In Vitro to In Vivo Extrapolation (IVIVE) for Predicting Drug-Induced Liver Toxicity

Omkar Nanekar, Akash Parekh, Nilanjan Chattopadhyay, Nitin Singhal
AIRAMATRIX, Mumbai, India

EUROTOX 2024 COPENHAGEN | DENMARK 8-11 SEPTEMBER 2024

Objective

Bridging the Gap: The genome of an organism plays an intricate role in toxicity. Various factors, such as genetic variances, gene expression levels, and interactions between different components, all contribute to the final, potentially toxic composition. The objective here is to develop a predictive model for drug-induced liver toxicity by identifying key genes associated with cell toxicity in laboratory conditions and evaluating their ability to predict toxicity in living organisms. **This has the potential to minimize the need for animal experimentation in future studies.**

Experiment Setup

Data Summary	TG-GATEs In-Vitro	TG-GATEs In-Vivo		DrugMatrix [5]
	Rats	Rat-Single	Rat-Repeat	In-Vivo Single+Repeat
Dosage Level	High, Medium, Low	High, Medium, Low	High, Medium, Low	6hr, 24hr, 3 day, 5 day & 7 day
Dosage Duration	2hr, 8hr and 24hr	3hr, 6hr, 9hr and 24hr	4 day,8 day,15 day & 29 day	
Study Level	Hepatocyte	Liver Tissues	Liver Tissues	Liver Tissues
# Compounds	145	143	158	70
# Samples	3370	6765	7378	205

Background

Toxicogenomics examines the connection between toxicology and genomics, providing an in-depth assessment of how genes and cellular mechanisms react to toxins. We investigate the efficacy of this method in explaining the processes of toxicity, identifying indicators of exposure, and facilitating the development of safer drugs and informed risk evaluations.

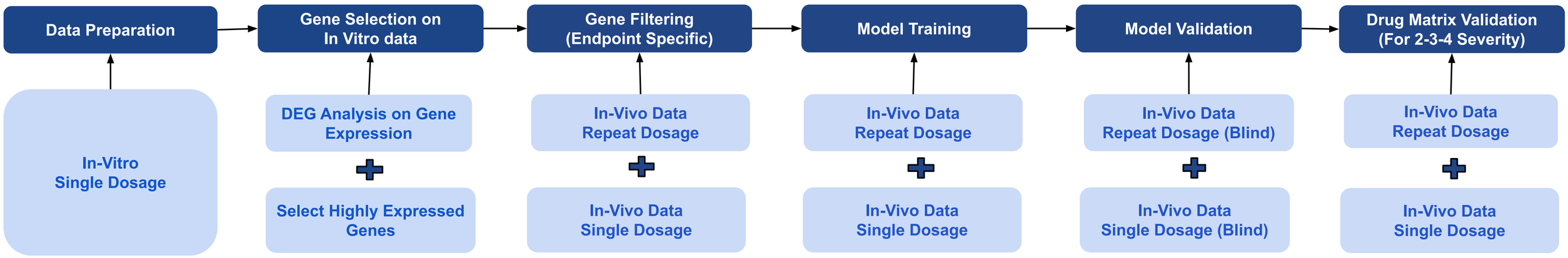
Differential Gene Expression (DEG) Analysis compares the levels of gene expression between two samples, usually a control group and a dosage group that has been exposed to a potentially toxic substance. Analyze alterations in gene expression following exposure to chemicals at specific dosage level. By identifying genes exhibiting significant disparities in expression levels between the groups, it aids in narrowing down the set of genes that may be involved in the toxic response.

In Vitro to In Vivo Extrapolation (IVIVE) plays an essential part in advancing toxicogenomics by connecting the controlled laboratory conditions with the complex structure of living organisms. By integrating In Vitro-In Vivo Extrapolation (IVIVE) and employing the use of DEG in our toxicological toolbox, we can aim for a future characterised by more ethical testing methods, safer pharmaceuticals, and a complete understanding of the effects of toxins on our well-being.

Methodology

- **Data Filtering:** Excluded severity class 1 compounds from both TG-GATES and Drug Matrix datasets to focus on more severe cases.
- **Preprocessing:** Combined in vivo single and repeat dose studies to create a unified dataset. Preprocessed the gene expression data of In Vivo and In Vitro using the RMA (Robust Multi-array Average) method, which includes background correction and normalization.
- **DEG Analysis:** Applied the Limma model to In vitro data to identify differentially expressed genes. Selected the top differentially expressed genes for further analysis.
- **Gene Selection and Model Training:** Split the data ensuring that compounds in the training set were not present in the testing set. Optimized gene selection with the help of machine learning techniques for specific endpoints (necrosis, mitosis and hypertrophy). Trained the model using this genes (AIRA gene set) on In vivo data for individual endpoints.
- **Blind Validation:** Internal validated the model on unseen in vivo compounds. Blind validated the model on Drug Matrix data.
- **Literature Review:** Reviewed research papers to identify genes related to necrosis and other injuries. Compared the selected genes (AIRA gene set) from the study with those cited in the literature.

IVIVE Pipeline



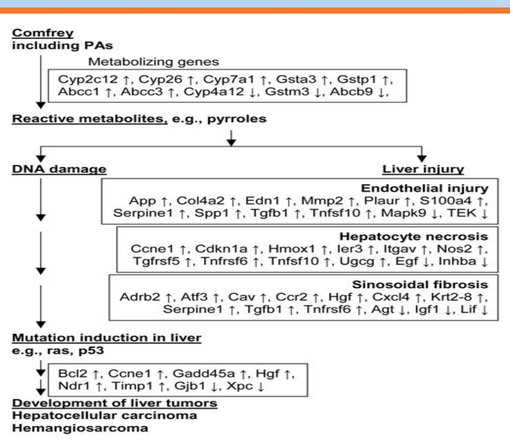
Performance (Necrosis)

Mean ROC AUC	Necrosis		
Dataset	AIRAMATRIX	Corton et al. [2]	Glaab et al. [3]
Validation	0.8218	0.6564	0.6658
Testing	0.8379	0.5809	0.6104
DrugMatrix [5]	0.7288	0.4500	0.4701

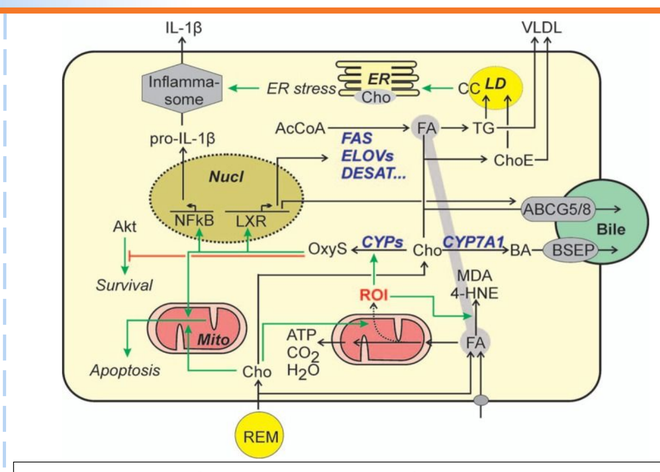
Performance (Other endpoints)

Mean ROC AUC	Increased Mitosis		Hypertrophy	
Dataset	AIRAMATRIX	Glaab et al. [3]	AIRA Gene	Glaab et al. [3]
Validation	0.9393 +/- 0.0178	0.8445 +/- 0.1144	0.8336 +/- 0.0854	0.6697 +/- 0.0817
Testing	0.9301 +/- 0.0081	0.8372 +/- 0.0262	0.8521 +/- 0.0100	0.5977 +/- 0.0219
DrugMatrix [5]	0.9807	0.9949	NA	NA

Network Analysis



Pathway Figures 1: Integration of gene expression with the proposed mechanistic pathway initiated by comfrey treatment leading to tumorigenesis



Pathway Figures 2: Cholesterol-dependent inflammatory response and cell death in the hepatocyte

References

[1] Yoshinobu Igarashi, Noriyuki Nakatsu, Tomoya Yamashita, Atsushi Ono, Yasuo Ohno, Tetsuro Urushidani and Hiroshi Yamada (2014) Open TG-GATEs: a large-scale toxicogenomics database. *Nucleic Acids Research* [online]. 43, D921-D927

[2] J. Christopher Corton, Thomas Hill, III, Jeffrey J. Sutherland, James L. Stevens, and John Rooney (2020) A Set of Six Gene Expression Biomarkers Identify Rat Liver Tumorigens in Short-Term Assays. *Toxicological Sciences* 2020 [online]. 177(1), 11–26

[3] Warren E Glaab, Daniel Holder, Yudong D He, Wendy J Bailey, David L Gerhold, Carolann Beare, Zoltan Erdos, Pamela Lane, Laura Michna, Nagaraja Muniappa, Jeffrey W Lawrence, Keith Q Tanis, Joseph F Sina, Thomas R Skopek, Frank D Sistare (2021) Universal Toxicity Gene Signatures for Early Identification of Drug-Induced Tissue Injuries in Rats. *Toxicological Sciences* 2021 [online]. 181(2), 148–159

[4] Shannon P, Markiel A, Ozier O, Baliga NS, Wang JT, Ramage D, Amin N, Schwikowski B, Ideker T. (2003) Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome Research 2003* [online]. 13(11), 2498-504

[5] Auerbach, S. S. (2012). *The DrugMatrix database*. In National Toxicology Program at NIEHS [online]

solutions@airamatrix.com
www.airamatrix.com