

SPxLINCS: a GWAS-guided, tissue-resolved framework for drug repurposing in hepatocellular carcinoma

Jiahuizi Peng

jp2824@ic.ac.uk

Imperial College London, the United Kingdom

Research Article

Keywords: hepatocellular carcinoma, drug repurposing, genome-wide association study, reverse transcriptomic signature, network pharmacology, drug set enrichment analysis

Posted Date: October 27th, 2026

DOI: <https://doi.org/10.21203/rs.3.rs-11267412/v1>

License: © ⓘ This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Additional Declarations: No competing interests reported.

SPxLINCS: a GWAS-guided, tissue-resolved framework for drug repurposing in hepatocellular carcinoma

Jiahuizi Peng^{1*}

¹ Department of Mathematics, Imperial College London, London, the United Kingdom, SW72AZ; Email: jp2824@ic.ac.uk

* correspondence to: Jiahuizi Peng, email: jp2824@ic.ac.uk

Abstract

Background Hepatocellular carcinoma (HCC) is the third leading cause of cancer-related mortality, yet effective therapeutic options remain limited. Although Genome-Wide Association Studies (GWAS) have identified multiple HCC susceptibility loci, translating these discoveries into therapeutic leads has proven challenging. Here we present a reverse transcriptomic repurposing pipeline that combines evidence from multiple GWAS, tissue-resolved transcriptomics, and small-molecule perturbational profiles to identify candidate drugs for HCC.

Materials and methods We developed a reproducible drug-repositioning pipeline linking 12 HCC-related GWAS datasets with multi-tissue GTEx regulatory profiles and LINCS L1000 perturbational signatures through S-PrediXcan and LINCS-based transcriptomic connectivity analysis (SPxLINCS). Previously calculated compound-level evidence scores were retained for ranking and visualization. Candidate networks were visualized in Cytoscape as bipartite and tripartite graphs, and clinical plausibility was assessed by Drug Set Enrichment Analysis (DrugEnrichr vs. SIDER) with Benjamini–Hochberg correction.

Results Across the 12 datasets, the SPxLINCS pipeline prioritized compounds that could be grouped into five mechanism-related categories after key-tissue and reverse-signature filtering: cell-cycle and mitotic kinase inhibitors, including CDK- and PLK1-related compounds; focal adhesion kinase inhibitors; DNA-dependent protein kinase and DNA-damage-response agents; topoisomerase I inhibitors; and translation-related inhibitors. Network visualization distinguished compounds that recurred across multiple GWAS datasets from compounds primarily associated with individual phenotypes. Drug-set enrichment analysis identified enrichment of terms related to hepatocellular and other hepatic malignancy indications in the SIDER database. Importazole, ER-27319, and LDN-193189 were recurrently prioritized across multiple datasets and tissues despite limited prior characterization in HCC. These compounds should therefore be regarded as computationally nominated and hypothesis-generating candidates requiring experimental validation.

Conclusions This multi-GWAS reverse transcriptomic approach yields a mechanism-guided

candidate list for drug repurposing in HCC. Signal consistency across independent datasets, combined with indication-level enrichment, supports advancing prioritized compounds into biomarker-stratified preclinical and early clinical evaluation.

Keywords hepatocellular carcinoma; drug repurposing; genome-wide association study; reverse transcriptomic signature; network pharmacology; drug set enrichment analysis

Introduction

Hepatocellular carcinoma (HCC) is the predominant histological subtype of primary liver cancer, accounting for an estimated 661,000 new cases -- approximately 80% of all liver cancers -- globally in 2018 and ranks as the third leading cause of cancer-related mortality worldwide [1,2]. Although systemic therapies including multikinase inhibitors and immune checkpoint inhibitors have been approved for HCC, responses are frequently limited by therapeutic resistance and tumor heterogeneity.

Genetic susceptibility contributes to HCC risk, and recent genome-wide association studies (GWAS) have identified multiple susceptibility loci across diverse etiologies and populations [3]. Most of these variants fall in non-coding sequence, complicating the identification of causal genes and effector mechanisms. The Genotype-Tissue Expression (GTEx) project was established to characterize the effects of genetic variation on gene expression across human tissues, providing a critical resource for mapping GWAS signals to tissue-specific regulatory mechanisms [4]. Transcriptome-wide association studies (TWAS), implemented through tools such as S-PrediXcan, leverage GTEx-derived prediction models to impute genetically regulated gene expression and identify disease-associated transcriptomic signatures at tissue resolution, thereby connecting GWAS loci to downstream biological pathways [5].

The Connectivity Map (CMap) framework can be performed to turn disease transcriptomics into drug hypotheses, which identifies functional relationships between diseases and drugs by comparing disease-associated gene expression signatures against a reference compendium of compound-induced transcriptional profiles from treated cell lines [6]. The core principle that compounds reversing a disease transcriptomic signature may exert therapeutic effects has been widely applied across computational drug repurpose and mechanistic pharmacology. Building on this basis, the LINCS L1000 project substantially expanded the scale and diversity of perturbational transcriptomic data by profiling thousands of small molecules across multiple cell lines and treatment conditions, enabling systematic inference of drug-gene-disease relationships at a significant scale [7-9]. The integration of genetically informed disease signatures with LINCS L1000 perturbational profiles thus offers a direct approach to nominating repurposable compounds for complex diseases such as HCC.

Conventional drug discovery is a costly and time-consuming process, typically requiring over

a decade and substantial financial investment to bring a single compound to market. Drug repurposing identifies new therapeutic indications for existing, clinically approved medications, offering a compelling alternative by leveraging established safety profiles, known pharmacokinetics, and existing regulatory data to accelerate clinical translation [10]. Drug repurposing has also attracted substantial interest in NSCLC, where compounds originally developed for non-oncological indications have been explored as potential anticancer agents using computational, experimental, and nanoparticle-based approaches [11]. For HCC, where the approved drug pipeline remains narrow, repurposing represents a practical route to identifying new treatment options with reduced developmental risk and then informing subsequent rounds of targeted drug development.

To evaluate the clinical plausibility of computationally nominated drug candidates, Drug Set Enrichment Analysis (DSEA) provides an independent validation step at the pharmacological level. Analogous to gene set enrichment analysis applied at the drug level, DSEA assesses whether a prioritized set of compounds is statistically enriched for shared therapeutic indications or mechanisms of action by integrating drug-induced transcriptional responses with curated indication and side-effect databases [12,13]. This approach enables systematic evaluation of whether nominated candidates converge onto disease-relevant clinical phenotypes, thereby supporting translational relevance beyond computational prediction.

In this study, we describe an integrative, multi-GWAS drug repurposing framework for HCC that extends the classical CMap paradigm by anchoring disease signatures in genetically regulated gene expression. By combining tissue-specific transcriptomic signatures derived from S-PrediXcan across 12 independent HCC GWAS datasets with LINCS L1000 perturbational profiles [14], we identify compounds whose transcriptional effects oppose the HCC disease signature. Candidate drugs are further evaluated through network-based visualization and DSEA to assess cross-dataset reproducibility and indication-level clinical plausibility. This workflow is scalable across GWAS datasets and provides a reproducible route from genetic association to mechanism-guided drug candidates, which may accelerate the identification of mechanism-guided therapeutic candidates for HCC.

Materials and methods

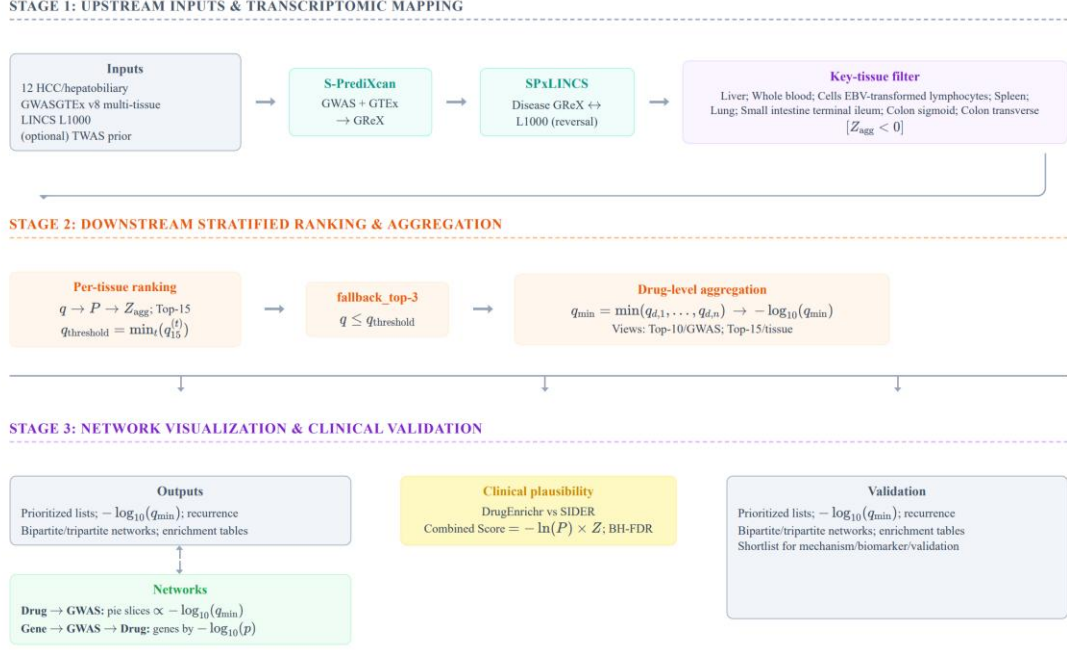


Figure 1. End-to-end HCC drug repurposing workflow (SPxLINCS). The pipeline integrates 12 HCC/hepatobiliary GWAS datasets with GTEx v8 multi-tissue transcriptomics via S-PrediXcan and LINCS L1000 perturbational profiles. Candidates are filtered by key-tissue selection and reverse transcriptomic signatures ($Z_{agg} < 0$), ranked by drug-level aggregation ($q_{\min}$), and visualized as bipartite and tripartite networks. Clinical plausibility is assessed by DrugEnrichr vs. SIDER with BH-FDR correction; all q -values are BH-FDR corrected.

Conceptual Framework

We employed the principle of reverse transcriptomic profiling for computational drug repurposing [15]. The overall SPxLINCS workflow is shown in Figure 1. Under this framework, a drug whose perturbational profile is strongly anti-correlated with the disease signature is nominated as a candidate to pharmacologically oppose the disease process. A strongly negative connectivity score therefore indicates that the compound has the transcriptional capacity to counteract HCC-associated oncogenic signaling, potentially restoring gene expression toward a physiologically normal state.

Data preparation

To ensure rigorous alignment between disease features and drug perturbational profiles, we intersected gene sets derived from 12 HCC GWAS summary statistics with the LINCS L1000 Level-5 dataset (GSE92742). Disease-level gene effects were first generated by processing each GWAS through S-PrediXcan across all 49 GTEx tissues, yielding tissue-specific Z -scores for genetically regulated gene expression [5]. Gene symbols were standardized to uppercase and deduplicated by averaging within each file.

For the LINCS L1000 dataset, gene identifiers were harmonized by constructing an Entrez-to-

symbol mapping table derived from the LINCS gene information file [16]. Row metadata fields were screened to identify the field with the greatest overlap with the S-PrediXcan gene set; duplicate entries were resolved by retaining the first occurrence to maintain one-to-one column mapping in the expression matrix. This procedure yielded a synchronized gene set of 11,358 overlapping genes mapped across 7,169 curated small-molecule signatures, establishing a standardized matrix for downstream connectivity scoring.

Data Processing and Filtering

To highlight high-quality and biologically robust signals, we applied a four-step curation procedure to the LINCS L1000 dataset. First, only small-molecule perturbations ($pert_type = trt_cp$) were retained; all vehicle and negative controls such as DMSO and PBS were excluded. Second, signatures were required to meet minimum quality thresholds: distillation signature strength ($distil_ss \geq 4.0$) and transcriptional activity score ($TAS \geq 0.35$). Third, the analysis was restricted to LINCS-defined Touchstone reference compounds, comprising well-annotated, clinical-grade drugs and standardized pharmacological tools. Fourth, where multiple experimental replicates existed for the same perturbation context (identical $pert_id \times cell_id \times pert_time$ combination), only the signature with the highest TAS was retained, eliminating redundancy without discarding biological signal.

Prior to connectivity scoring, each signature column was winsorized at the 1st and 99th percentiles to reduce the influence of extreme values, then subjected to per-column z-scoring (mean = 0, variance = 1) to harmonize expression scales across perturbational profiles.

Spatial correlation analysis

The screening and refinement workflow is summarized in Figure 2.

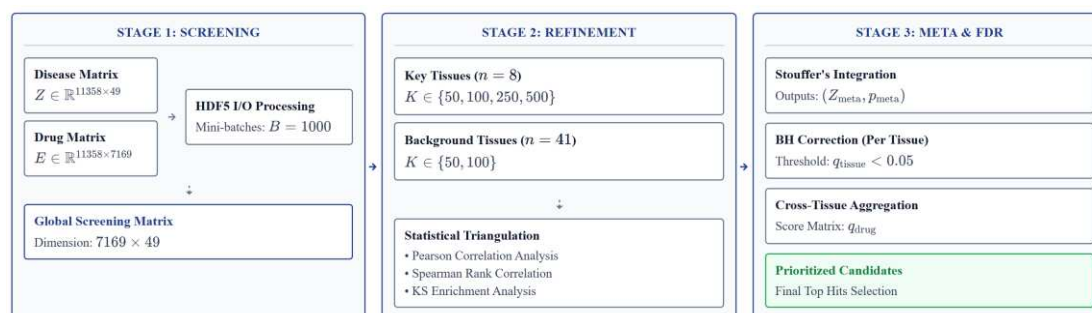


Figure 2. Spatial correlation analysis (SPxLINCS) workflow. Stage 1 screens all 7,169 drug signatures against the disease matrix ($Z \in \mathbb{R}^{11358 \times 49}$) in mini-batches of $B = 1,000$ via HDF5 I/O, yielding a $7,169 \times 49$ global screening matrix. Stage 2 applies tissue-stratified refinement: key tissues ($n = 8$, $K \in \{50, 100, 250, 500\}$) and background tissues ($n = 41$, $K \in \{50, 100\}$) are each evaluated by Pearson correlation, Spearman rank correlation, and KS enrichment. Stage 3 integrates per-analysis p-values via Stouffer's method, applies BH correction per tissue

($q_{tissue} < 0.05$), and aggregates across tissues to a drug-level score matrix q_{drug} .

Stage1: High-Throughput Connectivity Profiling

To identify candidate drug signatures capable of reversing the HCC transcriptomic program, we first performed a broad macro-scale screening across all 49 GTEx/LINCS tissues for the full set of 7,169 curated drug perturbation signatures.

Let $Z \in \mathbb{R}^{m \times n}$, with $m = 11,358$ genes and $n = 49$ tissues, denote the disease matrix, and $E \in \mathbb{R}^{m \times x}$, with $x = 7,169$, denote the normalized drug perturbation matrix. To keep computation tractable, signatures were processed in mini-batches of size $B = 1,000$ with data loading handled via HDF5 I/O. For each batch $E_{batch} \in \mathbb{R}^{m \times B}$, tissue-wise Pearson correlations against the reversed drug profiles were computed as:

$$r = \frac{Z^T(-E_{batch})}{m - 1}$$

yielding a correlation matrix $r \in \mathbb{R}^{n \times B}$ for each batch [17]. Each correlation coefficients was then converted to t -statistics using the standard transformation:

$$t = r \sqrt{\frac{m - 2}{1 - r^2}}$$

Two-tailed Student's t -tests were then applied to derive the corresponding p -value for each drug-tissue pair [18]. This procedure produced a $7,169 \times 49$ matrix of preliminary tissue-specific connectivity results for use in the subsequent refinement and meta-analysis.

Stage2: Tissue-Stratified Refinement and Meta-Analysis Phase

Signatures passing the initial screen were carried forward into a more stringent, multi-dimensional validation framework intended to suppress false positives and isolate high-confidence therapeutic leads. All 7,169 signatures were re-evaluated under this refined pipeline.

Tissues were divided into two biologically motivated strata; each assigned a distinct range of top-gene subset sizes K :

- **Key tissues:** liver (primary site of HCC), whole blood, EBV-transformed lymphocytes, spleen, lung, small intestine terminal ileum, colon sigmoid, and colon transverse. These tissues collectively represent the tumor immune microenvironment and the most clinically relevant metastatic sites. Applied subset sizes: $K \in \{50, 100, 250, 500\}$
- **Background tissues:** the remaining 41 tissues, serving as reference controls for systemic or non-specific signals: $K \in \{50, 100\}$

For every tissue-signature pair sharing at least $\text{min_genes} = 100$ overlapping genes, three

complementary analyses were carried out:

Analysis 1 — Full-gene Pearson correlation:

$$r_{full} = \text{corr}(Z, -E)$$

with a two-tailed p -value, quantifying linear transcriptomic reversal between the disease signature and the negated drug perturbation profile [17].

Analysis 2 — Full-gene Spearman rank correlation:

$$\rho_{full} = \text{corr}_{rank}(Z, -E)$$

with a two-tailed p -value. Gene expression values and reversed drug perturbation scores were converted to ranks over $[1, m]$, where $m = 11,358$. The resulting rank-based statistic is robust to distributional outliers and non-linear expression relationships that Pearson correlation alone may fail to capture [19].

Analysis 3 — Top- K connectivity with KS-style enrichment:

To detect localized reversal signals among the most strongly deregulated disease genes, a Kolmogorov-Smirnov (KS) enrichment score was applied [20]. The top and bottom K disease genes were selected to represent the upregulated and downregulated disease gene sets, respectively. The contribution of each gene in the drug-induced expression profile was weighted according to the absolute magnitude of its expression change:

$$w_i = |E_i|^\alpha,$$

where E_i denotes the drug-induced expression value for gene i . The primary analysis used $\alpha = 1$ as a simple linear weighting scheme, such that genes with larger absolute expression changes contributed proportionally more to the enrichment score. Weighted enrichment scores were calculated separately for the upregulated and downregulated disease gene sets, and the connectivity score was defined as:

$$CS = ES_{\text{up}} - ES_{\text{dn}}.$$

Empirical P values were estimated from 200 permutation iterations for each tissue–signature pair. Pearson and Spearman correlations were additionally calculated using the union of the top and bottom K disease genes at each value of K , providing complementary local correlation measures alongside the KS-style enrichment score.

Stage3: Meta-analysis and FDR correction:

Stouffer's Z -score method with signed directions was applied to integrate the Pearson, Spearman, and multiple KS p -values into a unified meta-score ($Z_{\text{meta}}, p_{\text{meta}}$) [21].

Correlation-based statistics contributed their natural signs, while KS scores contributed via $-\tau$. Within each tissue, multiple contexts per drug (*pert_id*) were aggregated by Stouffer's method using signs from Z_{meta} , yielding tissue-level aggregated statistics $p_{agg,tissue}$ and $Z_{agg,tissue}$. Benjamini-Hochberg FDR correction was then applied within each tissue to obtain q_{tissue} ($q < 0.05$). Across tissues, drug-level p-values ($p_{agg,tissue}$, signed by $Z_{agg,tissue}$) were further combined via Stouffer's method to derive a global $p_{agg,drug}$, with BH correction yielding q_{drug} . Results were reported as per-tissue top candidates using a default display size of $N = 15$, together with complete pairwise and aggregated tables for transparency and downstream prioritization.

Cross-dataset prioritization of drug repurposing candidates

For each of the 12 HCC or HCC-related GWAS datasets, drug prioritization was performed using the corresponding SPxLINCS output (Supplementary Tables 1-12). Because a compound could occur across multiple tissue-context combinations within a given GWAS-specific table, the previously calculated compound-level evidence score was retained unchanged for drug-level ranking and visualization.

Within each GWAS dataset, compounds were ranked according to the existing drug-level evidence score. The tenth-ranked compound served as the dataset-specific ranking boundary. For display, the corresponding statistical evidence was transformed to the $-\log_{10}(q)$ scale and shown in a faceted bar plot with one panel per GWAS dataset.

This procedure provides a standardized, drug-centered overview of the compounds prioritized within each dataset before tissue restriction or directionality filtering. The top 10 lists were used for descriptive comparison only. Recurrence of a compound among multiple GWAS-specific top 10 lists was treated as a descriptive indicator of consistency and was not interpreted as an additional significance test or as evidence of cross-GWAS global significance.

Selection of key-tissue drug repositioning signals across 12 GWAS datasets

GWAS-specific SPxLINCS tables (Supplementary Tables 1-12) from 12 hepatobiliary cancer-related datasets were grouped into four predefined categories: duct-related, FinnGen, Biobank-based HCC, and HCC-related.

To prioritize biologically relevant signals, analyses were restricted to a predefined tissue set: one strong-priority tissue (Liver) and intermediate-priority tissues (Whole blood, Cells EBV-transformed lymphocytes, Spleen, Lung, Small intestine terminal ileum, Colon sigmoid, and Colon transverse). Only records from these tissues were retained. To focus on reverse-signature evidence compatible with therapeutic reversal, we kept entries with negative aggregated tissue-level z-scores:

$$Z_{agg,tissue} < 0$$

Within each tissue, retained drugs were ranked by ascending adjusted significance (q value), then by ascending nominal significance (P value), and finally by ascending $Z_{\text{agg,tissue}}$. We initially extracted the top 15 drugs per tissue. For tissues contributing at least 15 eligible drugs, we recorded the q value at rank 15 and defined a GWAS-specific global threshold as the minimum of these tissue-specific 15th-rank q values:

$$q_{\text{threshold}} = \min_{t \in \mathcal{T}} (q_{15}^{(t)})$$

Drugs with $q \leq q_{\text{threshold}}$ were retained as the primary key-tissue set for that GWAS. For tissues with no drugs surviving this threshold, a fallback retained the top 3 reverse-signature drugs for that tissue. This produced one filtered output per GWAS. Across the 12 GWAS datasets, we summarized the count of retained drugs per tissue into a single table. Representative filtered drug tables from the two categories (East Asia, Europe) were shown in the main text.

GWAS-resolved drug-trait network with drug-centered pie-chart annotation

To summarize GWAS-specific repositioning signals across hepatobiliary phenotypes, we constructed a bipartite drug-GWAS network in Cytoscape using the key-tissue outputs (*top_drug* directory) [22]. For each GWAS-specific file, drug names and the previously calculated compound-level evidence scores were extracted. These scores were retained unchanged for visualization and displayed on a $-\log_{10}(q)$ scale.

Drug–GWAS edges were defined for every observed drug–trait pair and annotated with the number of occurrences within the corresponding GWAS-specific output. Drug nodes were assigned GWAS-specific attributes containing the corresponding evidence displayed scores, enabling pie-chart visualization of cross-GWAS profiles. Each pie-chart slice represented one GWAS dataset, and slice area was proportional to the displayed drug-level evidence score. GWAS nodes were displayed as labeled rectangles, and edge width was mapped to within-GWAS recurrence.

The network was used to visualize the distribution and recurrence of prioritized compounds across GWAS datasets. Network recurrence was treated as a descriptive measure of consistency.

Construction of an integrated tripartite gene-GWAS-drug network

To integrate genetically prioritized susceptibility genes with drug repositioning candidates, we constructed a three-layer network comprising gene, GWAS trait, and drug nodes. The 12 GWAS datasets were arranged into four predefined groups (duct-related, Biobank-based, FinnGen, and HCC-related), with one tripartite network constructed per group, each covering three GWAS datasets:

- **Duct-related:** Liver and bile duct cancer (Europe); Hepatic bile duct cancer (East Asia);

Hepatic bile duct cancer (Europe)

- **Biobank-based:** Self-reported hepatocellular carcinoma (Europe, UK Biobank); ICD-10 liver cell carcinoma (Europe, UK Biobank); Hepatocellular carcinoma (Japan, BioBank Japan)
- **FinnGen:** Benign liver and bile ducts (*exallc*, Europe, FinnGen); Benign liver (*exallc*, Europe, FinnGen); Liver intrahepatic bile ducts (*exallc*, Europe, FinnGen)
- **HCC-related:** Hepatic cancer (Europe); Hepatic cancer (East Asia); Alcohol related hepatocellular carcinoma (Europe)

Gene layer. TWAS gene signals were obtained from a cross-tissue TWAS of HCC [5]. For each of the three traits within a group, records were extracted, and duplicate gene entries were collapsed by retaining the entry with the highest association significance, defined as:

$$S_{gene} = -\log_{10}(p)$$

Genes were ranked within each trait by descending S_{gene} , and the top 15 genes per trait were retained, preserving p -value, Z -score, and tissue annotation.

Drug layer. For each GWAS-specific drug table, drug names and the previously calculated compound-level evidence scores were extracted. These scores were retained unchanged for network construction and visualization. Drug recurrence was recorded as the number of appearances within the corresponding GWAS-specific table, and the top 15 drugs per GWAS dataset were retained according to the existing drug-level ranking.

Network assembly. The final edge list comprised Gene $\rightarrow$ GWAS and Drug $\rightarrow$ GWAS edges. Nodes were typed as Gene, GWAS, or Drug, and arranged in a fixed three-layer layout: genes (upper layer), GWAS traits (middle layer), and drugs (lower layer). The three GWAS nodes within each group were positioned horizontally according to predefined coordinates. Display labels showed gene symbols, drug names, and trait names without prefixes. In Cytoscape, gene nodes were rendered as blue triangles, drug nodes as green circles, and GWAS trait nodes as red rectangles. Gene–GWAS edge width was mapped to $-\log_{10}(p)$; Drug–GWAS edges used uniform width for visual clarity; all edges were colored gray.

Downstream target-set pathway enrichment analysis

To elucidate the intracellular mechanisms associated with the 103 candidate small molecules in HCC, predicted human protein targets of the drug set were extracted and mapped against KEGG pathways via the L1000 Fast Ward (L1000FWD) annotation platform. Pathway enrichment was evaluated by Fisher's exact test, with pathways prioritized by P -value and enrichment score. To account for multiple comparisons, the Benjamini-Hochberg procedure was applied; both raw p -values and FDR-adjusted q -values are reported throughout.

Drug-set enrichment analysis (DSEA) for clinical indications

To independently validate the clinical relevance of nominated candidates independently of target- or pathway-based inference, drug-set enrichment analysis (DSEA) was conducted using DrugEnrichr against the SIDER indication database [23]. The curated list of unique drug names was directly cross-referenced to SIDER, and enriched indication terms were ranked by the Combined Score:

$$CS = -\ln(P) \times Z$$

where $-\ln(P)$ reflects the statistical significance of enrichment and Z represents the deviation from the expected rank. Multiple testing was controlled by Benjamini-Hochberg FDR correction.

Results

Table 1. Hepatic cancer (East Asia)

Drug	Cell line	Tissue	Rank in tissue	P value	q value
KU-0060648	AGS (stomach)	Cells EBV-transformed lymphocytes	1	2.49E-44	3.83E-41
erythrosine	PC3 (prostate)	Colon transverse	1	2.78E-34	4.27E-31
JNK-9L	HCC515 (lung)	Liver	1	1.46E-31	2.24E-28
erythrosine	HT29 (large intestine)	Liver	2	1.06E-30	8.15E-28
PHA-793887	MCF10A (breast)	Lung	1	2.53E-74	3.89E-71
triptolide	HA1E (kidney)	Lung	2	7.64E-67	5.87E-64
alvocidib	HEPG2 (liver)	Lung	3	1.28E-59	6.56E-57
AG-14361	HCC15 (lung)	Lung	4	3.31E-59	1.27E-56
PF-562271	HEPG2 (liver)	Lung	5	1.89E-52	5.80E-50
CGP-60474	BT20 (breast)	Lung	6	1.96E-44	5.03E-42
AT-7519	PC3 (prostate)	Lung	7	2.23E-36	4.89E-34
AS-601245	MCF10A (breast)	Lung	8	8.15E-36	1.57E-33
erythrosine	HT29 (large intestine)	Lung	9	6.31E-34	1.08E-31
JNK-9L	HS578T (breast)	Lung	10	1.75E-33	2.68E-31
ER-27319	A549 (lung)	Lung	11	2.42E-31	3.39E-29
epirubicin	PC3 (prostate)	Lung	12	1.07E-30	1.38E-28

dactinomycin	HEPG2 (liver)	Lung	13	1.38E-29	1.64E-27
KU-0060648	U937 (haematopoietic and lymphoid tissue)	Small intestine terminal ileum	1	2.71E-69	4.16E-66
piperlongumine	PC3 (prostate)	Small intestine terminal ileum	2	2.77E-59	2.13E-56
BI-2536	RMUGS (ovary)	Small intestine terminal ileum	3	2.86E-54	1.46E-51
TG-101348	HT115 (large intestine)	Small intestine terminal ileum	4	2.01E-38	7.73E-36
erythrosine	PC3 (prostate)	Small intestine terminal ileum	5	7.63E-31	2.35E-28
KU-0060648	JHUEM2 (endometrium)	Spleen	1	4.61E-99	7.09E-96
TG-101348	HS578T (breast)	Spleen	2	2.78E-76	2.14E-73
PI-103	PC3 (prostate)	Spleen	3	4.90E-59	2.51E-56
BI-2536	RMUGS (ovary)	Spleen	4	9.49E-57	3.65E-54
LY-294002	HA1E (kidney)	Spleen	5	1.27E-36	3.92E-34
piperlongumine	NPC (central nervous system)	Spleen	6	1.20E-35	3.08E-33
dorsomorphin	CL34 (large intestine)	Spleen	7	4.01E-35	8.81E-33
NVP-AUY922	U937 (haematopoietic and lymphoid tissue)	Spleen	8	2.20E-33	4.24E-31
ZSTK-474	SKBR3 (breast)	Spleen	9	1.36E-30	2.11E-28
mepacrine	JHUEM2 (endometrium)	Spleen	10	1.37E-30	2.11E-28
alvocidib	SKBR3 (breast)	Spleen	11	3.26E-30	4.56E-28
AZD-8055	BT20 (breast)	Spleen	12	5.16E-30	6.62E-28
triptolide	HA1E (kidney)	Spleen	13	7.85E-30	9.29E-28

AG-14361	HCC15 (lung)	Spleen	14	1.22E-29	1.34E-27
PHA-793887	HEPG2 (liver)	Spleen	15	1.54E-28	1.58E-26
KU-0060648	A549 (lung)	Whole blood	1	4.04E-70	6.21E-67
BI-2536	SKBR3 (breast)	Whole blood	2	5.47E-46	4.21E-43
LDN-193189	WSUDLCL2 (haematopoietic and lymphoid tissue)	Whole blood	3	1.50E-43	7.71E-41
TG-101348	H1299 (lung)	Whole blood	4	2.87E-40	1.10E-37
piperlongumine	MCF7 (breast)	Whole blood	5	2.36E-33	7.27E-31

Table 2. Benign liver exallc (Europe, FinnGen)

Drug	Cell line	Tissue	Rank in tissue	P value	q value
narciclasine	A549 (lung)	Cells EBV-transformed lymphocytes	1	2.82E-251	4.34E-248
homoharringtonine	PHH (liver)	Cells EBV-transformed lymphocytes	2	1.55E-140	1.19E-137
digitoxin	HA1E (kidney)	Cells EBV-transformed lymphocytes	3	3.71E-133	1.90E-130
anisomycin	HCC515 (lung)	Cells EBV-transformed lymphocytes	4	2.59E-123	9.96E-121
digitoxigenin	A549 (lung)	Cells EBV-transformed lymphocytes	5	5.31E-111	1.63E-108
digoxin	A549 (lung)	Cells EBV-transformed lymphocytes	6	1.46E-110	3.75E-108
strophanthidin	HUH7 (liver)	Cells EBV-transformed lymphocytes	7	1.01E-109	2.21E-107
cephaeline	A549 (lung)	Cells EBV-transformed lymphocytes	8	1.75E-100	3.36E-98
cycloheximide	A549 (lung)	Cells EBV-transformed lymphocytes	9	7.18E-99	1.23E-96
NSC-632839	PC3 (prostate)	Cells EBV-transformed	10	5.98E-96	9.20E-94

		lymphocytes			
emetine	PC3 (prostate)	Cells EBV-transformed lymphocytes	11	2.95E-89	4.12E-87
ouabain	MCF7 (breast)	Cells EBV-transformed lymphocytes	12	4.95E-85	6.34E-83
periplocymarin	A549 (lung)	Cells EBV-transformed lymphocytes	13	3.28E-80	3.88E-78
cinobufagin	MCF7 (breast)	Cells EBV-transformed lymphocytes	14	6.91E-79	7.59E-77
BI-2536	PC3 (prostate)	Cells EBV-transformed lymphocytes	15	1.82E-72	1.87E-70
NSC-632839	NCIH508 (large intestine)	Colon transverse	1	3.22E-48	4.95E-45
BI-2536	U937 (haematopoietic and lymphoid tissue)	Colon transverse	2	2.11E-41	1.63E-38
AG-14361	PL21 (haematopoietic and lymphoid tissue)	Liver	3	3.90E-40	2.00E-37
helveticoside	PC3 (prostate)	Liver	6	3.59E-24	9.21E-22
fluticasone	A549 (lung)	Liver	12	2.83E-18	3.62E-16
AG-14361	PL21 (haematopoietic and lymphoid tissue)	Small intestine terminal ileum	2	7.43E-21	5.71E-18
dorsomorphin	HT29 (large intestine)	Small intestine terminal ileum	5	9.37E-17	2.88E-14
narciclasine	OV7 (ovary)	Small intestine terminal ileum	6	4.51E-16	1.16E-13
BI-2536	SKLU1 (lung)	Spleen	10	8.25E-30	1.27E-27
dorsomorphin	PC3 (prostate)	Whole blood	12	1.89E-23	2.42E-21
CA-074-Me	HA1E (kidney)	Whole blood	15	4.78E-19	4.90E-17

Table 1-2. Representative key-tissue prioritized drug repositioning results from two GWAS

categories. Representative filtered outputs from two GWAS categories (East Asia, Europe); complete tables for all 12 GWAS are provided in Supplementary Tables 13–24. Candidates were restricted to key tissues and reverse-signature hits ($Z_{agg,tissue} < 0$), ranked within each tissue by adjusted significance (q), nominal significance (P), then $Z_{agg,tissue}$. The top 15 candidates per tissue were initially selected; a GWAS-specific threshold was defined as $q_{threshold} = \min_t (q_{15}^{(t)})$, and compounds satisfying $q \leq q_{threshold}$ were retained. Columns: Drug; Cell line; Tissue; Rank; P ; q ; Brief description; $n_{contexts}$; $Z_{agg,tissue}$; Selection.

Application of Key-Tissue Filtering Across 12 GWAS Datasets

Key-tissue filtering applied to the 12 GWAS-specific SPxLINCS tables produced one candidate list per dataset. Across the 12 GWAS datasets, the workflow recovered compounds with reverse-direction signatures, defined as $Z_{agg} < 0$, in biologically relevant tissues. The tissue-aware thresholding strategy preserved representation of both strongly and moderately supported key tissues.

Representative Results Across Two GWAS Categories

Two filtered tables were selected for the main text. Candidate compounds carry two annotation layers. Entries in **bold** denote compounds with existing HCC-related experimental or clinical evidence in the literature, regardless of whether that evidence directly corroborates the repurposing signal. The dual-annotation scheme makes explicit that computational prioritization and literature support are complementary but non-equivalent dimensions of candidate evaluation.

Table-specific computational findings

In Table 1, the strongest displayed liver-tissue signals were JNK-9L in the HCC515 context, ranked first, and erythrosine in the HT29 context, ranked second. The HCC515 cell line is a lung cancer cell line and HT29 is a large-intestine cancer cell line; therefore, the cell-line annotation should not be interpreted as evidence that either compound was tested in a liver-derived HCC model in these rows.

Other Table 1 candidates included KU-0060648, PHA-793887, triptolide, alvocidib, AG-14361, PF-562271, CGP-60474, AT-7519, AS-601245, ER-27319, epirubicin, dactinomycin, piperlongumine, BI-2536, TG-101348, PI-103, LY-294002, dorsomorphin, NVP-AUY922, ZSTK-474, mepacrine, AZD-8055, and LDN-193189. These candidates span several mechanistic classes, including CDK, PLK1, FAK, JNK, DNA-PK, mTOR/PI3K, HSP90, and other pathway-directed or cytotoxic compounds.

Most of the displayed Table 1 signals occurred in lung, spleen, whole-blood, terminal-ileum, or lymphocyte-related tissue panels rather than in a directly HCC-specific cell-line/tissue combination. For example, alvocidib and dactinomycin were shown with HEPG2 cells, but the

tissue annotation for those rows was lung. Similarly, a liver-labelled tissue panel does not necessarily indicate that the drug was tested in an HCC tissue model. These findings therefore support computational prioritization for follow-up, not a conclusion of clinical activity against HCC.

Table 2 represents a benign-liver-related GWAS category and should not be treated as an HCC-specific dataset. Its strongest displayed signals included narciclasine, homoharringtonine, digitoxin, anisomycin, digitoxigenin, digoxin, strophanthidin, cephaeline, cycloheximide, NSC-632839, emetine, ouabain, periplocymarin, cinobufagin, and BI-2536 in EBV-transformed lymphocytes. Additional displayed signals included AG-14361, helveticoside, fluticasone, dorsomorphin, narciclasine, BI-2536, and CA-074-Me in other listed tissue panels.

The cardiac-glycoside group in Table 2 comprises digitoxin, digitoxigenin, digoxin, strophanthidin, ouabain, periplocymarin, and helveticoside. Narciclasine and homoharringtonine are separate compounds and should not be classified as cardiac glycosides. The translation-related compounds shown in the table include anisomycin, cephaeline, cycloheximide, and emetine.

The liver-tissue rows in Table 2 included AG-14361 with PL21 cells, helveticoside with PC3 cells, and fluticasone with A549 cells. The presence of a liver tissue annotation does not establish HCC selectivity or clinical efficacy. The PHH annotation refers to primary human hepatocytes and not malignant HCC cells. The HUH7 annotation identifies a liver cancer cell line, but the corresponding tissue panel in the supplied row was EBV-transformed lymphocytes; this remains a computational association rather than a clinical result.

Candidates with HCC-related experimental evidence

The computational findings are also consistent with selected preclinical and translational evidence, although the literature findings remain separate from the reverse-signature statistics. Homoharringtonine (HHT) has been validated as an effective antitumor agent in HCC acting through an EphB4/ β -catenin-dependent mechanism [24]. Co-delivery of narciclasine and siULK1 in biocompatible micelles safely inhibited tumor growth and protective autophagy in HCC [25].

Among kinase-targeted agents, PF-562271, a focal adhesion kinase inhibitor, is notable, as the inverse correlation between GPD1L expression and treatment response highlights GPD1L as a potential predictive biomarker in HCC [26]. AT-7519, a CDK inhibitor, sensitized HCC cells to gefitinib and cabozantinib [27]. BI-2536, a PLK1 inhibitor, appeared in both supplied tables and showed signals across several tissue panels; insufficient intratumoral drug concentrations have been identified as a resistance mechanism, and PLK1 inhibitors achieving adequate intratumoral levels are considered promising for HCC [28, 29]. Alvocidib, AT-7519, and several additional candidates were independently identified as potential HCC therapeutics through

analysis of differentially expressed genes between tumor and non-tumor liver tissues [30].

In Table 1, AZD-8055, an mTOR inhibitor, was identified in the BT20 breast-cell-line context under the spleen tissue panel ($P = 5.16 \times 10^{-30}$, $q = 6.62 \times 10^{-28}$). Independently of this computational signal, PEGylated AZD-8055 was reported to suppress HCC tumor growth with optimized pharmacokinetics [31, 32]. The literature evidence therefore remains separate from the reverse-signature statistic.

Other HCC-related literature discussed in the supplied text includes KU-0060648, which was reported to inhibit HCC through DNA-PKcs-dependent and DNA-PKcs-independent mechanisms [33]; piperlongumine, which was reported to inhibit HCC cell migration and invasion through the ROS–ER–MAPKs–CHOP axis [34]; NVP-AUY922, an HSP90 inhibitor listed in Table 1, which has been reported to demonstrate therapeutic activity in HCC [35]; triptolide, which has been studied in combination with RSL3 in relation to apoptotic and ferroptotic pathways [36]; anisomycin, which has been reported to exert NK-dependent immunomodulatory effects [37]; and cinobufagin, which has been described as having multiple antitumor mechanisms with translational potential [38]. These compounds are present in the supplied tables, but the literature findings remain separate from the reverse-signature statistics.

Clinical and translational context of table-listed compounds

Several table-listed candidates also had clinical or translational relevance in related treatment contexts. Epirubicin appears in Table 1, although its appearance represents a computational reverse-signature signal rather than proof of predicted clinical benefit. A 2019 worldwide survey on TACE for HCC reported regional variation in the use of anthracycline agents, including epirubicin in Japan [39, 40]. The supplied literature discussion also reported sensitivity of selected HCC molecular subtypes to ouabain and TG-101348 [41]. Ouabain appears in Table 2 and TG-101348 appears in Table 1; however, these subtype-sensitivity findings remain separate from the reverse-signature statistics [41, 42]. LY-294002 enhanced oxaliplatin chemosensitivity via PI3K/AKT inhibition [43], and PI-103 combined with sonodynamic therapy suppressed HCC proliferation [44]. Both compounds appear in Table 1, but these preclinical combination findings do not establish them as standard HCC treatments.

Overall interpretation

The main mechanistic categories were cell cycle/mitotic kinase inhibitors, particularly CDK/PLK1 inhibitors, focal adhesion kinase (FAK) inhibitors, mTOR/AKT-related inhibitors, DNA-PK inhibitors, and translation-related compounds. These categories are represented in the tables by compounds such as alvocidib, AT-7519, PHA-793887, CGP-60474, BI-2536, PF-562271, AZD-8055, PI-103, LY-294002, KU-0060648, and anisomycin. Triptolide is also a prominent Table 1 candidate with HCC-related preclinical evidence [36]. This distribution aligns with the strongest reverse-signature signals, which were mainly observed in liver, whole

blood, spleen, and immune-related tissues.

By contrast, most of the other prioritized candidates are still supported mainly by preclinical evidence in HCC or related hepatobiliary disease contexts. At present, they are more suitable for biomarker-guided or combination-based evaluation than for direct clinical use. This distinction is particularly important because the reverse-signature results represent computational prioritization and do not independently establish clinical efficacy.

Overall, the representative tables indicate that inhibitors targeting CDK/PLK1, FAK, mTOR/AKT, DNA-PK, and the translation axis were statistically prominent reverse-signature candidates across the supplied hepatobiliary genetic contexts. Most of these candidates remain at the preclinical stage in HCC. The value of the dual-annotation framework lies in its ability to distinguish computational prioritization from the strength of literature support, thereby providing a clearer basis for subsequent mechanism-guided, biomarker-stratified, locoregional, or combination-focused studies.

Several candidates with limited or no HCC-related literature support in the supplied annotation produced strong reverse-signature statistics, including ER-27319, an FcεRI–Syk inhibitor, in lung tissue in Table 1 ($q = 3.39 \times 10^{-29}$), and LDN-193189, a BMP receptor inhibitor, in whole blood in Table 1 ($q = 7.71 \times 10^{-41}$). These compounds should be regarded as computationally prioritized candidates rather than established HCC or biliary tract cancer treatments. Because their appearance in the tables is based on reverse-signature statistics, these findings should not be interpreted as evidence of previously demonstrated antitumor activity.

Recurrent and trait-specific repurposing signals across HCC-related GWAS datasets

To compare drug repurposing signals across HCC and related GWAS datasets, compounds within each dataset were ranked using the minimum drug-level FDR-adjusted q -value (q_{min}) derived from the SPxLINCS output, and the top 10 compounds per dataset were visualized (Figure 3; Supplementary Tables 1-12). This cross-dataset comparison revealed a hierarchical structure of repurposing signals, comprising a core set of broadly replicated candidates and a periphery of trait-specific leads.

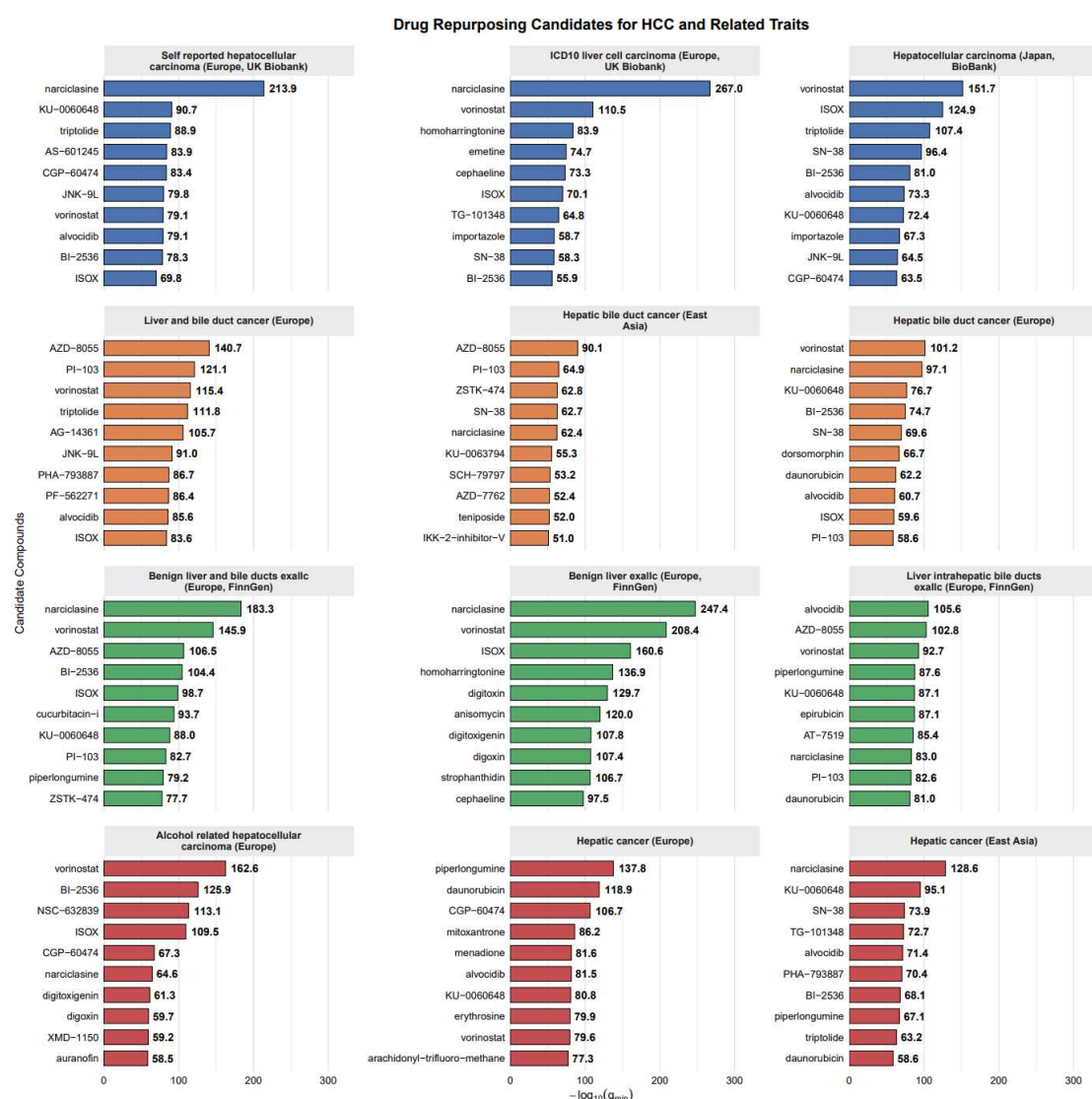


Figure 3. Top drug repurposing candidates across 12 HCC and related GWAS datasets. For each GWAS dataset, candidate compounds were identified from the corresponding SPxLINC output (Supplementary Tables 1-12). Drug-level significance was summarized as the minimum FDR-adjusted q -value across all entries per dataset (q_{min}), and the top 10 compounds with the smallest q_{min} were selected. Bar lengths represent $-\log_{10}(q_{min})$. Per-tissue top-15 bar charts are provided in Supplementary Figures 4-52.

Vorinostat (HDAC inhibitor) was the most frequently prioritized compound, appearing in 10 of 12 GWAS datasets. Its strongest signals were observed in alcohol-related HCC (Europe; $-\log_{10}(q_{min}) = 162.6$), hepatocellular carcinoma (Japan, BioBank Japan; 151.7), and benign liver and bile ducts *exallc* (Europe, FinnGen; 145.9). Narciclasine (cytoskeletal/translation axis inhibitor) was prioritized in 9 of 12 datasets, with the strongest association in ICD-10 liver cell carcinoma (Europe, UK Biobank; $-\log_{10}(q_{min}) = 267.0$), with additional strong signals in benign liver *exallc* (Europe, FinnGen; 247.4), self-reported HCC (Europe, UK Biobank; 213.9), benign liver and bile ducts *exallc* (Europe, FinnGen;

183.3), and hepatic cancer (East Asia; 128.6). ISOX and KU-0060648 (DNA-PK inhibitors) each appeared in 8 of 12 datasets, and alvocidib (CDK inhibitor) in 7 of 12, further defining a mechanistically coherent core of broadly replicated candidates. BI-2536 (PLK1 inhibitor), PI-103 (PI3K/mTOR inhibitor), and SN-38 (topoisomerase I inhibitor) were each prioritized in 6, 5, and 5 of 12 datasets, respectively, while AZD-8055 (mTOR inhibitor), piperlongumine, and triptolide each appeared in 4 of 12 datasets.

Notably, the recurrently prioritized compounds cluster into five mechanistic categories: HDAC inhibition (vorinostat), cytoskeletal/translation axis (narciclasine), DNA-PK/DNA damage response (ISOX, KU-0060648), cell cycle/mitotic kinase inhibition (alvocidib, BI-2536), and PI3K/mTOR/topoisomerase I targeting (PI-103, AZD-8055, SN-38). The convergence of independent GWAS signals onto these same mechanistic axes suggests that these pathways represent convergent transcriptomic vulnerabilities across HCC-related genetic backgrounds, rather than artifacts of any single dataset or cohort.

Alongside this broadly replicated core, marked trait-specific heterogeneity was evident. Homoharringtonine (translation inhibitor) ranked highly in ICD-10 liver cell carcinoma (Europe, UK Biobank; $-\log_{10}(q_{min}) = 83.9$) and benign liver *exallc* (Europe, FinnGen; 136.9), but was not consistently prioritized across other panels. PI-103 and AZD-8055 were disproportionately prominent in liver and bile duct-related phenotypes, whereas piperlongumine showed selective enrichment in intrahepatic bile duct and hepatic cancer datasets. JNK-9L appeared in 3 of 12 datasets, representing a more phenotype-restricted signal.

Collectively, these findings demonstrate that SPxLINCS identifies a mechanistically coherent, cross-dataset replicated shortlist of repurposing candidates alongside phenotype-specific leads, providing a principled basis for prioritizing compounds in mechanism-guided, biomarker-enriched validation studies.

Integrated visualization of GWAS-specific drug repositioning profiles

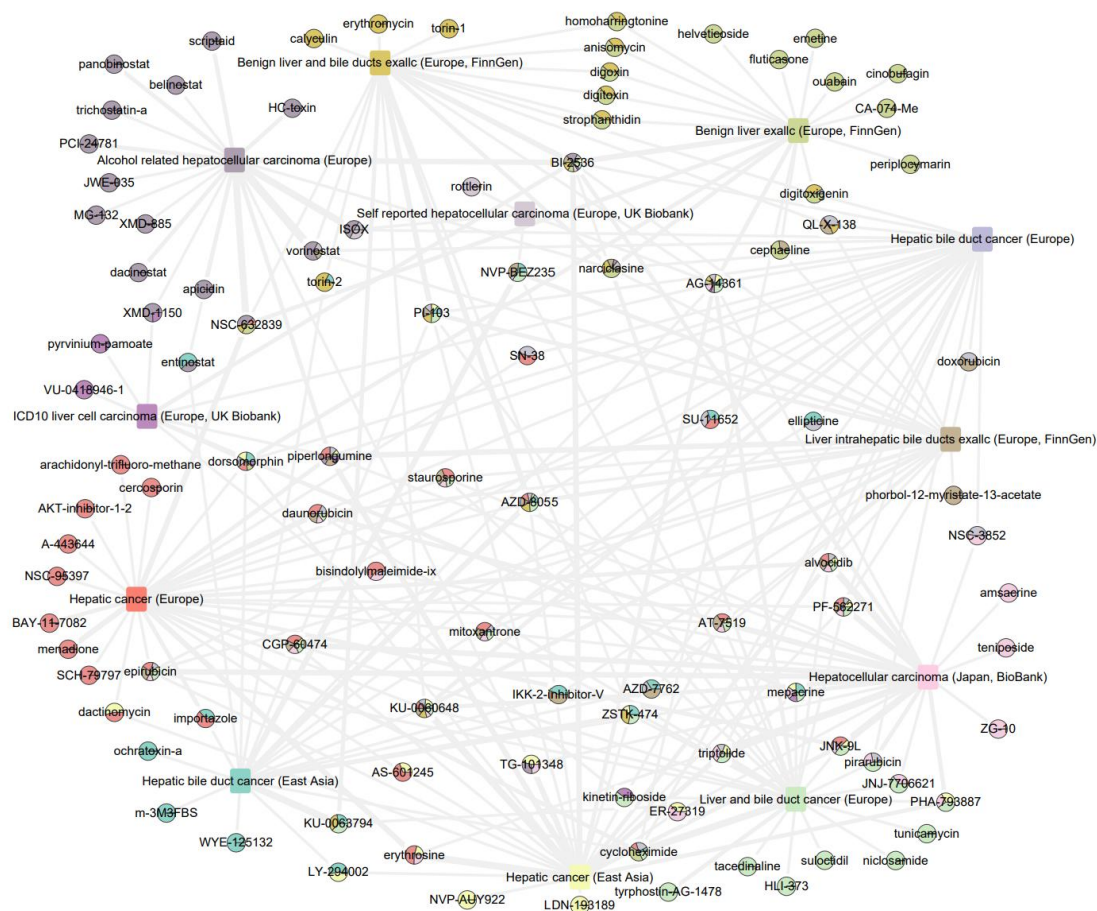


Figure 4. GWAS-resolved drug-trait network with drug-centered pie-chart annotation. Each drug is represented as a single circular node with a pie chart in which each slice corresponds to one GWAS dataset and slice area is proportional to $-\log_{10}(q_{min})$, where q_{min} is the minimum FDR-adjusted q -value for that drug within the corresponding GWAS-specific table. Each GWAS dataset is represented as a rectangular trait node. Edges indicate drug-GWAS associations; edge width reflects within-GWAS recurrence count.

The bipartite network provided a GWAS-resolved view of candidate repositioning compounds across hepatobiliary cancer-related traits (Figure 4), capturing two complementary properties of each candidate. Pie-chart composition reflected the breadth of cross-GWAS support: drugs with multiple large slices carried broad repositioning signals across datasets, whereas a single dominant slice indicated trait-specific candidacy. Edge width, scaled by within-GWAS recurrence count, reflected detection consistency within each dataset.

Several compounds reached extreme within-GWAS significance. Representative maxima included narciclasine ($-\log_{10}(q_{min}) \approx 183.3$), vorinostat/PCI-24781 (HDAC inhibitor; ≈ 162.6), BI-2536 (PLK1 inhibitor; ≈ 125.9), PI-103 (PI3K/mTOR inhibitor; ≈ 121.1), triptolide (≈ 111.8), KU-0060648 (DNA-PK inhibitor; ≈ 87.1), PF-562271 (FAK inhibitor; ≈ 86.4), and JNK-9L (≈ 91.0). Cardiotonic steroids also showed large values, including digitoxin (≈ 129.7), digoxin (≈ 107.4), and digitoxigenin (≈ 72.2).

In terms of network topology, multi-GWAS candidates clustered around three mechanistic axes: cell cycle/mitotic kinase inhibition (alvocidib, PHA-793887, BI-2536), DNA-PK/DNA damage response (KU-0060648), and focal adhesion signaling (PF-562271). These nodes combined large $-\log_{10}(q_{min})$ slice areas with high within-GWAS recurrence, reflected in thick edges across multiple trait nodes. Trait-specific patterns were also evident: topoisomerase I agents (e.g., SN-38) and select natural products showed dominant slices in duct- and HCC-related datasets, while cardiogenic steroids and translation inhibitors concentrated in benign liver FinnGen traits. Across the network, liver and immune-related tissue panels contributed the largest slice areas, consistent with the upstream key-tissue filtering strategy.

Across all datasets, five mechanistic axes recurred with the largest $-\log_{10}(q_{min})$ values and highest within-GWAS recurrence: cell cycle/mitotic kinases (CDK/PLK1), focal adhesion (FAK), DNA-PK/DNA damage response, topoisomerase I, and translation inhibition. Broadly replicated candidates such as BI-2536, KU-0060648, PF-562271, alvocidib and PHA-793887 are prioritized for cross-phenotype follow-up, while trait-specific leads like SN-38 are nominated for phenotype-restricted validation strategies.

Integrated tripartite network of hepatobiliary cancer GWAS traits, prioritized genes, and candidate drugs

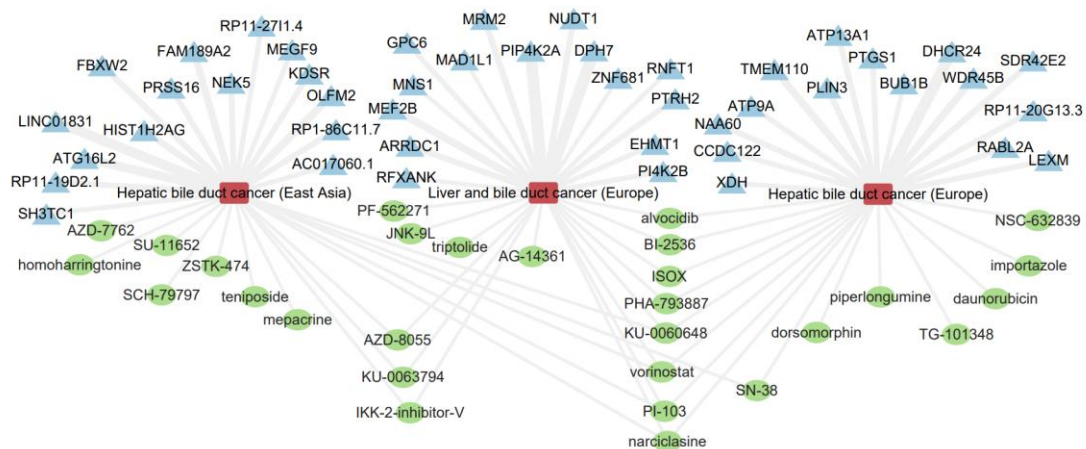


Figure 5. Tripartite network of hepatobiliary cancer GWAS traits, TWAS-prioritized genes and candidate drugs. A three-layer network was constructed for Liver and bile duct cancer (Europe), Hepatic bile duct cancer (East Asia) and Hepatic bile duct cancer (Europe). For each trait, the top 15 genes were selected from Supplementary Table 6 [5] based on maximum $-\log_{10}(p)$, and the top 15 drugs were selected from GWAS-specific repositioning results based on minimum q_{min} . Nodes denote genes (blue triangles), drugs (green circles) and GWAS traits (red rectangles). Gene-GWAS edge width is scaled by gene association significance, whereas drug-GWAS edges are shown at constant width for clarity. Network of the other three groups of traits are shown in Supplementary Figures 1-3.

To integrate TWAS-prioritized susceptibility genes with computational drug repositioning

candidates, a tripartite gene–GWAS–drug network was constructed for the duct-related trait group as a representative example (Figure 5). The resulting network comprised 78 nodes (3 GWAS traits, 45 genes, 30 drugs) connected by 75 edges (45 gene-GWAS and 30 drug-GWAS), with each GWAS trait serving as a central hub linking upper-layer genes to lower-layer drug candidates. The two bile duct cancer datasets from East Asian and European populations showed distinct but partially overlapping profiles, while the liver and bile duct cancer (Europe) node displayed its own trait-specific gene and drug neighborhood.

To characterize trait-specific genetic signals, gene-GWAS edges were scaled by $-\log_{10}(p)$, visually emphasizing the strongest TWAS associations within each branch. Distinct prioritized gene sets were observed across the three traits: hepatic bile duct cancer (East Asia) was dominated by LINC01831, ATG16L2, OLFM2, FAM189A2, and HIST1H2AG; liver and bile duct cancer (Europe) by DPH7, PIP4K2A, MEF2B, MRM2, and ARRDC1; hepatic bile duct cancer (Europe) by DHCR24, TMEM110, WDR45B, PLIN3, and ATP13A1. No overlap among the highest-ranked genes across three datasets reflects heterogeneity in the genetic architecture captured by population- and phenotype-specific GWAS.

To identify trait-resolved repositioning candidates, drugs were ranked by ascending $q_{\min}$ within each GWAS-specific table, retaining the top 15 per trait. For liver and bile duct cancer (Europe), top-ranked compounds included AZD-8055 (mTOR inhibitor), PI-103 (PI3K/mTOR inhibitor), vorinostat (HDAC inhibitor), triptolide, and AG-14361 (DNA-PK inhibitor), several reaching $-\log_{10}(q_{\min}) > 100$. For hepatic bile duct cancer (East Asia), prominent candidates included AZD-7762 (CHK1/2 inhibitor), SU-11652, ZSTK-474 (PI3K inhibitor), SCH-79797, and teniposide (topoisomerase II inhibitor). For hepatic bile duct cancer (Europe), representative compounds included NSC-632839, importazole (nuclear transport inhibitor), daunorubicin, piperlongumine, dorsomorphin (AMPK inhibitor), and TG-101348 (JAK2 inhibitor). Thicker edges represent higher recurrence within a given GWAS, indicating that these repositioning signals were detected more consistently for that trait.

This network view combines two complementary types of evidence within the same trait-specific framework. One layer represents disease biology supported by genetics, reflected by TWAS-prioritized genes, whereas the other captures computationally prioritized repositioning candidates. The marked trait-specificity observed in both gene and drug layers underscores the value of GWAS-stratified analysis for identifying phenotype-restricted molecular targets and repositioning leads for further validation.

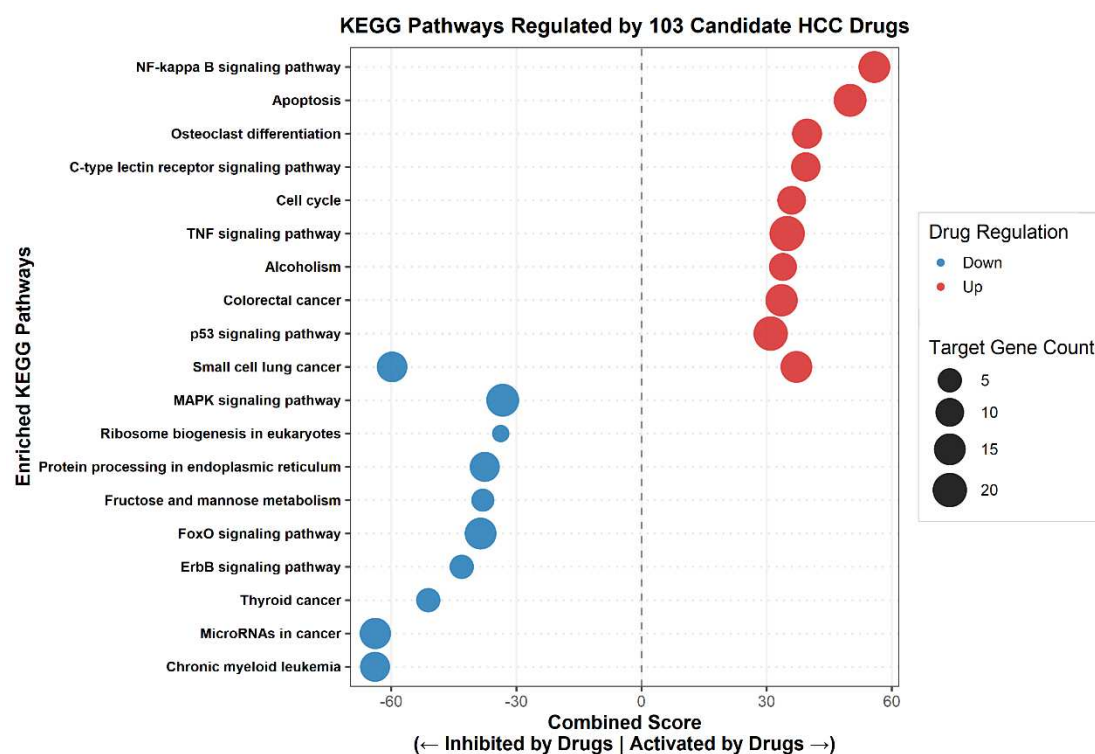


Figure 6. KEGG pathway enrichment analysis of downstream molecular targets of the 103 candidate drugs. Significant KEGG pathways were derived from downstream protein target networks of the 103 candidate compounds using the L1000FWD platform and are shown as a bubble plot. The x-axis shows the Combined Score: positive values (red) indicate activated pathways, whereas negative values (blue) indicate inhibited pathways. Bubble size reflects the number of overlapping target genes within each pathway.

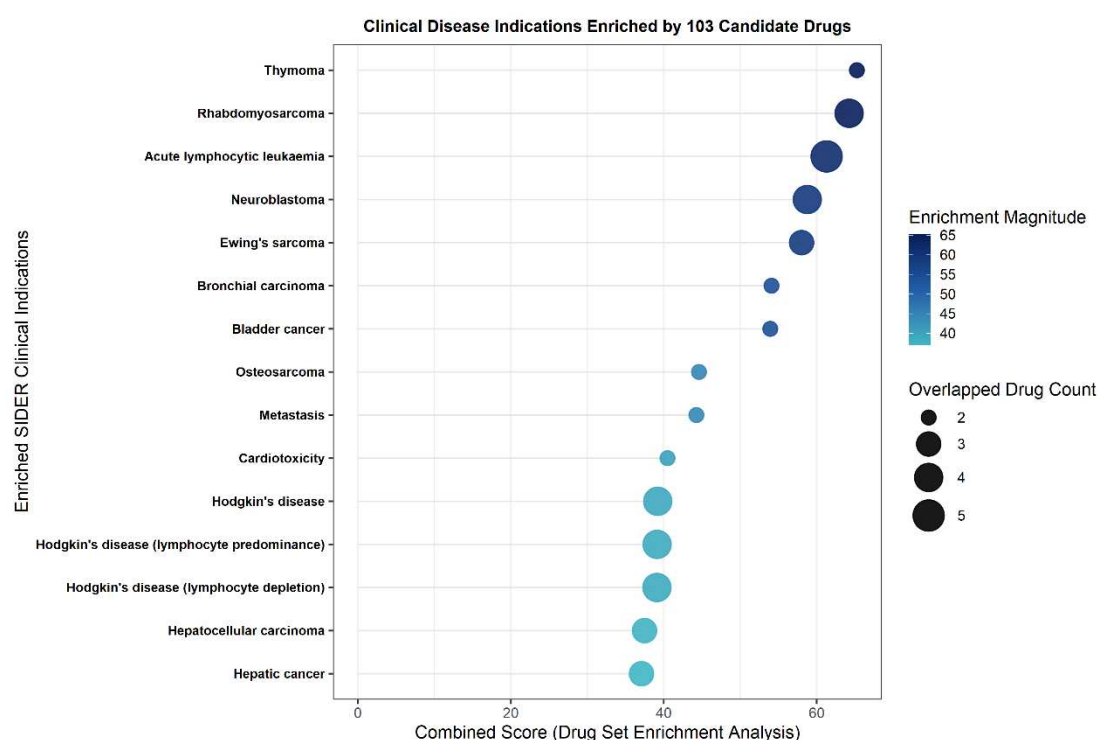


Figure 7. Drug Set Enrichment Analysis (DSEA) of the 103 candidate drugs against SIDER clinical indications.

The top 15 clinical indication terms enriched among the 103 candidate compounds were obtained from the SIDER database and are shown as a bubble plot. The x-axis shows the Combined Score, defined as

$$CS = -\ln(P) \times Z$$

bubble size indicates the number of overlapping candidate drugs in each indication category, and color intensity reflects enrichment magnitude. BH-FDR correction was applied throughout.

Identification of Core Signaling Pathways

To explore the mechanisms associated with the 103 candidate compounds, we mapped their downstream protein target networks to KEGG pathways using the L1000FWD platform (Figure 6). The results suggested a bidirectional regulatory pattern compatible with an antitumor effect. On the inhibitory side, where the Combined Score was negative, candidate drug targets were enriched in several pathways, including MAPK signaling (Combined Score approximately -40), ErbB signaling, FoxO signaling, and microRNAs in cancer (Combined Score approximately -62). Additional enriched terms included chronic myeloid leukemia, thyroid cancer, ribosome biogenesis in eukaryotes, protein processing in the endoplasmic reticulum, and fructose and mannose metabolism. On the activating side, where the Combined Score was positive, the candidate set was enriched for pathways related to tumor suppression, most notably the NF- κ B signaling pathway (Combined Score approximately $+60$), apoptosis, TNF signaling, p53 signaling, cell cycle, the C-type lectin receptor signaling pathway, osteoclast differentiation, colorectal cancer, and alcoholism. Together, these patterns suggest that the shortlisted compounds may exert anti-HCC effects by dampening oncogenic signaling while reactivating cell death-related pathways.

Multi-Dimensional Validation via Clinical Indications

We then evaluated the clinical relevance of the candidate set using DSEA against the SIDER indication database (Figure 7). Among the top 15 enriched terms, the highest Combined Scores were observed for thymoma (≈ 65), rhabdomyosarcoma (≈ 65), acute lymphocytic leukaemia (≈ 63), neuroblastoma (≈ 60), and Ewing's sarcoma (≈ 58). These rankings reflect the broad-spectrum antineoplastic activity of anthracyclines and topoisomerase inhibitors within the candidate set, which carry approved indications across multiple cancer types. Other enriched terms included bronchial carcinoma (≈ 52), bladder cancer (≈ 48), osteosarcoma (≈ 40), metastasis (≈ 40), cardiotoxicity (≈ 39), and Hodgkin's disease with its subtypes (lymphocyte predominance and lymphocyte depletion; ≈ 38 – 39).

Notably, hepatocellular carcinoma (Combined Score = 37.48) and hepatic cancer (≈ 35) were the only liver-specific indications to appear among the top 15 terms. This enrichment was mainly attributable to established HCC-active agents in the candidate set, including

mitoxantrone, epirubicin, and doxorubicin. Although their Combined Scores were lower than those of several hematologic malignancies, their presence as the only hepatic indications in an unbiased enrichment analysis still supports the relevance of the nominated compounds to HCC.

Discussion

In this study, we developed a reproducible, GWAS-resolved drug repurposing framework for hepatobiliary cancers by integrating genetic liability via S-PrediXcan, disease-reversing transcriptional signatures (SPxLINCS), and tissue biology. Three design features were particularly important. First, key-tissue selection and directionality filtering $Z_{agg} < 0$ anchored the hits to biological contexts relevant to HCC. Second, drug-level aggregation based on q_{min} and visualization using $-\log_{10}(q_{min})$ made it possible to identify both broadly recurrent and trait-specific candidates. Third, Cytoscape-based network summarization linked pharmacologic signals to genetically prioritized biology.

Several consistent patterns emerged across all 12 datasets. Recurrent signals across the 12 datasets were mainly concentrated in five mechanistic categories: CDK/PLK1 inhibitors, FAK inhibitors, DNA-PK/DNA damage response inhibitors, topoisomerase I inhibitors, and translation-axis inhibitors. These classes tended to show larger $-\log_{10}(q_{min})$ values and higher within-GWAS recurrence. Trait-specific signals, such as SN-38/irinotecan in duct-related datasets, highlight context-specific opportunities. DSEA against SIDER further supported enrichment for hepatocellular and related hepatic indications. KEGG pathway analysis revealed a bidirectional regulatory pattern: pathways related to proliferation, including MAPK, ErbB, cell cycle, and FoxO signaling, were suppressed, whereas pathways associated with tumor suppression, such as p53, apoptosis, and TNF signaling, were reactivated. Together, these findings provide mechanistic coherence for the nominated shortlist (Figures 6-7).

Dataset-level exemplars are instructive. In the East Asia duct-related analysis, alvocidib (CDK inhibitor), PF-562271 (FAK inhibitor), BI-2536 (PLK1 inhibitor), SN-38/irinotecan (topoisomerase I inhibitor), homoharringtonine, and narciclasine dominated under key-tissue constraints. While irinotecan-based regimens carry exploratory post-line activity in biliary tract cancers, most candidates remain preclinical for HCC and constitute hypothesis-generating leads best advanced through mechanism-guided, biomarker-enriched, or locoregional and combination strategies including TACE/HAIC and IO-anti-VEGF backbones.

Several compounds within the prioritized set carry established pharmacological profiles alongside significant translational constraints. AG-14361 (PARP-1 inhibitor) has established relevance in DNA damage response, cell cycle regulation, and apoptosis [45], with robust preclinical radiosensitization data; however, the clinical field has standardized on later-generation PARP inhibitors with superior PK/PD profiles. If PARP pathway biology is

genetically supported, pivot to olaparib-class agents is warranted. JNK-9L, a JNK inhibitor, has been linked to stress signaling and has been studied in cancer-related settings [46]. However, it remains an early research tool, with limited information on selectivity and ADME properties. Before any further scale-up or translational consideration, validation with more selective chemotypes or genetic perturbation would be necessary. Dorsomorphin has shown antitumor activity in several solid tumor models. In chemotherapy-resistant epithelial ovarian cancer, it resensitized cells to platinum agents by inducing autophagy and blocking cell-cycle progression [47]. In colorectal cancer, it also reduced ABCG2-mediated multidrug resistance by inhibiting transporter activity without changing ABCG2 protein expression [48]. Even so, its broad off-target effects limit its clinical potential. In this candidate set, LDN-193189 may provide a more selective way to interrogate BMP receptor signaling [49]. Mepacrine engages multiple oncogenic nodes including p53 stabilization, FACT complex disruption, autophagy modulation, NF- κ B suppression, and AKT inhibition [50]. In the HCC Japan dataset, mepacrine-associated signals were observed in spleen and transverse colon tissues. This pattern may reflect activity in immune-related compartments rather than direct hepatocellular killing. If so, mepacrine may be better considered as a combination agent than as a systemic monotherapy. Erythrosine, an antimicrobial photosensitizer [51], produced cross-tissue signals across the Hepatic Cancer East Asia cohort spanning liver, lung, small intestine terminal ileum, and colon transverse; in the absence of a mechanistic rationale for systemic administration, its computational signal is best interpreted within a photodynamic therapy framework. Where pathway biology is genetically supported, research-grade hits should be treated as directional evidence, with clinical-stage analogs substituted for translational follow-up.

Beyond pathway-level reorientation of established tool compounds, the present screen surfaced three candidates warranting specific attention for their combination of statistical robustness and absence of prior investigation in hepatic malignancy.

Importazole is a 2,4-diaminoquinazoline compound that specifically blocks importin- β -mediated nuclear import by interfering with the RanGTP-importin- β interaction, without disrupting transportin-mediated import or CRM1-mediated export [52]. It was identified through a FRET-based high-throughput screen and has been used exclusively as a chemical biology tool to dissect importin- β /RanGTP pathway function at defined cell cycle stages; no antitumor activity has been reported in any cancer type. In the Hepatic Bile Duct Cancer (East Asia) dataset, importazole ranked among the top candidates in EBV-transformed lymphocytes ($q = 2.80 \times 10^{-78}$), a signal of considerable strength relative to the broader candidate distribution in that tissue. This finding is mechanistically interpretable in the context of HCC biology: importin- β -dependent nuclear transport is required for the nuclear accumulation of STAT3 and additional transcription factors whose aberrant nuclear localization is well documented in hepatocellular carcinoma. Pharmacological blockade of this transport step has

not been explored as a therapeutic strategy in HCC or biliary tract cancers. The concentration of the reversal signal in immune-related tissue rather than hepatic parenchyma raises the additional possibility that importin- β inhibition disrupts nuclear import-dependent transcriptional programs in tumor-infiltrating immune populations. Genetic perturbation of importin- β in HCC-derived cell lines, combined with orthogonal chemotypes targeting the RanGTP-importin- β interaction, would constitute a tractable first experimental test of this hypothesis.

ER-27319 is an acridone-related compound that selectively inhibits Fc ϵ RI-mediated activation of Syk, a non-receptor tyrosine kinase critical for mast cell degranulation, through blocking tyrosine phosphorylation of Syk downstream of Fc ϵ RI gamma subunit ITAM engagement without affecting Lyn kinase activity or T cell- and B cell-associated Syk activation [53]. It was developed for allergic disease and has no reported antitumor activity in any cancer type. In the HCC Japan (BioBank) dataset, ER-27319 ranked fourth in spleen ($q = 3.25 \times 10^{-47}$), with a supporting signal in the lung panel of the Hepatic Cancer East Asia cohort. The spleen-dominant pattern is mechanistically coherent: Syk is broadly expressed in tumor-associated immune populations resident in the spleen, including mast cells, macrophages, and B cells, all of which contribute to the HCC immune microenvironment. Tumor-associated mast cell infiltration has been linked to HCC progression and poor prognosis, and Syk-dependent signaling in mast cells governs the release of pro-tumorigenic mediators including TNF- α and arachidonic acid metabolites, yet pharmacological targeting of the Fc ϵ RI–Syk axis in HCC has not been investigated. The recurrence of the signal across two geographically independent cohorts argues against a dataset-specific artifact. Whether the reversal signature reflects suppression of Syk-dependent mast cell activation within the hepatic tumor microenvironment, or broader Syk-mediated immune dysregulation, cannot be determined from the present data; either interpretation provides a biologically tractable hypothesis for experimental follow-up in HCC-relevant immune cell models.

LDN-193189 is a selective BMP type I receptor inhibitor (ALK2/ALK3) developed as a research tool to probe BMP signaling in skeletal and stromal cell contexts. Antitumor activity has been reported in epithelial ovarian cancer models, where LDN-193189 induced ROS-dependent, caspase 3-independent cell death and resensitized platinum-resistant cells to cisplatin and carboplatin in both in vitro and in vivo settings; dorsomorphin, a structurally distinct compound acting on an overlapping set of BMP receptors, produced complementary effects through autophagy induction and cell cycle blockade in the same models. Neither compound has been investigated in HCC or biliary tract cancers. In the Hepatic Cancer East Asia dataset, LDN-193189 ranked third in whole blood ($q = 7.71 \times 10^{-41}$). Dorsomorphin independently emerged as a top candidate in spleen ($q = 8.81 \times 10^{-33}$) and small intestine terminal ileum ($q = 2.88 \times 10^{-14}$) within the same dataset, and in whole blood in the FinnGen

cohort ($q = 2.42 \times 10^{-21}$). The independent prioritization of two chemically dissimilar inhibitors converging on the same molecular target across multiple tissues and independent GWAS cohorts substantially reduces the probability that either signal reflects a compound-specific artifact, and together they provide internal computational validation for BMP receptor involvement in the hepatic cancer transcriptome. BMP9-mediated ALK1 signaling in HCC-associated endothelial cells has been reported, and BMP4 has been implicated in hepatic cancer stem cell maintenance; pharmacologic inhibition of ALK2/ALK3 has not yet been examined in HCC. With its relatively favorable selectivity and its reported ability to induce tumor cell death in solid tumor models, LDN-193189 is a reasonable candidate for mechanistic testing in hepatic cancer systems.

Strengths and Validation Layers

GIN-DRIP (Genetics-Informed Network-based Drug Repurposing via in silico Perturbation) integrates genetics, genomics, perturbational data, and real-world electronic health records to convert GWAS findings into repurposing hypotheses [54]. The framework presented here builds on that general concept, but differs from it in three main respects. First, repurposing queries are anchored in genetically inferred liability through S-PrediXcan rather than in observational differential expression. Second, prioritization is carried out in parallel across 12 independent datasets and 49 tissues, allowing candidates to be ranked in a trait- and tissue-resolved manner at a scale that single-GWAS designs cannot provide. Third, genes, traits, and drugs are summarized within interpretable bipartite and tripartite network structures, rather than being returned simply as a flat ranked list. An additional screen for clinical plausibility was performed using DrugEnrichr against SIDER as an orthogonal validation step external to the transcriptomic pipeline [55].

Direct use of CMap is constrained by its reliance on observed differential expression. Framing the query around S-PrediXcan-derived genetic liability offers one way to reduce that dependence. In contrast to single-GWAS approaches such as DeSigN, SPxLINCS draws on multiple datasets at once, making it possible to compare results across studies and to prioritize candidates in a tissue-specific manner. DSEA provides a further layer of orthogonal support, helping narrow the output from a broad ranked list to a smaller set of candidates with clearer genetic and mechanistic grounding.

Even so, the robustness of the framework still needs to be tested more formally. Sensitivity analyses using Fisher, Stouffer, Tippett, and Robust Rank Aggregation (RRA), together with established positive controls, would provide a useful internal check.

Limitations

The most important limitation is the gap between in vitro signal and in vivo relevance. A reverse-signature result in LINCS L1000 cell lines does not necessarily translate into activity

in cirrhotic liver, where immune complexity and altered pharmacokinetic and pharmacodynamic conditions may substantially reshape drug response. Before any clinical implications can be considered, prospective validation in patient-derived HCC organoids, precision-cut liver slices, and immune-competent PDX models will be necessary. Several of the top-ranked classes, including PLK1, FAK, and DNA-PK inhibitors, also raise practical concerns because their exposure and toxicity profiles may be less well tolerated in the setting of chronic liver disease. In that context, these agents may be more realistically considered as locoregional sensitizers or combination partners than as systemic monotherapies. DSEA-based indication enrichment lends support to the overall convergence of the findings, but it does not replace evidence from randomized clinical studies.

There are also methodological issues that should be kept in mind. The current definition of key tissues may underrepresent stromal and immune components of the tumor microenvironment. Future iterations could refine Z_{agg} by incorporating priors from single-cell or spatial transcriptomic data. The use of q_{min} as the aggregation metric should be compared with alternatives such as Fisher, Stouffer, Tippet, and RRA. Likewise, recurrence thresholds such as requiring a signal in at least two key tissues or in at least two GWAS datasets should be examined as complementary filters rather than fixed assumptions. Some network edges may reflect uneven assay coverage or artefacts introduced during synonym harmonization, rather than true biological connectivity, and coverage-aware normalization has not yet been implemented.

The 12 GWAS datasets were intentionally chosen to capture a broad range of phenotypes, from histologically confirmed HCC to benign hepatobiliary disease, and to include both East Asian and European populations. Although this design helps test generalizability, it also creates clear interpretive limits. Signals arising from benign phenotypes may not carry over to malignant disease, and differences in linkage disequilibrium structure across populations may reduce the precision of S-PrediXcan estimates.

Future Directions

Bridging computation and experiment will require exposure-aware preclinical studies in patient-derived HCC organoids and precision-cut liver slices, with pharmacodynamic endpoints aligned to drug concentrations that are realistically achievable in the clinic. These candidates are likely to be most informative when advanced through biomarker-enriched designs, in which DDR, FAK, and CDK signatures are matched to molecular subgroups using TWAS/GReX profiles, and when tested in combination or locoregional settings rather than as systemic monotherapies.

Given the hepatic constraints on systemic dosing, pairing CDK or FAK inhibitors with IO-anti-VEGF regimens or with TACE, HAIC, or SBRT backbones represents a reasonable point of

entry [56]. Conventional randomized trial designs are often difficult to apply in advanced HCC complicated by impaired liver function; window-of-opportunity or neoadjuvant studies centered on pharmacodynamic endpoints may therefore offer a more feasible path for early clinical evaluation [57].

Conclusion

The GWAS-guided and tissue-constrained SPxLINCS framework described here provides a reproducible approach for prioritizing drug repurposing candidates in hepatobiliary cancers. By combining key-tissue selection, reverse-signature filtering, drug-level aggregation using $q_{\min}$ and $-\log(q_{\min})$, and network-based summarization, the pipeline consistently identified convergent mechanistic classes, including CDK/PLK1, FAK, DNA-PK/DNA damage response, topoisomerase I, and translation-axis inhibitors, while also distinguishing broadly replicated signals from those that appear to be trait-specific. Drug Set Enrichment Analysis against clinically curated indications provided an additional layer of support by linking these findings to hepatocellular and related hepatic cancers, thereby strengthening their translational plausibility.

Most of the prioritized compounds should be regarded as starting points rather than definitive therapeutic answers. The strength of transcriptomic reversal is not equivalent to demonstrated efficacy or tolerability in the diseased liver. The practical value of the framework lies in generating a shortlist with explicit mechanistic rationale and genetic grounding, replacing generic tool-compound signals, where possible, with pathway-supported and clinically more tractable candidates.

Turning genetic liability and perturbational transcriptomic data into testable clinical hypotheses for HCC requires a structured, multi-layer pipeline of this kind. The compound classes identified here are best treated as experimental leads, not as established therapies. Whether they have clinical potential will need to be examined in exposure-aware preclinical models and in early-phase studies built around realistic pharmacologic and pharmacodynamic endpoints. Although the present analysis focused on HCC, the framework itself should also be usable in other diseases with available GWAS and perturbational transcriptomic data.

Declarations

Abbreviations

HCC, hepatocellular carcinoma; GWAS, genome-wide association study; GTEx, Genotype-Tissue Expression; LINCS, Library of Integrated Network-Based Cellular Signatures; TWAS, transcriptome-wide association study; DSEA, drug set enrichment analysis; FAK, focal adhesion kinase; DDR, DNA damage response.

Ethics approval and consent to participate

Not applicable. This study used publicly available datasets and did not involve new human participants, patient recruitment, or animal experiments.

Consent for publication

Not applicable.

Availability of data and materials

Publicly available datasets and resources were analysed in this study. GWAS summary statistics for hepatocellular carcinoma and related hepatobiliary traits were obtained from the public sources described in the manuscript, including GWAS Catalog (<https://www.ebi.ac.uk/gwas/>), OpenGWAS (<https://gwas.mrcieu.ac.uk/>), and BioBank Japan summary statistics (<https://humandbs.biosciencedbc.jp/files/hum0197/hum0197.v3.BBJ.LiC.v1.zip>). GTEx transcriptomic resources were accessed through the GTEx Portal (<https://www.gtexportal.org/>). S-PrediXcan prediction models and covariance files were obtained from PredictDB (<https://predictdb.org/>). LINCS perturbational resources were accessed through the CLUE platform (<https://clue.io/>). L1000FWD resources were accessed through <https://lincsproject.org/LINCS/dmoa>. KEGG pathway resources were accessed through <https://www.kegg.jp/kegg/>. Drug set enrichment resources were accessed through DrugEnrichr (<https://maayanlab.cloud/DrugEnrichr/>). Processed prioritization outputs are provided in the Supplementary Information. Code and analysis resources are available at: <https://github.com/Jasmine-math/GWAS-Resolved-Tissue-Constrained-Drug-Repositioning-for-Hepatocellular-Carcinoma-via-SPxLINCS>

Competing interests

The author declares that there are no competing interests.

Funding

This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

Authors' contributions

J.P. conceived the study, developed the SPxLINCS framework, performed the computational and statistical analyses, interpreted the results, prepared the figures and tables, and wrote and revised the manuscript.

Acknowledgements

The author thanks the developers and maintainers of the public resources used in this study, including GWAS Catalog, OpenGWAS, BioBank Japan, GTEx, PredictDB, CLUE, L1000FWD, DrugEnrichr, KEGG, Cytoscape, and related databases and software tools.

Reference

- [1] Harriet Rungay, Jacques Ferlay, Catherine de Martel. *et al.* Global, regional and national burden of primary liver cancer by subtype, European Journal of Cancer. (2022). <https://doi.org/10.1016/j.ejca.2021.11.023>.
(<https://www.sciencedirect.com/science/article/pii/S0959804921012430>).
- [2] Sung, H. *et al.* Global Cancer Statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J. Clin.* **71**, 209–249 (2021).
- [3] Michael D. Gallagher, Alice S. Chen-Plotkin. The Post-GWAS Era: From Association to Function, The American Journal of Human Genetics. (2018). <https://doi.org/10.1016/j.ajhg.2018.04.002>.
(<https://www.sciencedirect.com/science/article/pii/S0002929718301344>)
- [4] The GTEx Consortium. The GTEx Consortium atlas of genetic regulatory effects across human tissues. *Science* **369**,1318-1330(2020). DOI: 10.1126/science.aaz1776
- [5] Jiahuizi Peng, Wen Zhang. Cross-tissue transcriptome-wide association study prioritizes druggable targets for hepatocellular carcinoma. (2025).
doi: <https://doi.org/10.64898/2025.12.18.695261>
- [6] Lamb, J. *et al.* The connectivity map: Using gene-expression signatures to connect small molecules, genes, and disease. *Science* **313**, 1929–1935 (2006).
- [7] Lim, N., Pavlidis, P. Evaluation of connectivity map shows limited reproducibility in drug repositioning. *Sci Rep* **11**, 17624 (2021). <https://doi.org/10.1038/s41598-021-97005-z>
- [8] Baek, SJ., Lee, H., Park, SM. *et al.* Multidimensional transcriptome dataset for systematic evaluation of Jakyakgamcho-tang-induced cell signatures. *Sci Data* **13**, 367 (2026). <https://doi.org/10.1038/s41597-026-06759-6>
- [9] Luo, S., Li, H., Mo, Z. *et al.* Connectivity map identifies luteolin as a treatment option of ischemic stroke by inhibiting *MMP9* and activation of the PI3K/Akt signaling pathway. *Exp Mol Med* **51**, 1–11 (2019). <https://doi.org/10.1038/s12276-019-0229-z>
- [10] So, HC., Chau, CL., Chiu, WT. *et al.* Analysis of genome-wide association data highlights candidates for drug repositioning in psychiatry. *Nat Neurosci* **20**, 1342–1349 (2017). <https://doi.org/10.1038/nm.4618>
- [11] Inci TG, Acar S, Turgut-Balik D. Nonsmall-cell lung cancer treatment: current status of drug repurposing and nanoparticle-based drug delivery systems. *Turk J Biol.* 2024 Apr 3;48(2):112-132. doi: 10.55730/1300-0152.2687. PMID: 39051063; PMCID: PMC11265851.

- [12] Napolitano Francesco, Sirci F., Carrella Diego. (2015). Drug-set Enrichment Analysis: a novel tool to investigate drug mode of action. *Bioinformatics* (Oxford, England). 32. 10.1093/bioinformatics/btv536.
- [13] Luo Zhihui, Zhu Li-Da, Wang Ya-Min. *et al.* (2022). DSEATM: drug set enrichment analysis uncovering disease mechanisms by biomedical text mining. 23. 10.1093/bib/bbac228. *Briefings in bioinformatics*.
- [14] Alexandra B. Keenan, Sherry L. Jenkins, Kathleen M. *et al.* The Library of Integrated Network-Based Cellular Signatures NIH Program: System-Level Cataloging of Human Cells Response to Perturbations, Cell System. (2018). <https://doi.org/10.1016/j.cels.2017.11.001>. (<https://www.sciencedirect.com/science/article/pii/S2405471217304908>)
- [15] Voloudakis G, Lee KM, Vicari JM. A genetically based computational drug repurposing framework for rapid identification of candidate compounds: application to COVID-19. *medRxiv* [Preprint]. (2025). doi: 10.1101/2025.01.10.25320348. PMID: 39867394; PMCID: PMC11759241.
- [16] Han, HW., Hahn, S., Jeong, H.Y. *et al.* LINCS L1000 dataset-based repositioning of CGP-60474 as a highly potent anti-endotoxemic agent. *Sci Rep* **8**, 14969 (2018). <https://doi.org/10.1038/s41598-018-33039-0>
- [17] (2008). Pearson's Correlation Coefficient. In: Kirch, W. (eds) *Encyclopedia of Public Health*. Springer, Dordrecht. https://doi.org/10.1007/978-1-4020-5614-7_2569
- [18] Kalpić, D., Hlupić, N., Lovrić, M. (2025). Student's *t*-Tests . In: Lovric, M. (eds) *International Encyclopedia of Statistical Science*. Springer, Berlin, Heidelberg. https://doi.org/10.1007/978-3-662-69359-9_665
- [19] (2008). Spearman Rank Correlation Coefficient. In: *The Concise Encyclopedia of Statistics*. Springer, New York, NY. https://doi.org/10.1007/978-0-387-32833-1_379
- [20] (2008). Kolmogorov–Smirnov Test. In: *The Concise Encyclopedia of Statistics*. Springer, New York, NY. https://doi.org/10.1007/978-0-387-32833-1_214
- [21] Mullen, B. A BASIC program for metaanalysis using Stouffer's formula. *Behavior Research Methods & Instrumentation* **14**, 551 (1982). <https://doi.org/10.3758/BF03203424>
- [22] Reimand, J., Isserlin, R., Voisin, V. *et al.* Pathway enrichment analysis and visualization of omics data using g:Profiler, GSEA, Cytoscape and EnrichmentMap. *Nat Protoc* **14**, 482–517 (2019). <https://doi.org/10.1038/s41596-018-0103-9>
- [23] Nicholls, H.L., Vargas, J.D., Sanghvi, M.M. *et al.* Genome-wide analysis of cardiac ventricular phenotypes reveals novel loci and therapeutic targets for heart failure. *Nat*

Commun **17**, 3293 (2026).

[24] Zhu M, Gong Z, Wu Q. *et al.* Homoharringtonine suppresses tumor proliferation and migration by regulating EphB4-mediated β -catenin loss in hepatocellular carcinoma. *Cell Death Dis.* (2020). doi: 10.1038/s41419-020-02902-2. Erratum in: *Cell Death Dis.* (2024). doi: 10.1038/s41419-024-07034-5. PMID: 32801343; PMCID: PMC7429962.

[25] Wang X, Wu F, Li G, Zhang N. *et al.* Lipid-modified cell-penetrating peptide-based self-assembly micelles for co-delivery of narcielazine and siULK1 in hepatocellular carcinoma therapy. *Acta Biomater.* (2018). doi: 10.1016/j.actbio.2018.05.030. Epub 2018 May 19. PMID: 29787814.

[26] Leung PKH, Das B, Cheng X. *et al.* Prognostic and Predictive Utility of GPD1L in Human Hepatocellular Carcinoma. *Int J Mol Sci.* (2023). doi: 10.3390/ijms241713113. PMID: 37685919; PMCID: PMC10487989.

[27] Karaś K, Karwaciak I, Chałaśkiewicz K. *et al.* Anti-hepatocellular carcinoma activity of the cyclin-dependent kinase inhibitor AT7519. *Biomed Pharmacother.* (2023). doi: 10.1016/j.biopha.2023.115002. Epub 2023 Jun 11. PMID: 37311277.

[28] Jörg Hauptenthal, Verena Bihrer, Huedayi Korkusuz. *et al.* Reduced Efficacy of the Plk1 Inhibitor BI 2536 on the Progression of Hepatocellular Carcinoma due to Low Intratumoral Drug Levels, *Neoplasia.* (2012). <https://doi.org/10.1596/neo.111366>. (<https://www.sciencedirect.com/science/article/pii/S1476558612800169>)

[29] Zhu D, Zhu Y, Liu L. *et al.* Metabolomic analysis of vascular cognitive impairment due to hepatocellular carcinoma. *Front Neurol.* (2023). doi: 10.3389/fneur.2022.1109019. PMID: 37008043; PMCID: PMC10062391.

[30] Caputo WL, de Souza MC, Basso CR. *et al.* Comprehensive Profiling and Therapeutic Insights into Differentially Expressed Genes in Hepatocellular Carcinoma. *Cancers (Basel).* (2023). doi: 10.3390/cancers15235653. PMID: 38067357; PMCID: PMC10705715.

[31] Wing Ki Chau, Terence Kin Wah Lee. *et al.* Covalent conjugation of AZD8055 with unsaturated fatty acids for the development of mTOR nanoblockers: increasing the therapeutic efficacy of hepatocellular carcinoma treatment, *eBioMedicine.* (2024). <https://doi.org/10.1016/j.ebiom.2024.105113>. (<https://www.sciencedirect.com/science/article/pii/S2352396424001488>)

[32] Hong Tang, Dilinuer Dilimulati, Zhentao Yang. *et al.* Chemically engineered mTOR-nanoparticle blockers enhance antitumour efficacy, *eBioMedicine.* (2024). <https://doi.org/10.1016/j.ebiom.2024.105099>. (<https://www.sciencedirect.com/science/article/pii/S2352396424001348>)

- [33] Chen MB, Zhou ZT, Yang L. *et al.* KU-0060648 inhibits hepatocellular carcinoma cells through DNA-PKcs-dependent and DNA-PKcs-independent mechanisms. *Oncotarget*. (2016). doi: 10.18632/oncotarget.7742. PMID: 26933997; PMCID: PMC4941370.
- [34] Chen Y, Liu JM, Xiong XX. *et al.* Piperlongumine selectively kills hepatocellular carcinoma cells and preferentially inhibits their invasion via ROS-ER-MAPKs-CHOP. *Oncotarget*. (2015). doi: 10.18632/oncotarget.3444. PMID: 25788268; PMCID: PMC4467445.
- [35] Cheng W, Ainiwaer A, Xiao L. *et al.* Role of the novel HSP90 inhibitor AUY922 in hepatocellular carcinoma: Potential for therapy. *Mol Med Rep*. (2015). doi: 10.3892/mmr.2015.3725. Epub 2015 May 4. PMID: 25955495; PMCID: PMC4464165.
- [36] Liu W, Wu G, Wang J. *et al.* Co-treatment with triptolide and RSL3 induces hepatocellular carcinoma cell apoptosis and ferroptosis. *Mol Med Rep*. (2025). doi: 10.3892/mmr.2025.13567. Epub 2025 May 16. PMID: 40376993; PMCID: PMC12105467.
- [37] Kim M, Lee SJ, Shin S. *et al.* Novel natural killer cell-mediated cancer immunotherapeutic activity of anisomycin against hepatocellular carcinoma cells. *Sci Rep*. (2018). doi: 10.1038/s41598-018-29048-8. PMID: 30006566; PMCID: PMC6045618.
- [38] de Lima Coelho, N.S.; Rodrigues, V.D.; Girotto, O.S. *et al.* Cinobufagin as a Potential Intervention Against Liver Cancer—A Comprehensive Review. *Pharmaceuticals* **2026**, *19*, 158. <https://doi.org/10.3390/ph19010158>
- [39] Guiu, B. (2023). Drugs for TACE of HCC. In: Lucatelli, P. (eds) Transarterial Chemoembolization (TACE) . Springer, Cham. https://doi.org/10.1007/978-3-031-36261-3_11.
- [40] Zhu, HD., Li, X., Sun, JH. *et al.* Transarterial Chemoembolization with Epirubicin-Loaded Microspheres for Hepatocellular Carcinoma: A Prospective, Single-Arm, Multicenter, Phase 2 Study (STOPPER Trial). *Cardiovasc Intervent Radiol* **47**, 325–336 (2024). <https://doi.org/10.1007/s00270-024-03666-4>.
- [41] Su H, Zhou X, Lin G. *et al.* Deciphering the Oncogenic Landscape of Hepatocytes Through Integrated Single-Nucleus and Bulk RNA-Seq of Hepatocellular Carcinoma. *Adv Sci (Weinh)*. 2025 Apr;12(14):e2412944. doi: 10.1002/advs.202412944. (2025). PMID: 39960344; PMCID: PMC11984907.
- [42] Song, Y., Lee, SY., Kim, S. *et al.* Inhibitors of Na⁺/K⁺ ATPase exhibit antitumor effects on multicellular tumor spheroids of hepatocellular carcinoma. *Sci Rep* **10**, 5318 (2020). <https://doi.org/10.1038/s41598-020-62134-4>
- [43] Xu R, Zhang Y, Li A. *et al.* LY-294002 enhances the chemosensitivity of liver cancer to oxaliplatin by blocking the PI3K/AKT/HIF-1 α pathway. *Mol Med Rep*. (2021). doi: 10.3892/mmr.2021.12147. Epub 2021 May 13. PMID: 33982772; PMCID: PMC8134878.

- [44] Yang H, Jing H, Han X. *et al.* Synergistic Anticancer Strategy of Sonodynamic Therapy Combined with PI-103 Against Hepatocellular Carcinoma. *Drug Des Devel Ther.* (2021). doi: 10.2147/DDDT.S296880. PMID: 33603343; PMCID: PMC7886098.
- [45] Ji X, Williams KP, Zheng W. *et al.* Applying a Gene Reversal Rate Computational Methodology to Identify Drugs for a Rare Cancer: Inflammatory Breast Cancer. *Cancer Inform.* (2023). doi: 10.1177/11769351231202588. PMID: 37846218; PMCID: PMC10576937.
- [46] Kamenecka T, Jiang R, Song X. *et al.* Synthesis, biological evaluation, X-ray structure, and pharmacokinetics of aminopyrimidine c-jun-N-terminal kinase (JNK) inhibitors. *J Med Chem.* (2010). doi: 10.1021/jm901351f. Erratum in: *J Med Chem.* 2010 Feb 25;53(4):1882. PMID: 19947601; PMCID: PMC2804074.
- [47] Li N, Wang T, Li Z. *et al.* Dorsomorphin induces cancer cell apoptosis and sensitizes cancer cells to HSP90 and proteasome inhibitors by reducing nuclear heat shock factor 1 levels. *Cancer Biol Med.* (2019). doi: 10.20892/j.issn.2095-3941.2018.0235. PMID: 31516744; PMCID: PMC6713636.
- [48] Li XP, Cao LQ, Yu ZZ. *et al.* Dorsomorphin attenuates ABCG2-mediated multidrug resistance in colorectal cancer. *Front Pharmacol.* (2024). doi: 10.3389/fphar.2024.1393693. PMID: 38855753; PMCID: PMC11157230.
- [49] Ali JL, Lagasse BJ, Minuk AJ. *et al.* Differential cellular responses induced by dorsomorphin and LDN-193189 in chemotherapy-sensitive and chemotherapy-resistant human epithelial ovarian cancer cells. *Int J Cancer.* (2015). doi: 10.1002/ijc.29220. Epub 2014 Sep 26. PMID: 25227893.
- [50] Oien DB, Pathoulas CL, Ray U. *et al.* Repurposing quinacrine for treatment-refractory cancer. *Semin Cancer Biol.* (2021). doi: 10.1016/j.semcancer.2019.09.021. Epub 2019 Sep 25. PMID: 31562955.
- [51] Gonçalves MLL, Santos EM, Renno ACM. *et al.* Erythrosine as a photosensitizer for antimicrobial photodynamic therapy with blue light-emitting diodes - An in vitro study. *Photodiagnosis Photodyn Ther.* (2021). doi: 10.1016/j.pdpdt.2021.102445. Epub 2021 Jul 18. PMID: 34284146.
- [52] Soderholm JF, Bird SL, Kalab P. *et al.* Importazole, a small molecule inhibitor of the transport receptor importin- β . *ACS Chem Biol.* (2011). doi: 10.1021/cb2000296. Epub 2011 Apr 21. PMID: 21469738; PMCID: PMC3137676.
- [53] Moriya K, Rivera J, Odom S. *et al.* ER-27319, an acridone-related compound, inhibits release of antigen-induced allergic mediators from mast cells by selective inhibition of fcepsilon receptor I-mediated activation of Syk. *Proc Natl Acad Sci U S A.* (1997). doi:

10.1073/pnas.94.23.12539. PMID: 9356485; PMCID: PMC25030.

[54] Zhong, X., Wei, Q., Tiwari, A. *et al.* A genetics-guided integrative framework for drug repurposing: identifying antihypertensive drugs for type 2 diabetes. *npj Drug Discov.* **3**, 3 (2026). <https://doi.org/10.1038/s44386-025-00032-3>

[55] Kuleshov MV, Jones MR, Rouillard AD, Fernandez NF, Duan Q, Wang Z, Koplev S, Jenkins SL, Jagodnik KM, Lachmann A. Enrichr: a comprehensive gene set enrichment analysis web server 2016 update. *Nucleic Acids Res.* doi: 10.1093/nar/gkw377. (2016). PMID: 27141961; PMCID: PMC4987924.

[56] Yuen CA, Zheng M, Saint-Germain MA. *et al.* Meningioma: Novel Diagnostic and Therapeutic Approaches. *Biomedicines.* (2025). doi: 10.3390/biomedicines13030659. PMID: 40149634; PMCID: PMC11940373.

[57] Li H, Ahmad M, Zhou D. Relaxin-2 Inhibits Cervical Cancer Growth by Downregulating Kruppel-Like Factor 5 (KLF5) Through the Inhibition of Hippo-YAP and JAK2/STAT3 Signaling Pathways. *J Biochem Mol Toxicol.* (2026). doi: 10.1002/jbt.70687. PMID: 41499346.