclasses to the Event class. C) Attributes of the Cell subclass. D) Attributes of the Cell Lineage Path subclass. E) Attributes of the Cell Development Step subclass.

Developmental relationships among Cells are described by two new event subclasses: Cell Development Step and Cell Lineage Path (Figure 1B), each tied to a Gene Ontology (GO) biological process (The Gene Ontology Consortium 2019). The Cell Development Step, similar to the Reaction subclass, has inputs and outputs defined, which are instances of the Cell class (an input representing the cell of origin, and an output representing the destination cell type). Cell Development Step attributes include regulators (molecules promoting or inhibiting the step) and required input components (input cell biomarkers required for the action of regulators) (Figure 1E). The Cell Lineage Path, resembling the existing Pathway subclass, consists of Cell Development Steps connected through shared input/output Cells (Figure 1D).

PATHWAY BROWSER

Cytomics Reactome webpages will enable users to search and browse the curated lineages through tissue-level view (Figure 2A) or lineage path view (Figure 2B). Based on annotated regulators and required input components of Cell Development Steps that constitute a Cell Lineage Path, and on annotated biomarkers of participating Cells, we will generate a sub-network of Reactome's gene functional-interaction (FI) network (Wu and Stein 2012; Wu and Haw 2017) for each Path. These networks will first be extracted programmatically from annotation and then manually checked to ensure quality. This lineage path gene regulatory network, available to users as a diagram and as an FI table, will be free of cell boundaries and arranged in an approximate time progression order from left to right (Figure 2C).

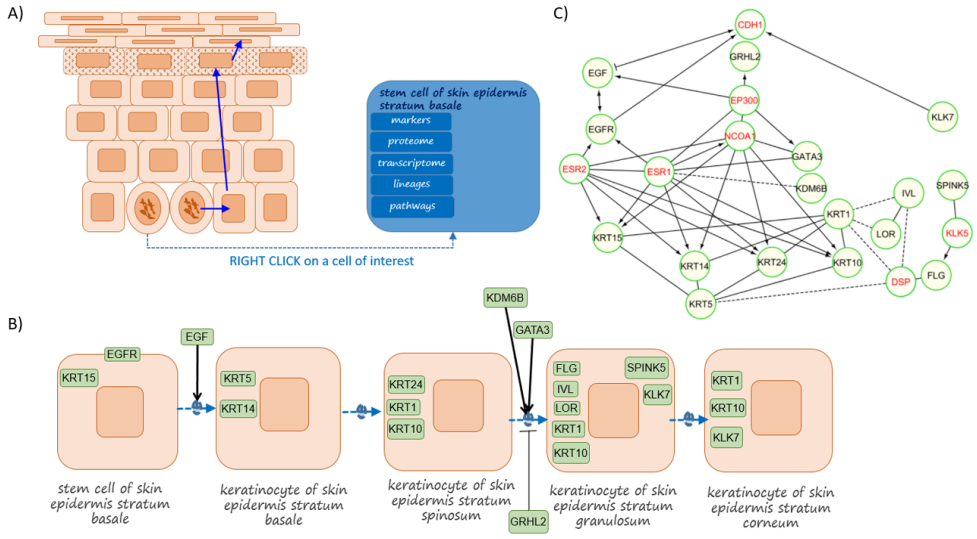


Figure 2. Development of keratinocytes in skin epidermis. A) Tissue-level view, where users can select a tissue's cell lineages and individual cells to access lists of annotated biomarkers, omics sets, and cell type-specific Reactome pathway viewer. B) Individual cell lineage path view where users can visualize and access detailed information on cell types and development steps that constitute a lineage, including biomarkers and development step regulators. C) Gene regulatory network view of the cell lineage path. Black labeled genes are path-derived biomarkers and regulators. Red labeled genes are linker genes.

Cell lineage paths are grouped into an expanded Developmental Biology chapter and are available among the browsable pathways on Reactome's website (currently available only on www.curator.reactome.org).

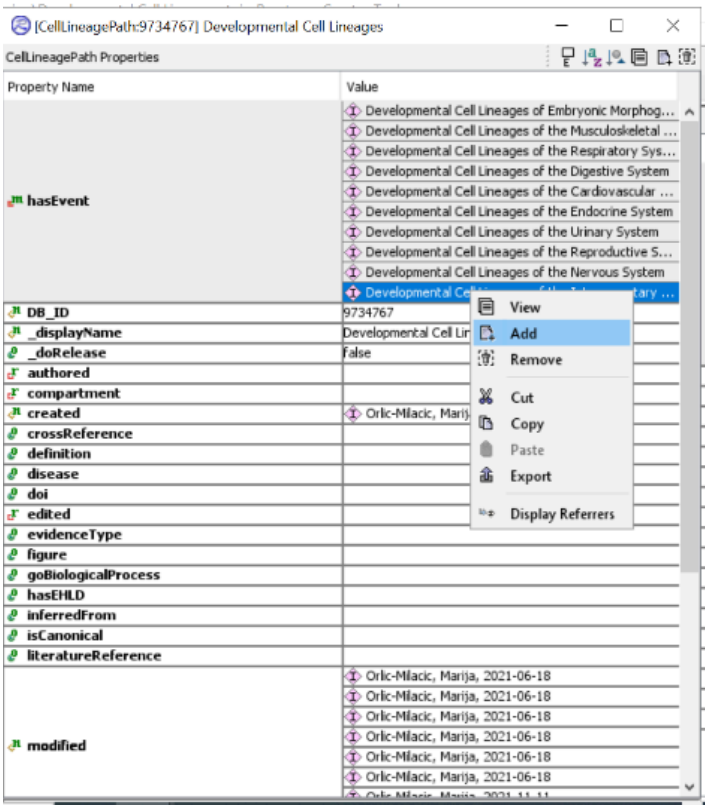
A detailed description of visual display and new website features is available in Cytomics Documentation (Appendix C9) written by Joel Weiser who worked on the new visual features with Guanming Wu.

Step-by-Step Annotation

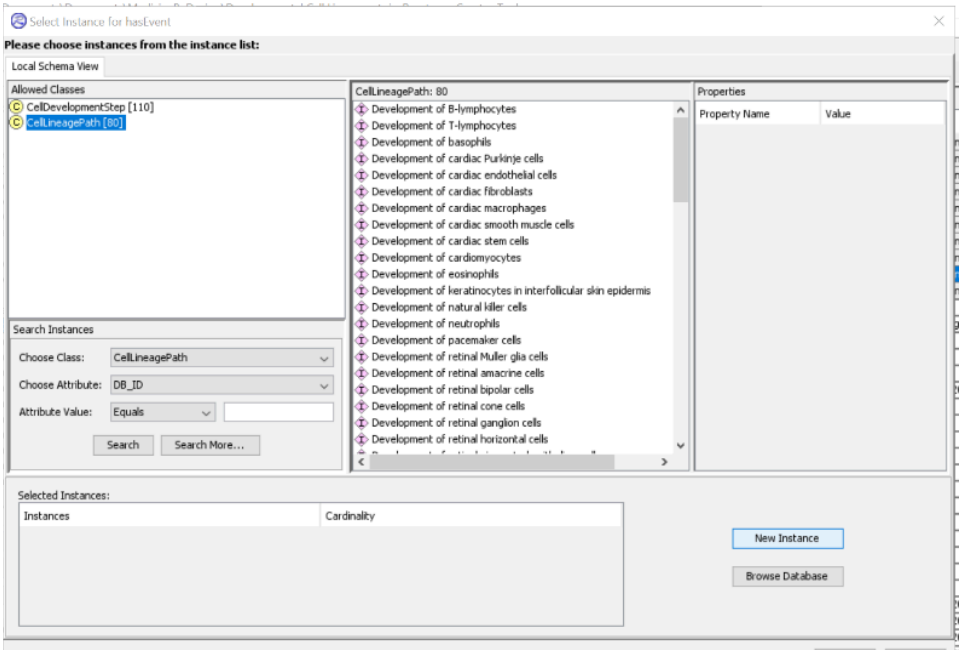
TissueLayer, Tissue and Organ attributes of Cells are populated using terms from public open source databases, the Cell Ontology (Sarntivijai et al. 2014; Osumi-Sutherland 2017) and UBERON (Haendel et al. 2014), and cross-referenced to these databases. ProteinMarker and RNAMarker attributes were intended to be semi-automatically populated from published immunohistochemistry and RNA in situ hybridization data, using the Reach tool natural language processing system (Valenzuela-Escárcega et al. 2018) and manual curation. Currently, only manual annotation of protein and RNA markers is functional. A good reference resource is the CellMarker database (Zhang et al. 2019; http://biocc.hrbmu.edu.cn/CellMarker/).

Annotation of Cell Lineage Paths

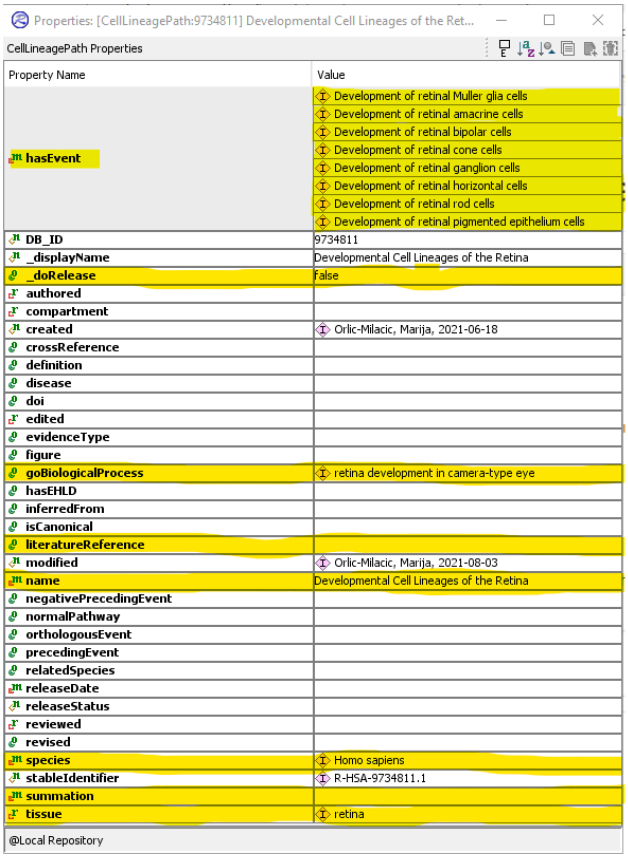
The first step in the creation of a new cell lineage path is to check out a parent cell lineage path, either Developmental Cell Lineages (DB_ID = 9734767) or any of its children into your local Curator Tool project. Open the parent cell lineage path record, select the hasEvent attribute field, right click and choose Add.



In the window that pops up, select CellLineagePath and click on New Instance.



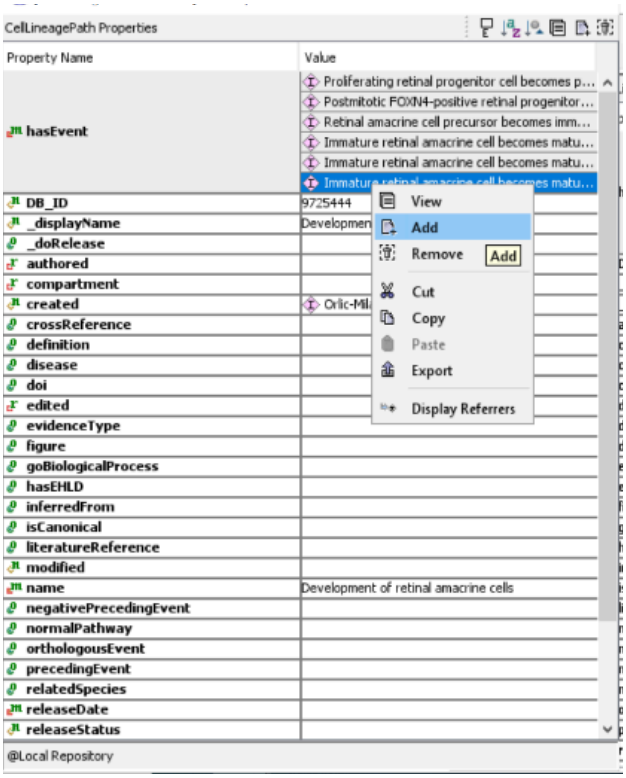
In the new window that appears, fill in the attributes of the cell lineage path, such as **name**, **species**, **summation**, **literatureReference**, **_doRelease**, **goBiologicalProcess**, **tissue**, and **hasEvent**. The hasEvent attribute(s) can be cellLineagePath(s) and cellDevelopmentStep(s):



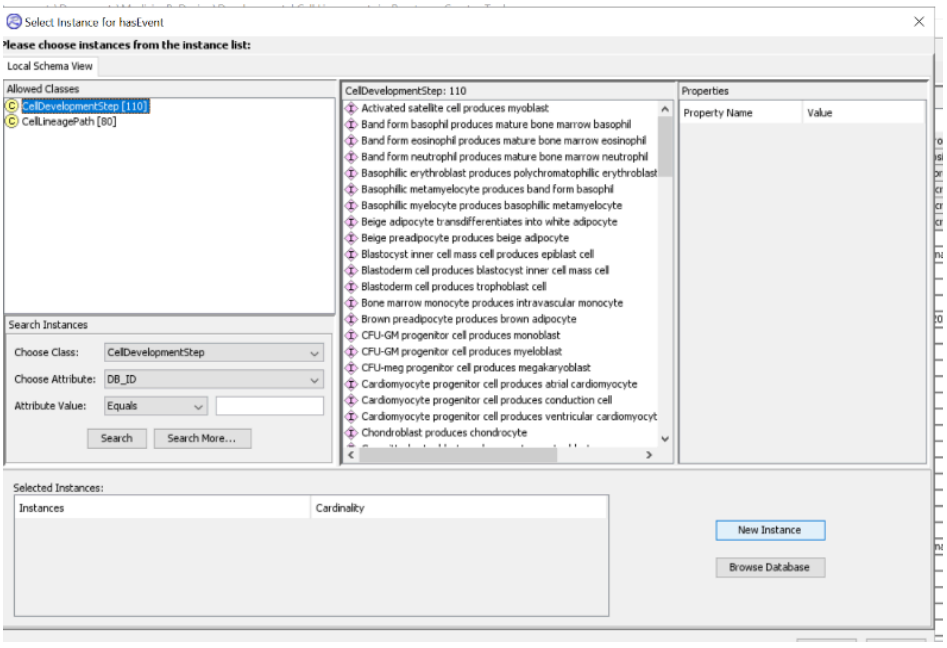
Annotation of Cell Development Steps

The first step in the annotation of a new cell development step is to identify, from published literature, the input cell type, the output cell type, regulators of the development step (such as growth factors), and required components of the input cell (molecules that regulators act on, such as growth factor receptors).

Open the parent cell lineage path record that a new cell development step belongs to, select the hasEvent attribute field, right click and choose Add.

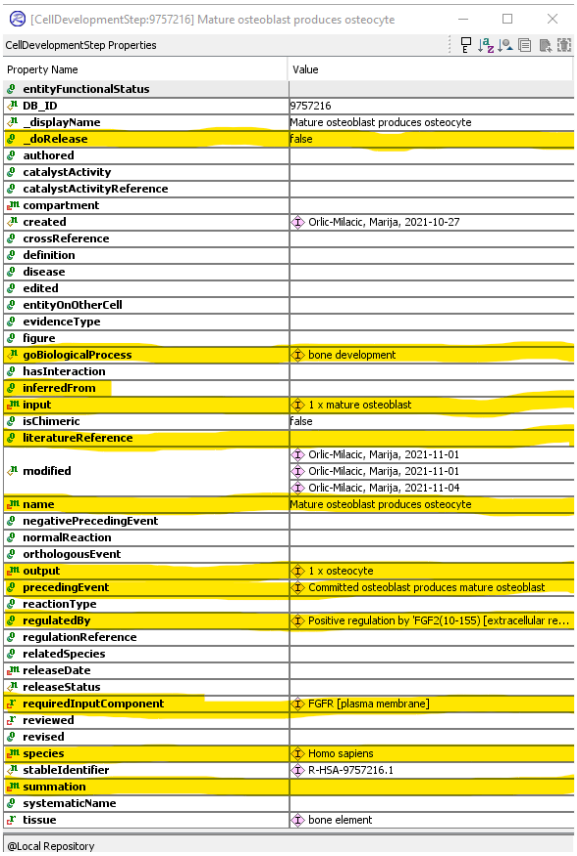


In the window that pops up, select CellDevelopmentStep and click on New Instance.



In the new window that appears, fill in the attributes of the cell development step, such as **name**, **species**, **summation**, **literatureReference**, **_doRelease**,

goBiologicalProcess, **tissue**, **input**, **output**, **precedingEvent**, **regulatedBy**, **regulationReference**, and **requiredInputComponent**.

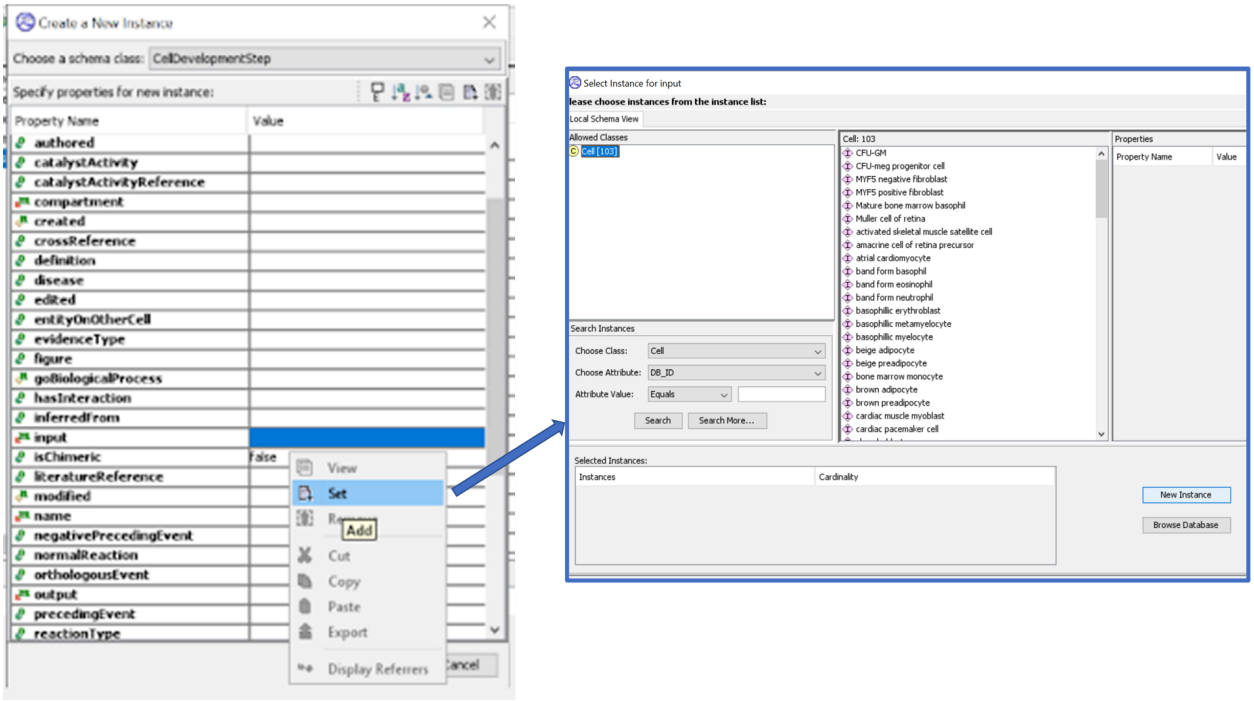


If there is no sufficient human-based experimental evidence available, a cell development step can be inferred from a model organism, similar to a reactionLikeEvent.

i. Inputs

A single input of the Cell class is to be defined for each cell development step.

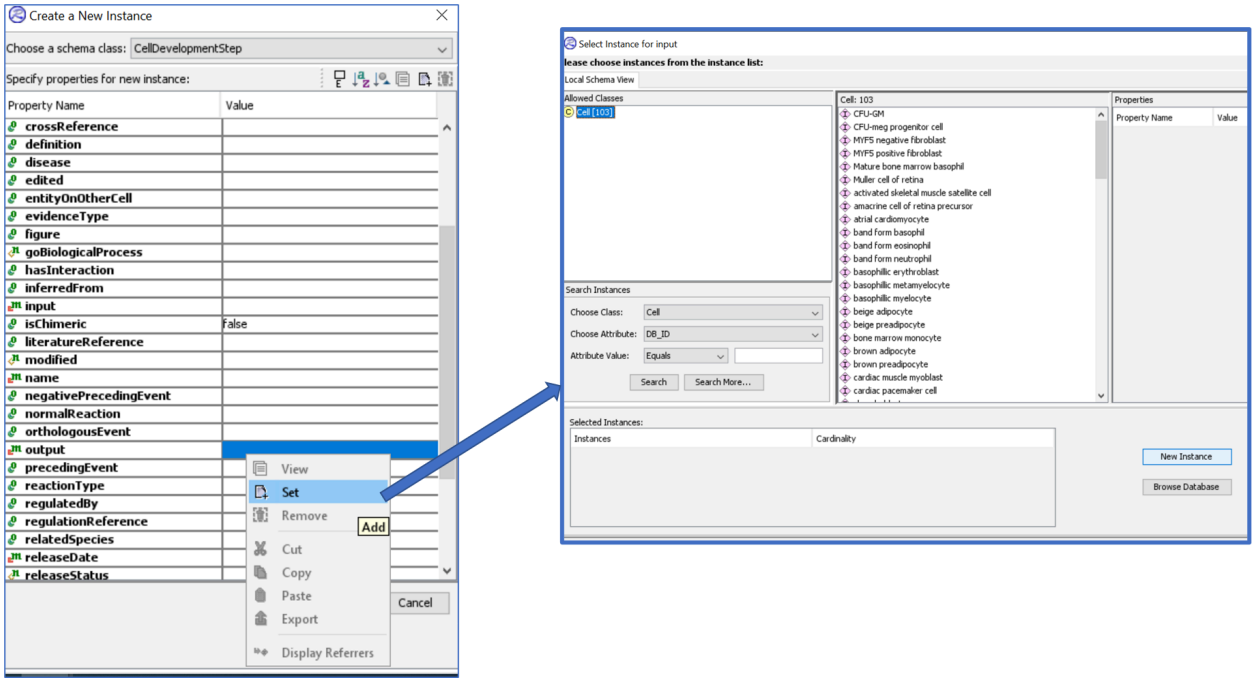
In a cellDevelopmentStep record, right click on the **input** field and select “Set”. From the window that pops up, you can choose from already annotated Cells or you can create a new cell instance.



ii. Outputs

A single output of the Cell class is to be defined for each cell development step. If an input cell can produce more than one output cell type, each input-output cell relationship is to be defined in a separate cell development step.

In a cellDevelopmentStep record, right click on the **output** field and select “Set”. From the window that pops up, you can choose from already annotated Cells or you can create a new cell instance.



iii. Regulators

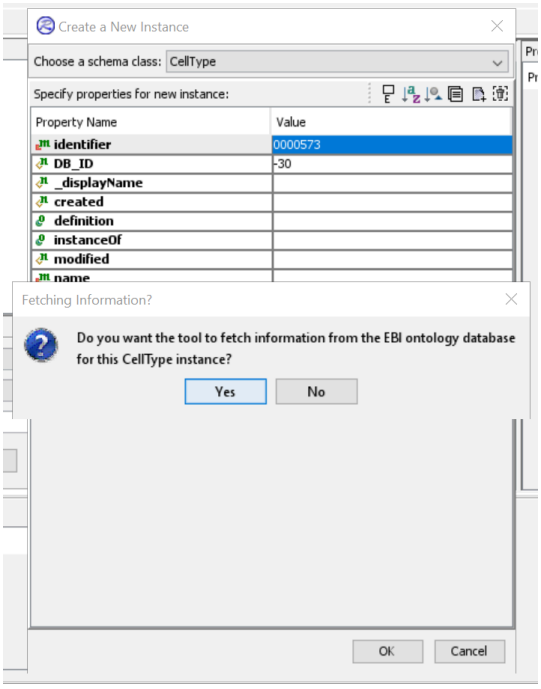
Any number of positive and negative regulators can be annotated for a cell development step using the **Regulation** subclasses - either **NegativeRegulation** or **PositiveRegulation**. Create the Regulation instances just like you would create them for a ReactionLikeEvent or use existing ones.

iv. Required Input Components

A required input component can be any PhysicalEntity present in the input cell that is needed for the action of the regulator or that was experimentally shown to be essential for the development step (e.g. a product of a gene whose conditional knockout abrogates the development step).

Annotation of Cells

A cell as a PhysicalEntity subclass is annotated by using a curator-selected Cell Ontology (https://www.ebi.ac.uk/ols4/ontologies/cl; CL) entry (identifiers starting with “CL_” prefix) as a reference **cellType**. When creating a new CellType instance, a curator needs to enter only the numerical part of the CL identifier, and the CellType record will be automatically populated:



One cellType can serve as a reference for multiple Cells, just like a referenceGeneProduct can serve as referenceEntity for multiple protein EWASs. The reason for this can be:

a) Cell Ontology does not have a more specific cellType term, so the curator has to use the closest match. For example, retinal horizontal cell (CL_0000745) is a reference cellType for three Cell instances: horizontal cell of retina precursor (DB_ID = 9725431), immature horizontal cell of retina (DB_ID = 9725440), and mature horizontal cell of retina (DB_ID = 9725415)

b) Two cells are closely related, but differ in their state, reflected in differential expression of key protein or RNA marker(s). The difference, although biologically relevant, is too miniscule to be captured by the CL. For example, retinal progenitor cell (CL_0002672) is a reference cellType for two Cell instances: postmitotic retinal progenitor cell PAX6+/FOXN4+ (DB_ID = 9725412) and postmitotic retinal progenitor cell PAX6+/FOXN4- (DB_ID = 9725454).

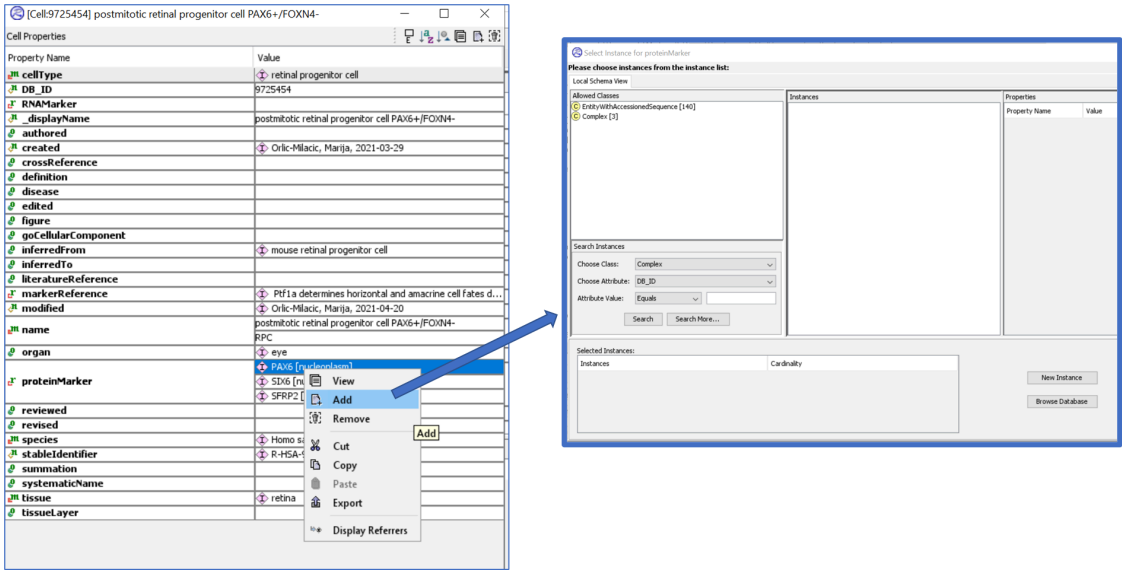
v. Protein Markers

Protein markers are proteins that are specific to the cell type and/or the cell state of the Cell or are differentially upregulated in the Cell. In immunohistochemical studies, protein markers can be used to distinguish the Cell from the neighboring cells. In

fluorescence-associated cell sorting (FACS), protein markers localizing to the cell surface can be used to isolate the cells of a desired type.

A useful search term for finding literature references with information on cell type-specific protein markers is [Cell / cellType _displayName] + marker; additional terms, such as “staining” or “sorting”, can be added to narrow down the search. The CL database rarely lists protein markers in a cell type description. The CellMarker database contains 13605 human cell markers (protein and RNA markers combined) across 467 human cell types, but markers are almost exclusively derived from high throughput studies (mostly scRNAseq) and are not reconciled between different studies for the same cell type (markers derived from each study are listed separately). CellMarker entries are, therefore, a good starting point to identify the most prominent markers, which overlap between studies (and, if possible, between human and mouse), and then use these markers + cellType _displayName as search terms. Please note that CellMarker also does not standardize gene names - they derive names directly from the study, and therefore an identical marker can be named several different ways on the cell type webpage (that is, the overlap between the studies may not be obvious - UniProt has to be consulted to retrieve the synonyms). Finally, the content of the CellMarker database has not been updated since it first went live in November 2018. It is uncertain if the database is actively curated or just maintained on the programmatic side (there have been some updates in terms of cross-linking to the CL database, which was not mentioned in their manuscript from January 2019).

To annotate a protein marker, open the Cell record. Right click on the proteinMarker field and select Add. You can currently choose a protein marker from the EWAS class or from the Complex class.

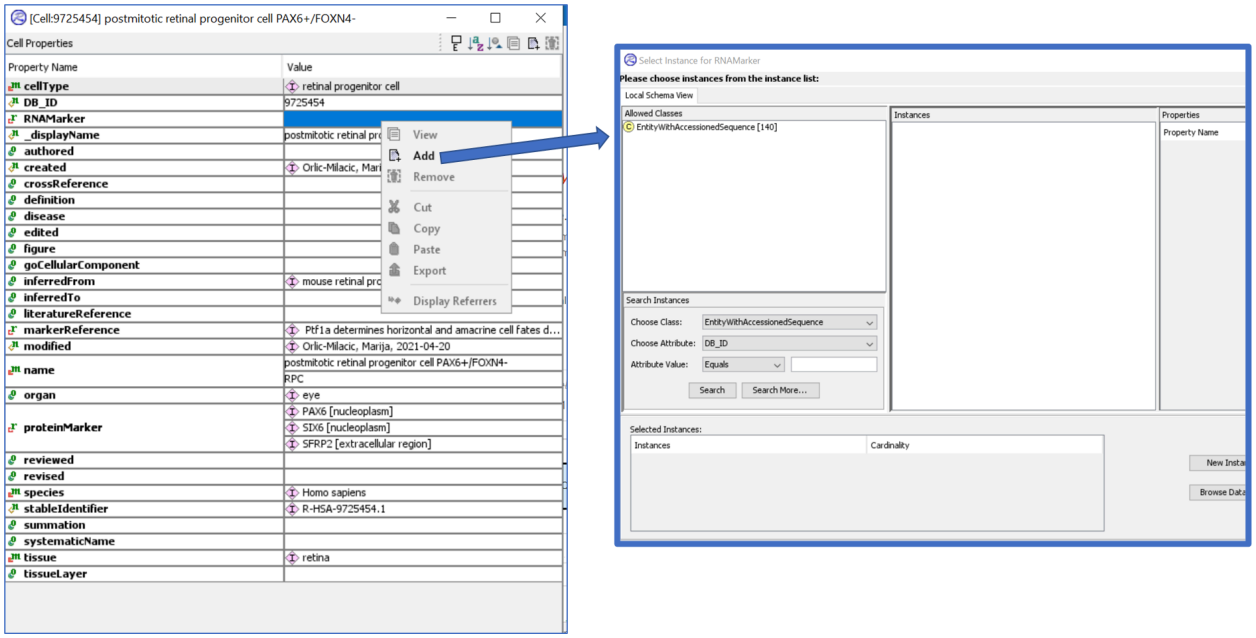


vi. RNA Markers

RNA markers are transcripts that are specific to the cell type and/or the cell state of the Cell or are differentially upregulated in the Cell. In RNA in situ hybridization studies, RNA markers can be used to distinguish the Cell from the neighboring cells. Increasingly, RNA markers are used to identify cells in scRNAseq experiments and to identify novel markers. Curators should evaluate if a novel RNA marker identified via scRNAseq or microarray expression analysis is robust enough to be annotated as an RNAMarker attribute of a Cell.

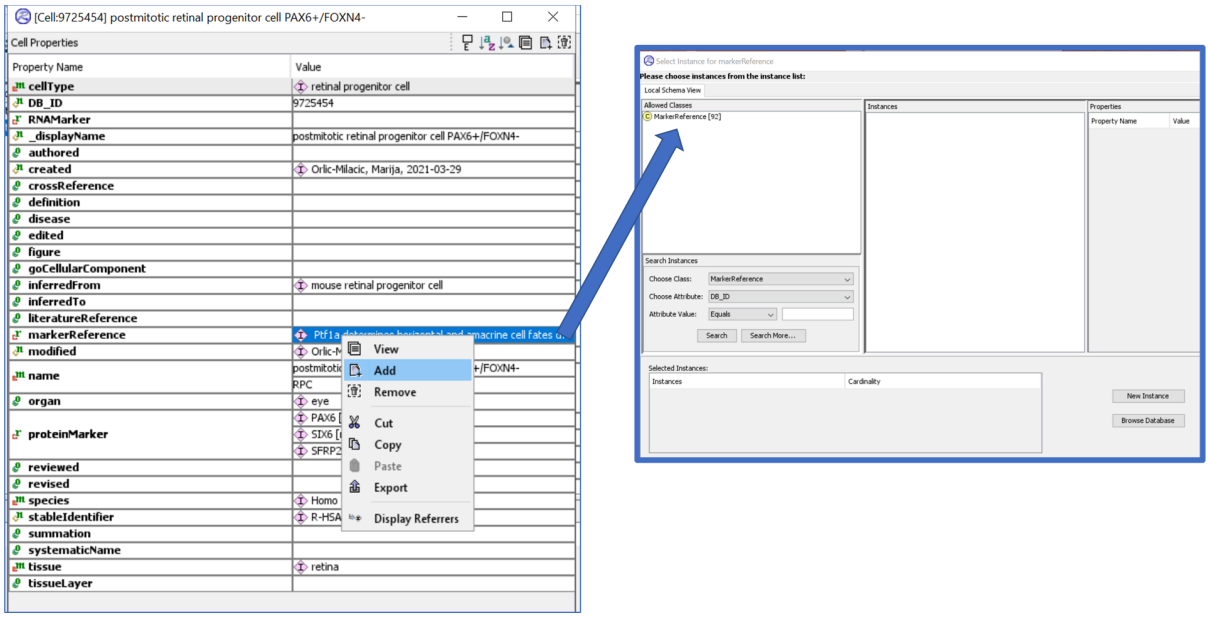
A useful search term for finding literature references with information on cell type-specific RNA markers is [Cell / cellType _displayName] + marker; additional terms, such as “hybridization” can be added to narrow down the search. The CellMarker database can be used as a starting point, as described above.

To annotate an RNA marker, open the Cell record. Right click on the RNAMarker field and select Add. RNA markers are restricted to the EWAS class.



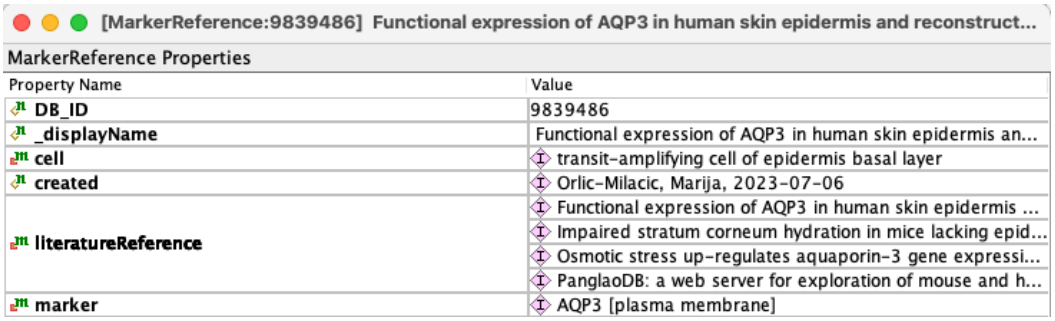
vii. Marker References

To associate literature references with individual protein or RNA markers, use the markerReference attribute of the Cell. Right clicking on the markerReference and selecting Add opens a new window in which MarkerReference is the only permitted class:



Select MarkerReference and click on New Instance. A new window will appear in which you can specify a **marker** (an RNA or a protein

EWAS), the **cell** it is a marker for and the **literatureReference(s)** supporting the assertion.. Note that a separate MarkerReference record is required for each marker annotated. Currently, marker in the MarkerReference record cannot be a protein complex, which is inconsistent with the option of specifying a complex as a proteinMarker in the Cell instance record.


[MarkerReference:9839486] Functional expression of AQP3 in human skin epidermis and reconstruct...

MarkerReference Properties

Property Name	Value
DB_ID	9839486
_displayName	Functional expression of AQP3 in human skin epidermis an...
cell	transit-amplifying cell of epidermis basal layer
created	Orlic-Milacic, Marija, 2023-07-06
literatureReference	Functional expression of AQP3 in human skin epidermis ...
	Impaired stratum corneum hydration in mice lacking epid...
	Osmotic stress up-regulates aquaporin-3 gene expressi...
	PanglaoDB: a web server for exploration of mouse and h...
marker	AQP3 [plasma membrane]

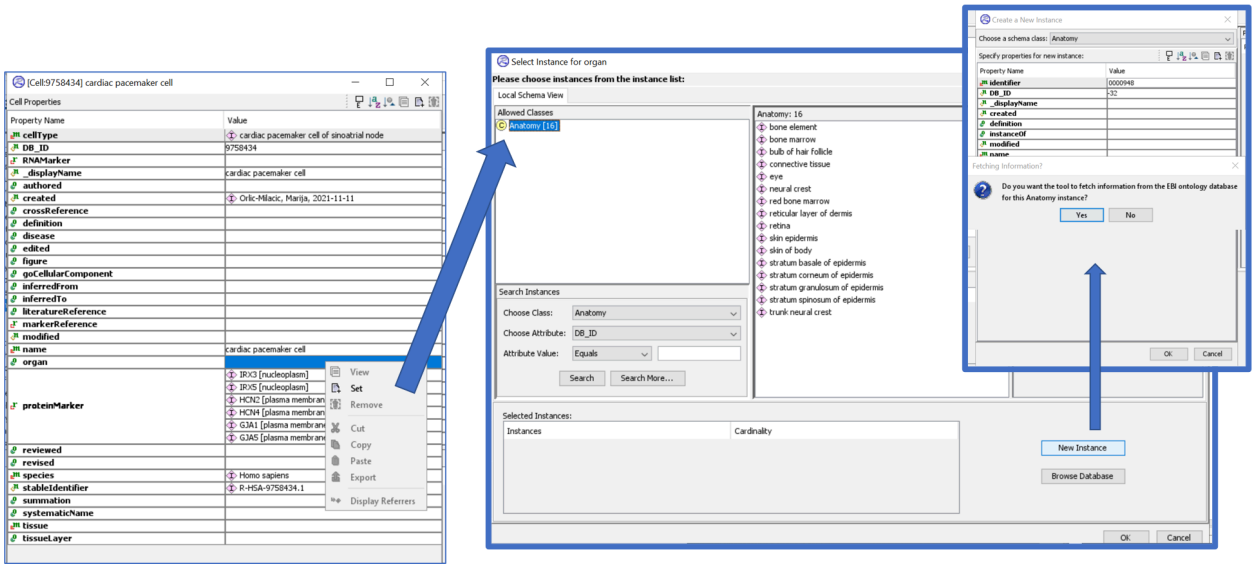
MarkerReference _displayName needs to be improved in the future editions of the Curator Tool. At the moment, the _displayName corresponds to the first literatureReference listed in the record, which makes it impossible to find relevant MarkerReference instances for a given marker by scanning the list of MarkerReferences. The optimal _displayName would be: marker._displayName: literatureReference.pubMedIdentifiers (separated by "," or "|" in case of multiple literatureReferences).

viii. Anatomical Localization

The UBERON Ontology (https://www.ebi.ac.uk/ols4/ontologies/uberon) is used as a reference database for anatomical localization. Curators can browse UBERON or use the search engine to find anatomical terms they are interested in. UBERON identifiers have a prefix in the form of either "UBERON_" or "UBERON:".

To annotate any anatomical attribute (tissue, tissueLayer or organ), a curator has to right click on the attribute in the instance record (e.g. Cell) and select Set (each anatomical attribute is

single-valued). In the new window that appears, a curator can create a new Anatomy class instance if a relevant instance is not available. Paste the numerical part of the UBERON identifier to the identifier field of the Anatomy record. The Curator Tool will offer to automatically populate the record.



1. *Tissue*
   Tissue is the mandatory attribute of Cell instances, and a required attribute of CellDevelopmentStep and CellLineagePath instances. When looking for a reference Anatomy term in UBERON, curators should make sure that the Anatomy term is defined as a tissue by UBERON, and not an organ or a tissue layer. For a reference, curators can also consult histology textbooks.
2. *Tissue Layer*
   TissueLayer is an optional attribute of Cell instances. For some highly organized tissues, such as skin epidermis or retina, it makes sense to fill in this attribute. For some other tissues, this attribute may be irrelevant. When looking for a reference Anatomy term in UBERON, curators should make sure that the Anatomy term is defined as a tissue layer by UBERON, and not a tissue or an organ For a reference, curators can also consult histology textbooks, as mentioned above.
3. *Organ*

Organ is an optional attribute of Cell instances. When looking for a reference Anatomy term in UBERON, curators should make sure that the Anatomy term is defined as an organ by UBERON, and not a tissue or a tissue layer. For a reference, curators can also consult histology textbooks, as mentioned above.

References

1. Burgess, Darren J. 2018. “Tracing Cell-Lineage Histories.” *Nature Reviews. Genetics*.
2. Croft, David, Gavin O’Kelly, Guanming Wu, Robin Haw, Marc Gillespie, Lisa Matthews, Michael Caudy, et al. 2011. “Reactome: A Database of Reactions, Pathways and Biological Processes.” *Nucleic Acids Research* 39 (Database issue): D691–97.
3. Haendel, Melissa A., James P. Balhoff, Frederic B. Bastian, David C. Blackburn, Judith A. Blake, Yvonne Bradford, Aurelie Comte, et al. 2014. “Unification of Multi-Species Vertebrate Anatomy Ontologies for Comparative Biology in Uberon.” *Journal of Biomedical Semantics* 5 (May): 21.
4. Joshi-Tope, G., M. Gillespie, I. Vastrik, P. D’Eustachio, E. Schmidt, B. de Bono, B. Jassal, et al. 2005. “Reactome: A Knowledgebase of Biological Pathways.” *Nucleic Acids Research* 33 (Database issue): D428–32.
5. Osumi-Sutherland, David. 2017. “Cell Ontology in an Age of Data-Driven Cell Classification.” *BMC Bioinformatics* 18 (Suppl 17): 558.
6. Saelens, Wouter, Robrecht Cannoodt, Helena Todorov, and Yvan Saeys. 2019. “A Comparison of Single-Cell Trajectory Inference Methods.” *Nature Biotechnology* 37 (5): 547–54.
7. Sarntivijai, Sirarat, Yu Lin, Zuoshuang Xiang, Terrence F. Meehan, Alexander D. Diehl, Uma D. Vempati, Stephan C. Schürer, et al. 2014. “CLO: The Cell Line Ontology.” *Journal of Biomedical Semantics* 5 (August): 37.
8. The Gene Ontology Consortium. 2019. “The Gene Ontology Resource: 20 Years and Still GOing Strong.” *Nucleic Acids Research* 47 (D1): D330–38.
9. Valenzuela-Escárcega, Marco A., Özgün Babur, Gus Hahn-Powell, Dane Bell, Thomas Hicks, Enrique Noriega-Atala, Xia Wang, Mihai Surdeanu, Emek Demir, and Clayton T. Morrison. 2018. “Large-Scale Automated Machine Reading Discovers New Cancer-Driving Mechanisms.” *Database: The Journal of Biological Databases and Curation* 2018 (January). https://doi.org/10.1093/database/bay098.
10. Wu, Guanming, and Robin Haw. 2017. “Functional Interaction Network Construction and Analysis for Disease Discovery.” *Methods in Molecular Biology* 1558: 235–53.
11. Wu, Guanming, and Lincoln Stein. 2012. “A Network Module-Based Method for

Identifying Cancer Prognostic Signatures." *Genome Biology* 13 (12): R112.
12. Zhang, Xinxin, Yujia Lan, Jinyuan Xu, Fei Quan, Erjie Zhao, Chunyu Deng, Tao Luo, et al. 2019. "CellMarker: A Manually Curated Resource of Cell Markers in Human and Mouse." *Nucleic Acids Research* 47 (D1): D721–28.

Disease Pathway Curation Guide

Introduction

Reactome disease pathways are contained in a separate top-level pathway called "Disease". Consistent with Reactome's pathway-centric view, disease pathways are not comprehensive descriptions of given diseases (bladder cancer, diabetes etc), but rather show the perturbations of a given normal pathway that have been characterized, and may include tags for multiple different diseases within one pathway.

Reactome's disease annotations are grouped by general biological processes ("Diseases of signal transduction by growth factor receptors and second messengers", "Diseases of metabolism", "Infectious diseases", etc.) rather than by specific diseases. The groupings generally mirror the pathway groupings of the normal event hierarchy.

At a high level, there are two main classes of Reactome disease pathway.

The first type of disease pathway has a "normal" counterpart and shows perturbations to a normal biological process that result from disease entities with altered behavior relative to the WT entities. This type includes, for instance, cancer pathways that arise as the result of gain- or loss-of-function mutations in oncogenes or tumor suppressor genes.

Both gain- and loss-of-function reactions with a corresponding normal biological pathway will be displayed (automatically highlighted in red by virtue of the disease tag, see below) in the context of the greyed-out WT pathway. Loss-of-function and some gain-of-function disease events are automatically overlaid on the corresponding normal reactions by virtue of their "normal reaction" attribute (see below). Other gain-of-function disease events must be manually laid out in the disease pathway, also described below.

The second type of disease pathway has no "normal" counterpart. This type includes, for instance, events that occur after the introduction and expression of foreign proteins encoded by genomes of infectious agents like viruses and intracellular parasites. These novel events must be manually laid out in the curator tool. Foreign/disease entities are automatically highlighted in red by virtue of the added disease tag (see below).

Because there is no normal WT pathway for these events, the ELV is unique for the disease pathway and does not have the greyed-out background.

While the bulk of disease annotation is the same as normal curation, there are a couple of disease-specific classes and fields that need to be filled out on entity and/or event records. These include:

- FailedReaction (class)
- Disease (field needed for disease entities and events)
- Functional status (field needed for disease RLEs involving genetic disease variants)
- Normal pathway (field needed for non-infectious disease pathways that have a normal pathway counterpart)
- Normal reaction (field needed for non-infectious disease RLEs that will be displayed as an automatic overlay on the corresponding normal reaction)

Relative to normal curation, disease curation also makes heavy use of sets.

Disease curation also requires extra layers of subpathways to make the display work properly, and a corresponding increase in the number of EHLDs. This is discussed more fully below.

This disease chapter has three main sections. The first, “Researching a disease”, briefly describes some useful resources for accessing disease information for curation. The second “Structuring the disease hierarchy and disease display” describes how to structure a disease pathway and how this shows up on the live site. The third section “annotating disease events and entities” has the nuts and bolts of disease event and entity curation.

If you are starting with disease curation for the first time, it is important to work closely with a curator that has experience in this area as you get started. It is also important to preserve the normal biology that has been curated and not try to adjust it to fit a disease process. This is because the curation process to date has gone through many changes to standardise it to be done in a consistent way. Any changes to the consensus could result in QA errors and/or affect scripts, download files and analyses based on curation. In the vast majority of cases, the normal biology curated is correct for basing its disease counterpart on it. However, the data model does not support annotation of mutations at the nucleic acid level so, for example, it is not possible to model a failedReaction on a normal gene expression event. Instead, the failedReaction should be modelled on the most proximal normal RLE that uses the gene product (protein) from the output of the gene expression event. Consult with a curator experienced in disease curation before undertaking such curation as it may avoid errors in this complicated process.

A detailed introductory video describing the annotation of disease variants can be found here (https://www.youtube.com/watch?v=0irXpoqiDp4).

Researching a disease

- Focus on biology you know well
- Make sure the biology is well characterized at the molecular level with good human biochemical evidence.
- It is generally useful to start off with a few good, high-level reviews to get an overview.
- For non-infectious disease curation, make sure that the normal biological process is adequately covered in its own non-disease pathway before starting the disease annotations. This is because disease pathways and events are overlaid on the normal events for display in the curator tool and the website. It is possible to add a few normal events to the normal pathway if you discover they are needed to support the disease curation, but simultaneously curating both the normal and the disease events is a big job and should be avoided.
- Infectious diseases are novel events that have their own diagrams, independent of underlying normal biology and curated pathways. If you plan to annotate host-pathogen interactions, it can be useful to check that the impacted normal human pathways show the biology you need to reference.

Resources

OMIM

OMIM is the classic resource. The entry for SERPINC1 (antithrombin III) is a good example. The text rambles (it's a cumulative historical narrative of relevant published results) but as an annotated bibliography it is generally quite complete, with very good coverage of papers describing molecular and cellular biology of the gene and molecular bases of associated disease. The entry for each gene includes a list of mutations with phenotypic effects (under the "table of contents" link on the right of the page, choose "allelic variants") and links to locus-specific databases which sometimes have much more extensive catalogues. (For a while, the catalogue that had the most extensive coverage for the most human genes was HGMD, accessible from OMIM "table of contents" -> "external links" -> "variation". It may still, but it's hard to tell because a license is needed to access it now.) It looks like searches in OMIM for a UniProt ID fail (e.g., searching for P01008 in UniProt does not return SERPINC1) but the UniProt record for a protein includes links to its OMIM record.

COSMIC

For cancer mutations, COSMIC is useful, since it provides information on mutations detected in individual tumors (far from ideal, not all types of mutations are equally covered). In addition, it is good to start with one or two good reviews just to get a bird's eye perspective of the landscape, and then to focus on 1-3 research publications where lots of patients were analyzed, which gives you an idea on the frequency of different mutations. This helps you create your mutant top list, since we aren't capturing all of the mutants, of course. After that, you follow your favourites through PubMed. If you're lucky, you may find a group that maintains a database of mutations for a specific gene.

Gene Reviews

Gene Reviews provides concise “reviews” of the genetics and clinical biology of many human diseases, as well as links to clinical testing laboratories, support groups, and other clinical resources that are not of much use to us. Its coverage is less comprehensive than OMIM’s and much more narrowly focused on material useful to the working doctor who has just encountered an affected patient and is wondering how to proceed.

IUPHAR/Guide to Pharmacology

Citing other resources

In certain cases, there are external databases that include information that is relevant to an annotated pathway.
One example is rare disease database such as RettBase, which describes Rett syndrome-related mutations in MECP2, FOXG1 and CDKL5: http://mecp2.chw.edu.au/. This database provides information about variants described in the disease pathway: Loss of function of MECP2 in Rett syndrome.
We are not able to cross-reference RettBase for individual MECP2, variants in the way that we cross-reference COSMIC for cancer variants because RettBase is not organized in a way that allows us to do this. However, we can mention RettBase in the pathway summation and cite their most recent publication (https://www.ncbi.nlm.nih.gov/pubmed/28544139) as a literatureReference for the Loss of function of MECP2 in Rett syndrome pathway. In addition, we can also create a second Publication instance of class "URL" and use this to cite database itself.

**NOTE that you should NOT not cite disease databases at the reaction level unless these disease databases capture biochemical/functional information.

Building a rough outline

After some preliminary research and reading, you will have a sense of what your disease pathway will look like:

- what normal pathway it is based on (in the case of non-infectious disease annotations)
- what kinds of mutations have been characterized (gain-of-function mutations, loss-of-function mutations)

- what normal functions are perturbed as a result of these genetic changes (ie, kinase-domain mutants in a receptor tyrosine kinase (RTK) don't trans-autophosphorylate and therefore downstream signaling is abrogated; constitutively active kinases phosphorylate their downstream substrates in an unregulated manner; binding domain mutants of a protein don't interact with their protein partners etc)
- drug:target interactions

Note that Reactome does not attempt to catalog every variant that has been identified for a given gene/protein. The focus is on capturing a few prevalent, well studied examples of each of the ways that the biological function of a protein can be altered in disease. Your disease pathway will likely involve sets of variants, each of which has been shown to act in a similar manner. Note that all variants with the same molecular consequence should be grouped together in a set, even if they occur in different diseases. See below for a more thorough description.

After this groundwork, you should be able to mock-up your disease pathway, grouping major classes of mutants under sub-headings/subpathways in your primary disease pathway.

As a hypothetical example, if you are annotating disease events in a RTK pathway, you might have in your outline:

SIGNALING BY RTK IN CANCER (parent pathway)

- constitutive signaling by overexpressed RTK (gain-of-function subpathway)
- constitutive signaling by RTK kinase domain mutants (gain-of-function subpathway with set of RTK mutants that have ligand-independent dimerization and activation)
- defective signaling by RTK kinase domain mutants (loss-of-function subpathway with set of RTK mutants that have lost kinase activity)
- signaling by RTK fusions (gain-of-function subpathway with set of fusion proteins that have ligand-independent dimerization and activation)
- RTK mutants bind tyrosine kinase inhibitors (essentially normal pathway showing binding of sensitive RTK mutants to known inhibitors)
- Drug resistance of RTK mutants (loss-of-function subpathway showing inability of RTK mutants to bind and be inhibited by specified TKIs)

Once you have a rough outline, you are ready to start annotating in the curator tool.

Assigning literature references for disease reactions

As with normal reactions, the references associated with a reaction MUST provide direct experimental evidence for the occurrence of that reaction in the species you are annotating (i.e. human for human Reactome). If there is no direct experimental evidence

in humans then you need to create an inferred human reaction as described in the section below. See Guidelines on evidence confidence here (Appendix C3).

Structuring the disease pathway

Ultimately, all disease pathways will end up under the “Disease” chapter on the website. You may have a general idea of where in the “Disease” hierarchy your pathway will end up, based on the process that you are annotating (“Diseases of Signal Transduction by growth factor receptors and second messengers” or “Diseases of metabolism”, for instance), or this may become more evident as your curation gets underway. This does not need to be decided before you start your curation.

To begin your disease curation, start by making your parent disease pathway. This parent pathway groups all the disease events related to a particular disease(s) that occur in the context of a given normal pathway (except in cases of infectious diseases, which are by-and-large independent of normal pathways).

For disease pathways without a corresponding “normal” pathway in Reactome

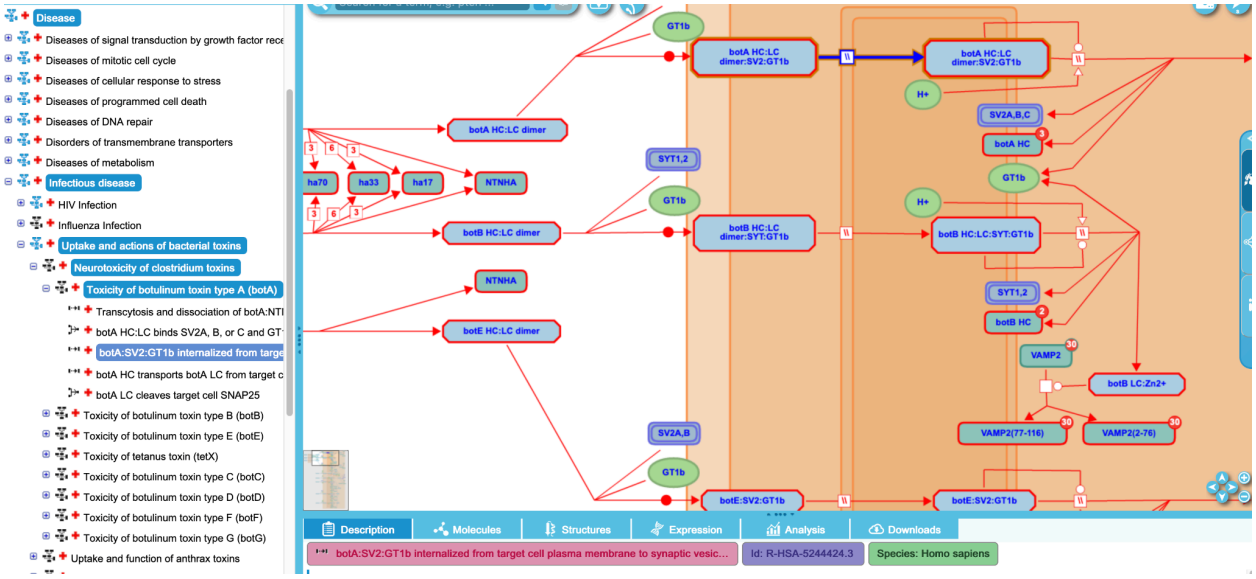
If the disease pathway will not have a corresponding "normal pathway" in Reactome, the pathway should still be placed under the disease chapter. Since these pathways are generally infectious processes, they will likely belong under the Disease child subpathway “Infectious disease”.

These pathways are labeled with an appropriate disease term (see below) and all species that contribute to the reactions and events (ie, generally, the human host and the infecting species, which is entered in the “related species slot, see below).

Pathway structure is much like that of normal curation. RLEs may be grouped into subpathways by the particular biology being described to facilitate comprehension and to manage diagram size, but because infectious disease pathways have their own diagrams and are not overlaid on the WT pathway, there are no particular requirements for how the pathways and subpathways are arranged.

Below is a screenshot of part of the “Toxicity of botulinum toxin type A” pathway, showing disease entities and disease reaction lines highlighted in red by virtue of the applied disease tag (described below). Host (human) proteins have no disease tags, are displayed bordered in black, but complexes of host and viral entities should be disease tagged and will therefore also appear bordered in red.

Because this is a disease-specific ELV, there is no greying out of a “normal pathway” background.



Below is a screenshot of the pathway record, showing the structure of the disease pathway. The parent pathway is labeled with a disease term (see below), and contains 5 disease events. The pathway is labeled with the two species that contribute to the reactions: Homo sapiens in the main species slot, and Clostridium botulinum in the “RelatedSpecies” slot. This class was introduced to prevent inappropriate inferring of pathways to the other species during website release. See below for more detailed description of the annotation of host:pathogen interactions.

Pathway Properties

Property Name	Value
deletedInstanceDB_ID	
hasEvent	Transcytosis and dissociation of botA:NTNHA:HA
	botA HC:LC binds SV2A, B, or C and GT1b on the target cell surface
	botA:SV2:GT1b internalized from target cell plasma membrane to synap...
	botA HC transports botA LC from target cell synaptic vesicle membrane i...
	botA LC cleaves target cell SNAP25
DB_ID	5250968
_displayName	Toxicity of botulinum toxin type A (botA)
_doRelease	true
authored	Krupa, S, Gopinathrao, G, 2006-06-15 10:36:11
compartment	
created	D'Eustachio, Peter, 2014-01-31
crossReference	
definition	
deletedStableIdentifier	
disease	botulism
name	Toxicity of botulinum toxin type A (botA)
negativePrecedingEvent	
normalPathway	
orthologousEvent	
precedingEvent	
previousReviewStatus	
relatedSpecies	Clostridium botulinum
releaseDate	2014-03-12
releaseStatus	
reviewStatus	five stars
reviewed	Ichtchenko, K, 2007-08-03 18:17:25
	Thirunavukkarasu, Nagarajan, Sharma, Shashi, 2014-11-17
revised	D'Eustachio, Peter, 2014-02-10
species	Homo sapiens
stableIdentifier	R-HSA-5250968.2
structureModified	
summation	Botulinum toxin type A (botA, also known as BoNT/A), a disul...

FOR DISEASE PATHWAYS WITH A NORMAL WT COUNTERPART

These disease pathways need to be linked to the WT ELV to allow proper display of these events. This is accomplished by structuring the disease pathway appropriately and by sharing the WT diagram with the disease pathway. Both of these are explained in more detail below.

To start, make a record in the curator tool for your top level pathway as usual.

- Give your pathway a name. The top level pathway name is often very general.
- Add a disease term from DOID in the appropriate pathway record slot. See below for guidelines on selecting an appropriate disease term. Adding the disease term triggers the automated coloring of the event or entity in red in the curator tool and in the pathway browser.
- Add the relevant wild-type pathway to the “normal pathway” field on the pathway record. This must be the “ELV-level” pathway (i.e it is listed as a RepresentedPathway for a PathwayDiagram) and contains the reaction-like events for which you intend to annotate corresponding disease events. This is needed to allow proper overlay of your disease events on the greyed-out background of the WT pathway.

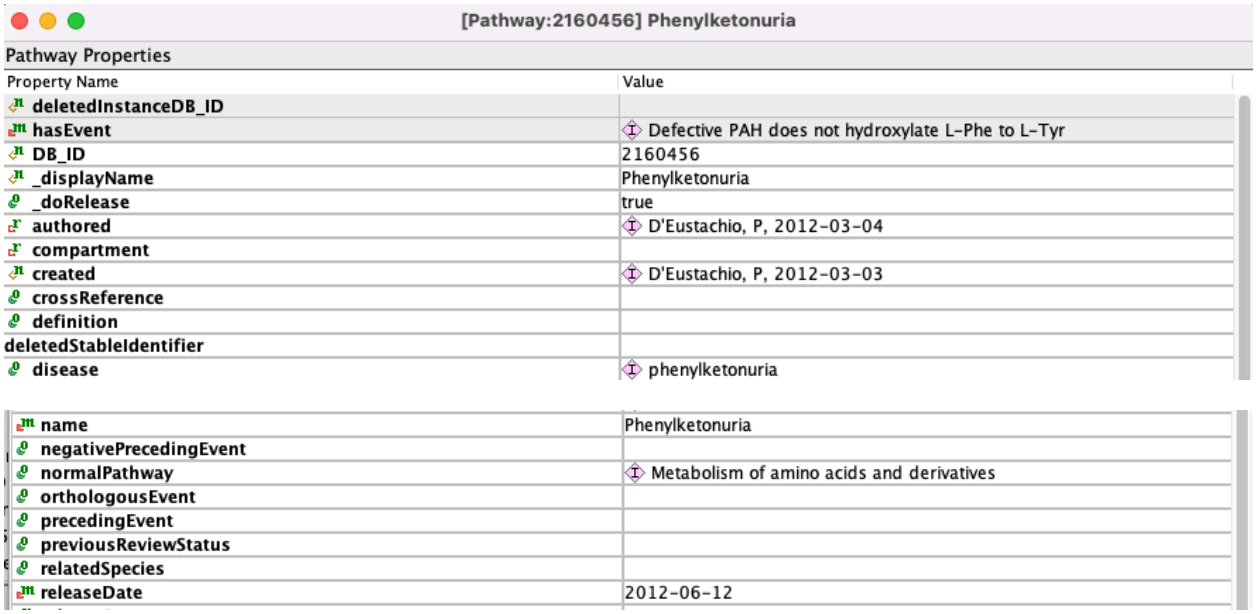
- (two other attributes, ‘entityFunctionalStatus’ and ‘normalReaction’ are also required for proper display of disease events, but these are applied at the RLE level not at the pathway level and will be described below.)

The sub-structure of the disease pathway (children, grandchildren) can be quite varied, depending on what the biology needs to represent.

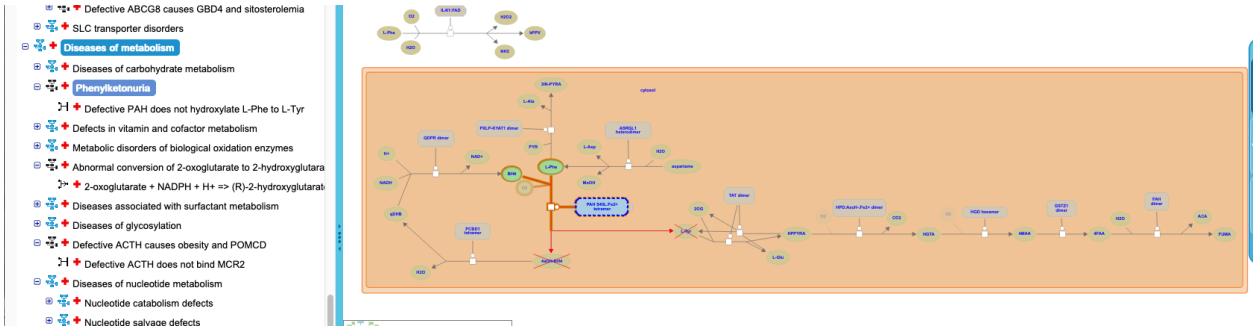
SIMPLEST PATHWAY STRUCTURE:

In the simplest case, there is one way a pathway can be altered during the course of a disease. This may be because a single variant in a single gene has been identified that affects the pathway in one way, or because there is a set of variants in a gene that have been identified that all have the same functional effect. In this case, all events can be shown directly under the parent disease pathway (grouped into subpathways if this grouping provides clarity to the user). All events will show up at the same time in a single view of the disease pathway.

As an example of this simple disease structure, look at the Phenylketonuria (R-HSA-2160456) pathway (under Diseases of metabolism).



Here, the parent disease pathway Phenylketonuria contains a single failed reaction showing the loss of function of the PAH variant PAH S40L. The pathway is tagged with the disease term “phenylketonuria” and the corresponding normal pathway (Metabolism of amino acids and derivatives) is added in the normalPathway attribute of the disease pathway record. The failed reaction (described more fully below) is shown in red automatically overlaid on the normal greyed-out pathway background (normal entities in the remainder of the WT pathway are non-interactive).



A simple disease pathway could also show the consequences of a gain-of-function mutation including the initiating event (for instance, ligand-independent dimerization and trans-autophosphorylation) and the downstream signaling events (binding to signaling partners, downstream phosphorylation events etc). All these RLE would be captured under the parent disease pathway and could be shown in the context of a single diagram, appropriately tagged with a disease term and the normal pathway.

More complex pathway structures

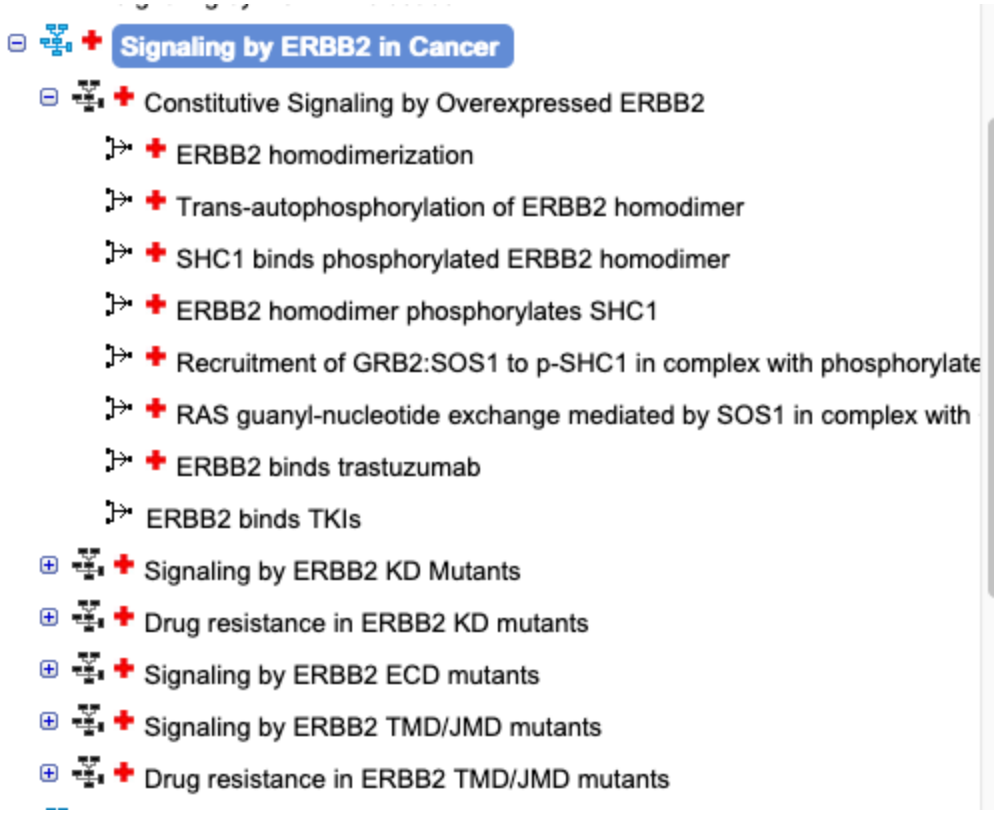
More often than not, however, a pathway can be altered in a variety of different ways: by mutation of a number of different proteins in the pathway, or by different types of mutations in the same protein that have different functional consequences, for instance. In these more complicated disease pathways, subpathways and sets of mutants become critical.

Reactome annotates the different functional consequences of different variants/sets of variants separately in their own view of the normal greyed-out WT pathway. For instance, a loss-of-function mutation in protein X would not be displayed in the same diagram view as a gain-of-function mutation in protein X that causes pathway activation- even though the affected pathway is (often, mostly) the same normal pathway.

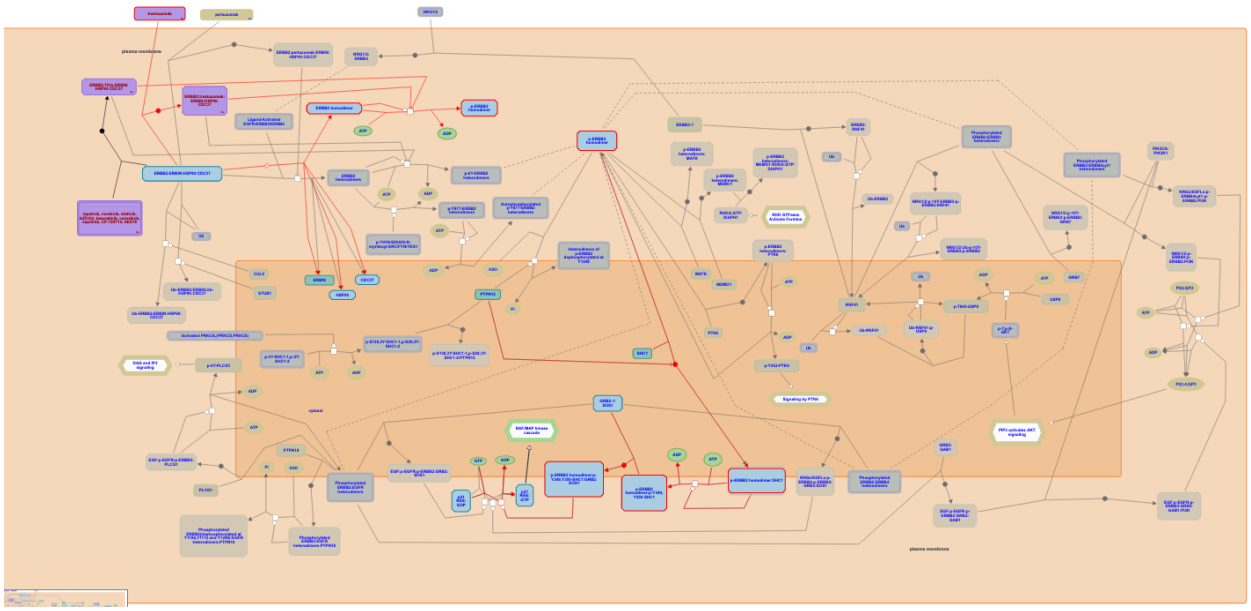
Variants (or sets of variants) that have the same functional consequence are grouped together and put in a sub-pathway that is distinct from sub-pathways showing the different functional consequences of other variants.

As an example, look at Signaling by ERBB2 in Cancer (R-HSA-1227990).

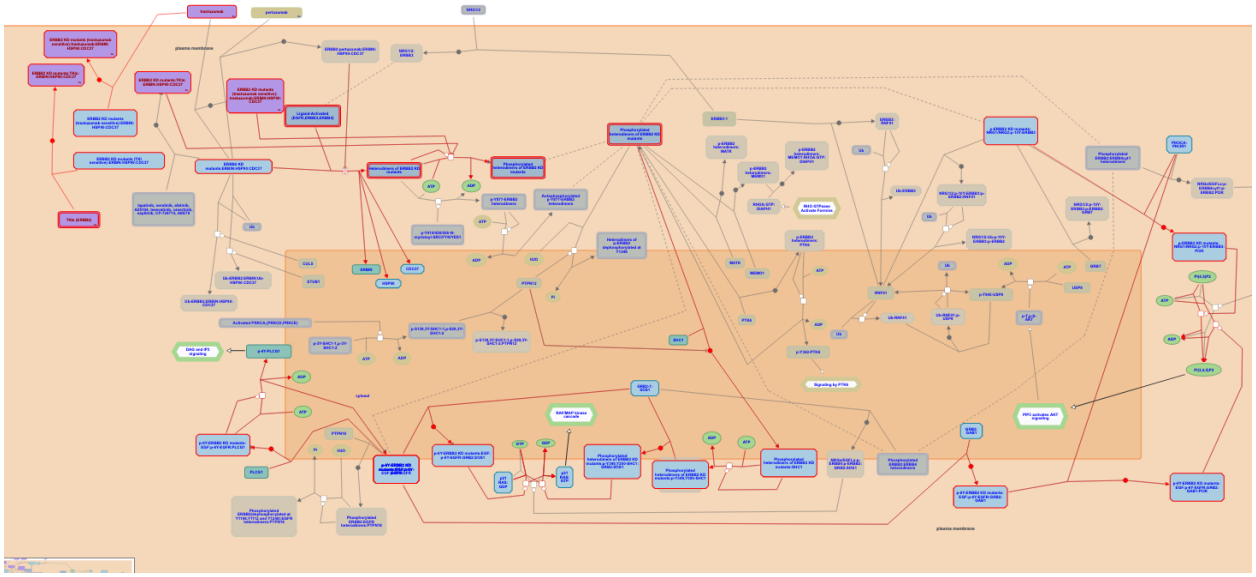
Here, the parent pathway is an EHLD. The pathway contains 6 child pathways, each describing different ways in which the same normal biology can be altered during cancer (Constitutive Signaling by Overexpressed ERBB2; Signaling by ERBB2 kinase domain (KD) mutants; Drug resistance in ERBB2 (KD) mutants (a special case, see below for structuring drug resistance annotations); Signaling by ERBB2 extracellular domain (ECD) mutants; Signaling by ERBB2 transmembrane domain/juxtamembrane domain (TMD/JMD) mutants; Drug resistance in ERBB2 TMD/JMD mutants).



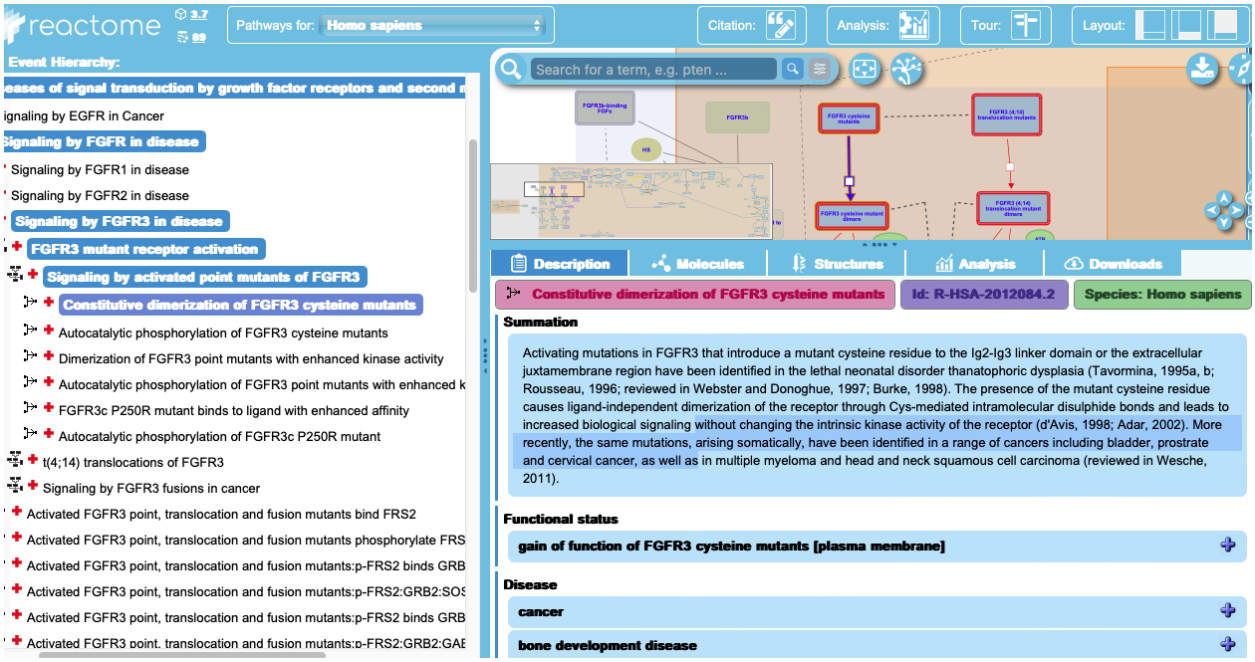
The disease reaction-like events of each of these child pathways (leaving aside the drug resistance pathways for the moment) are displayed in separate pathway views but overlaid on the same WT pathway diagram, as shown below for Constitutive Signaling by Overexpressed ERBB2:



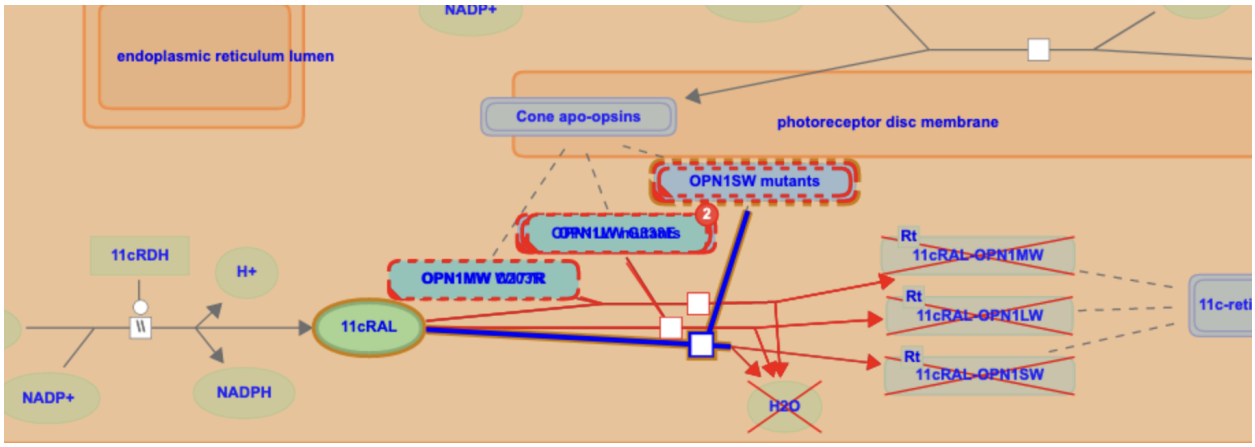
And for Signaling by ERBB2 KD mutants:



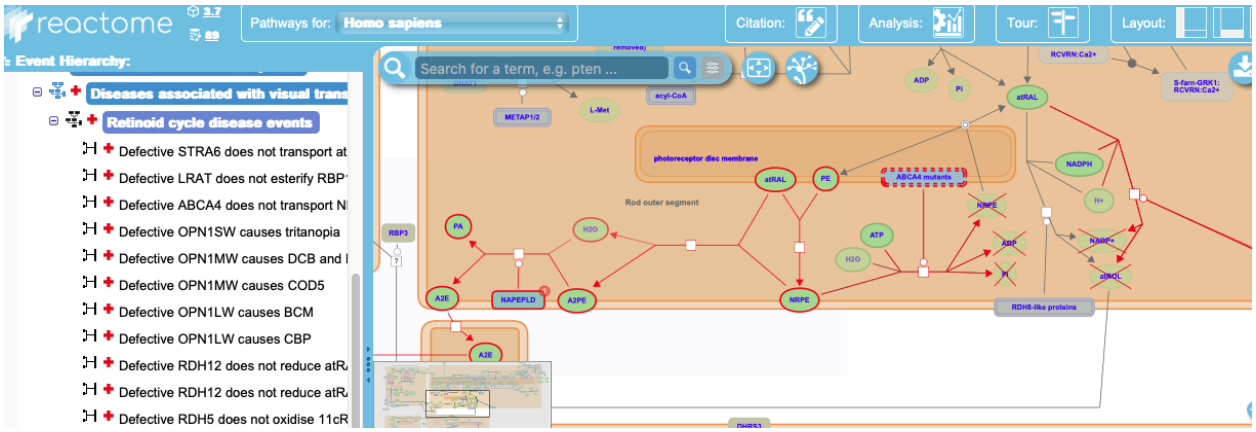
Separating functionally related variants into their own sets with their own diagram representations is clearer from a biological perspective, and should be done even if the variants in a given set arise in the context of multiple different diseases. For instance, mutations arise in FGFR3 in both bone development disorders and in cancer that allow the receptor to be constitutively active, dimerizing in the absence of ligands. These mutants, whether associated with bone disorders or cancer, are grouped together in a single set that is labeled with an appropriate disease term that encapsulates both diseases, or if that is so general a term as to be largely meaningless, with each of the relevant disease terms, as shown below:



Individual sets for groups of functionally related disease variants is also necessary to allow proper display of the disease events on the website. This is particularly true when the sets of variants affect the same WT RLE -in this case, separate sub-pathways are essential to allow proper display. If this is not done, the tool tries (successfully) to overlay multiple disease events onto a single reaction, making all of them illegible:



By consensus (as of August 2024), Reactome annotates loss-of-function failed reactions individually (ie each parent pathway containing a failed reaction should contain one and only one failed reaction). This is true even in cases where the failed reactions affect different reactions in the WT diagram and so could (unlike the case above) all be represented at the same time without ‘overlay clashing’. An old example with multiple failed reactions in a single parent pathway is shown in the pathway diagram below:



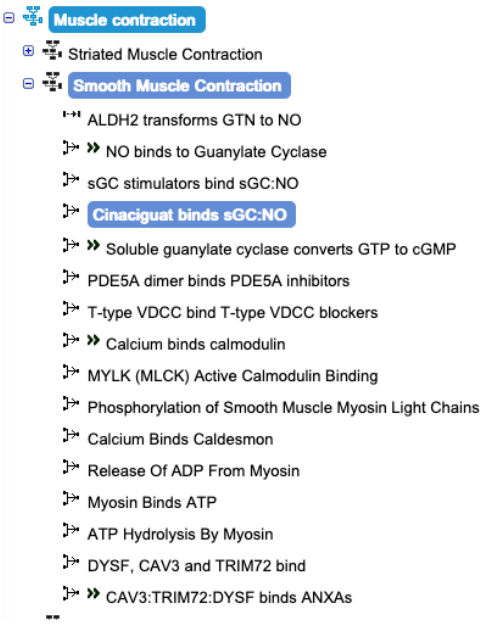
This convention was adopted to ensure consistent representation of failed reactions when taking into account the possibility of ‘overlay clashing’ when disease reactions need to be overlaid on the same WT reaction - and in spite of some user testing results that suggest that certain users would appreciate a way to combine different defective views into a single display to see all the weak points of a pathway.

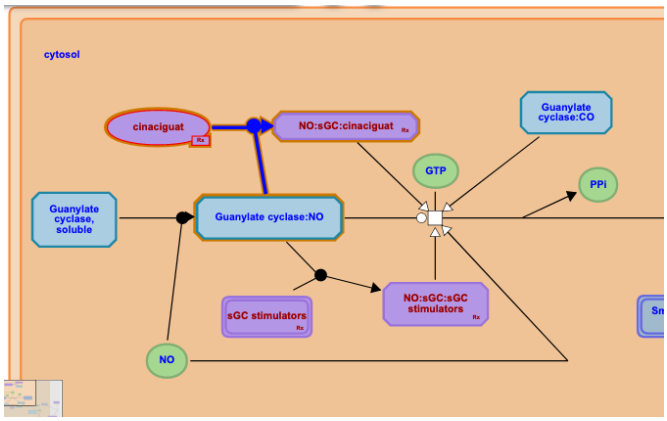
In contrast to loss-of-function reactions, a gain-of-function pathway can contain multiple reactions, as the effects of a gain-of-function mutation (or set of gain-of-function mutations) may propagate over many steps of the wild-type pathway. For instance, after ligand-independent dimerization and trans-autophosphorylation of an RTK due to mutation, the activated mutant receptor may recruit intracellular signaling proteins to propagate the signal. By convention, the disease pathway is continued until the mutant entity is no longer directly involved in the events that are being depicted.

Structure of pathways involving drug-binding

In Reactome, we annotate the action of drugs/therapeutics that affect pathway and protein activity. In general, drugs are shown binding to their target protein and then inhibiting or activating some other function of the target protein. See below for a full description of how to annotate drug:target interactions.

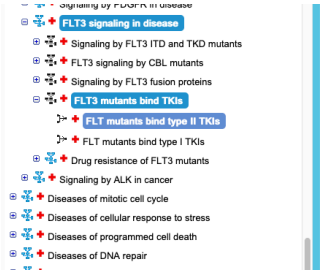
For drugs that affect the WT version of a protein, these binding and regulation events are shown in the context of the normal, WT pathway hierarchy and diagram. The drug physical entity and the binding event are labeled with an appropriate disease term, as described below. Drugs are also cross-referenced out to Guide to Pharmacology (GtoP) where possible, or to PubChem, as described below. Drug entities and related events are automatically colored in purple by virtue of these tags, and entities are decorated with the purple R_x icon at the bottom right, as shown below for the drug cinaciguat in the Smooth Muscle Contraction (R-HSA-9621179) pathway

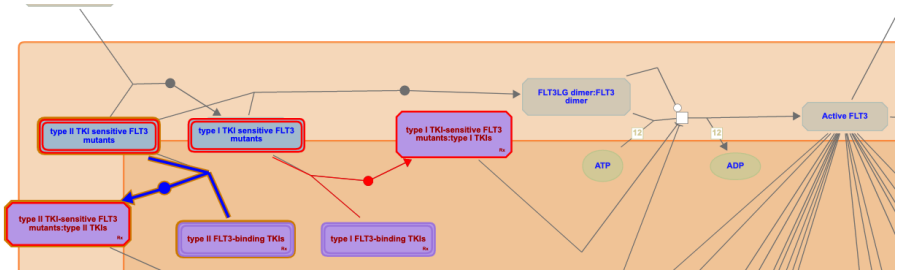




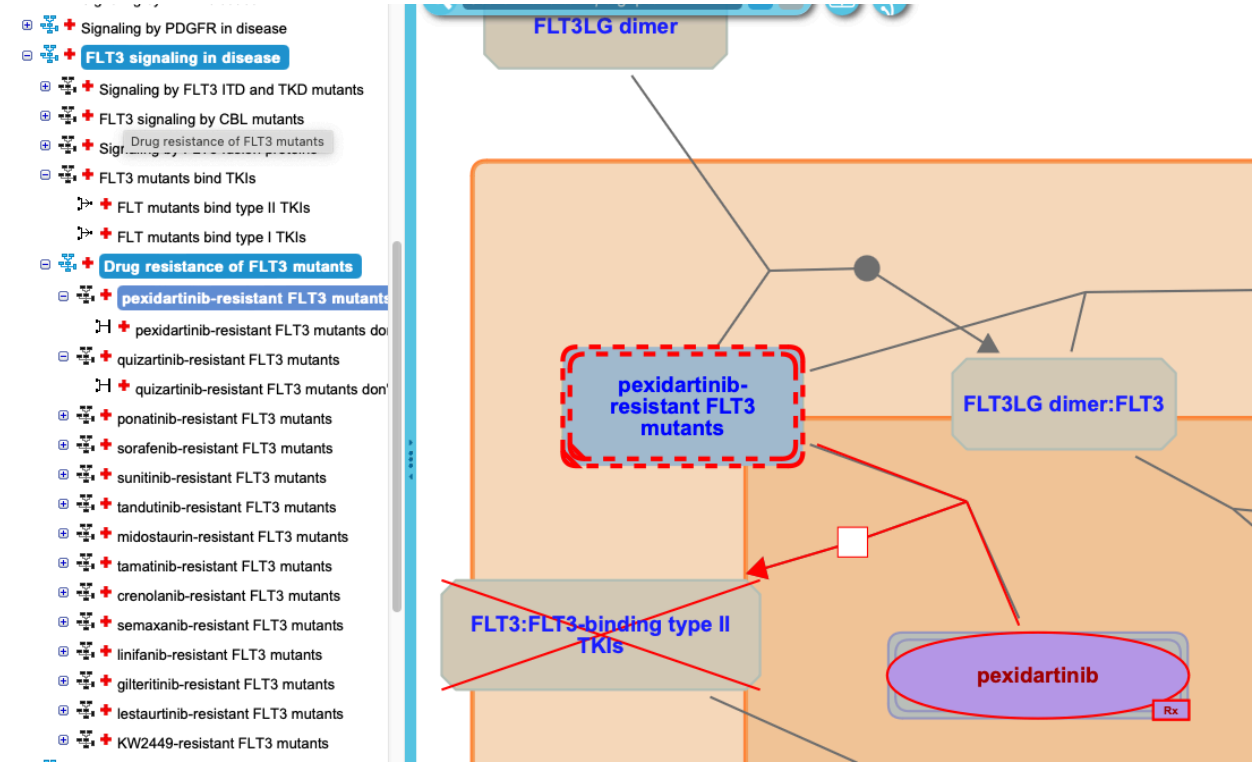
In cases where a drug has been shown to affect (or not affect) variant versions of proteins, Reactome disease pathways are able to capture this as well.

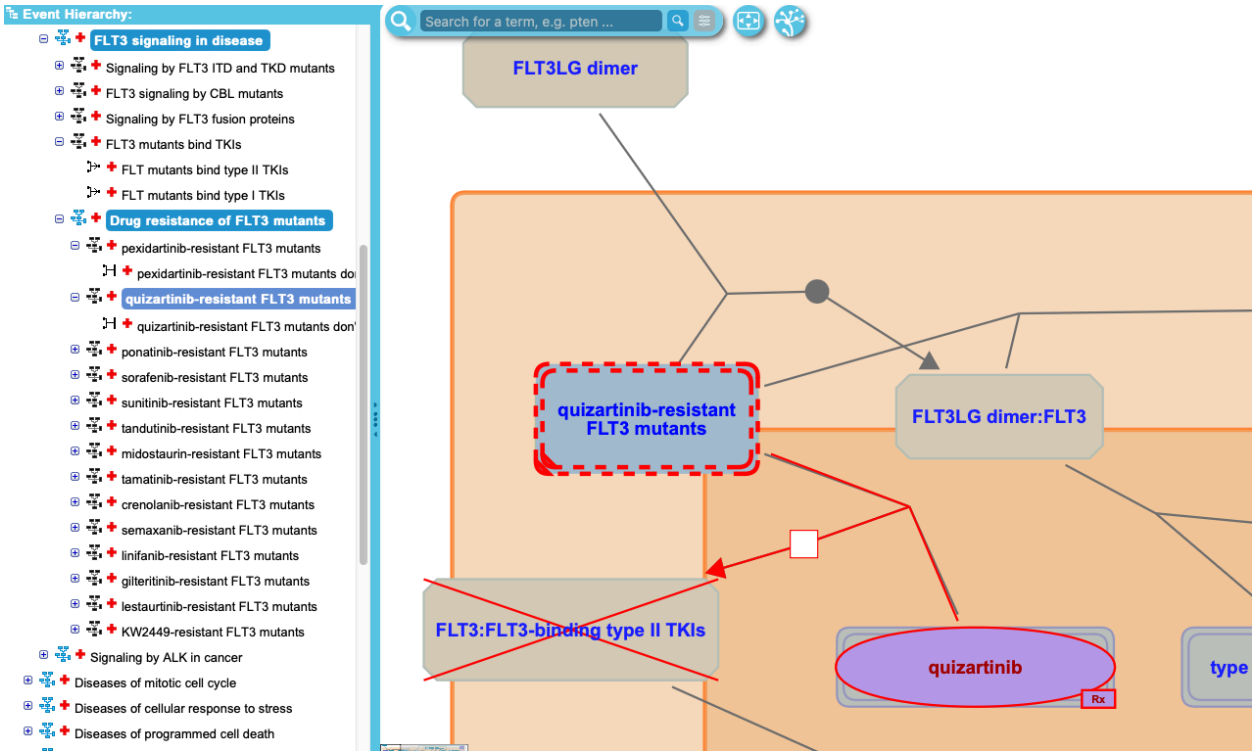
For drug-sensitive variants, the binding and regulation reactions are made in the same manner as for a drug that targets a wild-type protein, but the reactions are contained in the relevant disease pathway, as shown for FLT3 variants (R-HSA-9702510) in the screenshot below:



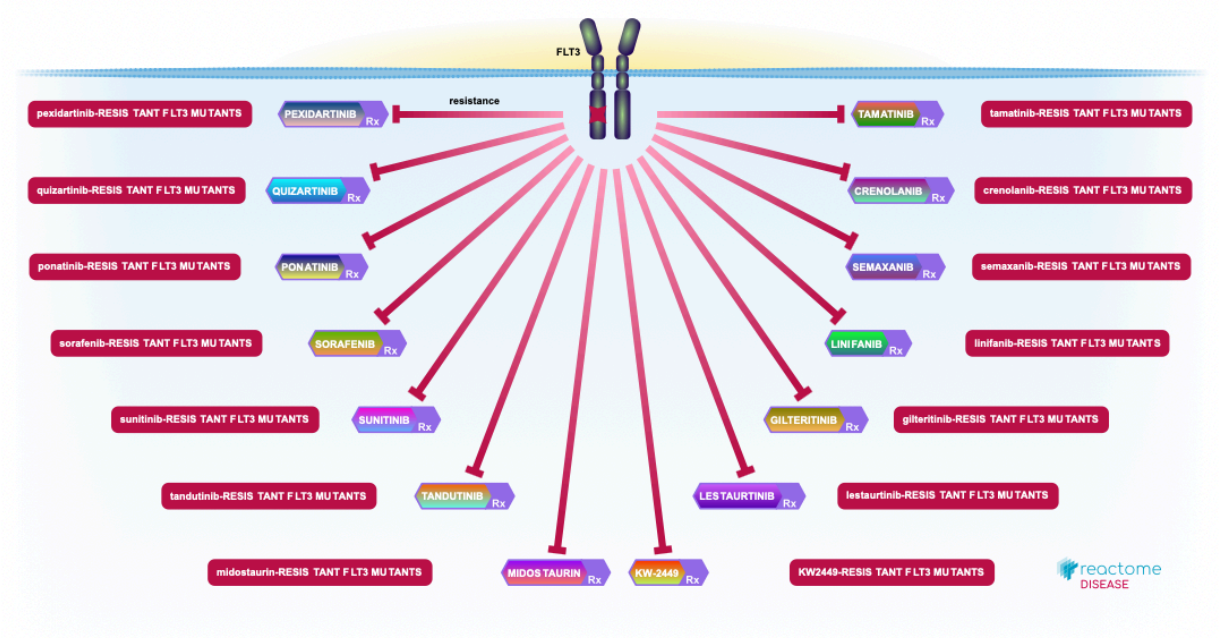


Drug resistant variants are shown as failed reactions, not binding to the relevant drug. Generally, sets of variants resistant to distinct drugs are represented individually in their own diagram views, as shown below for FLT3 variants resistant to pexidartinib and quizartinib, respectively:





In this case, the child pathway “Drug resistance of FLT3 mutants” is an EHLD level pathway.



The alternative to this representation would be to make a ‘superset’ of all the sub-sets of resistant mutants (“resistant to pexidartinib”, “resistant to quizartinib”, “resistant to ponitinib”) and make a single failed reaction where the superset was the input.

Sharing the diagram with disease pathways.

For disease pathways with no normal pathway

For infectious disease pathways, there is no sharing of WT diagrams. Each infectious disease process has its own diagram(s) with no ‘normal’ pathway slot filled in, and no need to share the WT pathway diagram.

For disease pathways with a normal WT counterpart

For disease pathways with a normal WT counterpart, you will need to share the WT pathway diagram with your disease pathway.

In principle, this is very simple. Go to the PathwayDiagram instance associated with your wild-type pathway and add your disease pathway name(s) to the multi-valued “representedPathway” slot.

In practice, which pathway to associate with the PathwayDiagram can be a bit more complicated. One way to think of the rule is: any disease pathways that directly contain disease RLE’s that need to be overlaid on the wild-type ELV need to be associated with the wild-type diagram.

Another way to think of this is: any RLEs that needs to be shown overlaid on the same WT reactions have to be in their own subpathway, and each of these subpathways need

to be added as a ‘representedPathway’ in the PathwayDiagram record for the wild-type pathway.

In practice, the disease pathways that are added to the pathwayDiagram record are generally children or grandchildren pathways of your parent pathway.

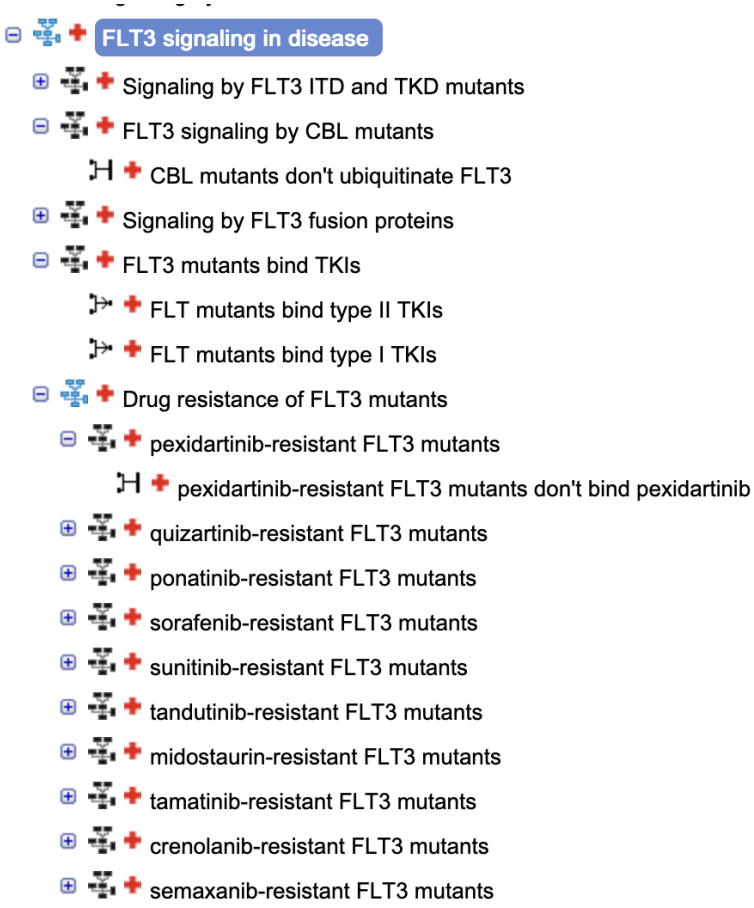
Consider the hypothetical pathway above:

SIGNALLING BY RTK IN CANCER (parent pathway): this will often be an EHLD depicting the clickable subpathways. The parent pathway is **not** added to the WT pathwayDiagram record as a represented pathway

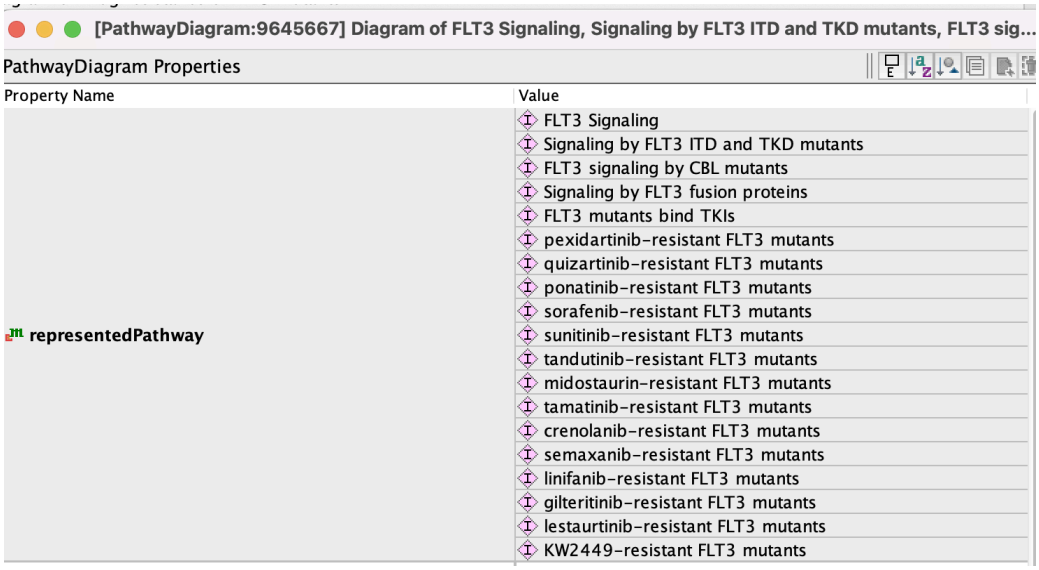
- constitutive signaling by overexpressed RTK (gain-of-function subpathway; pathway added to the WT pathwayDiagram record)
    - Individual RLEs, for instance
        - Constitutive dimerization of overexpressed RTK
        - Trans-autophosphorylation of overexpressed RTK etc
- constitutive signaling by RTK kinase domain mutants (gain-of-function subpathway with set of RTK mutants that have ligand-independent dimerization and activation; pathway added to the WT pathwayDiagram record)
    - Individual RLEs, for instance
        - Constitutive dimerization of KD mutants of RTK
        - Trans-autophosphorylation of KD mutants of RTK etc
- defective signaling by RTK kinase domain mutants (loss-of-function subpathway with set of RTK mutants that have lost kinase activity; pathway added to the WT pathwayDiagram record)
    - Individual RLEs (failedReactions), for instance
        - KD mutants of RTK don’t trans-autophosphorylate
- signaling by RTK fusions (gain-of-function subpathway with set of fusion proteins that have ligand-independent dimerization and activation; pathway added to the WT pathwayDiagram record)
    - Individual RLEs for instance
        - Fusion mutants of RTK dimerize
        - etc
- RTK mutants bind tyrosine kinase inhibitors (essentially normal pathway showing binding of sensitive RTK mutants to known inhibitors; pathway added to the WTpathwayDiagram record)

    - RLEs for instance
        - TKI-sensitive RTK mutants bind TKIs
- Drug resistance of RTK mutants (loss-of-function subpathways showing inability of RTK mutants to bind and be inhibited by specified TKIs; this will have a new EHLD with subpathway icons, pathway not added to the WT pathwayDiagram record)
    - Subpathway for each drug such as
        - Resistance of RTK mutants to TKI 1 (pathway added to the WT pathwayDiagram record)
            - Failed reaction, “TKI 1- resistant RTK mutants don’t bind TKI 1”
        - Resistance of RTK mutants to TKI 2 (pathway added to the WT pathwayDiagram record)
            - Failed reaction “TKI 2-resistant RTK mutants don’t bind TKI 2”

For a real example:

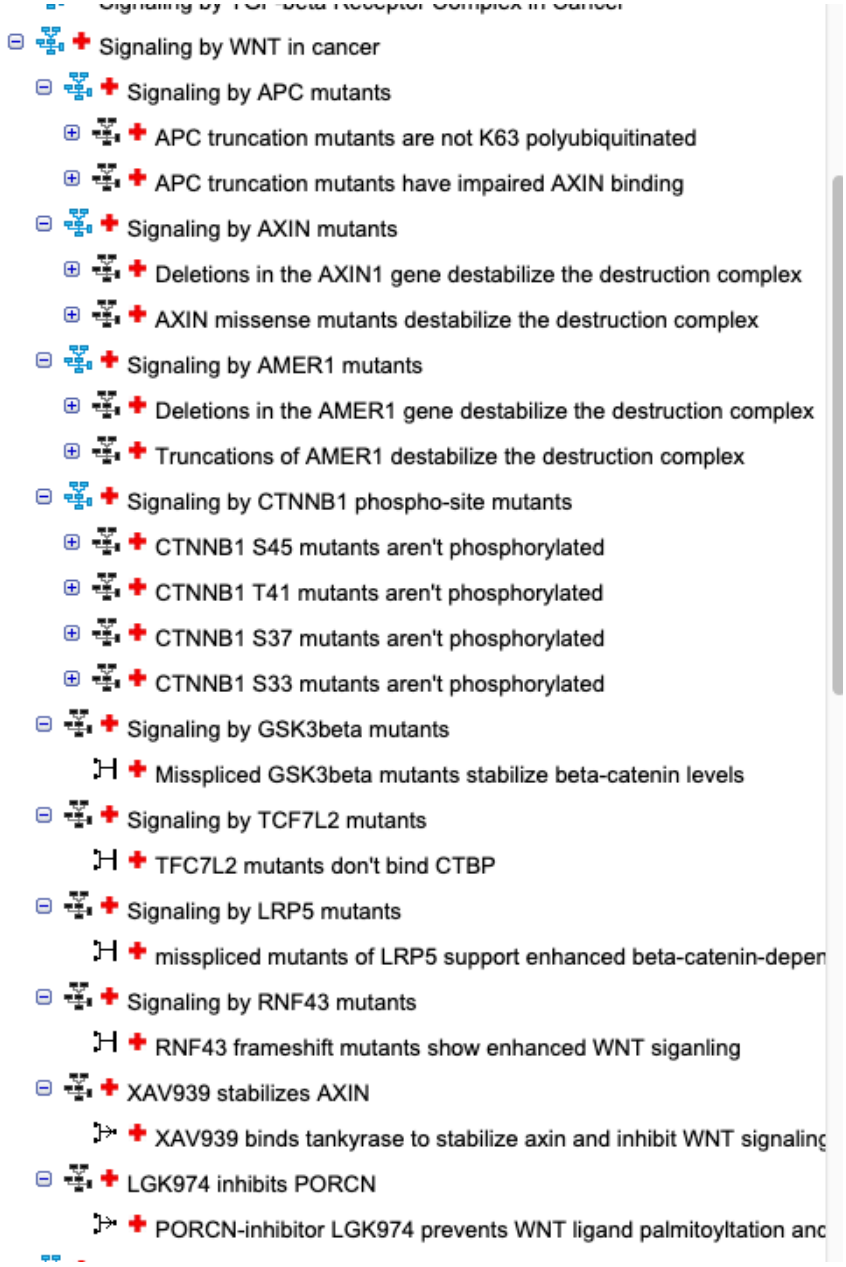


The corresponding PathwayDiagram record shown below lists first the WT pathway and then each of the disease RLE containing subpathways.



Note that the child pathway “Drug resistance of FLT3 mutants” has its own EHLD. It is a “grouping” pathway for all of the subpathways showing different drug resistant variants. This “grouping pathway itself does not have a normalpathway counterpart so is NOT associated with the WT pathway diagram. This is because each of the failed reactions (showing that the set of mutants is resistant to drug X, Y or Z etc) needs to be overlaid on the same WT reaction and therefore needs its own subpathway, as shown above.

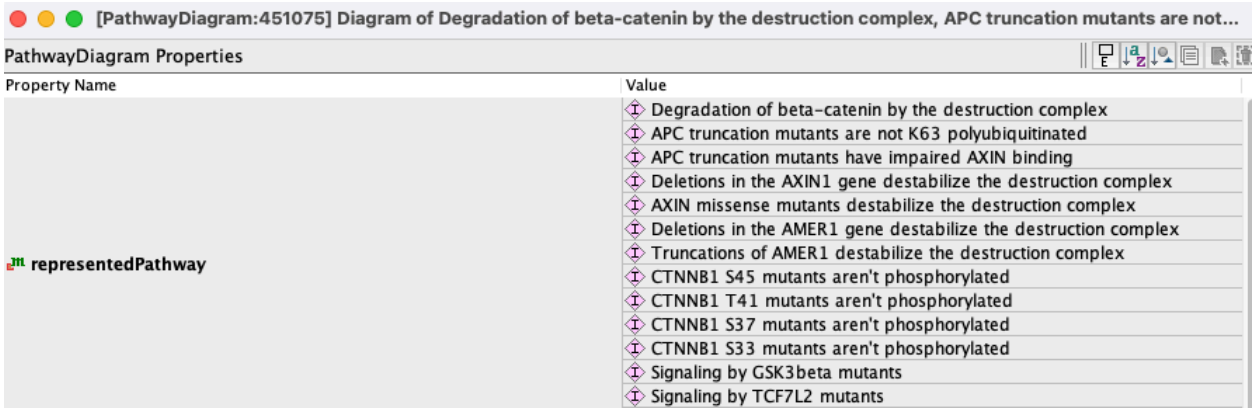
It is also possible to have different diagrams associated with different disease subpathways, as is shown in the “Signaling by WNT in cancer” pathway.



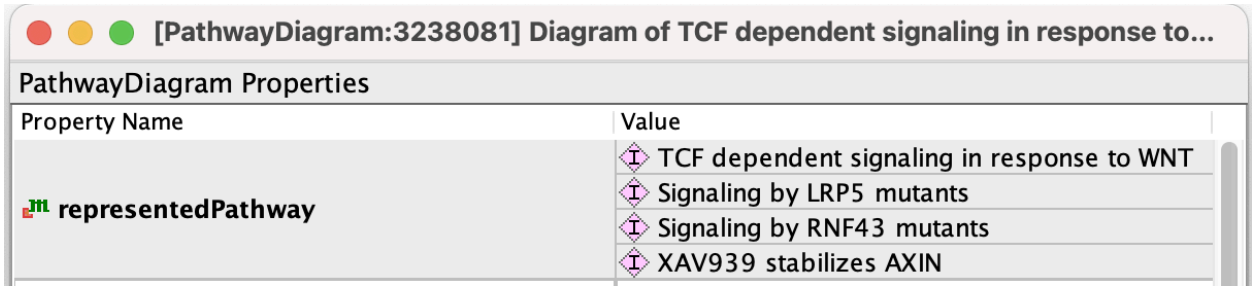
In this case, the wild-type WNT pathway has three associated normal pathway diagrams: Degradation of beta-catenin by the destruction complex, TCF-dependent signaling in response to WNT and WNT ligand biogenesis.

The parent disease WNT pathway, Signaling by WNT in cancer, has 10 subpathways containing a variety of gain-of-function and loss-of-function events.

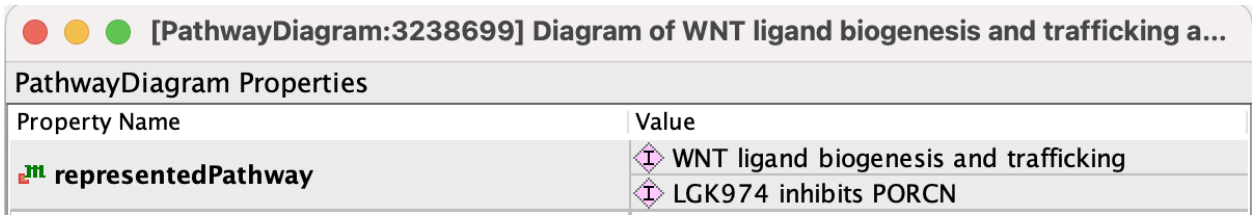
As shown below, the first 6 disease subpathways are all associated with the WT diagram “Degradation of beta-catenin by the destruction complex”:



The next 3 disease subpathways are associated with the WT diagram “TCF-dependent signaling in response to WNT”:



and the final disease subpathway is associated with the WT diagram “WNT ligand biogenesis and trafficking”:



Note: Any time you need to make revision to a "normal pathway" that has a disease counterpart, please be sure that you start off by "Building" (as opposed to"Checking

out") the full disease pathway into a local project. Then open the Disease diagram and make any necessary edits there.

Assigning Disease term attributes

We want to use “disease” in a broad sense: if a process has a specific bad outcome we can associate it with that outcome even if it hasn’t yet progressed to the point of causing major disease. The association of Amyloidosis with the Amyloids pathway is an example. It is possible, for instance, that amyloid deposition, that is going on in all of us, is pathological, even though most of us will die of something else before the amyloidosis progresses to a point where it’s symptomatic. Targeting these early pathological processes (propathology?) will likely be very important: good for us as patients because early intervention slows or stops the process and keeps us healthy longer; good for the drug companies because that intervention will likely be a continuing one – an additional drug to be taken for life rather than a one-time course.

Reactome disease terms are taken from the Disease Ontology (https://www.ebi.ac.uk/ols4/ontologies/doid). If gk-central doesn’t have the disease term you need, you can create a new one by opting to create a new instance. Browse the hierarchy here (https://www.ebi.ac.uk/ols4/ontologies/doid), then create a new record in Reactome by entering the corresponding DOID identifier (number only) in the identifier slot. You will be asked if you want to import the entry. Say yes. If the Disease Ontology itself doesn’t have an appropriate term, we can request that one be made.

To pathways:

Please be sure to label the top-level disease pathway and its sub-level disease pathway with a disease attribute.

Note that disease terms applied to pathways and events should describe *all entities and events that are captured in that event. For this reason, disease terms applied at the pathway level are often very general (“cancer”, not “lung adenocarcinoma”, for instance). Disease terms tend to get more specific at the entity level, as described below.

To reactions:

A disease attribute should be added to all reactions involving disease-related physical entities, such as proteins of bacterial/viral/fungal pathogens, disease variant human proteins, and drugs used in disease management. Use the same procedure as the one described for the addition of disease attributes to pathways. If there are several related disease tags that are applicable to reactions and pathways, using the most general tag is preferable, and even if specific disease attributes are added, always include the most general attribute for a given disease type. For example, for cancer-related reactions and pathways, always include the cancer tag.

To entities:

Disease attributes should be added to mutant proteins associated with disease and may be very specific, referring to the specific disease type(s) in which a particular mutation

was found. For example, EGFR L861Q mutant (DB_ID 1177542), in which L-leucine at position 861 is replaced with L-glutamine has been detected in non-small cell lung carcinoma and adult glioblastoma multiforme. Besides these specific disease tags, it may be advisable to also add more general disease tags to an EWAS, in this case, lung cancer, and cancer, to enable search of mutant EWASs using these general terms. When it comes to entity sets, such as EGFR KD mutants (DB_ID 1182966) which includes all kinase domain mutants of EGFR in cancer, a very general disease tag, such as cancer, is appropriate. This is because each member of the set has its own range of cancer types (while EGFR L861Q is found in lung cancer and glioblastoma, EGFR L858R is found in lung cancer, thymoma, thyroid cancer, breast cancer, and ovarian cancer), but their biological behavior is identical/similar. For the same reasons, only the general cancer disease tag should be used for events in which cancer disease entities participate.

To drugs:

When annotating drugs used to treat a particular disease, curators should add an appropriate disease tag to a drug entity. For a cancer drug, the general cancer disease attribute should definitely be added, but when it comes to more specific disease tags, it may be advisable to add only those cancer types for which the treatment by a given drug is approved.

Annotating Disease Reaction-Like-Events (RLEs)

Disease reaction-like events have 3 disease-specific fields that may need to be filled in:

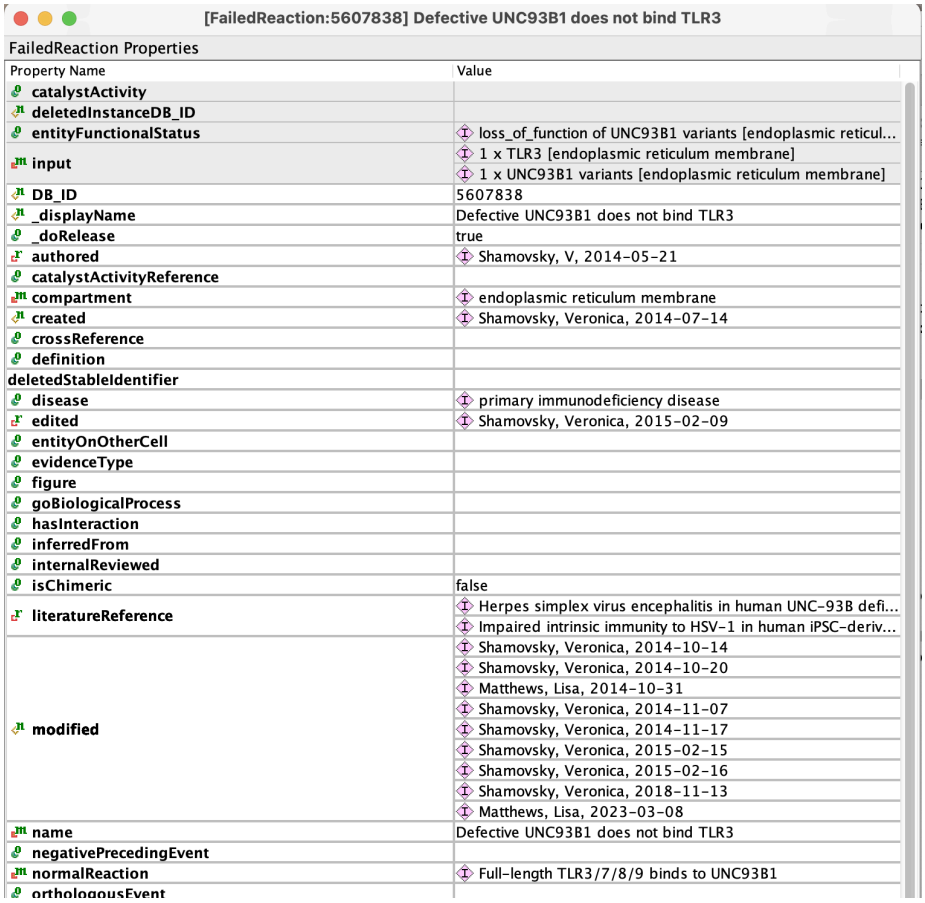
- disease (required for all disease RLE's; as described above)
- normalReaction (required for disease reactions that will be overlaid on a normal WT, described below)
- entityFunctionalStatus (EFS) (absolutely required for disease reactions that will be overlaid on a normal WT reaction, described below, but from a curation policy standpoint, required for *all disease reactions, regardless of overlay, as an additional source of information for users and to facilitate alignment with ACMG variant criteria).

Disease RLE's are either loss- or gain-of-function events. All loss-of-function are overlaid automatically on top of the corresponding WT reaction by virtue of their disease tag, normalReaction field and entityFunctionalStatus. Gain-of-function reactions may or may not be overlaid on normal WT reactions, as described below.

Annotating loss-of-function reactions

Loss-of-function 'reactions' are annotated as "FailedReaction". Failed reactions are labeled with a disease tag as described above but have only inputs (the disease variant entity or complex or set, plus any WT entities that are inputs in the 'normal' reaction), and do not have any outputs. See the note below about avoiding reaction participant mismatches.

The failedReaction record also lists the corresponding WT reaction in the “normalReaction” slot.

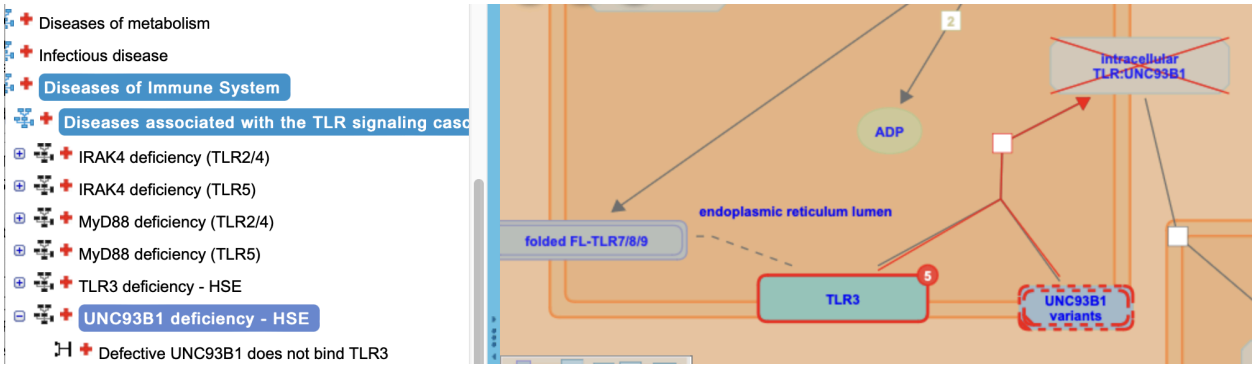


[FailedReaction:5607838] Defective UNC93B1 does not bind TLR3

FailedReaction Properties

Property Name	Value
catalystActivity	
deletedInstanceDB_ID	
entityFunctionalStatus	loss_of_function of UNC93B1 variants [endoplasmic reticul...
input	1 x TLR3 [endoplasmic reticulum membrane]
	1 x UNC93B1 variants [endoplasmic reticulum membrane]
DB_ID	5607838
_displayName	Defective UNC93B1 does not bind TLR3
_doRelease	true
authored	Shamovsky, V, 2014-05-21
catalystActivityReference	
compartment	endoplasmic reticulum membrane
created	Shamovsky, Veronica, 2014-07-14
crossReference	
definition	
deletedStableIdentifier	
disease	primary immunodeficiency disease
edited	Shamovsky, Veronica, 2015-02-09
entityOnOtherCell	
evidenceType	
figure	
goBiologicalProcess	
hasInteraction	
inferredFrom	
internalReviewed	
isChimeric	false
literatureReference	Herpes simplex virus encephalitis in human UNC-93B defi...
	Impaired intrinsic immunity to HSV-1 in human iPSC-deriv...
modified	Shamovsky, Veronica, 2014-10-14
	Shamovsky, Veronica, 2014-10-20
	Matthews, Lisa, 2014-10-31
	Shamovsky, Veronica, 2014-11-07
	Shamovsky, Veronica, 2014-11-17
	Shamovsky, Veronica, 2015-02-15
	Shamovsky, Veronica, 2015-02-16
	Shamovsky, Veronica, 2018-11-13
	Matthews, Lisa, 2023-03-08
name	Defective UNC93B1 does not bind TLR3
negativePrecedingEvent	
normalReaction	Full-length TLR3/7/8/9 binds to UNC93B1
orthologousEvent	

FailedReactions are placed in the appropriate disease pathway as usual, but are not dragged into the ELV by the curator. By virtue of the disease tag, the ‘normalReaction’ field, the EFS, and the appropriate sharing of pathway diagrams, FailedReactions are automatically overlaid on the background of the WT pathway and appear on top of the normal reaction with the disease input and the reaction node highlighted in red and the outputs of the WT reaction X’d out, as shown:



In general, FailedReactions should be named according to the structure “Defective [core protein name] doesn’t [some description of WT function]”, as in “Defective ALG9 does not add mannose to the N-glycan precursor”. The key is using the words ‘defective’ and ‘variant’ to describe the disease related entities.

To avoid reaction participant mismatches, all the inputs of the WT normal reaction must be present in the FailedReaction record (differences in stoichiometry are ignored). A good approach to making a FailedReaction can be to clone the normal WT reaction, switch type to FailedReaction, replace the WT entity with the relevant disease variant or set, and add disease tag, EFS and normal reaction to the relevant slots on the record. See the section below on avoiding participant mismatches for more considerations.

The molecular biology of the failed reaction should match the molecular biology of the normal reaction closely/exactly, both for biological coherence and to avoid mismatches. For instance, imagine a protein that binds to a substrate and then catalyzes a reaction. If there are subsets of variants identified that a) don’t bind or b) bind but are catalytically dead, these should be represented in two different failed reactions, the first showing a failed binding reaction and the second showing a failed catalysis. In some cases, this may require creating a new normal reaction for the WT pathway (if, for instance, the binding and catalysis are shown happening in a single step in the WT pathway). See here (Appendix C10, Multiple disease reactions and ENTITYFUNCTIONALSTATUS when using DISEASE VARIANT SUBSETS) for the original discussion of this).

Annotating gain-of-function reactions

There are two types of gain-of-function events in Reactome:

1. totally novel events that are not seen in the WT pathway. This may arise through the introduction of an infectious agent or by expression of a variant of a human protein that has an activity not displayed by the corresponding WT protein - for instance binding to a novel protein partner, or forming ligand-independent dimers.
2. gain-of-function events that happen at an elevated rate or efficiency compared to a corresponding event in a wild-type pathway.

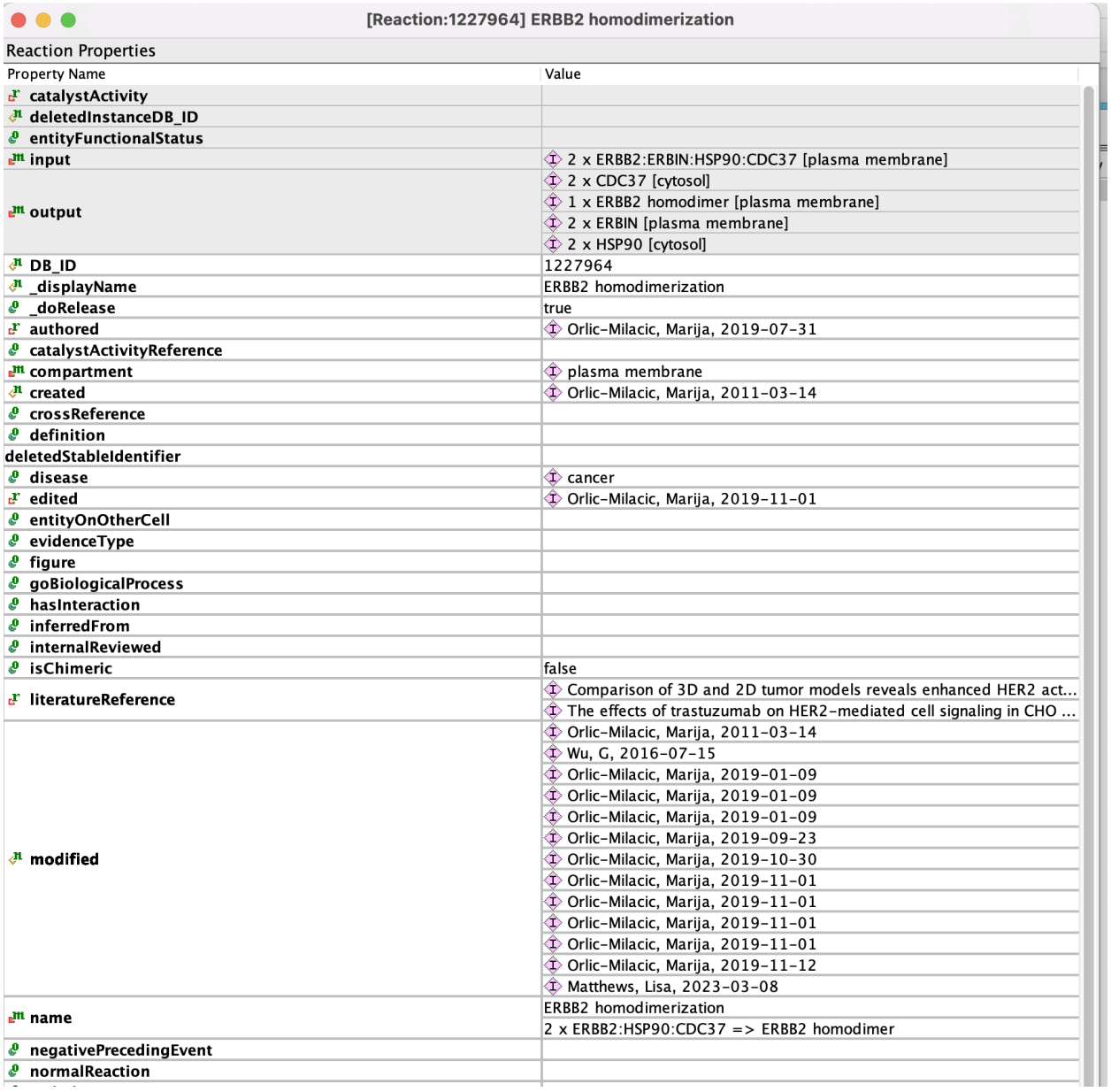
All gain-of-function RLE's are labeled with a disease tag as described above. This is needed for the automatic red line coloring in the ELV.

All gain-of-function RLEs should have their entity functional status slot filled out (described below).

Gain-of-function events not needing a 'normalReaction'

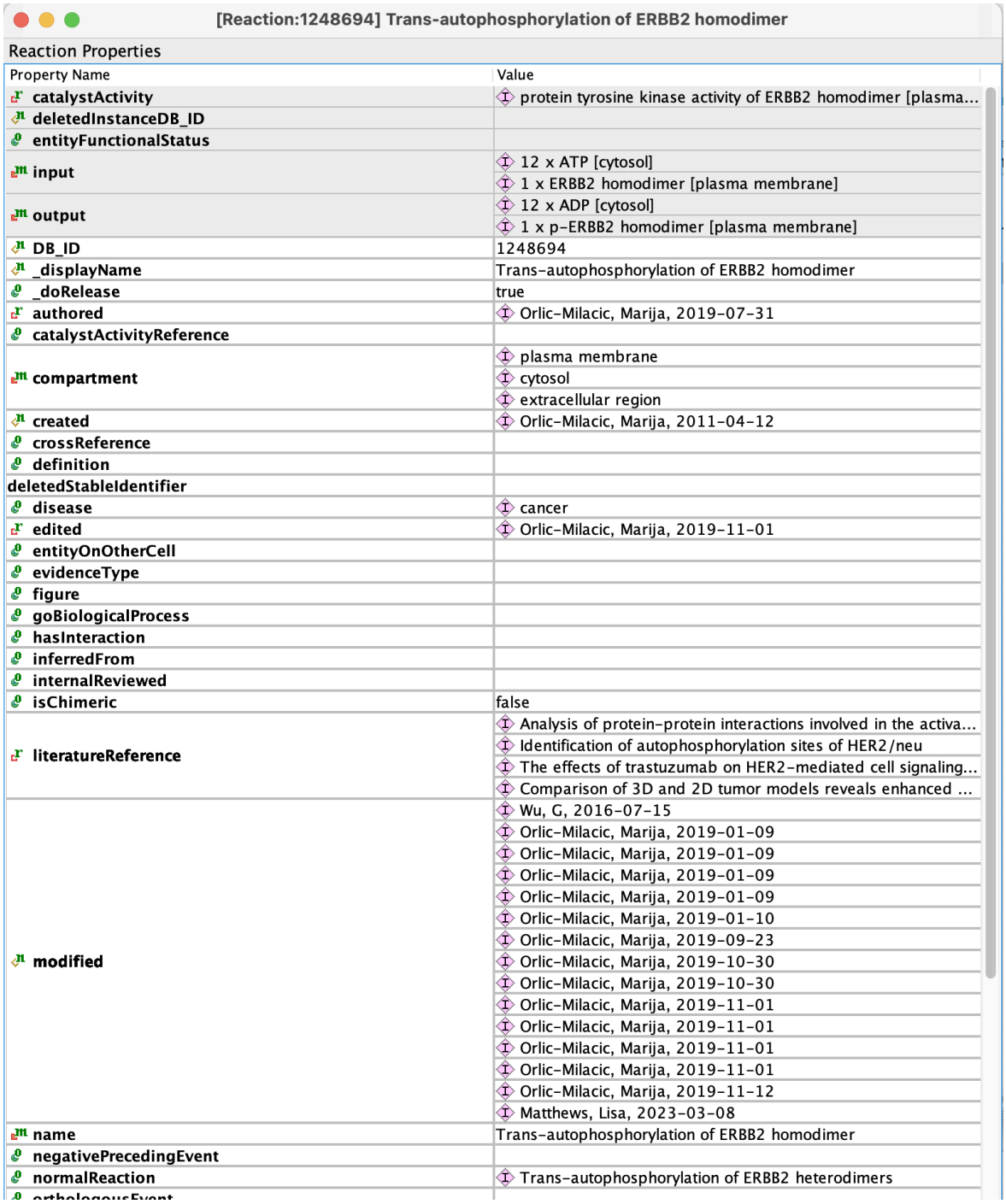
Some gain-of-function events represent totally novel activities that are not seen in the WT pathway. Novel gain-of-function RLEs of this kind do not have their 'normalReaction' slot filled out and are not overlaid on top of a normal WT reaction. These reactions are dragged into the pathway diagram manually by the curator as for WT reactions and occupy their own space in the pathway diagram.

Although the EntityFunctionalStatus slot is not required for display purposes for these RLEs, it should be filled out in any case. This is a new policy as of Feb 2023, so older records may not reflect this. This was introduced for consistency and to allow for better mapping of Reactome variants to ACMG variant criteria (see below).



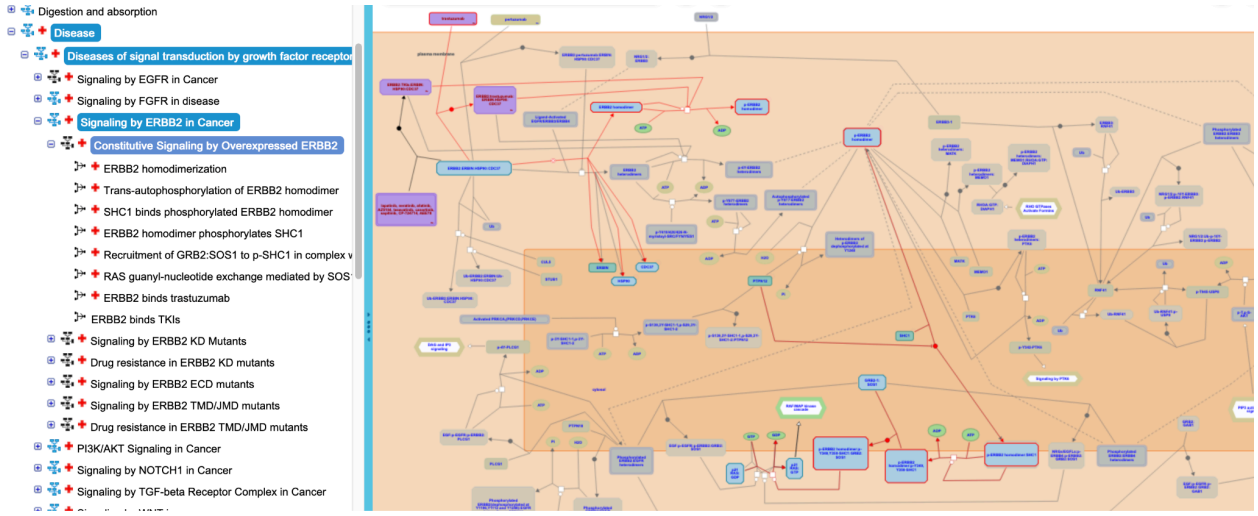
Gain-of-function events with a ‘normalReaction’

If a disease variant (or set of variants) performs the same reaction as the WT counterpart at an elevated rate or efficiency, the disease RLE can be overlaid on the normal reaction, as for loss-of-function FailedReactions above. In this case, the ‘normalReaction’ field on the disease RLE must be filled in to allow for proper display.



Often, a signaling pathway for a particular variant or set of variants will start with a novel gain-of-function event (ligand-independent binding, for instance) and then be followed by a series of reactions that mirror the steps performed by the WT protein downstream of the initiating event. These 'conserved' downstream events are also tagged with the corresponding 'normalReaction' and an EFS, and are automatically overlaid on the normal pathway events.

This can be seen in the “Constitutive signaling by Overexpressed ERBB2” (R-HSA-9634285) pathway. In the first step, a novel homodimerization of ERBB2 is depicted, resulting from overexpression of the protein in cancer. This reaction has no ‘normalReaction’. Downstream of dimerization, the receptor trans-autophosphorylates, binds and phosphorylates SHC1, recruits GRB2:SOS1 and activates RAS, among other actions. Each of these downstream events has a normal counterpart in the WT pathway that is captured in the disease RLE record and each of these is automatically overlaid on the WT pathway diagram. All the RLEs related to overexpressed ERBB2 homodimers are shown at the same time in the same diagram representation:



FunctionalStatus

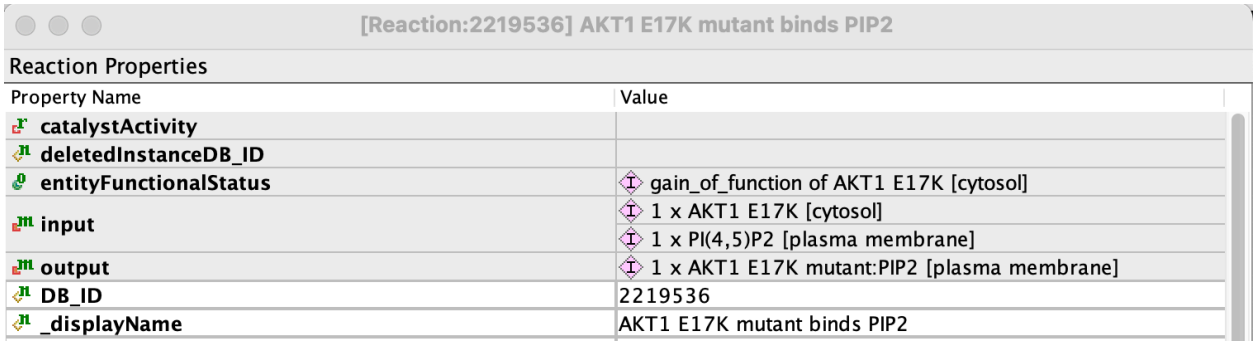
Overriding the automatic overlay

In some cases, a curator may wish to manually lay out a disease gain-of-function RLE even though it could theoretically be overlaid on a WT reaction. This may be to avoid reaction participant mismatches, described below, or for another reason. This is fine, and a curator would simply leave the ‘normalReaction’ slot empty and drag the reaction into the pathway diagram as usual. Disease tag and EntityFunctionalStatus fields should be filled out as usual.

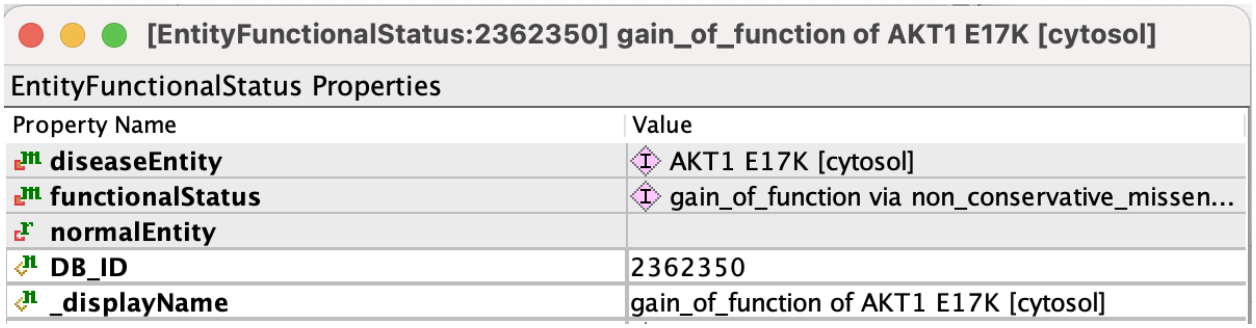
EntityFunctionalStatus

EntityFunctionalStatus is used to mark a disease event as either a gain- or a loss-of-function, and to describe the underlying reason for the change in behavior. This tag, along with the “NormalReaction” attribute, is especially important for loss-of-function events, as it is needed for deploying the red lines and X’ing out of WT outputs in the website display. Note that EFS’s are applied at the reaction-like-event level and not to entities themselves. This is because a given variant/set of variants may play different functional roles in different disease or reaction contexts.

The screenshot below shows the EFS of AKT1 E17K in the disease RLE “AKT1 E17K mutant binds PIP2”.



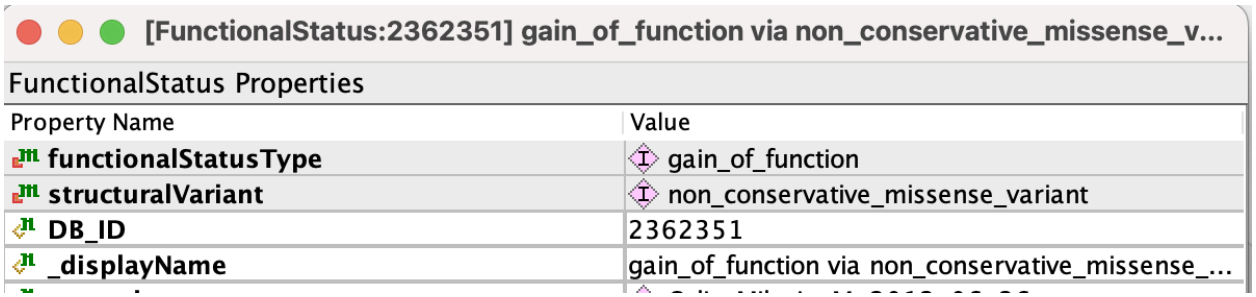
The EntityFunctionalStatus has three fields, diseaseEntity, functionalStatus and normalEntity:



- The disease entity is the disease variant (or complex or set containing the relevant disease variant(s)) that is responsible for the phenotype.

- The normalEntity is the corresponding normal protein that takes part in the RLE captured in the ‘normalReaction’ slot of the disease RLE. This field must be filled out if a ‘normalReaction’ is specified in the disease RLE. This field helps the overlay find the appropriate normal entity to overlay the disease variant on and helps to avoid reaction participant mismatches, described below.

- The FunctionalStatus itself consists of 2 fields, FunctionalStatusType, and structural variant:



The structuralVariant field takes terms from SequenceOntology (https://www.ebi.ac.uk/ols4/ontologies/so) and describes the nature of the underlying genetic change that gives rise to the variant.

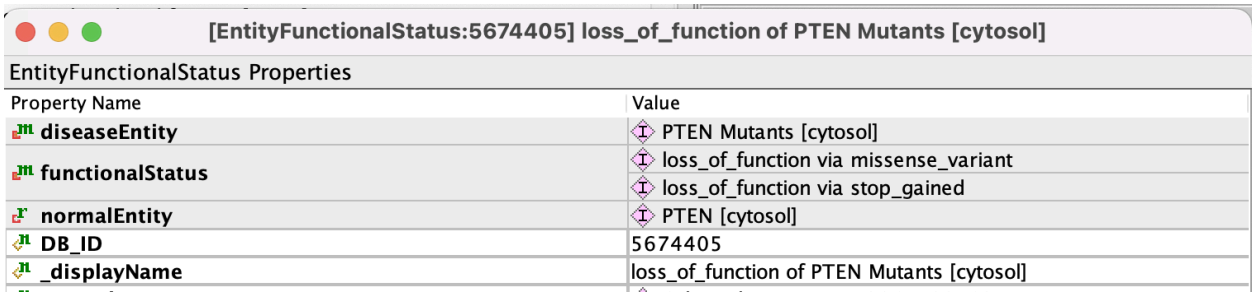
The functionalStatusType describes the functional consequence of this genetic change (gain-of-function, loss-of-function etc)

Many of the commonly used FunctionalStatus already exist in gk_central, but new combinations may be generated as required.

A given EFS may be used on as many disease RLEs as appropriate..

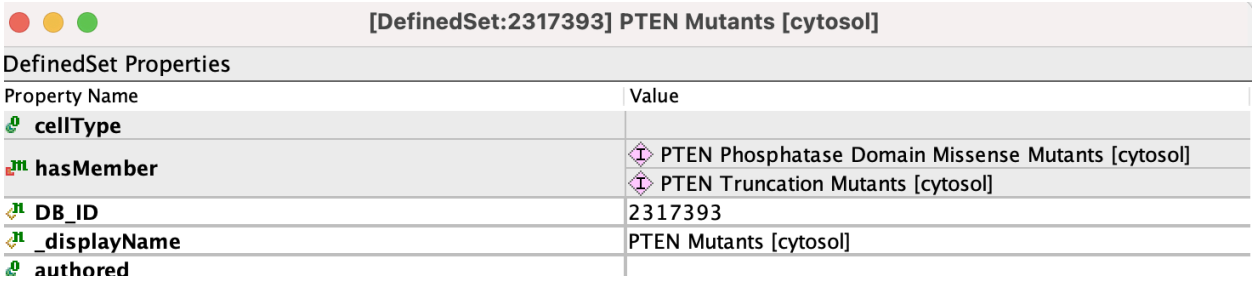
EFSs for variant sets:

The EFS shown above for AKT1 E17K is a simple case, where the disease variant is a single EWAS. Disease annotations often describe sets of mutants that function similarly. These sets can also be represented by a single EFS record that has multiple FunctionalStatus attributes to cover all the variants contained in the set, as shown in the screenshot below:



Here, the disease entity is a set of PTEN mutants, some of which arise as a consequence of missense mutations and some as a result of novel stop mutations.

Best practice says that the set of PTEN mutants should itself contain subsets corresponding to each of the functional status types indicated in the EFS record, as shown below:



(Original discussion of this is here; Appendix C10)

AVOIDING REACTION PARTICIPANT MISMATCHES IN DISEASE CURATION

In order for the overlay function to work properly, there must be careful attention paid to aligning the inputs and outputs of disease and WT reaction participants. If this is not done properly, a Reaction Participant mismatch error will be generated during QA.

Tips to avoid reaction participant mismatches:

1. A disease reaction must have exactly one or zero ‘normalReaction’ counterparts.
2. The curator can always draw disease reactions directly onto pathway diagrams to bypass auto-generating reaction diagrams. However, to make sure auto-generation work, the curator must specify values in the entityFunctionalStatus slot. The disease PE wrapped in EFS will be mapped to normal reaction participant.
3. Mapping between disease reaction participants and normal reaction participants are always based on participants’ referenceEntity values. If two normal reaction participants refer to the same ReferenceEntity instance, but having different roles (e.g. one input, and another catalyst), the mapping will not be reliable. Manual drawing is highly recommended for such a case.
4. If it is to be overlaid, a disease reaction, including FailedReaction and GoF reaction, must have exactly the same number of input entities as its normal reaction counterpart without considering stoichiometries (e.g. a disease reaction has input A, B, and its normal counterpart has A, A, B. This is fine, or vice versa). Creating failedReaction instances by cloning the normal event, switching the type of the clone to failedReaction, and then editing further as needed might be a good way to minimize this kind of mistake.
5. If a protein normally functions as part of a complex, there are two ways to annotate the disease reaction that occurs when a mutant form of the protein is expressed.

   a) If the mutant form of the protein still associates with the other members of the complex, but the complex does not function, then create a complex instance identical to the normal one except that the one affected normal protein is replaced with its disease counterpart. Then use that variant complex to create an entityFunctionalStatus instance.

   b) If the mutant form of the protein is incapable of forming a complex at all, then create a preceding normal reaction to annotate formation of the normal complex (if that doesn’t already exist), and create a failedReaction version of the normal complex-formation one, not of the later one in which the normal complex carries out its normal function. That yields the needed exact matching between the disease reaction and its normal counterpart.

Annotating Disease Entities

General approaches to disease entity annotation and naming

Naming

Reactome variant protein names are formulated in accordance with the HGVS nomenclature (https://hgvs-nomenclature.org/stable/), with the exception of using single letter abbreviations for amino acid residues.

When a descriptor is added (usually to set names), 'variant' is to be used rather than 'mutant' as in "PI3K variants". Specific rules for the naming of different variants are described more fully in the relevant sections below.

Aligning with ACMG variant guidelines

Where possible, we try to align our annotations with The American College of Medical Genetics and Genomics (ACMG) Standards and Guidelines for the interpretation of sequence variants (Richards et al. 2015). This framework classifies sequence variants into benign or pathogenic based on 1) population data, 2) computational and predictive data, 3) functional data, 4) segregation data, 5) de novo data, 6) allelic data, 7) other database, and 8) other data. Detailed criteria are listed for assessing the strength of evidence in each of the eight categories as supporting (BP1 to BP7) or strong (BS1 to BS4) for benign variants, and as supporting (PP1 to PP5), moderate (PM1 to PM6), strong (PS1 to PS4) or very strong (PVS1) for pathogenic variants.

Because Reactome's disease variant annotation focuses on annotating events at a molecular or biochemical level, (ie variants that have been functionally characterized) and on clinically relevant, pathogenic variants, benign variants are outside Reactome's scope.

In consequence, Reactome-annotated disease variants correspond to only 6 of the 27 ACMG criteria (PM1, PM4, PM5, PS1, PS3, and PVS1), as indicated in Table 1 and described more fully below. Of the remaining 21 ACMG criteria, BS1-4 and BP1-7, used for categorization of variants as benign, can be used for exclusion of potentially pathogenic variants from consideration when the relevant data is available. The criteria PP1-PP5, PM2, PM3, PM6, PS2, and PS4 are used for categorization of clinically relevant pathogenic variants but are not used in Reactome because they do not provide sufficient information to functionally annotate disease variants. These criteria can only be used to corroborate pathogenicity of a disease variant that conforms to criteria PM1, PM4, PM5, PS1, PS3 or PVS1.

Criterion	**Evidence strength**	**Pathogenicity**	**Evidence type**	**Evidence definition**	**Suitability as direct Reactome evidence**
BS1	strong	benign	populatio	variant represents minor allele whose	-

			n data	frequency is too high for disorder	
BS2	strong	benign	population data	variant incidence in controls is inconsistent with disease penetrance	-
BS3	strong	benign	functional data	well-established functional studies show no deleterious effect for variant	-
BS4	strong	benign	segregation data	variant does not segregate with disease	-
BP1	supporting	benign	computational and predictive data	missense variant is found in gene in which only truncating variants cause disease	-
BP2	supporting	benign	allelic data	variant is observed in trans with dominant variant or in cis with pathogenic variant	-
BP3	supporting	benign	computational and predictive evidence data	variant is an in-frame indel in repeat without known function	-
BP4	supporting	benign	computational and predictive evidence data	multiple lines of computational evidence suggest variant has no impact on gene or gene product	-
BP5	supporting	benign	other data	variant is found in disease case with an alternate cause	-
BP6	supporting	benign	other database	a reputable source without shared data defines variant as benign	-
BP7	supporting	benign	computational and predictive evidence data	variant is silent variant with non predicted splice impact	-
PP1	supporting	pathogenic	segregation data	variant co-segregates with disease in multiple affected family members	-
PP2	supporting	pathogenic	functional data	variant represents missense variant in gene in which there is low rate of benign missense variants and in which pathogenic missense variants are common	-
PP3	supporti	pathogenic	computational and	multiple lines of computational evidence support a deleterious effect of variant on	-

	ng		predictive data	gene/gene product	
PP4	supporting	pathogenic	other data	variant is detected in a patient whose phenotype of family history are highly specific for disease gene	-
PP5	supporting	pathogenic	other database	a reputable source identifies variant as pathogenic	-
PM1	moderate	pathogenic	functional data	variant represents mutational hotspot or is within functional domain without benign variation	✓
PM2	moderate	pathogenic	population data	variant is absent in population databases	-
PM3	moderate	pathogenic	allelic data	for recessive disorders variant is detected in trans with known pathogenic variant	-f
PM4	moderate	pathogenic	computational and predictive data	variant represents protein length changing variant	✓
PM5	moderate	pathogenic	computational and predictive data	variant represents novel missense change at amino acid residue previously reported to be affected by pathogenic missense change	✓✓
PM6	moderate	pathogenic	de novo data	de novo variant is detected (without paternity or maternity confirmed)	-
PS1	strong	pathogenic	computational and predictive data	variant leads to same amino acid change as an established pathogenic variant	✓✓✓
PS2	strong	pathogenic	de novo data	de novo variant is detected (with paternity and maternity confirmed)	-
PS3	strong	pathogenic	functional data	well-established functional studies show variant's deleterious effect	✓✓✓
PS4	strong	pathogenic	population data	prevalence of variant in affected patients is statistically significantly increased over controls	-
PVS1	very strong	pathogenic	computational and predictive data	variant is predicted to be null variant in gene where LOF is known disease mechanism	✓✓

PS3, PVS1 and PS1

Of the 6 ACMG criteria suitable as direct evidence for annotation of disease variants in Reactome, PS3 is the gold standard - variants whose deleterious effect(s) are supported by well-established functional studies. An additional requirement for Reactome is that these functional studies are at the molecular mechanism resolution level. Variants with this classification may be of any type - missense, fusion, in-frame insertion-deletion (in-frame indel), nonsense, frameshift, etc. - so long as they have been functionally characterized. PS3 variants are annotated as direct disease reaction participants or as members of disease variant sets.

PM5

ACMG classifies as PM5 any variant that represents a novel missense change at an amino acid previously reported to be affected by a pathogenic missense change. Reactome extends this classification to include variants of any type (missense, fusion, in-frame indel, etc.) that have not been directly functionally characterized but contain mutations analogous or equivalent to those found in annotated PS3 variants. PM5 variants are annotated as "candidates" within Reactome's disease variant sets.

PM1 and PM4

Nonsense and frameshift variants not directly functionally studied that conform to the criteria PM1 (moderate pathogenic based on functional data where variant represents mutational hotspot or is within functional domain without benign variation) and PM4 (moderate pathogenic based on computational and predictive data where variant represents protein length changing variant) so that these variants are protein length changing variants that result in truncation or complete ablation of functional domains key to a particular protein function, are annotated as candidates of disease variants sets.

References

1. Richards S, Aziz N, Bale S, Bick D, Das S, Gastier-Foster J, Grody WW, Hegde M, Lyon E, Spector E, Voelkerding K, Rehm HL; ACMG Laboratory Quality Assurance Committee. Standards and guidelines for the interpretation of sequence variants: a joint consensus recommendation of the American College of Medical Genetics and Genomics and the Association for Molecular Pathology. Genet Med. 2015 May;17(5):405-24. doi: 10.1038/gim.2015.30. Epub 2015 Mar 5. PMID: 25741868; PMCID: PMC4544753.

DISEASE ENTITY OVERVIEW

When making a record for a disease variant, it is often most efficient to start by cloning the WT record and then changing the relevant fields and adding the disease-specific classes.

Like disease events, disease entities must be labeled with a disease tag (described above).

Where possible, disease entities will have a cross reference to an external database like COSMIC, OMIM, ClinGen or other, as shown. An entity can have as many

cross-references as needed. COSMIC cross-referencing should make use of the Genomic Mutation ID where available, not the Legacy Identifier. ClinGen Canonical Allele (CA) identifiers are used wherever possible, but ClinVar IDs may be used if CA IDs are not available.

Disease entities should also have their literature field filled in citing the work where the variant was identified or characterized unless they are annotated as disease variant candidates.

[EntityWithAccessionedSequence:4791272] CTNNB1 S45P [cytosol]

EntityWithAccessionedSequence Properties

Property Name	Value
cellType	
compartment	cytosol
endCoordinate	781
hasModifiedResidue	L-serine 45 replaced with L-proline
referenceEntity	UniProt:P35222 CTNNB1
startCoordinate	1
DB_ID	4791272
_displayName	CTNNB1 S45P [cytosol]
authored	
created	Rothfels, K, 2013-10-28
crossReference	COSMIC:COSM12639 COSMIC:COSV62687880
definition	
disease	cancer hepatocellular carcinoma kidney cancer colorectal cancer adrenal gland cancer connective tissue cancer
edited	
figure	
goCellularComponent	
inferredFrom	
inferredTo	
literatureReference	AXIN1 mutations in hepatocellular carcinomas, and growt... Mutational spectrum of beta-catenin, AXIN1, and AXIN2 i... Nuclear accumulation of beta-catenin protein in Wilms' tu...
modified	Rothfels, K, 2013-11-01 Rothfels, K, 2014-02-04
name	CTNNB1 S45P Catenin beta-1 CTNB1_HUMAN
reviewed	
revised	
species	Homo sapiens
stableIdentifier	R-HSA-4791272.1
summation	
systematicName	

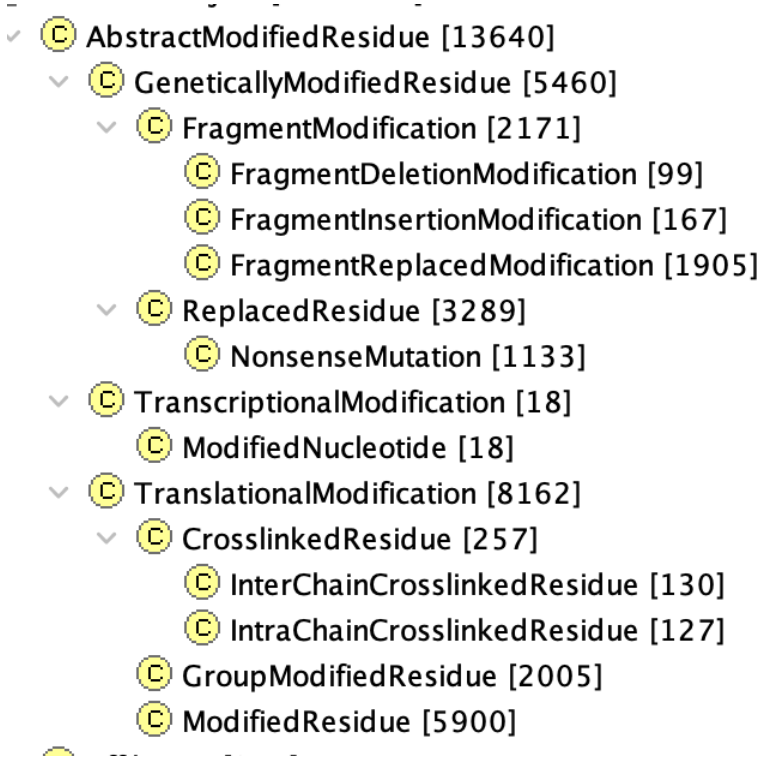
Data model classes for ‘hasModifiedResidue’

The changes in sequence that lead to the formation of a variant protein are captured by the field “hasModifiedResidue” using the modified residue class “GeneticallyModifiedResidue” to distinguish genetically encoded variants from

post-translational modifications (TranslationalModifications class) and RNA editing (TranscriptionalModification class).

Note that in both its variant naming convention and in the annotation of 'hasModifiedResidue', Reactome focuses on amino acid (protein level) changes and does not capture changes at the nucleic acid level.

The data model includes the following classes of modified residues to capture the amino acid changes leading to a protein variant (under "GeneticallyModifiedResidue" to distinguish them from post-translational modifications).



The following sections will go through annotation and naming strategies for common classes of amino acid variations leading to variant protein forms.

Missense mutations

These are perhaps the most straightforward variants to annotate. These variants are the consequence of the replacement of a single amino acid codon with a non-synonymous amino acid codon without change in the reading frame.

Naming

Format: "prefix" + "wild-type amino_acid" + "position" + "new_amino_acid"

Where "prefix" is the HGNC gene name and the amino acids are represented by single letter codes

example, MTR P1173L: substitution of proline 1173 with leucine

How to annotate

Missense mutations are captured with the “ReplacedResidue” class of GeneticallyModifiedResidue.

Use PSI-MOD IDs to construct the replaced residue, identifying both the normal residue with an [amino acid] removal instance and its mutant replacement with an [amino acid] residue instance. Here, L-proline is replaced by L-leucine at coordinate 1173.

[ReplacedResidue:3321905] L-proline 1173 replaced with L-leucine

ReplacedResidue Properties

Property Name	Value
coordinate	1173
psiMod	L-proline removal [MOD:01645] L-leucine residue [MOD:00020]
referenceSequence	UniProt:Q99707 MTR
DB_ID	3321905
_displayName	L-proline 1173 replaced with L-leucine

[EntityWithAccessionedSequence:3321944] MTR P1173L [cytosol]

EntityWithAccessionedSequence Properties

Property Name	Value
cellType	
compartment	cytosol
endCoordinate	1265
hasModifiedResidue	L-proline 1173 replaced with L-leucine
referenceEntity	UniProt:Q99707 MTR
startCoordinate	1
DB_ID	3321944
_displayName	MTR P1173L [cytosol]
authored	
created	Jassal, B, 2013-05-02
crossReference	
definition	
disease	methylmalonic aciduria and homocystinuria type cblG
edited	
figure	
goCellularComponent	
inferredFrom	
inferredTo	
literatureReference	
modified	Jassal, B, 2013-06-07
name	MTR P1173L Methionine synthase
reviewed	
revised	
species	Homo sapiens
stableIdentifier	R-HSA-3321944.1

Nonsense mutations

Nonsense variants are the consequence of mutations that replace a single amino acid codon with a nonsense codon.

Naming

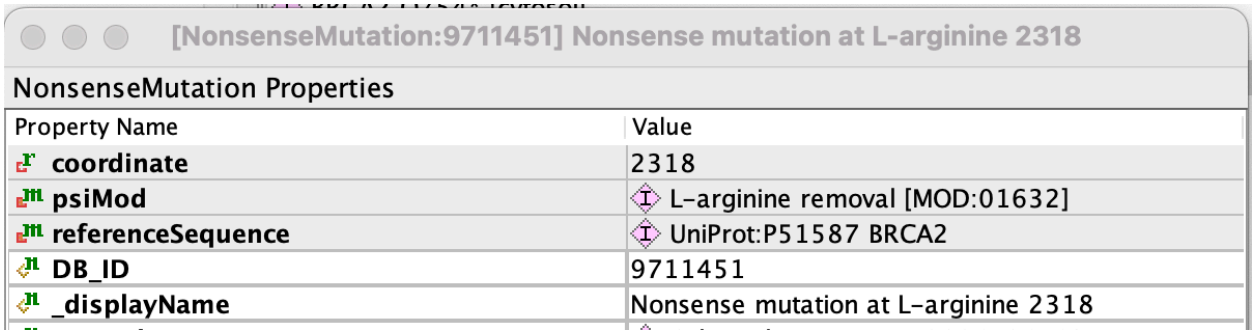
Format: “prefix” + “amino_acid” + “position” + “*” (stop codon asterisk shown in red just so it doesn’t get lost in the background)

Where “prefix” is the HGNC gene name, the amino acid is represented by single letter code and the stop codon is represented by an asterisk (*)

Example : BRCA2 R2318* - arginine 2318 replaced with stop codon

How to annotate

Nonsense mutations are captured using the “NonsenseMutation” subclass of “ReplacedResidue”. A single PSI MOD term is entered in the record indicating the removal of the WT amino acid. On the EWAS record, the endCoordinate is changed to be the last amino acid encoded by the variant (one less than the position of the stop codon).



On the EWAS record, the endCoordinate is changed to be the last amino acid encoded by the variant (one less than the position of the stop codon).

[EntityWithAccessionedSequence:9711457] BRCA2 R2318* [cytosol]

EntityWithAccessionedSequence Properties

Property Name	Value
cellType	
compartment	cytosol
endCoordinate	2317
hasModifiedResidue	Nonsense mutation at L-arginine 2318
referenceEntity	UniProt:P51587 BRCA2
startCoordinate	1
DB_ID	9711457
_displayName	BRCA2 R2318* [cytosol]
authored	
created	Orlic-Milacic, Marija, 2021-01-13
crossReference	COSMIC:COSV66459219
definition	
disease	cancer
edited	
figure	
goCellularComponent	
inferredFrom	
inferredTo	
literatureReference	
modified	
name	BRCA2 R2318*
reviewed	

FragmentModifications

The next section describes slightly more complicated variants involving changes to more than a single amino acid in the reference protein. These variants are captured under the AbstractModifiedResidue class “FragmentModification” that itself has three subclasses, as shown below:



In-frame deletions:

These variants are the consequence of the removal of a stretch of the WT amino acids between the initiator methionine and the stop codon without changing the reading frame of the protein.

Naming

Format: “prefix” + “first deleted amino_acid” +”first deleted amino acid position” + “_” +”last deleted amino acid” +”last amino acid position” + “del”

Where “prefix” is the HGNC gene name, the deleted amino acids are represented by single letter codes separated by underscore.

Example: HK1 H577_C672del: a deletion of residues H577 to C672, inclusive

How to annotate

Deletions are captured using the FragmentDeletionModification class specifying the start and stop coordinate of the deleted stretch of amino acids as shown below:

[FragmentDeletionModification:5621917] Deletion of residues 577 to 672

FragmentDeletionModification Properties

Property Name	Value
endPositionInReferenceSequence	672
referenceSequence	UniProt:P19367 HK1
startPositionInReferenceSequence	577
DB_ID	5621917
_displayName	Deletion of residues 577 to 672

The corresponding EWAS record. Note that in this case, the end coordinate of the protein is not adjusted relative to the WT protein to account for the deleted stretch of amino acids.

[EntityWithAccessionedSequence:5621873] HK1 H577_C672del [cytosol]

EntityWithAccessionedSequence Properties

Property Name	Value
cellType	
compartment	cytosol
endCoordinate	917
hasModifiedResidue	Deletion of residues 577 to 672
referenceEntity	UniProt:P19367 HK1
startCoordinate	1
DB_ID	5621873
_displayName	HK1 H577_C672del [cytosol]
authored	
created	Jassal, Bijay, 2014-09-05
crossReference	
definition	
disease	congenital nonspherocytic hemolytic anemia

For deletion of a single amino acid, the start and end position of the FragmentDeletionModification are the same:

[FragmentDeletionModification:3325558] Deletion of residues 88 to 88

FragmentDeletionModification Properties

Property Name	Value
endPositionInReferenceSequence	88
referenceSequence	UniProt:Q9NPF0 CD320
startPositionInReferenceSequence	88
DB_ID	3325558
_displayName	Deletion of residues 88 to 88

And the EWAS is named without a range of amino acids (CD320 E88del) as shown below:

[EntityWithAccessionedSequence:3325560] CD320 E88del [plasma membrane]

EntityWithAccessionedSequence Properties

Property Name	Value
cellType	
compartment	plasma membrane
endCoordinate	282
hasModifiedResidue	Deletion of residues 88 to 88
referenceEntity	UniProt:Q9NPF0 CD320
startCoordinate	36
DB_ID	3325560
_displayName	CD320 E88del [plasma membrane]

In-frame insertions

These variants are the consequence of adding a stretch of the amino acids between the initiator methionine and the stop codon without changing the reading frame of the protein. The inserted amino acids are NOT a copy of the sequence immediately upstream of the insertion site. If this is the case, these are annotated as duplications (below).

For example, for a sequence “YMVA”,

An insertion is YYMVAMVA - the YMVA in red is NOT a direct repeat of the upstream because the upstream sequence is interrupted by the insertion.

YMVA with a duplication is YMVAYMVA - the YMVA in red IS a direct repeat of the upstream sequence.

Naming

Format:

“prefix” + “N-terminal flanking amino acid” +”position of N-terminal flanking amino acid” + ”_” +”C-terminal flanking amino acid” + “position of C-terminal flanking amino acid” +”ins” + ”inserted sequence”

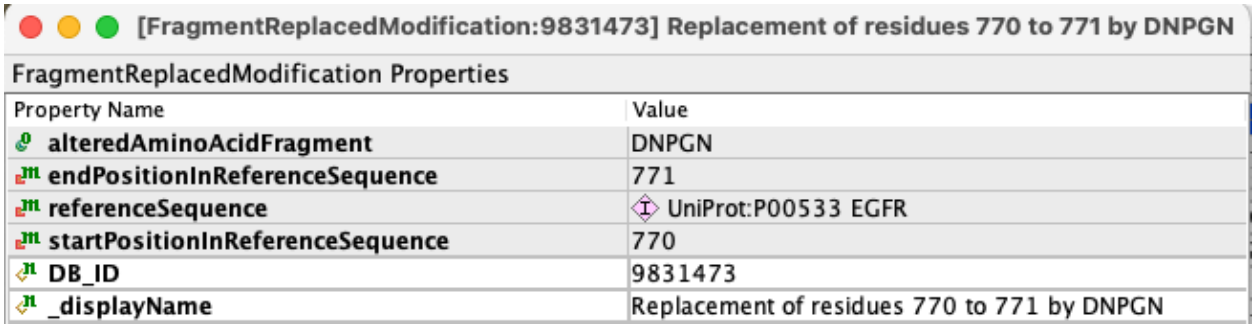
Where “prefix” is the HGNC gene name, the N- and C-terminal flanking amino acids are represented by single letter codes separated by underscore, and the inserted sequence is listed as single letter amino acid codes.

Example: - EGFR D770_N771insNPG - in this variant the three amino acids NPG are inserted between D770 and N771.

The WT protein sequence in this region is 770-DNPHV-774 and the variant sequence is 770-DNPGNPHV-(original sequence in blue, insertion in red). NPG is inserted at position 771.

How to annotate

Insertions are captured using FragmentReplacedModification class specifying the start and stop coordinates which overlap with positions of the N-terminal and C-terminal flanking amino acids, respectively, and specifying the alteredAminoAcidFragment, so that the first letter in the string is the one letter code for the N-terminal flanking amino acid, the last letter in the string is the one letter code for the D-terminal amino acid and the string in between correspond to the one letter codes of the inserted amino acids. The insertion site should be (StartPositioninReferenceSequence+1).


[FragmentReplacedModification:9831473] Replacement of residues 770 to 771 by DNPGN

FragmentReplacedModification Properties

Property Name	Value
alteredAminoAcidFragment	DNPGN
endPositionInReferenceSequence	771
referenceSequence	UniProt:P00533 EGFR
startPositionInReferenceSequence	770
DB_ID	9831473
_displayName	Replacement of residues 770 to 771 by DNPGN

In the EWAS record, the end coordinate of the WT protein is not changed to accommodate the insertion.

[EntityWithAccessionedSequence:1996322] EGFR D770_N771insNPG [plasma membrane]

EntityWithAccessionedSequence Properties

Property Name	Value
cellType	
compartment	plasma membrane
endCoordinate	1210
hasModifiedResidue	Replacement of residues 770 to 771 by DNPGN
referenceEntity	UniProt:P00533 EGFR
startCoordinate	25
DB_ID	1996322
_displayName	EGFR D770_N771insNPG [plasma membrane]
authored	
created	Orlic-Milacic, Marija, 2011-11-19
crossReference	
definition	
disease	cancer
	lung cancer
	non-small cell lung carcinoma
edited	
figure	
goCellularComponent	
inferredFrom	
inferredTo	
literatureReference	Structural, biochemical, and clinical characterization of...
modified	Orlic-Milacic, M, 2013-06-06
	Orlic-Milacic, Marija, 2023-03-20
name	EGFR D770_N771insNPG
	Epidermal Growth Factor Receptor D770_N771insNPG
	EGFR Asp770_Asn771insAsnProGly

In-frame duplications

These variants are the consequence of in-frame insertion of a stretch of the amino acid codons between the initiator methionine and the stop codon without changing the reading frame of the protein. The inserted amino acids ARE a copy of the sequence immediately upstream of the insertion site.

Naming

Format: “prefix” + “first amino acid in the duplicated string” + “position of the first amino acid in the duplicated string” + “_” +”last amino acid in the duplicated string” + “position of the last amino acid in the duplicated string” + “dup”

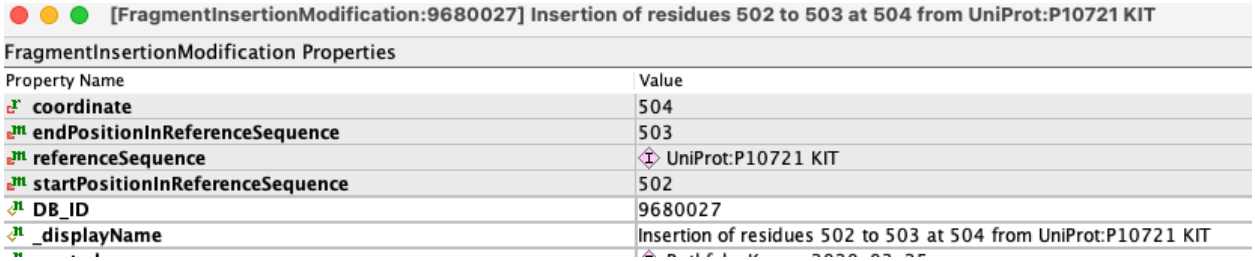
Where “prefix” is the HGNC gene name, the amino acids flanking the duplication are represented by single letter codes.

Example: KIT A502_Y503dup: a duplication of residues A502 and Y503. Because duplications are defined to happen at the C-terminal end of the sequence, the duplicated amino acids are located at coordinate 504, as shown below. For duplications, the site of insertion in the reference EWAS should always be (insertion end coordinate +1).

So in this example, WT KIT sequence is A502,Y503,F504,N505 etc and the variant sequence in A502,Y503,A502,Y503,F504,N505 etc, with duplicated residues shown in red and located at position 504 of WT sequence

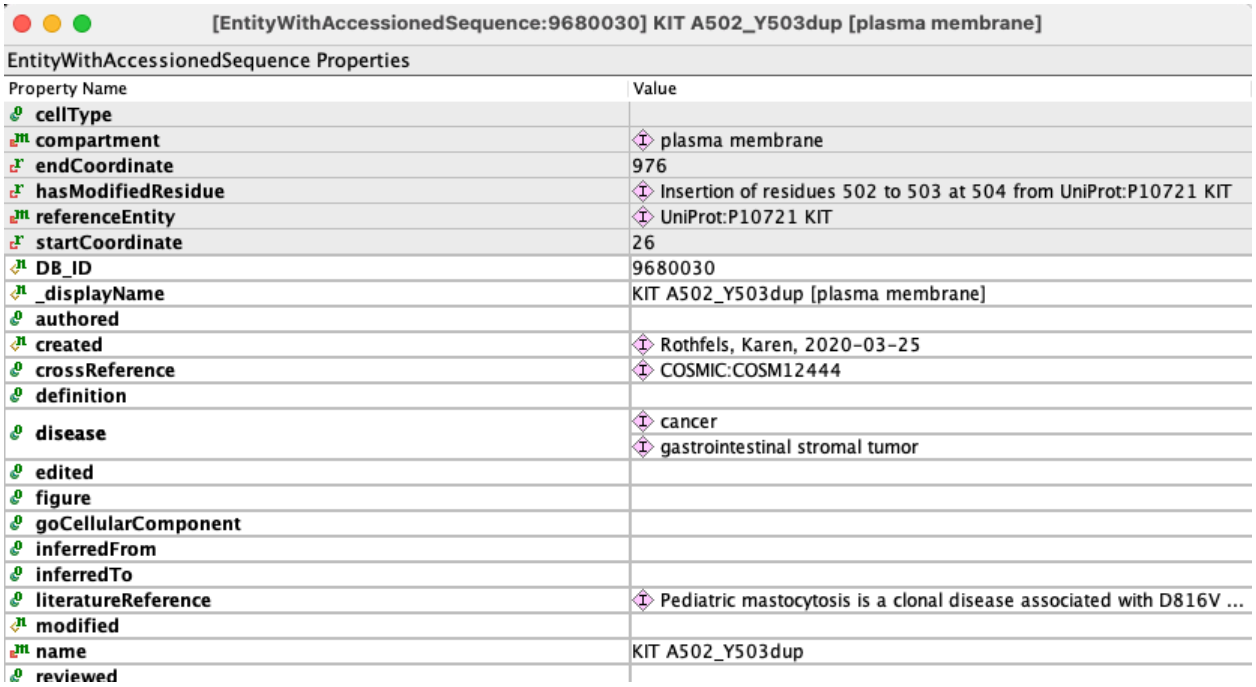
How to annotate

Duplications are captured using the FragmentInsertionModification class specifying the start and end coordinates of the duplicated stretch of amino acids in the reference sequence, and with the coordinate of the insertion at (endPositioninReferenceSequence+1) as shown below:



Notice that the referenceSequence of the FragmentInsertionModification is the same as the referenceEntity of the disease variant EWAS itself. This is in contrast to the annotation of fusion proteins, which are also annotated with FragmentInsertionModfications, as described below - but in the case of fusion proteins, the referenceEntity of the FragmentInsertionModification is different from that of the disease variant EWAS referenceEntity.

As is the case with in-frame deletions and in-frame insertions, the end coordinate of the duplication variant EWAS is NOT changed relative to the WT protein.



[EntityWithAccessionedSequence:9680030] KIT A502_Y503dup [plasma membrane]

EntityWithAccessionedSequence Properties

Property Name	Value
cellType	
compartment	plasma membrane
endCoordinate	976
hasModifiedResidue	Insertion of residues 502 to 503 at 504 from UniProt:P10721 KIT
referenceEntity	UniProt:P10721 KIT
startCoordinate	26
DB_ID	9680030
_displayName	KIT A502_Y503dup [plasma membrane]
authored	
created	Rothfels, Karen, 2020-03-25
crossReference	COSMIC:COSM12444
definition	
disease	cancer gastrointestinal stromal tumor
edited	
figure	
goCellularComponent	
inferredFrom	
inferredTo	
literatureReference	Pediatric mastocytosis is a clonal disease associated with D816V ...
modified	
name	KIT A502_Y503dup
reviewed	

Fusion proteins

Fusion proteins are the consequence of an in-frame fusion of coding sequences from two different genes, usually through chromosomal translocation. By convention, Reactome annotates these fusions as the insertion of a stretch of the C-terminal partner as a FragmentInsertionModification into the hasModifiedResidue slot of the N-terminal partner.

Naming

Format: “prefix (N-terminal protein)” + “(included amino acids from N-terminal protein)” + ”-” “prefix (C-terminal protein)” + “(included amino acids from C-terminal protein)” +”fusion”.

Where “prefix” is the HGNC gene name and (included amino acids) represents the start and end coordinates- but not the amino acid codes- of each fusion partner relative to their WT reference sequences, in accordance with Reactome’s controlled vocabulary for naming of protein fragments.

Example: ZMYM2(1-913)-FGFR1(429-822) - this is the insertion of FGFR1 residues 429-822 (relative to full-length WT FGFR1) after residue 913 of ZMYM2.

How to annotate

Fusions are captured as “FragmentInsertionModifications” class, specifying the start and stop codon of the C-terminal partner with coordinates from the WT C-terminal partner as shown below:

[FragmentInsertionModification:1604543] Insertion of residues 429 to 822 at 914 from UniProt:P11362 FGFR1

FragmentInsertionModification Properties

Property Name	Value
coordinate	914
endPositionInReferenceSequence	822
referenceSequence	UniProt:P11362 FGFR1
startPositionInReferenceSequence	429
DB_ID	1604543
_displayName	Insertion of residues 429 to 822 at 914 from UniProt:P11362 FGFR1

In the EWAS record, the end coordinate of the N-terminal partner is changed to the last amino acid of the N-terminal partner sequence that is contained in the fusion.

[EntityWithAccessionedSequence:1604544] ZMYM2(1-913)-FGFR1(429-822) fusion [cytosol]

EntityWithAccessionedSequence Properties

Property Name	Value
cellType	
compartment	cytosol
endCoordinate	913
hasModifiedResidue	Insertion of residues 429 to 822 at 914 from UniProt:P11362 FGFR1
referenceEntity	UniProt:Q9UBW7 ZMYM2
startCoordinate	1
DB_ID	1604544
_displayName	ZMYM2(1-913)-FGFR1(429-822) fusion [cytosol]
authored	
created	Rothfels, K, 2011-09-16
crossReference	
definition	
disease	subacute leukemia
	myelodysplastic myeloproliferative cancer
	precursor lymphoblastic lymphoma/leukemia
edited	
figure	
goCellularComponent	
inferredFrom	
inferredTo	
literatureReference	FGFR1 is fused with a novel zinc-finger gene, ZNF198, in the t(8;13) leukae...
	Consistent fusion of ZNF198 to the fibroblast growth factor receptor-1 in the...
	Fibroblast growth factor receptor 1 is fused to FIM in stem-cell myeloprolifer...
modified	Rothfels, K, 2011-09-21
	Rothfels, K, 2011-10-05
	Rothfels, K, 2011-11-22
	Rothfels, Karen, 2014-11-20
	Rothfels, Karen, 2021-09-13
name	ZMYM2(1-913)-FGFR1(429-822) fusion
	ZMYM2-FGFR1 fusion
	t(8;13) ZMYM2-FGFR1 fusion
	ZNF198-FGFR1 fusion
	ZNF198(1-913)-FGFR1(429-822) fusion
	t(8;13) ZNF198-FGFR1 fusion
	FIM-FGFR1 fusion
	RAMP-FGFR1 fusion

Post-translational modifications to either partner in the fusion protein are numbered according to each respective WT reference protein sequence and do not reflect the amino acid positions in the fusion. For instance, if in the context of the fusion protein, the FGFR1 partner is phosphorylated at (WT FGFR1 position) Y766, the fusion EWAS would be ZMYM2-pY766-FGFR1, despite the fact that in linear sequence the phosphorylation occurs at residue 1250 of the fusion (913+(822-429).

Fusion proteins with in-frame non-coding insertion sequences

If a fusion protein includes an in-frame insertion of amino acids encoded by an intronic or UTR sequence that are not present in the reference protein sequences of either of the two fusion partners, the insertion is defined in the name of the fusion protein in the format:

Naming

“prefix (N-terminal protein)”+ “(included amino acids from N-terminal protein)” + “ins” + “inserted sequence” + “-” + “prefix (C-terminal protein)” + “(included amino acids from C terminal protein)”+ fusion,

where the inserted sequence is listed as single letter amino acid codes.

Example: FIP1L1(1-391)insAQCP-PDGFRA(577-1089) fusion - this is an insertion of the amino acid string AQCP between the amino acid 391 of FIP1L1 and amino acid 577 of PDGFRA.

How to annotate:

An in-frame insertion of an amino acid string encoded by an intronic or UTR sequence that is not present in the reference protein sequences of either of the two fusion partners is annotated using the replacedFragmentModification, specifying that the last amino acid of the N-terminal fusion partner (proline at position 391 of FIP1L1 in this example) is replaced by a string of amino acids, where the first amino acid in the string is identical to the last amino acid in the N-terminal fragment, followed by the actual inserted string - PAQCP in this example, with the actual inserted string shown in bold.

[FragmentReplacedModification:9672470] Replacement of residues 391 to 391 b...

FragmentReplacedModification Properties

Property Name	Value
alteredAminoAcidFragment	PAQCP
endPositionInReferenceSequence	391
referenceSequence	UniProt:Q6UN15 FIP1L1
startPositionInReferenceSequence	391
DB_ID	9672470
_displayName	Replacement of residues 391 to 391 by PAQCP

In the EWAS record, the end coordinate of the N-terminal partner is changed to the last amino acid of the N-terminal partner sequence that is contained in the fusion.

[EntityWithAccessionedSequence:9672502] FIP1L1(1-391)insAQCP-PDGFRA(577-1089) fusion [cytosol]

EntityWithAccessionedSequence Properties

Property Name	Value
cellType	
compartment	cytosol
endCoordinate	391
hasModifiedResidue	Replacement of residues 391 to 391 by PAQCP Insertion of residues 577 to 1089 at 392 from UniProt:P16234 P...
referenceEntity	UniProt:Q6UN15 FIP1L1
startCoordinate	1
DB_ID	9672502
_displayName	FIP1L1(1-391)insAQCP-PDGFRA(577-1089) fusion [cytosol]

Indels

In-frame indels are variants that are the consequence of simultaneous insertion and removal of a stretch of amino acids from the WT sequence without change in the reading frame..

Naming

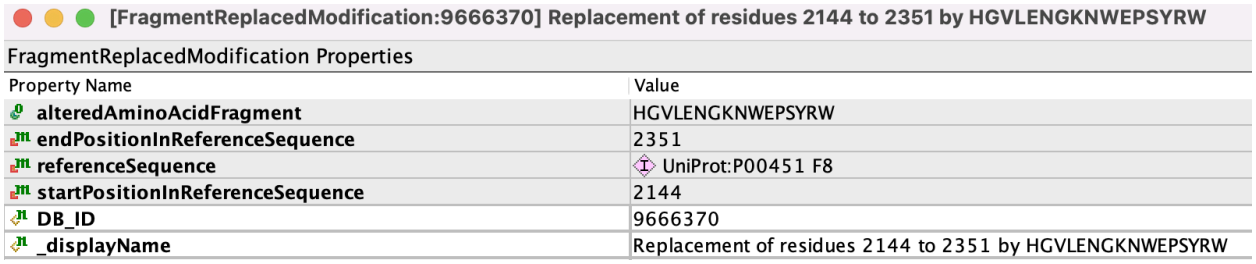
Format: “prefix” + “first deleted amino acid” + “position of first deleted amino acid” + “_” + “last deleted amino acid” + “position of last deleted amino acid” + “delins” + “inserted sequence”

Where “prefix” is the HGNC gene name, amino acids+positions deleted specifies position and single letter code of the removed amino acids and inserted sequence specifies the new amino acids with their single letter codes.

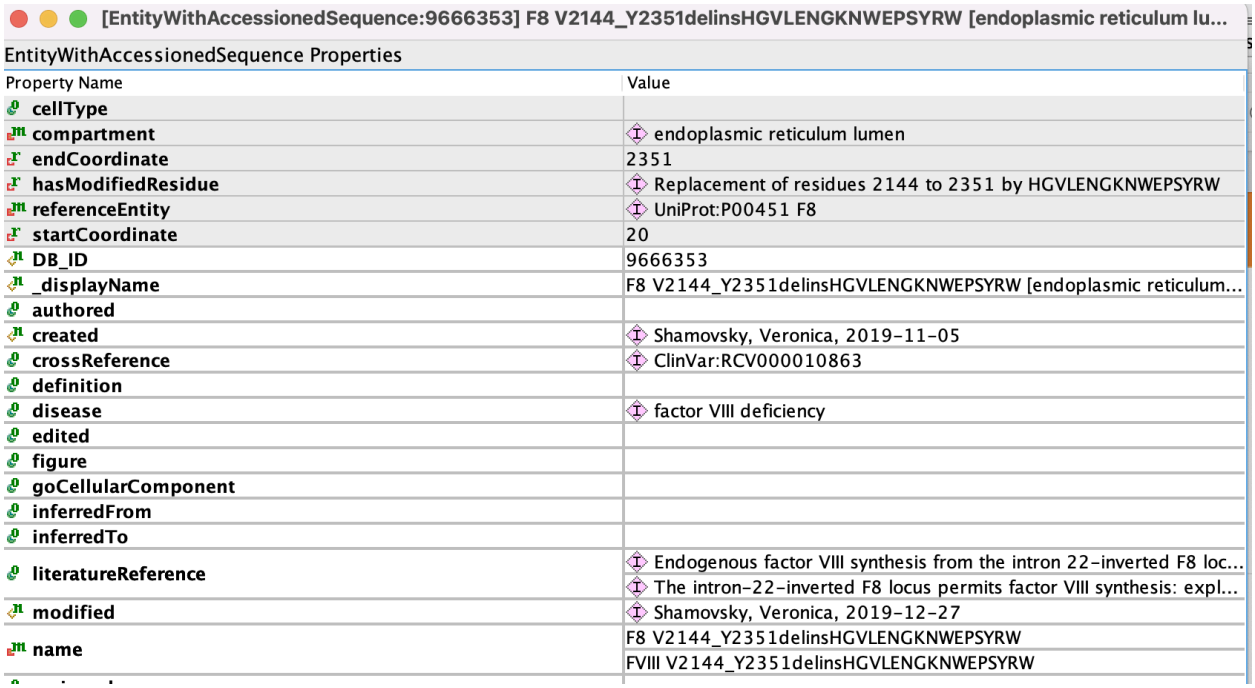
Example: F8 V2144_Y2351delinsHGVLENGKNWEPSYRW - this is removal of residues V2144 to Y2351 from F8 and their replacement with the specified stretch of amino acids

How to annotate

In-frame indels are captured with the FragmentReplacedModification class, specifying the start and end coordinates of the deleted amino acid stretch and specifying the inserted amino acid stretch that replaces them as an alteredAminoAcidfragment, using single letter codes entered as the start and stop positions, as shown below.



The end coordinate of the variant protein on the EWAS record is not changed relative to the WT protein to accommodate the indel.


[EntityWithAccessionedSequence:9666353] F8 V2144_Y2351delinsHGVLENGKNWEPSYRW [endoplasmic reticulum lu...

EntityWithAccessionedSequence Properties

Property Name	Value
cellType	
compartment	endoplasmic reticulum lumen
endCoordinate	2351
hasModifiedResidue	Replacement of residues 2144 to 2351 by HGVLENGKNWEPSYRW
referenceEntity	UniProt:P00451 F8
startCoordinate	20
DB_ID	9666353
_displayName	F8 V2144_Y2351delinsHGVLENGKNWEPSYRW [endoplasmic reticulum...
authored	
created	Shamovsky, Veronica, 2019-11-05
crossReference	ClinVar:RCV000010863
definition	
disease	factor VIII deficiency
edited	
figure	
goCellularComponent	
inferredFrom	
inferredTo	
literatureReference	Endogenous factor VIII synthesis from the intron 22-inverted F8 loc...
	The intron-22-inverted F8 locus permits factor VIII synthesis: expl...
modified	Shamovsky, Veronica, 2019-12-27
name	F8 V2144_Y2351delinsHGVLENGKNWEPSYRW
	FVIII V2144_Y2351delinsHGVLENGKNWEPSYRW

Frameshifts

Perhaps the most challenging to annotate! These are often incompletely described in papers, and in order to capture the amino acid changes that result, the curator may have to do quite a bit of leg work.

Frameshift variants are the consequence of mutations that lead to the insertion or deletion of nucleotides where the total number of inserted or deleted nucleotides is not a multiple of three, leading to a change in the reading frame.The published literature generally provides an exact description of the mutation at the *DNA* level, from which the curator may need to infer the altered C-terminal sequence of the mutated protein to create a “FragmentReplacedModifcation” in Reactome.

Before embarking on this exercise, however, consider the biology. Most (perhaps all) nonsense mutations that occur anywhere in a gene except its final exon trigger the process of nonsense-mediated decay. This process is part of mRNA maturation and its effect is to cause the mutant RNA to be degraded before it is ever translated. Thus, at the protein level annotated by Reactome, the result of the mutation is almost certainly disappearance of the protein, not appearance of an altered one. If the goal is to annotate the existence of a non-functional protein a better choice, other things being equal, would be to choose a missense mutation that has been shown to encode a protein whose normal activity (binding, catalysis, transport, etc.) is lost.

Naming

Format: “prefix” + “first amino acid altered due to frameshift” + “position of the first altered amino acid” + “fs” + “*” + “position of the stop codon relative to the first altered amino acid”

Where “prefix” is the HGNC gene name, and is followed by a single letter code description of the first affected amino acid, “fs” indicates frameshift, and “*#” indicates the length of amino acids before the first stop codon is reached.

Example: EXT1 L490Rfs*9 - in this protein, a frameshift mutation causes conversion of L490 to R, and is followed by a novel stretch of 7 amino acids (8 altogether, counting from the first substituted amino acid) before a stop codon is reached at the 9th position.

How to annotate

Frameshifts are captured using the FragmentReplacedModification class, specifying the start and end coordinates of the altered sequence, where the startPositionInReferenceSequence matches the position of the first amino acid substituted due to frameshift (490 in this example) and the endPositionInReferenceSequence is calculated by the formula:

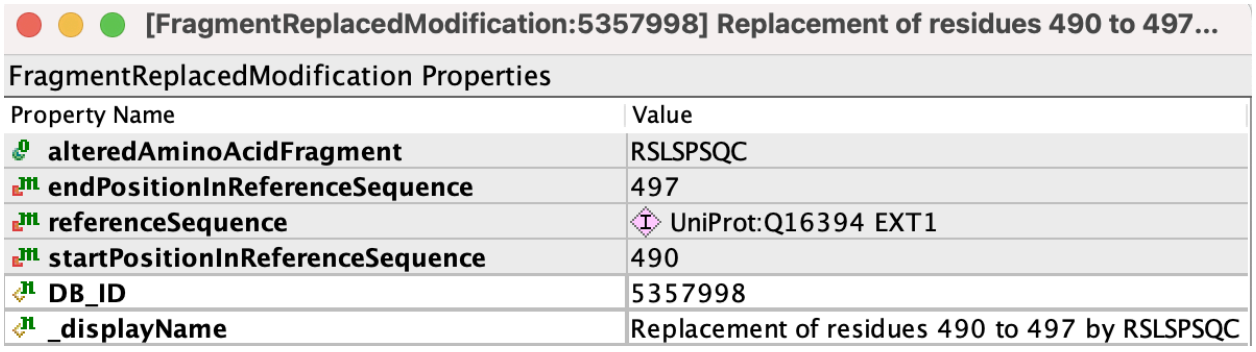
endPositionInReferenceSequence = startPositionInReferenceSequence + frameshift length - 2

The frameshift length is specified in the name of the frameshift variant (the number after the asterisk and equals the length of the substituted amino acid string + 1 (as it takes into account the stop codon).

In the EXT1 L490Rfs*9 example,

endPositionInReferenceSequence = 490 + 9 - 2 = 497

The altered amino acid sequence itself is specified using the alteredAminoAcidFragment attribute. The end coordinate of the EWAS is changed to reflect the new C-terminal amino acid.



If the paper does not completely describe all the details needed to annotate a frameshift variant, and if such details cannot be obtained from a disease variant database, they have to be deciphered.

The basic steps are

1. Alter the reference coding sequence for the wild-type protein to match the coding sequence of the variant. The reference coding sequence can be obtained from teh RefSeq mRNA cross-referenced by the UniProt record that is used as the reference for the wild-type proteins.
2. Translate the altered coding sequence in silico using EMBOSS TransSeq.
3. Align the resultant mutant peptide sequence with wild type protein sequence using EMBOSS Needle open software to determine the alteredAminoAcidFragment and the anime of the frameshift protein variant if not provided in the research article.

UniProt records provide the normal peptide sequence in FASTA format. They also provide a link to RefSeq (under the “Sequence databases” section) which links to the NCBI reference mRNA sequence. This is the starting point from where the nucleotide sequence is copied to Transeq below.

Several EMBOSS programs are available to translate and compare sequences.

*Transeq (https://www.ebi.ac.uk/jdispatcher/st/emboss_transeq) translates nucleic acid sequences to their corresponding peptide sequences.

*Needle (https://www.ebi.ac.uk/jdispatcher/psa/emboss_needle) can align two peptide sequences to provide a global sequence alignment.

An example below shows how these steps are utilized to decipher a frameshift mutation.

The disorder “Hereditary multiple exostoses 1” (EXT1) is caused by loss of function mutations in Exostosin 1 (EXT1). One mutation is a 1-bp deletion at nucleotide 1469 in the EXT1 gene, resulting in a frameshift mutation with a premature stop codon (Philippe et al. 1997 , Ahn et al. 1995).

Go to the UniProt record for human EXT1 (https://www.uniprot.org/uniprotkb/Q16394/entry), and scroll to the FASTA sequence for the WT peptide sequence.



Sequence[i]

Tools ▾ Download Add Highlight ▾ Copy sequence

Length 746
Mass (Da) 86,255

Last updated 2002-03-27 v2
Checksum[i] 842CD7E6C1312C1A

```
        10         20         30         40         50         60         70         80         90
MQAKKRYFIL LSAGSCLALL FYFGGLQFRA SRSHSRREEH SGRNGLHHPS PDHFWPRFPD ALRPFVPWDQ LENEDSSVHI SPRQKRDANS

       100        110        120        130        140        150        160        170        180
SIYKGKKCRM ESCFDFTLCK KNGFKVYVYP QQKGEKIAES YQNILAAIEG SRFYTSDPSQ ACLFVLSLDT LDRDQLSPQY VHNLRSKVQS

       190        200        210        220        230        240        250        260        270
LHLWNNGRNH LIFNLYSGTW PDYTEDVGFD IGQAMLAKAS ISTENFRPNF DVSIPLFSKD HPRTGGERGF LKFNTIPPLR KYMLVFKGKR

       280        290        300        310        320        330        340        350        360
YLTGIGSDTR NALYHVHNGE DVVLLTTCKH GKDWQKHKDS RCDRDNTEYE KYDYREMLHN ATFCLVPRGR RLGSFRFLEA LQAACVPVML

       370        380        390        400        410        420        430        440        450
SNGWELPFSE VINWNQAAVI GDERLLLQIP STIRSIHQDK ILALRQQTQF LWEAYFSSVE KIVLTTLEII QDRIFKHISR NSLIWNKHPG

       460        470        480        490        500        510        520        530        540
GLFVLPQYSS YLGDFPYYYA NLGLKPPSKF TAVIHAVTPL VSQSQPVLKL LVAAAKSQYC AQIIVLWNCD KPLPAKHRWP ATAVPVVVIE

       550        560        570        580        590        600        610        620        630
GESKVMSSRF LPYDNIITDA VLSLDEDTVL STTEVDFAFT VWQSFPERIV GYPARSHFWD NSKERWGYTS KWTNDYSMVL TGAAIYHKYY

       640        650        660        670        680        690        700        710        720
HYLYSHYLPA SLKNMVDQLA NCEDILMNFL VSAVTKLPPI KVTQKKQYKE TMMGQTSRAS RWADPDHFAQ RQSCMNTFAS WFGYMPLIHS

       730        740
QMRLDPVLFK DQVSILRKKY RDIERL
```

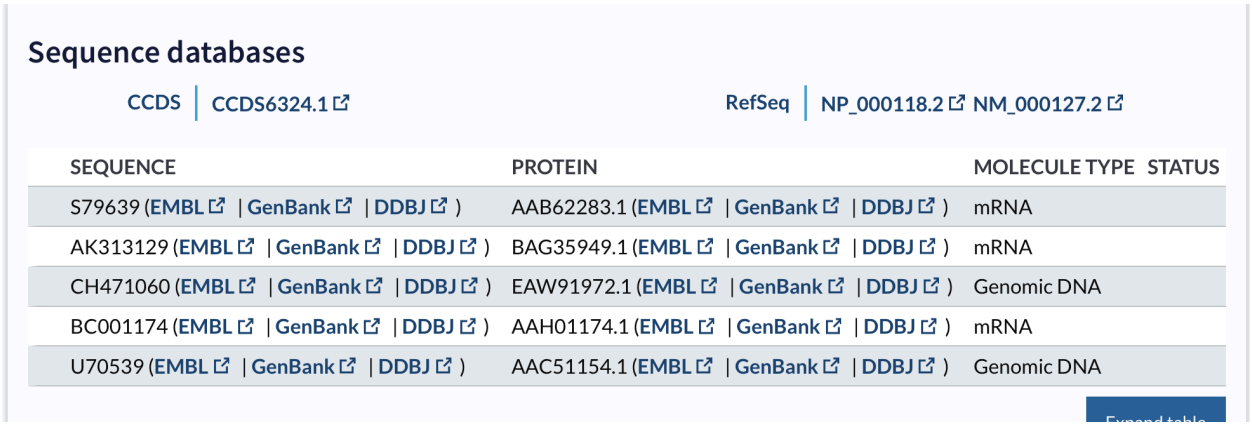
Click the “Download” link to obtain a text version of the sequence.

```
>sp|Q16394|EXT1_HUMAN Exostosin-1 OS=Homo sapiens OX=9606 GN=EXT1 PE=1 SV=2
MQAKKRYFILLSAGSCLALLFYFGGLQFRASRSHSRREEHSGRNGLHHPSPDHFWPRFPD
ALRPFVPWDQLENEDSSVHISPRQKRDANSSIYKGKKCRMESCFDFTLCKKNGFKVYVYP
QQKGEKIAESYQNILAAIEGSRFYTSDPSQACLFVLSLDTLDRDQLSPQYVHNLRSKVQS
LHLWNNGRNHLIFNLYSGTWPDYTEDVGFDIGQAMLAKASISTENFRPNFDVSIPLFSKD
HPRTGGERGFLKFNTIPPLRKYMLVFKGKRYLTGIGSDTRNALYHVHNGEDVVLLTTCKH
GKDWQKHKDSRCDRDNTEYEKYDYREMLHNATFCLVPRGRRLGSFRFLEALQAACVPVML
SNGWELPFSEVINWNQAAVIGDERLLLQIPSTIRSIHQDKILALRQQTQFLWEAYFSSVE
KIVLTTLEIIQDRIFKHISRNSLIWNKHPGGLFVLPQYSSYLGDFPYYYANLGLKPPSKF
TAVIHAVTPLVSQSQPVLKLLVAAAKSQYCAQIIVLWNCDKPLPAKHRWPATAVPVVVIE
GESKVMSSRFLPYDNIITDAVLSLDEDTVLSTTEVDFAFTVWQSFPERIVGYPARSHFWD
NSKERWGYTSKWTNDYSMVLTGAAIYHKYYHYLYSHYLPASLKNMVDQLANCEDILMNFL
VSAVTKLPPIKVTQKKQYKETMMGQTSRASRWADPDHFAQRQSCMNTFASWFGYMPLIHS
QMRLDPVLFKDQVSILRKKYRDIERL
```

Copy the peptide sequence to EMBOSS Needle and paste it into the first window.

A mutated nucleotide sequence needs to be created from which the translated mutant peptide sequence will be obtained.

In the UniProt record for EXT1, scroll to Sequence databases and select the nucleotide sequence from RefSeq (format NM_xxxxxx.x)



Sequence databases

CCDS | CCDS6324.1 RefSeq | NP_000118.2 NM_000127.2

SEQUENCE	PROTEIN	MOLECULE TYPE	STATUS
S79639 (EMBL | GenBank | DDBJ)	AAB62283.1 (EMBL | GenBank | DDBJ)	mRNA	
AK313129 (EMBL | GenBank | DDBJ)	BAG35949.1 (EMBL | GenBank | DDBJ)	mRNA	
CH471060 (EMBL | GenBank | DDBJ)	EAW91972.1 (EMBL | GenBank | DDBJ)	Genomic DNA	
BC001174 (EMBL | GenBank | DDBJ)	AAH01174.1 (EMBL | GenBank | DDBJ)	mRNA	
U70539 (EMBL | GenBank | DDBJ)	AAC51154.1 (EMBL | GenBank | DDBJ)	Genomic DNA	

Expand table

This opens the NCBI nucleotide record for EXT1 mRNA. (https://www.ncbi.nlm.nih.gov/nuccore/NM_000127.2). Scroll down to the “ORIGIN” section where you will find the nucleotide sequence conveniently numbered. Copy the sequence and paste it into EMBOSS Transeq.

```
ORIGIN
        1 ggcgaccgaa cgcggcggtc ggcagcgttc gcgcgggggc ctgcgaagcg ctgctcgggg
       61 ccggcactgc ccgcggggag gacgcgccgc cgccgccacc cagcgccgcc gccgccgccg
      121 cctccagccg ggccgccgcg cgtcccgggg gccggccccg cgagcgcagg agtaaacacc
      181 gccggagtct tggagccgct gcagaaggga ataaagagag atgcagggat ttgtgaggtt
      241 acggcgcccc agctgcaaga tgcactagcc ggctgaaccc gggatcggct gacttgttgg
      301 aaccggagtg ctctgcacgg agagtggtgg atgagttgaa gttgccttcc cggggctcat
      361 tttccacgct gccgagagga atccgagagg caaggcaatc acttcgtctt gccattgatt
      421 gggtatcggg agctttttt ttctcccctc tctctttctt ttcctccgtc ttgttgcatg
      481 caagaaaatt acagtccgct gctcgcccgc cctgggtgcg agatattcag ccccgctctc
      541 tcccgtgcat tgtgcaaccc aaagatgaaa gaccgaaggg gagaaagtta aagaaatcgc
      601 ccacatgcgc tggatcagtc cacggcttgg ggaaaggcat ccagagaagg tgggagcgga
      661 gagtttgaag tctttacagg cgggaagatg gcggactgga gctgaaagtg ttgattggga
      721 aacttgggtg attcttgtgt ttatttacaa tcctcttgac ccaggcagga cacatgcagg
      781 ccaaaaaacg ctatttcatc ctgctctcag ctggctcttg tctcgccctt ttgttttatt
      841 tcggaggctt gcagtttagg gcatcgagga gccacagccg gagagaagaa cacagcggta
      901 ggaatggctt gcaccacccc agtccggatc atttctggcc ccgcttcccg gacgctctgc
      961 gccccttcgt tccttgggat caattggaaa acgaggattc cagcgtgcac atttcccccc
     1021 ggcagaagcg agatgccaac tccagcatct acaaaggcaa gaagtgccgc atggagtcct
     1081 gcttcgattt cacccttttgc aagaaaaacg gcttcaaagt ctacgtatac ccacagcaaa
     1141 aaggggagaa aatcgccgaa agttaccaaa acattctagc ggccatcgag ggctccaggt
     1201 tctacacctc ggaccccagc caggcgtgcc tctttgtcct gagtctggat actttagaca
     1261 gagaccagtt gtcacctcag tatgtgcaca atttgagatc caaagtgcag agtctccact
     1321 tgtggaacaa tggtaggaat catttaattt ttaatttata ttccggcact tggcctgact
     1381 acaccgagga cgtggggttt gacatcggcc aggcgatgct ggccaaagcc agcatcagta
     1441 ctgaaaactt ccgacccaac tttgatgttt ctattcccct cttttctaag gatcatccca
     1501 ggacaggagg ggagaggggg tttttgaagt tcaacaccat ccctcctctc aggaagtaca
     1561 tgctggtatt caaggggaag aggtacctga cagggatagg atcagacacc aggaatgcct
     1621 tatatcacgt ccataacggg gaggacgttg tgctcctcac cacctgcaag catggcaaag
     1681 actggcaaaa gcacaaggat tctcgctgtg acagagacaa caccgagtat gagaagtatg
     1741 attatcggga aatgctgcac aatgccactt tctgtctggt tcctcgtggt cgcaggcttg
     1801 ggtccttcag attcctggag gctttgcagg ctgcctgcgt ccctgtgatg ctcagcaatg
     1861 gatgggagtt gccattctct gaagtgatta attggaacca agctgccgtc ataggcgatg
     1921 agagattgtt attacagatt ccttctacaa tcaggtctat tcatcaggat aaaatcctag
     1981 cacttagaca gcagacacaa ttcttgtggg aggcttattt ttcttcagtt gagaagattg
     2041 tattaactac actagagatt attcaggaca gaatattcaa gcacatatca cgtaacagtt
     2101 taatatggaa caaacatcct ggaggattgt tcgtactacc acagtattca tcttatctgg
     2161 gagattttcc ttactactat gctaatttag gtttaaagcc cccctccaaa ttcactgcag
     2221 tcatccatgc ggtgaccccc ctggtctctc agtcccagcc agtgttgaag cttctcgtgg
     2281 ctgcagccaa gtcccagtac tgtgcccaga tcatagttct atggaattgt gacaagcccc
     2341 taccagccaa acaccgctgg cctgccactg ctgtgcctgt cgtcgtcatt gaaggagaga
     2401 gcaaggttat gagcagccgt tttctgccct acgacaacat catcacagac gccgtgctca
     2461 gccttgacga ggacacggtg ctttcaacaa cagaggtgga tttcgccttc acagtgtggc
     2521 agagcttccc tgagaggatt gtggggtacc ccgcgcgcag ccacttctgg gataactcta
     2581 aggagcggtg gggatacaca tcaaagtgga cgaacgacta ctccatggtg ttgacaggag
     2641 ctgctattta ccacaaatat tatcactacc tatactccca ttacctgcca gccagcctga
     2701 agaacatggt ggaccaattg gccaattgtg aggacattct catgaacttc ctggtgtctg
     2761 ctgtgacaaa attgcctcca atcaaagtga cccagaagaa gcagtataag gagacaatga
     2821 tgggacagac ttctcgggct tcccgttggg ctgaccctga ccactttgcc cagcgacaga
     2881 gctgcatgaa tacgtttgcc agctggtttg gctacatgcc gctgatccac tctcagatga
     2941 ggctcgaccc cgtcctcttt aaagaccagg tctctatttt gaggaagaaa taccgagaca
     3001 ttgagcgact ttgaggaatc cggctgagtg ggggagggga agcaagaagg gatgggggtc
     3061 aagctgctct ctcttcccag tgcagatcca ctcatcagca gagccagatt gtgccaacta
     3121 tccaaaaact tagatgagca gaatgacaaa aaaaaaaagg ccaatgagaa ctcaactcct
     3181 ggctcctggg actgcaccag actgctccaa actcacctca ctggcttctg tgtcccaaga
     3241 ctaggttgtg tacagtttaa ttatggaaca ttaaataatt atttttgaaa tgattgctat
     3301 gcaggtttaa actttttaa tgatcaaaac tattaaaaac cagagttctt tgtttaatca
     3361 aaaaaaaaaa aaaaaa
//
```

In the Transeq window where the sequence is copied to, edit the nucleotide sequence according to evidence from literature. The coding sequence in the RefSeq record starts at position 774. The nucleotide change, from literature, is a 1-bp thymidine deletion at 1469 (Philippe et al. 1997] (table 2), Ahn et al. 1995] (Fig.6)).

Table 2

Characteristics of EXT1 and EXT2 Gene Defects in Patients with Multiple Exostoses

Family	Location of Mutation	Mutation	Consequence	Name of Mutation[a]
Familial cases:				
1	EXT1/Intron 5	TAGgta→TAGata	Altered splicing	1417 (+1G→A)
2	EXT2/Exon 4	TAC→TAG(Tyr→Stop)	Nonsense	Y222X
6	EXT2/Exon 4	GAT→AAT (Asp→Asn)	Missense	D227N
9	EXT1/Exon 3	TGG→TGA (Trp→Stop)	Nonsense	W374X
19	EXT2/Exon 2	TATATC→TATTATC	Frameshift	77- or 78insT
21	EXT1/Exon2	GGT→GAT (Gly→Asp)	Missense	G339D
25	EXT2/Exon4	GAT→AAT (Asp→Asn)	Missense	D227N
36	EXT1/Exon 6	CCCCCCTGG→CCCCCCGG	Frameshift	1469delT[b]
Isolated cases (*de novo* mutations):				
5	EXT1/Exon 6	CCCCCCTGG→CCCCCCGG	Frameshift	1469delT[b]
7	EXT2/ Exon 8	CAG→TAG (Glu→Stop)	Nonsense	Q401X
13	EXT1/Exon 1	GCTTCCCGG→GCTTCCGG	Frameshift	174-, 175- or 176delC
35	EXT1/Exon 2	CGC→TGC (Arg→Cys)	Missense	R340C

[a] According to the nomenclature defined by Beaudet and Tsui (1993). For both EXT1 and EXT2, base no. 1 corresponds to the first base of the initiation codon in the cDNA sequences (Ahn et al. 1995; Stickens et al. 1996; Wuyts et al. 1996).

[b] Already reported by Ahn et al. (1995) and by Wells et al. (1997).

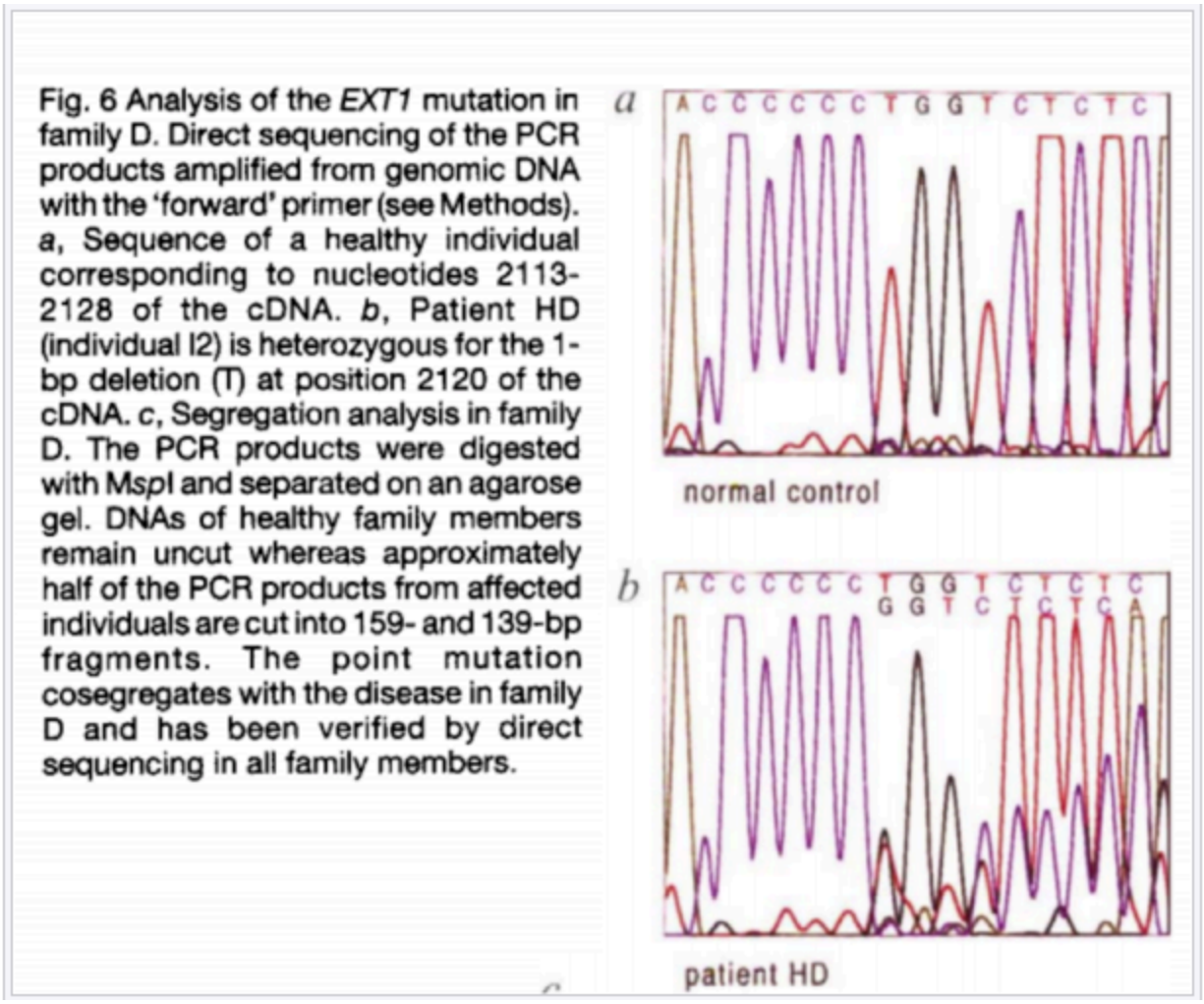
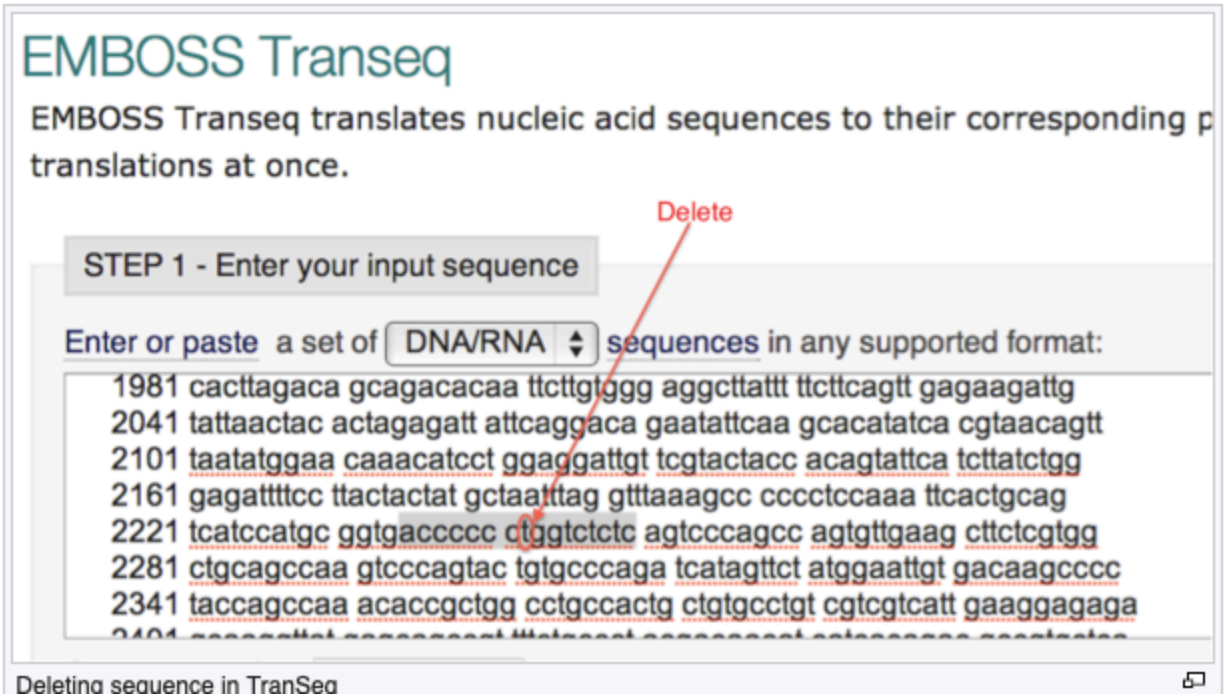


Fig. 6 Analysis of the *EXT1* mutation in family D. Direct sequencing of the PCR products amplified from genomic DNA with the 'forward' primer (see Methods). *a*, Sequence of a healthy individual corresponding to nucleotides 2113-2128 of the cDNA. *b*, Patient HD (individual I2) is heterozygous for the 1-bp deletion (T) at position 2120 of the cDNA. *c*, Segregation analysis in family D. The PCR products were digested with *Msp*I and separated on an agarose gel. DNAs of healthy family members remain uncut whereas approximately half of the PCR products from affected individuals are cut into 159- and 139-bp fragments. The point mutation cosegregates with the disease in family D and has been verified by direct sequencing in all family members.

Many times in literature, researchers only use the coding sequence in their experiments and count from that point onwards to the mutation site. As the coding sequence starts at 774, add that to 1469 (the deletion position) which equals 2243. Check this region in the sequence pasted into Transeq. The region matches that found in literature. Delete the t after the 6 c's and submit the job.

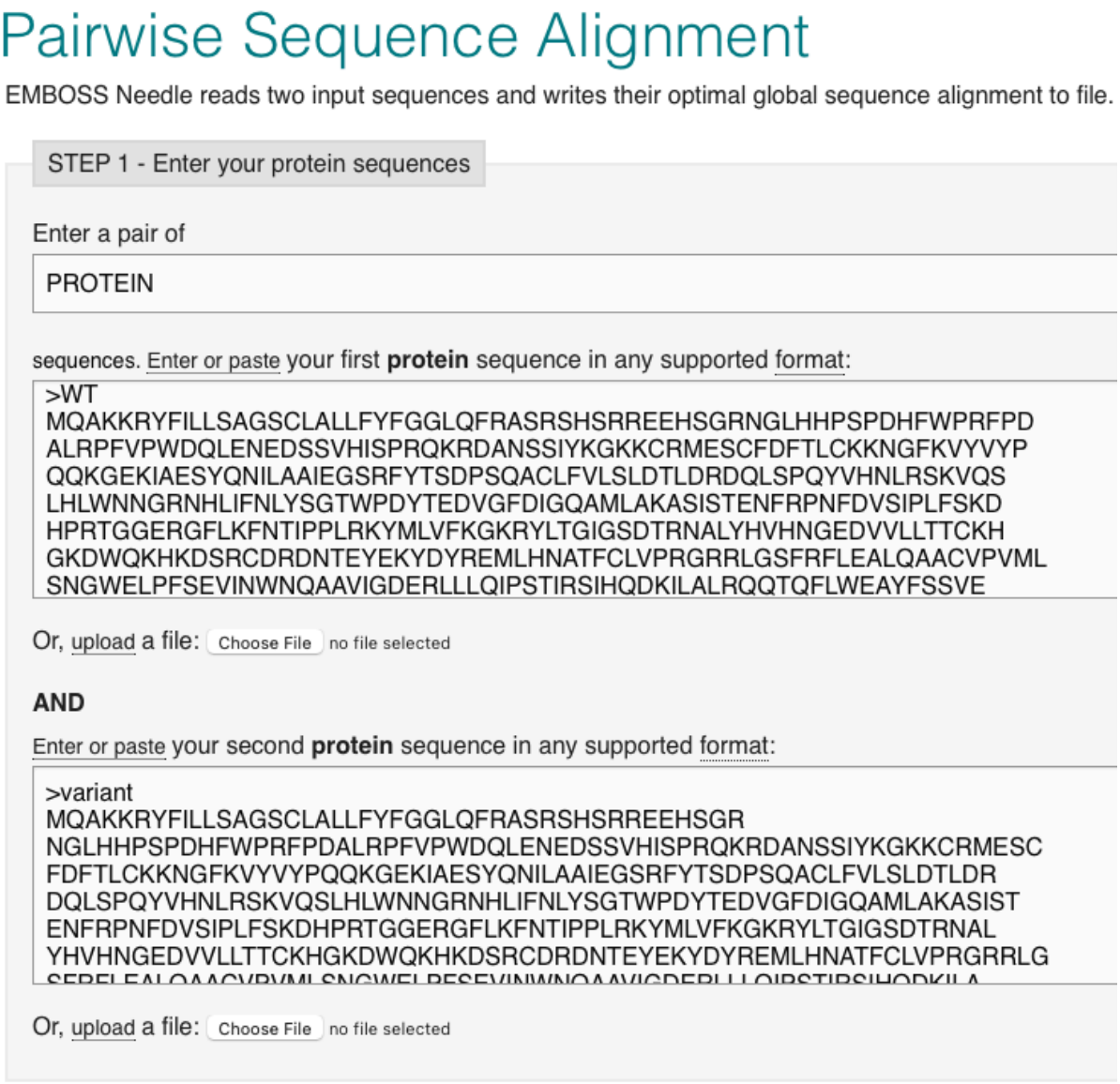
(Note, TranSeq can translate your sequence in any of the 6 reading frames. If you don't know which frame your initiator methionine is, you can translate the WT seq in all 3 forward reading frames and search the results for the amino acid sequence of the WT protein to orient yourself before submitting the variant allele DNA as described in the figure below.)



The result is below. Copy the mutant peptide sequence from the start to the first stop ("*") in the sequence

```
>EMBOSS_001_3
RPNAAVGSVRAGACEALLGAGTARGEDAPPPPPSAAAAAASSRAAARPGGRPRERRSKHR
RSLGAAAEGNKERCRDL*GYGAPAARCTSRLNPGSADLLEPECSARRVVDELKLPSRGSF
STLPRGIREARQSLRLAIDWVSGAFFFSPLSFFSSVLLHARKLQSAARPPWVRDIQPRSL
PCIVQPKDERPKGRKLKKSPTCAGSVHGLGKGIQRRWERRV*SLYRREDGGLELKVLIGK
LG*FLCLFTILLTQAGHMQAKKRYFILLSAGSCLALLFYFGGLQFRASRSHSRREEHSGR
NGLHHPSPDHFWPRFPDALRPFVPWDQLENEDSSVHISPRQKRDANSSIYKGKKCRMESC
FDFTLCKKNGFKVYVYPQQKGEKIAESYQNILAAIEGSRFYTSDPSQACLFVLSLDTLDR
DQLSPQYVHNLRSKVQSLHLWNNGRNHLIFNLYSGTWPDYTEDVGFDIGQAMLAKASIST
ENFRPNFDVSIPLFSKDHPRTGGERGFLKFNTIPPLRKYMLVFKGKRYLTGIGSDTRNAL
YHVHNGEDVVLLTTCKHGKDWQKHKDSRCDRDNTEYEKYDYREMLHNATFCLVPRGRRLG
SFRFLEALQAACVPVMLSNGWELPFSEVINWNQAAVIGDERLLLQIPSTIRSIHQDKILA
LRQQTQFLWEAYFSSVEKIVLTTLEIIQDRIFKHISRNSLIWNKHPGGLFVLPQYSSYLG
DFPYYYANLGLKPPSKFTAVIHAVTPRSLSPSQC*SFSWLQPSPSTVPRS*FYGIVTSPY
QPNTAGLPLLCLSSSLKERARL*AAVFCPTTTSSQTPCSALTRTRCFQQQRWISPSQCGR
ASLRGLWGTPRAATSGITLRSGGDTHQSGRTTTPWC*QELLFTTNIITTYTPITCQPA*R
TWWTNWPIVRTFS*TSWCLL*QNCLQSK*PRRSSIRRQ*WDRLLGLPVGLTLTTLPSDRA
A*IRLPAGLATCR*STLR*GSTPSSLKTRSLF*GRNTETLSDFEESG*VGEGKQEGMGVK
LLSLPSADPLISRARLCQLSKNLDEQNDKKKKANENSTPGSWDCTRLLQTHLTGFCVPRL
GCVQFNYGTLNNYF*NDCYAGLNFFNDQNY*KPEFFV*SKKKKKX
```

Paste into the second window of the Needle program. You had already pasted the w/t sequence (FASTA format) at the beginning of this exercise.

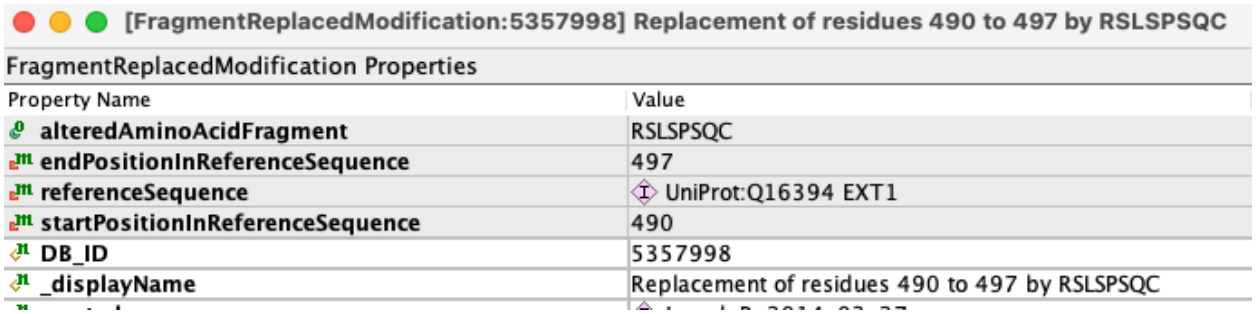


Click submit. An alignment result is returned that compares w/t to mutant sequences that you had input into Needle. Follow the sequence (vertical bars between the two sequences indicate similarity) till you reach the first dot (a difference between the two sequences).

```
WT          401 ILALRQQTQFLWEAYFSSVEKIVLTTLEIIQDRIFKHISRNSLIWNKHPG    450
                ||||||||||||||||||||||||||||||||||||||||||||||||||
variant     401 ILALRQQTQFLWEAYFSSVEKIVLTTLEIIQDRIFKHISRNSLIWNKHPG    450

WT          451 GLFVLPQYSSYLGDFPYYYANLGLKPPSKFTAVIHAVTP-LVSQSQPVLK    499
                |||||||||||||||||||||||||||||||||||| .:|.||.
variant     451 GLFVLPQYSSYLGDFPYYYANLGLKPPSKFTAVIHAVTPRSLSPSQC---    497
```

The first amino acid to change due to the frameshift is leucine to arginine at position 490. The altered mutant sequence after the frameshift is “RSLSPSQC” then the mutant sequence terminates. We can now fill in our FragmentReplacedModification instance.



After 8 altered amino acids, the 9th position is a stop (“*”). The name of the EWAS will reflect the mutation and fill in the hasModifiedResidue slot with the FragmentReplacedModification. Note that here, the end coordinate of the EWAS is changed to reflect the new C-terminal amino acid.

[EntityWithAccessionedSequence:3730975] EXT1 L490Rfs*9 [Golgi membrane]

EntityWithAccessionedSequence Properties

Property Name	Value
cellType	
compartment	Golgi membrane
endCoordinate	497
hasModifiedResidue	Replacement of residues 490 to 497 by RSLSPSQC
referenceEntity	UniProt:Q16394 EXT1
startCoordinate	1
DB_ID	3730975
_displayName	EXT1 L490Rfs*9 [Golgi membrane]
authored	
created	Jassal, B, 2013-06-17
crossReference	
definition	
disease	hereditary multiple exostoses
edited	
figure	
goCellularComponent	
inferredFrom	
inferredTo	
literatureReference	
modified	Jassal, B, 2013-06-17
	Jassal, B, 2013-06-20
	Jassal, B, 2014-03-27
	Jassal, B, 2014-03-28
name	EXT1 L490Rfs*9
	Exostosin-1
	EXT1_HUMAN

Frameshifts that produce a stop in the first affected codon affected:

It is possible for a frameshift mutation to introduce an immediate stop codon. Although these should properly be annotated as frameshifts, Reactome is not able to capture a stop codon as a replaced residue, as would be required for FragmentReplacedModification as described above for lengthier frameshifts.

For this reason, frameshifts that produce a stop codon as the first affected codon are annotated as nonsense mutations as described above.

Frameshifts that produce a stop codon in the second affected codon:

Annotated as usual for a frameshift, as described above.

Mutations affecting the initiator methionine

Mutations creating a new downstream initiation site

A mutation that eliminates the initiator methionine and results in the aberrant initiation of translation from a downstream methionine is annotated as an in-frame deletion, as per HGVS recommendations, with the variant protein name specifying a deletion that stretches from the first amino acid in the reference sequence to the alternative initiator methionine.

Naming

Named in accordance with the deletion naming rules.

Format: “prefix” + “amino acid at position 2” + “2” + “_” + “M” +”position of methionine residue now used as initiator” + “del”

Where “prefix” is the HGNC gene name, the deleted amino acids are represented by single letter codes separated by underscore.

Example: SERPING1 A2_M53del - in this variant, a missense mutation substitutes the WT M1 with a leucine codon, leading to the usage of methionine at position 53 as an alternative start codon.

How to annotate

[FragmentDeletionModification:9651392] Deletion of residues 1 to 52

FragmentDeletionModification Properties

Property Name	Value
endPositionInReferenceSequence	52
referenceSequence	UniProt:P05155 SERPING1
startPositionInReferenceSequence	1
DB_ID	9651392
_displayName	Deletion of residues 1 to 52

Note that although by HGVS naming convention the variant EWAS is named SERPING1 A2_M53del, the FragmentDeletionModification record captures the biologically relevant deletion of residues M1-K52.

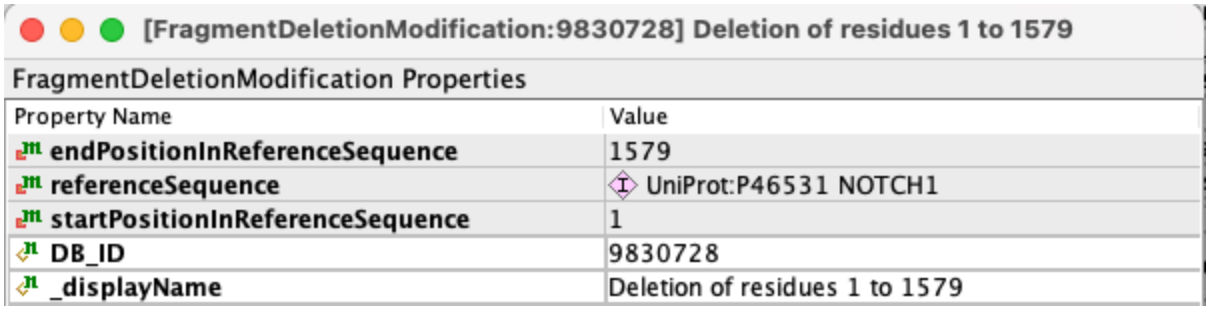
The protein variant EWAS record specifies the start coordinate as 53, without changing the end coordinate relative to the WT reference sequence.

[EntityWithAccessionedSequence:9651435] SERPING1 A2_M53del [endoplasmic reticu...

EntityWithAccessionedSequence Properties

Property Name	Value
cellType	
compartment	endoplasmic reticulum lumen
endCoordinate	500
hasModifiedResidue	Deletion of residues 1 to 52
referenceEntity	UniProt:P05155 SERPING1
startCoordinate	53
DB_ID	9651435
_displayName	SERPING1 A2_M53del [endoplasmic reticulum lumen]
authored	
created	Shamovsky, Veronica, 2019-06-25
crossReference	ClinVar:VCV000626353
definition	
disease	C1 inhibitor deficiency
edited	
figure	
goCellularComponent	
inferredFrom	
inferredTo	
literatureReference	Targeted next-generation sequencing for the mol...
	F12-46C/T polymorphism as modifier of the clini...
modified	Shamovsky, Veronica, 2020-01-10
	Orlic-Milacic, Marija, 2023-03-16
	Orlic-Milacic, Marija, 2023-03-17
	Orlic-Milacic, Marija, 2023-03-22
name	SERPING1 A2_M53del
	C1Inh A2_M53del
	plasma protease C1 inhibitor A2_M53del
	C1-In A2_M53del
	SERPING1 M1L
	SERPING1 Met1Leu

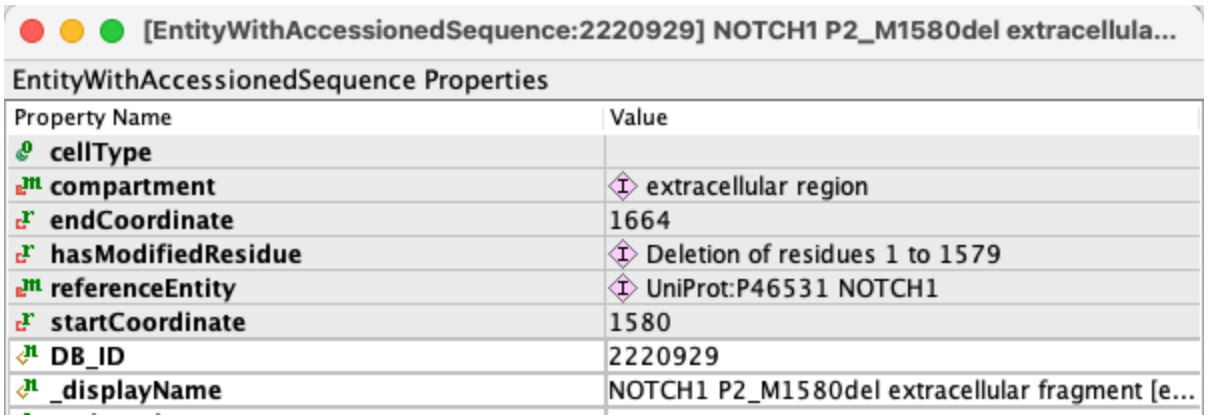
If a gene translocation fuses the promoter of one gene with the coding sequence of another gene, with accompanying truncation of the N' terminal coding sequence of the C-terminal fusion partner, the resulting disease variant is annotated as described above for mutations affecting the initiator methionine. The only such example in Reactome is the NOTCH1 variant NOTCH1 P2_M1580del (commonly known as TAN-1), which results from a translocation t(7;9) that fuses the promoter of T cell receptor beta to intron 24 of NOTCH1, resulting in the usage of an alternative start codon at position 1580 of the reference NOTCH1 sequence.



[FragmentDeletionModification:9830728] Deletion of residues 1 to 1579

FragmentDeletionModification Properties

Property Name	Value
endPositionInReferenceSequence	1579
referenceSequence	UniProt:P46531 NOTCH1
startPositionInReferenceSequence	1
DB_ID	9830728
_displayName	Deletion of residues 1 to 1579



[EntityWithAccessionedSequence:2220929] NOTCH1 P2_M1580del extracellula...

EntityWithAccessionedSequence Properties

Property Name	Value
cellType	
compartment	extracellular region
endCoordinate	1664
hasModifiedResidue	Deletion of residues 1 to 1579
referenceEntity	UniProt:P46531 NOTCH1
startCoordinate	1580
DB_ID	2220929
_displayName	NOTCH1 P2_M1580del extracellular fragment [e...

Mutations activating an upstream methionine initiation site.

A mutation that eliminates the initiator methionine and results in the aberrant initiation of translation from an upstream methionine is annotated as an insertion.

Naming

Named in accordance with insertion rules, above.

Format: “prefix” “amino_acids+positions_flanking” “ins” “inserted_sequence”

Where “prefix” is the HGNC gene name, the amino acids flanking the insertion are represented by single letter codes separated by underscore, and the inserted sequence is listed as single letter amino acid codes.

How to annotate

Annotate as for an insertion, with a FragmentInsertionModification, but preserve methionine 1 from the WT sequence and delete the novel initiator methionine to accommodate HGVS guidelines as described below:

As an example (made up, not annotated in Reactome)

WT coding sequence *MRST*MLRSQ… (WT initiating M in blue, upstream sequence in small italics)

Mutation in the initiator M changes it to V, and activates the methionine four amino acids upstream:

MRSTVLRSQ - this is what really happens, blue M converted to blue V, and inclusion in the coding sequence of upstream MRST.

By HGVS convention, this is actually annotated as:

MRSTVLRSQ - insertion of RSTV between the ‘original’ M and the second position amino acid (L).

Extension variants

The only currently annotated extension variant is IKBKG *420Wext*447, which represents a 3’ extension, with replacement of the stop codon at position 420 with tryptophan codon and extension of the protein sequence by a total of 27 amino acids, with a new stop codon at position 447, counting from the start codon of the reference wild-type protein sequence.

How to annotate

Both 3’ and 5’ extensions can be annotated using FragmentReplacedModification class. In the case of 3’ extensions, the wild-type stop codon coordinate is specified as the start and end coordinate of the replacement, and the added stretch of amino acids, in single letter abbreviations, is defined in the alteredAminoAcidFragment attribute. In the case of 5’ extensions, the -1 would be specified as the start and end coordinate of the replacement.

[FragmentReplacedModification:5262861] Replacement of residues 420 to 420 by WGRPVQGHCLPRTCPGPCSLRPFLPPA

FragmentReplacedModification Properties

Property Name	Value
alteredAminoAcidFragment	WGRPVQGHCLPRTCPGPCSLRPFLPPA
endPositionInReferenceSequence	420
referenceSequence	UniProt:Q9Y6K9 IKBKG
startPositionInReferenceSequence	420
DB_ID	5262861
_displayName	Replacement of residues 420 to 420 by WGRPVQGHCLPRTCPGPCSLRPFLPPA
created	Shamovsky, V, 2014-02-09

On the EWAS record, the stop codon is changed relative to the WT reference sequence to reflect the extension.

[EntityWithAccessionedSequence:5262978] IKBKG *420Wext*447 [cytosol]

EntityWithAccessionedSequence Properties

Property Name	Value
cellType	
compartment	cytosol
endCoordinate	446
hasModifiedResidue	Replacement of residues 420 to 420 by WGRPVQGHCLPRTCPGPCSLRPFLPPA
referenceEntity	UniProt:Q9Y6K9 IKBKG
startCoordinate	1
DB_ID	5262978
_displayName	IKBKG *420Wext*447 [cytosol]
authored	
created	Shamovsky, V, 2014-02-09
crossReference	
definition	
disease	primary immunodeficiency disease
edited	
figure	
goCellularComponent	
inferredFrom	
inferredTo	
literatureReference	X-linked anhidrotic ectodermal dysplasia with immunodeficiency is caused b...
modified	Shamovsky, Veronica, 2014-07-14
	Shamovsky, Veronica, 2014-12-08
	Orlic-Milacic, Marija, 2023-02-06
name	IKBKG *420Wext*447
	NEMO *420Wext*447
	IKKG X420WextX447
	IKK-gamma X420W
	Inhibitor of nuclear factor kappa-B kinase gamma subunit
	IKBKG *420Trpext*447
	IKBKG Ter420TrpextTer447

Splicing variants

Splicing variants, if the exact effect of the mutation has been determined experimentally at the protein level, are annotated in one of the previously described categories, depending on the consequence of the splicing mutation on the amino acid sequence, and named accordingly, as HGVS currently does not provide guidelines for nomenclature of splicing variants at the protein level. This also applies to inversion variants. Of the 6 splicing variants currently annotated in Reactome, four are annotated as frameshift variants (EXT2 A361Pfs*44, NF1 Q616Gfs*4, NF1 R69Nfs*7, and UNC93B1 L230Afs*188), one is annotated as an in-frame insertion (POMT2 I444_N445insLLWQ), and one as an in-frame deletion (SLC11A2 L104_P142del). Two inversion variants currently annotated in Reactome, F8 A20_M2143delinsMRIQDPGK and F8 V2144_Y2351delinsHGVLENGKNWEPSYRW, are annotated as in-frame indels, as their names imply.

Variants with multiple in-cis mutations

These usually contain a primary mutation and a secondary mutation that was selected for during the course of treatment with target therapy as a drug-resistance mechanism. Each mutation is annotated independently and they are separated with a semicolon in the disease variant name, e.g. EGFR E746_A750del;T790M, where T790M is a mutation associated with resistance to small tyrosine kinase inhibitors.

Guideline for automated annotation of disease variants

Automated annotation of disease variants in Reactome proceeds through the following steps:

1) Starting from manually annotated disease variants grouped into functional sets, define external database-derived and/or literature-derived variants with matching disease association that conform to Reactome-expanded ACMG criterion PM5 or to combined ACMG criteria PM1 and PM4.

2) Categorize variants of interest by mutation type (missense, nonsense, etc.).

3) Run the script that creates individual disease variant records for automatic upload to the Reactome Curator Tool.

4) Import automatically annotated disease variants into existing disease variant functional sets.

5) Expanded disease variant functional sets are readily displayed in interactive pathway diagrams.

References

1. Philippe C, Porter DE, Emerton ME, Wells DE, Simpson AH, Monaco AP. Mutation screening of the EXT1 and EXT2 genes in patients with hereditary multiple exostoses. Am J Hum Genet. 1997 Sep;61(3):520-8. doi: 10.1086/515505. PMID: 9326317; PMCID: PMC1715939.
2. Ahn J, Lüdecke HJ, Lindow S, Horton WA, Lee B, Wagner MJ, Horsthemke B, Wells DE. Cloning of the putative tumour suppressor gene for hereditary multiple exostoses (EXT1). Nat Genet. 1995 Oct;11(2):137-43. doi: 10.1038/ng1095-137. PMID: 7550340.

Adding cross-references to disease entities

We cross reference disease entities to COSMIC (cancer-associated mutations; https://cancer.sanger.ac.uk/cosmic) and/or ClinGen (https://reg.clinicalgenome.org/redmine/projects/registry/genboree_registry/landing) where appropriate. An entity can have as many cross-references as needed. COSMIC cross-referencing should make use of the Genomic Mutation ID where available, not the Legacy Identifier. ClinGen Canonical Allele (CA) identifiers are used wherever possible, but ClinVar IDs may be used if CA IDs are not available.

Final steps

Once your curation is done, you will need to add your pathway into the Disease hierarchy in the relevant subpathway. Check out the disease chapter, add your pathway to the proper spot in the subpathways (or create a new child of Disease pathway, if appropriate) and commit the Disease chapter back into gk_central.

You will also have to check if any parent pathway EHLDs or the Disease parent EHLD requires updating to accommodate your new disease pathway. Coordinate with Cris to

incorporate your updates to any EHLDs. Update any affected pathway summations to incorporate your new pathway.

Infectious Disease and Innate Immune System Pathway Curation Guide

The Reactome's concept of host-pathogen interactions refers to the relationship between a host organism and a pathogen (which can be any microorganism) at the molecular level. These interactions involve the recognition of the pathogen by the host's immune system, the activation of immune responses, and the evasion strategies employed by pathogens to avoid detection and destruction by the host's immune system.

We have set the disease / normal boundary by asking what the outcome of the host-pathogen interaction is. If the host wins and the pathogen is really neutralized, that counts as normal. If the pathogen wins to cause overt infection or to establish a stable latent state that could result in disease later, that counts as disease.

Entry and growth of pathogenic microorganisms within the host cell (e.g., HIV life cycle) also counts as disease and is annotated as a part of the "Infectious Diseases" module (see above).

In most cases, interactions involving host and pathogen entities are captured as a binding between the host PE (e.g., EWAS or a complex) and a pathogen-associated molecular pattern (PAMP), which could be a wide range of molecules found in pathogens. Examples of PAMPs include bacterial lipopolysaccharides, viral nucleic acids, and fungal cell wall components.

Annotating physical entities

1. Genomic, transcript, and protein sequences, including those of microbial origin, are annotated as EWASs which have their referenceEntity slot filled in with referenceSequence instance containing NCBI nucleotide, ENSEMBL, EMBL or UniProt identifier.
    - Note: Microbial nucleic acids can be annotated as direct children of GenomeEncodedEntitiy (GEE) class with microbial species assigned.
2. A simpleEntity should be used to annotate molecules such as bacterial lipopolysaccharides,peptidoglycans or fungal mannans. A referenceEntity is filled with a referenceMolecule instance containing a ChEBI identifier.

Assigning compartment for entities of microbial origin

We assign location on Reactome instances using GO compartment terms. The GO terms allow to distinguish Pathogen associated molecular patterns (PAMP) molecule locations in the pathogen compartments (e.g., "cell wall" or "virion membrane") and locations in infected host cells (e.g.,"host cell endoplasmic reticulum membrane", "host cell cytosol"). The GO terms distinguish these locations of PAMPs from locations of

endogenous host proteins in a normal cell (“endoplasmic reticulum membrane”, "cytosol" etc.).

In contrast to GO, we use normal cell terms (for example, “endoplasmic reticulum membrane” and **not** “host cell endoplasmic reticulum membrane”) to localize virus- or bacteria- or other pathogen-derived entities in an infected cell.

The GO terms that are used to localize proteins in a normal cell are listed in the Compartment class. Thus,

1. The compartment slot must be filled with a GO term from the Compartment list which includes those involving pathogen/symbiont patterns.

2. When the species slot value for a PE is a pathogen species and the compartment slot value is a normal cell term, then the **goCellularComponent** slot of that PE should also be filled with the GO "host cell [compartment]" term corresponding to the one chosen for the entity.

3. The Compartment list also includes GO terms that describe parts of bacteria or viruses ("cell wall", "viral capsid"). These terms can be also used to assign compartments for pathogen-derived entities.
    - Note: Individual pathogen proteins/entities localized in pathogen compartments (“viral capsid”, “virion membrane”) may form a complex, which functions in the host compartment. In this case, the record of the complex with microbial components should have a normal cell compartment term (e.g., “extracellular region”).

- The goCellularComponent slots in PE records of these pathogen-derived entities can be filled in with a GO “host cell

[Complex:9694500] S3:M:E:encapsidated SARS-CoV-2 genomic RNA: 7a:O-glycosyl 3a tetramer [extracellular region]

Complex Properties

Property Name	Value
cellType	
hasComponent	1 x 7a:O-glycosyl 3a tetramer [virion membrane] 1 x S3:M lattice:E protein [virion membrane] 1 x encapsidated SARS-CoV-2 genomic RNA [viral capsid]
DB_ID	9694500
_displayName	S3:M:E:encapsidated SARS-CoV-2 genomic RNA: 7a:O-glycosyl 3a tetramer [extracellular region]
authored	
compartment	extracellular region
created	Cook, Justin, 2020-07-07
crossReference	
definition	
disease	COVID-19
edited	
entityOnOtherCell	
figure	
goCellularComponent	
includedLocation	viral capsid virion membrane
inferredFrom	S3:M:E:encapsidated SARS coronavirus genomic RNA: 7a:O-glycosyl 3a tetramer [extracellular region]
inferredTo	
isChimeric	false
literatureReference	
modified	Orlic-Milacic, Marija, 2020-07-28 Gillespie, Marc E, 2020-08-14 Rothfels, Karen, 2020-08-28 Rothfels, Karen, 2021-09-01
name	S3:M:E:encapsidated SARS-CoV-2 genomic RNA: 7a:O-glycosyl 3a tetramer
relatedSpecies	
reviewed	
revised	
species	Severe acute respiratory syndrome coronavirus 2
stableIdentifier	R-COV-9694500.2
stoichiometryKnown	

compartment” term to depict their roles as invaders of a host cell.

- Note: A complex involving a pathogen protein/entity and host PE(s) should have a host compartment assigned, while microbial compartments assigned to individual components will be automatically added in the includedLocation slot.

ASSIGNING SPECIES IN PES, PATHWAYS AND REACTIONS THAT INVOLVE HOST-PATHOGEN INTERACTIONS

1. The “species” slot on the event identifies the environment where the reaction is happening. This should almost always be human for human Reactome curation. The exceptions are when we have annotated reactions (RLEs) in other species for inference (e.g Mus) or when we are specifically annotating another species pathways as part of another project (e.g Gallus and Mycobacterium tuberculosis).

In disease or immune system processes that involve host-pathogen interactions the host would be listed as the “species” in the records of RLE and complexes, which are formed by host and pathogen entities.

- Note: When creating a chimeric reaction to be used for inference (i.e one that involves entities from multiple species and (usually) observed in vitro) ALL species should be listed in the species slot.
- Note: A complex involving only simpleEntities should have NO species listed. A complex involving a pathogen protein/entity and simpleEntity or an OtherEntity (i.e not having a species) should have the pathogen species listed in the species attribute and no relatedSpecies.

2. The “related species” is used to tag the source (or sources) of genome-encoded entities from “elsewhere” that act as “bystanders” in a process. In disease or immune system processes, the pathogen(s) would be listed as “relatedSpecies" in RLEs and complex instances.

- Note: relatedSpecies value(s) specified in RLEs should be assigned to a first parent pathway.

3. The use of "species" and "relatedSpecies" in sets follows the same rule as for complexes and events.

- Note: When the set members contain participants from different non-host pathogen species, ALL microbial species should be listed in the species slot.
- Note: If your annotations involve a species that is new to Reactome, please contact Peter D'Eustachio as it will need to have an abbreviation assigned by him.

Assigning a disease term for entities and reactions that involve host-pathogen interactions

1. For all kinds of pathogen entities in human (or other host) pathways the disease slots are filled in with a value corresponding to the disease they would cause if not neutralized. The effect of this rule is to automatically color all pathogenic entities red, both when they are involved in a disease and also when they are in the process of being detected and neutralized by the innate immune system.

2. Assigning a disease value for complexes and reactions depends on the outcome of the host-pathogen interaction.

2a. In the disease-preventing innate immune reactions (e.g., E. coli flagellin binds to human TLR5 (R-HSA-188025) as part of the human TLR cascade) the disease slot is not filled thus leaving a reaction with its default value (black, normal). Complexes formed between pathogen entities and the host ones that detect and neutralize them

have empty disease slot, so in a normal reaction a red (disease) pathogen gets incorporated into a default (black, normal) pathogen : human receptor complex.

2b. In a disease-causing event (e.g., uptake of diphtheria toxin, Corynephage beta, which enters and destroys the human cell), a pathogen-derived entity (a disease slot is filled in) interacts with a normal host cell protein (default) to form an abnormal complex (a disease slot is filled in). A reaction-like event for an infectious disease process has a non-null disease slot value thus coloring its edges and nodes red. The entities will be automatically colored red if they are from a pathogen or complexes of human and pathogen proteins. Normal human entities that participate in these reactions are default (black).

2c. Disease-causing events are placed under disease, either infectious disease or diseases of a corresponding normal pathway. Disease host-pathogen interaction pathways have their own diagrams.

Examples of host-pathogen interactions:

1. C-ter TLR7 dimer binds TLR7, TLR8 ligands (reaction) (R-HSA-167983)

Here a defined set of viral single stranded RNAs binds a human TLR7, setting off events that neutralize viruses and leave the human healthy – in our human-centric view, a normal process.

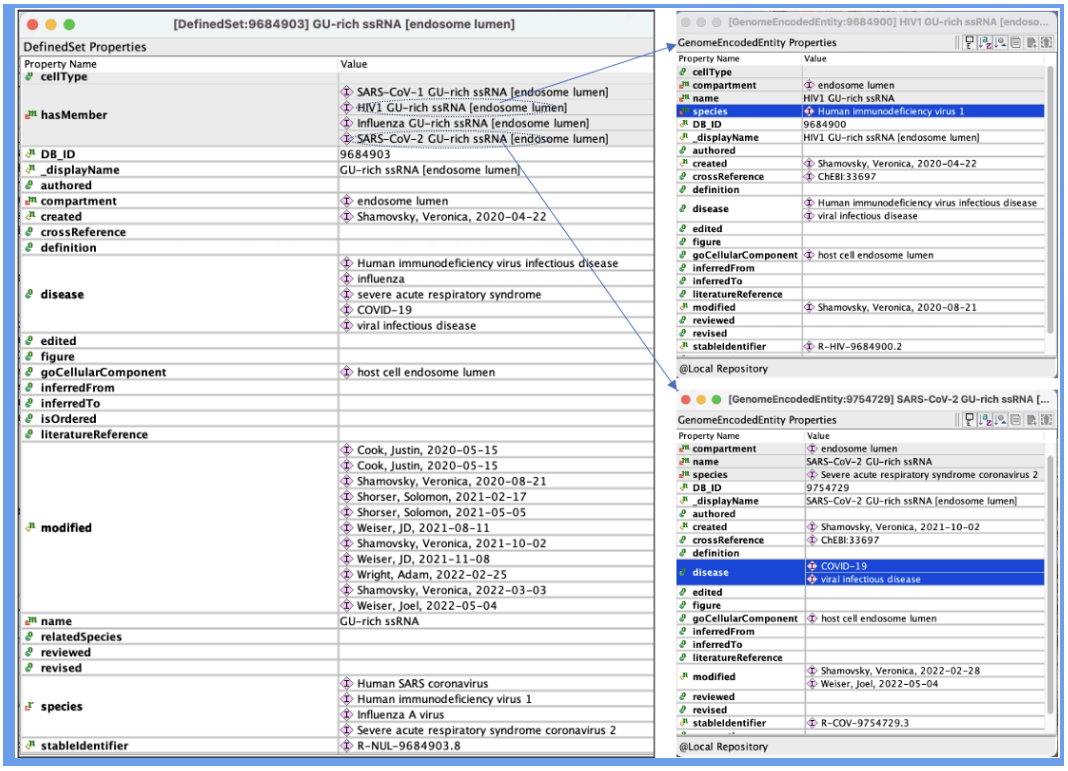
The reaction record lists all viruses in the ‘relatedSpecies’ slot, while human is indicated as the host in the species slot. The disease value is NULL.

Reaction Properties

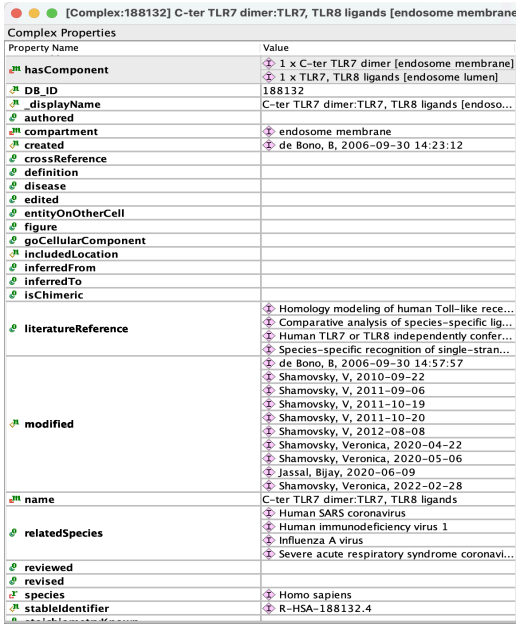
Property Name	Value
catalystActivity	
deletedInstanceDB_ID	
entityFunctionalStatus	
input	1 x C-ter TLR7 dimer [endosome membrane] 1 x TLR7, TLR8 ligands [endosome lumen]
output	1 x C-ter TLR7 dimer:TLR7, TLR8 ligands [endos...
DB_ID	167983
_displayName	C-ter TLR7 dimer binds TLR7, TLR8 ligands
_doRelease	true
authored	Luo, F, 2005-11-10 11:23:18
catalystActivityReference	
compartment	endosome membrane endosome lumen
created	Gillespie, ME, 2005-11-10 16:24:09
crossReference	
definition	
deletedStableIdentifier	
disease	
edited	Shamovsky, V, 2011-08-12
entityOnOtherCell	
evidenceType	

name	C-ter TLR7 dimer binds TLR7, TLR8 ligands
negativePrecedingEvent	
normalReaction	
orthologousEvent	
precedingEvent	
previousReviewStatus	
reactionType	
regulatedBy	Negative regulation by 'C-ter TLR7 dimer:HCQ [e...
regulationReference	Negative regulation by 'C-ter TLR7 dimer:HCQ [e...
relatedSpecies	Human SARS coronavirus Human immunodeficiency virus 1 Influenza A virus Severe acute respiratory syndrome coronavirus 2
releaseDate	2010-12-14
releaseStatus	
requiredInputComponent	
reverseReaction	
reviewStatus	five stars
reviewed	Gay, NJ, 2006-04-24 16:48:17
revised	
species	Homo sapiens
stableIdentifier	R-HSA-167983.5
structureModified	

In this reaction, a viral ssRNA is annotated as GenomeEncodedEntity (GEE) for each virus. Viral EWASs are combined in the defined set, listing all microbial species in the “spesies” slot, and all disease values in the disease slot. The compartment slot is filled in with “endosome lumen”, while goCellularComponent indicates that this PE is located in “host cell endosome lumen”.



The resulting complex of the human TLR7 homodimer and viral ssRNA has values similar to the reaction record:



[Complex:188132] C-ter TLR7 dimer:TLR7, TLR8 ligands [endosome membrane

Complex Properties

Property Name	Value
hasComponent	1 x C-ter TLR7 dimer [endosome membrane] 1 x TLR7, TLR8 ligands [endosome lumen]
DB_ID	188132
_displayName	C-ter TLR7 dimer:TLR7, TLR8 ligands [endoso...
authored	
compartment	endosome membrane
created	de Bono, B, 2006-09-30 14:23:12
crossReference	
definition	
disease	
edited	
entityOnOtherCell	
figure	
goCellularComponent	
includedLocation	
inferredFrom	
inferredTo	
isChimeric	
literatureReference	Homology modeling of human Toll-like rece... Comparative analysis of species-specific lig... Human TLR7 or TLR8 independently confer... Species-specific recognition of single-stran...
modified	de Bono, B, 2006-09-30 14:57:57 Shamovsky, V, 2010-09-22 Shamovsky, V, 2011-09-06 Shamovsky, V, 2011-10-19 Shamovsky, V, 2011-10-20 Shamovsky, V, 2012-08-08 Shamovsky, Veronica, 2020-04-22 Shamovsky, Veronica, 2020-05-06 Jassal, Bijay, 2020-06-09 Shamovsky, Veronica, 2022-02-28
name	C-ter TLR7 dimer:TLR7, TLR8 ligands
relatedSpecies	Human SARS coronavirus Human immunodeficiency virus 1 Influenza A virus Severe acute respiratory syndrome coronavi...
reviewed	
revised	
species	Homo sapiens
stableIdentifier	R-HSA-188132.4

2. Translocation of Vpr to the mitochondria (Reaction: R-HSA-180855)

Here an HIV protein is translocated to the mitochondria and while no human proteins are involved, the **species** listed for this reaction should still be human because it is a RLE occurring in the context of a human process.

The **relatedSpecies** should be Human immunodeficiency virus 1.

The **disease slot** is filled in with ‘human immunodeficiency virus infectious disease’:

3. Uptake and function of diphtheria toxin (Pathway: R-HSA-5336415)

As in case 1, the process is occurring in a human, so the pathway “species” is human involving a diphtheria protein (Corynephage beta). In this case, the result is that the bacterial protein enters and destroys the human cell – in our human-centric view, a disease process. The pathway records are filled in as follows:

species: Homo sapiens **relatedSpecies:** Corynephage beta **disease:** diphtheria

4. Host physical entity is transferred into a pathogen cell.

Consistent with the logic in case 1, 2 and 3, the following reactions are part of the human process, so it is reasonable to label them as human reactions, even in cases in which reactions are occurring inside the bacterium.

a. Cathelicidin LL-37 localizes to bacterial cell wall (Reaction: R-HSA-1222685)

species: Homo sapiens **relatedSpecies**: Mycobacterium tuberculosis **disease:** NULL

b. Nitric oxide enters the bacterium (Reaction: R-HSA-1222662)

species: Homo sapiens **relatedSpecies**: Mycobacterium tuberculosis **disease:** NULL

- Note: When a reaction involves only SimpleEntities (e.g., case 4b), RLE instance goes on skip list because QA check will detect a conflict because there are no human entities specified as reaction components.

5. Cry proteins stabilize unphosphorylated Bmal1:Clock,Npas2 (Reaction: R-MMU-549451)

   Here, the reaction describes a process occurring in a non-human species (Mus) and *only* mouse proteins are involved.

   species: Mus musculus **relatedSpecies**: NULL **disease:** NULL

   The environment is mouse, the proteins are mouse, and success in the event keeps the mouse healthy. So this is a normal reaction whose species is mouse and whose related species is NULL.

[Additional information about these attributes and examples of how the species and relatedSpecies attributes are used in disease and normal curation can be found here.]

Final steps

Once your curation is done, you will need to add your pathway into the Disease hierarchy in the relevant subpathway. Check out the disease chapter, add your pathway to the proper spot in the subpathways (or create a new child of Disease pathway, if appropriate) and commit the Disease chapter back into gk_central.

You will also have to check if any parent pathway EHLDs or the Disease parent EHLD requires updating to accommodate your new disease pathway. Coordinate with Cris to incorporate your updates to any EHLDs. Update any affected pathway summations to incorporate your new pathway.

Drug Curation Guide

General points

When annotating drugs used to treat a particular disease, curators should add an appropriate disease tag to a drug entity. For a cancer drug, the general cancer disease attribute should definitely be added, but when it comes to more specific disease tags, it

may be advisable to add only those cancer types for which the treatment by a given drug is approved.

In ***most*** cases, a drug will affect a target protein after it binds to it. Therefore, it will usually be the case that two steps are needed for any drug curated; a drug binding event to the target then whatever effect (in the form of a regulation) the drug has on the target (usually at an existing event). Sometimes (rarely), drugs will behave in other ways, for example, as group donors. The appropriate reaction must then be created.

The drug interaction file

The drug interaction file used in curation is a set of autogenerated drug reactions from GtP in an .rtpj. Bijay and Guanming devised a solution to create auto-generated drug reactions from GtP data. In a nutshell, the criteria used to create the file was to collect any approved drugs that bind the same human target, which is the primary target of those drugs, and create a Reactome-type reaction that has input of drug(s) and target and output of the drug(s): target complex. Also, collect all references from each drug page that describe them binding to this target into the Reactome reaction litRef slot.

See this file below.

Latest drug:target auto-generated GtP reaction file **IUPHAR_reactions_051821.rtpj** (in Reactome Team Drive > Curation > Drug curation).

Note most of these reactions are NOT checked into Reactome. Instead, identify the reaction you wish to curate and check that in ONLY from this file. Once in gk_central, check out into your local project and work on revising it.

This file provides a good starting point to revise any chosen reaction from this list you want to curate into Reactome. You'd still need to investigate whether these drugs are still in use or are experimental, for instance. As long as strong evidence exists to confirm pharmacological action, that drug can be curated.

Where you've found a drug in the literature that doesn't have a GtP ID, you can request new drug entries from GtP as long as you can provide evidence for the pharmacological effect of that drug. By consensus, Reactome primarily annotates drugs that have been approved for human use. If an exception is made in cases where a drug is biologically interesting / distinctive but not universally approved, it is crucial to explain in the summation why we have nevertheless chosen to annotate it.

Example - drug binds target, resultant complex regulates that target's event(s)

The drug **rivaroxaban** (Xarelto) functions as an anticoagulant that binds to and inhibits Factor Xa in the blood clotting cascade. This has the effect of slowing down the rate of blood coagulation, helping to prevent the formation of blood clots.

Find the pathway of the target protein where this drug acts and download the ELV-level pathway into your tool. The relevant pathway here is Formation of Fibrin Clot (Clotting Cascade R-HSA-140877)

You need to determine what type of drug this is, so you can create the relevant drug instance. Reading some literature and/or checking its identifiers (Chebi, GtP) will determine this.

There are two ways to create the drug reaction; either use the auto-generated reaction (recommended, Steps 1a, 2a) or create from scratch (Steps 1b, 2b).

STEP 1A

Find the reaction from the auto-generated drug reaction file (above) which should have the entity and target information already filled. Check it into gk_central then check out into your local project.

STEP 1B

Below is the process to create new drug instances (steps 1b and 2b) if there is no auto-generated reaction.

This example is a chemical drug so create an instance of chemicalDrug in the curator tool. The instance has mandatory slots

name - enter the common name and a brand name (optional)
rivaroxaban, Xarelto

compartment - same as the target protein in most cases (exception for example target is a membrane protein, drug will be extracellular)
extracellular region

disease - A DOID identifier that most closely matches the disease this drug is involved in
pulmonary embolism (DOID:9477)

referenceEntity - a referenceTherapeutic instance containing an GtP identifier. Any drug you create **must** have an GtP identifier so check their website first for one before curating
GtP:6388

(On the rare occasion that an GtP ID is not available for your drug, you can use other resources like PubChem Compound, DrugBank or Chebi as alternative identifiers. Remember to fill in the 'approved' slot in the referenceEntity instance when using these alternatives as this slot is mandatory. If you can't find a suitable identifier, see instructions below describing how to request a Chebi identifier.).

crossReference - currently optional but desirable - Add a Chebi identifier. If one is not available, easy to request one (see Best Practices- Using Chemical Compound Databases, Appendix C8).
Chebi:68579

literatureReference - currently optional but desirable - Indicate the specific references for the drug instance. This is useful for data dumps as well as users wanting specific references for a drug, rather than trying to match a specific reference from the summation for a set of drugs

reactionType - currently optional but desirable. Choose an item from this list (taken from the IUPHAR / Guide to Pharmacology glossary, https://www.guidetopharmacology.org/helpPage.jsp#glossary) to classify the effect of the drug on the reaction:

Activator - "A ligand that activates an ion channel by binding directly to it, including ligands that cause a change in the voltage dependence of channel gating that favours activation."

Agonist - "A ligand that binds to a receptor and alters the receptor state resulting in a biological response. Conventional agonists stabilise an active conformation of a receptor and increase receptor activity. When the receptor stimulus induced by an agonist reaches the maximal response capability of the system (tissue), then it will produce the system maximal response and be a full agonist in that system."
Allosteric modulator (is_a modulator) - "A ligand that increases (positive) or decreases (negative) the action of an agonist or antagonist by combining with a distinct site on the receptor macromolecule. Neutral allosteric ligands bind to an allosteric site without affecting the binding or function of orthosteric ligands but can still block the action of other allosteric modulators that act via the same allosteric site."
Antagonist - "A drug that reduces the action of another drug, generally an agonist"

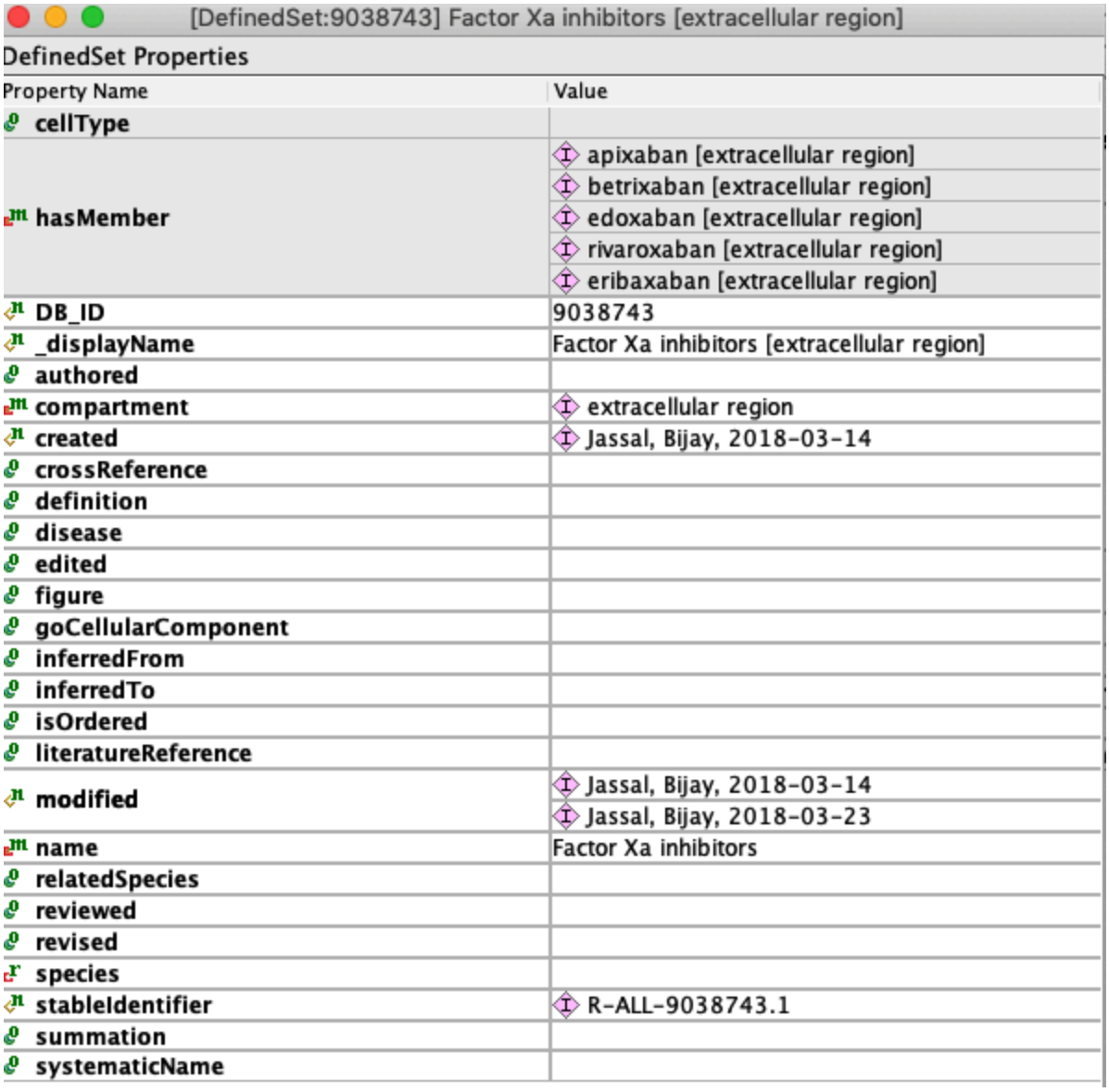
Antibody binding - A drug that is an antibody (immunoglobulin) and acts by binding the epitope for which it is specific.*
Channel blocker - “A ligand that blocks ion movement through an ion channel.”
Channel opener - A ligand that facilitates ion movement through an ion channel.*
Inhibitor - “A ligand that inhibits ion channel gating by directly binding to the channels, including ligands that inhibit ion channel activation and those that inhibit ion channel inactivation.”
Modulator - A ligand that changes the activity of its target protein on binding
Positive modulator (is_a modulator) - A ligand that increases the activity of its target protein on binding*.
Allosteric antagonist - (is_a antagonist) A ligand that decreases the action of an agonist or antagonist by combining with a distinct site on the receptor macromolecule*.
Antisense inhibitor - (is_a inhibitor) An RNA oligonucleotide ligand that base-pairs with an endogenous mRNA, to form a rapidly degraded complex.*
Gating inhibitor - (is_a inhibitor) ???
Inverse agonist - (is_a agonist) A ligand that by binding to unoccupied receptors reduces the fraction of them in an active conformation. For nuclear receptors, an inverse agonist refers to ligands that can promote co-repressor recruitment. A drug that binds to the same receptor as an agonist but induces a pharmacological response opposite to that of the agonist.
Negative allosteric modulator - (is_a allosteric modulator) A ligand that decreases the action of an agonist or antagonist by combining with a distinct site on the receptor macromolecule.*
Partial agonist - (is_a agonist) Agonists that in a given tissue, under specific conditions, cannot elicit as large an effect (even when applied at high concentration, so that all the receptors should be occupied) as can another agonist acting through the same receptors in the same tissue.
Positive allosteric modulator - (is_a allosteric modulator) A ligand that increases the action of an agonist or antagonist by combining with a distinct site on the receptor macromolecule.*

rivaroxaban chemicalDrug entry

ChemicalDrug Properties

Property Name	Value
cellType	
compartment	extracellular region
disease	pulmonary embolism
referenceEntity	rivaroxaban [IUPHAR:6388]
DB_ID	9015055
_displayName	rivaroxaban [extracellular region]
authored	
created	Jassal, Bijay, 2017-08-08
crossReference	ChEBI:68579
definition	
edited	
figure	
goCellularComponent	
inferredFrom	
inferredTo	
literatureReference	Discovery of the novel antithrombotic ...
	Rivaroxaban: a novel, oral, direct fact...
	The discovery and development of riv...
	Pharmacology of anticoagulants used i...
modified	Jassal, Bijay, 2017-08-09
	Jassal, Bijay, 2017-08-15
	Weiser, JD
	Jassal, Bijay, 2018-03-14
	Weiser, JD
	Shorser, Solomon, 2018-08-13
	Wu, G, 2018-09-19
	Jassal, Bijay, 2018-10-18
	Jassal, Bijay, 2018-10-18
	Wu, G, 2018-11-05
	Wu, G, 2018-12-12
name	rivaroxaban
	Xarelto
reviewed	
revised	
stableIdentifier	R-ALL-9015055.1
summation	
systematicName	

Rivaroxaban is one of a number of ‘xabans’ that are direct oral anticoagulants (DOACs) and all act on Factor Xa in an inhibitory way. Since they all act in the same way, a set of these drugs can be constructed.



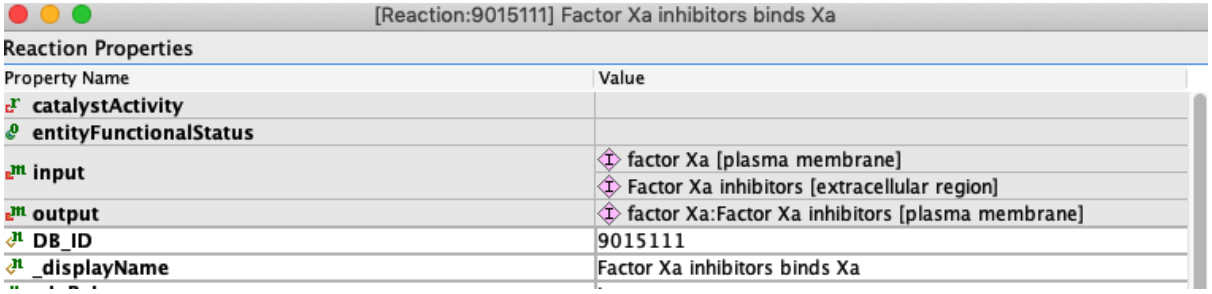
STEP 2A

In the autogenerated reaction, check the ‘functional form’ of the target. In pharmacological databases, the target is described as a single protein. This is what will be populated in the autogenerated reaction. However, in Reactome, the target protein may be in a complex, a cleaved form or PTM’d, etc etc. You will need to replace the target protein in the autogenerated reaction IF IT IS DIFFERENT to what is curated in Reactome. Check the compartments of all entities at this stage.

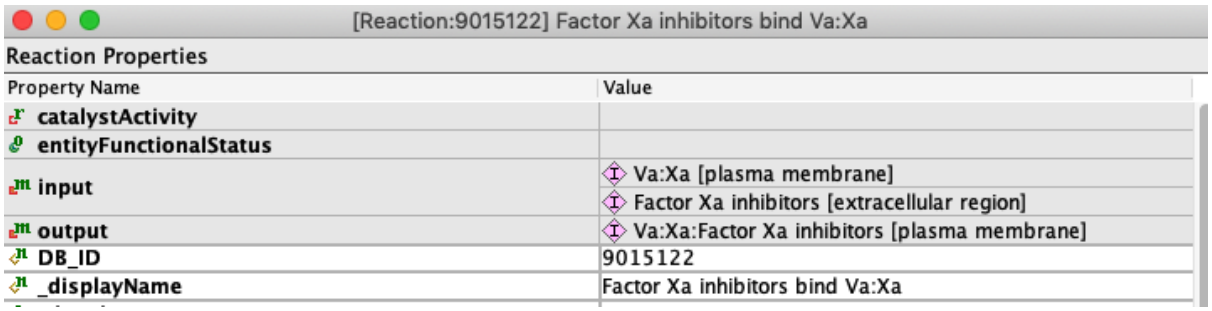
STEP 2B

If creating a new reaction from scratch, find the 'functional form' of the target in the relevant Reactome reaction to be used in the drug reaction.

Factor Xa inhibitors ("xabans") binds Factor Xa. The resulting complex will act on Factor Xa event. Fill in the disease slot (same as for the drug).



Factor Xa inhibitors also bind complexed Factor Xa (Va:Xa). The resulting complex will act on the Va:Xa event. Fill in the disease slot (same as for the drug).



STEP 3

(will most likely be existing event/s or if not, you'll need to create one to be the regulated event).

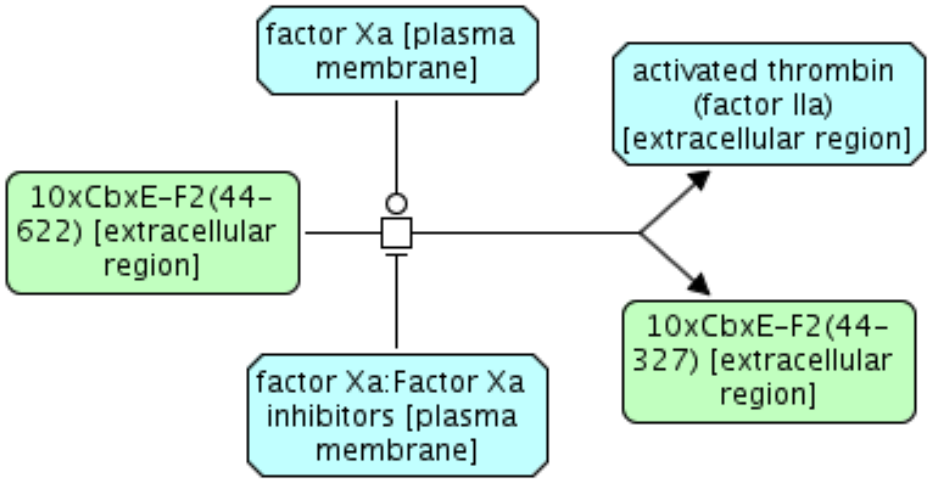
Find the event that Factor Xa is involved in and create a negative regulation of that event with the complex "Factor Xa:Factor Xa inhibitors".

Create the negative regulation instance as below

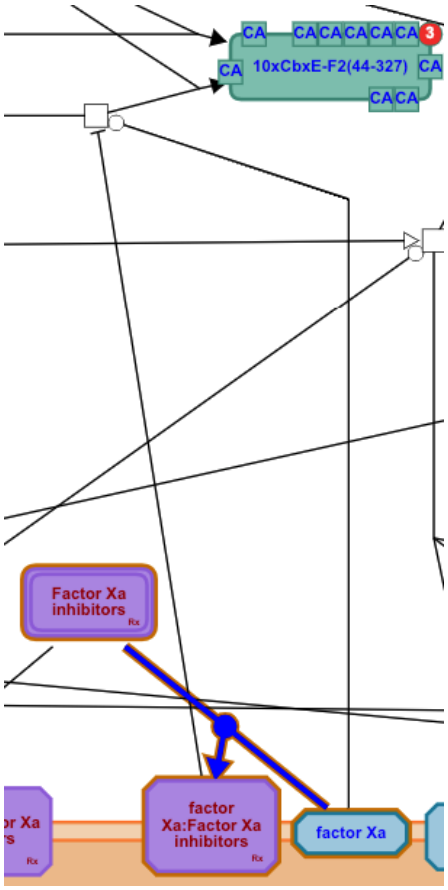
NegativeRegulation Properties

Property Name	Value
regulator	factor Xa:Factor Xa inhibitors [plasma membrane]
DB_ID	9015052
_displayName	Negative regulation by 'factor Xa:Factor Xa inhibitors [plasma membrane]'
activeUnit	
activity	
created	Jassal, Bijay, 2017-08-08
goBiologicalProcess	
literatureReference	
modified	Jassal, Bijay, 2017-08-08 Jassal, Bijay, 2017-08-08 Jassal, Bijay, 2017-08-08 Jassal, Bijay, 2017-10-25 Jassal, Bijay, 2018-03-02 Jassal, Bijay, 2018-03-14 Wu, G, 2018-05-08
stableIdentifier	R-HSA-9015052.3
summation	Factor Xa is a target for direct oral anticoagulant (DOAC) d...

...as seen in the curator tool central panel



...and as seen in the web display. Any entity containing a drug will be coloured purple.



Note - the actual regulation ‘line’ will be color-sensitive in the future* (ie purple for drugs).

*[This feature is currently on hold]

Add appropriate supporting references to both events.

Curator Tool Protocol

To support the new gk_central schema for drug annotation, a new feature has been added to the tool: auto-fetch the required information for the new ReferenceTherapeutic class based on the drug identifier in GtP (https://www.guidetopharmacology.org/). Here is an example of how to use it:

1. On the GtP web page, http://www.guidetopharmacology.org, search for imatinib in their search box.
2. On the result page, imatinib is the first entry. Click it.
3. On the imatinib page, find Ligand ID for it: 5687.
4. In the curator tool, create a new ReferenceTherapeutic instance, and enter 5687 as the identifier.
    - The curator tool will ask you if you want to fetch information from GtP.

- Choose “Yes”. Information for 5687 will be fetched automatically.

New molecule request in ChEBI

To request new identifiers for chemicals (and drugs) from ChEBI, follow this submission process. To submit to Guide to Pharmacology, you need to contact one of their curators to ask if the drug you want curated in GtP has an appropriate pharmacological action they are willing to add. Literature references always help your cause. The person to contact is Elena Faccenda, a senior curator at GtP.

Checklist

Here is a checklist for the whole drug curation process:

- Drugs in Guide to Pharmacology (GtP, formerly GtP) must be target-specific and approved (can be special cases eg. cancer therapeutics). See the GtP file here (Appendix C11).
- Check if the drug already exists in gk_central (NOT Reactome’s public website)
- Build projects (multiple at once if necessary) for all diagrams to be changed
- Create Drug entities, if necessary for different compartments
- Add ChEBI cross-references
- Drug entities have their disease attribute filled in
- Literature goes in the (binding) reaction
- Set the reaction’s doRelease=false
- In the reaction summary, mention all participating drugs
- In the ELV, the drug reactions need to be type changed to “Association”
- Only after review can the Regulation/Reference be added to a reviewed physiological reaction
- Fill in the tracking sheet, one item for every new drug reaction

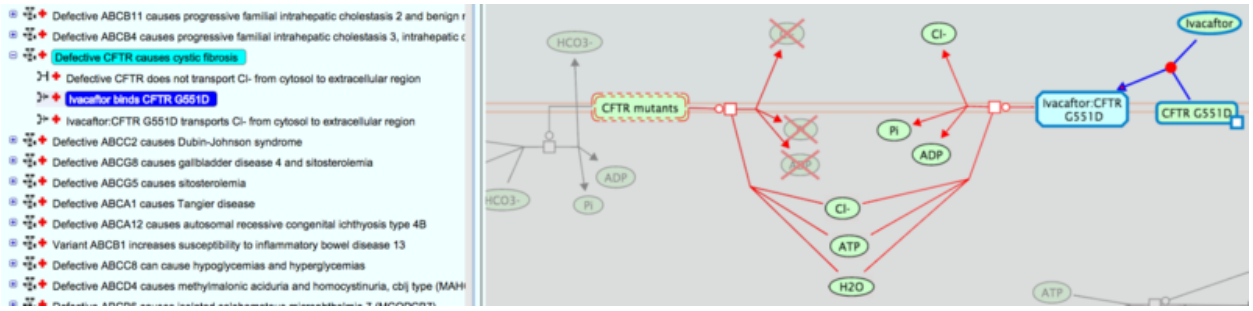
Complexities of Disease and Drug curation: Special cases

Example 1 - drug-target interaction in disease:

The cystic fibrosis transmembrane conductance regulator (CFTR) is a low conductance chloride-selective channel that mediates the transport of chloride ions in human airway epithelial cells. Chloride ions plays a key role in maintaining homeostasis of epithelial secretions in the lungs. Defects in CFTR can cause cystic fibrosis (CF), resulting in an ionic imbalance that impairs clearance of secretions, not only in the lung, but also in the pancreas, gastrointestinal tract and liver. More than 1500 mutations in the CFTR gene

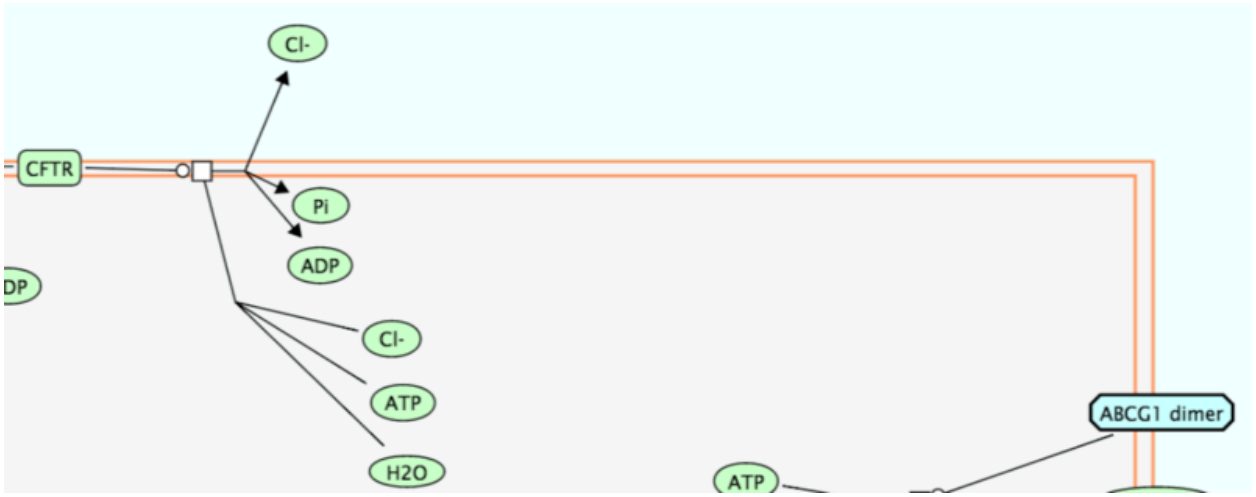
have been identified. CF patients with a particular mutation, G551D, have shown lung function improvements when given the drug Ivacaftor.

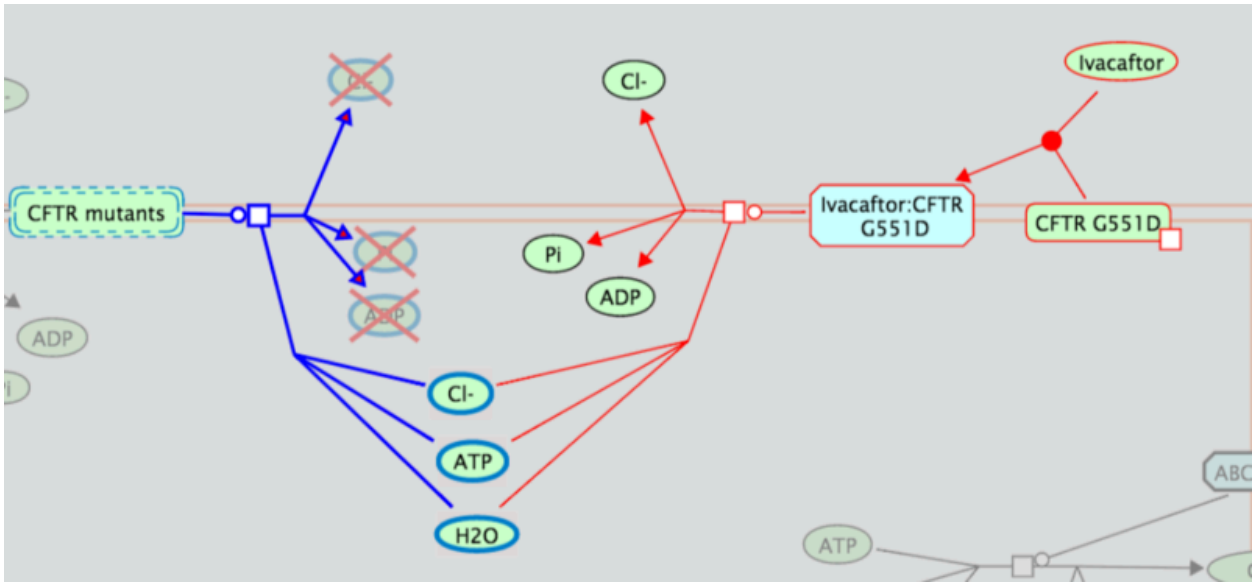
To curate the mutation reversal effect of the drug, create two events; first the drug binding to mutant protein and second, this complex reviving the function equivalent to the w/t protein.



Reactions showing ivacaftor binding the mutant protein (highlighted blue) and the following reaction (centre) showing channel functionality restored. The left-side reaction shows defective CFTR which cannot transport Cl- ions.

The construct in the hierarchy is complicated by having the loss-of-function (LoF) reaction (the left side) as part of the display. LoF reactions are generated automatically on the website. To add the reactions involving drugs, they must be drawn manually on the diagram shared by the disease pathway and the normal pathway, in this case, the PathwayDiagram instance lists the normal pathway (ABC-family proteins mediated transport) and the disease pathway (Defective CFTR causes cystic fibrosis), together with other disease pathways relevant to this diagram as representedPathway. The two disease drug reactions are drawn in this diagram and their edges coloured red. The cystic fibrosis disease attribute is associated with both.



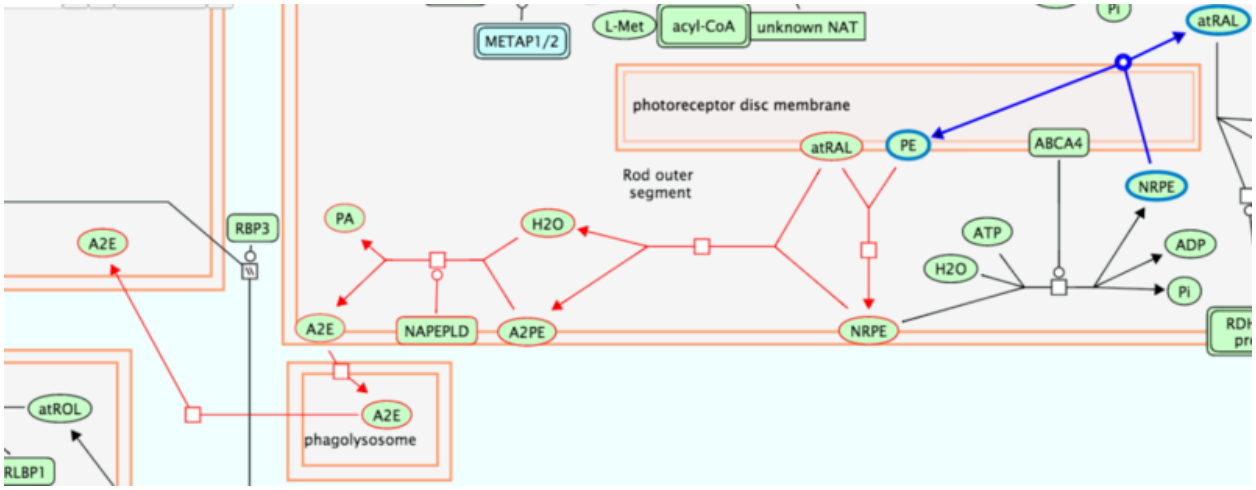


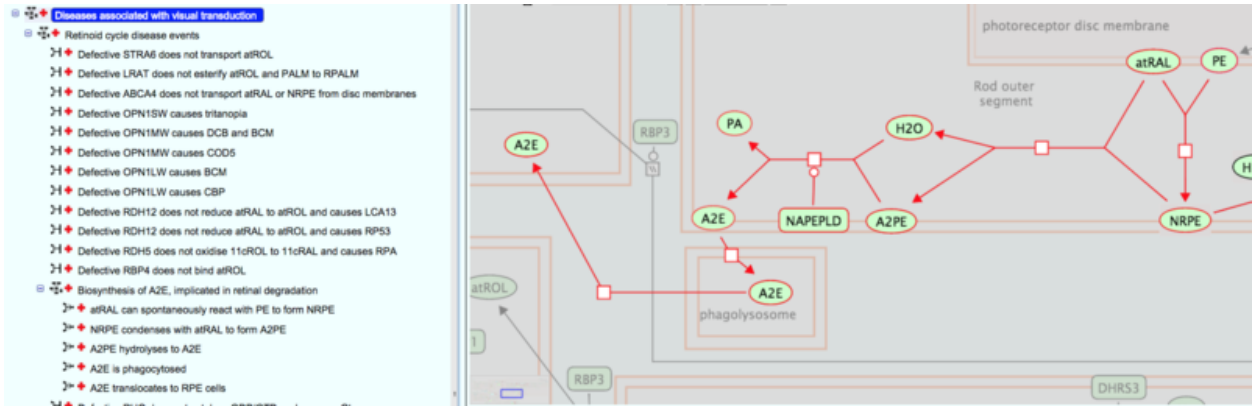
Now the LoF reaction can be viewed being switched between it and its normal counterpart and will work independently of these two drug-mutant reactions which only appear when looking at the disease counterpart.

Example 2 - diseases caused by accumulation of substrate over time:

Special case in Reactome where the reaction describing a particular substrate's metabolism occurs in normal physiology but over time, accumulation of the substrate leads to toxic consequences which can lead to disease. Examples are neurodegenerative diseases such as Alzheimer's and Parkinson's diseases, chronic obstructive pulmonary disease and retinal macular degeneration.

The bisretinoid A2E (di-retinoid-pyridinium-ethanolamine) is a major component of lipofuschin, a yellow-brown pigment grain composed mainly of lipids but also sugars and certain metals whose accumulation is associated with degenerative diseases. In the eye, A2E is the end-product of the condensation of 2 molecules of all-trans-retinal and phosphatidylethanolamine in photoreceptor outer disc membranes. Once formed, A2E is phagocytosed, together with outer segments, to retinal pigment epithelial (RPE) cells where it accumulates. There is no evidence as yet to indicate that A2E can be catabolised. The relationship between lipofuscin accumulation and retinal degeneration is illustrated by Stargardt disease type 1. Because the reactions can be considered "normal" (they occur as part of the normal metabolism of retinoids) as well as disease-causing, they are drawn in the normal pathway diagram and coloured red. To display in the disease view of the diagram, these reactions are added as components of the disease container pathway "Diseases associated with visual transduction". These reactions will now display in both normal and disease diagrams.



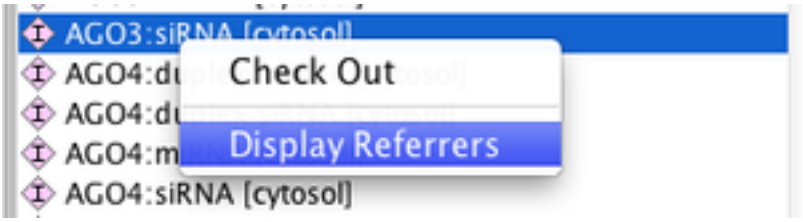


Example 3 - Disease-Specific Gene Expression

See discussion here (Appendix 12) of how to handle cases where gene expression is stimulated downstream of a disease-specific event, but the gene expression doesn't inherently include a disease entity.

Modifying, Deleting and Merging instances

If you locally modify or delete an instance you checked out from gk_central, it will also be **modified/deleted in gk_central** when you synchronize. You must **ALWAYS CHECK FIRST** that the instance you intend to modify/delete is not in use elsewhere, outside your local project. The best way to do this is to search for it in gk_central using the Database Browser Schema View, and select "display referrers". If it has any you didn't know about, do not modify or delete it!



Always use the curatortool database browser to check gk_central for referrers. Right clicking on a database object will allow you to look for referrers.
There may be circumstances when you think something should be modified or removed but if it has been used by another curator, check with them first, or contact an experienced curator for advice.

Procedure for deleting instances (Entities, Events or Regulation) that have been released:

A _deleted instances **MUST** be created any time a **released** Event or Entity is retired/deleted from the database. For example, when event restructuring occurs, any previously released events should NOT simply be flipped to do_release false. When this happens. this instance "disappears without a trace" to users. Before removing the instance, it should be replaced (by a replacement instance if available) in all referrers (where relevant) AND THEN deleted with the creation of a "_deleted" instance that gives a reason for deletion, a replacement instance (where relevant) and an explanation of the deletion. The procedure for choosing the appropriate "reason" and "replacement" instances is described below. This rule applies not only if an EWAS is retired or deleted, but also if the referenceGeneSequence associated with the EWAS is changed, e.g., because a UniProt entry is made obsolete and a different UniProt entry is to be used in its place.

Deleted Instance Options - When you opt to delete an instance you will be presented with a "Deleted Instance Dialog". If the instance has been released, a _deleted instance MUST be created. You can determine if the entity/event has been released by checking if the "released" attribute of its stableID is set to "True". **Note:** _deleted instances MUST have "reason" attribute filled. There are several options:

- **Killed and not replaced** - Used for events or physical entities that are simply wrong and must be removed. The event/entity is deleted and killed and not replaced in all of its referrers. It is CRITICAL that you also identify all diagrams in which the deleted instance has been used and make sure that the entity/event is deleted there as well. If you have the diagram open at the time that you delete the deprecated instance the diagram may updated automatically, but you should verify this. When selecting Killed and not replaced as "reason", the curator must also fill in a "curatorComments" that provides an explanation.
- **Duplicate** - When two separate but duplicate events or physical entities (A and B) are discovered within the data set. These events or physical entities must be

exactly the same. Chose the duplicate instance with the greatest number of referrers (A)as the one to retain. B will be deleted. BEFORE deleting B, you will need to replace it with "A" in all of the referrers of "B". It is CRITICAL that you also identify all diagrams in which the deleted instance has been used and make sure that the entity/event is replaced there as well. If you have the diagram open at the time that you make the replacement of the deprecated instance the diagram may be updated automatically, but you should verify this.
When ready to delete B, use "Duplicate" as the reason in the _deleted instance. In this case, you MUST specify the replacement instance DB_ID in the "replacement" attribute. ****Note for discussion: it is also possible to resolve duplicate instances via a merge in the curator tool as long as ALL of the referrers of both A and B are in the project during the merge. However, in the case of the merge, it is not possible to specify which instance is preserved as far as I know. Also, with a merge the history is lost (there is no record of the deleted instance). Question: should merge be used for "unreleased" duplicates?

- **Split** - When a single entity is found to represent two separate and distinct entities. If the entity is to be split as a result of an external database change (UniProt) this is handled by the managing editor (L.Mattthews), and clean up instructions will be provided. If splitting an event such as a high level pathway, the deletion must be discussed with the editors. If the instance to be split is an event, the common practice is to duplicate and then partition the components accordingly. Generally we feel that split should not be used for reactions and low-level pathway events at this time.
- **Merged** - Very similar to duplicate. When a number of similar (but not exactly the same) events/entities A and B are merged to a single instance A and B is deleted. These could be steps in a pathway that are collapsed to a single step. When Uniprot merges entries (say 2 proteins thought to be different turn out to be the same), this is reported in the Uniprot update and is managed by the release manager. Again, BEFORE deleting B, you will need to replace it with "A" in all of the referrers of "B"! It is CRITICAL that you also identify all diagrams in which the deleted instance has been used and make sure that the entity/event is replaced there as well. If you have the diagram open at the time that you make the replacement of the deprecated instance the diagram may be updated automatically, but you should verify this.
  When ready to delete B, use "Merged" as the reason in the _deleted instance. In this case, you MUST specify the replacement instance (A) DB_ID in the "replacement" attribute.
- **Obsolete** - Instance A is "killed" but has a replacement B. BEFORE deleting A, you will need to replace it with "B" in all of the referrers of "B"! It is CRITICAL that you also identify all diagrams in which the deleted instance has been used and make sure that the entity/event is replaced there as well. If you have the diagram open at the time that you make the replacement of the deprecated instance the diagram may be updated automatically, but you should verify this. When ready to delete A, use "Obsolete" as the reason. In this case, you MUST specify the

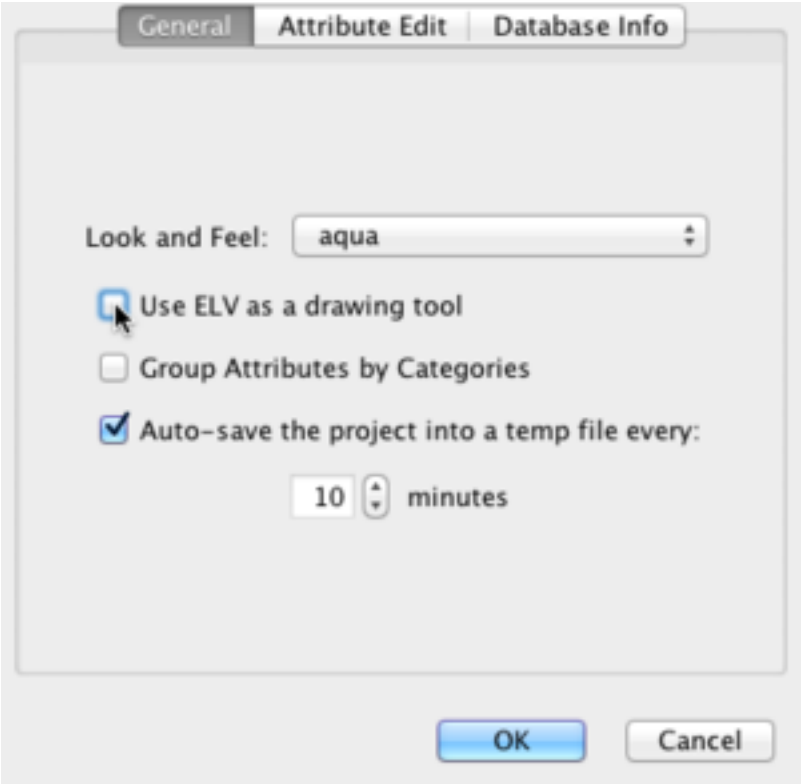
replacement instance (B) in the "replacement" attribute. Provide an explanation for replacement in the "curatorComments" attribute.

Updating modified reactions in the ELV (critical!)

When modifying existing reactions that have already been drawn into diagrams and deployed, be aware that not all changes made to participants in reactions will be automatically reflected in a diagram. For example, if you plan to replace the input of a reaction that is an EWAS with a set, the set will be drawn into the diagram, but the original EWAS will not disappear. It will have to be manually removed as described in the example below.

Example:
The original reaction had a single EWAS (FDX1L) as an input. I then made a defined set of 2 FDX proteins (FDX1, FDX1L) and substituted the defined set for the single EWAS as an input for the reaction. The ELV correctly showed the change (use of the set as an input) but also continued to show FDX1L (as a member of the defined set). I don't really want the EWAS FDX1L on the ELV anymore so I tried to delete it but received the error message "In the ELV drawing mode an entity cannot be deleted if it is linked to a reaction...").

To remove the EWAS (From the diagram only!!!) deselect the "Use ELV as a drawing tool" within the curatortool preferences panel, select to delete the EWAS from the ELV but retain the EWAS in the project. **After you delete these unwanted objects, please turn on this checkbox again to avoid any side-effects in the future.**



Toggling the "Use ELV as a drawing tool" option to off allows the drawing tool to directly modify the data in your project. Be very careful and be sure to reselect this option once you are done.

Please note that if an unused entity icon is still connected by lines to other entities (e.g., an orphan EWAS is connected to a set that includes it), you need to delete these connecting lines before the curator tool will allow you to delete the unused entity icon.

The bottom line is that you should always review and update the pathway diagram after modifying the reaction, and if the reaction is used in multiple diagrams ALL diagrams will need to be updated, so check for referrers of any reaction that is being modified.

Merging duplicate entities

To perform merging (for a complex for example), do the following:

1). In the database schema view, get these two complexes together in the instance list (the central panel) by doing REGEXP search: list DB_IDs together as this: DB_ID|DB_ID (Note | between two DB_IDs)

2). Check out these two instances.

3). While still in the view from step 1, search for Referrers for one complex, and check out all referrers in the Referrer dialog. Do this for the other instance.

4). In the local project schema view, merge these two complexes.

5). Find all instances that have been changed based on flag “>”.

6). Find pathway diagrams containing changed reactions. You may have to do this manually for the time being. In the db schema view, get pathways containing these changed reactions and see if they have diagrams (if the pathway listed does NOT have a diagram, look upwards at parent pathways until you find one that DOES have a diagram. Check these diagrams out by using “Build Project”. Re-open diagrams should fix references automatically. But you may need to re-position auto-inserted complex in the diagrams. So please check it manually to make sure.

7). Commit all changes into gk_central.
8). Delete the deprecated instance AFTER all other changes. Do NOT commit the deletion first. By committing the deletion first, the curator tool automatically updated referrers of the deleted instance by removing the deleted instance from their attributes and attached an InstanceEdit to these referrers. Consequently, you will see conflicts if the deletion is done first.

Deleting diagrams

If you are deleting a pathway diagram, you must delete the corresponding PathwayDiagram instance, you don't need to delete SearchableVertex instances that are associated with that diagram. In the future, this will be handled automatically.

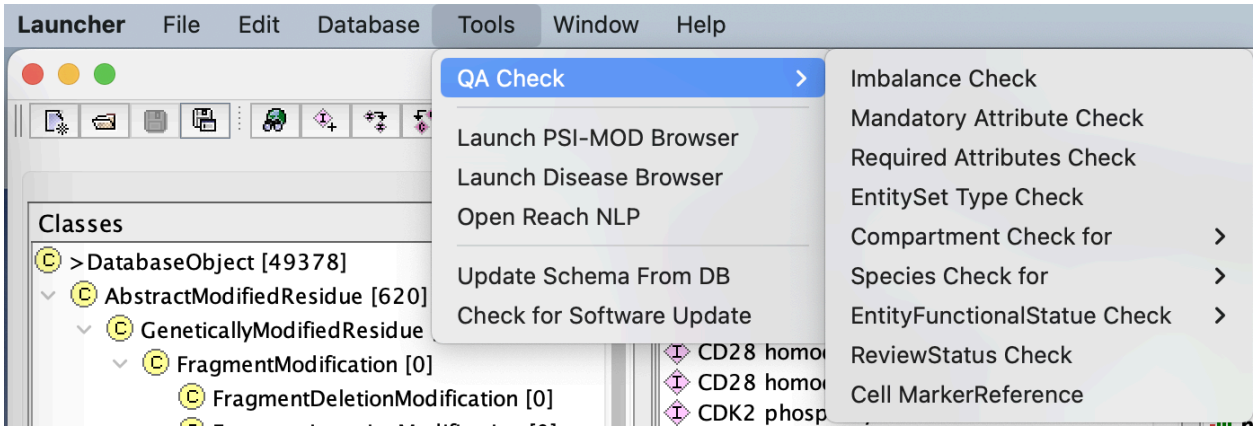
Diagram checks after deleting entities or reactionlikeEvents

If and when you need to delete an entity from gk_central, you must run the deleted object in diagram check over gk_central to find any diagrams that have used those instances.

QA

Project QA using the Curator Tool QA checks

Within the Tools menu the "QA Check" menu can be found.



QA Check Menu

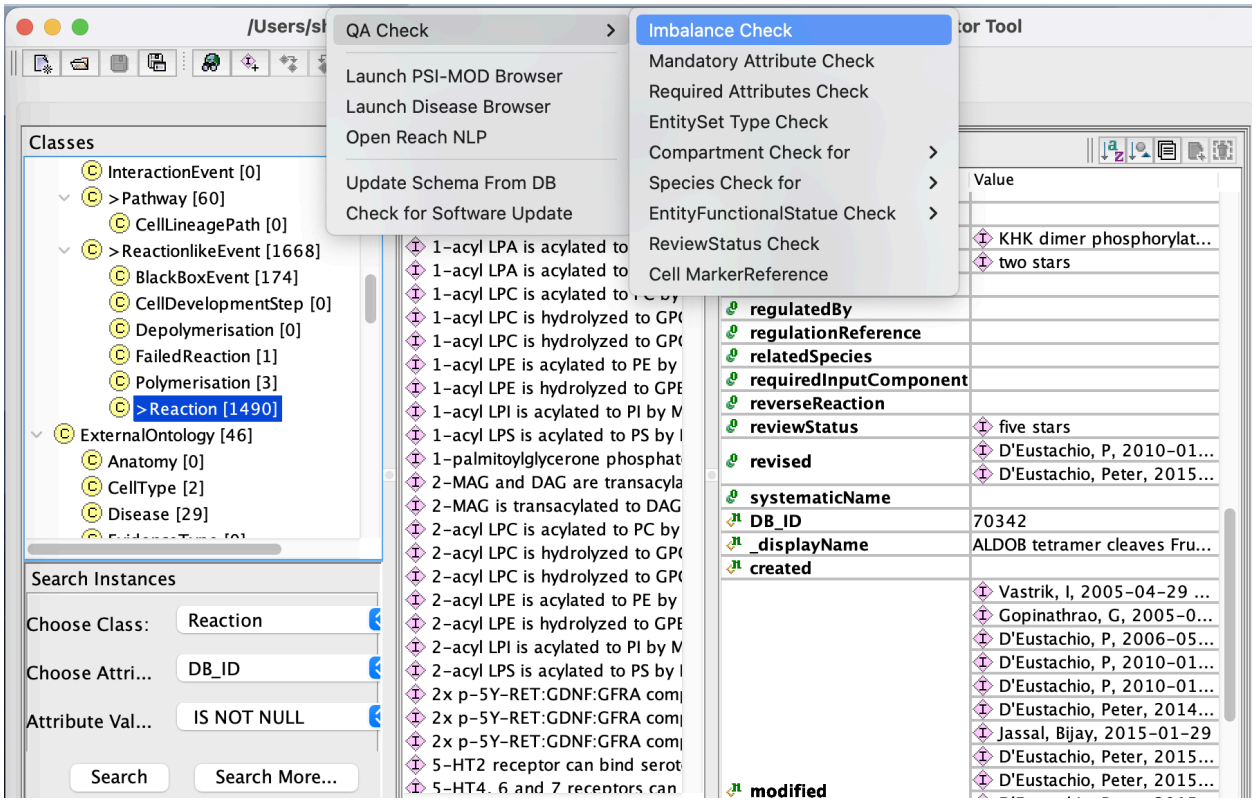
This menu has nine separate QA script items within it.

- Imbalance Check: checks that the EWAS referenceEntities present as input are also present as output and in the correct stoichiometry.
- Mandatory Attributes Check: Checks that the mandatory attributes for a class have been entered.
- Required Attributes Check: Checks that the required attributes for a class have been entered.
- EntitySet Type Check: Checks that set members belong to the same data model class.
- Compartment Check:
    - EntitySet: Looks for conflicts between compartment of a set and the compartment of members of that set.
    - Complex: Looks for conflicts between compartment of a complex and the compartment of individual components.
    - Reaction: Looks for conflicts between compartment of a reaction and the compartment(s) of individual participants.
- Species Check for
    - Entity Set: Looks for conflicts between the species of set members and the species of the set containing them..
    - Complex: Looks for conflicts between species of a complex and the species of individual components of that complex.
    - Reaction: Looks for conflicts between species of reactionLikeEvent and the species of all reaction participants.
    - Pathway: Looks for conflicts between the species of a pathway and the species of its contained events.
- EntityFunctionalStatus check
    - DiseaseEntity: Verifies that the diseaseEntity specified in the EntityFunctional status of an disease RLE is a participant of the disease RLE
    - NormalEntity:Verifies that the normalEntity specified in the EntityFunctional status of the disease RLE is a participant of the normalReaction specified in the disease RLE.
    -
- ReviewStatus Check: Flags Events have not been reviewed after structureModified instanceEdit date stamp
- Cell MarkerReference:

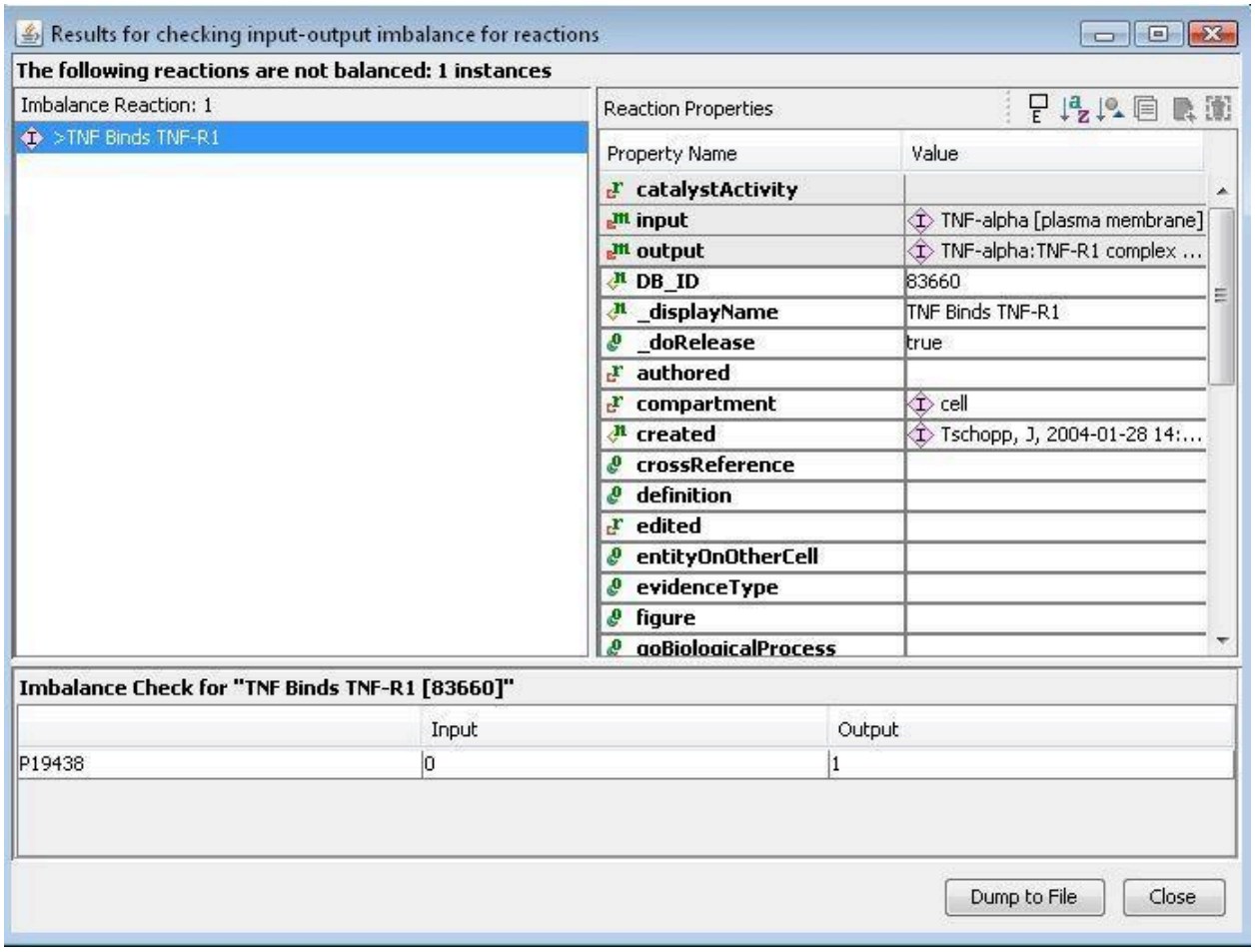
Note: In order for the QA checks to effectively pick up errors, the project that you are working on must be fully extracted from the database. Instructions on how to do a full extraction can be found here (Appendix C13).

IMBALANCE CHECK

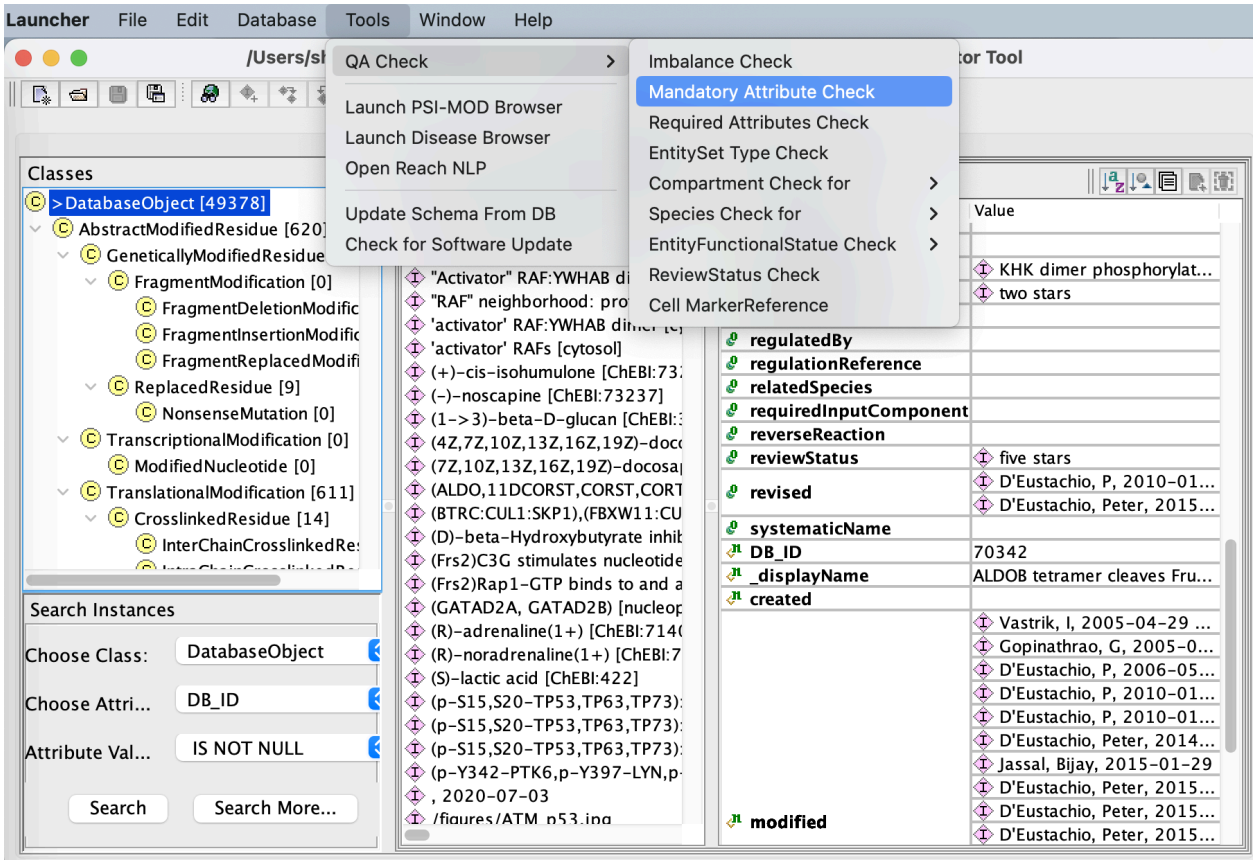
You must select Reactions in the hierarchy to perform this check.



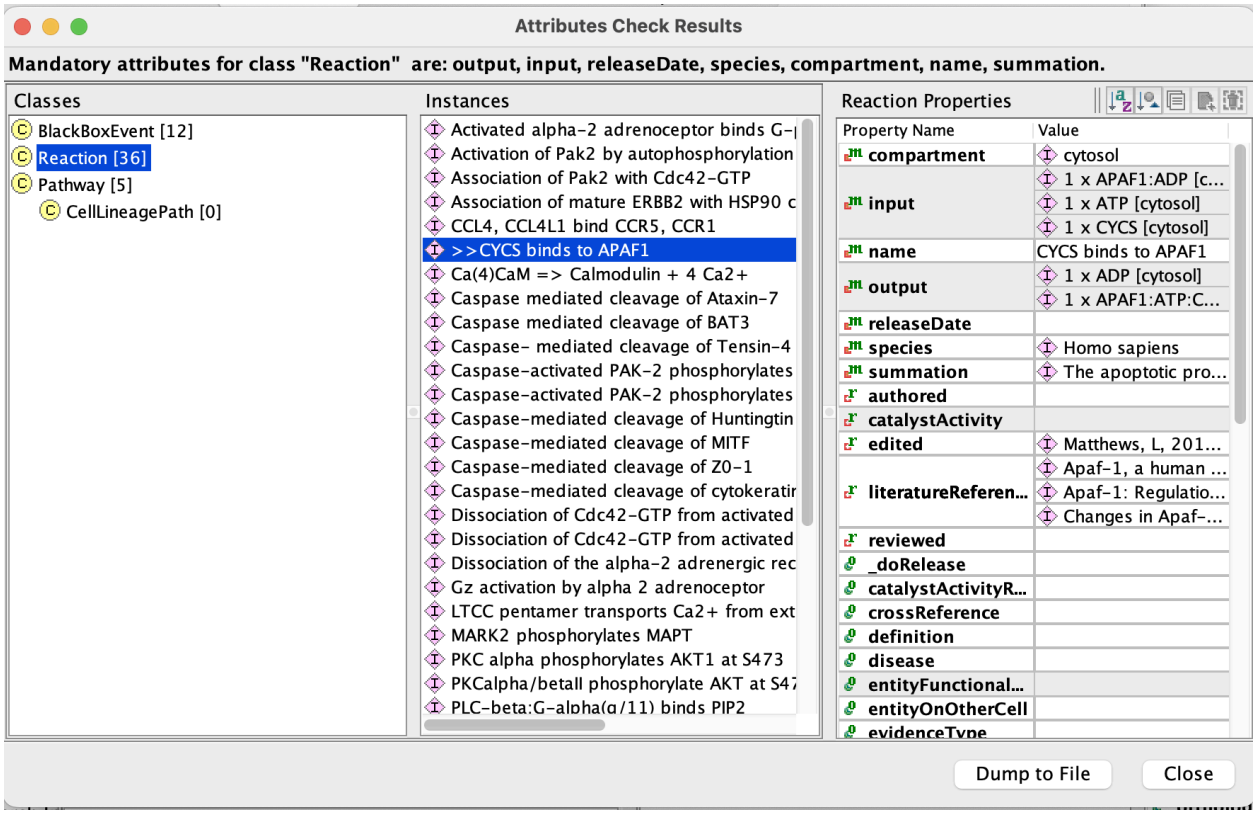
Reactions are flagged as cleavage reactions if the output differs from the input only that the output contains "fragments" of the input molecule. A true imbalance is shown below:



MANDATORY ATTRIBUTE CHECK



A list of instances missing mandatory attributes (by class) is shown. To make the missing attributes of the instance easier to see, you can use the "order attribute" button (downward arrow with circle and triangle) in the upper right side of the tool. This orders the instances by type (mandatory, required, optional...etc)

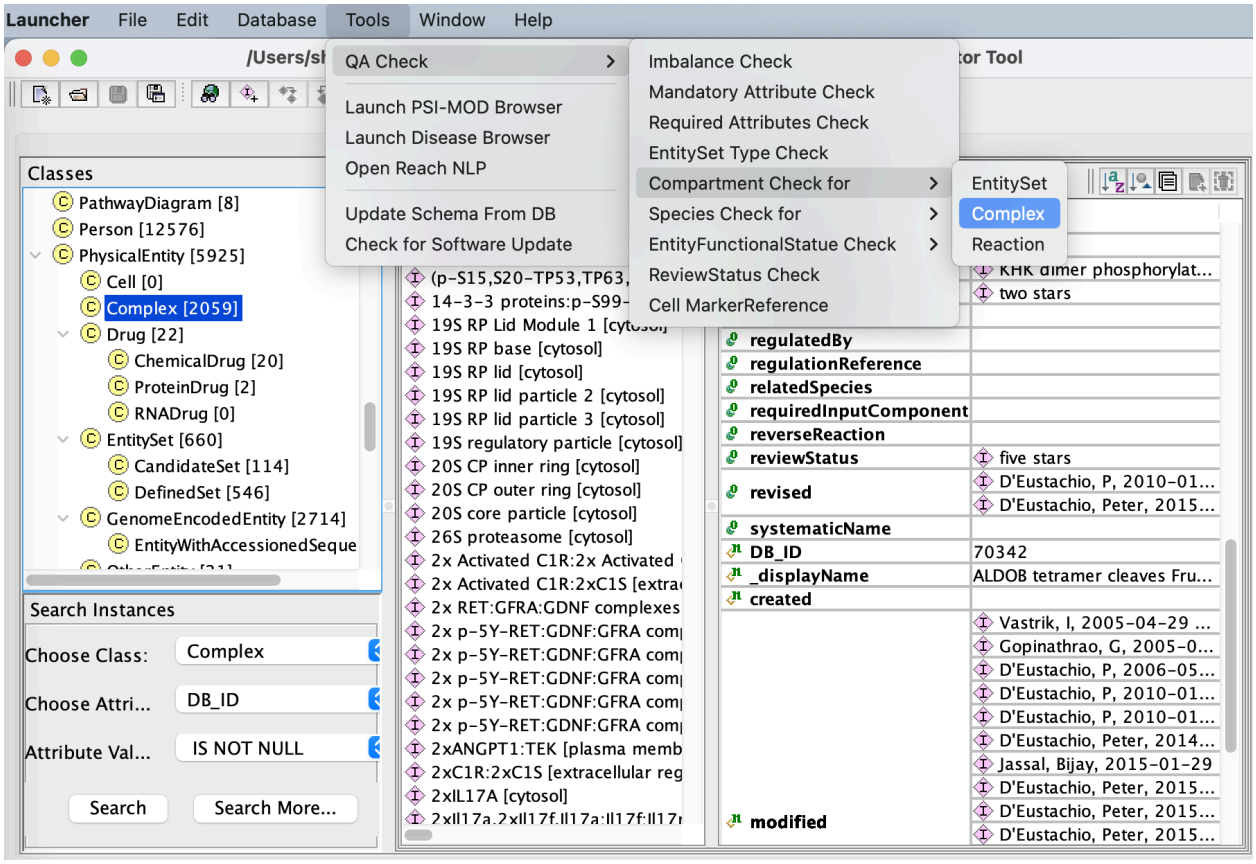


Required attribute check

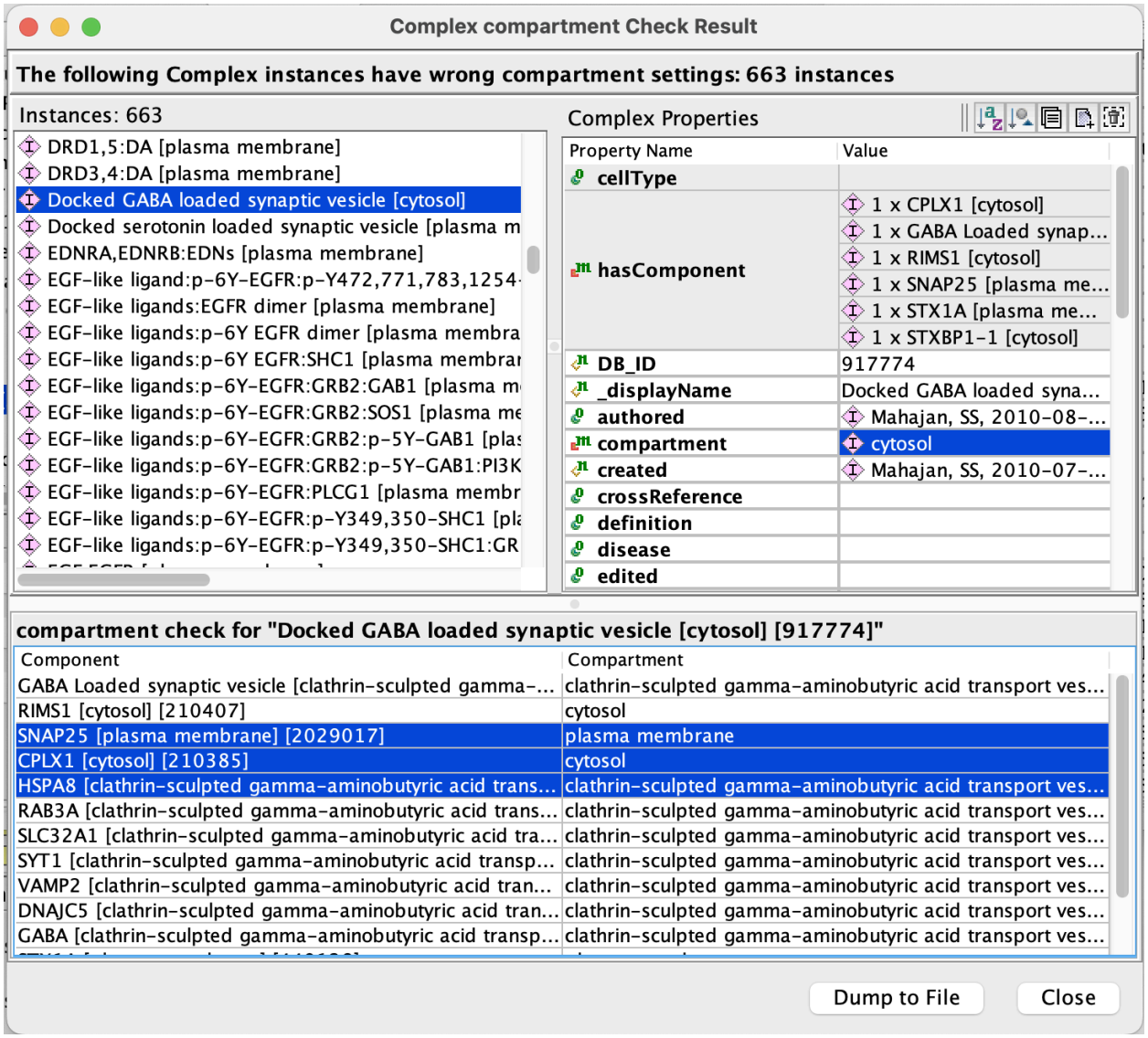
This checks work in the same way that the mandatory attribute check works.

Compartment check

Select the class you want to check in the hierarchy panel



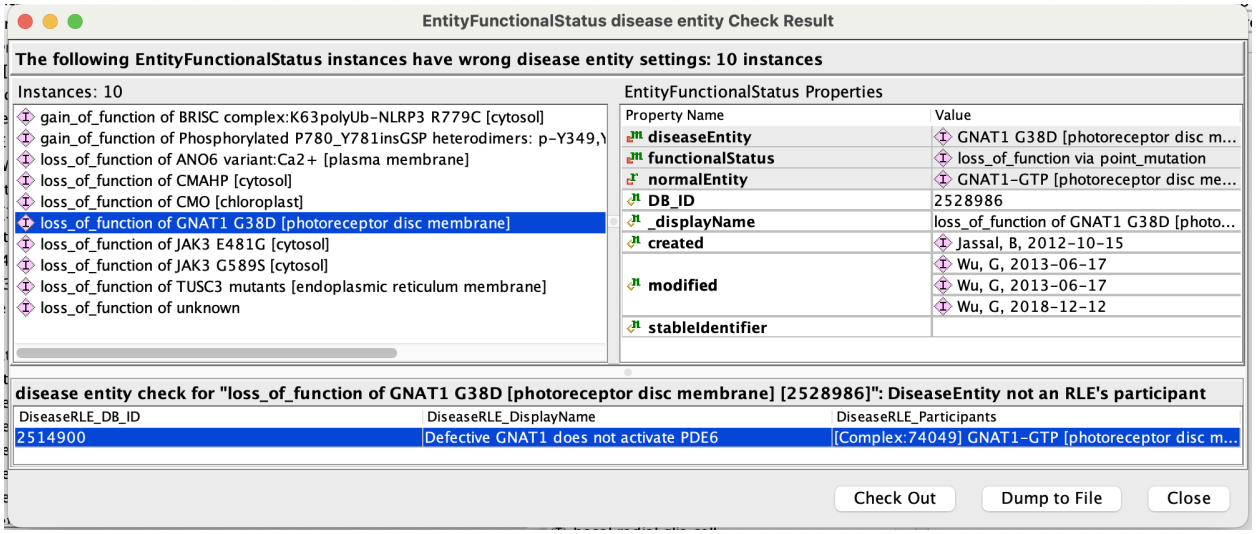
Compartment conflicts are indicated in the bottom of the dialog box.

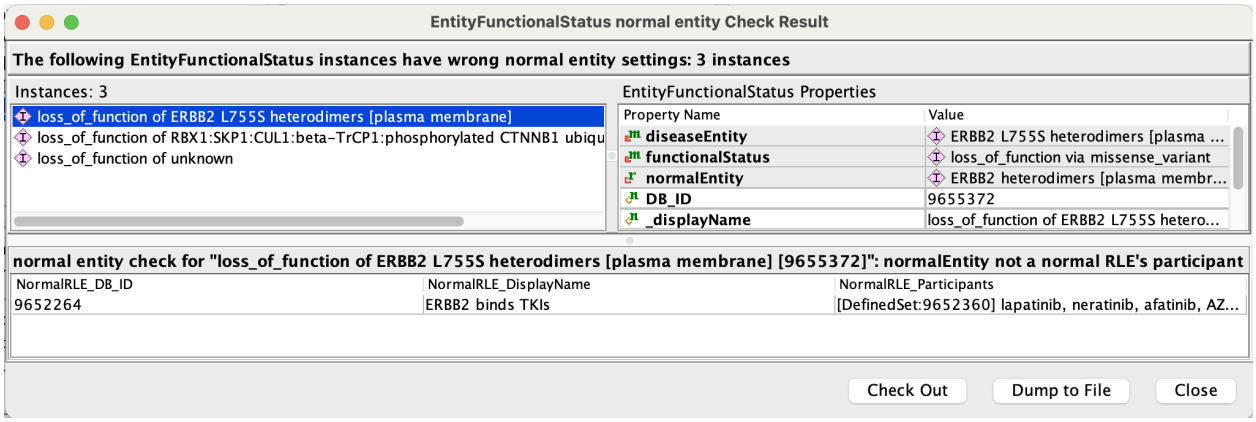


Species checks

These checks work in the same way that compartment checks work with the exception that Pathways are also checked for species conflicts.

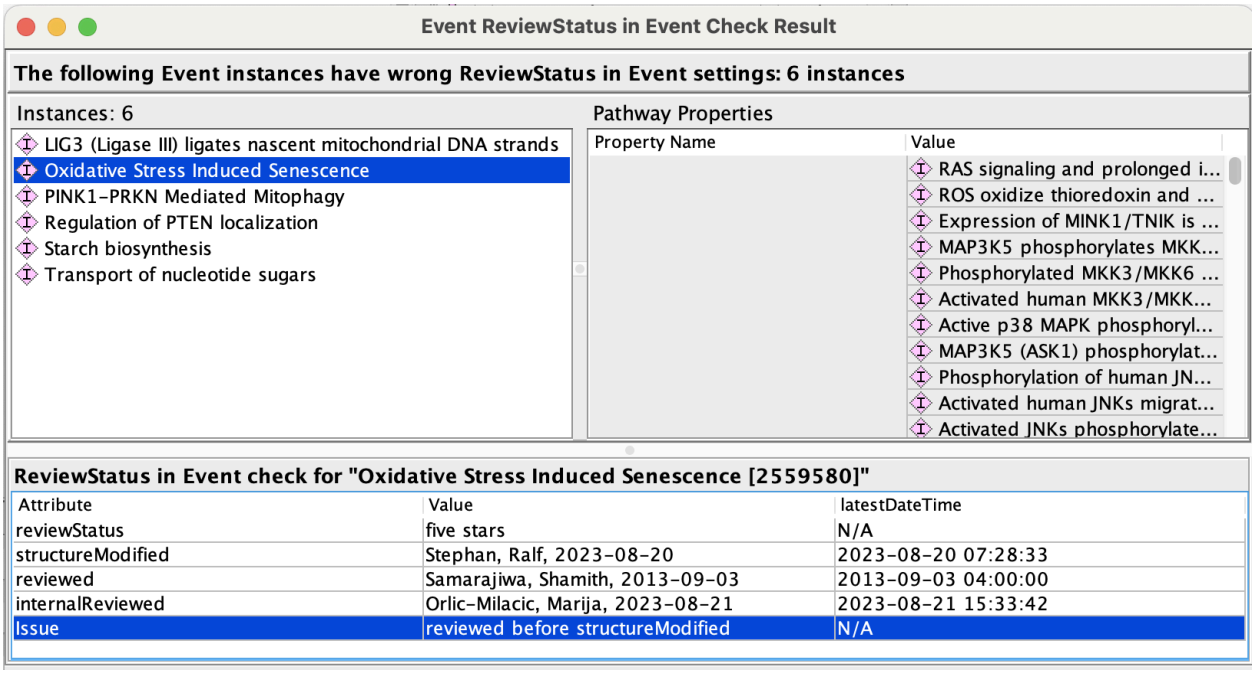
EntityFunctionalStatus checks





ReviewStatus checks

This Event check flags instances that have a structureModified date after the last reviewed date as displayed in the dialog box below:

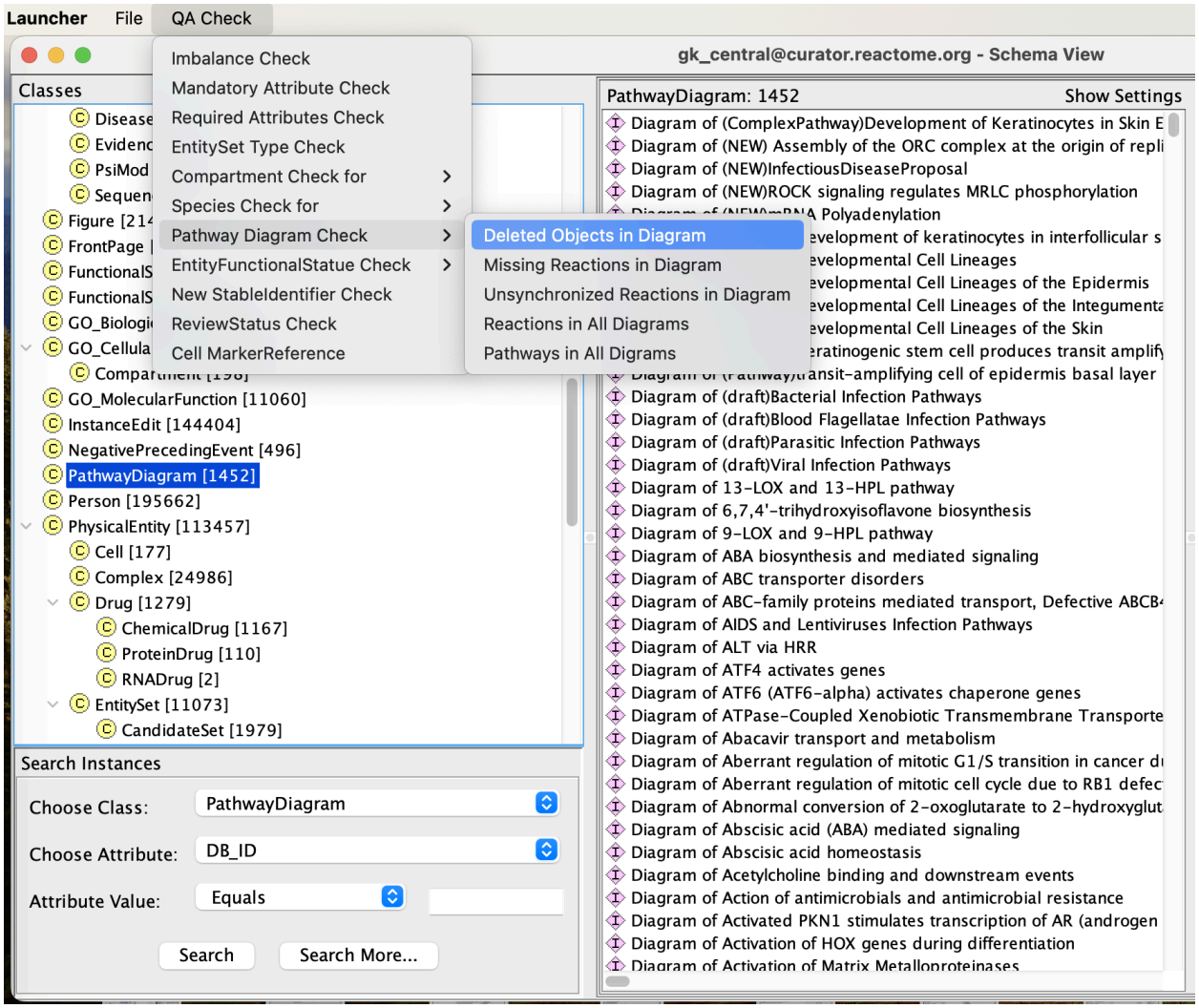


Cell MarkerReference

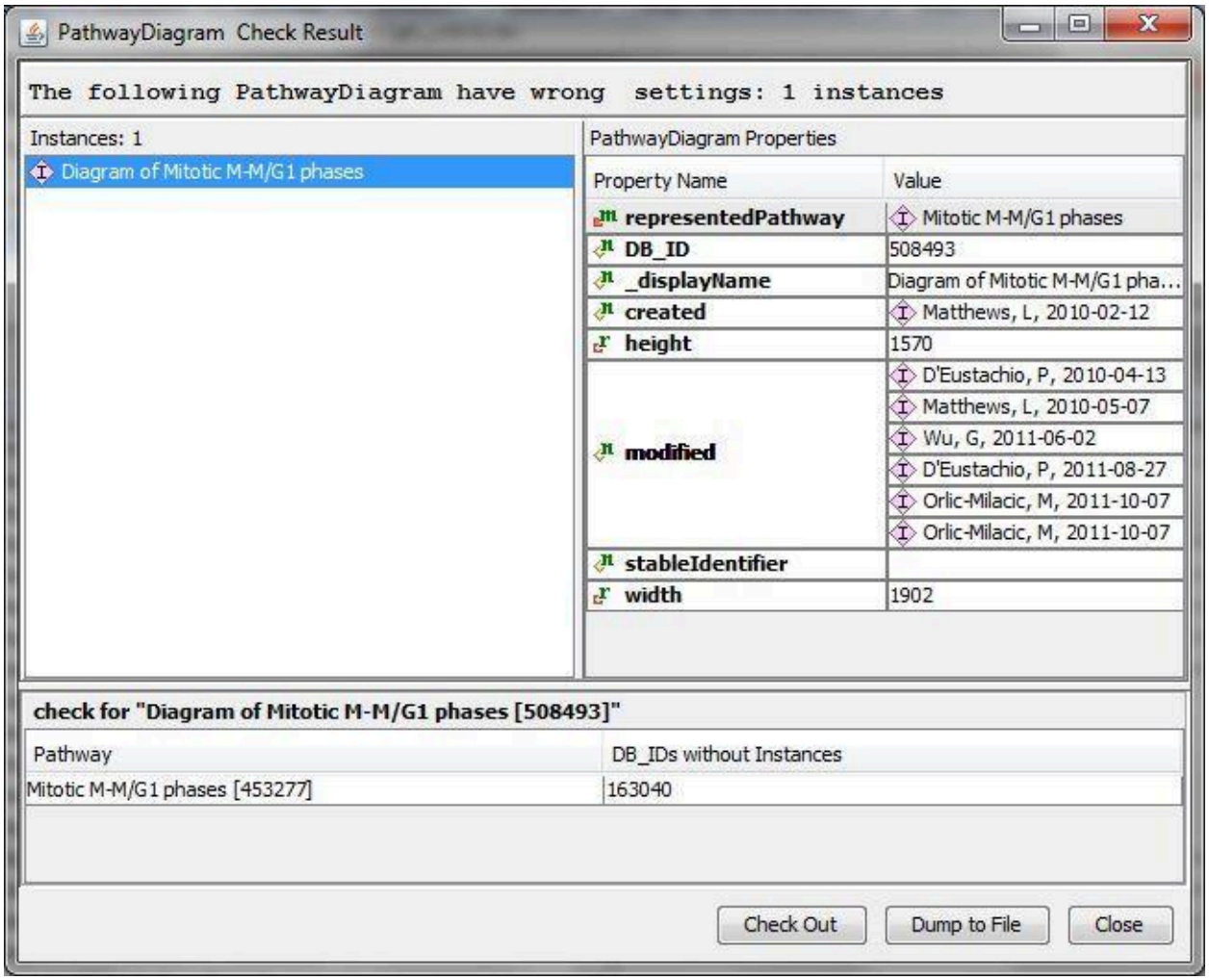
Diagram checks

- Deleted objects in diagrams

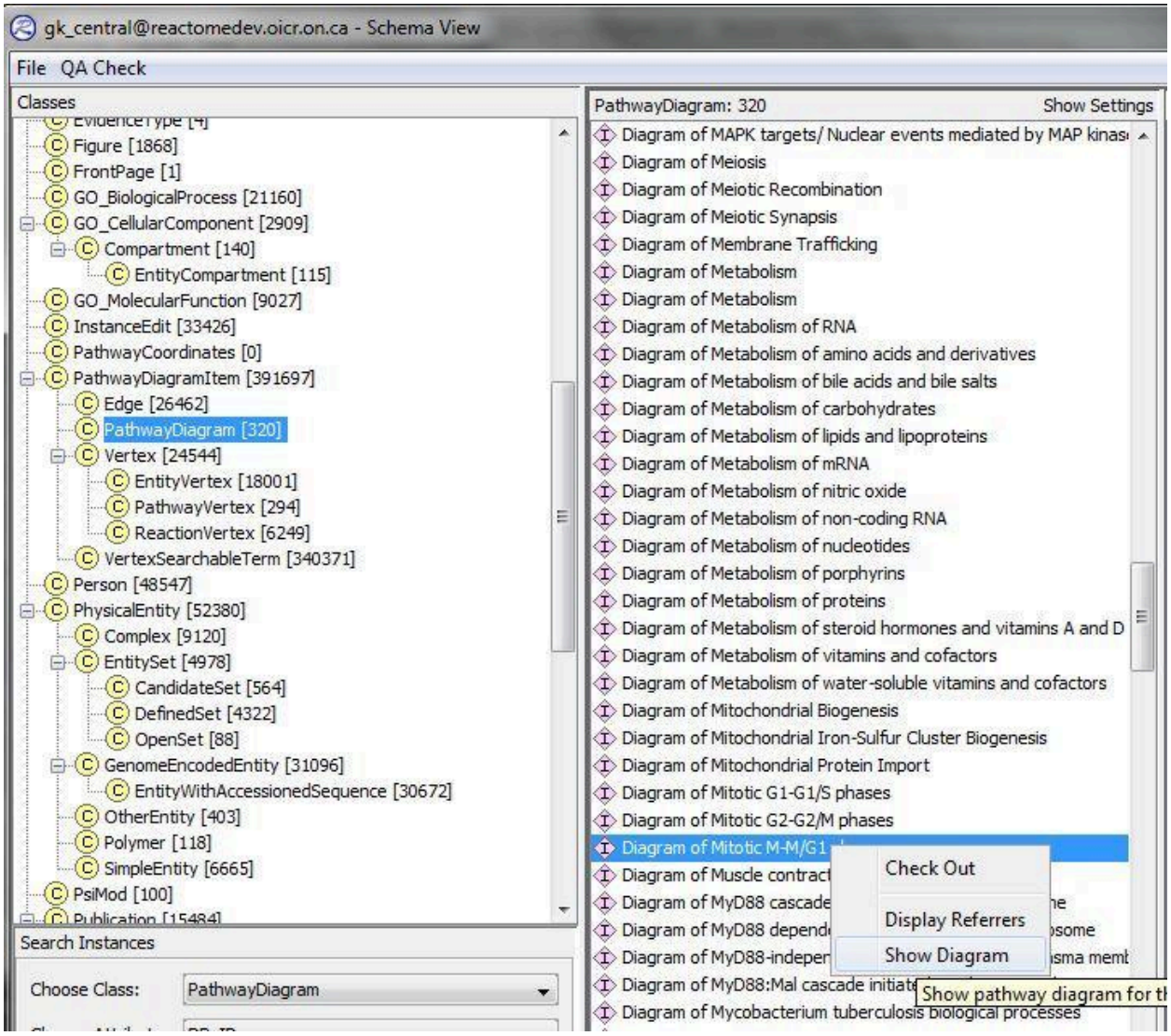
When an entity or reaction is deleted in the instanceview of the curator tool, it must also be removed manually from any diagram that it has been drawn into. This does not happen automatically. This check is run over gk_central and will look for any diagrams that are affected by the deletion of a reaction or ReactionLikeEvent. This check MUST be run after any deletions of reactionlike events or entities have been committed to gk_central so that the affected diagrams can be identified and the appropriate changes made in any affected diagrams.



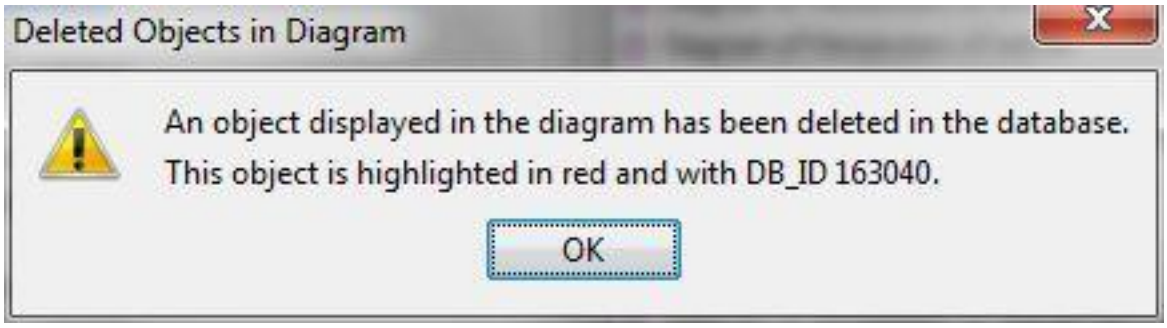
Any affected diagrams will be flagged. The DB_ID of the deleted instance will be displayed, but to see the affected "objects" in the diagram, you will need to view the diagram in gk_central.



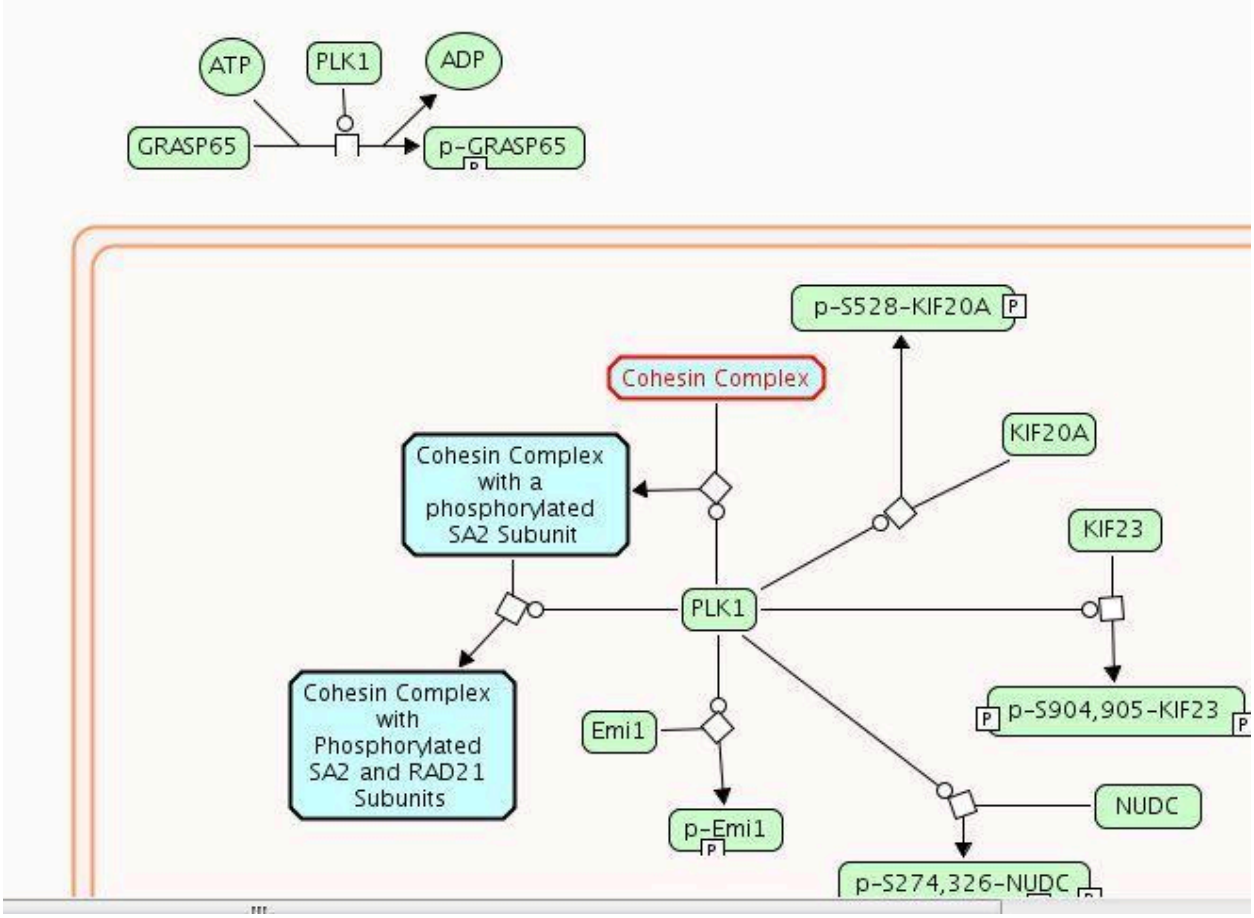
Select that diagram in gk_central, right click and opt to "Show diagram".



Affected objects will be flagged



and the objects in the diagrams highlighted in red.



QA OF PROJECTS BEFORE RELEASE

Because of the nature of the release process and the growing number of curators submitting projects the QA load has become greater. One of the solutions to this problem is for curators enter this data right from the beginning and to run the QA checks in the curator tool before finishing their projects

- Top Six List Of Problems Identified During the Slice
  - Species
  - Complex Balances
  - UniProt IDs
  - Complex Compartment Checks
  - Entity Compartments
  - Balancing Of Reactions

QA REMINDERS

1. All of the instances must be updated from gk_central in order for these checks to be meaningful

2. QA Checks should be run regularly, once you have created a reaction or even a bunch of EWASs.

Everyday QA includes:

- Complete check-outs (No shell instances)
- Match instance in DB
- QA Tools

Curation Tools/Automation

Remote Attribute Search Tool

See here (Appendix C14) for description of the remote attribute search tool.

To search Reactome curator, use:
http://curator.reactome.org/cgi-bin/remoteattsearch2?DB=gk_central

To search reactomerelease, use:
https://release.reactome.org/cgi-bin/remoteattsearch2?DB=slice_test

Using Reach NLP feature in the Curator Tool to process a personalized list of PMCIDs.

The Reach tool embedded in the curator tool uses automated NLP technology to extract pathways information from journal articles and suggest potentially annotateable interactions. The below protocol allows the curator to specify the list of PMCIDs through which REACH will search.

1. Formulate a PubMed search term, e.g. “(CDH2 OR "cadherin-2" OR "N-cadherin") AND (expression or transcription or promoter or binding or complex or interaction or regulation) and adhesion”

2. Note the number of results: e.g. 2,312 results for the above search term (on June 16, 2022). If a gene is poorly studied, it’s best to just use gene name + synonyms as a search term. If it’s a well-studied, popular gene, there will be too many results to process, most of which will be of no use for Reactome. In the latter case, it’s best to add additional words that imply some mechanistic evidence (second bracket in the above search term) and the annotation context (in the case of CDH2 it’s cell adhesion). You can build the search term interactively, depending on whether the number of results seems manageable and if the content of results seems appropriate.

3. Filter the results with the Free Full Text filter, the left-hand tab in PubMed: 1,230 results.

4. Change Display options FORMAT in the upper right corner from Summary (default) to PMID.

5. Save the list of PMIDs for your reference.

6. Go to https://www.ncbi.nlm.nih.gov/pmc/tools/idconv/ . This tool allows conversion of PMIDs to PMCIDs that need to be fed to NLP Reach. Paste the list of PMIDs and select csv as the output format (underneath the entry window). Unfortunately, the tool can process no more than 200 PMIDs at a time, unless an API is used – perhaps curators could get some training on how to do this. Otherwise, a long list of PMIDs will have to be processed in more than one run.

7. Look at the output file – if no PMCID is available for a PMID, then the PMCID column will stay blank. The PMIDs with no matching PMCID have to be manually evaluated for relevance by a curator.

8. Copy-paste the list of PMCIDs into a text editor such as Notepad and submit for processing (at the moment, forward to Guanming; the processing time is substantial, so it's best to do it well in advance).

9. After obtaining the results file in the .rrtf format (the file can be quite big, so a Dropbox needs to be used for sharing), follow Guanming's instructions for opening them in the Curator Tool:

    a. In the Curator Tool, use menu Tools/Open Reach NLP to open the window for Reach
    b. Find a button called "Import" at the bottom. Please use Import to load the results from the rrtf file

10. After importing the NLP results, you can visualize them using different filters:
    a. The number of rows will appear overwhelming. However, most rows do not contain proteins/genes/RNAs as their Entities (please note Entity A and Entity B). Filter the results to contain your gene of interest or one of its synonyms as either Entity A or Entity B. Currently, it's not possible to filter for several things at once using the OR statement, so you have to filter the results for keywords one-by-one. Please note that the filter is case-sensitive. For example, if you filter for "cdh2", you may get 0 results, while you will get quite a few if you filter with "CDH2".
    b. You can filter by the number of statements that support a relationship. Please note that more than 1 statement can originate from a single publication. Manual inspection is required to verify that a relationship is supported by at least 2 publications.

Examples of how to use the remoteattsearch tool can be found here (Appendix C14).

Identifying list members that are unique to one of two lists using Microsoft Excel

Here (Appendix C14) is a procedure that describes how to take two lists and compare entries to identify those that are present in only one of the two lists. This procedure can be useful, for example, to compare the list of proteins in the gk_central vs. live site to find those that are unreleased.

Preparing a pathway report for Internal and external review

When you prepare a word document for external review, include (or highlight within the document) ONLY the events that you are asking to be reviewed. Longer documents tend to be harder to get reviewed. If you need to include previously reviewed content for context, "gray" this out (using lighter gray text or highlighting) and make reference to this fact out at the top of the word document. Also, point this fact out to the managing editor sending out reviewer requests so that this fact can be pointed out in the letter as well.

Reporting Bugs

Please use JIRA to report bugs in the tool

Appendices

Appendix C1: Overview of Curation

Appendix C2: Quick Guide for New Curators

Appendix C3: Evidence Confidence

Appendix C4: UpdateTracker

Appendix C5: Data Model Glossary - no doc yet, link is to currently being worked on version - in progress

Appendix C6: Small_molecule_renaming - small mols short names

Appendix C7: Review of the PrecedingEvent Suggestion Feature

Appendix C8: Best Practices: Using Chemical Compound Databases

Appendix C9: Cytomics documentation

Appendix C10: Multiple disease and EntityFunctionalStatus when using disease variant subsets

Appendix C11: GtP_all_interactions

Appendix C12. Disease-specific gene expression

Appendix C13: Deep Extraction of Events

Appendix C14: Remote Attribute Search and examples

Appendix C15: Remove duplicates MSExcel

V90_Sept 2024

- update to Disease pathway curation guide
- incorporation of Cell lineage curation guide
- general updates and reorganization

C1. Overview of Curation

The Reactome Knowledgebase

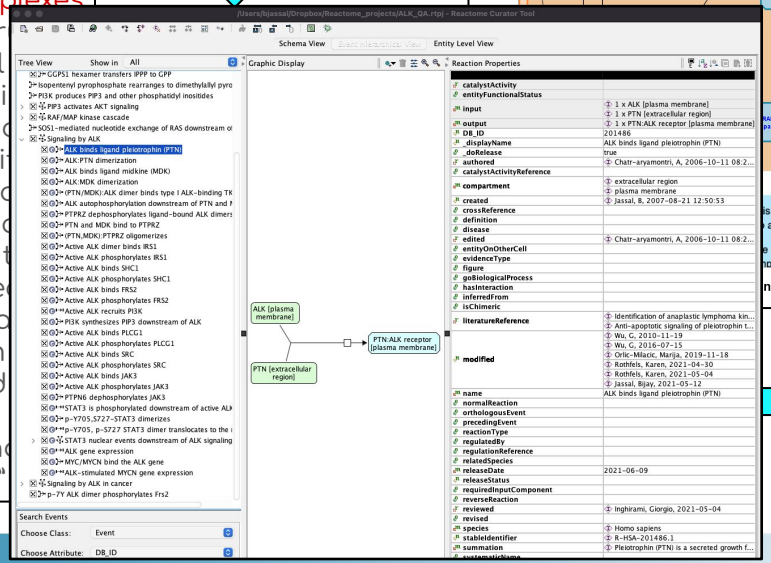
Reactome curation & annotation

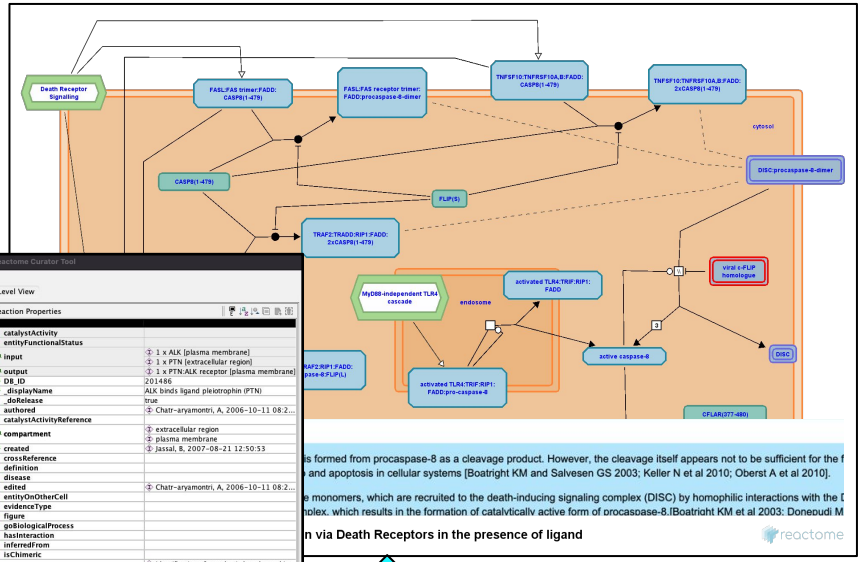
reactome

From text to computer-accessible pathway data

Nature 407(6805):770-6.The Biochemistry of Apoptosis.

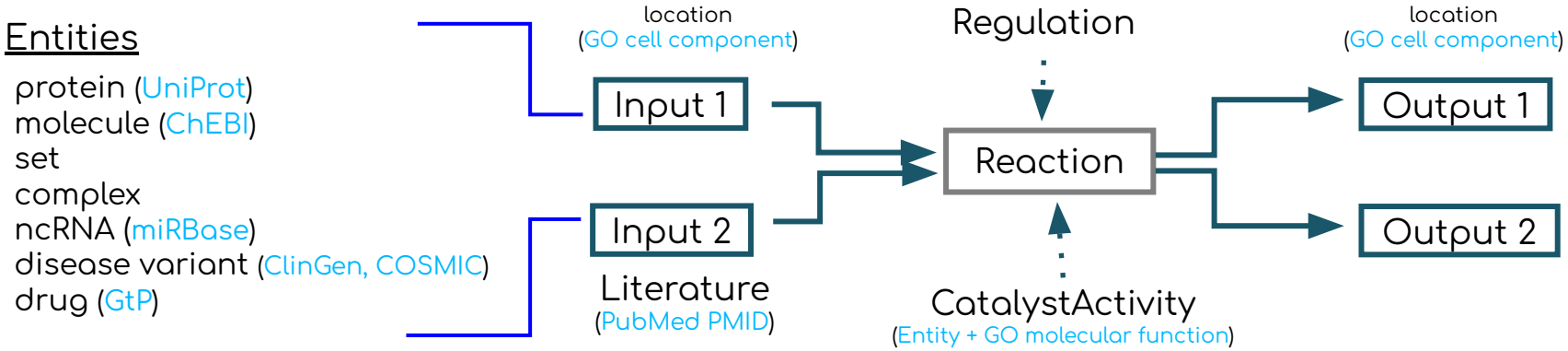
"Caspase-8 is the key initiator caspase in the death-receptor pathway. Upon ligand binding, death receptors such as CD95 (Apo-1/Fas) aggregate and form membrane-bound signalling complexes (Box 3). These complexes then recr through adapter proteins, several molecules of procaspase-8, resulti a high local concentration of zymo The induced proximity model posi that under these crowded conditi the low intrinsic protease activity procaspase-8 (ref. 20) is sufficient allow the various proenzyme mole to mutually cleave and activate ea other (Box 2). A similar mechanism action has been proposed to med the activation of several other caspases, including caspase-2 an nematode caspase CED-3 (ref. 21)."



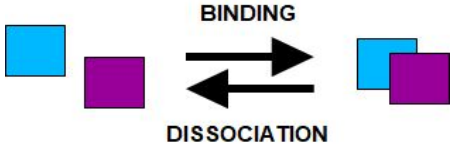


Data Model

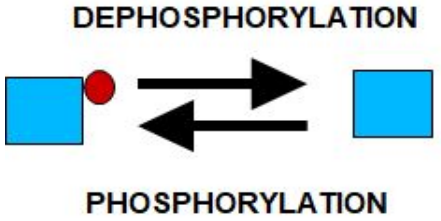
- Reactome is a reaction-based database
- Explicitly describes HUMAN biological processes as a series of biochemical reactions and events

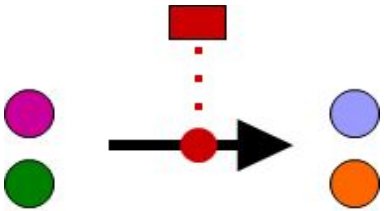


Reactome supports many types of events

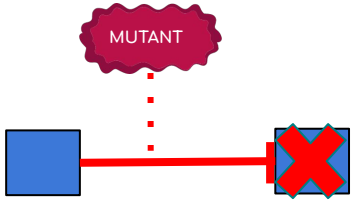




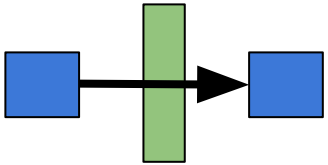




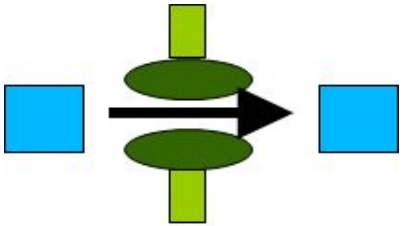
CLASSIC
BIOCHEMICAL



DISEASE

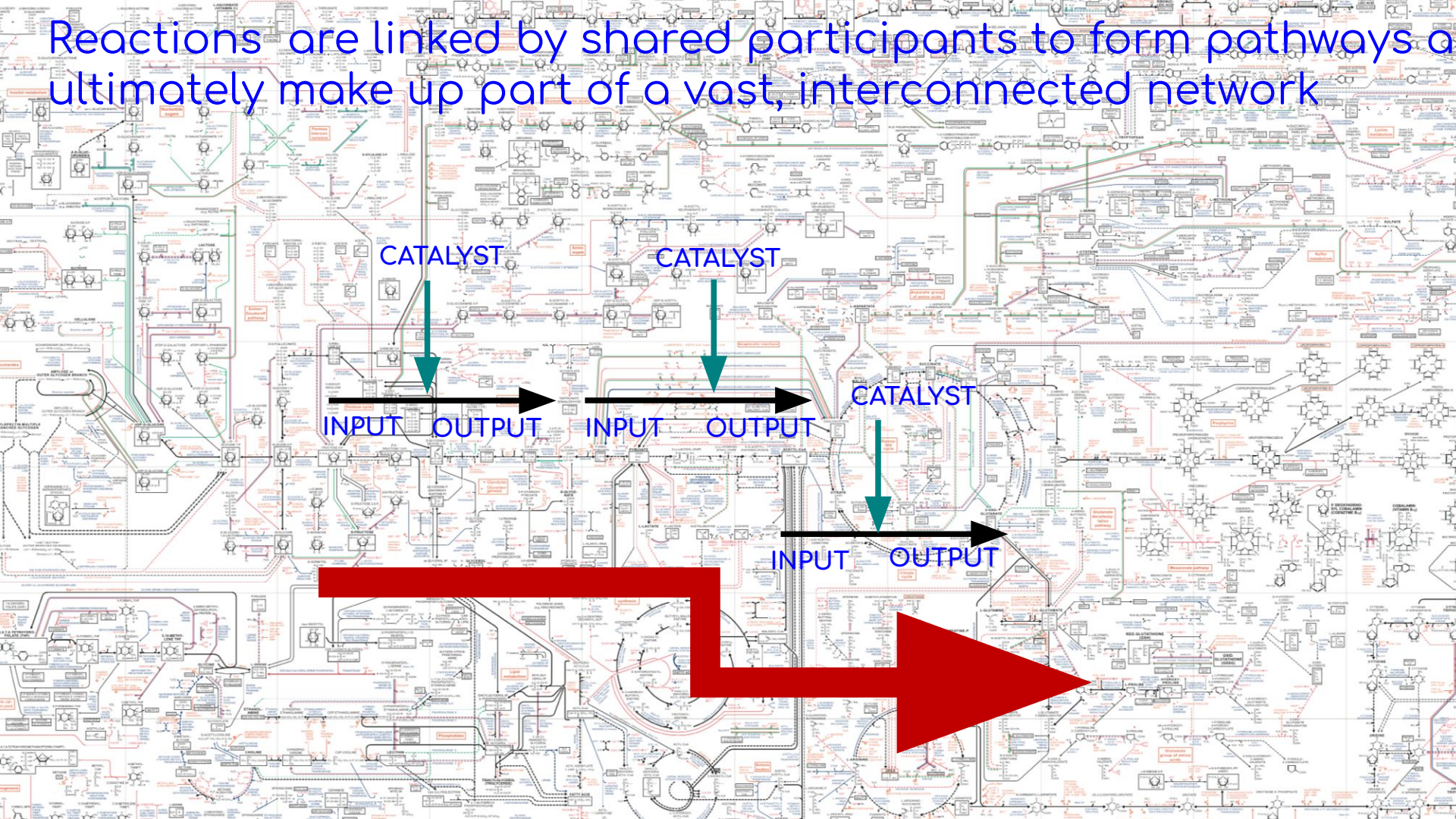


DIFFUSION

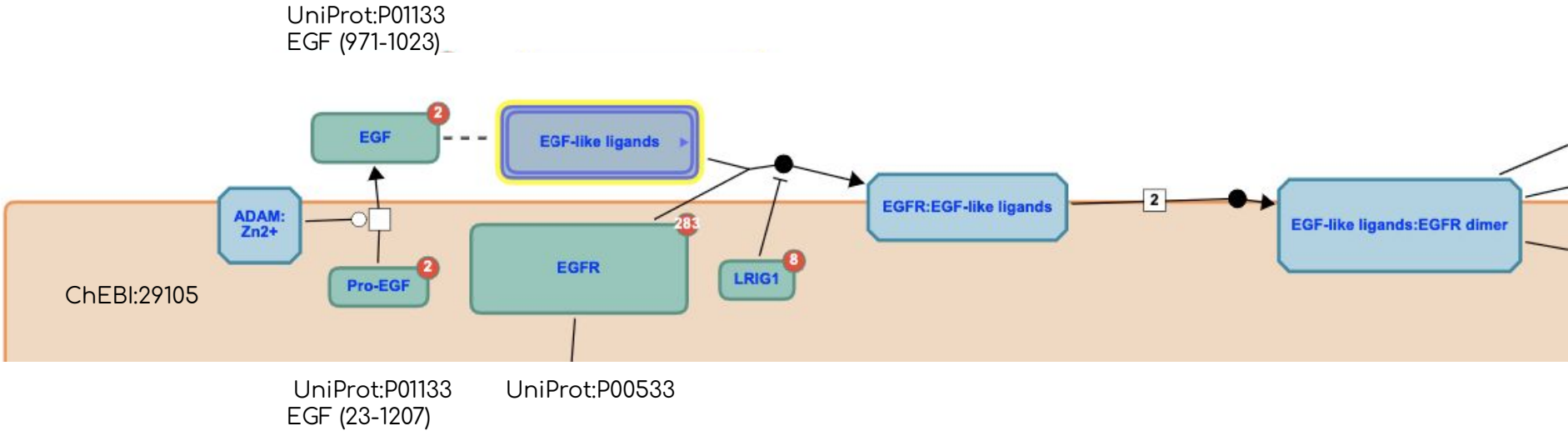


TRANSPORT

Reactions are linked by shared participants to form pathways a ultimately make up part of a vast, interconnected network



Example of linked reactions forming part of the Signaling by EGFR pathway



The curation process



Experts
(recruited)



Curators
(staff)



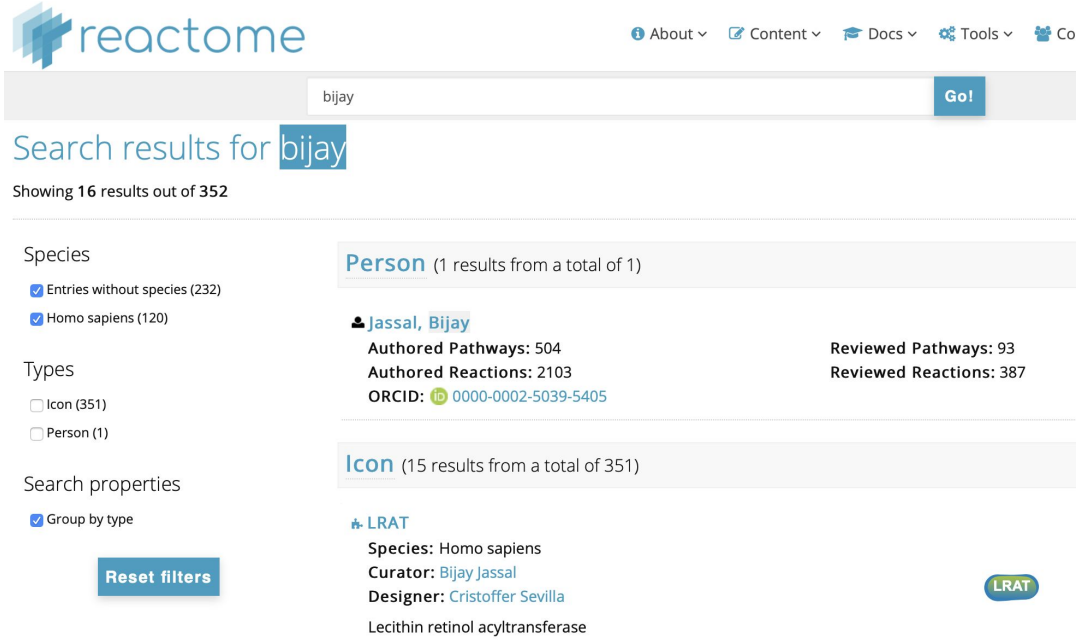
Reviewers
(recruited)

Crediting Contributors

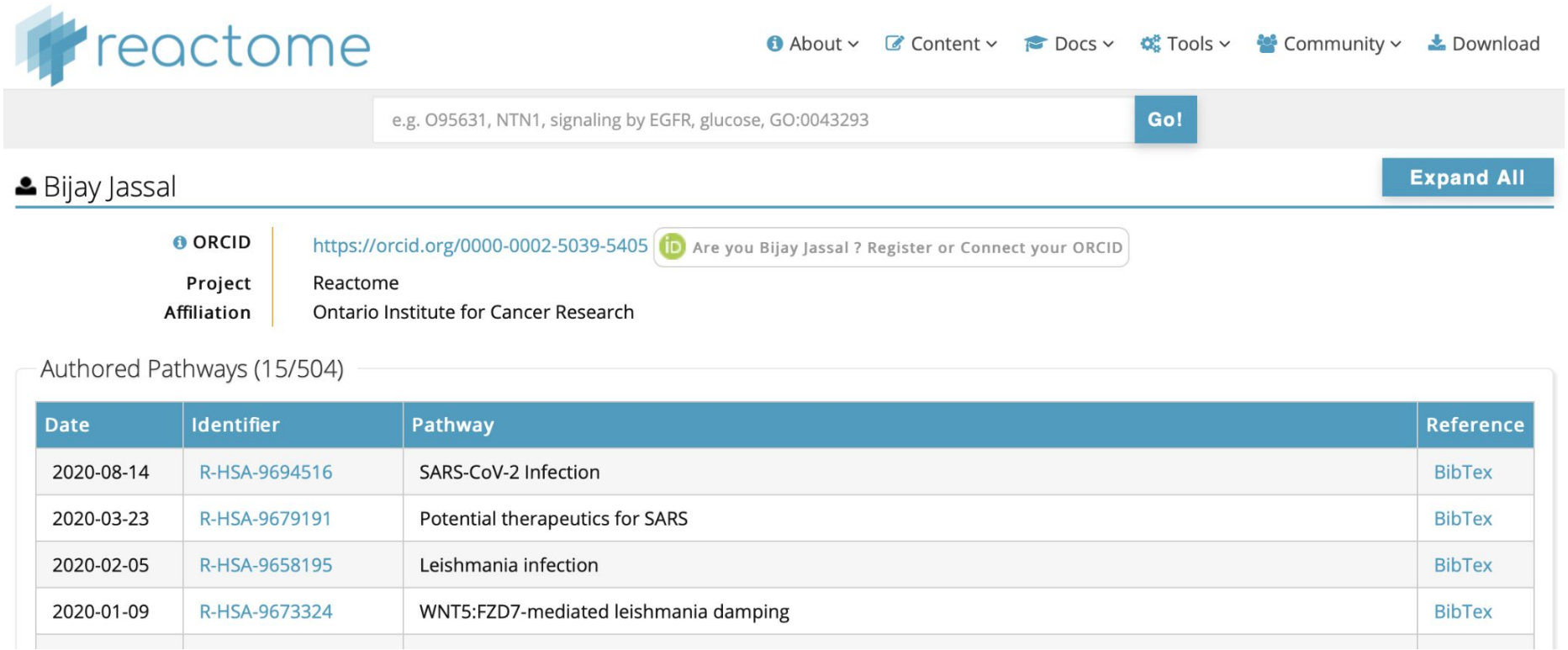
- 817 Reactome contributors since 2002
- Reactome has always recorded and displayed contributors' names as part of the record
- Any pathway revision is associated with a DOI and the DOI is associated with the name of all contributors
- DOIs make pathways citable

Crediting Contributors

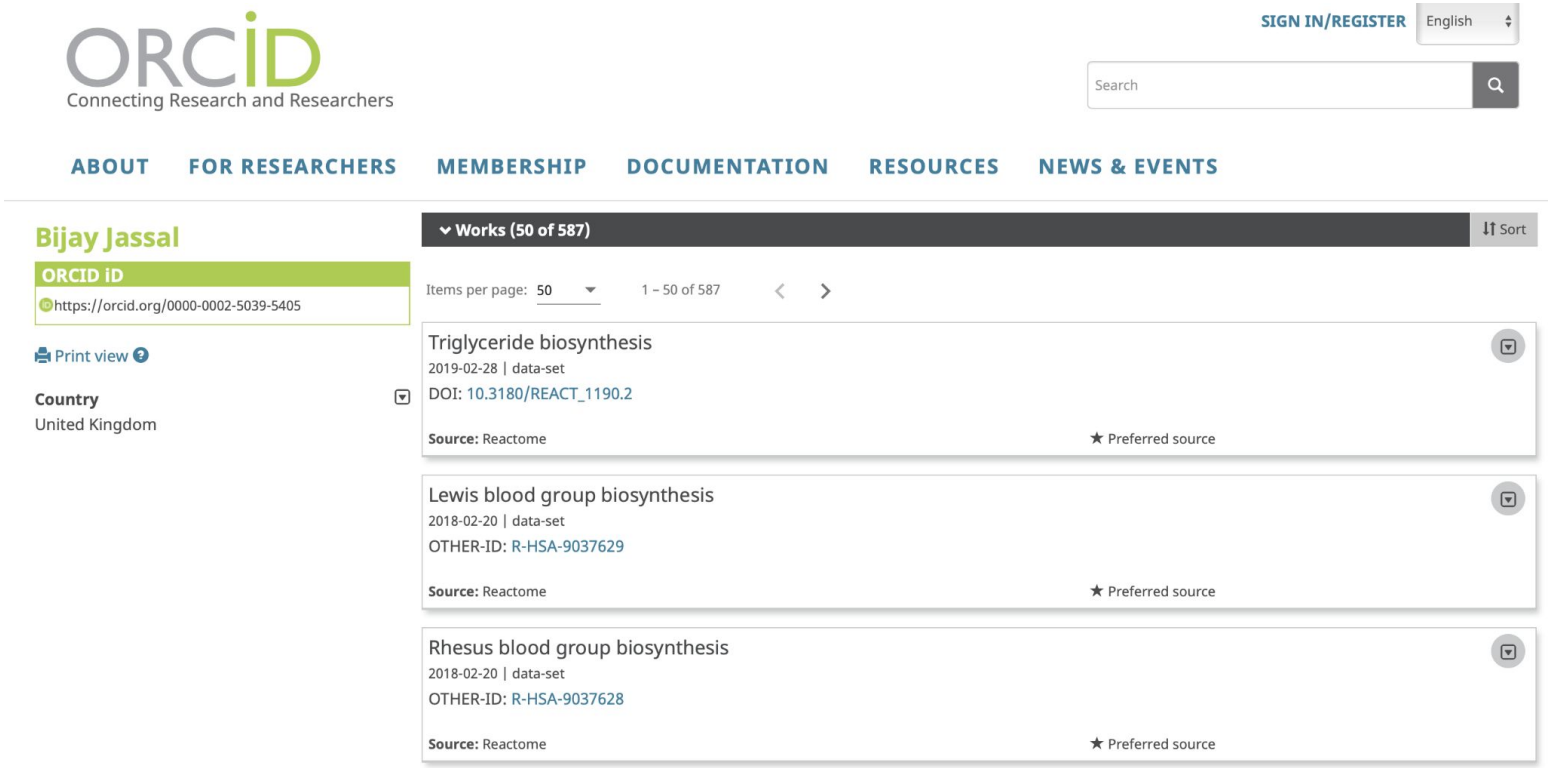
Reactome now is fully searchable by contributor names



Crediting Contributors



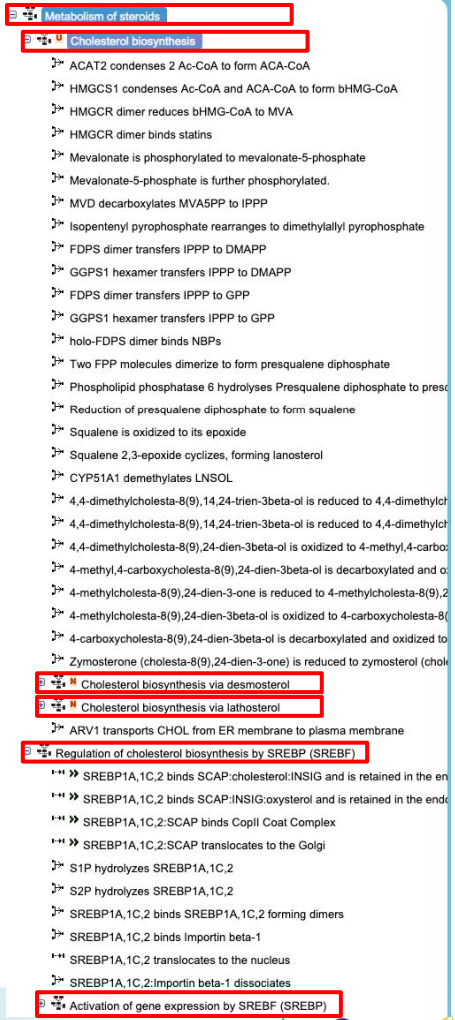
Crediting Contributors



Getting started: Creating a pathway outline

1 Metabolism of steroids **Pathway**

- 2 Cholesterol biosynthesis **Pathway**
  - 3 reactions **Reaction**
  - 3 Cholesterol biosynthesis via lathosterol **Pathway**
  - 3 Cholesterol biosynthesis via desmosterol **Pathway**
- 2 Regulation of cholesterol biosynthesis by SREBP (SREBF) **Pathway**
  - 3 reactions **Reaction**
  - 3 Activation of gene expression by SREBF (SREBP) **Pathway**



Expert knowledge capture - plan structure
Pathway hierarchy

Level description	Reactome stId	Pathway Name
top pathway	R-HSA-191273	Cholesterol biosynthesis
subpathway of Cholesterol biosynthesis	R-HSA-6807047	Cholesterol biosynthesis via desmosterol
subpathway of Cholesterol biosynthesis	R-HSA-6807062	Cholesterol biosynthesis via lathosterol
top pathway	R-HSA-1655829	Regulation of cholesterol biosynthesis by SREBP (SREBF)
subpathway Regulation of cholesterol biosynthesis by SREBP (S	R-HSA-2426168	Activation of gene expression by SREBF (SREBP)

D	E	F	G
Summation	**GO_BP**	**GO_BP name**	**PubMed refs**
Cholesterol is synthesized de novo from acetyl CoA. The overall synthetic p	GO:0006695	cholesterol biosynthetic process	11969204|12668600|3524618|1390320|16054061
The transformation of zymosterol into cholesterol can follow either of routes,	GO:0033489	cholesterol biosynthetic process via desmosterol	26114596
The transformation of zymosterol into cholesterol can follow either of routes,	GO:0033490	cholesterol biosynthetic process via lathosterol	14404284|26114596
Sterol regulatory element binding proteins (SREBPs, SREBFs) respond to l	GO:0045540	regulation of cholesterol biosynthetic process	18974038|19933148|15457548
After transiting to the nucleus SREBPs (SREBP1A/1C/2, SREBFs) bind sho	GO:0019216	regulation of lipid metabolic process	15589694|9748295|15457548

Expert knowledge capture - Reaction details

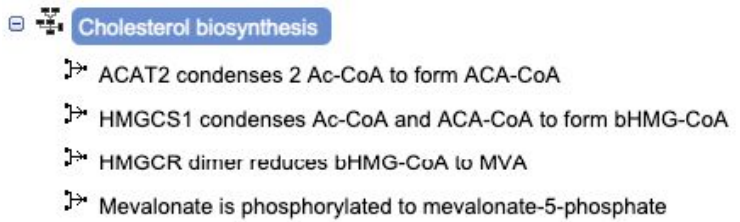
add new or revise?	Pathway Name	Reaction name	Reaction type	Tissue/Cell type	Summation	input n
Revise	top pathway	existing reaction 1	Reaction	(eg Human kidney/HEK-293)	revise rxn1	
New	subpathway 1	new reaction 2	BlackBoxEvent		describe rxn2	
New	subpathway 1	new reaction 3	Reaction		describe rxn3	
New	subpathway N	new reaction 4	Reaction		describe rxn4	

Reaction naming - catalysed reactions

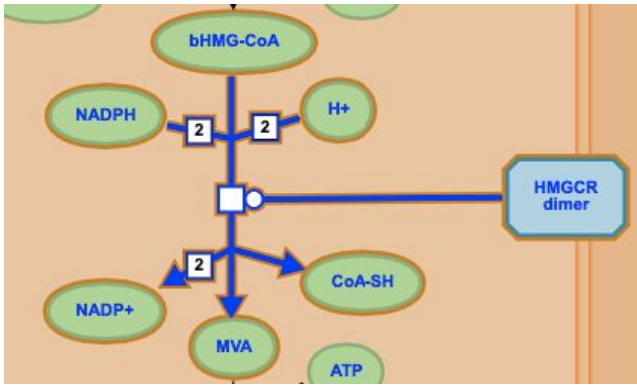
PROTEIN "does something" to SUBSTRATE (to PRODUCT)

HMGCR dimer reduces bHMG-CoA to MVA

3-hydroxy-3-methylglutaryl-CoA reductase dimer reduces beta-hydroxy-beta-methylglutaryl-CoA to mevalonate



"Dimeric 3-hydroxy-3-methylglutaryl-CoA reductase (HMGCR dimer) catalyzes the four-electron reduction of beta-hydroxy-beta-methylglutaryl-CoA (bHMG-CoA) to mevalonate (MVA). MVA concentrations in the cell are tightly controlled through the activity of HMGCR dimer, which is one of the most highly regulated enzymes in metabolism (Goldstein & Brown 1990)."



Reaction naming - non-catalysed reactions

SUBSTRATE 1 "binds" SUBSTRATE 2

TTPA binds alpha-TOH

TTPA + alpha-TOH => TTPA:alpha-TOH

A COMPLEX "dissociates"

VD3 dissociates from GC

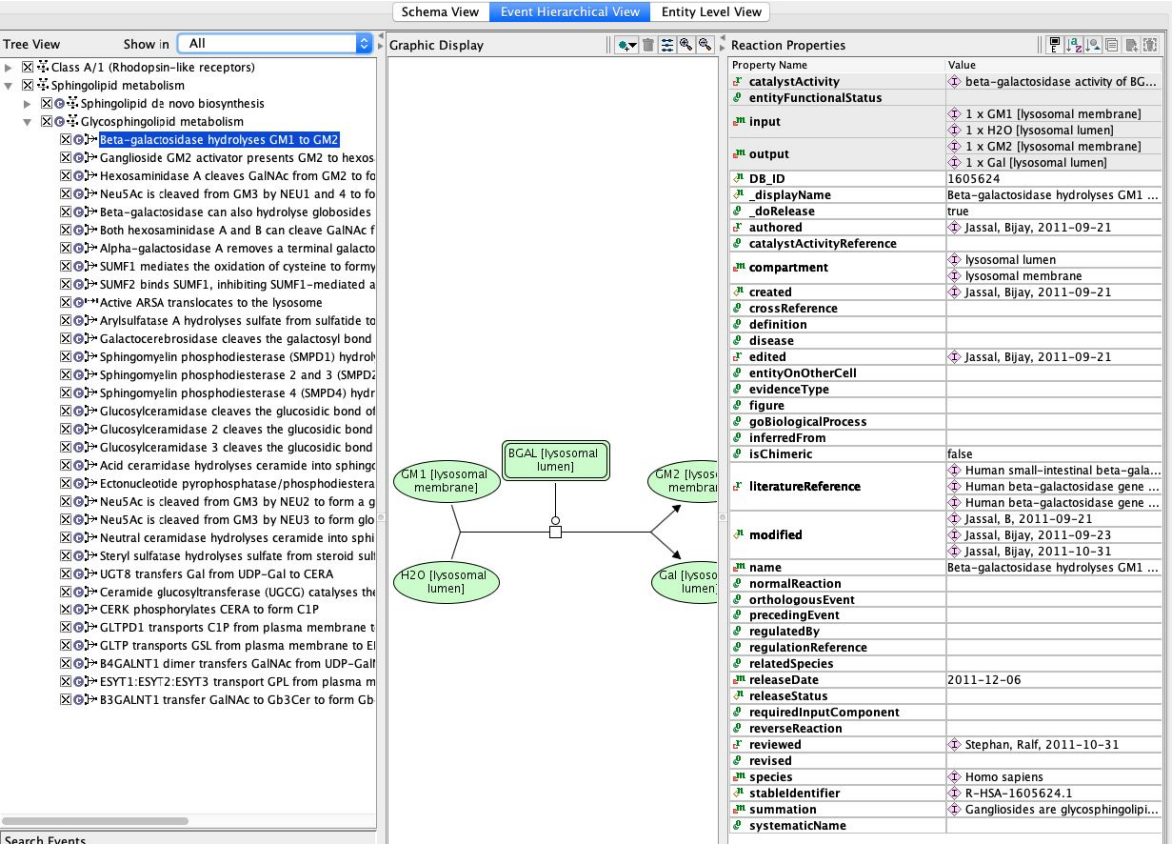
VD3:GC => VD3 + GC

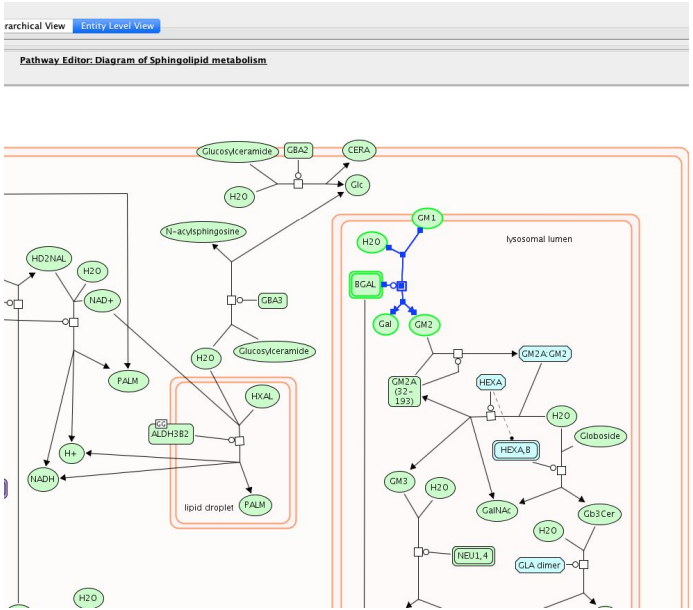
SUBSTRATE "translocates" from compartment 1 to compartment 2

VD3 translocates from ER membrane to extracellular region

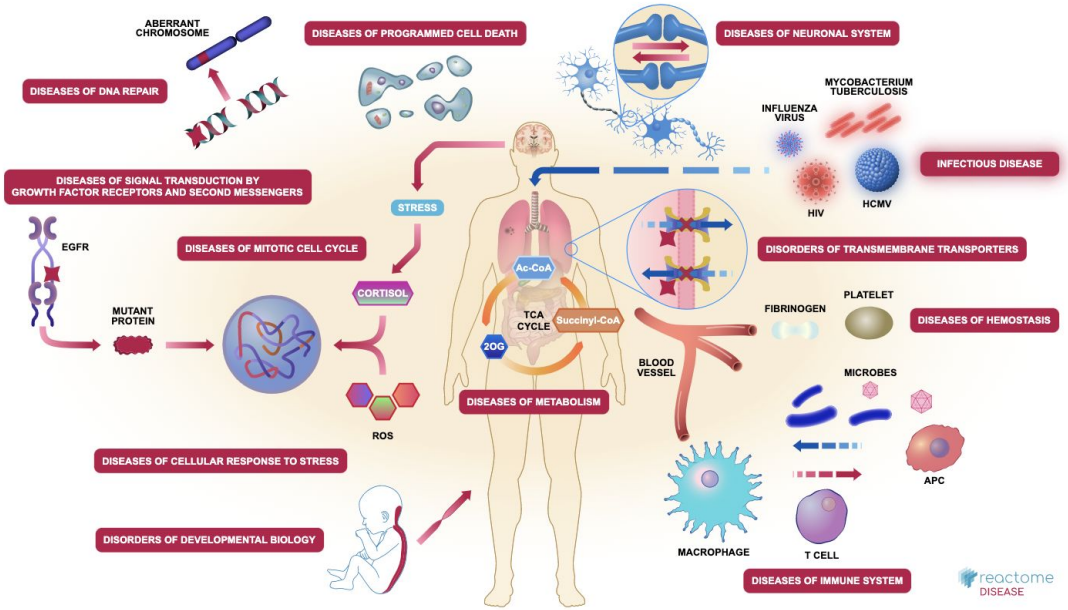
VD3 [ER membrane] => VD3 [extracellular region]

Data entry - Curator Tool





Disease pathways and processes in Reactome



- Reactome can annotate disease pathways such as cancer, metabolic, immune and infectious diseases
- Disease is overlaid normal pathways to show context
- Disease events may be loss-of-function or gain-of-function compared to the normal process
- Infection pathways are captured as host:pathogen interactions

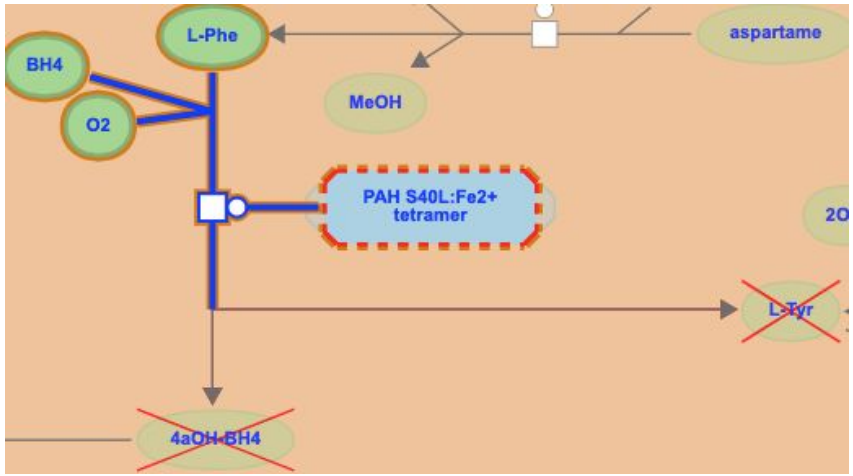
Expert knowledge for disease annotation

- Experts for normal pathways may already know details of disease events.
- Popular resources can be used as points of reference - OMIM, COSMIC (for cancer), Gene Reviews, Scriver encyclopedia of human genetic diseases
- For non-infectious disease pathways, the disease pathway should be constructed in the context of an annotated normal process
- Key/prevalent pathogenic variants are created for use in disease reactions

Disease variants

- Used in disease reactions that can be overlaid normal diagrams to show the effect of the disease

- Reactome can annotate missense, nonsense, fragment deletion, insertion and replacement variants as well as fusion proteins.

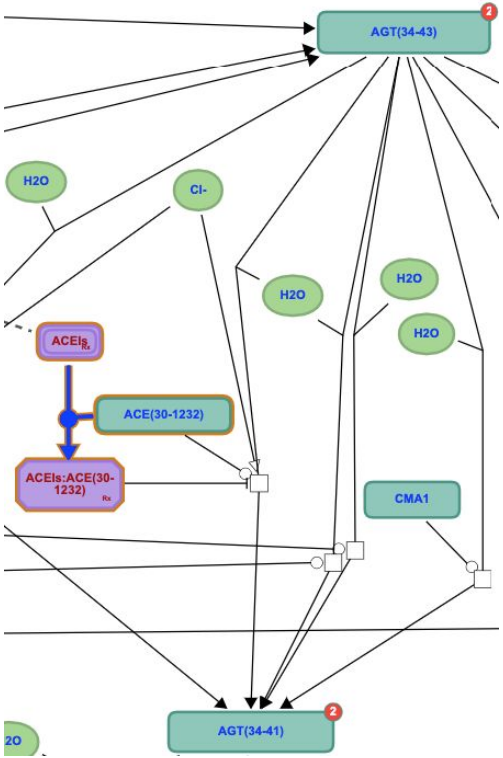


The disease phenylketonuria is caused by a missense mutation in the enzyme phenylalanine hydroxylase which normally converts the amino acid phenylalanine to tyrosine. The defective enzyme allows the build-up of phenylalanine to toxic levels which can cause intellectual disability and impair cognitive development.

Drug curation in Reactome

- The majority of drugs curated in Reactome “bind then regulate” a target
- Regulation can be described as either positive or negative (agonist/stimulatory/channel opener or antagonist/inhibitory/channel blocker drugs respectively)
- “Drug” class - chemical, protein and RNA subclasses
- Guide to Pharmacology IDs used as the main “reference” identifier for drugs
- Can group drugs into sets which have same MoA for a given target
- Any entity containing a drug will be filled in purple - visual impact in a diagram

Drug visualisation



Angiotensin-converting enzyme (ACE) mediates blood pressure by catalysing the cleavage of the hormone angiotensin to its vasoconstrictor form.

ACE inhibitors (ACEIs) bind to and inhibit the actions of ACE -> angiotensin cleavage is blocked -> ultimately results in the lowering of blood pressure.

Data Requirement Overview

	Level	Name	Summation	GO: BP	GO: MF	GO: CC	PMID	Input	Output	Catalyst	Regulator	Preceding Event	Mutation	Drug
Pathway														
Normal Reaction														
Disease Reaction														
Drug Reaction														

Mandatory

When available/relevant

Quick curation guide/reference

Overview of the Curation process (slides)

Full guide is available here for fine details.

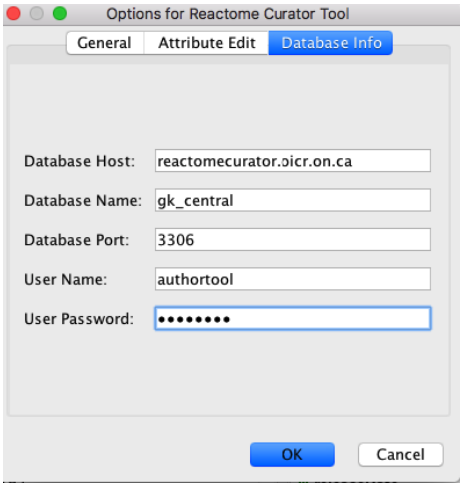
The Curator Tool – download, use, upload
Choosing a topic
The main data classes and attributes
Guide to naming entities and events
Drawing diagrams

The Curator Tool (CT)

The CT provides a standalone GUI for curation. It provides all the classes from which you can create instances to use in your pathway annotations. It also allows access to the central repository (gk_central) from which you can import instances already created to use in your annotation. Once your pathways are thoroughly curated and have all their must-have attributes filled, you can upload them to gk_central via synchronisation. This will check for duplicates before uploading.

Download the CT from the "Downloads" page here.
Install for your particular platform from here. Follow installation instructions on this page.

Open tool and go to "Preferences" to add connection details on the "Database info" tab.



Database Host: the server you upload new content to

Database Name: gk_central, the central data repository
User Name: ask
User Password: ask
With these username and password credentials, you are able to create and download content from gk_central but you can't upload material to it yet.

Choosing a topic

1. Focus on biology you know well.
2. Make sure reactions are well characterized at the molecular level with good HUMAN biochemical evidence that you should **NOT** include events based solely on microarray, proteomics or similar high throughput data
3. Check the Projects sheet to check you are not duplicating another curator's work

Main Data Classes for curation

For more information, see full data model glossary here.

Event
Pathway, Reaction, BlackBoxEvent, Depolymerization, Polymerization, FailedReaction

PhysicalEntity
EntityWithAccessionedSequence (EWAS), SimpleEntity, Complex, Polymer, EntitySet

ReferenceEntity
ReferenceGroup, ReferenceMolecule, ReferenceSequence, ReferenceDNASequence, ReferenceGeneProduct

CatalystActivity
Summation
Taxon (Species)
Publication – Book, LiteratureReference, URL
GO_BiologicalProcess
GO_CellularComponent
GO_MolecularFunction
Disease

Attributes

Data classes have a number of details that must be completed before being made public.

m - Mandatory, you must provide this information
r – Required, provide this is possible. Example is species, required for a set of proteins but not for a set of small molecules
o – Optional, not available for every class instance but may be a procedural requirement eg. inferredFrom should be filled in if evidence for an event comes from another species

Naming entities and events

For detailed background reading, please refer to this paper.

We have created a controlled vocabulary (cv) to consistently name entities and the events they participate in. Names are based on familiar biological names and terms, minimize ambiguity and support data mining.

Proteins (represented as EWASs in Reactome) are constructed around a 'core' Human Gene Organization (HUGO) Gene Nomenclature Committee (HGNC) approved gene symbol. The HGNC is the only worldwide authority that assigns standardized nomenclature to human genes, and as such is widely recognized as the authoritative source of names for genes. No such similar body exists for protein names; consequently, HGNC is the only practical source of short unambiguous names that can be used to name the translated products of those genes. Reactome obtains HGNC gene symbols indirectly from UniProt, an attribute associated with every peptide EWAS.

The CV also caters for coordinates and PTMs. For coordinates, to describe cleaved products, we use UniProt's 'Chain' feature. Example

Caspase-9 precursor, with peptide coordinates start:1 end:416 is named CASP9. The large and small subunits of caspase-9 are respectively named **CASP9(1–315)** and **CASP9(316–416).**

For PTMs, we add a prefix and hyphen to the gene symbol. Reactome defines PTMs with hasModifiedResidue annotations. These use terms obtained from the Protein Standards Initiative modifications (PSI-MOD) ontology. Examples

p-Y139-DAPP1 is DAPP1 phosphorylated on tyrosine-139.

p-Y150,S343,T346-WASF2 is WASF2 phosphorylated on tyrosine-150, serine-343 and threonine-346. Note that phosphorylations are ordered by coordinate.

p-Y55,S112,S121,Y227-SPRY2 —note that the ordering is by coordinate, not grouped by phosphorylation subtype.

p-Y-GAB2 is GAB2 phosphorylated on a tyrosine; the coordinate position of this tyrosine is unknown.

p-GLI3 is GLI3 phosphorylated; both the phosphorylation subtype and position are unknown.

When combinations of phosphorylation and another PTM occur, the phosphorylations are represented after everything else:

2xPalmC-MyrG-p-S1177-NOS3(2-1203) is NOS3 fragment 2-1203 with two palmitoylated cysteines, one myristoylated glycine and a phosphorylation on serine-1177.

Small molecules (represented by SimpleEntity in Reactome) have abbreviated names in most cases created by us. A big benefit of these 'short' names is reduced space and avoids clutter.

Complexes and sets also catered for.

Reactions

Reactome reaction names are created using free text, which can lead to ambiguity and inconsistency. Use of a controlled vocabulary for reaction names creates consistency and allows users to predict, search and understand these names more easily. They also facilitate bulk computational queries and data mining.

Reactions can be transformations (catalysed), binding, dissociation, transport, diffusion, (de)polymerisation and it is important to name these types consistently.

When a catalyst is involved, that should be the focus of the reaction name followed by the type of activity it performs followed by the substrate it acts upon.

CATALYST verb **SUBSTRATE** (to PRODUCT) – the product name is optional

PLB1 hydrolyses RPALM to atROL
PTGDS isomerizes PGH2 to PGD2

See naming doc for details of these types of reaction names.

When there is no catalyst involved in a reaction (binding, dissociation, translocation) then the focus is on the substrate. These reactions styles are

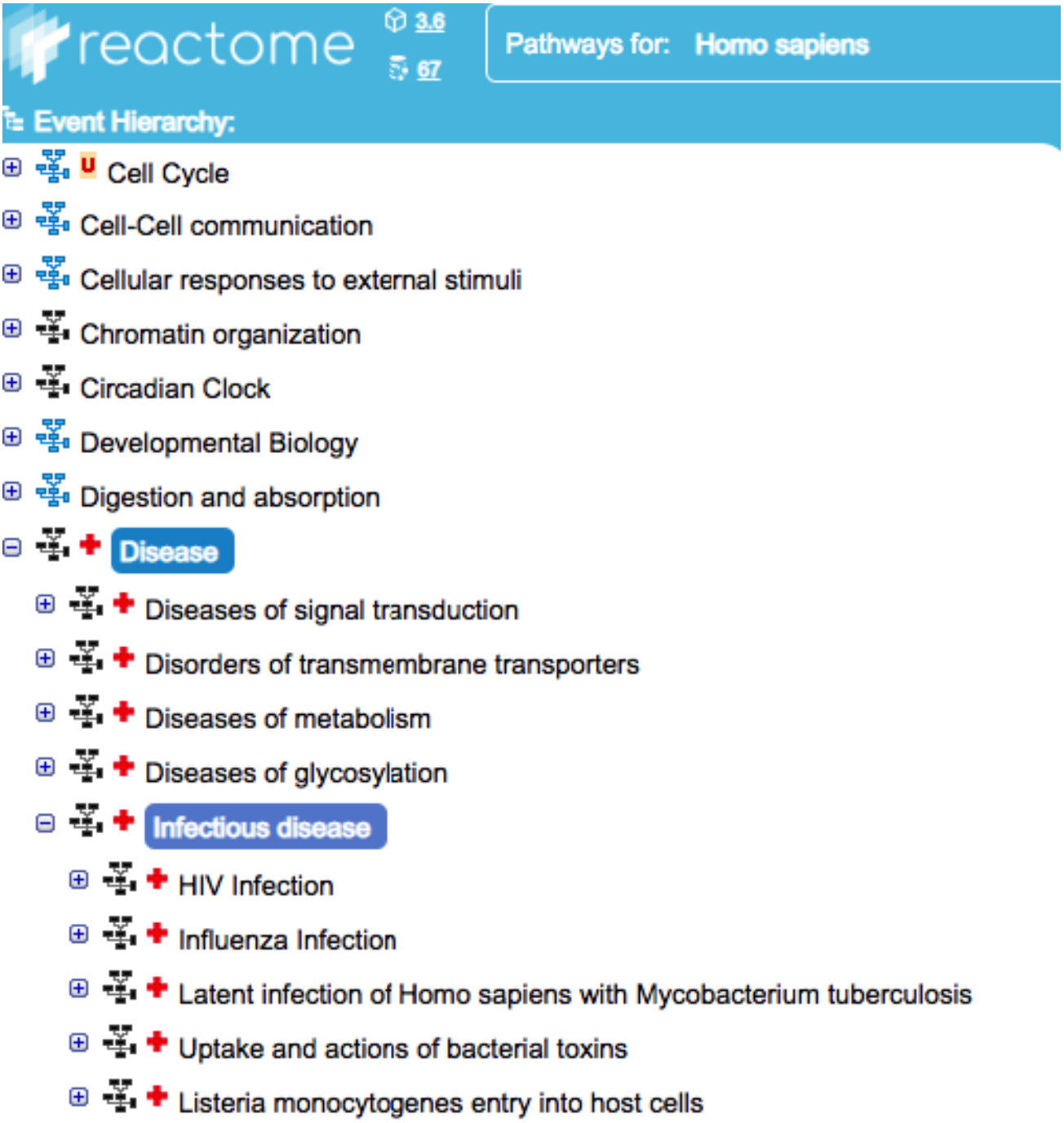
a **BINDS** b forming c
c **DISSOCIATES** TO a AND b
a **TRANSLOCATES FROM** [compartment x] **TO** [compartment y]

Drawing diagrams

To display the molecular details annotated in pathways and reactions, entity-level views (ELVs) are created using the Entity Level View tab in the CT. Refer to the curator guide for full details.

Annotations infectious disease

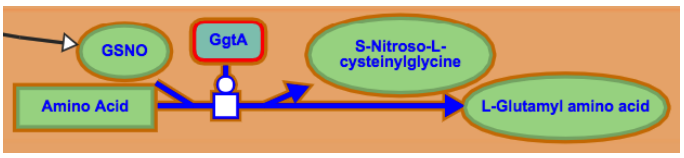
Disease pathways are housed under the ‘Disease’ chapter in Reactome’s hierarchy.



Reactions for your pathway will be created in the CT as below

Reaction Properties	
Property Name	Value
catalystActivity	glutathione hydrolase activity of GgtA [periplasmic space]
entityFunctionalStatus	
input	GSNO [periplasmic space]
	Amino Acid [periplasmic space]
output	L-Glutamyl amino acid [periplasmic space]
	S-Nitroso-L-cysteinylglycine [periplasmic space]
DB_ID	1222712
_displayName	Nitrosoglutathione gets cleaved to Cys(NO)-Gly
_doRelease	true
authored	Stephan, R, 2011-01-10
compartment	periplasmic space
created	Jassal, B, 2011-02-28
crossReference	
definition	
disease	tuberculosis
edited	Jassal, B, 2011-02-28
entityOnOtherCell	
evidenceType	
figure	
goBiologicalProcess	
inferredFrom	
isChimeric	false
literatureReference	Characterization of a glutathione metabolic mutant of My...
modified	Jassal, B, 2011-08-08
	Jassal, B, 2011-10-11
	Jassal, B, 2011-11-03
	Matthews, L, 2011-11-18
	Jassal, Bijay, 2011-11-23
	Jassal, B, 2012-02-08
	Jassal, B, 2012-02-10
	Jassal, Bijay, 2012-02-25
	Jassal, B, 2012-04-20
	Jassal, B, 2012-04-30
	Jassal, B, 2013-05-20
	Jassal, Bijay, 2014-07-14
name	Nitrosoglutathione gets cleaved to Cys(NO)-Gly
normalReaction	
orthologousEvent	
precedingEvent	
regulatedBy	
relatedSpecies	Mycobacterium tuberculosis
releaseDate	2012-06-12
releaseStatus	
requiredInputComponent	
reverseReaction	
reviewed	Warner, D, 2012-04-30
revised	
species	Homo sapiens
stableIdentifier	R-HSA-1222712.1
summation	Most gamma-glutamyl transpeptidases (GGT) cleave bot...
systematicName	

The “species” slot on the event identifies the environment where the reaction is happening. This should almost always be human for human Reactome curation. The pathogen(s) would be listed in the “relatedSpecies”. The ELV you need to create would be the same as a normal ELV but with pathogenic proteins in them. When drawing infection reactions that involve pathogenic proteins, the node and edges need to be coloured red but leave human proteins in their original colours



V90_Sept 2024

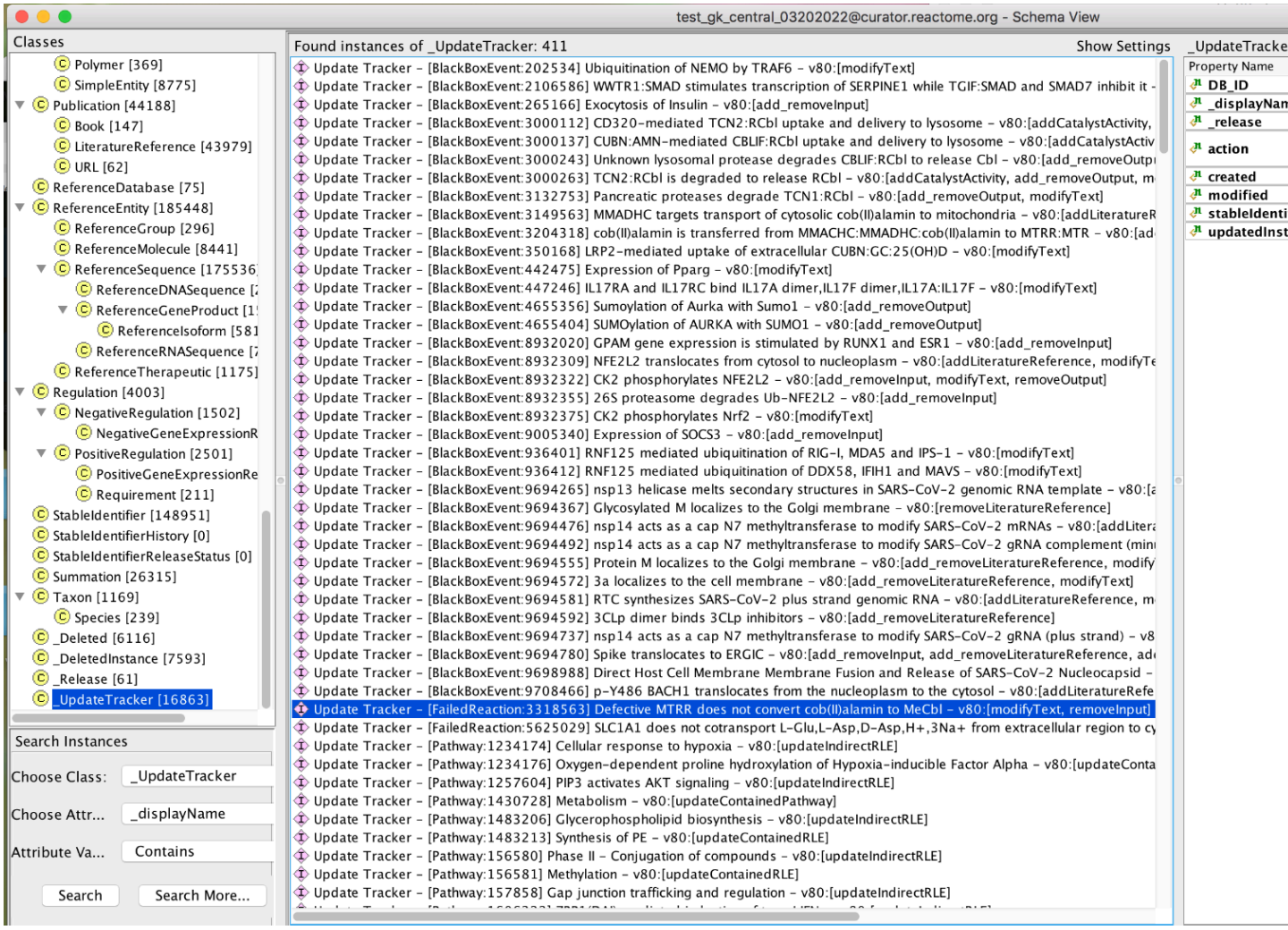
- updated links to the V90 Curator Guide and V90 Data Model Glossary

C3. Evidence confidence

Evidence	Comment	Examples	% Confidence
Two experimental papers from independent groups demonstrating the same mechanism	Gold standard, can be included as reaction unless reviewer expresses doubt.	879925	100
One experimental paper and external expert author indicates that they believe the event is proven by this	Can be included as reaction unless reviewer expresses doubt	1067640	90
One experimental paper with more than one experimental method determining the mechanism of the event	Not sufficient		
One experimental paper, reported as established in a review paper	Can be included unless Reactome reviewer expresses doubt		70
Two instances of correlated observations but no proof of mechanism, e.g. addition of recombinant kinase A, observation that protein X is phosphorylated, or addition of novel protein ligand leads to upregulation of signal, but no demonstration of ligand binding or mechanism that leads to signal	Black-box	8950269	50
Ligand binds receptor leading to some kind of signal. Ligand and/or receptor clearly belong to known family, other family members have well characterized mechanism of receptor activation and signaling. Inference made by curator that mechanism of signaling is the same for other single family member	Reaction for demonstrated molecular events e.g. ligand binding, specific phosphorylations or other activating events. Black box events should be included to represent events equivalent to the full series of molecular events demonstrated for the well characterized system.		80
Molecular details obtained for one member of a set used to infer that the same generic mechanism exists for the other set members, where set represents a family.	This is the intended usage for candidate sets and such sets can be used in reactions. Also sufficient for defined sets used in reactions that are part of a generic pathway such as GPCR signaling, where multiple literature sources suggest without question that the actions of all members of the family are mechanistically similar.Would not be sufficient for defined set used in a non-generic pathway event.	379048	80
Two reports of gene expression stimulated by transcription factor binding to specific promoter region	Reaction		100
One reports of gene/ expression stimulated by transcription factor binding to specific promoter region	Not sufficient		
Two reports of protein expression regulated following activation of a specific factor but no evidence of TF binding to gene promoter	Black box with regulation		60
Two reports of gene/protein expression regulation following activation of a specific transcription factor, but no evidence of binding to promoters	Black box with regulation		60
One reports of gene/protein expression regulation following activation of a specific transcription factor, but no evidence of binding to promoters	Not sufficient		
Dissociation events where dissociation is inferred from subsequent involvement in another event such as binding or translocation	Black box		70
Evidence that describes actions of a complex component where evidence is in vitro, not experimentally demonstrating activity while part of a complex	This is typically considered sufficient for a reaction as long as complex formation has been demonstrated (would be a preceding event).	446694	90
Single experiment evidence obtained with relative used to infer similar mechanism	Not sufficient	157644	70
Evidence of translocation but mechanism not identified	Black box unless small molecule capable of passive movement		
Event inferred from model organism			score of human event - 20
Note - in all cases, if publications refute the evidence, the event is contraversial and should not be inlcuded			

C4. Update Tracker class_use and logic

The UpdateTracker Class instances describe revisions to events that have occurred between releases. These instances are created automatically during database slicing and require a comparison between the current slice and previous slice databases.



The following changes constitute a revision in an RLE or Pathway.

I. RLE:

An RLE is considered revised if:

-A Catalyst added/removed/changed

-A regulator added/removed/changed

-A change is made in inputs/outputs

-A significant change in summation (?) (curator discretion(?))

II. Pathway:

A pathway is considered revised if::

1. An immediate child RLE is added/removed
2. An immediate child RLE revised
    -A Catalyst added/removed/changed
    -A regulator added/removed/changed
    -A change is made in inputs/outputs
    -A significant change in summation (?) (curator discretion(?))
3. An immediate child Pathway added/removed
4. An immediate child Pathway is revised as in **#1, #2, #3**

UpdateTracker instances register changes as “action” attributes.
The possible action types are:

"**add**"
"**remove**"
"**add_remove**" (i.e. both addition and removal on a multi-valued **instance** attribute)
"**modify**" (i.e. changed value on a multi-valued **text** attribute)
"**update**" (i.e. a contained child event in a pathway has changed)

They are paired with specifically what changed:

For Pathways:
"**hasEvent**"
"**literatureReference**"
"**summation**" (the actual instances in the slot)
"**text**" (the summation text content rather than the summation instance)
"**updateContainedRLE**" (a direct RLE child has a change)
"**updateIndirectRLE**" (a descendant, non-immediate RLE in a sub-pathway for a child has a change)
"**updateContainedPathway**" (a direct pathway child has a change)

*NOTE: If "updateIndirectRLE" is indicated for a pathway, this means one or more of its sub-pathways should have "updateContainedPathway".

For ReactionlikeEvents:
"**input**"
"**output**"

"**catalystActivity**"
"**literatureReference**"
"**summation**" (the actual instances in the slot)
"**text**" (the summation text content rather than the summation instance)

**Note; If an event has the same type of action occurs more than once in a release cycle, it is only listed once but multiple actions will be recorded for a single release if more than one type of action occurred (for example, "addInput", "addOutput" in an update tracker for an RLE can occur together but if 2 inputs were added "addInput" would only show up once).

PhysicalEntity Update types:

UpdateTracker instances register changes as “action” attributes.
The possible action types are:

"**add**"
"**remove**"
"**add_remove**" (i.e. both addition and removal on a multi-valued **instance** attribute)
"**modify**" (i.e. changed value on a multi-valued **text** attribute)
"**update**" (i.e. a contained child event in a pathway has changed)

The update types checked for all physical entities are:
"cellType",
"compartment",
"disease",
"literatureReference",
"name",
"summation",
"systematic name"

In addition, the following are checked for each type of Physical Entities:
Cell: "markerReference", "organ", "proteinMarker", "RNAMarker", "species", "tissue", "tissueLayer"
Complex: "hasComponent", "entityOnOtherCell", "species", "stoichiometryKnown"
Drug: "species"
EWAS: "referenceEntity", "endCoordinate", "startCoordinate", "hasModifiedResidue"

GEE: "inferredFrom", "species"
Polymer: "maxUnitCount", "minUnitCount", "repeatedUnit"
SimpleEntity: "species"
Set: "hasCandidate", "hasMember", "species"

C5. Data Model Glossary

Overview

Reactome is a slot-and-frame knowledge representation of a network of molecular transformations mediated by gene products. Classes (frames) are categories like reaction and protein. Related classes are grouped into superclasses to allow inheritance of shared attributes. For example, the "PhysicalEntity" superclass has child classes such as Complex and GenomeEncodedEntity. Classes have slots (attributes) whose values define properties of instances (e.g., the names and numbers of copies of the molecules that make up a complex). "Defining" attributes are used in the data model to distinguish instances of a class, to ensure that identical information, e.g., a water molecule localized to the cytosol, or a reaction by which a particular molecule from extracellular space is translocated into the cytosol mediated by a particular transport protein, is represented exactly once in Reactome. Many attributes of a class instance are instances of other classes. For example, a **Person** instance has a *surname* attribute that is a text string not taken from anyplace else but can also have an *affiliation* attribute that is an instance of the **affiliation** class, a class of *names* and *addresses* of institutions. In this glossary, **boldface** type indicates use of a word or phrase as a class name (the **CatalystActivity** of a particular enzyme) and italic type indicates use as an attribute (the *catalystActivity* that mediates a chemical transformation reaction). Words that can be either class names or attributes are defined under “Reactome Classes”; ones that are only attributes are defined under “Reactome attributes”.

Glossary of Key Reactome Classes

AbstractModifiedResidue [superclass]

Used to maintain logical integrity of data model, not used for manual annotation. The superclass in the Reactome data model from which the GeneticallyModifiedResidue, TranscriptionalModification, and TranslationalModification classes inherit their shared attributes. Describes a modification to reference sequence. The three subclasses of modifications are GeneticallyModifiedResidue, TranscriptionalModification, TranslationalModification.

GeneticallyModifiedResidue [superclass]

Used to maintain logical integrity of data model, not used for manual annotation. Subclass of abstractModifiedResidue from which the FragmentModification and ReplacedResidue classes inherit their shared attributes. Describes a protein sequence change resulting from a genetic modification

FragmentModification [superclass]

Used to maintain logical integrity of data model, not used for manual annotation. Children of this class (FragmentDeletionModification, FragmentInsertionModification and FragmentReplacedModification) are used to capture the deletion, insertion or replacement of consecutive amino acid residues in an EWAS in a disease variant of a gene product.

FragmentDeletionModification

The deletion of a continuous fragment of an EntityWithAccessionedSequence, e.g. deletion of amino acid residues 30 to 297 in the EGFR mutant protein EGFRvIII in glioblastoma. Attributes of the FragmentDeletionModification instance are the *referenceSequence* (UniProt P00533 EGFR in the case of EGFRvIII), the *startPositionInReferenceSequence*, which is the first amino acid of the deleted fragment (30 in the case of EGFRvIII), and the *endPositionInReferenceSequence*, which equals the last amino acid of the deleted fragment (297 in the case of EGFRvIII).

FragmentInsertionModification

Used to describe the modification of an EntityWithAccessionedSequence through insertion of a continuous fragment. This class is used to annotate variants EWASs containing duplications or fusions. A duplication is defined as the insertion of a stretch of WT amino-acid(s) that is the direct repeat of the amino-acid(s) immediately 5’ to the inserted sequence. For duplications, the *referenceSequence* specified in the FragmentInsertionModification record is the same as the *referenceEntity* specified in the EWAS record.
Fusions are the result of the insertion of a continuous stretch of amino-acids derived from a second EWAS at a specified residue in a first EWAS. For fusions, the *referenceSequence* specified in the FragmentInsertionModification record is the *referenceEntity* for the C-terminal fusion partner, and is different from the *referenceEntity* specified in the EWAS record, which represents the N-terminal partner.

FragmentReplacedModification

The modification of an EntityWithAccessionedSequence through the replacement of the WT amino-acid sequence with a specified alternate sequence of amino-acids as the result of an insertion, fusion, indel, frameshift or extension of the WT sequence. Details of these annotations vary depending on the type of replacement, and are outlined in the Curator Guide. Unlike FragmentDeletionModification and FragmentInsertionModification, FragmentReplacementModification requires that the curator explicitly annotate the single letter amino-acid sequence that arises in the variant protein. Note also that, by Reactome convention, insertions are captured by FragmentReplacedModification, not FragmentInsertionModification. In the FragmentReplaceModification record for an insertion, the *alteredAminoAcidFragment* includes both the flanking WT 5’- and 3’ amino-acids as well as the inserted amino-acids. For example, in the variant EGFR D770_N771insNPG, amino acids NPG are inserted between

D770 and N771, and the specified *alteredAminoAcidFragment* in the FragmentReplacedModification is DNPGN.

ReplacedResidue

A subclass of GeneticallyModifiedResidue, ReplacedResidue describes the replacement of an amino acid residue of a UniProt *referenceSequence* with a different amino acid residue or a stop codon as the result of a mutation, germline or somatic, in the corresponding gene. Class attributes identify the *coordinate* of the affected residue in the protein sequence, the identity of the residue specified in the UniProt sequence (first *psiMOD* attribute value), and the identity of the replacement residue (second *psiMOD* attribute value).

NonsenseMutation

A subclass of the **ReplacedResidue** class, **NonsenseMutation** describes the *coordinate* at which a protein *referenceSequence* is truncated due to a nonsense mutation. In the NonsenseMutation record, a single *psiMod* is specified, indicating as an amino acid removal the amino acid that is altered to generate a stop codon.

TranscriptionalModification [superclass]

Used to maintain logical integrity of data model, not used for manual annotation. Subclass of AbstractModifiedResidue. Describes an RNA sequence modification.

ModifiedNucleotide

Describes the location and chemical nature of a nucleotide modification in an RNA sequence.

TranslationalModification [superclass]

Used to maintain logical integrity of data model, not used for manual annotation. Subclass of **AbstractModifiedResidue**. Includes children **CrosslinkedResidue**, **GroupModifiedResidue** and **ModifiedResidue**. Describes a post translational chemical modification to a protein. Modifications are cross referenced to the psi-MOD ontology of terms that describe protein chemical modifications.

CrosslinkedResidue [superclass]

Used to maintain logical integrity of data model, not used for manual annotation. Subclass of **TranslationalModification**. Two subclasses, **Inter**- and **IntraChainCrosslinkedResidue**, are used to annotate disulfide, thioester or other protein crosslinks, including to ie sumoyl or ubiquitin groups.

InterChainCrosslinkedResidue

Subclass of **CrosslinkedResidue**. Used to annotate cross-linking between two different peptide chains.

IntraChainCrosslinkedResidue

Subclass of **CrosslinkedResidue**. Used to annotate cross-linking between amino acid residues within the same peptide chain. **IntraChainCrosslinkedResidue** instances are made in pairs, one associated with the EWAS record of each of the cross-linked proteins, and the **IntraChainCrosslinkedResidue** records are associated with each other through their *equivalentTo* attribute.

GroupModifiedResidue

Subclass of **TranslationalModification**. Describes the modification of an amino acid residue in protein with a chemical entity that cannot be specified in atomic detail, e.g., the attachment of a dextrin or glycogen moiety to a tyrosine side chain in the protein glycogenin. Such incompletely specified chemical entities are beyond the scope of the psi-MOD ontology but are available in the ChEBI ontology. The psiMod attribute would take a value such as MOD:00166 O4'-glucosyl-L-tyrosine, which describes the linkage between the protein and the modifying group. The modifying group is then specified with a ReferenceGroup instance (ChEBI terms) such as CHEBI:28912 "limit dextrin", as in the example: limit dextrins on L-Tyrosine [ChEBI:17895] 194 of UniProt:P46976 GYG1.

ModifiedResidue

Subclass of TranslationalModification. Used to describe the translational or post-translational modification of an amino-acid, such as phosphorylation, gamma-carboxylation etc. The **ModifiedResidue** record defines the associated EWAS reference sequence along with the coordinate of the modified amino-acid and the nature of the modification, taking terms from psiMOD. **ModifiedResidue** class is used when the chemical natures of the modifying group can be specified in atomic detail with a psiMod term. This is in contrast to **GroupModifiedResidue**, which is used when the modification can not be completely described by a term from psi-MOD.

Affiliation

The name and address of an institution, e.g. Cold Spring Harbor Laboratory / 1 Bungtown Road, Cold Spring Harbor NY USA. Any **Person** instance can have an *affiliation* attribute, but at present this information is recorded only for Reactome authors, reviewers, and curators.

CatalystActivity

Describes the catalyst in a **ReactionLikeEvent**: associates one *physicalEntity* with one *GO_molecularFunction*., e.g., cysteine-type endopeptidase activity of active caspase-8. If the

physicalEntity is a **Complex** and a component of the complex mediates the molecularFunction, that component should be identified as the *activeUnit* of the **CatalystActivity**.

If a **PhysicalEntity** can enable multiple **MolecularFunctions,** a separate **CatalystActivity** instance is created for each..

ControlReference [superclass]

Used to maintain logical integrity of data model, not used for manual annotation. Associates one or more **Publications** with an instance (e.g., catalystActivity) of another class.

CatalystActivityReference

Associates a **CatalystActivity** instance with one or more **Publications**.

MarkerReference

Associates an RNA or protein **EntityWithAccessionedSequence** that has been identified as a *proteinMarker* or *RNAmarker* of a **Cell** with one or more **Publications**. *MarkerReference* is an attribute of the Cell class of PhysicalEntity.

RegulationReference

Associates any kind of **Regulation** instance with one or more **Publications**.

ControlledVocabulary [superclass]

Used to maintain logical integrity of data model, not used for manual annotation. Lists of words that can be used in attributes (slot values) in specific classes, optionally connected to one another with is_a relationships, created by Reactome (as opposed to ones imported from external reference resources like the Gene Ontology, ChEBI, or NCBI Taxonomy)
Has three subclasses **DeletedControlledVocabulary** (for use in “reason” attribute of _Deleted instances), **NegativePrecedingEventReason** (for use in “reason” attribute of NegativePrecedingEvent instances) and ReactionType (for use in *reactionType* attribute of **ReactionLikeEvent**).

NegativePrecedingEvent

Used to flag an **Event** that should not be annotated as a value of another **Event**’s *precedingEvent* attribute.

NegativePrecedingEventReason

A subclass of ControlledVocabulary used to indicate why a computationally inferred causal connection (*precedingEvent*) between two **ReactionLikeEvents** was manually rejected by a curator (e.g. Abundant Molecules).

ReactionType

Terms used as attributes to classify **ReactionLikeEvents** in ways that support data mining and classification, e.g., in connection with the conversion of Reactome events into GO-CAM models (under development).

DrugActionType

Terms used as attributes to classify **ReactionLikeEvent** involving **drugs** to indicate the mode of action of the drug in the **ReactionLikeEvent**.

DatabaseObject [superclass]

Used to maintain logical integrity of data model, not used for manual annotation. The root term of the Reactome data model; parent to all classes of instances; not edited manually.

DatabaseIdentifier

Refers to an external database identifier. An instance of this class is used in for the *databaseIdentifier* attribute as an optional *crossReference* attribute of a Reactome instance to associate it with the representation of that instance in an external *ReferenceDatabase,* in addition to the required crossReferences to external referenceDatabases*.* . E.g., the Reactome **SimpleEntity** instance PQQ (pyroquinoline quinone), crossReferenced to ReferenceMolecule:29568 pyrroloquinoline quinone [ChEBI:18315] has a *crossReference* attribute to the *databaseIdentifier* COMPOUND:C00113 that links the Reactome instance to the record for this molecule maintained by the KEGG “compound” database.

_Deleted

Created when an instance is deleted from the database. Includes attributes for the deletedInstanceDB_ID, replacementInstanceDB_IDs (if relevant), reason that the instance was deleted (see DeletedControlledVocabulary below for reason definitions), as well as optional curatorComment about the deletion. Additional information about the deleted Instance is provided in the _DeletedInstance attribute.

DeletedControlledVocabulary

An attribute of the _Deleted instance (which is created when a curator opts to delete an instance). Describes the reason an instance was deleted from the database. Reasons are: Duplicated (instance was a duplicate of an existing instance), Killed not replaced (instance was found to be invalid and removed without replacement), Merged (an instances was merged with another instance), Obsoleted (no longer in use: see specified replacement or comment), Split (the instance was split into multiple instances).

_DeletedInstance

Attribute of _Deleted that includes the class of the deleted instance, deletedInstanceDB_ID, deletedInstanceStableIdentifier, name of the deleted instance, and species of the deleted instance (if relevant).

Disease

Disease encompasses terms that describe impairments of normal functions caused by pathogenic genetic mutations or environmental factors, such as microbial pathogens from an ExternalOntology, the Disease Ontology (DO) (https://disease-ontology.org/). Terms from the Disease class are used to populate the disease attribute which identifies any abnormal physicalEntity (**GenomeEncodedEntitiy**, **Complexy**, EntitySets) and any abnormal event (**ReactionLikeEvent**, **Pathway)**, and specifies the disease(s) with which these abnormal instances are associated. **Disease** instances include *identifier* and *name* attributes taken from DO.

EntityFunctionalStatus

EntityFunctionalStatus is used to mark a disease ReactionLikeEvent as either a gain- or a loss-of-function, and to describe the underlying reason for the change in behavior. The EntityFunctionalStatus has three attributes, *diseaseEntity*, the variant **EWAS** or **Complex** or *set* of variants, responsible for the phenotype, *normalEntity*, the wild-type instance corresponding to the disease **PhysicalEntity** and *functionalStatus*, that relates the molecular change to the functional one.

FunctionalStatus

Class used to describe the molecular basis of a change in function; combines a term from the *functionalStatusType* controlled vocabulary maintained by Reactome that identifies alterations in gene function (e.g., gain-of-function) and a *structuralVariant* term from the externalOntology Sequence Ontology that describes the nature of the genetic change (e.g., copy number gain).

Class **FunctionalStatus** is used to populate the *functionalStatus* attribute of **EntityFunctionalStatus**

Event [superclass]

Used to maintain logical integrity of data model; not used for manual annotation. Any transformation of input entities to output entities in one step (**ReactionlikeEvent**) or two or more steps (**Pathway**).

Pathway

A sequence of two or more causally connected **ReactionlikeEvents** and/or other **Pathways**, identified as *hasEvent* attributes of the pathway being annotated.

CellLineagePath

A **CellLineagePath**, organizationally similar to the **Pathway** subclass, is composed of other **CellLineagePath** instances or **CellDevelopmentSteps** as subevents that are connected through shared input/output Cells.
Both Cell Development Step and Cell Lineage Path are associated with a Gene Ontology (GO) biological process term and with a “tissue” term from UBERON, an ExternalOntology of Anatomy terms.

ReactionLikeEvent [superclass]

Used to maintain logical integrity of data model; not used for manual annotation. **ReactionLikeEvent** is a superclass that is used to organize other concrete reaction types, such as **Reaction**, **Polymerization** and **BlackBoxEvent.** A molecular process in which one or more input *physicalEntities* are transformed in a single step into output *physicalEntities*, optionally mediated by a *catalystActivity* and subject to *regulation*. Kinds of transformation include chemical reactions, translocation of molecules from one **compartment** to another, association of molecules to form **complexes** and **polymers**, dissociation of **complexes** and polymers, and developmental/differentiation relationships among Cell instances.

BlackBoxEvent

This class can be used to describe an incompletely specified **ReactionLikeEvent** or series of RLEs. A BlackBoxEvent for a reaction is for one that is known to happen but whose molecular details are incompletely known, e.g., a partially purified protein catalyzes a reaction so no EWAS can be identified, or degradation of a protein yields unspecified polypeptide products. BBEs may also be used to represent a multistep process (such as the transcription and

translation of a gene into a protein) when the details don't need to be specified in full. A fully annotated sequence of events for a BBE may be depicted in full elsewhere in Reactome.

CellDevelopmentStep

A **CellDevelopmentStep** is a **ReactionlikeEvent** that contains **Cell** instances as its inputs and outputs. The input attribute represents the cell of origin, the starting point of this development event, and the output attribute represents the destination cell type, the outcome of the event. Additional CellDevelopmentStep attributes include regulators (molecules promoting or inhibiting the step) and required input components (input cell biomarkers required for the action of regulators). **CellDevelopmentStep** is also associated with a relevant Gene Ontology (GO) biological process term and with a “tissue” term from the UBERON, an ExternalOntology of Anatomy terms.

Depolymerisation

A subclass of **ReactionlikeEvent** that is used to annotate depolymerisation and follows the pattern: Polymer -> Polymer + Unit (reverse situation of Polymerisation).

FailedReaction

A **FailedReaction** instance is a step in a disease process that is directly affected by a loss-of-function (LoF) mutation (germline or somatic). This type of disease event has its normal **ReactionLikeEvent** counterpart (the reaction mediated by the un-mutated gene product) specified at its normalReaction attribute and is represented as having *inputs* (an abnormal physicalEntity plus any wild-type entities that are inputs in the normal reaction), but no *outputs*.

FailedReaction instances are labeled with disease term(s) that are used to populate the disease attribute of the associated abnormal **PhysicalEntity** (**GenomeEncodedEntitiy**, **Complex** or **EntitySet**).

.

Polymerisation

A subclass of reactionlikeEvent in which two or more identical molecules or complexes are assembled into a polymer.

Reaction

The transformation of input physical entities into output ones in a single step. Transformations include chemical changes, as in metabolism, the transport of an entity between locations, the association of entities to form a complex, and the dissociation of complexes. These transformations can be catalyzed and regulated.

EvidenceType

EvidenceType is a subclass of ExternalOntology, manually populated with Gene Ontology evidence code terms. Used only to flag inferred orthologous events created computationally as part of the release process (IEA, inferred by electronic annotation) and not for human manual annotation by curators.

ExternalOntology

An ExternalOntology is one developed and maintained outside of Reactome and used as-is by Reactome to align our annotation with the community vocabularies and logical relationships incorporated into these ontologies. This is a superclass and used for generalizing sharing attributes among multiple types of external ontologies. This class should not be used for instance creation.

Anatomy

Anatomy encompasses terms representing body parts, organs and tissues from an ExternalOntology, Uberon (https://obophenotype.github.io/uberon/). Anatomy terms are used to populate organ, tissue, and tissueLayer attributes of the Cell subclass of PhysicalEntity and tissue attribute of CellDevelopmentStep and CellLineagePath subclasses of Event.

CellType

CellType terms represent different categories of animal cells, characterized by their specific structure, function, and biochemistry, derived from an ExternalOntology, Cell Ontology (https://obofoundry.org/ontology/cl.html). CellType terms are used to populate the *cellType* attribute of Cell and EntityWithAccessionedSequence subclasses of PhysicalEntity.

Disease

Disease encompasses terms that describe impairments of normal functions caused by pathogenic genetic mutations or environmental factors, such as microbial pathogens. Terms are derived from EBI Disease Oontology (DO) (https://disease-ontology.org/) and are used to populate the disease attribute of abnormal physicalEntities (genomeEncodedEntitiy, complex, entitySet) and abnormal events (**r**eactionLikeEvents, and pathways**)** to specify the disease(s) with which these abnormal instances are associated. Disease instances include *identifier* and *name* attributes taken from DO.

PsiMod

PsiMod encompasses terms representing post-translational chemical modifications of proteins derived from an ExternalOntology Protein Modification Ontology (MOD:

https://ontobee.org/ontology/MOD). The psiMod is an attribute of all the subclasses of TranslationalModification.

SequenceOntology

SequenceOntology is a local copy of portions of the ExternalOntology Sequence Ontology (http://www.sequenceontology.org/). It encompasses terms representing the variant features and attributes of biological sequence, used to populate the structuralVariant attribute of the subclasses of FunctionalStatus.

Figure

Figures for Reactome events are created in collaboration with Reactome's graphic illustrator. Curators needing to add a figure should consult with the illustrator and plans entered into to the illustration calendar. The *url* attribute of the **figure** instance holds the address of the image file in the Reactome CVS website repository, and the *figure* attribute of the data object to be illustrated holds the name of the **figure** instance. Example: /figures/apoptotic_factor_responses.jpg is this file location for the figure used to illustrate the pathway "Apoptotic factor-mediated response".

FrontPage

A list of the top-level pathways of the Reactome event hierarchy (autophagy, cell cycle, cell-cell communication, …) listed on the Reactome pathway browser page. The single instance of this class is maintained by the Editor-in-Chief and Managing Editor.

GO_BiologicalProcess

A local copy of the Gene Ontology Biological_Process ontology. Instances of this class provide *goBiologicalProcess* attributes for Reactome **Events** and **Regulation** instances.

GO_CellularComponent

A local copy of the Gene Ontology Cellular_Component ontology. Instances of this class provide the subcellular location or locations of Reactome **Events** and **PhysicalEntities** .

Compartment

A subset of GO_CellularComponent terms manually curated by Reactome to specify non-overlapping subcellular locations (*compartment* attribute values) of **Events** or **physicalEntities**. If a **physicalEntity** occurs in more than one compartment, different instances of the entity are created for each location, distinguished from one another by their *compartment* attributes.

GO_MolecularFunction

A local copy of the Gene Ontology Molecular_Function ontology. Instances of this class provide the *activity* of an instance of **catalystActivity**.

InstanceEdit

Records the date and time when a Reactome instance was *created* or *modified* (and, in the case of events, when they were authored, reviewed, or edited). It identifies the *person* responsible for the creation or modification and a datestamp. InstanceEdit instances are automatically generated when the central database is updated from the curator tool, and should not be manually edited.

PathwayDiagram

An SBML-compliant graphical representation of one or more **pathways** (*representedPathway* attribute values). Every **ReactionlikeEvent** in the included pathways is represented in the diagram, with icons that represent PhysicalEntities and edges that represent causal connections (provides input / output, catalyzes, regulates) among the physicalEntities. These are superimposed on icons that represent **compartments** (subcellular locations).

Person

The name of a person. Used to identify curators, authors, reviewers, and **literatureReference** authors. Any **Person** instance can have an *affiliation* attribute, but at present this information is recorded only for Reactome contributors, i.e., authors, reviewers, and curators. A contributor identifier such as ORCID can be associated with a person instance via the CrossReference attribute allowing to claim their contributions to Reactome in ORCID.

PhysicalEntity [superclass]

Used to maintain logical integrity of data model, not used for manual annotation. **PhysicalEntity** contains the classes **Cell**, **Complex**, **Drug**, **EntitySet**, **GenomeEncodedEntity**, **OtherEntity**, **Polymer**, and **SimpleEntity**. A PhysicalEntity is a physical substance that can interact with other substances. PhysicalEntities include all kinds of small molecules, proteins, nucleic acids,

chemical compounds, complexes, larger macromolecular assemblies, atoms (including ionized atoms), electrons, and photons.

Cell

A subclass of **PhysicalEntity**, represents a type of cell in a particular state of development/differentiation. A Cell instance has *cellType*, *tissueLayer*, *tissue*, and *organ* as single-valued attributes which are populated using terms from the Cell Ontology. Anatomy terms from the UBERON ExternalOntology are used to populate *tissueLayer*, *tissue*, and *organ* attributes. Cell instances also have multi-valued *ProteinMarker* and *RNAMarker* attributes, which are specific to the cell type and/or the cell state or are differentially upregulated in the Cell. The markers (protein or RNA) are manually curated and associated with literature references for evidence.

Complex

A **PhysicalEntity** that represents a larger functional unit formed by the covalent or noncovalent association of two or more **PhysicalEntity** *components* (which can themselves be **Complexes** or **Sets**). Usually, complex components are proteins, but they can also be RNA or DNA molecules, or simpleEntity chemicals, e.g. metal ions, nucleotides, lipids that function as enzyme cofactors. Complexes include homomultimers, e.g., ESR1 homodimer (nucleoplasm), heteromultimers, e.g., FANCD2:FANCI [nucleoplasm], and complexes composed of other complexes, e.g., PPARG:RXRA:Corepressor Complex [nucleoplasm] .

Drug

A superclass which is not used to instantiate specific instances, the **Drug** superclass contains subclasses **ChemicalDrug**, **ProteinDrug**, and **RNADrug**. A drug may be defined as any substance that is administered as a medicine to alter metabolic processes in the body, most often processes involved in diseases or abnormal states.
Reactome does not classify as drugs those substances, such as insulin or epinephrine, that are produced naturally by the human body, though these substances may be administered as medicines and may be considered drugs in other databases.
Artificial derivatives of naturally produced human substances, such as glargine (a derivative of insulin), are considered drugs, as are natural non-human substances such as quinine and caffeine.

ChemicalDrug

A subclass of the **Drug** class. A ChemicalDrug is a synthetic substance or non-human natural substance that is not a protein, RNA, or DNA molecule and is administered to the body to modify a metabolic reaction or reactions, usually to treat a disease or abnormal state.

Mandatory attributes of ChemicalDrug include one or more *diseases* treated by a drug, and *referenceEntity*, which is a reference to a corresponding entity in a database such as Guide to Pharmacology.

ProteinDrug

A **ProteinDrug** is a substance that contains an active component that is a protein, whether modified or unmodified, and is administered to the body to modify a metabolic reaction or reactions, usually to treat a disease or abnormal state. A subclass of the **Drug** class. Mandatory attributes of ProteinDrug include *disease*, which is a reference to the disease treated by a drug, and *referenceEntity*, which is a reference to a corresponding entity in a database such as Guide to Pharmacology. Examples of **ProteinDrugs** are semaglutide, a chemically modified form of GLP1, and etanercept, a fusion between the TNFalpha receptor and the Fc portion of IgG1. Naturally occurring human proteins such as insulin are considered instances of the class EntityWithAccessionedSequence rather than instances of the class ProteinDrug. Artificially modified versions of human proteins, such as glargine or lispro insulin, are considered **ProteinDrugs**.

RNADrug

A subclass of the **Drug** class. Mandatory attributes of RNADrug include *disease*, which is a reference to the disease treated by a drug, and *referenceEntity*, which is a reference to a corresponding entity in a database such as Guide to Pharmacology. A RNADrug is a substance that contains an active component that is a ribonucleic acid, whether modified or unmodified, and is administered to the body to modify a metabolic reaction or reactions, usually to treat a disease or abnormal state. **EntityWithAccessionedSequence** is the non-drug class that corresponds to **RNADrug**.
Examples of RNADrugs are inclisiran and asvasiran, which are non-natural small interfering RNAs (siRNAs). Naturally occurring human RNAs are considered instances of the class EntityWithAccessionedSequence rather than instances of the class RNADrug.

EntitySet [superclass]

Used to maintain logical integrity of data model, not used for manual annotation. Two or more **PhysicalEntities** grouped because of a shared molecular feature or function. The superclass for **CandidateSet** and **DefinedSet**. While sets are, by default, homogeneous (members having the same subclass of PhysicalEntity), they are not required to be. For example, the defined set platelet alpha granule contents contains, as members, EntitiesWithAccessionedSequences, Complexes and Sets.

CandidateSet

A collection of physical entities that are functionally indistinguishable for the purpose of Reactome annotation, some of which are well-characterized (the “members” of the set) and

some of which are incompletely characterized (the “candidates” of the set), as judged by the curator who assembles the set and the outside expert reviewers who evaluate it.

Members of a set are related by an "OR" function. That is, either one member OR another member can participate in a reaction.

Sets may be ordered or unordered as specified in the *isOrdered* attribute. A specific member of an ordered set has a correspondence with a specific member of another ordered set, as specified by their positions within the sets. For example, consider an ordered set containing substrate1 and substrate2 that reacts to yield an ordered set containing product1 and product2. In this case, substrate1 will yield product1 and substrate2 will yield product2. In the case of unordered sets, any member in an input set can correspond to any member in an output set. Sets in Reactome are considered ordered by default.

Example:

RIP2 ubiquitin ligases

One (*member*) has been shown to participate as a ubiquitin ligase (ITCH); the other (*candidates*) are thought to do so. https://reactome.org/content/detail/R-HSA-1248659

DefinedSet

A collection of well-characterized physical entities that are functionally indistinguishable for the purpose of Reactome annotation, e.g., a collection of isoforms of a protein that all mediate the identical metabolic reaction. A set is formally a list of instances linked by logical ORs.

Sets may be ordered or unordered as specified in the *isOrdered* attribute. A specific member of an ordered set has a correspondence with a specific member of another ordered set, as specified by their positions within the sets. For example, consider an ordered set containing substrate1 and substrate2 that reacts to yield an ordered set containing product1 and product2. In this case, substrate1 will yield only product1 and substrate2 will yield only product2. In the case of unordered sets, any member in an input set can correspond to any member in an output set. Sets in Reactome are considered ordered by default.

GenomeEncodedEntity

Any informational macromolecule (DNA, RNA, protein) or entity derived from one by post-synthetic modifications, for example, covalently modified forms of proteins. **GenomeEncodedEntity** is a subclass of **PhysicalEntity** and has one subclass: **EntityWithAccessionedSequence**. Unlike an EntityWithAccessionedSequence, a GenomeEncodedEntity is not required to have a reference entity, though a reference to an entity in a database can be specified in the *crossReference* attribute.

EntityWithAccessionedSequence

The subclass of **GenomeEncodedEntities** that can be associated through the *referenceEntity* attribute with reference molecules in UniProt (proteins) or ENSEMBL (DNA, RNA) databases. A

protein that has been partially purified and whose enzymatic properties are known but whose amino acid sequence is not, is a GenomeEncodedEntity but not an EntityWithAccessionedSequence.

OtherEntity

PhysicalEntities that we are unable or unwilling to describe in chemical detail and which, therefore, cannot be put in any other class. OtherEntity can be used to represent complex structures in the cell that take part in a reaction but which we can't or don't want to define molecularly.

Example 1: Cell membrane. In a case in which protein X associates with the membrane, but the actual membrane component(s) with which protein X interacts are unknown, the membrane can be represented as an "OtherEntity.

Example 2: kinesin-1, a microtubule motor protein, is involved in many kinds of movement in the cell, by 'walking' along microtubules, while dragging things like mitochondria, secretory vesicles, parts of the golgi, etc. Kinesin-1 binds to these complicated structures that we would not want to describe molecularly and which we can designate as "otherEntities".

Example 3: Holliday structure \[nucleoplasm\]
https://reactome.org/content/detail/R-ALL-75220#Homo%20sapiens

Polymer

Molecules that consist of indeterminate numbers of repeated units, and complexes that contain repeated units whose stoichiometry is variable or unknown. The repeated unit(s) (identified in the *repeatedUnit* attribute) can be any **PhysicalEntity**. The presence of more than one *repeatedUnit* value implies that the relative numbers of units in the polymer are unknown. If the units are present in known proportions, a **complex** of the appropriate numbers of units is used as the *repeatedUnit*. The size range of a **polymer** can be specified with *minUnitCount* and *maxUnitCount* values. Examples:

-\'glycogen\' with \'glucose\' as *repeatedUnit*.

-\'fibrin multimer\' with \'fibrin "monomer"\' (itself a Complex) as *repeatedUnit*.

-A microtubule consisting of equal amounts of alpha and beta tubulin would be constructed as a **polymer** containing a Complex of alpha and beta tubulins in the *repeatedUnit* attribute.

-Completely hypothetical Example: A complex consisting of 1 "part" of A and "4 "parts" of B (i.e. 1:4 ratio) would be represented as a **polymer** with a **complex** of one A and 4 B as its *repeatedUnit*.

-Another hypothetical Example: a complex where the ratio of individual building blocks A and B is unknown or variable is represented as a **polymer** containing A and B directly in the *repeatedUnit* attribute.

SimpleEntity

A defined chemical species not encoded directly or indirectly in the genome, typically a small molecule such as ATP or ethanol. The detailed structure of a **simpleEntity** is specified by linking it to the information provided for the molecule in the ChEBI external database via the *referenceEntity* slot. Separate **simpleEntity** instances are needed for each subcellular location (*compartment*) in which a molecule is found, e.g., ATP [cytosol] and ATP [nucleoplasm]. SimpleEntities such as ATP that occur in many species are not assigned a species.

Publication [superclass]

Used to maintain logical integrity of data model, not used for manual annotation. The Publication superclass encompasses any document, paper or electronic, that makes information available to the public. The subclasses of Publication are Book (chapters of books or entire books), LiteratureReference (journal articles), and URL (uniform resource locators that allow navigation of the internet to an electronic document).

Book

A book may be used in the literatureReference attribute of the relevant data class instances (ie pathway, reaction, entity). *Author*, *title*, and *year* of publication are required attributes. Chapters of books can also be cited within a book by using the *chapterTitle* attribute of the Book class.

LiteratureReference

A journal article, typically abstracted in PubMed, cited in a **summation** or linked to an **entity** or **event** instance. The attribute *pubMedIdentifier* contains the PubMed identifier (PMID) obtained from PubMed (https://pubmed.ncbi.nlm.nih.gov/) and is mandatory. Further information about the literature reference is also obtained from PubMed and is found in the attributes *author*, *year*, *title*, *journal*, *volume*, and *pages*.

URL

A website cited as a literature reference in data class instances such as **summation**, **entity,** or an **event** instance. The URL for the website is listed in the uniformResourceLocator attribute. The URL should contain all information, including transfer protocols such as “http://”, “https://”, or “ftp://”, required to navigate to the resource.

psiMod

A local copy of the PSI-MOD ontology, used to create descriptions of *modifiedResidues* of proteins.

ReferenceDatabase

This class describes the source (external database) of an identifier in the **DatabaseIdentifier** (and SequenceDatabaseIdentifier) instance, e.g., *DatabaseIdentifier* COMPOUND:C00113 links the Reactome instance to the record maintained by the KEGG “compound” database.

ReferenceEntity [superclass]

Used to maintain logical integrity of data model, not used for manual annotation. ReferenceEntity captures the invariant features of a molecule such as its names, molecular structure and links to external community reference databases like UniProt, ENSEMBL or ChEBI, and is created by importing content from those databases. A ReferenceEntity is an attribute of the PhysicalEntity subclasses Drug, EWAS and SimpleEntity, and enables grouping of Reactome PhysicalEntities like ’’Glc (alpha-D-glucose), extracellular region’ and ’Glc (alpha-D-glucose), cytosol’ (all linked to the record of alpha-D-glucose maintained by ChEBI database, indicating their identical chemical nature. ReferenceEntities are not used in Reactions directly; they are attached to the PhysicalEntities involved. ReferenceEntities derived from Uniprot and ChEBI usually don’t need to be created by curators; they are imported automatically. In the case of as yet unexisting ChEBI entries, contact the managing Editor.

ReferenceGroup

ReferenceGroup is a chemical group instance (children of CHEBI:24433) from the ChEBI database, used to populate the *modification* attribute of classes such as **GroupModifiedResidue.** The ReferenceGroup record consists of an *identifier* and a *referenceDatabase*.

ReferenceMolecule

ReferenceMolecule is used to cross-reference individual chemical molecules to the ChEBI database, and is used to populate the *referenceEntity* attribute of **SimpleEntity**. Consists of an *identifier* and a *referenceDatabase*.

ReferenceSequence [superclass]

Used to maintain logical integrity of data model, not used for manual annotation. Molecules with an accessionedSequence. This class is a subclass of **ReferenceEntity**, and is the superclass of **ReferenceDNASequence**, **ReferenceGeneProduct**, and **ReferenceRNASequence**, so instances of it should not be created manually.

ReferenceDNASequence

Used to populate the *referenceEntity* attribute for EWASs describing DNA molecules (for instance EWAS's of human genes). The **ReferenceDNASequence** record consists of a *referenceDatabase* (generally ENSEMBL), the *identifier*, and a *species*.

ReferenceGeneProduct

Used to populate the *referenceEntity* attribute for EWASs describing proteins. The **ReferenceGeneProduct** record consists of a *referenceDatabase* (generally Uniprot), the *identifier* and a *species*. The canonical form of the protein from UniProt is used as the **ReferenceGeneProduct** unless the literature supports the role of a particular isoform of a protein in the described function. **ReferenceGeneProduct** is also used to populate the *referenceSequence* attribute of **ModifiedResidue**.

ReferenceIsoform

Used to populate the *referenceEntity* attribute for EWASs describing proteins, where the experimental literature supports the role of a specified isoform of a protein in a particular function. The **ReferenceIsoform** record consists of a *referenceDatabase* (generally Uniprot), the *identifier* and a *species*.

ReferenceRNASequence

Used to populate the *referenceEntity* attribute for EWASs describing RNA molecules (for instance EWAS's of mRNAs). The **ReferenceRNASequence** record consists of a *referenceDatabase* (generally ENSEMBL), the *identifier*, and a *species*.

ReferenceTherapeutic

Used to populate the *referenceEntity* attribute for **PhysicalEntities** of the **Drug** class, including **ChemicalDrug**, **ProteinDrug**, and **RNADrug**. The **ReferenceTherapeutic** record consists of a *referenceDatabase* (generally Guide to Pharmacology) and the *identifier*.

Regulation [superclass]

Used to maintain logical integrity of data model, not used for manual annotation. Consists of subclasses **NegativeRegulation** and **PositiveRegulation**, along with their subclasses, **NegativeGeneExpressionRegulation** and **PositiveGeneExpressionRegulation** and **Requirement**, respectively. Instances of Regulation subclasses are added to the *regulatedBy* attribute of **ReactionLikeEvents** to capture the negative or positive effect of a **PhysicalEntity** on the described event.

NegativeRegulation

NegativeRegulation is an optional class used to populate the *regulatedBy* attribute of a **ReactionLikeEvent** that describes an inhibitory effect of a Regulator, a **PhysicalEntity**, on that event. The mechanism of action of a Regulator, if known, can be described through the optional *activity* attribute which is populated by terms from an ExternalOntology GO_MolecularFunction. If the Regulator is a Complex, the specific Complex component(s) that play the regulatory role can be specified as activeUnit(s) of the NegativeRegulation instance.

NegativeGeneExpressionRegulation

NegativeGeneExpressionRegulation is a subclass of **NegativeRegulation** that describes a direct inhibitory effect of a Regulator on a **BlackBoxEvent** that represents gene expression. The Regulator is a complex of regulatory component(s) such as transcription factors with the target gene or mRNA molecule. The regulatory component(s) are defined as *activeUnit(s)* of the **NegativeGeneExpressionRegulation** instance.

PositiveRegulation

PositiveRegulation is an optional class used to populate the *regulatedBy* attribute of a **ReactionLikeEvent** that describes a stimulatory effect of a Regulator, a **PhysicalEntity**, on that event. The mechanism of action of a Regulator, if known, can be described through the optional *activity* attribute which is populated by terms from an External Ontology GO_MolecularFunction. If the Regulator is a Complex, the specific Complex component(s) that play the regulatory role can be specified as activeUnit(s) of the PositiveRegulation instance.

PositiveGeneExpressionRegulation

PositiveGeneExpressionRegulation is a subclass of PositiveRegulation that describes a direct stimulatory effect of a Regulator on a BlackBoxEvent that represents gene expression, so that the Regulator represents a complex of regulatory component(s) such as transcription factors with the target gene or mRNA molecule. The regulatory component(s) are defined as *activeUnit(s)* of the PositiveGeneExpressionRegulaton instance.

sequence-related comments that are provided by the reference database. It is filled in automatically. Curators should not alter this.

Requirement

A subclass of **PositiveRegulation** that denotes a regulatory **PhysicalEntity** without which the regulated event cannot occur.

ReviewStatus

An attribute of event. Describes the level of internal/external review that the event has undergone.
One star (1): Restructured after internal review NOT RELEASED.
Two star (2) Restructured after external review NOT RELEASED.
Three star (3) Internally reviewed, awaiting external review RELEASED.
Four star (4) Restructured after external review, then internally re-reviewed RELEASED.
Five star. 5 Externally reviewed RELEASED. A more detailed description can be found in the user guide here (https://reactome.org/userguide/review-status).

StableIdentifier

A unique identifier associated with a database object that does not change over time. Stable Identifiers have the format R-XXX-YYYYYY.Z where R stands for Reactome and invariant. XXX is the three letter species code (see appendix XX for species codes). If the instance applies to all species (e.g. a simple entity such as ATP), the code is “ALL”. If the instance applies to none or multiple but not all species, the code is “NUL”. YYYYYY is the database identifier of the instance. If the instance is modified between database releases, the version number (Z) in the IdentifierVersion attribute is incremented.

StableIdentifierHistory

This class is used for tracking, and each instance represents a stable identifier with a particular minor version that was present at one time in the history (either still existing currently on at one time in the past and no longer). The class is used with "StableIdentifierReleaseStatus" to indicate which release numbers had this minor version and if/when it was changed (i.e. replaced, removed, incremented, etc.)

StableIdentifierReleaseStatus

Instances of this class are used as values for the attribute of StableIdentifier used to indicate whether the stableIdentifier has been released publicly.

Summation

A free text description of an **Event** or **physicalEntity**. Citations that provide useful background information, but that are not sources of primary data for the event or entity, can be linked to summations. (**LiteratureReferences** providing experimental data for the details of an event should be linked directly to the event). Attributes: *Text*- description of the event *LiteratureReference*- paper(s) that support the text description, but not the primary evidence for the event. See Summation example for FAM83B, (FAM83A, FAM83D) bind EGFR.

Taxon

A controlled vocabulary populated manually by Reactome from the NCBI Taxonomy Database, to provide terms to identify the *species* of events and genome-encoded physical entities. Taxonomy terms of higher rank than species, and their relationships, are maintained in Reactome to facilitate computational reasoning and data mining. Each taxon instance includes the taxon's *name* as assigned by NCBI, its NCBI identifier (in the form of a Reactome *crossReference*), and its NCBI-assigned *supertaxon*. This class is maintained by the Editor-in-Chief and Managing Editor. Taxon terms of species rank are used for annotation of gene products and other relevant PhysicalEntities, and events.

Species

Gene products and events have **species**-level NCBI-derived **taxon** terms as attributes. This class is maintained by the Editor-in-Chief and Managing Editor.

_Release

A class which indicates a Reactome version number and its date of release. Also, an attribute of the class _updateTracker that indicates the release number and date of a specified Reactome database release.

_UpdateTracker

Describes the changes that were made to an instance at a specified Reactome release relative to the previous release. Specifies the DBID and name of the instance that has changed (an event or PhysicalEntity), the release number and date of that release that the change occurred, and a list of changes made to the entity (see appendix R9 for a list of changes that are tracked)

Reactome Attributes

_displayName

This attribute contains the name of the instance that will be displayed on the Reactome website. It is automatically filled in by the curator tool each time a new instance is created, using attribute values supplied for the instance by the curator. For example, the _displayName of an event consists of the event's manually generated *name* and its manually selected *species*.

accession

Holds the numerical portions of identifiers assigned to GO terms in the local copies of the GO ontologies. Should not be manually edited.

accessUrl

Template that is used to form the URL that will be used to link to a particular record in an external database. Generally, this should not need to be touched. Contact developerss if you think any changes are necessary.

activity

contains the GO Molecular Function ontology. An Activity term is used to cross-reference a Reactome CatalystActivity. Curators should never create instances of class Activity directly. If curators want to use a GO term that is not in Reactome (yet) contact Managing Editor.

address

This is an attribute for the affiliation class. Enter the address of the research institute (etc).

affiliation

An Affiliation is the name and address of an institution, e.g., Cold Spring Harbor Laboratory 1 Bungtown Road, Cold Spring Harbor NY USA.

author

This attribute must be filled in for instances of LiteratureReference, (where it refers to name of the author of a particular reference), for InstanceEdit (where it refers to the curator that has generated the instanceEdit), and for events.

comment

This attribute is used for SequenceDatabaseIdentifier class instances and contains any cation in the cell. Is an attribute of PhysicalEntities and Events. Note that complexes with components that span multiple cellular locations should specify one primary location as the compartment and the location of the other components will be automatically filled into the “includedLocation” attribute during the database release process. For example, a transmembrane complex with components in the extracellular space, plasma membrane, and cytosol would have plasma membrane specified as the compartment.

componentOf

An attribute of GO_CellularComponent that is specified by the Gene Ontology and describes the GO Cellular Component instances that are parent terms. Imported from GO.

coordinate

Refers to the amino acid residue location at which a modification occurs within a protein.

created

Automatically filled in with the name of the curator/developer creating the instance and date/time stamp that the instance was created.

crossReference

This attribute holds references to the equivalent instance in other databases. For example, the small molecule 4MAA has the KEGG COMPOUND identifier COMPOUND:C01036 in its *crossReference* attribute.

databaseIdentifier

Holds database identifiers from external databases like UniProt and ChEBI (but not GO – see *Accession* attribute). Has attributes for the *referenceDatabase* and the *identifier* in that database.

dateTime (on InstanceEdit)

A timestamp generated automatically when the central database is updated from a curator tool project. Do not edit it manually.

DB_ID

These are the unique stable identifiers applied to each instance in the Reactome database. They are generated automatically when newly created instances are first submitted to the database. They should never be modified or manually created by a curator.

definition

This attribute is used in Event, PhysicalEntity, Activity, Evidencetype and GO_BiologicalProcess instances. The definition of a Reactome Activity is the official Gene Ontology definition for the equivalent GO Molecular Function term. The GO_BiologicalProcess definition is the official Gene Ontology definition for the corresponding Biological process term. The Evidencetype definition is taken from the GO evidence code definitions.

description

This is found in the class \'SequenceDBI\'. This is for the \'Description DE\' line from SwissProt.

editor

Choose or create a Person instance/s for the editor/s. Used to determine the information for the frontpage.

containedIn

endCoordinate

The amino acid residue location at which an EntityWithAccessionedSequencepart ends.

figure

Certain events in Reactome are associated with a descriptive illustration. See, for example, Autophagy. When such an illustration exists, the figure attribute of the event is populated with a figure instance containing the location of the image file on the server in its URL attribute. The URL has the format: figures/…/[FigureFileName].svg

firstname

The first name of a person. Optional attribute – can be left blank. Reactome uses only the first letter of the first name (entered separately in the *initial* attribute) to create names of people for display on the web site.

formula

The chemical formula of a **SimpleEntity**. For example, ATP = C10H16N5O13P3.

functionalStatusType

Describes the functional consequence of this genetic change (gain-of-function, loss-of-function, decreased_transcript_level etc). With sequenceVariant, functionalStatusType makes up **FunctionalStatus**, which is used to fill the *functionalStatus* attribute of EntityFunctionalStatus

geneIdentifier

This attribute is for storing information about what is the gene that this transcript/protein originates from. This is really just a shortcut. If the current instance is a gene identifier then this should be left empty. Accepts multiple values in order to be able to point to the SAME gene in different databases, e.g. EMBL, HUGO, Ensembl.

geneName

GN lines from sequence record.

hasComponent

Holds instances of **PhysicalEntity** that are components of a **Complex**.

hasEvent

Holds instances of **Event** that are grouped into a **Pathway**. They should be entered in the order as they appear in the **Pathway**. On **BlackBoxEvent** this attribute holds **Event** instances to describe known steps within the **BlackBoxEvent**, e.g. the reactions of the fatty acid beta oxidation cycle.

hasMember

Takes individual reactions for an instance of **ReactionlikeEvent** that represents a group of reactions via the use of an **EntitySet**. Such individual reactions do not need to be spelled out, unless they have a distinct feature like a literature reference or specific regulation.

identifier

This attribute holds the actual database identifier number for a given **DatabaseIdentifie**r instance.

inferredFrom

Points to the event or entity in another species that this event/entity has been inferred from. If the inference is based on computation only, this is indicated under evidenceType (= IEA).

initial

This attribute holds the initials of a person represented in a **Person** instance.

input

The input **PhysicalEntities** of a given **ReactionLikeEvent** are each entered individually.

instanceOf

Only used for GO classes and is populated/updated automatically.

InteractionEvent

This class is designed to model simple interactions between any physical entities. It is introduced to be used for data imported from other data sources including protein-protein interactions or other types of interactions.

journal (on LiteratureReference)

The name of the scientific journal in which the reference was published (or the title of a book). For references retrieved from PubMed with the curator and author tools, this slot is filled automatically.

keyword

Contains keywords associated with the sequence. This is pulled in (if available) automatically from the external database.

literatureReference (on summation)

Paper/s that support the text description, but not the primary evidence for the event. This shows up as hyperlinked lines on a web browser.

modification

For a given modifiedResidue instance, enter the simpleEntity that represents the specific modification of the residue within the modified protein. For example, for the "ModifiedResidue" instance: "Orthophosphate on Serine [nucleus] 428 of SPTREMBL:O46469", the modification would be "Orthophosphate".

modified

Filled in automatically when a database instance is modified.

name

A short textual description of the event/entity/etc. For some classes it is a defining attribute, so it should be chosen carefully to be unique.

Ontology

Internal definition and not instantiable for holding data. It is the content of the protegé file defining the database.

orthologousEvent

Points to equivalent events in other species. In contrast to \'inferredFrom\' this attribute is attached to the events in both species - it only indicates equivalence, not inference.

output

The output **PhysicalEntities** of a given event are each entered individually.

pages

The inclusive page numbers of a **LiteratureReference**. For references retrieved from PubMed with the curator and author tools, this attribute is filled automatically.

precedingEvent

Points to the preceding event(s), which is usually events whose output is used as input for the present event. The preceding event can also point to a pathway, but make sure the connection is always given on the reaction level as well (this is important when it comes to visualization).

proteinIdentifier

This attribute is for storing information about what the protein this transcript/gene produces. This is really just a shortcut. If the current instance is a protein identifier then this should be left empty. Accepts multiple values in order to be able to point to the same proteins in different databases, e.g. EMBL, HUGO, Ensembl, RefSeq.

pubMedIdentifier (on literatureReference)

The pubmed identifier (number only). This information must be manually supplied by the curator.

reactionType

A *controlledVocabulary* that provides attributes to classify *reactionLikeEvents*. Currently in use only to distinguish reactions involving drugs by modes of action.

drugActionType

Reactions that have *drugs* and *sets or complexes containing drugs* as inputs and output have a drugAction attribute taken from the IUPHAR / Guide to Pharmacology glossary to classify the effect of the drug on the reaction (agonist, channel blocker, etc.)..*

requiredInputComponent

Takes an input component that is essential for this event to happen, e.g. a defined domain of a protein. Use Domain, with the appropriate coordinates filled in.

Note: To be used only for \'non-trivial statements\'. For example, repeating an entity here that is identical to a PhysicalEntity given for input is not helpful.

relatedSpecies (on event)

This attribute is used in events involving more than one species (e.g event involving host-pathogen or symbiotic interactions) to denote the "bystander" species in the event. For example, in the event such as flagellin of Escherichia coli binds to human TLR5, the process is occurring on the human cell membrane and involves protein from the bystander bacterial species. Thus, human would be entered in the *species* attribute and E.coli would be entered as *relatedSpecies*. For additional information and use cases, see this document.

residue

This is to hold the residue (like serine or tyrosine) which is modified. The value is an instance of class ReferenceGroup or ReferenceMolecule, i.e. an instance representing serine.

resourceIdentifier

This attribute is used to add identifiers registered in the identifiers.org web site for ReferenceDatabase instances. It is added for QA purposes and should not affect curation work. The attribute is string, single-valued, and optional.

reverseReaction

Choose a Reaction that is the reverse of the current Reaction.

reviewer (on Event)

The name of the person who reviewed the event. These attribute values are the source of the reviewers given for high-level events listed on the Reactome table of contents.

secondaryIdentifier

Generated automatically.

sequenceOntology

SequenceOntology encompasses terms representing the variant features and attributes of biological sequence derived from an ExternalOntology Sequence Ontology (http://www.sequenceontology.org/). SequenceOntology terms are used to populate the *structuralVariant* attribute of **FunctionalStatus**.

species

This attribute holds the name of the species in which the described physicalEntity or Event is occurring. The *species* attribute on event identifies the environment where the reaction is happening. This should almost always be human for human Reactome curation. The exceptions are when we have annotated pathways in other species for inference (e.g Mus) or when we are specifically annotating another species pathways as part of another project (e.g Gallus and Mycobacterium tuberculosis). In disease or immune system processes, the host would be listed as the “species”.

When annotating a process involving host-pathogen interactions, the r*elatedSpecies* attribute should be used to list the pathogen (bystander) species. See the *relatedSpecies* attribute definition and use cases for a more detailed description. **Note: When creating a chimeric reaction to be used for inference (i.e one that involves entities from multiple species and (usually) observed in vitro) ALL species should be listed in the species attribute.

startCoordinate

The amino acid residue location at which an EntityWithAccessionedSequence starts.

structuralVariant

Field takes terms from SequenceOntology and describes the nature of the underlying genetic change that gives rise to the variant. With FunctionalStatusType, structuralVariant makes up FunctionalStatus, which is used to fill the *functionalStatus* attribute of EntityFunctionalStatus.

summation

A succinct free-text description of the instance, typically an event, with which it is associated.

superTaxon

This refers to the "parent" for a given taxon within the NCBI taxonomy hierarchy. For example the SuperTaxon for both Rattus and Mus is Murina and the SuperTaxon for Murina is Muridae.

surname

The family name of a **person**.

text

Holds the text of the summation.

templateEvent

Allows making a reference to a general process that underlies a **BlackBoxEvent**, e.g. 'Gene Expression' as template for the synthesis of a specific protein. This attribute has never been used - should it be removed from the data model?

title

The title of a **literatureReference**. For references retrieved from PubMed with the curator and author tools, this attribute is filled automatically.

transcriptIdentifier

This attribute is for storing information about what is the transcript that this protein originates from or that this gene produces. This is really just a shortcut. If the current instance is a transcript identifier then this should be left empty. Accepts multiple values in order to be able to point to the same gene and proteins in different databases, e.g. EMBL, Ensembl, RefSeq.

url

Holds URL for site which gives "summary information" about the database including a description of what information the db contains. This attribute should not need to be modified manually.

volume

The volume number of the *journal* in a **LiteratureReference**. For references retrieved from PubMed with the curator and author tools, this attribute is filled automatically.

year

The year a **LiteratureReference** was published. For references retrieved from PubMed with the curator and author tools, this slot is filled automatically.

Graph database specific classes

DBInfo

Generic information about the current version of the graph database used.

GO_Term

Gene Ontology term that is used by Reactome events as cross references. Can be GO_BiologicalProcess, GO_CellularConponent or GO_MolecularFunction.

Interaction

Parent label for all kind of molecular interactions brought into Reactome automatically from other databases, based on referenced evidence in a reviewed publication.

TopLevelPathway

A specific kind of Pathway which covers a broad topic of cellular biology. The top-level pathways represent the highest level of the hierarchical organisation of events in Reactome, as displayed in Pathway Browser.

UndirectedInteraction

A kind of molecular interaction for which no information is available about the directionality of the interaction that forms a bond between two molecules.

C6. Small_molecule_renaming- small mols short names

name	short name	location	Notes
(1,4-alpha-glycosyl)(n-2)		cytosol	
(1,4-alpha-glycosyl)n		cytosol	
(24R, 25R) 3alpha,7alpha,12alpha,24-tetrahydro:		peroxisomal matrix	
(24R, 25R) 3alpha,7alpha,24-trihydroxy-5beta-CH		peroxisomal matrix	
(2R) Pristanoyl-CoA		peroxisomal matrix	
(2S) Pristanoyl-CoA		peroxisomal matrix	
(Glc)1 (GlcNAc)2 (Man)7bc		endoplasmic reticulum lumen	
(Glc)1 (GlcNAc)2 (Man)8b		endoplasmic reticulum lumen	
(Glc)1 (GlcNAc)2 (Man)8c		endoplasmic reticulum lumen	
(Glc)1 (GlcNAc)2 (Man)9		endoplasmic reticulum lumen	
(Glc)1 (GlcNAc)2 (Man)9		endoplasmic reticulum quality control compartmer	
(Glc)1 (GlcNAc)2 (Man)9		ER to Golgi transport vesicle membrane	
(Glc)1 (GlcNAc)2 (Man)9		Golgi membrane	
(Glc)1 (GlcNAc)2 (Man)9 (PP-Dol)1		integral to lumenal side of endoplasmic reticulum	
(Glc)2 (GlcNAc)2 (Man)9 (Asn)1		endoplasmic reticulum lumen	
(Glc)2 (GlcNAc)2 (Man)9 (PP-Dol)1		integral to lumenal side of endoplasmic reticulum	
(Glc)3 (GlcNAc)2 (Man)9 (Asn)1		Golgi lumen	
(Glc)3 (GlcNAc)2 (Man)9 (Asn)1		integral to lumenal side of endoplasmic reticulum	
(Glc)3 (GlcNAc)2 (Man)9 (PP-Dol)1		integral to lumenal side of endoplasmic reticulum	
(GlcNAc)2 (Man)2 (PP-Dol)1		integral to cytosolic side of endoplasmic reticulum	
(GlcNAc)2 (Man)3 (PP-Dol)1		integral to cytosolic side of endoplasmic reticulum	
(GlcNAc)2 (Man)5 (Asn)1		Golgi lumen	
(GlcNAc)2 (Man)5 (PP-Dol)1		integral to cytosolic side of endoplasmic reticulum	
(GlcNAc)2 (Man)5 (PP-Dol)1		integral to lumenal side of endoplasmic reticulum	
(GlcNAc)2 (Man)6 (PP-Dol)1		integral to lumenal side of endoplasmic reticulum	
(GlcNAc)2 (Man)7 (PP-Dol)1		integral to lumenal side of endoplasmic reticulum	
(GlcNAc)2 (Man)7aa		cytosol	
(GlcNAc)2 (Man)7aa		endoplasmic reticulum quality control compartmer	
(GlcNAc)2 (Man)7bc		endoplasmic reticulum lumen	
(GlcNAc)2 (Man)7bc		Golgi lumen	
(GlcNAc)2 (Man)8 (Asn)1		Golgi lumen	
(GlcNAc)2 (Man)8 (PP-Dol)1		integral to lumenal side of endoplasmic reticulum	
(GlcNAc)2 (Man)8a		cytosol	
(GlcNAc)2 (Man)8a		endoplasmic reticulum quality control compartmer	
(GlcNAc)2 (Man)8a		Golgi lumen	
(GlcNAc)2 (Man)8b		endoplasmic reticulum lumen	
(GlcNAc)2 (Man)8b		endoplasmic reticulum quality control compartmer	
(GlcNAc)2 (Man)8b		Golgi lumen	
(GlcNAc)2 (Man)8c		endoplasmic reticulum lumen	
(GlcNAc)2 (Man)8c		endoplasmic reticulum quality control compartmer	
(GlcNAc)2 (Man)8c		Golgi lumen	
(GlcNAc)2 (Man)9		endoplasmic reticulum lumen	
(GlcNAc)2 (Man)9		endoplasmic reticulum quality control compartmer	
(GlcNAc)2 (Man)9		Golgi lumen	
(GlcNAc)2 (Man)9 (PP-Dol)1		integral to lumenal side of endoplasmic reticulum	
(GlcNAc)3 (Man)3 (Asn)1		Golgi lumen	
(GlcNAc)3 (Man)5 (Asn)1		Golgi lumen	
(GlcNAc)4 (Man)3 (Asn)1		Golgi lumen	
(R)-2-hydroxyglutarate	2HG	cytosol	
(R)-2-hydroxyglutarate	2HG	mitochondrial matrix	
(R)-3-aminoisobutyric acid	3AIB	cytosol	
(R)-3-aminoisobutyric acid	3AIB	mitochondrial matrix	
(S)-1-Pyrroline-5-carboxylate		mitochondrial matrix	
(S)-2-hydroxyglutarate	2HG	mitochondrial matrix	
(S)-3-Hydroxybutanoyl-CoA	3HB-CoA	mitochondrial matrix	
(S)-3-Hydroxydodecanoyl-CoA		mitochondrial matrix	
(S)-3-Hydroxyhexadecanoyl-CoA		mitochondrial matrix	
(S)-3-Hydroxytetradecanoyl-CoA		mitochondrial matrix	
(S)-dihydroorotate		cytosol	
(S)-Hydroxydecanoyl-CoA		mitochondrial matrix	
(S)-Hydroxyhexanoyl-CoA		mitochondrial matrix	
(S)-Hydroxyoctanoyl-CoA		mitochondrial matrix	
(S)-Lactate	LACT	cytosol	Looked at LA but linoleic acid and lauric acid can also be LA
(S)-Lactate	LACT	extracellular region	Looked at LA but linoleic acid and lauric acid can also be LA
(S)-Malate	MAL	cytosol	
(S)-Malate	MAL	mitochondrial matrix	
[4Fe-4S](1+)		cytosol	
[4Fe-4S](2+)		cytosol	

name	short name	location	Notes
[4Fe-4S](2+)		cytosol	
1,24,25-trihydroxyvitamin D3	vitD3(OH)3	mitochondrial matrix	
1,25-dihydroxyvitamin D3	vitD3(OH)2	mitochondrial matrix	
1,2-diacylglycerol	DAG	cytosol	
1,2-diacyl-glycerol 3-phosphate	PA	cytosol	
1,2-Dihydroxy-3-oxo-methylthiopentene (1-)		cytosol	
1,3-Bisphospho-D-glycerate		cytosol	
1,3-diaminopropane	Dap	cytosol	
10-formyltetrahydrofolate	10-fthp	cytosol	
10-formyltetrahydrofolate polyglutamate		cytosol	
11,12-EET	11,12EET	endoplasmic reticulum lumen	
11,14,17-eicosatrienoic acid	ETE	extracellular region	http://en.wikipedia.org/wiki/Polyunsaturated_fatty_acid
11-deoxycorticosterone	DOC	cytosol	http://www.ebi.ac.uk/chebi/searchId.do?chebiId=CHEBI:16973
11-deoxycorticosterone	DOC	mitochondrial matrix	http://www.ebi.ac.uk/chebi/searchId.do?chebiId=CHEBI:16973
11-deoxycortisol		cytosol	
11-deoxycortisol		mitochondrial matrix	
12-hydroxy-5,8,11,14-eicosatetraenoic acid	20-HETE	endoplasmic reticulum lumen	http://www.chemspider.com/Chemical-Structure.4446281.html
13(S')-hydroxyoctadecadienoic acid	13-HODE	nucleoplasm	
14,15-EET	14,15EET	endoplasmic reticulum lumen	
17alpha-hydroxypregnenolone		cytosol	
17alpha-hydroxyprogesterone		cytosol	
18-hydroxyarachidonate		endoplasmic reticulum lumen	
18-hydroxycorticosterone		mitochondrial matrix	
19-hydroxyarachidonic acid		endoplasmic reticulum lumen	
19-hydroxyprostaglandin H2	19-OH PGH2	endoplasmic reticulum lumen	
1-acyl-glycerol 3-phosphate		cytosol	
1-chloro-2,4-dinitrobenzene	DNCB	cytosol	http://www.inchem.org/documents/icsc/icsc/eics0416.htm
1D-myo-Inositol 1,3,4,5,6-pentakisphosphate	I(1,3,4,5,6)P5	cytosol	
1D-myo-inositol 1,3,4,5-tetrakisphosphate	I(1,3,4,5)P4	cytosol	
1D-myo-inositol 1,3,4,6-tetrakisphosphate	I(1,3,4,6)P4	cytosol	
1D-myo-inositol 1,3,4-trisphosphate	I(1,3,4)P3	cytosol	
1D-myo-Inositol 1,4,5-trisphosphate	I(1,4,5)P3	cytosol	http://www.ebi.ac.uk/chebi/searchId.do?chebiId=CHEBI:16595
1D-myo-Inositol 1,4,5-trisphosphate	I(1,4,5)P3	platelet dense tubular network membrane	http://www.ebi.ac.uk/chebi/searchId.do?chebiId=CHEBI:16595
1D-myo-inositol 3,4,5,6-tetrakisphosphate	I(3,4,5,6)P4	cytosol	
1-hexadecylglycerol-3-phosphate	HXDG3P	cytosol	
1L-myo-Inositol-1-phosphate	I1P	cytosol	
1-methyl-4-phenyl-pyridinium	MPP	cytosol	
1-methyl-4-phenyl-pyridinium	MPP	extracellular region	
1-Palmitoyl dihydroxyacetone phosphate		peroxisomal matrix	
2,3-Dioxo-methiopentane-1P (2-)		cytosol	
2,4-dinitrophenyl-S-glutathione	DNP-S-GSH	cytosol	http://www.ebi.ac.uk/chebi/searchId.do?chebiId=CHEBI:8927
20alpha,22beta dihydroxyCHOLesterol	DHCHOL	mitochondrial inner membrane	
20-OH-leukotriene B4	20HLTB4	endoplasmic reticulum lumen	http://www.ebi.ac.uk/chebi/searchId.do?chebiId=CHEBI:15646
22beta-hydroxyCHOLesterol	22HCHOL	mitochondrial inner membrane	
24-hydroxyCHOLesterol	24HCHOL	endoplasmic reticulum membrane	
24-hydroxyCHOLesterol	24HCHOL	extracellular region	
25(R) DHCA-CoA		cytosol	
25(R) DHCA-CoA		peroxisomal matrix	
25(R) TetraHCA-CoA		cytosol	
25(R) TetraHCA-CoA		peroxisomal matrix	
25(R) THCA-CoA		cytosol	
25(R) THCA-CoA		peroxisomal matrix	
25(S) 3alpha,7alpha,12alpha-trihydroxy-5beta-Cł		peroxisomal matrix	
25(S) 3alpha,7alpha-dihydroxy-5beta-CHOLest-24		peroxisomal matrix	
25(S) DHCA-CoA		peroxisomal matrix	
25(S) THCA-CoA		peroxisomal matrix	
25-hydroxyCHOLesterol	25HCHOL	cytosol	
27-hydroxyCHOLesterol	27HCHOL	endoplasmic reticulum membrane	
27-hydroxyCHOLesterol	27HCHOL	extracellular region	
27-hydroxyCHOLesterol	27HCHOL	mitochondrial matrix	
27-hydroxyCHOLesterol 3-sulfate	27HCHOL-SO4	cytosol	
2-acetylaminofluorene-N-sulfate	AAFNS	cytosol	http://ctd.mdibl.org/detail.go?type=chem&acc=C047043
2-acyl-1-palmitoyl-sn-glycero-3-phosphoCHOLine		endoplasmic reticulum membrane	
2-acylglycerols	2AGs	extracellular region	
2-acyl-sn-glycero-3-phosphoCHOLine		endoplasmic reticulum membrane	
2-Amino-3-carboxymuconate semialdehyde	ACS	cytosol	http://www.hmdb.ca/metabolites/HMDB01330
2-Aminomuconate	2AM	cytosol	
2-Aminomuconate semialdehyde	2AMA	cytosol	

name	short name	location	Notes
2-Arachidonoylglycerol	2AG	extracellular region	http://www.ebi.ac.uk/chebi/searchId.do?chebiId=CHEBI:52392
2-Arachidonoylglycerol	2AG	plasma membrane	http://www.ebi.ac.uk/chebi/searchId.do?chebiId=CHEBI:52392
2-Chloroethylene oxide	2CEO	endoplasmic reticulum lumen	
2'-deoxyadenosine	dA	cytosol	
2'-deoxyadenosine	dA	extracellular region	
2'-deoxyadenosine	dA	mitochondrial matrix	
2'-deoxyadenosine 5'-diphosphate	dADP	cytosol	
2'-deoxyadenosine 5'-diphosphate	dADP	mitochondrial matrix	
2'-deoxyadenosine 5'-diphosphate	dADP	nucleoplasm	
2'-deoxyadenosine 5'-monophosphate	dAMP	cytosol	
2'-deoxyadenosine 5'-monophosphate	dAMP	extracellular region	
2'-deoxyadenosine 5'-monophosphate	dAMP	mitochondrial matrix	
2'-deoxycytidine	dC	cytosol	
2'-deoxycytidine	dC	mitochondrial matrix	
2'-deoxycytidine 5'-diphosphate	dCDP	cytosol	
2'-deoxycytidine 5'-diphosphate	dCDP	mitochondrial matrix	
2'-deoxycytidine 5'-diphosphate	dCDP	nucleoplasm	
2'-deoxycytidine 5'-monophosphate	dCMP	mitochondrial matrix	
2-deoxy-D-ribose 1-phosphate	dRibP	cytosol	
2'-deoxyguanosine	dG	cytosol	
2'-deoxyguanosine	dG	mitochondrial matrix	
2'-deoxyguanosine 5'-diphosphate	dGDP	cytosol	
2'-deoxyguanosine 5'-diphosphate	dGDP	mitochondrial matrix	
2'-deoxyguanosine 5'-diphosphate	dGDP	nucleoplasm	
2'-deoxyguanosine 5'-monophosphate	dGMP	cytosol	
2'-deoxyguanosine 5'-monophosphate	dGMP	mitochondrial matrix	
2'-deoxyinosine	dI	cytosol	
2'-deoxyinosine	dI	mitochondrial matrix	
2'-deoxyinosine 5'-monophosphate	dIMP	cytosol	
2'-deoxyinosine 5'-monophosphate	dIMP	mitochondrial matrix	
2'-deoxyuridine	dU	cytosol	
2'-deoxyuridine	dU	mitochondrial matrix	
2'-deoxyuridine 3'-monophosphate	dU3MP	cytosol	
2'-deoxyuridine 5'-diphosphate	dUDP	cytosol	
2'-deoxyuridine 5'-diphosphate	dUDP	mitochondrial matrix	
2'-deoxyuridine 5'-diphosphate	dUDP	nucleoplasm	
2'-deoxyuridine 5'-monophosphate	dU5MP	mitochondrial matrix	
2-Hydroxyphytanoyl-CoA	3S2HPhy-CoA	peroxisomal matrix	http://www.hmdb.ca/metabolites/HMDB01295
2-lysophosphatidylcholine	LPC	extracellular region	http://www.ebi.ac.uk/chebi/searchId.do?chebiId=CHEBI:17504
2-mercaptoethanol	BME	cytosol	http://www.sigmaaldrich.com/catalog/ProductDetail.do?D7=0&N5=SEARCH_CONCAT_PNO|BRAND_KEY&N4=M7154|SIGMA&N25=0&QS=ON&F=SPEC
2-Methyl-3-oxopropanoate	MOPP	mitochondrial matrix	http://www.chemspider.com/Chemical-Structure.290.html
2-oxobutanoate	2-Ketobutyrate	cytosol	
2-oxoglutaramic acid	2OGA	cytosol	
2-oxoglutarate	2OG	cytosol	
2-Oxoglutarate	2OG	extracellular region	http://pubs.rsc.org/en/content/articlelanding/2011/CS/C0CS00203H
2-Oxoglutarate	2OG	mitochondrial matrix	http://pubs.rsc.org/en/content/articlelanding/2011/CS/C0CS00203H
2-Oxoglutarate	2OG	nucleoplasm	http://pubs.rsc.org/en/content/articlelanding/2011/CS/C0CS00203H
2-Oxoglutarate	2OG	peroxisomal matrix	http://pubs.rsc.org/en/content/articlelanding/2011/CS/C0CS00203H
2-Phenylethylamine	PEA	cytosol	
2-Phenylethylamine	PEA	extracellular region	
2-Phospho-D-glycerate	2PG	cytosol	
2-trans-4-cis-decadienoyl-CoA		mitochondrial matrix	
2-trans-Dodecenoyl-CoA		mitochondrial matrix	
3-(4-Hydroxyphenyl)pyruvate		cytosol	
3-(indol-3-yl)pyruvic acid		cytosol	
3,16-dihydroxypalmitate	DHPALM	endoplasmic reticulum lumen	
3,3'-diiodothyronine	(I)2-T4	cytosol	
3,3'-diiodothyronine 4-sulfate	(I)2-T4-SO4	cytosol	
3,4-dihydroxybenzoate	DHBENZ	cytosol	
3,5,3'-triiodothyronine	(I)3-T4	cytosol	
3,5,3'-triiodothyronine	(I)3-T4	extracellular region	
3,5,3'-triiodothyronine 4-sulfate	(I)3-T4-SO4	cytosol	
3',5'-Cyclic AMP	cAMP	cytosol	
3',5'-Cyclic AMP	cAMP	plasma membrane	
3',5'-Cyclic GMP	cGMP	cytosol	
3,7,24THCA		cytosol	
3,7,24THCA		mitochondrial matrix	
3,7,24THCA-CoA		cytosol	

name	short name	location	Notes
3,7,24THCA-CoA		peroxisomal matrix	
3-acetamidopropanal		peroxisomal matrix	
3alpha,7alpha,12alpha,24(S)-tetrahydroxy-5beta		mitochondrial matrix	
3alpha,7alpha,12alpha-trihydroxy-5beta-CHOLes		peroxisomal matrix	
3alpha,7alpha,12alpha-trihydroxy-5beta-CHOLes		mitochondrial matrix	
3alpha,7alpha,24(S)-trihydroxy-5beta-CHOLestar		mitochondrial matrix	
3alpha,7alpha-dihydroxy-5beta-CHOLest-24-one-		peroxisomal matrix	
3alpha,7alpha-dihydroxy-5beta-CHOLestan-26-al		mitochondrial matrix	
3-aminopropanal	APrOH	peroxisomal matrix	
3-Dehydroquinate	3DHQ	cytosol	
3-Dehydroshikimate	3DHSK	cytosol	
3-hydroxy 4-methoxybenzoate		cytosol	
3-Hydroxyanthranilate	3HAA	cytosol	
3-hydroxyhexacosanoyl-CoA		peroxisomal matrix	
3-Hydroxy-L-kynurenine		cytosol	
3-Hydroxy-N6,N6,N6-trimethyl-L-lysine	HTM-K	cytosol	
3-Hydroxy-N6,N6,N6-trimethyl-L-lysine	HTM-K	mitochondrial matrix	
3-hydroxyoctadecanoyl-CoA		cytosol	
3-hydroxypalmitate	HPALM	endoplasmic reticulum lumen	
3-hydroxypristanoyl-CoA		peroxisomal matrix	
3-ketohexacosanoyl-CoA		peroxisomal matrix	
3-ketopristanoyl-CoA		peroxisomal matrix	
3-ketosphinganine		endoplasmic reticulum membrane	
3-Methoxy-4-hydroxyphenylglycol		cytosol	
3-Methyladenine	MADE	nucleoplasm	
3-methyleneindolenine	3MEIN	endoplasmic reticulum lumen	http://www.ncbi.nlm.nih.gov/pubmed/11592314
3-methylguanosine cap		cytosol	
3-methylguanosine cap		nucleoplasm	
3-methylindole	MIND	endoplasmic reticulum lumen	
3-oxo-(7,10,13,16)-docosatetraenoyl-CoA		cytosol	
3-oxobehenoyl-CoA		cytosol	
3-oxocerotoyl-CoA		cytosol	
3-Oxodecanoyl-CoA		mitochondrial matrix	
3-Oxododecanoyl-CoA		mitochondrial matrix	
3-Oxohexanoyl-CoA		mitochondrial matrix	
3-oxooctadecanoyl-CoA		cytosol	
3-Oxooctanoyl-CoA		mitochondrial matrix	
3-Oxopalmitoyl-CoA		mitochondrial matrix	
3-oxopropanoate	OPROP	mitochondrial matrix	
3-Oxotetradecanoyl-CoA		mitochondrial matrix	
3-Phospho-D-glycerate	3PG	cytosol	
3-trans-decenoyl-CoA		mitochondrial matrix	
3-ureidoisobutyric acid	UIBA	cytosol	
3-Ureidopropionate	UPROP	cytosol	
4' phosphopantetheine	PPANT	cytosol	
4-(2-aminophenyl)-2,4-dioxobutanoic acid		cytosol	
4-(2-aminophenyl)-2,4-dioxobutanoic acid		mitochondrial matrix	
4-(4-(dimethylamino)styryl)-N-methylpyridinium	4-Di-2-ASP	cytosol	http://pubs.rsc.org/en/content/articlelanding/2011/CS/C0CS00203H
4-(4-(dimethylamino)styryl)-N-methylpyridinium	4-Di-2-ASP	extracellular region	http://pubs.rsc.org/en/content/articlelanding/2011/CS/C0CS00203H
4,4-dimethylCHOLesta-8(9),14,24-trien-3beta-ol	4,4DMCHtOL	endoplasmic reticulum membrane	
4,4-dimethylCHOLesta-8(9),14,24-trien-3beta-ol	4,4DMCHtOL	nuclear envelope	
4,4-dimethylCHOLesta-8(9),24-dien-3beta-ol	14DMLANOL	endoplasmic reticulum membrane	
4,4-dimethylCHOLesta-8(9),24-dien-3beta-ol	14DMLANOL	nuclear envelope	
4,8,12-trimethyltridecanoyl-CoA	4,8,12-TMTD-CoA	peroxisomal matrix	http://www.ebi.ac.uk/chebi/searchId.do;jsessionid=428A7ADEA1A391863B0A2A35062EE56D?chebiId=CHEBI%3A15495
4,8-dimethylnonanoylcarnitine		peroxisomal matrix	
4,8-dimethylnonanoyl-CoA		peroxisomal matrix	
4a-hydroxytetrahydrobiopterin		cytosol	
4-aminobenzoate	ABENZ	cytosol	
4-Aminobiphenyl	ABPH	endoplasmic reticulum lumen	
4-aminosalicylate	ASAL	cytosol	
4-Androstene-3, 17-dione	ADION	cytosol	
4-carboxyCHOLesta-8(9),24-dien-3beta-ol	4CZYMOL	endoplasmic reticulum membrane	
4-CHOLesten-7alpha, 12alpha-diol-3-one		endoplasmic reticulum membrane	
4-CHOLesten-7alpha,12alpha,24(S)-triol-3-one		endoplasmic reticulum membrane	
4-CHOLesten-7alpha,12alpha,27-triol-3-one		endoplasmic reticulum membrane	
4-CHOLesten-7alpha,24(S)-diol-3-one		endoplasmic reticulum membrane	
4-CHOLesten-7alpha,27-diol-3-one		endoplasmic reticulum membrane	
4-CHOLesten-7alpha-ol-3-one		endoplasmic reticulum membrane	

name	short name	location	Notes
4-cis-decenoyl-CoA		mitochondrial matrix	
4-Fumarylacetoacetate	4FAA	cytosol	
4-hydroxy-9-cis-retinoic acid	HcRA	endoplasmic reticulum lumen	
4-hydroxy-all-trans-retinoic acid	HatRA	endoplasmic reticulum lumen	
4-hydroxybutyrate	HBUT	mitochondrial matrix	
4-hydroxycyclophosphamide	HCPM	endoplasmic reticulum lumen	
4-hydroxydebrisoquine	HDBQ	endoplasmic reticulum lumen	
4-hydroxyestradiol-17beta	4-OH-Estradiol	cytosol	http://pubchem.ncbi.nlm.nih.gov/summary/summary.cgi?cid=1743#Synonyms
4-hydroxy-L-proline	HPRO	cytosol	
4-hydroxy-L-proline	HPRO	extracellular region	
4-Hydroxyphenylacetaldehyde	HPHAC	cytosol	
4-hydroxyphenytoin	HPHT	endoplasmic reticulum lumen	
4-hydroxytolbutamide	HTOLB	endoplasmic reticulum lumen	
4-Imidazolone-5-propanoate		cytosol	
4-Maleylacetoacetate	4MAA	cytosol	
4-methyl,4-carboxyCHOLesta-8(9),24-dien-3beta	4C4MZYMOL	endoplasmic reticulum membrane	
4-methylCHOLesta-8(9),24-dien-3beta-ol	4MZYMOL	endoplasmic reticulum membrane	
4-methylCHOLesta-8(9),24-dien-3-one	dh4MZYMOL	endoplasmic reticulum membrane	
4-methylpentan-1-ol	MePeOH	cytosol	
4-Methylthio-2-oxobutanoate (1-)		cytosol	
4-Trimethylammoniobutanal	TEABL	cytosol	
4-Trimethylammoniobutanoate	TEABT	cytosol	
5,10-methenyltetrahydrofolate polyglutamate		cytosol	same as entry 297 below?
5,10-methylenetetrahydrofolate polyglutamate	MTHFPG	cytosol	
5,6-Dihydrothymine	DHT	cytosol	
5,6-Dihydrouracil	DHU	cytosol	
5,6-Dihydrouracil	DHU	nucleoplasm	
5,6-EET		endoplasmic reticulum lumen	
5-acetylamino-6-formylamino-3-methyluracil	AFMU	cytosol	http://www.chemspider.com/Chemical-Structure.97287.html
5-alpha-dihydrotestosterone	DHTEST	cytosol	
5-aminolevulinate	dALA	cytosol	
5-aminolevulinate	dALA	mitochondrial matrix	
5beta-CHOLestan-3alpha, 7alpha, 12alpha-triol		mitochondrial matrix	
5beta-CHOLestan-3alpha, 7alpha-diol		mitochondrial matrix	
5beta-CHOLestan-3alpha,7alpha,12alpha,24(S),2		mitochondrial matrix	
5beta-CHOLestan-3alpha,7alpha,12alpha,24(S)-t		cytosol	
5beta-CHOLestan-3alpha,7alpha,12alpha,24(S)-t		mitochondrial matrix	
5beta-CHOLestan-3alpha,7alpha,12alpha,27-tetr		cytosol	
5beta-CHOLestan-3alpha,7alpha,12alpha,27-tetr		mitochondrial matrix	
5beta-CHOLestan-3alpha,7alpha,12alpha-triol		cytosol	
5beta-CHOLestan-3alpha,7alpha,24(S),27-tetrol		mitochondrial matrix	
5beta-CHOLestan-3alpha,7alpha,24(S)-triol		cytosol	
5beta-CHOLestan-3alpha,7alpha,24(S)-triol		mitochondrial matrix	
5beta-CHOLestan-3alpha,7alpha,26-triol		cytosol	
5beta-CHOLestan-3alpha,7alpha,26-triol		mitochondrial matrix	
5beta-CHOLestan-3alpha,7alpha-diol		cytosol	
5beta-CHOLestan-7alpha,12alpha,24(S)-triol-3-or		cytosol	
5beta-CHOLestan-7alpha,12alpha,27-triol-3-one		cytosol	
5beta-CHOLestan-7alpha,24(S)-diol-3-one		cytosol	
5beta-CHOLestan-7alpha,27-diol-3-one		cytosol	
5beta-CHOLestan-7alpha-ol-3-one		cytosol	
5beta-CHOLesten-7alpha, 12alpha-diol-3-one		cytosol	
5'-Deoxyadenosine	dADE	cytosol	
5-formiminotetrahydrofolate	FITHF	cytosol	
5-HpETE		cytosol	
5-hydroxyindole acetaldehyde	HIALD	mitochondrial matrix	
5-Hydroxyindole acetic acid	HIAA	mitochondrial matrix	
5-Hydroxyindoleacetaldehyde	HIALD	cytosol	same as 329
5-hydroxyomeprazole	HOMZ	endoplasmic reticulum lumen	
5-Hydroxytryptamine	5HT	cytosol	
5-hydroxytryptamine N-methyl conjugate	5HT-N-CH3	cytosol	
5-hydroxytryptophan	5HTP	cytosol	http://en.wikipedia.org/wiki/5-Hydroxytryptophan
5-hydroxyuracil	5OHU	nucleoplasm	
5-L-gamma-glutamyl amino acid	GGAA	extracellular region	
5-methyltetrahydrofolate	MTHF	cytosol	http://en.wikipedia.org/wiki/Levomefolic_acid
5-methyltetrahydrofolate	MTHF	extracellular region	http://en.wikipedia.org/wiki/Levomefolic_acid
5-methyltetrahydrofolate polyglutamate	MTHFPG	cytosol	
5'-Methylthio-ribose-1P (2-)		cytosol	

name	short name	location	Notes
5'-Methylthio-ribulose-1P (2-)		cytosol	
5-oxo-ETE		extracellular region	
5-oxoproline	OPRO-	cytosol	
5-phosphoribosylamine	PRA	cytosol	
5-phosphoribosylglycinamide	PRGA	cytosol	
6beta-hydroxytestosterone	6BHT	endoplasmic reticulum lumen	
6-hydroxypaclitaxel	HPTXL	endoplasmic reticulum lumen	
6-Mercaptopurine	6MP	cytosol	
6-Mercaptopurine	6MP	endoplasmic reticulum lumen	
6-Methylmercaptopurine	6MMP	cytosol	
6-Methylmercaptopurine	6MMP	endoplasmic reticulum lumen	
6-Phospho-D-gluconate	PDG	cytosol	
7 alpha-hydroxyCHOLesterol	7HCHOL	endoplasmic reticulum membrane	
7-dehydroCHOLesterol	7DHCHOL	cytosol	
7-methylguanosine 5'-diphosphate	7-MeGDP	cytosol	
7-methylguanosine 5'-monophosphate	7-MeGMP	cytosol	
7-methylguanosine cap		cytosol	
7-methylguanosine cap		nucleoplasm	
8,11,14-Eicosatrienoic acid	8,11,14ETA	extracellular region	
8,9-EET	8,9-EET	endoplasmic reticulum lumen	
8-oxoguanine	OGUA	nucleoplasm	
9(S')-HODE	9S-HODE	nucleoplasm	
9-cis-retinoic acid	9cisRA	endoplasmic reticulum lumen	
9-O-acetylneuraminic acid	OANA-	Golgi lumen	
Acetaldehyde	CH3CHO	cytosol	
Acetaldehyde	CH3CHO	endoplasmic reticulum lumen	
Acetaldehyde	CH3CHO	mitochondrial matrix	
Acetaldehyde	CH3CHO	nucleoplasm	
Acetaminophen	ACAP	cytosol	
Acetaminophen (Paracetamol, TN TYLENOL)	ACAP	endoplasmic reticulum lumen	
Acetaminophen glutathione conjugate	ACAP-GSH	cytosol	
Acetaminophen glutathione conjugate	ACAP-GSH	endoplasmic reticulum lumen	
Acetaminophen O-glucuronide	ACAP-OGLU	endoplasmic reticulum lumen	
Acetaminophen O-sulfate conjugate	ACAP-SO4	cytosol	
Acetate	CH3COO-	cytosol	
Acetate	CH3COO-	extracellular region	
Acetate	CH3COO-	mitochondrial matrix	
Acetate	CH3COO-	peroxisomal matrix	
Acetoacetate	ACA	cytosol	
Acetoacetate	ACA	extracellular region	
acetoacetate	ACA	mitochondrial matrix	
acetoacetyl-CoA	ACA-CoA	cytosol	
acetoacetyl-CoA	ACA-CoA	mitochondrial matrix	
acetophenone	APN	cytosol	
acetophenone-mercapturic acid	APN-MTA	cytosol	
acetophenone-mycothiol conjugate	APN-MYT	cytosol	
Acetylcarnitine	ACAR	peroxisomal matrix	
Acetyl-CoA	Ac-CoA	cytosol	
Acetyl-CoA	Ac-CoA	Golgi lumen	
Acetyl-CoA	Ac-CoA	mitochondrial matrix	
Acetyl-CoA	Ac-CoA	nucleoplasm	
Acetyl-CoA	Ac-CoA	peroxisomal matrix	
Acyclovir	ACV	cytosol	
Acyclovir	ACV	extracellular region	
adenine	Ade	cytosol	
adenine	Ade	extracellular region	
adenine	Ade	lysosomal lumen	
adenosine	Ade-Rib	cytosol	
Adenosine	Ade-Rib	extracellular region	
adenosine	Ade-Rib	lysosomal lumen	
adenosine 5'-monophosphate	AMP	extracellular region	
Adenosine 5'-phosphosulfate	APS	cytosol	
adenylosuccinate	ADS	cytosol	
ADP		cytosol	
ADP		endoplasmic reticulum membrane	
ADP		endosome	
ADP		extracellular region	
ADP		mitochondrial matrix	

name	short name	location	Notes
ADP		nucleoplasm	
ADP		plasma membrane	
ADP		platelet dense granule lumen	
Adrenaline	ADR	clathrin sculpted monoamine transport vesicle lum	
Adrenaline	ADR	cytosol	
Adrenaline	ADR	extracellular region	
Adrenaline	ADR	extracellular region	
Aflatoxin B1	AFTB1	endoplasmic reticulum lumen	
Aflatoxin exo-8,9-oxide	AFBO	endoplasmic reticulum lumen	
Aflatoxin exo-8,9-oxide glutathione conjugate	AFBO-SG	endoplasmic reticulum lumen	
Agmatine	AGM	cytosol	
Agmatine	AGM	extracellular region	
AICAR	AICAR	cytosol	
AIR	AIR	cytosol	
Ala-tRNA(Ala)	Ala-tRNA(Ala)	cytosol	
Ala-tRNA(Ala)	Ala-tRNA(Ala)	mitochondrial matrix	
Aldosterone	ALDO	mitochondrial matrix	
all-cis-icosa-pentaenoic acid		extracellular region	
all-trans-retinal	atRAL	cytosol	
All-trans-retinoic acid	atRA	endoplasmic reticulum lumen	
all-trans-retinol	atROL	plasma membrane	
alpha,alpha'-trehalose		cytosol	
alpha,alpha-trehalose		extracellular region	
alpha,alpha'-trehalose-6-phosphate		cytosol	
alpha-D-Galactose 1-phosphate		cytosol	
alpha-D-glucose	Glc	cytosol	Glc is better
alpha-D-glucose	Glc	endoplasmic reticulum lumen	Glc is better
alpha-D-glucose	Glc	extracellular region	Glc is better
alpha-D-glucose	Glc	Golgi lumen	Glc is better
alpha-D-glucose	Glc	nucleoplasm	Glc is better
alpha-D-Glucose 6-phosphate	Glc6P	cytosol	
alpha-D-Glucose 6-phosphate	Glc6P	endoplasmic reticulum lumen	
alpha-ketoadipate		cytosol	
alpha-ketoadipate		mitochondrial matrix	
alpha-keto-beta-methylvalerate	KMVA	cytosol	http://ctd.mdibl.org/detail.go?type=chem&acc=C016211
alpha-keto-beta-methylvalerate	KMVA	mitochondrial matrix	http://ctd.mdibl.org/detail.go?type=chem&acc=C016211
alpha-ketoisocaproate	KIC	cytosol	http://www.nutriline.org/article/8
alpha-ketoisocaproate	KIC	mitochondrial matrix	http://www.nutriline.org/article/8
alpha-ketoisovalerate	KIV	cytosol	
alpha-ketoisovalerate	KIV	mitochondrial matrix	
alpha-Linolenic acid	ALA	extracellular region	http://en.wikipedia.org/wiki/Polyunsaturated_fatty_acid
alpha-linolenic acid	ALA	nucleoplasm	http://en.wikipedia.org/wiki/Polyunsaturated_fatty_acid
alpha-methylacetoacetyl-CoA		mitochondrial matrix	
alpha-methyl-beta-hydroxybutyryl-CoA		mitochondrial matrix	
alpha-methylbutyryl-CoA		mitochondrial matrix	
AMP		cytosol	
AMP		mitochondrial matrix	
AMP		nucleoplasm	
AMP		peroxisomal matrix	
Anandamide	AEA	extracellular region	http://en.wikipedia.org/wiki/Anandamide
arachidic acid	AA	cytosol	
arachidic acid	AA	extracellular region	
Arachidonate	AA	cytosol	
Arachidonate	AA	endoplasmic reticulum lumen	
Arachidonate	AA	plasma membrane	
arachidonic acid	ARA	nucleoplasm	http://en.wikipedia.org/wiki/Arachidonic_acid
arachidonoyl-CoA		cytosol	
arachidoyl-CoA		cytosol	
Arg-tRNA(Arg)		cytosol	
Arg-tRNA(Arg)		mitochondrial matrix	
Asbestos	Asb	cytosol	
Ascorbate	VitC	cytosol	
ascorbate	VitC	extracellular region	
Ascorbate	VitC	stored secretory granule	
Asn-tRNA(Asn)		cytosol	
Asn-tRNA(Asn)		mitochondrial matrix	
Aspartame	ASP	extracellular region	
Asp-tRNA(Asp)		cytosol	

name	short name	location	Notes
Asp-tRNA(Asp)		mitochondrial matrix	
ATP		cytosol	
ATP		endosome	
ATP		extracellular region	
ATP		mitochondrial matrix	
ATP		nucleoplasm	
ATP		peroxisomal matrix	
ATP		plasma membrane	
ATP		platelet dense granule lumen	
Benzene	C6H6	endoplasmic reticulum lumen	
Benzo(a)pyrene-7,8-diol 9,10-epoxide	BPDE	cytosol	
Benzo(a)pyrene-7,8-diol-glutathione conjugate	BPD-SG	cytosol	
Benzoic acid	BENZA	mitochondrial matrix	
Benzoyl-CoA thioester	BEZ-CoA	mitochondrial matrix	
beta-Alanine	b-Ala	cytosol	
beta-Alanine	b-Ala	extracellular region	
beta-Alanine	b-Ala	mitochondrial matrix	
Beta-carotene	b-CARO	plasma membrane	
beta-D-mannosyldiacetylchitobiosyldiphosphodol	MDCDD	integral to cytosolic side of endoplasmic reticulum	
beta-estradiol	E2	cytosol	
beta-estradiol 3-sulfate	E2-SO4	cytosol	
beta-hydroxy-beta-methylglutaryl-CoA	bHMG-CoA	cytosol	
beta-hydroxy-beta-methylglutaryl-CoA	bHMG-CoA	mitochondrial matrix	
Beta-hydroxybutyrate	bHBA	cytosol	
Beta-hydroxybutyrate	bHBA	extracellular region	
beta-hydroxyisobutyrate	bHIBA	mitochondrial matrix	
beta-hydroxyisobutyryl-CoA	bHIB-CoA	mitochondrial matrix	
betaine	BET	cytosol	
betaine	BET	extracellular region	
beta-methylcrotonyl-CoA	bMC-CoA	mitochondrial matrix	general comment about greek letters. Where we have a, b, d, g etc to denote greek characters, I hope we can fill with the actual character.
beta-methylglutaconyl-CoA	bMC-CoA	mitochondrial matrix	
Bilirubin	BIL	cytosol	
Bilirubin	BIL	endoplasmic reticulum lumen	
Bilirubin diglucuronide	BDG	endoplasmic reticulum lumen	
Bilirubin monoglucuronide	BMG	endoplasmic reticulum lumen	
biliverdin	BV	cytosol	
Biotin	BTN	cytosol	
Biotin	BTN	extracellular region	
Biotin	BTN	mitochondrial matrix	
Bromide	Br-	cytosol	
Bromide	Br-	transport vesicle	
bromosulphophthalein	BSP	cytosol	
bromosulphophthalein	BSP	extracellular region	
Butanoyl-CoA	BT-CoA	mitochondrial matrix	
Butyrate	BUT	cytosol	
Butyrate	BUT	extracellular region	
Caffeine	CAF	endoplasmic reticulum lumen	
CAIR	CAIR	cytosol	
Calcidiol	CDL	cytosol	
Calcidiol	CDL	extracellular region	
Calcidiol	CDL	lysosomal lumen	
Calcitriol	CDL	cytosol	
Calcitroic acid	CTA	cytosol	
Calcium	Ca2+	cytosol	
Calcium	Ca2+	endoplasmic reticulum lumen	
Calcium	Ca2+	extracellular region	
Calcium	Ca2+	Golgi lumen	
Calcium	Ca2+	nucleoplasm	
Calcium	Ca2+	plasma membrane	
Calcium	Ca2+	platelet dense granule lumen	
Calcium	Ca2+	platelet dense tubular network lumen	
Calcium	Ca2+	platelet dense tubular network membrane	
Calcium	Ca2+	stored secretory granule	
carbamoyl phosphate	CAP	cytosol	
Carbamoyl phosphate	CAP	mitochondrial matrix	
carbon monoxide	CO	cytosol	
Carbon tetrachloride	CCl4	endoplasmic reticulum lumen	
Carnitine	CAR	cytosol	

name	short name	location	Notes
Carnitine	CAR	extracellular region	
Carnitine	CAR	mitochondrial intermembrane space	
Carnitine	CAR	mitochondrial matrix	
Carnitine	CAR	peroxisomal matrix	
Ceramide	CERA	cytosol	
ceramide	CERA	endoplasmic reticulum membrane	
ceramide	CERA	Golgi membrane	
ceramide	CERA	plasma membrane	
chenodeoxyCHOLate	CDCA	cytosol	
chenodeoxyCHOLate	CDCA	extracellular region	
ChenodeoxyCHOLic acid	CDCA	extracellular region	
chenodeoxyCHOLoyl-CoA	CDC-CoA	cytosol	
chenodeoxyCHOLoyl-CoA	CDC-CoA	peroxisomal matrix	
Chloride	Cl-	cytosol	
Chloride	Cl-	endoplasmic reticulum lumen	
Chloride	Cl-	extracellular region	
Chloride	Cl-	transport vesicle	
Chloroquine	CLQ	cytosol	
Chloroquine	CLQ	extracellular region	
CHOLate	CCA	cytosol	
CHOLate	CCA	extracellular region	
CHOLate	CCA	peroxisomal matrix	
CHOLest-5-ene-3beta,7alpha,24(S)-triol		endoplasmic reticulum membrane	
CHOLest-5-ene-3beta,7alpha,25-triol		cytosol	
CHOLest-5-ene-3beta,7alpha,27-triol		endoplasmic reticulum membrane	
CHOLesta-5,7,24-trien-3beta-ol	7dhDESOL	endoplasmic reticulum membrane	
CHOLesta-7,24-dien-3beta-ol	CHdOL	endoplasmic reticulum membrane	
CHOLesta-8(9),24-dien-3-one	ZYMONE	endoplasmic reticulum membrane	
CHOLesterol	CHOL	clathrin-coated endocytic vesicle membrane	
CHOLesterol	CHOL	cytosol	
CHOLesterol	CHOL	endoplasmic reticulum lumen	
CHOLesterol	CHOL	endoplasmic reticulum membrane	
CHOLesterol	CHOL	endosome lumen	
CHOLesterol	CHOL	endosome membrane	
CHOLesterol	CHOL	extracellular region	
CHOLesterol	CHOL	lipid particle	
CHOLesterol	CHOL	mitochondrial inner membrane	
CHOLesterol	CHOL	mitochondrial intermembrane space	
CHOLesterol	CHOL	mitochondrial matrix	
CHOLesterol	CHOL	plasma membrane	
CHOLesterol	CHOL	transport vesicle membrane	
CHOLesterol esters	CHEST	clathrin-coated endocytic vesicle membrane	
CHOLesterol esters	CHEST	endoplasmic reticulum lumen	
CHOLesterol esters	CHEST	endosome lumen	
CHOLesterol esters	CHEST	endosome membrane	
CHOLesterol esters	CHEST	extracellular region	
CHOLesterol esters	CHEST	plasma membrane	
CHOLesterol sulfate	CHOL-SO4	cytosol	
CHOLic acid	CCA	extracellular region	
CHOLine	Cho	cytosol	
CHOLine	Cho	extracellular region	
CHOLine phosphate	ChoP	cytosol	
CHOLoyl-CoA	CHO-CoA	cytosol	
CHOLoyl-CoA	CHO-CoA	peroxisomal matrix	
Chondroitin sulfate	CHS	extracellular region	
Chorismate	CRSM	cytosol	
Cimetidine	CIM	cytosol	
Cimetidine	CIM	extracellular region	
cis,cis-3,6-Dodecadienoyl-CoA		mitochondrial matrix	
Citrate	CIT	cytosol	
Citrate	CIT	extracellular region	
Citrate	CIT	mitochondrial matrix	
Clonidine	CLON	cytosol	
Clonidine	CLON	extracellular region	
CMP	CMP	nucleoplasm	
CMP-sialic acid	CMP-SA	cytosol	
CMP-sialic acid	CMP-SA	Golgi lumen	
CO2		cytosol	

name	short name	location	Notes
CO2		endoplasmic reticulum lumen	
CO2		mitochondrial intermembrane space	
CO2		mitochondrial matrix	
CO2		nucleoplasm	
CO2		peroxisomal matrix	
Co2+		mitochondrial matrix	
CoA-SH		cytosol	
CoA-SH		endoplasmic reticulum lumen	
CoA-SH		mitochondrial matrix	
CoA-SH		peroxisomal matrix	
Cobamide coenzyme	VitB12 coenzyme	mitochondrial matrix	
Copper	Cu2+	cytosol	
Copper	Cu2+	extracellular region	
Copper	Cu2+	Golgi lumen	
Copper	Cu2+	plasma membrane	
Copper	Cu2+	stored secretory granule	
Copper A	CuA	mitochondrial inner membrane	
Coproporphyrinogen I	COPROI	cytosol	
Coproporphyrinogen III	COPROII	cytosol	
Coproporphyrinogen III	COPROIII	mitochondrial intermembrane space	
Corticosterone	CORST	mitochondrial matrix	
Cortisol	CORT	cytosol	
Cortisol	CORT	mitochondrial matrix	
Cortisone	COR	cytosol	
Coumarin	COUM	endoplasmic reticulum lumen	
Creatine	CRET	cytosol	
Creatine	CRET	extracellular region	
Creatinine	CREAT	cytosol	
Creatinine	CREAT	extracellular region	
Crotonoyl-CoA		mitochondrial matrix	
crotonyl-CoA		mitochondrial matrix	
CTP		cytosol	
CTP		mitochondrial matrix	
CTP		nucleoplasm	
Cyclophosphamide	CPM	endoplasmic reticulum lumen	
cystathionine	CYSTA	cytosol	
cysteinylglycine	Cys-Gly	cytosol	
cysteinylglycine	Cys-Gly	extracellular region	
Cys-tRNA(Cys)		cytosol	
Cys-tRNA(Cys)		mitochondrial matrix	
cytidine	Cyt-Rib	cytosol	
cytidine	Cyt-Rib	extracellular region	
cytidine	Cyt-Rib	lysosomal lumen	
cytidine 5'-diphosphate	CDP	cytosol	
cytidine 5'-diphosphate	CDP	mitochondrial matrix	
cytidine 5'-diphosphate	CDP	nucleoplasm	
cytidine 5'-monophosphate	CMP	cytosol	
cytidine 5'-monophosphate	CMP	extracellular region	
cytidine 5'-monophosphate	CMP	Golgi lumen	
cytosine	Cyt	cytosol	
cytosine	Cyt	extracellular region	
DAHP(3-)	DAHP(3-)	cytosol	
D-Alanine	D-Ala	cytosol	
D-Alanine	D-Ala	extracellular region	
D-Aspartate	D-Asp	cytosol	
D-Aspartate	D-Asp	extracellular region	
dATP	dATP	cytosol	
dATP	dATP	mitochondrial matrix	
dATP	dATP	nucleoplasm	
D-beta hydroxybutyrate	bHBA	mitochondrial matrix	
D-chiro-inositol	CHI	cytosol	
D-chiro-inositol	CHI	extracellular region	
dCMP	dCMP	cytosol	
dCTP	dCTP	cytosol	
dCTP	dCTP	mitochondrial matrix	
dCTP	dCTP	nucleoplasm	
Deamino-NAD+	DA-NAD+	cytosol	
Deamino-NAD+	DA-NAD+	nucleoplasm	

name	short name	location	Notes
Debrisoquine	DSQ	endoplasmic reticulum lumen	
Debrisoquine O-glucuronide	DSQ-OG	endoplasmic reticulum lumen	
decanoic acid	DECA	cytosol	
Decanoic acid	DECA	endoplasmic reticulum lumen	
decanoic acid	DECA	extracellular region	
Decanoyl-CoA	DEC-CoA	mitochondrial matrix	
Dehydroascorbate	DHvitC	cytosol	
dehydroascorbate	DHvitC	extracellular region	
Dehydroascorbate	DHvitC	stored secretory granule	
dehydroepiandrosterone	DHEA	cytosol	
Dehydroepiandrosterone	DHEA	cytosol	
dehydroepiandrosterone sulfate	DHEA-SO4	cytosol	
Dehydroepiandro-sterone sulfate	DHEA-SO4	cytosol	
Dehydroepiandro-sterone sulfate	DHEA-SO4	extracellular region	
delta3-isopentenyl pyrophosphate	IPPP	cytosol	
DeoxyCHOLic acid	DOCA	extracellular region	
Dephospho-CoA	DP-CoA	cytosol	
D-Erythrose 4-phosphate	E4P	cytosol	
desacetylmycothiol	DAMT	cytosol	
Desipramine	DESI	cytosol	
Desipramine	DESI	extracellular region	
desmosterol	DESMOL	endoplasmic reticulum membrane	
Dexamethasone	DEXA	nucleoplasm	
Dextromethorphan	DEXM	endoplasmic reticulum lumen	
Dextrorphan	DEXT	endoplasmic reticulum lumen	
D-fructose	Fru	cytosol	
D-fructose	Fru	extracellular region	
D-Fructose 1,6-bisphosphate	Fru(1,6)P2	cytosol	
D-Fructose 1-phosphate	Fru(1)P	cytosol	
D-Fructose 2,6-bisphosphate	Fru(2,6)P2	cytosol	
D-Fructose 6-phosphate	Fru(6)P	cytosol	
D-galactose	Gal	cytosol	
D-galactose	Gal	extracellular region	
D-Glucono-1,5-lactone 6-phosphate		cytosol	
D-glucosamine 6-phosphate	GlcN6P	cytosol	
D-Glucose 1-phosphate	Glc1P	cytosol	
D-Glyceraldehyde	GA	cytosol	
D-Glyceraldehyde 3-phosphate	GA3P	cytosol	
dGTP		cytosol	
dGTP		mitochondrial matrix	
dGTP		nucleoplasm	
DHCA		cytosol	
DHCA		mitochondrial matrix	
diacylglycerols	DAGs	extracellular region	
diacylglycerols	DAGs	Golgi membrane	
diacylglycerols	DAGs	plasma membrane	
diaminobutyrate	DAB	cytosol	
diaminobutyrate	DAB	extracellular region	
digoxin	DIGX	cytosol	
digoxin	DIGX	extracellular region	
dihydroceramide	DHCE	endoplasmic reticulum membrane	
Dihydrofolate	DHF	cytosol	
Dihydroxyacetone phosphate	DHAP	cytosol	
Dihydroxyacetone phosphate	DHAP	mitochondrial matrix	
Dihydroxyacetone phosphate	DHAP	peroxisomal matrix	
Diiodotyrosine	DIT	extracellular region	
Dimethylallyl pyrophosphate	DMAPP	cytosol	
dipyrromethane	DIPY	cytosol	
D-mannose	Man	cytosol	
D-mannose	Man	endoplasmic reticulum quality control compartmer	
D-mannose	Man	extracellular region	
D-mannose	Man	Golgi lumen	
D-Mannose 1-phosphate	Man1P	cytosol	
D-Mannose 6-phosphate	Man6P	cytosol	
D-methylmalonyl-CoA	MEMA-CoA	mitochondrial matrix	
Docosahexaenoic acid	DHA	extracellular region	http://en.wikipedia.org/wiki/Polyunsaturated_fatty_acid
Docosapentaenoic acid	DPA	extracellular region	http://en.wikipedia.org/wiki/Polyunsaturated_fatty_acid
Docosatetraenoic acid	DTTA	extracellular region	

name	short name	location	Notes
Dodecylcarboxylate	DDCX	cytosol	
Dodecylcarboxylate	DDCX	endoplasmic reticulum lumen	
Dodecylcarboxylate	DDCX	extracellular region	
DoliCHOL	DOL	integral to cytosolic side of endoplasmic reticulum	
Dolichyl beta-D-glucosyl phosphate	DbGP	endoplasmic reticulum quality control compartmer	
Dolichyl beta-D-glucosyl phosphate	DbGP	integral to cytosolic side of endoplasmic reticulum	
Dolichyl beta-D-glucosyl phosphate	DbGP	integral to lumenal side of endoplasmic reticulum	
Dolichyl diphosphate	DOLDP	integral to cytosolic side of endoplasmic reticulum	
Dolichyl phosphate	DOLP	endoplasmic reticulum quality control compartmer	
Dolichyl phosphate	DOLP	integral to cytosolic side of endoplasmic reticulum	
Dolichyl phosphate	DOLP	integral to lumenal side of endoplasmic reticulum	
Dolichyl phosphate D-mannose	DOLPM	integral to lumenal side of endoplasmic reticulum	
Dopa	Dopa	cytosol	
Dopamine	DA	clathrin sculpted monoamine transport vesicle lum	
Dopamine	DA	cytosol	
Dopamine	DA	extracellular region	
Dopamine 3-O-sulfate	DAOS	cytosol	
D-ribose 1-phosphate	R1P	cytosol	
D-ribose 5-phosphate	R5P	cytosol	
D-Ribulose 5-phosphate	RU5P	cytosol	
D-Serine	D-Ser	cytosol	
D-Serine	D-Ser	extracellular region	
D-sorbitol 6-phosphate	Sor6P	cytosol	
D-tryptophan	D-Trp	extracellular region	
dTTP		nucleoplasm	
dUMP		cytosol	
dUMP		nucleoplasm	
dUTP		cytosol	
dUTP		mitochondrial matrix	
dUTP		nucleoplasm	
D-Xylulose 5-phosphate	XY5P	cytosol	
e-		endosome membrane	
e-		extracellular region	
eicosapentaenoic acid	EPA	nucleoplasm	http://en.wikipedia.org/wiki/Polyunsaturated_fatty_acid
Elaidic acid	ELDA	extracellular region	
electron	e-	cytosol	
electron	e-	endoplasmic reticulum lumen	
EPSP(4-)	EPSP(4-)	cytosol	
Ergothioneine	ERGT	cytosol	
Ergothioneine	ERGT	extracellular region	
Estradiol-17beta	EST17b	cytosol	
Estrogen	ESTG	extracellular region	Is a collective name
estrone	E1	cytosol	http://www.gfmer.ch/Endo/Lectures_08/sexual_hormones.htm
Estrone	E1	cytosol	http://www.gfmer.ch/Endo/Lectures_08/sexual_hormones.htm
estrone 3-sulfate	E1S	cytosol	http://www.gfmer.ch/Endo/Lectures_08/sexual_hormones.htm
estrone 3-sulfate	E1S	extracellular region	http://www.gfmer.ch/Endo/Lectures_08/sexual_hormones.htm
Ethanol	EtOH	cytosol	
Ethanol	EtOH	endoplasmic reticulum lumen	
ethenoadenine	EtAD	nucleoplasm	
ethenocytosine	EtCYS	nucleoplasm	
Ethylene	C2H4	endoplasmic reticulum lumen	
Ethylene oxide	EtO	endoplasmic reticulum lumen	
ezetimibe	EZE	extracellular region	
ezetimibe	EZE	plasma membrane	
FAD		cytosol	
FAD		endoplasmic reticulum membrane	
FAD		mitochondrial inner membrane	
FAD		mitochondrial intermembrane space	
FAD		mitochondrial matrix	
FAD		mitochondrial outer membrane	
FAD		peroxisomal matrix	
FADH2		mitochondrial inner membrane	
FADH2		mitochondrial matrix	
FAICAR		cytosol	
Farnesyl pyrophosphate	FAPP	cytosol	
Fatty Acids	FAs	extracellular region	
fatty acyl CoA	FACoA	cytosol	
Fe++	Fe2+	peroxisomal matrix	

name	short name	location	Notes
Fe2+	Fe2+	phagocytic vesicle	
ferriheme	FeHM	cytosol	
ferriheme	FeHM	extracellular region	
ferriheme	FeHM	mitochondrial outer membrane	
FGAM	FGAM	cytosol	
Flavin mononucleotide	FMN	cytosol	
flavin mononucleotide	FMN	mitochondrial inner membrane	
Flavin mononucleotide	FMN	peroxisomal matrix	
Fluoride	Fl-	cytosol	
Fluoride	Fl-	transport vesicle	
fMet-Leu-Phe	FMLP	extracellular region	
Folate	FOLA	cytosol	
Folate	FOLA	extracellular region	
formaldehyde	CH2O	cytosol	
Formaldehyde	CH2O	endoplasmic reticulum lumen	
Formaldehyde	CH2O	nucleoplasm	
formamidopyrimidine	FAPY	nucleoplasm	
Formate	HCOOH	cytosol	
Formate	HCOOH	extracellular region	
Formate	HCOOH	peroxisomal matrix	
Formyl-CoA	FOR-CoA	peroxisomal matrix	
Formylkynurenine	NFK	cytosol	
formylmycothiol	FMYC	cytosol	
Fumarate	FUMA	cytosol	
Fumarate	FUMA	mitochondrial matrix	
GABA		clathrin sculpted gamma-aminobutyric acid transp	
GABA		cytosol	
GABA		extracellular region	
GABA		mitochondrial matrix	
GABA		synaptic vesicle	
gamma-glutamyl group	gGLUT	cytosol	
gamma-glutamyl group	gGLUT	extracellular region	
gamma-glutamyl-L-cysteine	gGluCys	cytosol	
gamma-Linolenic acid	GLA	extracellular region	http://en.wikipedia.org/wiki/Polyunsaturated_fatty_acid
GDP		cytosol	
GDP		endosome	
GDP		extracellular region	
GDP		Golgi lumen	
GDP		mitochondrial matrix	
GDP		nucleoplasm	
GDP		plasma membrane	
GDP		platelet dense granule lumen	
GDP-alpha-D-mannose	GDP-Man	cytosol	
GDP-alpha-D-mannose	GDP-Man	Golgi lumen	
GDP-L-fucose	GDP-Fuc	cytosol	
GDP-L-fucose	GDP-Fuc	Golgi lumen	
Geranyl pyrophosphate	GPP	cytosol	
GlcNAc-Ins	GlcNAc-Ins	cytosol	
GlcNAc-Ins-P	GlcNAc-Ins-P	cytosol	
Gln-tRNA(Gln)	Gln-tRNA(Gln)	cytosol	
Gln-tRNA(Gln)	Gln-tRNA(Gln)	mitochondrial matrix	
glucosaminyl-inositol	GlcNI	cytosol	
glutaryl-CoA	GL-CoA	mitochondrial matrix	
Glutathione	GSH	cytosol	
Glutathione	GSH	endoplasmic reticulum lumen	
Glutathione	GSH	extracellular region	
Glu-tRNA(Glu)		cytosol	
Glu-tRNA(Glu)		mitochondrial matrix	
Glycerol		cytosol	
glycerol	Glycerol	extracellular region	
Glycerol-3-phosphate	G3P	cytosol	
Glycerol-3-phosphate	G3P	mitochondrial matrix	
glycine	Gly	endoplasmic reticulum lumen	
Glycine	Gly	extracellular region	
Glycine	Gly	lysosomal lumen	
Glycine	Gly	mitochondrial intermembrane space	
Glycine	Gly	mitochondrial matrix	
glycine	Gly	peroxisomal matrix	

name	short name	location	Notes
Glycine	Gly	synaptic vesicle	
glycochenodeoxyCHOLate	GCDC	cytosol	
glycochenodeoxyCHOLate	GCDC	extracellular region	
glycochenodeoxyCHOLate	GCDC	peroxisomal matrix	
glycoCHOLate	GC	cytosol	
glycoCHOLate	GC	extracellular region	
glycoCHOLate	GC	peroxisomal matrix	
glycogen	Glyg		
Glycolate	Glycolate	peroxisomal matrix	
Glycosaminoglycan	GAG	extracellular region	http://en.wikipedia.org/wiki/Glycosaminoglycan
Glyoxylate		mitochondrial matrix	
Glyoxylate		peroxisomal matrix	
Gly-tRNA(Gly)		cytosol	
Gly-tRNA(Gly)		mitochondrial matrix	
GMP		cytosol	
GMP		extracellular region	
GMP		Golgi lumen	
GMP		nucleoplasm	
GTP		cytosol	
GTP		endosome	
GTP		ER to Golgi transport vesicle membrane	
GTP		extracellular region	
GTP		mitochondrial matrix	
GTP		nucleoplasm	
GTP		plasma membrane	
GTP		platelet dense granule lumen	
Guanidine	Gu	cytosol	
Guanidine	Gu	extracellular region	
Guanidinoacetate	GAA	cytosol	
Guanidinoacetate	GAA	mitochondrial intermembrane space	
guanine	Gua	cytosol	
guanine	Gua	extracellular region	
guanosine	Gua-Rib	cytosol	
guanosine	Gua-Rib	extracellular region	
guanosine	Gua-Rib	lysosomal lumen	
H+		clathrin sculpted acetylCHOLine transport vesicle l	
H+		clathrin sculpted monoamine transport vesicle lum	
H+		cytosol	
H+		early endosome lumen	
H+		endoplasmic reticulum lumen	
H+		endosome lumen	
H+		extracellular region	
H+		Golgi lumen	
H+		late endosome lumen	
H+		lysosomal lumen	
H+		mitochondrial intermembrane space	
H+		mitochondrial matrix	
H+		nuclear envelope	
H+		peroxisomal matrix	
H+		phagocytic vesicle	
H+		platelet dense tubular network lumen	
H+		synaptic vesicle	
H2O		cytosol	
H2O		endoplasmic reticulum lumen	
H2O		extracellular region	
H2O		mitochondrial matrix	
H2O		nucleoplasm	
H2O		peroxisomal matrix	
H2O		plasma membrane	
H2O		stored secretory granule	
H2O2		cytosol	
H2O2		extracellular region	
H2O2		mitochondrial intermembrane space	
H2O2		peroxisomal matrix	
Halothane	HALO	endoplasmic reticulum lumen	
HCO3-		cytosol	
HCO3-		extracellular region	
HCO3-		mitochondrial matrix	

name	short name	location	Notes
heme		cytosol	
heme		extracellular region	
heme		mitochondrial matrix	
heme		mitochondrial outer membrane	
heme		peroxisomal matrix	
heme		plasma membrane	
Heparan 4-sulfate	H4S		
Heparan 6-sulfate	H6S		
Heparan 4,6-disulfate	HSE		
Heparin	HPIN		
Heparan sulfate alpha-D-glucosaminide		extracellular region	
Heparan sulfate N-acetyl-alpha-D-glucosaminide		extracellular region	
Hexacosanoyl-CoA	C26:0 CoA	cytosol	http://www.hmdb.ca/metabolites/HMDB06459
Hexacosanoyl-CoA		peroxisomal matrix	http://www.hmdb.ca/metabolites/HMDB06459
hexadec-2-enal		cytosol	
hexadecanal	HXAL	cytosol	
Hexadecanol	HXOL	peroxisomal matrix	
Hexanoate	HXA	extracellular region	
Hexanoyl-CoA	HX-CoA	mitochondrial matrix	
Hippuric acid	HIPA	mitochondrial matrix	
Histamine	Hist	clathrin sculpted monoamine transport vesicle lum	
Histamine	Hist	cytosol	
Histamine	Hist	extracellular region	
His-tRNA(His)		cytosol	
His-tRNA(His)		mitochondrial matrix	
HMG CoA		mitochondrial matrix	
HNK-1 carbohydrate		plasma membrane	
Homocysteine	HCYS	cytosol	
Homogentisate	HGTA	cytosol	
Hyaluronic acid	HUA	cytosol	
Hydrogen	H2	endoplasmic reticulum lumen	
Hydrogen bromide	HBr	endoplasmic reticulum lumen	
hydrogen iodide	HI	extracellular region	
Hydroxycoumarin	HCOU	endoplasmic reticulum lumen	
Hydroxymethylbilane	HMBL	cytosol	
hypoxanthine	Hyp	cytosol	
hypoxanthine	Hyp	extracellular region	
Hypoxanthine	Hyp	nucleoplasm	
iE-DAP		cytosol	
Ile-tRNA(Ile)		cytosol	
Ile-tRNA(Ile)		mitochondrial matrix	
Imipramine	IMIP	cytosol	
Imipramine	IMIP	extracellular region	
Imiquimod	IMIQ	cytosol	
Imiquimod	IMIQ	endosome membrane	
IMP	IMP	cytosol	
Indole	INDOL	endoplasmic reticulum lumen	
inorganic sulfate	SO4(2-)	cytosol	
inorganic sulfate	SO4(2-)	extracellular region	
inosine	Ino	cytosol	
inosine	Ino	extracellular region	
inosine	Ino	lysosomal lumen	
inosine 5'-monophosphate	IMP	extracellular region	
Inositol	Ins	cytosol	
Inositol	Ins	extracellular region	
Iodide	I-	cytosol	
Iodide	I-	extracellular region	
Iodide	I-	transport vesicle	
Iodine	I2	extracellular region	
iron (3+)	Fe3+	cytosol	
iron (3+)	Fe3+	extracellular region	
iron (ferric)	Fe3+	cytosol	
iron (ferric)	Fe3+	endosome lumen	
iron (ferric)	Fe3+	endosome membrane	
iron (ferric)	Fe3+	extracellular region	
iron (ferric)	Fe3+	mitochondrial inner membrane	
iron (ferrous)	Fe2+	cytosol	
iron (ferrous)	Fe2+	endosome lumen	

name	short name	location	Notes
iron (ferrous)	Fe2+	extracellular region	
iron (ferrous)	Fe2+	mitochondrial inner membrane	
iron (ferrous)	Fe2+	mitochondrial matrix	
iron (ferrous)	Fe2+	nucleoplasm	
iron (oxidation state not specified)	Fe	cytosol	
isobutyryl-CoA	ISB-CoA	mitochondrial matrix	
Isocaproic aldehyde	ISCAL	cytosol	
Isocaproic aldehyde	ISCAL	mitochondrial matrix	
Isocitrate	ISCIT	cytosol	
Isocitrate	ISCIT	mitochondrial matrix	
Isocitrate	ISCIT	peroxisomal matrix	
Isoniazid	ISNZ	cytosol	
Isoprotenerol	ISPOL	extracellular region	
isovaleryl-CoA	ISV-CoA	mitochondrial matrix	
kynurenic acid	KYNA	cytosol	
kynurenic acid	KYNA	mitochondrial matrix	
L-2-Aminoadipate	2AMA	mitochondrial matrix	
L-2-Aminoadipate 6-semialdehyde	2AMAS	mitochondrial matrix	
lactose	Lac	extracellular region	
L-Alanine	L-Ala	cytosol	
L-Alanine	L-Ala	endoplasmic reticulum lumen	
L-Alanine	L-Ala	extracellular region	
L-Alanine	L-Ala	lysosomal lumen	
L-Alanine	L-Ala	mitochondrial matrix	
L-Alanine	L-Ala	peroxisomal matrix	
lanosterol	LNSOL	endoplasmic reticulum membrane	
L-Arginine	L-Arg	cytosol	
L-Arginine	L-Arg	endoplasmic reticulum lumen	
L-Arginine	L-Arg	extracellular region	
L-Arginine	L-Arg	lysosomal lumen	
L-Arginine	L-Arg	mitochondrial intermembrane space	
L-Arginine	L-Arg	mitochondrial matrix	
L-Asparagine	L-Asn	cytosol	
L-Asparagine	L-Asn	endoplasmic reticulum lumen	
L-Asparagine	L-Asn	extracellular region	
L-Asparagine	L-Asn	lysosomal lumen	
L-Asparagine	L-Asn	mitochondrial matrix	
L-Aspartate	L-Asp	cytosol	
L-Aspartate	L-Asp	endoplasmic reticulum lumen	
L-Aspartate	L-Asp	extracellular region	
L-Aspartate	L-Asp	lysosomal lumen	
L-Aspartate	L-Asp	mitochondrial matrix	
Lauroyl-CoA	LAU-CoA	mitochondrial matrix	
L-Citrulline	L-Cit	cytosol	
L-Citrulline	L-Cit	mitochondrial matrix	
L-cysteine	L-Cys	cytosol	
L-cysteine	L-Cys	endoplasmic reticulum lumen	
L-cysteine	L-Cys	extracellular region	
L-cysteine	L-Cys	lysosomal lumen	
L-cysteine	L-Cys	mitochondrial matrix	
L-Cystine	L-Cyst	cytosol	
L-Cystine	L-Cyst	extracellular region	
Leukotriene A4	LTA4	cytosol	
Leukotriene B4	LTB4	cytosol	
Leukotriene B4	LTB4	endoplasmic reticulum lumen	
Leukotriene B4	LTB4	extracellular region	
Leukotriene C4	LTC4	cytosol	
Leukotriene C4	LTC4	extracellular region	
Leukotriene D4	LTD4	extracellular region	
Leukotriene E4	LTE4	extracellular region	
Leu-tRNA(Leu)		cytosol	
Leu-tRNA(Leu)		mitochondrial matrix	
L-Glutamate	L-Glu	clathrin sculpted glutamate transport vesicle lume	
L-Glutamate	L-Glu	cytosol	
L-Glutamate	L-Glu	endoplasmic reticulum lumen	
L-Glutamate	L-Glu	extracellular region	
L-Glutamate	L-Glu	lysosomal lumen	
L-Glutamate	L-Glu	mitochondrial matrix	

name	short name	location	Notes
L-Glutamate	L-Glu	synaptic vesicle	
L-Glutamate 5-semialdehyde	L-Glu5S	mitochondrial matrix	
L-Glutamine	L-Gln	cytosol	
L-Glutamine	L-Gln	endoplasmic reticulum lumen	
L-Glutamine	L-Gln	extracellular region	
L-Glutamine	L-Gln	lysosomal lumen	
L-Glutamine	L-Gln	mitochondrial matrix	
L-glycine	Gly	cytosol	
L-Histidine	L-His	cytosol	
L-Histidine	L-His	endoplasmic reticulum lumen	
L-Histidine	L-His	extracellular region	
L-Histidine	L-His	lysosomal lumen	
L-Histidine	L-His	mitochondrial matrix	
Li+	Li+	cytosol	
Li+	Li+	extracellular region	
lignoceric acid	LGCA	cytosol	
lignoceric acid	LGCA	extracellular region	
limit dextrins		extracellular region	
linoleic acid	LINA	cytosol	
linoleic acid	LINA	extracellular region	
linoleic acid	LINA	nucleoplasm	
linoleoyl-CoA	LIN-CoA	mitochondrial matrix	
Lipoamide	LIPAM	mitochondrial matrix	
Lipoate	LIPA	cytosol	
Lipoate	LIPA	extracellular region	
Lipoxin A4	LXA4	extracellular region	
L-Isoleucine	L-Ile	cytosol	
L-Isoleucine	L-Ile	endoplasmic reticulum lumen	
L-Isoleucine	L-Ile	extracellular region	
L-Isoleucine	L-Ile	lysosomal lumen	
L-Isoleucine	L-Ile	mitochondrial matrix	
lithoCHOLate	LCHA	cytosol	
lithoCHOLate sulfate	LCHAS	cytosol	
LithoCHOLic acid	LCHA	extracellular region	
L-Kynurenine	L-KYN	cytosol	
L-Kynurenine	L-KYN	mitochondrial matrix	
L-Leucine	L-Leu	cytosol	
L-Leucine	L-Leu	endoplasmic reticulum lumen	
L-Leucine	L-Leu	extracellular region	
L-Leucine	L-Leu	lysosomal lumen	
L-Leucine	L-Leu	mitochondrial matrix	
L-Lysine	L-Lys	cytosol	
L-Lysine	L-Lys	endoplasmic reticulum lumen	
L-Lysine	L-Lys	extracellular region	
L-Lysine	L-Lys	lysosomal lumen	
L-Lysine	L-Lys	mitochondrial matrix	
L-Methionine	L-Met	cytosol	
L-Methionine	L-Met	endoplasmic reticulum lumen	
L-Methionine	L-Met	extracellular region	
L-Methionine	L-Met	lysosomal lumen	
L-Methionine	L-Met	mitochondrial matrix	
L-methylmalonyl-CoA	L-MM-CoA	mitochondrial matrix	
long-chain fatty acids	LCFA	cytosol	
long-chain fatty acids	LCFA	extracellular region	
long-chain fatty acids	LCFA	peroxisomal matrix	
Loperamide	LPAM	endoplasmic reticulum lumen	
L-Ornithine	L-Orn	cytosol	
L-Ornithine	L-Orn	extracellular region	
L-Ornithine	L-Orn	mitochondrial intermembrane space	
L-Ornithine	L-Orn	mitochondrial matrix	
LPA		extracellular region	
L-Palmitoylcarnitine		mitochondrial intermembrane space	
L-Palmitoylcarnitine		mitochondrial matrix	
L-Phenylalanine	L-Phe	cytosol	
L-Phenylalanine	L-Phe	endoplasmic reticulum lumen	
L-Phenylalanine	L-Phe	extracellular region	
L-Phenylalanine	L-Phe	lysosomal lumen	
L-Phenylalanine	L-Phe	mitochondrial matrix	

name	short name	location	Notes
L-Proline	L-Pro	cytosol	
L-Proline	L-Pro	endoplasmic reticulum lumen	
L-Proline	L-Pro	extracellular region	
L-Proline	L-Pro	lysosomal lumen	
L-Proline	L-Pro	mitochondrial matrix	
L-serine	L-Ser	cytosol	
L-serine	L-Ser	endoplasmic reticulum lumen	
L-serine	L-Ser	extracellular region	
L-serine	L-Ser	lysosomal lumen	
L-serine	L-Ser	mitochondrial matrix	
L-Threonine	L-Thr	cytosol	
L-Threonine	L-Thr	endoplasmic reticulum lumen	
L-Threonine	L-Thr	extracellular region	
L-Threonine	L-Thr	lysosomal lumen	
L-Threonine	L-Thr	mitochondrial matrix	
L-Tryptophan	L-Trp	cytosol	
L-Tryptophan	L-Trp	endoplasmic reticulum lumen	
L-Tryptophan	L-Trp	extracellular region	
L-Tryptophan	L-Trp	lysosomal lumen	
L-Tryptophan	L-Trp	mitochondrial matrix	
L-Tyrosine	L-Tyr	cytosol	
L-Tyrosine	L-Tyr	endoplasmic reticulum lumen	
L-Tyrosine	L-Tyr	extracellular region	
L-Tyrosine	L-Tyr	lysosomal lumen	
L-Tyrosine	L-Tyr	mitochondrial matrix	
L-Valine	L-Val	cytosol	
L-Valine	L-Val	endoplasmic reticulum lumen	
L-Valine	L-Val	extracellular region	
L-Valine	L-Val	lysosomal lumen	
L-Valine	L-Val	mitochondrial matrix	
Lys-tRNA(Lys)		cytosol	
Lys-tRNA(Lys)		mitochondrial matrix	
Magnesium	Mg2+	cytosol	
Magnesium	Mg2+	endoplasmic reticulum membrane	
Magnesium	Mg2+	extracellular region	
Magnesium	Mg2+	Golgi membrane	
Magnesium	Mg2+	peroxisomal matrix	
Magnesium	Mg2+	plasma membrane	
Magnesium	Mg2+	platelet dense granule lumen	
Malonyl-CoA		cytosol	
maltose	Mal	cytosol	
maltose	Mal	extracellular region	
maltotriose		extracellular region	
Manganese	Mn2+	cytosol	
Manganese	Mn2+	extracellular region	
MDP	MDP	cytosol	
Melatonin	MLT	cytosol	
Melatonin	MLT	extracellular region	
Metformin	MTF	cytosol	
Metformin	MTF	extracellular region	
methacrylyl-CoA		mitochondrial matrix	
Methimazole	MTZ	endoplasmic reticulum lumen	
Methimazole S-oxide	MTZ-SOX	endoplasmic reticulum lumen	
Met-tRNA(Met)		cytosol	
Met-tRNA(Met)		mitochondrial matrix	
Mevalonate	MVA	cytosol	http://en.wikipedia.org/wiki/Mevalonate
Mevalonate-5-phosphate	MVA5P	cytosol	
Mevalonate-5-pyrophosphate	MVA5PP	cytosol	
Mg++	Mg2+	mitochondrial matrix	
Mn2+	Mn2+	mitochondrial matrix	
Mn2+	Mn2+	phagocytic vesicle	
MoCo	MoCo	cytosol	
molybdate(2-)	MoO4(2-)	cytosol	
Molybdopterin	MPT	cytosol	
Molybdopterin	MPT	cytosol	
Monoiodotyrosine	MIT	extracellular region	
mycothiol	MSH	cytosol	
mycothione	MSSM	cytosol	

name	short name	location	Notes
Myristic acid	MYSA	cytosol	
Myristic acid	MYSA	extracellular region	
myristoyl-CoA	MYS-CoA	cytosol	
myristoyl-CoA	MYS-CoA	mitochondrial matrix	
N-(L-Arginino)succinate		cytosol	
N,N'-diacetylchitobiosyldiphosphodoliCHOL	DCDPDOL	integral to cytosolic side of endoplasmic reticulum	
N-1-methylnicotinamide	MNA	cytosol	
N-1-methylnicotinamide	MNA	extracellular region	
N4-acetylsulfanilamide		cytosol	
N6,N6,N6-Trimethyl-L-lysine		mitochondrial matrix	
N-acetyl-5-hydroxytryptamine	Ac5HT	cytosol	
N-Acetylbenzoquinoneimine	NABQI	cytosol	
N-Acetylbenzoquinoneimine	NABQI	endoplasmic reticulum lumen	
N-Acetyl-D-glucosamine 1-phosphate	AcGlcN1P	cytosol	
N-Acetyl-D-glucosamine 6-phosphate	AcGlcN6P	cytosol	
N-acetyl-D-glucosaminyl-diphosphodoliCHOL		integral to cytosolic side of endoplasmic reticulum	
N-acetylisoniazid	NAcISNZ	cytosol	
N-Acetyl-L-glutamate	NAcGlu	mitochondrial matrix	
N-acetyl-p-aminobenzoate	NAcPAB	cytosol	
N-acetyl-p-aminosalicylate	NAcPAS	cytosol	
N-acetylspermidine	NASPM	cytosol	
N-acetylspermidine	NASPM	peroxisomal matrix	
N-acetylspermine	NASPN	cytosol	
N-acetylspermine	NASPN	peroxisomal matrix	
NAD+		cytosol	
NAD+		mitochondrial matrix	
NAD+		peroxisomal matrix	
NADH		cytosol	
NADH		mitochondrial matrix	
NADH		peroxisomal matrix	
NADP+		cytosol	
NADP+		endoplasmic reticulum lumen	
NADP+		extracellular region	
NADP+		mitochondrial matrix	
NADP+		nuclear envelope	
NADP+		peroxisomal matrix	
NADPH		cytosol	
NADPH		endoplasmic reticulum lumen	
NADPH		extracellular region	
NADPH		mitochondrial matrix	
NADPH		nuclear envelope	
NADPH		peroxisomal matrix	
N-arachidonyl glycine	NAGLY	extracellular region	
N-carbamoyl-L-aspartate		cytosol	
N-demethylated loperaminde		endoplasmic reticulum lumen	
NECML	NECML	extracellular region	
N-epsilon-(1-(1-carboxy)ethyl)lysine		extracellular region	
N-ethylsuccinimide	NETSU	cytosol	
N-ethylsuccinimide-mercapturic acid		cytosol	
N-ethylsuccinimide-mycothiol conjugate		cytosol	
N-Formimino-L-glutamate		cytosol	
N-Formyl-GAR	N-Formyl-GAR	cytosol	
N-glycoprotein	NGP	Golgi lumen	
NH3		cytosol	
NH3		nucleoplasm	
NH4+		cytosol	
NH4+		extracellular region	
NH4+		mitochondrial matrix	
NH4+		peroxisomal matrix	
N-hydroxy-2-acetylaminofluorene	NH2AAF	cytosol	
N-Hydroxy-4-aminobiphenyl	NHABP	cytosol	
N-Hydroxy-4-aminobiphenyl		endoplasmic reticulum lumen	
N-hydroxy-4-aminobiphenyl O-acetylated conjug		cytosol	
N-hydroxy-4-aminobiphenyl O-glucuronide		endoplasmic reticulum lumen	
N-hydroxy-4-aminobiphenyl O-sulfate		cytosol	
Niacin	VitB3	extracellular region	
Nicotinamide	NAM	cytosol	
Nicotinate	NCA	cytosol	

name	short name	location	Notes
Nicotinate D-ribonucleotide		cytosol	
Nicotinate D-ribonucleotide		nucleoplasm	
Nicotine		endoplasmic reticulum lumen	
Nicotine N-glucuronide		endoplasmic reticulum lumen	
Nitrate	HNO3	cytosol	
Nitrate	HNO3	transport vesicle	
Nitric oxide	NO	cytosol	
N-methylpyridinium ion	MP+	cytosol	
n-Oleoylethanolamide		extracellular region	
Noradrenaline	NAd	clathrin sculpted glutamate transport vesicle lume	http://en.wikipedia.org/wiki/Norepinephrine
Noradrenaline	NAd	clathrin sculpted monoamine transport vesicle lum	http://en.wikipedia.org/wiki/Norepinephrine
Noradrenaline	NAd	cytosol	http://en.wikipedia.org/wiki/Norepinephrine
Noradrenaline	NAd	extracellular region	http://en.wikipedia.org/wiki/Norepinephrine
Noradrenaline	NAd	stored secretory granule	http://en.wikipedia.org/wiki/Norepinephrine
O-acetylCHOLine	AcCho	clathrin sculpted acetylCHOLine transport vesicle l	
O-acetylCHOLine	AcCho	cytosol	
O-acetylCHOLine	AcCho	extracellular region	
O-acetylCHOLine	AcCho	plasma membrane	
O-acetylserine	AcSer	cytosol	
Octanoic acid		endoplasmic reticulum lumen	
Octanoyl-CoA		mitochondrial matrix	
Octanoyl-CoA		peroxisomal matrix	
O-hexadecylglycerone phosphate		cytosol	
O-hexadecylglycerone phosphate		peroxisomal matrix	
Oleic acid	OLEA	cytosol	
Oleic acid	OLEA	extracellular region	
omega-Hydroxydodecanoic acid		endoplasmic reticulum lumen	
Omeprazole	Losec	endoplasmic reticulum lumen	
O-phospho-L-serine		cytosol	
orotate		cytosol	
orotidine 5'-monophosphate		cytosol	
Orthophosphate	Pi	cytosol	
Orthophosphate	Pi	endoplasmic reticulum lumen	
Orthophosphate	Pi	extracellular region	
Orthophosphate	Pi	mitochondrial matrix	
Orthophosphate	Pi	nucleoplasm	
Orthophosphate	Pi	plasma membrane	
Orthophosphate	Pi	platelet dense granule lumen	
Oxalate	OX	peroxisomal matrix	http://en.wikipedia.org/wiki/Oxalate
Oxaliplatin		cytosol	
Oxaliplatin		extracellular region	
Oxaloacetate	OA	cytosol	
Oxaloacetate	OA	mitochondrial matrix	
oxidized glutathione	OGSH	cytosol	
Oxindole		endoplasmic reticulum lumen	
Oxygen	O2	cytosol	
Oxygen	O2	endoplasmic reticulum lumen	
Oxygen	O2	extracellular region	
Oxygen	O2	mitochondrial intermembrane space	
Oxygen	O2	mitochondrial matrix	
Oxygen	O2	nucleoplasm	
Oxygen	O2	peroxisomal matrix	
Oxygen	O2	stored secretory granule	
Paclitaxel	PTXL	endoplasmic reticulum lumen	
PAF	PAF	extracellular region	
Palmitate	Palm	cytosol	
Palmitate	Palm	extracellular region	
Palmitate	Palm	peroxisomal matrix	
Palmitate	Palm	plasma membrane	
palmitic acid	Palm	nucleoplasm	
Palmitoleic acid	Pmoa	extracellular region	
palmitoyl-CoA	PalmCoA	cytosol	
palmitoyl-CoA	PalmCoA	mitochondrial matrix	
palmitoyl-CoA	PalmCoA	peroxisomal matrix	
p-aminohippurate	pAHP	cytosol	
p-aminohippurate	pAHP	extracellular region	
Pantothenate	PanK	cytosol	
Pantothenate	PanK	extracellular region	

name	short name	location	Notes
PAP [adenosine 3',5'-bismonophosphate]	PAP	cytosol	
PAP		Golgi lumen	
PAPS 3'-phospho-5'-adenylyl sulfate]	PAPS	cytosol	
PAPS	PAPS	Golgi lumen	
Paraxanthine		cytosol	
Paraxanthine		endoplasmic reticulum lumen	
Pb++	Pb2+	cytosol	
Pb++	Pb2+	mitochondrial matrix	
Pentadecanoic acid		extracellular region	
Pentanoate	Valerate	extracellular region	
Phenol	PhOH	cytosol	
Phenol	PhOH	endoplasmic reticulum lumen	
Phenyl sulfate		cytosol	
Phenylacetaldehyde		cytosol	
phenylacetate		mitochondrial matrix	
phenylacetyl glutamine		mitochondrial matrix	
phenylacetyl-CoA		mitochondrial matrix	
Phenytoin		endoplasmic reticulum lumen	
Phe-tRNA(Phe)		cytosol	
Phe-tRNA(Phe)		mitochondrial matrix	
Phosphatidyl CHOLine	PC	endoplasmic reticulum membrane	
Phosphatidyl CHOLine	PC	Golgi membrane	
Phosphatidyl CHOLine	PC	plasma membrane	
phosphatidylCHOLines	PCs	extracellular region	
phosphatidylethanolamine	PE	external side of plasma membrane	
phosphatidylethanolamine	PE	internal side of plasma membrane	
Phosphatidylinositol-3,4,5-trisphosphate	PI(3,4,5)P3	endocytic vesicle membrane	
Phosphatidylinositol-3,4,5-trisphosphate	PI(3,4,5)P3	endosome membrane	
Phosphatidylinositol-3,4,5-trisphosphate	PI(3,4,5)P3	plasma membrane	
Phosphatidylinositol-4,5-bisphosphate	PI(4,5)P2	plasma membrane	
phosphatidylserine	PS	external side of plasma membrane	
phosphatidylserine	PS	internal side of plasma membrane	
PhosphoCHOLine	PCho	plasma membrane	
Phosphocreatine	Pcr	cytosol	
Phosphoenolpyruvate	PEP	cytosol	http://www.ebi.ac.uk/chebi/searchId.do?chebiId=CHEBI:18021
Phosphoenolpyruvate	PEP	mitochondrial matrix	http://www.ebi.ac.uk/chebi/searchId.do?chebiId=CHEBI:18021
phosphoethanolamine	PETA	cytosol	http://www.ebi.ac.uk/chebi/searchId.do?chebiId=CHEBI:17553
phospholipids	PL	clathrin-coated endocytic vesicle membrane	
phospholipids		endoplasmic reticulum lumen	
phospholipids		endoplasmic reticulum membrane	
phospholipids		endosome lumen	
phospholipids		endosome membrane	
phospholipids		extracellular region	
phospholipids		plasma membrane	
phospholipids		transport vesicle membrane	
Phospho-pantothenate		cytosol	
Phosphopantothenoylcysteine		cytosol	
Phytanate		peroxisomal matrix	
Phytanoyl-CoA		peroxisomal matrix	
phytoceramide		endoplasmic reticulum membrane	
phytosphingosine		endoplasmic reticulum membrane	
phytosterols		cytosol	
phytosterols		extracellular region	
PI(3,4)P2		plasma membrane	
Pioglitazone	Actos	nucleoplasm	
p-nitrophenol		cytosol	
p-nitrophenol sulfate		cytosol	
Porphobilinogen		cytosol	
Potassium	K+	cytosol	
Potassium	K+	extracellular region	
PQQ [pyrroloquinoline quinone]	PQQ	stored secretory granule	
Precursor Z		cytosol	
pregn-5-ene-3,20-dione		cytosol	
pregn-5-ene-3,20-dione-17-ol		cytosol	
pregnenolone	PREG	cytosol	
Pregnenolone	PREG	cytosol	
Pregnenolone	PREG	mitochondrial matrix	
pregnenolone sulfate	PREGS	cytosol	

name	short name	location	Notes
presqualene diphosphate	PSQPP	cytosol	
Pristanal		peroxisomal matrix	
Pristanic acid		peroxisomal matrix	
Procainamide		cytosol	
Procainamide		extracellular region	
Progesterone	P4	cytosol	
Propionate	EtCOO- or C2H5COO-	cytosol	
Propionate	EtCOO- or C2H5COO-	extracellular region	
propionyl CoA		peroxisomal matrix	
Propionylcarnitine		peroxisomal matrix	
propionyl-CoA		mitochondrial matrix	
Prostacyclin	PGI2	endoplasmic reticulum lumen	
Prostacyclin	PGI2	extracellular region	
Prostaglandin D2	PGD2	cytosol	
Prostaglandin D2	PGD2	extracellular region	
Prostaglandin E1	PGE1	cytosol	
Prostaglandin E1	PGE1	extracellular region	
Prostaglandin E2	PGE2	cytosol	
Prostaglandin E2	PGE2	extracellular region	
Prostaglandin F2alpha	PGF2a	cytosol	
Prostaglandin F2alpha	PGF2a	extracellular region	
Prostaglandin F2-alpha	PGF2a	extracellular region	
Prostaglandin G2	PGG2	endoplasmic reticulum lumen	
Prostaglandin H2	PGH2	cytosol	
Prostaglandin H2	PGH2	endoplasmic reticulum lumen	
Protoporphyrin IX		mitochondrial intermembrane space	
Protoporphyrin IX		mitochondrial matrix	
Protoporphyrinogen IX		mitochondrial intermembrane space	
Pro-tRNA(Pro)		cytosol	
Pro-tRNA(Pro)		mitochondrial matrix	
PRPP	PRPP	cytosol	
Putrescine		cytosol	
Putrescine		peroxisomal matrix	
Pyridine	PY	cytosol	
Pyridoxal	PXL	cytosol	
Pyridoxal phosphate	PXLP	cytosol	
Pyridoxamine	PXA	cytosol	
Pyridoxamine phosphate	PXAP	cytosol	
Pyridoxine	PDX	cytosol	
Pyridoxine phosphate	PDXP	cytosol	
pyrophosphate	PPi	cytosol	
pyrophosphate	PPi	extracellular region	
pyrophosphate	PPi	mitochondrial matrix	
pyrophosphate	PPi	nucleoplasm	
Pyrophosphate	PPi	peroxisomal matrix	
pyrophosphate	PPi	platelet dense granule lumen	
Pyruvate	PYR	cytosol	
Pyruvate	PYR	extracellular region	
Pyruvate	PYR	mitochondrial matrix	
Pyruvate	PYR	peroxisomal matrix	
q-dihydrobiopterin	qDHB	cytosol	
Quinidine	QND	cytosol	
Quinidine	QND	extracellular region	
quinine	QN	cytosol	
quinine	QN	extracellular region	
Quinolinate	QUIN	cytosol	
R-848		cytosol	
R-848		endosome membrane	
Retinyl palmitate	RPALM	cytosol	
Retinyl palmitate	RPALM	extracellular region	
Riboflavin	RIB	cytosol	
ROS		cytosol	
Rosiglitazone	RGZ	extracellular region	
Rosiglitazone	RGZ	nucleoplasm	
S-2-(hydroxyethyl)-glutathione		endoplasmic reticulum lumen	
Saccharin	SACC	extracellular region	
saccharopine	SACN	mitochondrial matrix	
S-adenosylhomocysteine	SAH	cytosol	

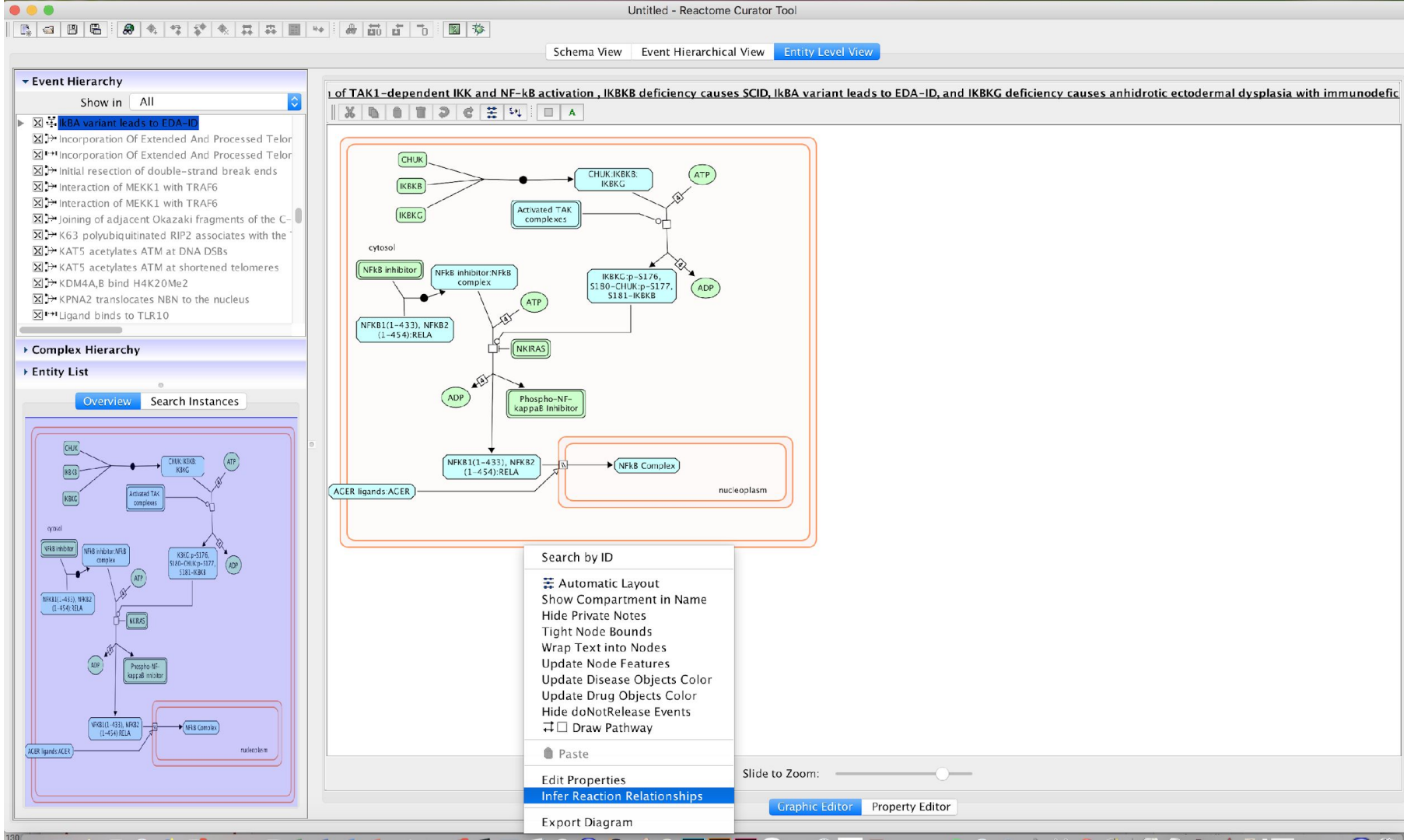
name	short name	location	Notes
S-adenosylhomocysteine	SAH	nucleoplasm	
S-Adenosylmethioninamine	SAMA	cytosol	
S-adenosylmethionine	SAM	cytosol	
S-adenosylmethionine	SAM	nucleoplasm	
SAICAR	SAICAR	cytosol	
Salicylate	SAL	mitochondrial matrix	
Salicylate-CoA thioester		mitochondrial matrix	
Salicyluric acid	SUA	mitochondrial matrix	
Sedoheptulose 7-phosphate		cytosol	
semidehydroascorbate	SHAS	cytosol	
Serotonin	5HT	clathrin sculpted monoamine transport vesicle lum	
Serotonin	5HT	extracellular region	
Serotonin	5HT	platelet dense granule lumen	
Ser-tRNA(Ser)		cytosol	
Ser-tRNA(Ser)		mitochondrial matrix	
Shikimate	SKM	cytosol	
Shikimate 3-phosphate	SKMP	cytosol	
sialic acid	SA	cytosol	
sialic acid		endocytic vesicle membrane	
Sialic acid		Golgi membrane	
sialic acid		lysosomal lumen	
sialic acid		plasma membrane	
Silicon dioxide	SiO2	cytosol	
siroheme	SRHM	cytosol	
S-methyl-2-mercaptoethanol	MMETOH	cytosol	
S-methylthio adenosine	MTAD	cytosol	
sodium ion	Na+	cytosol	
Sodium ion	Na+	early endosome lumen	
sodium ion	Na+	extracellular region	
sodium ion	Na+	Golgi lumen	
Sodium ion	Na+	late endosome lumen	
Spermidine	SPM	cytosol	
Spermidine	SPM	extracellular region	
Spermidine	SPM	peroxisomal matrix	
Spermine	SPN	cytosol	
Spermine	SPN	extracellular region	
Spermine	SPN	peroxisomal matrix	
sphinganine	SPA	cytosol	
sphinganine	SPA	endoplasmic reticulum membrane	
sphinganine 1-phosphate	SPAP	cytosol	
Sphingomyelin	SPHM	cytosol	
Sphingomyelin	SPHM	extracellular region	
sphingomyelins	SPHM	Golgi membrane	
sphingomyelins	SPHM	plasma membrane	
sphingosine	SPG	cytosol	
sphingosine	SPG	endoplasmic reticulum membrane	
sphingosine	SPG	extracellular region	
sphingosine	SPG	Golgi membrane	
Sphingosine-1-phosphate	S1P	cytosol	
Sphingosine-1-phosphate	S1P	extracellular region	
Squalene	SQNE	cytosol	
Squalene 2,3 epoxide	SQOX	cytosol	
stearate	STEA	cytosol	
Stearate	STEA	extracellular region	
stearic acid	STEA	extracellular region	
stearoyl-CoA	ST-CoA	cytosol	
Succinate	SUCCA	cytosol	
Succinate	SUCCA	extracellular region	
Succinate	SUCCA	mitochondrial matrix	
Succinate	SUCCA	nucleoplasm	
Succinate	SUCCA	peroxisomal matrix	
succinate semialdehyde	SUCCSA	mitochondrial matrix	
Succinyl-CoA	SUCC-CoA	mitochondrial matrix	
sucrose	Suc	extracellular region	
Sugar phosphate	SUGP	nucleoplasm	
Sulfanilamide	SAMD	cytosol	
sulfate	SO4(2-)	plasma membrane	
sulfide	S(2-)	cytosol	

name	short name	location	Notes
sulfide (2-)	S(2-)		
sulfite	H2SO3	cytosol	
sulfur	S	cytosol	
sulfurated MoCo		cytosol	
Tamoxifen	TAM	endoplasmic reticulum lumen	
Tamoxifen N-glucuronide	TAMNG	endoplasmic reticulum lumen	
Tamoxifen N-oxide	TAMO	endoplasmic reticulum lumen	
taurine	TAU	cytosol	
taurine		extracellular region	
taurine		peroxisomal matrix	
taurochenodeoxyCHOLate	TCCA	cytosol	
taurochenodeoxyCHOLate		extracellular region	
taurochenodeoxyCHOLate		peroxisomal matrix	
tauroCHOLate	TCA	cytosol	
tauroCHOLate		extracellular region	
tauroCHOLate		peroxisomal matrix	
taurolithoCHOLate	TLCA	cytosol	
taurolithoCHOLate sulfate	TLCAS	cytosol	
taurolithoCHOLate sulfate		extracellular region	
Testosterone	TES	cytosol	
Testosterone	TES	endoplasmic reticulum lumen	
tetracosanoyl-CoA		cytosol	
tetracosanoyl-CoA		peroxisomal matrix	
TetraHCA		cytosol	
TetraHCA		mitochondrial matrix	
Tetrahydrobiopterin	THBP	cytosol	
Tetrahydrofolate	THF	cytosol	http://en.wikipedia.org/wiki/Folic_acid
Tetrahydrofolate	THF	mitochondrial matrix	http://en.wikipedia.org/wiki/Folic_acid
Tetramethylammonium	TMAM	cytosol	
Tetramethylammonium	TMAM	extracellular region	
THCA		cytosol	
THCA		mitochondrial matrix	
THF-polyglutamate		cytosol	
THF-polyglutamate		mitochondrial matrix	
Thiamin	THMN	cytosol	
Thiamin	THMN	extracellular region	
Thiamin diphosphate	TDP	cytosol	
Thiamin diphosphate	TDP	mitochondrial matrix	
Thiamin pyrophosphate	TPP	peroxisomal matrix	
Thiamin triphosphate	TTP	cytosol	
Thromboxane A2	TXA2	endoplasmic reticulum lumen	
Thromboxane A2	TXA2	extracellular region	
Thr-tRNA(Thr)		cytosol	
Thr-tRNA(Thr)		mitochondrial matrix	
thymidine	Thy-dRib	cytosol	
thymidine	Thy-dRib	extracellular region	
thymidine	Thy-dRib	lysosomal lumen	
thymidine	Thy-dRib	mitochondrial matrix	
thymidine 5'-diphosphate	TDP	cytosol	
thymidine 5'-diphosphate	TDP	mitochondrial matrix	
thymidine 5'-monophosphate	TMP	extracellular region	
thymidine 5'-monophosphate	TMP	mitochondrial matrix	
thymine	Thy	cytosol	
thymine	Thy	extracellular region	
Thymine	Thy	nucleoplasm	
Thyroxine	T4	cytosol	
Thyroxine	T4	extracellular region	
tiglyl-CoA		mitochondrial matrix	
TMP thymidine 5'-monophosphate	TMP	cytosol	
Tolbutamide	TBM	endoplasmic reticulum lumen	
torcetrapib		extracellular region	
trans,cis-Lauro-2,6-dienoyl-CoA		mitochondrial matrix	
trans-2,3-dehydrohexacosanoyl-CoA		peroxisomal matrix	
trans-2,3-dehydropristanoyl-CoA		peroxisomal matrix	
trans-Dec-2-enoyl-CoA		mitochondrial matrix	
trans-Hex-2-enoyl-CoA		mitochondrial matrix	
trans-Hexadec-2-enoyl-CoA		mitochondrial matrix	
trans-Oct-2-enoyl-CoA		mitochondrial matrix	

name	short name	location	Notes
trans-octadec-2-enoyl-CoA		cytosol	
trans-Tetradec-2-enoyl-CoA		mitochondrial matrix	
triacylglycerol	TAG	cytosol	
triacylglycerols	TAGs	clathrin-coated endocytic vesicle membrane	
triacylglycerols	TAGs	endoplasmic reticulum lumen	
triacylglycerols	TAGs	endoplasmic reticulum membrane	
triacylglycerols	TAGs	endosome lumen	
triacylglycerols	TAGs	endosome membrane	
triacylglycerols	TAGs	extracellular region	
triacylglycerols	TAGs	plasma membrane	
Trifluoroacetylchloride	TFAC	endoplasmic reticulum lumen	
Trimethyamine N-oxide	TMAO	endoplasmic reticulum lumen	
Trimethylamine	TMA	endoplasmic reticulum lumen	
tRNA(Ala)		cytosol	
tRNA(Ala)		mitochondrial matrix	
tRNA(Arg)		cytosol	
tRNA(Arg)		mitochondrial matrix	
tRNA(Asn)		cytosol	
tRNA(Asn)		mitochondrial matrix	
tRNA(Asp)		cytosol	
tRNA(Asp)		mitochondrial matrix	
tRNA(Cys)		cytosol	
tRNA(Cys)		mitochondrial matrix	
tRNA(Gln)		cytosol	
tRNA(Gln)		mitochondrial matrix	
tRNA(Glu)		cytosol	
tRNA(Glu)		mitochondrial matrix	
tRNA(Gly)		cytosol	
tRNA(Gly)		mitochondrial matrix	
tRNA(His)		cytosol	
tRNA(His)		mitochondrial matrix	
tRNA(Ile)		cytosol	
tRNA(Ile)		mitochondrial matrix	
tRNA(Leu)		cytosol	
tRNA(Leu)		mitochondrial matrix	
tRNA(Lys)		cytosol	
tRNA(Lys)		mitochondrial matrix	
tRNA(Met)		cytosol	
tRNA(Met)		mitochondrial matrix	
tRNA(Phe)		cytosol	
tRNA(Phe)		mitochondrial matrix	
tRNA(Pro)		cytosol	
tRNA(Pro)		mitochondrial matrix	
tRNA(Ser)		cytosol	
tRNA(Ser)		mitochondrial matrix	
tRNA(Thr)		cytosol	
tRNA(Thr)		mitochondrial matrix	
tRNA(Trp)		cytosol	
tRNA(Trp)		mitochondrial matrix	
tRNA(Tyr)		cytosol	
tRNA(Tyr)		mitochondrial matrix	
tRNA(Val)		cytosol	
tRNA(Val)		mitochondrial matrix	
Trp-tRNA(Trp)		cytosol	
Trp-tRNA(Trp)		mitochondrial matrix	
TTP thymidine 5'-triphosphate	TTP	cytosol	
TTP thymidine 5'-triphosphate	TTP	mitochondrial matrix	
Tyramine	TYR	cytosol	
Tyr-tRNA(Tyr)		cytosol	
Tyr-tRNA(Tyr)		mitochondrial matrix	
ubiquinol	QH2	mitochondrial inner membrane	
ubiquinone	CoQ	mitochondrial inner membrane	
UDP	UDP	extracellular region	
UDP	UDP	Golgi lumen	
UDP-D-galactose	UDP-Gal	cytosol	
UDP-D-galactose	UDP-Gal	Golgi lumen	
UDP-GalNAc		cytosol	
UDP-GalNAc		Golgi lumen	

name	short name	location	Notes
UDP-GlcNAc		cytosol	
UDP-GlcNAc		Golgi lumen	
UDP-glucose	UDP-Glc	cytosol	
UDP-glucose	UDP-Glc	extracellular region	
UDP-glucose	UDP-Glc	Golgi lumen	
UDP-glucuronic acid	UDP-GCU	cytosol	http://www.hmdb.ca/metabolites/HMDB00127
UDP-glucuronic acid	UDP-GCU	endoplasmic reticulum lumen	http://www.hmdb.ca/metabolites/HMDB00127
UDP-N-acetyl-D-glucosamine	UDP-AcGlcN	cytosol	
UDP-N-acetyl-D-glucosamine	UDP-AcGlcN	Golgi lumen	
UDP-xylose	UDP-Xyl	cytosol	
UDP-xylose	UDP-Xyl	Golgi lumen	
UMP	UMP	cytosol	
UMP	UMP	Golgi lumen	
UMP	UMP	nucleoplasm	
uracil	Ura	cytosol	
uracil	Ura	extracellular region	
Uracil	Ura	nucleoplasm	
Urate		cytosol	
Urate		extracellular region	
Urea		cytosol	
Urea		extracellular region	
Urea		mitochondrial matrix	
uridine	Ura-Rib	cytosol	
uridine	Ura-Rib	extracellular region	
uridine	Ura-Rib	lysosomal lumen	
uridine	Ura-Rib	mitochondrial matrix	
uridine 2'-monophosphate	U2MP	cytosol	
uridine 2'-monophosphate	U2MP	mitochondrial matrix	
uridine 3'-monophosphate	U3MP	cytosol	
uridine 3'-monophosphate	U3MP	mitochondrial matrix	
uridine 5'-diphosphate	U5DP	cytosol	
uridine 5'-diphosphate	U5DP	endoplasmic reticulum lumen	
uridine 5'-diphosphate	U5DP	mitochondrial matrix	
uridine 5'-diphosphate	U5DP	nucleoplasm	
uridine 5'-monophosphate	U5MP	extracellular region	
uridine 5'-monophosphate	U5MP	mitochondrial matrix	
Urocanate	UCA	cytosol	
Uroporphyrinogen I	URO1	cytosol	
Uroporphyrinogen III	URO3	cytosol	
UTP uridine 5'-triphosphate	UTP	cytosol	
UTP uridine 5'-triphosphate	UTP	extracellular region	
UTP uridine 5'-triphosphate	UTP	mitochondrial matrix	
UTP uridine 5'-triphosphate	UTP	nucleoplasm	
Val-tRNA(Val)		cytosol	
Val-tRNA(Val)		mitochondrial matrix	
Vinyl chloride	VC	endoplasmic reticulum lumen	
Vitamin D3	VD3	cytosol	
Vitamin D3	VD3	extracellular region	
vitamin K epoxide	VKO	endoplasmic reticulum lumen	
vitamin K hydroquinone	VKHQ	endoplasmic reticulum lumen	
Xanthine	XAN	cytosol	
XMP	XMP	cytosol	
Zinc 2+ [zinc (2+)]	Zn2+	Golgi lumen	
Zinc 2+	Zn2+	Golgi membrane	
Zinc 2+	Zn2+	stored secretory granule	
Zn++	Zn2+	cytosol	
Zn++	Zn2+	endocytic vesicle membrane	
Zn++	Zn2+	endoplasmic reticulum lumen	
Zn++	Zn2+	endosome lumen	
Zn++	Zn2+	extracellular region	
Zn++	Zn2+	nucleoplasm	
Zn++	Zn2+	plasma membrane	
Zn++	Zn2+	synaptic vesicle	
Zn2+		mitochondrial matrix	
Zn2+		phagocytic vesicle	
zymosterol	ZSOL	endoplasmic reticulum membrane	

C7. Review of the Curator Tool PrecedingEvent Suggestion Feature


Untitled - Reactome Curator Tool
Schema View
Event Hierarchical View
Entity Level View
Event Hierarchy
Show in
All
IkBA variant leads to EDA-ID
Incorporation Of Extended And Processed Telor
Incorporation Of Extended And Processed Telor
Initial resection of double-strand break ends
Interaction of MEKK1 with TRAF6
Interaction of MEKK1 with TRAF6
Joining of adjacent Okazaki fragments of the C-
K63 polyubiquitinated RIP2 associates with the
KAT5 acetylates ATM at DNA DSBs
KAT5 acetylates ATM at shortened telomeres
KDM4A,B bind H4K20Me2
KPNA2 translocates NBN to the nucleus
Ligand binds to TLR10
Complex Hierarchy
Entity List
Overview
Search Instances
of TAK1-dependent IKK and NF-kB activation , IKBKB deficiency causes SCID, IkBA variant leads to EDA-ID, and IKBKG deficiency causes anhidrotic ectodermal dysplasia with immunodefic
CHUK
IKBKB
IKBKG
CHUK:IKBKB: IKBKG
ATP
Activated TAK complexes
cytosol
NFkB inhibitor
NFkB inhibitor:NFkB complex
IKBKG:p-S176, S180-CHUK:p-S177, S181-IKBKB
ADP
ATP
NFKB1(1-433), NFKB2 (1-454):RELA
NKIRAS
ADP
Phospho-NF-kappaB Inhibitor
NFKB1(1-433), NFKB2 (1-454):RELA
NFkB Complex
nucleoplasm
AGER ligands:AGER
Search by ID
Automatic Layout
Show Compartment in Name
Hide Private Notes
Tight Node Bounds
Wrap Text into Nodes
Update Node Features
Update Disease Objects Color
Update Drug Objects Color
Hide doNotRelease Events
Draw Pathway
Paste
Edit Properties
Infer Reaction Relationships
Export Diagram
Slide to Zoom:
Graphic Editor
Property Editor

Untitled - Reactome Curator Tool

Schema View | Event Hierarchical View | Entity Level View

Event Hierarchy

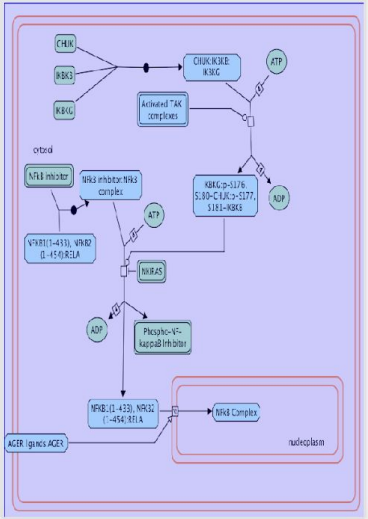
Show in: All

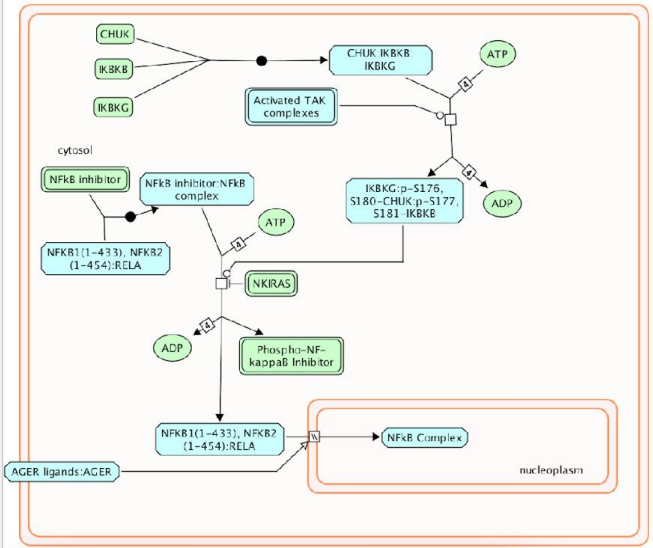
- IkBA variant leads to EDA-ID
- Incorporation Of Extended And Processed Telor
- Incorporation Of Extended And Processed Telor
- Initial resection of double-strand break ends
- Interaction of MEKK1 with TRAF6
- Interaction of MEKK1 with TRAF6
- Joining of adjacent Okazaki fragments of the C-
- K63 polyubiquitinated RIP2 associates with the
- KAT5 acetylates ATM at DNA DSBs
- KAT5 acetylates ATM at shortened telomeres
- KDM4A,B bind H4K20Me2
- KPNA2 translocates NBN to the nucleus
- Ligand binds to TLR10

Complex Hierarchy

Entity List

Overview | Search Instances





Escape Entities

Edit entities that should be escaped for analysis:

ADP, ATP, CO2, GDP, GTP, H+, H2O, Pi

☑ Include entity set members for analysis

☐ Include drugs for analysis

OK | Cancel

Slide to Zoom:

Graphic Editor | Property Editor

Untitled - Reactome Curator Tool

Schema View | Event Hierarchical View | Entity Level View

Event Hierarchy

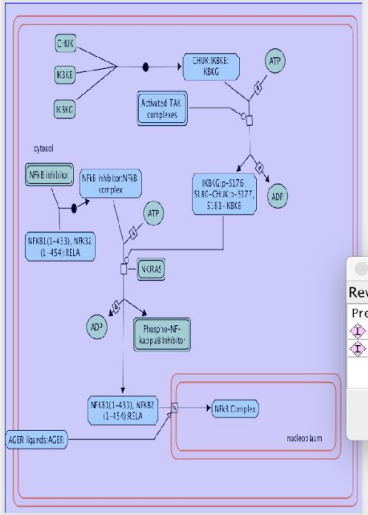
Show in: All

- Second phosphorylation of IRAK1 by IRAK4 bou
- Secreted CD14 binds LPS
- TAK1 is activated
- TAK1-dependent IKK and NF-kB activation
  - CHUK, IKBKB and IKBKG form IKK comple
  - Activated TAK1 mediates phosphorylation
  - Activated TAK1 phosphorylates IKBKB
  - Ubiquitination of IKBKG by TRAF6
  - NFkB inhibitor binds NFkB complex
  - Active IKK Complex phosphorylates NF-kB
  - NFkB complex is transported from cytosol
  - Regulation of NF-kB signaling
- TIFA oligomerization

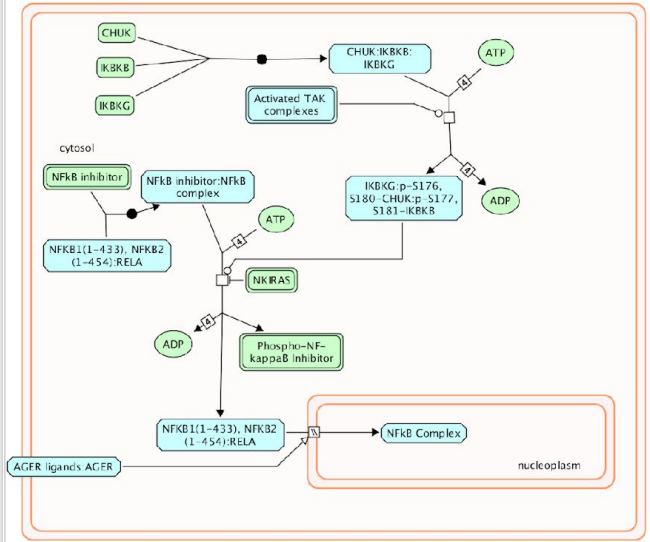
Complex Hierarchy

Entity List

Overview | Search Instances



1 of TAK1-dependent IKK and NF-kB activation , IKBKB deficiency causes SCID, IkBA variant leads to EDA-ID, and IKBKG deficiency causes anhidrotic ectodermal dysplasia with immunodefic



Reaction Relationship Inference

Review the following inferred relationships and check ones you want to approve:

PrecedingReaction	FollowingReaction	SharedEntity	Action	Reject Reason
Active IKK Complex phosphorylates NF-kB inhibi...	NFkB inhibitor binds NFkB complex	NFKB1(1-433), NFKB2(1-454):RELA [cytosol]	No_Action	
CHUK, IKBKB and IKBKG form IKK complex	Activated TAK1 mediates phosphorylation of the IKK Complex	CHUK:IKBKB:IKBKG [cytosol]	No_Action	

OK | Cancel

Slide to Zoom:

Graphic Editor | Property Editor

Untitled - Reactome Curator Tool

Schema View | Event Hierarchical View | Entity Level View

Event Hierarchy

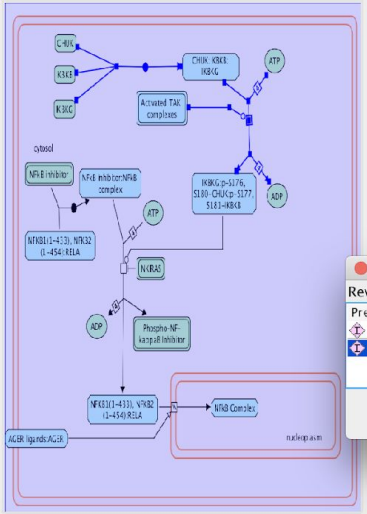
Show in: All

- Second phosphorylation of IRAK1 by IRAK4 bour
- Secreted CD14 binds LPS
- TAK1 is activated
- TAK1-dependent IKK and NF-kB activation
  - CHUK, IKBKB and IKBKG form IKK complex
  - Activated TAK1 mediates phosphorylation
  - Activated TAK1 phosphorylates IKBKB
  - Ubiquitination of IKBKG by TRAF6
  - NFkB inhibitor binds NFkB complex
  - Active IKK Complex phosphorylates NF-kB
  - NFkB complex is transported from cytosol t
  - Regulation of NF-kB signaling
- TIFA oligomerization

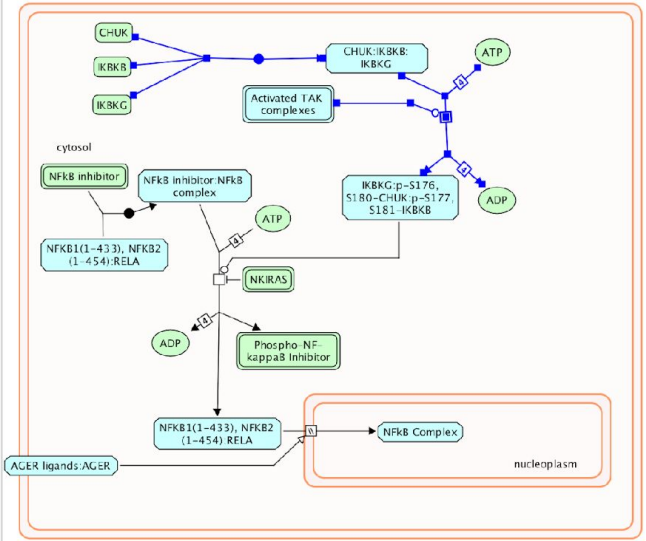
Complex Hierarchy

Entity List

Overview | Search Instances



of TAK1-dependent IKK and NF-kB activation , IKBKB deficiency causes SCID, IkBA variant leads to EDA-ID, and IKBKG deficiency causes anhidrotic ectodermal dysplasia with immunodefic



Reaction Relationship Inference

Review the following inferred relationships and check ones you want to approve:

PrecedingReaction	FollowingReaction	SharedEntity	Action	Reject Reason
Active IKK Complex phosphorylates NF-kB inhibi...	NFkB inhibitor binds NFkB complex	NFKB1(1-433), NFKB2(1-454):RELA [cytosol]	No_Action	
CHUK, IKBKB and IKBKG form IKK complex	Activated TAK1 mediates phosphorylation of the IKK Complex	CHUK:IKBKB:IKBKG [cytosol]	Approve	

OK | Cancel

Slide to Zoom:

Graphic Editor | Property Editor

Untitled - Reactome Curator Tool

Schema View | Event Hierarchical View | Entity Level View

Event Hierarchy

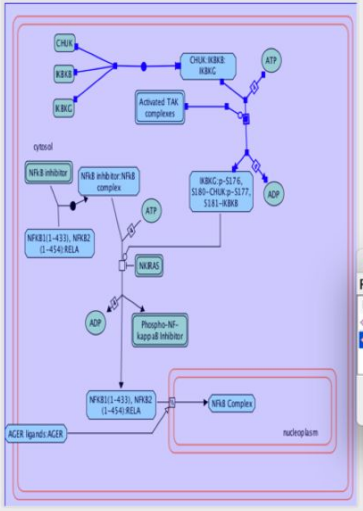
Show in: All

- SENP1,2,5 proteolytically process SUMO1
- SIRT6 deacetylates RBBP8
- Second phosphorylation of IRAK1 by IRAK4 bound to M
- Second phosphorylation of IRAK1 by IRAK4 bound to a
- Secreted CD14 binds LPS
- TAK1 is activated
- TAK1-dependent IKK and NF-kB activation
- CHUK, IKBKB and IKBKG form IKK complex
- Activated TAK1 mediates phosphorylation of the
- Activated TAK1 phosphorylates IKBKB
- Ubiquitination of IKBKG by TRAF6
- NFkB inhibitor binds NFkB complex
- Active IKK Complex phosphorylates NF-kB inhibit

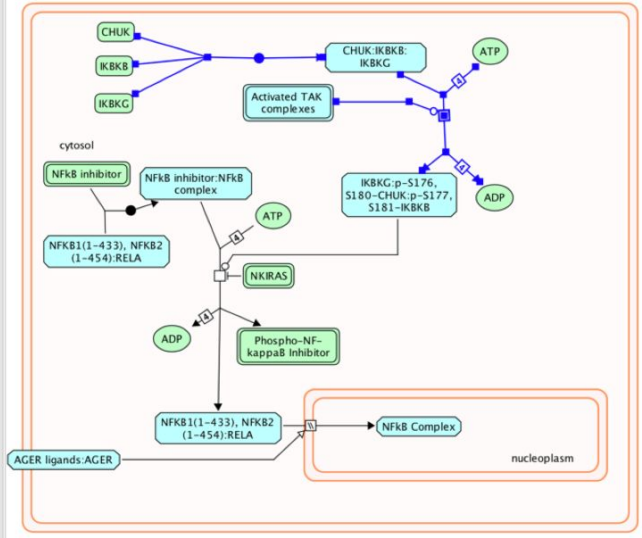
Complex Hierarchy

Entity List

Overview | Search Instances



1 of TAK1-dependent IKK and NF-kB activation , IKBKB deficiency causes SCID, IkBA variant leads to EDA-ID, and IKBKG deficiency causes anhidrotic ectodermal dysplasia with immunodefic



Reaction Relationship Inference

Review the following inferred relationships and check ones you want to approve:

PrecedingReaction	FollowingReaction	SharedEntity	Reject Reason
Active IKK Complex phosphorylates NF-kB inhibitor	NFkB inhibitor binds NFkB complex	NFKB1(1-433), NFKB2(1-454):RELA [cyt...	
CHUK, IKBKB and IKBKG form IKK complex	Activated TAK1 mediates phosphorylation of the IKK Complex	CHUK:IKBKB:IKBKG [cytosol]	

Approve
Reject
✓ No_Action

OK | Cancel

Slide to Zoom:

Graphic Editor | Property Editor

Untitled - Reactome Curator Tool

Schema View | Event Hierarchical View | Entity Level View

Event Hierarchy

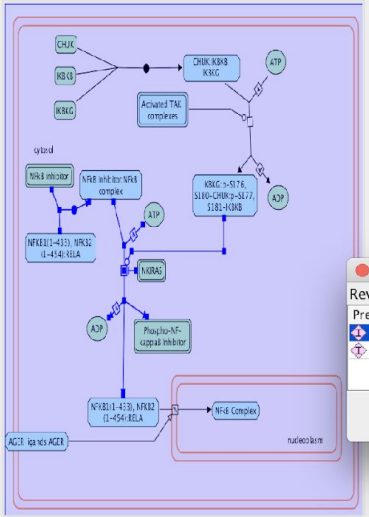
Show in: All

- Second phosphorylation of IRAK1 by IRAK4 bour
- Secreted CD14 binds LPS
- TAK1 is activated
- TAK1-dependent IKK and NF-kB activation
  - CHUK, IKBKB and IKBKG form IKK complex
  - Activated TAK1 mediates phosphorylation
  - Activated TAK1 phosphorylates IKBKB
  - Ubiquitination of IKBKG by TRAF6
  - NFkB inhibitor binds NFkB complex
  - Active IKK Complex phosphorylates NF-kB
  - NFkB complex is transported from cytosol t
  - Regulation of NF-kB signaling
- TIFA oligomerization

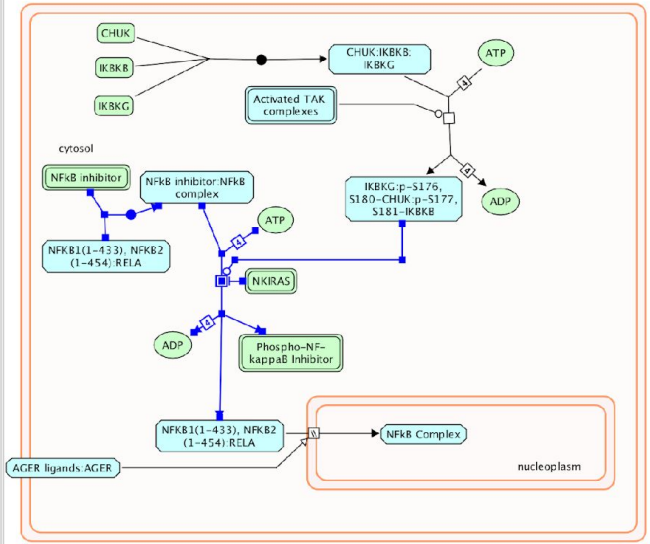
Complex Hierarchy

Entity List

Overview | Search Instances



...of TAK1-dependent IKK and NF-kB activation , IKBKB deficiency causes SCID, IkBA variant leads to EDA-ID, and IKBKG deficiency causes anhidrotic ectodermal dysplasia with immunodefic



Reaction Relationship Inference

Review the following inferred relationships and check ones you want to approve:

PrecedingReaction	FollowingReaction	SharedEntity		Reject Reason
Active IKK Complex phosphorylates NF-kB inhibi...	NFkB inhibitor binds NFkB complex	NFKB1(1-433), NFKB2(1-454):RELA [cytosol]	Approve	
CHUK, IKBKB and IKBKG form IKK complex	Activated TAK1 mediates phosphorylation of the IKK Complex	CHUK:IKBKB:IKBKG [cytosol]	Approve	

Approve
Reject
✓ No_Action

OK | Cancel

Slide to Zoom:

Graphic Editor | Property Editor

Untitled - Reactome Curator Tool

Schema View | Event Hierarchical View | Entity Level View

Event Hierarchy

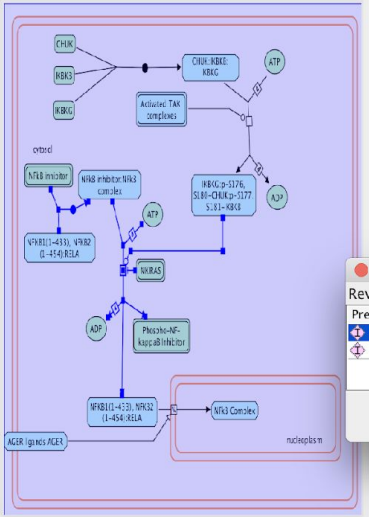
Show in: All

- Second phosphorylation of IRAK1 by IRAK4 bour
- Secreted CD14 binds LPS
- TAK1 is activated
- TAK1-dependent IKK and NF-kB activation
  - CHUK, IKBKB and IKBKG form IKK complex
  - Activated TAK1 mediates phosphorylation
  - Activated TAK1 phosphorylates IKBKB
  - Ubiquitination of IKBKG by TRAF6
  - NFkB inhibitor binds NFkB complex
  - Active IKK Complex phosphorylates NF-kB
  - NFkB complex is transported from cytosol t
- Regulation of NF-kB signaling
- TIFA oligomerization

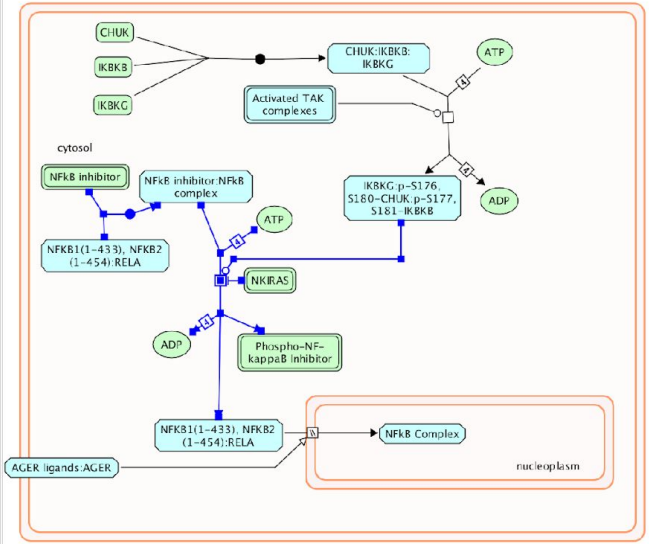
Complex Hierarchy

Entity List

Overview | Search Instances



1 of TAK1-dependent IKK and NF-kB activation , IKBKB deficiency causes SCID, IkBA variant leads to EDA-ID, and IKBKG deficiency causes anhidrotic ectodermal dysplasia with immunodefic



Reaction Relationship Inference

Review the following inferred relationships and check ones you want to approve:

PrecedingReaction	FollowingReaction	SharedEntity	Action	Reject Reason
Active IKK Complex phosphorylates NF-kB inhibi...	NFkB inhibitor binds NFkB complex	NFKB1(1-433), NFKB2(1-454):RELA [cytosol]	Reject	
CHUK, IKBKB and IKBKG form IKK complex	Activated TAK1 mediates phosphorylation of the IKK Complex	CHUK:IKBKB:IKBKG [cytosol]	Approve	

- Abundant Molecules
- Different Cell Type
- Drug molecule
- Futile Cycle
- Incorrect Order
- Negative relationship

Slide to Zoom:

Graphic Editor | Property Editor

Untitled - Reactome Curator Tool

Schema View | Event Hierarchical View | Entity Level View

Event Hierarchy

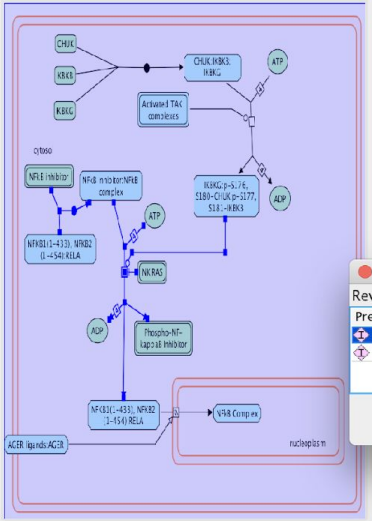
Show in: All

- S100A8:S100A9 binds TLR4:LY96
- SENP1,2,5 proteolytically process SUMO1
- SIRT6 deacetylates RBBP8
- Second phosphorylation of IRAK1 by IRAK4 bour
- Second phosphorylation of IRAK1 by IRAK4 bour
- Secreted CD14 binds LPS
- TAK1 is activated
- TAK1-dependent IKK and NF-kB activation
  - CHUK, IKBKB and IKBKG form IKK complex
  - Activated TAK1 mediates phosphorylation
  - Activated TAK1 phosphorylates IKBKB
  - Ubiquitination of IKBKG by TRAF6
  - NFkB inhibitor binds NFkB complex

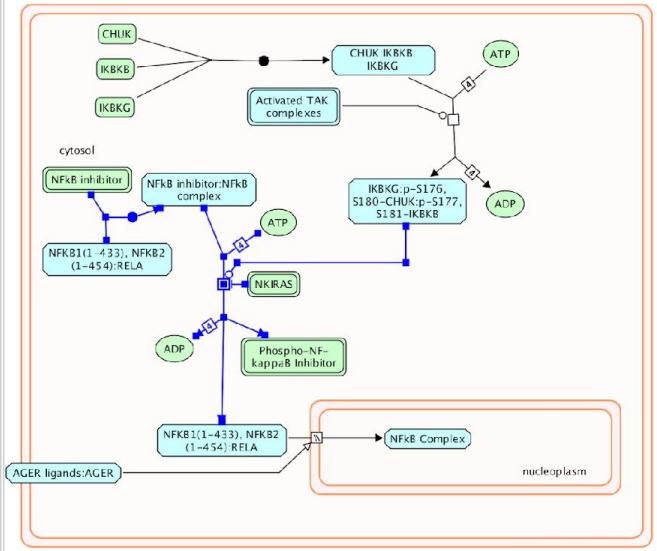
Complex Hierarchy

Entity List

Overview | Search Instances



ı of TAK1-dependent IKK and NF-kB activation , IKBKB deficiency causes SCID, IkBA variant leads to EDA-ID, and IKBKG deficiency causes anhidrotic ectodermal dysplasia with immunodefic



Reaction Relationship Inference

Review the following inferred relationships and check ones you want to approve:

PrecedingReaction	FollowingReaction	SharedEntity	Action	Reject Reason
Active IKK Complex phosphorylates NF-kB inhibi...	NFkB inhibitor binds NFkB complex	NFKB1(1-433), NFKB2(1-454):RELA [cytosol]	Reject	Incorrect Order
CHUK, IKBKB and IKBKG form IKK complex	Activated TAK1 mediates phosphorylation of the IKK Complex	CHUK:IKBKB:IKBKG [cytosol]	Approve	

OK | Cancel

Slide to Zoom:

Graphic Editor | Property Editor

Untitled - Reactome Curator Tool

Schema View | Event Hierarchical View | Entity Level View

Event Hierarchy

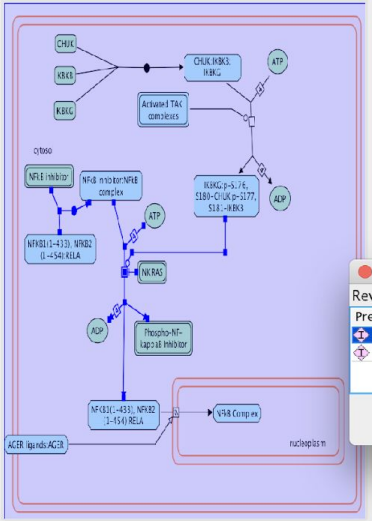
Show in All

- S100A8:S100A9 binds TLR4:LY96
- SENP1,2,5 proteolytically process SUMO1
- SIRT6 deacetylates RBBP8
- Second phosphorylation of IRAK1 by IRAK4 bour
- Second phosphorylation of IRAK1 by IRAK4 bour
- Secreted CD14 binds LPS
- TAK1 is activated
- TAK1-dependent IKK and NF-kB activation
  - CHUK, IKBKB and IKBKG form IKK complex
  - Activated TAK1 mediates phosphorylation
  - Activated TAK1 phosphorylates IKBKB
  - Ubiquitination of IKBKG by TRAF6
  - NFkB inhibitor binds NFkB complex

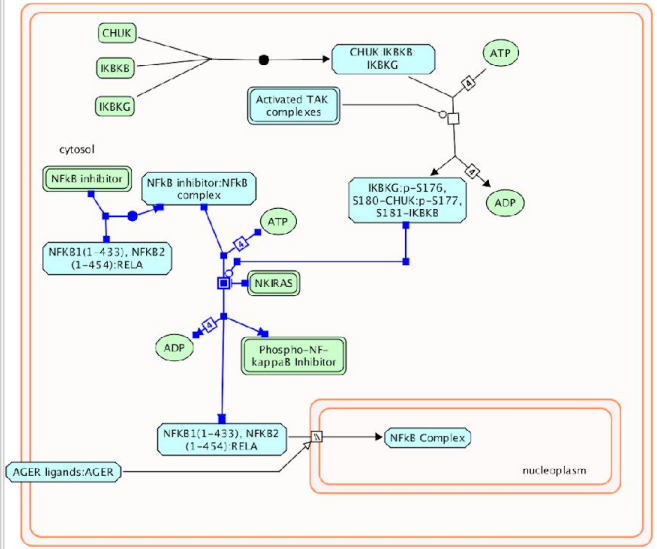
Complex Hierarchy

Entity List

Overview | Search Instances



1 of TAK1-dependent IKK and NF-kB activation , IKBKB deficiency causes SCID, IkBA variant leads to EDA-ID, and IKBKG deficiency causes anhidrotic ectodermal dysplasia with immunodefic



Reaction Relationship Inference

Review the following inferred relationships and check ones you want to approve:

PrecedingReaction	FollowingReaction	SharedEntity	Action	Reject Reason
Active IKK Complex phosphorylates NF-kB inhibi...	NFkB inhibitor binds NFkB complex	NFKB1(1-433), NFKB2(1-454):RELA [cytosol]	Reject	Incorrect Order
CHUK, IKBKB and IKBKG form IKK complex	Activated TAK1 mediates phosphorylation of the IKK Complex	CHUK:IKBKB:IKBKG [cytosol]	Approve	

OK | Cancel

Slide to Zoom:

Graphic Editor | Property Editor

Untitled - Reactome Curator Tool

Schema View | Event Hierarchical View | Entity Level View

Event Hierarchy

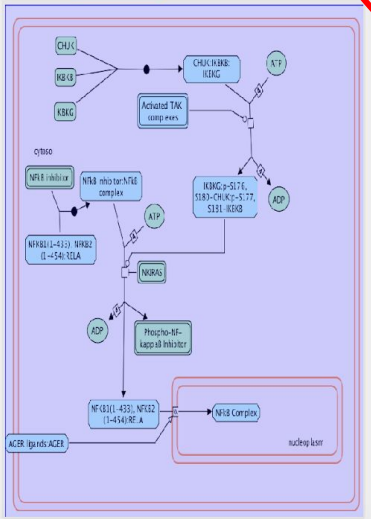
Show in All

- TAK1 is activated
- TAK1-dependent IKK and NF-kB activation
  - CHUK, IKBKB and IKBKG form IKK comple
  - > Activated TAK1 mediates phosphorylatic
  - Activated TAK1 phosphorylates IKBKB
  - Ubiquitination of IKBKG by TRAF6
  - > NFkB inhibitor binds NFkB complex
  - Active IKK complex phosphorylates NF-kE
  - NFkB complex is transported from cytosol
  - Regulation of NF-kB signaling
- TIFA oligomerization
- TIRAP is phosphorylated by BTK
- TLR1/2 ligand binds to CD14

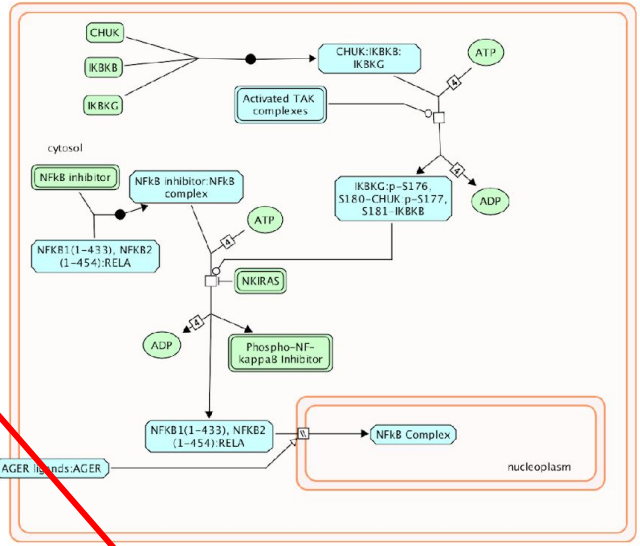
Complex Hierarchy

Entity List

Overview | Search Instances



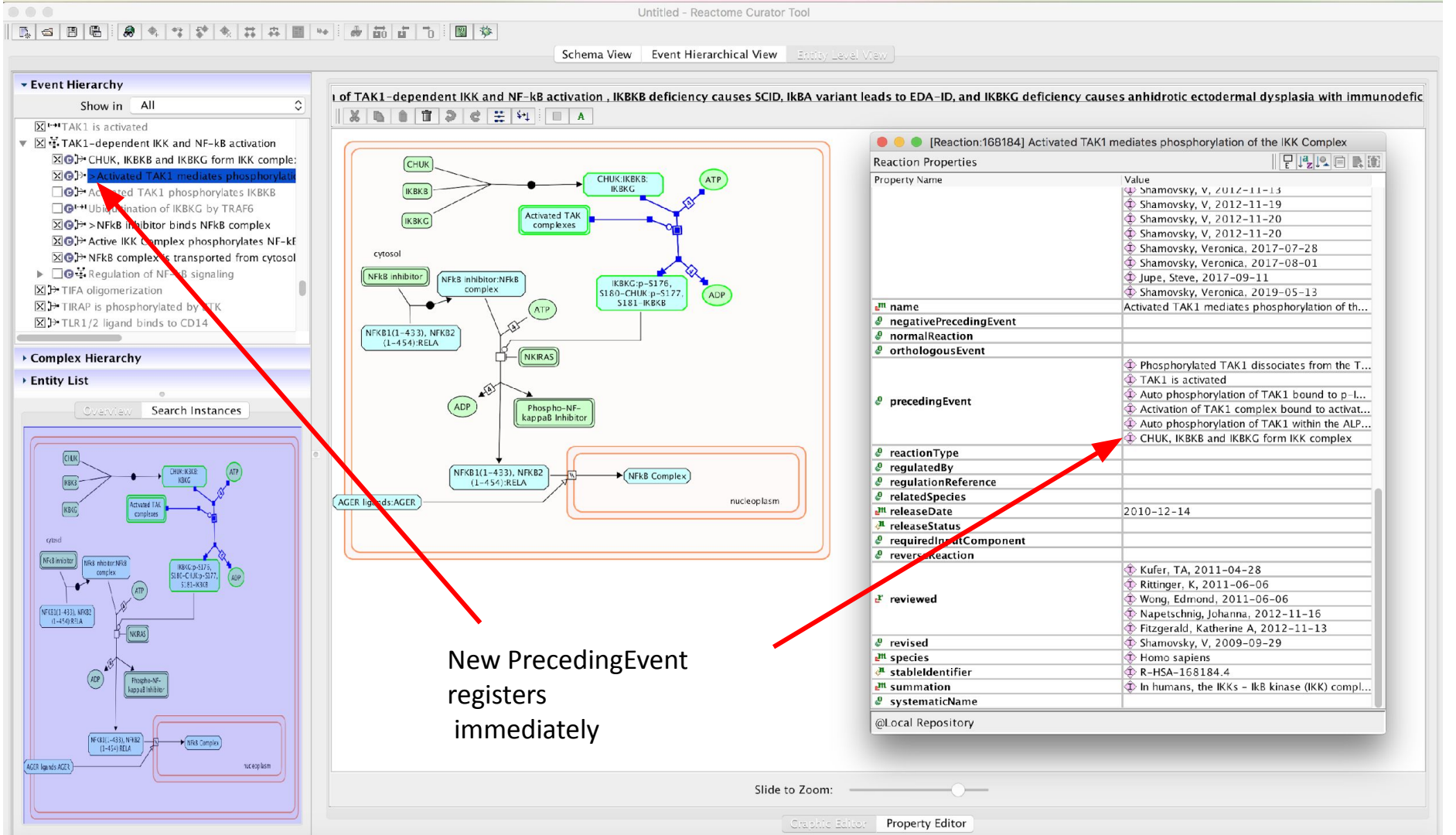
of TAK1-dependent IKK and NF-kB activation , IKBKB deficiency causes SCID, IkBA variant leads to EDA-ID, and IKBKG deficiency causes anhidrotic ectodermal dysplasia with immunodefic



New PrecedingEvent registers immediately

Slide to Zoom:

Graphic Editor | Property Editor


Untitled - Reactome Curator Tool
Schema View
Event Hierarchical View
Entity Level View
Event Hierarchy
Show in
All
TAK1 is activated
TAK1-dependent IKK and NF-kB activation
CHUK, IKBKB and IKBKG form IKK comple:
> Activated TAK1 mediates phosphorylati
Ac ated TAK1 phosphorylates IKBKB
Ubiq ination of IKBKG by TRAF6
> NFkB inhibitor binds NFkB complex
Active IKK complex phosphorylates NF-kE
NFkB complex is transported from cytosol
Regulation of NF-kB signaling
TIFA oligomerization
TIRAP is phosphorylated by BTK
TLR1/2 ligand binds to CD14
Complex Hierarchy
Entity List
Overview
Search Instances
of TAK1-dependent IKK and NF-kB activation , IKBKB deficiency causes SCID, IkBA variant leads to EDA-ID, and IKBKG deficiency causes anhidrotic ectodermal dysplasia with immunodefic
CHUK
IKBKB
IKBKG
CHUK:IKBKB: IKBKG
ATP
Activated TAK complexes
cytosol
NFkB inhibitor
NFkB inhibitor:NFkB complex
IKBKG:p-S176, S180-CHUK p-S177, S181-IKBKB
ADP
NFKB1(1-433), NFKB2 (1-454):RELA
NKIRAS
Phospho-NF-kappaB Inhibitor
NFKB1(1-433), NFKB2 (1-454):RELA
NFkB Complex
AGER ligands:AGER
nucleoplasm
[Reaction:168184] Activated TAK1 mediates phosphorylation of the IKK Complex
Reaction Properties
Property Name
Value
Shamovsky, V, 2012-11-19
Shamovsky, V, 2012-11-20
Shamovsky, V, 2012-11-20
Shamovsky, Veronica, 2017-07-28
Shamovsky, Veronica, 2017-08-01
Jupe, Steve, 2017-09-11
Shamovsky, Veronica, 2019-05-13
name
Activated TAK1 mediates phosphorylation of th...
negativePrecedingEvent
normalReaction
orthologousEvent
precedingEvent
Phosphorylated TAK1 dissociates from the T...
TAK1 is activated
Auto phosphorylation of TAK1 bound to p-I...
Activation of TAK1 complex bound to activat...
Auto phosphorylation of TAK1 within the ALP...
CHUK, IKBKB and IKBKG form IKK complex
reactionType
regulatedBy
regulationReference
relatedSpecies
releaseDate
2010-12-14
releaseStatus
requiredInputComponent
reversibleReaction
reviewed
Kufer, TA, 2011-04-28
Rittinger, K, 2011-06-06
Wong, Edmond, 2011-06-06
Napetschnig, Johanna, 2012-11-16
Fitzgerald, Katherine A, 2012-11-13
revised
Shamovsky, V, 2009-09-29
species
Homo sapiens
stableIdentifier
R-HSA-168184.4
summation
In humans, the IKKs - IkB kinase (IKK) compl...
systematicName
@Local Repository
New PrecedingEvent registers immediately
Slide to Zoom:
Graphic Editor
Property Editor

Untitled - Reactome Curator Tool

Schema View | Event Hierarchical View | Entity Level View

Event Hierarchy

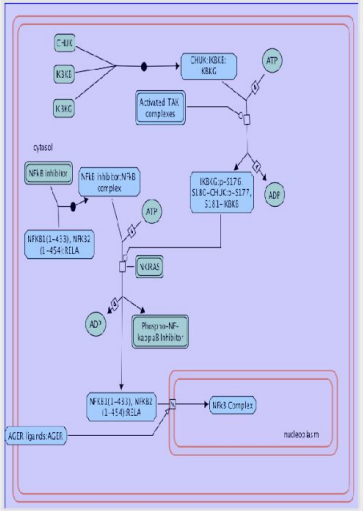
Show in: All

- IκBA variant leads to EDA-ID
- Incorporation Of Extended And Processed Telo
- Incorporation Of Extended And Processed Telo
- Initial resection of double-strand break ends
- Interaction of MEKK1 with TRAF6
- Interaction of MEKK1 with TRAF6
- Joining of adjacent Okazaki fragments of the C-
- K63 polyubiquitinated RIP2 associates with the
- KAT5 acetylates ATM at DNA DSBs
- KAT5 acetylates ATM at shortened telomeres
- KDM4A,B bind H4K20Me2
- KPNA2 translocates NBN to the nucleus
- Ligand binds to TLR10

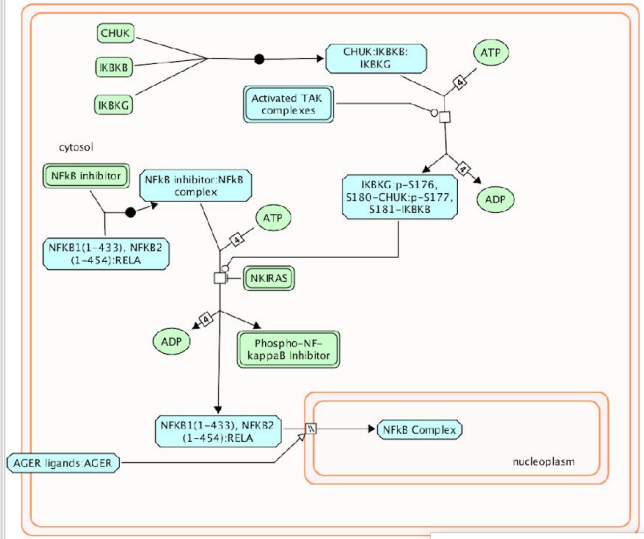
Complex Hierarchy

Entity List

Overview | Search Instances



ı of TAK1-dependent IKK and NF-kB activation , IKBKB deficiency causes SCID, IkBA variant leads to EDA-ID, and IKBKG deficiency causes anhidrotic ectodermal dysplasia with immunodefic



Search by ID

- Automatic Layout
- Show Compartment in Name
- Hide Private Notes
- Tight Node Bounds
- Wrap Text into Nodes
- Update Node Features
- Update Disease Objects Color
- Update Drug Objects Color
- Hide doNotRelease Events
- Draw Pathway

Paste

Edit Properties

Infer Reaction Relationships

Export Diagram

Property Editor

Untitled - Reactome Curator Tool

Schema View | Event Hierarchical View | Entity Level View

Event Hierarchy

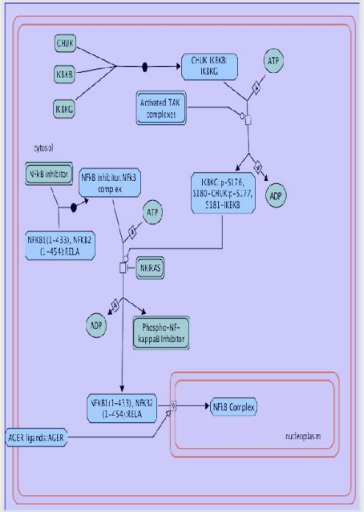
Show in: All

- IkBA variant leads to EDA-ID
- Incorporation Of Extended And Processed Telo
- Incorporation Of Extended And Processed Telo
- Initial resection of double-strand break ends
- Interaction of MEKK1 with TRAF6
- Interaction of MEKK1 with TRAF6
- Joining of adjacent Okazaki fragments of the C-
- K63 polyubiquitinated RIP2 associates with the
- KAT5 acetylates ATM at DNA DSBs
- KAT5 acetylates ATM at shortened telomeres
- KDM4A,B bind H4K20Me2
- KPNA2 translocates NBN to the nucleus
- Ligand binds to TLR10

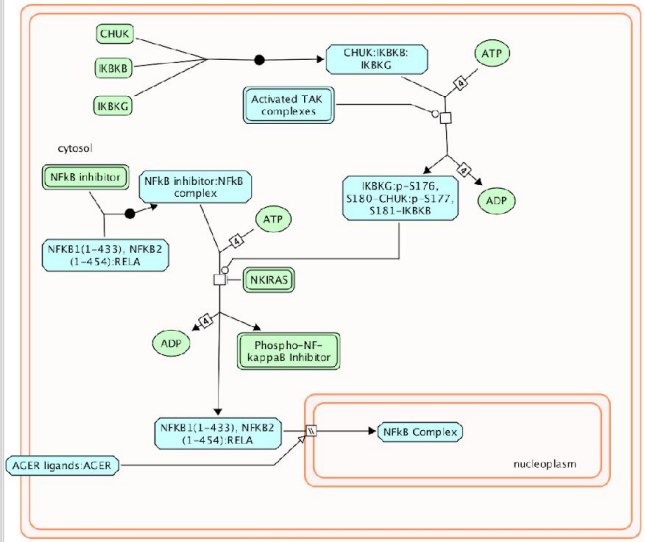
Complex Hierarchy

Entity List

Overview | Search Instances



of TAK1-dependent IKK and NF-kB activation , IKBKB deficiency causes SCID, IkBA variant leads to EDA-ID, and IKBKG deficiency causes anhidrotic ectodermal dysplasia with immunodefic



No Relationship

No relationship can be inferred in this pathway diagram.

OK

Slide to Zoom:

Graphic Editor | Property Editor

Rejection reason definitions:

1. Abundant molecule : small/common molecules through which you don't want to make a connection
2. Different cell type: shared molecules occur in different cell type
3. Drug molecule:
4. Futile cycle:creation of preceding event creates a loop that is not biologically meaningful.
5. Incorrect order: relationship would create an event order that does biologically relevant.
6. Negative relationship

*We have agreed to reject all recycling loops as futile cycles – for example, ubiquitin comes up quite often, same with any enzyme that we annotated through substrate binding-substrate modification-substrate dissociation steps. This is not to say that in a real cell such loops do not affect processes that consume some of the loop components, but since we don't have any kinetic parameters, we err on the side of assuming that reactants are present in excess.

Questions for discussion:

1. Which of the reasons for rejection should be chosen for a molecule entering an epithelial cell from one side and leaving it on the other side? (so that the export reaction cannot be preceding the import reaction).

2. What if we know for reasons outside the annotated pair of reactions, that they should never be connected because one is normally “off” when the other is “on”? Do we add a term for blocking that preceding – following instance, or do we NOT add it and instead check that all the coordinate regulation that causes this outcome is correctly annotated. My immediate hunch is that we do both – redundant but reliable and, as far as I can tell, also informative to users.

C8. Best Practices- Using Chemical Compound Databases
Which general small molecule databases do we use?

For Reactome, the ChEBI representation of a small molecule is our definitive reference. If ChEBI lacks an entry for your molecule, request one as explained elsewhere in this document. It is appropriate to create cross-references to PubChem and KEGG Compound, but these are not an acceptable substitute for ChEBI.

Where do we use small molecules?

Normal metabolites can be small molecules (ReferenceMolecule), also the majority of drugs are (ReferenceTherapeutic). For enzymatic reactions it is recommended to refer to Rhea which always links to ChEBI. Approved drugs can be found reliably in Guide to Pharmacology (GtP), which always links to PubChem (which most of the time has a ChEBI link if it exists).

In chemical databases I just search the name, no?

What can happen if you think the name suffices:

- there can be many entries with same name in PubChem, this can also happen in ChEBI
- in ChEBI the name that you have can also be the name of a generic structural class; they differ by the wording in the definition and by having asterisks in the SMILES string
- the drug compound may be a salt and the database may have both salt and ion
- the compound may have neutral or charged entries (ions), you need to know the metabolic form to decide what entry you want
- the drug stereochemistry may not be as fully specified as in the database, or vice versa; the drug may also be a mixture, with the database not having the effective component
- databases may not list your molecule under the name, and don't have it as synonym

So you better forget about names, also because there are better ways. First, as Rhea directly links to ChEBI the only case that poses problems is if GtP links to PubChem but PubChem has no ChEBI link, but see below.

Now that I know the exact molecule, how do I find it in ChEBI?

In the case your drug from GtP has a link to PubChem you know the exact molecule. To find it in ChEBI there are two possibilities:

- PubChem has links to ChEBI (however, see Caveat)
- GtP and PubChem have the InChi key of the molecule. This string is unique to any molecule (see Technical Details below). Just copy the InChi key and put it in the ChEBI search box.
- Caveat: PubChem links to ChEBI even if the molecules are not exactly the same, so you need to check if the InChi keys match

ChEBI does not have the molecule, how to submit?

The safest and most efficient way to submit a molecule at ChEBI is to

- find the exact entry in PubChem (see above)
- copy the 'isomeric SMILES' string of the structure
- if there is no 'isomeric SMILES' copy the 'canonical SMILES' string
- start submission at ChEBI by clicking on the "submit" button at the top right of the ChEBI home page and logging in (anyone can request login credentials)
- Click on "edit" in the page that opens once you've logged in to get a fresh structure input tab click Open:



- paste string, click OK, no need to draw the molecule
- continue submission as usual

Technical Details

What is SMILES and InChi?

Molecules are 3-dimensional objects that can be accurately depicted 2-dimensionally, i.e. as structural formula. To copy/paste/search this formula (as picture) is difficult, so to work with molecules computationally, textual codes were developed that can equally accurately define the 3-dimensional molecule, and at the same time being short strings that are easy to handle. Today, these two are used:

- **SMILES**, see https://en.wikipedia.org/wiki/Simplified_molecular-input_line-entry_system
- **InChi**, see https://en.wikipedia.org/wiki/International_Chemical_Identifier

The chemical databases used in Reactome (ChEBI and PubChem) always have both a SMILES and an InChi string for any specific molecule.

Canonical or isomeric SMILES?

Molecules with stereogenic centers cannot be uniquely defined using canonical SMILES and need an extended code called isomeric SMILES. In short, if there is an isomeric SMILES given by the database, always use that.

The InChi strings always have every stereochemistry specifications that are present in the molecule.

What is an InChi key?

SMILES and InChi can be very long strings. Also, they are not unique, i.e., for the same molecule there can be different SMILES and different InChi strings. So, comparing the strings of two molecules cannot tell you if both molecules are identical. The InChi key solves that problem.

- see https://en.wikipedia.org/wiki/International_Chemical_Identifier#InChIKey

Every InChi key is unique to its molecule, so you can use it to search for the exact molecule in databases, or even with Google in the web. The key is a hash, and so the molecule cannot be reconstructed from it alone, but databases always have both InChi and key associated. The uniqueness is not perfect -- amino acid zwitterions have the same key as the acid, and there can be very rare duplicate keys, see https://www.inchi-trust.org/technical-faq-2/#13.3.

V90_Sept 2024 updates

- moved from DevWiki
- all referenceMolecule instances MUST have ChEBI as the data source - PubChem and KEGG are not acceptable as a matter of editorial policy.
- update the procedure for requesting a new ChEBI term.

C9. Cytomics Documentation

Project Purpose

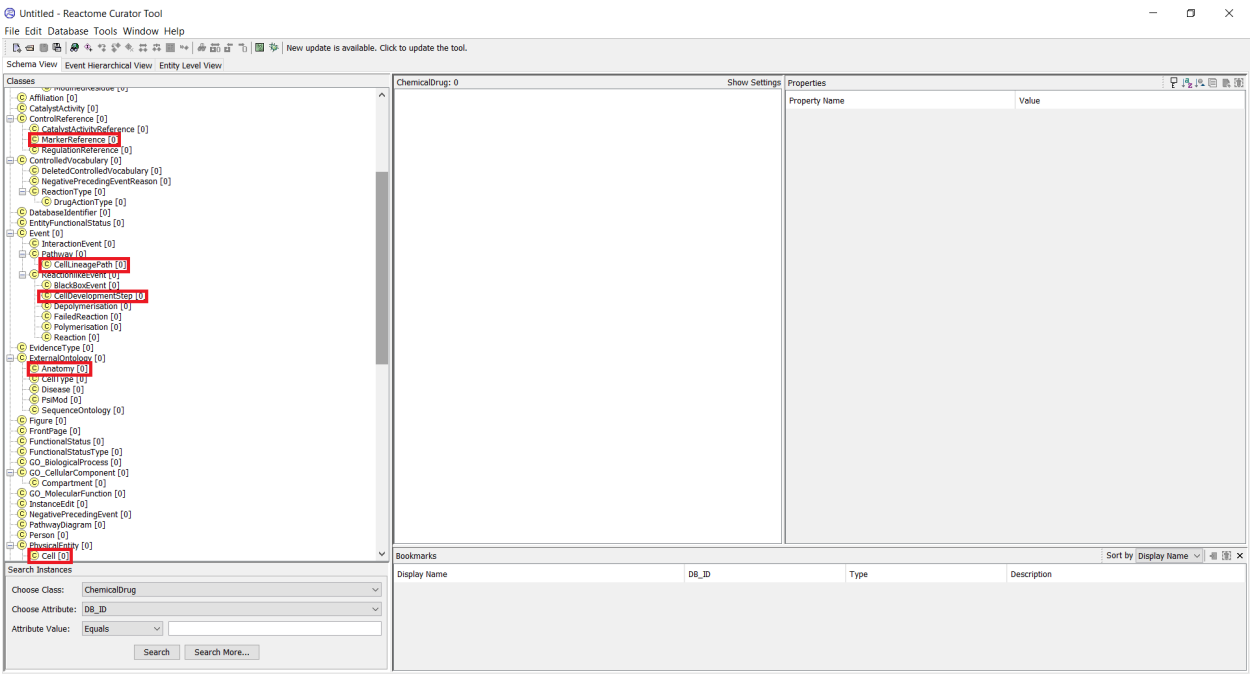
The purpose of the Cytomics Project is to allow curation and display of specialized pathways showing the development steps of a cell lineage (i.e. development for a particular cell type).

Project Work Summary (Data and Code Changes)

The project involves updates to 5 areas:

1. Data model changes

- 5 new classes have been added to gk_central (in order of appearance in screenshot below)
    - **Marker Reference**: A type of Control Reference that associates one or more EWAS markers(i.e. protein or RNA marker) with one or more literature references for evidence of the marker(s) in a Cell instance
    - **Cell Lineage Path**: A specialized Pathway which contains only other Cell Lineage Path instances or Cell Development Step as events in the "hasEvent" attribute and has a "tissue" attribute containing an Anatomy instance
    - **Cell Development Step**: A specialized ReactionlikeEvent which contains Cell instances for its inputs and outputs and has a "tissue" attribute containing an Anatomy instance
    - **Anatomy**: A type of External Ontology from UBERON (http://uberon.org) describing a particular type of organ, tissue type
    - **Cell**: A specialized Physical Entity representing a type of cell in a particular state of development



2. Updates to Pathway Diagram Curator code (i.e. pathway diagram visualization on curator.reactome.org) - https://github.com/reactome/PathwayDiagram/tree/feature/cell-lineage-path

- Convert project to use maven (creation of pom.xml file)
- Exposes a method to set a "CellLineagePath" by its db_id
- Updating hover tooltip with "CellDevelopmentStep" node in diagram
- Processing XML and rendering icons for "Cell" Physical Entities
- Add participating molecules menu for "Cell" Physical Entity to show marker physical entities
    - Change RESTful API endpoint for participating molecules to "entitySubunits" to get marker Physical Entities associated with a "Cell" Physical Entity (as well as continuing to get Physical Entities in a compositional node like a complex, set, polymer)

3. Updates to Pathway Browser Curator code (i.e. the curator website pathway browser - https://curator.reactome.org/PathwayBrowser) - https://github.com/reactome/browser-curator/tree/feature/cell-lineage-paths

- Adds model classes: *CellDevelopmentStep*, *CellLineagePath*, *Cell*, *Anatomy*, *ControlReference* and *MarkerReference* (includes new cases in *DatabaseObjectFactory* for creating instances of these, new *SchemaClass* and *PropertyType* constants, and *DatabaseObjectImages* for icons)
- Adds new DetailsPanel and Table classes: *CellPanel, MarkerReferencePanel, CellDevelopmentStepTable, CellLineagePathTable, CellTable*
- Updates DetailsPanel and Table classes:
    - *EntityWithAccessionedSequencePanel* (handles if an EWAS doesn't have a reference entity)
    - *PhysicalEntityPanel* (handles when the PhysicalEntity is a Cell)
    - *OverviewTableFactory* (handles creation of CellDevelopmentStep, CellLineagePath, Cell tables)
    - *TableRowFactory* (handles rows for a MarkerReference and an EWAS)
- Adds a method for RESTfulClient to obtain all "CellLineagePath" instances
- Updates a series of classes to work for any Event - i.e. "Pathway" or "CellLineagePath" rather than just a Pathway
    - In Selection, DiagramOpenRequestEvent, DiagramOpenedEvent, PathwayDiagramOpenRequestEvent, PathwayDiagramOpenedEvent, DiagramOpenRequestHandler, Ancestors, Path, OrthologousEventPathway, DownloadsTab, DownloadsTabPresenter, DownloadMolecule, ExpressionTab, ExpressionTabDisplay, ExpressionTabPresenter, MoleculesTab, MoleculesTabDisplay, MoleculesTabPresenter, Hierarchy, HierarchyDisplay, HierarchyPresenter, HierarchyPathLoader, HierarchyItem, HierarchyTree, GAManager, OrthologyManager, State, StateHelper, StateManger, TitleManager, ViewportPresenter, Diagram, DiagramDisplay, DiagramPresenter

        - NOTE: Some of these changes could be reverted since "CellLineagePath" has become a sub-class of "Pathway" in the data model instead of Event sub-class type
    - Adds war and compiler plugins for pom.xml
    - Updates dev and curator server URLs to dev.reactome.org and curator.reactome.org

4. RESTful API endpoints for querying the curator relational database gk_central (NOTE: This is a separate API from the Content Service which queries the graph database) - https://github.com/reactome/RESTfulAPI/tree/feature/cell-lineage-paths

- Adds new methods in org.reactome.restfulapi.RESTfulAPIResource
    - **/entitySubunits/{dbId}**
        - Similar to /complexSubunits/{dbId}, but also returns Marker physical entities for Cell instances
    - **/cellLineagePaths/{speciesName}**
        - Returns CellLineagePath instances that exist for a given species ("Homo sapiens" by default)

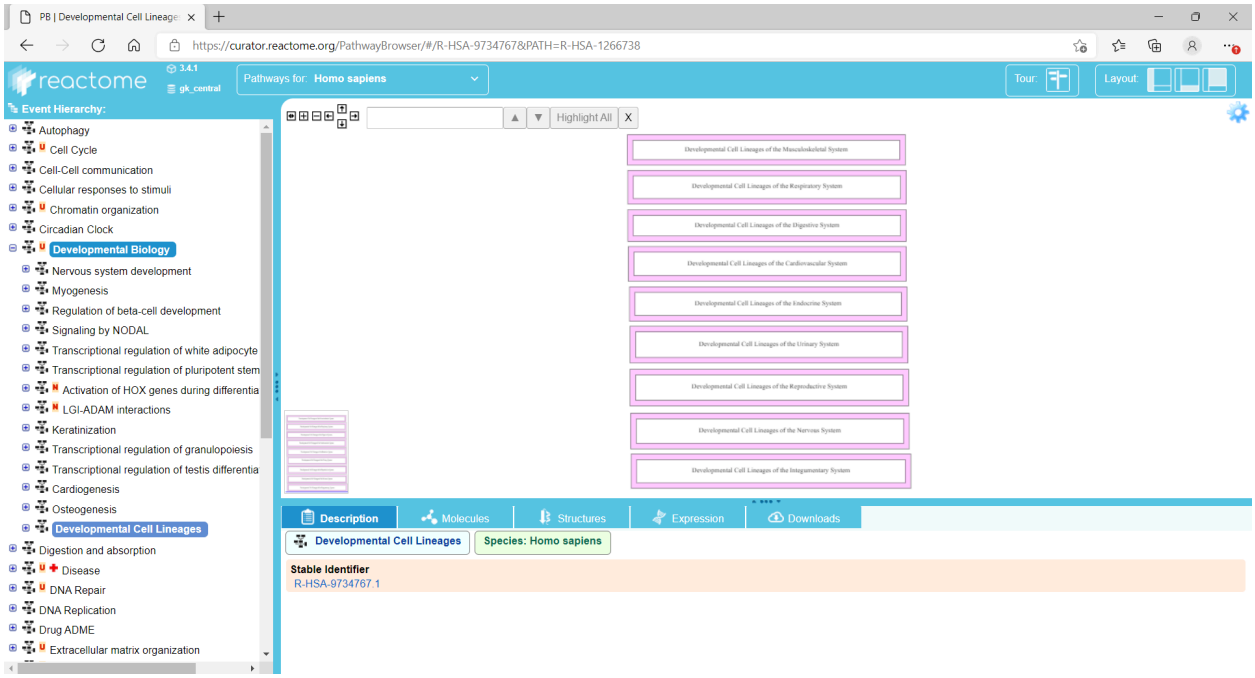
5. RTF pathway report downloader (i.e. RTF download button in "Download" tab of Pathway Browser Details Panel - https://curator.reactome.org/PathwayBrowser/#/DTAB=DT) - https://github.com/reactome/Release/tree/feature/cell-lineage-path

- Updates rtfexporter in the Release repository (website/cgi-bin/rtfexporter)
    - Changes call to ReactomeDatabaseReader from "init_pathways" to "init_events" to work for CellLineagePath instances
- Updates Perl module files
    - DocumentGeneration/ReactomeDatabaseReader.pm
        - Methods added to initialize and set events (i.e. both Pathways and CellLineagePaths) and get events by db_id
        - "get_chapters_from_events_text_units", "get_instance_text_units" method updated to process "CellLineagePath" instances
    - ReactomeReaction/Node.pm
        - Adds case and constant for color and rendering parameters for "Cell" diagram node
    - Graphics/SimpleDraw/Box.pm
        - Adds method for rendering "Cell"
    - Utils.pm
        - Updates method for obtaining "Pathway" events to obtaining sub-events in Events in "CellLineagePath" also

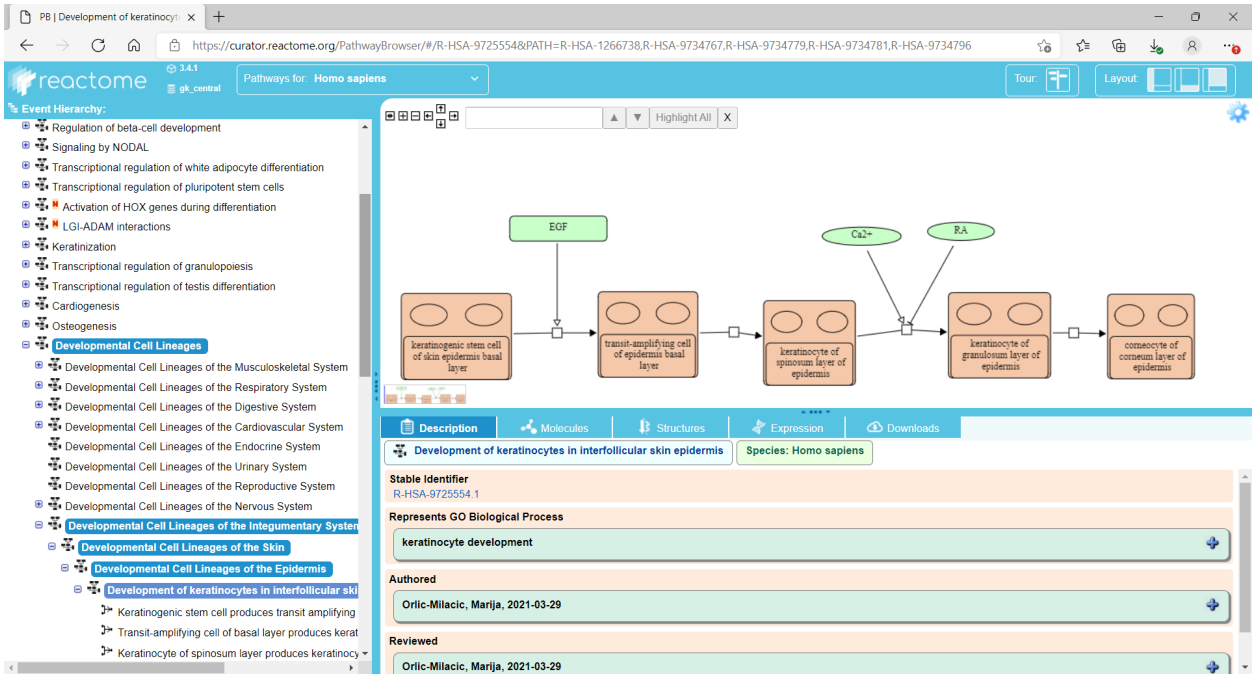
- Creates method to get top level Events (i.e. both Pathways and CellLineagePaths) and top level CellLineagePath instances

Viewing the Project

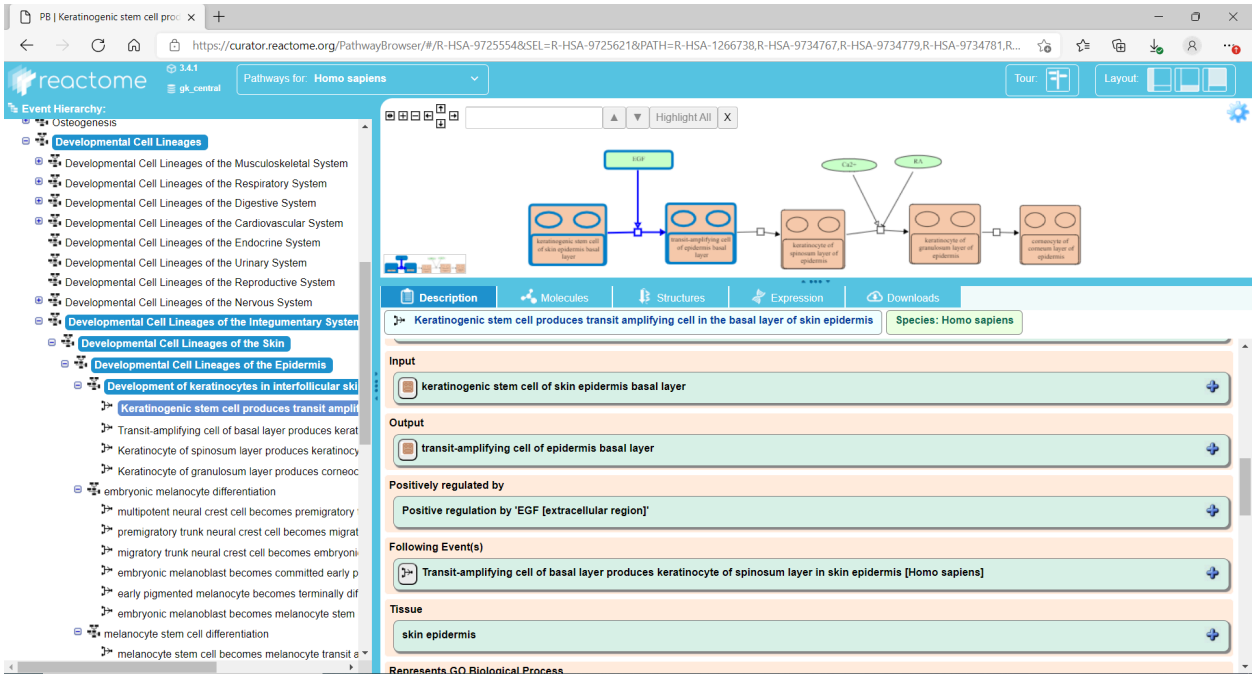
The project in its current state is deployed at https://curator.reactome.org. The entry point for all Cytomics work is under the "Developmental Biology" pathway in the Pathway Browser in the subpathway "Developmental Cell Lineages" https://curator.reactome.org/PathwayBrowser/#/R-HSA-9734767&PATH=R-HSA-1266738



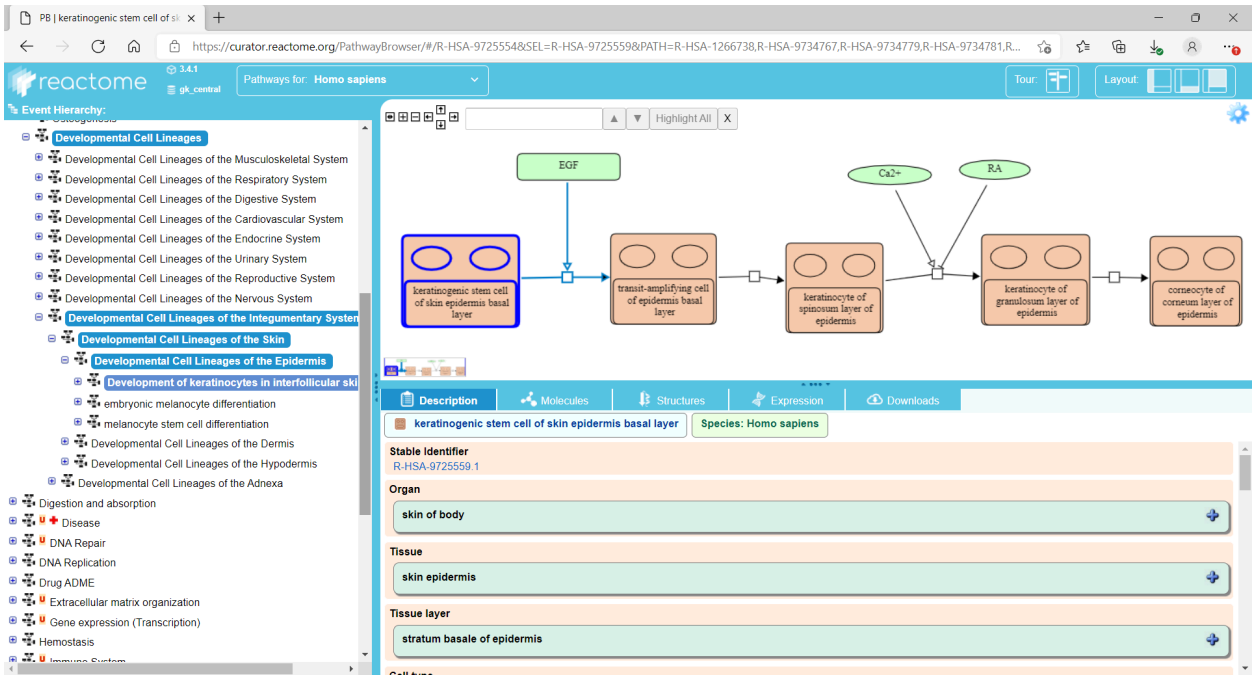
Selecting a "CellLineagePath" (from the Pathway Hierarchy):



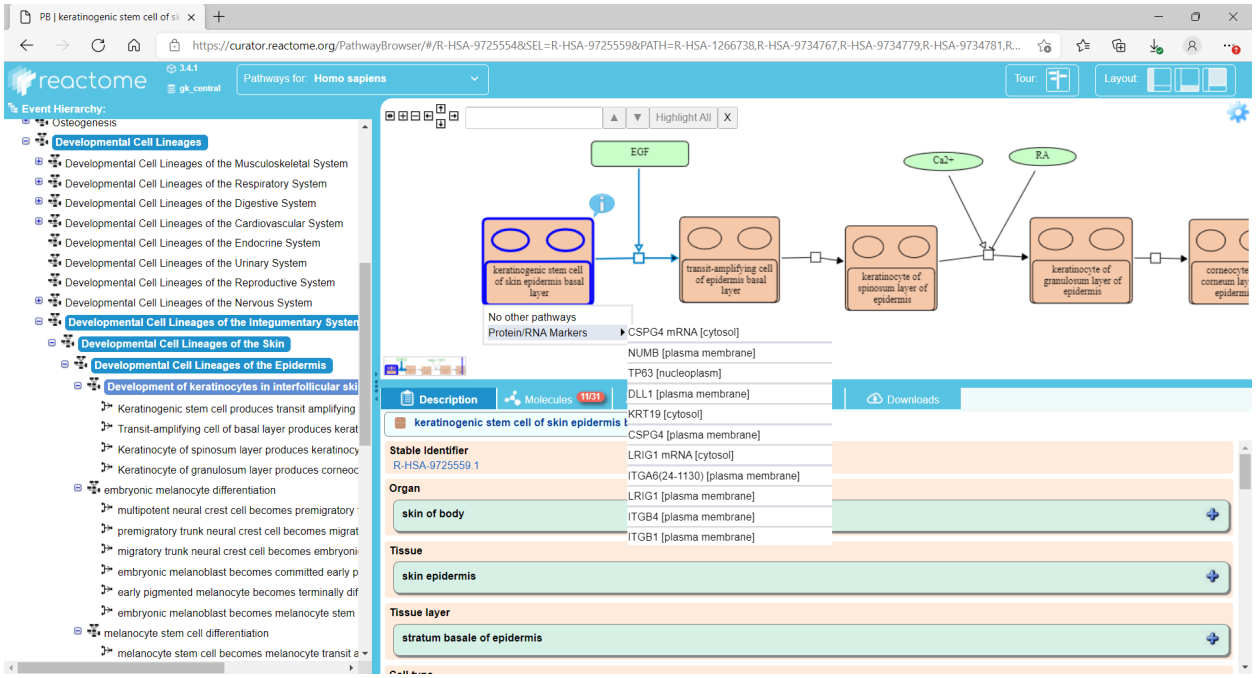
Selecting a "Cell Development Step" (from the Pathway Hierarchy or Pathway Diagram):



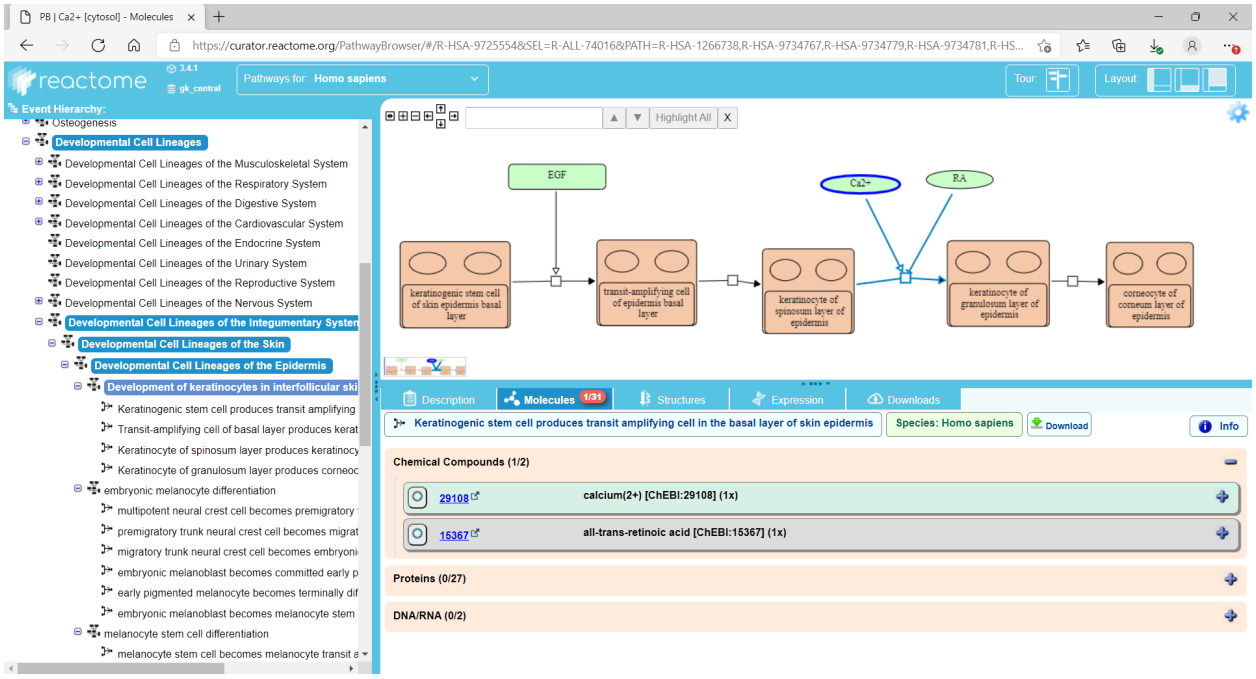
Selecting a "Cell" Physical Entity (from the Pathway Hierarchy or Pathway Diagram):



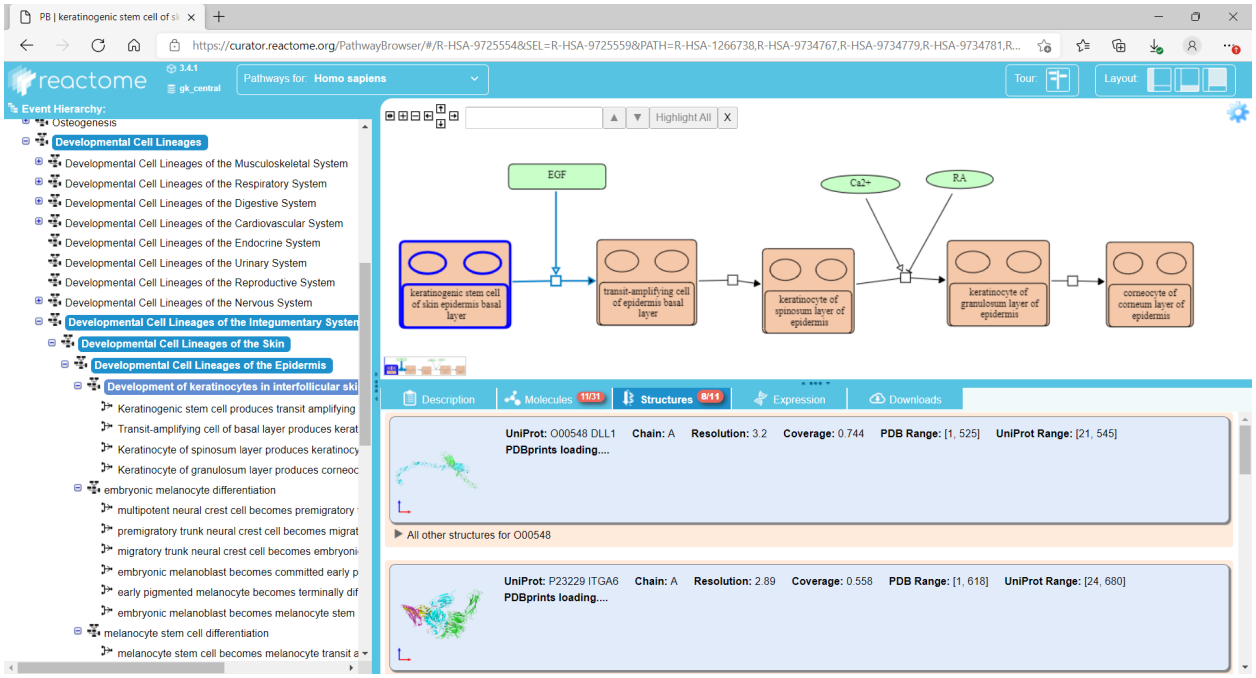
Opening the Markers menu for a "Cell" Physical Entity (by right-clicking or clicking the blue info icon when hovering over a "Cell" glyph in the Pathway Diagram)



Displaying the "Molecules Tab" (in the "Details Panel"):



Displaying the "Structures Tab" (for a selected "Cell" in the "Details Panel"):



Running the Project Locally

1. Create a new directory for the project in the location you choose and open a terminal at this location

2. Git clone the Pathway Diagram Curator, Pathway Browser Curator, and (optionally) RESTful API repositories

    git clone https://github.com/reactome/browser-curator

    git clone https://github.com/reactome/PathwayDiagram

    git clone https://github.com/reactome/RESTfulAPI (optional)

3. Run RESTfulAPI connected to a Reactome database (optional)

    - Navigate to the RESTful API directory and open a terminal
    - Run “git checkout cell-lineage-paths”
    - Open "web/WEB-INF/applicationContext.xml" and update the constructor arguments for MySQLAdaptor bean (replace the values in bold) to connect to a running Reactome database

```
<bean id="dba" class="org.gk.persistence.MySQLAdaptor">
  <constructor-arg index="0" value="DB_HOST"/
  <constructor-arg index="1" value="DB_NAME"/
  <constructor-arg index="2" value="DB_USER"/>
  <constructor-arg index="3" value="DB_PASS"/>
```

- Compile the RESTful API as a war file and deploy to a local Tomcat Server or run the project through Eclipse as a Tomcat application

3. Maven install the PathwayDiagram jar so it is available locally

    cd PathwayDiagram

    git checkout cell-lineage-path

    mvn clean install

    cd ..

4. Edit the browser-curator web.xml file to configure the RESTful API used

    - Open a terminal and navigate to the browser-curator directory
    - Run “git checkout cell-lineage-paths”
    - Open the file "src/main/webapp/WEB-INF/web.xml"
    - Change the proxyhost value (in bold below) to the server hosting the RESTful API (leave as "localhost" if running a local version of the RESTful API)

```
<!-- RESTFul WS PROXY Configuration (ONLY USED IN DEV MODE) -->
<servlet>
    <servlet-name>ReactomeRESTfulAPI</servlet-name>
    <servlet-class>org.reactome.server.utils.proxy.ProxyServlet</servlet-class>
    <init-param>
        <param-name>proxyHost</param-name>
        <param-value>localhost</param-value>
    </init-param>
    <init-param>
        <param-name>proxyPort</param-name>
        <param-value>8080</param-value>
    </init-param>
    <!-- Comment the following one to connect to https -->
    <init-param>
        <param-name>proxyProtocol</param-name>
        <param-value>http</param-value>
    </init-param>
    <init-param>
        <param-name>proxyPath</param-name>
```

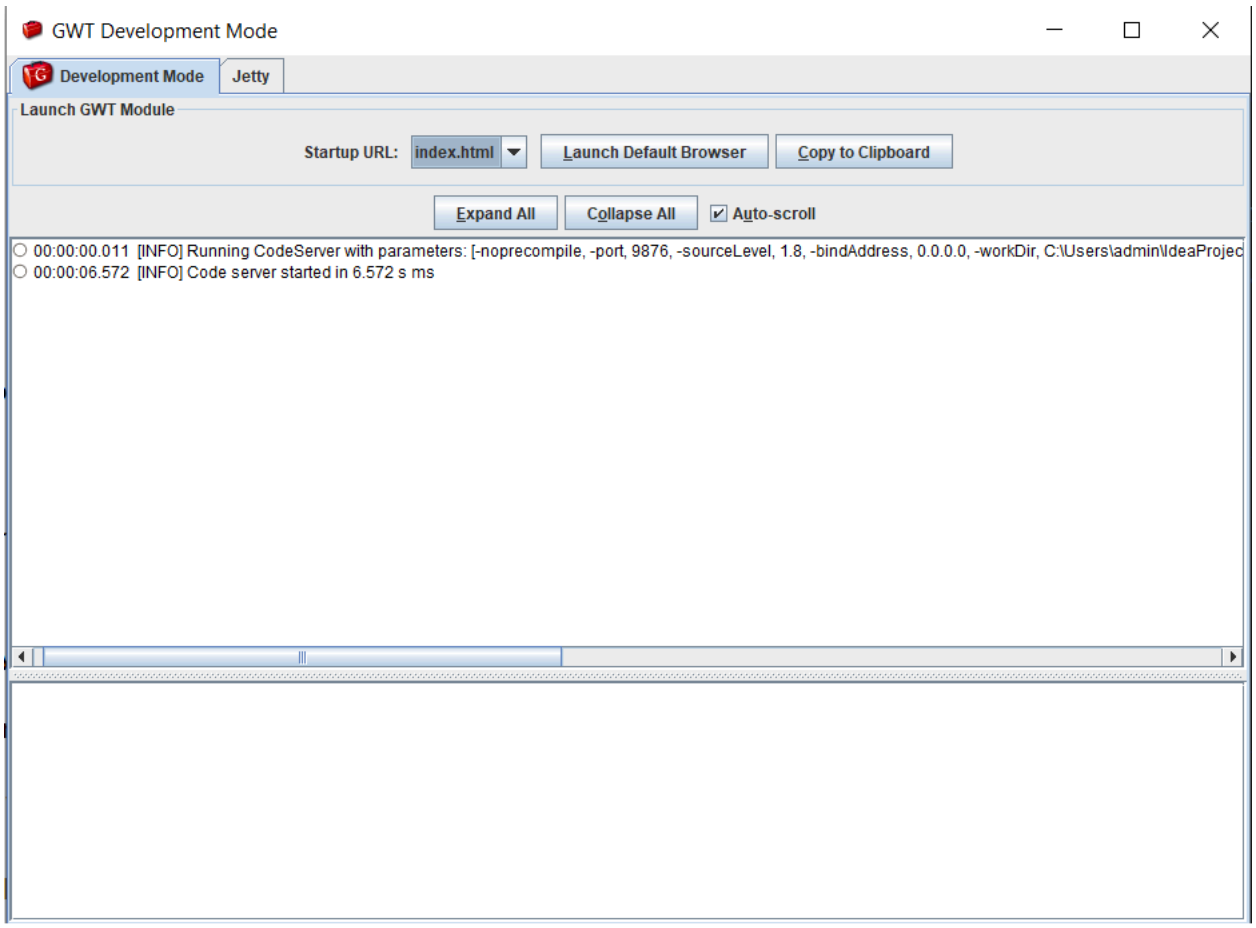
```
            <param-value>/ReactomeRESTfulAPI/RESTfulWS</param-value>
        </init-param>
    </servlet>
```

5. Compile and run the browser-curator

    cd browser-curator

    mvn clean compile; mvn gwt:devmode

6. Launch the Pathway Browser by clicking "Launch Default Browser"



Definitions

Anatomy - See "Data Model Changes" description for "Project Work Summary" section

Cell - See "Data Model Changes" description for "Project Work Summary" section

Cell Development Step - See "Data Model Changes" description for "Project Work Summary" section

Cell Lineage Path - See "Data Model Changes" description for "Project Work Summary" section

Marker Reference - See "Data Model Changes" description for "Project Work Summary" section

Pathway Browser Curator - An older version of the Reactome Pathway Browser to navigate curated Reactome pathways (i.e. both publicly released and not released). It includes a Pathway Hierarchy (i.e. a collection of general, top-level pathways with nested specific, lower-level pathways) to select a pathway to view, the Pathway Diagram to visualize and interact with the pathway components, and the Pathway Details Panel to provide more information for a selected pathway, reaction, physical entity from the Pathway Hierarchy or Pathway Diagram.

Pathway Diagram Curator - An older version of the Reactome Pathway Diagram display used as a component part of the Pathway Browser for visualizing curation data

RESTful API - A Reactome repository and service which, when hosted on a server, provides URLs to obtain data from a Reactome database in XML and JSON format

RTF (Rich-text format) - A text file with a specialized markup for describing visual elements in the doc

Known Issues/Bugs

- Selection of a CellLineagePath or CellDevelopmentStep in the Pathway Diagram does not highlight the corresponding selection in the Pathway Hierarchy
- Selection of a CellLineagePath or CellDevelopmentStep in the PathwayHierarchy jumps in the scroll to centre the higher level event

Opportunity for Feedback/Future Work

-

C10. Multiple disease reactions and EntityFunctionalStatus when using disease variant subsets

Should I create one or two disease reactions?

1. The question is, when different mutation affects say enzyme function in different ways (binding to substrate vs loss of enzyme activity) can the two types be represented as a single disease event in which the catalysis reaction is defective?

Here's the story:
I recently finished annotating loss-of-function mutations in MUTYH, a DNA glycosylase involved in base excision repair. For more details, please see here. To see failed reaction in gk_central, the DB_ID is 9605313.

There are about 30 MUTYH disease variants that have been functionally studied, most of them are missense, some are frameshift-caused truncations, and there's also an in-frame insertion.

All loss-of-function mutants lose, to a various extent, the catalytic ability. For some mutants, in addition to catalytic ability, studies also looked at substrate binding, with findings that some mutants can still bind the substrate but cannot process it, some cannot bind, some have reduced binding.

Three questions:

1) ONE OR TWO FAILED EVENTS?
I decided to annotate only one failed event, describing the loss of catalytic activity: "Defective MUTYH mutants do not cleave adenine mispaired with 8-oxoguanine". I believe this failed event captures and displays crucial information about MUTYH loss-of-function.
If I were to annotate an additional failed event: "Defective MUTYH mutants do not recognize and bind adenine opposite to an 8-oxoguanine", another disease pathway would need to be created, involving only mutants that do not bind to the substrate, in order for this failed event and the failed catalysis event to be displayed properly. I do not think that this addition would provide a significant extra information to the users.

Do you agree or do you think that another failed event would add value?

CONCLUSION #1: When there is evidence that a mutation affects both substrate binding and catalytic activity, it is beneficial to point this out by creating two failed reactions, one for binding, one for catalytic activity. This approach provides a more detailed mechanistic annotation of mutant proteins.

2) MULTI-VALUED ENTITY FUNCTIONAL STATUS?
Entity functional status is a multi-valued attribute. Therefore, I have added loss-of-function entity functional status due to various types of mutations to the failed reaction, as was done before for

already published reactions. Users should be able to tell from disease EWAS names which disease variant is affected by what type of mutation.
This is how the functional status displays:

This is similar to PTEN functional status released many years ago:

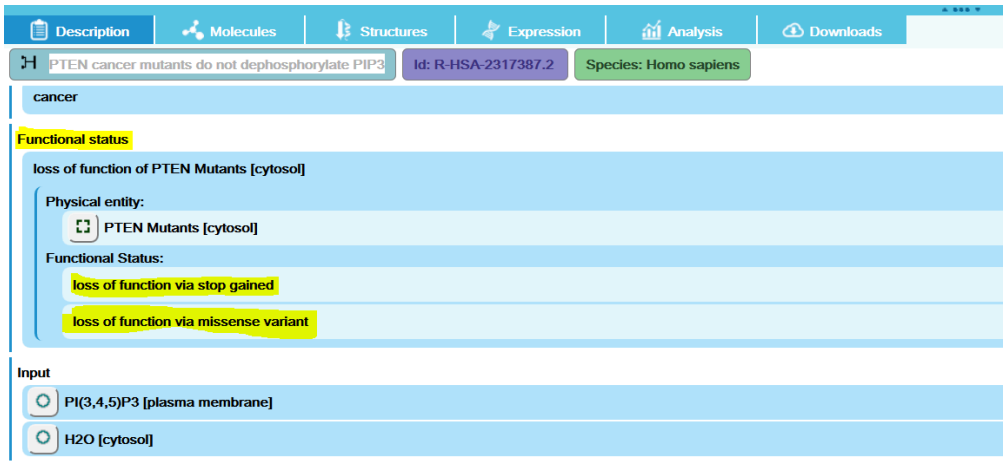
and BRAF gain-of-function:

and CFTR loss-of-function mutants:

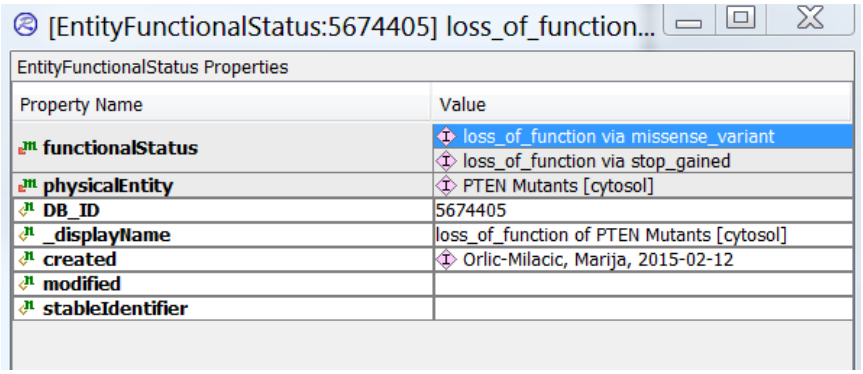
And ABCB4 loss-of-function mutants:

And many others.

CONCLUSION #2: EntityFunctionalStatus,which is used to categorize a disease RLE (gain- or a loss-of-function, or and to describe the underlying reason for the change in behavior), is a multivalued attribute. In a disease reaction, when a wt protein is replaced with a *set* (set A) of its disease variants, such that the function of each disease set member is affected in a *similar way* by different types of underlying mutations (such as loss of function due to gain of stop or loss of function due to missense mutation) the set can all be described with a single EntityFunctionalStatus (i.e loss of function of "set"), associating with it as many different sequence ontology terms as applicable in the "functionalStatus". For example, in the reaction R-HSA-2317387 "PTEN cancer mutants do not dephosphorylate PIP3", EntityFunctionalStatus is "loss of function of PTEN Mutants [cytosol], and includes two sequence ontology terms: "loss of function via stop gained" and "loss of function via missense variant".



This is how this annotation looks in the Curator Tool:



3) SHOULD THERE BE SUBSETS IN A SET OF MUTANTS THAT CORRESPOND TO MUTATION TYPE?

In the case of PTEN mutants, mutants were separated into subsets based on the type of mutation. However, in the case of MUTYH, BRAF, CFTR and ABCB4, there was no subgrouping of mutants based on mutation type.

Is subgrouping of mutants necessary, does it add anything, or should users simply figure out who is who from standardized mutant names?

CONCLUSION #3: Although not formally required so far, it is beneficial for the users to group disease set members into subsets that correspond to underlying mutation types, so that the number of subsets is equal to the numerical value of the EntityFunctionalStatus of a reaction in which this mutant set participates (e.g. if three different sequence ontology terms are listed under EntityFunctional status, the mutant set should have three subsets; for example, if the EntityFunctionalStatus lists the following: loss_of_function_via_missense_variant and loss_of_function_via_stop_gained, the mutants set “X mutants” that participates in the reactions should have subsets X_missense_variants and X_nonsense_variants.

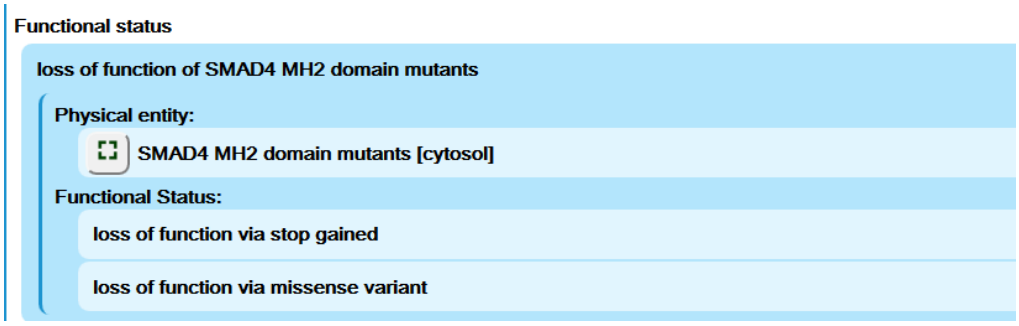
So far, we have not been consistent with creating subsets of mutants based on sequence variant type. In the majority of cases, no subsets have been created.

Here are some examples to illustrate who this looks in practice.

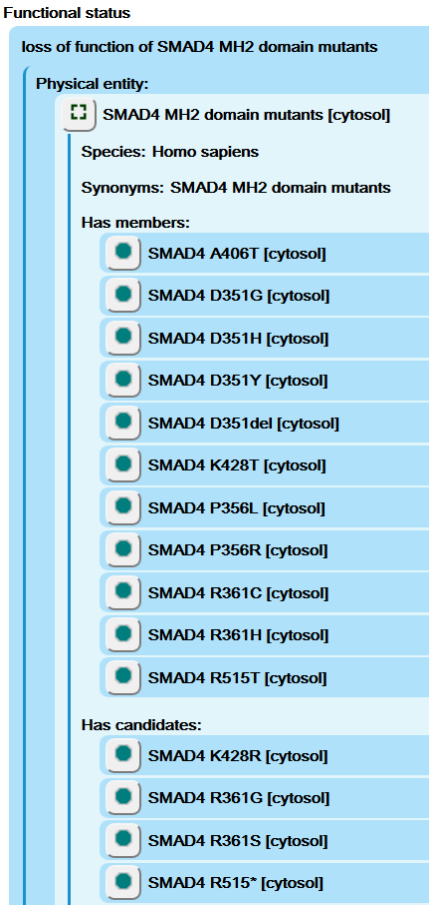
1) In the above-mentioned reaction R-HSA-2317387 “PTEN cancer mutants do not dephosphorylate PIP3”, EntityFunctionalStatus is “loss of function of PTEN Mutants [cytosol], and includes two sequence ontology terms: “loss of function via stop gained” and “loss of function via missense variant”. The mutant set “PTEN Mutants” includes two subsets, “PTEN Phosphatase Domain Missense Mutants” and “PTEN Truncation Mutants”. The first subsets (missense mutants) corresponds to the functional status “loss of function via missense variant”, while the second subset (truncation mutants) corresponds to the functional status “loss of function via stop gained”.

In order to implement the rule for creating subsets, we should come up with an unambiguous list of sequence variant-related naming terms for mutant subsets.

2) In the reaction R-HSA-3311014 “SMAD4 MH2 Domain Mutants do not bind phosphorylated SMAD2 and SMAD3”, EntityFunctionalStatus lists two terms: “loss of function via stop gained” and “loss of function via missense variant”.



The mutant set,”SMAD4 MH2 domain mutants”, however, contains no subsets:



Users can tell who is who if they known naming rules for mutations at the protein level.

C11. GtP_all_interactions_dump_110121 - IUPHAR_all_interactions_dump_110121

target	target_id	target_gene	target_uniprot	target_ensembl_gene_id	target_species	ligand	ligand_id	ligand_pubchem	approved_drug	type	action	action_comment	selectivity	endogenous	primary_target	concentration_r	affinity_units	affinity_high	affinity_median	affinity_low	original_affinity	original_affinity	original_affinity	original_affinity	assay_descripti	receptor_site	ligand_context	pubmed_id
4-aminobutyrate aminotransferase	2464	ABAT	P80404	ENSG00000183044	Human	vigabatrin	4821	178101523	t	Inhibitor	Irreversible inhibition		Not Determined	f	t		pKi		3.069999933		Ki		850000		#ERROR!				8565821	21686767
5-HT_{1A} receptor	1	HTR1A	P08908	ENSG00000178394	Human	flibanserin	8182	249565862	t	Agonist			Not Determined	f	t		pKi		9		Ki		1		#ERROR!				12177684
5-HT_{1A} receptor	1	HTR1A	P08908	ENSG00000178394	Human	buspirone	36	136560035	t	Agonist	Partial agonist		Not Determined	f	t		pKi	8		7.699999809	-				#ERROR!				1386736|9760039|9550290
5-HT_{1A} receptor	1	HTR1A	P08908	ENSG00000178394	Human	vilazodone	7427	178103999	t	Agonist	Partial agonist		Not Determined	f	t		pKi		9.699999809		Ki		0.200000003		#ERROR!	Binding of vilazodone to human 5-HT_{1A}			19499624
5-HT_{1A} receptor	1	HTR1A	P08908	ENSG00000178394	Human	vilazodone	7427	178103999	t	Agonist	Partial agonist		Not Determined	f	t		pIC50		9.520000458		IC50		0.300000012		#ERROR!				15341484
5-HT_{1A} receptor	1	HTR1A	P08908	ENSG00000178394	Human	vortioxetine	7351	178103923	t	Agonist	Partial agonist		Non-selective	f	t		pKi		7.820000172		Ki		15		#ERROR!				21486038
5-HT_{1B} receptor	2	HTR1B	P28222	ENSG00000135312	Human	frovatriptan	7191	178103766	t	Agonist	Agonist		Not Determined	f	t		pKi		7.989999771		Ki		10.30000019		#ERROR!				9986723
5-HT_{1B} receptor	2	HTR1B	P28222	ENSG00000135312	Human	eletriptan	40	136650227	t	Agonist	Full agonist		Not Determined	f	t		pKi		8		Ki		10		#ERROR!				10193663
5-HT_{1B} receptor	2	HTR1B	P28222	ENSG00000135312	Human	naratriptan	45	136650691	t	Agonist	Partial agonist		Not Determined	f	t		pKi		8.100000381		-				#ERROR!				10193663
5-HT_{1B} receptor	2	HTR1B	P28222	ENSG00000135312	Human	rizatriptan	51	136650905	t	Agonist	Partial agonist		Not Determined	f	t		pKi		6.900000095		-				#ERROR!				10193663
5-HT_{1B} receptor	2	HTR1B	P28222	ENSG00000135312	Human	sumatriptan	54	136651095	t	Agonist	Partial agonist		Not Determined	f	t		pKi	8.100000381		6.5	Ki	316		7.940000057		#ERROR!				10611634|10193663|1565658|8863519|10513577|1
5-HT_{1B} receptor	2	HTR1B	P28222	ENSG00000135312	Human	vortioxetine	7351	178103923	t	Agonist	Partial agonist		Non-selective	f	t		pKi		7.480000019		Ki		33		#ERROR!				21486038
5-HT_{1B} receptor	2	HTR1B	P28222	ENSG00000135312	Human	zolmitriptan	60	136651366	t	Agonist	Partial agonist		Not Determined	f	t		pKi		7.699999809		-				#ERROR!				10193663
5-HT_{1B} receptor	2	HTR1B	P28222	ENSG00000135312	Human	risperidone	96	136650903	t	Antagonist	Antagonist		Not Determined	f	t		pKi	7.300000191		6.599999905	-				#ERROR!				8935801|9605573|18308814
5-HT_{1D} receptor	3	HTR1D	P28221	ENSG00000179546	Human	ergotamine	149	136650239	t	Agonist	Agonist		Not Determined	f	t		pKi		9.100000381		Ki		0.800000012		#ERROR!				10715164
5-HT_{1D} receptor	3	HTR1D	P28221	ENSG00000179546	Human	frovatriptan	7191	178103766	t	Agonist	Agonist		Not Determined	f	t		pKi		8.359999657		Ki		4.400000095		#ERROR!				9986723
5-HT_{1D} receptor	3	HTR1D	P28221	ENSG00000179546	Human	dihydroergotamine	121	136650182	t	Agonist	Full agonist		Not Determined	f	t		pKi	9.899999619		9.199999809	-				#ERROR!				9605573|8967979|1652050
5-HT_{1D} receptor	3	HTR1D	P28221	ENSG00000179546	Human	eletriptan	40	136650227	t	Agonist	Full agonist		Selective	f	t		pKi		8.899999619		Ki		1.2599999		#ERROR!				10193663
5-HT_{1D} receptor	3	HTR1D	P28221	ENSG00000179546	Human	naratriptan	45	136650691	t	Agonist	Full agonist		Not Determined	f	t		pKi	9		8.399999619	Ki				#ERROR!				10193663|9303769|7780656
5-HT_{1D} receptor	3	HTR1D	P28221	ENSG00000179546	Human	rizatriptan	51	136650905	t	Agonist	Full agonist		Not Determined	f	t		pKi		7.900000095		-				#ERROR!				10193663
5-HT_{1D} receptor	3	HTR1D	P28221	ENSG00000179546	Human	sumatriptan	54	136651095	t	Agonist	Full agonist		Not Determined	f	t		pKi	8.699999809		8	Ki	10		1.99000001		#ERROR!				10611634|10193663|1565658|8967979|1652050
5-HT_{1D} receptor	3	HTR1D	P28221	ENSG00000179546	Human	zolmitriptan	60	136651366	t	Agonist	Full agonist		Not Determined	f	t		pKi		8.899999619		-				#ERROR!				10193663
5-HT_{1D} receptor	3	HTR1D	P28221	ENSG00000179546	Human	risperidone	96	136650903	t	Antagonist	Antagonist		Not Determined	f	t		pKi	8		7.800000191	-				#ERROR!				8935801|8967979
5-HT_{1F} receptor	5	HTR1F	P30939	ENSG00000179097	Human	eletriptan	40	136650227	t	Agonist	Full agonist		Not Determined	f	t		pKi		8		Ki		10		#ERROR!				10193663
5-HT_{2A} receptor	6	HTR2A	P28223	ENSG00000102468	Human	asenapine	22	136650779	t	Antagonist	Antagonist		Not Determined	f	t		pKi		10.22000027		Ki		0.059999999		#ERROR!				18308814
5-HT_{2A} receptor	6	HTR2A	P28223	ENSG00000102468	Human	blonanserin	7670	223366005	t	Antagonist	Antagonist		Not Determined	f	t		pKi		9.090000153		Ki		0.810000002		#ERROR!				15686911
5-HT_{2A} receptor	6	HTR2A	P28223	ENSG00000102468	Human	flibanserin	8182	249565862	t	Antagonist	Antagonist		Not Determined	f	t		pKi		7.309999943		Ki		49		#ERROR!				12177684
5-HT_{2A} receptor	6	HTR2A	P28223	ENSG00000102468	Human	haloperidol	86	136650345	t	Antagonist	Antagonist		Not Determined	f	t		pKi	7.300000191		6.699999809	-				#ERROR!				8935801|9732398|12629531|15102927|18308814
5-HT_{2A} receptor	6	HTR2A	P28223	ENSG00000102468	Human	mianserin	135	136650606	t	Antagonist	Antagonist		Not Determined	f	t		pKi	9.600000381		7.699999809	-				#ERROR!				10611634|15322733|15107597
5-HT_{2A} receptor	6	HTR2A	P28223	ENSG00000102468	Human	mirtazapine	7241	178103815	t	Antagonist	Antagonist		Not Determined	f	t		pKi		7.159999847		Ki		69		#ERROR!				15771415
5-HT_{2A} receptor	6	HTR2A	P28223	ENSG00000102468	Human	molindone	207	136650621	t	Antagonist	Antagonist		Not Determined	f	t		pKi		6.5		-				#ERROR!				12629531
5-HT_{2A} receptor	6	HTR2A	P28223	ENSG00000102468	Human	olanzapine	7247	178103821	t	Antagonist	Antagonist		Not Determined	f	t		pKi		8.239999771		Ki		5.800000191		#ERROR!				21823597
5-HT_{2A} receptor	6	HTR2A	P28223	ENSG00000102468	Human	olanzapine	47	136650756	t	Antagonist	Antagonist		Not Determined	f	t		pKi	8.899999619		8.600000381	-				#ERROR!				8935801|12629531|12176106|18308814
5-HT_{2A} receptor	6	HTR2A	P28223	ENSG00000102468	Human	pimozide	90	136650825	t	Antagonist	Antagonist		Not Determined	f	t		pKi	7.699999809		7.099999905	-				#ERROR!				12629531|12176106
5-HT_{2A} receptor	6	HTR2A	P28223	ENSG00000102468	Human	quetiapine	50	136650869	t	Antagonist	Antagonist		Not Determined	f	t		pKi	7		6.400000095	-				#ERROR!				12629531|12176106
5-HT_{2A} receptor	6	HTR2A	P28223	ENSG00000102468	Human	sertindole	98	136651037	t	Antagonist	Antagonist		Not Determined	f	t		pKi	9.399999619		9.199999809	-				#ERROR!				8935801|12629531|12176106
5-HT_{2A} receptor	6	HTR2A	P28223	ENSG00000102468	Human	thioridazine	100	136651259	t	Antagonist	Antagonist		Not Determined	f	t		pKi	8		7.400000095	-				#ERROR!				12629531|12176106
5-HT_{2A} receptor	6	HTR2A	P28223	ENSG00000102468	Human	thiothixene	212	136651260	t	Antagonist	Antagonist		Not Determined	f	t		pKi		7.300000191		-				#ERROR!				12629531
5-HT_{2A} receptor	6	HTR2A	P28223	ENSG00000102468	Human	trazodone	213	136651275	t	Antagonist	Antagonist		Not Determined	f	t		pKi		7.400000095		-				#ERROR!				15322733
5-HT_{2A} receptor	6	HTR2A	P28223	ENSG00000102468	Human	ziprasidone	59	136651359	t	Antagonist	Antagonist		Not Determined	f	t		pKi	9.510000229		8.800000191	-				#ERROR!				8935801|12629531|12176106|18308814
5-HT_{2A} receptor	6	HTR2A	P28223	ENSG00000102468	Human	chlorpromazine	83	136650088	t	Antagonist	Inverse agonist		Not Determined	f	t		pKi		8.100000381		-				#ERROR!				12629531
5-HT_{2A} receptor	6	HTR2A	P28223	ENSG00000102468	Human	clozapine	38	136650115	t	Antagonist	Inverse agonist		Not Determined	f	t		pKi	9		7.599999905	-				#ERROR!				8935801|9732398|15322733|12629531|15102927
5-HT_{2A} receptor	6	HTR2A	P28223	ENSG00000102468	Human	loxapine	205	136650533	t	Antagonist	Inverse agonist		Not Determined	f	t		pKi		8.100000381		-				#ERROR!				12629531
5-HT_{2A} receptor	6	HTR2A	P28223	ENSG00000102468	Human	pimavanserin	8423	252166633	t	Antagonist	Inverse agonist		Selective	f	t		pKi		9.300000191		-				#ERROR!	Ki determined in membrane radioligand binding assa	15102927|7519387
5-HT_{2A} receptor	6	HTR2A	P28223	ENSG00000102468	Human	risperidone	96	136650903	t	Antagonist	Inverse agonist		Not Determined	f	t		pKi	10		9.300000191	-				#ERROR!				8935801|12629531|12176106
5-HT_{2B} receptor	7	HTR2B	P41595	ENSG00000135914	Human	methysergide	134	136650596	t	Agonist	Partial agonist		Not Determined	f	t		pKi	9.399999619		8	-				#ERROR!				15322733|11104741|9459568
5-HT_{2B} receptor	7	HTR2B	P41595	ENSG00000135914	Human	mianserin	135	136650606	t	Antagonist	Antagonist		Not Determined	f	t		pKi	8.800000191		7.900000095	-				#ERROR!				15322733|9459568|10455251
5-HT_{2C} receptor	8	HTR2C	P28335	ENSG00000147246	Human	lorcaserin	2941	136650530	t	Agonist	Agonist		Selective	f	t		pEC50		8.239999771		IC50		5.800000191		#ERROR!	IP accumulation assay.		NULL
5-HT_{2C} receptor	8	HTR2C	P28335	ENSG00000147246	Human	lorcaserin	2941	136650530	t	Agonist	Full agonist		Selective	f	t		pEC50		8.100000381		-				#ERROR!				18252809
5-HT_{2C} receptor	8	HTR2C	P28335	ENSG00000147246	Human	lorcaserin	2941	136650530	t	Agonist	Full agonist		Selective	f	t		pKi		7.800000191		-				#ERROR!				18252809
5-HT_{2C} receptor	8	HTR2C	P28335	ENSG00000147246	Human	methysergide	134	136650596	t	Antagonist	Antagonist		Not Determined	f	t		pKi	9.100000381		8.600000381	-				#ERROR!				15322733|10611640
5-HT_{2C} receptor	8	HTR2C	P28335	ENSG00000147246	Human	mirtazapine	7241	178103815	t	Antagonist	Antagonist		Not Determined	f	t		pKi		7.499999847		Ki		39		#ERROR!				15771415
5-HT_{2C} receptor	8	HTR2C	P28335	ENSG00000147246	Human	thioridazine	100	136651259	t	Antagonist	Antagonist		Not Determined	f	t		pKi	7.300000191		7.199999809	-				#ERROR!				12629531|10991983
5-HT_{2C} receptor	8	HTR2C	P28335	ENSG00000147246	Human	trazodone	213	136651275	t	Antagonist	Antagonist		Not Determined	f	t		pKi		6.599999905		-				#ERROR!				15322733
5-HT_{2C} receptor	8	HTR2C	P28335	ENSG00000147246	Human	loxapine	205	136650533	t	Antagonist	Inverse agonist		Not Determined	f	t		pKi	8		7.800000191	-				#ERROR!				12629531|10991983
5-HT_{2C} receptor	8	HTR2C	P28335	ENSG00000147246	Human	mianserin	135	136650606	t	Antagonist	Inverse agonist		Not Determined	f	t		pKi	9.199999809		8.300000191	-				#ERROR!				10611634|15322733|10217294
5-HT_{2C} receptor	8	HTR2C	P28335	ENSG00000147246	Human	olanzapine	47	136650756	t	Antagonist	Inverse agonist		Not Determined	f	t		pKi	8.399999619		8.100000381	-				#ERROR!				12629531|10991983|18308814
5-HT_{2C} receptor	8	HTR2C	P28335	ENSG00000147246	Human	sertindole	98	136651037	t	Antagonist	Inverse agonist		Not Determined	f	t		pKi	9.199999809		9	-				#ERROR!				12629531|10991983
5-HT_{2C} receptor	8	HTR2C	P28335	ENSG00000147246	Human	ziprasidone	59	136651359	t	Antagonist	Inverse agonist		Not Determined	f	t		pKi	9.010000229		7.900000095	-				#ERROR!				12629531|10991983|18308814
5-HT₃ receptor	379	HTR3A	P46098	ENSG00000166736	Human	alosetron	2296	136649921	t	Antagonist	Antagonist		Selective	f	t		pKi		9.5		Ki		0.316000015		#ERROR!				17652911
5-HT₃ receptor	379	HTR3A	P46098	ENSG00000166736	Human	granisetron	2300	136650322	t	Antagonist	Antagonist		Selective	f	t		pKi	8.800000191		8.600000381	Ki	2.5		1.5	~	#ERROR!				7565620|8818349
5-HT₃ receptor	379	HTR3A	P46098	ENSG00000166736	Human	metoclopramide	241	136650598	t	Antagonist	Antagonist		Not Determined	f	t		pKi	6.449999809		6	Ki	1030		354	~	#ERROR!				11488465|8818349
5-HT₃ receptor	379	HTR3A	P46098	ENSG00000166736	Human	ondansetron	2290	136650761	t	Antagonist	Antagonist		Selective	f	t		pKi	8.300000191		7.800000191	Ki	15		5	~				7565620|11488465|8818349
5-HT₃ receptor	379	HTR3A	P46098	ENSG00000166736	Human	palonosetron	7486	178051789	t	Antagonist	Antagonist		Selective	f	t		pKi		10.48999977		Ki		0.032000002		#ERROR!				20724042
5-HT₃ receptor	379	HTR3A	P46098	ENSG00000166736	Human	tropisetron	260	136651284	t	Antagonist	Antagonist		Selective	f	t		pKi	8.800000191		8.5	Ki	3		1.5		#ERROR!				7565620|9463477
5-HT₃ receptor	379	HTR3A	P46098	ENSG00000166736	Human	vortioxetine	7351	178103923	t	Antagonist	Antagonist		Non-selective	f	t		pKi		8.430000305		Ki		3.700000048		#ERROR!				21486038
5-HT₃/sub>AB	378	HTR3B|HTR3A	P46098|O95264	ENSG00000166736|ENS	Human	metoclopramide	241	136650598	t	Antagonist	Antagonist		Not Determined	f	t		pKi		5.699999809		-				#ERROR!				11489465
5-HT₃/sub>AB	378	HTR3B|HTR3A	P46098|O95264	ENSG00000166736|ENS	Human	ondansetron	2290	136650761	t	Antagonist	Antagonist		Not Determined	f	t		pKi		7.800000191		-				#ERROR!				11489465
5-HT₄ receptor	9	HTR4	Q13639	ENSG00000164270	Human	cisapride	240	136650104	t	Agonist	Partial agonist		Not Determined	f	t		pKi	7.400000095		6.400000095	-				#ERROR!				9349523|7796807|10464498|11218067|10683202|1
5-HT₇ receptor	12	HTR7	P34969	ENSG00000148680	Human	lurasidone	7461	187051764	t	Antagonist	Antagonist		Not Determined	f	t		pKi		9.31000042		Ki		0.495000005		#ERROR!				20404009
5-HT₇ receptor	12	HTR7	P34969	ENSG00000148680	Human	vortioxetine	7351	178103923	t	Antagonist	Antagonist		Non-selective	f	t		pKi		6.349999905		Ki		450		#ERROR!				21486038
5-LOX	1385	ALOX5	P09917	ENSG00000012779	Human	zileuton	5297	178101979	t	Inhibitor	Inhibition		Selective	f	t		pIC50		6.519999981		IC50		300		#ERROR!	Inhibition of recombinant 5-LOX in vitro.			24508141
5-LOX	1385	ALOX5	P09917	ENSG00000012779	Human	zileuton	5297	178101979	t	Inhibitor	Inhibition		Selective	f	t		pIC50		6.400000095		IC50		400		#ERROR!	Inhibition of leukotriene B4 biosynthesis by human p			1848634
5-LOX	1385	ALOX5	P09917	ENSG00000012779	Human	zileuton	5297	178101979	t	Inhibitor	Inhibition		Selective	f	t		pIC50		5.820000172		IC50		1500		#ERROR!	Inhibition of purified recombinant human 5-LOX in a			30199704
ABCA1	756	ABCA1	O95477	ENSG00000165029	Human	probucol	7277	178103851	t	Inhibitor	Inhibition		Selective	f	t		-				-				#ERROR!	Inhibition of cholesterol efflux from normal human sk	15514211|15140889	
ABL proto-oncogene 1, non-receptor tyros	1923	ABL1	P00519	ENSG00000097007	Human	bosutinib	5710	178102336	t	Inhibitor	Inhibition		Not Determined	f	t		pIC50		9		IC50		1		#ERROR!				12543790
ABL proto-oncogene 1, non-receptor tyros	1923	ABL1	P00519	ENSG00000097007	Human	dasatinib	5678	163923305	t	Inhibitor	Inhibition		Not Determined	f	t		pIC50		9.600000381		IC50		0.270000011		#ERROR!				23279183
ABL proto-oncogene 1, non-receptor tyros	1923	ABL1	P00519	ENSG00000097007	Human	dasatinib	5678	178102305	t	Inhibitor	Inhibition		Not Determined	f	t		pKd		10.48999977		Kd		0.032000002		#ERROR!	Binding affinity for the unphosphorylated kinase dom			31539241
ABL proto-oncogene 1, non-receptor tyros	1923	ABL1	P00519	ENSG00000097007	Human	imatinib	5687	178102314	t	Inhibitor	Inhibition		Not Determined	f	t		pIC50		6.699999809		IC50		220		#ERROR!				17376680
ABL proto-oncogene 1, non-receptor tyros	1923	ABL1	P00519	ENSG00000097007	Human	nilotinib	5697	178102324	t	Inhibitor	Inhibition		Not Determined	f	t		pIC50		7.800000191		IC50		15		#ERROR!				15930265
ABL proto-oncogene 1, non-receptor tyros	1923	ABL1	P00519	ENSG00000097007	Human	ponatinib	5890	178102513	t	Inhibitor	Inhibition		Not Determined	f	t		pIC50		8.100000381		IC50		8.600000381		#ERROR!				20513156
acetylcholinesterase (Cartwright blood grc	2465	ACHE	P22303	ENSG00000087085	Human	donepezil	6599	178103212	t	Inhibitor	Inhibition		Selective	f	t		pIC50	8.300000191		7.699999809	-				#ERROR!				16570913|8039546|1738151
acetylcholinesterase (Cartwright blood grc	2465	ACHE	P22303	ENSG00000087085	Human	physostigmine	6598	178103211	t	Inhibitor	Inhibition		Family selective	f	t		pIC50	7.800000191		7.599999905	-				#ERROR!				16570913
acetylcholinesterase (Cartwright blood grc	2465	ACHE	P22303	ENSG00000087085	Human	pralidoxime	7597	223365932	t	Inhibitor	Inhibition		Not Determined	f	t		pIC50		3.099999905		IC50		878000		#ERROR!	Inhibition of human recombinant AChE by modified E			21215642
acetylcholinesterase (Cartwright blood grc	2465	ACHE	P22303	ENSG00000087085	Human	rivastigmine	6602	178103215	t	Inhibitor	Inhibition		Not Determined	f	t		pIC50		5.400000095		IC50		4150		#ERROR!				16570913
acetylcholinesterase (Cartwright blood grc	2465	ACHE	P22303	ENSG00000087085	Human	tacrine	6687	178103299	t	Inhibitor	Inhibition		Not Determined	f	t		pKi		7.5		Ki		30		#ERROR!				18479118
Adenosine deaminase	1230	ADA	P00813	ENSG00000196839	Human	pentostatin	4805	178101507	t	Inhibitor	Inhibition		Selective	f	t		pIC50		10.80000019		IC50				#ERROR!				849300
aldo-keto reductase family 1 member B	2768	AKR1B1	P15121	ENSG00000085662	Human	tolrestat	7404	178103976	t	Inhibitor	Inhibition		Not Determined	f	t		pIC50		7.329999924		IC50		47		#ERROR!				10737739
ALK receptor tyrosine kinase	1839	ALK	Q9UM73	ENSG00000171094	Human	alectinib	7739	223366072	t	Inhibitor	Inhibition		Not Determined	f	t		pIC50		8.720000267		IC50		1.899999976		#ERROR!				22225917
ALK receptor tyrosine kinase	1839	ALK	Q9UM73	ENSG00000171094	Human	brigatinib	7741	223366074	t	Inhibitor	Inhibition		Not Determined	f	t		pIC50	9.430000305		9.210000038	IC50	0.620000005		0.370000005	#ERROR!				21502504|27144831
ALK receptor tyrosine kinase	1839	ALK	Q9UM73	ENSG00000171094	Human	brigatinib	7741	223366074	t	Inhibitor	Inhibition		Not Determined	f	t		pKi		10.05000019		Ki		0.090000004		#ERROR!	Wild type ALK			21502504
ALK receptor tyrosine kinase	1839	ALK	Q9UM73	ENSG00000171094	Human	ceritinib	7397	137033969	t	Inhibitor	Inhibition		Selective	f	t		pIC50		9.699999809		IC50		0.200000003		#ERROR!				23742252
ALK receptor tyrosine kinase	1839	ALK	Q9UM73	ENSG00000171094	Human	crizotinib	4903	178101602	t	Inhibitor	Inhibition		Not Determined	f	t		pIC50		9		IC50		1		#ERROR!				21812414
ALK receptor tyrosine kinase	1839	ALK	Q9UM73	ENSG00000171094	Human	loriatinib	7476	178051779	t	Inhibitor	Inhibition		Not Determined	f	t		pKi		9.149999619		Ki		0.699999988		#ERROR!	Note that this is inhibition of the ALK L1196M mutant			24819116
α_{1A}-adrenoceptor	22	ADRA1A	P35348	ENSG00000120907	Human	methoxamine	483	136650587	t	Agonist	Full agonist		Not Determined	f	t		pKi	5.199999809		5	-				#ERROR!				7651358|10433504
α_{1A}-adrenoceptor	22	ADRA1A	P35348	ENSG00000120907	Human	oxymetazoline	124	136650785	t	Agonist	Full agonist		Non-selective	f	t		pKi	8.199999809		8	-				#ERROR!				7651358|8719417|10433504|8564227
α_{1A}-adrenoceptor	22	ADRA1A	P35348	ENSG00000120907	Human	phenylephrine	485	136650819	t	Agonist	Full agonist		Not Determined	f	t		pKi	5.400000095		5.199999809	-				#ERROR!				10433504
α_{1A}-adrenoceptor	22	ADRA1A	P35348	ENSG00000120907	Human	alfuzosin	7109	178103686	t	Antagonist	Antagonist		Not Determined	f	t		pKi		8.100000381		Ki		8.199999809		#ERROR!				10812954
α_{1A}-adrenoceptor	22	ADRA1A	P35348	ENSG00000120907	Human	doxazosin	7170	178103745	t	Antagonist	Antagonist		Not Determined	f	t		pKi		9.300000191		Ki		0.537		#ERROR!			NULL	
α_{1A}-adrenoceptor	22	ADRA1A	P35348	ENSG00000120907	Human	phentolamine	502	136650818	t	Antagonist	Antagonist		Not Determined	f	t		pKi		8.600000381		-				#ERROR!				7651358
α_{1A}-adrenoceptor	22	ADRA1A	P35348	ENSG00000120907	Human	silodosin	493	136650458	t	Antagonist	Antagonist		Selective	f	t		pKi		10.39999962		-				#ERROR!				7651358
α_{1A}-adrenoceptor	22	ADRA1A	P35348	ENSG00000120907	Human	tamsulosin	488	136651350	t	Antagonist	Antagonist		Selective	f	t		pKi	10.69999981		10	-				#ERROR!				9249248|7651358|9490024|10369480|10334511
α_{1A}-adrenoceptor	22	ADRA1A	P35348	ENSG00000120907	Human	terazosin	7302	178103875	t	Antagonist	Antagonist		Not Determined	f	t		pKi		8.699999809		Ki		2		#ERROR!				9379432
α_{1A}-adrenoceptor	22	ADRA1A	P35348	ENSG00000120907	Human	prazosin	503	136650849	t	Antagonist	Inverse agonist		Not Determined	f	t		pKi	9.899999619		9	-				#ERROR!				9249248|7651358|9490024|10369480|10334511
α_{1B}-adrenoceptor	23	ADRA1B	P35368	ENSG00000170214	Human	phenylephrine	485	136650819	t	Agonist	Full agonist		Not Determined	f	t		pIC50	7.5		6.300000191	-				#ERROR!				9249248|7969082
α_{1B}-adrenoceptor	23	ADRA1B	P35368	ENSG00000170214	Human	alfuzosin	7109	178103686	t	Antagonist	Antagonist		Not Determined	f	t		pKi		8.600000381		Ki		2.799999952		#ERROR!				7658428
α_{1B}-adrenoceptor	23	ADRA1B	P35368	ENSG00000170214	Human	doxazosin	7170	178103745	t	Antagonist	Antagonist		Not Determined	f	t		pKi		9.100000381		Ki		0.813000023		#ERROR!			NULL	
α_{1B}-adrenoceptor	23	ADRA1B	P35368	ENSG00000170214	Human	phentolamine	502	136650818	t	Antagonist	Antagonist		Not Determined	f	t		pKi		7.5		-				#ERROR!				7651358
α_{1B}-adrenoceptor	23	ADRA1B	P35368	ENSG00000170214	Human	terazosin	7302	178103875	t	Antagonist	Antagonist		Not Determined	f	t		pKi		8.600000381		Ki		2.680000067		#ERROR!				9379432
α_{1B}-adrenoceptor	23	ADRA1B	P35368	ENSG00000170214	Human	prazosin	503	136650849	t	Antagonist	Inverse agonist		Not Determined	f	t		pKi	9.899999619		9.600000381	-				#ERROR!				9249248|7651358|10369480
α_{1B}-adrenoceptor	23	ADRA1B	P35368	ENSG00000170214	Human	tamsulosin	488	136651350	t	Antagonist	Inverse agonist		Not Determined	f	t		pKi	9.699999809		9.5	-				#ERROR!				9249248|7651358|10369480
α_{1D}-adrenoceptor	24	ADRA1D	P25100	ENSG00000171873	Human	alfuzosin	7109	178103686	t	Antagonist	Antagonist		Not Determined	f	t		pKi		8.399999619		Ki		3.599999905		#ERROR!				10812954
α_{1D}-adrenoceptor	24	ADRA1D	P25100	ENSG00000171873	Human	dapiprazole	7155	178103730	t	Antagonist	Antagonist		Not Determined	f	t		pKi		8.399999619		Ki		4.090000153		#ERROR!			NULL	
&al

target	target_id	target_gene	target_uniprot	target_ensembl_gene_id	target_species	ligand	ligand_id	ligand_pubchen	approved_drug	type	action	action_commen	selectivity	endogenous	primary_target	concentration_r	affinity_units	affinity_high	affinity_median	affinity_low	original_affinity	original_affinity	original_affinity	original_affinity	assay_descripti	receptor_site	ligand_context	pubmed_id
α_{2A}-adrenoceptor	25	ADRA2A	P08913	ENSG00000150594	Human	apraclonidine	7117	178103694	t	Agonist	Agonist		Not Determined	f	t		pKi		8.5		Ki		2.900000095		#ERROR!				8784451
α_{2A}-adrenoceptor	25	ADRA2A	P08913	ENSG00000150594	Human	guanabenz	5443	178102094	t	Agonist	Agonist		Not Determined	f	t		pIC50		7.400000095		IC50		41		#ERROR!			NULL
α_{2A}-adrenoceptor	25	ADRA2A	P08913	ENSG00000150594	Human	brimonidine	520	135650022	t	Agonist	Full agonist		Not Determined	f	t		pKi	8.699999809		6.699999809	-				#ERROR!			9824686|9605427|9760042|9227000
α_{2A}-adrenoceptor	25	ADRA2A	P08913	ENSG00000150594	Human	clonidine	516	135650112	t	Agonist	Partial agonist		Not Determined	f	t		pKi	9.199999809		7.199999809	-				#ERROR!			9824686|9605427|9760042
α_{2A}-adrenoceptor	25	ADRA2A	P08913	ENSG00000150594	Human	dexmedetomidine	521	135650174	t	Agonist	Partial agonist		Not Determined	f	t		pKi	9.600000381		7.599999905	-				#ERROR!			9824686|9605427|9760042|9227000
α_{2A}-adrenoceptor	25	ADRA2A	P08913	ENSG00000150594	Human	guanfacine	522	135650327	t	Agonist	Partial agonist		Not Determined	f	t		pKi	7.300000191		7.099999905	-				#ERROR!			9605427|9283709
α_{2A}-adrenoceptor	25	ADRA2A	P08913	ENSG00000150594	Human	oxymetazoline	124	135650785	t	Agonist	Partial agonist		Selective	f	t		pKi	8.600000381		8	Ki	10		2.799999952	#ERROR!			9605427|7996470|9227000
α_{2A}-adrenoceptor	25	ADRA2A	P08913	ENSG00000150594	Human	mirtazapine	7241	178103815	t	Antagonist	Antagonist		Not Determined	f	t		pIC50		7.099999905		IC50		85.11000061		#ERROR!			10636247
α_{2A}-adrenoceptor	25	ADRA2A	P08913	ENSG00000150594	Human	mirtazapine	7241	178103815	t	Antagonist	Antagonist		Not Determined	f	t		pKi		7.699999809		Ki		20		#ERROR!	Inhibition of [³H]rauwolscine binding.		15771415
α_{2A}-adrenoceptor	25	ADRA2A	P08913	ENSG00000150594	Human	tolazoline	7310	178103883	t	Antagonist	Antagonist		Not Determined	f	t		pKi		6.699999809		Ki		199		#ERROR!	Inhibition of agonist-stimulated [³⁵S]GT		9605427
α_{2A}-adrenoceptor	25	ADRA2A	P08913	ENSG00000150594	Human	yohimbine	102	135652727	t	Antagonist	Antagonist		Non-selective	f	t		pKi	9.199999809		8.399999619	-				#ERROR!			7996470|13532477|908642
α_{2B}-adrenoceptor	26	ADRA2B	P18089	ENSG00000222040	Human	guanabenz	5443	178102094	t	Agonist	Agonist		Not Determined	f	t		pIC50		6.800000191		IC50		149		#ERROR!			NULL
α_{2B}-adrenoceptor	26	ADRA2B	P18089	ENSG00000222040	Human	dexmedetomidine	521	135650174	t	Agonist	Full agonist		Not Determined	f	t		pKi	9.699999809		7.5	-				#ERROR!			9824686|9605427|9760042|9227000
α_{2B}-adrenoceptor	26	ADRA2B	P18089	ENSG00000222040	Human	guanfacine	522	135650327	t	Agonist	Full agonist		Not Determined	f	t		pKi		6.5		5.800000191	-			#ERROR!			9605427
α_{2B}-adrenoceptor	26	ADRA2B	P18089	ENSG00000222040	Human	brimonidine	520	135650222	t	Agonist	Partial agonist		Not Determined	f	t		pKi	8.300000191		6	-				#ERROR!			9824686|9605427|9760042
α_{2B}-adrenoceptor	26	ADRA2B	P18089	ENSG00000222040	Human	clonidine	516	135650112	t	Agonist	Partial agonist		Not Determined	f	t		pKi	9.5		6.699999809	-				#ERROR!			9824686|9605427|9760042
α_{2B}-adrenoceptor	26	ADRA2B	P18089	ENSG00000222040	Human	mirtazapine	7241	178103815	t	Antagonist	Antagonist		Not Determined	f	t		pIC50		6.699999809		IC50		223.8699951		#ERROR!			10636247
α_{2B}-adrenoceptor	26	ADRA2B	P18089	ENSG00000222040	Human	phenoxybenzamine	7268	178103842	t	Antagonist	Antagonist		Not Determined	f	t		pKi		8.5		Ki		3.5		#ERROR!			1834671
α_{2B}-adrenoceptor	26	ADRA2B	P18089	ENSG00000222040	Human	tolazoline	7310	178103883	t	Antagonist	Antagonist		Non-selective	f	t		pKi		5.5		Ki		3162		#ERROR!	Inhibition of agonist-stimulated [³⁵S]GT		9605427
α_{2B}-adrenoceptor	26	ADRA2B	P18089	ENSG00000222040	Human	yohimbine	102	135652727	t	Antagonist	Antagonist		Non-selective	f	t		pKi	8.899999619		7.900000095	-				#ERROR!			7996470|13532477|908642
α_{2C}-adrenoceptor	27	ADRA2C	P18825	ENSG00000184160	Human	apraclonidine	7117	178103694	t	Agonist	Agonist		Not Determined	f	t		pKi		7.519999981		Ki		30		#ERROR!			8784451
α_{2C}-adrenoceptor	27	ADRA2C	P18825	ENSG00000184160	Human	guanabenz	5443	178102094	t	Agonist	Agonist		Not Determined	f	t		pIC50		5.960000038		IC50		1106		#ERROR!			NULL
α_{2C}-adrenoceptor	27	ADRA2C	P18825	ENSG00000184160	Human	dexmedetomidine	521	135650174	t	Agonist	Full agonist		Not Determined	f	t		pKi	9.300000191		7	-				#ERROR!			9824686|9605427|9760042
α_{2C}-adrenoceptor	27	ADRA2C	P18825	ENSG00000184160	Human	brimonidine	520	135650022	t	Agonist	Partial agonist		Not Determined	f	t		pKi	7.599999905		5.699999809	-				#ERROR!			9824686|9605427|9760042|9227000
α_{2C}-adrenoceptor	27	ADRA2C	P18825	ENSG00000184160	Human	guanfacine	522	135650327	t	Agonist	Partial agonist		Not Determined	f	t		pKi	6.199999809		5.400000095	-				#ERROR!			9605427
α_{2C}-adrenoceptor	27	ADRA2C	P18825	ENSG00000184160	Human	mirtazapine	7241	178103815	t	Antagonist	Antagonist		Not Determined	f	t		pIC50		6.699999809		IC50		199.5299988		#ERROR!			10636247
α_{2C}-adrenoceptor	27	ADRA2C	P18825	ENSG00000184160	Human	mirtazapine	7241	178103815	t	Antagonist	Antagonist		Not Determined	f	t		pKi		7.739999771		Ki		18		#ERROR!	Inhibition of [³H]rauwolscine binding.		15771415
α_{2C}-adrenoceptor	27	ADRA2C	P18825	ENSG00000184160	Human	tolazoline	7310	178103883	t	Antagonist	Antagonist		Non-selective	f	t		pKi		5.429999828		Ki		3715		#ERROR!	Inhibition of agonist-stimulated [³⁵S]GT		9605427
α_{2C}-adrenoceptor	27	ADRA2C	P18825	ENSG00000184160	Human	yohimbine	102	135652727	t	Antagonist	Antagonist		Non-selective	f	t		pKi	9.5		8.5	-				#ERROR!			7996470|13532477|908642
Androgen receptor	628	AR	P10275	ENSG00000169083	Human	dromostanolone propi	6947	178103529	t	Agonist	Agonist		Selective	f	t		-				-				#ERROR!			NULL
Androgen receptor	628	AR	P10275	ENSG00000169083	Human	ethylestrenol	6948	178103530	t	Agonist	Agonist		Non-selective	f	t		-				-				#ERROR!			NULL
Androgen receptor	628	AR	P10275	ENSG00000169083	Human	fluoxymesterone	2861	135650271	t	Agonist	Agonist		Selective	f	t		pKi		8.239999771		Ki		5.699999809		#ERROR!			17439112
Androgen receptor	628	AR	P10275	ENSG00000169083	Human	nandrolone	6949	178103531	t	Agonist	Agonist		Non-selective	f	t		-				-				#ERROR!			NULL
Androgen receptor	628	AR	P10275	ENSG00000169083	Human	testosterone propiona	7100	178103678	t	Agonist	Agonist		Selective	f	t		pKi		9.600000381		Ki		0.25		#ERROR!			17574413
Androgen receptor	628	AR	P10275	ENSG00000169083	Human	apalutamide	9043	315661128	t	Antagonist	Antagonist		Not Determined	f	t		pIC50		6.909999847		IC50		122		#ERROR!			20218717
Androgen receptor	628	AR	P10275	ENSG00000169083	Human	apalutamide	9043	315661128	t	Antagonist	Antagonist		Not Determined	f	t		pIC50		6.699999809		IC50		200		#ERROR!	Antagonism of testosterone-induced activation of wi		26137992
Androgen receptor	628	AR	P10275	ENSG00000169083	Human	bicalutamide	2863	135649993	t	Antagonist	Antagonist		Selective	f	t		pKi		7.71999979		Ki		19		#ERROR!			16420057
Androgen receptor	628	AR	P10275	ENSG00000169083	Human	bicalutamide	2863	135649993	t	Antagonist	Antagonist		Not Determined	f	t		pIC50				6		IC50		1000			20218717
Androgen receptor	628	AR	P10275	ENSG00000169083	Human	clascoterone	11215	434122275	t	Antagonist	Antagonist		Selective	f	t		pKi		7.400000095		Ki		40		#ERROR!	Determined by measuring displacement of [3H]meth		NULL
Androgen receptor	628	AR	P10275	ENSG00000169083	Human	cyproterone acetate	2865	135650153	t	Antagonist	Antagonist		Non-selective	f	t		pKi		7.849999905		Ki		14		#ERROR!			9484511
Androgen receptor	628	AR	P10275	ENSG00000169083	Human	darolutamide	10439	385612199	t	Antagonist	Antagonist		Not Determined	f	t		pIC50		7.590000153		IC50		26		#ERROR!	Antagonism of testosterone-induced activation of wi		26137992
Androgen receptor	628	AR	P10275	ENSG00000169083	Human	enzalutamide	6812	178103418	t	Antagonist	Antagonist		Selective	f	t		pIC50		7.440000057		IC50		36		#ERROR!			19359544
Androgen receptor	628	AR	P10275	ENSG00000169083	Human	enzalutamide	6812	178103418	t	Antagonist	Antagonist		Not Determined	f	t		pIC50		6.659999847		IC50		219		#ERROR!	Antagonism of testosterone-induced activation of wi		26137992
Androgen receptor	628	AR	P10275	ENSG00000169083	Human	flutamide	6943	178103525	t	Antagonist	Antagonist	Displacement of	Selective	f	t		pKi		5.429999828		Ki		3700		#ERROR!			18571420
Androgen receptor	628	AR	P10275	ENSG00000169083	Human	nilutamide	2864	135650719	t	Antagonist	Antagonist		Selective	f	t		pIC50	7.119999886		7.070000172	IC50	85		75	#ERROR!			9111629
Angiotensin-converting enzyme	1613	ACE	P12821	ENSG00000159640	Human	benazepril	6374	178102992	t	Inhibitor	Inhibition		Not Determined	f	t		pIC50		8.800000191		IC50		1.700000048		#ERROR!			2999394
Angiotensin-converting enzyme	1613	ACE	P12821	ENSG00000159640	Human	benazeprilat	6375	178102993	t	Inhibitor	Inhibition		Selective	f	t		pIC50		6.599999905		IC50		280		#ERROR!			17506720
Angiotensin-converting enzyme	1613	ACE	P12821	ENSG00000159640	Human	captopril	5158	178101850	t	Inhibitor	Inhibition		Not Determined	f	t		pKi		8.399999619		Ki		3.980000019		#ERROR!			9187274
Angiotensin-converting enzyme	1613	ACE	P12821	ENSG00000159640	Human	cilazapril	6459	178103074	t	Inhibitor	Inhibition		Not Determined	f	t		-				-				#ERROR!			NULL
Angiotensin-converting enzyme	1613	ACE	P12821	ENSG00000159640	Human	enalapril	6322	178102940	t	Inhibitor	Inhibition		Not Determined	f	t		pIC50		4.300000191		IC50		50000		#ERROR!			7527095
Angiotensin-converting enzyme	1613	ACE	P12821	ENSG00000159640	Human	enalaprilat	6332	178102950	t	Inhibitor	Inhibition		Selective	f	t		pIC50		7.5		IC50		34		#ERROR!			7527095
Angiotensin-converting enzyme	1613	ACE	P12821	ENSG00000159640	Human	fosinopril	6456	178103071	t	Inhibitor	Inhibition		Not Determined	f	t		-				-				#ERROR!			NULL
Angiotensin-converting enzyme	1613	ACE	P12821	ENSG00000159640	Human	imidapril	6377	178102995	t	Inhibitor	Inhibition		Not Determined	f	t		-				-				#ERROR!			NULL
Angiotensin-converting enzyme	1613	ACE	P12821	ENSG00000159640	Human	imidaprilat	6378	178102996	t	Inhibitor	Inhibition		Selective	f	t		pIC50		8.699999809		IC50		2		#ERROR!			17547476
Angiotensin-converting enzyme	1613	ACE	P12821	ENSG00000159640	Human	lisinopril	6360	178102978	t	Inhibitor	Inhibition		Not Determined	f	t		pKi		9.399999619		Ki		0.360000014		#ERROR!			9187274
Angiotensin-converting enzyme	1613	ACE	P12821	ENSG00000159640	Human	moexipril	6571	178103185	t	Inhibitor	Inhibition		Not Determined	f	t		-				-				#ERROR!			NULL
Angiotensin-converting enzyme	1613	ACE	P12821	ENSG00000159640	Human	perindopril	6367	178102985	t	Inhibitor	Inhibition		Not Determined	f	t		-				-				#ERROR!			NULL
Angiotensin-converting enzyme	1613	ACE	P12821	ENSG00000159640	Human	perindoprilat	6373	178102991	t	Inhibitor	Inhibition		Selective	f	t		pIC50				9		IC50		1.049999952		#ERROR!			17716647
Angiotensin-converting enzyme	1613	ACE	P12821	ENSG00000159640	Human	quinaprilat	6352	178102970	t	Inhibitor	Inhibition		Not Determined	f	t		pIC50		8.5		IC50		2.900000095		#ERROR!			1995872
Angiotensin-converting enzyme	1613	ACE	P12821	ENSG00000159640	Human	ramipril	6339	178102957	t	Inhibitor	Inhibition		Not Determined	f	t		pIC50		7.599999905		IC50		26		#ERROR!			6097265
Angiotensin-converting enzyme	1613	ACE	P12821	ENSG00000159640	Human	ramiprilat	6344	178102962	t	Inhibitor	Inhibition		Not Determined	f	t		pIC50		8.399999619		IC50		4.199999809		#ERROR!			6097265
Angiotensin-converting enzyme	1613	ACE	P12821	ENSG00000159640	Human	spirapril	6575	178103189	t	Inhibitor	Inhibition		Not Determined	f	t		-				-				#ERROR!			NULL
Angiotensin-converting enzyme	1613	ACE	P12821	ENSG00000159640	Human	spiraprilat	6576	178103190	t	Inhibitor	Inhibition		Not Determined	f	t		pIC50		9.100000381		IC50		0.800000012		#ERROR!			2544729
Angiotensin-converting enzyme	1613	ACE	P12821	ENSG00000159640	Human	trandolapril	6453	178103069	t	Inhibitor	Inhibition		Not Determined	f	t		pIC50		7.800000191		IC50		15		#ERROR!			7527095
Angiotensin-converting enzyme	1613	ACE	P12821	ENSG00000159640	Human	trandolaprilat	6455	178103070	t	Inhibitor	Inhibition		Not Determined	f	t		pIC50		8.5		IC50		3.200000048		#ERROR!			7527095
Angiotensin-converting enzyme	1613	ACE	P12821	ENSG00000159640	Human	zofenopril	6462	178103076	t	Inhibitor	Inhibition		Not Determined	f	t		-				-				#ERROR!			NULL
A₁ receptor	18	ADORA1	P30542	ENSG00000163485	Human	caffeine	407	135652675	t	Antagonist	Antagonist		Non-selective	f	t		pKi	4.96999979		4.349999905	Ki	44900		10700	#ERROR!			8469419|14563788
A_{2A} receptor	19	ADORA2A	P29274	ENSG00000128271	Human	regadenoson	5996	176102231	t	Agonist	Agonist		Selective	f	t		pKi		6.539999962		Ki		290		#ERROR!			16818376
A_{2A} receptor	19	ADORA2A	P29274	ENSG00000128271	Human	caffeine	407	135652675	t	Antagonist	Antagonist		Non-selective	f	t		pKi	5.599999905		4.599999905	-				#ERROR!			9920286|14563788
A_{2A} receptor	19	ADORA2A	P29274	ENSG00000128271	Human	theophylline	413	135651254	t	Antagonist	Antagonist		Not Determined	f	t		pKi	5.800000191		5.170000076	-				#ERROR!			9459566|16219300|9179373
A_{2B} receptor	20	ADORA2B	P29275	ENSG00000170425	Human	caffeine	407	135652675	t	Antagonist	Antagonist		Non-selective	f	t		pKi	4.980000019		4.46999979	Ki	33800		10400	#ERROR!			19569717|12014951|18404493
A_{2B} receptor	20	ADORA2B	P29275	ENSG00000170425	Human	theophylline	413	135651254	t	Antagonist	Antagonist		Not Determined	f	t		pKi	5.039999962		4.130000114	Ki	74000		9070	#ERROR!			16219300|11705449|12014951|18404493
A₃ receptor	21	ADORA3	P0DMS8	ENSG00000121933	Human	caffeine	407	135652675	t	Antagonist	Antagonist		Non-selective	f	t		pKi		4.880000114		Ki		13300		#ERROR!			NULL
ATP4A	849	ATP4A	P20648	ENSG00000105675	Human	esomeprazole	5488	178102135	t	Inhibitor	Inhibition		Selective	f	t		-				-				#ERROR!			NULL
ATP4A	849	ATP4A	P20648	ENSG00000105675	Human	omeprazole	4279	178101096	t	Inhibitor	Inhibition		Selective	f	t		-				-				#ERROR!			NULL
ATP-binding cassette, sub-family C (CFTR	2594	ABCC8	Q09428	ENSG00000006071	Human	repaglinide	6841	178103447	t	Inhibitor	Inhibition		Selective	f	t		pIC50		6.96999979		IC50		106		#ERROR!			15380228
AT₁ receptor	34	AGTR1	P30556	ENSG00000144891	Human	candesartan	587	135650050	t	Antagonist	Antagonist		Selective	f	t		pIC50	9.720000267		9.489999771	IC50	0.319999993		0.189999998	#ERROR!			10193788
AT₁ receptor	34	AGTR1	P30556	ENSG00000144891	Human	eprosartan	588	135650236	t	Antagonist	Antagonist		Selective	f	t		pIC50	8.800000191		8.399999619	IC50	3.980000019		1.580000043	#ERROR!			1309870
AT₁ receptor	34	AGTR1	P30556	ENSG00000144891	Human	irbesartan	589	135650414	t	Antagonist	Antagonist		Selective	f	t		pIC50	8.800000191		8.699999809	IC50	1.99000001		1.580000043	#ERROR!			10193788
AT₁ receptor	34	AGTR1	P30556	ENSG00000144891	Human	losartan	590	135650532	t	Antagonist	Antagonist		Selective	f	t		pIC50	8.699999809		7.400000095	IC50	39.79999924		1.99000001	#ERROR!			8372104
AT₁ receptor	34	AGTR1	P30556	ENSG00000144891	Human	olmesartan	591	135650757	t	Antagonist	Antagonist		Selective	f	t		pIC50		8.100000381		-				#ERROR!			11451212
AT₁ receptor	34	AGTR1	P30556	ENSG00000144891	Human	telmisartan	592	135651242	t	Antagonist	Antagonist		Selective	f	t		pIC50		8.399999619		-				#ERROR!			9878991
AT₁ receptor	34	AGTR1	P30556	ENSG00000144891	Human	valsartan	3937	178100765	t	Antagonist	Antagonist		Selective	f	t		pIC50		8.609999657		IC50		2.430000067		#ERROR!			8577935
BCL2 apoptosis regulator	2844	BCL2	P10415	ENSG00000171791	Human	venetoclax	8318	252166530	t	Antagonist	Antagonist		Selective	f	t		pKi		11		Ki		0.01	<	Note that this Ki is below the detection limit of the as		23291630
β₁-adrenoceptor	28	ADRB1	P08588	ENSG00000043591	Human	isoprenaline	536	135650431	t	Agonist	Full agonist		Non-selective	f	t		pKi	7		6.599999905	-				#ERROR!			8982677|2849109
β₁-adrenoceptor	28	ADRB1	P08588	ENSG00000043591	Human	dobutamine	535	135650199	t	Agonist	Partial agonist		Not Determined	f	t		pKi		5.5		-				#ERROR!			10531390
β₁-adrenoceptor	28	ADRB1	P08588	ENSG00000043591	Human	pindolol	91	135650827	t	Agonist	Partial agonist		Non-selective	f	t		pKi	9.300000191			-			0.519999981	#ERROR!			17804228
β₁-adrenoceptor	28	ADRB1	P08588	ENSG00000043591	Human	acebutolol	7107	178103684	t	Antagonist	Antagonist		Selective	f	t		pKi	6.400000095			Ki			422	#ERROR!			NULL
β₁-adrenoceptor	28	ADRB1	P08588	ENSG00000043591	Human	atenolol	548	135649971	t	Antagonist	Antagonist		Selective	f	t		pKi	7.599999905		6.699999809	-				#ERROR!			15060759|10079020|15655528
β₁-adrenoceptor	28	ADRB1	P08588	ENSG00000043591	Human	betaxolol	549	135649987	t	Antagonist	Antagonist		Selective	f	t		pKi	8.800000191			-				#ERROR!			10079020
β₁-adrenoceptor	28	ADRB1	P08588	ENSG00000043591	Human	bupranolol	550	135650033	t	Antagonist	Antagonist		Non-selective	f	t		pKi	9		7.300000191	-				#ERROR!			10079020|10411574
β₁-adrenoceptor	28	ADRB1	P08588	ENSG00000043591	Human	carvedilol	551	135650057	t	Antagonist	Antagonist		Non-selective	f	t		pKi		9.5		-				#ERROR!			10411574
β₁-adrenoceptor	28	ADRB1	P08588	ENSG00000043591	Human	esmolol	7178	178103753	t	Antagonist	Antagonist		Not Determined	f	t		pKi		6.900000095		Ki		113.6999969		#ERROR!			NULL
β₁-adrenoceptor	28	ADRB1	P08588	ENSG00000043591	Human	labetalol	7207	178103782	t	Antagonist	Antagonist		Non-selective	f	t		pKi		8.199999809		Ki		5.840000153		#ERROR!			NULL
β₁-adrenoceptor	28	ADRB1	P08588	ENSG00000043591	Human	levobetaxolol	8035	249565715	t	Antagonist	Antagonist		Selective	f	t		pKi		9.100000381		Ki		0.75999999		#ERROR!			11572462
β₁-adrenoceptor	28	ADRB1	P08588	ENSG00000043591	Human	levobunolol	570	135650031	t	Antagonist	Antagonist		Non-selective	f	t		pKi		8.399999619		Ki		3.99000001		#ERROR!			NULL
β₁-adrenoceptor	28	ADRB1	P08588	ENSG00000043591	Human	metoprolol	553	135650599	t	Antagonist	Antagonist		Not Determined	f	t		pKi	7.599999905			7	-			#ERROR!			10079020|15655528|10411574|14730417
β₁-adrenoceptor	28	ADRB1	P08588	ENSG00000043591	Human	nadolol	554	135650678	t	Antagonist	Antagonist		Non-selective	f	t		pKi	6.900000095			-				#ERROR!			10411574
β₁-adrenoceptor	28	ADRB1	P08588	ENSG00000043591	Human	practolol	555	135650846	t	Antagonist	Antagonist		Not Determined	f	t		pKi	6.800000191		6.099999905	-				#ERROR!			10079020|15655528
β₁-adrenoceptor	28	ADRB1	P08588	ENSG00000043591	Human	propafenone	2561	135650856	t	Antagonist	Antagonist		Non-selective	f	t		pKi		6.699999809		Ki		205		#ERROR!			NULL
β₁-adrenoceptor	28	ADRB1	P08588	ENSG00000043591	Human	sotalol	7297	178103870	t	Antagonist	Antagonist		Non-selective	f	t		pKi		6.099999905		Ki		833		#ERROR!			NULL
β₂-adrenoceptor	29	ADRB2	P07550	ENSG00000169252	Human	arformoterol	7479	187051782	t	Agonist	Agonist		Not Determined	f	t		pKi		8.599999153		Ki		2.599999905		#ERROR!			15324892
β₂-adrenoceptor	29	ADRB2	P07550	ENSG00000169252	Human	fenoterol	557	135650257	t	Agonist	Agonist		Selective	f	t		pKi		6.900000095		Ki		126		#ERROR!			24454993
β₂-adrenoceptor	29	ADRB2	P07550	ENSG00000169252	Human	formoterol	3465	178100459	t	Agonist	Agonist		Selective	f	t		pEC50		10.07999992		-				#ERROR!			20590599
β₂-adrenoceptor	29	ADRB2	P07550	ENSG00000169252	Human	indacaterol	7455	187051758	t	Agonist	Agonist		Selective	f	t		pKi		7.800000191		Ki		15.89999962		#ERROR!			22932315
β₂-adrenoceptor	29	ADRB2	P07550	ENSG00000169252	Human	olodaterol	7543	187051844	t	Agonist	Agonist		Selective	f	t		pIC50				10		IC50		0.100000001		#ERROR!			20371707
β₂-adrenoceptor	29	ADRB2	P07550	ENSG00000169252	Human	orciprenaline	7250	178103824	t	Agonist	Agonist		Selective	f	t		pKd		5.320000172		Ki		4810		#ERROR!			19168263
β₂-adrenoceptor	29	ADRB2	P07550	ENSG00000169252	Human	procaterol	3464	53801069	t	Agonist	Agonist		Selective	f	t		pEC50		8.430000305		-				#ERROR!			20590599
β₂-adrenoceptor	29	ADRB2	P07550	ENSG00000169252	Human	vilanterol	7353	178103925	t	Agonist	Agonist		Selective	f	t		pEC50		9.399999619		EC50		0.398000002		#ERROR!			20462258
β₂-adrenoceptor	29	ADRB2	P07550	ENSG00000169252	Human	isoprenaline	536	135650431	t	Agonist	Full agonist		Non-selective	f	t		pKi		6.400000095		-				#ERROR!			8982677
β₂-adrenoceptor	29	ADRB2	P07550	ENSG00000169252	Human	salmeterol	559	135650978	t	Agonist	Full agonist		Selective	f	t		pEC50		9.920000076		EC50		0.119999997		#ERROR!			20590599
β₂-adrenoceptor	29	ADRB2	P07550	ENSG00000169252	Human	ephedrine	556	135650234	t	Agonist	Partial agonist		Not Determined	f	t		pKi		5.599999905		-				#ERROR!			9295336
β₂-adrenoceptor	29	ADRB2	P07550	ENSG00000169252	Human	pindolol	91	135650827	t	Agonist	Partial agonist		Non-selective	f	t		pKi		9.399999619		Ki		0.400000006		#ERROR!			17804228
β₂-adrenoceptor	29	ADRB2	P07550	ENSG00000169252	Human	salbutamol	558	135650977	t	Agonist	Partial agonist		Selective	f	t		pKi	6.099999905		5.800000191	Ki	1580		794	#ERROR!			10531390|15655528
β₂-adrenoceptor	29	ADRB2	P07550	ENSG00000169252	Human	terbutaline	560	135651244	t	Agonist	Partial agonist		Selective	f	t		pKi		5.599999905		Ki		2510		#ERROR!			15655528
β₂-adrenoceptor	29	ADRB2	P07550	ENSG00000169252	Human	alprenolol	563	135649926	t	Antagonist	Antagonist		Not Determined	f	t		pKi		9		-				#ERROR!			15655528
β₂-adrenoceptor	29	ADRB2	P07550	ENSG00000169252	Human	bupranolol	550	135650033	t	Antagonist	Antagonist		Non-selective	f	t		pKi	9.100000381		8.300000191	-				#ERROR!			10079020|10411574
β₂-adrenoceptor	29	ADRB2	P07550	ENSG00000169252	Human	carvedilol	551	135650057	t	Antagonist	Antagonist		Non-selective	f	t		pKi	9.899999619		9.399999619	-				#ERROR!			15655528|10411574

target	target_id	target_gene	target_uniprot	target_ensembl_gene_id	target_species	ligand	ligand_id	ligand_pubchem	approved_drug	type	action	action_comment	selectivity	endogenous	primary_target	concentration_r	affinity_units	affinity_high	affinity_median	affinity_low	original_affinity	original_affinity	original_affinity	original_affinity	original_affinity	assay_descripti	receptor_site	ligand_context	pubmed_id
β₂-adrenoceptor	29	ADRB2	P07550	ENSG00000169252	Human	levobunolol	570	135650031	t	Antagonist	Antagonist		Non-selective	f	t		pKi		9.260000229		Ki		0.550000012		#ERROR!				NULL
β₂-adrenoceptor	29	ADRB2	P07550	ENSG00000169252	Human	nadolol	554	135650678	t	Antagonist	Antagonist		Non-selective	f	t		pKi	8.600000381		7	-			#ERROR!				15655528|10411574
β₂-adrenoceptor	29	ADRB2	P07550	ENSG00000169252	Human	propranolol	564	135650658	t	Antagonist	Antagonist		Non-selective	f	t		pKi	9.5		9.100000381	-			#ERROR!				10079020|10531390|15655528|12920204
β₂-adrenoceptor	29	ADRB2	P07550	ENSG00000169252	Human	sotalol	7297	178103870	t	Antagonist	Antagonist		Non-selective	f	t		pKi		6.53000021		Ki		298		#ERROR!				NULL
β₂-adrenoceptor	29	ADRB2	P07550	ENSG00000169252	Human	timolol	565	135651263	t	Antagonist	Antagonist		Not Determined	f	t		pKi		9.699999809		-				#ERROR!				15655528
β₃-adrenoceptor	30	ADRB3	P13945	ENSG00000188778	Human	mirabegron	7445	178104017	t	Agonist	Agonist		Selective	f	t		pEC50		7.650000095		EC50		22.39999962		#ERROR!				17293563
β₃-adrenoceptor	30	ADRB3	P13945	ENSG00000188778	Human	bupranolol	550	135650033	t	Antagonist	Antagonist		Non-selective	f	t		pKi	7.300000191		6.800000191	-			#ERROR!				10079020|10411574|8719421|7903415
β₃-adrenoceptor	30	ADRB3	P13945	ENSG00000188778	Human	levobunolol	570	135650031	t	Antagonist	Antagonist		Non-selective	f	t		pKi		6.800000191		Ki		160		#ERROR!				14744619
β₃-adrenoceptor	30	ADRB3	P13945	ENSG00000188778	Human	propranolol	564	135650658	t	Antagonist	Antagonist		Non-selective	f	t		pKi	7.199999809		6.300000191	-			#ERROR!				10079020|14744619
B-Raf proto-oncogene, serine/threonine ki	1943	BRAF	P15056	ENSG00000157764	Human	dabrafenib	6494	178103108	t	Inhibitor	Inhibition		Not Determined	f	t		pIC50		8.489999771		IC50		3.200000048		#ERROR!				NULL
B-Raf proto-oncogene, serine/threonine ki	1943	BRAF	P15056	ENSG00000157764	Human	regorafenib	5891	178102514	t	Inhibitor	Inhibition		Not Determined	f	t		pIC50		7.550000191		IC50		28		#ERROR!				22222036
B-Raf proto-oncogene, serine/threonine ki	1943	BRAF	P15056	ENSG00000157764	Human	vemurafenib	5893	178102516	t	Inhibitor	Inhibition		Not Determined	f	t		pIC50			7	IC50		100		#ERROR!				22808911
Bruton tyrosine kinase	1948	BTK	Q06187	ENSG00000010671	Human	acalabrutinib	8912	310264693	t	Inhibitor	Inhibition		Selective	f	t		pIC50		8.289999962		IC50		5.099999905		#ERROR!				26641137
Bruton tyrosine kinase	1948	BTK	Q06187	ENSG00000010671	Human	acalabrutinib	8912	310264693	t	Inhibitor	Inhibition		Selective	f	t		pEC50		8		EC50		10	<				NULL
Bruton tyrosine kinase	1948	BTK	Q06187	ENSG00000010671	Human	ibrutinib	6912	178103494	t	Inhibitor	Inhibition		Not Determined	f	t		pIC50		9.300000191		IC50		0.5		#ERROR!				17154430
Bruton tyrosine kinase	1948	BTK	Q06187	ENSG00000010671	Human	ibrutinib	6912	178103494	t	Inhibitor	Inhibition		Not Determined	f	t		pIC50		8.819999695		IC50		1.5		#ERROR!				26641137
B₂ receptor	42	BDKRB2	P30411	ENSG00000168398	Human	icatibant	667	135652196	t	Antagonist	Antagonist		Selective	f	t		pKi		10.18999958		Ki		0.064000003		#ERROR!				10514288
B₂ receptor	42	BDKRB2	P30411	ENSG00000168398	Human	icatibant	667	135652196	t	Antagonist	Antagonist		Not Determined	f	t		pA2		8.399999619		-				#ERROR!				8735629
B₂ receptor	42	BDKRB2	P30411	ENSG00000168398	Human	icatibant	667	135652196	t	Antagonist	Antagonist		Not Determined	f	t		pIC50	9.399999619		8	-			#ERROR!				8302267|9203620
butyrylcholinesterase	2471	BCHE	P06276	ENSG00000114200	Human	rivastigmine	6602	178103215	t	Inhibitor	Inhibition		Not Determined	f	t		pIC50		7.400000095		IC50		37		#ERROR!				16570913
butyrylcholinesterase	2471	BCHE	P06276	ENSG00000114200	Human	tacrine	6687	178103299	t	Inhibitor	Inhibition		Not Determined	f	t		pKi		7.199999809		Ki		70		#ERROR!				18479118
CaCC	708	ANO1	Q5XXA6	ENSG00000131620	Human	crofelemer	7453	187051756	t	Inhibitor	Inhibition		Selective	f	t		pIC50		5.190000057		IC50		6500		#ERROR!				19808995
carbonic anhydrase 14	2598	CA14	Q8ULX7	ENSG00000168298	Human	ethoxzolamide	6814	178103420	t	Inhibitor	Inhibition		Not Determined	f	t		pKi		7.599999905		Ki		25		#ERROR!				16039848
carbonic anhydrase 7	2749	CA7	P43166	ENSG00000168748	Human	acetazolamide	6792	178103398	t	Inhibitor	Inhibition		Not Determined	f	t		pKi		8.600000381		Ki		2.5		#ERROR!				20605094
carbonic anhydrase 7	2749	CA7	P43166	ENSG00000168748	Human	brinzolamide	6797	178103403	t	Inhibitor	Inhibition		Not Determined	f	t		pKi		8.550000191		Ki		2.799999952		#ERROR!				23965175
carbonic anhydrase 7	2749	CA7	P43166	ENSG00000168748	Human	chlorthalidone	7147	178103722	t	Inhibitor	Inhibition		Not Determined	f	t		pKi		8.550000191		Ki		2.799999952		#ERROR!				19119014
carbonic anhydrase 7	2749	CA7	P43166	ENSG00000168748	Human	methazolamide	6828	178103434	t	Inhibitor	Inhibition		Not Determined	f	t		pKi		8.680000305		Ki		2.099999905		#ERROR!				23965175
CaS receptor	54	CASR	P41180	ENSG00000036828	Human	cinacalcet	3308	178100343	t	Allosteric modula	Positive		Not Determined	f	t		pEC50		7.300000191		EC50		51		#ERROR!				14593085
CaS receptor	54	CASR	P41180	ENSG00000036828	Human	cinacalcet	3308	178100343	t	Allosteric modula	Positive		Not Determined	f	t		pKB	6.599999905		5.900000095	-			#ERROR!	Ca²⁺_i release			22211744|25220431|27002221|23372019
Ca_v1.1	528	CACNA1S	Q13698	ENSG00000081248	Human	diltiazem	2298	135650185	t	Channel blocker	Antagonist		Non-selective	f	t		-				-				#ERROR!				NULL
Ca_v1.1	528	CACNA1S	Q13698	ENSG00000081248	Human	verapamil	2406	135651305	t	Channel blocker	Antagonist		Non-selective	f	t		-				-				#ERROR!				NULL
Ca_v1.2	529	CACNA1C	Q13936	ENSG00000151067	Human	verapamil	2406	135651305	t	Channel blocker	Antagonist		Non-selective	f	t		pIC50			6.5	-				#ERROR!				8913371
Ca_v1.3	530	CACNA1D	Q01668	ENSG00000157388	Human	nitrendipine	2334	135650723	t	Gating inhibitor	Inhibition		Not Determined	f	t		pIC50		8.399999619		IC50		3.589999914		#ERROR!	Recombinant Cav1.3 calcium channel complexes ex			19029287
Catechol-O-methyltransferase	2472	COMT	P21964	ENSG00000093010	Human	entacapone	6647	178103260	t	Inhibitor	Inhibition	soluble enzyme	Not Determined	f	t		pKi		9.520000458		Ki		0.300000012		#ERROR!	soluble enzyme			7703232
Catechol-O-methyltransferase	2472	COMT	P21964	ENSG00000093010	Human	entacapone	6647	178103260	t	Inhibitor	Inhibition	membrane-boun	Not Determined	f	t		pKi		8.699999809		Ki		2		#ERROR!	membrane-bound enzyme			7703232
Catechol-O-methyltransferase	2472	COMT	P21964	ENSG00000093010	Human	opicapone	8988	310264766	t	Inhibitor	Inhibition		Selective	f	t		pKi		9		Ki		1		#ERROR!				24847974
Catechol-O-methyltransferase	2472	COMT	P21964	ENSG00000093010	Human	tolcapone	6646	178103259	t	Inhibitor	Inhibition	soluble enzyme	Not Determined	f	t		pKi		9.569999695		Ki		0.270000011		#ERROR!	soluble enzyme			7703232
Catechol-O-methyltransferase	2472	COMT	P21964	ENSG00000093010	Human	tolcapone	6646	178103259	t	Inhibitor	Inhibition	membrane-boun	Not Determined	f	t		pKi		9.539999962		Ki		0.289999962		#ERROR!	membrane-bound enzyme			7703232
CB₁ receptor	56	CNR1	P21554	ENSG00000118432	Human	Δ⁹	2424	135651250	t	Agonist	Partial agonist		Not Determined	f	t		pKi	7.400000095		7.300000191	-			#ERROR!				7565624|8819477
CCR4	61	CCR4	P51679	ENSG00000183813	Human	mogamulizumab	6477	178103091	t	Antibody	Inhibition		Selective	f	t		-				-				#ERROR!				21154168
CCR5	62	CCR5	P51681	ENSG00000160791	Human	maraviroc	806	135650567	t	Antagonist	Antagonist		Selective	f	t		pIC50		8.100000381		IC50		7.940000057		#ERROR!				16298345
CD19	2764	CD19	P15391	ENSG00000177455	Human	inebilizumab	7992	249565673	t	Antibody	Binding		Selective	f	t		pKd	10.30000019		9.300000191	Kd	0.5		0.050000001	#ERROR!				8913371
CD19	2764	CD19	P15391	ENSG00000177455	Human	tafasitamab	11131	134122191	t	Antibody	Binding		Selective	f	t		-				-				#ERROR!				18829563
CD2	2600	CD2	P06729	ENSG00000116824	Human	alefacept	6787	178103393	t	Inhibitor	Inhibition		Selective	f	t		-				-				#ERROR!				12089534|11970990
CD20 (membrane-spanning 4-domains, s	2628	MS4A1	P11836	ENSG00000156738	Human	ibritumomab tiuxetan	6777	178103383	t	Antibody	Binding		Not Determined	f	t		-				-				#ERROR!				NULL
CD20 (membrane-spanning 4-domains, s	2628	MS4A1	P11836	ENSG00000156738	Human	obinutuzumab	6941	178103523	t	Antibody	Binding		Not Determined	f	t		-				-				#ERROR!				23537278|21378274
CD20 (membrane-spanning 4-domains, s	2628	MS4A1	P11836	ENSG00000156738	Human	ocrelizumab	7580	223365915	t	Antibody	Binding		Selective	f	t		pKd		9.079999924		Kd		0.839999974		#ERROR!				NULL
CD20 (membrane-spanning 4-domains, s	2628	MS4A1	P11836	ENSG00000156738	Human	ofatumumab	6778	178103384	t	Antibody	Binding		Not Determined	f	t		pKd		9.920000076		Kd		0.120999999		#ERROR!	Binding affinity for clone 2F2 ((HuMax-CD20).			NULL
CD20 (membrane-spanning 4-domains, s	2628	MS4A1	P11836	ENSG00000156738	Human	rituximab	6780	178103386	t	Antibody	Binding		Not Determined	f	t		pKd		8.510000229		Kd		3.099999905		#ERROR!				15102696
CD20 (membrane-spanning 4-domains, s	2628	MS4A1	P11836	ENSG00000156738	Human	tositumomab	6781	178103387	t	Antibody	Binding		Not Determined	f	t		-				-				#ERROR!				NULL
CD22	2786	CD22	P20273	ENSG00000012124	Human	inotuzumab ozogamic	8266	252166478	t	Antibody	Binding		Not Determined	f	t		pKd		9.699999809		Kd		0.200000003		#ERROR!				14615373
CD22	2786	CD22	P20273	ENSG00000012124	Human	moxetumomab pasud	7675	223366010	t	Antibody	Binding		Not Determined	f	t		pKd		8.220000267		Kd		6		#ERROR!				22350553|22003067
CD30	1877	TNFRSF8	P28908	ENSG00000120949	Human	brentuximab vedotin	6772	178103378	t	Antibody	Binding		Selective	f	t		-				-				#ERROR!				NULL
CD33	2601	CD33	P20138	ENSG00000105383	Human	gemtuzumab ozogam	6775	178103381	t	Antibody	Binding		Selective	f	t		-				-				#ERROR!				10720144
CD38	2766	CD38	P28907	ENSG00000004468	Human	daratumumab	7395	178103967	t	Antibody	Antagonist		Selective	f	t		pKd		8.220000267		Kd		6.059999943		#ERROR!				NULL
CD38	2766	CD38	P28907	ENSG00000004468	Human	isatuximab	8719	252827377	t	Antibody	Binding		Selective	f	t		pKd	10		8.520000458	Kd	3		0.100000001	#ERROR!				NULL
CD3e	2742	CD3E	P07766	ENSG00000198851	Human	catumaxomab	7385	178103957	t	Antibody	Binding		Not Determined	f	t		-				-				#ERROR!				20190561
CD3e	2742	CD3E	P07766	ENSG00000198851	Human	muromonab-CD3	6889	178103471	t	Antibody	Binding		Selective	f	t		-				-				#ERROR!				3105134
CD4	2995	CD4	P01730	ENSG00000010610	Human	ibalizumab	10208	381744890	t	Antibody	Binding		Selective	f	t		-				-				#ERROR!				1380539
CD52	2602	CD52	P31358	ENSG00000169442	Human	alemtuzumab	6770	178103376	t	Antibody	Binding		Not Determined	f	t		-				-				#ERROR!				9593475
CD6	2917	CD6	P30203	ENSG00000013725	Human	itolizumab	9421	328083521	t	Antibody	Binding		Selective	f	t		pKi		7.219999779		Ki		60		#ERROR!				18707547
CD79B	2852	CD79B	P40259	ENSG00000007312	Human	polatuzumab vedotin	8404	252166614	t	Antibody	Binding		Selective	f	t		-				-				#ERROR!				NULL
CD80	2744	CD80	P33681	ENSG00000121594	Human	abatacept	6891	178103473	t	Inhibitor	Inhibition		Not Determined	f	t		pKd		7.920000076		Kd		12	~		Measuring binding of ¹²⁵I-labeled B7(	23899231|1714933		
CD80	2744	CD80	P33681	ENSG00000121594	Human	belatacept	6892	178103474	t	Inhibitor	Inhibition		Not Determined	f	t		-				-				#ERROR!	IC₅₀ = 0.009µg/mL	19300198		
CD86	2745	CD86	P42081	ENSG00000114013	Human	abatacept	6891	178103473	t	Inhibitor	Inhibition		Selective	f	t		pKd		7.920000076		Kd		12	~		Measuring binding of ¹²⁵I-labeled B7(	23899231|1714933		
CD86	2745	CD86	P42081	ENSG00000114013	Human	belatacept	6892	178103474	t	Inhibitor	Inhibition		Selective	f	t		-				-				#ERROR!	IC₅₀ = 0.120µg/mL	19300198		
CFTR	707	CFTR	P13569	ENSG00000001626	Human	crofelemer	7453	187051756	t	Inhibitor	Inhibition		Selective	f	t		pIC50		5.150000095		IC50		7000		#ERROR!				19808995
CGRP receptor	48	RAMP1|CALCRL	Q16602|O60894	ENSG00000064989|ENSG	Human	rimegepant	10704	134859151	t	Antagonist	Antagonist		Not Determined	f	t		pKi		10.56999969		Ki		0.027000001		#ERROR!	Dissociation constant determined in a radioligand di			23153230
CGRP receptor	48	RAMP1|CALCRL	Q16602|O60894	ENSG00000064989|ENSG	Human	erenumab	9250	315661325	t	Antibody	Antagonist		Selective	f	t		pKi		10.69999981		Ki		0.02		#ERROR!	Displacement binding of [(125)I]-CGRP at the huma			26559125
CIC-2	699	CLCN2	P51788	ENSG00000114859	Human	lubiprostone	4242	178101059	t	Activator	None		Non-selective	f	t		-				-				#ERROR!				NULL
coagulation factor II, thrombin	2362	F2	P00734	ENSG00000180210	Human	argatroban	6385	178103001	t	Inhibitor	Inhibition		Not Determined	f	t		pKi		7.699999809		Ki		19		#ERROR!				9361377
coagulation factor II, thrombin	2362	F2	P00734	ENSG00000180210	Human	bivalirudin	6470	178103084	t	Inhibitor	Inhibition		Not Determined	f	t		pKi		8.600000381		Ki		2.559999943		#ERROR!				1290488
coagulation factor II, thrombin	2362	F2	P00734	ENSG00000180210	Human	desirudin	6458	178103073	t	Inhibitor	Inhibition		Not Determined	f	t		pKi		12.69999981		Ki		0.0002		#ERROR!				1894196
coagulation factor II, thrombin	2362	F2	P00734	ENSG00000180210	Human	lepirudin	6469	178103083	t	Inhibitor	Inhibition		Not Determined	f	t		pKi		13		Ki		1.00E-04		#ERROR!				16363236
coagulation factor V	2606	F5	P12259	ENSG00000198734	Human	drotrecogin alfa	6788	178103394	t	Inhibitor	Inhibition	Antithrombotic ef	Selective	f	t		-				-				#ERROR!				11463021|11714212
coagulation factor VIII	2607	F8	P00451	ENSG00000185010	Human	drotrecogin alfa	6788	178103394	t	Inhibitor	Inhibition	Antithrombotic ef	Selective	f	t		-				-				#ERROR!				11463021|11714212
coagulation factor X	2359	F10	P00742	ENSG00000126218	Human	apixaban	6390	10300606	t	Inhibitor	Inhibition		Not Determined	f	t		pKi		10.10000038		Ki		0.079999998		#ERROR!				18315548
coagulation factor X	2359	F10	P00742	ENSG00000126218	Human	betrixaban	9602	340590232	t	Inhibitor	Inhibition		Selective	f	t		pKi		9.930000305		Ki		0.116999999		#ERROR!	Inhibition of human factor Xa.			19297154
coagulation factor X	2359	F10	P00742	ENSG00000126218	Human	edoxaban	7575	223365911	t	Inhibitor	Inhibition		Not Determined	f	t		pKi		9.25		Ki		0.560000002		#ERROR!				20503967
coagulation factor X	2359	F10	P00742	ENSG00000126218	Human	rivaroxaban	6388	178103004	t	Inhibitor	Inhibition		Not Determined	f	t		pKi		9.399999619		Ki		0.400000006		#ERROR!				20139357
colony stimulating factor 1 receptor	1806	CSF1R	P07333	ENSG00000182578	Human	pexidartinib	8710	252827368	t	Inhibitor	Inhibition		Not Determined	f	t		pIC50		7.889999866		IC50		13		#ERROR!				26222558
complement C1r	2334	C1R	P00736	ENSG00000159403	Human	nafamostat	4262	178101079	t	Inhibitor	Inhibition		Not Determined	f	t		pIC50		4.920000076		IC50		12000		#ERROR!				9544206
COX-1	1375	PTGS1	P23219	ENSG00000095303	Human	aspirin	4139	178100961	t	Inhibitor	Inhibition		Not Determined	f	t		pIC50			3.25	IC50		560100		#ERROR!				19884005
COX-1	1375	PTGS1	P23219	ENSG00000095303	Human	bromfenac	7131	178103707	t	Inhibitor	Inhibition		Not Determined	f	t		pIC50		8.069999695		IC50		8.479999542		#ERROR!				NULL
COX-1	1375	PTGS1	P23219	ENSG00000095303	Human	diclofenac	2714	135650177	t	Inhibitor	Inhibition		Not Determined	f	t		pIC50		7.920000076		IC50		12		#ERROR!				16252917
COX-1	1375	PTGS1	P23219	ENSG00000095303	Human	fenoprofen	4820	178101522	t	Inhibitor	Inhibition		Non-selective	f	t		pIC50		6.840000153		IC50		145		#ERROR!				NULL
COX-1	1375	PTGS1	P23219	ENSG00000095303	Human	flurbiprofen	4194	178101011	t	Inhibitor	Inhibition		Non-selective	f	t		pIC50		7.119999886		IC50		75		#ERROR!				10377455
COX-1	1375	PTGS1	P23219	ENSG00000095303	Human	ibuprofen	2713	135650365	t	Inhibitor	Inhibition		Not Determined	f	t		pIC50		5.519999981		IC50		3000		#ERROR!				9544212
COX-1	1375	PTGS1	P23219	ENSG00000095303	Human	ketoprofen	4795	178101497	t	Inhibitor	Inhibition		Non-selective	f	t		pIC50		6.480000019		IC50		330		#ERROR!	Inhibition of COX1 in human whole blood.			18667313
COX-1	1375	PTGS1	P23219	ENSG00000095303	Human	ketorolac	6661	178103273	t	Inhibitor	Inhibition		Selective	f	t		pIC50		9.720000267		IC50		0.189999998		#ERROR!				10377455
COX-1	1375	PTGS1	P23219	ENSG00000095303	Human	ketorolac	6661	178103273	t	Inhibitor	Inhibition		Non-selective	f	t		pIC50			4.5	IC50		31500		#ERROR!	Instantaneous inhibition of human COX1 by compou	22091869		
COX-1	1375	PTGS1	P23219	ENSG00000095303	Human	meclofenamic acid	7219	178103793	t	Inhibitor	Inhibition		Not Determined	f	t		pIC50		7.300000191		IC50		50		#ERROR!				11844663
COX-1	1375	PTGS1	P23219	ENSG00000095303	Human	mefenamic acid	2593	53801036	t	Inhibitor	Inhibition		Not Determined	f	t		pIC50		4.599999905		IC50		25000		#ERROR!				23454516
COX-1	1375	PTGS1	P23219	ENSG00000095303	Human	naproxen	5230	178101914	t	Inhibitor	Inhibition		Non-selective	f	t		pIC50		5.489999771		IC50		3200		#ERROR!				9057869
COX-1	1375	PTGS1	P23219	ENSG00000095303	Human	nimesulide	7401	178103973	t	Inhibitor	Inhibition		Not Determined	f	t		pIC50		4.760000229		IC50		17400		#ERROR!				15993594
COX-1	1375	PTGS1	P23219	ENSG00000095303	Human	oxaprozin	7252	178103826	t	Inhibitor	Inhibition		Not Determined	f	t		pIC50		5.659999847		IC50		2200		#ERROR!				9650852
COX-1	1375	PTGS1	P23219	ENSG00000095303	Human	phenylbutazone	7270	178103844	t	Inhibitor	Inhibition		Not Determined	f	t		pIC50		5.519999981		IC50		3000		#ERROR!				NULL
COX-1	1375	PTGS1	P23219	ENSG00000095303	Human	piroxicam	7273	178103847	t	Inhibitor	Inhibition		Not Determined	f	t		pIC50		5.899999866		IC50		1300		#ERROR!				9083488
COX-1	1375	PTGS1	P23219	ENSG00000095303	Human	suprofen	7298	178103871	t	Inhibitor	Inhibition		Not Determined	f	t		pIC50			6.25	IC50		560		#ERROR!				18667313
COX-2	1376	PTGS2	P35354	ENSG00000073756	Human	aspirin	4139	178100961	t	Inhibitor	Inhibition		Not Determined	f	t		pIC50		5.619999886		IC50		2400		#ERROR!				22386242
COX-2	1376	PTGS2	P35354	ENSG00000073756	Human	benzydamide	7124	178103701	t	Inhibitor	Inhibition		Not Determined	f	t		pIC50		8.319999695		IC50		4.800000191		#ERROR!				NULL
COX-2	1376	PTGS2	P35354	ENSG00000073756	Human	carprofen	7141	178103717	t	Inhibitor	Inhibition		Not Determined	f	t		pIC50		6.989999771		IC50		102		#ERROR!	Inhibition of COX-2-induced conversion of arachado	21873070		
COX-2	1376	PTGS2	P35354	ENSG00000073756	Human	celecoxib	2892	178100241	t	Inhibitor	Inhibition		Selective	f	t		pIC50		8.699999809		IC50		2		#ERROR!				12643942
COX-2	1376	PTGS2	P35354	ENSG00000073756	Human	celecoxib	2892	178100241	t	Inhibitor	Inhibition		Selective	f	t		pIC50		6.519999981		IC50		300		#ERROR!				29031075
COX-2	1376	PTGS2	P35354	ENSG00000073756	Human	diclofenac	2714	135650177	t	Inhibitor	Inhibition		Not Determined	f	t		pIC50		7.699999809		IC50		20		#ERROR!				17341061
COX-2	1376	PTGS2	P35354	ENSG00000073756	Human	etodolac	7185	178103760	t	Inhibitor	Inhibition		Selective	f	t		pIC50		4.28000021		IC50		53000		#ERROR!	Measured as LPS-induced PGE₂ produ	11804398		
COX-2	1376	PTGS2	P35354	ENSG00000073756	Human	etoricoxib	2896	135700245	t	Inhibitor	Inhibition		Selective	f	t		pIC50		6		IC50		1100		#ERROR!				11160644
COX-2	1376	PTGS2	P35354	ENSG00000073756	Human	flurbiprofen	4194	178101011	t	Inhibitor	Inhibition		Non-selective	f	t		pIC50			8	IC50		10		#ERROR!				10091674
COX-2	1376	PTGS2	P35354	ENSG00000073756	Human	ibuprofen	2713	135650365	t	Inhibitor	Inhibition		Not Determined	f	t		pIC50		5.869999886		IC50		1350		#ERROR!				22091869
COX-2	1376	PTGS2	P35354	ENSG00000073756	Human	ketoprofen	4795	178101497	t	Inhibitor	Inhibition		Non-selective	f	t		pIC50		6.159999847		IC50		690		#ERROR!	Inhibition of COX2 in human whole blood.	18667313		
COX-2	1376	PTGS2	P35354	ENSG00000073756	Human	ketorolac	6661	178103273	t	Inhibitor	Inhibition		Not Determined	f	t		pIC50		4.219999979		IC50		60500		#ERROR!	Instantaneous inhibition of human COX2 by compou	22091869		
COX-2	1376	PTGS2	P35354	ENSG00000073756	Human	ketorolac	6661	178103273	t	Inhibitor	Inhibition		Not Determined	f	t		pIC50		6.920000076		IC50		120		#ERROR!	Inhibition of human COX2 measured after pre-incub	22091869		
COX-2	1376	PTGS2	P35354	ENSG00000073756	Human	lumiracoxib	2897	178100246	t	Inhibitor	Inhibition		Selective	f	t		pKi		6.489999771		Ki		320		#ERROR!				17434872
COX-2	1376	PTGS2	P35354	ENSG00000073756	Human	meclofenamic acid	7219	178103793	t	Inhibitor	Inhibition		Not Determined	f	t		pIC50		7.400000095		IC50		40		#ERROR!				11844663
COX-2	1376	PTGS2	P35354	ENSG00000073756	Human	mefenamic acid	2593	53801036	t	Inhibitor	Inhibition		Not Determined	f	t		pIC50		5.539999962		IC50		2900		#ERROR!				23454516
COX-2	1376	PTGS2	P35354	ENSG00000073756	Human	meloxicam	7220	178103794	t	Inhibitor	Inhibition		Selective	f	t		pIC50		6.309999943		IC50		490		#ERROR!				9083488
COX-2	1376	PTGS2	P35354	ENSG00000073756	Human	naproxen	5230	178101914	t	Inhibitor	Inhibition		Non-selective	f	t		pIC50		5.599999905		IC50		2500		#ERROR!				9057869
COX-2	1376	PTGS2	P35354	ENSG00000073756	Human	nimesulide	7401	178103973	t	Inhibitor	Inhibition		Not Determined	f	t		pIC50		6.219999979		IC50		600		#ERROR!				15993594
COX-2	1376	PTGS2	P35354	ENSG00000073756	Human	oxaprozin	7252	178103826	t	Inhibitor	Inhibition		Not Determined	f	t		pIC50		4.440000057		IC50		36000		#ERROR!				9650852

target	target_id	target_gene	target_uniprot	target_ensembl_gene_id	target_species	ligand	ligand_id	ligand_pubchen	approved_drug	type	action	action_commen	selectivity	endogenous	primary_target	concentration_r	affinity_units	affinity_high	affinity_median	affinity_low	original_affinity	original_affinity	original_affinity	original_affinity	original_affinity	assay_descripti	receptor_site	ligand_context	pubmed_id
COX-2	1376	PTGS2	P35354	ENSG00000073756	Human	phenylbutazone	7270	178103844	t	Inhibitor	Inhibition		Not Determined	f	t		pIC50		3.549999952		IC50		284000		#ERROR!				7562922
COX-2	1376	PTGS2	P35354	ENSG00000073756	Human	piroxicam	7273	178103847	t	Inhibitor	Inhibition		Not Determined	f	t		pIC50		3.660000086		IC50		218000		#ERROR!				7562922
COX-2	1376	PTGS2	P35354	ENSG00000073756	Human	rofecoxib	2893	178100242	t	Inhibitor	Inhibition		Selective	f	t		pIC50	6.5		6.099999905	-				#ERROR!				10377455
COX-2	1376	PTGS2	P35354	ENSG00000073756	Human	suprofen	7298	178103871	t	Inhibitor	Inhibition		Not Determined	f	t		pIC50		5.559999943		IC50		2750		#ERROR!				18667313
COX-2	1376	PTGS2	P35354	ENSG00000073756	Human	valdecoxib	2894	178100243	t	Inhibitor	Inhibition		Selective	f	t		pIC50		8.300000191		IC50		5		#ERROR!				10715145
c-ros oncogene 1, receptor tyrosine kinas	1840	ROS1	P08922	ENSG00000047936	Human	loriatinib	7476	187051779	t	Inhibitor	Inhibition		Not Determined	f	t		pKi		11.30000019		Ki		0.005		#ERROR!				NULL
CS1 (CD319)	2850	SLAMF7	Q9NQ25	ENSG00000026751	Human	elotuzumab	8361	252166573	t	Antibody	Binding		Selective	f	t		pKd		8.050000191		Kd		8.859999657		#ERROR!				NULL
CT receptor	43	CALCR	P30988	ENSG00000004948	Human	calcitonin (salmon)	6973	178103552	t	Agonist	Full agonist		Not Determined	f	t		pEC50	11.10000038		7.599999905	-				#ERROR!				14623894|7769107|7588285|8828514
CXCR4	71	CXCR4	P61073	ENSG00000121966	Human	plerixafor	844	135649932	t	Antagonist	Antagonist		Not Determined	f	t		pKi		6.699999809		Ki		199		#ERROR!				30476826
CXCR4	71	CXCR4	P61073	ENSG00000121966	Human	plerixafor	844	135649932	t	Antagonist	Antagonist		Not Determined	f	t		pIC50		6.190000057		IC50		651		#ERROR!	Antagonism of SDF-1/CXCL12 ligand binding to CXI			16815309
CXCR4	71	CXCR4	P61073	ENSG00000121966	Human	plerixafor	844	135649932	t	Antagonist	Antagonist		Not Determined	f	t		pKi		7		Ki		100		#ERROR!				11923301
cyclin dependent kinase 4	1976	CDK4	P11802	ENSG00000135446	Human	palbociclib	7380	178103952	t	Inhibitor	Inhibition		Not Determined	f	t		pIC50		7.960000038		IC50		11		#ERROR!				15542782
cyclin dependent kinase 4	1976	CDK4	P11802	ENSG00000135446	Human	ribociclib	7383	178103955	t	Inhibitor	Inhibition		Selective	f	t		pIC50		8		IC50		10		#ERROR!			NULL	
cyclin dependent kinase 6	1978	CDK6	Q00534	ENSG00000105810	Human	palbociclib	7380	178103952	t	Inhibitor	Inhibition		Not Determined	f	t		pIC50		7.800000191		IC50		16		#ERROR!				15542782
CYP11B1	1359	CYP11B1	P15538	ENSG00000160882	Human	metyrapone	5224	178101909	t	Inhibitor	Inhibition		Not Determined	f	t		pIC50		7.840000153		IC50		14.60000038		#ERROR!				21129965
CYP19A1	1362	CYP19A1	P11511	ENSG00000137869	Human	aminoglutethimide	7054	178103633	t	Inhibitor	Inhibition		Not Determined	f	t		-				-				#ERROR!				19470632
CYP19A1	1362	CYP19A1	P11511	ENSG00000137869	Human	anastrozole	5137	178101829	t	Inhibitor	Inhibition		Not Determined	f	t		pIC50		7.820000172		IC50		15		#ERROR!				20413308
CYP19A1	1362	CYP19A1	P11511	ENSG00000137869	Human	exemestane	7073	178103652	t	Inhibitor	Inhibition		Selective	f	t		pIC50		7.300000191		IC50		50.09999847		#ERROR!	Inhibition of aromatase (CYP19A1) extracted to hom			2951074
CYP19A1	1362	CYP19A1	P11511	ENSG00000137869	Human	fadrozole	8311	252166523	t	Inhibitor	Inhibition		Non-selective	f	t		pIC50		8.350000381		IC50		4.5		#ERROR!				1825337
CYP19A1	1362	CYP19A1	P11511	ENSG00000137869	Human	letrozole	5209	178101895	t	Inhibitor	Inhibition		Selective	f	t		pKi		10.69999981		Ki		0.02		#ERROR!				22386564
CYP19A1	1362	CYP19A1	P11511	ENSG00000137869	Human	testolactone	7303	178103876	t	Inhibitor	Inhibition		Selective	f	t		pKi		4.460000038		Ki		35000		#ERROR!				7083195
CysLT₁ receptor	269	CYSLTR1	Q9Y271	ENSG00000173198	Human	montelukast	3340	178100370	t	Antagonist	Antagonist		Not Determined	f	t		pKi		8.550000191		-				#ERROR!	against [³H]LTD₄ in huma			11996896
CysLT₁ receptor	269	CYSLTR1	Q9Y271	ENSG00000173198	Human	montelukast	3340	178100370	t	Antagonist	Antagonist	against 33nM LTI	Not Determined	f	t		pIC50		8.640000343		IC50		2.299999952		#ERROR!				10462554
CysLT₁ receptor	269	CYSLTR1	Q9Y271	ENSG00000173198	Human	montelukast	3340	178100370	t	Antagonist	Antagonist		Selective	f	t		pIC50	8.640000343		8.31000042	IC50	4.900000095		2.299999952		#ERROR!	against [³H]LTD4 in COS-7 or HEK-29	10391245|10462554		
CysLT₁ receptor	269	CYSLTR1	Q9Y271	ENSG00000173198	Human	pranlukast	3634	178100541	t	Antagonist	Antagonist		Not Determined	f	t		pKi		8.800000191		7.099999905	-			#ERROR!	against [³H]LTD₄ in huma	11996896|9504401		
CysLT₁ receptor	269	CYSLTR1	Q9Y271	ENSG00000173198	Human	pranlukast	3634	178100541	t	Antagonist	Antagonist	against 33nM LTI	Not Determined	f	t		pIC50		10		IC50		0.100000001		#ERROR!				10462554
CysLT₁ receptor	269	CYSLTR1	Q9Y271	ENSG00000173198	Human	pranlukast	3634	178100541	t	Antagonist	Antagonist	against [<sup>3<	Not Determined	f	t		pIC50		8.369999886		8.140000343	IC50	7.199999809		4.300000191		#ERROR!				10391245|10462554
CysLT₁ receptor	269	CYSLTR1	Q9Y271	ENSG00000173198	Human	zafirlukast	3322	178100355	t	Antagonist	Antagonist	zafirlukast is only	Selective	f	t		pKi		8.899999619		-				#ERROR!	against [³H]LTD₄ in huma	11996896|9504401		
CysLT₁ receptor	269	CYSLTR1	Q9Y271	ENSG00000173198	Human	zafirlukast	3322	178100355	t	Antagonist	Antagonist	against 10nM LTI	Not Determined	f	t		pIC50		7.699999809		IC50		20		#ERROR!				20423349
CysLT₁ receptor	269	CYSLTR1	Q9Y271	ENSG00000173198	Human	zafirlukast	3322	178100355	t	Antagonist	Antagonist		Selective	f	t		pIC50	8.739999771		8.590000153	IC50	2.599999905		1.799999952		#ERROR!	against [³H]LTD₄ in COS-	10391245|10462554		
CysLT₁ receptor	269	CYSLTR1	Q9Y271	ENSG00000173198	Human	zafirlukast	3322	178100355	t	Antagonist	Antagonist	against 33nM LTI	Not Determined	f	t		pIC50		9.590000153		IC50		0.25999999		#ERROR!				10462554
cytotoxic T-lymphocyte-associated protein	2743	CTLA4	P16410	ENSG00000163599	Human	ipilimumab	6888	178103470	t	Antibody	Binding		Selective	f	t		pKd		9		Kd		1	<					NULL
DAT	927	SLC6A3	Q01959	ENSG00000142319	Human	bupropion	7135	178103711	t	Inhibitor	Inhibition		Non-selective	f	t		pIC50		6.349999905		IC50		443		#ERROR!				19821577
DAT	927	SLC6A3	Q01959	ENSG00000142319	Human	dexamfetamine	2147	135650979	t	Inhibitor	Inhibition		Non-selective	f	t		pKi		6.960000038		Ki		109		#ERROR!				23602445
DAT	927	SLC6A3	Q01959	ENSG00000142319	Human	dexmethylphenidate	7554	223365890	t	Inhibitor	Inhibition		Selective	f	t		pKi		7.599999905		Ki		25		#ERROR!	Displacement of [³H]WIN-35428 from t			21129986
DAT	927	SLC6A3	Q01959	ENSG00000142319	Human	mazindol	4797	178101499	t	Inhibitor	Inhibition		Selective	f	t		pKi		8		Ki		10		#ERROR!			NULL	
DAT	927	SLC6A3	Q01959	ENSG00000142319	Human	methylphenidate	7236	178103810	t	Inhibitor	Inhibition		Selective	f	t		pIC50		7.099999905		IC50		79		#ERROR!				17228864
DAT	927	SLC6A3	Q01959	ENSG00000142319	Human	modafenane	4792	178101494	t	Inhibitor	Inhibition		Non-selective	f	t		-				-				#ERROR!			NULL	
DAT	927	SLC6A3	Q01959	ENSG00000142319	Human	sibutramine	2586	135651043	t	Inhibitor	Inhibition		Family selective	f	t		pKi		6.300000191		Ki		502		#ERROR!			NULL	
DAT	927	SLC6A3	Q01959	ENSG00000142319	Human	trimipramine	7317	178103890	t	Inhibitor	Inhibition		Non-selective	f	t		pKi		5.420000076		Ki		3780		#ERROR!				9537821
δ receptor	317	OPRD1	P41143	ENSG00000116329	Human	nalbuphine	1663	135650680	t	Agonist	Agonist		Not Determined	f	t		pKi		6.239999771		Ki		580		#ERROR!				19282177
δ receptor	317	OPRD1	P41143	ENSG00000116329	Human	naloxone	1638	135650684	t	Antagonist	Antagonist		Not Determined	f	t		pKi		7.199999809		-				#ERROR!				9686407
δ receptor	317	OPRD1	P41143	ENSG00000116329	Human	naltrexone	1639	135650687	t	Antagonist	Antagonist		Not Determined	f	t		pKi				8	-			#ERROR!				9686407
dihydrofolate reductase	2603	DHFR	P00374	ENSG00000228716	Human	methotrexate	4815	178101517	t	Inhibitor	Inhibition		Selective	f	t		pKi		8.920000076		Ki		1.200000048		#ERROR!				7877140
dihydrofolate reductase	2603	DHFR	P00374	ENSG00000228716	Human	pemetrexed	6837	178103443	t	Inhibitor	Inhibition		Not Determined	f	t		pKi		8.149999619		Ki		7		#ERROR!				16078850|9762351
dihydrofolate reductase	2603	DHFR	P00374	ENSG00000228716	Human	pralatrexate	6840	178103446	t	Inhibitor	Inhibition		Not Determined	f	t		pKi		7.349999905		Ki		45		#ERROR!				19221750
dihydroorotate dehydrogenase (quinone)	2604	DHODH	Q02127	ENSG00000102967	Human	leflunomide	6825	178103431	t	Inhibitor	Inhibition		Not Determined	f	t		pKi		4.899999866		Ki		13000		#ERROR!				9719606
dihydroorotate dehydrogenase (quinone)	2604	DHODH	Q02127	ENSG00000102967	Human	teriflunomide	6844	178103450	t	Inhibitor	Inhibition		Not Determined	f	t		pKi		7.519999981		Ki		30		#ERROR!				17228860
dipeptidyl peptidase 4	1612	DPP4	P27487	ENSG00000197635	Human	alogliptin	6319	178102937	t	Inhibitor	Inhibition		Not Determined	f	t		pIC50		8.5		IC50		3.400000095		#ERROR!				22475866
dipeptidyl peptidase 4	1612	DPP4	P27487	ENSG00000197635	Human	linagliptin	6318	178102936	t	Inhibitor	Inhibition		Not Determined	f	t		pKi		9		Ki		1		#ERROR!				18052023
dipeptidyl peptidase 4	1612	DPP4	P27487	ENSG00000197635	Human	saxagliptin	6316	178102934	t	Inhibitor	Inhibition		Not Determined	f	t		pKi		9.199999809		Ki		0.600000024		#ERROR!				19149538
dipeptidyl peptidase 4	1612	DPP4	P27487	ENSG00000197635	Human	sitagliptin	6286	178102906	t	Inhibitor	Inhibition		Not Determined	f	t		pIC50		8.100000381		IC50		8.600000381		#ERROR!				20927248
dipeptidyl peptidase 4	1612	DPP4	P27487	ENSG00000197635	Human	teneligliptin	9906	363894200	t	Inhibitor	Inhibition		Not Determined	f	t		pIC50		9.430000305		IC50		0.370000005		#ERROR!				22695656
dipeptidyl peptidase 4	1612	DPP4	P27487	ENSG00000197635	Human	vildagliptin	6310	178102929	t	Inhibitor	Inhibition		Not Determined	f	t		pKi		7.800000191		Ki		17		#ERROR!				19149538
DNA topoisomerase II alpha	2637	TOP2A	P11388	ENSG00000131747	Human	etoposide	6815	178103421	t	Inhibitor	Inhibition		Not Determined	f	t		pIC50		7.300000191		IC50		50.40000153		#ERROR!	Note: enzyme subunit is not specified in the referenc			
D₁ receptor	214	DRD1	P21728	ENSG00000184845	Human	fenoldopam	939	135651054	t	Agonist	Full agonist		Not Determined	f	t		pKi	7.900000095		6.5	-				#ERROR!				7525564
D₁ receptor	214	DRD1	P21728	ENSG00000184845	Human	flupentixol	948	135650273	t	Antagonist	Antagonist		Not Determined	f	t		pKi	8.399999619		7	-				#ERROR!	Note that this assay uses the cis form of the c	1826762|7525564		
D₂ receptor	215	DRD2	P14416	ENSG00000149295	Human	benzquinamide	7124	178103701	t	Agonist	Agonist		Not Determined	f	t		pKi		5.400000095		Ki		3964		#ERROR!				22711801
D₂ receptor	215	DRD2	P14416	ENSG00000149295	Human	ropinirole	7295	178103868	t	Agonist	Agonist		Not Determined	f	t		pKi		8.140000343		Ki		7.199999809		#ERROR!				9057850
D₂ receptor	215	DRD2	P14416	ENSG00000149295	Human	rotigotine	941	135650929	t	Agonist	Agonist		Not Determined	f	t		pKi		10.22000027		Ki		0.059999999		#ERROR!				12086487
D₂ receptor	215	DRD2	P14416	ENSG00000149295	Human	bromocriptine	35	135650026	t	Agonist	Full agonist		Not Determined	f	t		pKi	8.300000191		7.300000191	-				#ERROR!				12388666|8301582|7756621
D₂ receptor	215	DRD2	P14416	ENSG00000149295	Human	pramipexole	953	135650847	t	Agonist	Full agonist		Not Determined	f	t		pKi	7.400000095		5.099999905	-				#ERROR!				7756621|7664822
D₂ receptor	215	DRD2	P14416	ENSG00000149295	Human	apomorphine	33	135649950	t	Agonist	Partial agonist		Not Determined	f	t		pKi	7.5		5.699999809	-				#ERROR!				12388666|8301582|7756621|7576010|1354163
D₂ receptor	215	DRD2	P14416	ENSG00000149295	Human	aripiprazole	34	135649956	t	Agonist	Partial agonist		Not Determined	f	t		pKi		9.100000381		Ki		0.800000012		#ERROR!				23279866
D₂ receptor	215	DRD2	P14416	ENSG00000149295	Human	aripiprazole	34	135649956	t	Agonist	Partial agonist		Not Determined	f	t		pEC50		7.420000076		EC50		38		#ERROR!	Measuring cAMP production via the Gi-coupled sign	22025698		
D₂ receptor	215	DRD2	P14416	ENSG00000149295	Human	cabergoline	37	135652674	t	Agonist	Partial agonist		Not Determined	f	t		pKi	9.199999809		9	-				#ERROR!				12388666
D₂ receptor	215	DRD2	P14416	ENSG00000149295	Human	amisulpride	963	135649938	t	Antagonist	Antagonist		Not Determined	f	t		pKi	8		7.900000095	-				#ERROR!				1354163|8099194
D₂ receptor	215	DRD2	P14416	ENSG00000149295	Human	amisulpride	963	135649938	t	Antagonist	Antagonist		Not Determined	f	t		pKi		7.800000191		-				#ERROR!				1975644
D₂ receptor	215	DRD2	P14416	ENSG00000149295	Human	asenapine	22	135650779	t	Antagonist	Antagonist		Not Determined	f	t		pKi		8.920000076		Ki		1.200000048		#ERROR!				18308814
D₂ receptor	215	DRD2	P14416	ENSG00000149295	Human	blonanserin	7670	223366005	t	Antagonist	Antagonist		Not Determined	f	t		pKi		9.850000381		Ki		0.140000001		#ERROR!				15686911
D₂ receptor	215	DRD2	P14416	ENSG00000149295	Human	clozapine	38	135650115	t	Antagonist	Antagonist		Not Determined	f	t		pKi	6.900000095		5.800000191	-				#ERROR!				8301582|7907989|1354163|8301592|18308814
D₂ receptor	215	DRD2	P14416	ENSG00000149295	Human	domperidone	965	135650206	t	Antagonist	Antagonist		Selective	f	t		pKi	8.399999619		7.900000095	Ki	12.60000038		3.980000019		#ERROR!				8301582|1354163
D₂ receptor	215	DRD2	P14416	ENSG00000149295	Human	fluspirilene	948	135650273	t	Antagonist	Antagonist		Not Determined	f	t		pKi		8.819999695		Ki		1.5		#ERROR!				8301582
D₂ receptor	215	DRD2	P14416	ENSG00000149295	Human	fluphenazine	204	135650275	t	Antagonist	Antagonist		Not Determined	f	t		pKi		8.840000153		Ki		1.440000057		#ERROR!				17826096
D₂ receptor	215	DRD2	P14416	ENSG00000149295	Human	haloperidol	86	135650345	t	Antagonist	Antagonist		Family selective	f	t		pKi	8.800000191		7.400000095	-				#ERROR!				8301582|7907989|1354163|7664822|7862709
D₂ receptor	215	DRD2	P14416	ENSG00000149295	Human	loxapine	205	135650533	t	Antagonist	Antagonist		Non-selective	f	t		pKi	8.300000191		7.920000076	Ki		12		5	#ERROR!				12629531|9015795
D₂ receptor	215	DRD2	P14416	ENSG00000149295	Human	olanzapine	47	135650796	t	Antagonist	Antagonist		Not Determined	f	t		pKi		8.680000305		Ki		2.099999905		#ERROR!				9430133
D₂ receptor	215	DRD2	P14416	ENSG00000149295	Human	perospirone	7556	223365892	t	Antagonist	Antagonist		Not Determined	f	t		pKi		9.220000267		Ki		0.600000024		#ERROR!				1060115
D₂ receptor	215	DRD2	P14416	ENSG00000149295	Human	perphenazine	209	135650810	t	Antagonist	Antagonist		Not Determined	f	t		pKi	9.590000153		8.850000381	Ki	1.399999976		0.25999999		#ERROR!				12629531
D₂ receptor	215	DRD2	P14416	ENSG00000149295	Human	pimozide	90	135650825	t	Antagonist	Antagonist		Not Determined	f	t		pKi	8.800000191		7	-				#ERROR!				8301582|1354163
D₂ receptor	215	DRD2	P14416	ENSG00000149295	Human	pipotiazine	7557	223365893	t	Antagonist	Antagonist		Not Determined	f	t		pKi		9.699999809		Ki		0.200000003		#ERROR!				1586393
D₂ receptor	215	DRD2	P14416	ENSG00000149295	Human	prochlorperazine	7279	178103853	t	Antagonist	Antagonist		Not Determined	f	t		pKi				Ki		3.609999895		#ERROR!			NULL	
D₂ receptor	215	DRD2	P14416	ENSG00000149295	Human	promazine	7281	178103855	t	Antagonist	Antagonist		Not Determined	f	t		pKi		6.519999981		Ki		300		#ERROR!				16135699
D₂ receptor	215	DRD2	P14416	ENSG00000149295	Human	quetiapine	50	135650869	t	Antagonist	Antagonist		Not Determined	f	t		pKi		7.199999809		Ki		69		#ERROR!				9430133
D₂ receptor	215	DRD2	P14416	ENSG00000149295	Human	risperidone	96	135650903	t	Antagonist	Antagonist		Not Determined	f	t		pKi		9.359999657		Ki		0.439999998		#ERROR!				9430133
D₂ receptor	215	DRD2	P14416	ENSG00000149295	Human	sertindole	98	135651037	t	Antagonist	Antagonist		Not Determined	f	t		pKi	8.920000076		8.039999962	Ki	9.100000381		1.200000048		#ERROR!				12629531|12176106
D₂ receptor	215	DRD2	P14416	ENSG00000149295	Human	sulpiride	5501	135652147	t	Antagonist	Antagonist		Not Determined	f	t		pKi	7.920000076		7.21999979	Ki	60		12		#ERROR!				7576010
D₂ receptor	215	DRD2	P14416	ENSG00000149295	Human	(-)-sulpiride	958	135651093	t	Antagonist	Antagonist		Not Determined	f	t		pKi	8		6.28000021	Ki	520		10		#ERROR!				8301582|1354163|8301592
D₂ receptor	215	DRD2	P14416	ENSG00000149295	Human	trifluoperazine	214	135651279	t	Antagonist	Antagonist		Not Determined	f	t		pKi	9.020000458		8.890000343	Ki	1.299999952		0.959999979		#ERROR!				12629531|9015795
D₂ receptor	215	DRD2	P14416	ENSG00000149295	Human	ziprasidone	59	135651359	t	Antagonist	Antagonist		Not Determined	f	t		pKi		8.550000191		Ki		2.799999952		#ERROR!				9430133
D₃ receptor	216	DRD3	P35462	ENSG00000151577	Human	ropinirole	7295	178103868	t	Agonist	Agonist		Not Determined	f	t		pKi		7.71999979		Ki		19		#ERROR!				9057850
D₃ receptor	216	DRD3	P35462	ENSG00000151577	Human	pramipexole	953	135650847	t	Agonist	Full agonist		Not Determined	f	t		pKi	8.699999809		8.399999619	-				#ERROR!				7756621|7664822
D₃ receptor	216	DRD3	P35462	ENSG00000151577	Human	apomorphine	33	135649950	t	Agonist	Partial agonist		Not Determined	f	t		pKi	7.599999905		6.099999905	-				#ERROR!				12388666|8301582|7756621|7576010|1354163
D₃ receptor	216	DRD3	P35462	ENSG00000151577	Human	bromocriptine	35	135650026	t	Agonist	Partial agonist		Not Determined	f	t		pKi	8.199999809		7.099999905	-				#ERROR!				12388666|8301582|7756621
D₃ receptor	216	DRD3	P35462	ENSG00000151577	Human	domperidone	965	135650206	t	Antagonist	Antagonist		Not Determined	f	t		pKi	7.599999905		7.099999905	-				#ERROR!				8301582|1354163
D₃ receptor	216	DRD3	P35462	ENSG00000151577	Human	haloperidol	86	135650345	t	Antagonist	Antagonist		Family selective	f	t		pKi	8.600000381		7.5	-				#ERROR!				8301582|1354163|7862709|18308814
D₃ receptor	216	DRD3	P35462	ENSG00000151577	Human	loxapine	205	135650533	t	Antagonist	Antagonist		Non-selective	f	t		pKi		7.679999828		Ki		21		#ERROR!			NULL	
D₃ receptor	216	DRD3	P35462	ENSG00000151577	Human	perospirone	7556	223365892	t	Antagonist	Antagonist		Not Determined	f	t		pKi		9.550000191		Ki		0.280000001		#ERROR!				1354163
D₃ receptor	216	DRD3	P35462	ENSG00000151577	Human	prochlorperazine	7279	178103853	t	Antagonist	Antagonist		Not Determined	f	t		pKi		8.350000381		Ki		4.449999809		#ERROR!			NULL	
D₃ receptor	216	DRD3	P35462	ENSG00000151577	Human	promazine	7281	178103855	t	Antagonist	Antagonist		Not Determined	f	t		pKi		6.800000191		Ki		159		#ERROR!				16135699
D₃ receptor	216	DRD3	P35462	ENSG00000151577	Human	sertindole	98	135651037	t	Antagonist	Antagonist		Not Determined	f	t		pKi	8.800000191		8	Ki	10		1.600000024		#ERROR!				8935801|9430133
D₃ receptor	216	DRD3	P35462	ENSG00000151577	Human	(-)-sulpiride	958	135651093	t	Antagonist	Antagonist		Not Determined	f	t		pKi	7.699999809		6.679999828	Ki	207		20		#ERROR!				8301582|1354163|8301592
D₄ receptor	217	DRD4	P21917	ENSG00000069696	Human	apomorphine	33	135649950	t	Agonist	Partial agonist		Not Determined	f	t		pKi	8.399999619		-						#ERROR!				12388666
D₄ receptor	217	DRD4	P21917	ENSG00000069696	Human	bromocriptine	35	135650026	t	Antagonist	Antagonist		Not Determined	f	t		pKi	6.400000095		-						#ERROR!				12388666
D₄ receptor	217	DRD4	P21917	ENSG00000069696	Human	haloperidol	86	135650345	t	Antagonist	Antagonist		Family selective	f	t		pKi	8.800000191		8.699999809	-					#ERROR!				8102973|7862709|18308814
D₄ receptor	217	DRD4	P21917	ENSG00000069696	Human	loxapine	205	135650533	t	Antagonist	Antagonist		Non-selective	f	t		pKi		8.109999657		Ki		7.800000191		#ERROR!				9015795
D₄ receptor	217	DRD4	P21917	ENSG00000069696	Human	perospirone	7556	223365892	t	Antagonist	Antagonist		Not Determined	f	t		pKi		10.05000019		Ki		0.090000004		#ERROR!				9577836
D₄ receptor	217	DRD4	P21917	ENSG00000069696	Human	sertindole	98	135651037	t	Antagonist	Antagonist		Not Determined	f	t		pKi	9.069999695		7.800000191	Ki	16		0.850000024		#ERROR!				16135699|9577836|9015795
elastase, neutrophil expressed	2358	ELANE	P08246	ENSG00000197561	Human	alpha1-proteinase inh	7487	187051790	t	Inhibitor	Inhibition		Not Determined	f	t		-				-				#ERROR!			NULL	
elastase, neutrophil expressed	2358	ELANE	P08246	ENSG00000197561	Human	sivelestat	6441	178103057	t	Inhibitor	Inhibition		Not Determined	f	t		pIC50		7.400000095		IC50		44		#ERROR!				21741848
ENaCαβγ	742	SCNN1B|SCNN1	P37088|P51168|	ENSG00000111319|ENSG	Human	amiloride	2421	135649936	t	Channel blocker	None		Non-selective	f	t		pIC50	7		6.699999809	IC50	200		100		#ERROR!			NULL	
ENaCαβγ	742	SCNN1B|SCNN1	P37088|P51168|	ENSG00000111319|ENSG	Human	triamterene	4329	178101146	t	Channel blocker	None		Not Determined	f	t		pIC50		5.300000191		IC50		5000	–					8107805|14500741
enhancer of zeste 2 polycomb repressive	2654	EZH2	Q15910	ENSG00000106462	Human	tazemetostat	7011	178103590	t	Inhibitor	Inhibition		Not Determined	f	t		pIC50		7.960000038		IC50		11		#ERROR!				23620515
epidermal growth factor receptor	1797	EGFR	P00533	ENSG00000146648	Human	panitumumab	6883	178103465	t	Antibody	Antagonist		Selective	f	t		pKd		10.30000019		Kd		0.050000001		#ERROR!				14967460
epidermal growth factor receptor	1797	EGFR	P00533	ENSG00000146648	Human	cetuximab	6882	178103464	t	Antibody	Binding		Selective	f	t		pKd		9.409999847		Kd		0.389999986		#ERROR!				NULL
epidermal growth factor receptor	1797	EGFR	P00533	ENSG00000146648	Human	necitumumab	8090	240565770	t	Antibody	Binding		Selective	f	t		pKd		9.489999771		Kd		0.319999993		#ERROR!	Binding affinity of whole IMC-11F8 IgG using BIAcor			
epidermal growth factor receptor	1797	EGFR	P00533	ENSG00000146648	Human	afatinib	5667	178102295	t	Inhibitor	Inhibition		Not Determined	f	t		pIC50	9.300000191		7.960000038	IC50	11		0.5	

target	target_id	target_gene	target_uniprot	target_ensembl_gene_id	target_species	ligand	ligand_id	ligand_pubchen	approved_drug	type	action	action_commen	selectivity	endogenous	primary_target	concentration_r	affinity_units	affinity_high	affinity_median	affinity_low	original_affinity	original_affinity	original_affinity	original_affinity	original_affinity	assay_descripti	receptor_site	ligand_context	pubmed_id
epidermal growth factor receptor	1797	EGFR	P00533	ENSG00000146648	Human	erlotinib	4920	176101619	t	Inhibitor	Inhibition		Not Determined	f	t		pKd		9.170000076		Kd	0.670000017		#ERROR!					22037378
epidermal growth factor receptor	1797	EGFR	P00533	ENSG00000146648	Human	gefitinib	4941	176101639	t	Inhibitor	Inhibition		Not Determined	f	t		pKd		9		Kd	1		#ERROR!					22037378
epidermal growth factor receptor	1797	EGFR	P00533	ENSG00000146648	Human	gefitinib	4941	176101639	t	Inhibitor	Inhibition		Not Determined	f	t		pKi		8.300000191		Ki	5.5		#ERROR!					17416531
epidermal growth factor receptor	1797	EGFR	P00533	ENSG00000146648	Human	icotinib	7641	223365976	t	Inhibitor	Inhibition		Not Determined	f	t		pIC50		7.349999905		IC50	45		#ERROR!		Inhibition of EGFR-mediated intracellular tyrosine ph			23959248
epidermal growth factor receptor	1797	EGFR	P00533	ENSG00000146648	Human	lapatinib	5692	178102319	t	Inhibitor	Inhibition		Not Determined	f	t		pIC50		8		IC50	10.19999981		#ERROR!					12467226
EP₂ receptor	341	PTGER2	P43116	ENSG00000125384	Human	PGE₁	1882	135650814	t	Agonist	Full agonist		Not Determined	f	t		pKi		8.039999962		Ki	9.100000381		#ERROR!					8163486
EP₃ receptor	342	PTGER3	P43115	ENSG00000050628	Human	misoprostol (methyl e	1936	135650611	t	Agonist	Full agonist	EP₃	Not Determined	f	t		pKi		6.5		Ki	319		#ERROR!					10634944
Equilibrative nucleoside transporter 1	1117	SLC29A1	Q99808	ENSG00000122759	Human	dipyridamole	4807	178101509	t	Inhibitor	Inhibition		Non-selective	f	t		pKi		8.800000191		Ki	1.5		#ERROR!					14634039
erb-b2 receptor tyrosine kinase 2	2019	ERBB2	P04626	ENSG00000141736	Human	pertuzumab	5046	178101741	t	Antibody	Inhibition		Not Determined	f	t		pIC50		8		IC50	10	<			Preferred value for inhibition of heregulin β1 bi	NULL		
erb-b2 receptor tyrosine kinase 2	2019	ERBB2	P04626	ENSG00000141736	Human	trastuzumab	5082	178101775	t	Antibody	Inhibition		Not Determined	f	t		-				-							NULL	
erb-b2 receptor tyrosine kinase 2	2019	ERBB2	P04626	ENSG00000141736	Human	lapatinib	5692	178102319	t	Inhibitor	Inhibition		Not Determined	f	t		pKd		8.149999619		Kd	7		#ERROR!					22037378
erb-b2 receptor tyrosine kinase 2	2019	ERBB2	P04626	ENSG00000141736	Human	lapatinib	5692	178102319	t	Inhibitor	Inhibition		Not Determined	f	t		pIC50		8		IC50	9.800000191		#ERROR!					12467226
erb-b2 receptor tyrosine kinase 2	2019	ERBB2	P04626	ENSG00000141736	Human	neratinib	5686	178102313	t	Inhibitor	Inhibition		Not Determined	f	t		pIC50		7.230000019		IC50	59		#ERROR!		Inhibition of autophosphorylation of the cytoplasmic			15173008
erb-b2 receptor tyrosine kinase 2	2019	ERBB2	P04626	ENSG00000141736	Human	neratinib	5686	178102313	t	Inhibitor	Inhibition		Not Determined	f	t		pKd		8.220000267		Kd	6		#ERROR!					22037378
erb-b2 receptor tyrosine kinase 2	2019	ERBB2	P04626	ENSG00000141736	Human	tucatinib	9922	363894216	t	Inhibitor	Inhibition		Selective	f	t		pIC50		8.159999847		IC50	6.900000095		#ERROR!		Inhibition of enzyme activity in a biochemical assay.			32241871
Estrogen receptor-α	620	ESR1	P03372	ENSG00000091831	Human	diethylstilbestrol	2801	135651472	t	Agonist	Agonist		Not Determined	f	t		pKi		10.39999962		-			#ERROR!					9048584
Estrogen receptor-α	620	ESR1	P03372	ENSG00000091831	Human	ethinylestradiol	2801	178103650	t	Agonist	Agonist		Selective	f	t		pIC50		8.699999809		IC50	2		#ERROR!					16722623
Estrogen receptor-α	620	ESR1	P03372	ENSG00000091831	Human	tamoxifen	2801	135651233	t	Agonist	Agonist		Not Determined	f	t		pKi		7.820000172		-			#ERROR!					9048594
Estrogen receptor-α	620	ESR1	P03372	ENSG00000091831	Human	bazedoxifene	2801	178103227	t	Antagonist	Antagonist		Family selective	f	t		pIC50		7.639999866		IC50	23		#ERROR!					11356100
Estrogen receptor-α	620	ESR1	P03372	ENSG00000091831	Human	clomiphene	2801	178100979	t	Antagonist	Antagonist		Selective	f	t		pKi		8.890000343		Ki	1.299999952		#ERROR!				NULL	
Estrogen receptor-α	620	ESR1	P03372	ENSG00000091831	Human	fulvestrant	2801	135650370	t	Antagonist	Antagonist		Not Determined	f	t		pKi		8.979999542		Ki	1.039999962		#ERROR!					12672240
Estrogen receptor-α	620	ESR1	P03372	ENSG00000091831	Human	raloxifene	2801	135650883	t	Antagonist	Antagonist		Not Determined	f	t		pKi		9.5		-			#ERROR!					9048584
Estrogen receptor-α	620	ESR1	P03372	ENSG00000091831	Human	tamoxifen	2801	135651233	t	Antagonist	Antagonist		Not Determined	f	t		pKi		7.800000191		-			#ERROR!					9048584
Estrogen receptor-β	621	ESR2	Q92731	ENSG00000140009	Human	bazedoxifene	7355	178103927	t	Antagonist	Antagonist		Family selective	f	t		pIC50		7.070000172		IC50	85		#ERROR!					11356100
Estrogen receptor-β	621	ESR2	Q92731	ENSG00000140009	Human	fulvestrant	1015	135650370	t	Antagonist	Antagonist		Not Determined	f	t		pKi		8.859999657		Ki	1.389999986		#ERROR!					12672240
Estrogen receptor-β	621	ESR2	Q92731	ENSG00000140009	Human	raloxifene	2820	135650883	t	Antagonist	Antagonist		Not Determined	f	t		pKi		8		Ki	10		#ERROR!					11459665
ET_A receptor	219	EDNRA	P25101	ENSG00000151617	Human	ambrisentan	3951	178100778	t	Antagonist	Antagonist		Selective	f	t		pA2		7.099999905		-			#ERROR!					15139756
ET_A receptor	219	EDNRA	P25101	ENSG00000151617	Human	ambrisentan	3951	178100778	t	Antagonist	Antagonist		Selective	f	t		pIC50		7.659999847		IC50	21.70000076		#ERROR!					15139756
ET_A receptor	219	EDNRA	P25101	ENSG00000151617	Human	macitentan	7352	178103924	t	Antagonist	Antagonist		Selective	f	t		pIC50		9.300000191		IC50	0.5		#ERROR!					23862294
ET_A receptor	219	EDNRA	P25101	ENSG00000151617	Human	sitaxsentan	3950	178100777	t	Antagonist	Antagonist		Selective	f	t		pA2		8		-			#ERROR!					9171878
ET_B receptor	220	EDNRB	P24530	ENSG00000136160	Human	bosentan	3494	178100482	t	Antagonist	Antagonist		Not Determined	f	t		pKi		7.099999905		Ki	80		#ERROR!					12502366
Erythropoietin receptor	1718	EPOR	P19235	ENSG00000187266	Human	peginesatide	7447	178104019	t	Agonist	Agonist		Not Determined	f	t		pIC50		10.43000031		IC50	0.037		#ERROR!					16982323
Farnesoid X receptor	603	NR1H4	Q96RI1	ENSG00000012504	Human	obeticholic acid	3435	176100437	t	Agonist	Agonist		Selective	f	t		pEC50		7		EC50	99		#ERROR!					12166927
farnesyl diphosphate synthase	644	FDPS	P14324	ENSG00000160752	Human	alendronate	3141	135649795	t	Inhibitor	Inhibition		Selective	f	t		pIC50		7.300000191		IC50	50		#ERROR!		In vitro inhibition of recombinant human prote			11160603
farnesyl diphosphate synthase	644	FDPS	P14324	ENSG00000160752	Human	alendronate	3141	135649795	t	Inhibitor	Inhibition		Not Determined	f	t		pIC50		6.340000153		IC50	460		#ERROR!					10620343
farnesyl diphosphate synthase	644	FDPS	P14324	ENSG00000160752	Human	ibandronic acid	3059	135649664	t	Inhibitor	Inhibition		Selective	f	t		pKi		6.710000038		Ki	195		#ERROR!					18327899
farnesyl diphosphate synthase	644	FDPS	P14324	ENSG00000160752	Human	pamidronic acid	7259	178103833	t	Inhibitor	Inhibition		Selective	f	t		pIC50		6.699999809		IC50	200		#ERROR!		Inhibition of recombinant human FDPS expressed in			18327899
farnesyl diphosphate synthase	644	FDPS	P14324	ENSG00000160752	Human	risedronate	3176	135650902	t	Inhibitor	Inhibition		Not Determined	f	t		pKi		7.090000153		Ki	82.19999695		#ERROR!		Inhibition of human recombinant FPPS expressed in			18327899
farnesyl diphosphate synthase	644	FDPS	P14324	ENSG00000160752	Human	risedronate	3176	135650902	t	Inhibitor	Inhibition		Not Determined	f	t		pIC50		8.409999847		IC50	3.900000095		#ERROR!					10620343
farnesyl diphosphate synthase	644	FDPS	P14324	ENSG00000160752	Human	risedronate	3176	135650902	t	Inhibitor	Inhibition		Not Determined	f	t		pIC50		8.239999771		IC50	5.699999809		#ERROR!		Inhibition of human recombinant FPPS expressed in			18327899
farnesyl diphosphate synthase	644	FDPS	P14324	ENSG00000160752	Human	risedronate	3176	135650902	t	Inhibitor	Inhibition		Not Determined	f	t		pIC50		8.239999771		IC50	5.699999809		#ERROR!		in vitro assay, inhibition of recombinant human enzy			11160603
farnesyl diphosphate synthase	644	FDPS	P14324	ENSG00000160752	Human	risedronate	3176	135650902	t	Inhibitor	Inhibition		Not Determined	f	t		pIC50		6.340000153		IC50	452.8999939		#ERROR!		Inhibition of human recombinant FPPS expressed in			18327899
farnesyl diphosphate synthase	644	FDPS	P14324	ENSG00000160752	Human	risedronate	3176	135650902	t	Inhibitor	Inhibition		Selective	f	t		pIC50		8		IC50	10		#ERROR!		In vitro inhibition of recombinant human FDPS			11160603
farnesyl diphosphate synthase	644	FDPS	P14324	ENSG00000160752	Human	risedronate	3176	135650902	t	Inhibitor	Inhibition		Not Determined	f	t		pKi		9.43999958		Ki	0.360000014		#ERROR!		Inhibition of human recombinant FPPS expressed in			18327899
farnesyl diphosphate synthase	644	FDPS	P14324	ENSG00000160752	Human	zoledronic acid	3177	135651365	t	Inhibitor	Inhibition		Not Determined	f	t		pIC50		8.390000343		IC50	4.099999905		#ERROR!		Inhibition of human recombinant FPPS expressed in			18327899
farnesyl diphosphate synthase	644	FDPS	P14324	ENSG00000160752	Human	zoledronic acid	3177	135651365	t	Inhibitor	Inhibition		Not Determined	f	t		pKi		7.070000172		Ki	85.90000153		#ERROR!		Inhibition of human recombinant FPPS expressed in			18327899
farnesyl diphosphate synthase	644	FDPS	P14324	ENSG00000160752	Human	zoledronic acid	3177	135651365	t	Inhibitor	Inhibition		Not Determined	f	t		pKi		10.14999962		Ki	0.07		#ERROR!		Inhibition of human recombinant FPPS expressed in			18327899
farnesyl diphosphate synthase	644	FDPS	P14324	ENSG00000160752	Human	zoledronic acid	3177	135651365	t	Inhibitor	Inhibition		Not Determined	f	t		pIC50		6.320000172		IC50	475.2999878		#ERROR!		Inhibition of human recombinant FPPS expressed in			18327899
farnesyl diphosphate synthase	644	FDPS	P14324	ENSG00000160752	Human	zoledronic acid	3177	135651365	t	Inhibitor	Inhibition		Not Determined	f	t		pIC50		8.390000343		IC50	4.099999905		#ERROR!		Inhibition of human FPP synthase expressed in Escl			18937434
fibroblast growth factor receptor 2	1809	FGFR2	P21802	ENSG00000066468	Human	palifermin	6954	178103536	t	Agonist	Agonist		Not Determined	f	t		-				-							NULL	
FKBP prolyl isomerase 1A	2609	FKBP1A	P62942	ENSG00000088832	Human	sirolimus	6031	178102653	t	Inhibitor	Inhibition		Not Determined	f	t		pKi		9.699999809		Ki	0.200000003		#ERROR!		Measuring ability to inhibit the proteins rotamase act			9857082
FKBP prolyl isomerase 1A	2609	FKBP1A	P62942	ENSG00000088832	Human	tacrolimus	6784	178103390	t	Inhibitor	Inhibition		Selective	f	t		pKi		9.399999619		Ki	0.400000006		#ERROR!					9857082
fms related receptor tyrosine kinase 1	1812	FLT1	P17948	ENSG00000102755	Human	ruxolitinib	9428	328083528	t	Inhibitor	Inhibition		Selective	f	t		pIC50		7.480000019		IC50	33		#ERROR!		Evaluated in a [³²]p-ATP incorporation			25482937
fms related receptor tyrosine kinase 1	1812	FLT1	P17948	ENSG00000102755	Human	pazopanib	5698	178102325	t	Inhibitor	Inhibition		Family selective	f	t		pIC50		8		IC50	10		#ERROR!					18620382
fms related receptor tyrosine kinase 1	1812	FLT1	P17948	ENSG00000102755	Human	tivozanib	6058	178102679	t	Inhibitor	Inhibition		Not Determined	f	t		pIC50		9.680000305		IC50	0.209999993		#ERROR!					23098265
fms related receptor tyrosine kinase 3	1807	FLT3	P36888	ENSG00000122025	Human	midostaurin	5702	178102329	t	Inhibitor	Inhibition		Not Determined	f	t		pIC50		6.28000021		IC50	528		#ERROR!					12124173
fms related receptor tyrosine kinase 3	1807	FLT3	P36888	ENSG00000122025	Human	midostaurin	5702	178102329	t	Inhibitor	Inhibition		Not Determined	f	t		pKd		7.960000038		Kd	11		#ERROR!					22037378
fms related receptor tyrosine kinase 3	1807	FLT3	P36888	ENSG00000122025	Human	sorafenib	5711	178102337	t	Inhibitor	Inhibition		Not Determined	f	t		pIC50		7.239999771		IC50	58		#ERROR!					15466206
fms related receptor tyrosine kinase 3	1807	FLT3	P36888	ENSG00000122025	Human	sorafenib	5711	178102337	t	Inhibitor	Inhibition		Not Determined	f	t		pIC50		7.480000019		IC50	33		#ERROR!					21613405
fms related receptor tyrosine kinase 4	1814	FLT4	P35916	ENSG00000037280	Human	fruquintinib	9428	328083528	t	Inhibitor	Inhibition		Selective	f	t		pIC50		9.300000191		IC50	0.5		#ERROR!		Evaluated in a [³²]p-ATP incorporation			25482937
fms related receptor tyrosine kinase 4	1814	FLT4	P35916	ENSG00000037280	Human	lenvatinib	7426	178103998	t	Inhibitor	Inhibition		Not Determined	f	t		pIC50		8.279999733		IC50	5.199999809		#ERROR!					23098265
fms related receptor tyrosine kinase 4	1814	FLT4	P35916	ENSG00000037280	Human	nintedanib	5936	178102559	t	Inhibitor	Inhibition		Not Determined	f	t		pIC50		7.899999866		IC50	13		#ERROR!					18559524
fms related receptor tyrosine kinase 4	1814	FLT4	P35916	ENSG00000037280	Human	pazopanib	5698	178102325	t	Inhibitor	Inhibition		Family selective	f	t		pIC50		7.329999924		IC50	47		#ERROR!					18620382
fms related receptor tyrosine kinase 4	1814	FLT4	P35916	ENSG00000037280	Human	sunitinib	5713	178102339	t	Inhibitor	Inhibition		Not Determined	f	t		pIC50		8.050000191		IC50	8.899999619		#ERROR!					20570526
fms related receptor tyrosine kinase 4	1814	FLT4	P35916	ENSG00000037280	Human	tivozanib	6058	178102679	t	Inhibitor	Inhibition		Not Determined	f	t		pIC50		9.619999886		IC50	0.239999995		#ERROR!					23098265
FP receptor	344	PTGFR	P43088	ENSG00000122420	Human	tafluprost	7451	187051754	t	Agonist	Agonist		Selective	f	t		-				-								15037111
FP receptor	344	PTGFR	P43088	ENSG00000122420	Human	bimatoprost	1958	135649995	t	Agonist	Full agonist		Not Determined	f	t		pIC50		5.300000191		IC50	5011		#ERROR!					12606640
FP receptor	344	PTGFR	P43088	ENSG00000122420	Human	latanoprost (free acid	1960	135650508	t	Agonist	Full agonist		Selective	f	t		pKi		8.600000381		-			#ERROR!					10634944
FP receptor	344	PTGFR	P43088	ENSG00000122420	Human	latanoprost (isopropyl	1961	135650509	t	Agonist	Full agonist		Not Determined	f	t		pKi		6.300000191		-			#ERROR!					10634944
GABA_A receptor α1 su	404	GABRA1	P14867	ENSG00000022355	Human	flumazenil	4192	178101009	t	Allosteric modula	Antagonist		Non-selective	f	t		pKi		9.100000381		Ki	0.800000012		#ERROR!		Affinity measurec benzodiazepine site			9767648
GABA_A receptor α1 su	404	GABRA1	P14867	ENSG00000022355	Human	clonazepam	6963	178103544	t	Allosteric modula	Positive		Non-selective	f	t		pKi		8.899999619		Ki	1.269999981		#ERROR!		Assay using recombinant GABA_A chan			2551039
GABA_A receptor α1 su	404	GABRA1	P14867	ENSG00000022355	Human	diazepam	3364	178100389	t	Allosteric modula	Positive		Non-selective	f	t		pKi		7.789999962		Ki	16.29999924		#ERROR!		Assay using recc benzodiazepine site			2551039
GABA_A receptor α1 su	404	GABRA1	P14867	ENSG00000022355	Human	flunitrazepam	4193	178101010	t	Allosteric modula	Positive		Non-selective	f	t		pKi		8.279999733		Ki	5.199999809		#ERROR!		Assay using recc benzodiazepine site			8632757
GABA_A receptor α1 su	404	GABRA1	P14867	ENSG00000022355	Human	triazolam	7313	178103886	t	Allosteric modula	Positive		Not Determined	f	t		pKd		9		Kd	1		#ERROR!		Binding affinity to GABA_A α1&be			18537233
GABA_A receptor α1 su	404	GABRA1	P14867	ENSG00000022355	Human	triazolam	7313	178103886	t	Allosteric modula	Positive		Not Determined	f	t		pKi		9.100000381		Ki	0.800000012		#ERROR!		Binding affinity for human recombinant GABA<sub>A			10633039
GABA_A receptor α1 su	404	GABRA1	P14867	ENSG00000022355	Human	zolpidem	4348	178101165	t	Allosteric modula	Positive		Selective	f	t		pKi	7.679999828		7.389999866	Ki	41	20.70000076		#ERROR!		Assay using recombinant GABA_A chan	871941	18813598
GABA_A receptor α2 su	405	GABRA2	P47869	ENSG00000151834	Human	flumazenil	4192	178101009	t	Allosteric modula	Antagonist		Non-selective	f	t		pKi		9.050000191		Ki	0.899999976		#ERROR!		Affinity measurec benzodiazepine site			9767648
GABA_A receptor α2 su	405	GABRA2	P47869	ENSG00000151834	Human	clonazepam	6963	178103544	t	Allosteric modula	Positive		Non-selective	f	t		pKi		8.770000458		Ki	1.700000048		#ERROR!		Assay using recombinant GABA_A chan			2551039
GABA_A receptor α2 su	405	GABRA2	P47869	ENSG00000151834	Human	diazepam	3364	178100389	t	Allosteric modula	Positive		Non-selective	f	t		pKi		7.769999981		Ki	16.89999962		#ERROR!		Assay using recc benzodiazepine site			2551039
GABA_A receptor α2 su	405	GABRA2	P47869	ENSG00000151834	Human	flunitrazepam	4193	178101010	t	Allosteric modula	Positive		Non-selective	f	t		pKi		8.279999733		Ki	5.199999809		#ERROR!		Assay using recc benzodiazepine site			8632757
GABA_A receptor α2 su	405	GABRA2	P47869	ENSG00000151834	Human	triazolam	7313	178103886	t	Allosteric modula	Positive		Not Determined	f	t		pKi		9.229999542		Ki	0.589999974		#ERROR!		Binding affinity to human recombinant GABA<sub>A			10633039
GABA_A receptor α3 su	406	GABRA3	P34903	ENSG00000011677	Human	flumazenil	4192	178101009	t	Allosteric modula	Antagonist		Non-selective	f	t		pKi		8.979999542		Ki	1.049999952		#ERROR!		Affinity measurec benzodiazepine site			9767648
GABA_A receptor α3 su	406	GABRA3	P34903	ENSG00000011677	Human	clonazepam	6963	178103544	t	Allosteric modula	Positive		Non-selective	f	t		pKi		8.699999809		Ki	2		#ERROR!		Assay using recombinant GABA_A chan			2551039
GABA_A receptor α3 su	406	GABRA3	P34903	ENSG00000011677	Human	diazepam	3364	178100389	t	Allosteric modula	Positive		Non-selective	f	t		pKi		7.769999981		Ki	17		#ERROR!		Assay using recc benzodiazepine site			2551039
GABA_A receptor α3 su	406	GABRA3	P34903	ENSG00000011677	Human	flunitrazepam	4193	178101010	t	Allosteric modula	Positive		Non-selective	f	t		pKi		7.800000191		Ki	15.69999981		#ERROR!		Assay using recc benzodiazepine site			8632757
GABA_A receptor α3 su	406	GABRA3	P34903	ENSG00000011677	Human	triazolam	7313	178103886	t	Allosteric modula	Positive		Not Determined	f	t		pKi		8.840000153		Ki	1.429999948		#ERROR!		Binding affinity to human recombinant GABA<sub>A			10633039
GABA_A receptor α5 su	408	GABRA5	P31644	ENSG00000186297	Human	flumazenil	4192	178101009	t	Allosteric modula	Antagonist		Non-selective	f	t		pKi		9.220000267		Ki	0.600000024		#ERROR!		Affinity measurec benzodiazepine site			9767648
GABA_A receptor α5 su	408	GABRA5	P31644	ENSG00000186297	Human	flunitrazepam	4193	178101010	t	Allosteric modula	Positive		Non-selective	f	t		pKi		8.260000229		Ki	5.5		#ERROR!		Assay using recc benzodiazepine site			8632757
GABA_A receptor α5 su	408	GABRA5	P31644	ENSG00000186297	Human	triazolam	7313	178103886	t	Allosteric modula	Positive		Not Determined	f	t		pKd		8.399999619		Kd	4		#ERROR!		Binding affinity to GABA_A receptor &alp			18537233
GABA_A receptor α5 su	408	GABRA5	P31644	ENSG00000186297	Human	triazolam	7313	178103886	t	Allosteric modula	Positive		Not Determined	f	t		pKi		8.81000042		Ki	1.539999962		#ERROR!		Binding affinity for human recombinant GABA<sub>A			10633039
GABA_B receptor	242	KCTD16|KCTD8	Q9UBS5|Q7589	ENSG00000232632|ENS(	Human	baclofen	1084	135649976	t	Agonist	Full agonist		Not Determined	f	t		pKi	6.199999809		4.300000191	-				#ERROR!				11082110|12663046
γ-Glutamyl carboxylase	1268	GGCX	P38435	ENSG00000115486	Human	anisindione	6960	178103541	t	Inhibitor	Inhibition		Not Determined	f	t		-				-							NULL	
GAT1	929	SLC6A1	P30531	ENSG00000157103	Human	tiagabine	4818	178101520	t	Inhibitor	Inhibition		Selective	f	t		pIC50		7.199999809		IC50	70		#ERROR!					7851497
ghrelin receptor	246	GHSR	Q92847	ENSG00000121853	Human	macimorelin	9745	354702218	t	Agonist	Agonist		Selective	f	t		pIC50		7.809999943		IC50	15.60000038		#ERROR!		In a radioligand displacement assay using <sup>125			12240910
GHRH receptor	247	GHRHR	Q02643	ENSG00000106128	Human	sermorelin	6998	178103577	t	Agonist	Agonist		Not Determined	f	t		-				-							NULL	
GHRH receptor	247	GHRHR	Q02643	ENSG00000106128	Human	tesamorelin	6959	178103540	t	Agonist	Agonist		Selective	f	t		-				-							NULL	
GLP-1 receptor	249	GLP1R	P43220	ENSG00000112164	Human	exendin-4	1135	135652150	t	Agonist	Agonist		Selective	f	t		pIC50		9.180000305		IC50	0.660000026		#ERROR!		Displacement of GLP-1 from hGLP-1 receptor expre			18412318
GLP-1 receptor	249	GLP1R	P43220	ENSG00000112164	Human	lixisenatide	7387	178103959	t	Agonist	Agonist		Not Determined	f	t		pKi		8.880000114		Ki	1.330000043		#ERROR!		CHO-K1 cells overexpressing the human GLP-1 rec			20570597
GLP-1 receptor	249	GLP1R	P43220	ENSG00000112164	Human	semaglutide	9724	354702201	t	Agonist	Agonist		Selective	f	t		pIC50		9.420000076		IC50	0.379999995		#ERROR!		Membrane radioligand displacement assay <sup>1			26308095
GLP-1 receptor	249	GLP1R	P43220	ENSG00000112164	Human	semaglutide	9724	354702201	t	Agonist	Agonist		Selective	f	t		pEC50		11.21000004		EC50	0.0062		#ERROR!		In vitro potency assessed in BHK cells expres			26308095
GLP-1 receptor	249	GLP1R	P43220	ENSG00000112164	Human	exendin-4	1135	135652150	t	Agonist	Full agonist		Selective	f	t		pKi	9		8.699999809	Ki	1.99000001		1	#ERROR!				15526268
GLP-1 receptor	249	GLP1R	P43220	ENSG00000112164	Human	lixaglutide	1133	135652269	t	Agonist	Full agonist		Not Determined	f	t		pEC50		10.19999981		-			#ERROR!					10794683
GLP-2 receptor	250	GLP2R	O95838	ENSG00000065325	Human	teduglutide	7049	178103628	t	Agonist	Agonist		Not Determined	f	t		-				-			#ERROR!					25859983
Glucocorticoid receptor	625	NR3C1	P04150	ENSG00000113580	Human	budesonide	7434	178104006	t	Agonist	Agonist		Selective	f	t		pEC50		7.909999847		EC50	12.39999962		#ERROR!					21880489
Glucocorticoid receptor	625	NR3C1	P04150	ENSG00000113580	Human	budesonide	7434	178104006	t	Agonist	Agonist		Non-selective	f	t		pEC50		7.420000076		EC50	38		#ERROR!					28937774
Glucocorticoid receptor	625	NR3C1	P04150	ENSG00000113580	Human	ciclesonide	7469	187051772	t	Agonist	Agonist		Selective	f	t		pKi		7.429999828		Ki	37		#ERROR!					15878996
Glucocorticoid receptor	625	NR3C1	P04150	ENSG00000113580	Human	clobetasol propionate	7062	178103641	t	Agonist	Agonist		Selective	f	t		pKi		9.180000305		Ki	0.662		#ERROR!				NULL	
Glucocorticoid receptor	625	NR3C1	P04150	ENSG00000113580	Human	desoximetasone	7067	178103646	t	Agonist	Agonist		Selective	f	t		pKi		8.930000305		Ki	1.169999957		#ERROR!				NULL	
Glucocorticoid receptor	625	NR3C1	P04150	ENSG00000113580	Human	dexamethasone	2768	135651471	t	Agonist	Agonist		Non-selective	f	t		pIC50	9		8.100000381	-				#ERROR!				8282004|10611474|10747884
Glucocorticoid receptor	625	NR3C1	P04150	ENSG00000113580	Human	diflorasone diacetate	7068	178103647	t	Agonist	Agonist		Selective	f	t		pKi		8.539999962		Ki	2.910000086		#ERROR!				NULL	
Glucocorticoid receptor	625	NR3C1	P04150	ENSG00000113580	Human	difluprednate	7474	187051777	t	Agonist	Agonist		Selective	f	t		-				-								21182429
Glucocorticoid receptor	625	NR3C1	P04150	ENSG00000113580	Human	flunisolide	7076	178103655	t	Agonist	Agonist		Selective	f	t		pKi		8.619999886		Ki	2.420000076		#ERROR!				NULL	
Glucocorticoid receptor	625	NR3C1	P04150	ENSG00000113580	Human	fluocinolone acetonid	7077	178103656	t	Agonist	Agonist		Selective	f	t		pIC50		8.5		IC50	3.130000114		#ERROR!				NULL	
Glucocorticoid receptor	625	NR3C1	P04150	ENSG00000113580	Human	fluocinonide	7078	178103657	t

target	target_id	target_gene	target_uniprot	target_ensembl_gene_id	target_species	ligand	ligand_id	ligand_pubchen	approved_drug	type	action	action_commen	selectivity	endogenous	primary_target	concentration_r	affinity_units	affinity_high	affinity_median	affinity_low	original_affinity	original_affinity	original_affinity	original_affinity	original_affinity	assay_descripti	receptor_site	ligand_context	pubmed_id
Glucocorticoid receptor	625	NR3C1	P04150	ENSG00000113580	Human	mometasone	9029	315661114	t	Agonist	Agonist		Not Determined	f	t		-			-					#ERROR!				NULL
Glucocorticoid receptor	625	NR3C1	P04150	ENSG00000113580	Human	triamcinolone	2870	135651276	t	Agonist	Agonist		Selective	f	t		pIC50		7.699999809		-					#ERROR!				10747884
Glucocorticoid receptor	625	NR3C1	P04150	ENSG00000113580	Human	beclometasone	7059	178103638	t	Agonist	Full agonist		Selective	f	t		pKi		8.430000305		Ki		3.74000001			#ERROR!				NULL
Glucocorticoid receptor	625	NR3C1	P04150	ENSG00000113580	Human	betamethasone	7061	178103640	t	Agonist	Full agonist		Selective	f	t		pIC50		8.109999657		IC50		7.739999771			#ERROR!				NULL
Glucocorticoid receptor	625	NR3C1	P04150	ENSG00000113580	Human	prednisolone	2866	135650850	t	Agonist	Full agonist		Not Determined	f	t		pIC50		8.149999619		IC50		7			#ERROR!	Binding to human GR in vitro.			29424542
Glucocorticoid receptor	625	NR3C1	P04150	ENSG00000113580	Human	prednisolone	2866	135650850	t	Agonist	Full agonist		Not Determined	f	t		pKi		8.619999886		Ki		2.400000095			#ERROR!				11520196
Glucocorticoid receptor	625	NR3C1	P04150	ENSG00000113580	Human	prednisolone	2866	135650850	t	Agonist	Full agonist		Selective	f	t		pIC50		7.800000191		IC50		15.80000019			#ERROR!				10747884
Glucocorticoid receptor	625	NR3C1	P04150	ENSG00000113580	Human	prednisone	7096	178103674	t	Agonist	Full agonist		Selective	f	t		pKi		6.28000021		Ki		521			#ERROR!				NULL
Glucocorticoid receptor	625	NR3C1	P04150	ENSG00000113580	Human	mifepristone	2805	135650608	t	Antagonist	Antagonist		Non-selective	f	t		pKd		9.399999619		-					#ERROR!				8282004\3560943
glucosidase alpha, acid	2611	GAA	P10253	ENSG00000171298	Human	miglitol	4842	178101544	t	Inhibitor	Inhibition		Not Determined	f	t		pIC50		6.460000038		IC50		350			#ERROR!				18595718
GluN2A	456	GRIN2A	Q12879	ENSG00000183454	Human	amantadine	4128	178100950	t	Channel blocker	None	GluN2C > GluN2	Non-selective	f	t		-			-					#ERROR!				NULL
GluN2A	456	GRIN2A	Q12879	ENSG00000183454	Human	ketamine	4233	178101050	t	Channel blocker	None		Non-selective	f	t		-			-					#ERROR!				NULL
GluN2B	457	GRIN2B	Q13224	ENSG00000150086	Human	amantadine	4128	178100950	t	Channel blocker	None	GluN2C > GluN2	Non-selective	f	t		-			-					#ERROR!				NULL
GluN2B	457	GRIN2B	Q13224	ENSG00000150086	Human	ketamine	4233	178101050	t	Channel blocker	None		Non-selective	f	t		-			-					#ERROR!				NULL
GluN2C	458	GRIN2C	Q14957	ENSG00000161509	Human	amantadine	4128	178100950	t	Channel blocker	None	GluN2C > GluN2	Non-selective	f	t		pIC50		4.650000095		IC50		22500			#ERROR!	Measured vs. exogenously expressed zeta1/a			17303642
GluN2C	458	GRIN2C	Q14957	ENSG00000161509	Human	ketamine	4233	178101050	t	Channel blocker	None		Non-selective	f	t		pIC50		6.179999828		IC50		660			#ERROR!	Measured vs. exogenously expressed zeta1/a			17303642
GluN2D	459	GRIN2D	O15399	ENSG00000105464	Human	amantadine	4128	178100950	t	Channel blocker	None	GluN2C > GluN2	Non-selective	f	t		-			-					#ERROR!				NULL
GluN2D	459	GRIN2D	O15399	ENSG00000105464	Human	ketamine	4233	178101050	t	Channel blocker	None		Non-selective	f	t		-			-					#ERROR!				NULL
glutathione-disulfide reductase	2613	GSR	P00390	ENSG00000104687	Human	oxiglutatione	6835	178103441	t	Inhibitor	Inhibition		Not Determined	f	t		-			-					#ERROR!				NULL
GnRH ₁ receptor	256	GNRHR	P30968	ENSG00000109163	Human	goserelin	3879	178100715	t	Agonist	Full agonist		Selective	f	t		-			-					#ERROR!				NULL
GnRH ₁ receptor	256	GNRHR	P30968	ENSG00000109163	Human	leuprolide	1175	135652260	t	Agonist	Full agonist		Selective	f	t		pKi	9.100000381		8.5	Ki	3.160000086		0.790000021		#ERROR!				17095587
GnRH ₁ receptor	256	GNRHR	P30968	ENSG00000109163	Human	nafarelin	3902	178100736	t	Agonist	Full agonist		Selective	f	t		-			-					#ERROR!				NULL
GnRH ₁ receptor	256	GNRHR	P30968	ENSG00000109163	Human	abarelix	1188	135651965	t	Antagonist	Antagonist		Selective	f	t		pKi	9.489999771		9.100000381	Ki	0.790000021		0.319999993		#ERROR!				17095587
GnRH ₁ receptor	256	GNRHR	P30968	ENSG00000109163	Human	cetrorelix	1190	135652056	t	Antagonist	Antagonist		Not Determined	f	t		pIC50		8.699999809		IC50		1.980000019			#ERROR!				11861490
GnRH ₁ receptor	256	GNRHR	P30968	ENSG00000109163	Human	cetrorelix	1190	135652055	t	Antagonist	Antagonist		Selective	f	t		pKi	10		9.300000191	-					#ERROR!				17095587|9300077|7649152
GnRH ₁ receptor	256	GNRHR	P30968	ENSG00000109163	Human	degarelix	5585	178102220	t	Antagonist	Antagonist		Selective	f	t		pKi		8.770000458		Ki		1.700000048			#ERROR!				21188095
GnRH ₁ receptor	256	GNRHR	P30968	ENSG00000109163	Human	elagolix	8362	252166574	t	Antagonist	Antagonist		Not Determined	f	t		pKi		9.050000191		Ki		0.899999976			#ERROR!				19006286
Granulocyte colony-stimulating factor rece	1719	CSF3R	Q99062	ENSG00000119535	Human	pegfilgrastim	6969	178103549	t	Agonist	Agonist		Not Determined	f	t		-			-					#ERROR!				NULL
Growth hormone receptor	1720	GHR	P10912	ENSG00000112964	Human	pegvisomant	7485	178051798	t	Antagonist	Antagonist		Not Determined	f	t		-			-					#ERROR!				10770982
Guanylyl cyclase, α<sub>1</sub&gt	1287	GUCY1A1|GUCY	Q02108|Q02153	ENSG00000061918|ENS(	Human	NO	2509	178100228	t	Activator	Activation		Selective	f	t		-			-					#ERROR!				NULL
Guanylyl cyclase-C	1750	GUCY2C	P25092	ENSG00000070019	Human	linaclotide	5017	178101714	t	Agonist	Agonist		Selective	f	t		pKi		8.909999847		Ki		1.24000001			#ERROR!	Competitive-inhibition of [¹²⁵I]-STa bin	17694454|20863829		
histone deacetylase 1	2658	HDAC1	Q13547	ENSG00000116478	Human	romidepsin	7006	178103585	t	Inhibitor	Inhibition		Not Determined	f	t		pKi		11.81999969		Ki		0.0015			#ERROR!				20139990
histone deacetylase 1	2658	HDAC1	Q13547	ENSG00000116478	Human	tucidinostat	8305	25216651 7	t	Inhibitor	Inhibition		Not Determined	f	t		-			-					#ERROR!				20060381
histone deacetylase 1	2658	HDAC1	Q13547	ENSG00000116478	Human	vorinostat	6852	178103458	t	Inhibitor	Inhibition		Not Determined	f	t		pKi		8.890000343		Ki		1.299999952			#ERROR!				20139990
H₁ receptor	262	HRH1	P35367	ENSG00000196639	Human	amitriptyline	200	135649939	t	Antagonist	Antagonist		Not Determined	f	t		pKi		9.300000191		Ki		0.5			#ERROR!	Antagonism of [³H]mepyramine bindin			16782354
H₁ receptor	262	HRH1	P35367	ENSG00000196639	Human	astemizole	2603	135649968	t	Antagonist	Antagonist		Selective	f	t		pKi		8.520000458		Ki		3			#ERROR!				12747773
H₁ receptor	262	HRH1	P35367	ENSG00000196639	Human	azelastine	7121	178103698	t	Antagonist	Antagonist		Selective	f	t		pKi		8.899999619		Ki		1.25999999			#ERROR!				21381763
H₁ receptor	262	HRH1	P35367	ENSG00000196639	Human	chlorpheniramine	6976	178103555	t	Antagonist	Antagonist		Selective	f	t		pKi		8.149999619		Ki		7			#ERROR!				21381763
H₁ receptor	262	HRH1	P35367	ENSG00000196639	Human	clemastine	6063	178102684	t	Antagonist	Antagonist		Selective	f	t		pKi		10.31000042		Ki		0.048999999			#ERROR!				NULL
H₁ receptor	262	HRH1	P35367	ENSG00000196639	Human	cyclizine	7151	178103726	t	Antagonist	Antagonist		Selective	f	t		pKi		8.350000381		Ki		4.440000057			#ERROR!				NULL
H₁ receptor	262	HRH1	P35367	ENSG00000196639	Human	cyproheptadine	277	135650152	t	Antagonist	Antagonist		Not Determined	f	t		pKi		10.19999981		-					#ERROR!				7925364
H₁ receptor	262	HRH1	P35367	ENSG00000196639	Human	desloratadine	7157	178103732	t	Antagonist	Antagonist		Selective	f	t		pKi		9.010000229		Ki		0.970000029			#ERROR!				15482930
H₁ receptor	262	HRH1	P35367	ENSG00000196639	Human	diphenhydramine	1224	135650192	t	Antagonist	Antagonist		Not Determined	f	t		pKi		7.900000095		-					#ERROR!				12065734
H₁ receptor	262	HRH1	P35367	ENSG00000196639	Human	epinastine	7176	178103751	t	Antagonist	Antagonist		Not Determined	f	t		pIC50		7.599999905		IC50		25			#ERROR!	Antagonism of 5 nM [³H]mepyramine b			18446005
H₁ receptor	262	HRH1	P35367	ENSG00000196639	Human	fexofenadine	4819	178101521	t	Antagonist	Antagonist		Selective	f	t		pKi		7.570000172		Ki		27			#ERROR!				19660947
H₁ receptor	262	HRH1	P35367	ENSG00000196639	Human	hydroxyzine	7199	178103774	t	Antagonist	Antagonist		Not Determined	f	t		pKi		8.699999809		Ki		1.99000001			#ERROR!				11809864
H₁ receptor	262	HRH1	P35367	ENSG00000196639	Human	ketotifen	7206	178103781	t	Antagonist	Antagonist		Not Determined	f	t		pKi		8.600000381		Ki		2.5			#ERROR!				21470866
H₁ receptor	262	HRH1	P35367	ENSG00000196639	Human	loratadine	7216	135653791	t	Antagonist	Antagonist		Selective	f	t		pKi		7.426999828		Ki		37			#ERROR!				15771458
H₁ receptor	262	HRH1	P35367	ENSG00000196639	Human	mizolastine	10102	381118812	t	Antagonist	Antagonist		Selective	f	t		pKi		8.569999695		Ki		2.700000048			#ERROR!	Radioligand displacement membrane binding assay			19553115
H₁ receptor	262	HRH1	P35367	ENSG00000196639	Human	promethazine	7282	178103856	t	Antagonist	Antagonist		Not Determined	f	t		pKi		9.619999886		Ki		0.239999995			#ERROR!				16782354
H₁ receptor	262	HRH1	P35367	ENSG00000196639	Human	rupatadine	10103	381118813	t	Antagonist	Antagonist		Selective	f	t		pIC50		8.409999847		IC50		3.900000095			#ERROR!				16220969
H₁ receptor	262	HRH1	P35367	ENSG00000196639	Human	terfenadine	2608	135651245	t	Antagonist	Antagonist		Selective	f	t		pKi		7.400000095		Ki		40			#ERROR!				19660947
H₁ receptor	262	HRH1	P35367	ENSG00000196639	Human	tripelennamine	7318	178103891	t	Antagonist	Antagonist		Selective	f	t		pIC50		7.400000095		IC50		39.79999924			#ERROR!				14667234
H₁ receptor	262	HRH1	P35367	ENSG00000196639	Human	tripelennamine	7318	178103891	t	Antagonist	Antagonist		Not Determined	f	t		pKd		8		Kd		10			#ERROR!				14667234
H₁ receptor	262	HRH1	P35367	ENSG00000196639	Human	triprolidine	1228	135651281	t	Antagonist	Antagonist		Selective	f	t		pKi	9		8.5	Ki	3.160000086		1		#ERROR!				7925364|12065734
H₁ receptor	262	HRH1	P35367	ENSG00000196639	Human	cetirizine	1222	135650064	t	Antagonist	Inverse agonist		Not Determined	f	t		pKi		8.199999809		-					#ERROR!				7925364
H₁ receptor	262	HRH1	P35367	ENSG00000196639	Human	(R)-cetirizine	1214	135650416	t	Antagonist	Inverse agonist		Not Determined	f	t		pKi		8.5		-					#ERROR!				15206929
H₁ receptor	262	HRH1	P35367	ENSG00000196639	Human	mepyramine	1227	135650867	t	Antagonist	Inverse agonist		Not Determined	f	t		pKi	9		8.699999809	-					#ERROR!				12065734|15206929
H₂ receptor	263	HRH2	P25021	ENSG00000113749	Human	cimetidine	1231	135650099	t	Antagonist	Antagonist		Selective	f	t		pKi		6.849999905		Ki		140			#ERROR!				23466604
H₂ receptor	263	HRH2	P25021	ENSG00000113749	Human	famotidine	7074	178103653	t	Antagonist	Antagonist		Selective	f	t		-			-					#ERROR!				NULL
H₂ receptor	263	HRH2	P25021	ENSG00000113749	Human	ranitidine	1234	135650887	t	Antagonist	Antagonist		Selective	f	t		pKi		7.099999905		Ki		79.40000153			#ERROR!				7921611
H₃ receptor	264	HRH3	Q9Y5N1	ENSG00000101180	Human	pitolisant	8924	310264706	t	Antagonist	Antagonist		Selective	f	t		pKi	8.569999695		8.06000042	Ki	2.700000048				#ERROR!	Measured in a [¹²⁵I]iodoproxyfan bindi	26084539|19329325		
hydroxymethylglutaryl-CoA reductase	639	HMGCR	P04035	ENSG00000113161	Human	atorvastatin	2949	135649972	t	Inhibitor	Competitive		Not Determined	f	t		pIC50		8.100000381		IC50		8			#ERROR!				11349148
hydroxymethylglutaryl-CoA reductase	639	HMGCR	P04035	ENSG00000113161	Human	atorvastatin	2949	135649972	t	Inhibitor	Competitive		Not Determined	f	t		pKi		7.849999905		Ki		14			#ERROR!	In vitro inhibition of HMG-CoA reductase			16128575
hydroxymethylglutaryl-CoA reductase	639	HMGCR	P04035	ENSG00000113161	Human	cerivastatin	2950	135650063	t	Inhibitor	Competitive		Not Determined	f	t		pKi		8.239999771		Ki		5.699999809			#ERROR!	In vitro inhibition of HMG-CoA reductase			16128575
hydroxymethylglutaryl-CoA reductase	639	HMGCR	P04035	ENSG00000113161	Human	cerivastatin	2950	135650063	t	Inhibitor	Competitive		Not Determined	f	t		pIC50		8		IC50		10			#ERROR!	Inhibition of HMG-CoA reductase			15686906
hydroxymethylglutaryl-CoA reductase	639	HMGCR	P04035	ENSG00000113161	Human	fluvastatin	2951	135650279	t	Inhibitor	Competitive		Not Determined	f	t		pIC50		7.550000191		IC50		28			#ERROR!	Inhibition of HMG-CoA reductase			11349148
hydroxymethylglutaryl-CoA reductase	639	HMGCR	P04035	ENSG00000113161	Human	fluvastatin	2951	135650279	t	Inhibitor	Competitive		Not Determined	f	t		pKi		7.550000191		Ki		28			#ERROR!				12773150
hydroxymethylglutaryl-CoA reductase	639	HMGCR	P04035	ENSG00000113161	Human	fluvastatin	2951	135650279	t	Inhibitor	Competitive		Not Determined	f	t		pKi		6.559999943		Ki		256			#ERROR!	In vitro inhibition of HMG-CoA reductase			16128575
hydroxymethylglutaryl-CoA reductase	639	HMGCR	P04035	ENSG00000113161	Human	lovastatin	2739	135652701	t	Inhibitor	Competitive		Not Determined	f	t		pIC50		7.690000057		IC50		20			#ERROR!	in vitro inhibition of HMG-CoA reductase			1597859
hydroxymethylglutaryl-CoA reductase	639	HMGCR	P04035	ENSG00000113161	Human	lovastatin	2739	135652701	t	Inhibitor	Competitive		Not Determined	f	t		pIC50		7.300000191		IC50		50			#ERROR!	in vitro inhibition of HMG-CoA reductase			2153213
hydroxymethylglutaryl-CoA reductase	639	HMGCR	P04035	ENSG00000113161	Human	lovastatin	2739	135652701	t	Inhibitor	Competitive		Not Determined	f	t		pKi		9.220000267		Ki		0.600000024			#ERROR!	Inhibition of HMG-CoA reductase			6933445
hydroxymethylglutaryl-CoA reductase	639	HMGCR	P04035	ENSG00000113161	Human	pravastatin	2953	135650848	t	Inhibitor	Competitive		Not Determined	f	t		pIC50		7.519999981		IC50		30			#ERROR!	In vitro inhibition of HMG-CoA reductase			1597859
hydroxymethylglutaryl-CoA reductase	639	HMGCR	P04035	ENSG00000113161	Human	pravastatin	2953	135650848	t	Inhibitor	Competitive		Not Determined	f	t		pIC50		5.860000134		IC50		1370			#ERROR!	In vitro inhibition of HMG-CoA reductase			11392538
hydroxymethylglutaryl-CoA reductase	639	HMGCR	P04035	ENSG00000113161	Human	pravastatin	2953	135650848	t	Inhibitor	Competitive		Not Determined	f	t		pKi		6.989999771		Ki		103			#ERROR!	In vitro inhibition of HMG-CoA reductase			16128575
hydroxymethylglutaryl-CoA reductase	639	HMGCR	P04035	ENSG00000113161	Human	rosuvastatin	2954	135650928	t	Inhibitor	Competitive		Not Determined	f	t		pIC50		8.300000191		IC50		5			#ERROR!				11349148
hydroxymethylglutaryl-CoA reductase	639	HMGCR	P04035	ENSG00000113161	Human	rosuvastatin	2954	135650928	t	Inhibitor	Competitive		Not Determined	f	t		pIC50		8.510000229		IC50		3.099999905			#ERROR!	Inhibition of HMG-CoA reductase			17574412
hydroxymethylglutaryl-CoA reductase	639	HMGCR	P04035	ENSG00000113161	Human	rosuvastatin	2954	135650928	t	Inhibitor	Competitive		Not Determined	f	t		pKi		8.640000343		Ki		2.299999952			#ERROR!				16128575
hydroxymethylglutaryl-CoA reductase	639	HMGCR	P04035	ENSG00000113161	Human	rosuvastatin	2954	135650928	t	Inhibitor	Competitive		Not Determined	f	t		pIC50		8.270000458		IC50		5.400000095			#ERROR!	in vitro inhibition of cholesterol synthesis			11256847
hydroxymethylglutaryl-CoA reductase	639	HMGCR	P04035	ENSG00000113161	Human	simvastatin	2955	135651045	t	Inhibitor	Competitive		Not Determined	f	t		pIC50		7.960000038		IC50		11			#ERROR!				11349148
hydroxymethylglutaryl-CoA reductase	639	HMGCR	P04035	ENSG00000113161	Human	simvastatin	2955	135651045	t	Inhibitor	Competitive		Not Determined	f	t		pIC50		7.920000076		IC50		12.10000038			#ERROR!	In vitro inhibition of HMG-CoA reductase			1433193
immunoglobulin heavy constant epsilon	2741	IGHE	P01854	ENSG00000211891	Human	omalizumab	6890	178103472	t	Antibody	Binding		Selective	f	t		pKd		7.590000153		Kd		26			#ERROR!	F(ab) apparent binding affinity of e25 measured by s			25200415
immunoglobulin heavy constant epsilon	2741	IGHE	P01854	ENSG00000211891	Human	omalizumab	6890	178103472	t	Antibody	Binding		Selective	f	t		pKd		8.199999847		Kd		6.980000019			#ERROR!	in vitro binding affinity			25200415
IMPase 1	1463	IMPA1	P29218	ENSG00000133731	Human	Li⁺	5212	178101897	t	Inhibitor	Inhibition		Not Determined	f	t		pKi		3.5		-					#ERROR!				1377913
inosine monophosphate dehydrogenase 1	2624	IMPDH1	P20839	ENSG00000106348	Human	mycophenolate mofet	6831	178103437	t	Inhibitor	Inhibition		Not Determined	f	t		-			-					#ERROR!				NULL
inosine monophosphate dehydrogenase 1	2624	IMPDH1	P20839	ENSG00000106348	Human	mycophenolic acid	6832	178103438	t	Inhibitor	Inhibition		Not Determined	f	t		pIC50		7.699999809		IC50		20			#ERROR!	Reference does not distinguish which enzyme isofor			1967654
inosine monophosphate dehydrogenase 1	2624	IMPDH1	P20839	ENSG00000106348	Human	ribavirin	6842	178103448	t	Inhibitor	Inhibition		Not Determined	f	t		pIC50	6		5.599999905	IC50	2500		1000		#ERROR!	in vitro inhibition of IMPDH (isoform not speci			22555152
inosine monophosphate dehydrogenase 1	2624	IMPDH1	P20839	ENSG00000106348	Human	thioguanine	6845	178103451	t	Inhibitor	Inhibition		Not Determined	f	t		-			-					#ERROR!				18609073|14529546
inosine monophosphate dehydrogenase 2	2625	IMPDH2	P12268	ENSG00000178035	Human	mycophenolate mofet	6831	178103437	t	Inhibitor	Inhibition	See Inhibitor Con	Not Determined	f	t		-			-					#ERROR!				NULL
inosine monophosphate dehydrogenase 2	2625	IMPDH2	P12268	ENSG00000178035	Human	mycophenolic acid	6832	178103438	t	Inhibitor	Inhibition		Not Determined	f	t		pIC50		7.699999809		IC50		20			#ERROR!	Reference does not distinguish which enzyme isofor			1967654
inosine monophosphate dehydrogenase 2	2625	IMPDH2	P12268	ENSG00000178035	Human	ribavirin	6842	178103448	t	Inhibitor	Inhibition		Not Determined	f	t		pIC50	6		5.599999905	IC50	2500		1000		#ERROR!	in vitro inhibition of IMPDH (isoform not speci			22555152
inosine monophosphate dehydrogenase 2	2625	IMPDH2	P12268	ENSG00000178035	Human	thioguanine	6845	178103451	t	Inhibitor	Inhibition		Not Determined	f	t		-			-					#ERROR!				18609073|14529546
Insulin-like growth factor I receptor	1801	IGF1R	P08069	ENSG00000140443	Human	teprotumumab	10612	404859059	t	Antibody	Antagonist		Selective	f	t		pKd		9.25		Kd		0.564999998			#ERROR!	Dissociation constant for antibody 22; determined by			NULL
Insulin-like growth factor I receptor	1801	IGF1R	P08069	ENSG00000140443	Human	teprotumumab	10612	404859059	t	Antibody	Antagonist		Selective	f	t		pKd		12.15999985		Kd		0.000695			#ERROR!	Dissociation constant for antibody 18; determined by			NULL
integrin α4β7	2770	ITGA4|ITGB7	P13612|P26010	ENSG00000115232|ENS(	Human	vedolizumab	7437	178104009	t	Antibody	Antagonist		Selective	f	t		pIC50		8.270000458		IC50		5.340000153			#ERROR!	Binding to Hut-78 cells			NULL
integrin, alpha 4 subunit (antigen CD49D,	2443	ITGA4	P13612	ENSG00000115232	Human	natalizumab	6591	178103204	t	Antibody	Inhibition		Not Determined	f	t		-			-					#ERROR!				15293871
integrin αIIbβ3	2579	ITGA2B|ITGB3	P05106|P08514	ENSG00000005961|ENS(	Human	abciximab	6584	178103197	t	Antibody	Binding		Not Determined	f	t		-			-					#ERROR!				NULL
integrin αIIbβ3	2579	ITGA2B|ITGB3	P05106|P08514	ENSG00000005961|ENS(	Human	eptifibatide	6585	178103198	t	Inhibitor	Inhibition		Not Determined	f	t		pIC50	6.849999905		6.239999771	IC50	570		140		#ERROR!	inhibition of human platelet aggregation. The IC50 v			10999999
integrin αIIbβ3	2579	ITGA2B|ITGB3	P05106|P08514	ENSG00000005961|ENS(	Human	tirofiban	6586	178103199	t	Inhibitor	Inhibition		Not Determined	f	t		pIC50		9.430000305		IC50		0.370000005			#ERROR!	Displacement of biotinylated fibrinogen from human			23644213
integrin, alpha L subunit (antigen CD11A (	2451	ITGAL	P20701	ENSG00000005844	Human	lifitegrast	7533	178051834	t	Antagonist	Inhibition		Not Determined	f	t		-			-					#ERROR!				24900456
integrin, alpha L subunit (antigen CD11A (	2451	ITGAL	P20701	ENSG00000005844	Human	efalizumab	6593	178103206	t	Antibody	Binding		Selective	f	t		pKd		11.39999962		Kd		0.004			#ERROR!	Binding of humanized MHM24 to human CD11a on	NULL		
Interferon α/β receptor	1898	IFNAR1|IFNAR2	P17181|P48551	ENSG00000142166|ENS(	Human	peginterferon alfa-2b	7462	178051765	t	Agonist	Agonist		Selective	f	t		-			-					#ERROR!				10607680
interferon α/β receptor 1	1723	IFNAR1	P17181	ENSG00000142166	Human	peginterferon alfa-2b	7462	178051765	t	Agonist	Agonist		Selective	f	t		-			-					#ERROR!				10607680
Interleukin-11 receptor	2311	IL6ST|IL11RA	P40189|Q14626	ENSG00000137070|ENS(	Human	oprelvekin	6971	178103550	t	Agonist	Full agonist		Not Determined	f	t		-			-					#ERROR!				18020592|11033834
Interleukin 17 receptor A	1738	IL17RA	Q96F46	ENSG00000177663	Human	brodalumab	7540	178051841	t	Antibody	Binding		Selective	f	t		pKd		9.159999847		Kd		0.689999998			#ERROR!				NULL
Interleukin-1 receptor, type I	1905	IL1RAP|IL1R1	Q9NPH3|P14778	ENSG00000115594|ENS(	Human	anakinra	6972	178103551	t	Inhibitor	Inhibition		Not Determined	f	t		pKd		7.820000172		Kd		15			#ERROR!				1834644
Interleukin-2 receptor subunit α	1695	IL2RA	P01589	ENSG00000134460	Human	daclizumab	6880	178103462	t	Antibody	Binding		Selective	f	t		pKd				Kd		8			#ERROR!				NULL
Interleukin-4 receptor subunit α	1697	IL4R	P24394	ENSG00000077238	Human	dupilumab	7574	223365910	t	Antibody	Binding		Selective	f	t		pIC50		11.06000042		IC50		0.0088			#ERROR!	Value derived from an ELISA-based solution compet	NULL		
Interleukin-6 receptor	2310	IL6ST|IL6R	P40189|P08887	ENSG00000160712|ENS(	Human	satralizumab	9093	315661178	t	Antibody	Binding		Selective	f	t		pKd		8.850000381		Kd		1.399999976			#ERROR!				NULL
Interleukin-6 receptor	2310	IL6ST|IL6R	P40189|P08887	ENSG00000160712|ENS(	Human	tocilizumab	6881	178103463	t	Antibody	Binding		Selective	f	t		pKd		8.600000381		Kd		2.539999962			#ERROR!				NULL
Interleukin-6 receptor, α subunit	1708	IL6R	P08887	ENSG00000160712	Human	sarilumab	7999	249565679	t	Antibody	Binding		Not Determined	f	t		pKd	11.05000019		10.64000034	Kd	0.023		0.009		#ERROR!	Dissociation equilibrium constant calculated using te			NULL
IP receptor	345	PTGIR	P43119	ENSG00000160013	Human	iloprost	1895	135650378	t	Agonist	Full agonist		Not Determined	f	t		pKi	8		7.5	-					#ERROR!				16604093|10634944
IP receptor	345	PTGIR	P43119	ENSG00000160013	Human	treprostinil	5820	178102445	t	Agonist	Full agonist		Not Determined	f	t		pEC50		8.720000267		-					#ERROR!				22480736
IP receptor	345	PTGIR	P43119	ENSG00000160013	Human	treprostinil	5820	178102445	t	Agonist	Full agonist		Not Determined	f	t		pKi		7.489999771		-					#ERROR!				22480736
isocitrate dehydrogenase (NADP(+)) 1	2884	IDH1	O75874	ENSG00000138413	Human	ivosidenib	9217	315661293	t	Inhibitor	Inhibition		Selective	f	t		-			-					#ERROR!				NULL
Janus kinase 1	2047	JAK1	P23458	ENSG00000162434	Human	baricitinib	7792	223366123	t	Inhibitor	Inhibition		Not Determined	f	t

target	target_id	target_gene	target_uniprot	target_ensembl_gene_id	target_species	ligand	ligand_id	ligand_pubchen	approved_drug	type	action	action_commen	selectivity	endogenous	primary_target	concentration_r	affinity_units	affinity_high	affinity_median	affinity_low	original_affinity	original_affinity	original_affinity	original_affinity	original_affinity	assay_descripti	receptor_site	ligand_context	pubmed_id
Janus kinase 1	2047	JAK1	P23458	ENSG00000162434	Human	filgotinib	7913	249565596	t	Inhibitor	Inhibition		Not Determined	f	t		pIC50		7.960000038		IC50			10		#ERROR!				24006460
Janus kinase 1	2047	JAK1	P23458	ENSG00000162434	Human	ruxolitinib	5688	178102315	t	Inhibitor	Inhibition		Not Determined	f	t		pIC50	10.05000019		8.479999542	IC50	3.299999952		0.090000004		#ERROR!				20130243|23061660
Janus kinase 1	2047	JAK1	P23458	ENSG00000162434	Human	ruxolitinib	5688	178102315	t	Inhibitor	Inhibition		Not Determined	f	t		pIC50		8.220000267		IC50			6		#ERROR!				24417533
Janus kinase 1	2047	JAK1	P23458	ENSG00000162434	Human	tofacitinib	5677	178102304	t	Inhibitor	Inhibition		Not Determined	f	t		pIC50		7.820000172		IC50			15		#ERROR!	In a biochemical enzyme assay.			24417533
Janus kinase 1	2047	JAK1	P23458	ENSG00000162434	Human	upadacitinib	9246	315661321	t	Inhibitor	Inhibition		Selective	f	t		pIC50		7.369999886		IC50			43		#ERROR!				NULL
Janus kinase 1	2047	JAK1	P23458	ENSG00000162434	Human	upadacitinib	9246	315661321	t	Inhibitor	Inhibition		Selective	f	t		pIC50		8.520000458		IC50			3		#ERROR!	at 1uM ATP in a biochemical assay			NULL
Janus kinase 2	2048	JAK2	O60674	ENSG00000096968	Human	baricitinib	7792	223366123	t	Inhibitor	Inhibition		Not Determined	f	t		pIC50		8.149999619		IC50			7		#ERROR!				24417533
Janus kinase 2	2048	JAK2	O60674	ENSG00000096968	Human	baricitinib	7792	223366123	t	Inhibitor	Inhibition		Not Determined	f	t		pIC50		8.239999771		IC50		5.699999809			#ERROR!				20363976
Janus kinase 2	2048	JAK2	O60674	ENSG00000096968	Human	delgocitinib	9619	340592249	t	Inhibitor	Inhibition		Non-selective	f	t		pIC50		8.590000153		IC50		2.599999905			#ERROR!				32511913
Janus kinase 2	2048	JAK2	O60674	ENSG00000096968	Human	fedratinib	5716	178102342	t	Inhibitor	Inhibition		Not Determined	f	t		pIC50		8.520000458		IC50			3		#ERROR!				21106455|18394554
Janus kinase 2	2048	JAK2	O60674	ENSG00000096968	Human	ruxolitinib	5688	178102315	t	Inhibitor	Inhibition		Not Determined	f	t		pIC50	9.619999886		8.550000191	IC50	2.799999952		0.239999995		#ERROR!				20130243|23061660
Janus kinase 2	2048	JAK2	O60674	ENSG00000096968	Human	ruxolitinib	5688	178102315	t	Inhibitor	Inhibition		Not Determined	f	t		pIC50		8.050000191		IC50			9		#ERROR!				24417533
Janus kinase 2	2048	JAK2	O60674	ENSG00000096968	Human	tofacitinib	5677	178102304	t	Inhibitor	Inhibition		Not Determined	f	t		pIC50		7.110000134		IC50			77		#ERROR!	In a biochemical enzyme assay.			24417533
Janus kinase 3	2049	JAK3	P52333	ENSG00000105639	Human	tofacitinib	5677	178102304	t	Inhibitor	Inhibition		Not Determined	f	t		pIC50		8.220000267		IC50			6		#ERROR!				27572962
Janus kinase 3	2049	JAK3	P52333	ENSG00000105639	Human	tofacitinib	5677	178102304	t	Inhibitor	Inhibition		Not Determined	f	t		pIC50		7.260000229		IC50			55		#ERROR!	In a biochemical enzyme assay.			24417533
Janus kinase 3	2049	JAK3	P52333	ENSG00000105639	Human	tofacitinib	5677	178102304	t	Inhibitor	Inhibition		Not Determined	f	t		pIC50		6.679999828		IC50		210		1	#ERROR!				21105711|14593182
kallikrein B1	2379	KLKB1	P03952	ENSG00000164344	Human	lanadelumab	9094	315661179	t	Antibody	Binding		Selective	f	t		pIC50		10.05000019		IC50			0.090000004		#ERROR!				NULL
κ receptor	318	OPRK1	P41145	ENSG00000082556	Human	nalbuphine	1663	135650680	t	Agonist	Full agonist		Not Determined	f	t		pKi	7.5		7.400000095	-					#ERROR!				9262330
κ receptor	318	OPRK1	P41145	ENSG00000082556	Human	butorphanol	7591	223366926	t	Agonist	Partial agonist		Not Determined	f	t		pKi		9.920000076		Ki			0.119999997		#ERROR!	Displacement of [³H]U69593 from hum			18674902
κ receptor	318	OPRK1	P41145	ENSG00000082556	Human	buprenorphine	1670	135652673	t	Antagonist	Antagonist		Not Determined	f	t		pKi	10.19999981		9.100000381	-					#ERROR!				9686407|9262330
κ receptor	318	OPRK1	P41145	ENSG00000082556	Human	nalmefene	1628	135650681	t	Antagonist	Antagonist		Not Determined	f	t		pKi		9.5		-					#ERROR!				9686407
κ receptor	318	OPRK1	P41145	ENSG00000082556	Human	naloxone	1638	135650684	t	Antagonist	Antagonist		Not Determined	f	t		pKi	8.600000381		7.599999905	-					#ERROR!				9686407|7869844|7624359|9262330
κ receptor	318	OPRK1	P41145	ENSG00000082556	Human	naltrexone	1639	135650687	t	Antagonist	Antagonist		Not Determined	f	t		pKi	9.399999619		8.399999619	-					#ERROR!				9686407|7624359
Kidney-specific Na-K-Cl symporter	968	SLC12A1	Q13621	ENSG00000074803	Human	bumetanide	4837	121015139	t	Inhibitor	Inhibition		Non-selective	f	t		pIC50		6.480000019		-					#ERROR!				11882915
Kidney-specific Na-K-Cl symporter	968	SLC12A1	Q13621	ENSG00000074803	Human	furosemide	4839	178101541	t	Inhibitor	Inhibition		Non-selective	f	t		pIC50		5.150000095		-					#ERROR!				11882915
kinase insert domain receptor	1813	KDR	P35968	ENSG00000128052	Human	ramucirumab	7390	178103962	t	Antibody	Antagonist		Selective	f	t		pIC50		9		IC50			1		#ERROR!				12917408
kinase insert domain receptor	1813	KDR	P35968	ENSG00000128052	Human	ramucirumab	7390	178103962	t	Antibody	Antagonist		Selective	f	t		pKd		10.30000019		Kd			0.050000001		#ERROR!				12917408
kinase insert domain receptor	1813	KDR	P35968	ENSG00000128052	Human	axitinib	5659	178102287	t	Inhibitor	Inhibition		Not Determined	f	t		pIC50		9.600000381		IC50			0.25		#ERROR!				20869793
kinase insert domain receptor	1813	KDR	P35968	ENSG00000128052	Human	axitinib	5659	178102287	t	Inhibitor	Inhibition		Not Determined	f	t		pKd		8.229999542		Kd			5.900000095		#ERROR!				20137378
kinase insert domain receptor	1813	KDR	P35968	ENSG00000128052	Human	cabozantinib	5887	178102511	t	Inhibitor	Inhibition		Not Determined	f	t		pIC50		10.46000004		IC50			0.035		#ERROR!				21926191
kinase insert domain receptor	1813	KDR	P35968	ENSG00000128052	Human	fruquintinib	9428	328083528	t	Inhibitor	Inhibition		Selective	f	t		pIC50		7.460000038		IC50			35		#ERROR!	Evaluated in a [³²p-ATP incorporation			25482937
kinase insert domain receptor	1813	KDR	P35968	ENSG00000128052	Human	fruquintinib	9428	328083528	t	Inhibitor	Inhibition		Selective	f	t		pIC50		7.599999905		IC50			25		#ERROR!	Evaluated in a Z-lyte assay.			25482937
kinase insert domain receptor	1813	KDR	P35968	ENSG00000128052	Human	lenvatinib	7426	178103998	t	Inhibitor	Inhibition		Not Determined	f	t		pIC50		8.399999619		IC50			4		#ERROR!				23098265
kinase insert domain receptor	1813	KDR	P35968	ENSG00000128052	Human	pazopanib	5698	178102325	t	Inhibitor	Inhibition		Family selective	f	t		pIC50		7.519999981		IC50			30		#ERROR!				18620382
kinase insert domain receptor	1813	KDR	P35968	ENSG00000128052	Human	pazopanib	5698	178102325	t	Inhibitor	Inhibition		Family selective	f	t		pKd		7.849999905		Kd			14		#ERROR!				22037378
kinase insert domain receptor	1813	KDR	P35968	ENSG00000128052	Human	sorafenib	5711	178102337	t	Inhibitor	Inhibition		Not Determined	f	t		pIC50		7.050000191		IC50			90		#ERROR!				15466206
kinase insert domain receptor	1813	KDR	P35968	ENSG00000128052	Human	sorafenib	5711	178102337	t	Inhibitor	Inhibition		Not Determined	f	t		pIC50		7.050000191		IC50			90		#ERROR!				21613405
kinase insert domain receptor	1813	KDR	P35968	ENSG00000128052	Human	sorafenib	5711	178102337	t	Inhibitor	Inhibition		Not Determined	f	t		pKd		7.230000019		Kd			59		#ERROR!				22037378
kinase insert domain receptor	1813	KDR	P35968	ENSG00000128052	Human	tivozanib	6058	178102679	t	Inhibitor	Inhibition		Not Determined	f	t		pIC50		9.800000191		IC50			0.159999996		#ERROR!				23098265
kinase insert domain receptor	1813	KDR	P35968	ENSG00000128052	Human	vandetanib	5717	178102343	t	Inhibitor	Inhibition		Not Determined	f	t		pIC50		8		IC50			9		#ERROR!				15743202
KIT proto-oncogene, receptor tyrosine kin	1805	KIT	P10721	ENSG00000157404	Human	masitinib	5656	178102284	t	Inhibitor	Inhibition		Not Determined	f	t		pIC50		6.699999809		IC50			200		#ERROR!				19789626
KIT proto-oncogene, receptor tyrosine kin	1805	KIT	P10721	ENSG00000157404	Human	masitinib	5656	178102284	t	Inhibitor	Inhibition		Not Determined	f	t		pKd		8.090000153		Kd			8.100000381		#ERROR!				22037378
KIT proto-oncogene, receptor tyrosine kin	1805	KIT	P10721	ENSG00000157404	Human	pexidartinib	8710	25228768	t	Inhibitor	Inhibition		Not Determined	f	t		pIC50		7.570000172		IC50			27	<		#ERROR!				26222558
KIT proto-oncogene, receptor tyrosine kin	1805	KIT	P10721	ENSG00000157404	Human	sorafenib	5711	178102337	t	Inhibitor	Inhibition		Not Determined	f	t		pIC50		7.170000076		IC50			68		#ERROR!				15466206
KIT proto-oncogene, receptor tyrosine kin	1805	KIT	P10721	ENSG00000157404	Human	sorafenib	5711	178102337	t	Inhibitor	Inhibition		Not Determined	f	t		pIC50		7.170000076		IC50			68		#ERROR!				21613405
KIT proto-oncogene, receptor tyrosine kin	1805	KIT	P10721	ENSG00000157404	Human	sunitinib	5713	178102339	t	Inhibitor	Inhibition		Not Determined	f	t		pKd		9.430000305		Kd			0.370000005		#ERROR!				22037378
KIT proto-oncogene, receptor tyrosine kin	1805	KIT	P10721	ENSG00000157404	Human	sunitinib	5713	178102339	t	Inhibitor	Inhibition		Not Determined	f	t		pIC50		7.880000114		IC50			13.10000038		#ERROR!				20570526
K_{2P}10.1	521	KCNK10	P57789	ENSG00000100433	Human	halothane	2401	135650346	t	Activator	None	studied at 1-5 mM	Non-selective	f	t		-				-					#ERROR!				10880510
K_{2P}9.1	520	KCNK9	Q9NPC2	ENSG00000169427	Human	halothane	2401	135650346	t	Activator	None	studied at 1-5 mM	Non-selective	f	t		-				-					#ERROR!				11886861
K_{Ir}6.2	442	KCNJ11	Q14654	ENSG00000187486	Human	glibenclamide	2414	135650306	t	Inhibitor	None		Non-selective	f	t		-				-					#ERROR!			NULL	
K_{Ir}6.2	442	KCNJ11	Q14654	ENSG00000187486	Human	tolbutamide	6848	178103454	t	Inhibitor	None		Non-selective	f	t		-				-					#ERROR!			NULL	
K_v1.1	572	KCNH2	Q12809	ENSG00000055118	Human	dofetilide	2604	135650201	t	Channel blocker	Inhibition		Selective	f	t		pKi		8.189999568		Ki			6.400000095		#ERROR!				17536794
K_v11.1	572	KCNH2	Q12809	ENSG00000055118	Human	ibutilide	7200	178103775	t	Channel blocker	None		Selective	f	t		pIC50	8		7.599999905	IC50	25.10000038		10		#ERROR!				12873512|15266014
K_v1.8	545	KCNA10	Q16322	ENSG00000143105	Human	fampridine	2416	135649796	t	Channel blocker	None		Non-selective	f	t		pIC50		2.799999952		IC50			1580000		#ERROR!				10836990
K_v2.2	547	KCNB2	Q92953	ENSG00000182674	Human	fampridine	2416	135649796	t	Channel blocker	None		Non-selective	f	t		pIC50		2.799999952		IC50			1500000		#ERROR!				9612272
K_v4.1	552	KCND1	Q9NSA2	ENSG00000102057	Human	fampridine	2416	135649796	t	Channel blocker	None		Non-selective	f	t		pIC50		2		-					#ERROR!				10729221
K_v7.2	561	KCNQ2	O43526	ENSG00000075043	Human	retigabine	2601	135650896	t	Activator	None		Non-selective	f	t		pEC50		5.599999905		EC50			2500		#ERROR!				11466425
K_v7.3	562	KCNQ3	O43525	ENSG00000184156	Human	retigabine	2601	135650896	t	Activator	None		Non-selective	f	t		pEC50		6.199999809		EC50			631		#ERROR!				11466425
K_v7.4	563	KCNQ4	P56996	ENSG00000117013	Human	retigabine	2601	135650896	t	Activator	None		Non-selective	f	t		pEC50		5.199999809		EC50			6310		#ERROR!				11466425
K_v7.5	564	KCNQ5	Q9NR82	ENSG00000185760	Human	retigabine	2601	135650896	t	Activator	None		Non-selective	f	t		pEC50		5		EC50			10000		#ERROR!				11890000
L-Aromatic amino-acid decarboxylase	1271	DDC	P20711	ENSG00000132437	Human	carbidopa	5159	178101851	t	Inhibitor	Inhibition		Selective	f	t		-				-					#ERROR!			NULL	
L-Phenylalanine hydroxylase	1240	PAH	P00439	ENSG00000171759	Human	sapropterin	5276	178101959	t	Activator	Activation		Selective	f	t		pKi		5.400000095		Ki			4000		#ERROR!	Binding affinity for human PAH.			15537351
L-Tyrosine hydroxylase	1243	TH	P07101	ENSG00000180176	Human	alpha-methyltyrosine	6956	178103538	t	Inhibitor	Inhibition		Selective	f	t		-				-					#ERROR!			NULL	
L-Tyrosine hydroxylase	1243	TH	P07101	ENSG00000180176	Human	methyltyrosine	5094	178101787	t	Inhibitor	Inhibition		Non-selective	f	t		-				-					#ERROR!			NULL	
maltase-glucoamylase	2627	MGAM	O43451	ENSG00000257335	Human	miglitol	4842	178101544	t	Inhibitor	Inhibition		Not Determined	f	t		pKi		6		Ki			1000		#ERROR!	Inhibition of human recombinant N-terminal subunit			20970346
MC₂ receptor	283	MC2R	Q01718	ENSG00000185231	Human	corticotropin zinc hydr	6967	178103547	t	Agonist	Agonist		Selective	f	t		-				-					#ERROR!			NULL	
MC₂ receptor	283	MC2R	Q01718	ENSG00000185231	Human	cosyntropin	6965	178103545	t	Agonist	Agonist		Selective	f	t		-				-					#ERROR!			NULL	
mechanistic target of rapamycin kinase	2109	MTOR	P42345	ENSG00000198793	Human	everolimus	5889	135651512	t	Inhibitor	Inhibition		Selective	f	t		pIC50		8.699999809		IC50			2		#ERROR!				9723437
mechanistic target of rapamycin kinase	2109	MTOR	P42345	ENSG00000198793	Human	temsirolimus	5892	178102515	t	Inhibitor	Inhibition		Selective	f	t		pIC50		5.800000191		IC50			1760		#ERROR!				21438579
MET proto-oncogene, receptor tyrosine ki	1815	MET	P08581	ENSG00000105976	Human	cabozantinib	5887	178102511	t	Inhibitor	Inhibition		Not Determined	f	t		pIC50		8.899999619		IC50			1.299999952		#ERROR!				21926191
MET proto-oncogene, receptor tyrosine ki	1815	MET	P08581	ENSG00000105976	Human	capmatinib	7904	249565587	t	Inhibitor	Inhibition		Not Determined	f	t		pIC50		9.890000343		IC50			0.129999995		#ERROR!				21918175
Mineralocorticoid receptor	626	NR3C2	P08235	ENSG00000151623	Human	fludrocortisone	2873	135650268	t	Agonist	Agonist		Non-selective	f	t		pIC50		9.920000076		-					#ERROR!				8252004
Mineralocorticoid receptor	626	NR3C2	P08235	ENSG00000151623	Human	drospirenone	2874	135650211	t	Antagonist	Antagonist		Not Determined	f	t		-				-					#ERROR!				1493716
Mineralocorticoid receptor	626	NR3C2	P08235	ENSG00000151623	Human	eplerenone	2876	135650235	t	Antagonist	Antagonist		Selective	f	t		pKi		6.909999847		Ki			124		#ERROR!				18038968
Mineralocorticoid receptor	626	NR3C2	P08235	ENSG00000151623	Human	nimodipine	2523	53801057	t	Antagonist	Antagonist	inhibition of aldos	Non-selective	f	t		pIC50		6.800000191		IC50			160		#ERROR!				18250364
Mineralocorticoid receptor	626	NR3C2	P08235	ENSG00000151623	Human	spironolactone	2875	135651065	t	Antagonist	Antagonist		Not Determined	f	t		pKi		8.630000114		Ki			2.319999933		#ERROR!				18038968
mitogen-activated protein kinase kinase 1	2062	MAP2K1	Q02750	ENSG00000169032	Human	binimetinib	7921	249565604	t	Allosteric modula	Negative		Not Determined	f	t		pIC50		7.920000076		IC50			12		#ERROR!			NULL	
mitogen-activated protein kinase kinase 1	2062	MAP2K1	Q02750	ENSG00000169032	Human	cobimetinib	7626	223365961	t	Allosteric modula	Negative		Selective	f	t		pIC50		9.050000191		IC50			0.899999976		#ERROR!	Biochemical assay measuring MEK-induced phosph			24900486
mitogen-activated protein kinase kinase 1	2062	MAP2K1	Q02750	ENSG00000169032	Human	selumetinib	5665	178102293	t	Allosteric modula	Negative		Not Determined	f	t		pIC50	7.920000076		7.849999905	IC50	14.10000038		12		#ERROR!				17332304|17237274
mitogen-activated protein kinase kinase 1	2062	MAP2K1	Q02750	ENSG00000169032	Human	trametinib	6495	178103109	t	Inhibitor	Inhibition		Not Determined	f	t		pIC50	9.149999619		9.039999962	IC50	0.920000017		0.099999988		#ERROR!	Inhibition of unphosphorylated MEK1.			21245089|21523318
mitogen-activated protein kinase kinase 2	2063	MAP2K2	P36507	ENSG00000126934	Human	binimetinib	7921	249565604	t	Allosteric modula	Negative		Not Determined	f	t		pIC50		7.920000076		IC50			12		#ERROR!			NULL	
mitogen-activated protein kinase kinase 2	2063	MAP2K2	P36507	ENSG00000126934	Human	cobimetinib	7626	223365961	t	Allosteric modula	Negative		Selective	f	t		pIC50		6.699999809		IC50			199		#ERROR!				24900486
mitogen-activated protein kinase kinase 2	2063	MAP2K2	P36507	ENSG00000126934	Human	trametinib	6495	178103109	t	Inhibitor	Inhibition		Not Determined	f	t		pIC50		8.739999771		IC50			1.799999952		#ERROR!				21523318
MMP7	1631	MMP7	P09237	ENSG00000137673	Human	doxycycline	6464	178103078	t	Inhibitor	Inhibition		Non-selective	f	t		pKd		4.139999866		Kd			73000		#ERROR!				15665254
Monoamine oxidase A	2489	MAOA	P21397	ENSG00000189221	Human	moclobemide	7428	178104000	t	Inhibitor	Inhibition		Not Determined	f	t		pKi		8.300000191		Ki			5		#ERROR!				21680183
Monoamine oxidase A	2489	MAOA	P21397	ENSG00000189221	Human	tranylcypromine	5281	178101964	t	Inhibitor	Inhibition		Not Determined	f	t		pIC50		4.699999809		IC50			20000		#ERROR!				15110846
Monoamine oxidase A	2489	MAOA	P21397	ENSG00000189221	Human	phenelzine	7266	178103840	t	Inhibitor	Irreversible inhibition		Not Determined	f	t		pKi		7.329999924		Ki			47		#ERROR!				18426226
Monoamine oxidase B	2490	MAOB	P27338	ENSG00000069535	Human	rasagiline	6641	178103254	t	Inhibitor	Inhibition		Not Determined	f	t		pIC50		7.849999905		IC50			14		#ERROR!				11159700
Monoamine oxidase B	2490	MAOB	P27338	ENSG00000069535	Human	safinamide	8291	252166503	t	Inhibitor	Inhibition		Selective	f	t		pIC50		8.109999657		IC50			7.699999809		#ERROR!				24139167
Monoamine oxidase B	2490	MAOB	P27338	ENSG00000069535	Human	safinamide	8291	252166503	t	Inhibitor	Inhibition		Selective	f	t		pKi		6.349999905		Ki			450		#ERROR!				15027868
Monoamine oxidase B	2490	MAOB	P27338	ENSG00000069535	Human	selegiline	6639	178103252	t	Inhibitor	Inhibition		Not Determined	f	t		pKi	6		6.699999809	Ki		1960	970		#ERROR!				15974574|21377879
Monoamine oxidase B	2490	MAOB	P27338	ENSG00000069535	Human	tranylcypromine	5281	178101964	t	Inhibitor	Inhibition		Non-selective	f	t		pIC50		4.71999979		IC50			19000		#ERROR!				15110846
Monoamine oxidase B	2490	MAOB	P27338	ENSG00000069535	Human	phenelzine	7266	178103840	t	Inhibitor	Irreversible inhibition		Not Determined	f	t		pKi		7.820000172		Ki			15		#ERROR!				18426226
M₁ receptor	13	CHRM1	P11229	ENSG00000168539	Human	amitriptyline	200	135649939	t	Antagonist	Antagonist		Not Determined	f	t		pKi		7.829999924		Ki			14.69999981		#ERROR!				8100134
M₁ receptor	13	CHRM1	P11229	ENSG00000168539	Human	atropine	320	135652667	t	Antagonist	Antagonist		Not Determined	f	t		pKi	9.600000381		8.5	-					#ERROR!				3443095|9614217|12049493|16439611|11303071|9
M₁ receptor	13	CHRM1	P11229	ENSG00000168539	Human	biperiden	7128	178103704	t	Antagonist	Antagonist		Selective	f	t		pKd		9.319999695		Kd			0.479999989		#ERROR!				1346637
M₁ receptor	13	CHRM1	P11229	ENSG00000168539	Human	dicyclomine	355	135650178	t	Antagonist	Antagonist		Not Determined	f	t		pKi		9.079999924		Ki			0.829999983		#ERROR!			NULL	
M₁ receptor	13	CHRM1	P11229	ENSG00000168539	Human	glycopyrrolate	7459	187051762	t	Antagonist	Antagonist		Not Determined	f	t		pIC50		9.899999619		IC50			0.126000002		#ERROR!	Assay uses glycopyrronium bromide			21036043
M₁ receptor	13	CHRM1	P11229	ENSG00000168539	Human	N-methyl scopolamine	316	135650725	t	Antagonist	Antagonist		Not Determined	f	t		pKi		9.899999619		-					#ERROR!				16439611
M₁ receptor	13	CHRM1	P11229	ENSG00000168539	Human	oxybutynin	359	135650784	t	Antagonist	Antagonist		Not Determined	f	t		pKi		8.600000381		Ki			2.400000095		#ERROR!				22243489|21212494|20590605
M₁ receptor	13	CHRM1	P11229	ENSG00000168539	Human	pirenzepine	328	135650832	t	Antagonist	Antagonist		Not Determined	f	t		pKi	8.300000191		7.800000191	-					#ERROR!				12049493|2704370|7925952|1994002|9113359|204
M₁ receptor	13	CHRM1	P11229	ENSG00000168539	Human	propantheline	329	135650857	t	Antagonist	Antagonist		Not Determined	f	t		pKi		9.699999809		-					#ERROR!				12049493
M₁ receptor	13	CHRM1	P11229	ENSG00000168539	Human	scopolamine	330	135651031	t	Antagonist	Antagonist		Not Determined	f	t		pKi		9		-					#ERROR!				12049493
M₁ receptor	13	CHRM1	P11229	ENSG00000168539	Human	trihexyphenidyl	7315	178103888	t	Antagonist	Antagonist		Not Determined	f	t		pKi		8.899999886		Ki			1.350000024		#ERROR!			NULL	
M₂ receptor	14	CHRM2	P08172	ENSG00000181072	Human	bethanechol	297	135649988	t	Agonist	Full agonist		Not Determined	f	t		pKi		4		-					#ERROR!				9224827
M₂ receptor	14	CHRM2	P08172	ENSG00000181072	Human	amitriptyline	200	135649939	t	Antagonist	Antagonist		Not Determined	f	t		pKi		7.929999828		Ki			11.80000019		#ERROR!				8100134
M₂ receptor	14	CHRM2	P08172	ENSG00000181072	Human	atropine	320	135652667	t	Antagonist	Antagonist		Not Determined	f	t		pKi	9.199999809		7.800000191	-					#ERROR!				3443095|12235229|12049493|2704370|11303071|9
M₂ receptor	14	CHRM2	P08172	ENSG00000181072	Human	dexetimide	354	135650172	t	Antagonist	Antagonist		Not Determined	f	t		pKi		8.899999619		-					#ERROR!				9454790
M₂ receptor	14	CHRM2	P08172	ENSG00000181072	Human	glycopyrrolate	7459	187051762	t	Antagonist	Antagonist		Not Determined	f	t		pIC50		9.300000191		IC50			0.500999987		#ERROR!	Assay uses glycopyrronium bromide			21036043
M₂ receptor	14	CHRM2	P08172	ENSG00000181072	Human	propantheline	329	135650857	t	Antagonist	Antagonist		Not Determined	f	t		pKi		9.5		-					#ERROR!				12049493
M₂ receptor	14	CHRM2	P08172	ENSG00000181072	Human	solifenacin	7483	187051786	t	Antagonist	Antagonist		Not Determined	f	t		pIC50		6.239999771		IC50			573		#ERROR!				21524581
M₂ receptor	14	CHRM2	P08172	ENSG00000181072	Human	solifenacin	7483	187051786	t	Antagonist	Antagonist		Non-selective	f	t		pKi	7.099999905		6.900000095	-					#ERROR!				22122494|20590605
M₃ receptor	15	CHRM3	P20309	ENSG00000133019	Human	bethanechol	297	135649988	t	Agonist	Full agonist		Not Determined	f	t		pKi		4.199999809		-					#ERROR!				9224827
M₃ receptor	15	CHRM3	P20309	ENSG00000133019	Human	carbachol	298	135650055	t	Agonist	Full agonist		Not Determined	f	t		pKi	4.400000095		4	-					#ERROR!				12235229|9224827|10323594
M₃ receptor	15	CHRM3	P20309	ENSG00000133019	Human	pilocarpine	305	135650824	t	Agonist	Partial agonist		Not Determined	f	t		pKi		5.099999905		-					#ERROR!				9224827
M₃ receptor	15	CHRM3	P20309	ENSG00000133019	Human	aclidinium	7449	178104021	t	Antagonist	Antagonist		Non-selective	f	t		pIC50		9.770000458		IC50			0.170000002		#ERROR!				19653626
M₃ receptor	15	CHRM3	P20309	ENSG00000133019	Human	amitriptyline	200	135649939	t	Antagonist	Antagonist		Not Determined	f	t		pKi		7.899999886		Ki			12.80000019		#ERROR!				8100134
M₃ receptor	15	CHRM3	P20309	ENSG00000133019	Human	atropine	320	135652667	t	Antagonist	Antagonist		Not Determined	f	t		pKi	9.800000191		8.899999619	-					#ERROR!				3443095|12049493|2704370|11303071|9029489|16
M₃ receptor	15	CHRM3	P20309	ENSG00000133019	Human	clidinium

target	target_id	target_gene	target_uniprot	target_ensembl_gene_id	target_species	ligand	ligand_id	ligand_pubchem	approved_drug	type	action	action_comment	selectivity	endogenous	primary_target	concentration_r	affinity_units	affinity_high	affinity_median	affinity_low	original_affinity	original_affinity	original_affinity	original_affinity	original_affinity	assay_descripti	receptor_site	ligand_context	pubmed_id
M₃ receptor	15	CHRM3	P20309	ENSG00000133019	Human	darifenacin	321	135650160	t	Antagonist	Antagonist		Non-selective	f	t		pKi	9.100000381		8.600000381	-				#ERROR!				11303071|9113599|12122494|20590605|9671109
M₃ receptor	15	CHRM3	P20309	ENSG00000133019	Human	dicyclomine	355	135650178	t	Antagonist	Antagonist		Not Determined	f	t		pKi		9.029999733				Ki	0.930000007		#ERROR!				NULL
M₃ receptor	15	CHRM3	P20309	ENSG00000133019	Human	glycopyrrolate	7459	187051762	t	Antagonist	Antagonist		Not Determined	f	t		pIC50		9.600000381		IC50		0.250999987			#ERROR!	Assay uses glycopyrronium bromide		21036043
M₃ receptor	15	CHRM3	P20309	ENSG00000133019	Human	ipratropium	325	135650410	t	Antagonist	Antagonist		Not Determined	f	t		pKi	9.800000191		9.300000191	-				#ERROR!				11303071|16847442
M₃ receptor	15	CHRM3	P20309	ENSG00000133019	Human	N-methyl scopolamine	316	135650725	t	Antagonist	Antagonist		Not Determined	f	t		pKi		10.39999962		-					#ERROR!				16847442
M₃ receptor	15	CHRM3	P20309	ENSG00000133019	Human	oxybutynin	359	135650784	t	Antagonist	Antagonist		Not Determined	f	t		pKi	8.899999619		8.800000191	-				#ERROR!				22243489|12122494|20590605
M₃ receptor	15	CHRM3	P20309	ENSG00000133019	Human	propantheline	329	135650857	t	Antagonist	Antagonist		Not Determined	f	t		pKi		10		-					#ERROR!				12049493
M₃ receptor	15	CHRM3	P20309	ENSG00000133019	Human	solifenacin	7483	187051786	t	Antagonist	Antagonist		Not Determined	f	t		pIC50		6.880000114		IC50		130.8000031			#ERROR!				21524581
M₃ receptor	15	CHRM3	P20309	ENSG00000133019	Human	solifenacin	7483	187051786	t	Antagonist	Antagonist		Non-selective	f	t		pKi	8		7.699999809	-				#ERROR!				12122494|20590605
M₃ receptor	15	CHRM3	P20309	ENSG00000133019	Human	tiotropium	367	135651266	t	Antagonist	Antagonist		Not Determined	f	t		pKi	11.10000038		9.5	-				#ERROR!				16847442|8441333
M₃ receptor	15	CHRM3	P20309	ENSG00000133019	Human	tropicamide	7319	178103892	t	Antagonist	Antagonist		Not Determined	f	t		pKi		7.519999981		Ki		30			#ERROR!				NULL
M₄ receptor	16	CHRM4	P08173	ENSG00000180720	Human	amitriptyline	200	135649939	t	Antagonist	Antagonist		Not Determined	f	t		pKi		8.140000343		Ki		7.199999809			#ERROR!				8100134
M₄ receptor	16	CHRM4	P08173	ENSG00000180720	Human	atropine	320	135652667	t	Antagonist	Antagonist		Not Determined	f	t		pKi		9.850000381		Ki		0.140000001			#ERROR!				29680464
M₄ receptor	16	CHRM4	P08173	ENSG00000180720	Human	atropine	320	135652667	t	Antagonist	Antagonist		Not Determined	f	t		pKi	9.5		8.699999809	-				#ERROR!				34430951204949|27043701|11303071
M₄ receptor	16	CHRM4	P08173	ENSG00000180720	Human	glycopyrrolate	7459	187051762	t	Antagonist	Antagonist		Not Determined	f	t		pIC50		9.800000191		IC50		0.158000007			#ERROR!	Assay uses glycopyrronium bromide		21036043
M₅ receptor	17	CHRM5	P08912	ENSG00000184984	Human	amitriptyline	200	135649939	t	Antagonist	Antagonist		Not Determined	f	t		pKi		7.800000191		Ki		15.69999981			#ERROR!				8100134
M₅ receptor	17	CHRM5	P08912	ENSG00000184984	Human	atropine	320	135652667	t	Antagonist	Antagonist		Not Determined	f	t		pKi	9.699999809		9.300000191	-				#ERROR!				2704370|11303071
M₅ receptor	17	CHRM5	P08912	ENSG00000184984	Human	glycopyrrolate	7459	187051762	t	Antagonist	Antagonist		Not Determined	f	t		pIC50		9.699999809		IC50		0.200000003			#ERROR!				21036043
MT₁ receptor	287	MTNR1A	P48039	ENSG00000168412	Human	agomelatine	198	135649906	t	Agonist	Full agonist		Non-selective	f	t		pKi	10.39999962		10	-				#ERROR!				12764576|9618428
MT₁ receptor	287	MTNR1A	P48039	ENSG00000168412	Human	ramelteon	1356	135651228	t	Agonist	Full agonist		Non-selective	f	t		pKi		10.89999962		-					#ERROR!				15695169
MT₁ receptor	287	MTNR1A	P48039	ENSG00000168412	Human	tasimelteon	7393	178103965	t	Agonist	Full agonist		Non-selective	f	t		pKi		9.5		Ki		0.3039999			#ERROR!				19054552
MT₂ receptor	288	MTNR1B	P49286	ENSG00000134640	Human	agomelatine	198	135649906	t	Agonist	Full agonist		Non-selective	f	t		pKi	10.5		9.899999619	-				#ERROR!				12764576|9618428
MT₂ receptor	288	MTNR1B	P49286	ENSG00000134640	Human	ramelteon	1356	135651228	t	Agonist	Full agonist		Non-selective	f	t		pKi		10		-					#ERROR!				15695169|21182402
MT₂ receptor	288	MTNR1B	P49286	ENSG00000134640	Human	tasimelteon	7393	178103965	t	Agonist	Agonist		Non-selective	f	t		pKi		10.19999981		Ki		0.069200002			#ERROR!				19054552
μ receptor	319	OPRM1	P35372	ENSG00000112038	Human	hydrocodone	7081	178103660	t	Agonist	Agonist		Not Determined	f	t		pKi		8.699999809		Ki		2			#ERROR!				22439881
μ receptor	319	OPRM1	P35372	ENSG00000112038	Human	hydromorphone	7082	178103661	t	Agonist	Agonist		Not Determined	f	t		pKi		9.550000191		Ki		0.280000001			#ERROR!				19282177
μ receptor	319	OPRM1	P35372	ENSG00000112038	Human	levorphanol	7595	223365930	t	Agonist	Agonist		Not Determined	f	t		pIC50		9.890000343		IC50		0.129999995			#ERROR!				12502358
μ receptor	319	OPRM1	P35372	ENSG00000112038	Human	loperamide	7215	178103790	t	Agonist	Agonist		Selective	f	t		pKi		9.279999733		Ki		0.529999971			#ERROR!				15454210
μ receptor	319	OPRM1	P35372	ENSG00000112038	Human	methadone	5458	125210108	t	Agonist	Agonist		Not Determined	f	t		pIC50		8.399000343		IC50		4.099999905			#ERROR!	Binding affinity against μ-opioid receptor (humar		11585443
μ receptor	319	OPRM1	P35372	ENSG00000112038	Human	nalbuphine	1663	135650680	t	Agonist	Agonist		Not Determined	f	t		pKi		8.800000191		Ki		1.600000024			#ERROR!				19282177
μ receptor	319	OPRM1	P35372	ENSG00000112038	Human	pethidine	7221	178103795	t	Agonist	Agonist		Not Determined	f	t		pIC50		6.5		IC50		315			#ERROR!				11585443
μ receptor	319	OPRM1	P35372	ENSG00000112038	Human	tapentadol	7477	187051780	t	Agonist	Agonist		Not Determined	f	t		pKi		6.800000191		Ki		160			#ERROR!				17656655
μ receptor	319	OPRM1	P35372	ENSG00000112038	Human	codeine	1673	135650117	t	Agonist	Full agonist		Not Determined	f	t		pKi		6.900000095		-					#ERROR!				9686407
μ receptor	319	OPRM1	P35372	ENSG00000112038	Human	fentanyl	1626	135650258	t	Agonist	Full agonist		Not Determined	f	t		pKi		9.199999809		-					#ERROR!				9686407
μ receptor	319	OPRM1	P35372	ENSG00000112038	Human	morphine	1627	135650624	t	Agonist	Full agonist		Selective	f	t		pKi		9		-					#ERROR!				9686407|2549383
μ receptor	319	OPRM1	P35372	ENSG00000112038	Human	sufentanil	3534	178100513	t	Agonist	Full agonist		Selective	f	t		pKi		9.899999619		Ki		0.137999997			#ERROR!				21215785
μ receptor	319	OPRM1	P35372	ENSG00000112038	Human	buprenorphine	1670	135652673	t	Agonist	Partial agonist		Not Determined	f	t		pKi		8.800000191		-					#ERROR!				9686407
μ receptor	319	OPRM1	P35372	ENSG00000112038	Human	butorphanol	7591	223365926	t	Agonist	Partial agonist		Not Determined	f	t		pKi		9.920000076		Ki		0.119999997			#ERROR!	Displacement of [^{3H}]DAMGO from hun		19674902
μ receptor	319	OPRM1	P35372	ENSG00000112038	Human	alvimopan	7471	187051774	t	Antagonist	Antagonist		Selective	f	t		pKi		9.329999924		Ki		0.469999999			#ERROR!				18313920
μ receptor	319	OPRM1	P35372	ENSG00000112038	Human	levallorphan	7209	178103784	t	Agonist	Antagonist		Selective	f	t		pKi	9.319999695		8.770000458	Ki	1.690000057		0.479999989	#ERROR!	Competition binding assay- the calculated K<sub>i<		21621410
μ receptor	319	OPRM1	P35372	ENSG00000112038	Human	methylnaltrexone	7563	223365899	t	Antagonist	Antagonist		Not Determined	f	t		pKi		8.699999809		Ki		2			#ERROR!	Displacement of [³H]DAMGO from hun		19282177
μ receptor	319	OPRM1	P35372	ENSG00000112038	Human	nalidemedine	9150	315661235	t	Antagonist	Antagonist		Selective	f	t		pKi		8.949999809		Ki		1.129999995			#ERROR!				30446313
μ receptor	319	OPRM1	P35372	ENSG00000112038	Human	naltrexone	1628	135650681	t	Agonist	Antagonist		Not Determined	f	t		pKi				-		9.5			#ERROR!				9686407
μ receptor	319	OPRM1	P35372	ENSG00000112038	Human	nalorphine	1629	135650682	t	Antagonist	Antagonist		Not Determined	f	t		pKi		8.899999619		-					#ERROR!				9686407
μ receptor	319	OPRM1	P35372	ENSG00000112038	Human	naltrexone	1639	135650687	t	Antagonist	Antagonist		Not Determined	f	t		pKi	9.699999809		9.149999619	Ki	0.709999979		0.199000001	#ERROR!				9686407|24973897
Na-Cl symporter	970	SLC12A3	P55017	ENSG00000070915	Human	chlorothiazide	4835	178101537	t	Inhibitor	Inhibition		Non-selective	f	t		-				-					#ERROR!				NULL
Na-Cl symporter	970	SLC12A3	P55017	ENSG00000070915	Human	cyclothiazide	4167	178100987	t	Inhibitor	Inhibition		Non-selective	f	t		-				-					#ERROR!				NULL
Na-Cl symporter	970	SLC12A3	P55017	ENSG00000070915	Human	hydrochlorothiazide	4836	178101538	t	Inhibitor	Inhibition		Non-selective	f	t		-				-					#ERROR!				NULL
Na-Cl symporter	970	SLC12A3	P55017	ENSG00000070915	Human	metolazone	4838	178101540	t	Inhibitor	Inhibition		Non-selective	f	t		-				-					#ERROR!				NULL
Na_v1.2	579	SCN2A	Q99250	ENSG00000136531	Human	phenytoin	2624	53801032	t	Channel blocker	Pore blocker		Not Determined	f	t		pIC50		4.900000095		-					#ERROR!				1658608
Na_v1.4	581	SCN4A	P35499	ENSG00000007314	Human	mexiletine	2629	135650603	t	Channel blocker	Antagonist		Not Determined	f	t		pIC50		3.400000095		-					#ERROR!				14608007
Na_v1.5	582	SCN5A	Q14524	ENSG00000183873	Human	lidocaine	2623	135650521	t	Channel blocker	Pore blocker		Not Determined	f	t		pKd		4.800000191		-					#ERROR!				8786356
Na_v1.7	584	SCN9A	Q15858	ENSG00000169432	Human	lacosamide	7472	135651775	t	Channel blocker	Antagonist		Not Determined	f	t		pIC50		3.700000048		IC50		182000			#ERROR!				18378801
nectin cell adhesion molecule 4	3112	NECTIN4	Q96NY8	ENSG00000143217	Human	enfortumab vedotin	10738	405067331	t	Antibody	Binding		Selective	f	t		-				-					#ERROR!				31526130
NET	926	SLC6A2	P23975	ENSG00000103546	Human	amitriptyline	200	135649939	t	Inhibitor	Inhibition		Non-selective	f	t		pKi		6.480000019		Ki		333			#ERROR!				23602445
NET	926	SLC6A2	P23975	ENSG00000103546	Human	amoxapine	201	135649941	t	Inhibitor	Inhibition		Non-selective	f	t		pKi		7.889999866		Ki		13			#ERROR!				NULL
NET	926	SLC6A2	P23975	ENSG00000103546	Human	atomoxetine	7118	178103695	t	Inhibitor	Inhibition		Non-selective	f	t		pKd		8.689999558		Kd		2.029999971			#ERROR!				9871604
NET	926	SLC6A2	P23975	ENSG00000103546	Human	bupropion	7135	178103711	t	Inhibitor	Inhibition		Not Determined	f	t		pKi		6.360000134		Ki		441			#ERROR!				19442525
NET	926	SLC6A2	P23975	ENSG00000103546	Human	clomipramine	2398	135650111	t	Inhibitor	Inhibition		Not Determined	f	t		pKd		7.420000076		Kd		38			#ERROR!				9537821
NET	926	SLC6A2	P23975	ENSG00000103546	Human	desipramine	2399	135650171	t	Inhibitor	Inhibition		Non-selective	f	t		pKi		8.680000305		Ki		2.099999905			#ERROR!				20462211
NET	926	SLC6A2	P23975	ENSG00000103546	Human	desvenlafaxine	7158	178103733	t	Inhibitor	Inhibition		Family selective	f	t		pKi		6.190000057		Ki		650	>		#ERROR!				24900709
NET	926	SLC6A2	P23975	ENSG00000103546	Human	dexamfetamine	2147	135650979	t	Inhibitor	Inhibition		Non-selective	f	t		pKi				Ki		101			#ERROR!				23602445
NET	926	SLC6A2	P23975	ENSG00000103546	Human	dosulepin	7549	223365886	t	Inhibitor	Inhibition		Non-selective	f	t		pKi		7.340000153		Ki		46			#ERROR!				9537821
NET	926	SLC6A2	P23975	ENSG00000103546	Human	doxepin	1225	135650208	t	Inhibitor	Inhibition		Not Determined	f	t		pKi		7.539999962		Ki		29			#ERROR!				NULL
NET	926	SLC6A2	P23975	ENSG00000103546	Human	duloxetine	202	135650212	t	Inhibitor	Inhibition		Not Determined	f	t		pKi		8.220000267		Ki		5.9699979			#ERROR!				23084899
NET	926	SLC6A2	P23975	ENSG00000103546	Human	imipramine	357	135650384	t	Inhibitor	Inhibition		Non-selective	f	t		pKi		7.800000191		Ki		16			#ERROR!				NULL
NET	926	SLC6A2	P23975	ENSG00000103546	Human	levomilnacipran	7435	178104007	t	Inhibitor	Inhibition		Not Determined	f	t		pIC50		7.400000095		IC50		40			#ERROR!				18468895
NET	926	SLC6A2	P23975	ENSG00000103546	Human	lofepramine	7551	223365887	t	Inhibitor	Inhibition		Not Determined	f	t		pKi		8.270000458		Ki		5.400000095			#ERROR!				9537821
NET	926	SLC6A2	P23975	ENSG00000103546	Human	maprotiline	2402	135650566	t	Inhibitor	Inhibition		Selective	f	t		pKi		7.920000076		Ki		12			#ERROR!				19836247
NET	926	SLC6A2	P23975	ENSG00000103546	Human	methylphenidate	7236	178103810	t	Inhibitor	Inhibition		Selective	f	t		pIC50		7.210000038		IC50		61			#ERROR!				NULL
NET	926	SLC6A2	P23975	ENSG00000103546	Human	milnacipran	7436	178104008	t	Inhibitor	Inhibition		Not Determined	f	t		pIC50		9.100000381		IC50		0.7940003			#ERROR!				23347683
NET	926	SLC6A2	P23975	ENSG00000103546	Human	nefazodone	7247	178103821	t	Inhibitor	Inhibition		Non-selective	f	t		pKd		6.440000057		Kd		360			#ERROR!	Dissociation constant of [3H]- nisoxatine vs. r		9871604
NET	926	SLC6A2	P23975	ENSG00000103546	Human	nomifensine	4792	178101494	t	Inhibitor	Inhibition		Selective	f	t		pKi		8.100000381		Ki		7.940000057			#ERROR!				NULL
NET	926	SLC6A2	P23975	ENSG00000103546	Human	nortriptyline	2404	135650736	t	Inhibitor	Inhibition		Non-selective	f	t		pKi		8.199999809		Ki		6.300000191			#ERROR!				10715164
NET	926	SLC6A2	P23975	ENSG00000103546	Human	protriptyline	7285	178103859	t	Inhibitor	Inhibition		Selective	f	t		pKd		8.850000381		Kd		1.409999967			#ERROR!				9537821
NET	926	SLC6A2	P23975	ENSG00000103546	Human	protriptyline	7285	178103859	t	Inhibitor	Inhibition		Selective	f	t		pKi		8.170000076		Ki		6.800000191			#ERROR!				18983139
NET	926	SLC6A2	P23975	ENSG00000103546	Human	protriptyline	7285	178103859	t	Inhibitor	Inhibition		Selective	f	t		pIC50		8.770000458		IC50		1.700000048			#ERROR!	Inhibition of [³H]nisoxetine binding to C		22420944
NET	926	SLC6A2	P23975	ENSG00000103546	Human	reboxetine	4808	178101510	t	Inhibitor	Inhibition		Selective	f	t		pKi		8		Ki		10			#ERROR!				10812041
NET	926	SLC6A2	P23975	ENSG00000103546	Human	sibutramine	2586	135651043	t	Inhibitor	Inhibition		Family selective	f	t		pKi		5.25		Ki		5619			#ERROR!				NULL
NET	926	SLC6A2	P23975	ENSG00000103546	Human	tapentadol	7477	187051780	t	Inhibitor	Inhibition		Not Determined	f	t		pKi		5.059999943		Ki		8800			#ERROR!				17656655
NET	926	SLC6A2	P23975	ENSG00000103546	Human	trimipramine	7317	178103890	t	Inhibitor	Inhibition		Non-selective	f	t		pKi		5.610000134		Ki		2450			#ERROR!				9537821
NET	926	SLC6A2	P23975	ENSG00000103546	Human	venlafaxine	7321	178103894	t	Inhibitor	Inhibition		Not Determined	f	t		pIC50		6.269999981		IC50		535			#ERROR!				20378347
nicotinic acetylcholine receptor α1 r	462	CHRNA1	P02708	ENSG00000138435	Human	succinylcholine	4004	178100830	t	Agonist	Full agonist	selective for (&al	Selective	f	t		-				-					#ERROR!				NULL
nicotinic acetylcholine receptor α1 r	462	CHRNA1	P02708	ENSG00000138435	Human	pancuronium	4001	178100827	t	Antagonist	Antagonist		Selective	f	t		-				-					#ERROR!				NULL
nicotinic acetylcholine receptor α3 r	464	CHRNA3	P32297	ENSG00000080644	Human	mecamylamine	3990	178100816	t	Channel blocker	None	α3β4	Non-selective	f	t		pIC50		6.409999847		IC50		390			#ERROR!				NULL
nicotinic acetylcholine receptor α3 r	464	CHRNA3	P32297	ENSG00000080644	Human	mecamylamine	3990	178100816	t	Channel blocker	None	α3β2	Non-selective	f	t		pIC50		5.119999886		IC50		7600			#ERROR!				NULL
nicotinic acetylcholine receptor α4 r	465	CHRNA4	P43681	ENSG00000101204	Human	varenicline	5459	135452109	t	Agonist	Agonist		Selective	f	t		pKi		9.960000038		Ki		0.109999999			#ERROR!	Binding affinity to human nicotinic acetylcholine rece		16171993
NK₁ receptor	360	TACR1	P25103	ENSG00000115353	Human	aprepitant	3490	178100479	t	Antagonist	Antagonist		Selective	f	t		pKi		10.05000019		Ki		0.090000004			#ERROR!				9804700|10737756
NPC1 like intracellular cholesterol transpc	2629	NPC1L1	Q9UHC9	ENSG00000015520	Human	ezetimibe	6816	178103422	t	Antagonist	Inhibition		Selective	f	t		pKd		6.659999847		Kd		220			#ERROR!	Measure of ezetimibe glucuronide binding to NPC1L		15928087
Organic anion transporter 1	1025	SLC22A6	Q4U2R8	ENSG00000197901	Human	probenecid	4357	178101174	t	Inhibitor	Inhibition	Inhibition of urate	Non-selective	f	t		pIC50		4.900000095		IC50		12500			#ERROR!				12472777
Ornithine decarboxylase	1276	ODC1	P11926	ENSG00000115758	Human	eflornithine	5176	178101866	t	Inhibitor	Inhibition		Selective	f	t		pKd		4.899999866		Kd		13000			#ERROR!				12859253
OX₁ receptor	321	HCRTR1	O43613	ENSG00000121764	Human	lemborexant	9302	318164821	t	Antagonist	Antagonist		Not Determined	f	t		pKi		8.220000267		Ki		6			#ERROR!	In a radioligand binding assay		25953512
OX₁ receptor	321	HCRTR1	O43613	ENSG00000121764	Human	suvorexant	2890	135650615	t	Antagonist	Antagonist		Not Determined	f	t		pKi	9.300000191		8.699999809	Ki		0.500999987		#ERROR!				20565075|24376396|23692283
OX₂ receptor	322	HCRTR2	O43614	ENSG00000137252	Human	lemborexant	9302	318164821	t	Antagonist	Antagonist		Not Determined	f	t		pKi		8.520000458		Ki		3			#ERROR!	In vitro radioligand binding assay		25953512
OX₂ receptor	322	HCRTR2	O43614	ENSG00000137252	Human	suvorexant	2890	135650615	t	Antagonist	Antagonist		Not Determined	f	t		pKi	9.5		8.899999619	Ki				#ERROR!				20565075|24376396|23692283
P2Y₁₂ receptor	328	P2RY12	Q9H244	ENSG00000169313	Human	cangrelor	1776	135650051	t	Antagonist	Antagonist		Not Determined	f	t		pIC50		8		-					#ERROR!				11502873
P2Y₁₂ receptor	328	P2RY12	Q9H244	ENSG00000169313	Human	cangrelor	1776	135650051	t	Antagonist	Antagonist		Not Determined	f	t		pIC50		9.399999619		IC50		0.400000006			#ERROR!				12213051
P2Y₁₂ receptor	328	P2RY12	Q9H244	ENSG00000169313	Human	ticagrelor	1765	135649975	t	Antagonist	Antagonist		Selective	f	t		pKi		7.849999905		Ki		14			#ERROR!				22984835
P2Y₁₂ receptor	328	P2RY12	Q9H244	ENSG00000169313	Human	ticagrelor	1765	135649975	t	Antagonist	Antagonist		Selective	f	t		pKi		8.699999809		Ki		1.99000001			#ERROR!	Binding data obtained using washed platelets in 96-v		17827008
P2Y₁₂ receptor	328	P2RY12	Q9H244	ENSG00000169313	Human	ticagrelor	1765	135649975	t	Antagonist	Antagonist		Selective	f	t		pKb		8.590000153		-					#ERROR!				25186974
pancreatic lipase	2590	PNLIP	P16233	ENSG00000175535	Human	orlistat	5277	178101960	t	Inhibitor	Inhibition		Not Determined	f	t		pIC50		8.899999619		IC50		1.299999952			#ERROR!				NULL
PAR1	347	F2R	P25116	ENSG00000181104	Human	vorapaxar	4047	178100871	t	Antagonist	Antagonist		Selective	f	t		pKi		8.090000153		Ki		8.100000381			#ERROR!				18447380
peptidylprolyl isomerase A	2751	PPIA	P62937	ENSG00000196262	Human	cyclosporin A	1024	135652074	t	Inhibitor	Inhibition	Inhibition of the p	Not Determined	f	t		pIC50		8.300000191		IC50		5			#ERROR!				1373887
Peroxisome proliferator-activated receptor	593	PPARA	Q07869	ENSG00000186951	Human	bezafibrate	2668	135649989	t	Agonist	Agonist		Not Determined	f	t		pEC50		4.300000191		-					#ERROR!				15293980|11779144
Peroxisome proliferator-activated receptor	593	PPARA	Q07869	ENSG00000186951	Human	ciprofibrate	3438	178100438	t	Agonist	Agonist		Not Determined	f	t		pEC50		6.050000191		EC50		900			#ERROR!	Transactivation of a GAL4-PPARα ligand bind		22381047
Peroxisome proliferator-activated receptor	593	PPARA	Q07869	ENSG00000186951	Human	clofibrate	2667	135650110	t	Agonist	Agonist		Selective	f	t		pIC50		4.25		-					#ERROR!				9836620|15293980|10691680
Peroxisome proliferator-activated receptor	593	PPARA	Q07869	ENSG00000186951	Human	fenofibrate	7186	135703761	t	Agonist	Agonist		Selective	f	t		pEC50		4.519999981		EC50		30000			#ERROR!	In a cell based hPPAR-GAL4 transactivation assay.		15293980
Peroxisome proliferator-activated receptor	593	PPARA	Q07869	ENSG00000186951	Human	gemfibrozil	3439	178100439	t	Agonist	Agonist		Selective	f	t		pEC50		4.230000019		EC50		59000			#ERROR!				21889235
Peroxisome proliferator-activated receptor	595	PPARG	P37231	ENSG00000132170	Human	rosiglitazone	1056	135650927	t	Agonist	Agonist		Selective	f	t		pKi		6.920000076		Ki		120			#ERROR!				11720854
Peroxisome proliferator-activated receptor	595	PPARG	P37231	ENSG00000132170	Human	mesalazine	2700	135649824	t	Agonist	Full agonist		Non-selective	f	t		pIC50		1.820000052		-					#ERROR!				15824083
Peroxisome proliferator-activated receptor	595	PPARG	P37231	ENSG00000132170	Human	pioglitazone	2694	135650829	t	Agonist	Full agonist		Selective	f	t		pIC50		6.199999809		-					#ERROR!				9836620|11095972|9454824
Peroxisome proliferator-activated receptor	595	PPARG	P37231	ENSG00000132170	Human	rosiglitazone	1056	135650927	t	Agonist	Full agonist		Selective	f	t		pKd		7.400000095		-					#ERROR!				9836620|9013583|9454824
Peroxisome proliferator-activated receptor	595	PPARG	P37231	ENSG00000132170	Human	troglitazone	2693	135651283	t	Agonist	Full agonist		Selective	f	t		pIC50		6.300000191		IC50		501.2000122			#ERROR!				9836620|9454824
Peroxisome proliferator-activated receptor	595	PPARG	P37231	ENSG00000132170	Human	troglitazone	2693	135651283	t	Agonist	Full agonist		Selective	f	t		pKi		5.800000191		Ki		1585			#ERROR!				8828476
phosphatidylinositol-4,5-bisphosphate 3-k	2153	PIK3CA	P42336	ENSG00000121879	Human	alpelisib	7955	249565638	t	Inhibitor	Inhibition		Not Determined	f	t		pIC50		8.300000191		IC50		5			#ERROR!				23726034
phosphatidylinositol-4,5-bisphosphate 3-k	2155	PIK3CD	O00329	ENSG00000171608	Human	duvelisib	7795	223366126	t	Inhibitor	Inhibition		Not Determined	f	t		pIC50		8.600000381		IC50		2.5			#ERROR!				24211136
phosphatidylinositol-4,5-bisphosphate 3-k	2155	PIK3CD	O00329	ENSG00000171608	Human	idelalisib	6741	178103352	t	Inhibitor	Inhibition	in vitro act	Not Determined	f	t		pIC50		8.600000381		IC50		2.5			#ERROR!				20959606
phosphatidylinositol-4,5-bisphosphate 3-k	2156	PIK3CG	P48736	ENSG00000105851	Human	duvelisib	7795	223366126	t	Inhibitor	Inhibition		Not Determined	f	t		pIC50		7.570000172		IC50		27			#ERROR!				24211136
phosphodiesterase 3A	1298	PDE3A	Q14432	ENSG00000172572	Human	anagrelide	7114	178103691	t	Inhibitor	Inhibition		Selective	f	t		pIC50	7.269999981		7.090000153	IC50		82	54	#ERROR!				1321910|30273381|1311763

target	target_id	target_gene	target_uniprot	target_ensembl_gene_id	target_species	ligand	ligand_id	ligand_pubchem	approved_drug	type	action	action_comment	selectivity	endogenous	primary_target	concentration_r	affinity_units	affinity_high	affinity_median	affinity_low	original_affinity	original_affinity	original_affinity	original_affinity	original_affinity	assay_descripti	receptor_site	ligand_context	pubmed_id
phosphodiesterase 3A	1298	PDE3A	Q14432	ENSG00000172572	Human	enoximone	9063	315661148	t	Inhibitor	Inhibition		Selective	f	t		pIC50		5.420000076		IC50		3800		#ERROR!	Measuring inhibition of human platelet PDE			3027339
phosphodiesterase 3A	1298	PDE3A	Q14432	ENSG00000172572	Human	inamrinone	7202	178103777	t	Inhibitor	Inhibition		Not Determined	f	t		pIC50		4.800000191		IC50		16000		#ERROR!	Inhibition of PDE3 purified from human platelets, an			2536438
phosphodiesterase 3A	1298	PDE3A	Q14432	ENSG00000172572	Human	milrinone	5225	178101910	t	Inhibitor	Inhibition		Selective	f	t		pIC50	6.360000134		6.349999905	IC50	450		440	#ERROR!				10644042|14592490
phosphodiesterase 3B	1299	PDE3B	Q13370	ENSG00000152270	Human	cilostazol	7148	178103723	t	Inhibitor	Inhibition		Selective	f	t		pIC50		6.420000076		IC50		380		#ERROR!				10644042
phosphodiesterase 3B	1299	PDE3B	Q13370	ENSG00000152270	Human	inamrinone	7202	178103777	t	Inhibitor	Inhibition		Selective	f	t		pIC50		4.510000229		IC50		31200		#ERROR!				10644042
phosphodiesterase 4A	1300	PDE4A	P27815	ENSG00000065989	Human	apremilast	7372	178103944	t	Inhibitor	Inhibition		Selective	f	t		pIC50		7.849999905		IC50		14		#ERROR!	Potency against recombinant human PDE4A1A isofo			24882690
phosphodiesterase 4A	1300	PDE4A	P27815	ENSG00000065989	Human	ibudilast	7399	178103971	t	Inhibitor	Inhibition		Not Determined	f	t		pIC50		7.269999981		IC50		54		#ERROR!				18686943
phosphodiesterase 4B	1301	PDE4B	Q07343	ENSG00000184588	Human	apremilast	7372	178103944	t	Inhibitor	Inhibition		Selective	f	t		pIC50	7.570000172		7.369999886	IC50	43		27	#ERROR!	Potency against recombinant human PDE4B1 and E			24882690
phosphodiesterase 4B	1301	PDE4B	Q07343	ENSG00000184588	Human	ibudilast	7399	178103971	t	Inhibitor	Inhibition		Not Determined	f	t		pIC50		7.190000057		IC50		65		#ERROR!				18686943
phosphodiesterase 4B	1301	PDE4B	Q07343	ENSG00000184588	Human	roflumilast	6962	178103543	t	Inhibitor	Inhibition		Not Determined	f	t		pIC50		9.399999619		-				#ERROR!				19195882
phosphodiesterase 4C	1302	PDE4C	Q08493	ENSG00000105650	Human	apremilast	7372	178103944	t	Inhibitor	Inhibition		Selective	f	t		pIC50		6.929999828		IC50		118		#ERROR!	Potency against recombinant human PDE4C1 isofor			24882690
phosphodiesterase 4C	1302	PDE4C	Q08493	ENSG00000105650	Human	ibudilast	7399	178103971	t	Inhibitor	Inhibition		Not Determined	f	t		pIC50		6.619999886		IC50		239		#ERROR!				18686943
phosphodiesterase 4D	1303	PDE4D	Q08499	ENSG00000113448	Human	apremilast	7372	178103944	t	Inhibitor	Inhibition		Selective	f	t		pIC50		7.480000019		IC50		33		#ERROR!	Potency against recombinant human PDE4D2 isofor			24882690
phosphodiesterase 5A	1304	PDE5A	O76074	ENSG00000138735	Human	sildenafil	4743	178101454	t	Inhibitor	Inhibition		Selective	f	t		pIC50	9		8.409999847	IC50	3.900000095		1	#ERROR!				10385692|23137303
phosphodiesterase 5A	1304	PDE5A	O76074	ENSG00000138735	Human	tadalafil	7299	178103872	t	Inhibitor	Inhibition		Selective	f	t		pIC50		8.520000458		IC50		3		#ERROR!				21189023
phosphodiesterase 5A	1304	PDE5A	O76074	ENSG00000138735	Human	vardenafil	7320	178103893	t	Inhibitor	Inhibition		Selective	f	t		pIC50		9.699999809		IC50		0.200000003		#ERROR!				15837326
plasminogen	2394	PLG	P00747	ENSG00000122194	Human	6-aminocaproic acid	6574	178103188	t	Inhibitor	Inhibition		Selective	f	t		pIC50		4.400000095		IC50		40000		#ERROR!	Measuring inhibition of fibrinolysis in plasma.			24900876
plasminogen	2394	PLG	P00747	ENSG00000122194	Human	aprotinin	6570	178103184	t	Inhibitor	Binding		Not Determined	f	t		pIC50		6.800000191		IC50		160		#ERROR!				17666018
plasminogen	2394	PLG	P00747	ENSG00000122194	Human	tranexamic acid	6573	178103187	t	Inhibitor	Binding		Not Determined	f	t		pIC50		3.599999905		IC50		241000		#ERROR!				17666018
platelet derived growth factor receptor alpi	1803	PDGFRA	P16234	ENSG00000134853	Human	olaratumab	9172	315661255	t	Antibody	Binding		Selective	f	t		pIC50		9.239999771		IC50		0.579999983		#ERROR!	Cellular assay		NULL	
platelet derived growth factor receptor alpi	1803	PDGFRA	P16234	ENSG00000134853	Human	olaratumab	9172	315661255	t	Antibody	Binding		Selective	f	t		pKd		10.39999962		Kd		0.039999999		#ERROR!	BIAcore assay		NULL	
platelet derived growth factor receptor bet	1804	PDGFRB	P09619	ENSG00000113721	Human	sunitinib	5713	178102339	t	Inhibitor	Inhibition		Not Determined	f	t		pIC50		8.199999809		IC50		5.699999809		#ERROR!				23279183
platelet derived growth factor receptor bet	1804	PDGFRB	P09619	ENSG00000113721	Human	sunitinib	5713	178102339	t	Inhibitor	Inhibition		Not Determined	f	t		pKi		8.100000381		Ki		8		#ERROR!				12538485
poly(ADP-ribose) polymerase 1	2771	PARP1	P09874	ENSG00000143799	Human	niraparib	8275	25216487	t	Inhibitor	Binding		Not Determined	f	t		pIC50		8.420000076		IC50		3.799999952		#ERROR!				19873981
poly(ADP-ribose) polymerase 1	2771	PARP1	P09874	ENSG00000143799	Human	olaparib	7519	187051821	t	Inhibitor	Inhibition		Not Determined	f	t		pIC50		8.300000191		IC50		5		#ERROR!	Assayed using PARP-1 isolated from HeLa cells, in		18800822|12857389	
poly(ADP-ribose) polymerase 1	2771	PARP1	P09874	ENSG00000143799	Human	rucaparib	7736	223366069	t	Inhibitor	Inhibition		Not Determined	f	t		pKi		8.850000381		Ki		1.399999976		#ERROR!				17363489
poly(ADP-ribose) polymerase 1	2771	PARP1	P09874	ENSG00000143799	Human	talazoparib	8313	252166525	t	Inhibitor	Inhibition		Not Determined	f	t		-				-				#ERROR!				25195882
poly(ADP-ribose) polymerase 2	2772	PARP2	Q9UGN5	ENSG00000129484	Human	niraparib	8275	25216487	t	Inhibitor	Binding		Not Determined	f	t		pIC50		8.680000305		IC50		2.099999905		#ERROR!				19873981
poly(ADP-ribose) polymerase 2	2772	PARP2	Q9UGN5	ENSG00000129484	Human	olaparib	7519	187051821	t	Inhibitor	Inhibition		Not Determined	f	t		pIC50		9		IC50		1		#ERROR!	Assayed using PARP-2 isolated from HeLa cells, in		18800822|12857389	
poly(ADP-ribose) polymerase 2	2772	PARP2	Q9UGN5	ENSG00000129484	Human	talazoparib	8313	252166525	t	Inhibitor	Inhibition		Not Determined	f	t		-				-				#ERROR!				25195882
Progesterone receptor	627	PGR	P06401	ENSG00000082175	Human	dydrogesterone	2878	135650214	t	Agonist	Agonist	Binding to PRs o	Not Determined	f	t		pKi		6.900000095		Ki		126		#ERROR!				14670641|21376746
Progesterone receptor	627	PGR	P06401	ENSG00000082175	Human	levonorgestrel	2881	135650516	t	Agonist	Agonist	Selective	f	t		-				-				#ERROR!				14670641|6645495
Progesterone receptor	627	PGR	P06401	ENSG00000082175	Human	medroxyprogesterone	2879	135650577	t	Agonist	Agonist	Affinity at human	Selective	f	t		pKi		9.470000267		Ki		0.340000004		#ERROR!				9464360
Progesterone receptor	627	PGR	P06401	ENSG00000082175	Human	norethisterone	2880	135650731	t	Agonist	Agonist		Not Determined	f	t		pEC50		8.659999847		EC50		2.200000048		#ERROR!				9464361
Progesterone receptor	627	PGR	P06401	ENSG00000082175	Human	progesterone	2377	135651551	t	Agonist	Agonist	Affinity at human	Not Determined	f	t		pKi		8.460000038		Ki		3.5		#ERROR!				9464360
Progesterone receptor	627	PGR	P06401	ENSG00000082175	Human	ulipristal acetate	7460	187051763	t	Agonist	Partial agonist		Not Determined	f	t		pEC50		7		EC50		100	>					18243712
Progesterone receptor	627	PGR	P06401	ENSG00000082175	Human	ulipristal acetate	7460	187051763	t	Antagonist	Antagonist		Selective	f	t		pIC50		9.699999809		IC50		0.200000003		#ERROR!				18243712
Progesterone receptor	627	PGR	P06401	ENSG00000082175	Human	mifepristone	2805	135650608	t	Antagonist	Mixed		Selective	f	t		pKi		8.960000038		Ki		1.100000024		#ERROR!				12781197
programmed cell death 1 (CD279)	2760	PDCD1	Q15116	ENSG00000188389	Human	cemiplimab	10090	381118800	t	Antibody	Binding		Selective	f	t		pKd		8.210000038		Kd		6.110000134		#ERROR!	Immobilised antibody binding to mobile phase mono			28265006
programmed cell death 1 (CD279)	2760	PDCD1	Q15116	ENSG00000188389	Human	cemiplimab	10090	381118800	t	Antibody	Binding		Selective	f	t		pKd		9.199999809		Kd		0.628000021		#ERROR!	Immobilised antibody binding to mobile phase dimer			28265006
programmed cell death 1 (CD279)	2760	PDCD1	Q15116	ENSG00000188389	Human	nivolumab	7335	178103907	t	Antibody	Binding		Selective	f	t		pKd		9.140000343		Kd		0.730000019		#ERROR!				21154417|23302904
programmed cell death 1 (CD279)	2760	PDCD1	Q15116	ENSG00000188389	Human	pembrolizumab	7499	187051801	t	Antibody	Binding		Selective	f	t		pKd		10		Kd		0.100000001	~					NULL
proprotein convertase subtilisin/kexin type	2388	PCSK9	Q8NBP7	ENSG00000169174	Human	alirocumab	6744	178103355	t	Inhibitor	Inhibition		Not Determined	f	t		pKd		9.420000076		Kd		0.377000004		#ERROR!				NULL
proprotein convertase subtilisin/kexin type	2388	PCSK9	Q8NBP7	ENSG00000169174	Human	evolocumab	7343	178103915	t	Inhibitor	Inhibition		Selective	f	t		pKd		9.720000267		Kd		0.189999998		#ERROR!	Dissociation equilibrium constant for wild-type hPCS		NULL	
proteasome 20S subunit beta 5	2406	PSMB5	P28074	ENSG00000100804	Human	bortezomib	6391	178103007	t	Inhibitor	Inhibition		Not Determined	f	t		pIC50		7.699999809		IC50		20		#ERROR!				19428245
proteasome 20S subunit beta 5	2406	PSMB5	P28074	ENSG00000100804	Human	ixazomib	8450	252166660	t	Inhibitor	Inhibition		Selective	f	t		pKi		9.029997733		Ki		0.930000007		#ERROR!				20160034
protein kinase C delta	1485	PRKCD	Q05655	ENSG00000163932	Human	ingenol mebutate	7443	178104015	t	Activator	Activation		Not Determined	f	t		pKi		9.420000076		Ki		0.375999987		#ERROR!				15126366
PTH1 receptor	331	PTH1R	Q03431	ENSG00000160801	Human	abaloparatide	8299	252166511	t	Agonist	Agonist		Selective	f	t		-				-				#ERROR!				24078470
PTH1 receptor	331	PTH1R	Q03431	ENSG00000160801	Human	teriparatide	4448	135652417	t	Agonist	Full agonist		Not Determined	f	t		pIC50		7.400000095		-				#ERROR!				8702701
renin	2413	REN	P00797	ENSG00000143839	Human	aliskiren	4812	178101514	t	Inhibitor	Inhibition		Not Determined	f	t		pIC50		9.199999809		IC50		0.600000024		#ERROR!				18307734
Retinoic acid receptor-α	590	RARA	P10276	ENSG00000131759	Human	alitretinoin	2645	135651418	t	Agonist	Agonist		Non-selective	f	t		pKd		9.5		-				#ERROR!				3184125|7791761
Retinoic acid receptor-α	590	RARA	P10276	ENSG00000131759	Human	tamibarotene	2648	135649930	t	Agonist	Agonist		Selective	f	t		pIC50		6.900000095		-				#ERROR!				3184125|7791761|10637371
Retinoic acid receptor-α	590	RARA	P10276	ENSG00000131759	Human	tamibarotene	2648	135649930	t	Agonist	Agonist		Selective	f	t		pEC50		7.349999905		EC50		Agonist	45	#ERROR!				20925433
Retinoic acid receptor-α	590	RARA	P10276	ENSG00000131759	Human	tazarotene	6952	178103534	t	Agonist	Agonist		Family selective	f	t		pEC50		7.199999809		EC50		63		#ERROR!				19058965
Retinoic acid receptor-α	590	RARA	P10276	ENSG00000131759	Human	tretinoin	2644	135651435	t	Agonist	Agonist		Not Determined	f	t		pEC50		7.769999981		EC50		17		#ERROR!				19058965
Retinoic acid receptor-β	591	RARB	P10826	ENSG00000077092	Human	adapalene	5429	178102082	t	Agonist	Agonist		Family selective	f	t		pKi		7.46999979		Ki		34		#ERROR!	Binding affinity to RARβ using [³<sup>H			8544175
Retinoic acid receptor-β	591	RARB	P10826	ENSG00000077092	Human	alitretinoin	2645	135651418	t	Agonist	Agonist		Not Determined	f	t		pKi		9.699999809		-				#ERROR!				10964567|7791761
Retinoic acid receptor-β	591	RARB	P10826	ENSG00000077092	Human	tazarotene	6952	178103534	t	Agonist	Agonist		Family selective	f	t		pEC50		8.800000212		EC50		0.100000381		#ERROR!				19058965
Retinoic acid receptor-β	591	RARB	P10826	ENSG00000077092	Human	tretinoin	2644	135651435	t	Agonist	Agonist		Non-selective	f	t		pEC50		7.920000076		EC50		12		#ERROR!				19058965
Retinoic acid receptor-γ	592	RARG	P13631	ENSG00000172819	Human	adapalene	5429	178102082	t	Agonist	Agonist		Family selective	f	t		pKi		6.899999866		Ki		130		#ERROR!	Binding affinity to RARγ using [³<su			8544175
Retinoic acid receptor-γ	592	RARG	P13631	ENSG00000172819	Human	alitretinoin	2645	135651418	t	Agonist	Agonist		Selective	f	t		pKd		9		-				#ERROR!				10964567|7791761
Retinoic acid receptor-γ	592	RARG	P13631	ENSG00000172819	Human	tazarotene	6952	178103534	t	Agonist	Agonist		Family selective	f	t		pEC50		7.400000095		EC50		40		#ERROR!				19058965
Retinoic acid receptor-γ	592	RARG	P13631	ENSG00000172819	Human	tretinoin	2644	135651435	t	Agonist	Agonist		Non-selective	f	t		pEC50		9.699999809		EC50		0.200000003		#ERROR!				19058965
Retinoic acid receptor-γ	592	RARG	P13631	ENSG00000172819	Human	trifarotene	9962	374883849	t	Agonist	Agonist		Selective	f	t		pEC50		8.109999657		EC50		7.699999809		#ERROR!	Ligand-induced activation in a cellular reporter assa			29706423
Retinoid X receptor-α	610	RXRA	P19793	ENSG00000186350	Human	bexarotene	2807	135650520	t	Agonist	Agonist		Family selective	f	t		pIC50		7.400000095		-				#ERROR!				10637371|10052980|8071941
Retinoid X receptor-β	611	RXRB	P28702	ENSG00000228333	Human	bexarotene	2807	135650520	t	Agonist	Agonist		Family selective	f	t		pIC50		7.699999809		-				#ERROR!				10637371|10052980|8071941
Retinoid X receptor-γ	612	RXRG	P48443	ENSG00000143171	Human	bexarotene	2807	135650520	t	Agonist	Agonist		Family selective	f	t		pIC50		7.539999962		-				#ERROR!				10637371|10052980|8071941
ret proto-oncogene	2185	RET	P07949	ENSG00000165731	Human	vandetanib	5717	178103343	t	Inhibitor	Inhibition		Not Determined	f	t		pIC50		8.399999619		IC50		4		#ERROR!	In a biochemical assay.			29657135
ret proto-oncogene	2185	RET	P07949	ENSG00000165731	Human	vandetanib	5717	178102343	t	Inhibitor	Inhibition		Not Determined	f	t		pIC50		7		IC50		100		#ERROR!				12499271
ret proto-oncogene	2185	RET	P07949	ENSG00000165731	Human	vandetanib	5717	178102343	t	Inhibitor	Inhibition		Not Determined	f	t		pKd		7.46999979		Kd		34		#ERROR!				22037378
Rho associated coiled-coil containing prot	1503	ROCK1	Q13464	ENSG00000067900	Human	fasudil	5181	178101871	t	Inhibitor	Inhibition		Not Determined	f	t		pKi		7		Ki		100		#ERROR!				21145740
Rho associated coiled-coil containing prot	1503	ROCK1	Q13464	ENSG00000067900	Human	fasudil	5181	178101871	t	Inhibitor	Inhibition		Not Determined	f	t		pIC50	5.570000172		5.5	IC50	3160		2670	#ERROR!				21145740|19597037
Rho associated coiled-coil containing prot	1503	ROCK1	Q13464	ENSG00000067900	Human	ripasudil	10423	385612183	t	Inhibitor	Inhibition		Not Determined	f	t		pIC50		7.289999962		IC50		51		#ERROR!				24025505
Rho associated coiled-coil containing prot	1504	ROCK2	O75116	ENSG00000134318	Human	fasudil	5181	178101871	t	Inhibitor	Inhibition		Not Determined	f	t		pIC50	5.889999866		5.889999886	IC50	1349		1290	#ERROR!				21145740|19597037
Rho associated coiled-coil containing prot	1504	ROCK2	O75116	ENSG00000134318	Human	fasudil	5181	178101871	t	Inhibitor	Inhibition		Not Determined	f	t		pKi		6.570000172		Ki		271		#ERROR!				19597037
Rho associated coiled-coil containing prot	1504	ROCK2	O75116	ENSG00000134318	Human	ripasudil	10423	385612183	t	Inhibitor	Inhibition		Not Determined	f	t		pIC50		7.71999979		IC50		19		#ERROR!				24025505
ribonucleotide reductase catalytic subunit	2630	RRM1	P23921	ENSG00000167325	Human	fludarabine	4802	101504	t	Inhibitor	Inhibition		Not Determined	f	t		pIC50		6		IC50		1000		#ERROR!	Inhibition of ADP reduction by ribonucleotide reducta			7048062
ribonucleotide reductase catalytic subunit	2630	RRM1	P23921	ENSG00000167325	Human	gemcitabine	4793	178101495	t	Inhibitor	Inhibition		Not Determined	f	t		-				-				#ERROR!				2233693
ribonucleotide reductase catalytic subunit	2630	RRM1	P23921	ENSG00000167325	Human	hydroxyurea	6822	178103428	t	Inhibitor	Inhibition		Not Determined	f	t		pIC50		3.829999924		IC50		148000		#ERROR!	Dissolved in 1% DMSO, in vitro assay using r			15670581
ribonucleotide reductase regulatory subun	2631	RRM2	P31350	ENSG00000171848	Human	clofarabine	6802	178103408	t	Inhibitor	Inhibition		Not Determined	f	t		pIC50		8.300000191		IC50		5		#ERROR!	Inhibition of growth of K562 cells			1707752
ribonucleotide reductase regulatory subun	2631	RRM2	P31350	ENSG00000171848	Human	fludarabine	4802	178101504	t	Inhibitor	Inhibition		Not Determined	f	t		pIC50		6		IC50		1000		#ERROR!	Inhibition of ADP reduction by ribonucleotide reducta			7048062
ribonucleotide reductase regulatory subun	2631	RRM2	P31350	ENSG00000171848	Human	gemcitabine	4793	178101495	t	Inhibitor	Inhibition		Not Determined	f	t		-				-				#ERROR!				2233693
ribonucleotide reductase regulatory subun	2631	RRM2	P31350	ENSG00000171848	Human	hydroxyurea	6822	178103428	t	Inhibitor	Inhibition		Not Determined	f	t		pIC50		3.829999924		IC50		148000		#ERROR!	Dissolved in 1% DMSO, in vitro assay using r			15670581
RyR1	747	RYR1	P21817	ENSG00000196218	Human	dantrolene	4172	178100992	t	Antagonist	Antagonist		Non-selective	f	t		-				-				#ERROR!			NULL	
S1P₁ receptor	275	S1PR1	P21453	ENSG00000170989	Human	fingolimod	2407	135650261	t	Agonist	Agonist		Not Determined	f	t		pIC50		6.079999924		IC50		840		#ERROR!				15615513
S1P₁ receptor	275	S1PR1	P21453	ENSG00000170989	Human	ozanimod	8709	252827367	t	Agonist	Agonist		Selective	f	t		pEC50		9.800000191		EC50		0.159999996		#ERROR!	In a cAMP assay.			18708635|26990079
S1P₁ receptor	275	S1PR1	P21453	ENSG00000170989	Human	ozanimod	8709	252827367	t	Agonist	Agonist		Selective	f	t		pEC50		9.360000343		EC50		0.409999996		#ERROR!	In a GTPγS assay.			26990079
S1P₁ receptor	275	S1PR1	P21453	ENSG00000170989	Human	siponimod	9289	318164808	t	Agonist	Agonist		Not Determined	f	t		pEC50		9.399999619		EC50		0.400000006		#ERROR!	In a GTPγS binding assay			24900670|29735753
S1P₅ receptor	279	S1PR5	Q9H228	ENSG00000180739	Human	fingolimod	2407	135650261	t	Agonist	Agonist		Selective	f	t		pIC50		5.679999828		IC50		2100		#ERROR!				15615513
S1P₅ receptor	279	S1PR5	Q9H228	ENSG00000180739	Human	ozanimod	8709	252827367	t	Agonist	Agonist		Selective	f	t		pEC50		7.960000038		EC50		11		#ERROR!	In a GTPγS assay.			26990079
S1P₅ receptor	279	S1PR5	Q9H228	ENSG00000180739	Human	siponimod	9289	318164808	t	Agonist	Agonist		Not Determined	f	t		pEC50		9.010000229		EC50		0.980000019		#ERROR!	In a GTPγS binding assay			24125884|22464698|9735753
serine protease 1	2397	PRSS1	P07477	ENSG00000204983	Human	camostat	6432	178103048	t	Inhibitor	Inhibition		Not Determined	f	t		pIC50		7.300000191		IC50		50.59999847		#ERROR!				7629790
serine protease 1	2397	PRSS1	P07477	ENSG00000204983	Human	nafamostat	4262	178101079	t	Inhibitor	Inhibition		Not Determined	f	t		pIC50		7.769999981		IC50		17		#ERROR!				9544206
serpin family C member 1	2632	SERPINC1	P01008	ENSG00000117601	Human	dalteparin	6803	178103409	t	Activator	Activation		Selective	f	t		-				-				#ERROR!	Catalyses the inactivation of coagulation proteases b			3744129
serpin family C member 1	2632	SERPINC1	P01008	ENSG00000117601	Human	danaparoid	6804	178103410	t	Activator	Activation		Selective	f	t		-				-				#ERROR!	Catalyses the inactivation of coagulation proteases b		8137606|19398784	
serpin family C member 1	2632	SERPINC1	P01008	ENSG00000117601	Human	enoxaparin	6811	178103417	t	Activator	Activation		Selective	f	t		-				-				#ERROR!	Catalyses the inactivation of coagulation proteases b			7667822
serpin family C member 1	2632	SERPINC1	P01008	ENSG00000117601	Human	fondaparinux	6819	178103425	t	Activator	Activation		Selective	f	t		pKd		7.489999771		Kd		32		#ERROR!	Catalyses the inactivation of coagulation proteases b			12383040
serpin family C member 1	2632	SERPINC1	P01008	ENSG00000117601	Human	heparin	4214	178101031	t	Activator	Activation		Selective	f	t		pKd		7.849999905		Kd		14.19999981		#ERROR!	Catalyses the inactivation of coagulation proteases b			23598032
serpin family C member 1	2632	SERPINC1	P01008	ENSG00000117601	Human	tinzaparin	6846	178103452	t	Activator	Activation		Selective	f	t		-				-				#ERROR!	Catalyses the inactivation of coagulation proteases b			7528134
SERT	928	SLC6A4	P31645	ENSG00000108576	Human	amitriptyline	200	135649939	t	Inhibitor	Inhibition		Selective	f	t		pKi		6.849999905		Ki		141		#ERROR!				23602445
SERT	928	SLC6A4	P31645	ENSG00000108576	Human	amoxapine	201	135649941	t	Inhibitor	Inhibition		Non-selective	f	t		pKi		7.739999771		Ki		18		#ERROR!			NULL	
SERT	928	SLC6A4	P31645	ENSG00000108576	Human	citalopram	7547	223355883	t	Inhibitor	Inhibition		Selective	f	t		pKi		8.359999657		Ki		4.380000114		#ERROR!	Displacement of [³<sup>H]citalopram from h			18487050
SERT	928	SLC6A4	P31645	ENSG00000108576	Human	clomipramine	2398	135650111	t	Inhibitor	Inhibition		Selective	f	t		pKd		9.550000191		Kd		0.280000001		#ERROR!				9537821
SERT	928	SLC6A4	P31645	ENSG00000108576	Human	clomipramine	2398	135650111	t	Inhibitor	Inhibition		Selective	f	t		pKi		9.680000305		Ki		0.209999993		#ERROR!				9537821
SERT	928	SLC6A4	P31645	ENSG00000108576	Human	dapoxetine	7901	249565584	t	Inhibitor	Inhibition		Selective	f	t		pIC50		8.949999809		IC50		1.120000005		#ERROR!	Inhibition of the reuptake of [³<sup>H]5-HT		NULL	
SERT	928	SLC6A4	P31645	ENSG00000108576	Human	desipramine	2399	135650171	t	Inhibitor	Inhibition		Non-selective	f	t		pKi		6.789999962		Ki		163		#ERROR!	Assay using HEK-293 cells transfected with hSERT			9400906
SERT	928	SLC6A4	P31645	ENSG00000108576	Human	desvenlafaxine	7158	178103732	t	Inhibitor	Inhibition		Not Determined	f	t		pKi		7.820000172		Ki		15		#ERROR!				24900709
SERT	928	SLC6A4	P31645	ENSG00000108576	Human	dosulepin	7549	223365885	t	Inhibitor	Inhibition		Non-selective	f	t		pKi		8.069999695		Ki		8.600000381		#ERROR!				9537821
SERT	928	SLC6A4	P31645	ENSG00000108576	Human	doxepin	1225	135650208	t	Inhibitor	Inhibition		Not Determined	f	t		pKi		7.170000076		Ki		68		#ERROR!				9537821
SERT	928	SLC6A4	P31645	ENSG00000108576	Human	duloxetine	202	135650212	t	Inhibitor	Inhibition		Not Determined	f	t		pKi		8.300000191		Ki		5		#ERROR!				18387300
SERT	928	SLC6A4	P31645	ENSG00000108576	Human	escitalopram	7177	178103752	t	Inhibitor	Inhibition		Not Determined	f	t		pKd		8.680000305		Kd		2.099999905		#ERROR!				27049939
SERT	928	SLC6A4	P31645	ENSG00000108576	Human	escitalopram	7177	178103752	t	Inhibitor	Inhibition		Selective	f	t		pIC50		9		IC50		1		#ERROR!				19720528
SERT	928	SLC6A4	P31645	ENSG00000108576	Human	fluoxetine	203	135650270	t	Inhibitor	Inhibition		Selective	f	t		pKi		8.5		Ki		3		#ERROR!				9537821
SERT	928	SLC6A4	P31645	ENSG00000108576	Human	fluvoxamine	7189	178103764	t	Inhibitor	Inhibition		Selective	f	t		pKd		8.659999847		Kd		2.200000048		#ERROR!				9537821
SERT	928	SLC6A4	P31645	ENSG00000108576	Human	imipramine	357	135650384	t	Inhibitor	Inhibition		Non-selective	f	t		pKi		7.690000057		Ki		20.29999924		#ERROR!	Displacement of [³<sup>H]citalopram from h			8496700
SERT	928	SLC6A4	P31645	ENSG00000108576	Human	levomilnacipran	7435	178104007	t	Inhibitor	Inhibition		Not Determined	f	t		pIC50		6.489999771		IC50		320		#ERROR!				18468895
SERT	928	SLC6A4	P31645	ENSG00000108576	Human	lofepramine	7551	223365887	t	Inhibitor	Inhibition		Not Determined	f	t		pKi		7.21999979		Ki		60		#ERROR!				9537821
SERT	928	SLC6A4	P31645	ENSG00000108576	Human	milnacipran	7436	178104008	t	Inhibitor	Inhibition		Not Determined	f	t		pIC50		7.300

target	target_id	target_gene	target_uniprot	target_ensembl_gene_id	target_species	ligand	ligand_id	ligand_pubchen	approved_drug	type	action	action_commen	selectivity	endogenous	primary_target	concentration_r	affinity_units	affinity_high	affinity_median	affinity_low	original_affinity	original_affinity	original_affinity	original_affinity	original_affinity	assay_descripti	receptor_site	ligand_context	pubmed_id
SERT	928	SLC6A4	P31645	ENSG00000108576	Human	nortriptyline	2404	135650736	t	Inhibitor	Inhibition		Non-selective	f	t		pKi		8.159999847		Ki		6.980000019		#ERROR!				NULL
SERT	928	SLC6A4	P31645	ENSG00000108576	Human	paroxetine	4790	178101492	t	Inhibitor	Inhibition		Selective	f	t		pKi		9.600000381		Ki		0.25		#ERROR!				9537821
SERT	928	SLC6A4	P31645	ENSG00000108576	Human	paroxetine	4790	178101492	t	Inhibitor	Inhibition		Not Determined	f	t		pKd		10.10000038		Kd		0.079999998		#ERROR!				27049939
SERT	928	SLC6A4	P31645	ENSG00000108576	Human	protriptyline	7285	178103859	t	Inhibitor	Inhibition		Selective	f	t		pKd		7.710000038		Kd		19.60000038		#ERROR!				9537821
SERT	928	SLC6A4	P31645	ENSG00000108576	Human	sertraline	4798	178101500	t	Inhibitor	Inhibition		Selective	f	t		pKi		9.100000381		Ki		0.790000021		#ERROR!				NULL
SERT	928	SLC6A4	P31645	ENSG00000108576	Human	trimipramine	7317	178103890	t	Inhibitor	Inhibition		Non-selective	f	t		pKi		6.829999924		Ki		149		#ERROR!				9537821
SERT	928	SLC6A4	P31645	ENSG00000108576	Human	venlafaxine	7321	178103894	t	Inhibitor	Inhibition		Non-selective	f	t		pIC50		7.570000172		IC50		27		#ERROR!				20378347
SERT	928	SLC6A4	P31645	ENSG00000108576	Human	vilazodone	7427	178103999	t	Inhibitor	Inhibition		Non-selective	f	t		pIC50	9.300000191		8.800000191	IC50	1.5800000043		0.5	#ERROR!				19499624|15341484
SERT	928	SLC6A4	P31645	ENSG00000108576	Human	vortioxetine	7351	178103923	t	Inhibitor	Inhibition		Non-selective	f	t		pKi		8.800000191		Ki		1.600000024		#ERROR!				21486038
sigma non-opioid intracellular receptor 1	2552	SIGMAR1	Q99720	ENSG00000147955	Human	(-)-pentazocine	1606	135650806	t	Antagonist	Antagonist		Non-selective	f	t		-				-								NULL
SMO	239	SMO	Q99835	ENSG00000128602	Human	glasdegib	8201	249565881	t	Antagonist	Antagonist		Not Determined	f	t		pIC50		8.300000191		IC50		5		#ERROR!				24900436
SMO	239	SMO	Q99835	ENSG00000128602	Human	sonidegib	8199	249565879	t	Antagonist	Antagonist		Not Determined	f	t		pKi		8.220000267		Ki		6		#ERROR!				23063522
SMO	239	SMO	Q99835	ENSG00000128602	Human	vismodegib	6975	178103554	t	Antagonist	Antagonist		Selective	f	t		pKi		7.799999962		Ki		16.20000076		#ERROR!				23063522
Sodium/glucose cotransporter 1	915	SLC5A1	P13866	ENSG00000100170	Human	sotagliflozin	8312	252166524	t	Inhibitor	Inhibition		Not Determined	f	t		pIC50		7.440000057		IC50		36		#ERROR!				22739142
Sodium/glucose cotransporter 2	916	SLC5A2	P31639	ENSG00000140675	Human	canagliflozin	4582	178101303	t	Inhibitor	Inhibition		Non-selective	f	t		pIC50		8.399999619		IC50		4.400000095		#ERROR!				22355316
Sodium/glucose cotransporter 2	916	SLC5A2	P31639	ENSG00000140675	Human	dapagliflozin	4594	178101315	t	Inhibitor	Inhibition		Selective	f	t		pIC50		9.300000191		IC50		0.49000001		#ERROR!				20637636
Sodium/glucose cotransporter 2	916	SLC5A2	P31639	ENSG00000140675	Human	empagliflozin	4754	178101464	t	Inhibitor	Inhibition		Non-selective	f	t		pIC50		8.5		IC50		3.200000048		#ERROR!				21985634
Sodium/glucose cotransporter 2	916	SLC5A2	P31639	ENSG00000140675	Human	ertugliflozin	8376	252166588	t	Inhibitor	Inhibition		Selective	f	t		pIC50		9.06000042		IC50		0.876999974		#ERROR!				21449606
Sodium/glucose cotransporter 2	916	SLC5A2	P31639	ENSG00000140675	Human	ipragliflozin	9394	328033496	t	Inhibitor	Inhibition		Selective	f	t		pIC50		8.130000114		IC50		7.380000114		#ERROR!	Measuring inhibition of sodium-dependent [<sup>14			22139434
Sodium/glucose cotransporter 2	916	SLC5A2	P31639	ENSG00000140675	Human	sotagliflozin	8312	252166524	t	Inhibitor	Inhibition		Not Determined	f	t		pIC50		8.739999771		IC50		1.799999952		#ERROR!				22739142
Sodium/glucose cotransporter 2	916	SLC5A2	P31639	ENSG00000140675	Human	tofogliflozin	9395	328083497	t	Inhibitor	Inhibition		Selective	f	t		pIC50		8.539999962		IC50		2.900000095		#ERROR!	Measuring inhibition of sodium-dependent [<sup>14			22889351
sodium/potassium-transporting ATPase su	833	ATP1A1	P05023	ENSG00000163399	Human	acetyldigitoxin	6794	178103400	t	Inhibitor	Inhibition		Family selective	f	t		-				-				#ERROR!				NULL
sodium/potassium-transporting ATPase su	833	ATP1A1	P05023	ENSG00000163399	Human	deslanoside	6806	178103412	t	Inhibitor	Inhibition		Family selective	f	t		-				-				#ERROR!				NULL
sodium/potassium-transporting ATPase su	833	ATP1A1	P05023	ENSG00000163399	Human	digitoxin	6782	178103388	t	Inhibitor	Inhibition		Family selective	f	t		-				-				#ERROR!				NULL
sodium/potassium-transporting ATPase su	833	ATP1A1	P05023	ENSG00000163399	Human	digoxin	4726	178101437	t	Inhibitor	Inhibition		Family selective	f	t		-				-				#ERROR!				NULL
SST₁ receptor	355	SSTR1	P30872	ENSG00000139874	Human	pasireotide	2018	135652473	t	Agonist	Full agonist		Not Determined	f	t		pIC50		8		-				#ERROR!				15477717
SST₂ receptor	356	SSTR2	P30874	ENSG00000180616	Human	lanreotide	2031	135652022	t	Agonist	Full agonist		Selective	f	t		pKi		9.600000381		8.699999809	-			#ERROR!				8769372|9650799|10598788|7988476|9652348
SST₂ receptor	356	SSTR2	P30874	ENSG00000180616	Human	pasireotide	2018	135652473	t	Agonist	Full agonist		Not Determined	f	t		pIC50		9		-				#ERROR!				15477717
SST₂ receptor	356	SSTR2	P30874	ENSG00000180616	Human	vapreotide	2054	135652433	t	Agonist	Full agonist		Not Determined	f	t		pKi		10.10000038		8.300000191	-			#ERROR!				8769372|7988476
SST₃ receptor	357	SSTR3	P32745	ENSG00000183473	Human	pasireotide	2018	135652473	t	Agonist	Full agonist		Not Determined	f	t		pIC50		8.800000191		-				#ERROR!				15477717
SST₃ receptor	357	SSTR3	P32745	ENSG00000183473	Human	vapreotide	2054	135652433	t	Agonist	Full agonist		Not Determined	f	t		pKi		7.900000095		7.400000095	-			#ERROR!				8769372|9650799|7988476
SST₅ receptor	359	SSTR5	P35346	ENSG00000162009	Human	lanreotide	2031	135652022	t	Agonist	Full agonist		Not Determined	f	t		pKi		9.300000191		7.400000095	-			#ERROR!				8769372|9650799|10598788|9600011|7988476|965
SST₅ receptor	359	SSTR5	P35346	ENSG00000162009	Human	pasireotide	2018	135652473	t	Agonist	Full agonist		Not Determined	f	t		pIC50		9.800000191		-				#ERROR!				15477717
SST₅ receptor	359	SSTR5	P35346	ENSG00000162009	Human	vapreotide	2054	135652433	t	Agonist	Full agonist		Not Determined	f	t		pKi		9.199999809		7.300000191	-			#ERROR!				8769372|9650799|10598788|9600011|7988476|965
steroid 5 alpha-reductase 2	2633	SRD5A2	P31213	ENSG00000049319	Human	finasteride	6818	178103424	t	Inhibitor	Inhibition		Not Determined	f	t		pIC50		7.840000153		IC50		14.60000038		#ERROR!				3783591
Thrombopoietin receptor	1722	MPL	P40238	ENSG00000117400	Human	eltrombopag	6961	178103542	t	Agonist	Agonist		Selective	f	t		pEC50		7.420000076		EC50		38		#ERROR!				18783949
Thrombopoietin receptor	1722	MPL	P40238	ENSG00000117400	Human	lusutrombopag	10032	375973220	t	Agonist	Agonist		Selective	f	t		-				-				#ERROR!				29274361
Thrombopoietin receptor	1722	MPL	P40238	ENSG00000117400	Human	romiplostim	6974	178103553	t	Agonist	Agonist		Not Determined	f	t		-				-				#ERROR!				NULL
thymidylate synthetase	2642	TYMS	P04818	ENSG00000176890	Human	capecitabine	6799	178103405	t	Inhibitor	Inhibition		Not Determined	f	t		-				-				#ERROR!				10631692|15709193
thymidylate synthetase	2642	TYMS	P04818	ENSG00000176890	Human	raltitrexed	7403	178103975	t	Inhibitor	Inhibition		Selective	f	t		pIC50		6.539999962		IC50		290		#ERROR!				22739090
Thyroid hormone receptor-α	588	THRA	P10827	ENSG00000126351	Human	dextrothyroxine	6951	178103533	t	Agonist	Agonist		Not Determined	f	t		-				-				#ERROR!				6777394
Thyroid hormone receptor-β	589	THRB	P10828	ENSG00000151090	Human	dextrothyroxine	6951	178103533	t	Agonist	Agonist		Not Determined	f	t		-				-				#ERROR!				6777394
thyroid peroxidase	2526	TPO	P07202	ENSG00000115705	Human	methimazole	6649	178103262	t	Inhibitor	Inhibition		Not Determined	f	t		-				-				#ERROR!				748042
thyroid peroxidase	2526	TPO	P07202	ENSG00000115705	Human	propylthiouracil	6650	178103263	t	Inhibitor	Inhibition		Not Determined	f	t		-				-				#ERROR!				748042
TLR7	1757	TLR7	Q9NYK1	ENSG00000196664	Human	imiquimod	5003	178101700	t	Agonist	Agonist		Non-selective	f	t		pEC50		4.940000057		EC50		11400		#ERROR!				29152046
TLR7	1757	TLR7	Q9NYK1	ENSG00000196664	Human	hydroxychloroquine	7198	178103773	t	Antagonist	Antagonist		Not Determined	f	t		pIC50		5.559999943		IC50		2780		#ERROR!	Inhibition of hTLR7 R848-induced luciferase reporte			24342772
TLR9	1759	TLR9	Q9NR96	ENSG00000173366	Human	hydroxychloroquine	7198	178103773	t	Antagonist	Antagonist		Not Determined	f	t		pIC50		7.099999905		IC50		80		#ERROR!	Inhibition of hTLR9 CpG2006-induced luciferase rep			24342772
TRPM8	500	CZZ2W7		ENSG00000144481	Human	(-)-menthol	2430	135651519	t	Activator	None	inhibited by intrac	Not Determined	f	t		pEC50		4.599999905		EC50		25100		#ERROR!				15306801
TRPV1	507	TRPV1	Q8NER1	ENSG00000196689	Human	capsaicin	2486	53801090	t	Activator	Agonist		Non-selective	f	t		pEC50		7.5		EC50		31.60000038		#ERROR!				15306801
tryptase alpha/beta 1	2424	TPSAB1	Q15661	ENSG00000172236	Human	gabexate	7863	223366193	t	Inhibitor	Inhibition		Selective	f	t		pIC50		8.470000267		IC50		3.400000095		#ERROR!				11172730
tryptase alpha/beta 1	2424	TPSAB1	Q15661	ENSG00000172236	Human	nafamostat	4262	178101079	t	Inhibitor	Inhibition		Not Determined	f	t		pIC50		10		IC50		0.096299996		#ERROR!				12939527
tubulin beta class I	2640	TUBB	P07437	ENSG00000196230	Human	cabazitaxel	6798	178103404	t	Inhibitor	Inhibition		Not Determined	f	t		-				-				#ERROR!				NULL
tubulin beta class I	2640	TUBB	P07437	ENSG00000196230	Human	colchicine	2367	135650118	t	Inhibitor	Inhibition		Not Determined	f	t		pIC50		7.960000038		IC50		11		#ERROR!				16504507
tubulin beta class I	2640	TUBB	P07437	ENSG00000196230	Human	docetaxel	6809	178103415	t	Inhibitor	Inhibition		Not Determined	f	t		-				-				#ERROR!				NULL
tubulin beta class I	2640	TUBB	P07437	ENSG00000196230	Human	eribulin	6813	178103419	t	Inhibitor	Inhibition		Not Determined	f	t		pIC50		8.229999542		IC50		5.900000095		#ERROR!				21324687
tubulin beta class I	2640	TUBB	P07437	ENSG00000196230	Human	ixabepilone	6824	178103430	t	Inhibitor	Inhibition		Not Determined	f	t		-				-				#ERROR!				NULL
tubulin beta class I	2640	TUBB	P07437	ENSG00000196230	Human	paclitaxel	2770	135652721	t	Inhibitor	Inhibition	Mitotic cell cycle i	Not Determined	f	t		pEC50		8.050000191		EC50		9		#ERROR!				16377187
tubulin beta class I	2640	TUBB	P07437	ENSG00000196230	Human	vinblastine	6851	178103457	t	Inhibitor	Inhibition		Non-selective	f	t		pIC50		9		IC50		1		#ERROR!				NULL
tubulin beta class I	2640	TUBB	P07437	ENSG00000196230	Human	vincristine	6785	178103391	t	Inhibitor	Inhibition		Non-selective	f	t		-				-				#ERROR!				NULL
tumor associated calcium signal transduci	2837	TACSTD2	P09758	ENSG00000184292	Human	sacituzumab govitteca	8253	249565933	t	Antibody	Binding		Not Determined	f	t		-				-				#ERROR!				NULL
UDP-glucose ceramide glucosyltransferas	2528	UGCG	Q16739	ENSG00000148154	Human	miglustat	4841	178101543	t	Inhibitor	Inhibition		Not Determined	f	t		pKi		5.130000114		Ki		7400		#ERROR!				NULL
Urate anion exchanger 1	1031	SLC22A12	Q96S37	ENSG00000197891	Human	sulfinpyrazone	5826	178102451	t	Inhibitor	Inhibition		Selective	f	t		pIC50		4		IC50		100000		#ERROR!				17325024
Vesicular monoamine transporter 2	1012	SLC18A2	Q05940	ENSG00000165646	Human	deutetrabenazine	8707	252827365	t	Inhibitor	Inhibition		Not Determined	f	t		-				-				#ERROR!				NULL
Vesicular monoamine transporter 2	1012	SLC18A2	Q05940	ENSG00000165646	Human	reserpine	4823	178101525	t	Inhibitor	Inhibition		Non-selective	f	t		pKi		7.900000095		Ki		12		#ERROR!				8643547
Vesicular monoamine transporter 2	1012	SLC18A2	Q05940	ENSG00000165646	Human	tetrabenazine	4834	178101536	t	Inhibitor	Inhibition		Non-selective	f	t		pKi		7		Ki		97		#ERROR!				8643547
Vitamin D receptor	605	VDR	P11473	ENSG00000111424	Human	calcipotriol	2778	135650047	t	Agonist	Agonist		Not Determined	f	t		pKd		9.300000191		-				#ERROR!				7976510
Vitamin D receptor	605	VDR	P11473	ENSG00000111424	Human	paricalcitol	2791	135651536	t	Agonist	Agonist		Not Determined	f	t		-				-				#ERROR!				NULL
vitamin K epoxide reductase complex sub	2645	VKORC1	Q9BQB6	ENSG00000167397	Human	alitretinoin	2645	135651418	t	Inhibitor	Inhibition		Not Determined	f	t		-				-				#ERROR!				10327215|18374197
vitamin K epoxide reductase complex sub	2645	VKORC1	Q9BQB6	ENSG00000167397	Human	dicumarol	6808	178103414	t	Inhibitor	Inhibition		Not Determined	f	t		-				-				#ERROR!				18610395
vitamin K epoxide reductase complex sub	2645	VKORC1	Q9BQB6	ENSG00000167397	Human	phenindione	6838	178103444	t	Inhibitor	Inhibition		Not Determined	f	t		-				-				#ERROR!				12978288|13037763
vitamin K epoxide reductase complex sub	2645	VKORC1	Q9BQB6	ENSG00000167397	Human	phenprocoumon	6839	178103445	t	Inhibitor	Inhibition		Not Determined	f	t		-				-				#ERROR!				13037763
vitamin K epoxide reductase complex sub	2645	VKORC1	Q9BQB6	ENSG00000167397	Human	warfarin	6853	178103459	t	Inhibitor	Inhibition		Selective	f	t		pKi		5.579999924		Ki		2633		#ERROR!				23618698
V_{1A} receptor	366	AVPR1A	P37288	ENSG00000166148	Human	conivaptan	2203	135651340	t	Antagonist	Antagonist		Not Determined	f	t		pKi	8.399999619		8.199999809	-				#ERROR!				9884194|9957
V₂ receptor	368	AVPR2	P30518	ENSG00000126895	Human	conivaptan	2203	135651340	t	Antagonist	Antagonist		Selective	f	t		pKi		9.43999958		Ki		0.360000014		#ERROR!				20471258
V₂ receptor	368	AVPR2	P30518	ENSG00000126895	Human	tolvaptan	2226	135651270	t	Antagonist	Antagonist		Selective	f	t		pKi		9.369999886		Ki		0.430000007		#ERROR!				9864265
xanthine dehydrogenase	2646	XDH	P47989	ENSG00000158125	Human	allopurinol	6795	178103401	t	Inhibitor	Inhibition		Not Determined	f	t		pKi		5.150000095		Ki		7000		#ERROR!				8691450
xanthine dehydrogenase	2646	XDH	P47989	ENSG00000158125	Human	allopurinol	6795	178103401	t	Inhibitor	Inhibition		Not Determined	f	t		pIC50		5.380000114		IC50		4200		#ERROR!				8691450
NULL		NULL	NULL		Human	eptinezumab	10694	404859141	t	Antibody	Binding		Selective	f	t		pEC50		9.989999771		EC50		0.103		#ERROR!	Binding affinity for α-CGRP determined in an I	NULL		NULL
NULL		NULL	NULL		Human	fremanezumab	9208	315661284	t	Antibody	Binding		Selective	f	t		pKd		10.30000019		Kd		0.050000001	<			NULL		NULL
NULL		NULL	NULL		Human	galcanezumab	8967	310264747	t	Antibody	Binding		Selective	f	t		pKd		8.659999847		Kd		2.200000048		#ERROR!				NULL
NULL		NULL	NULL		Human	florbetaben	7769	223366102	t	None	Binding	PET imaging age	Selective	f	t		pKi		8.649999619		Ki		2.220000029		#ERROR!	Inhibition of ¹⁸F-AV-45 binding to A&be			19837759
NULL		NULL	NULL		Human	florbetapir	7344	178103916	t	None	Binding	PET imaging age	Selective	f	t		pKi		8.539999962		Ki		2.869999886		#ERROR!				23047223
NULL		NULL	NULL		Human	flutemetamol	7342	178103914	t	None	Binding	PET imaging age	Selective	f	t		pKi		9.130000114		Ki		0.74000001		#ERROR!				22011144
NULL		NULL	NULL		Human	belimumab	6887	178103469	t	Antibody	Binding		Selective	f	t		pIC50		10.61999989		EC50		0.024		#ERROR!	Binding affinity to immobilised BLyS			14613291
NULL		NULL	NULL		Human	eculizumab	6884	178103466	t	Antibody	Binding		Selective	f	t		pKd		10.52000046		Kd		0.029999999		#ERROR!				NULL
NULL		NULL	NULL		Human	eculizumab	6884	178103466	t	Antibody	Binding		Selective	f	t		pKd		9.220000267		Kd		0.600000024		#ERROR!	At pH6.0 in a SPR assay.			29649283
NULL		NULL	NULL		Human	eculizumab	6884	178103466	t	Antibody	Binding		Selective	f	t		pKd		10.52000046		Kd		0.029999999		#ERROR!	At pH7.4 in a SPR assay.			29649283
NULL		NULL	NULL		Human	ravulizumab	10168	381118878	t	Antibody	Binding		Selective	f	t		pKd		7.659999847		Kd		22		#ERROR!	At pH6.0 in a SPR assay.			29649283
NULL		NULL	NULL		Human	ravulizumab	10168	381118878	t	Antibody	Binding		Selective	f	t		pKd		9.31000042		Kd		0.49000001		#ERROR!	At pH7.4 in a SPR assay.			29649283
NULL		NULL	NULL		Human	burosumab	9292	318164811	t	Antibody	Binding		Selective	f	t		-				-				#ERROR!				NULL
NULL		NULL	NULL		Human	emapalumab	9295	318164814	t	Antibody	Binding		Selective	f	t		pKd		9.140000343		Kd		0.731999993		#ERROR!	Measured using Biacore technology			NULL
NULL		NULL	NULL		Human	ustekinumab	6885	178103467	t	Antibody	Inhibition		Selective	f	t		pKd	11.14999962		9	Kd	1		0.007	<					NULL
NULL		NULL	NULL		Human	ixekizumab	7541	187051842	t	Antibody	Binding		Selective	f	t		pKd		11.69999981		Kd		0.002	<		Binding affinity of mAbs 104, 121 and 126 from pate	NULL		NULL
NULL		NULL	NULL		Human	secukinumab	8078	249565758	t	Antibody	Binding		Selective	f	t		pIC50		8.680000305		IC50		2.099999905		#ERROR!	Inhibition of IL-17-induced IL-6 production by human	NULL		NULL
NULL		NULL	NULL		Human	secukinumab	8078	249565758	t	Antibody	Binding		Selective	f	t		pKd		9.909999847		Kd		0.122000001		#ERROR!	Binding affinity for recombinant human IL-17	NULL		NULL
NULL		NULL	NULL		Human	canakinumab	6773	178103379	t	Antibody	Binding		Selective	f	t		pKd		10.46000004		Kd		0.035		#ERROR!				NULL
NULL		NULL	NULL		Human	ribonacept	6790	178103396	t	Antibody	Binding		Selective	f	t		-				-				#ERROR!				18668535|19649332
NULL		NULL	NULL		Human	risankizumab	8922	310264703	t	Antibody	Binding		Selective	f	t		pKd		11		Kd		0.01	<					25905918
NULL		NULL	NULL		Human	tildrakizumab	8093	249565773	t	Antibody	Binding		Selective	f	t		pKd		9.529999733		Kd		0.296999991		#ERROR!	Binding affinity constant for clone 13B8-b CDRH2-b.	NULL		NULL
NULL		NULL	NULL		Human	mepolizumab	7685	223366020	t	Antibody	Binding		Selective	f	t		pKd		11.38000011		Kd		0.0042		#ERROR!	Binding affinity for recombinant human IL-5.			11496242
NULL		NULL	NULL		Human	reslizumab	8091	249565771	t	Antibody	Binding		Selective	f	t		pKd		8.650000191		Kd		8.899999619		#ERROR!				NULL
NULL		NULL	NULL		Human	siltuximab	7396	178103968	t	Antibody	Binding		Not Determined	f	t		pKd		10.23999977		Kd		0.057		#ERROR!				NULL
NULL		NULL	NULL		Human	atezolizumab	7990	249565671	t	Antibody	Binding		Selective	f	t		pKd	10		9	Kd	1		0.100000001	#ERROR!				NULL
NULL		NULL	NULL		Human	avelumab	8735	252827393	t	Antibody	Binding		Selective	f	t		pIC50		10.14999962		IC50		0.071000002		#ERROR!	Measuring inhibition of PD-1 binding to PD-L1.			NULL
NULL		NULL	NULL		Human	durvalumab	7985	249565666	t	Antibody	Binding		Selective	f	t		pKd		8.699999809		Kd		2		#ERROR!				NULL
NULL		NULL	NULL		Human	denosumab	6886	178103468	t	Antibody	Binding		Not Determined	f	t		pIC50		9.640000343		IC50		0.230000004	<					NULL
NULL		NULL	NULL		Human	romosozumab	8092	249565772	t	Antibody	Binding		Selective	f	t		pKd	12.22000027		11.22000027	Kd	0.006		0.0006	#ERROR!	Affinity across a range of antibodies tested, as the le			NULL
NULL		NULL	NULL		Human	IL-2	4985	178101682	t	None	Binding		Not Determined	f	t		pIC50		7.21999979		IC50		60		#ERROR!				12656598
NULL		NULL	NULL		Human	etanercept	6789	178103395	t	Antagonist	Inhibition		Selective	f	t		-				-				#ERROR!				9865320|10090426
NULL		NULL	NULL		Human	adalimumab	4860	178101561	t	Antibody	Binding		Selective	f	t		pKd		8		Kd		10		#ERROR!				NULL
NULL		NULL	NULL		Human	certolizumab pegol	6774	178103380	t	Antibody	Binding		Selective	f	t		-				-				#ERROR!	Note IC90 value of 3ng/ml is provided			17636564
NULL		NULL	NULL		Human	golimumab	6776	178103382	t	Antibody	Binding		Selective	f	t		pIC50		10.73999977		IC50		0.017999999		#ERROR!	Affinity for soluble TNFα measured by surface			20519961
NULL		NULL	NULL		Human	infliximab	5004	178101701	t	Antibody	Binding		Selective	f	t		pKd		8.739999771		Kd		1.799999952		#ERROR!	Binding affinity of antibody cA2 to recombinant huma	NULL		NULL
NULL		NULL	NULL		Human	aflibercept	6786	178103392	t	Antagonist	Inhibition		Selective	f	t		-				-				#ERROR!				24179482|24144450|19236257
NULL		NULL	NULL		Human	pegaptanib	6836	178103442	t	Antagonist	Inhibition		Selective	f	t		-				-				#ERROR!				15623532|15660527
NULL		NULL	NULL		Human	bevacizumab	6771	178103377	t	Antibody	Binding		Selective	f	t		pIC50	8.300000191		8	IC50	10		5	#ERROR!				NULL
NULL		NULL	NULL		Human	brolucizumab	8713	252827371	t	Antibody	Binding		Selective	f	t		pKd		8.800000191		Kd		1.590000033		#ERROR!				NULL
NULL		NULL	NULL		Human	ranibizumab	6779	178103385	t	Antibody	Binding		Selective	f	t		pKd		9.819999695		Kd		0.150000006		#ERROR!				NULL
NULL		NULL	NULL		Human	caplacizumab	9123	315661208	t	Antibody	Binding		Selective	f	t		pKd		8.520000458		Kd		3.099999905		#ERROR!				NULL

C12. Disease-specific gene expression

How to handle disease-specific gene expression

Here is the short question, before my lengthy explanation- can there be a disease reaction (ie it has a disease tag on it) that has no specific disease participants in it ie again -all the participants are genetically WT? Can a complex or an entity be labeled with a disease tag if there is no genetic modification to it?

In this particular case, my fusion protein NPM-ALK (disease tagged) is known to phosphorylate STAT3, which goes on to activate (disease-specific) gene expression. STAT3 itself is wild-type, and it is also believed to be activated by WT ALK.

In the WT ALK signaling pathway, STAT3 is shown as being phosphorylated, dimerizing and translocating to nucleus, where it does nuclear gene expression things.

In the disease pathway, STAT 3 is shown as binding and being phosphorylated by ALK fusions , but the dimerization and translocation are not specifically shown. In earlier annotations, STAT3 is just shown as a regulator of gene expression events, which avoids the issue that I am now trying to work around, below.

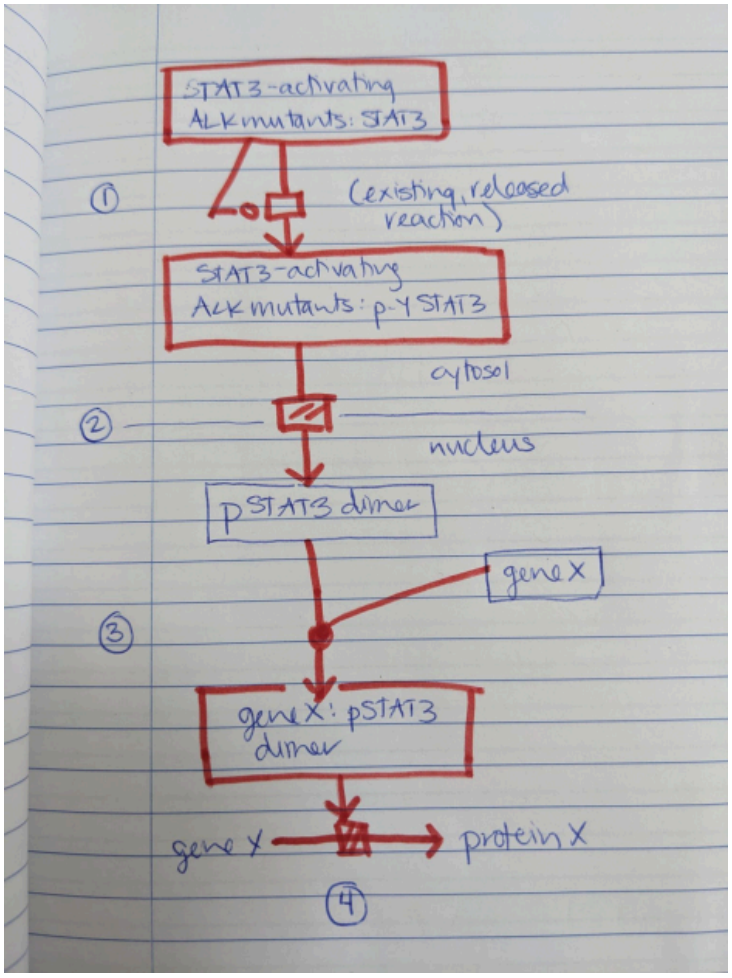
I now need to annotate an event where, in the context of NPM-ALK signaling, phosphorylated STAT3 binds to a gene promoter to promote expression, so I need the physical entity (which is still just WT) in the nucleus. I was just about to make a BBE showing that after phosphorylation in the cytosol by ALK fusions (disease tagged reaction), STAT3 is released, dimerizes and translocates to the nucleus, where it will be free to bind to any gene I want it to regulate.

The problem is that none of those components of my proposed downstream reactions (nuclear p-STAT3, the genes it is going to regulate) are inherently disease in themselves.

1. Can I make my BBE showing a cytosolic ALK mutant: p_Y STAT3 complex (disease tagged) yielding nuclear pY-STAT3 (in reality WT, but needed as a disease entity)? 2. After that, can I use this WT-but-disease activated pY-STAT3 to activate WT-but disease-activated genes? This reaction would need to be tagged with a disease tag even though all the entities are WT.

2. Alternately, I could make a disease-tagged p-STAT3 EWAS and dimer complex, even though they are genetically WT, and use that in my disease-specific reactions. In this model, the disease reactions are consistent (the reactions contain an entity with a disease tag on it), but the entity is inconsistent (no genetic modification, but tagged with a disease tag).

3. As I was doing this curation, I found myself reconsidering whether the complex formed by the (WT) gene and (WT) transcription factor (in this specific instance, the p-Y STAT3:geneX complex) should really be disease labelled, as indicated in both schematics, or whether just the gene expression that is regulated by that complex should be.



Discussion:

Why do you need this phosphorylated STAT3 and its downstream reactions to be disease tagged:

1. Is it phosphorylated on particular residues by the NMP-ALK fusion protein that aren't normally phosphorylated by ALK?
2. Is such a high quantity of phosphorylated STAT3 produced that it saturates all STAT3 binding sites in the genome and starts binding to regulatory elements of genes that it normally does not regulate?

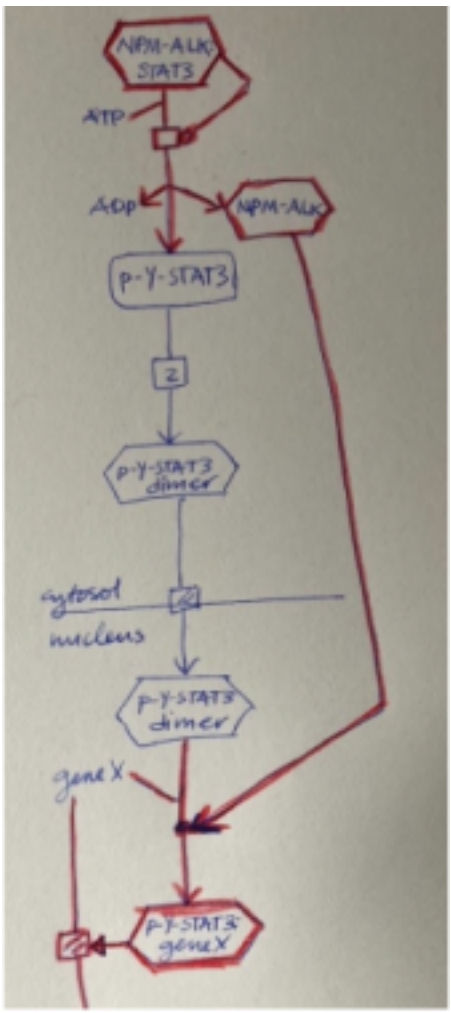
If reason 1 is in place, then a disease-specific phosphorylated STAT3 could be tagged with a disease tag, and any aberrant transcriptional events that it regulates would be tagged by a disease tag (but not products of these transcriptional events, as they are normal proteins). I had a similar case in "Deregulated CDK5 triggers multiple neurodegenerative pathways in Alzheimer's disease models", where e.g. transcription factor FOXO3 is aberrantly phosphorylated (reaction 8870558) by a disease-specific complex CDK5:p25, producing a disease-specific FOXO3 phosphorylation product, which then stimulates the expression of pro-apoptotic genes, e.g. BCL2L11.

If reason 2 is in place, then a complex of STAT3 and a target gene that it normally does not bind to would be tagged with a disease tag. The black box event of gene expression regulated by this complex would also be disease tagged, but not the product of this black box event. This

would be similar to tagging the ERBB2 homodimer and the reactions it takes part in with disease tags. Although ERBB2 homodimer consists of two normal ERBB2 proteins, it is disease-tagged because it is not normally present in the cell (normally low level of ERBB2 expression prevents homodimerization; homodimerization happens when ERBB2 is amplified or overexpressed in cancer).

If neither 1) nor 2) are true, then I don't see the need for disease tags after STAT3 phosphorylation. It would be the case of disease-specific pathway merging with the normal pathway, as is the case with other gain-of-function pathways. If Reactome does not have annotations of downstream STAT3 targets that are relevant for ALK signaling, they should be included as part of the normal pathway (an imaginary "Transcriptional regulation by STAT3", and the disease pathway should point to this pathway. If the comprehensive annotation of such normal pathway is out of scope, perhaps a portion of it can be annotated as a starter and the remainder filled in later, or they should be included in the normal ALK signaling diagram for the time being.

Conclusion:
A disease event must involve at least one entity that is disease-tagged (and that is not genetically normal). The best solution in this case is to have the NPM-ALK fusion (legitimately disease tagged) as a positive regulator of either or both of the downstream (WT entities only) STAT3-gene binding event and the (WT entities only) gene expression event that it stimulates, as shown, and to mark the events with a disease tag.



C13.
Deep extraction of events

Deep extracting events in a local project

In the **Event Browser** view, select the pathway that contains all of the events that you want to include. It is best to select the smallest pathway that still contains all the events that you'd like to review.

Check this out into a local project.
To extract “top down”:

1. Select all ReactionLikeEvents (RLE) right click and select “update from database”(*make a note of # of RLE instances*)
2. If number of RLEs increases in the “update” you'll need to select all and update again.
3. Repeat this process until the number of RLEs stabilizes *A total of 2 rounds of updates should extract all component reactions as well as their immediate referrers ...e.g reactions in the “inferrredFrom” slots.*
4. Select all CatalystActivities, right click and select “update from database”
5. Select all EntitySets (*note starting #*) right click and select “update from database”. Repeat until the number of sets stabilizes (*note the **final** # of sets*).
6. Select all Complexes and update as described above. After you've updated complexes, recheck # of set. If sets have increased, repeat step 5 on sets. Then recheck complexes. Repeat this loop until neither number increases (*note the **final** # of sets*).
7. Select all GenomeEncodedEntity update from DB.
8. Select all Polymer and update (check number of EntitySets and Complexes to see if they'e increased. If so, repeat step 5or 6 as necessary).
9. Select “OtherEntity” and update from DB.
10. Select SimpleEntity and update from DB.

Project extraction should now be complete. Don't forget to save!

V90_Sept 2024

- Moved from DevWIki

C14. Remote Attribute Search Tool

This tool can be used to search within a data class for instances with specified attributes (or with specified attributes of attributes). For example you can search for all reactions whose input includes complexes containing a given ReferenceEntity.

To search Reactome curator, use (new):
http://curator.reactome.org/cgi-bin/remoteattsearch2?DB=gk_central

To search reactomerelease, use (new);
https://release.reactome.org/cgi-bin/remoteattsearch2?DB=slice_test

Non-functional Remote Attribute Search on gk_central and live site (2024)
[To search gk_central, use:
http://reactomerelease.oicr.on.ca/cgi-bin/remoteattsearch2

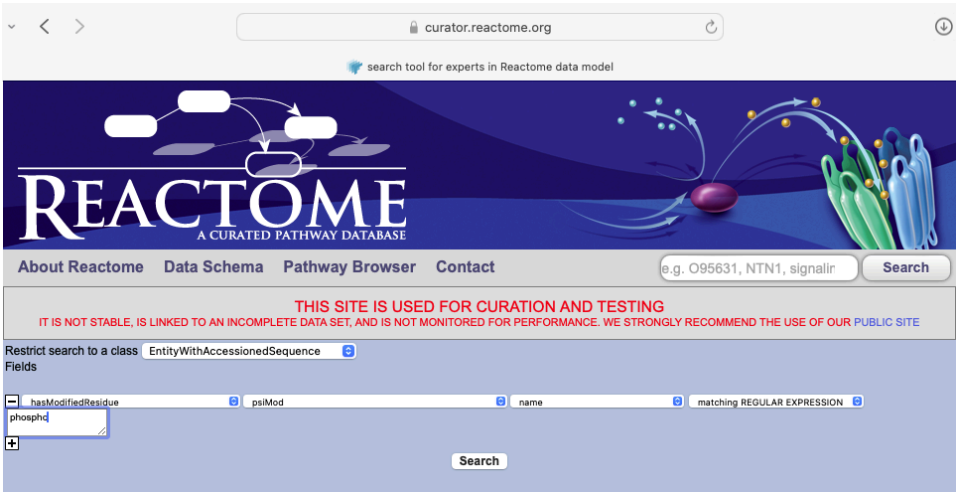
To search the live site, use:
http://www.reactome.org/cgi-bin/remoteattsearch2?DB=gk_current]

Examples of how to use the remoteattsearch tool can be found below.

Remoteattsearch examples

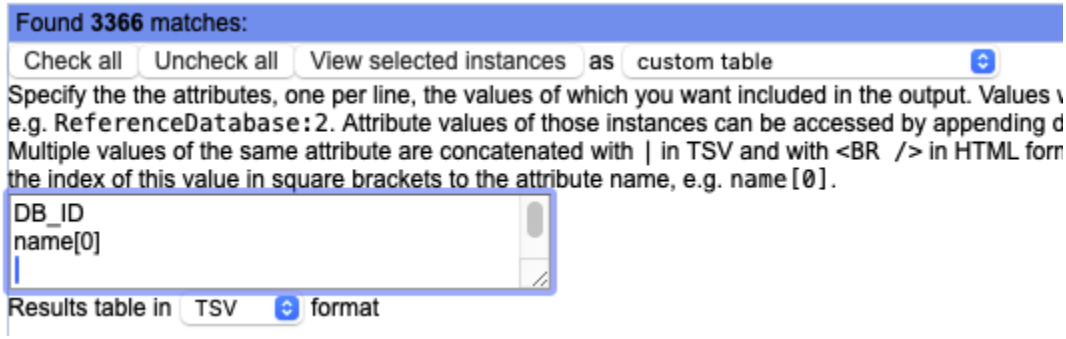
1. Simple

class - EWAS
hasModifiedResidue
psiMod
name (regex) phospho



Produces 3366 matches on curator (Sept 2024).

This lists EWASs as that is the class you queried.
Choose to display results as **custom table** showing:
DB_ID
name[0] (this limits to 1st name) *or*
name (this prints off all names separated by pipes)



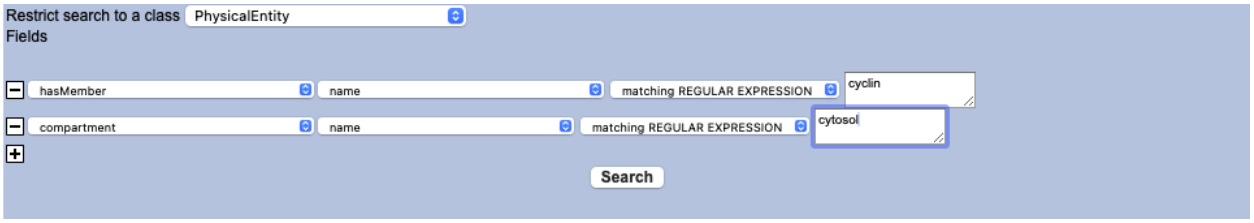
Once you have listed the attributes you wish to have displayed, click the “View selected instances” to generate the results.

Like this:

```
#DB_ID  #name[0]
9700461 p-7Y ALK Y1278S
9909362 p-S-PER3
206197  Phospho (2 sites) Erk2
206159  Phospho ARM
2197668 p-Y291-WAS
442924  phosphorylated MAPKAPK3
9656067 p-S338,Y341,T492,S494,S621 RAF1 S259T
420350  pFyn(Y531)
9860512 p-Y464-PIK3R2
2454244 p-Y65-FCER1G
1996341 p-6Y-EGFR A289T
8957037 p-AMELX
9009184 p-S286,S304-Runx2-P2
```

2. ADD an additional filter example

class - PhysicalEntity
hasMember
name (regex) - cyclin -> 83 matches
AND
compartment
name (regex) - cytosol -> 30 matches



3. Upstream example

class EWAS
hasComponent:Complex
name regex polymerase
-> 139 matches
AND
species name regex sapiens
-> 63 matches

V90_Sept 2024

- moved from DevWiki
- examples updated with images
- grey out references to gk_central and live site remote attribute search

C15. Remove duplicates MSExcel

Identifying list members that are unique to one of two lists using Microsoft Excel

This is a procedure describing how to take two lists and compare entries to identify those that are present in only one of the two lists. This procedure can be useful, for example, to compare the list of proteins in the gk_central vs. live site to find those that are unreleased.

Identifying list members that are unique

1. Create a new file that contains two columns to be compared.
2. Select the cell ranges (both columns).
3. Under conditonal formatting, choose "Highlight Cell Rules" then "Duplicate Values"
4. Format cell that contain...."Duplicate" values with a colour. Click OK
5.Under "Sort and Filter" Select Filter.
6. For each column, Select "Sort by Color" under the little arrow choosing the colour of the unique values above (this is no color by default)
7. This should produce a list with uncoloured (unique) values at the top of the column(s).