



Reactome

This document introduces Reactome, the user interface and the database content. Further information can be found in the online Reactome user guide at <http://www.reactome.org/userguide/Usersguide.html>.

Reactome	1
<i>What is Reactome?.....</i>	2
<i>What is this Tutorial For?.....</i>	2
<i>Exercises</i>	<i>Error! Bookmark not defined.</i>
<i>The Reactome Home Page</i>	3
<i>Searching Reactome.....</i>	3
Simple text search.....	3
<i>The Pathway Browser</i>	6
<i>Pathway Diagrams.....</i>	10
<i>Overview diagrams and Subpathway icons</i>	13
<i>Subpathway Shading.....</i>	13
<i>Navigating Pathway Diagrams</i>	14
<i>Details Panel.....</i>	15
The Structures Tab.....	18
The Expression Tab	19
Analysis Tab	19
Downloads Tab	19
<i>Disease in Reactome</i>	21
Reactome Tools.....	23
<i>Analyse Data</i>	23
Results for Identifier lists without associated numeric values.....	25
Results for Identifier lists with associated numeric values (expression representation)	28
<i>Species Comparison</i>	31
<i>Molecular Interaction Overlay</i>	33
Details Views.....	38

What is Reactome?

Reactome is a curated database of pathways and reactions in human biology. Reactions can be considered as pathway ‘steps’. Reactome defines a ‘reaction’ as any events in biology that changes the state of a biological molecule. Binding, activation, translocation, degradation and classical biochemical events involving a catalyst are all reactions. Information in the database is authored by expert biologists, entered and maintained by Reactome’s team of Curators and Editorial staff. Reactome content frequently cross-references other resources e.g. [NCBI](#), [Ensembl](#), [UniProt](#), KEGG ([Gene](#) and [Compound](#)), [ChEBI](#), [PubMed](#) and [GO](#). [Inferred orthologous reactions](#) are available for 17 non-human species including mouse, rat, chicken, puffer fish, worm, fly, yeast, rice, and Arabidopsis

What is this Tutorial For?

This tutorial introduces features of the Reactome website using a combination of short explanations and exercises. You will learn how to search Reactome, interpret the views, use the tools and if necessary find documentation or contact us for help.

The Reactome Home Page

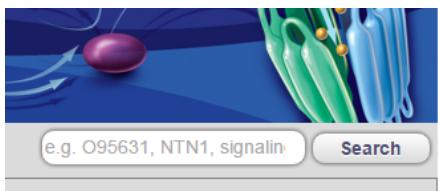
Reactome's Home page is <http://www.reactome.org/>. This page contains links, tools, search functions and documentation that will help you use Reactome.

The screenshot shows the Reactome Home Page. At the top is a banner with the Reactome logo and a molecular diagram. Below the banner is a navigation bar with links: About, Content, Documentation, Tools, Community, Download, Contact, a search bar containing "e.g. O95631, NTN1, signalin", and a "Search" button. Below the navigation is a grid of six buttons: "Browse Pathways", "Analyze Data", "Reactome FI Network", "User Guide", "Data Download", and "Contact Us". To the right is a sidebar titled "Tweets" showing three recent tweets from @reactome, @GrameneDatabase, and @BioGPS. At the bottom is a footer with links to various sections like About, Content, Documentation, Tools, and Community, along with social media icons for Facebook, Twitter, and YouTube.

Searching Reactome

Simple text search

The simple text search tool is located top right of the Home page. Most other Reactome website pages include this search tool. To search type a word, phrase or identifier in the search box. The search has an auto-complete function; if the text you wanted to use appears in the drop-down list, select it and results will be displayed. For other text click **Search**.



Search results for *protein kinase*

Showing 30 of 22544

Species

- Homo sapiens (21574)
- Entries without species (970)
- Danio rerio (29759)
- Mus musculus (22221)
- Canis familiaris (20583)
- Drosophila melanogaster (20036)

More...

Types

- Protein (15773)
- Reaction (4081)
- Pathway (929)
- Regulation (781)
- Complex (537)
- Set (374)

More...

Compartments

- cytosol (6018)
- plasma membrane (5062)
- extracellular region (3824)
- nucleoplasm (3094)
- endoplasmic reticulum lumen (993)
- Golgi lumen (825)

More...

Reaction types

- regulates (755)
- binds (667)
- phosphorylates (239)
- activates (89)
- transports (54)
- ubiquitinates (35)

More...

Search properties

- clustered Search

[Reset filters](#)

Reaction (3 results from a total of 4081)

- cGMP stimulates Protein Kinase G (Homo sapiens)**

Protein Kinase G (PKG) is a homodimer held together by a leucine zipper present in the N terminus. Each

- Activation of conventional Protein Kinase C (Homo sapiens)**

The conventional Protein Kinase C (cPKC) isoforms have two membrane-targetting domains, a C1 domain

- Activation of Protein Kinase C novel isoforms (Homo sapiens)**

Activation of the novel Protein Kinase C (nPKC) isoforms (delta, epsilon, eta and theta) requires binding to the

Set (3 results from a total of 374)

- Activated conventional protein kinase C (Homo sapiens)**

- Mitogen-activated protein kinase (Homo sapiens)**

This CandidateSet contains sequences identified by William Pearson's analysis of Reactome catalyst entities.

- Protein Kinase A_catalytic subunits (Homo sapiens)**

Genes and Transcripts (3 results from a total of 24)

- unidentified protein tyrosine kinase (Homo sapiens)**

- unidentified protein tyrosine kinase (Homo sapiens)**

- unknown kinase (Homo sapiens)**

Complex (3 results from a total of 537)

- Protein kinase C_novel isoforms:DAG (Homo sapiens)**

- Activated cGMP-dependent protein kinase (PKGs) (Homo sapiens)**

- Protein Kinase C_alpha type: DAG (Homo sapiens)**

Regulation (3 results from a total of 781)

'DAGs [plasma membrane]' positively regulates 'DAG stimulates protein kinase C-delta'

Search results are presented in categories with grey headers, representing molecular entity or event types.

Below the header is a list of items from the named type that match your search terms. Matching words are highlighted in grey. Click on the name of an item to go to a page of details.

On the left side of the search results page are groups of check boxes. These can be used to change defaults and add filters. Check or uncheck boxes for types that you do/don't want to see, e.g. if you want to see results for Mus musculus, under Species uncheck Homo sapiens and check Mus musculus. In the Type, Compartment and Reaction Type filter groups, select to only see results that fall into the selected category, e.g.

to see only reactions, select Reaction in the Types filter group. To see only Reactions that occur in the cytosol, also select Cytosol in Compartments.

Click on any of the results to go to a page of information about it. The contents of these Results Detail pages will depend on the type, but all have a section named ‘Locations in the PathwayBrowser’. This contains an expandable hierarchy that identifies the locations of the item in Reactome pathway diagrams. Click the plus symbol to open the hierarchy.

Locations in the PathwayBrowser

- Hemostasis(Homo sapiens)
 - Hemostasis(Homo sapiens)
 - └ Platelet homeostasis(Homo sapiens)
 - └ Nitric oxide stimulates guanylate cyclase(Homo sapiens)
 - └ cGMP effects(Homo sapiens)
 - └ cGMP stimulates Protein Kinase G(Homo sapiens)

This hierarchy reflects the organisation of events in Reactome. ‘Top-level’ pathways representing a broad area of biology typically contain one or more hierarchical levels of subpathways, becoming more specific with each level. The most general pathways are closest to the left-hand edge of the screen; the most specific subpathways have the greatest indentation to the right. Note that some search result items (e.g. proteins) can be in more than one pathway, which will be reflected in search results by multiple hierarchies. Select any event name to open the corresponding Pathway Diagram. The diagram will open via an animation, which shows you where the pathway is located in the Pathway Overview. The selected item will be highlighted in blue within the pathway diagram.

The Pathway Browser

The Pathway Browser is the primary means of viewing and interacting with pathways in Reactome. It includes tools for analysing datasets and exploring pathways. These tools allow several types of analysis:

- Pathway over-representation analysis and pathway topology-based analysis
- Comparison of a pathway with its equivalent in another species
- The overlay of user-supplied expression data onto a pathway
- The overlay of protein-protein or protein-compound interaction data from external databases or user-supplied data onto a pathway (coming soon).

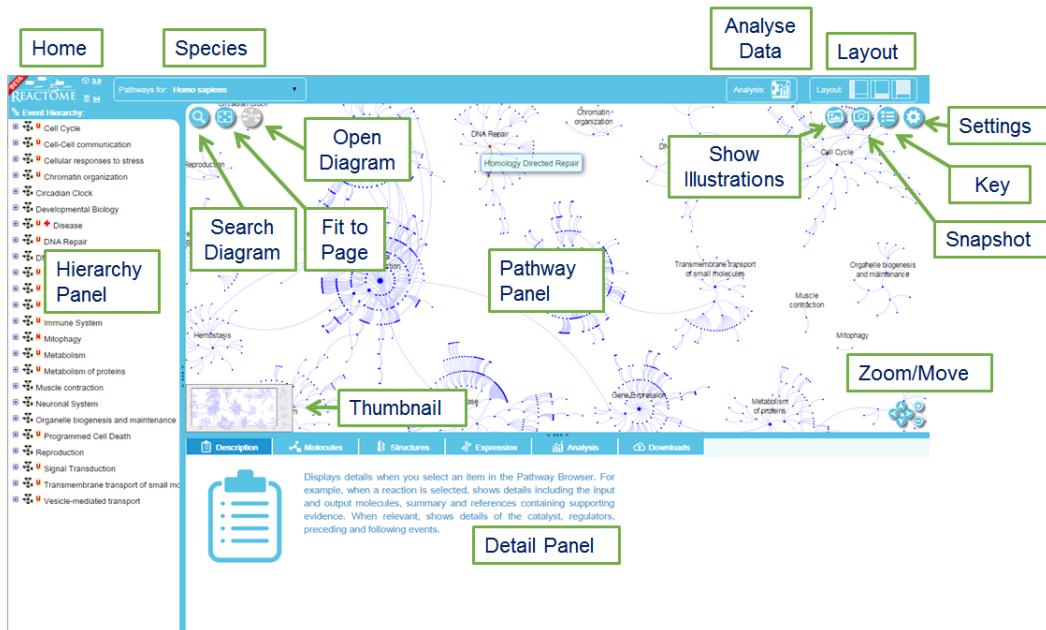
See the section [Reactome Tools](#) for details of the analysis tools.

The Pathway Browser is launched by clicking the [Browse Pathways](#) button on the Homepage:



When the Pathway Browser opens, it displays an overview of all Reactome pathways. The hierarchy of pathways is represented as nodes (circles) arranged concentrically (in rings) around a central node. The central nodes represent the top-level, broad topics available in Reactome. Most are connected by lines to a surrounding concentric ring of 'child' subpathways, which may themselves be connected to further concentric rings of sub-sub pathways. This overview gives an overview of all pathways and allows rapid zooming to details. As you zoom in, labels will appear for the subpathway nodes. This view is particularly useful for displaying data analysis results for all pathways. To see the details of a specific pathway, select and click (or double-click) the node representing the pathway. Alternatively select and click on the name in the Pathway Hierarchy panel on the left.

Features of the Pathway Browser are labelled in the diagram below.



Home – the Reactome logo is a button linked to the homepage.

Species – Reactome is a database of curated human biological pathways. These human pathways are used to computationally infer equivalent pathways in model organisms (described in detail [here](#)). Use the Species drop-down to select a species and view the predicted pathway. Note that infectious disease pathways that show the interaction of pathogens with human proteins are included in Homo sapiens.

Layout – The pathway browser is divided into 3 main sections. These are: the Hierarchy Panel, on the left, the Details Panel, bottom right, the Pathway Panel, top right. The layout buttons allows you to show or hide these panels.

Analyse Data – Opens a panel where analysis can be performed.

Key – this opens a key explaining the objects present in the pathway panel.

Zoom/Move Toolbar - Click on the arrows to move the diagram. Click on the plus and minus buttons to zoom in and out. Alternatively, click and drag the diagram, and use your mouse scroll wheel to zoom.

Search Diagram – click the button to open the search panel. Note that this searches the diagram. To search all of Reactome, use the Search on the Homepage.

Fit to Page - resizes the Pathway Overview or Diagram to fit the available space.

Open Diagram – when a pathway node is selected in the Pathway Overview, this button opens the corresponding Pathway Diagram

Show Illustrations – some Reactome pathways have a corresponding high-quality graphic. Select this button to view.

Snapshot – use this button to download the visible region of the Pathway Overview or Diagram, including any analysis overlay, as a PNG file or upload it to GenomeSpace.

Thumbnail – Provides an overview of the entire pathway when zoomed in so that only a region of the pathway is displayed.

Pathway Panel - This is where Pathway Diagrams are displayed when they are selected in the Pathway Hierarchy or in search results. If no pathway is selected, this panel displays a brief Reactome help guide.

Details Panel - Contains details of objects when they are selected in the Pathway Diagram. The content depends on the type of object, i.e. pathway, reaction, complex, set, protein or small molecule (see the Details section).

Hierarchy Panel – Includes the hierarchical organisation of Reactome pathways and pathway events or steps, known as reactions. When the Pathway Browser is opened, only the top level of the hierarchy is shown, as an alphabetical list of Reactome's main topics. Most topics are too large to be represented as a single diagram. Instead they are split into sub-pathways, which may be further divided into sub-sub-pathways. Most topics in Reactome are organised in this hierarchical manner. Sub-pathways can be revealed by clicking on the + icon to the left of the pathway name, and hidden by clicking on the - icon.

Pathway names are preceded by the icon . As sub-pathways are revealed, at some point a level is reached where the reactions will become visible. Reactions names are preceded by the icon . You may also see the icon representing “black-box” reactions, where the full molecular details have been omitted for clarity, or the details are incompletely determined.

The order of reactions from top to bottom in the Hierarchy Panel usually follows their order as steps in the pathway so that the preceding reaction is above and the subsequent reaction below, but this is not always the case. Pathways may not be linear, they can be circular or branched.

Consequently, reactions often have multiple preceding or subsequent reactions. These connected events are often visible within the pathway diagram but can be part of a different pathway diagram. To view a complete list of the events that precede or follow a reaction, refer to the Preceding and Following Events section in the Details Panel, as described in the Details Panel chapter below.

Pathway Diagrams

Pathway Diagrams represent pathways as a series of connected molecular events, known in Reactome as ‘reactions’, which can be considered as steps in the pathway.

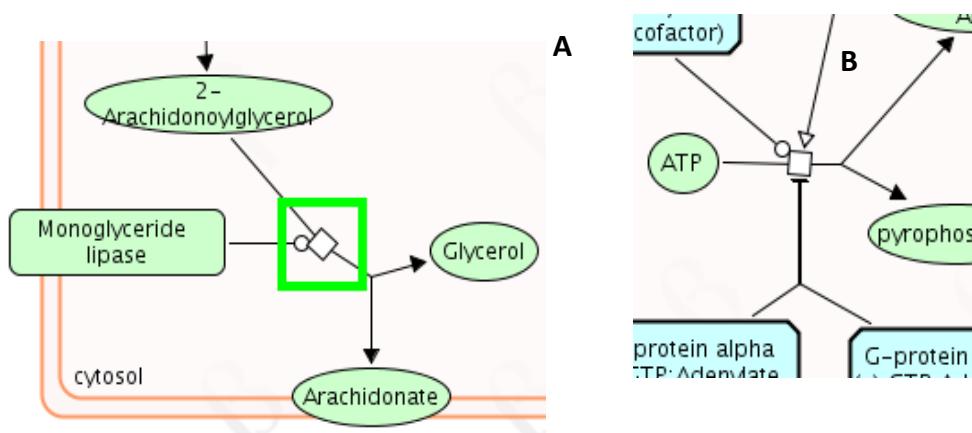
Cellular compartments are represented as pink/orange boxes with a double boundary. A typical diagram has a box to represent the cytosol, bounded by a double-line that represents the plasma membrane. The white area outside this box represents the extracellular space. Other organelles are represented as additional labelled boxes within the cytosol.

Molecules, represented in diagrams as blue/green boxes, are placed in the physiologically-correct cellular compartment, or lie on the boundary of a compartment to indicate that they are in the corresponding membrane, e.g. a plasma membrane protein will be placed on the boundary of the cytosol.

Reactions typically include:

- Input and output molecules, and when relevant a catalyst (see diagram A below)
- Inputs, outputs and catalyst are represented as boxes or ovals
- Green boxes with rounded corners are proteins
- Green boxes with square corners are proteins that have no Uniprot accession (or did not at the time the reaction was created).
- Green ovals are small molecules or sets of small molecules
- Blue boxes with a double boundary are sets, i.e. proteins or small molecules that are functionally equivalent.
- Blue boxes with cut corners are complexes, i.e. proteins and/or other molecules that are bound in a multimolecular entity.
- Green boxes with a white inner box are sub-pathways.
- Reaction input and output molecules are joined by lines to a central ‘reaction node’ (surrounded by a green box in Figure A below). Clicking this node selects the reaction.
- The outputs of a reaction have an arrowhead on the line connecting them to the reaction node.
- Reaction inputs/output molecules are often connected by arrows to preceding or subsequent reactions (i.e. the preceding/subsequent steps in the pathway).
- Catalysts are connected to the reaction node by a line ending in a circle.
- Numbered boxes on the line between an input/output and the reaction node indicates the number of molecules of this type in the reaction (when $n > 1$).

- Molecules that regulate a reaction are connected to the reaction node by a line ending in an open triangle for positive regulation or a 'T'-shaped head for negative regulation (see B below).
- A white box labelled P on the boundary of an object indicates a phosphorylation event.
- Proteins or small molecules that are also part of a displayed set may be connected to the set object by a line of short dashes.
- Sets with overlapping content may be connected by a line of long dashes.
- Reactions that represent a disease process use red connecting lines. Objects associated with a disease are bordered in red.



There are 5 subtypes of reaction node, indicating reaction subclasses.



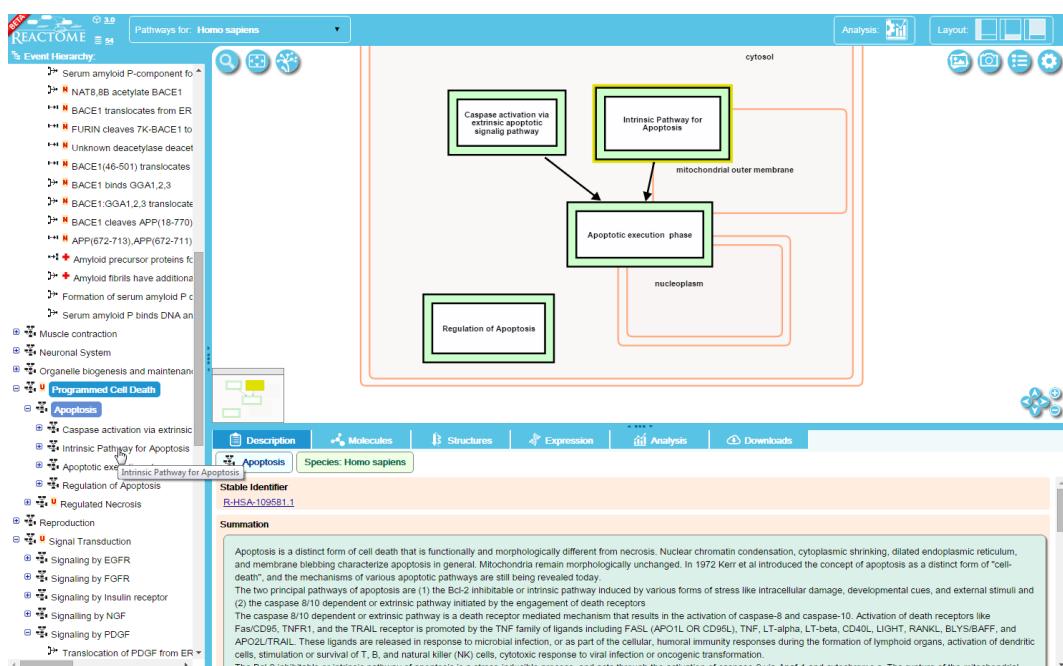
- Open squares represent 'transition', i.e. a change of state that is not one of the defined subclasses.
- Solid circles represent 'association', i.e. binding
- Double-bordered circles represent 'dissociation'
- A square with two slashes represents an 'omitted process'. This is used when the full details of a reaction have been deliberately omitted. This is most commonly used for events that include representative members of a large family to illustrate the general behaviour of the group. It can be used for reactions that occur with

no fixed order or stoichiometry, or for degradation events where the output is a random set of fragments.

- Squares containing a question mark represent an ‘uncertain process’, where some details of the reaction are known, but the process is thought to be more complex than represented.
Explanatory details are typically included in the Description.

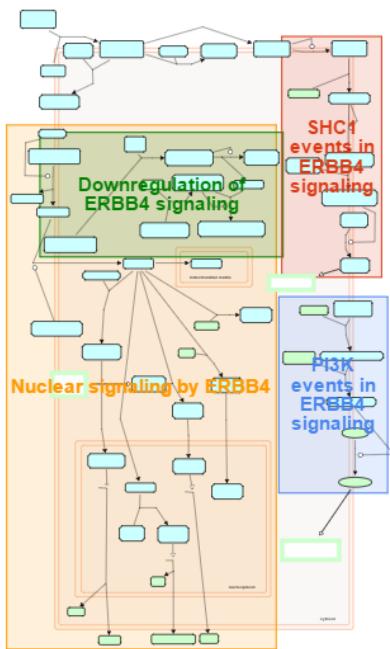
Overview diagrams and Subpathway icons

Topics such as Apoptosis are too large to be displayed in a single Pathway Diagram. Instead they are divided into sub-pathways, each with its own Pathway Diagram. Large topics are represented in the Pathway Panel as an Overview Diagram, containing one or more Subpathway Icons, boxes with a green boundary. Click on a Subpathway Icon to select it. Its name is highlighted in pale blue in the Hierarchy. Its parent pathway is highlighted in dark blue. Details of the pathway will appear in the Details Panel, including a summary, which may help you decide whether you want to open the Pathway Diagram. If you do, either double-click on the Subpathway Icon, or on the highlighted name in the Hierarchy. In the figure below, the pathway Apoptosis is selected, and the Overview Diagram for this pathway displayed. The mouse pointer is hovering over the subpathway name 'Intrinsic Pathway for Apoptosis' in the Hierarchy, causing the corresponding Subpathway Icon to be highlighted in yellow in the Diagram. If the name of the subpathway is clicked in the Hierarchy the Pathway Diagram will open.



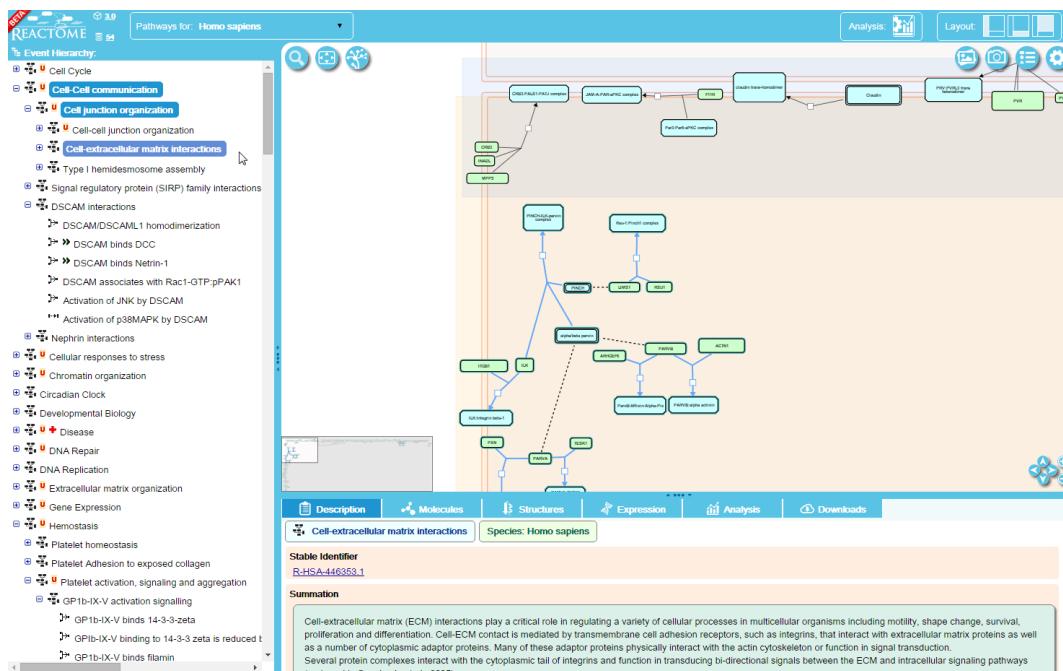
Subpathway Shading

When a Pathway Diagram contains subpathways, the area containing each subpathway is overlaid with a coloured box, labelled with the subpathway name, to help locate it in the larger diagram. Below is the diagram for ERBB4 signaling. The four subpathways are represented as 4 shaded areas.



Navigating Pathway Diagrams

The Pathway Diagram, Hierarchy and Details Panels are interactively connected. Selecting an entity (molecular object) or event in the diagram will reveal relevant information in the Details Panel. Clicking a pathway name in the Hierarchy will open the corresponding Pathway Diagram in the Pathway Diagram Panel, and reveal details of the pathway in the Details Panel. Where a subpathway is contained within a diagram that is already displayed, clicking the sub-pathway name in the hierarchy will cause all the reactions in that pathway to be highlighted in blue. In the figure below, the diagram for Cell junction organisation is displayed in the Pathway Panel, and the sub-pathway Cell-extracellular matrix interactions has been selected in the Hierarchy, causing all the reactions in this sub-pathway to be highlighted (blue) on the Pathway Diagram.



Similarly, selecting a reaction in the Hierarchy causes that reaction to be highlighted in the pathway diagram and updates the Details Panel to show details of the reaction. If the reaction is not in the visible region of the diagram, the view will re-centre and zoom to show it.

Details Panel

The Details Panel is the bottom right panel of the Pathway Browser. It displays the details of molecular objects or events when they are selected in the Pathway Diagram. The details shown depend on the type of object or event that is selected. The Details Panel can be revealed/hidden using the middle Layout button.

The figure below is an example for a reaction. Details for pathways, complexes, sets, proteins and small molecules are similar - see the Details Views section in the chapter Other Reactome Features.

The Details Panel has several tabs:

The Description tab

This tab contains the defining details of the selected diagram object. For example, reaction details always describe the input and output molecules, and where relevant the enzyme catalyst. Other items in the Description tab include:

References – reactions always include links to publications that contain experimental data demonstrating the reaction.

Summation – a summary of the molecular event, often with background information and additional literature citations

Authored - an expert biologist who contributed materials allowing construction of the reaction/pathway. All Reactome pathways are authored by an expert biologist.

Reviewed – All Reactome pathways are reviewed by an expert biologist who verifies the content provided by the Author.

Cellular Compartment - Identifies the cellular compartment for the reaction, with the associated Gene Ontology (GO) term.

Preceding/Following Events - Links to reactions that precede/follow this reaction.

Inferred from another species - Reactome represents human biology and wherever possible cites experimental data obtained using human reagents. When this is unavailable, a human reaction may be inferred from data obtained from model organisms, if the Author of the reaction believes this is valid. When this is the case, the reaction details will include an

'Inferred from another species' link, indicating that the human reaction has been manually inferred from experimental data obtained in another species. Selecting the link will reveal details of the reaction in the model organism including the literature references.

Computationally Inferred To – Provides a drop-down list of model organism species, select one to see computationally predicted pathways, inferred from Reactome's human pathway.

Details are organised into panels. Panels that have a plus icon on the right hand side can be expanded to reveal further levels of detail by clicking on the icon or the detail name.

The screenshot shows a software interface for a pathway. At the top, there is a header with the word 'Output'. Below it is a panel with a title 'TRAIL:TRAIL receptor-2:FADD [plasma membrane]'. To the right of the title is a small blue square icon with a white plus sign. The main body of the panel is currently collapsed, showing only the title and the plus sign icon.

Clicking the plus icon reveals further details:

The screenshot shows the same pathway interface after the plus icon was clicked. The panel is now fully expanded, revealing detailed information. The title 'TRAIL:TRAIL receptor-2:FADD [plasma membrane]' is still at the top. Below it, there is a section labeled 'Species: Homo sapiens' and 'Synonyms: TRAIL:TRAIL receptor-2:FADD'. Underneath this, there is a section labeled 'Has components:' which contains two items: 'TRAIL receptor-2:TRAIL Trimer [plasma membrane]' and 'FADD [cytosol]'. The 'FADD [cytosol]' item is expanded, showing its own detailed information: 'Species: Homo sapiens', 'Synonyms: FADD, FADD protein (FAS-associating death domain-containing protein) (Mediator of receptor induced toxicity), FADD protein, FAS-associating death domain-containing protein, Mediator of receptor induced toxicity', 'Coordinates in the Reference Sequence: 1 .. 208', 'Reference Entity: UniProt:Q13158 FADD', and 'External cross-references'.

Panels that represent molecules have an icon on the left hand side that represents the molecule subtype, e.g. the protein icon is a green circle. Mouse-over the icon to see an explanation of the subtype. The icon is also a selection button; clicking the icon causes details of the molecule to be displayed, and selects it on the pathway diagram.

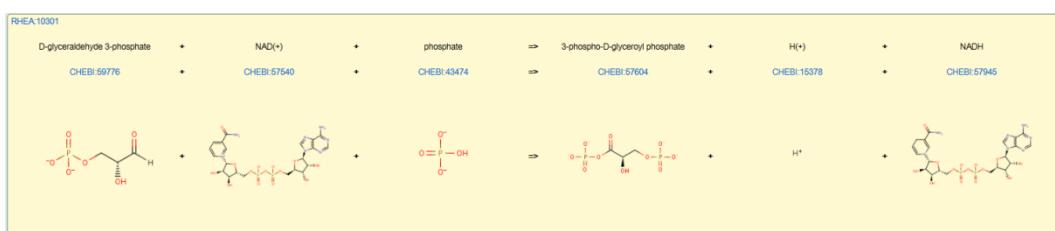
The Molecules Tab

Details all the molecules involved in the pathway displayed in the panel above. If an item is selected in the diagram, the corresponding molecules are highlighted.

Molecules are grouped into the subtypes Protein, Chemical Compound, Sequences (for DNA/RNA) and Other. The Download button allows details of the molecules to be downloaded in several formats. Click on the buttons to select the fields you want included in the output file, and the format of the file. Click Download to save the file.

The Structures Tab

The content displayed in this tab depends on the type of object or event selected in the pathway diagram. For Reactions it will display equivalent reaction diagrams from the [Rhea](#) database when available:



The panel has a link top-left to details in the Rhea database.

The molecular structures are linked to the [ChEBI](#) database:

If no Rhea diagram is available it will display:

Overview Molecules (4) **Structures (0)** Expression Processes Downloads

This event does not contain Rhea structures to be shown

For Proteins or pathway objects that represent sets or complexes, corresponding structure information from [PDBe](#) is represented if available:

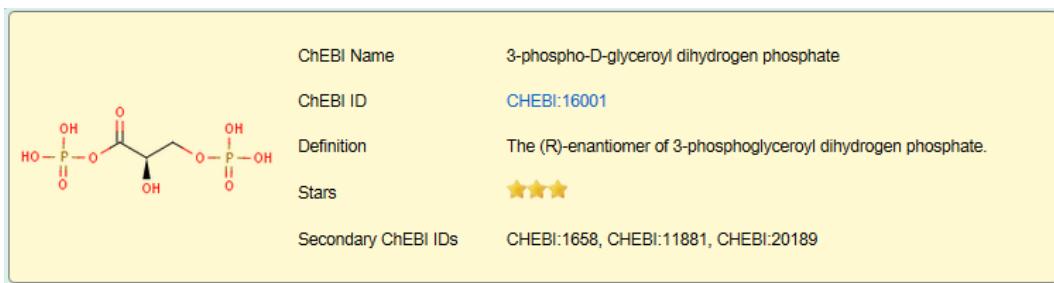
Overview Molecules Structures (1/1) Expression Processes Downloads

UniProt O15519 CFLAR Chain: A Resolution: 1.90 Coverage: 0.57 PDB Range: [1, 272]
 UniProt Range: [209, 480]

3h11 X-RAY

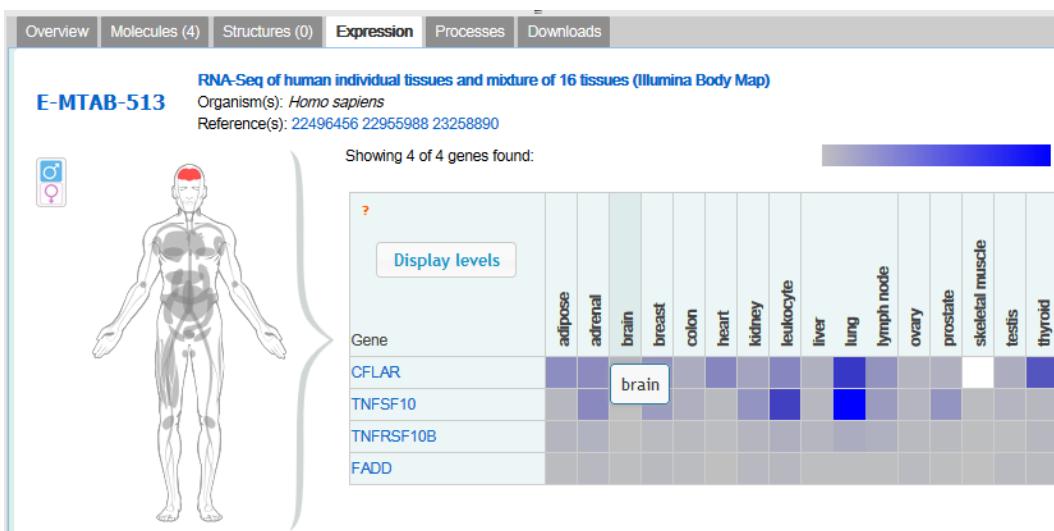
► All other structures for O15519

For simple molecules, the structures tab represents information from the ChEBI database:



The Expression Tab

This tab represents expression information obtained from the [Gene Expression Atlas](#):



Analysis Tab

This tab displays the results of analyses – see the section Reactome Tools for further details

Downloads Tab

The Downloads Tab contains buttons which start a download of the currently displayed pathway.

Overview Molecules Structures Expression Processes Downloads

 **Extrinsic Pathway for Apoptosis**

The download options below are for the selected pathway, not individual events or entities selected in it.

 SBML
 SBGN
 BIOPAX 2
 BIOPAX 3
 PDF

 Word
 Protege

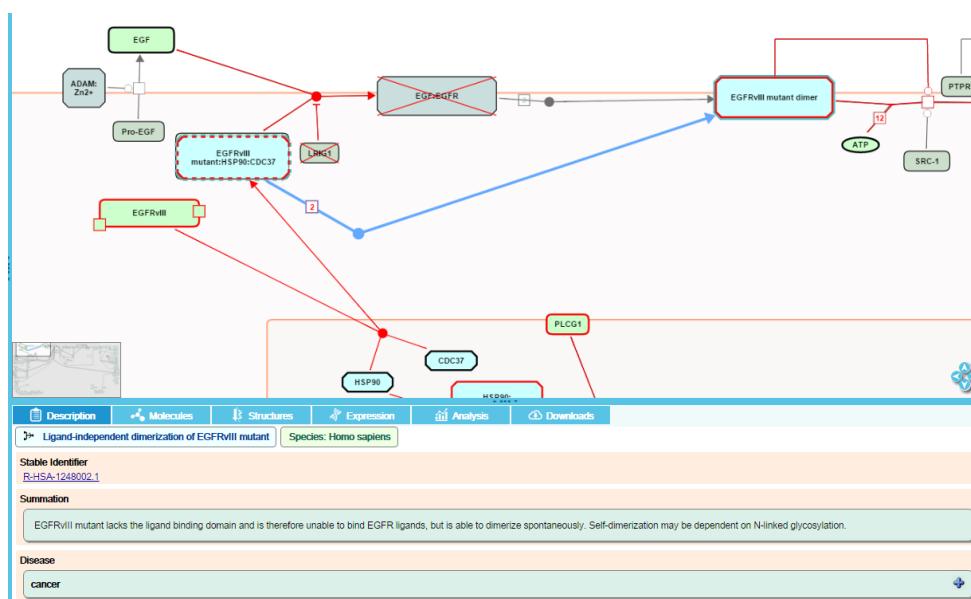
Disease in Reactome

Reactome can annotate and display pathways associated with disease.

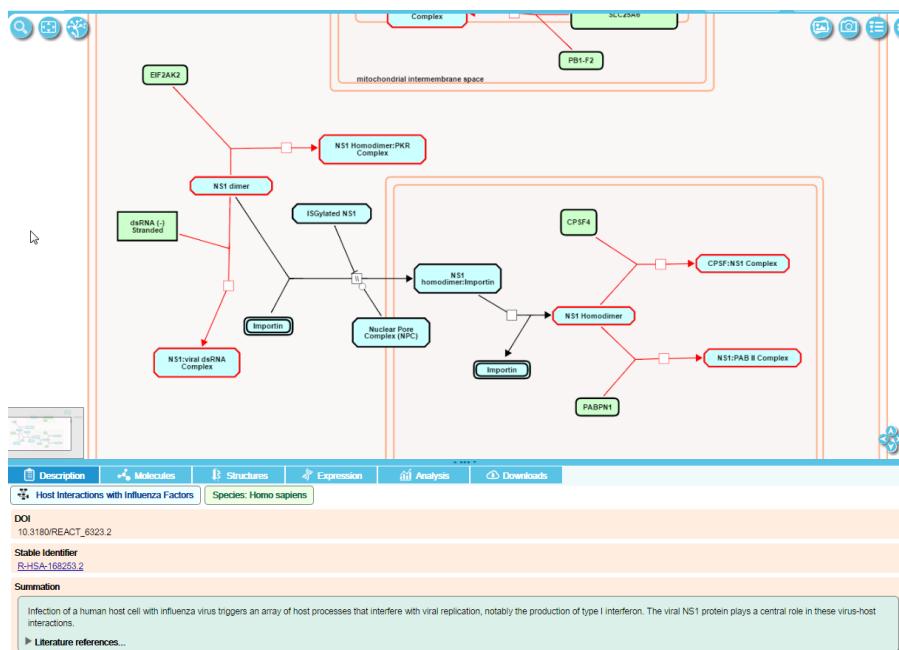
Disease events or processes are overlaid onto their normal counterparts to provide a pathway context to the disease event.

Disease events have red connecting lines; abnormal molecules involved in these processes are outlined in red.

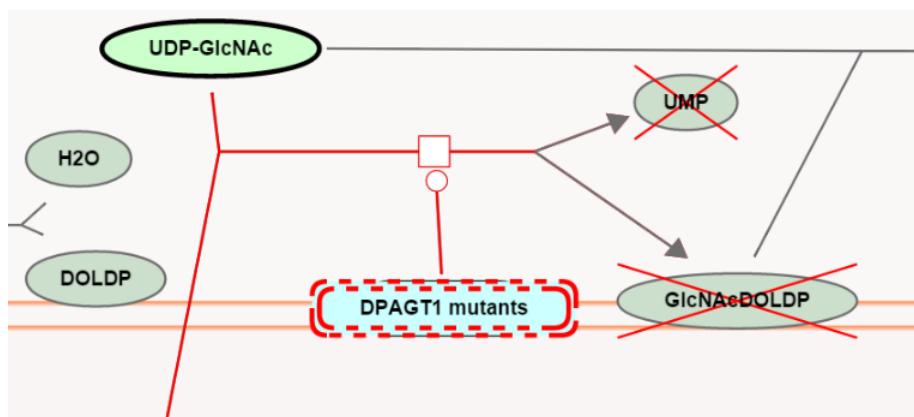
For cancer-related processes, it is useful to view the altered disease events alongside their normal counterparts in the same diagram. The user can then see where normal processes diverge into ones that are implicated in cancer.



For infection processes, events involved in the life cycle of the infecting agent and host interactions with the infecting agent have red connecting lines. Molecules derived from the infecting agent are outlined in red, while those derived from the host are outlined as normal in black.



For metabolic processes, proteins that have lost all or most of their functional activity towards a substrate causes the majority of defects. The diagram displays these disease reactions as 'stop points' in the pathway. Defective enzyme catalysts are outlined by a red dashed line; products that are no longer made are shown greyed-out with a superimposed red cross.



Reactome Tools

Analyse Data

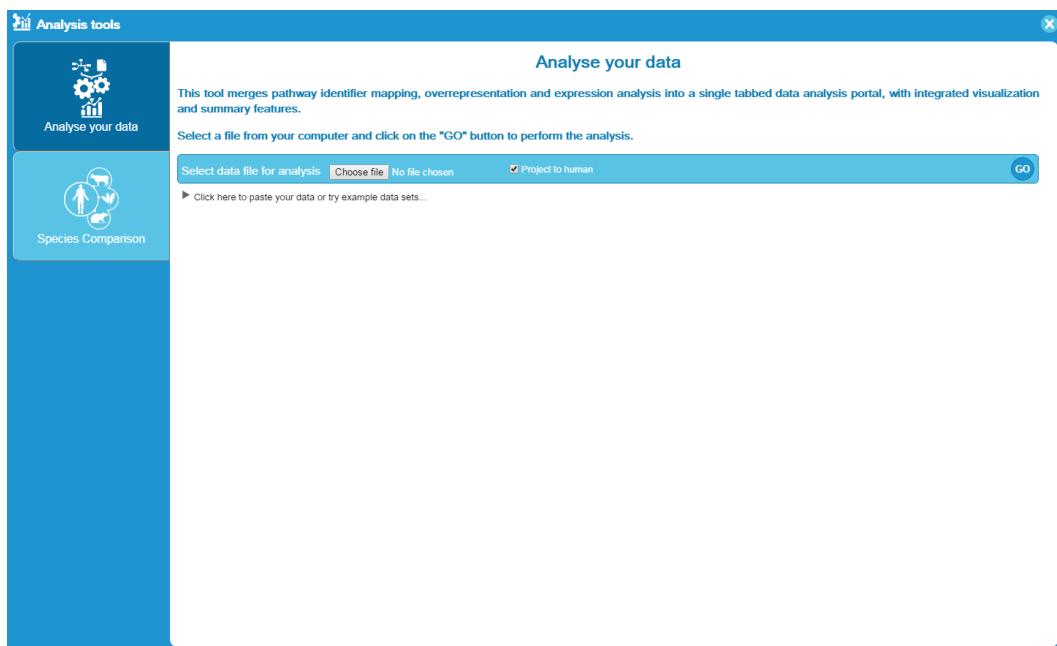
The Analyse data panel is opened in the Pathway Browser by clicking on the Analysis button on the Pathway Browser header bar:



Alternatively, from the Homepage select the Analyze data button:



This opens a submission form, where you can select the analysis you want to perform, paste in or browse to a file containing your data, or use an example data set.



A screenshot of a web-based analysis tool. The title bar says "Analysis tools". On the left, there's a sidebar with two buttons: "Analyse your data" (selected) and "Species Comparison". The main area has a heading "Analyse your data" with a sub-instruction: "This tool merges pathway identifier mapping, overrepresentation and expression analysis into a single tabbed data analysis portal, with integrated visualization and summary features." Below that, it says "Select a file from your computer and click on the "GO" button to perform the analysis." There's a "Choose file" button, a "No file chosen" message, a checked checkbox for "Project to human", and a "GO" button. At the bottom, there's a note: "► Click here to paste your data or try example data sets...".

There are two sections to the submission form. Use the Analyse your data section to submit your data. Several different analyses can be performed, depending on the format of your data.

If your data is a single column of identifiers such as UniProt IDs, gene symbols or ChEBI IDs, they are mapped to pathways and over-representation and pathway-topology analyses are run. Over-representation analysis is a statistical test that determines whether certain Reactome pathways are over-represented (enriched) in the submitted data. It answers the question ‘Does my list contain more proteins for pathway X than would be expected by chance?’

Pathway topology analysis considers the connectivity between molecules that is represented by the pathway steps (which we refer to as reactions) in the pathway. It groups all the molecules represented in each reaction as a pathway ‘unit’. If any of these molecules is represented in your query set, this is considered a match to that reaction. This may give a better indication of the proportion of the pathway that matches your data, rather than the number of molecules that are common between your data and the pathway.

If your data has one or more additional columns of numbers, and a row of column headers where the first column header starts with the # symbol, it will be recognised as expression data and expression overlay will be run. The numbers are used to produce a coloured overlay for Reactome pathway diagrams.

Identifier mapping

The submission process recognises many types of identifiers. As part of the pre-analysis, they are mapped to Reactome molecules. The ideal identifiers to use are UniProt IDs for proteins, ChEBI IDs for small molecules, and either HGNC gene symbols or ENSEMBL IDs for DNA/RNA molecules, as these are our main external reference sources for proteins and small molecules. Many other identifiers are recognised and mapped to appropriate Reactome molecules. Accepted identifiers include HUGO gene symbols, GenBank/EMBL/DDBJ, RefPep, RefSeq, EntrezGene, MIM, InterPro, EnsEMBL protein, EnsEMBL gene, EnsEMBL transcript, and some Affymetrix and Agilent probe IDs. Uniprot isoforms may be specified using the format P12345-2. If the –n suffix is omitted this canonical form and all isoforms of it will be matched. Mixed identifier lists may be used. Identifiers must be one per line. Protein-specific identifiers will typically map to protein entities, while gene-specific identifiers will map to the gene, transcript and derived proteins. If desired results can be filtered to show protein-specific or gene/transcript-specific results, details below.

By default, all non-human identifiers are converted to their human equivalents. If you want to use non-human identifiers to search a computationally-inferred non-human Reactome pathway, uncheck the box. You may also prefer to uncheck this box if your query consists of a mixture of human and microbial identifiers, representing an infection.

Below is an example of the identifiers-only format:

Analyse your data

This tool merges pathway identifier mapping, overrepresentation and expression analysis into a single tabbed data analysis portal, with integrated visualization and summary features.

Select a file from your computer and click on the "GO" button to perform the analysis.

Project to human

Paste the data to analyse

```
#GBM Uniprot
P01023
Q99758
O15439
O43184
Q13444
P82987
P04083
Q775R6
P27540
Q13315
P36543
Q13535
Q96GD4
```

Project to human

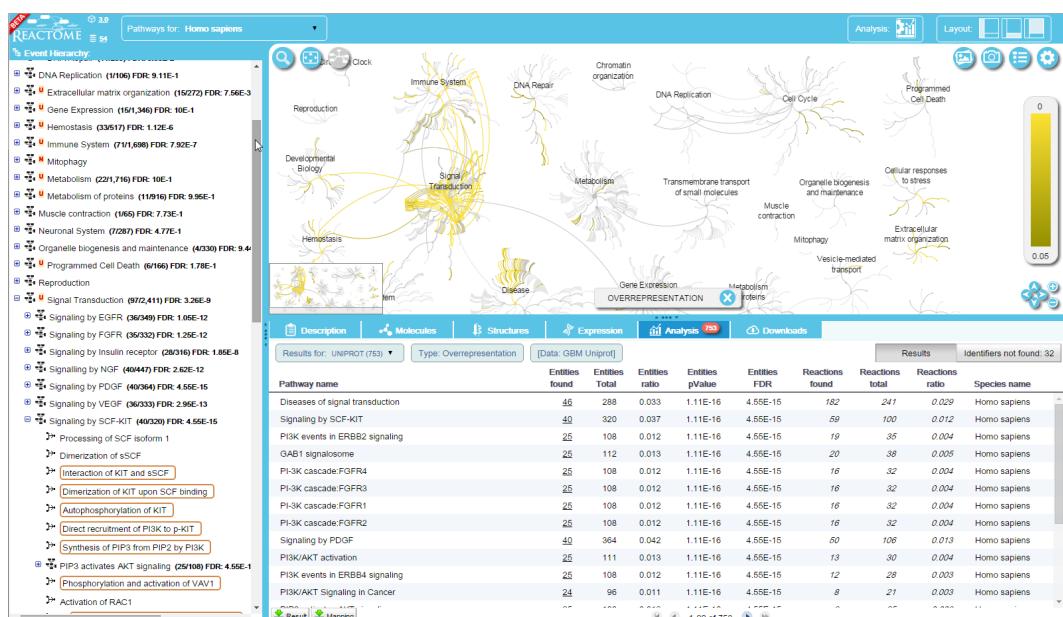
Some examples:

-
-
-
-
-
-
-

Click the Go button to start the analysis.

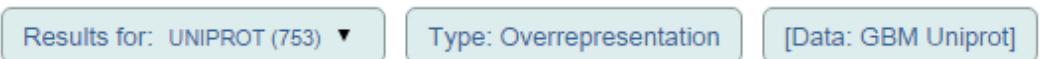
Results for Identifier lists without associated numeric values

If you submit a single column of protein or small molecule identifiers they are mapped to pathways, over-representation and pathway-topology analyses are performed. The results will resemble the example below.



Analysis results are shown in the Analysis tab, within the Details Panel. All Reactome pathways are shown, in blocks of 20 pathways, ranked by the FDR value obtained from over-representation analysis. In addition, the number of molecules matched/total number of molecules and FDR values are added to the right side of pathway names in the Pathway Hierarchy Panel. Reactions that contain any of the matching molecules are boxed in orange. When you click on the name of a pathway in the Analysis tab, the Pathway Hierarchy will expand to show it and its name will be highlighted in dark blue. If it has subpathways, you can use the plus buttons to the left of the pathway name to reveal subpathways and reactions.

By default, all molecules (protein, small molecules, genes, transcripts) are used for over-representation analysis. Subtypes of molecule can be selected from a drop-down list located top-left of the results table. Selecting one of these subsets will re-display results that consider only the selected molecular subtype.



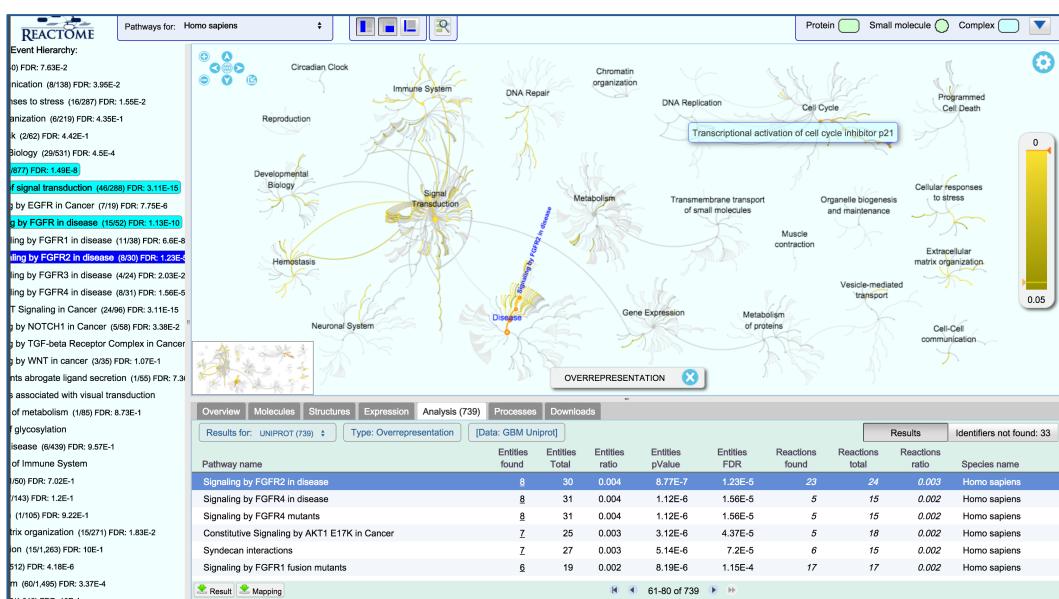
The columns in analysis details represent:

1. Pathway name. Click the name to open the pathway.
2. Entities found, the number of molecules of the type selected with Results Type that are common between the submitted data set and the pathway named in column 1. Click on this number to display the matched submitted identifiers and their mapping to Reactome molecules.
3. Entities total. The total number of molecules of the type selected with Results Type within the pathway named in column 1.
4. Entities ratio. Put simply, the proportion of Reactome pathway molecules represented by this pathway. Calculated as the ratio of entities from this pathway that are molecules of the type selected with Results Type Vs. all entities of the type selected with Results Type.
5. Entities pvalue. The result of the statistical test for over-representation, for molecules of the results type selected.
6. Entities FDR – False discovery rate. Corrected over-representation probability.
7. Reactions found. The number of reactions in the pathway that are represented by at least one molecule in the submitted data set, for the molecule type selected with Results Type.
8. Reactions Total. The number of reactions in the pathway that contain molecules of the type selected with Results Type.

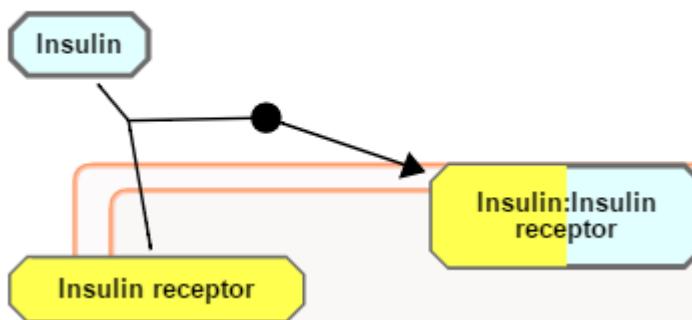
9. Reactions ratio. Put simply, the proportion of Reactome reactions represented by this pathway. Calculated as the the ratio of reactions from this pathway that contain molecules of the type selected with Results Type Vs. all Reactome reactions that contain molecules of the type selected with Results Type.

10. Species Name.

Following an analysis, the Pathway Bowser displays the Pathway Overview. All pathways that contain identifiers from your submitted list are highlighted, using a coloured scale to indicate the corrected probability (FDR). The colour scheme can be changed using the Settings button accessed by clicking the cogwheel symbol in the top right corner. The highlighting of pathways in the Pathway Overview provides a one-glance overview of the analysis result for all pathways. To see the detail of a specific pathway, select and double-click the node representing the pathway in the overview, or instead go to the Pathway Hierarchy on the left and select it there.



In Pathway Diagrams, entities are re-coloured (dark green in the default colour scheme) if they were represented in the submitted data set. Complexes, Sets and Subpathway Icons are coloured to represent the proportion that is represented in the submitted identifier list. In the figure below, Insulin receptor is yellow indicating that is was in the submitted list. Insulin was not in the submitted list so it is not re-coloured. The complex of insulin:Insulin receptor is part re-coloured, part not, indicating that some molecules in the complex were represented in the submitted dataset while others were not.



At the right side of the Analysis results details is a button indicating the number of submitted identifiers that were not successfully matched to molecules in Reactome. Click the button to produce a list.

Results for: TOTAL (1123)		Type: Overrepresentation	[Data: Uniprot_accession_list]	Results	Identifiers not found: 73
Not found identifiers					
Q96KC3 P3Q1308621 P162O95P533967 O15421 P43P00751 P106435 P1P017326 P18440861 P064937					

Results for Identifier lists with associated numeric values (expression representation)

To run expression analysis submit your data in a format that includes a first row of column headers. The header for column 1 must start with the # symbol. The first column must contain protein, compound or other suitable identifiers, such as probe IDs. All other columns must be numeric values, with no alphabetical characters. The analysis tool will interpret your data as expression data. The numeric values are used to colour objects in pathway diagrams. This view was created for microarray data, but any dataset that consists of a list of identifiers with associated numeric values can be used, e.g. quantitative proteomics, GWAS scores.

The tool is launched using the Analyse data button in the Pathway Browser header bar. Either paste your data into the submission form or browse to a saved file (or select an example file).

The figure below shows the correct data format. Each row must have an identifier in the first column (a header row is optional).

The submission process recognised many types of identifiers. As part of the pre-analysis, they are mapped to equivalent UniProt accessions or for small compounds to ChEBI IDs. These are the ideal identifiers to use with Reactome analysis tools. Other identifiers that are recognised and converted to Uniprot equivalents include HUGO gene symbols, GenBank/EMBL/DDBJ, RefPep, RefSeq, EntrezGene, MIM and InterPro IDs, some Affymetrix and Agilent probe IDs, Ensembl protein, transcript and gene identifiers. Identifiers that contain only numbers such as those from OMIM and EntrezGene must be prefixed by the source database name and a colon e.g. MIM:602544, EntrezGene:55718. Mixed identifier lists may be used. Identifiers must be one per line.

By default, all non-human identifiers are mapped to their human equivalents, unless the Project to human checkbox is unselected.

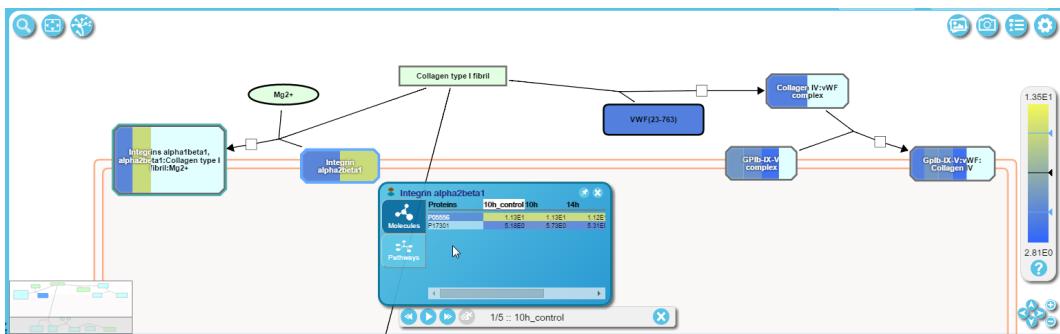
After column 1, all other columns must contain numbers, representing expression or other values. Microsoft Excel and tab separated value (TSV) files can be used. When submitted, columns of numbers are considered as separate samples or experimental conditions. The values are used to overlay colour onto pathway diagrams. An Experiment Browser tool allows you to select and view overlays for each submitted data column. This is particularly useful for visualising time-points or a disease progression.

	10h_control	10h	14h	18h	24h
1053_at	8.040878	7.147358	6.708705	6.794622	7.475157
1729_at	6.860688	6.99164	7.129922	7.11222	7.46721
1730_at	6.479505	6.500292	6.500292	6.500292	5.717815
200000_s_at	9.301569	9.710802	9.874874	9.934659	9.405911
200001_s_at	12.401575	12.511045	12.564418	12.538462	12.410574
200002_s_at	12.401575	12.511045	12.564418	12.538462	12.410574
200003_s_at	9.609852	9.999299	9.73072	9.530097	9.104303
200004_s_at	12.401575	12.511045	12.564418	12.538462	12.410574
200005_s_at	10.3127458	9.548578	9.278313	9.871472	8.753136
200006_s_at	12.110464	11.913288	11.938524	11.809943	11.458105
200007_s_at	12.110464	11.913288	11.938524	11.809943	11.458105
200008_s_at	10.377248	9.902753	9.998662	10.248412	9.513486
200009_s_at	12.61513	12.313568	12.422621	12.368137	12.170793

The results may take a few seconds to appear.

The results page is very similar to that seen following submission of a simple one-column list of identifiers, with extra columns in the Analysis details following column 9. These extra columns represent the submitted expression values.

Clicking on a pathway name launches the Pathway Browser and displays the relevant Pathway Diagram (see example below).



Objects in the Pathway Diagram are re-coloured according to the numeric values submitted. The colours are based on a scale as represented in a bar on the right hand side. There are several colour schemes, selected using the Settings button (cogwheel top right). The scale (on the right side) automatically adjusts to fit the range of values in the dataset.

Objects that were not represented in the input data are not re-coloured.

Objects with bands of colour represent complexes or sets containing more than one molecule. When zoomed out, the colour of the band reflects the average of the values submitted, for molecules represented in the dataset. The size of the band reflects the proportion of molecules that had submitted values. Zoom in to see individual bands, for each molecule that had a submitted value, arranged alphabetically by name. The bands of colour now reflect the values submitted. If multiple columns of values were submitted, representing multiple samples, e.g. time-points or a disease progression, the order of the bands will be the same for each sample.

To view details of the components of a complex or set right click it. This opens a Contextual Information Panel (see figure above). This has two tabs, Molecules and Pathways. Molecules shows the participating molecules, and if an expression analysis has been performed, their expression values. Pathways identifies whether the selected object is present in any other Reactome pathways, with links to the appropriate Pathway Diagrams.

Multiple CIPs can be opened and pinned so they remain visible when other entities are selected.

The browser remembers CIPs that were pinned on the last visit (for up to 5 diagrams).

The orange Experiment Browser toolbar (bottom left in the figure above) is used to step through the columns of your data, e.g. time-points or disease progression. Move between them by clicking the arrow buttons. The header of the data column (if present) is displayed between the arrows. The Pathway Diagram will re-colour to reflect the new values.

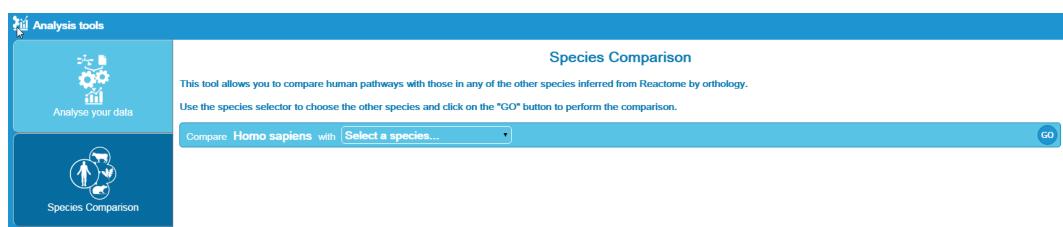
For an explanation of the results seen in the Identifier results tab see the explanation in the section Results for Identifier lists without associated numeric values.

Species Comparison

The manually-curated human pathways in Reactome are used to predict equivalent pathways in 20 other species. This automated, computational process is based on orthology. A full description of the inference process can be found on the Home page under Documentation, Orthology Prediction.

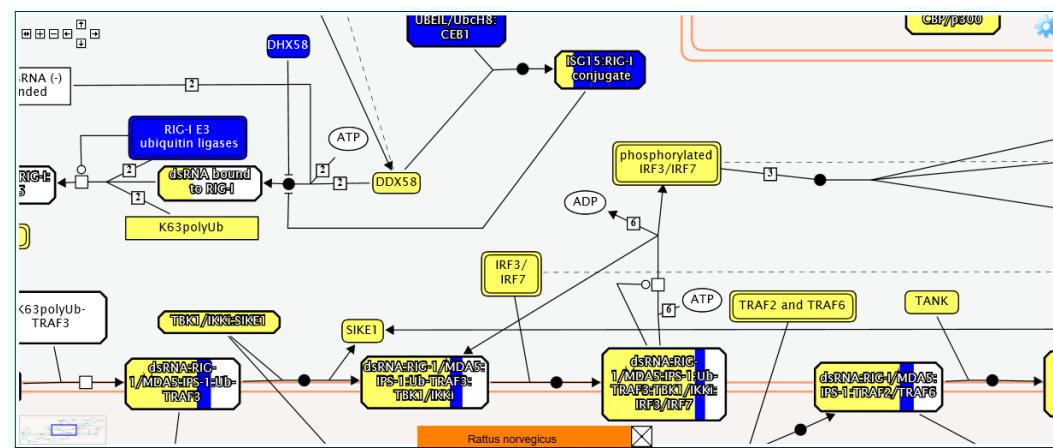
The Species Comparison tool allows you to compare human pathways with computationally-predicted pathways in model organisms, highlighting the elements of the pathway that are common to both species and those that may be absent in the model organism.

Species Comparison is launched using the Analysis button in the Pathway Browser header bar. In the Species Comparison section, select one of the species in the dropdown list. Click the Go button:



The results are ready to view when analysis results appear (or are updated if already present) in the Pathway Hierarchy.

Clicking on a pathway name launches the Pathway Browser and displays the relevant Pathway Diagram (see example below).



The colour of reaction objects indicates the result of the comparison:

- Yellow indicates that the protein has an inferred equivalent in the comparison species.
- Blue indicates that no equivalent was identified. This protein may not exist in the comparison species.

- White indicates that inference was not possible. This is always the case for small molecules, DNA and other objects that have no UniProt entry (or did not at the time the pathway was constructed).
- Objects with bands of colour represent complexes or sets containing more than one molecule. The bands of colour reflect the inference success for the molecules within the complex/set.

To view species comparison results for a complex or set right click it and select the option Display Participating Molecules.

This reveals a table representing all the proteins involved in the complex/set. Each square in the grid represents one component of the complex/set, coloured as described above.

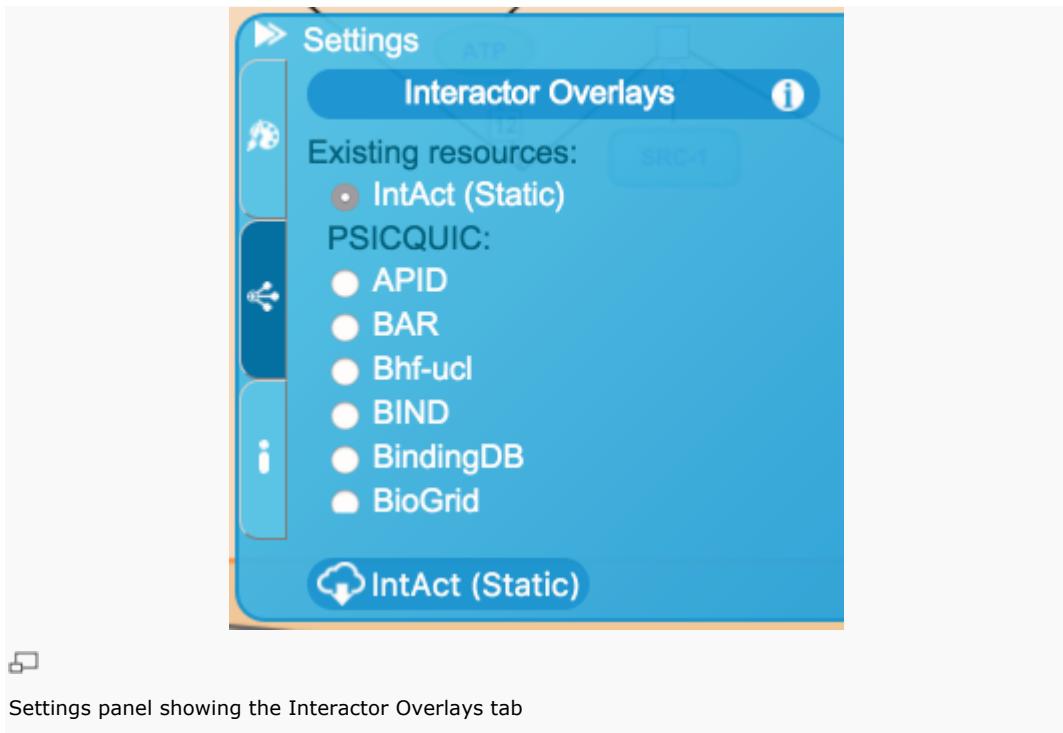
Component Name
Bound Myosin Actin
Complex
Component Name
Calcium [ChEBI:29108]
UniProt:P60660 MYL6
UniProt:P06753 TPM3
UniProt:P62736 ACTA2
UniProt:O60504 SORBS3
UniProt:P14649 MYL6B
UniProt:Q9Y490 TLN1
UniProt:P56199 ITGA1
UniProt:P18206 VCL

The orange species bar at the bottom of the Pathway Diagram (see example above) can be used to turn off species comparison colouring, by unchecking the box.

Refer to the Navigating Pathway Diagrams section for more information on the diagram content.

Molecular Interaction Overlay

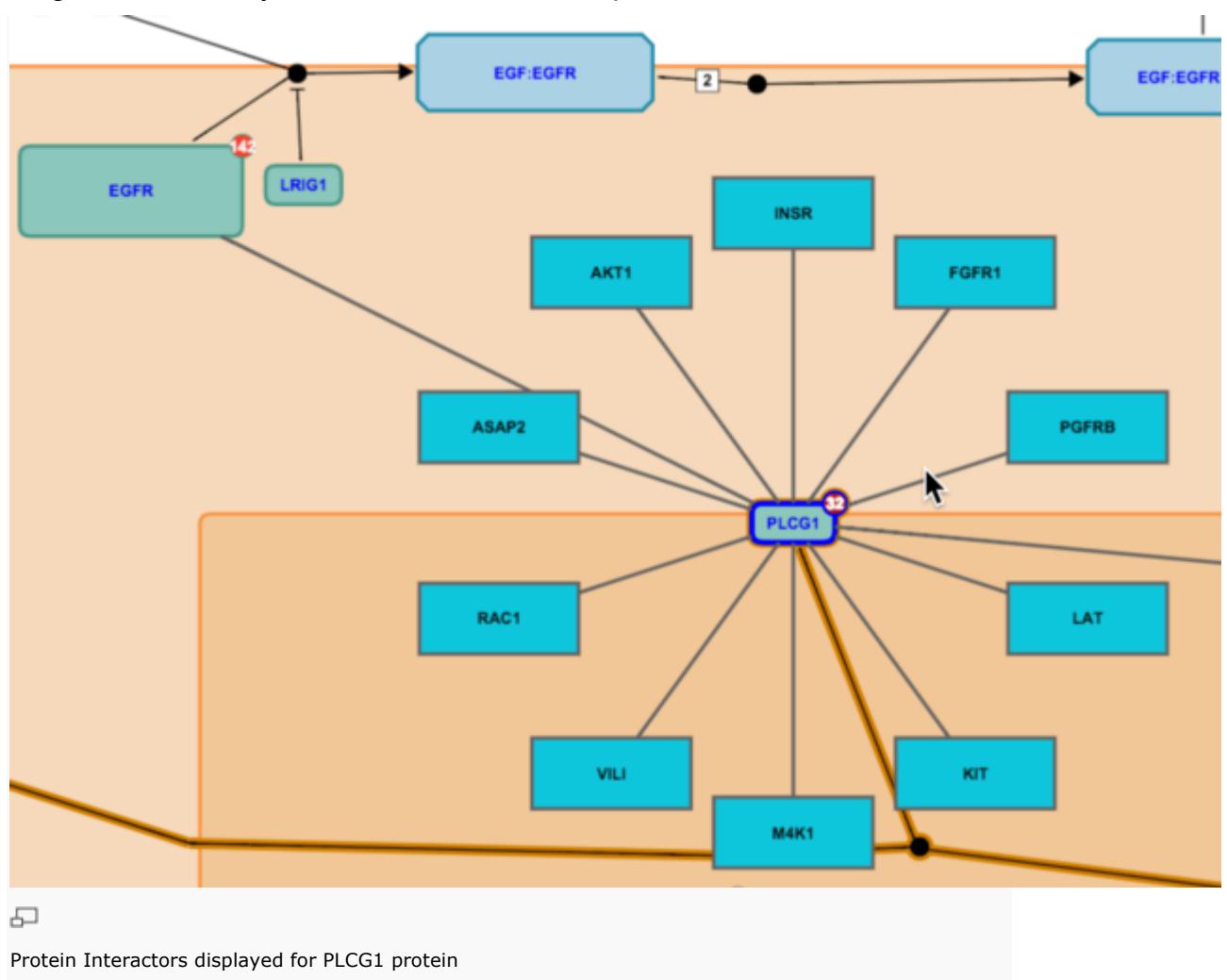
The Molecular Interaction (MI) overlay allows protein-protein or protein-small molecule interactions to be superimposed onto the pathway diagram. The source depends on the currently selected interaction database. The default interaction database is IntAct [IntAct (Static)], which provides fast access to a quarterly updated and locally-hosted version of IntAct data. Access to the IntAct-hosted dataset, and other PSICQUIC sources of interaction data (protein-protein and protein-small molecule) can be selected using the ‘PSICQUIC’ feature of the Interactor Overlays tab within the Settings panel in the middle-right of the Pathway diagram. A list allows selection of the source of interactors. This is automatically populated by querying the PSICQUIC Registry. If a new database is selected while interactors are displayed on the pathway diagram, the proteins represented by those entities will be used as queries in the new database and the display will automatically update. An ‘i’ (‘information’) button to the top right of the Interactor Overlays tab provides a quick tip about using the MI overlay. Selecting either ‘Cyan’ or ‘Teal’ in the ‘Interactors Colour Profile’ drop-down menu within the ‘Colour Palette’ feature of the Settings panel will change the interactor colour



Settings panel showing the Interactor Overlays tab

Once a pathway diagram is displayed in the viewport, a small red circle with a white-lettered number in the top right corner of the entity will automatically appear for individual protein and chemical entities that have

interactors. The number represents the number of known interactors for the given protein. Pressing the red circle with the cursor will display the interactors. A maximum of 10 interactors are displayed as a ring of blue-bordered boxes (protein interactors) or green-bordered ovals (small molecule interactors) connected by black lines to the selected protein. If the same interactor is connected to more than one entity it is re-used, i.e. connected to all the selected entities in the diagram. Furthermore, if the interactor is a protein or small molecule entity that pre-exists within the pathway diagram, a black line connects the two entities. Pressing the red circle a second time, will remove the interactors from the pathway diagram. If an entity interacts with itself, a loop-link is included.



A number of clickable actions provide additional information about the interactor and its relationship to the protein entity. These include:

- Hovering the mouse pointer over a protein interactor produces a tooltip containing the name and UniProt identifier of the protein.

- Clicking on an interactor opens the database entry for the selected protein or chemical in a new window.
- Clicking on the line that connects the interactor to the pathway item opens a new window containing details of the interaction at the source database.
- Right-clicking the protein entity in the pathway diagram will display a table of all the interactors for the selected entity. The interactive and scrollable table provides access to additional interactor information. Clicking the ‘Interactor’ name or accession identifier with connect to the UniProt or the interaction database, respectively. Clicking the ‘id’ button in the table will toggle the table to display either the interactor name or database identifier. Clicking the pin will lock the interactors table to the viewport. Clicking the ‘X’ in the table will close the table view.

The screenshot shows the IntAct Interaction Details page for interaction_id: EBI-9468482. The top navigation bar includes Home, Advanced Search, About, Resources, Download, and a search bar. A sidebar on the right shows interaction_id: EBI-9468482 and Example: BRCA2_Q09508 direct_1961411. Below the navigation, the page title is IntAct > Interaction Details. A tab menu at the top of the main content area includes Interactions (1), Interactors, Interaction Details (selected), and Graph. The main content is divided into sections: Publication, Experiment (1006 interactions), and Annotations.

Publication

PubMed Id: 24728074
Journal: Mol. Cell Proteomics (1535-9476)
Year of Publication: 2014
Cross References:

Database	Identifier	Secondary identifier	Qualifier
pubmed	24728074	-	primary-reference
imex	IM-22269	-	imex-primary

Experiment (1006 interactions)

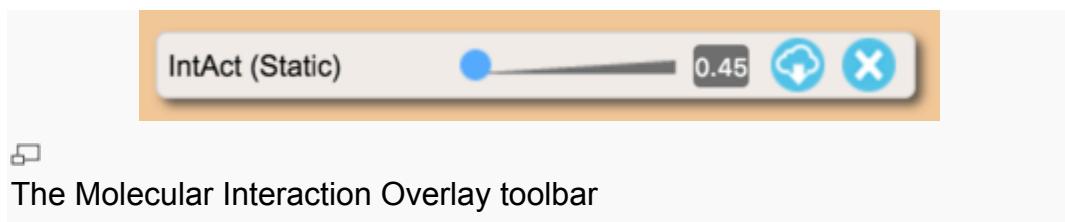
Accession: EBI-9451021 Host organism: In vitro
Name: leung-2014-1 Interaction Detection Method: fpo
Participant Identification Method: tag fluorescence
Cross References:
Annotations:

Topic	Text
data-processing	Some interactions in the experiment involve participants with identical sequences and may seem duplicated. They were kept separated by request of the authors, since they reflect independent replicates for unique pY sites despite being the same peptide sequence.
curation depth	imex curation

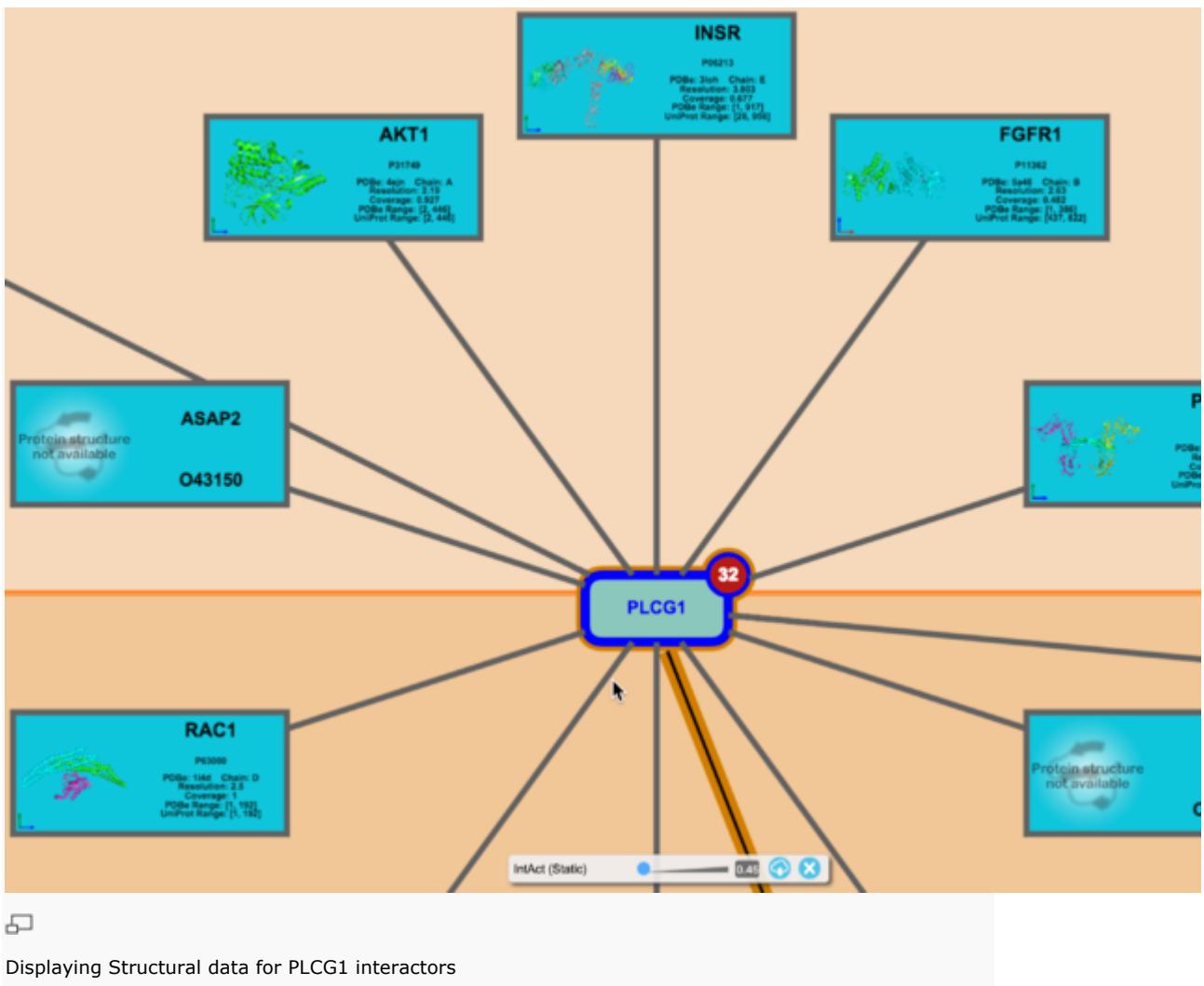
The IntAct entry for the interaction between PLGG1 and KIT

Each binary interaction is scored, using a simple addition of the cumulated value of a weighted score for the interaction detection method and the interaction type for each interaction evidence associated with that binary

pair. The default threshold is 0.45. Interactions with a confidence level equal to or above this threshold will be visible in the viewport. Moving the ‘Slider’ feature of the Molecular Overlay toolbar (light grey bar with dark grey slider at the bottom of the pathway diagram) to the right will adjust the confidence level and interactors below the displayed confidence level will disappear from the diagram. Moving the ‘Slider’ feature back to the left will result in the interactors reappearing with a confidence level equal to or above the threshold. Clicking the ‘X’ in the Molecular Overlay toolbar will remove the toolbar and all the interactors from the viewport.



Using the zoom in feature, to the bottom right corner of the Pathway Diagram or scrolling using the mouse, will increase relative size of the interactor icons. Once a size threshold is reached, a small image displaying the crystal structure of the protein or the chemical structure will appear in the interactor box or oval, respectively.



Viewing the Equivalent Pathway in another Species

Reactome is human-centric and aims to represent human biology.

Pathways in other species are electronically inferred from curated human pathways - a description of the inference process can be found [here](#).

To view the predicted conservation of a pathway in another species, select the species of interest in the Species dropdown menu in the top left corner of the Pathway Browser. The dropdown button will turn dark blue and the Pathway Browser will reset. Now select any pathway name from the Hierarchy to open the species-relevant computationally-predicted version of the human pathway. Any reaction events that could not be inferred from the human pathway will be absent. Use the dropdown again to switch to another species or back to human.

Details Views

Pathway Details

Pathway details appear in the Detail Panel at the bottom of the Pathway Browser.

The screenshot shows the Reactome Pathway Browser interface. At the top, there is a navigation bar with tabs: Overview, Molecules, Structures, Expression, Analysis, Processes, and Downloads. Below the navigation bar, the pathway name 'Platelet Adhesion to exposed collagen' is displayed, along with a 'Species: Homo sapiens' dropdown menu. The main content area is titled 'Stable Identifier' and shows 'REACT_1230.2'. A 'Summation' box contains a detailed text about platelet adhesion, mentioning integrins, collagen binding proteins, and the GpVI receptor. Below the summation box is a 'Literature references...' link. At the bottom of the panel, there are buttons for 'View computationally predicted event' and 'Select a species to go to...'. The overall layout is clean and organized, typical of scientific software interfaces.

The details are represented as boxes or panels, many contain additional levels of detail that can be expanded using a plus symbol on the right hand side. Note that there are often multiple levels of further details, which may not be apparent until the previous level is revealed. An icon to the left of molecule names will select that molecule in the diagram and reload the Details Panel with details of that molecule.

Pathway details always include a Summation, often with links to literature that provides general background for the pathway.

Other details within this section include:

Stable identifier – with the version number suffix removed, this can be used to refer to the pathway, providing a stable link that will not be affected by revisions. Click on the ID to see a revision history.

Authored - the expert biologists that contributed materials that allowed this pathway to be created in Reactome.

View computationally predicted event in – the dropdown list identifies the species that have been computationally predicted by the Reactome inference process to include this pathway.

References – this includes key literature sources that were used to create the pathway.

Reaction Details

Reaction details are similar to Pathway Details but have additional relevant sections:

Input/Output - identifies the input/output molecules, sets or complexes for this reaction. The composition of sets and complexes can be revealed using the Plus symbol on the right side of the panel.

Catalyst Activity – If the reaction has a catalyst this section identifies the molecule or complex, the associated GO Molecular Function term, and when relevant the molecule that has catalytic activity relevant to that function.

Cellular Compartment - identifies the cellular compartment for this reaction, with the associated GO term.

Preceding/Following Events - links to pathways that precede/follow this reaction.

View computationally predicted event in – provides a dropdown list of species that are computationally predicted to have a reaction that is functionally equivalent to the selected reaction.

Reviewed – names the expert that reviewed this reaction before it was added to the website.

Complex or Set Details

Details the members of the complex or set are displayed, similar to the details for reactions with some additional details:

Components - details the molecules that are part of the set or complex.

Produced by - details the reactions that generate the complex or set.

Consumed by - details the reactions that use the complex/set as an input.

Protein/Small molecule Details

Clicking on a single protein or small molecule in the Pathway Diagram, or where identified in the detail pane as part of a reaction, set or complex, updates the Detail Panel with details of that protein/small molecule.

The sections represented include:

External identifiers – lists cross-references in databases other than Reactome.

Coordinates in the Reference Sequence – identifies the peptide string that is represented, by reference to the peptide string available in the Primary External Reference database. For human Peptides, this is UniProt. Note that several peptide molecules may be derived from the same UniProt reference, allowing Reactome to represent the result of cleavage and post-translational modification events.