

Molecular Interaction Networks

Service providers at the BioHackathon:

- **DIP** (Lukasz Salwinski, UCLA)
- **STRING/STICH** (Michael Kuhn, EMBL)
- **IntAct** (Bruno Aranda, EBI)

The **IMEx consortium** is a group of major public interaction data providers sharing curation effort and exchanging completed records on molecular interaction data. DIP, IntAct and MINT are full-fledged members. Others coming soon.

Service consumers:

- **Cytoscape** (Keichiro Ono)
- **CellDesigner** (Akira Funahashi, Keio Univ. & SBI)

Exchange formats

Exchange formats is not an issue. Standards already exist.

- **PSI-MI XML 2.5** (aka MIF - XML based) – Complete information.
- **MITAB** (Tabulated data) – Contains reduced information but still useful to some user profiles

| | Accession number molecule A | Accession number molecule B | Alternative id molecule A | Alternative id molecule B | Names molecule A | Names molecule B | Species molecule A | Species molecule B | First Author | PubMed Identifier | Interaction type | Interaction detection method | Source database(s) | Interaction details |
|---|---|--|---------------------------|---------------------------|------------------|------------------|------------------------------|------------------------------|--|--|--|--|--|--|
| 1 | P60896 , EBI-79819 | P51507 , EBI-79792 | DSS1, SHFDG1 | FANCD1, FACD | SHFM1 | BRCA2 | 9605 (human) | 9605 (human) | Marston et al. (1998) Marston et al. (1998) Marston et al. (1999) Marston et al. (1999) | 10373512 10373512 10373512 10373512 | physical interaction physical interaction physical interaction physical interaction | 2 hybrid 2 hybrid coip coip | intact intact intact intact | EBI-79835 EBI-79842 EBI-79852 EBI-79847 |
| 2 | P51507 , EBI-79792 | P15927 , EBI-621404 | FANCD1, FACD | REPA2, RPA32 | BRCA2 | RPA2 | 9605 (human) | 9605 (human) | Wong et al. (2003) Wong et al. (2003) Wong et al. (2003) | 12527904 12527904 12527904 | physical interaction physical interaction physical interaction | anti bait coip pull down pull down | intact intact intact | EBI-1372805 EBI-1372823 EBI-1372896 |
| 3 | P51507 , EBI-79792 | P27694 , EBI-621369 | FANCD1, FACD | REPA1, RPA70 | BRCA2 | RPA1 | 9605 (human) | 9605 (human) | Wong et al. (2003) Wong et al. (2003) Wong et al. (2003) | 12527904 12527904 12527904 | physical interaction physical interaction physical interaction | anti bait coip pull down pull down | intact intact intact | EBI-1372805 EBI-1372823 EBI-1372896 |

Web Services

Some providers have web services already (e.g. IntAct, STRING/STICH, others not), and they are **not** compatible.

Molecular Interaction Networks

Standard WSDL -> PSICQUIC

- Define a **standard WSDL** for a common Web Service interface -> project **PSICQUIC**
 - Done: the WSDL has been created, containing a comprehensive set of methods that return data in the two available formats
 - Issue: returned XML can be big (up to hundreds of Mb). Results are paginated so each invocation returns a limited amount of Interaction Data.
 - Cytoscape can already understand the output of the service, as it works with the standard formats.
 -
 - So, we could have a common client to rule them all!
 -

Molecular Interaction Networks

PSICQUIC WSDL Methods:

- getByInteractor (dbRef)
- getByInteraction (dbRef)
- getByInteractorList (dbRefList)
- getByInteractionList (dbRefList)
- getByBetweenList (dbRefList)
- getByQuery (query)

- getVersion
- getSupportedDataTypes

All queries can be paginated, and exceptions have been defined too.

Molecular Interaction Networks

Common Query Language

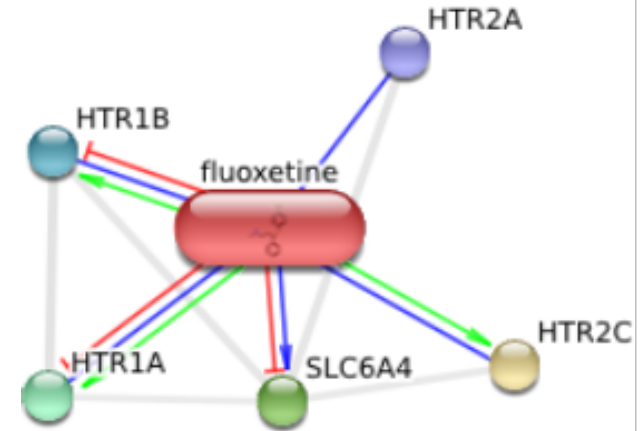
- Define a standard Common Query Language for interaction data retrieval.
- IntAct has created a generic Lucene based query language, which is currently used by Cytoscape (and others). It searches on a PSI MITAB file, so it could be used by anybody (and of course, open-source!)
- No further work has been done on this within the providers.
- Example: “id:P12345 species:human”

Molecular Interaction Networks

Use-case gathering from clients

- Many issues automatically resolved by using the standard formats.
- Problem (use-case from Cytoscape): cytoscape uses identifiers that the providers not necessarily contain (e.g. NCBI identifiers). The IMEx consortium defines Uniprotkb and Refseq as the recommended identifiers. Mapping would be needed.

STRING/STITCH



- REST API

- get interaction network in PSI-MI 2.5 format: can be used in Cytoscape
- resolve items; get abstracts

- SOAP API

1. with Soaplab2: partially done, works in Taverna
2. PSICQUIC WSDL: to be implemented

あしからず