BioRange SP 3.2.2.: Translational Medicine through Comparative Genomics and data integration

Wilco Fleuren[1], Raoul Frijters[1], Tim Hulsen[1], Peter Groenen[2], Jacob de Vlieg[1,2], Wynand Alkema[2]

[1]CDD, Centre for Molecular and Biomolecular Informatics (CMBI), Radboud University Nijmegen, Nijmegen.
[2]Molecular Design and Informatics (MDI), Organon, part of Schering-Plough Corporation, Oss.

Summary

One of the problems in drug development is the translation of data obtained in animal models to effects in clinical studies. In order to choose the most appropriate model organism for testing new drugs, knowledge of the critical differences between model organisms and human is essential. One approach for detecting these differences is a conservation analysis of the biology related to targets and drugs of interest. We have developed a method which integrates and visualizes gene conservation, literature and expression data. We applied the method to microarray data of methapyrilene, a known carcinogen in rats, but not in mice and humans. We identified Cyp2c12, a gene that is involved in methapyrilene metabolization and does not have orthologs in mouse or human. This might be an indication for alternative metabolite formation in rats and could explain the toxicity of methapyrilene in rats.

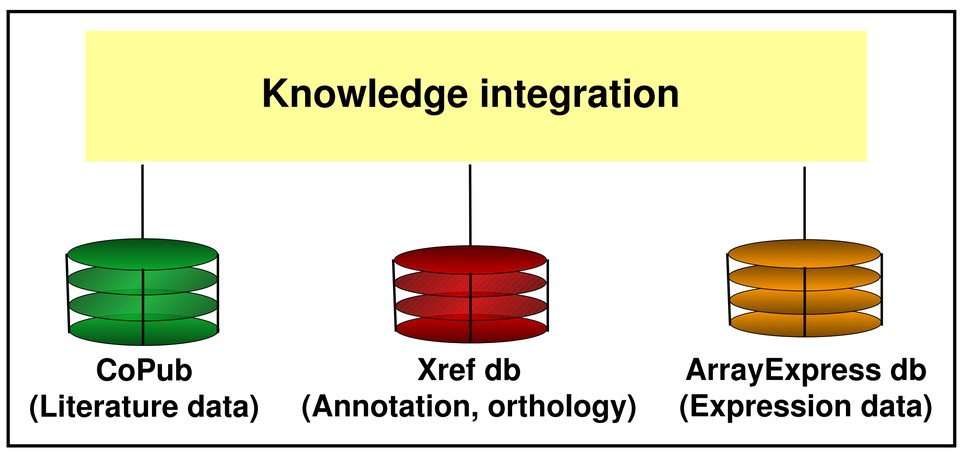


Figure 1 | Integration of literature data from CoPub, annotation and orthology data from the Xref db and expression data from the ArrayExpress db.

Materials & Methods

We integrated the data from three databases (**figure 1**): CoPub[1] (literature data), Xref db (gene annotation and orthology information) and ArrayExpress[2] (publicly available microarray data). Literature mining was performed with co-publication information from Medline abstracts using CoPub[1]. Publicly available microarray data from rats treated with methapyrilene were downloaded from the ArrayExpress[2] database for literature-based keyword enrichment analysis[3].

[1] http://services.nbic.nl/cgi-bin/copub/CoPub.pl
[2] http://www.ebi.ac.uk/microarray-as/aer/
[3] Frijters et al. Pharmacogenomics, Nov. 2007.

Results

Methapyrilene induced genes were analyzed for keyword enrichment and together with gene orthology mapping, visualized in a scalable vector graphics (SVG) format (**figure 2**). In the generated network, various areas representing distinct biological processes can be distinguished, e.g. oxidative stress and drug metabolism, which gives an overview of methapyrilene's mechanisms of action[3]. Zooming in on the network, we identified a gene, Cyp2c12, which lacks orthologs in mice and humans and is involved in methapyrilene metabolization. Therefore, Cyp2c12 is a candidate gene for explaining the interspecies differences in methapyrilene toxicity.

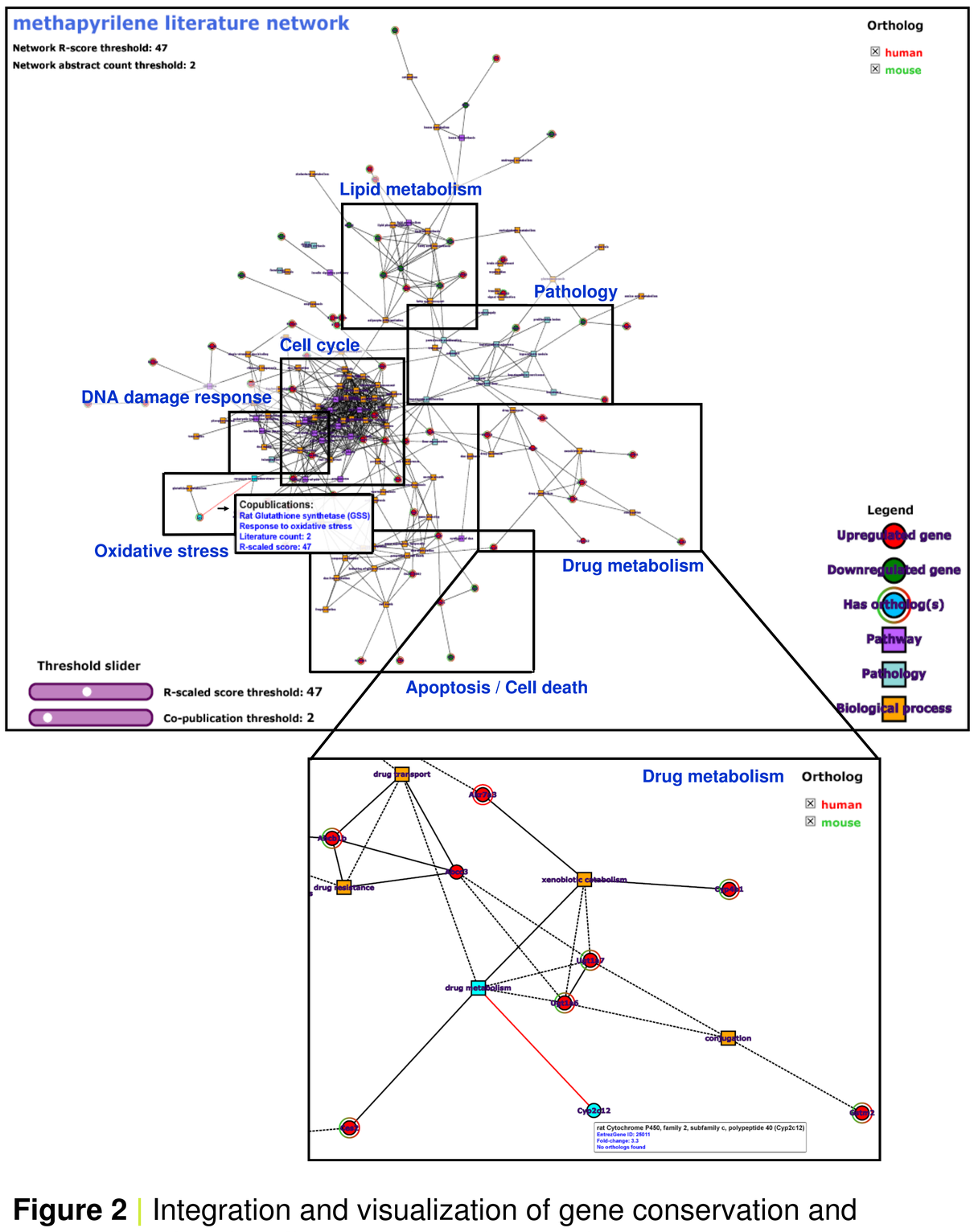
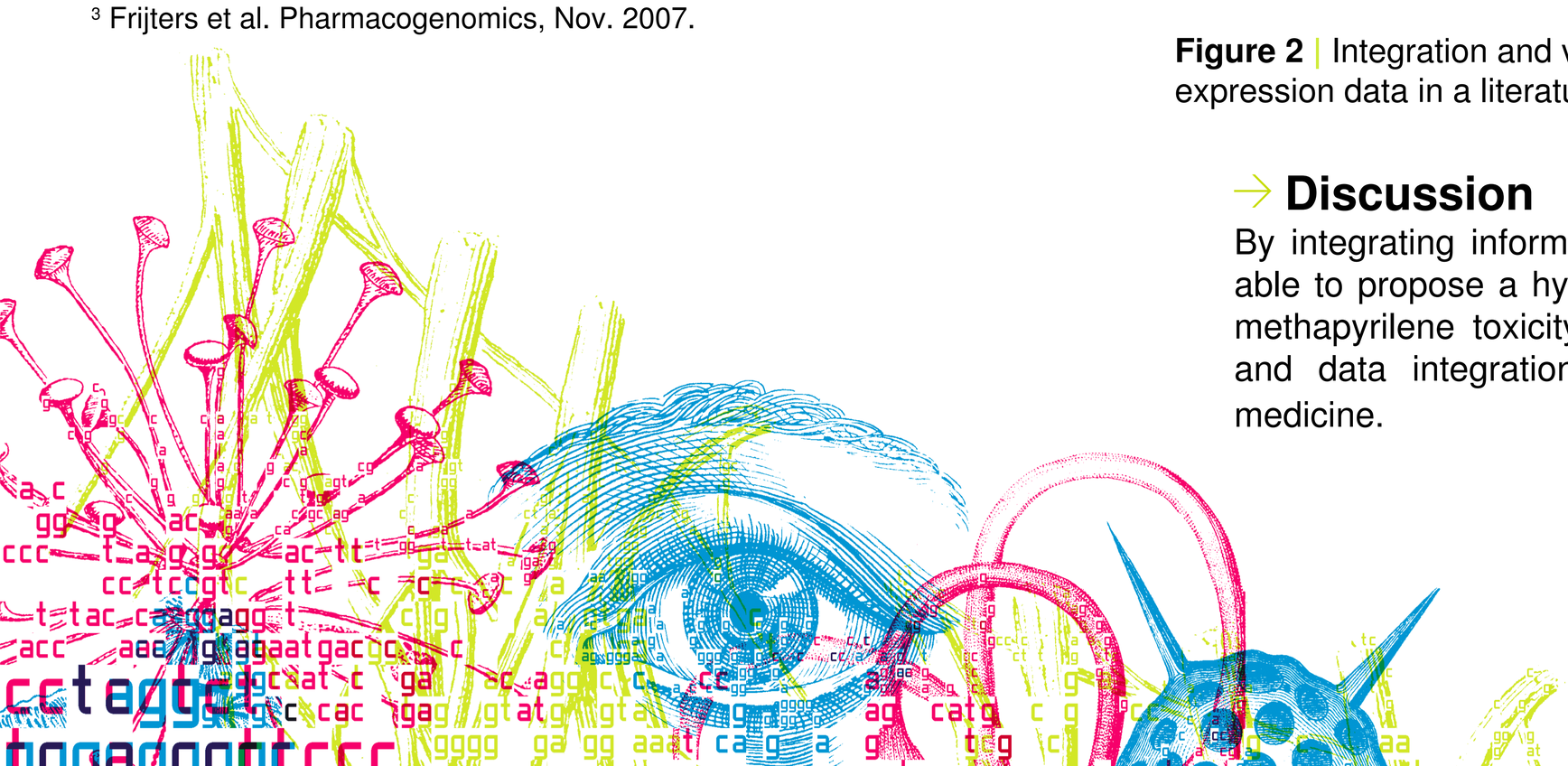


Figure 2 | Integration and visualization of gene conservation and expression data in a literature-based network.

Discussion

By integrating information from various data sources, we were able to propose a hypothesis for the interspecies differences in methapyrilene toxicity. This shows that comparative genomics and data integration is a useful approach for translational medicine.



Netherlands Bioinformatics Centre