DECO: Data-integration for Endpoints, Chemoinformatics and Omics

Joost van Delft[1], Jan Polman[1], Danyel Jennen[1], Gina Montoya-Parra[2], Hennicke Kamp [2], Eugene van Someren [3], André Boorsma [3], Rob Stierum [3], Dinant Kroese [3]

1 Department of Toxicogenomics, Maastricht University, The Netherlands
2 Experimental Toxicology and Ecology, BASF SE, Ludwigshafen, Germany
3 Research Group Microbiology and Systems Biology, TNO, Zeist, The Netherlands

Background

Chemical analogues identified by chemoinformatic tools are presumed to induce similar biological responses, such as similar toxicities. Therefore, toxicity may be predicted by read-across from members of the same chemical class. However, sometimes subtle structure changes result in relevant toxicological changes, making prediction solely based on chemoinformatics insufficient.
Any biologically based information demonstrating that structurally similar chemicals also share similar biology, would help to verify similar toxicity. High-information-content data generated with omics technologies from *in vivo* or *in vitro* testing or by high-throughput *in vitro* screening assays (HTS) appear promising in providing the *biological verification* of hypothesis generated by chemoinformatics.

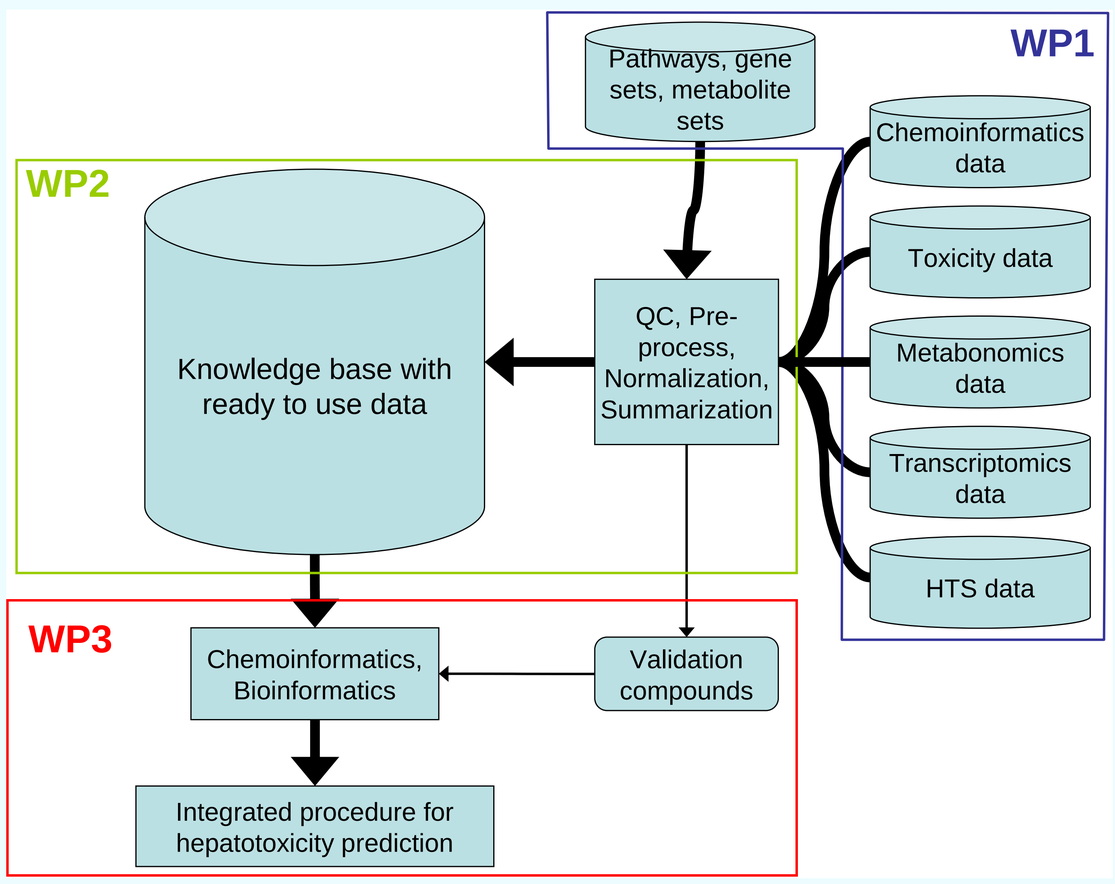
Hypothesis

Evaluation strategies based on integrating chemoinformatics approaches with mechanistic data from 'omics' and HTS technologies, will strengthen confidence that the respective compounds within a class behave similarly in terms of their toxicological profile.

Objectives

✓To develop a transparent framework that improves the prediction of the repeated dose toxicity – i.e. **hepatotoxicity** – of new chemicals by integrating chemoinformatic data with biological information from 'omics' and HTS technologies.
✓To generate a high-quality database that contains for a large number of chemicals, its chemical properties, *in vitro* omics, *in vivo* omics, HTS data, and (*in vivo*) toxicity test.
✓To develop and evaluate expert rules, useful for industry as well as regulators

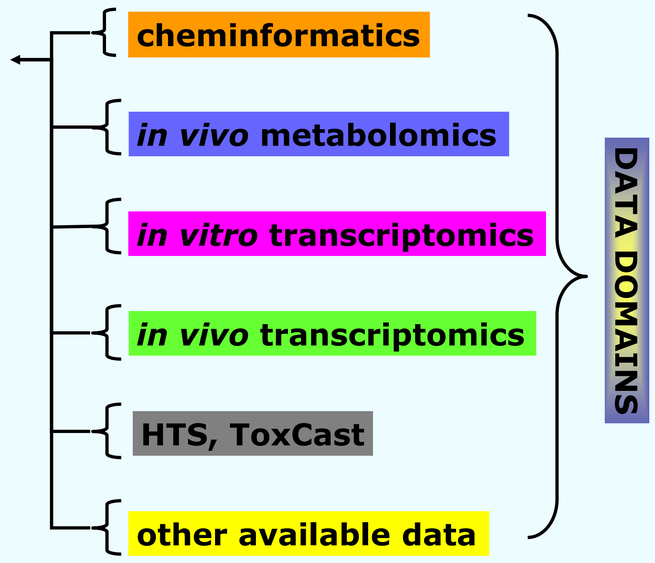
Project Organization in three Work Packages and the associated Workflows

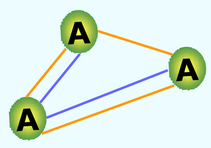


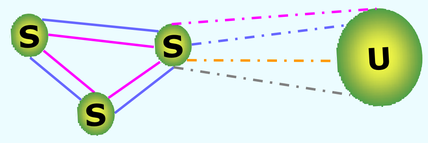
Based on a large number of chemicals, a database is generated which contains the chemical properties, *in vitro* and *in vivo* omics data (from public domain such as DrugMatrix & TG-Gates as well as in-house data from UM and BASF), HTS data (i.e. ToxCast) and *in vivo* hepatotoxicity test results. Tools are developed that integrate data from these different domains. Three levels of toxicity are predicted, namely pathology, mechanism of action, and hepatotoxicity potency. Eventually, we will determine if integration of omics and/or HTS indeed improves the prediction by chemoinformatics.

Toxicity prediction approaches based on similarity scores

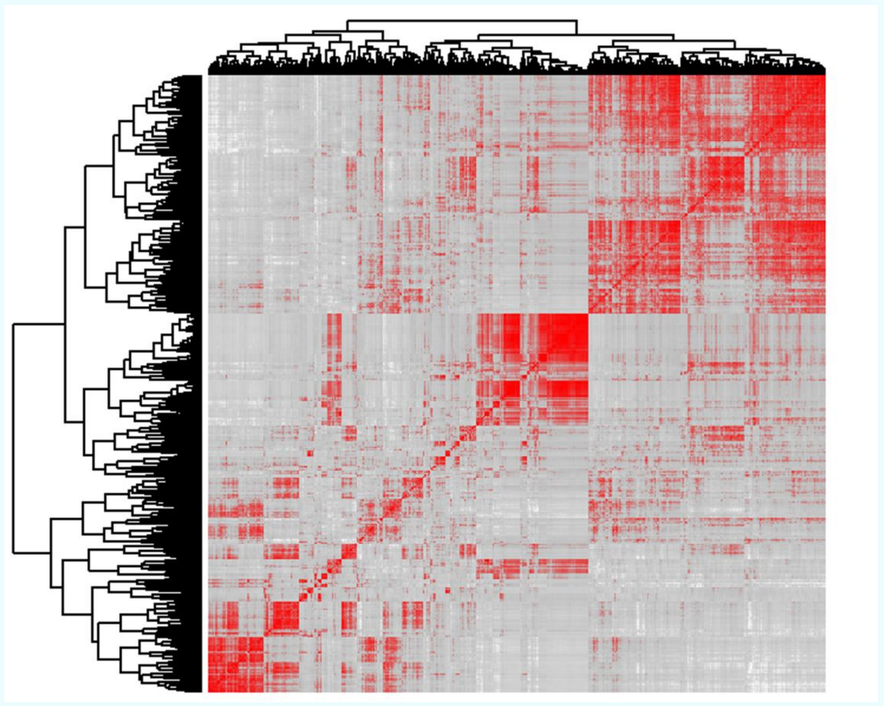
- Preparation of matrixes of similarity scores
- Grouping of chemicals based on these matrixes, e.g. by hierarchical clustering, Connectivity mapping and Cytoscape
- Interpreting these groups towards toxicological effects (enrichment analyses)
- Testing with the validation compounds



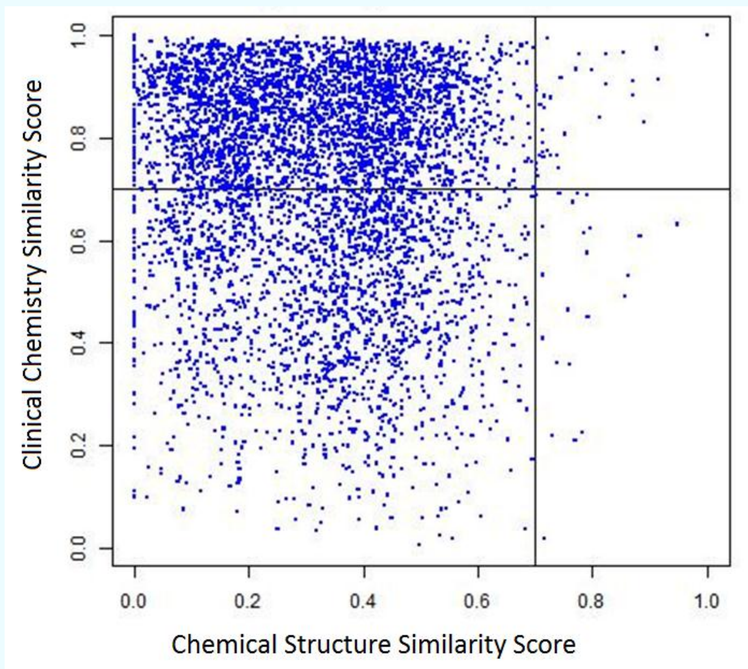




Results



Example of similarity matrix: Heatmap of correlations based on clinical chemistry for 583 compounds in DrugMatrix.



Comparison of similarity scores for compounds from TG-GATES.
→Structural similar is not equal to toxicological similar!
→Will inclusion of omics data improve Toxicity prediction?

Final outcome

A demonstration project where the advancement of integrating different domains, i.e. chemoinformatics, 'omics, and HTS, is clearly explained to regulators and industry.

Joost van Delft, PhD
Department of Toxicogenomics
Maastricht University

E: j.vandelft@maastrichtuniversity.nl
T: +3143 388 1092

P.O. Box 616
6200 MD Maastricht, The Netherlands
www.maastrichtuniversity.nl