



Reactome

This tutorial introduces Reactome, the user interface and the database content. Exercises help you practice what you have learned; you will need to refer to the details and screenshots in this document. Further information can be found in the online Reactome user guide at <http://www.reactome.org/userguide/Usersguide.html>.

You will learn:

- How pathways and reactions are visualized in Reactome
- Querying Reactome through the web interface
- How to overlay interaction and expression data on Reactome pathways
- How to perform Pathway Overrepresentation Analysis in Reactome
- How to map datasets to Reactome pathways

Reactome	1
<i>You will learn:</i>	1
<i>What is Reactome?</i>	3
<i>What is this Tutorial For?</i>	3
<i>Exercises</i>	3
<i>The Reactome Home Page</i>	4
Exercise 1	4
<i>Searching Reactome</i>	5
Simple text search.....	5
<i>The Pathway Browser</i>	6
Exercise 2	9
<i>Pathway Diagrams</i>	10
Exercise 3	12
<i>Overview diagrams and Subpathway icons</i>	13
<i>Navigating Pathway Diagrams</i>	13
Exercise 4	14
<i>Details Panel</i>	15
The Overview tab	16
The Molecules Tab.....	17
The Structures Tab	18
The Expression Tab	19

Analysis Tab.....	19
Processes Tab	20
Downloads Tab	20
Exercise 5	21
<i>Disease in Reactome</i>	21
Exercise 6	23
Reactome Tools	23
<i>Analyse Data</i>	23
Results for Identifier lists without associated numeric values.....	26
Exercise 7	29
Results for Identifier lists with associated numeric values (expression representation)	31
Exercise 8	33
<i>Species Comparison</i>	35
Exercise 9	37
<i>Molecular Interaction Overlay</i>	38
The Search & Tools Icon.....	40
Exercise 10	42
<i>Other Features and Tools in the Pathway Browser</i>	42
Details Views.....	44
<i>The BioMart interface</i>	48
Exercise 11	50
<i>Answers to Exercises</i>	51
Exercise 1	51
Exercise 2	51
Exercise 3	51
Exercise 4	52
Exercise 5	52
Exercise 6	53
Exercise 7	53
Exercise 8	53
Exercise 9	54
Exercise 10	54
Exercise 11	55
<i>Printable Key to Reactome Pathway Diagrams</i>	56

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.

Exercises

The exercises will help you understand Reactome content. Some of the questions are not directly addressed in the tutorial but can be answered with a little deduction - if you get stuck the answers are available at the end of this document in Answers to Exercises.

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, there is a navigation bar with links: About, Content, Documentation, Tools, Community, Download, Contact. A search bar contains the text "e.g. O95631, NTN1, signaling" and a "Search" button. Below the navigation bar are six large icons: Browse Pathways, Analyze Data, Reactome FI Network, User Guide, Data Download, and Contact Us. To the right is a "Reactome News" section with three news items from @reactome, dated 23 Sep, 22 Sep, and 22 Sep. Below the news section are logos for OICR, NYU Langone Medical Center, CSH Cold Spring Harbor Laboratory, and EMBL-EBI. A footer section contains a grid of links for About, Content, Documentation, Tools, Community, Download, and Contact.

Reactome
A CURATED PATHWAY DATABASE

About Content Documentation Tools Community Download Contact e.g. O95631, NTN1, signaling Search

Browse Pathways **Analyze Data** **Reactome FI Network**

User Guide **Data Download** **Contact Us**

Reactome News

reactome @reactome 23 Sep
Version 46 Released: Topics with new or revised events include Disease (Defects in biotin metabolism, Defects ... tinyurl.com/n2e9z4z
Expand

reactome @reactome 22 Sep
Manteia database and data mining tools related to embryonic development is using Reactome data, manteia.igbmc.fr, [#usereactome](https://twitter.com/usereactome)

reactome @reactome 22 Sep
Michael Yaffe - how dynamic re-wiring of DDR signaling could be used to treat cancer [#EMBOmtg](https://twitter.com/EMBOmtg)

Tweet to @reactome

About Reactome

Reactome is a free, open-source, curated and peer reviewed pathway database. Our goal is to provide intuitive bioinformatics tools for the visualization, interpretation and analysis of pathway knowledge to support basic research, genome analysis, modeling, systems biology and education. The current version (v45) of Reactome was released on June 11, 2013.

OICR **NYU Langone** **CSH** **Cold Spring Harbor Laboratory** **EMBL-EBI**

The development of Reactome is supported by a grant from the US National Institutes of Health (P41 HG003751), Ontario Research Fund, and the European Molecular Biology Laboratory.

About **Content** **Documentation** **Tools** **Community**

About Reactome
News
Reactome Team
Scientific Advisory Board
Other Reactomes
License Agreement
Reactome Disclaimer

Table of Contents
DOIs
Data Schema
Editorial Calendar
Statistics

User Guide
Data Model
Orthology Prediction
Object/Relational Mapping
Wiki
Linking to Reactome
Referencing Reactome

Pathway Browser
Analyze Data
Species Comparison
Reactome FI Network
Advanced Search
Small Molecule Search
BioMart

Reactome Outreach
Reactome Events
Reactome Training
Reactome Publications
Papers Citing Reactome
Resources Guide
Mailing List

Download **Contact**

f **t** **YouTube**

Exercise 1

This exercise encourages you to examine the Homepage and find information via the menu bar.

1. What's the latest news item on the Reactome Home page?
2. How many human proteins are represented in Reactome?
3. What date is the next release?
4. What is the first new topic that will be included?

Searching Reactome

Simple text search

The simple text search tool is located top right of the Home page. Several other Reactome pages include this search tool. To search Reactome type a word, phrase or identifier in the search box:



Click **Search**. The results will be displayed:

Search for: in

Results 1 to 10 of 6416

Pathways (888) Reactions (2380) Proteins (601) Others (2547)

Pathway: Metabolism of proteins (Homo sapiens)
 Protein metabolism comprises the pathways of translation, post-translational modification and protein folding
 Last changed: 2013-09-06 10:20:53

Pathway: MAP kinase activation in TLR cascade (Homo sapiens)
 The mitogen activated protein kinase (MAPK) cascade, one of the most ancient and evolutionarily conserved signaling pathways, is involved in many processes of immune responses. The MAP kinases cascade transduces signals from the cell membrane to the nucleus in response to a wide range of stimuli (Chang and Karin, 2001; Johnson et al, 2002). There are three major groups of MAP kinases the extr
 Last changed: 2013-09-06 06:00:42

Pathway: Thrombin signalling through proteinase activated receptors (PARs) (Homo sapiens)
 Thrombin activates proteinase activated receptors (PARs) that signal through heterotrimeric G proteins of the G12/13 and Gq families, thereby connecting to a host of intracellular signaling pathways. Thrombin activates PARs by cleaving an N-terminal peptide that then binds to the body of the receptor to effect transmembrane signaling. Intermolecular ligation of one PAR molecule by another can occur but
 Last changed: 2013-09-06 00:55:41

Pathway: RAF/MAP kinase cascade (Homo sapiens)
 The MAP kinase cascade describes a sequence of phosphorylation events involving serine/threonine-specific protein kinases. Used by various signal transduction pathways, this cascade constitutes a common 'module' in the transmission of an extracellular signal into the nucleus
 Last changed: 2013-09-06 06:00:42

Pathway: MAPK targets/ Nuclear events mediated by MAP kinases (Homo sapiens)
 MAPKs are protein kinases that, once activated, phosphorylate their specific cytosolic or nuclear substrates at serine and/or threonine residues. Such phosphorylation events can either positively or negatively regulate substrate, and thus entire signaling cascade activity. The major cytosolic target of activated ERKs are RSKs (90 kDa Ribosomal protein S6 Kinase). Active RSKs translocates to the nuc
 Last changed: 2013-09-06 06:00:42

Search results have a type: Reaction; Pathway; Protein or Other. The type is indicated by an icon on the left and a bold type name preceding the title of the result. Click on the title to go to the corresponding Reactome web page. Most results will have descriptive text details. The term used as a query will be highlighted if it appears within the title or descriptive text (see the example above).

At the top of the results page is a set of check boxes labelled Pathways, Reactions, Proteins, and Others, each with a bracketed number showing the number of results of that type. The check boxes allow you to limit results by type. Uncheck boxes for categories that you don't want to see, e.g. if you only want to see Protein results, uncheck Pathways, Reactions and Others, then click the Show button.

The Pathway Browser

The Pathway Browser is the primary means of viewing and interacting with pathways in Reactome. It is an interactive pathway viewer and associated tools for 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.

See the section Reactome Tools for analysis tools details.

The Pathway Browser is launched by clicking the Browse Pathways button on the Homepage:



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

The screenshot shows the Reactome pathway browser interface. At the top, there are navigation buttons: Home button, Species Selector, Show/Hide, Quick Tour, and Diagram Key. Below these, the main interface is divided into three main sections:

- Hierarchy Panel:** A tree view on the left showing biological pathways, with 'Muscle contraction' selected.
- Pathway Diagram Panel:** A large central area displaying a detailed pathway diagram for 'MLCK Active Calmodulin Binding'. It includes a 'Zoom/Move Toolbar' and an 'Analyse Data' button.
- Details Panel:** A panel on the bottom right providing information about the selected pathway, including its stable identifier (REACT_20521.1), a summation, and input molecules like MYLK(1-1914) [cytosol] and Active Calmodulin [cytosol].

Other labeled components include a 'Search & Tools' button in the top right and a 'Thumbnail' of the pathway diagram in the bottom left of the diagram panel.

Home button – the Reactome logo acts as a shortcut button. Click on it to return to the homepage.

Species Selector – Reactome is primarily 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 Selector to view the pathways inferred for the selected organism.

Show/Hide toolbar – The pathway browser is divided into 3 main sections. These are 1) the Pathway Hierarchy, on the left, 2) the Details Panel, bottom right, 3) the Diagram Panel, top right. The Show/Hide Toolbar allows you to show or hide these panels.

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

Quick Tour – click here to be taken on a quick tour of Reactome's key features.

Diagram Key – this is a mini key to the objects present in pathway diagrams. Click on the downward-pointing arrow to reveal a complete key to all pathway diagram features.




Zoom/Move Toolbar - Click on the arrows to move the diagram. The central button resizes the Pathway Diagram to fit the available space. Click on the plus and minus buttons to zoom in and out. You can also click and drag the diagram, and zoom with your mouse scroll wheel.

Pathway Diagram 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).

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

Hierarchy Panel - Contains a hierarchically organised list of the events, i.e. pathways and reactions, within Reactome. When first opened, the hierarchy shows only the top level of the hierarchy, an alphabetical list of Reactome's main topics. Topics such as apoptosis are too large to be considered or represented as a single pathway. Instead they are divided into sub-pathways, which may be further divided into sub-sub-pathways. Most pathway topics are divided into sub-pathways in this hierarchical manner. Individual events or steps in a pathway are known as reactions; connected reactions form a pathway. The pathway hierarchy represents Reactome's organisation of pathways, sub-pathways and reactions. 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 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 are not always 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 may 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.

Exercise 2

This exercise is to check that you understand the Pathway Hierarchy.

From the Home page, search for PDGF signaling.

Click on the top pathway hit. This will open it in the Pathway Browser.

Ignore the diagram, look at the Pathways tab on the left.

1. How many sub-pathways does this pathway have?
2. How many reactions are in the first sub-pathway?
3. What reaction(s) follow(s) Processing of classical PDGFs?

Hint: Look in the Details pane bottom-right. If it is hidden, use the panel buttons to the right of the Species selector, above the diagram, to reveal it.

Pathway Diagrams

Pathway Diagrams represent the steps of a pathway as a series of interconnected molecular events, known in Reactome as 'reactions'.

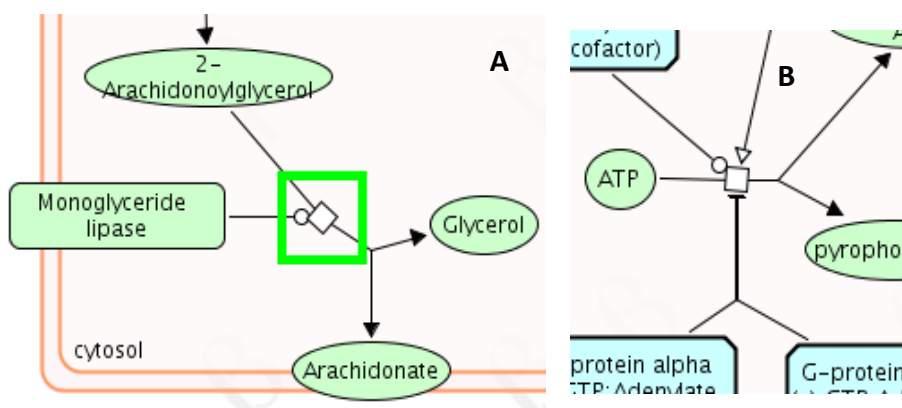
Cellular compartments are represented as pink boxes - a typical diagram has a box representing the cytosol, bounded by a double-line that represents the plasma membrane. The white background outside this represents the extracellular space. Other organelles are represented as additional labelled boxes within the cytosol.

Molecules 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 molecule on the boundary of the cytosol is in the plasma membrane.

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 are individual proteins, sets of proteins, or sets that contain proteins and small molecules.
- Green ovals are small molecules or sets of small molecules
- Blue boxes are complexes, i.e. proteins and/or small molecules that are bound to each other.
- 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 are connected to the set object by a line of short dashes.
- Sets with overlapping content are connected by a line of long dashes.
- Reactions that represent a disease process use red connecting lines. Objects associated with a disease state 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.

Exercise 3

This exercise is to check that you understand reaction subtypes.

From the Home page, search for the pathway 'Effects of PIP2 hydrolysis' and open it in the Pathway Browser.

1. What symbol represents the reaction for 'Binding of IP3 to the IP3 receptor'?
2. What symbol represents the reaction 'Transport of Ca⁺⁺ from platelet dense tubular system to cytoplasm'? What subtype of reaction is this?
3. Open the subpathway 'Arachidonate production from DAG'. What is the name of the catalyst for '2-AG hydrolysis to arachidonate by MAGL'? Can you find its UniProt ID and name the three outputs of this reaction?

Overview diagrams and Subpathway icons

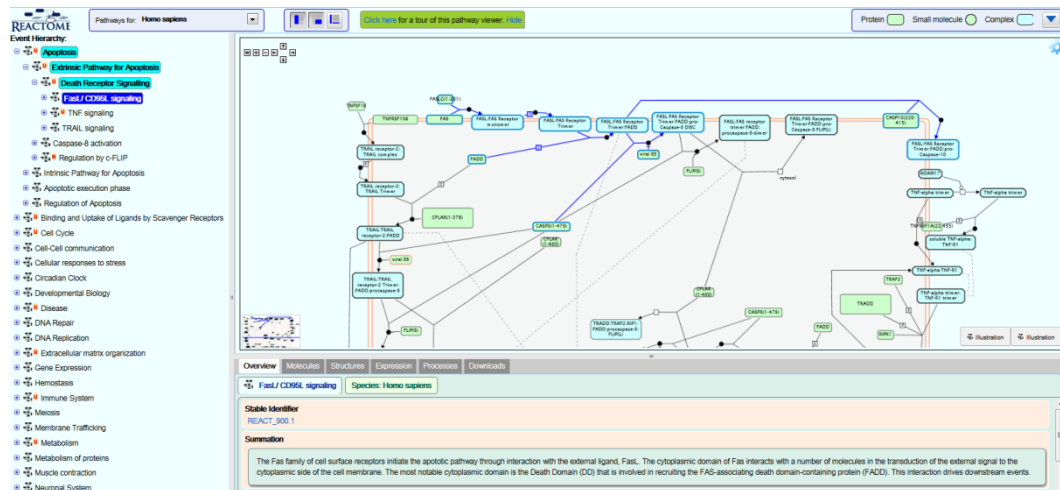
Topics such as Apoptosis are too large to be displayed in a single Pathway Diagram. Instead they are divided into sub-pathways, each with its own Pathway Diagram. Topics that are too large for a single pathway diagram are represented as an Overview Diagram, containing one or more Subpathway Icons, when the pathway is selected in the Hierarchy. Subpathway Icons are boxes with a green boundary. A right-click on the Subpathway Icons will produce a menu including the option 'Go To Pathway', select this to open the corresponding Pathway Diagram. Alternatively, click to select the Subpathway Icons, which will cause its name to be highlighted in pale blue in the Hierarchy. Its parent pathway will be highlighted in dark blue, and details of the pathway will appear in the Details Panel. These details will include a summary of the pathway, which may help you decide whether you want to open the Pathway Diagram. If you do, click on the pale blue highlighted name in the Hierarchy. In the figure below, the Subpathway Icon 'Extrinsic Pathway for Apoptosis' is selected and consequently highlighted in blue. Selection has also caused this Subpathway Icons to be highlighted in pale blue, and its parent pathway to be highlighted in dark blue, on the Pathway Hierarchy.

The screenshot displays the REACTOME pathway viewer interface. On the left, the 'Event Hierarchy' panel shows a tree structure of biological processes. The 'Apoptosis' category is expanded, and the 'Extrinsic Pathway for Apoptosis' is selected, highlighted in dark blue. Below it, 'Intrinsic Pathway for Apoptosis' and 'Apoptotic execution phase' are highlighted in pale blue. The main area shows an overview diagram with four subpathway icons: 'Extrinsic Pathway for Apoptosis' (blue border), 'Intrinsic Pathway for Apoptosis' (green border), 'Apoptotic execution phase' (green border), and 'Regulation of Apoptosis' (green border). Arrows indicate relationships between these icons. The bottom panel shows the 'Overview' tab for the 'Extrinsic Pathway for Apoptosis' in 'Homo sapiens'. It includes the DOI (10.3180/REACT_1059.1), the stable identifier (REACT_1059.1), and a summary text: 'Known as the "death receptor pathway" the extrinsic or caspase 8/10 dependent pathway is activated by ligand binding. The "death receptors" are specialized cell-surface receptors including Fas/CD95, tumor necrosis factor-alpha (TNF-alpha) receptor 1, and two receptors, DR4 and DR5, that bind to the TNF-alpha related apoptosis-inducing ligand (TRAIL). The extrinsic and intrinsic pathways unite in the activation of Caspase-3, though the two pathways communicate through the pro-apoptotic Bcl-2 family member Bid before uniting at the shared activation of Caspase-3.'

Navigating Pathway Diagrams

The Pathway Diagram, Hierarchy and Details Panel are interactively connected. Selecting any molecular object or event object 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 the diagram that is already displayed, clicking on 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 sub-pathway FasL/ CD95L signaling has been selected in the Hierarchy, causing all the reactions in this sub-pathway to be highlighted 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 to show it (this may take a few seconds).

You can zoom in or out of a diagram using the navigation tool in the top-left corner of the Pathway Diagram Panel, or by using the mouse scroll wheel (in most browsers). You can also click and drag the diagram.

Exercise 4

This exercise is to check that you understand Subpathway icons and the relationship between the Pathway Hierarchy and Pathway Diagram.

Open the pathway Apoptosis in the Pathway Browser.

Select the subpathway icon 'Intrinsic Pathway in Apoptosis'.

1. How do you open the Pathway Diagram for this sub-pathway?
2. With this Pathway Diagram open, what happens if you click on the sub-pathway 'BH3-only proteins associate with and inactivate anti-apoptotic BCL-2 members'?
3. What happens if you click on the reaction 'Sequestration of tBID by BCL-2' in the in the pathway hierarchy?

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 button on the Show/Hide Panel.

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 screenshot shows the 'Details Panel' for the reaction 'TRAIL:TRAIL receptor-2 Trimer Binds FADD' in the Reactome Pathway Browser. The panel is organized into several sections:

- Overview:** Shows the reaction name and the species 'Homo sapiens'.
- Stable Identifier:** REACT_1721.2
- Summation:** The trimeric complex of TRAIL and TRAIL receptor-2 (TRAIL:TRAIL receptor-2) binds FADD (Sprick et al. 2000).
- Input:** TRAIL receptor-2:TRAIL Trimer [plasma membrane] and 2 x FADD [cytosol].
- Output:** TRAIL:TRAIL receptor-2:FADD [plasma membrane].
- Preceding Event(s):** TRAIL Binds TRAIL-Receptor2 [Homo sapiens].
- Following Event(s):** TRAIL:TRAIL-Receptor2 Trimer:FADD complex binds Caspase-10 [Homo sapiens] and TRAIL:TRAIL-Receptor2 Trimer:FADD complex binds Caspase-8 [Homo sapiens].
- Computationally inferred to:** Select a species to go to...
- Negatively regulated by:** CFLAR(1-376) [cytosol].
- References:** FADD/MORT1 and caspase-8 are recruited to TRAIL receptors 1 and 2 and are essential for apoptosis mediated by TRAIL receptor 2; Receptor-mediated endocytosis is not required for tumor necrosis factor-related apoptosis-inducing ligand (TRAIL)-induced apoptosis.
- Reviewed:** Doi: 10.1093/nar/gkz1000

The Details Panel has several tabs:

The Overview 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 Details Panel 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.

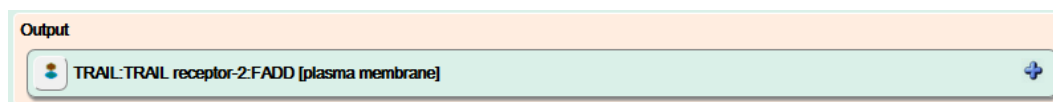
Cellular Compartment - Identifies the cellular compartment for the reaction, with the associated Gene Ontology (GO) term.

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.



Clicking the plus icon reveals further details:

Output

- TRAIL:TRAIL receptor-2:FADD [plasma membrane]
 - Species: Homo sapiens
 - Synonyms: TRAIL:TRAIL receptor-2:FADD
 - Has components:
 - TRAIL receptor-2:TRAIL Trimer [plasma membrane]
 - FADD [cytosol]
 - 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
 - External cross-references
 - FADD [cytosol]

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

This tab contains a List button which displays details of the molecules associated with the event or object selected in the pathway diagram, including links and accession numbers in various external databases.

List Download

- Proteins (4 molecules)
 - CFLAR(1-376) [cytosol] (4 external references)
 - UniProt >> [O15519]
 - GeneCards >> G [O15519]
 - UCSC human >> [O15519]
 - RefSeq >> [NP_003870]
 - FADD [cytosol] (6 external references)
 - TNFRSF10B [plasma membrane] (6 external references)
 - TNFSF10 [extracellular region] (6 external references)

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.

List

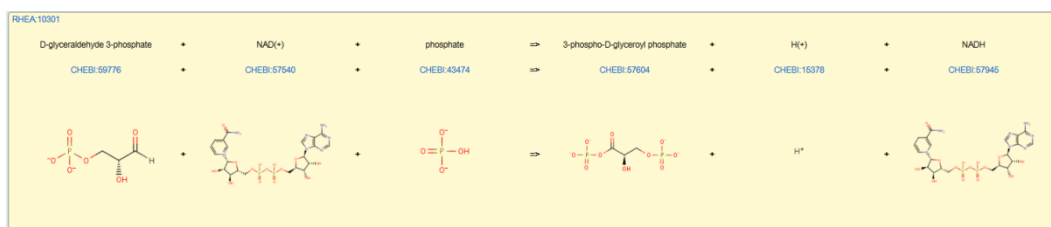
Molecules Type (select as required)

Fields (select as required)

Formats (select one)

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

PDBe

▶ All other structures for O15519

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

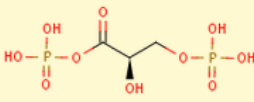
ChEBI Name 3-phospho-D-glyceroyl dihydrogen phosphate

ChEBI ID [CHEBI:16001](#)

Definition The (R)-enantiomer of 3-phosphoglyceroyl dihydrogen phosphate.

Stars ★★

Secondary ChEBI IDs CHEBI:1658, CHEBI:11881, CHEBI:20189



The Expression Tab

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

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

E-MTAB-513 RNA-Seq of human individual tissues and mixture of 16 tissues (Illumina Body Map)
 Organism(s): *Homo sapiens*
 Reference(s): 22496456 22955988 23258890

Showing 4 of 4 genes found:

Display levels

Gene	adipose	adrenal	brain	breast	colon	heart	kidney	leukocyte	liver	lung	lymph node	ovary	prostate	skeletal muscle	testis	thyroid
CFLAR			brain													
TNFSF10																
TNFRSF10B																
FADD																

Analysis Tab

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

Processes Tab

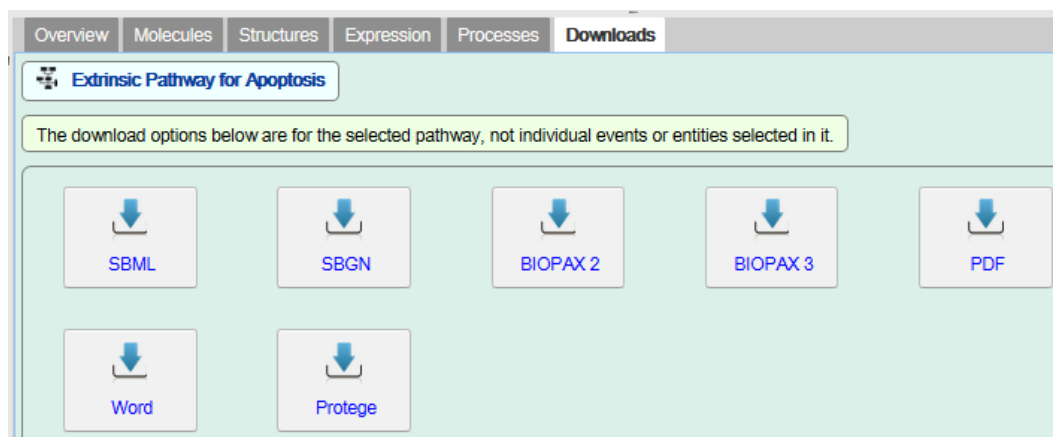
Events and molecular objects in Reactome are re-used if they can take part in more than one process. You can identify the involvement of a molecule or event in multiple processes by referring to the Processes Tab. The layout varies depending on the object or event that is selected in the pathway diagram, but will resemble the figure below.

This table displays all the Reactome processes that involve the protein CD47. It participates in 2 pathways, is an input to 3 reactions, is part of 4 complexes and is present within a set. Panels in the table behave as described above; clicking on the name or plus icon will open further details of the selected object or event, clicking on the icon on the left side will select the object/event and either highlight it in the currently displayed pathway diagram or open the relevant pathway diagram and select the object in this.

Overview	Molecules	Structures	Expression	Processes	Downloads
● CD47 [plasma membrane] Species: Homo sapiens					
Involved in pathways		Cell surface interactions at the vascular wall [Homo sapiens] +			
		Signal regulatory protein (SIRP) family interactions [Homo sapiens] +			
Is an input in reactions		CD47 binds SIRP [Homo sapiens] +			
		SIRP alpha binds CD47 [Homo sapiens] +			
		SIRP gamma binds CD47 [Homo sapiens] +			
Is present in complexes		CD47 bound to its ligand [plasma membrane] +			
		SIRP gamma:CD47 [plasma membrane] +			
		SIRPA:CD47 [plasma membrane] +			
		pSIRP alpha:CD47 [plasma membrane] +			
Is present in sets		Integrin alpha5beta1, Integrin alphaVbeta3, CD47 [plasma membrane] +			

Downloads Tab

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



Exercise 5

This exercise checks that you understand the Details Panel.

1. Find the reaction Activated type I receptor phosphorylates SMAD2/3 directly. What pathway does it belong to?
2. In which cellular compartment does this reaction take place?
3. What is the GO molecular function associated with the catalyst?
4. What references verify this reaction?
5. Is this reaction predicted to occur in *Canis familiaris*? In *Saccharomyces cerevisiae*?
6. Is this event likely to occur in liver?
7. Are 3D structures available for TGF-beta1?

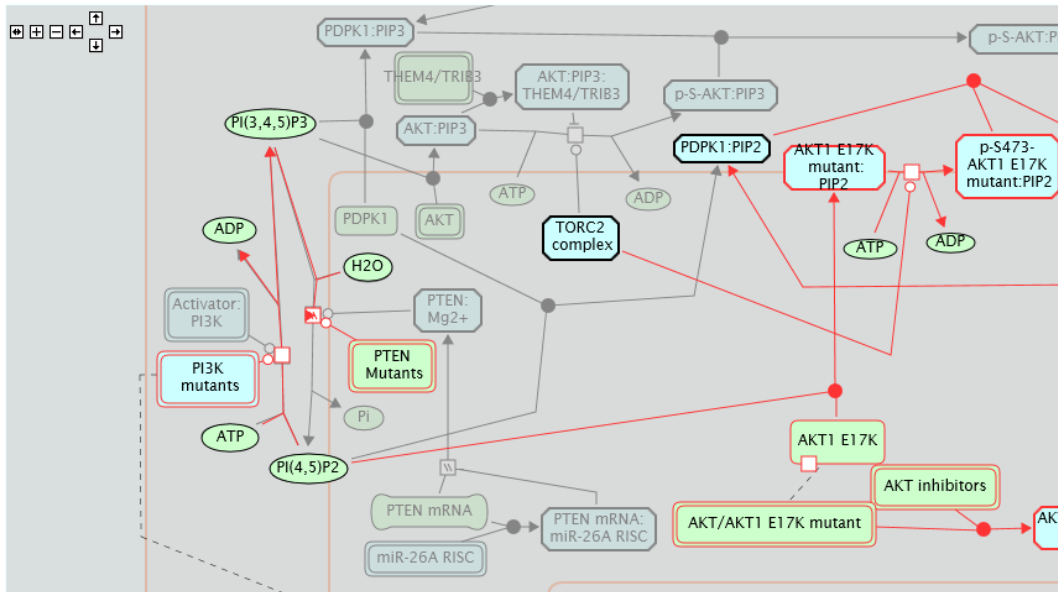
Disease in Reactome

A recent feature in Reactome is the annotation and display of diseases.

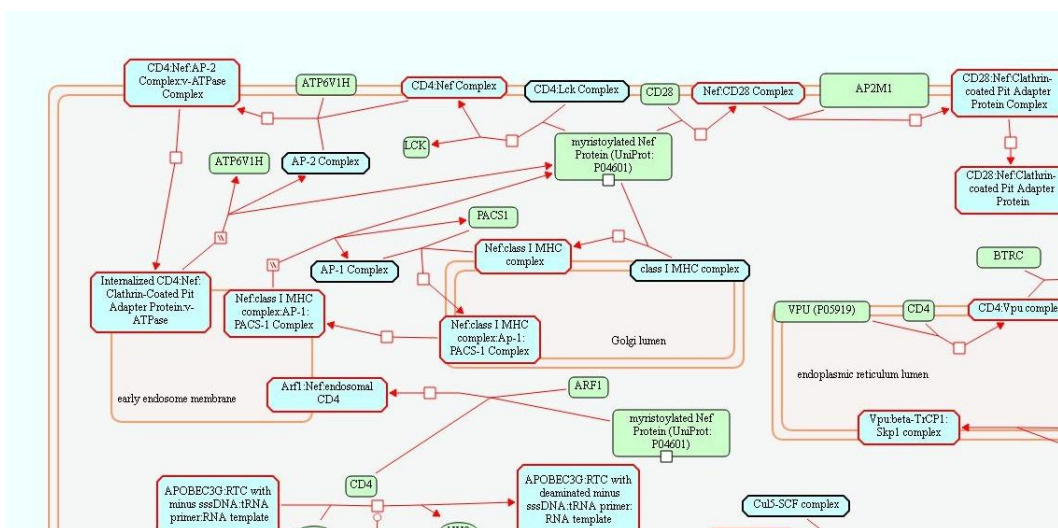
Disease events or processes are overlaid onto their normal counterparts to provide a pathway context to the disease event. Disease events have red connecting lines; abnormal molecules involved in these processes are outlined in red.

For cancer-related processes, it is useful to view the altered disease events alongside their normal counterparts in the same diagram. The user

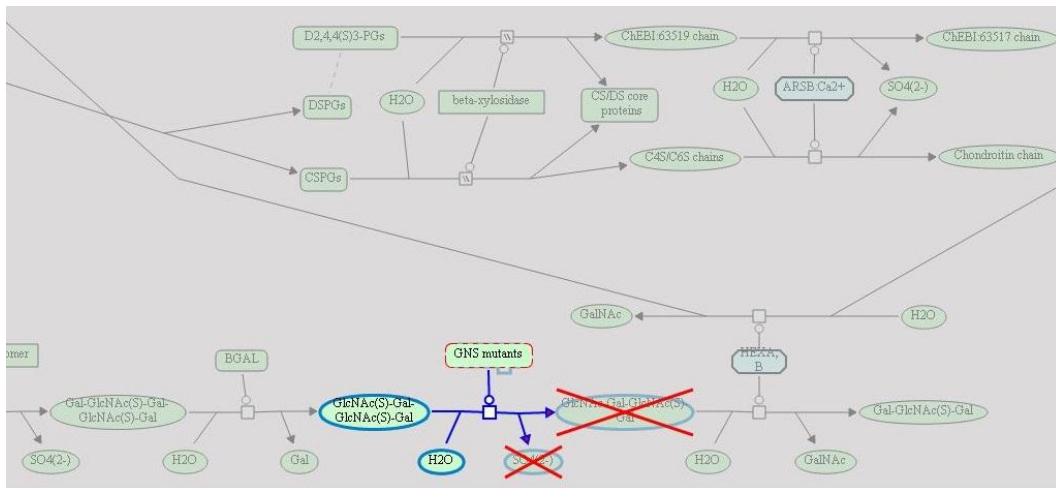
can then see where normal processes diverge into ones that are implicated in cancer.



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.



Exercise 6

1. In the Pathway Browser, find the Disease section. Open the pathway “Signaling by FGFR in disease” and then click on the sub-pathway “Signaling by FGFR mutants”. What disease(s) is/are associated with this pathway? (Hint: look in the Overview tab in the details section).
2. In the Pathway Browser, find the Disease section and under it, click the pathway “Abnormal metabolism in phenylketonuria”. In phenylketonuria, the enzyme that normally catalyses the conversion of phenylalanine to tyrosine is defective. What is the name of the defective enzyme? What is the residue mutation that causes the enzyme to become non-functional? (Hint: look at the diagram).
3. In the mucopolysaccharidoses disease section, open the “MPS1 – Hurler syndrome” sub-pathway and click on the defective reaction it contains. Mutations in the enzyme IDUA cause accumulation of its substrates. What are the substrates that accumulate when IDUA is defective? (Note that there are three reactions involving defective IDUA, each with an equivalent normal reaction. You may want to look at the Details of these).
4. From the homepage, search for the reaction “Defective MMADHC does not bind MMACHC:B12r”. The catalyst has lost its function. What has caused this loss of function? (Hint: look in the Overview tab at “Functional status”).
5. How many mutations of MMADHC are represented?

Reactome Tools

Analyse Data

Please Note – This tutorial refers to a Beta version of analysis tools at time of writing only available on our development site:

<http://reactomedev.oicr.on.ca/PathwayBrowser/>

This new version of analysis tools will shortly replace the tools found on the main Reactome website.

The Analyse data panel is opened in the Pathway Browser by clicking on the Analyse Data button on the Pathway Browser header bar, boxed in red in the image below:



When these tools are moved from the beta site to the main Reactome site
Alternatively, from the Homepage select the Analyse 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.

Analysis Tools

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 "Analyse" button to perform the analysis.

Select data file for analysis No file selected. Project to human

▶ [Click here to paste your data or try example data sets...](#)

Species Comparison

This tool allows you to compare human pathways with those in any of the other species inferred from Reactome by orthology.

Use the species selector to choose the other species and click on the "Compare" button to perform the comparison.

Compare Homo sapiens with Select a species...

There are two sections to the submission form. Use the Analysis Tools section to submit your own 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 overrepresentation and pathway-topology analyses are run. Overrepresentation 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.

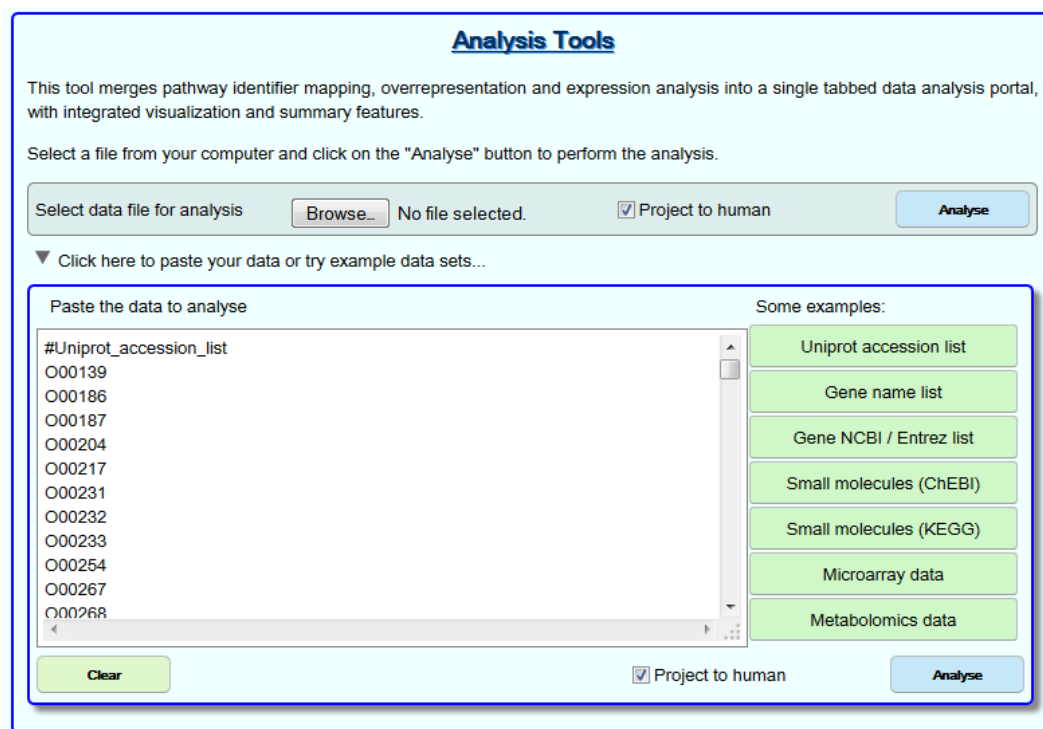
If your data has one or more additional columns of numbers, it will be recognised as expression data and expression overlay will be run. The numbers are used to produce a coloured overlay for Reactome pathway diagrams.

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/DBJ, 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.

Below is an example of the identifiers-only format:



Analysis Tools

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 "Analyse" button to perform the analysis.

Select data file for analysis No file selected. Project to human

▼ Click here to paste your data or try example data sets...

Paste the data to analyse

```
#Uniprot_accession_list
O00139
O00186
O00187
O00204
O00217
O00231
O00232
O00233
O00254
O00267
O00268
```

Some examples:

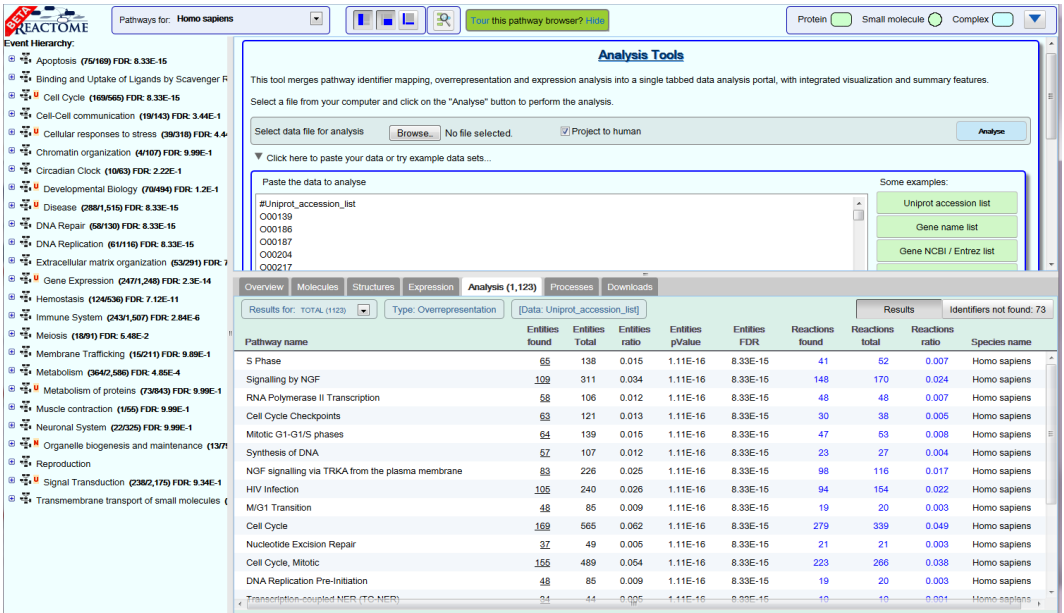
-
-
-
-
-
-
-

Project to human

Click the Analyse 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 Tools

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 "Analyse" button to perform the analysis.

Select data file for analysis: No file selected. Project to human

Click here to paste your data or try example data sets...

Paste the data to analyse

Some examples:

-
-
-

Pathway name	Entities found	Entities Total	Entities ratio	Entities pValue	Entities FDR	Reactions found	Reactions total	Reactions ratio	Species name
S Phase	65	138	0.015	1.11E-16	8.33E-15	41	82	0.007	Homo sapiens
Signalling by NGF	109	311	0.034	1.11E-16	8.33E-15	148	170	0.024	Homo sapiens
RNA Polymerase II Transcription	58	106	0.012	1.11E-16	8.33E-15	48	48	0.007	Homo sapiens
Cell Cycle Checkpoints	63	121	0.013	1.11E-16	8.33E-15	30	38	0.006	Homo sapiens
Mitotic G1-G1/S phases	64	139	0.015	1.11E-16	8.33E-15	47	63	0.008	Homo sapiens
Synthesis of DNA	67	107	0.012	1.11E-16	8.33E-15	23	27	0.004	Homo sapiens
NGF signalling via TRKA from the plasma membrane	83	226	0.025	1.11E-16	8.33E-15	98	116	0.017	Homo sapiens
HIV Infection	105	240	0.028	1.11E-16	8.33E-15	94	164	0.022	Homo sapiens
M/G1 Transition	48	85	0.009	1.11E-16	8.33E-15	19	20	0.003	Homo sapiens
Cell Cycle	169	565	0.062	1.11E-16	8.33E-15	279	339	0.049	Homo sapiens
Nucleotide Excision Repair	37	49	0.005	1.11E-16	8.33E-15	21	21	0.003	Homo sapiens
Cell Cycle, Mitotic	155	489	0.054	1.11E-16	8.33E-15	223	266	0.038	Homo sapiens
DNA Replication Pre-Initiation	48	85	0.009	1.11E-16	8.33E-15	19	20	0.003	Homo sapiens
Transcription-coupled NER (TC-NER)	24	44	0.025	1.11E-16	8.33E-15	10	10	0.001	Homo sapiens

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 (use the plus buttons to reveal subpathways and associated values).

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 of of these subsets will re-display results that consider only the selected molecular subtype.

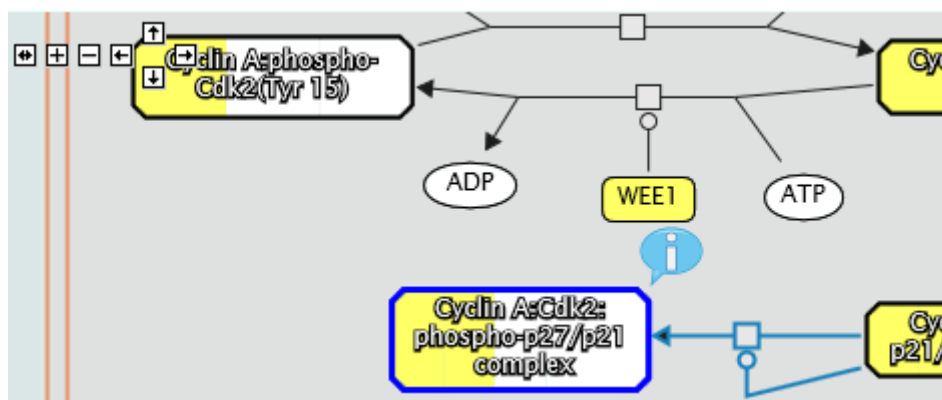
Results for: UNIPROT (1123) [Data: Uniprot_accession_list]

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. 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. 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.

Clicking on a pathway name opens the appropriate Pathway Diagram and selects it in the Pathway Hierarchy. Entities in the Diagram are coloured yellow if they were represented in the submitted data set, white if they were not. Complexes, Sets and Subpathway Icons are coloured to represent the proportion that is represented in the submitted identifier list. In the figure below, WEE1 is yellow indicating that it was in the submitted list. ADP was not in the submitted list so it is white. The complex Cyclin A:Cdk2:phospho-p27/p21 is part yellow, part white indicating that some molecules in the complex were represented in the submitted dataset while others were not.



To find out more hover the mouse pointer over a complex, and when the 'information' icon appears click it to see a menu of options. One of these is 'display participating molecules'. Click this to generate a colour-coded table of the molecules within the complex:

**Components for Cyclin
A:Cdk2:phospho-p27/p21
complex**

Component Name
p-T187-CDKN1B [nucleoplasm]
CDK2 [nucleoplasm]
CCNA2 [nucleoplasm]
CCNA1 [nucleoplasm]
p-S130-CDKN1A [nucleoplasm]

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

Q96KG3
P3Q1308521
P162095P533967
O15421
P43P00751
P106435
P1F017326
P18440861
P064937

Exercise 7

This exercise is to check that you understand the results of Pathway Analysis.

On the Home page, open the Pathway Browser. Click on the Analyse Data button. When the submission form appears, in the Analysis Tools section select 'Click here to paste your data or try example data sets...'

..and click on the Uniprot accession list button on the right side.

Your submission page should look like this:

Analysis Tools

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 "Analyse" button to perform the analysis.

Select data file for analysis No file selected. Project to human

▼ [Click here to paste your data or try example data sets...](#)

Paste the data to analyse

```
#Uniprot_accession_list
O00139
O00186
O00187
O00204
O00217
O00231
O00232
O00233
O00254
O00267
O00268
```

Some examples:

-
-
-
-
-
-
-

Project to human

Click the Analyse button.

1. What pathway is most over-represented in this dataset? *Hint – it may help if you sort by the column Ids in Pathway*
2. How many IDs match this pathway?
3. Roughly what proportion of molecules in the pathway are matched?
4. What proportion of reactions are matched?

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

If you submit data in a format that includes columns of numeric values, following a first column of protein, compound or other suitable identifiers such as probe IDs, 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.

Analysis Tools

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 "Analyse" button to perform the analysis.

Select data file for analysis No file selected. Project to human

▼ Click here to paste your data or try example data sets...

Paste the data to analyse

```
#Probeset 10h_control 10h 14h 18h 24h
1053_at 8.040078 7.147358 6.706705 6.794622 7.475157
1729_at 6.869688 6.99104 7.129922 7.112222 7.04721
1861_at 6.437999 6.620092 6.20117 6.407735 5.717815
200000_s_at 9.381569 9.710802 9.874874 9.934639 9.495911
200002_at 12.555275 12.511045 12.564419 12.538642 12.439174
200003_s_at 12.401259 12.054083 12.275169 12.206342 12.015476
200005_at 9.609852 9.099299 9.73072 9.530097 9.194303
200012_x_at 12.486269 12.402275 12.302666 12.256543 12.282444
200014_s_at 10.371458 9.548578 9.978313 9.871472 8.753136
200016_x_at 12.110468 11.913288 11.938524 11.899243 11.458105
200022_at 12.205038 11.927471 12.064725 12.031422 11.932256
```

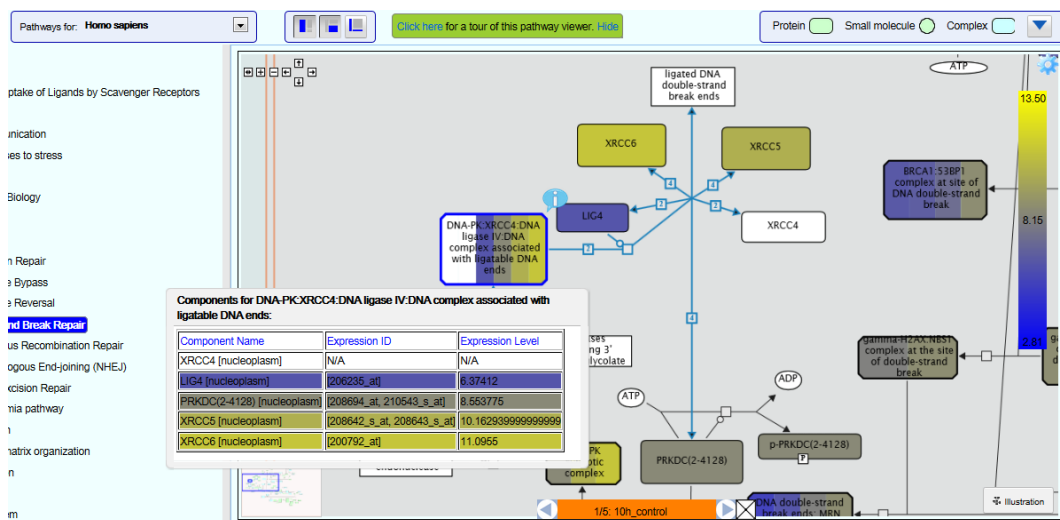
Some examples:

-
-
-
-
-
-
-

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 coloured according to the numeric values submitted. The colours are based on a scale as represented in a bar on the right hand side, with the highest values represented in bright yellow, through pale yellow to pale blue, to bright blue to represent the

lowest values. The scale (see figure above) automatically adjusts to fit the range of values in the dataset.

White is used for objects that were not represented in the input data.

Objects with bands of colour represent complexes or sets containing more than one molecule. The bands of colour reflect the values of the molecules within the complex/set.

To view details of the components of a complex or set right click it and select the menu option 'Display Participating Molecules'. The complex is represented as a table, each row representing a component of the complex/set (see figure above).

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.

Exercise 8

This exercise is to check that you can run and interpret the results of Expression Analysis.

Launch Expression Analysis and load the example dataset (*The data was obtained from ArrayExpress and comes from the processed results of an experiment studying the differentiation of KG1 cells.*). Click Analyse. When the results are displayed, find the pathway **DNA Repair**,

1. How many proteins are in this pathway?
2. What proportion of these had expression data?
3. Which pathway has the highest average expression value at 24h?

Click on the pathway name **DNA repair** to see it in the Pathway Browser.

4. Which subpathway has the greatest coverage?
5. Which subpathway has predominantly low expression levels for represented values?
6. Which has predominantly high expression levels for represented values?

Select the Subpathway Icon for **Nucleotide Excision Repair** (Right-click, Go to Pathway). Use the Experiment Browser toolbar to cycle through the timepoints.

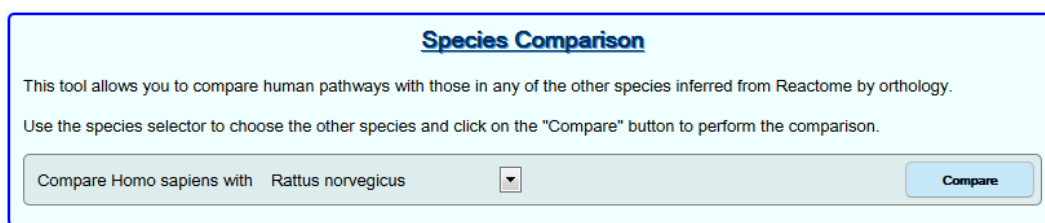
7. Which protein has the greatest change of expression between timepoints?
8. Find the complex 'Active Pol II complex with repaired DNA template:mRNA hybrid' (middle-right of the diagram). Which component of the complex has the highest expression at 24h?
9. What was the probe ID used to measure expression of this component?

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.

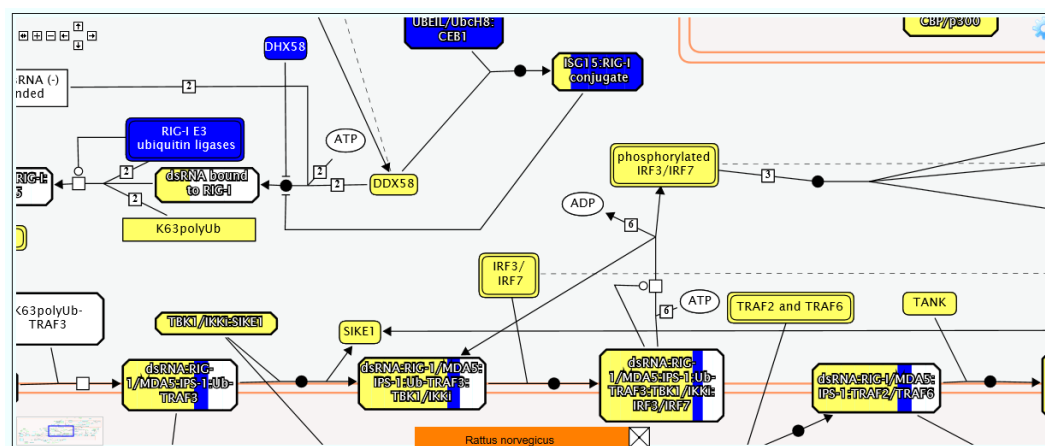
The Species Comparison tool allows you to compare human pathways with computationally-predicted pathways in model organisms, highlighting the elements of the pathway that are common to both species and those that may be absent in the model organism.

Species Comparison is launched using the Analyse Data button in the Pathway Browser header bar. In the Species Comparison section, select one of the species in the dropdown list. Click the Compare 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).



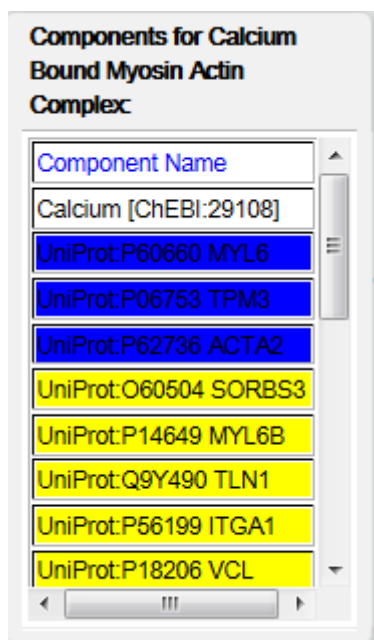
The colour of reaction objects indicates the result of the comparison:

- 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
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.

Exercise 9

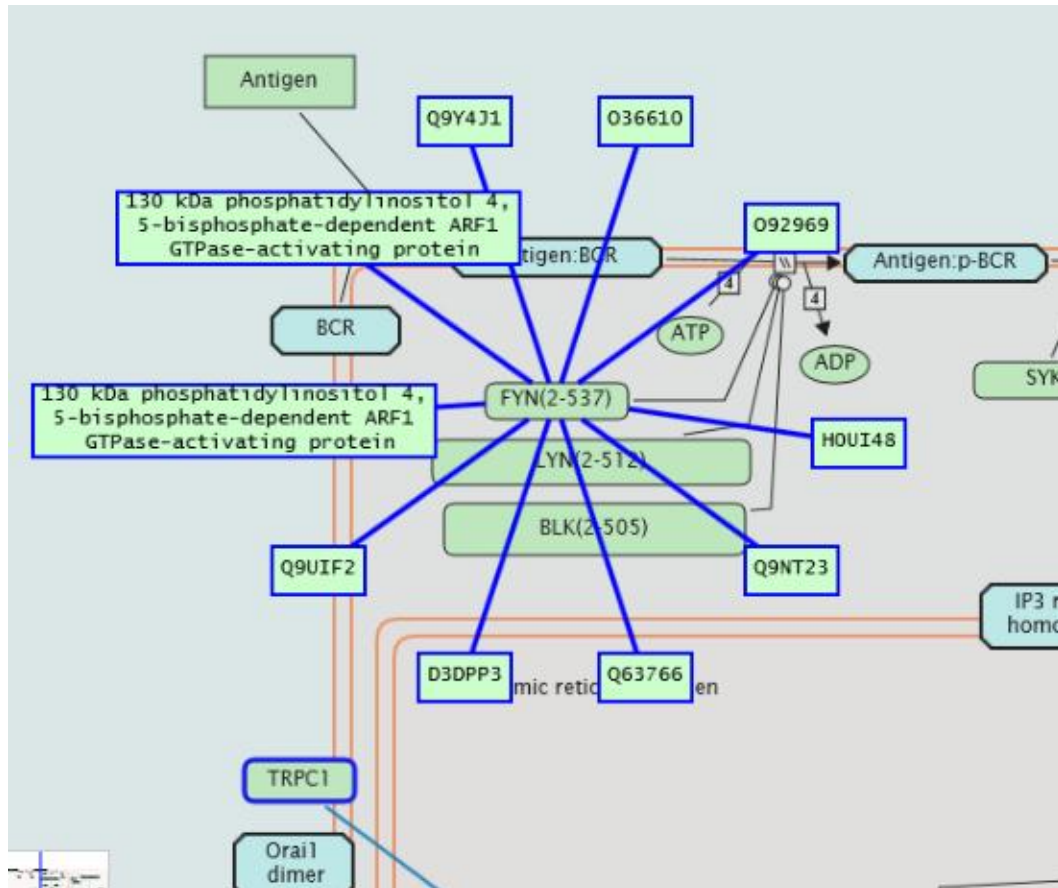
This exercise is to check that you can run and interpret Species Comparison results.

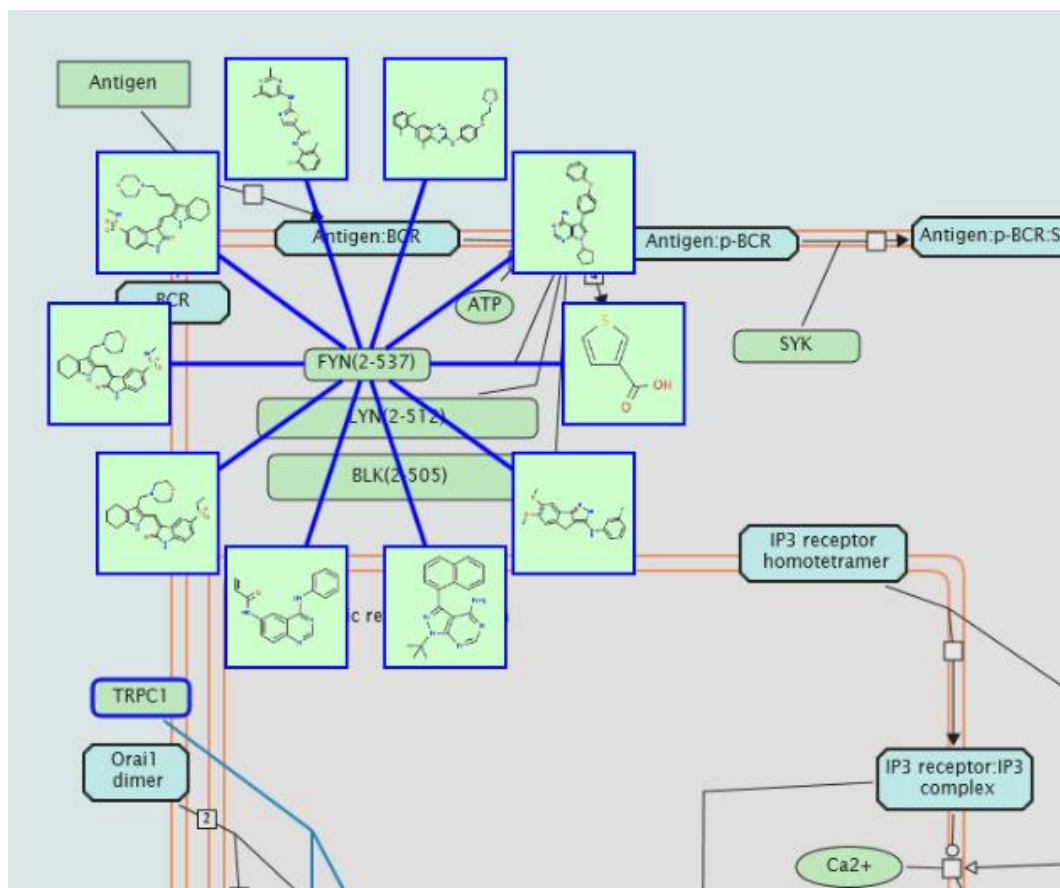
Launch Species Comparison and select the species *Rattus norvegicus*. When the results are displayed, open the pathway 'Immune System', to open the Pathway Browser. Use the Pathway Hierarchy to Navigate to 'Innate Immune System', 'Complement Cascade'.

1. Find Complement factor B (CFB(260-764), top right of diagram) - what colour is it and why?
2. Which other species are inferred to have this protein? *Hint: You can answer this question without rerunning Species comparison.*
3. Find Complement factor 2 (C2a) - why is it blue?
4. Find Cell surface:C4b:c2a How many proteins contribute to this object? Are they all predicted to exist in *Rattus norvegicus*?
5. Why is C2a+ (left side near the bottom) uncoloured?

Molecular Interaction Overlay

The Molecular Interaction (MI) overlay tool is built into the Pathway Browser. It allows protein-protein or protein-compound interactions to be overlaid (superimposed) onto the Pathway Diagram. To add interactors, right click on the protein and select 'Display Interactors'.





Up to 10 interactors are displayed as a ring of boxes connected by blue lines to the selected protein. A white box superimposed onto the selected protein indicates the total number of interactors available up to a value of 50, above this 50+ is displayed (see inset, figure above). Details of up to 100 interactors for every protein in the displayed pathway can be viewed as a table. Details of all available interactors can be exported.

Hover the mouse pointer over a protein interactor to produce a pop-up containing the name and ID of the protein. Hover over a small molecule interactor to displays its name. Click on the line that connects the interactor to the pathway item to view details of the interaction in the source database. If no interactors are available at the selected source a message indicating this will appear.

Interactors can be displayed for any number of proteins. If the same interactor is connected to more than one item the interactor is re-used, i.e. connected to all the selected items in the Pathway Diagram.

While interactors are displayed, right clicking on the selected protein produces two new options, Hide Interactors which removes them, and Export Interactors. Interactions are exported in PSI-MITAB format.

The default source of interactions is IntAct. Source can be configured using the 'Interaction Overlay Options', which shows when you press the cogwheel button (top-right of the diagram). The list of available interaction data sources includes protein-protein and protein-compound databases. In addition, a user-supplied list of interactions can be used.

The Search & Tools Icon

The Search and Tools icon (cogwheel) in the top right corner of the Pathway Browser Panel can be used to configure the Molecular Interaction overlay tool. Click on it and select the menu option Interaction overlay options. This produces a configuration panel:

The screenshot shows a configuration panel titled "Interaction Overlay Options". It contains several sections:

- Interaction Database:** A dropdown menu currently set to "IntAct".
- Upload a file:** A "Select File" button.
- Clear Overlay:** A "Clear" button.
- Submit a new PSICQUIC Service:** A "New Service" button.
- Set confidence level threshold:** A text input field containing "0.5".
- Coloring Mode:** A checkbox labeled "Turn on colouring" which is currently unchecked.
- Visual Indicators:** Two colored squares are shown: a green square labeled "Above Threshold" and a red square labeled "Below Threshold".
- Additional Actions:** Two buttons at the bottom: "Display table of all interactors for pathway" and "Export all interactors for pathway".
- Close:** A "Close" button in the bottom left corner.

- Interaction database - The source of interactions data can be selected using a dropdown list. The default is IntAct.
- Upload a file - use this to submit your own list of interaction data. The format required is PSI-MITAB. The simplest form of this has accession numbers in column 1, names in column 2. If you want to use the colour by confidence score feature (see below) these values must be in column 3. You will be prompted to supply a label for your dataset. If the upload is successful, this label will appear on the Interaction Database dropdown and can be selected as a source of interactions. Submitted datasets persist until the web page is closed.
- Clear overlay - removes all interactors from the Pathway Diagram.
- Submit a New PSICQUIC service - can be used to add a new source interactions database, see the Reactome User Guide for details.

- Set Confidence Level Threshold - allows the user to set a confidence threshold, used for colouring interactors. The default threshold is +/- 0.5 Interactions with associated confidence scores beyond the threshold are coloured according to the colours set for Above and Below, when you press the 'Colour' button.
- The Above and Below colours can be changed by clicking on the coloured squares, which produces a selection menu.
- Hide Table of all interactors for pathway - this button switches off the table of interactors.
- Clicking on the blue squares within this table will add/remove interactors from that protein.
- Clicking on the protein name will cause the Pathway Diagram to centre on the protein represented in the first column interactors.
- Export all interactors for pathway - exports all interactors (no 50 limit) in PSI-MITAB format.

Exercise 10

This exercise is to check that you can use and interpret Molecular Interaction Overlays

Open the Pathway Diagram for Netrin-1 Signaling.

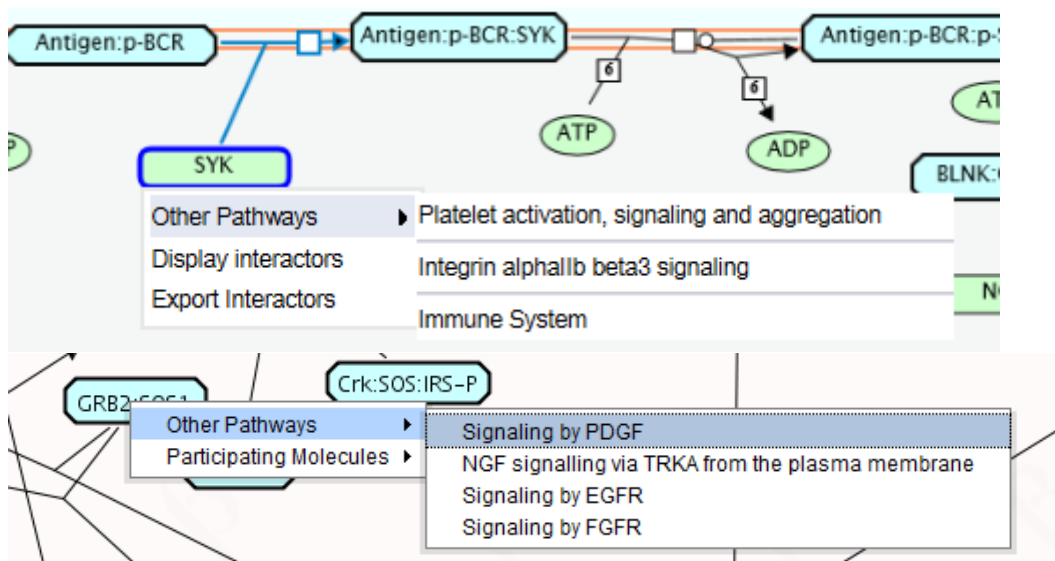
1. Find the protein NEO1 (top left of the cytosol). Right click on it and select Display Interactors. How many are there?
2. How many times has the interaction between NEO1 and TTN (Q8WZ42) been documented? *Hint: This detail is not in Reactome.*
3. Find the protein SRC-1 (to the right of NEO1). Display interactors for this protein. How many are there? Can you get a list of them?
4. What is the easiest way to remove interactors?

Other Features and Tools in the Pathway Browser

Within Pathway Diagrams, right-clicking on the box representing a molecule, complex or reaction presents the user with a menu or list of features relevant to the type of item selected. Note that only features available for the selected object will appear so very few objects have the full range of features.

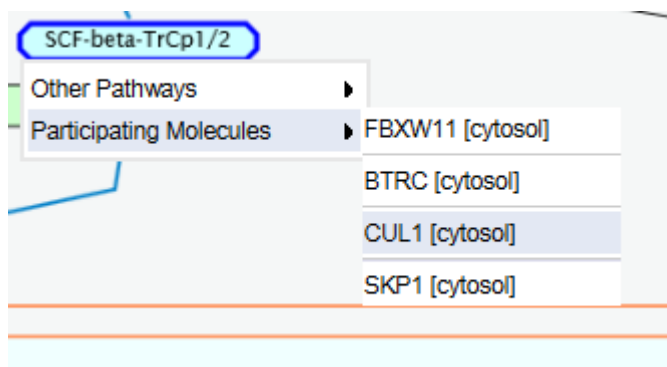
Other Pathways

This feature displays a list of the other pathways that include the selected item as a participant. Clicking on any of the pathway names will display that pathway in the Pathway Diagram Panel (see GRB2:SOS example below).



Participating Molecules

Lists the component molecules for a complex or set. Clicking on a molecule in the list opens a new window with details (example below for Activated Raf 1 complex:MEK2).



Go To Pathway

When a Sub-pathway Box is selected, this displays the associated Pathway Diagram.

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](#).

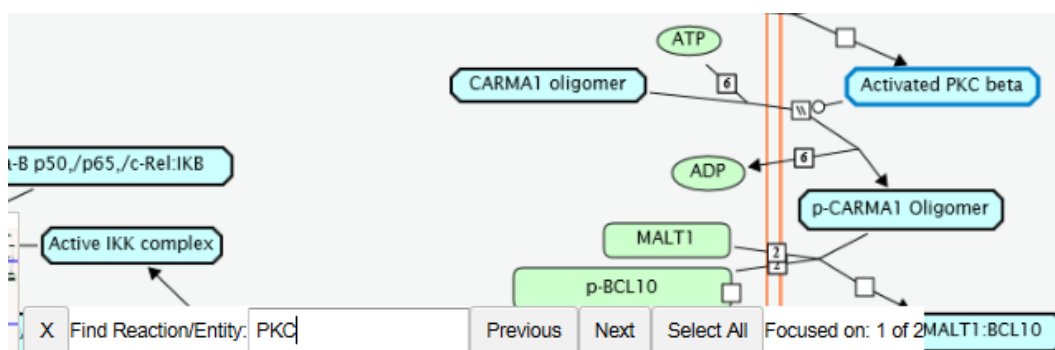
To view the predicted conservation of a pathway in another species, select the species of interest in the "Switch species" dropdown menu in the upper left corner of the window. The Pathway Diagram will re-draw (a revolving arrow icon on the Pathways tab indicates the diagram is re-drawing) in the Pathway Diagram Panel. Any reactions that were not inferred will be absent.

The Cogwheel Button

This button, located in the top right corner of the Pathway Diagram, produces a control panel that has several functions.

Search Diagram

Produces a panel that can be used for searching the displayed pathway diagram:



In the example above the search has identified 2 occurrences of the text PKC in the diagram. It is focused on the first of these, which is highlighted in blue.

Snapshot Current View

Selecting this option copies to the clipboard the currently visible region of the pathway diagram including any overlaid expression or other information.

Download Diagram

Selecting this option opens a new browser tab to display a PNG image of the entire currently displayed diagram. This can be saved to your computer. In Firefox right-click and select Save Image As, etc.

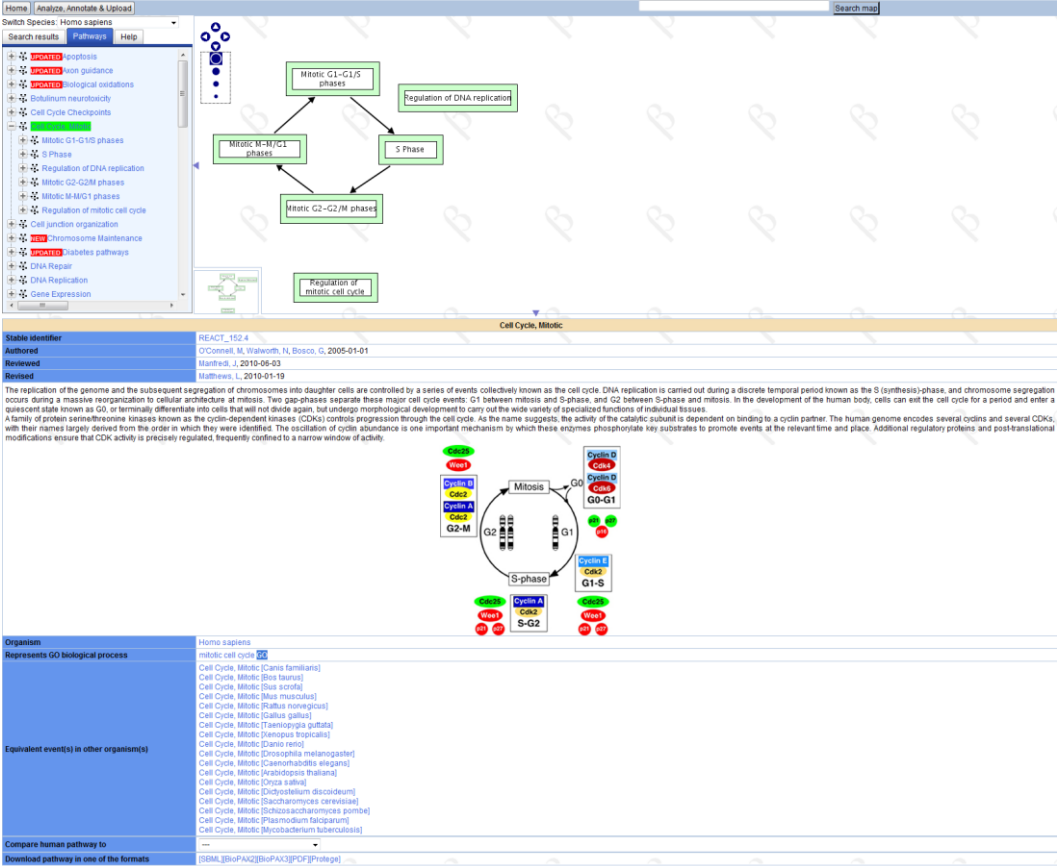
Interaction Overlay Options

This option produces a panel for configuring interaction overlays, as explained above.

Details Views

Pathway Details

Pathway details appear in the Detail Panel at the bottom of the Pathway Browser when a pathway is selected in the Pathways tab:



Stable identifier REACT_152_4

Author O'Connell, M, Walworth, N, Bosco, G, 2005-01-01

Reviewed Martini, J, 2010-06-03

Revised Matthews, L, 2010-01-19

The replication of the genome and the subsequent segregation of chromosomes into daughter cells are controlled by a series of events collectively known as the cell cycle. DNA replication is carried out during a discrete temporal period known as the S (synthesis)-phase, and chromosome segregation occurs during a massive reorganization to cellular architecture of mitosis. Two gap-phases separate these major cell cycle events: G1 between mitosis and S-phase, and G2 between S-phase and mitosis. In the development of the human body, cells can exit the cell cycle for a period and enter a quiescent state known as G0, or terminally differentiate into cells that will not divide again, but undergo morphological development to carry out the wide variety of specialized functions of individual tissues. A family of protein serine/threonine kinases known as the cyclin-dependent kinases (CDKs) controls progression through the cell cycle. As the name suggests, the activity of the catalytic subunit is dependent on binding to a cyclin partner. The human genome encodes several cyclins and several CDKs, with their names largely derived from the order in which they were identified. The oscillation of cyclin abundance is one important mechanism by which these enzymes phosphorylate key substrates to promote events, at the relevant time and place. Additional regulatory proteins and post-translational modifications ensure that CDK activity is precisely regulated, frequently confined to a narrow window of activity.

Organism Homo sapiens

Represents GO biological process mitotic cell cycle [GO](#)

Equivalent event(s) in other organism(s)

- Cell Cycle, Mitotic [Canis familiaris]
- Cell Cycle, Mitotic [Bos taurus]
- Cell Cycle, Mitotic [Sus scrofa]
- Cell Cycle, Mitotic [Mus musculus]
- Cell Cycle, Mitotic [Rattus norvegicus]
- Cell Cycle, Mitotic [Callus gallus]
- Cell Cycle, Mitotic [Taeniopogon guttatus]
- Cell Cycle, Mitotic [Drosophila melanogaster]
- Cell Cycle, Mitotic [Danio rerio]
- Cell Cycle, Mitotic [Drosophila melanogaster]
- Cell Cycle, Mitotic [Caenorhabditis elegans]
- Cell Cycle, Mitotic [Arabidopsis thaliana]
- Cell Cycle, Mitotic [Drosophila melanogaster]
- Cell Cycle, Mitotic [Drosophila melanogaster]
- Cell Cycle, Mitotic [Schistosoma mansoni]
- Cell Cycle, Mitotic [Plasmodium falciparum]
- Cell Cycle, Mitotic [Plasmodium falciparum]
- Cell Cycle, Mitotic [Mycobacterium tuberculosis]

Compare human pathway to ...


















Download pathway in one of the formats [SBML](#) [BioPAX](#) [BioPAX3](#) [Protege](#)

This provides a summary of the pathway and may contain a figure and/or links to literature at PubMed that provide background for the pathway.

- Other details within this section include:
- Author - the expert biologists that contributed materials that allowed this pathway to be created in Reactome.
- Reviewed - the expert biologists that verified the content for this pathway.
- Represents GO Biological process - identifies the GO term for the biological process associated with this pathway.
- Preceding/Following Events - links to pathways that precede/follow this pathway.
- Equivalent event(s) in other organism(s) - identifies whether this pathway has been inferred in other species included in the Reactome inference process.
- Download Pathway - links to download the pathway in several formats allowing reuse with other tools, or as a pdf for reference.

Reaction Details

Reaction details are displayed in the in the Detail Panel at the bottom of the Pathway Browser when a reaction is selected in the Pathways tab or a reaction node is selected on a Pathway Diagram:

SLP-76 stimulates PLC gamma 2	
Stable identifier	REACT_20573_1
Authoried	Aravamudan, J. 2009-06-03
Reviewed	Poole AW, 2009-11-02 Jones, M., 2009-11-02 Hartley, M., 2009-11-02
<p>SLP-76 has a well-established role in recruitment of PLC gamma 1 in immunoreceptor signalling; its role in the recruitment of PLC gamma 2 in integrin signalling is less clear. Results from SLP-76 null mice imply a functional role in GPIIb/IIIa signalling. Platelets from SLP-76 null mice exhibit a major reduction in spreading and a decrease in whole cell phosphotyrosine levels when adhered to a fibrinogen-coated surface. In vivo reconstitution of SLP-76 by retroviral gene transfer corrects bleeding diathesis and restores normal responses to both collagen and fibrinogen (Judd et al., 2000). [PubMed]</p>	
Input (present at start of reaction)	phosphorylated SLP-76 [cytosol]     Phospholipase C gamma 2 [plasma membrane]    
Output (present at end of reaction)	phosphorylated SLP-76 [cytosol]     Phospholipase C gamma 2, phosphorylated [plasma membrane]    
Preceding event(s)	Syk activation leads to SLP-76 activation [Homo sapiens]
Following event(s)	PLC gamma 2-mediated PI3K hydrolysis [Homo sapiens]
Organism	Homo sapiens
Cellular compartment	plasma membrane 
References	<p>Condeelis, J, Huggan, SC, Schwenkfelder, BS, Yap, CL, Yuan, Y, Jackson, SP Integrin alpha IIb beta 3-dependent calcium signals regulate platelet-fibrinogen interactions under flow. Involvement of phospholipase C gamma 2 2003 J Biol Chem PubMed Gross, BS, Marford SK, Watson, SP Evidence that phospholipase C-gamma2 interacts with SLP-76, Syk, Lyn, LAF and the Fc receptor gamma-chain after stimulation of the collagen receptor glycoprotein VI in human platelets 1999 Eur J Biochem PubMed Narain, P, Yuan, Y, Chakraborty, L, Eddy, A, Frensdorf, M, Cazenave, JP, Gachel, C, Jackson, SP, Lanza, F Signaling role for phospholipase C gamma 2 in platelet glycoprotein IIb/IIIa calcium flux and cytoskeletal reorganization. Involvement of a pathway distinct from FcR gamma chain and Fc gamma 2 2003 J Biol Chem PubMed</p>
Equivalent event(s) in other organisms	<p>SLP-76 stimulates PLC gamma 2 [Canis familiaris] SLP-76 stimulates PLC gamma 2 [Sus scrofa] SLP-76 stimulates PLC gamma 2 [Bos taurus] SLP-76 stimulates PLC gamma 2 [Mus musculus] SLP-76 stimulates PLC gamma 2 [Rattus norvegicus] SLP-76 stimulates PLC gamma 2 [Callus gallus] SLP-76 stimulates PLC gamma 2 [Taeniopygia guttata] SLP-76 stimulates PLC gamma 2 [Danio rerio]</p>
Compare human pathway to	---
Download pathway in one of the formats	SBML BioPAX SBML PDF Protege

This provides a summary of the reaction and may contain a figure. It will contain links to literature at PubMed that provide experimental data verifying the reaction, or an 'Inferred From' link

Other details within this section include:

Authoried - the expert biologists who contributed materials allowing the construction of the pathway.

Reviewed - the expert biologists who verified the pathway.

Input/Output - identifies the input/output molecules, sets or complexes for this reaction. Icons to the right of these named items link to further information in external resources, e.g. the red U on a grey background links to UniProt.

Input (present at start of reaction)	phosphorylated SLP-76 [cytosol]     Phospholipase C gamma 2 [plasma membrane]    
--------------------------------------	--

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.

Complex or Set Details

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

Is represented by generalisation(s) - identifies supersets that include this set/complex as a member

Catalyses events - is a catalyst for the events named on the right.

Produced by events - is an output of the reactions named on the right.

Hierarchical view of the components - identifies all the components of a complex in a hierarchical manner, identifying sub-complexes before splitting them into their components.

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 fields represented include:

Name - the name used in Reactome

Links to corresponding entries in external databases - the identifiers used in external databases, act as links to them.

Other identifiers related to this sequence - secondary identifiers from external databases.

Is represented by generalisation(s) - sets that include this protein/small molecule.

Component of - complexes that include this protein/small molecule

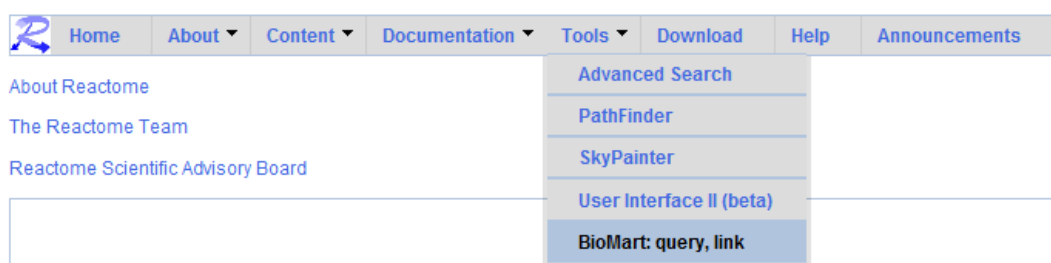
Catalyses events - reactions where this protein acts as a catalyst.

Other forms of this molecule - representation in Reactome of post-translationally modified forms of this protein.

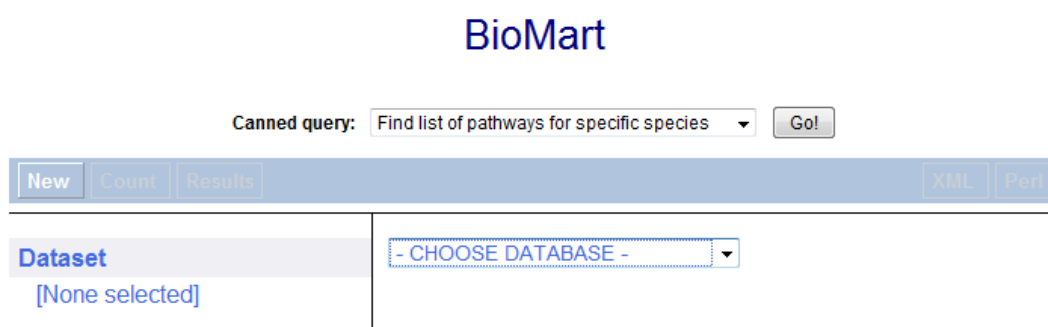
The BioMart interface

BioMart is the name of a project and a corresponding software suite that allows rapid bulk searches of large databases, and the ability to link between databases so that a single query can retrieve data from more than one database. There are several means to access BioMart data including a standardized web interface. Databases available in BioMart include Reactome, UniProt, HGNC, HapMap and ENSEMBL genome databases.

BioMart is accessible via the Navigation Bar. Click on the Tools, select BioMart:



Once the BioMart page has loaded it should look like this:



Click on Choose Database to open the dropdown list, and select Reactome. At this point, you need to decide which Reactome dataset you want to interaction, pathway, reaction, or complex. Set your choice using the Choose Dataset dropdown list:

Now that a Dataset has been selected, the panel on the left-hand side of the page will open out and display Filters and Attributes that you can set to configure your query:

Clicking on “Filters” opens a list of options in the panel on the right-hand side of the page. Filters limit the information that is returned by your query; if you don’t set filters, BioMart will return the entire database. These terms are essentially the query terms, e.g. you can search using a list of Entrez Gene IDs by checking the box for *Limit to Reactions containing these IDs*, and selecting *Entrez Sequence IDs* from the dropdown list. Submit your IDs by pasting into the box or Browsing for a file that contains them.

When you set a filter, its name appears underneath Filters in the left-hand panel.

Filters

Entrez Sequence ID(s) (e.g. 8772): [ID-list specified]

Attributes determine the information that is returned by your query. Some are set to be returned by default; these are displayed under Attributes in the left-hand panel. Click on Attributes to specify additional attributes to be returned in the results of your query.

Once you have set up the Filters and Attributes, click on Results to run the query. Results are presented in a tabular form and by default, only the first 10 results are shown. You can display up to 200 results lines in a single page. In the panel above the results table, you can export the entire results set in a number of different formats:

Export all results to Unique results only

Email notification to

TSV means tab-separated (delimited) values, XLS is the spreadsheet standard used by Excel. Check the Unique results only box to remove duplicates from the results.

Exercise 11

Go to the Reactome BioMart page.

How would you select the “pathway” dataset?

If you were interested in the protein “Nucleolar transcription factor 1” (UniProt ID P17480), how would you identify pathways in Reactome involving that protein? How many pathways does your query find?

How would you find the UniProt IDs of the other proteins in the first of the pathways that you discovered?

Save the results list from 3) in tab-delimited format. How many lines of output are there, excluding column headers?

Answers to Exercises

The answers may change as more data is added to Reactome; these are based on the September 2013 Reactome release.

Exercise 1

1. See Reactome News on the Homepage. At time of writing, the latest item is “Reactome, GMOD, and WormBase at Google”.
2. Found via the Menu bar under Content, Statistics. At time of writing, 7085.
3. Found via the Menu bar under Content, Editorial Calendar. Ver 47 : December 2013
4. At time of writing the first topic is Beta-catenin independent WNT signaling.

Exercise 2

1. The top pathway hit is ‘Signaling by PDGF’. Clicking on the + symbol to the left of the pathway name reveals 9 reactions, and 1 subpathway.
2. The subpathway ‘Downstream signal transduction’ has 16 reactions and 3 subpathways.
3. Usually the next event(s) can be identified by following the connections on the diagram, but not for this reaction. You might be tempted to assume it is the next reaction down in the pathway hierarchy, this is often the case, but not for this reaction. Both of these will only give you a partial answer, because one of the following events is represented in a different Pathway Diagram. The hint suggests opening the Details pane. One of the details shown is Following event(s). This will correctly identify all reactions that follow. The correct answer is ‘PDGF dimer binds two receptors simultaneously’ and ‘PDGF binds to extracellular matrix proteins’.

Exercise 3

1. The reaction node for this reaction is a solid circle, representing a binding reaction.
2. The reaction node symbol for this reaction is an open square, representing a ‘transition’ reaction.

3. The name of the catalyst is the label for the catalyst object, MGLL (this is the HUGO gene name). To find the UniProt ID right-click on MGLL to see its details in the Details pane. The second section of the Details Panel is Links to this entry in other databases, and it contains the UniProt ID, Q99685. If you click on this name it will expand and show more information about this protein, including alternative names (MGLL is also known as monoglyceride lipase). Click on the reaction node to see Input and Output details in the Details Panel, the outputs are AA (the Reactome [small molecule abbreviation](#) used for Arachidonate), Glycerol and H⁺.

Exercise 4

1. Left-click selects the sub-pathway object but does not open the diagram. To do this either a) right-click and select the option Go To Pathway, or b) click on the sub-pathway name, highlighted in the pathway hierarchy on the Pathways tab.
2. 'Intrinsic Pathway for Apoptosis' has sub-pathways, click on the + symbol to reveal them. When you click on the sub-pathway 'BH3-only proteins associate with and inactivate anti-apoptotic BCL-2 members' the 3 reactions in this sub-pathway are selected on the Pathway Diagram, indicated by pale blue borders around the reaction objects and dark blue reaction lines.
3. Clicking on 'Sequestration of tBID by BCL-2' will select just this reaction and centres the Pathway Diagram on this reaction (unless you are zoomed out - if you were try zooming in and drag the diagram so you can't see this reaction, then click on it in the hierarchy again).

Exercise 5

1. 'Activated type I receptor phosphorylates SMAD2/3 directly is part of the pathway 'TGF-beta receptor signaling activates SMADs'.
2. You can often answer this question by referring to the diagram but for a definitive answer look for the Compartment defined in the Details Panel. Click on the reaction in the pathway hierarchy or diagram, look at the Details panel for Cellular compartment - **early endosome membrane**.
3. Open the Catalyst section of the Details Panel. The associated GO molecular function is 'transmembrane receptor protein serine/threonine kinase activity'.
4. The reference is 'Phosphorylation of Smad signaling proteins by receptor serine/threonine kinases' by Souchelnytski et al. 2002.

5. The answer can be found by looking in the Detail Panel at the section 'View computationally predicted event in:' Click here for a dropdown list of species where this event has been computationally inferred. *Canis familiaris* is in the list, but *Saccharomyces* is not, so yes for dog, no for yeast.
6. Look at the Expression tab. None of the genes involved have significant expression in liver, so this event is unlikely to occur in liver.
7. Select the catalyst of this reaction, and then look in the Structures tab. Scroll down and you will find a link to an ePDB structure for TGFB1. Note that this is the 'best' structure available, based on measures set by ePDB. Other structures are also available via the link.

Exercise 6

1. The diseases include dwarfism syndromes (achondroplasia, hypochondroplasia and the neonatal lethal disorders thanatophoric dysplasia I and II) and craniosynostosis disorders such as Pfeiffer, Apert, Crouzon, Jackson-Weiss and Muenke syndromes.
2. The defective enzyme is phenylalanine hydroxylase (PAH). The mutation is a residue substitution of S to L at position 40 (S40L).
3. The 3 substrates that can build up are heparan sulphate chain(1), IdoA-GalNAc(4S)-GlcA-Gal-Gal-Xyl and heparin sulphate chain(6).
4. Expand the Functional status panel to see the details. The functional status of the enzyme is 'loss of function via point mutation'.
5. Still in the Functional status panel, expand the panel for the physical entity, MMADHC mutants. Seven different mutations are listed.

Exercise 7

1. DNA replication has the greatest coverage
2. 57, 56%.
3. Cell cycle is most enriched, with a probability score of 6.90E-10.

Exercise 8

1. 126 proteins in the pathway. You need to click on the pathway name and then select the Molecules tab in the details Panel.

2. 74 have expression values in the dataset.
3. Metabolism of proteins (10.0)
4. Nucleotide Excision Repair
5. DNA Damage Bypass
6. Nucleotide Excision Repair
7. It's easier to see this if you zoom out so all the pathway is visible. I think it's RAD23B, top left of the diagram.
8. POLR2H
9. 209302_at.

Exercise 9

1. Complement factor B is yellow, indicating that it has been inferred to exist in rat.
2. You can answer this by clicking on the box for Complement factor B and looking at its details, look for the 'Homologues of...' sections and open them, to find out which other species that have been inferred to have Complement factor B.
3. Complement factor 2 is blue, indicating that it has not been inferred to exist in rats.
4. Cell surface:C4b:c2a is a set or complex. Right-click on it and select the option 'View participating molecules' to see which are predicted to exist in rat. There are 3 components, two yellow and predicted to exist in rat and one blue and not predicted to exist in rat.
5. Calcium is not coloured because it is a small molecule, species comparison is not relevant. Other pathway objects that have no UniProt ID will also be uncoloured.

Exercise 10

1. NEO1 has two interactors (number may change as IntAct is updated).
2. Click on the line between NEO1 and TTN (may be labelled Q8WZ42) to open a page describing the interaction at the source database, IntAct. This interaction has been identified only once using a variation of the two-hybrid method.
3. SRC has more than 50 interactors, indicated by 50+ in the interaction count box (you need to be zoomed in to see this). You can see up to

100 in the table within the Interaction Overlay Options panel; if you export the interactions you will get all interactions available from the selected source, in this case IntAct.

4. There is a Clear button in the Interaction Overlay Options panel.

Exercise 11

Click on Choose Database and select Reactome, then click on Choose Dataset and select Pathway.

Add a Filter, limiting to pathways containing a list of IDs, select UniProt IDs from the dropdown list, and enter P17480 in the box. The query returns 9 pathways.

Copy the Pathway Stable ID, REACT_2232, change the filter so it limits by Pathway Stable ID, and enter REACT_2232 in the box. Add the Attribute Protein UniProt ID. You should have 2 results returned.

You will be asked to save your chosen download format to your computer. This produces 2 lines of output.

Printable Key to Reactome Pathway Diagrams

Key to Reactome Pathway Diagrams

