QTLNetMiner
Linking Crop Traits to Genes through Data Integration and Text Mining

Keywan Hassani-Pak
Computational and Systems Biology Department

Ondex – Data Integration and Network Generation

- Supports many use cases in Life Sciences and Systems Biology where similar requirement exist
- Integrates public databases with custom multi-omics datasets
- Provides a GUI for reproducible data integration workflows
- Free and open source

Ondex
www.ondex.org

Building knowledge networks

- Wheat Genes
- Homology/Domains
- Annotations

Ondex Text-Mining Plugin

Input data
- 27,416 Arabidopsis gene names from Phytozome
- 52,561 Abstracts from PubMed that contain Arabidopsis
- 22,201 curated citations from TAIR
- 1,349 Trait Ontology terms from Planteome

Text-mining output

These steps connect 5553 Arabidopsis genes to 409 TO terms based on 18,341 co-citations
Genome-scale knowledge networks

_efficient search and knowledge extraction_

1. Pre-define biologically plausible paths (semantic motifs) starting with a Gene node based on a metagraph schema
2. Query every Gene node in the graph with the union of these semantic motifs (e.g. 57 motifs in the wheat network)
3. Annotate the extracted graph based on user provided input data (i.e. search terms, gene list, QTL)

**QTLNetMiner – Gene Discovery Made Easy**

- Web application for very fast search of large genome-scale knowledge networks
- Ranking of candidate genes based on relevance to trait terms
- User-friendly visualisation of gene-evidence networks
- Facilitates hypothesis validation and generation

**QTLNetMiner**  
[https://ondex.rothamsted.ac.uk/QTLNetMiner](https://ondex.rothamsted.ac.uk/QTLNetMiner)

**QTLNetMiner – Data Mining and Gene Ranking**

- Search knowledge network using trait-based keywords
- Real-time user feedback and query suggestions
Visualisation of networks controlling complex traits

Systems Biology approaches are essential to understand complex traits

QTLNetMiner enables discovery and exploration of genes and biological processes controlling complex traits

Ontologies are used in cross-species annotation transfer, query term suggestion and text-mining methods

Need comprehensive list of trait synonyms to improve our text-mining methods

Conclusions

Acknowledgements

We welcome collaborations!