Protein Networks and Systems Biology

ELIXIR 2015, Bologna, Italy

**Protein-Protein Interactions and**

**Bioinformatics tools**

Luana Licata

Bioinformatics and Computational Biology Unit

Molecular Genetics Laboratory

University of Rome Tor Vergata, Italy

**Introduction**

**Protein-Protein Interactions and Molecular interactions databases**

A good knowledge of the biological systems complexity necessitates a complete understanding of protein-protein interactions and their molecular machinery. Molecular interaction data can be generated by many different techniques and there is a huge amount of data and large-scale dataset of molecular interactions constantly published.

Therefore, numerous databases have been established in order to archive, disseminate and manipulate molecular interaction data. Main purpose of these repositories is to capture molecular interaction data, to organize it in a structured format, which allows users to perform searches and bioinformatics analysis. Therefore, molecular interaction databases are a key means for the interpretation of cellular complexity.

**IntAct Database**

## IntAct is a publicly available repository of molecular interactions ([http://www.ebi.ac.uk/intact](http://www.ebi.ac.uk/intact" \t "pmc_ext)) containing molecular interaction data coming either from manually curated literature or from direct data depositions. The IntAct curation interface has been developed as a web-based platform in order to allow external curation teams to annotate data directly into the IntAct database. All data can be search, viewing and download. Data can be accessed through the IntAct ftp site in PSI-MI XML and MITAB 2.5, 2.6 and 2.7 formats. Alternatively, the complete dataset can be downloaded directly from the website in RDF and XGMML formats. Data can also be accessible through PSICQUIC web service IMEx website. IntAct is an active member of IMEx consortium adopting a common curation rules and data format.

**Training session: Searching IntAct database**

In this training session you will use IntAct database. You will perform different search, check the number of interaction evidences supporting each search, look for the experimental details of a specific interaction, and download data using different formats.

**Step1*:*** *Quick Search*

In the quick search panel, it is possible to type protein name, gene name, Accession Numbers, GO term, publication ID, authors name, experimental detection method, etc.

In these exercises you will perform a very simple search and look at your results in the IntAct viewer. You may perform the exercise below or alternatively, you may try a protein you are interested in.

Open the IntAct homepage in a web browser (http://www.ebi.ac.uk/intact).

Type in the Quick Search ‘**MEKK1**’ and click Search

**Q1:** How many binary interactions do you find?

Searching on a non-specific gene name will bring you up a mixed set of results. Refine your search using either **Q13233** or M3K1\_HUMAN.

**Q2:** How many binary interactions do you find by refining the search?

Go to the detailed view of the interaction of MEKK1 (**Q13233**) with BRAF (P15056) that it has as “interaction detection Method” : Protein Kinase assay

**Q3:** Which of these proteins acts as the enzyme in this reaction?

**Q4:** Which is the stoichiometry?

***Step2*:** Refining your search using the Advanced Search

Using the Molecular Interaction Query Language, available from the quick search panel, it is possible to perform more complex queries. This will allow you to write more complex queries, for example:

To discover if any interactions are known between Plasmodium falciparum (isolate 3D7) and human proteins, Clear any previous searches, then type ‘taxidA:36329 AND taxidB:9606’.

However, ‘Show Advanced Fields’ option is an easier way to perform a complex search in IntAct. Clicking on this to the right of the Quick Search box on the IntAct homepage will open up the Advanced Search, allowing you to specify one or more fields you wish to search in, and building the query for you as you progress.

Clear the previous search, then use Advanced Fields to search for ‘organism’ = human AND ‘Interaction detection method’ = surface plasmon resonance

**Q5:** How many interactions do see for ‘organism’ = human AND ‘Interaction detection method’ = two hybrid?

**Q6:** How many interactions do you see for P53\_HUMAN AND ‘Interaction type’ = physical

**Q7:** How many interactions do you see for ABL1-HUMAN AND interactor feature = direct

Go to the detailed view of the interaction of ABL1-HUMAN (P00519) with P41 (EBI-7094147) that it has as “interaction detection Method”: isothermal titration calorimetry

**Q8:** Which are the Interaction Parameters associated to this interaction?

Go back to the interaction list of the previous search and Go to the detailed view of the interaction of ABL1-HUMAN (P00519) with P41 (EBI-7094147) that it has as “interaction detection Method” : X-ray crystallography.

### Q9: Which are the Experimental Feature of this interaction for both the interactors?  Which binding site and which mutation?

**Step3:** Download data

Search for all the human PPIs present in IntAct using the quick or advanced search.

**Q10:** Download the interactions found as a PSI-MI xml and MITAB 2.7.

Go to the search page and type the PMID: 22014111.

**Q11:** How many interactions did you find? Download the all the PPIs associated to this PMID as XGMML format?

Go to the download page and download as PSI2.5 format all the PPIs of the dataset “Virus”.

**Q12:** How many papers belong to this dataset? How many interactions?