

GEPIA3: Enhanced drug sensitivity and interaction network analysis for cancer research

Yu-Jian Kang^{1,2,†}, Lingjie Pan^{2,†}, Yiyu Liu^{2,†}, Zhengqin Rong^{3,4}, Jiaxi Liu², Fenglin Liu^{2,*}

¹Chongqing Key Laboratory of Intelligent Oncology for Breast Cancer, Cancer Hospital, School of Medicine, Chongqing University, Chongqing 400030, China

²School of Life Sciences, Peking University, Beijing 100871, China

³Wuxi Yituo Biotechnology Co., Ltd., Wuxi 214000, China

⁴Biomedical Informatics & Genomics Center, Key Laboratory of Biomedical Information Engineering of Ministry of Education, School of Life Science and Technology, Xi'an Jiaotong University, Xi'an, Shaanxi 710049, P. R. China

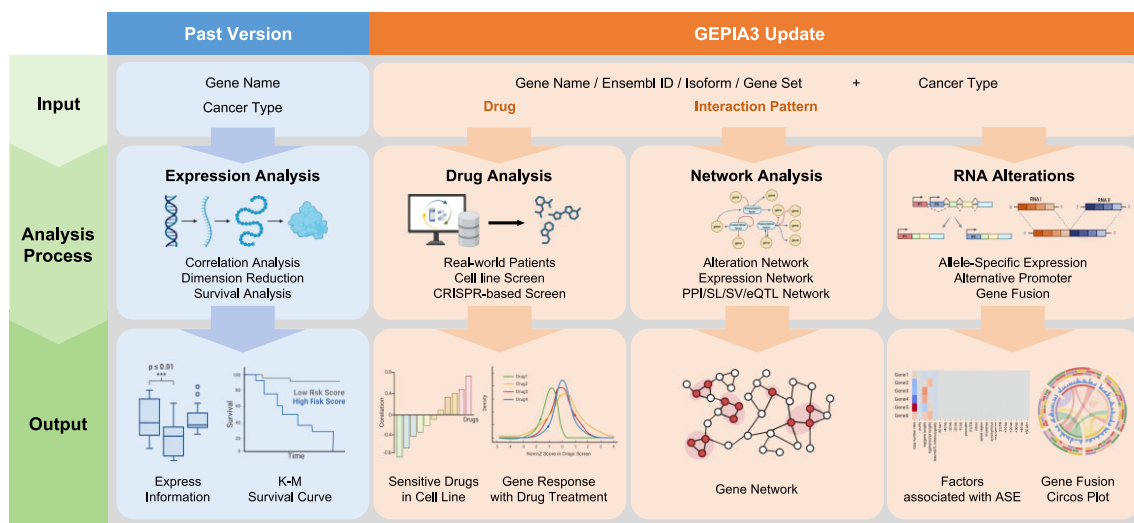
*To whom correspondence should be addressed. Email: flf@pku.edu.cn

†These authors contribute equally to this work.

Abstract

The GEPIA series has provided robust and widely used tools for pan-cancer analysis of gene expression data. In the post-genomic era, a major challenge lies in deconvoluting complex regulatory relationship influenced by multiple factors and discovering gene-based precision therapeutics. Here we present GEPIA3, an advanced version of GEPIA that provides a comprehensive analysis of gene/protein interactions across various cancer types. This version facilitates the investigation of treatment sensitivity utilizing both real-world patient data and cell line screens for over 1000 therapeutic agents, as well as the integration of RNA alterations derived from the pan-cancer analysis of whole genomes project. GEPIA3 represents a significant enhancement of the original platform, enabling in-depth exploration of gene regulation and cancer phenotypes, thereby supporting the identification of novel biomarkers and therapeutic targets. GEPIA3 is publicly accessible at <https://gepia3.bioinfo.pku.edu.cn>.

Graphical abstract



Introduction

The advent of high-throughput sequencing technologies has revolutionized cancer research by providing large-scale multi-omics data that enable detailed investigations of gene expression, mutation landscapes, and tumor heterogeneity. Notably, The Cancer Genome Atlas (TCGA) [1] and the Genotype-

Tissue Expression (GTEx) projects [2] have been pivotal in generating critical RNA sequencing data for understanding both cancerous and normal tissues. These resources have driven the development of multiple bioinformatic tools, including the Gene Expression Profiling Interactive Analysis (GEPIA) series [3–5]. The original [3] web server, launched in

Received: February 23, 2025. Revised: April 18, 2025. Editorial Decision: May 1, 2025. Accepted: May 7, 2025

© The Author(s) 2025. Published by Oxford University Press on behalf of Nucleic Acids Research.

This is an Open Access article distributed under the terms of the Creative Commons Attribution-NonCommercial License

(<https://creativecommons.org/licenses/by-nc/4.0/>), which permits non-commercial re-use, distribution, and reproduction in any medium, provided the original work is properly cited. For commercial re-use, please contact reprints@oup.com for reprints and translation rights for reprints. All other permissions can be obtained through our RightsLink service via the Permissions link on the article page on our site—for further information please contact journals.permissions@oup.com.

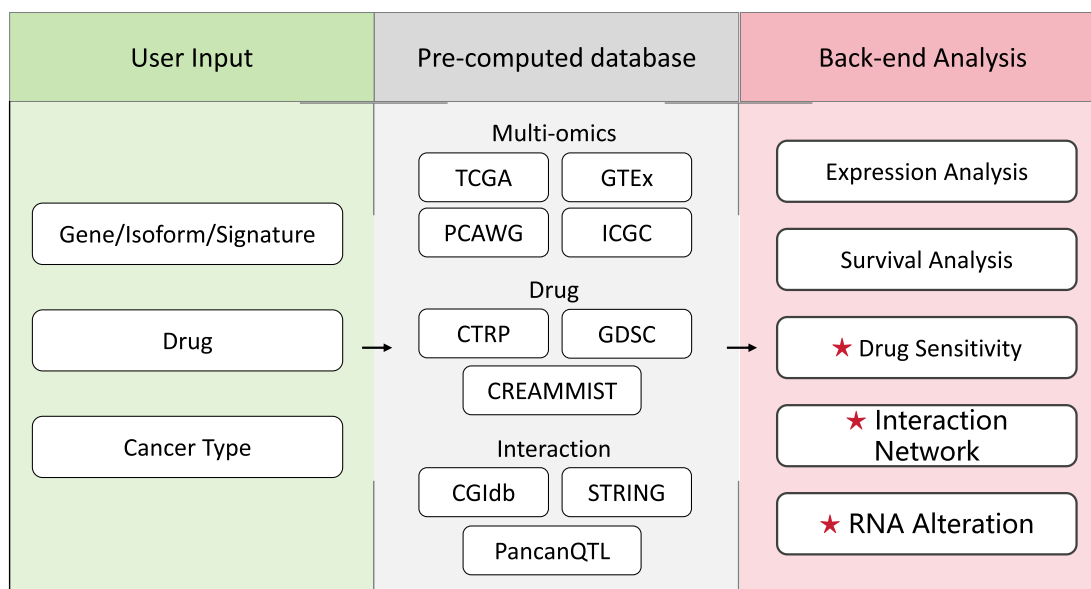


Figure 1. Schema describing the overall architecture of GEPIA3. Red stars mark the new functions in GEPIA3..

2017, provided users with a user-friendly interface to conduct differential gene expression analyses, survival assessments, and gene correlation studies across diverse cancer datasets derived from TCGA and GTEx. Subsequent iterations, GEPIA2 [4] and GEPIA2021 [5], introduced enhanced customization, advanced visualization capabilities, and deconvolution-based methods for investigating tumor microenvironment heterogeneity.

Nowadays, a significant challenge in cancer research is bridging the gap between genetic discoveries and clinical applications [6]. Earlier GEPIA versions enabled the discovery of diagnostic biomarkers, prognostic indicators and potential therapeutic targets, leaving a crucial need for linking these molecular discoveries to pharmacological implementation [7]. GEPIA3 developed new modules with gene-drug interaction analysis. Analyzing the relationship between gene expression and drug responses in TCGA cohorts uncovers predictive biomarkers for treatment decisions and candidate targets for drug development [8, 9]. Particularly for industrial users, the real-world cohort relevant markers enable precision clinical trial design through molecular stratification of patient sub-populations [10].

Differential gene expression in cancer is often driven by underlying genetic alterations [11]. Deciphering these disruptions is essential for uncovering regulatory relationships involved in tumorogenesis. Allele-specific expression (ASE) and expression quantitative trait loci (eQTL) analysis shed lights on identifying gene regulation in cancer [12, 13]. ASE represents the differences in transcript levels between two alleles within the same sample, enabling direct detection of *cis*-regulatory effects while minimizing the influence of extrinsic factors such as cellular environment or gene interactions [12]. An eQTL is a locus that accounts for a portion of the genetic variance in gene expression phenotype, typically involving SNPs associated with expression levels [13]. GEPIA3 incorporates these events to uncover the underlying mechanisms of gene expression variation in cancer.

Moreover, cancer-associated genes rarely act in isolation. Their functional impact is deeply influenced by interactions within complex biological networks [14, 15]. Revealing multi-scale genetic interactions, such as co-expression [16] and co-

occurrence/mutual exclusivity of alterations [17], elucidate critical oncogenic pathways and novel vulnerabilities. In addition, screening synthetic lethal partners of oncogenes offers a promising strategy for selective cancer cell killing [18], on the fundamental principle that simultaneous genetic perturbation of two targets induces cell death, whereas individual inactivation of either gene maintains cellular viability [19]. Incorporated with the multi-scale networks, GEPIA3 facilitates investigating comprehensive networks beyond expression and identifying functional pathways in cancer.

In response to these emerging challenges and user feedback, we present GEPIA3, the latest iteration of the GEPIA series. GEPIA3 introduces innovative modules that enhance analytical capabilities, including: (i) a comprehensive drug sensitivity analysis that integrates patient-derived data from TCGA with cell-line screening data for over 1000 therapeutic agents, (i) offering advanced analyses of gene/protein interactions, as well as regulatory networks across multiple cancer types, c) incorporating RNA alterations—such as allele-specific expression, alternative promoters, and RNA fusions—based on pan-cancer analysis of whole genomes (PCAWG) project [20]. These updates provide an effective and comprehensive platform for pan-cancer multi-omics analysis, empowering researchers to mine large-scale genomic datasets for exploring cancer mechanisms, therapeutic targets, and precision medicine strategies.

Materials and methods

Web server architecture

GEPIA3 features a more streamlined and explicit architecture for web server implementation, with all functions organized into five distinct modules: Expression Analysis, Survival Analysis, Drug Analysis, Network Analysis, and RNA Alterations. Datasets are stored in a PostgreSQL database. The server-side infrastructure is built using Nginx and Django to handle custom requests. The front-end architecture, based on the framework from GEPIA1/2/2021 (HTML5, CSS3, and JavaScript), has been enhanced with visualizations through Matplotlib [21], Seaborn [22], and ECharts [23] (Fig. 1).

Table 1. Data sources of GEPIA3

Functionalities	Resources	version	Download source
Expression analysis	TCGA	PANCAN 2016–09-01	https://xenabrowser.net
	GTEX	2016–04-19	https://xenabrowser.net
Drug cell line responses	GDSC	GDSC1, GDSC2	https://www.cancerrxgene.org
	CTRP	v2.1	https://ctd2-data.nci.nih.gov/Public/Broad/CTRPv2.1_2016_pub_NatChemBiol_12_109/
Drug CRISPR screen	CREAMMIST	-	https://creammist.mtms.dev
	BioGRID ORCS	1.1.17	https://downloads.thebiogrid.org/BioGRID-ORCS/Release-Archive/BIOGRID-ORCS-1.1.17/
Alteration network	PCAWG	-	https://docs.icgc-argo.org/docs/data-access/icgc-25k-data
	PCAWG	-	[34]
	Transcriptome Core Group <i>et al. Nature.</i> 2020		
Protein Interaction	Xiong <i>et al. Nat Biotechnol.</i> 2024	-	[41]
eQTL	PancanQTL	-	http://bioinfo.life.hust.edu.cn/PancanQTL/
SL/SV	CGIdb	CGIdb2.0	http://www.medsysbio.org/CGIdb2/
Allele-specific Expression	PCAWG	-	[34]
	Transcriptome Core Group <i>et al. Nature.</i> 2020		
Alternative Promoter	PCAWG	-	https://pcawg.xenahubs.net
Gene Fusion	PCAWG	-	https://docs.icgc-argo.org/docs/data-access/icgc-25k-data

Data resources

GEPIA3 newly incorporated pan-cancer multi-omics data including drug sensitivity, gene interaction networks, multi-level gene alterations (Table 1 and Table S1). We also updated TCGA and GTEx expression datasets into a new version that increased the sample size from 18 005 to 20 594. The data processing details were provided at supplementary materials.

Expression analysis updates

GEPIA3 newly incorporates DESeq2 as an optional method for differential expression analysis, which is performed based on gene-level read count data [24]. Given the widespread focus on gene signatures in tumor research, all expression-based functions in GEPIA3 support the use of multiple genes as input, and then compute the signature score using the first principal component value derived from Principal Component Analysis of the expression levels of the input genes. Additionally, by integrating DNA Single Nucleotide Variations (SNVs), GEPIA3 facilitates differential expression analysis both with and without the consideration of tumor hotspot mutations for TCGA datasets.

Survival analysis

To assess the combined expression effect of multiple genes on prognosis, GEPIA3 developed multivariable survival analysis module for user-defined genes. We employed the Python package lifelines for survival analysis, which included the generation of Kaplan–Meier curves, calculation of 95% confidence intervals, log-rank tests, and Cox regression analysis.

Drug sensitivity

GEPIA3 enables the discovery of associations between drug sensitivity and gene expression across multiple datasets. For real-world patient drug responses, GEPIA3 adopted a systematic standardization strategy for TCGA drug response data (see “Quality Control for TCGA Drug Response Data” in supplementary material) and analyzed differential gene expression in relation to therapy outcomes (compare survival between patients receiving specific treatments and those

without). In the context of cancer cell drug sensitivity screening, GEPIA3 calculates the correlation between query gene expression and drug responses across various cell line databases, including CREAMMIST [25], GDSC1 [26], GDSC2 [27], and CTRPv2.1 [28]. Furthermore, CRISPR screening was a powerful tool for identifying gene-drug interactions to uncover genetic vulnerabilities with agents, such as novel resistance genes with BRAF inhibitor [29, 30]. GEPIA3 integrates 4 studies collected from BioGRID database [31], and compares cell responses to gene perturbations under cancer therapeutic conditions versus untreated controls.

Network analysis

The network analysis module integrates the Expression Network from GEPIA2’s Similar Genes Detection, while also considering dissimilar genes. It incorporates oncogenic protein-protein interactions, synthetic lethality/synthetic viability and co-occurrence or mutual exclusivity of gene alterations to construct gene networks. Co-occurrence or mutual exclusivity of gene alterations was tested using DISCOVER [32], followed by Benjamini–Hochberg (BH) multiple testing correction. Only cancer types with more than 50 samples were included. Data with a false discovery rate ≤ 0.05 from OncoPPI were retained, and STRING interactions were filtered to include only those with a medium confidence score (≥ 0.4).

For Network visualization, the top 10 genes with the highest degree centrality were displayed. In the Expression Network, the top 10 genes most correlated with the input gene were visualized. Network annotations were enhanced by integrating the STRING [33] database, facilitating gene-level mapping and providing biologically relevant interaction data.

RNA alterations

In GEPIA3, the term “RNA alterations” refers to transcript-level regulatory or structural variations that go beyond total expression abundance [34]. GEPIA3 currently supports three types of RNA alterations: ASE, alternative promoter usage, and gene fusions. The ASE module visualizes the expression factors of query genes in heatmap, including copy number ratio, sample purity, gene and transcript length (both

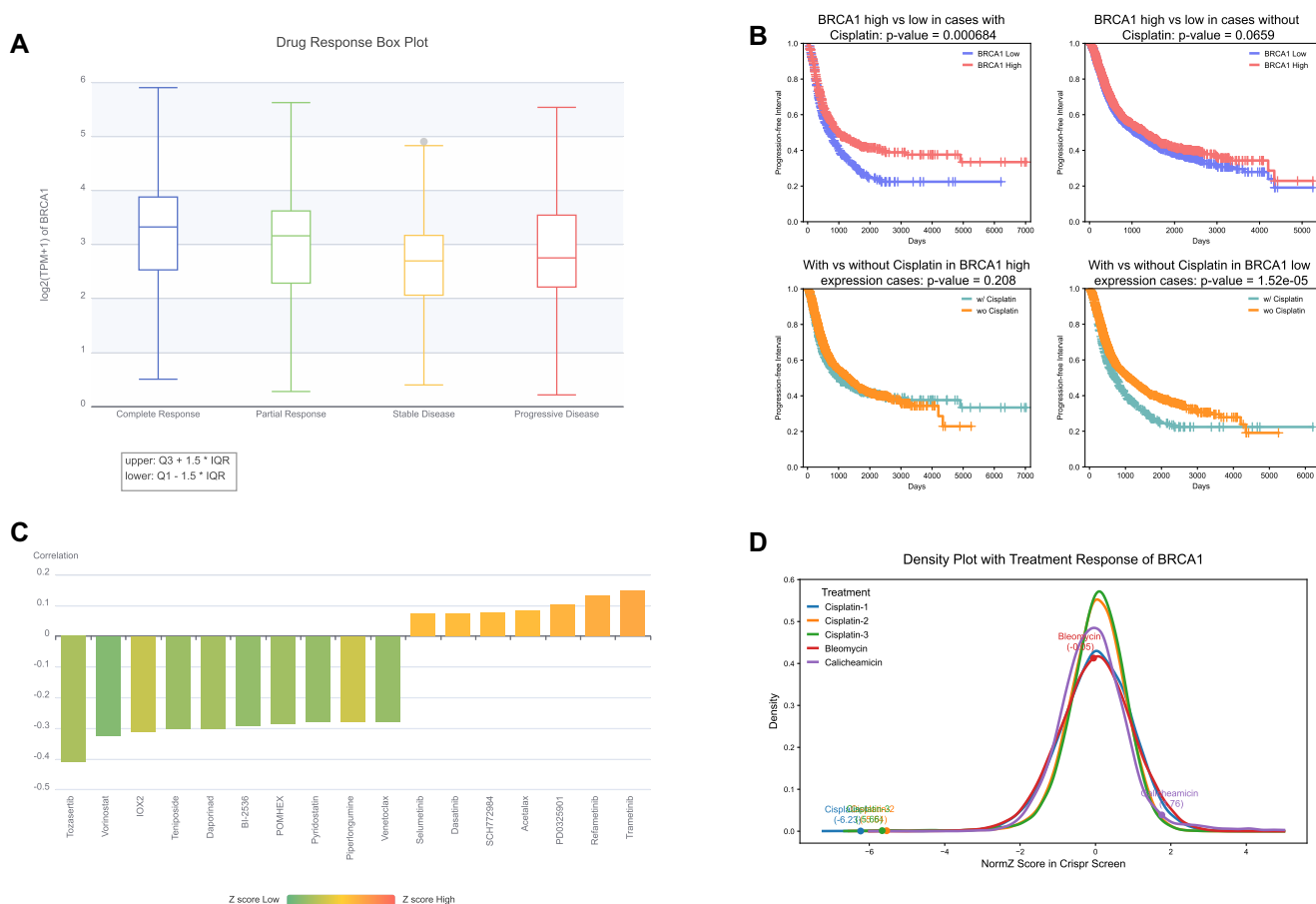


Figure 2. Examples of GEPIA3 outputs in Drug Analysis module. **(A)** BRCA1 expression level in TCGA patient primary tumor tissues with different Cisplatin responses. The y axis shows the expression level of query gene ($\log_2(TPM + 1)$), while x axis shows the drug responses annotated by TCGA (complete response/partial response/stable disease/progressive disease). **(B)** Comparative analysis of progression-free survival (PFS) stratified by Cisplatin treatment status and BRCA1 expression levels (high vs low). **(C)** Top10 sensitive drugs with BRCA1 high expression and low expression respectively in cell lines. The x-axis shows the drug names, while y axis shows the corresponding correlation coefficient between query gene expression abundance and drug response. The colors represent Fisher Z-score of correlation. Only significant gene-drug pairs were plotted. **(D)** The density plot of all genes' CRISPR Z-score (normalized sgRNA reads counts) between drug-treated and untreated cells. The dots on density line point the Z-score of BRCA1 (query gene). Negative scores represent genes whose mutation leads to their depletion from the cell population, whereas positive NormZ scores represent genes whose mutation leads to a selective growth advantage in the presence of the drug [35]. Thus, BRCA1 mutation exhibited Cisplatin resistance in the screen.

log-transformed), heterozygosity of the lead eQTL variant, and somatic mutational burden categories. The alternative promoter analysis displayed the mean relative promoter activity across different tissues (tumor-peritumor tissue pairs with p-value less than 0.05 were annotated on the graph). Gene fusion event circos plot was visualized using pyCirclize.

Results

Drug sensitivity

Genome-based identification of drug sensitivity is crucial for the development and application of genotoxic agents in cancer therapy. To this end, GEPIA3 introduced a new module to analyze the correlation between drug sensitivity and gene expression/variation using multiple data sources, including real-world patient data, cell line screening, and CRISPR-based screening.

For real-world patient data, GEPIA3 utilized survival and therapy response annotations from TCGA. Differential ex-

pression analysis identified drugs with distinct gene expression profiles across varying therapeutic responses (Fig. 2A). Survival analysis for each drug enabled users to compare gene expression and survival outcomes between patients with and without specific drug treatments (Fig. 2B).

In cell lines, GEPIA3 integrated data from the CREAMIST, GDSC, and CTRP databases [28], correlating gene expression or copy number variations with drug sensitivity scores (IC50, IC90, EC50, Einf, AUC) to identify genotoxic drugs across multiple cell types (Fig. 2C). Finally, GEPIA3 compared cell responses to CRISPR gene perturbations, distinguishing between DNA damage therapy-treated and untreated conditions [35], allowing users to identify selective growth advantages or depletions driven by specific gene mutations (Fig. 2D).

Network analysis

The complex interplay of gene interaction networks constitutes a fundamental mechanism driving oncogenesis and malignant transformation. By integrating multiple data types in-

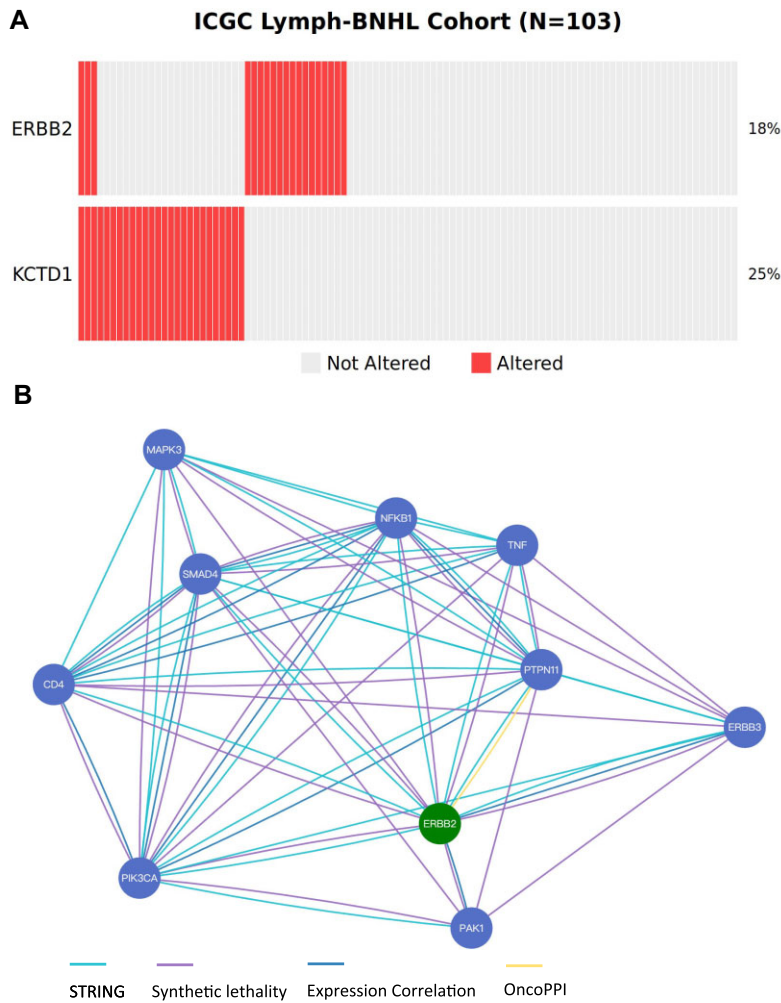


Figure 3. Example of GEPIA3 outputs in Network Analysis module. **(A)** Heatmap showing the mutually exclusive alteration of *ERBB2* and *KCTD1* at the sample level in the ICGC Lymph-BNHL cohort. The alterations include alternative promoters, expression outliers, variants, alternative splicing, allele-specific expression, and copy number variations. The heatmap visualizes the presence (red) or absence (gray) of alterations across different samples, with the percentage on the right indicating the number of samples with alterations for each gene. **(B)** Comprehensive analysis of the gene network centered around *ERBB2* in COAD. Nodes represent genes, and edges indicate gene interaction relationships. The network displays *ERBB2* and its neighboring nodes with the top 10 degree centrality scores, considering gene relationships including synthetic lethality, expression correlation, and OncoPPI (protein–protein interactions significantly enriched with somatic mutations in their interfaces). Different edge colors represent various types of interactions: purple edges denote synthetic lethality, dark blue edges indicate expression correlation, light yellow edges represent OncoPPI, and light blue edges correspond to medium-score associations in the STRING database for gene–protein interaction.

cluding gene alterations, the network analysis module aims to uncover the molecular mechanisms that drive tumorigenesis and identify potential therapeutic targets across various cancer types.

The alteration network allows users to explore co-mutation and mutual exclusivity patterns, with heatmaps visualizing mutation distributions across ICGC samples, offering insights into genetic alterations at the individual sample level (Fig. 3A). The expression network identifies genes correlated with the target gene across cancer and normal tissues. The oncoPPI module examines protein–protein interactions enriched with somatic mutations across TCGA cancer types, revealing potential oncogenic mechanisms. The SL/SV module identifies genes with synthetic lethality or viability relationships, aiding therapeutic exploration.

Additionally, the comprehensive analysis subsection integrates these modules using a degree centrality-based algorithm to uncover key genes linked to the gene of interest (Fig. 3B).

Graph-based visualizations, enriched with STRING database annotations, provide an intuitive representation of gene networks. GEPIA3 enables users to explore and visualize these complex relationships through a unified platform, facilitating the identification of driving node in tumorigenesis.

RNA alterations

RNA alterations expand the scope of transcriptome regulation beyond gene expression abundance. GEPIA3 developed a new module enabling on allelic expression imbalance and visualizes the effect sizes of multiple factors and somatic mutation burden categories via heatmaps (Fig. 4A). In the Alternative Promoter subsection, users can explore promoter activity across cancer types. GEPIA3 provides bar plots to compare promoter usage between tumor and peritumoral tissue, highlighting differences in activity (Fig. 4B). The Gene Fusion subsection facilitates the search for gene fusion events within spe-

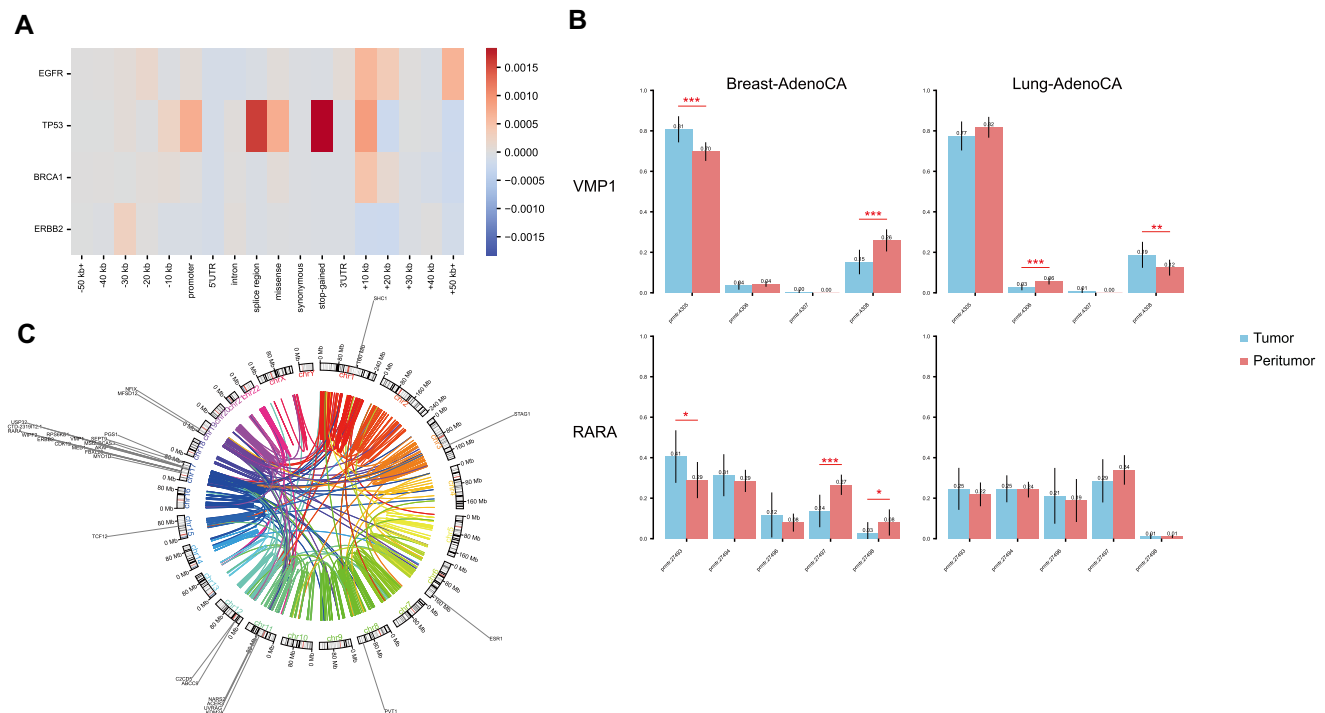


Figure 4. Example of GEPIA3 outputs in the RNA Alterations module. **(A)** Heatmaps showing the effect sizes of various factors on gene allelic expression imbalance, as well as the input somatic mutation burdens. **(B)** Relative promoter activities of *VMP1* and *RARA* in breast adenocarcinoma, lung adenocarcinoma, and their corresponding peritumor tissues. The numbers above the bars represent the mean promoter activity values (in black), with error bars indicating the standard deviation. *P*-values for tumor-peritumor tissue pairs with *P*-values < 0.05 are labeled in red. **(C)** Circos plot displaying all gene fusion events in breast adenocarcinoma. Only the top 20 most frequent genes are annotated on the plot.

cific cancer types, focusing on individual genes or all associated fusions. GEPIA3 generates circos plots to map fusion locations across chromosomes and bar plots to display the most frequent genes identified. The characterization of gene fusions in breast cancer (Fig. 4C) reflects the polysomy of chromosome 17.

Use cases

Lung cancer is the leading cause of cancer-related deaths worldwide [36]. The well-known oncogene *MYC* gene contributes to pathogenesis of various cancers [37], and shows drug resistance property through both overexpression and mutation in non-small cell lung cancer (NSCLC) [38].

Using GEPIA3, we found that high *MYC* expression in LUAD was significantly associated with poor prognosis in patients treated with carboplatin ($P = 0.0493$, log-rank test, Fig. 5A), but not in those untreated with carboplatin ($P = 0.673$, log-rank test), linking *MYC* overexpression to carboplatin resistance and chemotherapy efficacy reduction reported by previous study [39].

GEPIA3 also revealed interactions between *MYC* and key genes such as *JUN* as well as genes in ERK MAPK signaling pathway (Fig. 5B). Integrated analysis via GEPIA3 identifies synthetic lethality between *MYC* amplification and MAPK pathway dysregulation, with *MYC* overexpression conferring selective vulnerability to ERK inhibitors in lung cancer cell lines (Table 2). Considering previous studies showing that ERK inhibition suppresses *MYC* and impairs pancreatic cancer proliferation [40], similar mechanisms and potential therapeutic combinations in Table 2 are suggested in lung cancer,

offering a potential strategy to reverse drug resistance. To elucidate the molecular basis of *MYC* dysregulation, the eQTL module in GEPIA3 allows for the querying of SNPs associated with *MYC* expression levels across different cancer types.

We also revealed that among lung cancer patients with high *KRAS* expression, those receiving Gemcitabine treatment demonstrated significantly poor survival outcomes compared to those without Gemcitabine treatment, representing a candidate signature for drug resistance (Supplementary Fig. S3). In addition, cases with informative RNA alterations (see the *RASSF1* and *ALK* sections in supplementary material), together with *MYC*, highlight the significance of GEPIA3 updates in discovering novel oncogenic mechanisms and potential therapeutic strategies as extensions of gene expression.

Discussions

We developed GEPIA3 for exploring the relationship between gene expression, genetic alterations, and drug sensitivity in cancer research. By integrating diverse data sources, including real-world patient data, cell line screening, and CRISPR-based perturbations, GEPIA3 provides critical insights into the molecular mechanisms driving drug resistance and therapeutic efficacy across cancer types. Its key functionalities include survival analysis, gene expression profiling, regulatory network construction, and the investigation of RNA alterations such as ASE, alternative promoter usage, and gene fusions.

A major strength of GEPIA3 is its ability to analyze survival outcomes and gene expression profiles linked to drug responses, particularly in chemotherapy-treated patients. By

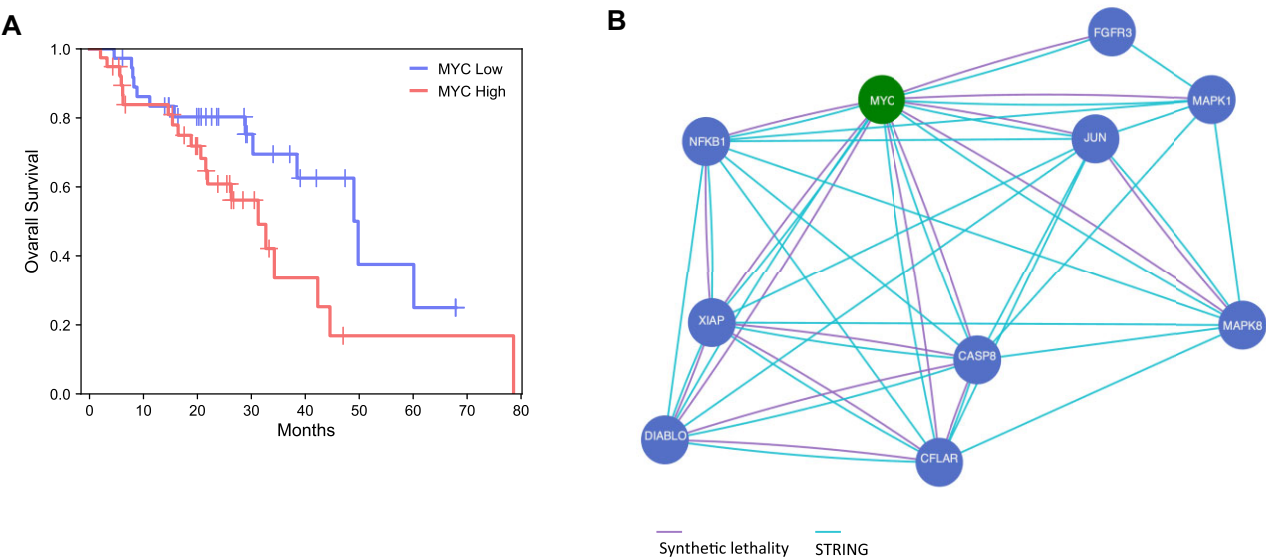


Figure 5. Comprehensive Examples Using GEPIA3. **(A)** Overall survival analysis comparing *MYC* high- and low-expression groups in LUAD (KM curves). **(B)** Synthetic lethality analysis of the gene network centered around *MYC* in LUAD. Nodes represent genes, with edges indicating gene interactions. Purple edges denote synthetic lethality, while light blue edges represent associations from the STRING database of gene/protein interactions.

Table 2. Top five sensitive drug with *MYC* expression in lung derived cell lines in GDSC2 output by GEPIA3 (ranked by q-value)

Drug name	Cell Line Number	Correlation (<i>MYC</i> expression vs IC50)	Z score	P-value	Q-value	Direction	Targeted pathways	Putative targets
WZ4003	150	−0.347	−4.39	1.37E-05	0.00 275	sensitivity with high expression	Other, kinases	NUAK1, NUAK2
VX-11e	165	−0.318	−4.19	3.13E-05	0.00 275	sensitivity with high expression	ERK MAPK signaling	ERK2
Refametinib	164	−0.31	−4.06	5.47E-05	0.00 275	sensitivity with high expression	ERK MAPK signaling	MEK1, MEK2
Ulixertinib	166	−0.306	−4.03	6.21E-05	0.00 275	sensitivity with high expression	ERK MAPK signaling	ERK1, ERK2
I-BET-762	150	−0.32	−4.02	6.62E-05	0.00 275	sensitivity with high expression	Chromatin other	BRD2, BRD3, BRD4

analyzing complex molecular networks, GEPIA3 aids in identifying prognostic biomarkers and potential therapeutic targets. The combination of these functionality aids in identifying variations that may influence drug sensitivity or resistance, enabling more personalized treatment approaches.

Acknowledgements

We thank National Teaching Center for Experimental Biology, Peking University.
Author contributions: F.L. and Y.K. designed the project. Y.K. developed the drug sensitivity module. L.P. and J.L. developed the regulatory network module. Y.L. developed the RNA alterations module. Z.R. reconstructed the basic functionalities of GEPIA and updated the data. F.L., Y.K., L.P., Y.L., J.L., and Z.R. wrote the manuscript. F.L. supervised the study.

Supplementary data

Supplementary data is available at NAR online.

Conflict of interest

R.Z. is an employee of Wuxi Yiou Biotechnology Company. The remaining authors declare no competing interests.

Funding

This work was supported by Peking University Undergraduate Teaching Reform Funding [grant number JG2025013] and Fundamental Research Funds for the Central Universities [grant number 2024CDJYXTD-010]. Funding to pay the Open Access publication charges for this article was provided by Peking University Undergraduate Teaching Reform Funding.

Data availability

GEPIA3 is publicly accessible at <https://gepia3.bioinfoliu.com/>. This website is free and open to all users and there is no login requirement.

References

1. Weinstein JN, Collisson EA, Mills GB *et al.* The cancer genome atlas pan-cancer analysis project. *Nat Genet* 2013;45:1113–20. <https://doi.org/10.1038/ng.2764>
2. Lonsdale J, Thomas J, Salvatore M *et al.* The Genotype-Tissue Expression (GTEx) project. *Nat Genet* 2013;45:580–5. <https://doi.org/10.1038/ng.2653>
3. Tang Z, Li C, Kang B *et al.* GEPIA: a web server for cancer and normal gene expression profiling and interactive analyses. *Nucleic Acids Res* 2017;45:W98–102. <https://doi.org/10.1093/nar/gkx247>

4. Tang Z, Kang B, Li C *et al.* GEPIA2: an enhanced web server for large-scale expression profiling and interactive analysis. *Nucleic Acids Res* 2019;47:W556–60. <https://doi.org/10.1093/nar/gkz430>
5. Li C, Tang Z, Zhang W *et al.* GEPIA2021: integrating multiple deconvolution-based analysis into GEPIA. *Nucleic Acids Res* 2021;49:W242–6. <https://doi.org/10.1093/nar/gkab418>
6. Elmore LW, Greer SF, Daniels EC *et al.* Blueprint for cancer research: critical gaps and opportunities. *CA Cancer J Clin* 2021;71:107–39. <https://doi.org/10.3322/caac.21652>
7. Collins FS, Varmus H. A new initiative on precision medicine. *N Engl J Med* 2015;372:793–5. <https://doi.org/10.1056/NEJMp1500523>
8. Dang CV, Reddy EP, Shokat KM *et al.* Drugging the ‘undruggable’ cancer targets. *Nat Rev Cancer* 2017;17:502–8. <https://doi.org/10.1038/nrc.2017.36>
9. Clayton EA, Pujol TA, McDonald JF *et al.* Leveraging TCGA gene expression data to build predictive models for cancer drug response. *BMC Bioinf* 2020;21:364. <https://doi.org/10.1186/s12859-020-03690-4>
10. Perez-Gracia JL, Sanmamed MF, Bosch A *et al.* Strategies to design clinical studies to identify predictive biomarkers in cancer research. *Cancer Treat Rev* 2017;53:79–97. <https://doi.org/10.1016/j.ctrv.2016.12.005>
11. Delgado MD, León J. Gene expression regulation and cancer. *Clin Transl Oncol* 2006;8:780–7. <https://doi.org/10.1007/s12094-006-0132-7>
12. Pastinen T. Genome-wide allele-specific analysis: insights into regulatory variation. *Nat Rev Genet* 2010;11:533–8. <https://doi.org/10.1038/nrg2815>
13. Gilad Y, Rifkin SA, Pritchard JK. Revealing the architecture of gene regulation: the promise of eQTL studies. *Trends Genet* 2008;24:408–15. <https://doi.org/10.1016/j.tig.2008.06.001>
14. Hu P, Bader G, Wigle DA *et al.* Computational prediction of cancer-gene function. *Nat Rev Cancer* 2007;7:23–34. <https://doi.org/10.1038/nrc2036>
15. de la Fuente A. From ‘differential expression’ to ‘differential networking’ - identification of dysfunctional regulatory networks in diseases. *Trends Genet* 2010;26:326–33. <https://doi.org/10.1016/j.tig.2010.05.001>
16. Savino A, Provero P, Poli V. Differential Co-expression analyses allow the identification of critical signalling pathways altered during tumour transformation and progression. *Int J Mol Sci* 2020;21, 9461. <https://doi.org/10.3390/ijms21249461>
17. El Tekle G, Bernasocchi T, Unni AM *et al.* Co-occurrence and mutual exclusivity: what cross-cancer mutation patterns can tell us. *Trends Cancer* 2021;7:823–36. <https://doi.org/10.1016/j.trecan.2021.04.009>
18. Chan DA, Giaccia AJ. Targeting cancer cells by synthetic lethality: autophagy and VHL in cancer therapeutics. *Cell Cycle* 2008;7:2987–90. <https://doi.org/10.4161/cc.7.19.6776>
19. O’Neil NJ, Bailey ML, Hieter P. Synthetic lethality and cancer. *Nat Rev Genet* 2017;18:613–23. <https://doi.org/10.1038/nrg.2017.47>
20. Aaltonen LA, Abascal F, Abeshouse A *et al.* Pan-cancer analysis of whole genomes. *Nature* 2020;578:82–93. <https://doi.org/10.1038/s41586-020-1969-6>
21. Hunter JD. Matplotlib: a 2D graphics environment. *Comput Sci Eng* 2007;9:90–5. <https://doi.org/10.1109/MCSE.2007.55>
22. Waskom ML. Seaborn: statistical data visualization. *JOSS* 2021;6:3021. <https://doi.org/10.21105/joss.03021>
23. Li D, Mei H, Shen Y *et al.* ECharts: a declarative framework for rapid construction of web-based visualization. *Visual Inform* 2018;2:136–46. <https://doi.org/10.1016/j.visinf.2018.04.011>
24. Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol* 2014;15:550. <https://doi.org/10.1186/s13059-014-0550-8>
25. Yingtaeweessittikul H, Wu J, Mongia A *et al.* CREAMMIST: an integrative probabilistic database for cancer drug response prediction. *Nucleic Acids Res* 2023;51:D1242–8. <https://doi.org/10.1093/nar/gkac911>
26. Yang W, Soares J, Greninger P *et al.* Genomics of drug sensitivity in cancer (GDSC): a resource for therapeutic biomarker discovery in cancer cells. *Nucleic Acids Res* 2013;41:D955–61. <https://doi.org/10.1093/nar/gks1111>
27. Iorio F, Knijnenburg TA, Vis DJ *et al.* A landscape of pharmacogenomic interactions in cancer. *Cell* 2016;166:740–54. <https://doi.org/10.1016/j.cell.2016.06.017>
28. Rees MG, Seashore-Ludlow B, Cheah JH *et al.* Correlating chemical sensitivity and basal gene expression reveals mechanism of action. *Nat Chem Biol* 2016;12:109–16. <https://doi.org/10.1038/nchembio.1986>
29. Shalem O, Sanjana NE, Hartenstein E *et al.* Genome-scale CRISPR-Cas9 knockout screening in human cells. *Science* 2014;343:84–7. <https://doi.org/10.1126/science.1247005>
30. Konermann S, Brigham MD, Trevino AE *et al.* Genome-scale transcriptional activation by an engineered CRISPR-Cas9 complex. *Nature* 2015;517:583–8. <https://doi.org/10.1038/nature14136>
31. Oughtred R, Rust J, Chang C *et al.* The BioGRID database: A comprehensive biomedical resource of curated protein, genetic, and chemical interactions.. *Protein Sci* 2021;30:187–200. <https://doi.org/10.1002/pro.3978>
32. Canisius S, Martens JW, Wessels LF. A novel independence test for somatic alterations in cancer shows that biology drives mutual exclusivity but chance explains most co-occurrence. *Genome Biol* 2016;17:261. <https://doi.org/10.1186/s13059-016-1114-x>
33. Szklarczyk D, Kirsch R, Koutrouli M *et al.* The STRING database in 2023: protein-protein association networks and functional enrichment analyses for any sequenced genome of interest. *Nucleic Acids Res* 2023;51:D638–46. <https://doi.org/10.1093/nar/gkac1000>
34. Calabrese C, Davidson NR, Demircioğlu D *et al.* Genomic basis for RNA alterations in cancer. *Nature* 2020;578:129–36. <https://doi.org/10.1038/s41586-020-1970-0>
35. Olivieri M, Cho T, Álvarez-Quilón A *et al.* A genetic map of the response to DNA damage in human cells. *Cell* 2020;182:481–96. <https://doi.org/10.1016/j.cell.2020.05.040>
36. Zhang Y, You P, Liu R *et al.* Artificial intelligence in clinical trials of lung cancer: current and future prospects. *Intell Oncol* 2025;1:34–51. <https://doi.org/10.1016/j.intonc.2024.11.003>
37. Baluapuri A, Wolf E, Eilers M. Target gene-independent functions of MYC oncoproteins. *Nat Rev Mol Cell Biol* 2020;21:255–67. <https://doi.org/10.1038/s41580-020-0215-2>
38. Dhanasekaran R, Deutzmann A, Mahauad-Fernandez WD *et al.* The MYC oncogene — the grand orchestrator of cancer growth and immune evasion. *Nat Rev Clin Oncol* 2022;19:23–36. <https://doi.org/10.1038/s41571-021-00549-2>
39. Dong S, Li A, Pan R *et al.* Carboplatin-resistance-related DNA damage repair prognostic gene signature and its association with immune infiltration in breast cancer. *Front Immunol* 2025;16:1522149. <https://doi.org/10.3389/fimmu.2025.1522149>
40. Waters AM, Khatib TO, Papke B *et al.* Targeting p130Cas- and microtubule-dependent MYC regulation sensitizes pancreatic cancer to ERK MAPK inhibition. *Cell Rep* 2021;35:109291. <https://doi.org/10.1016/j.celrep.2021.109291>
41. Xiong D, Qiu Y, Zhao J *et al.* A structurally informed human protein-protein interactome reveals proteome-wide perturbations caused by disease mutations. *Nat Biotechnol* 2024; <https://doi.org/10.1038/s41587-024-02428-4>