In-silico platform based on bioinformatic and chemoinformatic data to complement zebrafish embryo teratogenicity test

Xabier Cendoya[1], Celia Quevedo[2] , Ángel Rubio[1], Maitane Ipiñazar[2] and Francisco J. Planes[1]

[1] CEIT and Tecnun, University of Navarra, San Sebastian, Spain. *fplanes@tecnun.es*

[2] BIOBIDE; San Sebastian, Spain. *quevedo@biobide.es*

INTRODUCTION

Approximately 3% of newborns present congenital anomalies and around 5-10% of those are caused by exposure to teratogenic agents. For this reason, regulatory organisms and the industry demand for effective methods to test the developmental toxicity of drugs, industry chemicals or waste products. The use of the zebrafish embryotoxicity test is an attractive strategy to minimize in-vivo assays and animal models. Overall, this assay has a good predictability; however, the outcome is based on morphologic evaluation, which is subjective and subtle effects might be neglected. With the increasing amount of molecular databases, the development of *in-silico* tools that complement experimental assays is promising. QSAR models are typically the method of choice to predict teratogenicity. However, these computational methods are limited to the toxicity of the assay that is being tested and biological data is not considered. In this work, we present an *in-silico* platform that integrates both bioinformatics and chemoinformatics data in order to more accurately characterize mechanisms of action of teratogenic compounds. The ultimate goal of this approach is to complement the zebrafish embryotoxicity test using heterogeneous data (beyond zebrafish) and improve the sensitivity of the assay using biological markers.

METHODS

This work consists of an *in-silico* platform (TERATOOL) that integrates diverse sources of data to allow bioinformatics and chemoinformatics analysis (Figure 1). The information that was integrated included:

- ✓ A database of approximately 400 compounds with labels for their risk of teratogenicity. 290 of them were obtained from Enoch et al [1].
- ✓ Chemical structures and properties, together with biological target data, were obtained from ChEMBL [2], DrugBank [3] and HMDB [4].
- ✓ A number of publicly available transcriptomics experiments of the zebrafish embryotoxicity test (40 compounds) [5,6,7,8,9,10].
- ✓ A metabolic network reconstruction of the zebrafish was obtained from Bekaert et al [11].

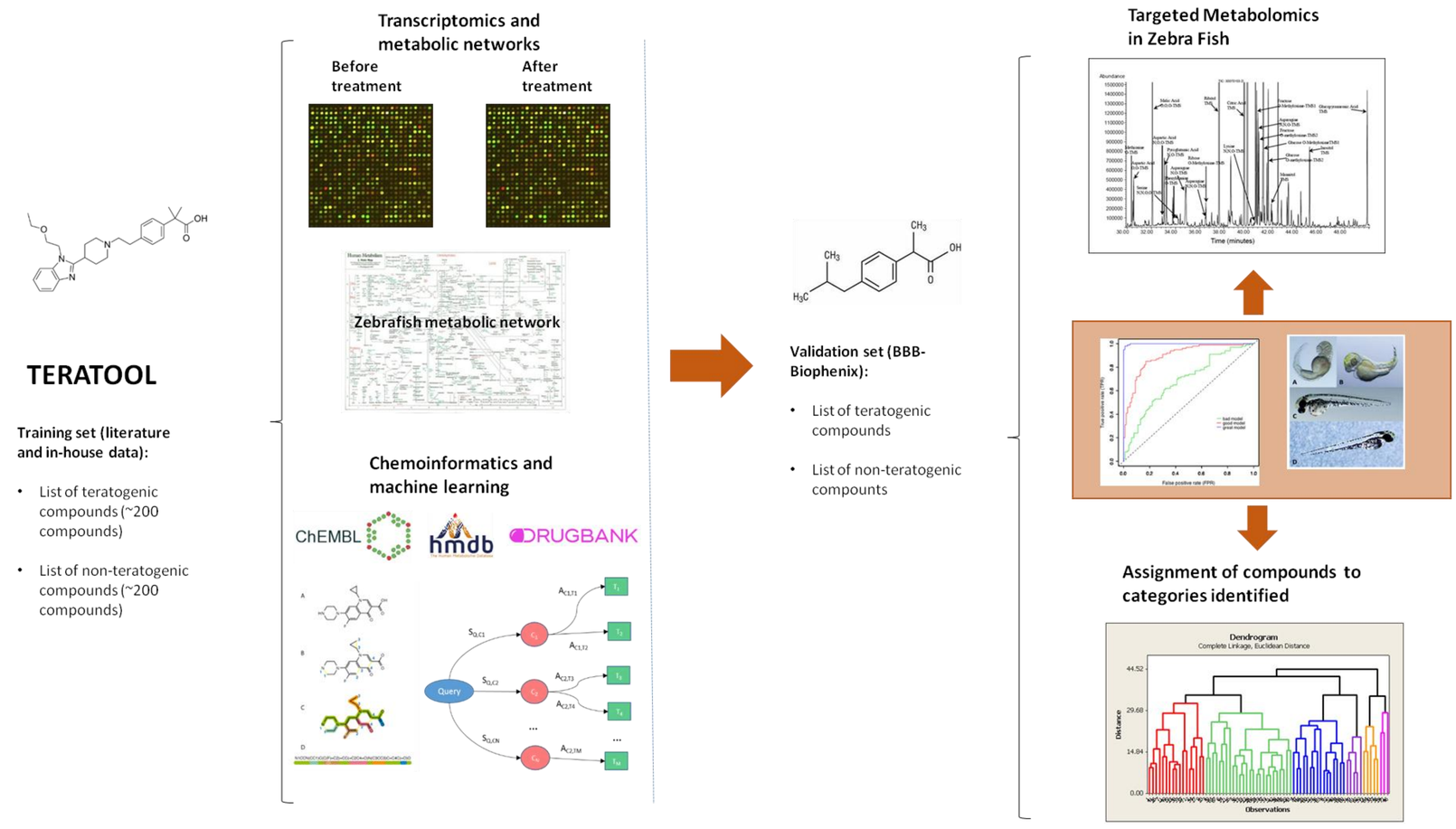
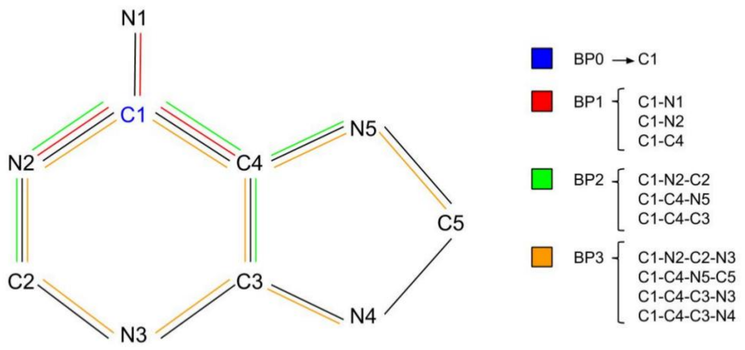


Figure 1. The TERATOOL platform integrates a list of approximately 400 molecules an includes chemoinformatics, transcriptomics and metabolic network data. With this, we can test several algorithms and identify potential biomarkers that can later be experimentally validated.

RESULTS

A differential expression analysis of the transcriptomics data was carried out. Selecting the genes that had an *adjusted p.value* < *0.05*, recurrent genes (differentially expressed for at least 4 of the compounds) that appeared differentially expressed in teratogenic molecules were identified.

Using molecular fingerprints and the Tanimoto equation, similarities to the molecules in ChEMBL, DrugBank and HMBD were calculated for the labeled compounds. This was done to understand the mechanism of action of teratogenic compounds beyond annotated interactions (Figure 2).



$$S_{AB} = \sqrt{\frac{\sum_{\forall i} CV_i \cdot \#common_i}{\sum_{\forall i} CV_i \cdot \#totalA_i} \cdot \frac{\sum_{\forall i} CV_i \cdot \#common_i}{\sum_{\forall i} CV_i \cdot \#totalB_i}}$$

Figure 2. Molecular fingerprint obtained by taking atom bond paths and Tanimoto equation.

Finally, in order to better understand the underlying metabolic mechanism of alteration when exposed to teratogenic agents, the reporter metabolites algorithm [12] was used. We searched for metabolites that indicated highly altered regions of the metabolic network, namely by establishing an integrative score based on differential expression analysis of neighbor genes (Figure 3). We found several metabolites potentially reporting teratogenic action with a substantially higher redundancy than gene biomarkers.

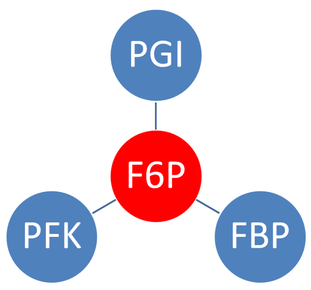


Figure 3. The idea behind the reporter metabolites algorithm is that, if the neighbor genes linked to a certain metabolite (fructose-6-phosphate in the example) are highly altered in a certain condition, the metabolite will be capable of reporting such condition.

REFERENCES

[1] Enoch SJ, Cronin MT, Madden JC, et al. Formation of Structural Categories to Allow for Read-Across for Teratogenicity. QSAR & Combinatorial Science 2009; 28:696–768.
[2] A.P. Bento, A. Gaulton, A. Hersey, L.J. Bellis, J. Chambers, M. Davies, F.A. Krüger, Y. Light, L. Mak, S. McGlinchey, M. Nowotka, G. Papadatos, R. Santos and J.P. Overington (2014) 'The ChEMBL bioactivity database: an update.' Nucleic Acids Res., 42 1083-1090.
[3] Law V, Knox C, Djoumbou Y, Jewison T, Guo AC, Liu Y, Maciejewski A, Arndt D, Wilson M, Neveu V, Tang A, Gabriel G, Ly C, Adamjee S, Dame ZT, Han B, Zhou Y, Wishart DS. DrugBank 4.0: shedding new light on drug metabolism. Nucleic Acids Res. 2014 Jan 1;42(1):D1091-7.
[4] Wishart DS, Jewison T, Guo AC, Wilson M, Knox C, et al., HMDB 3.0 — The Human Metabolome Database in 2013. Nucleic Acids Res. 2013. Jan 1;41(D1):D801-7.
[5] Schiller, V., Wichmann, A., Kriehuber, R., Schaefers, C., Fischer, R., & Fenske, M. (2013). Transcriptome alterations in zebrafish embryos after exposure to environmental estrogens and anti-androgens can reveal endocrine disruption. Reproductive Toxicology, 42, 210-223.
[6] Hermsen, S. A., Pronk, T. E., van den Brandhof, E. J., van der Ven, L. T., & Piersma, A. H. (2013). Transcriptomic analysis in the developing zebrafish embryo after compound exposure: individual gene expression and pathway regulation. Toxicology and applied pharmacology, 272(1), 161-171.
[7] Tzima, E., Serifi, I., Tsikari, I., Alzualde, A., Leonardos, I., & Papamarcaki, T. (2017). Transcriptional and Behavioral Responses of Zebrafish Larvae to Microcystin-LR Exposure. International Journal of Molecular Sciences, 18(2), 365.
[8] Choi, J. S., Kim, R. O., Yoon, S., & Kim, W. K. (2016). Developmental Toxicity of Zinc Oxide Nanoparticles to Zebrafish (Danio rerio): A Transcriptomic Analysis. PloS one, 11(8), e0160763.
[9] Haggard, D. E., Noyes, P. D., Waters, K. M., & Tanguay, R. L. (2016). Phenotypically anchored transcriptome profiling of developmental exposure to the antimicrobial agent, triclosan, reveals hepatotoxicity in embryonic zebrafish. Toxicology and applied pharmacology, 308, 32-45.
[10] GSE89780: https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE89780
[11] Bekaert, M. (2012). Reconstruction of Danio rerio metabolic model accounting for subcellular compartmentalisation. PloS one, 7(11), e49903.
[12] Patil, K. R., & Nielsen, J. (2005). Uncovering transcriptional regulation of metabolism by using metabolic network topology. *Proceedings of the National Academy of Sciences of the United States of America*, *102*(8), 2685-2689.

This work was supported by the Basque Government.