Journal Pre-proof

Machine learning classification of steatogenic compounds using toxicogenomics profiles

Brian Bwanya, Saad Lodhi, Theo M de Kok, Luiz Ladeira, Marcha CT Verheijen, Danyel GJ Jennen, Florian Caiment

PII: S0300-483X(25)00196-9

DOI: https://doi.org/10.1016/j.tox.2025.154237

Reference: TOX154237

To appear in: *Toxicology*

Received date: 10 June 2025
Revised date: 14 July 2025
Accepted date: 14 July 2025

Please cite this article as: Brian Bwanya, Saad Lodhi, Theo M de Kok, Luiz Ladeira, Marcha CT Verheijen, Danyel GJ Jennen and Florian Caiment, Machine learning classification of steatogenic compounds using toxicogenomics profiles, *Toxicology,* (2025) doi:https://doi.org/10.1016/j.tox.2025.154237

Machine learning classification of steatogenic compounds using toxicogenomics profiles

Brian Bwanya[a*], Saad Lodhi[a*], Theo M de Kok[a], Luiz Ladeira[b], Marcha CT Verheijen[a], Danyel GJ Jennen[a], and Florian Caiment[a]

* Contributed equally

[a] Department of Translational Genomics, GROW Research Institute for Oncology and Developmental Biology, Maastricht University, 6229 ER Maastricht, The Netherlands.

[b] Biomechanics Research Unit, GIGA Institute, University of Liège, Avenue de l'Hôpital, 11, B34 +5 4000 Liège Belgium.

Corresponding author: Florian Caiment
Email address: florian.caiment@maastrichtuniversity.nl

Abstract

The transition toward new approach methodologies for toxicity testing has accelerated the development of computational models that utilize transcriptomic data to predict chemical-induced adverse effects. Here, we applied supervised machine learning to gene expression data derived from primary human hepatocytes and rat liver models (*in vitro* and *in vivo*) to predict drug-induced hepatic steatosis. We evaluated five machine learning classifiers using microarray data from the Open TG-GATEs database. Among these, support vector machine (SVM) consistently achieved the highest performance, with area under the receiver operating characteristic curve (ROC-AUC) of 0.820 in primary human hepatocytes, 0.975 in the rat *in vitro* model, and 0.966 in the rat *in vivo* model. To gain mechanistic insights, we functionally profiled the top-ranked predictive genes. Enrichment analyses revealed strong associations with lipid metabolism, mitochondrial function, insulin signalling, oxidative stress, all biological processes central to steatosis pathogenesis. Key predictive genes such as *CYP1A1*, *PLIN2*, and *GCK* mapped to lipid metabolism networks and liver disease annotations, while others highlighted novel transcriptomics signals. Integration with differentially expressed genes and known steatosis markers highlighted both overlapping and distinct molecular features, suggesting that machine learning models capture biologically relevant signals. These findings demonstrate the potential of machine learning models guided by transcriptomic data to identify early molecular signatures of drug-induced hepatic steatosis. The

support vector machine model's strong predictive accuracy across species highlights its promise as a scalable and interpretable tool for chemical risk assessment. As data limitations in human toxicology persist, expanding high-quality transcriptomic resources will be critical to further advance non-animal approaches in regulatory toxicology.

Keywords: Drug-induced hepatic steatosis (DIHS), steatogenic prediction, machine learning (ML), Support Vector Machine (SVM)

Highlights

- Developed a ML framework for predicting DIHS using transcriptomic data from human and rat liver models
- SVM classifier achieved the highest predictive performance across all datasets
- SVM-derived predictive genes revealed mechanistic insights, including known and novel markers linked to lipid metabolism and liver disease
- Literature mining and enrichment analyses uncovered understudied genes potentially involved in early stress responses and progressive liver pathology
- Study supports the integration of ML with toxicogenomics for human-relevant, mechanism-based chemical safety assessment in line with new approach methodologies

1. Introduction

Advancing drug safety assessment is essential for reducing reliance on animal models for toxicity evaluation and improving the predictive accuracy of adverse drug reactions in humans. Among adverse drug reactions, drug-induced liver injury (DILI) remains a major clinical challenge, with a high risk of progression to acute liver failure (ALF) (Hosack et al., 2023). DILI is not only a major cause of ALF but also a primary driver of drug withdrawals from the market, resulting in substantial financial losses for pharmaceutical companies (Dirven et al., 2021). A notable manifestation of DILI is hepatic steatosis, commonly referred to as fatty liver disease. Hepatic steatosis is characterized by the excessive accumulation of triglycerides within hepatocytes, a process which can impair liver function and overall metabolic health (Idilman et al., 2016). Drug-induced hepatic steatosis (DIHS) arises from adverse reactions to pharmacologically

diverse medications, such as amiodarone, valproic acid, dexamethasone, tamoxifen, glucocorticoids, and other chemotherapeutic agents (Amacher & Chalasani, 2014). In some cases, DIHS can progress to drug-induced steatohepatitis, an inflammatory DILI subtype marked by lipid accumulation, hepatocyte degeneration, and significant liver damage (Cataldi et al., 2021).

Accurately identifying DIHS is complex due to the multifactorial nature of liver fat accumulation and the overlapping etiologies of hepatic steatosis (Cataldi et al., 2021). This challenge is exacerbated by the strong association between DIHS and the duration and dosage of medication (Satapathy et al., 2015). The subtlety of early-stage steatosis, which often lacks overt symptoms, complicates diagnosis even further. Although histological examination remains the gold standard for diagnosis, it involves invasive procedures like liver biopsy, which may not always be feasible or justifiable for every patient suspected of having DILI. While liver biopsies are valuable for establishing the stage and severity of steatosis, they offer limited insight into the underlying molecular mechanisms driving disease onset and progression. Additionally, individual variations in genetic makeup and metabolism add another layer of difficulty in identification of steatogenic drugs (Weiler et al., 2015). Notably, the Drug-Induced Liver Injury Network (DILIN) estimates that approximately 27% of DILI cases involve some degree of steatosis (Kolaric et al., 2021). This statistic underscores the clinical relevance of DIHS, and the need for improved tools for detection and monitoring steatogenicity of chemical compounds.

Despite advancements in alternative testing strategies, chemical safety evaluations continue to rely heavily on archaic *in vivo* animal toxicity assays, which remain the regulatory gold standard for assessing systemic toxicity (Caloni et al., 2022; Schmeisser et al., 2023). These evaluations involve measuring biological endpoints in animals exposed to the test compounds to assess potential adverse effects, which in turn inform predictions of human outcomes (Browne et al., 2024). However, animal-based toxicity assays are resource-intensive, ethically contentious, and often fail to fully recapitulate human-specific toxicological responses (Atkins et al., 2020). As a result, there is increasing momentum towards integrating toxicogenomics data with machine learning (ML) to enhance predictive toxicology, enabling more human-relevant, high-throughput chemical safety assessments. In alignment with this shift, the European Union established the Animal-free Safety assessment of chemicals: Project Cluster for Implementation of novel Strategies (ASPIS) to advance sustainable, animal-free, and reliable approaches for chemical risk assessment. Within this framework, the ASPIS cluster curated a reference list of steatogenic and

non-steatogenic compounds, serving as a standardized foundation for investigating chemical-induced steatosis. This availability of well-defined binary classifications facilitates the application of ML algorithms in toxicity prediction. Central to this approach is transcriptomics data, which captures molecular signatures of steatogenesis and provides a mechanistic basis for classifying compounds based on their potential to induce hepatic lipid accumulation and associated toxicological outcomes. For this study, transcriptomic data were sourced from the Open Toxicogenomics Project-Genomics Assisted Toxicity Evaluation System (TG-GATEs) database (Igarashi et al., 2015).

Machine learning has emerged as a powerful approach for predictive toxicology, enabling improved chemical hazard prediction through integrative analysis of toxicogenomic data. Jiang et al. (2023) demonstrated this potential by applying supervised ML algorithms to Primary Human Hepatocytes (PHH) transcriptomic data from TG-GATEs, yielding high predictive performance and identifying 13 key genes linked to drug-induced intrahepatic cholestasis (DIC). In a related approach, O'Donovan et al. (2023) employed transfer learning within deep neural networks to predict human gene expression profiles from rat data, demonstrating cross-species generalization. Aguayo et al. (2021) integrated ML within a qualitative gene-expression activity relationship framework to predict DILI from cancer cell line transcriptomes, achieving consistently high predictive performance across validation tests. Collectively, these studies underscore the growing utility of ML within toxicogenomics for advancing mechanism-based, human-relevant chemical safety assessment.

Unlike other hepatotoxicity endpoints, DIHS poses unique challenges due to its gradual onset and the multifactorial nature of lipid accumulation, which is influenced by genetic, metabolic, and environmental factors (Cataldi et al., 2021; Satapathy et al., 2015; Weiler et al., 2015). Although prior studies have successfully applied ML to TG-GATEs data, the approach used here diverges in several key aspects. Jiang et al. (2023) developed ML models for DIC using differentially expressed genes **(**DEGs) as input features. O'Donovan et al. (2023) applied transfer learning to predict human hepatic transcriptomic responses from rat *in vitro* and *in vivo* data. Aguayo et al. (2021) integrated ML with pathway enrichment to identify key biological processes underlying DILI. In contrast, our study directly uses the pre-processed gene expression data as input, without restricting features to DEGs or pathway-enriched subsets. This strategy enables the model to capture broader transcriptomic signatures of steatogenesis that might be missed by

targeted approaches. Additionally, we employ bootstrap resampling**,** with model performance evaluated independently across 1,000 iterations, to enhance robustness and reproducibility. This framework was used to address two key questions relevant to chemical risk assessment:

1. Can ML algorithms accurately predict the steatogenic potential of a compound from transcriptomic data?
2. Can predictive ML models identify both established and novel gene markers of DIHS, and do these features generalize across species to support translational relevance?

To address these questions, we evaluated five supervised ML algorithms: elastic net logistic regression (ENLR), extreme gradient boosting (XGBoost), k-nearest neighbors (KNN), random forest (RF), and support vector machines (SVM), for predicting chemical steatogenicity using human and rat TG-GATEs transcriptomic data.

2. Materials and methods

2.1 Data retrieval

A curated list of 31 steatogenic and 9 non-steatogenic compounds (Supplementary Table 1), endorsed by the ASPIS chemical selection working group, was used to query ArrayExpress (https://www.ebi.ac.uk/biostudies/arrayexpress) for liver transcriptomic datasets. To ensure consistency in experimental design, species, and exposure conditions, analyses were restricted to the Open TG-GATEs database, which includes data for 170 compounds (Igarashi et al., 2015). TG-GATEs comprises microarray-based gene expression profiles for ~19,914 genes in PHH, and ~12,153 genes in primary rat hepatocytes (*in vitro*; PRH) and rat liver tissue (*in vivo*; RLT). PHH and PRH samples were collected at 2, 8, and 24 h following low, medium, or high-dose treatment (two biological replicates per condition). RLT samples were collected at 3, 6, 9, and 24 h after a single acute dose (three replicates per condition) (Figure 1). Data on six ASPIS-listed compounds were available in TG-GATEs: four steatogenic (tamoxifen, valproic acid, amiodarone, carbon tetrachloride) and two non-steatogenic (colchicine, imipramine). Full dataset details are provided in Igarashi et al. (2015).

2.2 Data labeling

The TG-GATEs datasets for human and rat liver samples were systematically labeled and categorized as steatogenic or non-steatogenic, following classifications from the validated ASPIS chemical list. It is essential to underscore that data labeling was performed independently of dose and timepoint considerations.

2.3 Data preprocessing

The raw microarray data were preprocessed in R using the affy package (version 1.80.0) (Gautier et al., 2004). Probes were reannotated with Ensembl gene identifiers using CustomCDF version 25 from the BrainArray database (Dai et al., 2005). Preprocessing included background correction, normalization, and summarization of probe-level intensities to generate gene expression values. Genes with a mean expression level below 6 were excluded. The resulting gene expression data were subsequently used for downstream analyses, including differential gene expression analysis, ML, and pathway enrichment. Principal component analysis (PCA) was performed independently for rat and PHH datasets to assess sample-level variance (Supplementary Figures 1A-F).

2.4 Differential Gene Expression Analysis

The Omics Data Analysis Framework for Regulatory application (R-ODAF) (M. C. Verheijen et al., 2022) was used to identify DEGs between steatogenic and non-steatogenic conditions. Each dataset, PHH, PRH, and RLT was analyzed independently. The complete list of DEGs is available here https://github.com/TGX-UM/ASPIS_Omics_ML/tree/main/List_of_DEGs.

2.5 Machine learning classification of compounds

Five supervised ML algorithms, ENLR, XGBoost, KNN, RF, and SVM were used to classify compounds based on their steatogenic potential. Models were trained using TG-GATEs liver transcriptomic data for the six ASPIS-listed compounds. The complete ML code is publicly available at https://github.com/TGX-UM/ASPIS_Omics_ML/tree/main/Scripts. ML workflows were developed using the tidymodels package (version 1.2.0) (Kuhn & Wickham, 2020).The following steps were applied to develop the ML models.

i. Feature engineering and preprocessing

Near-zero variance features (genes) were removed to eliminate highly correlated or uninformative genes that may have bypassed mean expression filtering. Class imbalance was corrected using the Synthetic Minority Oversampling Technique (SMOTE). In PHH data, the non-steatogenic class was underrepresented, while in PRH and RLT, the steatogenic class was the minority.

ii. Model training and resampling

Data from PHH, PRH, and RLT were analyzed separately. To ensure statistical robustness, each dataset underwent 1,000 stratified bootstrap resampling iterations, preserving class balance. Each resample was split into a training (analysis) set and an out-of-bag (OOB) test set. Models were trained on the analysis set and evaluated on the corresponding OOB test set.

iii. Model evaluation

Model performance was assessed using six metrics: accuracy, sensitivity, specificity, area under the receiver operating characteristic curve (ROC AUC), Youden's J index, and Matthews correlation coefficient (MCC). These were computed across all 1,000 bootstrap iterations to ensure reliability.

iv. Model application to full TG-GATEs PHH dataset

Based on cross-validation performance, the linear SVM achieved the highest ROC AUC across all datasets (PHH, PRH, and RLT) and was selected as the best-performing model. The SVM model trained on PHH data was subsequently applied to the full TG-GATEs PHH dataset. Steatogenicity was predicted for all compounds, doses, and time points using a classification probability threshold > 0.7. Predictions were benchmarked against the next two best-performing models, RF and ENLR, based on their ROC AUC scores.

v. Feature importance and predictive gene selection

Gene importance scores were derived from the final linear SVM models trained on each dataset (PHH, PRH, and RLT). Scores were calculated from model coefficients, converted to absolute

values, and normalized (0–1). The top 100 ranked genes per dataset were selected for downstream enrichment analyses.

vi. Functional mapping

Top predictive genes were mapped onto the Liver Lipid Metabolism Physiological Map (LiverLipidPM) v3 (Ladeira et al., 2024; Ladeira et al., 2025) to investigate their roles in hepatic lipid metabolism. Genes not annotated for liver-specific functions or disease associations were subjected to additional pathway enrichment analysis using Enrichr-KG (Evangelista et al., 2023), focusing on Gene Ontology (GO) Biological Processes, Reactome, KEGG, and WikiPathways databases. The full analytical workflow, from data preprocessing to functional interpretation, is summarized in Figure 1. Overlapping gene lists from PHH, PRH, and RLT datasets were harmonized for integrated visualization. Rat genes were mapped to human orthologs using the R packages org.Hs.eg.db, org.Rn.eg.db, and biomaRt (Durinck et al., 2009). Gene identifiers that could not be automatically mapped were manually verified and annotated using the Bgee database (Bastian et al., 2025).

The LiverLipidPM comprises 639 unique HGNC-approved symbols represented as gene, RNA, or protein nodes within a comprehensive network of metabolic, signalling, and regulatory pathways. We programmatically accessed the LiverLipidPM through the minervaR package (Gawron et al., 2023) to extract node names for direct comparison with predictive gene sets. This enabled assessment of overlap between LiverLipidPM components and top predictive genes from PHH, PRH, and RLT datasets, evaluating mechanistic coverage. Overlay files were generated for each dataset to visualize matched genes interactively on the Molecular Interaction NEtwoRk VisuAlization (MINERVA) platform.

vii. Text-mining and functional inference for unannotated genes

Genes not captured in Disease Ontology (DO), LiverLipidPM, or pathway enrichment were profiled using literature mining. Queries were constructed using the rentrez package (Winter, 2017) and run against PubMed and National Center for Biotechnology Information (NCBI) Gene database. Searches combined gene names with liver pathology–related keywords (e.g., hepatic dysfunction, metabolic regulation) restricted to title and abstract fields (TIAB). Genes were

categorized based on evidence from PubMed, NCBI Gene, both, or neither. Keywords used are listed in Supplementary Table 2.

viii. Comparison with DEGs and known steatosis genes

The top 1,000 SVM-predictive genes were compared with DEGs between steatogenic and non-steatogenic groups, and a curated list of known steatosis-associated genes from the Rat Genome Database (RGD) (Vedi et al., 2023). The RGD list was filtered based on the term column to retain genes associated with steatotic liver disease, metabolic dysfunction-associated steatohepatitis, and metabolic dysfunction-associated steatotic liver disease. Overlapping genes were mapped to the LiverLipidPM v3 (Ladeira et al., 2024). All gene lists used in this analysis are available at https://github.com/TGX-UM/ASPIS_Omics_ML.

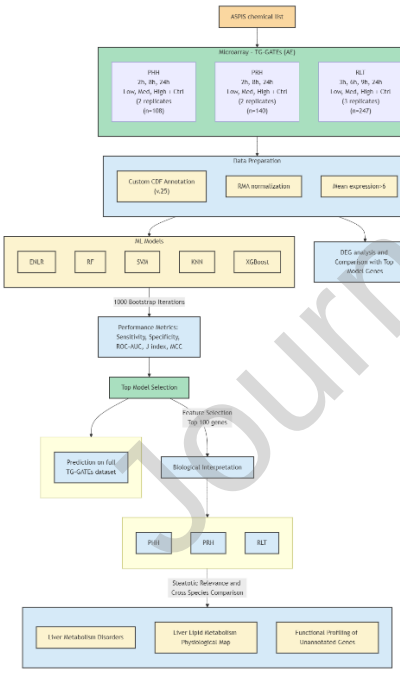


Figure 1. Workflow diagram illustrating the machine learning (ML) analysis pipeline.

A curated list of 31 steatogenic and 9 non-steatogenic compounds was used to query the ArrayExpress (AE) database. Microarray gene expression data from the TG-GATEs project were retrieved for primary human hepatocytes (PHH), primary rat hepatocytes (PRH), and rat liver tissue (RLT), and normalised using the Robust Multi-array Average (RMA) method via the *affy* R

package. Five supervised ML algorithms were trained on the normalised data, with performance evaluated across 1,000 bootstrap resamples using sensitivity, specificity, ROC AUC, Youden's J-index, and Matthews correlation coefficient (MCC). The best-performing model, based on ROC AUC was used for downstream analysis, including predictive gene selection, pathway enrichment, functional profiling, and prediction across the full TG-GATEs PHH dataset.

3. Results

This study aimed to develop ML models to predict the steatogenic potential of chemicals using microarray gene expression data, and to identify both established and novel genes associated with DIHS. We also evaluated the translatability of *in vitro* findings to the *in vivo* situation. To achieve this, we used gene expression data from the Open TG-GATEs database, which includes transcriptomic profiles from PHH, PRH, and RLT. After preprocessing and quality control, the final dataset comprised 108 PHH samples, 140 PRH samples, and 247 RLT samples. Table 1 summarizes the classification of these samples into steatogenic and non-steatogenic categories. The final number of genes used for model training was 8,272 for the PHH dataset, 9,038 for the PRH dataset, and 8,305 for the RLT dataset. Supplementary Figures 2A-C provides bar plots illustrating the number of genes removed by the mean expression filter and those retained for analysis.

Table 1. Summary classification of data based on the ASPIS-cluster chemical list

Dataset	Number of Steatogenic samples	Number of non-steatogenic samples	Total number of samples
TG GATEs PHH	62	46	108
TG GATEs PRH	68	72	140
TG GATEs RLT	112	135	247

Predicting the steatogenic potential of chemicals using gene expression data

Among the five ML models tested, SVM outperformed all other classifiers across datasets, achieving the highest ROC AUC scores: 0.820 (PHH), 0.975 (PRH), and 0.966 (RLT) (Figures 2A-C). It also demonstrated high sensitivity (>0.8) across all datasets. In terms of specificity, the

PHH model performed moderately, with values slightly above 0.6, while both PRH and RLT models showed strong specificity, both exceeding 0.85 (Table 2). For the PHH dataset, the SVM model produced a Youden's J index of 0.515 and a MCC of 0.545. In the PRH dataset, both the J index and MCC were greater than 0.80, with similarly strong values observed in the RLT dataset (Table 2).

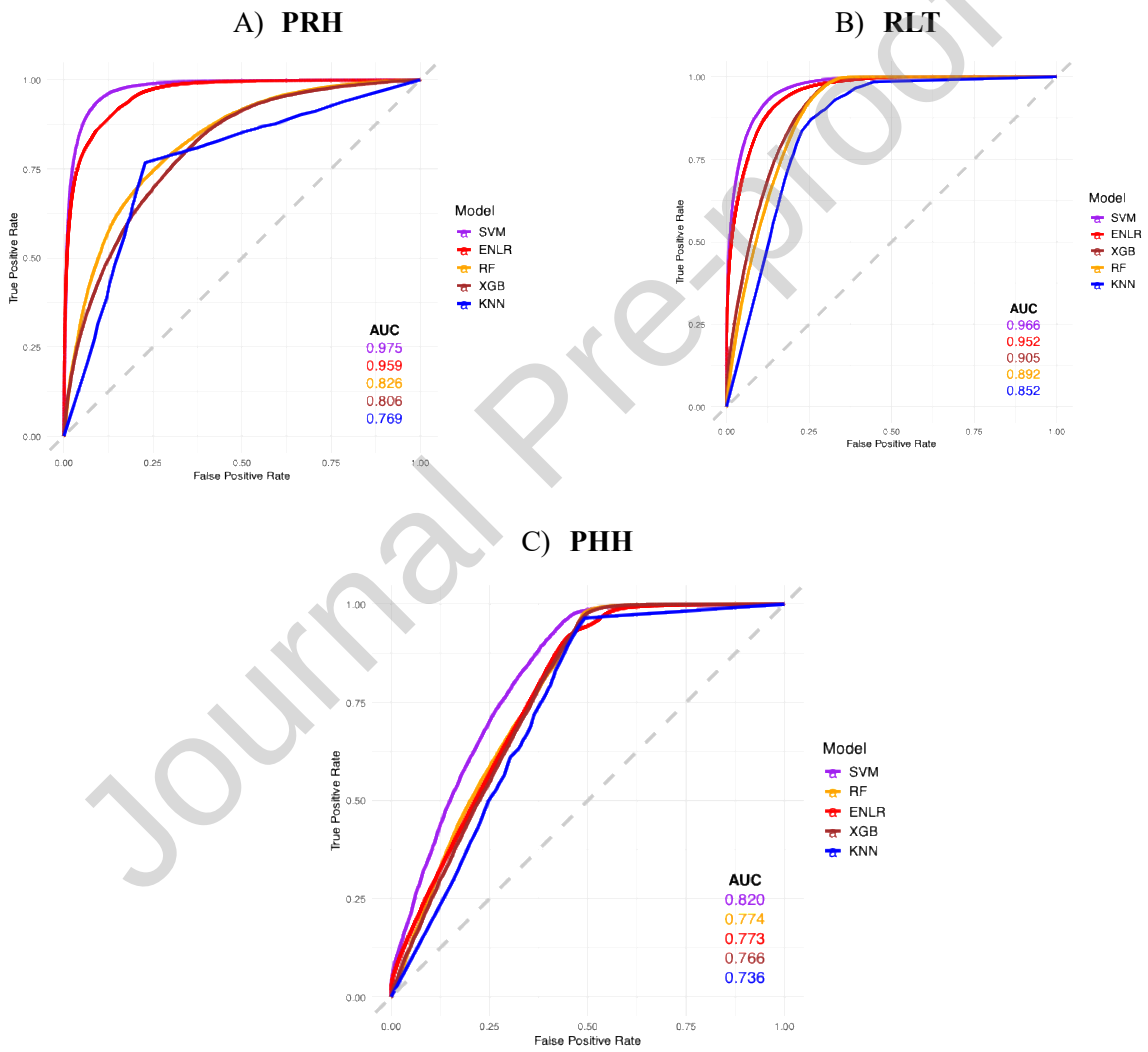


Figure 2. ROC AUC curves showing the classification performance of five machine learning models across three datasets.

Support vector machine (SVM), random forest (RF), elastic net logistic regression (ENLR), k-nearest neighbours (KNN), and extreme gradient boosting (XGBoost) were evaluated on A)

primary rat hepatocytes (PRH), B) rat liver tissue (RLT), and C) primary human hepatocytes (PHH). The colors of the ROC curves and AUC values correspond to their respective models as shown in the legend

Table 2. Additional performance metrics for all ML models across datasets. Metrics include sensitivity, specificity, Youden's J index, and MCC.

Models	Accuracy	Sensitivity	Specificity	J index	MCC
Primary Human Hepatocytes					
SVM	0.778	0.896	0.619	0.515	0.545
RF	0.740	0.866	0.571	0.437	0.464
ENLR	0.743	0.854	0.593	0.447	0.468
XGB	0.738	0.851	0.586	0.437	0.458
KNN	0.684	0.719	0.636	0.356	0.355
Primary Rat Hepatocytes					
SVM	0.923	0.930	0.916	0.847	0.846
ENLR	0.884	0.858	0.909	0.767	0.769
RF	0.746	0.763	0.731	0.493	0.493
XGB	0.725	0.736	0.715	0.451	0.451
KNN	0.770	0.768	0.772	0.539	0.539
Rat Liver Tissue					
SVM	0.900	0.927	0.878	0.805	0.802
ENLR	0.879	0.916	0.849	0.765	0.761
XGB	0.831	0.914	0.762	0.677	0.676
RF	0.828	0.914	0.756	0.671	0.671
KNN	0.801	0.836	0.772	0.608	0.605

SVM model evaluation on the full TG-GATEs PHH dataset

The SVM model trained on PHH data was then applied to the full TG-GATEs PHH dataset to predict compound steatogenicity. Out of 2,605 arrays evaluated, the SVM model classified 883 as steatogenic, compared to 953 by ENLR and only 70 by RF (Figure 3A). The comprehensive list of predictions from these three models is available here: https://github.com/TGX-UM/ASPIS_Omics_ML.

For non-steatogenic classifications, the SVM model identified 976 samples, RF identified 774, and ENLR yielded the highest number with 1,147 non-steatogenic predictions. The number of uncertain classifications varied substantially among the models, with SVM classifying 746

samples as uncertain, ENLR 505, and RF exhibited the highest uncertainty with 1,761 samples. Among control samples, those predicted as steatogenic by the SVM model had a median probability of 0.25, while those classified as non-steatogenic had a median probability of 0.75 (Figures 3B-D).

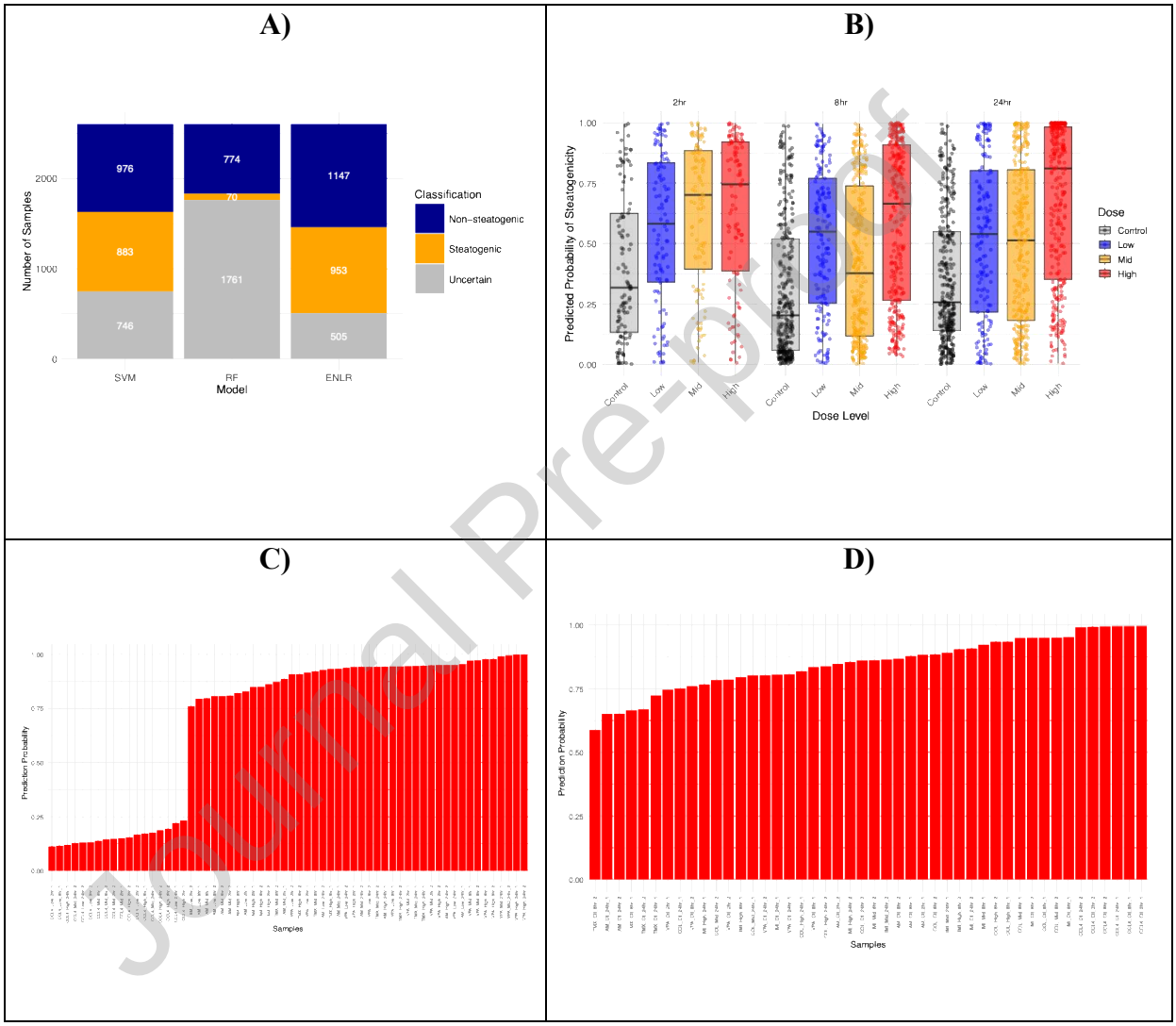


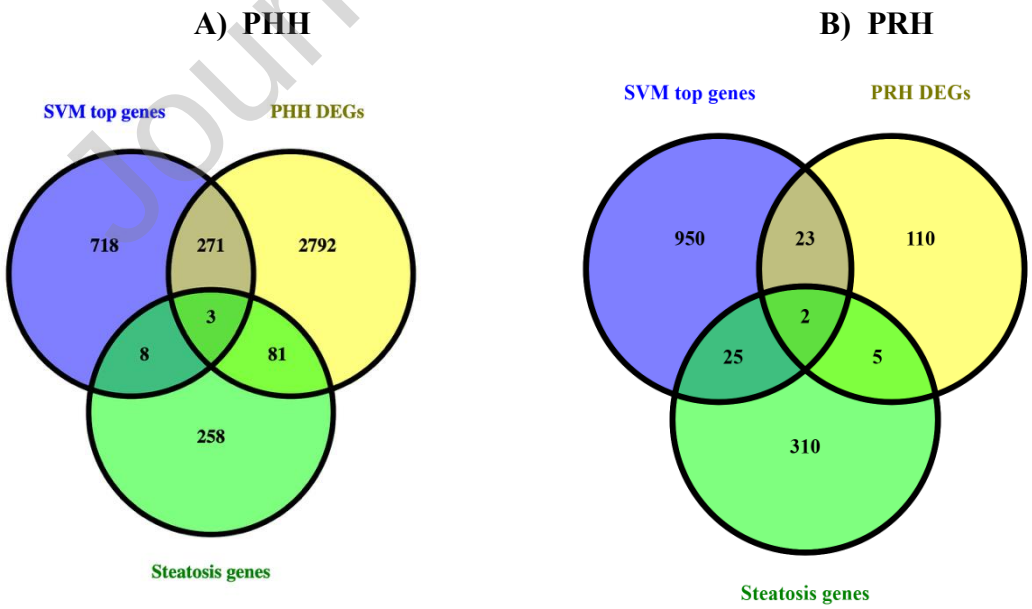
Figure 3. Machine learning-based predictions of steatogenic and non-steatogenic outcomes in PHH samples.

A) Stacked bar plot comparing prediction outcomes across SVM, RF, and ENLR models for PHH samples, with predictions categorized as steatogenic, non-steatogenic, or uncertain at a threshold > 0.7. B) Boxplot showing the dose and time-dependent effects on predicted steatogenicity across controls, low, mid, and high doses at 2 h, 8 h, and 24 h time points. Predicted probabilities of

steatogenicity generated using the SVM model applied to the TG-GATEs PHH dataset. C) Bar plot showing prediction probabilities generated by the SVM model for PHH samples treated with predefined steatogenic compounds. Each bar represents the probability of steatogenic classification across different doses and time points. D) Bar plot illustrating non-steatogenic prediction probabilities generated by the SVM model for control and treated samples originally labelled as non-steatogenic in the predefined chemical list. Each bar represents the predicted probability for an individual sample across different experimental conditions.

Comparison of SVM identified predictive genes with DEGs and known steatosis associated genes

To assess biological relevance, the top 1,000 SVM-predictive genes were compared with DEGs and a curated list of steatosis-associated genes. Overlap was limited across all datasets (Figures 4A-C). In PHH data, only three genes, *ACADM*, *INSR*, and *ULK1* were common to all three gene sets. Similarly, in PRH and RLT most SVM-predictive genes showed minimal overlap with both DEGs and the curated steatosis-associated gene list. Only two genes (*Gpt*, *Cd14*) overlapped across all three gene sets in PRH, while nine genes overlapped in RLT. Among these nine, notable shared genes include *Gck*, *Dhrs7*, and *Gpt2*. The full gene lists are available at https://github.com/TGX-UM/ASPIS_Omics_ML.



C) RLT

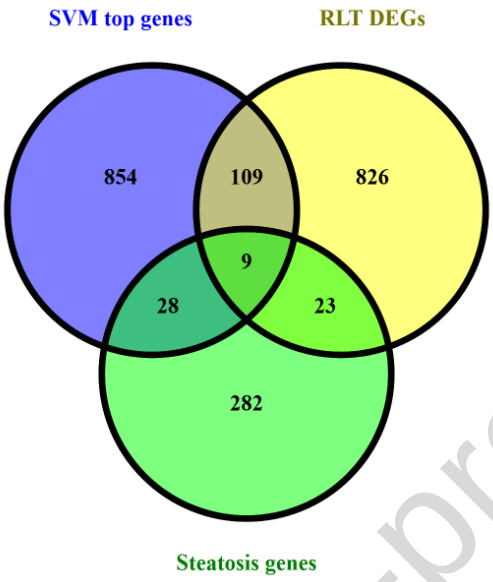


Figure 4. Overlap among the top 1,000 SVM-predictive genes, DEGs and known steatosis-associated genes.

Venn diagrams illustrate gene overlaps for (A) PHH, (B) PRH, and (C) RLT datasets. Known steatosis-associated genes were curated from the Rat Genome Database (RGD) based on liver disease terms related to steatosis.

Mapping Overlapping genes onto LiverLipidPM

Automated gene identifier conversion achieved a 98.8% success rate. In the human-derived PHH dataset, four genes (ENSG00000261829, ENSG00000267458, ENSG00000242861, ENSG00000286388) lacked HGNC-approved symbols and were identified as novel transcripts without functional annotation. In the rat-derived PRH and RLT datasets, three gene identifiers (ENSRNOG00000023828, ENSRNOG00000057601, ENSRNOG00000004147) lacked direct human orthologs and could not be converted. Additionally, three human and twelve rat genes were not automatically processed and were manually annotated. The full list of these unmapped gene symbols is available at the project GitHub repository (https://github.com/TGX-UM/ASPIS_Omics_ML).

Following this annotation, the predictive gene sets were mapped onto the LiverLipidPM (Figures 5A-D) to explore their functional roles in lipid metabolism. This analysis identified *ACADM* (acyl-CoA dehydrogenase medium chain) and *INSR* (insulin receptor) as the key mechanistic players. *ACADM* catalyzes mitochondrial beta-oxidation of saturated and unsaturated fatty acids and regulates acyl-CoA availability within fatty acid elongation pathways, redirecting metabolic flux away from triacylglyceride synthesis toward alternative biological processes. *INSR* plays a central role in glucose uptake, facilitating acetyl-CoA flux and thereby influencing both energy production and lipid metabolism.

Moving on to the PRH samples, only one gene, *Gpt* (glutamic-pyruvic transaminase), appeared in the central intersection and was present in the LiverLipidPM (Figures 5A-D). *Gpt* catalyzes the conversion of pyruvate into alanine and alpha-ketoglutarate, diverting pyruvate from the tricarboxylic acid cycle and into an anabolic pathway for amino acid generation (Figures 5A-D). In contrast, the RLT samples were more enriched in genes, with 9 hits in the intersections, though the overlap with the LiverLipidPM was limited. Notable genes such as *Gck* (glucokinase), *Dgat2* (diacylglycerol O-acyltransferase 2), and *Scarb1* were visualized (Figures 5A-D). *Gck* is involved in the first step of glucose metabolism, phosphorylating glucose into glucose-6-phosphate, which initiates the glycolysis pathway. *Dgat2* plays a crucial role in lipid droplet dynamics by converting diacylglycerols into triacylglycerols, allowing lipid storage in intracellular lipid droplets. *Scarb1* facilitates the uptake of high-density lipoproteins (HDL) in hepatocytes, increasing intracellular cholesterol and phospholipid levels, which can subsequently be stored in lipid droplets.

Figures 5A-D depict the data visualization in the LiverLipidPM, showcasing the SVM, steatosis, and DEG intersection gene list for PHH samples

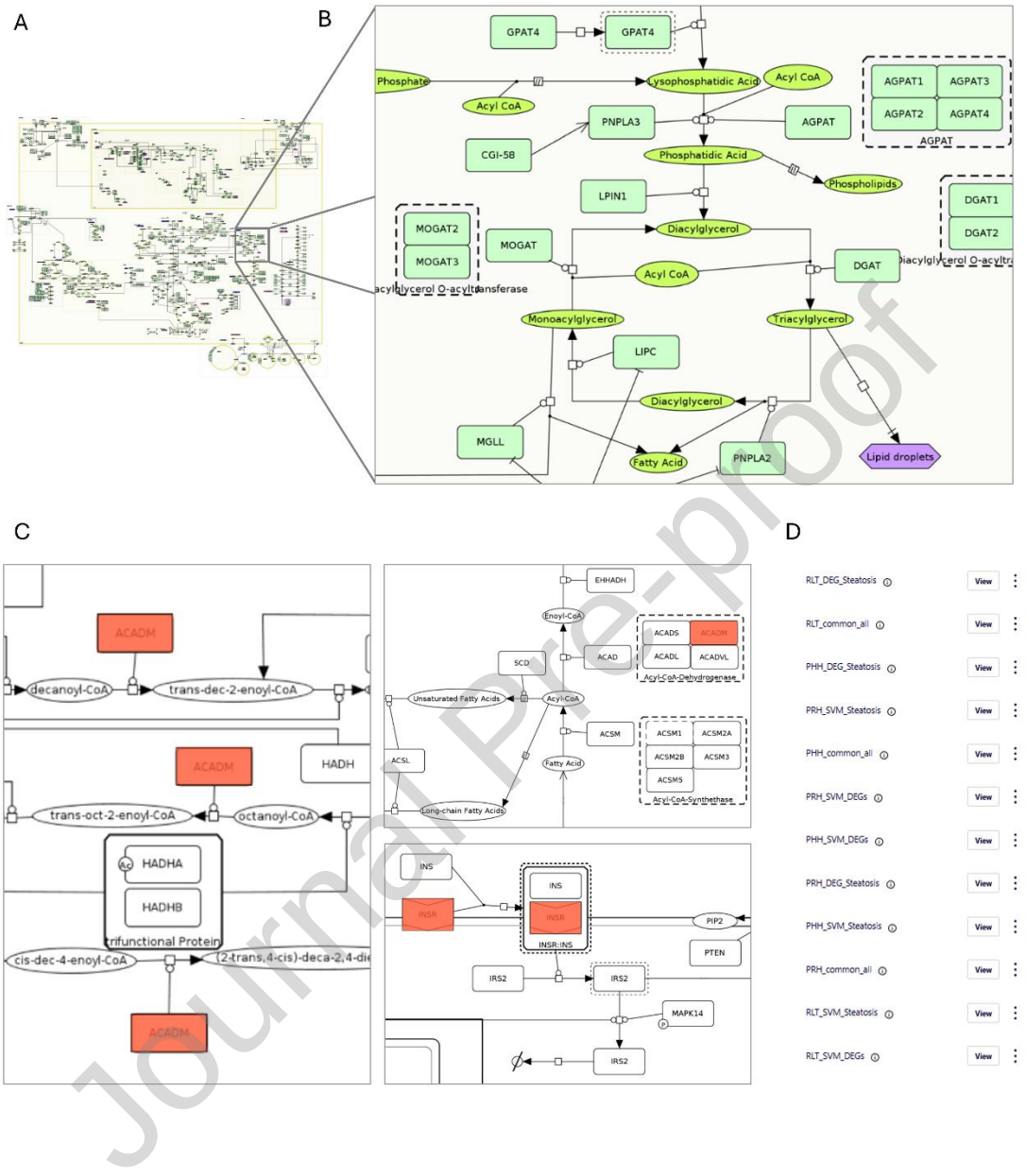


Figure 5. ***Visualization of predictive genes in the LiverLipidPM.***

A) Overview of the full LiverLipidPM showing the network of integrated metabolic, signaling, and regulatory pathways. B) Zoom-in on the fatty acid elongation pathway, illustrating molecular interactions represented in the map. C) Visualization of predictive genes from the PHH dataset, overlaid on the map. Distinct colors are used to distinguish data from different gene lists; colors do not convey intrinsic meaning. D) Gene list overlays available for interactive exploration on the MINERVA platform at https://ontox.elixir-

luxembourg.org/minerva/index.html?id=Liver_Lipid_Metabolism_Physiological_Map_v3, accessible via the "Overlays" tab.

Biological Interpretation of Predictive Gene Sets

To ensure a balance between predictive relevance and biological interpretability, the top 100 genes from each dataset were selected based on their importance scores derived from the SVM model, the best-performing model across all three datasets. This approach prioritized the most predictive features for subsequent biological analysis. Across all three datasets, the importance scores of the selected genes remained consistent, with the top 100 genes selected above the threshold of 0.45 in PHH, 0.47 in PRH, and 0.44 in RLT, indicating a similar ranking of predictive features.

Mapping the Top 100 SVM Genes to the LiverLipidPM

The selected top 100 SVM genes were mapped to the LiverLipidPM (Ladeira et al., 2024), a manually curated knowledge-based framework that integrates pathways and regulatory networks involved in hepatic lipid metabolism (e.g., fatty acid, triglyceride, and cholesterol synthesis). This map was used as a reference to identify which specific lipid-metabolic processes were represented by the most predictive genes thereby facilitating the identification of chemical-induced disruptions contributing to steatosis development. In total, 15 genes were identified: 3 in PHH (eg., *CYP1A1*), 4 in PRH (eg., *Mgll*), and 8 in RLT (eg., *Gck*). These genes are primarily associated with mitochondrial function, lipid and glucose metabolism, highlighting their relevance in lipid regulatory mechanisms. A comprehensive overview of the mapped genes and associated pathways across all datasets is presented in Table 3.

In PHH, *BPGM* and *NDUFA1* are associated with glucose and mitochondrial metabolism, respectively, while *CYP1A1* is linked to lipid metabolism. For PRH data, *Mgll* and *Plin2* are associated with lipid metabolism, with *Plin2* additionally linked to lipid droplet regulation. In the RLT dataset, *Aldh1b1* and *Gck* are linked to lipid metabolism, while *Atox1* and *Ech1* are involved in mitochondrial pathways.

Furthermore, comparison of the top 100 SVM genes with the steatosis adverse outcome pathway (AOP) gene list (Verhoeven et al., 2024), using the method described by Ladeira et al. (2025),

revealed six overlapping genes: *BAX* and *CYPA1A* in PHH, *HSPA4*, *MGLL* and *PLIN2* in PRH and *GOT1* in RLT.

Table 3. Mapping of the Top 100 SVM genes to the LiverLipidPM

Datasets	Genes	Processes	General Function
PHH	*CYP1A1*	Lipid Metabolism	Fatty acid metabolism
	BPGM	Glucose Metabolism	Steroid metabolism
	NDUFA1	Lipid Metabolism	Electron transport
PRH	*Mgll*	Lipid Metabolism	Lipid hydrolysis/Lipid signaling
	Plin2	Lipid droplet	Lipid storage
	Atp7a	Mitochondrial Metabolism Pathway	Copper transport
	Elovl2	Cholesterol Biosynthesis Pathway	Fatty acid/Lipid elongation
RLT	*Eci1*	Mitochondrial Metabolism Pathway	Fatty acid oxidation
	Aldh1b1	Glucose Metabolism	Aldehyde detoxification
	Gck	Glucose Metabolism and Insulin Signaling	Glucose phosphorylation
	Got1	Glucose Metabolism	Amino acid metabolism
	Slc2a1	Glucose Metabolism	Glucose Transport
	Atox1	Mitochondrial Metabolism Pathway	Copper transport
	Ech1	Mitochondrial Metabolism Pathway	Fatty acid/Lipid isomerization
	Timm10	Mitochondrial Metabolism Pathway	Protein translocation

Gene Annotation and Disease Ontology Analysis.

In parallel with the mapping to the LiverLipidPM, disease annotation analysis on the top 100 SVM-derived genes per dataset was performed to evaluate known links to liver disease and liver metabolic dysfunction. A total of 24 genes in PHH, and 29 genes each in PRH and RLT, were mapped to liver-related metabolic disorders, indicating substantial alignment between predictive gene signatures and established liver pathologies (Supplementary Table 3). Commonly enriched DO terms across all three datasets include hepatocellular carcinoma (HCC), experimental liver cirrhosis (ELC), chemical and drug induced liver injury (CDILI) and steatotic liver disease (SLD), reflecting overlap with known toxicant-induced liver conditions.

In PHH, several genes showed direct links to metabolic outcomes, including annotations to CDILI, HCC, SLD, and ELC. For instance, *CYP1A1*, *ANXA6*, *BAX* are associated with both CDILI and HCC, while *PHF6* is exclusively linked to HCC. Additional genes annotated to CDILI included *SOD2* and *IMMT*, with *SOD2* also annotated under SLD. Genes such as *CTSZ* and *IL17RC* are linked to ELC.

In PRH, several genes are annotated to ELC, including *Fgfr1*, *Plin2*, *Mme* and, *Cyp17a1*, with *Fgfr1* and *Plin2* also linked to SLD. *Fgfr1*, *Cyp17a1*, and *Mme* additionally show associations with HCC. Other HCC-related genes include *Dhfr*, which is also annotated to CDILI.

In RLT, several genes are linked to ELC, CDILI, SLD and HCC. *Ada*, *Got1*, *Dhrs7*, *Nqo1*, and *Dhfr* are associated with both ELC and CDILI, with *Dhrs7* and *Nqo1* also linked to SLD, and *Dhfr* additionally annotated to HCC. Interestingly, despite overlapping disease annotations, gene-level concordance across datasets was minimal, with only *Dhfr* and *Pskh1* shared between PRH and RLT datasets. A complete list of genes and their associated key liver-related terms is provided in Supplementary Table 3. Additional genes across all three datasets are also annotated to broader liver or systemic disorders including metabolic syndromes, hepatomegaly, Progressive Familial intrahepatic Cholestasis, and dyslipidemia.

To assess whether top-ranked genes not mapped to disease terms or lipid pathways were still functionally relevant, enrichment analysis was performed. The results revealed associations with core biological processes, including apoptosis, cellular stress response and spliceosome processing. In PHH, genes such as *NONO* and *PRPF3* contribute to these functions. In PRH, genes including *Atg10* and *Wipi2* are linked to autophagy, phospholipid metabolism, and lipoprotein

biosynthesis, mechanisms essential for maintaining hepatic lipid balance and cellular integrity. Supplementary Figures 3A-C and Supplementary Table 4 provide the full enrichment results.

In RLT, genes including *Fpgs*, *Hoga1*, and *Acot4* are associated to dicarboxylic acid metabolism, while broader metabolic processes are reflected by *Dera* and *Bco1*, both involved in fat-soluble vitamin metabolism. Additionally, *Pik3r5*, a gene implicated in Insulin Signaling and metabolic regulation, was identified. Although no enrichment results met the adjusted significance threshold ($q < 0.05$), numerous terms remained significant at nominal p-value threshold ($p < 0.05$), likely due to small number of input genes limiting statistical power for multiple testing correction.

Functional profiling of unannotated genes

To address the biological relevance of predictive genes that remained uncharacterized in prior annotation or enrichment analysis, a bottom-up literature mining approach was applied. This approach sought novel or less-characterized genes implicated in lipid metabolism, cellular stress, inflammation, and broader liver pathologies, such as fibrosis, cirrhosis and cholestasis, that may contribute indirectly to steatosis progression. Keyword-based profiling of PubMed title/abstracts and NCBI Gene summaries were performed to highlight under-studied genes and prioritize them for follow-up validation. A total of 66 genes in PHH, 52 in PRH, and 47 in RLT were screened (Supplementary Table 5) to explore potential mechanistic roles in liver dysfunction. Among these, 48 genes in PHH, 47 in PRH, and 37 in RLT returned hits, based on keyword-driven evidence from PubMed (TIAB) queries. Biological terms associated with hit genes included apoptosis, inflammation, oxidative stress and hepatocellular carcinoma, which appeared across multiple datasets. Dataset-specific enrichment included steatohepatitis and cholestasis in PHH, cholesterol metabolism and mitophagy in PRH and steatosis in RLT. Notably, several genes in each dataset are linked to both early stress responses and advanced liver pathologies. For example, *PIM1* in PHH is associated with apoptosis and fibrosis, while *AP3M2*, *CHP1*, *GLMN* and *LYRM2* are linked to hepatocellular carcinoma.

In PRH, *Igf2bp2* exhibited broad relevance, being linked to ferroptosis, fibrosis, inflammation and insulin resistance, suggesting potential involvement in both early stress responses and progressive pathological features. In RLT, *Cyp4a1* is involved in lipid metabolism, oxidative stress and steatosis, whereas *Ccz1* is connected to lipid droplet regulation, a hallmark of steatotic progression. NCBI Gene summaries provided additional context for two genes in PHH and 1 in RLT,

highlighting roles in transport, mitochondrial function, apoptosis and metabolic regulation. Altogether, 132 of the 165 predictive genes, 48 in PHH, 47 in PRH, and 37 in RLT, showed functional annotations related to cellular responses, metabolic regulation and liver disease progression (Supplementary Figures 4A-C; Supplementary Table 5). In contrast, 16 genes in PHH, 5 in PRH and 9 in RLT lacked PubMed or NCBI Gene annotations, reflecting a residual knowledge gap and highlighting candidates for future investigation (https://github.com/TGX-UM/ASPIS_Omics_ML/tree/main/Literature_Mining_Results).

4. Discussion

Machine learning is transforming predictive toxicology, aligning with the global shift towards New Approach Methodologies (NAMs), which prioritize ethical, human-relevant, and mechanistically driven strategies for next-generation risk assessment (NGRA) (Aguayo et al., 2021; Jiang et al., 2023; Lin & Chou, 2022; O'Donovan et al., 2023). In this study, we evaluated the performance of five supervised ML algorithms, ENLR, KNN, RF, SVM, and XGBoost, trained on the Open TG-GATEs transcriptomic data (Igarashi et al., 2015) to accurately predict the steatogenic potential of chemical compounds. These classifiers were selected to represent a diverse array of learning paradigms, encompassing penalized linear models (ENLR), distance-based methods (KNN), ensemble tree learners (RF, XGBoost), and margin-based classifiers (SVM). This diversity facilitates a comprehensive comparison across fundamentally different algorithmic approaches, ensuring robust evaluation under identical training conditions. Importantly, we prioritized models that are computationally efficient and interpretable, with the goal of developing tools that can run on standard laboratory workstations. Among these algorithms, SVM consistently achieved the highest performance, with ROC AUC values exceeding 0.8 across all three datasets (Figures 2A-C). These values significantly surpass the 0.5 baseline, indicating that the SVM model effectively captures consistent and robust transcriptomic patterns associated with steatogenicity. Consequently, SVM was selected for downstream prediction tasks and feature interpretation across PHH, PRH, and RLT datasets.

The consistently superior performance of the SVM model across all datasets can be attributed to its strong suitability for high-dimensional, low-sample size data, typical of toxicogenomic experiments. Microarray data live in the classic “largep, smalln” regime where SVMs have

historically excelled (Guyon et al., 2002; Statnikov et al., 2005). In such contexts, where the number of features (genes) greatly exceeds the number of samples, SVMs are particularly effective because they construct an optimal hyperplane that maximizes the margin between classes, thereby improving generalization and reducing overfitting (Cristianini & Shawe-Taylor, 2000; Hastie et al., 2009). Margin-based learning is especially effective in contexts where class boundaries are diffuse, as frequently observed in early transcriptomic perturbations from chemical exposure. Conversely, KNN relies on distance metrics that degrade in high dimensions due to the curse of dimensionality, where all points become equidistant, reducing discriminatory power (Biau & Devroye, 2015). Tree-based models such as RF and XGBoost require sufficient sample sizes to build informative decision nodes. In high-dimensional datasets with limited samples, tree-based models may rely on unstable or weakly informative splits, increasing the risk of overfitting and reducing model generalizability (Ancuceanu et al., 2020; Guo et al., 2023). ENLR, while interpretable and effective for feature selection in high-dimensional settings, assumes linear additive effects and may therefore miss nonlinear or higher-order gene-gene interactions, features commonly observed in toxicogenomic responses (Alexander-Dann et al., 2018; Zou & Hastie, 2005).

Interestingly, the SVM model achieved near-identical ROC AUC values in PRH (0.975) and RLT (0.966), reflecting consistently high predictive performance across both *in vitro* and *in vivo* rat models. This finding challenges the common assumption that *in vitro* systems lack the complexity to recapitulate *in vivo* outcomes (Williams et al., 2013). Our results suggest that PRH retain sufficient transcriptomic resolution to distinguish steatogenic from non-steatogenic exposures with high fidelity. A plausible explanation lies in the acute exposure design of the TG-GATEs dataset used, which captures early transcriptional responses to chemical insult. At these early time points, cell-autonomous stress responses may dominate, enabling *in vitro* models to mirror key *in vivo* transcriptional patterns (Liu et al., 2019).

Building on this, it is particularly striking that SVM models trained on rat data outperformed the PHH trained model despite identical data processing steps. This performance gap may be explained by a combination of biological and technical factors. First, inter-individual variability in the PHH dataset likely impacted the predictive performance. Human hepatocytes were sourced from multiple donors (Igarashi et al., 2015), introducing variation in gene expression and

metabolic activity. Donor differences in age, gender, genetic background, and health status contribute to this heterogeneity, which, while biologically meaningful and essential for human-relevant modelling, poses challenges for ML models trained on relatively small datasets. For instance, polymorphisms in *CYP3A4*, a key phase one metabolic enzyme, can result in 15–30% variation in drug metabolism between individuals (Godoy et al., 2013). In contrast, the rat data were generated from genetically homogeneous male Sprague-Dawley rats maintained under controlled experimental conditions (Igarashi et al., 2015). This genetic and environmental uniformity likely reduced background variability and improved the model's ability to detect consistent steatogenic transcriptional patterns. As such, the lower performance in PHH trained model may reflect limitations in sample scale rather than biological relevance, emphasizing the need for larger and more diverse human datasets to support robust predictive modeling.

In addition to donor variability, fundamental physiological differences between human and rat liver systems likely contribute to the observed performance gap. Species-specific differences in absorption, distribution, metabolism, and excretion (ADME) can substantially alter compound disposition and transcriptomic responses. Human and rat hepatocytes differ in the composition and basal expression levels of cytochrome P450 (CYP) isoforms. For example, the *CYP2C*, *CYP2E1*, and *CYP3A* subfamilies vary across species in both substrate specificity and inducibility (Hammer et al., 2021; Martignoni et al., 2006). Quantitative studies have shown that *CYP3A* enzymes are more abundant and strongly induced in rat liver under activation by the pregnane-X-receptor (*PXR*), whereas *CYP3A4* induction in human hepatocytes is more variable and donor-dependent (Hammer et al., 2021; Lawrence et al., 2001). Lipid metabolism is another key area of divergence. Rats exhibit higher peroxisomal β-oxidation capacity and greater sensitivity to peroxisome proliferator-activated receptor alpha (*PPARα*) agonists, supporting more efficient lipid clearance and a robust steatogenic transcriptomic profile (Lawrence et al., 2001; Tahri-Joutey et al., 2021). In contrast, while overall PPARα expression is comparable between human and rat livers, humans also express a truncated, low-abundance isoform with reduced functionality. This may dampen fatty acid oxidation and contribute to subtler transcriptomic shifts in human hepatocytes, especially under steatotic or inflammatory conditions such as non-alcoholic steatohepatitis (NASH). These metabolic distinctions likely result in more consistent and learnable steatogenic signatures in rat-

derived data, helping to explain the superior performance of the SVM model trained on rat samples compared to PHH.

Structural differences in the datasets, including class distribution, may have further contributed to the superior performance of the rat-trained SVM models. To enhance statistical robustness, all datasets underwent 1,000 bootstrap resampling iterations, and SMOTE was applied to balance class distributions prior to model training. However, SVM models trained on rat data still outperformed the PHH-trained model, suggesting that statistical harmonization alone could not overcome dataset-specific limitations. A key technical factor is sample size and class composition. The rat datasets, particularly RLT, included more replicates per condition and a more favourable initial class balance. In PRH and RLT, the majority class was non-steatogenic, while the PHH dataset was initially skewed toward steatogenic samples (Table 1). Although SMOTE corrected for class imbalance before training, it does not address variability in within-class consistency or signal strength. These structural differences likely contributed to the pronounced specificity gap observed between models: the PHH-trained SVM achieved 0.619, whereas PRH and RLT models exceeded 0.85 (Table 2). Across all datasets, SVM models showed high sensitivity (>0.80), confirming strong ability to detect steatogenic compounds. Yet the PHH-trained model lagged in overall performance (Table 2). It yielded a J index of 0.515 and a MCC of 0.545, reflecting moderate but reliable predictive capacity. In contrast, both PRH and RLT trained models produced J index and MCC values exceeding 0.80, indicating excellent classification performance. Notably, a J index > 0.5 reflects a favourable balance between sensitivity and specificity, while an MCC > 0.5 confirms robustness, even under class imbalance conditions.

To evaluate model applicability, the final SVM classifier trained on PHH data was deployed on the full TG-GATEs PHH dataset, which included samples treated with compounds of unknown steatogenic status. A classification probability threshold > 0.7 was applied to ensure high confidence in predictions. This threshold reflects the biological need for reliability in toxicological assessments, where both false positives and false negatives can carry significant consequences. Statistically, it reduces ambiguity by excluding samples near the decision boundary (i.e., around 0.5 probability). A notable example is carbon tetrachloride (CCl_4), which, despite being labeled as steatogenic in the ASPIS reference list and included as such during training, was predicted as non-steatogenic by the PHH SVM model. This discrepancy suggests that the model may have learned

robust transcriptomic signatures of steatogenicity, rather than merely memorizing compound labels or relying on superficial patterns. Alternatively, it is possible that CCl_4 does not induce steatosis under the specific experimental conditions captured in TG-GATEs, indicating that its prior classification may require re-evaluation. Indeed, CCl_4 is widely used to experimentally induce liver injury, particularly inflammation and fibrosis (Dong et al., 2016; Masubuchi & Ihara, 2022; Pilling et al., 2024), whereas steatosis often arises later as a secondary effect or typically depends on its combination with high-fat diets, which themselves promote lipid accumulation.

Several model predictions were supported by existing toxicological evidence, reinforcing the biological plausibility of the SVM classifier's output. Two notable examples include omeprazole and nitrofurantoin, both predicted as steatogenic with high confidence (probability ≥ 0.9) at high doses and 24-hour exposure. Omeprazole, a commonly used proton pump inhibitor, has been linked to an increased risk of non-alcoholic fatty liver disease (NAFLD). A nationwide cohort study reported a higher incidence of fatty liver among long-term PPI users compared to non-users (Pyo et al., 2021). Nitrofurantoin (NFT), primarily known for its antimicrobial properties, has been implicated in DILI. A recent analysis revealed that prolonged nitrofurantoin use was associated with severe liver injury, including fibrosis and cirrhosis, while even short-term exposure could result in hepatocellular damage (Chalasani et al., 2023). Notably, the RF model also uniquely predicted NFT as steatogenic at a low dose (8 h) and high dose (2 h), suggesting early transcriptional changes may be detectable even under shorter or lower-exposure conditions.

While several model predictions were supported by existing toxicological evidence, validation using histological endpoints from TG-GATEs presents key limitations. Although the *in vivo* dataset includes annotated liver pathology, two factors constrain its utility for confirming our steatogenic predictions. First, the highest dose tested corresponds to the lowest observed adverse effect level (LOAEL), which is unlikely to induce lipid accumulation within the 24 h exposure window. Second, our objective was not to detect overt steatosis but to classify molecular initiating events (MIEs) that may lead to steatosis under repeated or chronic exposure.

Selection of the Open TG-GATEs database as the primary data source was guided by both scientific merit and practical constraints. TG-GATEs remains one of the most comprehensive toxicogenomic resources available, containing standardized microarray-based gene expression

profiles for 170 compounds across human and rat liver models (Igarashi et al., 2015). Despite relying on legacy microarray technology, TG-GATEs is still considered a benchmark omics dataset due to its harmonized study design, consistent dosing protocols, and inclusion of both *in vitro* and *in vivo* systems. We explored alternative data sources, including DrugMatrix (Ganter et al., 2005) and additional public datasets (Supplementary Table 1). However, these had shortcomings we could not compromise. For example, the DrugMatrix contains disproportionately more control samples than exposure conditions. This imbalance poses challenges for supervised learning, as the model may become biased toward the majority class, reducing its ability to learn discriminative patterns associated with chemical-induced responses. In addition to dataset imbalance, compound coverage was also limited. Thirty-six of the forty ASPIS-listed chemicals were data-poor, lacking transcriptomic profiles under controlled experimental conditions.

We also attempted to incorporate RNA-sequencing data from the Hepatic and Cardiac Toxicity Systems modeling (HeCaToS) project (M. Verheijen et al., 2022), but differences in platform, dose selection, and time points precluded integration. While our results using TG-GATEs are promising and demonstrate clear potential for ML in mechanism-driven NGRA, key feasibility barriers remain. Microarrays are no longer the standard in transcriptomics, and the relatively small sample sizes in TG-GATEs limit model generalizability. Ideally, robust ML models require more samples than features, a criterion not met by current toxicogenomic datasets. Realizing the full potential of ML in NGRA will require new, large-scale, high-quality transcriptomic datasets spanning a diverse range of compounds, doses, and time points. These resources will be essential for building scalable, reproducible, and regulatory-ready models for next-generation chemical safety assessment.

Beyond evaluating model performance, we investigated whether the top 100 predictive genes from the best-performing model (SVM) align with known or novel mechanisms involved in DIHS. Mapping these top-ranked genes onto the LiverLipidPM (Ladeira et al., 2024) revealed key mechanistic processes including mitochondrial function, lipid droplet regulation, and energy metabolism, all of which are central to lipid metabolism. These molecular signatures were consistently observed across PHH, PRH, and RLT datasets.

Among the genes identified in the PHH trained model, *CYP1A1* emerged as a notable candidate gene associated with NAFLD. Huang et al. (2018) demonstrated that *CYP1A1* is

upregulated in oleic acid–treated HepG2 cells, linking its increased expression to lipid peroxidation in hepatocytes. Alongside *CYP1A1*, *BPGM* and *NDUFA1* also mapped onto the LiverLipidPM, indicating their roles in glucose and lipid metabolism. *NDUFA1* encodes a subunit of mitochondrial complex I, a critical component of the electron transport chain responsible for oxidative phosphorylation and ATP production. Dysfunction of complex I leads to electron leakage and reactive oxygen species generation, causing oxidative stress and mitochondrial impairment, both of which are implicated in DILI and metabolic diseases such as NAFLD (Mihajlovic & Vinken, 2022; Petrosillo et al., 2007). Although direct evidence linking *NDUFA1* to liver dysfunction is limited, a recent study in a maternal malnutrition pig model showed significant changes in *NDUFA1* expression in metabolically stressed liver tissue (Wang et al., 2023), suggesting its potential involvement in hepatic pathophysiology. Similarly, *BPGM* upregulation has been associated with disrupted amino acid and lipid metabolism during liver disease progression (Cai et al., 2020).

The SVM model trained on the PRH dataset identified several key genes associated with lipid metabolism, including *Mgll*, *Plin2*, *Atp7a*, and *Elov2*. *Plin2* plays a crucial role in lipid droplet dynamics and regulates hepatic lipid homeostasis by controlling lipolytic enzymes (Gluchowski et al., 2017; Tsai et al., 2017). A study showed that mice lacking *Plin2* exhibit a 60% reduction in triglyceride content in hepatocytes (Tsai et al., 2017). Mechanistically, *PLIN2* inhibits lipolytic enzymes, including PNPLA2 and *MGLL* (McIntosh et al., 2012). Elevated levels of *Plin2* in fatty liver tissue and its reduction in hepatocytes alleviate steatosis, underscoring its significant role in the development of steatosis (Imai et al., 2007). Similarly, *Mgll* deficiency has been associated with protection against hepatic steatosis due to enhanced lipid storage in adipose tissue and reduced lipid absorption in the intestine (Tardelli et al., 2019). *Atp7a*, involved in copper homeostasis, is critical for lipid metabolism (Blades et al., 2021). Disruption of *Atp7a* may lead to intracellular copper imbalance and increased superoxide levels leading to oxidative stress, which has been linked to metabolic dysfunctions including NAFLD (Morrell et al., 2017; Sudhahar et al., 2013). Moreover, copper deficiency contributes to the development of NAFLD (Aigner et al., 2010).

In the RLT dataset, the SVM model identified key genes, including *Eci1*, *Gck*, and *Aldh1b1*, which were mapped to the LiverLipidPM and linked to mitochondrial metabolism, insulin

signalling, and glucose metabolism. Increased *Gck* expression in the liver is associated with lipogenesis and fatty liver development (Peter et al., 2011). *Aldh1b1* has been reported to be downregulated in human NASH at the protein level (Li et al., 2018), indicating disruptions in metabolic pathways.

While only a limited number of top-ranked genes from each dataset mapped to the LiverLipidPM, the low number of mapped genes likely results from the map's focus on well-established lipid metabolism pathways, potentially overlooking genes that contribute to hepatic lipid dysregulation through indirect or less-characterized mechanisms. Additionally, the LiverLipidPM is primarily hepatocyte-centric, missing mechanisms from other liver cell populations, which may be relevant in the multicellular liver environment. Despite these limitations, the LiverLipidPM is continually updated, incorporating new mechanisms, including those identified in this study, thereby expanding its coverage. Nevertheless, six of the top 100 SVM identified genes, *BAX* and *CYP1A1* (PHH), *Hspa4*, *Mgll*, and *Plin2* (PRH), and *Got1* (RLT), mapped to key events in the AOP network described by Verhoeven et al. (2024). Specifically, these genes are associated with apoptosis, endoplasmic reticulum (ER) stress, and lipid droplet regulation, highlighting the model's ability to capture biological mechanisms aligned with established AOP frameworks and hepatic lipid physiology

Comparing the top 100 SVM-predictive genes across datasets revealed no overlap between the PHH and rat datasets, but *Dhfr* and *Pskh1* were common between the rat datasets. *Dhfr* plays a key role in folate metabolism, and abnormal folate levels have been linked to NAFLD, NASH, and fibrosis (Yang et al., 2024). *Pskh1*, a serine kinase, is associated with paediatric intrahepatic cholestasis (Maddirevula et al., 2024), suggesting that further investigation may provide valuable insights in understanding its role in liver pathophysiology.

Additionally, DO based gene annotation analysis provided broader context by revealing consistent associations with key liver conditions across all three datasets. In PHH, genes such as *CYP1A1* and *BAX* were linked to both hepatocellular carcinoma and DILI. *BAX*, a key regulator of apoptosis, promotes mitochondrial-mediated cell death, particularly in the context of DILI (Bajt et al., 2008). Other notable genes, such as *SOD2*, were also identified and linked to steatosis and DILI. *SOD2* is a mitochondrial enzyme critical for handling reactive oxygen species (Palma et al., 2020). In NAFLD, *SOD2* deficiency exacerbates mitochondrial generation of reactive oxygen species, which upregulates sterol regulatory element-binding protein 1c (*SREBP1c*), a

transcription factor driving fatty acid synthesis. The overexpression of *SREBP1c* also activates the NFκB inflammation pathway, worsening hepatic steatosis (Ma et al., 2021).

Although limited overlap was observed between PRH and RLT, consistent enrichment of liver disease-related terms across all datasets suggests a shared biological response to steatogenic compounds. The identification of genes related to liver dysfunction through both lipid pathway mapping and DO analysis further supports the model's ability to uncover expression patterns unique to steatogenicity. While only two genes overlapped between PRH and RLT, and none overlapped with PHH, each dataset captured distinct patterns of steatogenesis. Genes that lacked annotations in DO or lipid metabolism pathways were subjected to further enrichment analysis to explore their potential involvement in broader processes associated with hepatic dysfunction, including inflammation, cellular stress, and metabolic disruption. Although not statistically significant, several genes showed nominal enrichment ($p < 0.05$) across PHH, PRH, and RLT datasets, revealing involvement in processes related to apoptosis, lipid metabolism, and cellular stress regulation (Supplementary Figures 3A-C and Supplementary Table 4).

Genes that remained uncharacterized in prior annotation and enrichment analysis were further investigated using a bottom-up literature mining approach. This method aimed to identify novel or less-characterized genes with potential relevance to early stress responses and progressive liver pathologies. Out of the 165 predictive genes lacking prior annotations, 132 (48 in PHH, 47 in PRH, and 37 in RLT) were linked to processes related to oxidative stress, apoptosis, and lipid metabolism. For example, *PIM1* in PHH was identified and is linked to apoptosis (Gu et al., 2009). *Igf2bp2*, identified in PRH, is an m6A reader gene and has been linked to apoptosis, fatty liver, fibrosis, and HCC (Pu et al., 2020; Wang et al., 2021; Xu et al., 2022).

In RLT, genes such as *Cyp4a1* and *Ccz1* were identified; *Ccz1* has been reported to be a mediator in autophagy (Dong et al., 2015). These findings demonstrate that literature mining can uncover mechanistically relevant genes across all three datasets, spanning early stress responses to advanced liver pathologies. Furthermore, this approach has the potential to reveal model-specific signatures, distinguishing human, rat, and other species, and with further analysis, integrate these findings into species-relevant ontologies.

Together, these findings show that the top-ranked features from our SVM model not only recapitulate well-established genes linked to steatosis and other liver-related disorders but also highlight understudied genes that warrant further investigation as potential mediators of hepatic lipid dysfunction.

Conclusion

We present a machine learning framework for predicting DIHS using transcriptomic profiles from human and rat liver models. Among five classifiers, evaluated SVM consistently outperformed others, achieving high predictive accuracy and identifying biologically relevant gene features across *in vitro* and *in vivo* systems. Enrichment analyses confirmed that top-ranked genes were mechanistically linked to lipid metabolism, mitochondrial function, and hepatic stress pathways, hallmarks of steatogenic response.

Despite strong performance, the study is constrained by the limited sample size and legacy microarray platform of the TG-GATEs dataset. These factors restrict generalizability and underscore the need for large-scale, human-relevant transcriptomic resources. Future efforts should prioritize the generation of donor-diverse, high-quality transcriptomic datasets, ideally using RNA-sequencing or similarly scalable and well-annotated platforms. Such resources will be essential for training more robust and translational ML models. In parallel, active engagement with regulatory stakeholders will be critical to define validation criteria and facilitate the integration of omics-based models into next-generation risk

Our findings illustrate the potential of interpretable, data-driven models to detect early molecular signatures predictive of liver injury. As regulatory agencies shift toward mechanism-based safety assessment, this approach offers a scalable, biologically grounded alternative to traditional animal testing, supporting the broader adoption of NAMs in NGRA. To enable real-world implementation, active engagement with regulatory stakeholders will be essential to define validation criteria and ensure that omics-based ML models can be appropriately evaluated and integrated into regulatory frameworks.

Data Availability

All data used for training the machine learning models is available at https://github.com/TGX-UM/ASPIS_Omics_ML/tree/main/Data

Code availability

The code used in this study can be found here: https://github.com/TGX-UM/ASPIS_Omics_ML/tree/main/Scripts

Author contributions: CRediT

Brian Bwanya: Conceptualization, Formal Analysis, Investigation, Methodology, Software, Visualization, Writing - Original draft preparation, Writing - Review and editing.
Saad Lodhi: Conceptualization, Formal Analysis, Investigation, Methodology, Software, Visualization, Writing - Original draft preparation, Writing - Review and editing.
Theo M de Kok: Funding acquisition, Resources.
Luiz Ladeira: Formal Analysis, Methodology, Visualization, Writing - Review and editing.
Marcha CT Verheijen: Supervision, Visualization, Writing - Review and editing.
Danyel GJ Jennen: Funding acquisition, Resources, Supervision, Visualization, Writing – Review and editing.
Florian Caiment: Conceptualization, Data curation, Methodology, Project administration, Resources, Software, Supervision, Validation, Funding acquisition, Visualization, Writing - Review and editing.

Funding

This work was performed in the context of the ONTOX project (https://ontox-project.eu/) that has received funding from the European Union's Horizon 2020 Research and Innovation programme under grant agreement No 963845. ONTOX is part of the ASPIS project cluster (https://aspis-cluster.eu/). Additional support for this work was provided by the Virtual Human Platform for Safety Assessment (VHP4Safety) project, which is funded by the Netherlands Research Council

(NWO) Netherlands Research Agenda: Research on Routes by Consortia (NWA-ORC 1292.19.272).

References

Aguayo, A., Brunak, S., & Taboureau, O. (2021). Extrapolation of drug induced liver injury responses from cancer cell lines using machine learning approaches. *Computational Toxicology*, *17*, 100147. https://doi.org/10.1016/j.comtox.2020.100147

Aigner, E., Strasser, M., Haufe, H., Sonnweber, T., Hohla, F., Stadlmayr, A., Solioz, M., Tilg, H., Patsch, W., Weiss, G., Stickel, F., & Datz, C. (2010). A role for low hepatic copper concentrations in nonalcoholic Fatty liver disease. *Am J Gastroenterol*, *105*(9), 1978-1985. https://doi.org/10.1038/ajg.2010.170

Alexander-Dann, B., Pruteanu, L. L., Oerton, E., Sharma, N., Berindan-Neagoe, I., Modos, D., & Bender, A. (2018). Developments in toxicogenomics: understanding and predicting compound-induced toxicity from gene expression data. *Mol Omics*, *14*(4), 218-236. https://doi.org/10.1039/c8mo00042e

Amacher, D. E., & Chalasani, N. (2014). Drug-induced hepatic steatosis. *Semin Liver Dis*, *34*(2), 205-214. https://doi.org/10.1055/s-0034-1375960

Ancuceanu, R., Hovanet, M. V., Anghel, A. I., Furtunescu, F., Neagu, M., Constantin, C., & Dinu, M. (2020). Computational Models Using Multiple Machine Learning Algorithms for Predicting Drug Hepatotoxicity with the DILIrank Dataset. *Int J Mol Sci*, *21*(6). https://doi.org/10.3390/ijms21062114

Atkins, J. T., George, G. C., Hess, K., Marcelo-Lewis, K. L., Yuan, Y., Borthakur, G., Khozin, S., LoRusso, P., & Hong, D. S. (2020). Pre-clinical animal models are poor predictors of human toxicities in phase 1 oncology clinical trials. *Br J Cancer*, *123*(10), 1496-1501. https://doi.org/10.1038/s41416-020-01033-x

Bajt, M. L., Farhood, A., Lemasters, J. J., & Jaeschke, H. (2008). Mitochondrial bax translocation accelerates DNA fragmentation and cell necrosis in a murine model of acetaminophen hepatotoxicity. *J Pharmacol Exp Ther*, *324*(1), 8-14. https://doi.org/10.1124/jpet.107.129445

Bastian, F. B., Cammarata, A. B., Carsanaro, S., Detering, H., Huang, W. T., Joye, S., Niknejad, A., Nyamari, M., Mendes de Farias, T., Moretti, S., Tzivanopoulou, M., Wollbrett, J., & Robinson-Rechavi, M. (2025). Bgee in 2024: focus on curated single-cell RNA-seq datasets, and query tools. *Nucleic Acids Res*, *53*(D1), D878-D885. https://doi.org/10.1093/nar/gkae1118

Biau, G. r., & Devroye, L. (2015). *Lectures on the Nearest Neighbor Method* (1st ed.). Springer International Publishing : Imprint: Springer,. https://doi.org/10.1007/978-3-319-25388-6

Blades, B., Ayton, S., Hung, Y. H., Bush, A. I., & La Fontaine, S. (2021). Copper and lipid metabolism: A reciprocal relationship. *Biochim Biophys Acta Gen Subj*, *1865*(11), 129979. https://doi.org/10.1016/j.bbagen.2021.129979

Browne, P., Paul Friedman, K., Boekelheide, K., & Thomas, R. S. (2024). Adverse effects in traditional and alternative toxicity tests. *Regul Toxicol Pharmacol*, *148*, 105579. https://doi.org/10.1016/j.yrtph.2024.105579

Cai, F. F., Song, Y. N., Lu, Y. Y., Zhang, Y., Hu, Y. Y., & Su, S. B. (2020). Analysis of plasma metabolic profile, characteristics and enzymes in the progression from chronic hepatitis B to hepatocellular carcinoma. *Aging (Albany NY)*, *12*(14), 14949-14965. https://doi.org/10.18632/aging.103554

Caloni, F., De Angelis, I., & Hartung, T. (2022). Replacement of animal testing by integrated approaches to testing and assessment (IATA): a call for in vivitrosi. *Arch Toxicol*, *96*(7), 1935-1950. https://doi.org/10.1007/s00204-022-03299-x

Cataldi, M., Citro, V., Resnati, C., Manco, F., & Tarantino, G. (2021). New Avenues for Treatment and Prevention of Drug-Induced Steatosis and Steatohepatitis: Much More Than Antioxidants. *Adv Ther*, *38*(5), 2094-2113. https://doi.org/10.1007/s12325-021-01669-y

Chalasani, N., Li, Y. J., Dellinger, A., Navarro, V., Bonkovsky, H., Fontana, R. J., Gu, J., Barnhart, H., Phillips, E., Lammert, C., Schwantes-An, T. H., Nicoletti, P., Kleiner, D. E., Hoofnagle, J. H., & Drug Induced Liver Injury, N. (2023). Clinical features, outcomes, and HLA risk factors associated with nitrofurantoin-induced liver injury. *J Hepatol*, *78*(2), 293-300. https://doi.org/10.1016/j.jhep.2022.09.010

Cristianini, N., & Shawe-Taylor, J. (2000). *An Introduction to Support Vector Machines and Other Kernel-Based Learning Methods*. Cambridge University Press.

Dai, M., Wang, P., Boyd, A. D., Kostov, G., Athey, B., Jones, E. G., Bunney, W. E., Myers, R. M., Speed, T. P., Akil, H., Watson, S. J., & Meng, F. (2005). Evolving gene/transcript definitions significantly alter the interpretation of GeneChip data. *Nucleic Acids Res*, *33*(20), e175. https://doi.org/10.1093/nar/gni179

Dirven, H., Vist, G. E., Bandhakavi, S., Mehta, J., Fitch, S. E., Pound, P., Ram, R., Kincaid, B., Leenaars, C. H. C., Chen, M., Wright, R. A., & Tsaioun, K. (2021). Performance of preclinical models in predicting drug-induced liver injury in humans: a systematic review. *Sci Rep*, *11*(1), 6403. https://doi.org/10.1038/s41598-021-85708-2

Dong, S., Chen, Q. L., Song, Y. N., Sun, Y., Wei, B., Li, X. Y., Hu, Y. Y., Liu, P., & Su, S. B. (2016). Mechanisms of CCl4-induced liver fibrosis with combined transcriptomic and proteomic analysis. *J Toxicol Sci*, *41*(4), 561-572. https://doi.org/10.2131/jts.41.561

Dong, Y., Yu, Q., Chen, Y., Xu, N., Zhao, Q., Jia, C., Zhang, B., Zhang, K., Zhang, B., Xing, L., & Li, M. (2015). The Ccz1 mediates the autophagic clearance of damaged mitochondria in response to oxidative stress in Candida albicans. *Int J Biochem Cell Biol*, *69*, 41-51. https://doi.org/10.1016/j.biocel.2015.10.002

Durinck, S., Spellman, P. T., Birney, E., & Huber, W. (2009). Mapping identifiers for the integration of genomic datasets with the R/Bioconductor package biomaRt. *Nat Protoc*, *4*(8), 1184-1191. https://doi.org/10.1038/nprot.2009.97

Evangelista, J. E., Xie, Z., Marino, G. B., Nguyen, N., Clarke, D. J. B., & Ma'ayan, A. (2023). Enrichr-KG: bridging enrichment analysis across multiple libraries. *Nucleic Acids Res*, *51*(W1), W168-W179. https://doi.org/10.1093/nar/gkad393

Ganter, B., Tugendreich, S., Pearson, C. I., Ayanoglu, E., Baumhueter, S., Bostian, K. A., Brady, L., Browne, L. J., Calvin, J. T., Day, G. J., Breckenridge, N., Dunlea, S., Eynon, B. P., Furness, L. M., Ferng, J., Fielden, M. R., Fujimoto, S. Y., Gong, L., Hu, C., . . . Jarnagin, K. (2005). Development of a large-scale chemogenomics database to improve drug candidate selection and to understand mechanisms of chemical toxicity and action. *J Biotechnol*, *119*(3), 219-244. https://doi.org/10.1016/j.jbiotec.2005.03.022

Gautier, L., Cope, L., Bolstad, B. M., & Irizarry, R. A. (2004). affy--analysis of Affymetrix GeneChip data at the probe level. *Bioinformatics*, *20*(3), 307-315. https://doi.org/10.1093/bioinformatics/btg405

Gawron, P., Smula, E., Schneider, R., & Ostaszewski, M. (2023). Exploration and comparison of molecular mechanisms across diseases using MINERVA Net. *Protein Sci*, *32*(2), e4565. https://doi.org/10.1002/pro.4565

Gluchowski, N. L., Becuwe, M., Walther, T. C., & Farese, R. V., Jr. (2017). Lipid droplets and liver disease: from basic biology to clinical implications. *Nat Rev Gastroenterol Hepatol*, *14*(6), 343-355. https://doi.org/10.1038/nrgastro.2017.32

Godoy, P., Hewitt, N. J., Albrecht, U., Andersen, M. E., Ansari, N., Bhattacharya, S., Bode, J. G., Bolleyn, J., Borner, C., Bottger, J., Braeuning, A., Budinsky, R. A., Burkhardt, B., Cameron, N. R., Camussi, G., Cho, C. S., Choi, Y. J., Craig Rowlands, J., Dahmen, U., . . . Hengstler, J. G. (2013). Recent advances in 2D and 3D in vitro systems using primary hepatocytes, alternative hepatocyte sources and non-parenchymal liver cells and their use in investigating mechanisms of hepatotoxicity, cell signaling and ADME. *Arch Toxicol*, *87*(8), 1315-1530. https://doi.org/10.1007/s00204-013-1078-5

Gu, J. J., Wang, Z., Reeves, R., & Magnuson, N. S. (2009). PIM1 phosphorylates and negatively regulates ASK1-mediated apoptosis. *Oncogene*, *28*(48), 4261-4271. https://doi.org/10.1038/onc.2009.276

Guo, W., Liu, J., Dong, F., Song, M., Li, Z., Khan, M. K. H., Patterson, T. A., & Hong, H. (2023). Review of machine learning and deep learning models for toxicity prediction. *Exp Biol Med (Maywood)*, *248*(21), 1952-1973. https://doi.org/10.1177/15353702231209421

Guyon, I., Weston, J., Barnhill, S., & Vapnik, V. (2002). Gene Selection for Cancer Classification using Support Vector Machines. *Machine Learning*, *46*(1), 389-422. https://doi.org/10.1023/A:1012487302797

Hammer, H., Schmidt, F., Marx-Stoelting, P., Potz, O., & Braeuning, A. (2021). Cross-species analysis of hepatic cytochrome P450 and transport protein expression. *Arch Toxicol*, *95*(1), 117-133. https://doi.org/10.1007/s00204-020-02939-4

Hastie, T., Tibshirani, R., & Friedman, J. H. (2009). *The elements of statistical learning : data mining, inference, and prediction* (2nd ed.). Springer.

Hosack, T., Damry, D., & Biswas, S. (2023). Drug-induced liver injury: a comprehensive review. *Therap Adv Gastroenterol*, *16*, 17562848231163410. https://doi.org/10.1177/17562848231163410

Huang, B., Bao, J., Cao, Y. R., Gao, H. F., & Jin, Y. (2018). Cytochrome P450 1A1 (CYP1A1) Catalyzes Lipid Peroxidation of Oleic Acid-Induced HepG2 Cells. *Biochemistry (Mosc)*, *83*(5), 595-602. https://doi.org/10.1134/S0006297918050127

Idilman, I. S., Ozdeniz, I., & Karcaaltincaba, M. (2016). Hepatic Steatosis: Etiology, Patterns, and Quantification. *Semin Ultrasound CT MR*, *37*(6), 501-510. https://doi.org/10.1053/j.sult.2016.08.003

Igarashi, Y., Nakatsu, N., Yamashita, T., Ono, A., Ohno, Y., Urushidani, T., & Yamada, H. (2015). Open TG-GATEs: a large-scale toxicogenomics database. *Nucleic Acids Res*, *43*(Database issue), D921-927. https://doi.org/10.1093/nar/gku955

Imai, Y., Varela, G. M., Jackson, M. B., Graham, M. J., Crooke, R. M., & Ahima, R. S. (2007). Reduction of hepatosteatosis and lipid levels by an adipose differentiation-related

protein antisense oligonucleotide. *Gastroenterology*, *132*(5), 1947-1954. https://doi.org/10.1053/j.gastro.2007.02.046

Jiang, J., van Ertvelde, J., Ertaylan, G., Peeters, R., Jennen, D., de Kok, T. M., & Vinken, M. (2023). Unraveling the mechanisms underlying drug-induced cholestatic liver injury: identifying key genes using machine learning techniques on human in vitro data sets. *Arch Toxicol*, *97*(11), 2969-2981. https://doi.org/10.1007/s00204-023-03583-4

Kolaric, T. O., Nincevic, V., Kuna, L., Duspara, K., Bojanic, K., Vukadin, S., Raguz-Lucic, N., Wu, G. Y., & Smolic, M. (2021). Drug-induced Fatty Liver Disease: Pathogenesis and Treatment. *J Clin Transl Hepatol*, *9*(5), 731-737. https://doi.org/10.14218/JCTH.2020.00091

Kuhn, M., & Wickham, H. (2020). *Tidymodels: a collection of packages for modeling and machine learning using tidyverse principles*. Retrieved 12/03/2024 from https://www.tidymodels.org

Ladeira, L., Verhoeven, A., Ertvelde, J. v., Jiang, J., Sanz-Serrano, J., Gamba, A., Vanhaecke, T., Heusinkveld, H. J., Jover, R., Vinken, M., Geris, L., & Staumont, B. (2024). Unlocking liver physiology: comprehensive pathway maps for mechanistic understanding

Creators. https://doi.org/10.5281/zenodo.14515238

Ladeira, L., Verhoeven, A., van Ertvelde, J., Jiang, J., Gamba, A., Sanz-Serrano, J., Vanhaecke, T., Heusinkveld, H. J., Jover, R., Vinken, M., Geris, L., & Staumont, B. (2025). Unlocking liver physiology: comprehensive pathway maps for mechanistic understanding [Brief Research Report]. *Frontiers in Toxicology*, *Volume 7 - 2025*. https://doi.org/10.3389/ftox.2025.1619651

Lawrence, J. W., Li, Y., Chen, S., DeLuca, J. G., Berger, J. P., Umbenhauer, D. R., Moller, D. E., & Zhou, G. (2001). Differential gene regulation in human versus rodent hepatocytes by peroxisome proliferator-activated receptor (PPAR) alpha. PPAR alpha fails to induce peroxisome proliferation-associated genes in human cells independently of the level of receptor expresson. *J Biol Chem*, *276*(34), 31521-31527. https://doi.org/10.1074/jbc.M103306200

Li, H., Toth, E., & Cherrington, N. J. (2018). Alcohol Metabolism in the Progression of Human Nonalcoholic Steatohepatitis. *Toxicol Sci*, *164*(2), 428-438. https://doi.org/10.1093/toxsci/kfy106

Lin, Z., & Chou, W. C. (2022). Machine Learning and Artificial Intelligence in Toxicological Sciences. *Toxicol Sci*, *189*(1), 7-19. https://doi.org/10.1093/toxsci/kfac075

Liu, Y., Jing, R., Wen, Z., & Li, M. (2019). Narrowing the Gap Between In Vitro and In Vivo Genetic Profiles by Deconvoluting Toxicogenomic Data In Silico. *Front Pharmacol*, *10*, 1489. https://doi.org/10.3389/fphar.2019.01489

Ma, Y., Lee, G., Heo, S. Y., & Roh, Y. S. (2021). Oxidative Stress Is a Key Modulator in the Development of Nonalcoholic Fatty Liver Disease. *Antioxidants (Basel)*, *11*(1). https://doi.org/10.3390/antiox11010091

Maddirevula, S., Shagrani, M., Ji, A. R., Horne, C. R., Young, S. N., Mather, L. J., Alqahtani, M., McKerlie, C., Wood, G., Potter, P. K., Abdulwahab, F., AlSheddi, T., van der Woerd, W. L., van Gassen, K. L. I., AlBogami, D., Kumar, K., Muhammad Akhtar, A. S., Binomar, H., Almanea, H., . . . Alkuraya, F. S. (2024). Large-scale genomic investigation of pediatric cholestasis reveals a novel hepatorenal ciliopathy caused by PSKH1 mutations. *Genet Med*, *26*(11), 101231. https://doi.org/10.1016/j.gim.2024.101231

Martignoni, M., Groothuis, G. M., & de Kanter, R. (2006). Species differences between mouse, rat, dog, monkey and human CYP-mediated drug metabolism, inhibition and induction. *Expert Opin Drug Metab Toxicol*, *2*(6), 875-894. https://doi.org/10.1517/17425255.2.6.875

Masubuchi, Y., & Ihara, A. (2022). Protection of mice against carbon tetrachloride-induced acute liver injury by endogenous and exogenous estrogens. *Drug Metab Pharmacokinet*, *46*, 100460. https://doi.org/10.1016/j.dmpk.2022.100460

McIntosh, A. L., Senthivinayagam, S., Moon, K. C., Gupta, S., Lwande, J. S., Murphy, C. C., Storey, S. M., & Atshaves, B. P. (2012). Direct interaction of Plin2 with lipids on the surface of lipid droplets: a live cell FRET analysis. *Am J Physiol Cell Physiol*, *303*(7), C728-742. https://doi.org/10.1152/ajpcell.00448.2011

Mihajlovic, M., & Vinken, M. (2022). Mitochondria as the Target of Hepatotoxicity and Drug-Induced Liver Injury: Molecular Mechanisms and Detection Methods. *Int J Mol Sci*, *23*(6). https://doi.org/10.3390/ijms23063315

Morrell, A., Tallino, S., Yu, L., & Burkhead, J. L. (2017). The role of insufficient copper in lipid synthesis and fatty-liver disease. *IUBMB Life*, *69*(4), 263-270. https://doi.org/10.1002/iub.1613

O'Donovan, S. D., Cavill, R., Wimmenauer, F., Lukas, A., Stumm, T., Smirnov, E., Lenz, M., Ertaylan, G., Jennen, D. G. J., van Riel, N. A. W., Driessens, K., Peeters, R. L. M., & de Kok, T. (2023). Application of transfer learning to predict drug-induced human in vivo gene expression changes using rat in vitro and in vivo data. *PLoS One*, *18*(11), e0292030. https://doi.org/10.1371/journal.pone.0292030

Palma, F. R., He, C., Danes, J. M., Paviani, V., Coelho, D. R., Gantner, B. N., & Bonini, M. G. (2020). Mitochondrial Superoxide Dismutase: What the Established, the Intriguing, and the Novel Reveal About a Key Cellular Redox Switch. *Antioxid Redox Signal*, *32*(10), 701-714. https://doi.org/10.1089/ars.2019.7962

Peter, A., Stefan, N., Cegan, A., Walenta, M., Wagner, S., Konigsrainer, A., Konigsrainer, I., Machicao, F., Schick, F., Haring, H. U., & Schleicher, E. (2011). Hepatic glucokinase expression is associated with lipogenesis and fatty liver in humans. *J Clin Endocrinol Metab*, *96*(7), E1126-1130. https://doi.org/10.1210/jc.2010-2017

Petrosillo, G., Portincasa, P., Grattagliano, I., Casanova, G., Matera, M., Ruggiero, F. M., Ferri, D., & Paradies, G. (2007). Mitochondrial dysfunction in rat with nonalcoholic fatty liver Involvement of complex I, reactive oxygen species and cardiolipin. *Biochim Biophys Acta*, *1767*(10), 1260-1267. https://doi.org/10.1016/j.bbabio.2007.07.011

Pilling, D., Martinez, T. C., & Gomer, R. H. (2024). Inhibition of CCl4-induced liver inflammation and fibrosis by a NEU3 inhibitor. *PLoS One*, *19*(11), e0308060. https://doi.org/10.1371/journal.pone.0308060

Pu, J., Wang, J., Qin, Z., Wang, A., Zhang, Y., Wu, X., Wu, Y., Li, W., Xu, Z., Lu, Y., Tang, Q., & Wei, H. (2020). IGF2BP2 Promotes Liver Cancer Growth Through an m6A-FEN1-Dependent Mechanism. *Front Oncol*, *10*, 578816. https://doi.org/10.3389/fonc.2020.578816

Pyo, J. H., Kim, T. J., Lee, H., Choi, S. C., Cho, S. J., Choi, Y. H., Min, Y. W., Min, B. H., Lee, J. H., Kang, M., Lee, Y. C., & Kim, J. J. (2021). Proton pump inhibitors use and the risk of fatty liver disease: A nationwide cohort study. *J Gastroenterol Hepatol*, *36*(5), 1235-1243. https://doi.org/10.1111/jgh.15236

Satapathy, S. K., Kuwajima, V., Nadelson, J., Atiq, O., & Sanyal, A. J. (2015). Drug-induced fatty liver disease: An overview of pathogenesis and management. *Ann Hepatol*, *14*(6), 789-806. https://doi.org/10.5604/16652681.1171749

Schmeisser, S., Miccoli, A., von Bergen, M., Berggren, E., Braeuning, A., Busch, W., Desaintes, C., Gourmelon, A., Grafstrom, R., Harrill, J., Hartung, T., Herzler, M., Kass, G. E. N., Kleinstreuer, N., Leist, M., Luijten, M., Marx-Stoelting, P., Poetz, O., van Ravenzwaay, B., . . . Tralau, T. (2023). New approach methodologies in human regulatory toxicology - Not if, but how and when! *Environ Int*, *178*, 108082. https://doi.org/10.1016/j.envint.2023.108082

Statnikov, A., Aliferis, C. F., Tsamardinos, I., Hardin, D., & Levy, S. (2005). A comprehensive evaluation of multicategory classification methods for microarray gene expression cancer diagnosis. *Bioinformatics*, *21*(5), 631-643. https://doi.org/10.1093/bioinformatics/bti033

Sudhahar, V., Urao, N., Oshikawa, J., McKinney, R. D., Llanos, R. M., Mercer, J. F., Ushio-Fukai, M., & Fukai, T. (2013). Copper transporter ATP7A protects against endothelial dysfunction in type 1 diabetic mice by regulating extracellular superoxide dismutase. *Diabetes*, *62*(11), 3839-3850. https://doi.org/10.2337/db12-1228

Tahri-Joutey, M., Andreoletti, P., Surapureddi, S., Nasser, B., Cherkaoui-Malki, M., & Latruffe, N. (2021). Mechanisms Mediating the Regulation of Peroxisomal Fatty Acid Beta-Oxidation by PPARalpha. *Int J Mol Sci*, *22*(16). https://doi.org/10.3390/ijms22168969

Tardelli, M., Bruschi, F. V., Claudel, T., Fuchs, C. D., Auer, N., Kunczer, V., Stojakovic, T., Scharnagl, H., Habib, A., Grabner, G. F., Zimmermann, R., Lotersztajn, S., & Trauner, M. (2019). Lack of monoacylglycerol lipase prevents hepatic steatosis by favoring lipid storage in adipose tissue and intestinal malabsorption. *J Lipid Res*, *60*(7), 1284-1292. https://doi.org/10.1194/jlr.M093369

Tsai, T. H., Chen, E., Li, L., Saha, P., Lee, H. J., Huang, L. S., Shelness, G. S., Chan, L., & Chang, B. H. (2017). The constitutive lipid droplet protein PLIN2 regulates autophagy in liver. *Autophagy*, *13*(7), 1130-1144. https://doi.org/10.1080/15548627.2017.1319544

Vedi, M., Smith, J. R., Thomas Hayman, G., Tutaj, M., Brodie, K. C., De Pons, J. L., Demos, W. M., Gibson, A. C., Kaldunski, M. L., Lamers, L., Laulederkind, S. J. F., Thota, J., Thorat, K., Tutaj, M. A., Wang, S. J., Zacher, S., Dwinell, M. R., & Kwitek, A. E. (2023). 2022 updates to the Rat Genome Database: a Findable, Accessible, Interoperable, and Reusable (FAIR) resource. *Genetics*, *224*(1). https://doi.org/10.1093/genetics/iyad042

Verheijen, M., Sarkans, U., Wolski, W., Jennen, D., Caiment, F., Kleinjans, J., & HeCaTo, S. C. (2022). Multi-omics HeCaToS dataset of repeated dose toxicity for cardiotoxic & hepatotoxic compounds. *Sci Data*, *9*(1), 699. https://doi.org/10.1038/s41597-022-01825-1

Verheijen, M. C., Meier, M. J., Asensio, J. O., Gant, T. W., Tong, W., Yauk, C. L., & Caiment, F. (2022). R-ODAF: Omics data analysis framework for regulatory application. *Regul Toxicol Pharmacol*, *131*, 105143. https://doi.org/10.1016/j.yrtph.2022.105143

Verhoeven, A., van Ertvelde, J., Boeckmans, J., Gatzios, A., Jover, R., Lindeman, B., Lopez-Soop, G., Rodrigues, R. M., Rapisarda, A., Sanz-Serrano, J., Stinckens, M., Sepehri, S., Teunis, M., Vinken, M., Jiang, J., & Vanhaecke, T. (2024). A quantitative weight-of-evidence method for confidence assessment of adverse outcome pathway networks: A case study

on chemical-induced liver steatosis. *Toxicology*, *505*, 153814. https://doi.org/10.1016/j.tox.2024.153814

Wang, F., Man, C., Wang, X., Odle, J., Maltecca, C., & Lin, X. (2023). MicroRNA and mRNA sequencing analyses reveal key hepatic metabolic and signaling pathways responsive to maternal undernutrition in full-term fetal pigs. *J Nutr Biochem*, *116*, 109312. https://doi.org/10.1016/j.jnutbio.2023.109312

Wang, J., Chen, L., & Qiang, P. (2021). The role of IGF2BP2, an m6A reader gene, in human metabolic diseases and cancers. *Cancer Cell Int*, *21*(1), 99. https://doi.org/10.1186/s12935-021-01799-x

Weiler, S., Merz, M., & Kullak-Ublick, G. A. (2015). Drug-induced liver injury: the dawn of biomarkers? *F1000Prime Rep*, *7*, 34. https://doi.org/10.12703/P7-34

Williams, D. P., Shipley, R., Ellis, M. J., Webb, S., Ward, J., Gardner, I., & Creton, S. (2013). Novel in vitro and mathematical models for the prediction of chemical toxicity. *Toxicol Res (Camb)*, *2*(1), 40-59. https://doi.org/10.1039/c2tx20031g

Winter, D. J. (2017). rentrez: An R package for the NCBI eUtils API. *The R Journal*, *9*(2), 520-526. https://journal.r-project.org/archive/2017/RJ-2017-058/index.html

Xu, Z., He, B., Jiang, Y., Zhang, M., Tian, Y., Zhou, N., Zhou, Y., Chen, M., Tang, M., Gao, J., & Peng, F. (2022). Igf2bp2 knockdown improves CCl(4)-induced liver fibrosis and TGF-beta-activated mouse hepatic stellate cells by regulating Tgfbr1. *Int Immunopharmacol*, *110*, 108987. https://doi.org/10.1016/j.intimp.2022.108987

Yang, M., Wang, D., Wang, X., Mei, J., & Gong, Q. (2024). Role of Folate in Liver Diseases. *Nutrients*, *16*(12). https://doi.org/10.3390/nu16121872

Zou, H., & Hastie, T. (2005). Regularization and Variable Selection Via the Elastic Net. *Journal of the Royal Statistical Society Series B: Statistical Methodology*, *67*(2), 301-320. https://doi.org/10.1111/j.1467-9868.2005.00503.x

Declaration of interests

☒ The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

☐ The authors declare the following financial interests/personal relationships which may be considered as potential competing interests:

Graphical abstract