ConsensusPathDB

Backgroud

- **Big trend**:
- Functional interactions between cellular entities like genes, proteins, metabolites, etc. are the key drivers of cellular functions.
- past few years, the analysis of interaction networks has become crucial to understand biological processes .
- **Practical application**:
- Analyses combining expression and interaction data have recently been used to reveal previously unknown disease mechanisms and understand their dysfunctions in human diseases.
- Thus, collecting comprehensive human interaction data is the key to treat human disease and gain new insights into cell biology.

Problems

model organisms：

- While for several model organisms like *Saccharomyces cerevisiae* and *Caenorhabditis elegans*, such comprehensive functional interaction networks are available.

Human：

- larger part of the human interactome still remains undiscovered .
- the existing knowledge on human functional interactions is dispersed in over 200 interaction databases, each of which has a specific data format, focus and bias .（too much data，redundancies）
- Most integration efforts with respect to interaction data so far have focused on merging homogeneous interaction networks. For example, APID , MiMI and UniHI integrate protein–protein interaction networks from multiple sources.（ Incomplete）

Expect

- **We expect**:
- Find the complete information without redundancies for a certain physical entity or pathway in just one database.
- **In other words**:
- Find methods to build a database to integrated datas in a complementary manner and redundancies are avoided.

Citation

- Atanas Kamburov
- Christoph Wierling
- Hans Lehrach
- Ralf Herwig

Method： Mapping of functional interactions

- Feathers： straightforward and efficient for the integration of networks from any single species.
- Simple physical entities with the same identifier： first translated the annotations to a uniform identifier type, which is a UniProt entry name in case of proteins, Ensembl gene ID in case of genes and transcripts, and KEGG/ChEBI ID in case of metabolites.
- complexes with the same composition： compared according to their individual protein composition.
- Functional interactions of physical entities： flexible distinguish between primary and secondary interaction participants.
- Biological pathways :sets of interactions, whose compositions are adopted from the source databases.(eg KEGG's Glycolysis/gluconeogenesis pathway contains 31 reactions whereas Reactome's Glycolysis contains 10 reactions.)

Introduction

- **Object of study** :human, mouse and yeast
- **Types of functional interactions** :metabolic and signaling reactions, physical protein interactions and gene regulatory interactions.
- **The methods:**Physical entities from the external resources are mapped to each other on the basis of common identifiers like UniProt and Entrez. Interactions with matching primary participants are also mapped and grouped together according to similarity.
- **The manners:**in a complementary manner and redundancies are avoided.
- **The mission:assemble** a more complete and a less biased picture of cellular biology.
- **States** :ConsensusPathDB does not provide any additional quality control filters. All interactions provided by the different database sources are treated in the same way.（Without original）

Content

- **Human**(18 public resources)
- Release 16 (09.09.2010)
- unique physical entities: 41,271
- unique functional interactions: 155,432
- pathways: 2,205

- **yeast** (8 public resources)
- Release SC5 (21.06.2010)
- unique physical entities: 14,532
- unique functional interactions: 194,480
- pathways: 734

- **mouse** (8 public resources)
- Release MM5 (21.06.2010)
- unique physical entities: 21,946
- unique functional interactions: 13,648
- pathways: 1,381

- The database content is updated every three months with the newest available versions of the source databases.

The ConsensusPathDB database is available at：http://cpdb.molgen.mpg.de/

Thanks for your attention!