Reactome: Quick tour

Steve Jupe [1]
Systems
Beginner
0.5 hour

This quick tour provides a brief introduction to Reactome, the biological pathway database.

Learning objectives:

- Understanding of Reactome and how it can help you to find out more about your biological pathways of interest
- Know where to find out more about Reactome

What is Reactome?

Why do we need Reactome [2]?

Life at the molecular level is an intricate network of protein interactions and biochemical reactions. A pathway is a small part of this network – a series of connected events, typically characterised by the product or biological outcome. Researchers have defined many biological pathways. These pathways can be accessed via the scientific literature, textbooks, or electronic databases, but the information is of variable quality and is largely unavailable for computational re-use. Reactome provides access to a curated, quality-assured slice of this information, available as diagrams and descriptions on the Reactome website and as downloadable pathways in standard formats such as BioPAX [3] and SBML.

What does Reactome contain?

Reactome [4] aims to represent all human biological processes as interconnected molecular events or 'reactions'. Reactome broadly defines the term 'reaction' to include any molecular event in biology, such as binding, phosphorylation, catalysed biochemical events, spontaneous molecular events and transport. Every reaction stored in Reactome defines the precise form of the molecules that participate in the reaction and the cellular compartment in which the reaction takes place. Reactions can be considered as 'steps' in a pathway. Pathways are represented as diagrams containing a series of interconnected events, often organised into sub-pathways.

Quality control and attribution are key features of the data in Reactome. Pathways are authored by expert biologists and independently peer reviewed before inclusion. Reactions have literature citations that experimentally verify the described event using human reagents, or if these are not available, reagents from model organisms (as long as the author and reviewer agree that it is valid to infer that the equivalent event can take place in humans).

Examples of pathways available in Reactome include immune system signalling, the cell cycle, apoptosis, and the host-parasite interactions of HIV. These manually curated human pathways are used as templates to infer equivalent pathways in 20 other species, including commonly-used model organisms. The EGFR signalling pathway is shown below in Figure 1.

Reactome uses primary external sources and extensively cross-references many other relevant sources of biological, chemical and literature information. Primary external references include: UniProt [5] for proteins; ChEBI [6] for small molecules [7]; NCBI Taxonomy [8] IDs for species; Gene
Ontology [9] (GO) terms for catalytic activity, cellular compartment and biological process.

Reactome data and software are open source, freely available to download and reuse from the Reactome website [4], under a Creative Commons Attribution 3.0 Unported licence [10].

Figure 1  EGFR signalling pathway. EGFR is a tyrosine kinase receptor that plays a key role in the regulation of cell division and cell death.

What can I do with Reactome?

View your expression data in a pathway context

Reactome [2] can help you understand the biological significance of expression data. The 'Analyze Expression Data' tool overlays your expression data onto Reactome pathway diagrams, using colour to represent the expression level (Figure 2).
Figure 2 Using colour-coding to represent experimental expression values on a Reactome pathway diagram.

**Identify pathways represented by a list of gene, protein or compound identifiers**

The 'Analyze Data' tool takes your list of gene, protein or compound identifiers and maps them onto Reactome [2] pathways (Figure 3). It can determine whether any pathways are significantly over-represented (enriched) in your set, answering the question 'does my list contain more members of certain pathways than would be expected by chance?'.
The search tool (top right of Reactome [2] web pages) allows you to search for proteins, reactions and pathways using familiar names. In addition, it recognises many identifiers from external sources, including UniProt [5] IDs, Entrez Gene [11] IDs, and HUGO [12] gene symbols. The results are sorted into categories, Proteins, Reactions, Pathways etc. Results can be filtered using check boxes on the left. Select a hit to view it in the Pathway Browser.
Figure 5 Results from a text search for 'CCR5'. Note the check boxes on the left. These are used to filter results.

Getting data from Reactome

- All Reactome [2] pathways can be downloaded from the Reactome website in several formats. For reference use, Rich Text Format (Word compatible) and PDF formats are available. For computational re-use, SBML level 2.3 and BioPAX [3] level 2 and 3 formats are available.
- Individual pathways can be downloaded using links available at the bottom of the pathway browser (shown in Figure 6).
- Pathway diagrams are available in PDF and PNG formats.
- The Reactome MySQL [14] database, schema, code and web interface can be downloaded for local implementation and development.
Contributing to Reactome

The goal of Reactome [2] is to represent the molecular details of all human pathways. As of Reactome version 54 (Sept 2015), our coverage of human proteins is 8464 – roughly half of all proteins with an identifiable function. Coverage improves continuously; see the Editorial Calendar [15] for details of planned content for the next quarterly release. All our content is sourced from and reviewed by expert biologists who generously volunteer their time and effort. We welcome contributions of any size from the biological community. Significant additions can be assigned Digital Object Identifiers and therefore qualify as online publications. If you are interested please contact help@reactome.org [16] for an informal discussion.

Get help and support on Reactome

Support and find out more

- New users who want to learn more – see the Reactome user guide [17]
- Contact our support team [16]

References


Funding

The development of Reactome is supported by a grant from the US National Institutes of Health [20] (P41 HG003751)

Reactome is a collaboration between groups at the EMBL- European Bioinformatics Institute [21] (EBI), Ontario Institute for Cancer Research [22], New York University Medical Center [23] and Cold Spring Harbor Laboratory [24].

Contributors
Steve Jupe is the Biology Coordinator for the Reactome project in the Proteomics Services team [25] at the European Bioinformatics Institute. His role is primarily the identification and addition of new topics to Reactome, particularly in the areas of immune signalling and haemostasis. Steve has a BSc in Biology and for his PhD he researched the growth-regulating activity of the plant hormones gibberellin and ethylene at UCW Aberystwyth, obtaining his PhD in 1989. This was followed by postdoctoral positions at the Institute for Arable Crops Research, Rothamsted and Long Ashton, where he trained in molecular biology as part of the AFRC plant molecular biology initiative. He worked on plant-pathogen interactions at Royal Holloway, University of London before moving away from the lab in 1996 to join the Bioinformatics group at SmithKline Beecham, now GlaxoSmithKline, where he worked on numerous drug discovery projects and maintained a database of GPCRs. Steve joined the Reactome group in 2009.
Reactome: Quick tour

[16] mailto:help@reactome.org
[21] http://www.ebi.ac.uk
[22] http://oicr.on.ca/
[23] http://www.med.nyu.edu/