scientific reports

Check for updates

OPEN

Integrative analysis of thiamethoxam induced hepatocellular carcinoma toxicity mechanisms

Chenghao Wang[1], Zhichao Wang[1], Yinqin Hu[2], Shi Xiang[1], Hui Chen[1,3]✉ & Chao Yang[1,3]✉

Neonicotinoid (NEO) pesticides play a crucial role in agricultural production. However, their potential risks to human health and the environment cannot be overlooked. To gain a comprehensive understanding of the toxicity and mode of action of NEOs, thiamethoxam (THX), which exhibits the highest potential for carcinogenicity and hepatotoxicity, was selected as the subject of this study. We identified 61 intersection genes between THX targets and hepatocellular carcinoma (HCC)-related genes. These genes were then uploaded to the Metascape database for Gene Ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) analyses. The GO analysis indicated that the significant biological processes mainly involved the response to xenobiotic stimuli, cellular response to chemical stress, cellular response to biotic stimuli, and response to toxic substances. The KEGG enrichment analysis pinpointed several key pathways, primarily including the cell cycle and Glycolysis/Gluconeogenesis. Subsequently, the intersection genes were imported into the Gene Expression Profiling Interactive Analysis (GEPIA) and Gene Expression Omnibus (GEO) databases to analyze expression differences, leading to the identification of 15 significantly differentially expressed core genes (SDECGs). By applying the Support Vector Machine (SVM) machine—learning model, we screened out five feature genes (CYP2C19, CYP3A4, FBP1, THBS4, CYP7A1) and constructed a nomogram. Molecular docking of THX with these five feature genes showed binding energies of less than −5 kcal/mol. This study offers a theoretical foundation for understanding the underlying mechanisms of THX-induced HCC. The findings provide a scientific basis for the safety assessment of THX in agricultural applications and contribute to the establishment of pesticide safety standards.

Keywords Hepatocellular carcinoma, Neonicotinoid pesticide, Thiamethoxam, Network toxicology, Machine learning, Molecular docking

Neonicotinoid (NEO) pesticides are the fourth generation of insecticides, which have good effects on the control of stack-inhaling mouth pests, small Lepidopteran and Coleopteran pests[1]. The commonly used NEOs include Imidacloprid (IMI), Acetamiprid (ACE), Thiamethoxam (THX), Clothianidin (CLO) and Dinotefuran (DIN)[2]. Unlike previous generations of pesticides, such as organophosphorus pesticides and pyrethroids, which inhibit acetylcholinesterase and interfere with sodium channels, NEOs kill insects by directly binding to the insect's acetylcholine receptor and simulating acetylcholine acting on its nervous system[3–5]. Its insecticidal mechanism is more selective, systematic, and less residual, so it has been widely used worldwide[6]. As the utilization of neonicotinoid pesticides has increased, they have been detected in various environmental media, such as domestic soil, water resources, and residential areas[7–9]. Residues have also been found in foods, such as fruit and other products consumed by humans, which has prompted concerns about potential exposure risks[10]. Moreover, traces of these pesticides have been detected in biological samples like blood and urine, further emphasizing the necessity for discourse surrounding their safety[11,12]. Studies have demonstrated that NEOs can have adverse effects on human health, including contributing to chronic diseases such as neurological disorders

[1]Oncology Research Center, Jiangxi Provincial Key Laboratory of Traditional Chinese Medicine Diagnosis and Rehabilitation of Malignant Tumors, Jiangxi University of Chinese Medicine, Nanchang 330004, People's Republic of China. [2]Department of Cardiovascular Disease, ShuGuang Hospital Affiliated to Shanghai University of Traditional Chinese Medicine, Shanghai 200021, People's Republic of China. [3]Jiangxi Engineering Research Center for Translational Cancer Technology, Jiangxi University of Chinese Medicine, No. 1688 Meiling Road, Wanli District, Nanchang 330004, People's Republic of China. ✉email: ch_0102@163.com; yangchao@jxutcm.edu.cn

and cancer[13,14]. Recent clinical research suggested that exposure to neonicotinoids and their metabolites may increase the risk of liver cancer[15].

Hepatocellular carcinoma (HCC) is the most common type of primary liver cancer and ranks as the seventh most common frequently occurring malignant tumor and is the fourth leading cause of cancer-related mortality in the world[16,17]. Researches indicate that the development of HCC is associated with chronic liver diseases and pesticide exposure[18]. Nevertheless, few studies have undertaken a comparative analysis of the hepatic toxicity and carcinogenic potential of commonly used neonicotinoid insecticides. Furthermore, research on the mechanisms by which these pesticides may induce HCC is also limited.

To address the aforementioned, we initially utilized the ADMETlab3.0 and Protox3.0 databases to predict the toxicity of five commonly used NEOs and found the pesticide with the highest potential for carcinogenicity and liver damage. Subsequently, network toxicology, enrichment analysis, GEO database and four machine learning algorithms combined with molecular docking were used to investigate the mechanism of the pesticide inducing HCC. The above methods are valuable due to their capacity to predict and detect possible toxic effects by analyzing large datasets and biological data, thereby providing an in-depth understanding of the risks related to pesticide exposure at the molecular level. Moreover, assessing the potential carcinogenic and hepatotoxic effects of this pesticide is crucial for evaluating its safety in both agricultural and medical contexts, as well as for generating scientific evidence that can guide the establishment of safety protocols and protective guidelines.

Methods

Toxicity analysis of common neonicotinoid pesticides and pesticide selection

The chemical information and Simplified Molecular Input Line Entry System (SMILES) identifiers for five NEOs were obtained from PubChem (https://pubchem.ncbi.nlm.nih.gov/) and PubMed (https://pubmed.nc- bi.nlm.nih.gov/). The toxicity scores for these pesticides were assessed using the ProTox3.0 (https://tox.charite.de/protox3/index.php?site=home) and ADMETlab3.0 (https://admetlab3.scb- dd.com/server/screening) databases. The pesticide with the highest potential for carcinogenicity and hepatotoxicity was selected for further analysis.

Identification of pesticide and HCC-related targets

Relevant pesticide targets were retrieved from the SwissTarget Prediction (http://swisstarget-prediction.ch/) and CTD (https://ctdbase.org/) databases. Targets in the Swiss Target Prediction were chosen if their p-value exceeded 0. Protein names were standardized by using the database- UniProt (https://www.uniprot.org/), and duplicate entries were eliminated. The main HCC-related targets were obtained from GenCards (https://www.genecards.org/) and supplemented by OMIM (https://www.omi- m.org/). The intersection of pesticide and HCC-related targets were identified by VENNY2.1 (https://bioinfogp.cnb.csic.es/tools/venny/index.html) as potential toxicological targets for further investigation.

GO and KEGG functional analysis

The intersected protein names were input into the Metascape database (https://metascape.org/gp/index.html#/main/step1) for GO analysis and KEGG pathway enrichment[19]. For GO analysis, the top 20 biological processes (BP), cellular components (CC), and molecular functions (MF) were selected according to p-values. The 20 signaling pathways with the smallest p-values were selected in the KEGG analysis. The plot was then performed according to −log10(P).

GEPIA analysis and GEO data collection for validation of expression differences, chromosomal position, and correlation of SDECGs

In order to reveal the core mechanism of THX affecting HCC in more depth, we screened the intersection targets. The intersected targets were input into the GEPIA (http://gepia.cancer-pku.cn/) website to assess significant differential expression ($P<0.01$) of each target in HCC and normal groups, and the differential targets were identified as core genes for further analysis. Then a search of the GEO-database (https://www.ncbi.nlm.nih.gov/geo/) for datasets related to "Hepatocellular carcinoma" was conducted. GSE45267, GSE62232, GSE112790 and GSE54236 were used for further study. Gene symbol annotation and data correction were performed using Perl code to extract the expression levels of core genes in normal and HCC samples. Then these data were extracted for differential expression analysis using RStudio packages such as "limma", "ggpubr" and "pheatmap". The results were shown using boxplots and heatmaps, and then verified with the results obtained from GEPIA. Significantly differentially expressed core genes (SDECGs) were defined as genes with an adjusted p-value (adjP, FDR-corrected) < 0.01, consistent with GEPIA expression results[20]. The chromosomal location of core genes was determined using Perl code and visualized as circular plots using the 'Rcircos' R package. Additionally, the correlation of each SDECG was computed using the 'cor' function and visualized.

Immune cell infiltration, differences, and correlation in HCC samples

Immune cell infiltration was assessed using the CIBERSORT algorithm in R (4.4.1), with 1,000 iterations to estimate the relative abundance of various immune cell types. The results were visualized through bar plots representing immune cell content in each sample. To compare immune cell infiltration between normal and HCC groups, single-sample Gene Set Enrichment Analysis (ssGSEA) was conducted using the "GSVA" and "GSABase" R packages. The ssGSEA results were visualized using box plots to highlight differences in immune cell content. Furthermore, the correlation between SDECGs and ssGSEA scores was analyzed, and the correlation coefficients were visualized for better interpretation[21].

Machine learning models for pesticide-induced HCC and feature gene selection with external validation

Four predictive models were constructed using the expression data of significantly SDECGs, including Random Forest (RF), Support Vector Machine (SVM), Generalized Linear Model (GLM), and Extreme Gradient Boosting (XGB). All models were implemented in R (4.4.1) using the caret package, with model-specific algorithms provided by randomForest, kernlab, and xgboost. Each model was trained on a 70% randomly selected training set using five-fold repeated cross-validation to optimize performance and minimize overfitting. To further evaluate the predictive performance and interpretability of each model, model interpretation procedures were constructed using the DALEX package. A unified prediction function was first defined to ensure the compatibility of all models. Based on this function, the prediction residuals of each model for the test set were calculated. The residual distribution was visualized using inverse cumulative distribution plots and box plots. The receiver operating characteristic (ROC) curve and the area under the curve (AUC) were calculated by pROC software package. The receiver operating characteristic (ROC) curve was drawn for comparison visualization, and the AUC value was marked to comprehensively evaluate the discrimination ability of each model by the above method. After determining the optimal model, a nomogram was developed based on the feature genes and their expression levels in both normal and HCC groups. Decision curves and calibration curves were then used to assess the accuracy of the nomogram. For external validation, a separate dataset containing normal and HCC samples was obtained from GEO database. Machine learning models were then applied in R following the same methodology, and ROC curves were plotted to validate the model performance.

Molecular docking

The 3D structures of the selected pesticide components and feature genes were obtained from the PubChem, UniProt database and PDB database (http://www.rcsb.org/). If the required structure was not obtained in the PDB database, the AlphaFold structure was obtained in the UniProt database. Molecular docking was carried out using AutoDock Vina (1.1.2), and visualization was done with Pymol (2.6).

Results

Toxicity analysis of common neonicotinoid pesticides and pesticide selection

The chemical details of five plasticizers (IMI, ACE, THX, CLO, and DIN) are provided in Table 1 in Supplementary File, which includes their chemical formulas, molecular weights, and SMILES structures. Toxicity information related to NEOs was retrieved using ADMETlab 3.0 and Protox 3.0, as shown in Table 1. On the basis of carcinogenic and hepatotoxic properties, THX was selected for further investigation.

Identification of THX and HCC-related targets

The core biological significance of gene crossover in network toxicology lies in its ability to decipher the molecular interconnections between compound targets and disease mechanisms. This process provides a critical mechanistic foundation for elucidating how THX exerts its effects in the context of this study, thereby offering deeper insights into the underlying action pathways and contributing to a more comprehensive understanding of the therapeutic or toxicological mechanisms involved. A total of 11 and 124 THX-related targets were obtained from Swiss Target Prediction and CTD, respectively. After merging and removing duplicates, 84 THX-related targets were identified. A total of 11,938 and 192 HCC-related targets were retrieved from Genecards and OMIM, respectively. To ensure that the genes obtained were highly correlated with HCC, the GeneCards "score" threshold was set to ≥ 3 (determined by rounding the median score value upward[22]). Following the merging and removal of duplicates, 5672 targets were identified. A Venn diagram was constructed, resulting in the identification of 61 common targets (Fig. 1), which were subjected to further analysis to elucidate the molecular mechanisms underlying THX-induced HCC.

GO and KEGG functional analysis

The 61 common targets were imported into Metascape for GO and KEGG enrichment analysis[23]. The significant biological processes identified based on p-values included response to xenobiotic stimulus, cellular response to chemical stress, cellular response to biotic stimulus and response to toxic substance. The mainly involved cellular components were peroxisome, organelle outer membrane, endoplasmic reticulum lumen. The primary molecular functions included protein domain specific binding, protein homodimerization activity, protease binding. The KEGG enrichment analysis identified several key pathways, including IL-17, FoxO, HIF-1 signaling pathway, cell cycle and Glycolysis / Gluconeogenesis (Fig. 2).

GEPIA analysis, GEO data collection, and validation of expression differences, chromosome location, and correlation of SDECGs

Following the importation of the 61 common targets into the GEPIA website, 15 targets exhibited significant differential expression ($P < 0.01$) (Figs. 1 2 in Supplementary File). We combined three datasets (GSE45267, GSE62232, and GSE112790) for analysis and model construction. After removing the batch effect using the "ComBat" function, the combined dataset included 41, 10, and 15 normal samples, and 46, 80, and 183 hepatocellular carcinoma (HCC) samples from each respective dataset[21]. And GSE54236, which contained 80 normal and 81 HCC samples was used for external validation. A differential analysis between the HCC group and the normal group from the combined GEO dataset was performed. The results revealed the following genes were differentially expressed (adj$P < 0.01$): ANXA10, ATF3, BAX, BCL6, CDC25C, CDK1, CYP2C19, CYP3A4, CYP7A1, ESR1, FBP1, GADD45B, IMPDH2, S100P and THBS4. The aforementioned genes were identified as SDECGs, these core genes selected through human sample validation, and demonstrate high accuracy and clinical relevance. The genes BAX, CDC25C, CDK1, CYP7A1, IMPDH2, S100P and THBS4 exhibited high

Drug name	Property	Database	Probability
Imidacloprid	Carcinogenicity	ADMETlab 3.0	+ +
	Respiratory	ADMETlab 3.0/Protox3.0	+ + + / + + +
	Human Hepatotoxicity	ADMETlab 3.0/Protox3.0	+ / +
	Hematotoxicity	ADMETlab 3.0	+
	Immunotoxicity	Protox3.0	+ + +
	Genotoxicity	ADMETlab 3.0	+ + +
Acetamiprid	Carcinogenicity	ADMETlab 3.0	+
	Respiratory	ADMETlab 3.0/Protox3.0	+ + + / + + +
	Human Hepatotoxicity	Protox3.0	+
	Immunotoxicity	Protox3.0	+ + +
	Genotoxicity	ADMETlab 3.0	+ + +
Thiamethoxam	Carcinogenicity	ADMETlab 3.0	+ + +
	Respiratory	ADMETlab 3.0/Protox3.0	+ + + / + + +
	Human Hepatotoxicity	ADMETlab 3.0/Protox3.0	+ + / +
	Drug-induced Nephrotoxicity	ADMETlab 3.0	+ +
	Hematotoxicity	ADMETlab 3.0	+ +
	Immunotoxicity	Protox3.0	+ + +
	Genotoxicity	ADMETlab 3.0	+ + +
Clothianidin	Carcinogenicity	ADMETlab 3.0	+ +
	Respiratory	ADMETlab 3.0/Protox3.0	+ + + / + + +
	Human Hepatotoxicity	ADMETlab 3.0/Protox3.0	+ / +
	Hematotoxicity	ADMETlab 3.0	+ +
	Immunotoxicity	Protox3.0	+ + +
Dinotefuran	Carcinogenicity	ADMETlab 3.0	+ +
	Respiratory	Protox3.0	+ + +
	Human Hepatotoxicity	ADMETlab 3.0/Protox3.0	+ / +
	Immunotoxicity	Protox3.0	+ + +
	Genotoxicity	ADMETlab 3.0	+ + +

Table 1. The prediction toxicity of Thiamethoxam. Tip: 0.5–0.7(+),0.7–0.9(+ +),0.9–1.0(+ + +).Larger probability represents a greater risk of toxicity.

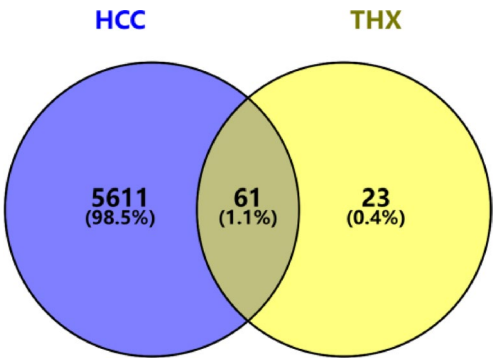

Fig. 1. Venn diagram of the intersection between HCC and THX target sets.

expression levels in HCC group, whereas the remaining genes showed low expression in HCC group, as illustrated in Figs. 3A and B. The expression patterns of these genes were consistent with the results obtained from GEPIA analysis. The exact chromosomal locations of the THX core genes are shown in Fig. 3C. As shown in Fig. 3D and E, the correlation analysis between each pair of SDECGs in the HCC samples demonstrated a robust, predominantly positive correlation between them.

Analysis of immune cell infiltration, differences, and correlation between normal and HCC samples

To investigate the underlying mechanisms distinguishing HCC from normal tissues, immune cell infiltration was analyzed from multiple perspectives. Initially, immune cell infiltration analysis was conducted, followed by

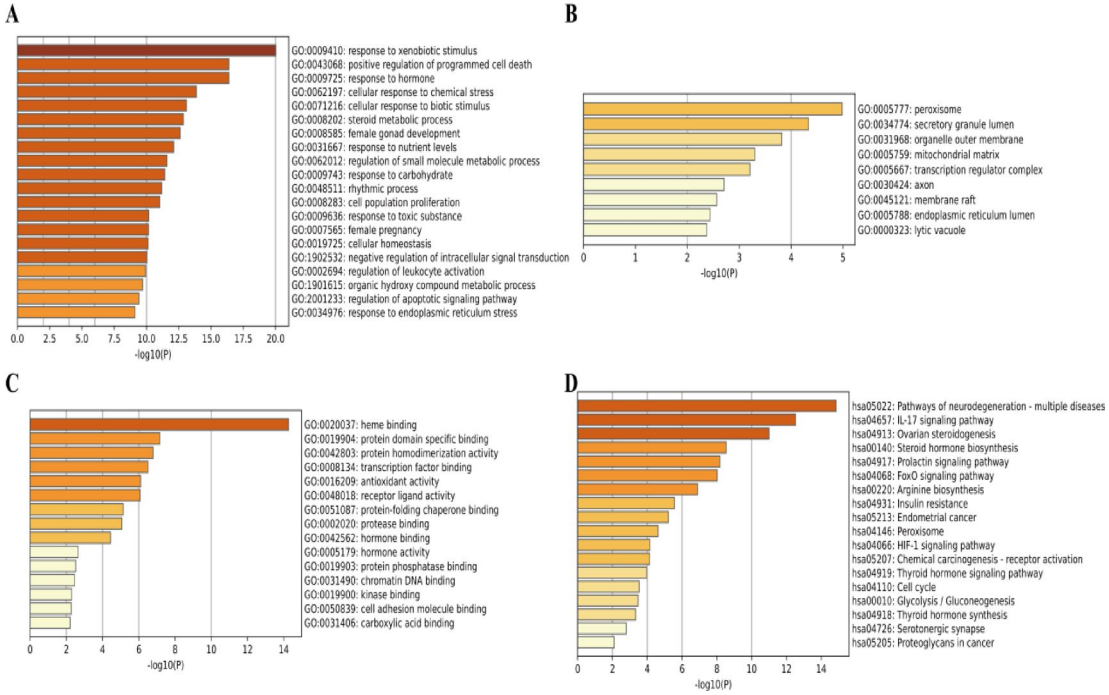

Fig. 2. GO and KEGG analysis of THX and HCC-Related Targets (**A**–**C**) Diagrams of BP, CC, and MF results in GO analysis; (**D**) KEGG analysis diagram.

ssGSEA to identify significantly altered immune cells in both normal and HCC samples. The results revealed that monocytes were predominantly expressed in the normal group, while T cells (CD8), T cells regulatory(Tregs), M1 macrophages, and activated dendritic cells were significantly upregulated in the HCC group (Fig. 4A). Further correlation analysis of immune cells demonstrated significant associations between immune cell infiltration and SDECGs ($P < 0.05$) (Fig. 4B). Specifically, resting dendritic cells, M0 macrophages, M1 macrophages, resting mast cells, and activated CD4 memory T cells showed a primarily negative correlation with SDECGs. In contrast, activated dendritic cells, M2 macrophages, activated mast cells, monocytes, activated NK cells, follicular helper T cells, and regulatory T cells (Tregs) exhibited a positive correlation with SDECGs.

Machine learning model for THX-induced HCC and feature gene selection with external validation

Four machine learning models—SVM, RF, XGB, GLM—were constructed using SDECGs data. Among these models, the SVM model demonstrated superior performance, as evidenced by its better ROC curves, residual box plots, and inverse cumulative distribution plots. Specifically, the SVM model exhibited the highest Area Under the Curve (AUC) and the lowest residuals and inverse cumulative values (Fig. 5A–C). Using the importance scores calculated from the SVM model, ten feature genes were identified from the SDECGs, including CYP2C19, CYP3A4, FBP1, THBS4, CYP7A1, ESR1, CDK1, BCL6, IMPDH2, and CDC25C (Fig. 5D). The top five feature genes were then used to construct nomograms, allowing for the derivation of individual score scales (Fig. 5E). These scores were employed to evaluate sensitivity to THX-induced HCC and predict the risk of developing HCC due to prolonged exposure to THX. Calibration curves and decision curves showed that the proximity of solid lines and dotted lines, as well as the distances between the red and grey lines, indicated a high degree of accuracy in the model's predictions (Fig. 5F–G). The ROC curve further confirmed the model's reliability, with an AUC of 0.951 (95% CI > 0.903) (Fig. 5H). To validate the model externally, the top five identified genes were incorporated into the SVM-based prediction model. The resulting ROC curve for the validation set demonstrated an AUC of 0.902, highlighting the model's high accuracy when applied to the GEO dataset (Fig. 5I).

Molecular Docking

In order to verify whether the proteins encoded by the top five THX feature genes are the optimal targets for inducing HCC, molecular docking was performed. The Autodock Vina docking results showed that the binding energy of the complexes formed by THX and the target proteins were (−6.2, −6.9, −6.8, −5.7, −6.0 kcal/mol) (Table 2), respectively, which were all less than -5 kcal/mol. These data suggested strong interactions between THX and the target proteins. Visualization with PyMOL revealed the formation of hydrogen bonds between the target proteins and THX which confirmed the stability of these interactions (Fig. 6).

Discussion

As a second-generation NEO featuring a mechanism analogous to its predecessor, THX has emerged as a staple in global agriculture—particularly across key farming nations including China, the United States, and

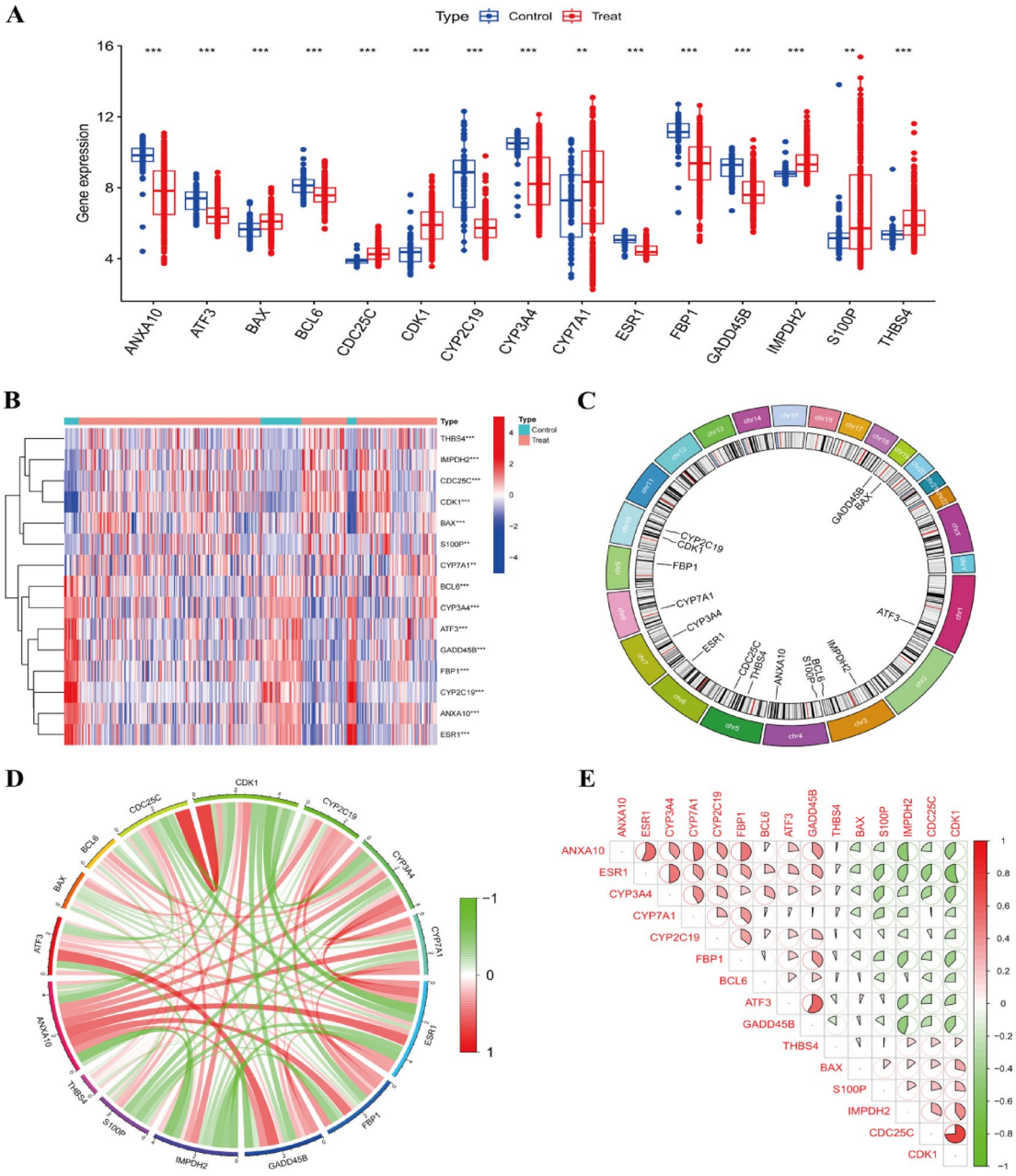

Fig. 3. Differential gene expression profiling and interaction analysis. (**A**) Box plot showing the expression differences of core genes between normal and HCC samples; (**B**) Heatmap of SDECG expression in normal and HCC samples; (**C**) Circular plot depicting the chromosomal locations of core genes; (**D**) Correlation network of SDECGs; (**E**) Correlation analysis between the two SDECGs. Tip: ** is $p < 0.01$; *** is $p < 0.001$.

numerous European countries—due to its exceptional efficacy, minimal toxicity, and broad-spectrum pesticidal activity. This versatility is further underscored by its dual application in seed treatment protocols and foliar spray formulations, enabling comprehensive pest management across diverse cropping systems such as grain, fruit, vegetables and so on[24,25]. The use of THX has also generated some controversy and challenges. In recent years, several countries and regions have implemented strict regulatory measures, or have limited its use on certain crops[26]. However, there is limited research on thiamethoxam's potential to cause disease, particularly cancer. Accordingly, in this study, we first predicted the toxicity of commonly used neonicotinoid pesticides through the toxicology prediction platform. The findings of this study indicated that the carcinogenic potential and liver injury toxicity of THX were higher than those of the other four commonly used neonicotinoid pesticides. Subsequently, the mechanism of thiamethoxam-induced hepatocellular carcinoma was studied using network toxicology methods. Additionally, ACE was found to have the lowest carcinogenic potential, while CLO and DIN exhibited lower hepatotoxicity, providing valuable insights for future pesticide usage.

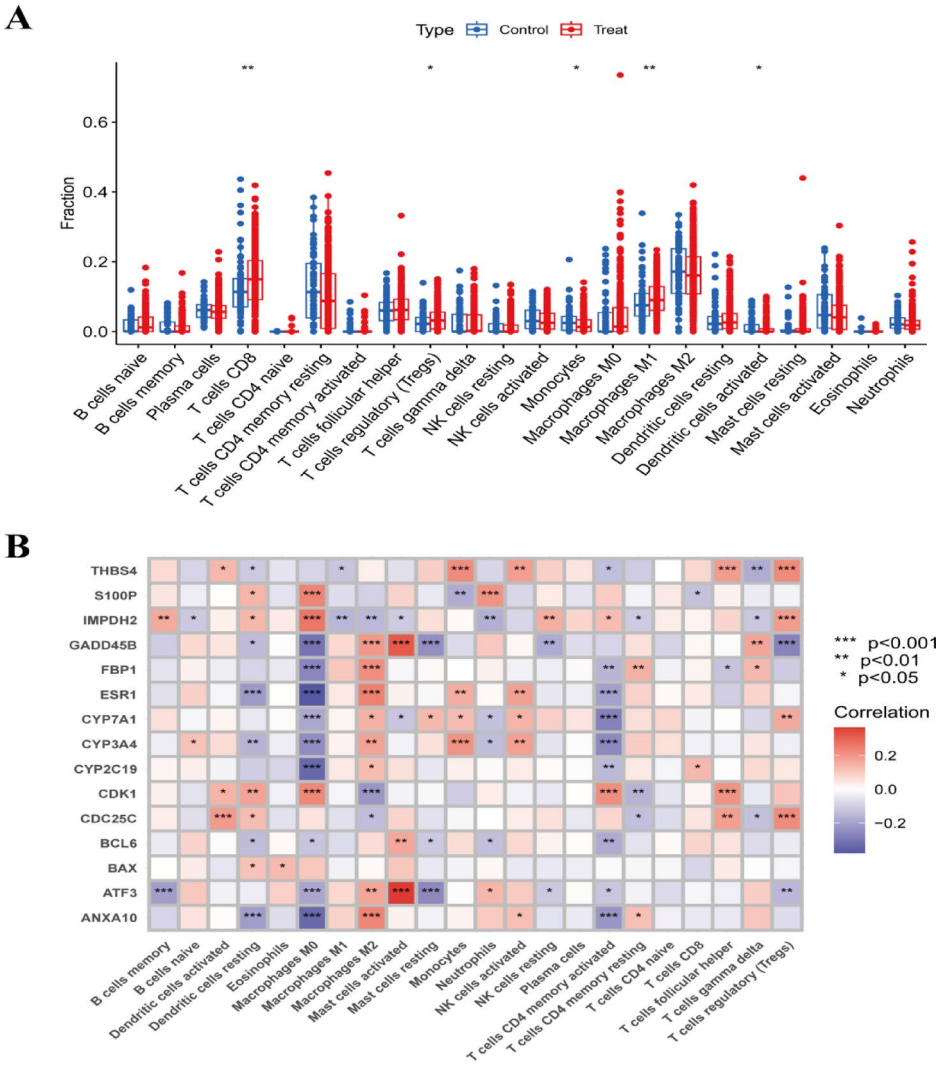

Fig. 4. Immune cell infiltration analysis of SDECGs in HCC. (**A**) Box plot comparing immune cell fractions between normal and HCC samples; (**B**) Heatmap of the correlation analysis between SDECGs and immune cells. Tip: * is $p < 0.05$; ** is $p < 0.01$; *** is $p < 0.001$.

GO analyses of the 61 intersecting targets revealed key biological processes, including cellular response to xenobiotic stimulus, chemical stress, biotic stimulus, and toxic substances. The result of KEGG enrichment analysis mainly included cell cycle and Glycolysis / Gluconeogenesis. The development of HCC is associated with chronic liver diseases such as chronic hepatitis, cirrhosis, and liver fibrosis. Prolonged exposure to external stimuli, including viral infections and chemicals, can lead to chronic liver damage, increasing the risk of HCC[27]. The onset of HCC is often accompanied by dysregulation of the cell cycle. HCC cells frequently accelerate the cell cycle by modulating cyclin-dependent kinases (CDKs), promoting tumor proliferation and transformation[28]. The up-regulation of CDK1 in HCC samples, as observed in our differential gene analysis, lends further support to support this conclusion. As shown in Fig. 3E, there was a strong positive correlation between CDK1 and CDC25C. The metabolic pathways of glycolysis and gluconeogenesis are vital in the development and progression of hepatocellular carcinoma (HCC)[29]. Gluconeogenesis is the reverse process of glycolysis and is inhibited in tumorous liver[30]. In clinical studies, synthetic glucocorticoid drugs such as dexamethasone have been reported to have therapeutic effects on HCC. The underlying mechanism is to enhance the gluconeogenic pathway to inhibit tumor growth and angiogenesis[31]. This is consistent with the results obtained from our KEGG enrichment analysis.

We identified differentially expressed core genes through integrated analysis of GEPIA and GEO datasets, revealing a consistent trend of dysregulation in 15 candidate targets (SDECGs). Leveraging these data-driven observations, we developed a Support Vector Machine (SVM) model to prioritize key diagnostic features, which identified 10 high-impact genes (CYP2C19, CYP3A4, FBP1, THBS4, CYP7A1, ESR1, CDK1, BCL6, IMPDH2, and CDC25C). Based on their ranked importance in the SVM model, the top five genes were selected to construct a clinical nomogram. This data-derived diagnostic model for chronic THX exposure gains biological plausibility from the established roles of its key genes (e.g., CYP2C19, CYP3A4, CYP7A1, FBP1) in xenobiotic

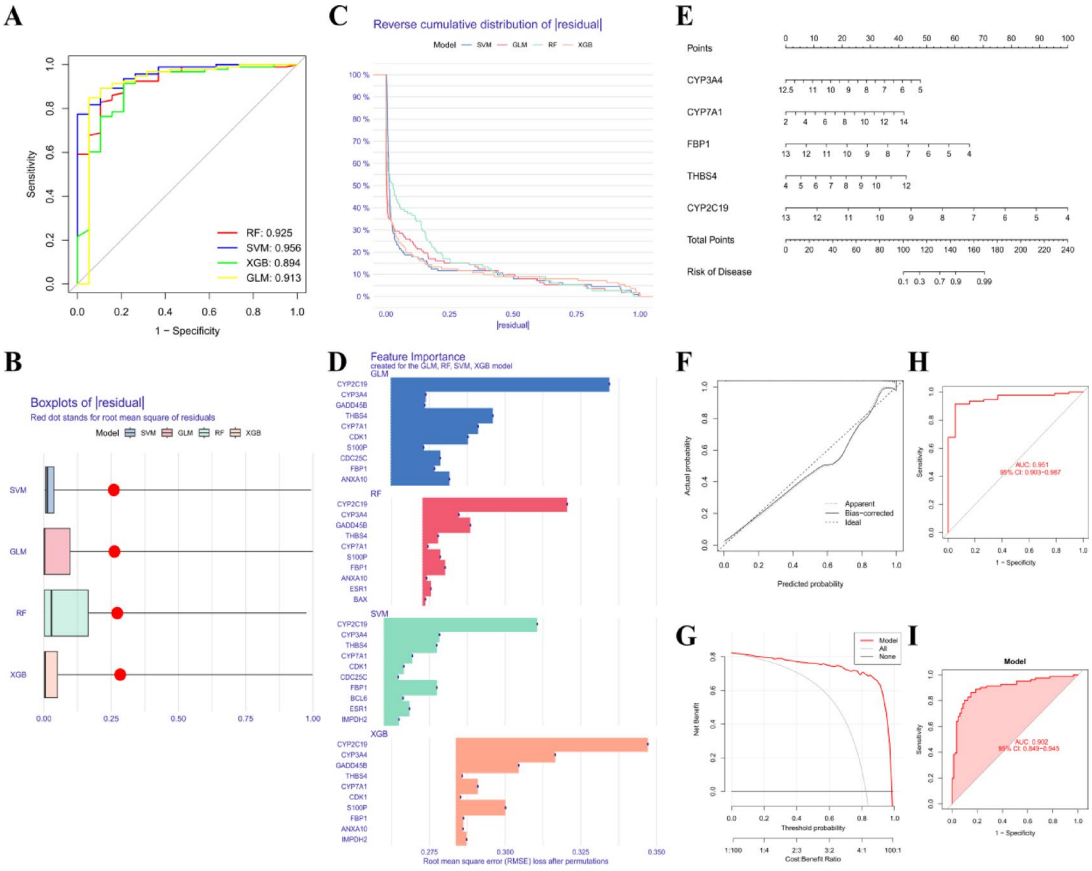

Fig. 5. Machine learning for screening the key genes. (**A**) ROC curves for SVM, RF, XGB, GLM; (**B**) Box plots of residuals for SVM, RF, XGB, GLM; (**C**) Reverse cumulative distribution of residuals for SVM, RF, XGB, GLM machine learning models; (**D**) Bar plot showing the feature importance of SVM, RF, XGB, GLM machine learning models; (**E**) Nomogram was constructed based on five feature genes; (**F**) Calibration curve of the feature gene nomogram for THX-induced HCC; (**G**) Decision curve for the feature gene nomogram for THX-induced HCC; (**H**) ROC curve for the test GEO dataset; (**I**) ROC curves for external validation using the GEO dataset.

Gene name	UniProt id	Binding energy (kcal/mol)
CYP2C19	P33261	− 6.2
CYP3A4	P08684	− 6.9
FBP1	P09467	− 6.8
THBS4	P35443	− 5.7
CYP7A1	P22680	− 6.0

Table 2. Results of molecular docking binding energy.

metabolism and cellular stress responses. Critically, these genes converge on liver metabolic pathways-a well-established hub influencing HCC pathogenesis[32]. Cytochrome P450 (CYP) is essential in the metabolism of drugs, chemicals, and endogenous compounds[33]. CYP2C19 is a drug-metabolizing enzyme, the metabolites produced during its metabolism may exert a significant influence on the liver's microenvironment[34]. Abnormal CYP2C19 metabolism may contribute to the exacerbation of liver inflammation, thereby promoting the development of HCC[35]. CYP3A4 expression has been found to be significantly reduced in liver tumor tissues of patients with hepatocellular carcinoma[36]. Several studies have indicated that reduced CYP3A4 expression is associated with a poorer prognosis in patients with hepatocellular carcinoma (HCC)[37]. Dysregulation of bile acid metabolism is involved in the development of liver cancer[38]. Multiple studies have suggested that decreased CYP3A4 expression correlates with worse prognosis in patients with HCC, this enzyme is expressed only in the endoplasmic reticulum of hepatocytes and is regulated by the end product bile acids[39]. Studies have shown a correlation between elevated levels of major circulating bile acids and an increased risk of HCC, which is consistent with our results of increased CYP7A1 expression[40]. Fructose-1,6-bisphosphatase (FBP1) is a crucial

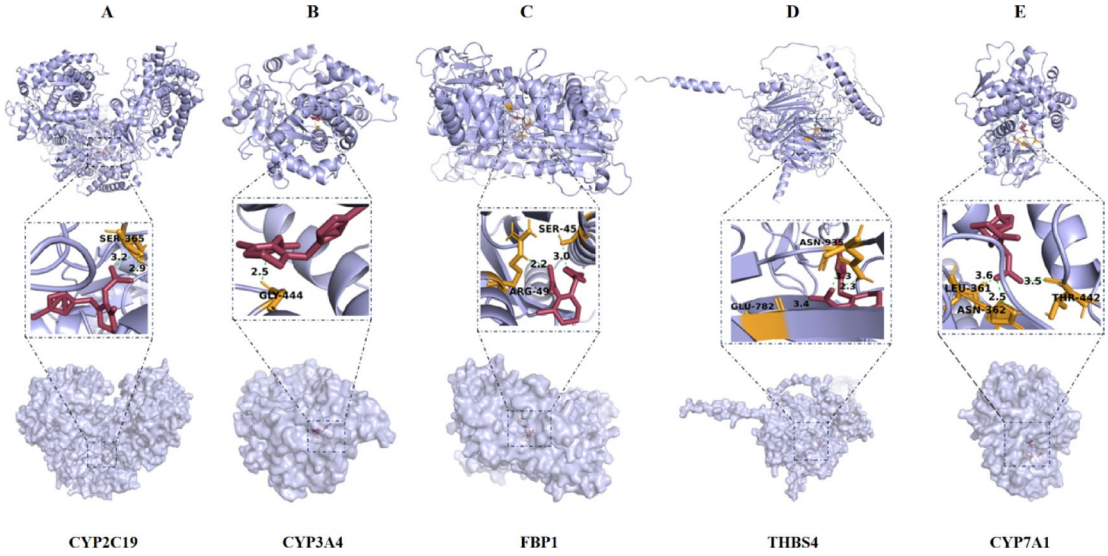

Fig. 6. The molecular docking results of THX with the top five genes. The molecular docking analysis of THX binding to CYP2C19 (**A**), CYP3A4 (**B**), FBP1 (**C**), THBS4 (**D**) and CYP7A1 (**E**) was conducted using AutoDock Vina software,

enzyme in the process of gluconeogenesis[41]. Experimental studies have shown that low expression of FBP1 promotes HCC progression[42]. At the same time, loss of FBP1 disrupts liver metabolic homeostasis and promotes tumor progression[43]. This verified the accuracy of our KEGG analysis results. Studies have demonstrated that overexpression of THBS4 can mediate the proliferation and metastasis of HCC[44]. The expression of THBS4 have been linked to the invasion and metastasis of HCC[45]. Additionally, in the liver cancer microenvironment, THBS4 may promote tumor growth and metastasis by regulating intercellular interactions and angiogenesis[46]. The results of molecular docking showed that THX had a stable binding to the five signature genes, which revealed the accuracy of the above machine learning results and provided certain evidence for our study, and elucidated that THX-induced HCC may be mainly through the five signature genes.

It is noteworthy that considerable discrepancies in ssGSEA were discerned between the HCC and normal groups with respect to various immune cell types, including Monocytes, T cells CD8, Tregs, Macrophages M1 and Dendritic cells activated. Furthermore, correlation analysis revealed significant associations between some SDECGs and immune cell types. These include M1 macrophages, Tregs, Monocytes. M1 macrophages are known to promote the activation of T cells and NK cells by secreting pro-inflammatory cytokines and chemical factors, and have been shown to enhance tumour-specific immune responses. These cytokines can induce tumor cell apoptosis, inhibit tumor angiogenesis, and even enhance immune surveillance of tumor cells[47,48]. In the tumor immune microenvironment, a higher proportion of M1 macrophages is often correlated with improved prognosis. Additionally, monocyte count has been identified as an independent prognostic factor for overall survival (OS) in patients with HCC. In the HCC microenvironment, monocytes can be polarized into tumor-promoting macrophages (M2 macrophages). M2 macrophages foster tumor growth and metastasis by enhancing angiogenesis and fibrosis within the tumor microenvironment[49]. However, in our study, there was no significant difference in M2 macrophages, which may be related to the number of samples. In addition, monocytes can trigger autophagy, induce epithelial-mesenchymal transition (EMT) of cancer cells, and promote tumor metastasis[50]. Tregs are the key mediators of HCC-related immunosuppression. Tregs can help cancer cells escape immune surveillance by inhibiting anti-tumor immune responses[51]. Our completed molecular docking results further imply potential mechanistic intersections: for instance, if THX binds to key regulatory proteins associated with these immune cells (such as receptors or signaling molecules on M1 macrophages, Tregs, or monocytes), it might modulate their functions. Binding to M1 macrophage-related targets could either enhance their anti—tumor cytokine secretion or disrupt it, while interactions with Treg-associated targets might impair their immunosuppressive capacity. These potential bindings, if validated, would deepen our understanding of THX's immunomodulatory role in HCC, aligning with KEGG identified IL-17 mediated immune microenvironment alterations and the observed immune cell discrepancies from ssGSEA.

This study elucidates the potential molecular mechanisms of THX-induced HCC using network toxicology and machine learning methods, providing support for disease diagnosis in populations exposed to THX or long-term use of this pesticide. However, it should be noted that this study is not without limitations. For example, this study solely investigates the molecular mechanisms underlying THX-induced HCC from a data analysis perspective. Therefore, additional experiments (such as animal studies) are required to validate these findings. Furthermore, due to the small number of clinical liver cancer samples, errors may have been introduced into the study results.

Conclusion

This study systematically investigated THX, a NEO with notable carcinogenic and hepatotoxic potential, by identifying 61 intersection genes between THX targets and hepatocellular carcinoma (HCC)-related genes. Through integrated bioinformatics analyses and machine learning modeling, five core feature genes (CYP2C19, CYP3A4, FBP1, THBS4, CYP7A1) were screened, and molecular docking confirmed strong binding affinities (≤ −5 kcal/mol) between THX and these genes. These findings provide mechanistic insights into THX-induced HCC, establishing a theoretical foundation for understanding its hepatotoxic and carcinogenic mechanisms.

Data availability

The datasets used and/or analysed during the current study are available from the corresponding author on reasonable request.

Received: 30 March 2025; Accepted: 14 July 2025
Published online: 22 July 2025

References

1. Zhang, X. et al. Environmental occurrence, toxicity concerns, and biodegradation of neonicotinoid insecticides. *Environ. Res.* **218**, 114953. https://doi.org/10.1016/j.envres.2022.114953 (2023).
2. Sass, J. B. & Raichel, D. Human acute poisoning incidents associated with neonicotinoid pesticides in the US. Incident data system (IDS) database from 2018–2022: Frequency and severity show public health risks, regulatory failures. *Environ Health* **23**, 102. https://doi.org/10.1186/s12940-024-01139-2 (2024).
3. Colombo, A., Orsi, F. & Bonfanti, P. Exposure to the organophosphorus pesticide chlorpyrifos inhibits acetylcholinesterase activity and affects muscular integrity in Xenopus laevis larvae. *Chemosphere* **61**, 1665–1671. https://doi.org/10.1016/j.chemosphere.2005.04.005 (2005).
4. Soderlund, D. M. et al. Mechanisms of pyrethroid neurotoxicity: implications for cumulative risk assessment. *Toxicology* **171**, 3–59. https://doi.org/10.1016/s0300-483x(01)00569-8 (2002).
5. Yin, C. et al. Knockdown of the nicotinic acetylcholine receptor β1 subunit decreases the susceptibility to five neonicotinoid insecticides in whitefly (*Bemisia tabaci*). *J Agric Food Chem* **71**, 7221–7229. https://doi.org/10.1021/acs.jafc.3c00782 (2023).
6. Okeke, E. S. et al. Ecotoxicological impact of dinotefuran insecticide and its metabolites on non-targets in agroecosystem: Harnessing nanotechnology- and bio-based management strategies to reduce its impact on non-target ecosystems. *Environ. Res.* **243**, 117870. https://doi.org/10.1016/j.envres.2023.117870 (2024).
7. Wang, R., Wang, Z., Yang, H., Wang, Y. & Deng, A. Highly sensitive and specific detection of neonicotinoid insecticide imidacloprid in environmental and food samples by a polyclonal antibody-based enzyme-linked immunosorbent assay. *J Sci Food Agric* **92**, 1253–1260. https://doi.org/10.1002/jsfa.4691 (2012).
8. Goedjen, G. J., Capel, P. D., Barry, J. D. & Arnold, W. A. Occurrence and distribution of neonicotinoids and fiproles within groundwater in Minnesota: Effects of lithology, land use and geography. *Sci Total Environ* **954**, 176411. https://doi.org/10.1016/j.scitotenv.2024.176411 (2024).
9. Zhou, Y. et al. Development of a fast and sensitive method for measuring multiple neonicotinoid insecticide residues in soil and the application in parks and residential areas. *Anal Chim Acta* **1016**, 19–28. https://doi.org/10.1016/j.aca.2018.02.047 (2018).
10. Li, X. et al. Neonicotinoid residues in fruits and vegetables in Shenzhen: Assessing human exposure and health risks. *Chemosphere* **364**, 143267. https://doi.org/10.1016/j.chemosphere.2024.143267 (2024).
11. Shang, N. et al. Exposure levels and health implications of fungicides, neonicotinoid insecticides, triazine herbicides and their associated metabolites in pregnant women and men. *Environ. Pollut.* **342**, 123069. https://doi.org/10.1016/j.envpol.2023.123069 (2024).
12. Yamamuro, T., Ohta, H., Aoyama, M. & Watanabe, D. Simultaneous determination of neonicotinoid insecticides in human serum and urine using diatomaceous earth-assisted extraction and liquid chromatography-tandem mass spectrometry. *J Chromatogr B Analyt Technol Biomed Life Sci* **969**, 85–94. https://doi.org/10.1016/j.jchromb.2014.06.008 (2014).
13. Loser, D. et al. Functional alterations by a subgroup of neonicotinoid pesticides in human dopaminergic neurons. *Arch Toxicol* **95**, 2081–2107. https://doi.org/10.1007/s00204-021-03031-1 (2021).
14. Wang, Y. et al. A comprehensive review on the pretreatment and detection methods of neonicotinoid insecticides in food and environmental samples. *Food Chem X* **15**, 100375. https://doi.org/10.1016/j.fochx.2022.100375 (2022).
15. Zhang, H. et al. Exposure to neonicotinoid insecticides and their characteristic metabolites: Association with human liver cancer. *Environ Res* **208**, 112703. https://doi.org/10.1016/j.envres.2022.112703 (2022).
16. Nguyen, T., Vennatt, J., Downs, L., Surabhi, V. & Stanietzky, N. Advanced imaging of hepatocellular carcinoma: A review of current and novel techniques. *J Gastrointest Cancer* **55**, 1469–1484. https://doi.org/10.1007/s12029-024-01094-8 (2024).
17. Huang, S. et al. Exploring extrahepatic metastasis of hepatocellular carcinoma based on methylation driver genes and establishing a prognostic model for hepatocellular carcinoma. *Gene* **933**, 148937. https://doi.org/10.1016/j.gene.2024.148937 (2025).
18. VoPham, T. et al. Pesticide exposure and hepatocellular carcinoma risk: A case-control study using a geographic information system (GIS) to link SEER-Medicare and California pesticide data. *Environ Res* **143**, 68–82. https://doi.org/10.1016/j.envres.2015.09.027 (2015).
19. Kanehisa, M., Furumichi, M., Sato, Y., Matsuura, Y. & Ishiguro-Watanabe, M. KEGG: Biological systems database as a model of the real world. *Nucleic Acids Res* **53**, D672–D677. https://doi.org/10.1093/nar/gkae909 (2025).
20. Liu, Y. et al. Exploring the potential mechanisms of Tongmai Jiangtang capsules in treating diabetic nephropathy through multi-dimensional data. *Front Endocrinol (Lausanne)* **14**, 1172226. https://doi.org/10.3389/fendo.2023.1172226 (2023).
21. Liu, Y., Wang, W., Cui, X., Lyu, J. & Xie, Y. Exploring genetic associations of 3 types of risk factors with ischemic stroke: An integrated bioinformatics study. *Stroke* **55**, 1619–1628. https://doi.org/10.1161/strokeaha.123.044424 (2024).
22. Cheng, M. et al. Exploring the mechanism of PPCPs on human metabolic diseases based on network toxicology and molecular docking. *Environ. Int.* **196**, 109324. https://doi.org/10.1016/j.envint.2025.109324 (2025).
23. Kanehisa, M. & Goto, S. KEGG: kyoto encyclopedia of genes and genomes. *Nucleic Acids Res* **28**, 27–30. https://doi.org/10.1093/nar/28.1.27 (2000).
24. Guo, Y. et al. Pesticide thiamethoxam in seed treatment: Uptake, metabolic transformation and associated synergistic effects against wheat aphids. *Sci Total Environ* **949**, 174955. https://doi.org/10.1016/j.scitotenv.2024.174955 (2024).
25. Nixon, L. J. & Leskey, T. C. Evaluation of insecticide residues against spotted lanternfly (*Hemiptera*: *Fulgoridae*). *J Econ Entomol* **117**, 1582–1587. https://doi.org/10.1093/jee/toae106 (2024).
26. Liu, C.-P. et al. An exceptional water stable terbium-based metal-organic framework for selective detection of pesticides. *RSC Adv* **14**, 35220–35226. https://doi.org/10.1039/d4ra06622g (2024).
27. Kariyama, K. et al. EZ-ALBI score for predicting hepatocellular carcinoma prognosis. *Liver Cancer* **9**, 734–743. https://doi.org/10.1159/000508971 (2020).

28. Shen, S., Dean, D. C., Yu, Z. & Duan, Z. Role of cyclin-dependent kinases (CDKs) in hepatocellular carcinoma: Therapeutic potential of targeting the CDK signaling pathway. *Hepatol Res* **49**, 1097–1108. https://doi.org/10.1111/hepr.13353 (2019).
29. Chen, D. et al. Identification of subclusters and prognostic genes based on glycolysis/gluconeogenesis in hepatocellular carcinoma. *Front Immunol* **14**, 1232390. https://doi.org/10.3389/fimmu.2023.1232390 (2023).
30. Shang, R.-Z., Qu, S.-B. & Wang, D.-S. Reprogramming of glucose metabolism in hepatocellular carcinoma: Progress and prospects. *World J Gastroenterol* **22**, 9933–9943. https://doi.org/10.3748/wjg.v22.i45.9933 (2016).
31. Shang, F. et al. The anti-angiogenic effect of dexamethasone in a murine hepatocellular carcinoma model by augmentation of gluconeogenesis pathway in malignant cells. *Cancer Chemother Pharmacol* **77**, 1087–1096. https://doi.org/10.1007/s00280-016-3030-x (2016).
32. Gao, B. et al. Metabolic reprogramming in hepatocellular carcinoma: mechanisms of immune evasion and therapeutic implications. *Front Immunol* **16**, 1592837. https://doi.org/10.3389/fimmu.2025.1592837 (2025).
33. Villeneuve, J.-P. & Pichette, V. Cytochrome P450 and liver diseases. *Curr Drug Metab* **5**, 273–282. https://doi.org/10.2174/1389200043335531 (2004).
34. Sun, H. et al. Up-Regulation of CYP2C19 expression by BuChang NaoXinTong via PXR activation in HepG2 cells. *PLoS ONE* **11**, e0160285. https://doi.org/10.1371/journal.pone.0160285 (2016).
35. Shoieb, S. M., El-Ghiaty, M. A., Alqahtani, M. A. & El-Kadi, A. O. S. Cytochrome P450-derived eicosanoids and inflammation in liver diseases. *Prostaglandins Other Lipid Mediat* **147**, 106400. https://doi.org/10.1016/j.prostaglandins.2019.106400 (2020).
36. Flannery, P. C., Abbott, K. L. & Pondugula, S. R. Correlation of PPM1A downregulation with CYP3A4 repression in the tumor liver tissue of hepatocellular carcinoma patients. *Eur J Drug Metab Pharmacokinet* **45**, 297–304. https://doi.org/10.1007/s13318-019-00595-3 (2020).
37. Ashida, R. et al. CYP3A4 gene is a novel biomarker for predicting a poor prognosis in hepatocellular carcinoma. *Cancer Genom Proteom* **14**, 445–453. https://doi.org/10.21873/cgp.20054 (2017).
38. Zhang, W., Zhang, Y., Wan, Y., Liu, Q. & Zhu, X. A bile acid-related prognostic signature in hepatocellular carcinoma. *Sci Rep* **12**, 22355. https://doi.org/10.1038/s41598-022-26795-7 (2022).
39. Zhang, J. et al. MicroRNA-185 modulates CYP7A1 mediated cholesterol-bile acid metabolism through post-transcriptional and post-translational regulation of FoxO1. *Atherosclerosis* **348**, 56–67. https://doi.org/10.1016/j.atherosclerosis.2022.03.007 (2022).
40. Stepien, M. et al. Prediagnostic alterations in circulating bile acid profiles in the development of hepatocellular carcinoma. *Int J Cancer* **150**, 1255–1268. https://doi.org/10.1002/ijc.33885 (2022).
41. Ghanem, A., Kitanovic, A., Holzwarth, J. & Wölfl, S. Mutational analysis of fructose-1,6-bis-phosphatase FBP1 indicates partially independent functions in gluconeogenesis and sensitivity to genotoxic stress. *Microb Cell* **4**, 52–63. https://doi.org/10.15698/mic2017.02.557 (2017).
42. Yang, J. et al. Inhibiting histone deacetylases suppresses glucose metabolism and hepatocellular carcinoma growth by restoring FBP1 expression. *Sci Rep* **7**, 43864. https://doi.org/10.1038/srep43864 (2017).
43. Li, F. et al. FBP1 loss disrupts liver metabolism and promotes tumorigenesis through a hepatic stellate cell senescence secretome. *Nat Cell Biol* **22**, 728–739. https://doi.org/10.1038/s41556-020-0511-2 (2020).
44. Cheng, Z. et al. Endoplasmic reticulum stress promotes hepatocellular carcinoma by modulating immunity: A study based on artificial neural networks and single-cell sequencing. *J Transl Med* **22**, 658. https://doi.org/10.1186/s12967-024-05460-9 (2024).
45. Guo, D. et al. THBS4 promotes HCC progression by regulating ITGB1 via FAK/PI3K/AKT pathway. *FASEB J* **34**, 10668–10681. https://doi.org/10.1096/fj.202000043r (2020).
46. Su, F. et al. Over-expression of Thrombospondin 4 correlates with loss of miR-142 and contributes to migration and vascular invasion of advanced hepatocellular carcinoma. *Oncotarget* **8**, 23277–23288. https://doi.org/10.18632/oncotarget.15054 (2017).
47. Wu, L. et al. Orchestrated metal-coordinated carrier-free celastrol hydrogel intensifies T cell activation and regulates response to immune checkpoint blockade for synergistic chemo-immunotherapy. *Biomaterials* **312**, 122723. https://doi.org/10.1016/j.biomaterials.2024.122723 (2025).
48. Yang, Y.-L. et al. T cells, NK cells, and tumor-associated macrophages in cancer immunotherapy and the current state of the art of drug delivery systems. *Front Immunol* **14**, 1199173. https://doi.org/10.3389/fimmu.2023.1199173 (2023).
49. Xu, Y. et al. TRAF2 promotes M2-polarized tumor-associated macrophage infiltration, angiogenesis and cancer progression by inhibiting autophagy in clear cell renal cell carcinoma. *J Exp Clin Cancer Res* **42**, 159. https://doi.org/10.1186/s13046-023-02742-w (2023).
50. Chen, D.-P. et al. Peritumoral monocytes induce cancer cell autophagy to facilitate the progression of human hepatocellular carcinoma. *Autophagy* **14**, 1335–1346. https://doi.org/10.1080/15548627.2018.1474994 (2018).
51. Qiu, H. et al. The positive feedback loop of furin and TGFβ1 enhances the immune responses of Tregs to hepatocellular carcinoma cells and hepatitis B virus in vitro. *Cell Biol Int* **46**, 1215–1226. https://doi.org/10.1002/cbin.11806 (2022).

Acknowledgements

This work was supported by the National Natural Science Foundation of China (Grant Nos. 81860492 and 81660491) and Jiangxi University of Chinese Medicine Research Fund (Grant No. 2020BSZR001).

Author contributions

Chenghao Wang: Formal analysis, Writing–original draft. Zhichao Wang: Formal analysis, Writing–review & editing. Yinqin Hu & Shi Xiang: Data curation. Hui Chen & Chao Yang: Conceptualization, Writing–review & editing.

Declarations

Competing 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.

Additional information

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1038/s41598-025-11792-3.

Correspondence and requests for materials should be addressed to H.C. or C.Y.