

LnCeVar 2.0: an updated resource and web tools for genomic variations disrupting ceRNA networks from single-cell/spatial transcriptomics data

Qiuyan Guo^{1,†}, Qian Liu^{2,†}, Mengyu Xin^{2,†}, Houxing Li^{2,†}, Jiatong Li², Yifan Dai², Rui Sun², Yujie Zhang², Yuting He², Borui Xu², Xinxin Shan², Yue Gao², Shangwei Ning^{2,*}, Hui Zhi^{2,*}, Peng Wang^{2,*}

¹Department of Gynecology, the First Affiliated Hospital of Harbin Medical University, Harbin 150081, China

²College of Bioinformatics Science and Technology, Harbin Medical University, Harbin 150081, China

*To whom correspondence should be addressed. Email: wpgqy@163.com

Correspondence may also be addressed to Hui Zhi. Email: zhihui@ems.hrbmu.edu.cn

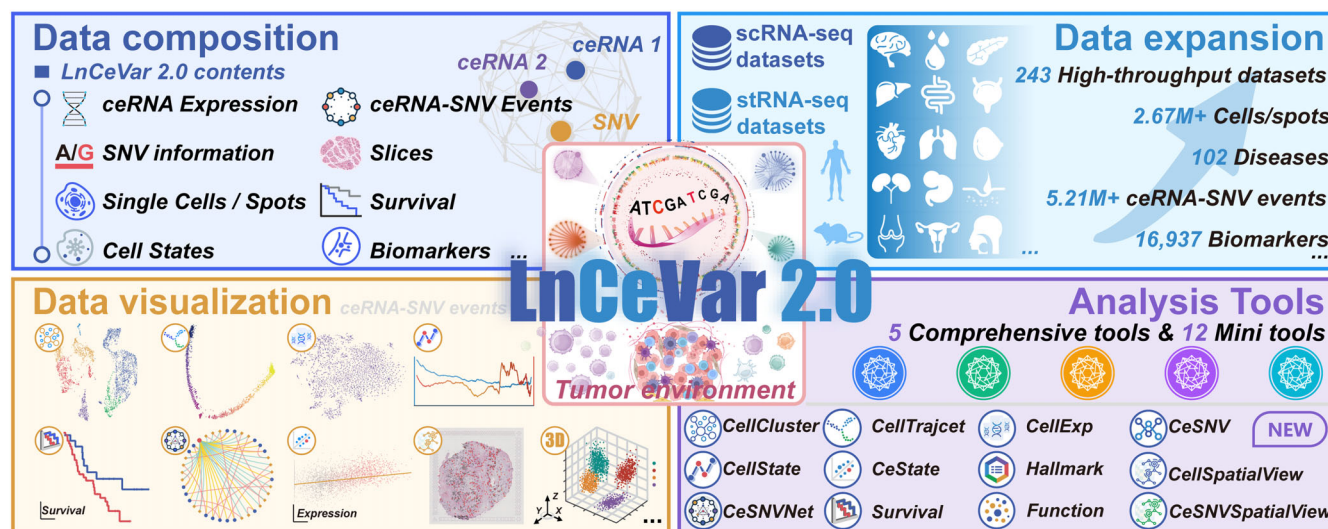
Correspondence may also be addressed to Shangwei Ning. Email: ningsw@ems.hrbmu.edu.cn

[†]The first four authors should be regarded as Joint First Authors.

Abstract

LnCeVar 2.0 (available at <http://bio-bigdata.hrbmu.edu.cn/LnCeVar> or <http://www.bio-bigdata.net/LnCeVar>) is an updated database investigating genomic variations that disrupt competing endogenous RNA (ceRNA) networks via single-cell and spatial transcriptomics. Enhancements include expanded data and improved features: (i) 16 937 experimentally supported cancer biomarkers as well as 5785 validated ceRNA interactions and single nucleotide variant (SNV)–ceRNA events, manually curated and linked to key cancer pathogenic processes; (ii) 812 single-cell RNA sequencing/spatial transcriptomics RNA sequencing datasets covering 102 diseases, clinical treatments (e.g. chemotherapy, immunotherapy), and normal tissues; (iii) 5 218 062 single-cell- and spatial-specific SNV–ceRNA events across 2 673 603 cells/spots, with cellular functional perturbation networks; (iv) 5 comprehensive and 12 mini tools for multilevel cross talk analysis and 3D visualization; and (v) novel inference of SNV effects on cell types, states, and functions at single-cell and spatial levels. LnCeVar 2.0 features a user-friendly interface for searching, browsing, and analyzing data. For instance, the *CeVarState* interface illustrates how SNV–ceRNA events influence cell states during developmental processes, revealing interactions that determine cell fate. The *CeVarSC3D* and *CeVarST3D* tools perform multilevel cross talk analyses of SNVs, ceRNA networks, and cell states in disease pathology, providing interactive 3D visualizations. Overall, we anticipate that the updated database will facilitate the high-resolution investigation of SNV–ceRNA networks and advance our understanding of the regulatory mechanisms in complex disease ecosystems.

Graphical abstract



Received: August 18, 2025. Revised: September 10, 2025. Accepted: September 17, 2025

© The Author(s) 2025. Published by Oxford University Press.

This is an Open Access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted reuse, distribution, and reproduction in any medium, provided the original work is properly cited.

Introduction

Single nucleotide variants (SNVs) exert diverse functional effects across genome regions, contributing to genetic diversity and individual phenotypic variation [1]. Those occurring in coding regions can alter protein structure, while variants in regulatory regions may disrupt gene regulation by impairing transcription factor–DNA binding or protein–RNA interactions [2, 3]. Notably, SNVs are far more prevalent in non-coding regions, particularly within microRNA (miRNA) binding sites on long non-coding RNAs (lncRNAs) [4]. Such variants can perturb lncRNA–miRNA interactions, thereby disrupting competing endogenous RNA (ceRNA) regulatory networks and influencing disease initiation and progression [5, 6]. For instance, a functional genetic variant, rs67311347 (G > A), increases the expression of ENTPD3-AS1 by forming a ZNF8 binding motif, and this upregulation inhibits renal cell carcinoma via the miR-155/HIF-1 α signaling pathway [7]. The lncRNA LOC146880, harboring the single nucleotide polymorphism (SNP) rs140618127, provides a binding site for miR-539-5p, which inhibits ENO1 phosphorylation and downregulates the phosphatidylinositol-3-kinase (PI3K)/Akt pathway, thereby reducing the risk of non-small cell lung cancer [8]. A genome-wide study of 2130 ovarian cancer patients identified a locus at 3p26.1 associated with overall survival and a likely causal variant that alters enhancer activity by interacting with lncRNA BHLHE40-AS1, influencing oncogenic expression and tumor growth [9]. The identification of such personalized variation events will enhance our understanding of individual disease pathology and further contribute to the advancement of precision medicine.

Our previous work, LnCeVar 1.0, curated SNV–ceRNA events from thousands of samples and cell lines, serving as a vital resource for investigating the functional impact of genome variations on ceRNA network regulation in human diseases [10]. Notably, single-cell RNA sequencing (scRNA-seq) and spatial transcriptomics RNA sequencing (stRNA-seq) technologies enable the characterization of cellular heterogeneity and dynamic states within the tumor microenvironment, facilitating the dissection of intercellular interactions and signaling pathways [11]. The stRNA-seq further captures cellular spatial locations within tissues, enabling precise analysis of cell distribution and interactions in the tumor microenvironment [12]. SNVs in the cellular genome disrupt gene regulatory networks, driving changes in cellular states and behaviors, such as metastasis, invasion, and immune escape [13]. Thus, identifying SNVs at the single-cell and spatial levels and further analyzing their impact on gene regulatory networks is crucial for understanding the determinants of cell fate and revealing dynamic changes.

Existing research has developed methods such as RESA [14], scSNV [15], and SComatic [16], which enable the identification of SNVs at single-cell and spatial resolutions. Meanwhile, large-scale SNV resources at the single-cell and spatial resolutions are rapidly emerging, such as CanCellVar [17] and stSNV [18]. To better understand the impact of SNVs within cells on ceRNA regulatory networks and further explore the driving factors behind different cell fates, we developed LnCeVar 2.0. This updated database will enable us to study complicated SNV–ceRNA networks at a single-cell and spatial resolution, helping us to gain a deeper understanding of disease pathology.

LnCeVar 2.0 is an enhanced resource for investigating SNV–ceRNA regulation at cellular/spatial resolutions. It in-

tegrates high-throughput scRNA-seq, stRNA-seq, and bulk-seq data from over 20 000 human/mouse normal/diseased samples, identifying 5 218 062 functional SNV–ceRNA events across 2 673 603 cells/spots. It also includes 5785 experimentally validated ceRNA interactions/SNV–ceRNA events and 16 937 manually curated cancer biomarker annotations (linked to pathology, diagnosis, and treatment). With a user-friendly web interface for searching, browsing, downloading, and analyzing, as well as tools for gene expression/SNV analysis, 3D cross talk studies, and more. This update will further enhance our understanding of individual disease pathology and drive advancements in precision medicine.

Materials and methods

Data collection and processing

High-throughput scRNA-seq and stRNA-seq datasets were compiled through manual curation of PubMed publications and retrieval from public repositories, including NCBI-GEO [19], 10x Genomics (<https://www.10xgenomics.com/>), and relevant databases [20–25]. Based on GENCODE annotations (version GRCh38/GRCm39), LnCeVar 2.0 systematically identifies human and mouse gene expression profiles across diverse gene types. Data preprocessing and normalization were performed using the R package Seurat (v4.3) [26], with the FindClusters, RunTSNE, and RunUMAP functions utilized for cell clustering and visualization. The R packages Monocle 2 (v2.18.0) [27] and Monocle 3 (v1.2.9) [28] were employed to calculate pseudotime, construct cell developmental trajectories, and determine cell states. The scSNV [15] and SComatic [16] methods were used to identify SNVs at the single-cell and spatial levels. A multivariate multiple regression (MMR) model was applied to examine the effects of SNVs on ceRNA expression [5, 10, 29]. The impact of functional SNV–ceRNA events was estimated using the corresponding coefficients (*coef*-values) for the SNVs, and the significance (*P*-values) of the MMR model was assessed via Pillai's trace statistics. All statistical analyses were conducted using the R programming language (v4.2.3). Further details on data collection and processing can be found in the [Supplementary Methods](#).

Database construction and interface organization

The LnCeVar 2.0 platform was built using MySQL (version 5.5) for data management, with web interfaces developed through Java Server Pages and deployed on the Tomcat web server (v6). Several JavaScript libraries were integrated to enable interactive data visualization and table rendering, including jQuery (v1.11.3), Datatables (v1.10.10), and ECharts (v4.0). The LnCeVar 2.0 database can be accessed at <http://bio-bigdata.hrbmu.edu.cn/LnCeVar> or <http://www.bio-bigdata.net/LnCeVar>. The previous version (LnCeVar 1.0, available at <http://www.bio-bigdata.net/LnCeVar1.0/>) remains accessible for users requiring legacy data or features.

Results

New collection of high-throughput scRNA-seq and stRNA-seq data

To enhance the comprehensiveness and functionality of the LnCeVar 2.0 database, substantial efforts have been devoted

to expanding its data content, specifically enriching the high-throughput scRNA-seq and stRNA-seq datasets (Fig. 1A–G). The database now encompasses gene expression profiles of 812 human and mouse samples, collectively covering 1 773 362 cells and 900 241 spatial spots. Following meticulous data collection and preprocessing procedures (Supplementary Methods), 400 distinct cell types were identified, substantially expanding the range of cell types represented in the database. To broaden the disease spectrum, 102 diseases and disease-related phenotypes with diverse clinical trials and outcomes (such as chemotherapy, immunotherapy, targeted gene therapy, and CRISPR gene knockout) have been incorporated. To conduct comparative analyses of the tumor microenvironment at single-cell and spatial levels, we compiled 244 gene expression profiles from normal human/mouse organs and tissues. The aim was to define the cellular composition of major healthy organs and create a landscape of human/mouse cells. Using GENCODE gene type and transcript annotations (version GRCh38/GRCm39), LnCeVar 2.0 systematically identifies human and mouse gene expression profiles across various gene types, including protein-coding genes, lncRNAs, pseudogenes, and other genes, resulting in the identification of over 20 000 coding and noncoding genes from the scRNA-seq and stRNA-seq data.

To fully leverage the high-resolution capabilities of single-cell and spatial omics technologies for advancing gene expression regulation studies in the tumor microenvironment, we have systematically curated and annotated the collected datasets. Each dataset is characterized by the following features: (i) gene expression profiles at single-cell and spatial-spot resolutions (Fig. 1A and B); (ii) multidimensional coordinates mapping cellular clusters, developmental trajectories, and spatial positions (Fig. 1C and D); (iii) functional state profiling, pathway activity quantification, and cancer hallmark enrichment scoring for individual cells and spatial spots (Fig. 1E); and (iv) histological images that enable visualization of gene expression within the context of spatially resolved tissue architecture and disease pathology (Fig. 1F). A panel of tools, by integrating high-resolution single-cell and spatial transcriptomics data, enhances the interpretability of histopathology and facilitates investigating mechanisms that disrupt gene networks (Fig. 1G). This represents a significant expansion compared to LnCeVar 1.0 in terms of dataset quantity, cell numbers, cell types, gene counts, and single-cell and spatial features (Table 1).

Newly identified functional single-cell and spatially specific SNV–ceRNA events

LnCeVar 2.0 provides a comprehensive profile of genomic SNVs in human and mouse, revealing the co-localization of SNVs in distinct cell clusters along with their unique characteristics and enabling exploration of the profound impact of SNVs on gene expression at the single-cell and spatial levels (Supplementary Fig. S1). Following the collection and processing of SNV data, a total of 1 839 720 and 707 433 SNVs were identified in over 70 human and mouse tissues, respectively. LnCeVar 2.0 compiles ceRNA candidates from existing databases [30–34], providing information on potential single-cell and spatially specific ceRNA interactions. A published method based on probability theory was applied to construct cell-specific networks and identify ceRNAs at single-cell and spatial resolutions [35]. A candidate SNV–ceRNA event

was identified if different genotypes of a variation could alter ceRNA regulations (e.g. gain, loss, or alternation of miRNA binding sites). To further identify functional SNV–ceRNA events at the expression level, the MMR model was employed to examine whether a given SNV regulates the expression of both ceRNA genes [5, 10, 29]. The effect of functional SNV–ceRNA events was estimated using the corresponding coefficients for the SNV (details in Supplementary Methods and Supplementary Fig. S2). Ultimately, a total of 5 218 062 functional SNV–ceRNA events were identified across 2 673 603 single cells and spatial spots. Collectively, these functional single-cell and spatially specific SNV–ceRNA events will facilitate investigations into fine-tuned SNV–ceRNA networks at single-cell and spatial resolutions, aiding to the understanding of regulatory mechanisms underlying complex biological ecosystems.

Curation of validated SNV–ceRNA events, biomarkers, and functional annotations

To provide more detailed and reliable information regarding the mechanisms of SNV and ceRNA regulation, LnCeVar 2.0 has been updated to include an expanded set of experimentally verified ceRNA interactions and SNV–ceRNA events, which were collated from our previous studies [30–34, 36] and manual curation (Supplementary Methods). Specifically, a total of 5785 validated ceRNA interactions and SNV–ceRNA events were confirmed using high-confidence experimental techniques, including polymerase chain reaction, western blotting, luciferase reporter assays, and other methodologies. Additionally, LnCeVar 2.0 conducted manual curation of ceRNA and SNV annotations related to cancer biomarkers associated with disease pathology, diagnosis, and treatment. Following manual curation, a total of 16 937 experimentally supported biomarkers were incorporated into LnCeVar 2.0. Furthermore, LnCeVar 2.0 contains 18 100 and 14 039 functional gene sets for the human and mouse genomes, respectively. These gene sets encompass a wide range of functional annotations, including Gene Ontology [37], canonical pathways [38], cancer cell states [39], and classical cancer hallmarks [40], thereby facilitating comprehensive functional analysis (Supplementary Fig. S3). To infer the regulatory impact of SNV–ceRNA events on clinical characteristics, LnCeVar 2.0 integrated bulk-seq data and clinical profiles from 22 515 tumor samples across various types. These datasets, sourced from TCGA [41] and NCBI-GEO [19], encompass ceRNA expression, genome variation across individuals, clinicopathological features, clinical treatment details, and follow-up data. The incorporation of cancer biomarkers, functional annotations, and clinical features into LnCeVar 2.0 provides a more robust framework for investigating ceRNA regulatory mechanisms through pathological analysis.

Novel data features and enhanced analysis tools

Recent advances in single-cell and spatial transcriptomics have revolutionized the dissection of molecular heterogeneity in disease, enabling unprecedented resolution to explore how genomic variations disturb gene expression, regulatory networks, and cellular phenotypes in the tumor microenvironment. To better leverage such data, LnCeVar 2.0 has been updated with several new features for exploring the regulatory roles of SNV–ceRNA events at the single-cell and spatial levels, thereby facilitating the study of fine-tuned SNV–ceRNA

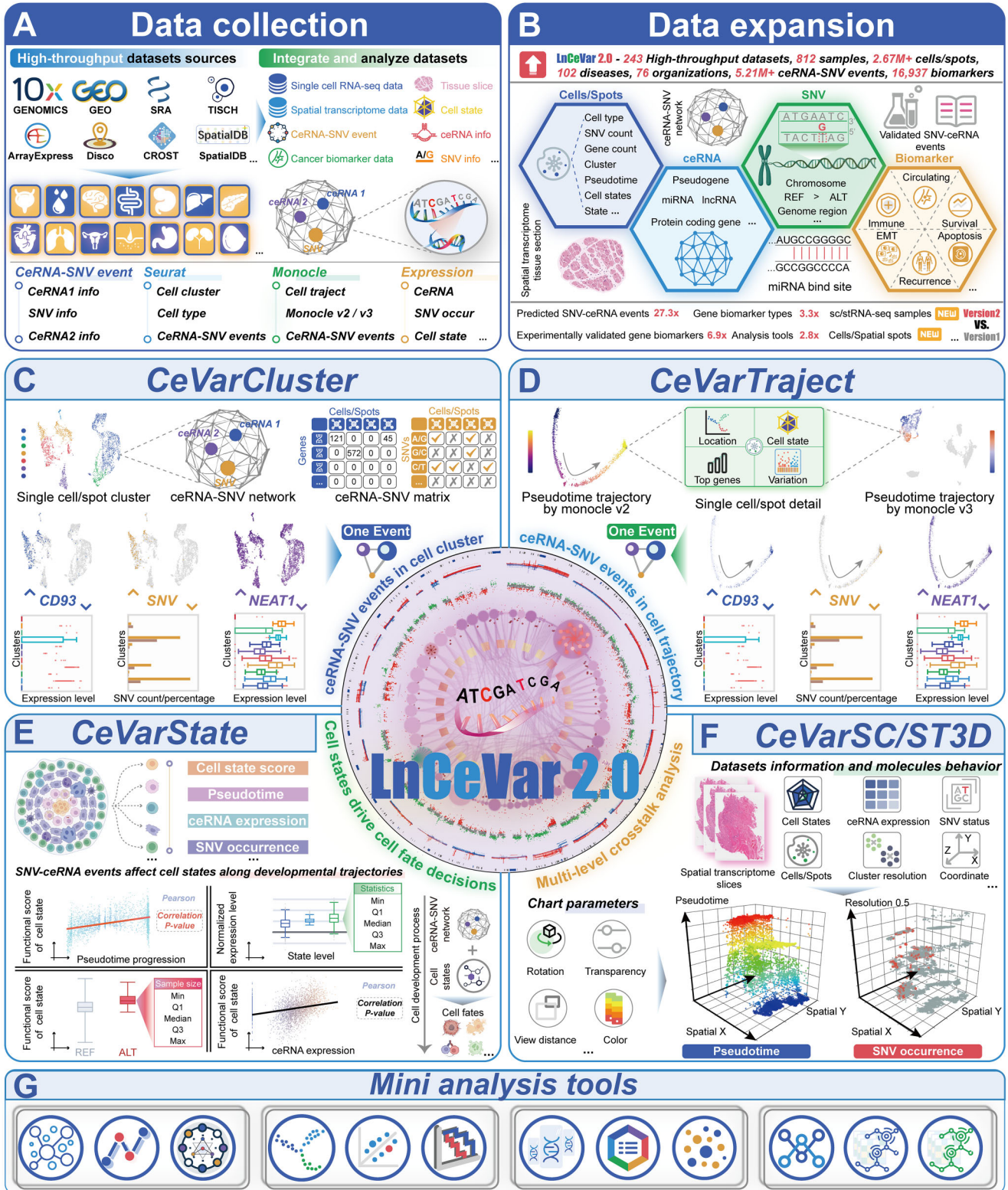


Figure 1. Data expansion and new features of LnCeVar 2.0. **(A, B)** Collection and expansion of high-throughput scRNA-seq, stRNA-seq, and bulk-seq datasets of human and mouse. **(C, D)** The multidimensional mapping of cellular clusters, developmental trajectories, and spatial positions. **(E)** Functional state profiling, pathway activity quantification, and cancer hallmark enrichment scoring for individual cells/spots. **(F)** Visualizing gene expression in single-cell and spatially resolved tissue architecture and pathology.

Table 1. Data expansion and functional enhancements in LnCeVar 2.0

Datasets and features		LnCeVar 1.0	LnCeVar 2.0	Fold increase↑
High-throughput datasets	scRNA-seq samples	–	357	New
	stRNA-seq samples	–	455	New
	Bulk-seq samples	11 432	22 515	1.97↑
	Single cells	–	1 773 362	New
	Cell types	–	400	New
	Spatial cells/spots	–	900 241	New
	Spatial slices	–	478	New
	Diseases/phenotypes	58	102	1.76↑
	Normal tissues/organs	28	80	2.86↑
	Species	Human	Human, mouse	New
Genetic variations, gene annotations, and regulatory events	Variations	191 158	2 547 153	13.32↑
	Coding genes	–	21 639	New
	Noncoding genes	1845	3272	1.77↑
	Predicted SNV–ceRNA events	191 158	5 218 062	27.30↑
	Validated SNV–ceRNA events and ceRNA interactions	54	5785	107.13↑
	Experimentally validated	2444	16 937	6.93↑
	diagnostics and therapy biomarkers	–	–	–
	Category of diagnostics and therapy biomarkers	3	10	3.33↑
	Functional gene sets	7246	32 139	4.44↑
	Mini analysis tools	6	12	2.00↑
Function features and analysis tools	Comprehensive analysis tools	–	5	New
	Cell cluster analysis	–	✓	New
	Cell trajectory analysis	–	✓	New
	Cell type annotation	–	✓	New
	Cell state annotation	–	✓	New
	scRNA-seq data analysis	–	✓	New
	stRNA-seq data analysis	–	✓	New
	Multilevel correlation analysis	–	✓	New
	3D-data visualization	–	✓	New

effects to unravel the regulatory mechanisms underlying complex ecosystems. A range of flexible tools, including 5 comprehensive and 12 mini analysis tools, have been developed to facilitate data retrieval and analysis (Supplementary Figs S4 and S5). For example, the *CeVarCluster* tool examines the patterns of gene expression and SNV mutations in SNV–ceRNA events while analyzing the distribution of cell clusters, types, states, and trajectories. The *CeVarTraject* tool illustrates the dynamic changes in SNV genotype and gene expression variation along cell differentiation trajectories. The *CeVarState* tool evaluates the driver effect of SNV–ceRNA events on cell states along developmental trajectories, revealing the dynamic interactions that determine cell fate. The *CeVarSC3D* and *CeVarST3D* tools enable multilevel cross talk analyses of SNVs, ceRNA networks, and cell states contributing to disease pathology at single-cell and spatial levels, with results presented via an interactive 3D view. Each of the 12 mini tools provides fast, user-friendly analysis including functional, hallmark and cell state annotations, cell clustering, survival and correlation analyses, and SNV–ceRNA network construction.

A usage example of LnCeVar 2.0 database

LnCeVar 2.0 provides a convenient, user-friendly web interface for data searching, browsing, analysis, and downloading (Fig. 2A–E). Its “HOME” page includes search and browse interfaces (Supplementary Figs S6 and S7) with keywords such as different gene types (lncRNAs, mRNAs, pseudogenes), diseases, organs, and SNV genomic regions (Fig. 2A and B). To illustrate the database’s functionality, a case study on *NEAT1*-

related data is presented through interactive panels and tables, enabling users to filter and reorder results via table headers (Fig. 2A). All *NEAT1*-related basic gene information (including gene name, annotated gene IDs from external databases, etc.), as well as the number of associated ceRNAs, SNV–ceRNA events, related disease information, and species information, are displayed in the search results (Fig. 2C and Supplementary Figs S8 and S9). The “Detail” column directs users to a dedicated page containing comprehensive information on disease–SNV–ceRNA associations, showcasing how different SNV genotypes affect ceRNA regulatory status and further impact downstream gene expression across single-cell and spatial omics data from various diseases (Fig. 2D and E). Additionally, information is provided on whether *NEAT1* and its downstream targets have been validated in experiments as cancer biomarkers associated with cancer cell growth, metastasis, disease recurrence, survival, drug resistance, and immune escape (Supplementary Fig. S10).

Moreover, to further characterize SNV–ceRNA events involving cellular heterogeneity and dynamic states within the high-grade serous ovarian carcinoma (HGSOC) microenvironment, thereby facilitating the dissection of intercellular interactions and signaling pathways, a panel of online tools has been developed to study ceRNA-regulated mechanisms at the single-cell and spatial levels (Fig. 3A–D). The *CeVarCluster* tool was used to investigate the comprehensive mapping of *NEAT1*-related SNV–ceRNA events across diverse cell populations (Fig. 3A). Within this tool, cells and spatial spots can be stratified into distinct clusters, enabling visualization of ceRNA expression profiles and SNV genotype

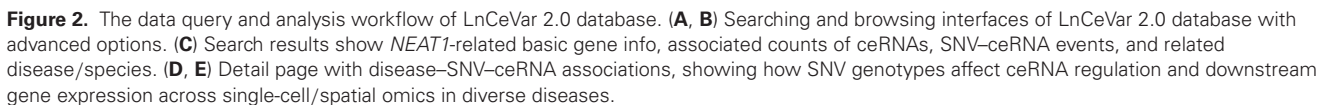


Figure 2. The data query and analysis workflow of LnCeVar 2.0 database. **(A, B)** Searching and browsing interfaces of LnCeVar 2.0 database with advanced options. **(C)** Search results show *NEAT1*-related basic gene info, associated counts of ceRNAs, SNV-ceRNA events, and related disease/species. **(D, E)** Detail page with disease-SNV-ceRNA associations, showing how SNV genotypes affect ceRNA regulation and downstream gene expression across single-cell/spatial omics in diverse diseases.

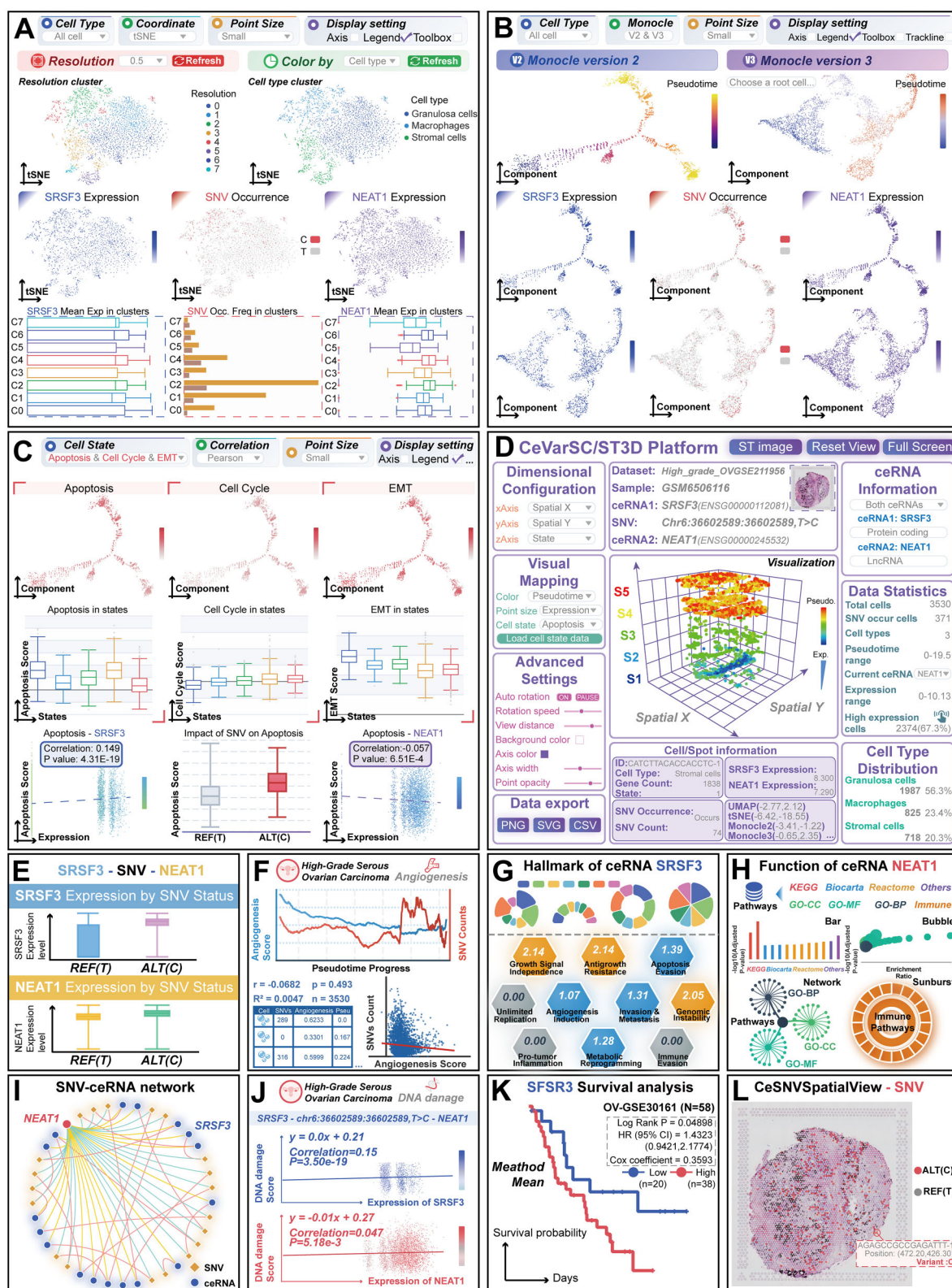


Figure 3. A case study for utilization of analysis tools in LnCeVar 2.0. **(A)** *CeVarCluster* maps *NEAT1*-related SNV-ceRNA events across cell populations, stratifying cells/spatial spots visualization by cell type, state, gene content. **(B)** *CeVarTraject* constructs and visualizes developmental cell trajectories to elucidate dynamics of *NEAT1*-related SNV-ceRNA events and their lineage associations. **(C)** *CeVarState* examines SNV-ceRNA impact on cellular function heterogeneity by assessing cell states (e.g. apoptosis, cell cycle, EMT) and their links to SNV genotypes/ceRNA expression across contexts. **(D)** *CeVarSC3D* and *CeVarST3D* enable multilevel cross talk analyses of SNVs, ceRNA networks, and cell states in disease pathology at single-cell/spatial levels via interactive 3D views. **(E-L)** A panel of mini tools sheds light on the regulatory role of SNVs in *NEAT1*-related ceRNAs, revealing potential therapeutic implications for clinical practice in HGSOc.

occurrences based on cell type, cell state, gene content, and other relevant features (Supplementary Fig. S11). To elucidate the dynamics of *NEAT1*-related SNV–ceRNA events and their associations across different lineages, the *CeVarTraject* tool was used to construct and visualize the developmental cell trajectories (Fig. 3B and Supplementary Fig. S12). To explore the regulatory impact of SNV–ceRNA events on heterogeneity in cellular functions, the *CeVarState* tool was used to evaluate cell states (such as apoptosis, cell cycle, and epithelial mesenchymal transition) and their associations with SNV genotypes and ceRNA expression across diverse functional contexts (Fig. 3C and Supplementary Fig. S13). The *CeVarSC3D* and *CeVarST3D* tools perform multilevel cross talk analyses of SNVs, ceRNA networks, and cell states underlying individual disease pathology at the single-cell and spatial levels, presented in an interactive 3D view (Fig. 3D and Supplementary Fig. S14). These tools enable the examination of the complex interplay between *NEAT1*-associated ceRNA expression and genome variations, which influences individual cell functions and disease pathology.

Several mini tools offer quick and easy-to-use analyses, including functional, hallmark and cell state annotations, cell clustering, survival and correlation analyses, and SNV–ceRNA network construction (Fig. 3E–L). Using these tools, we found that distinct SNV genotypes modulate the ceRNA interaction between *NEAT1* and *SRSF3* in HGSOC cells (Fig. 3E). Furthermore, the cumulative burden of genome mutations is closely associated with HGSOC angiogenesis (Fig. 3F). Previous studies have demonstrated that *NEAT1* knockdown inhibits homologous recombination capacity and increases Olaparib-induced DNA damage in HGSOC, thereby enhancing their sensitivity to Olaparib and conferring a critical therapeutic advantage [42]. Additionally, *SRSF3* overexpression is essential for ovarian cancer cell growth and survival, while its knockdown impairs DNA repair activity, with underlying mechanisms explored in prior research [43]. Using the analysis tools in LnCeVar 2.0, we revealed that the *NEAT1*–*SRSF3* interaction is involved in key biological processes driving cancer initiation and progression, such as genome instability and mutation (Fig. 3G–I). We also observed a positive correlation between *NEAT1*–*SRSF3* expression levels and DNA damage (Fig. 3J) and found that high *SRSF3* expression is associated with poor prognosis in HGSOC (Fig. 3K). These findings, integrating existing knowledge with novel insights derived from LnCeVar 2.0, unraveling the roles of SNV–ceRNA networks in disease pathology, thereby facilitating a deeper understanding of ovarian cancer mechanisms. The spatial map of ceRNA expression and SNV mutations enables visualization and investigation of their impact on disease characteristics through integration with spatial pathological section data (Fig. 3L). In summary, the web-based tools of LnCeVar 2.0 provide efficient means of analyzing SNV–ceRNA regulatory mechanisms, facilitating detailed integration, comparison, and visualization.

Discussion

The development of LnCeVar 2.0 represents a significant advancement in the field of ceRNA and SNV research, addressing the growing demands posed by the expansion of single-cell and spatial transcriptomics datasets. By integrating a vast array of high-throughput scRNA-seq, stRNA-seq, and bulk-seq data across over 20 000 normal and disease samples from hu-

man and mouse, this updated database provides an enhanced resource for investigating SNV and gene expression regulation at cellular and spatial resolutions. The inclusion of over 1.7 million cells, 900 thousand spatial spots, and 400 distinct cell types enables comprehensive analyses of the tumor microenvironment and healthy tissue landscapes, facilitating comparative studies that were previously constrained by data limitations.

Compared to LnCeVar 1.0, LnCeVar 2.0 clarifies the mechanistic links between SNVs, ceRNA dynamics, and cell fate decisions, particularly in the context of disease progression and treatment responses. For instance, it identifies *NEAT1*-related SNV–ceRNA events as potential targets by linking them to cancer hallmarks (e.g. genomic instability) and drug sensitivity (e.g. *NEAT1* enhancing HGSOC's response to Olaparib). It addresses traditional single-gene target limitations via multidimensional data integration, aiding targeted therapy development and overcoming drug resistance. Although LnCeVar 2.0 has identified a large number of SNV–ceRNA events, it is limited in that only a small proportion of these events have been experimentally validated and found to contribute to the development of specific diseases. Many of the remaining events require further analysis and validation. LnCeVar 2.0 will continuously improve research methods and tools for analysing candidate SNV–ceRNA events to address this limitation.

Looking to the future, LnCeVar 2.0 will continue to expand its data coverage, focusing on integrating multimodal datasets to capture the full complexity of genome variation and gene regulation. By bridging genome variations, single-cell and spatial dynamics, and clinical outcomes, LnCeVar 2.0 ultimately aims to accelerate discoveries in disease mechanisms and facilitate the development of personalized therapeutic strategies.

Acknowledgements

The authors gratefully thank the following public databases and platforms for providing the public datasets used in this study: The Cancer Genome Atlas (TCGA), Gene Expression Omnibus (GEO), Molecular Signatures Database (MsigDB), Tumor Immune Single-cell Hub 2 (TISCH2), Deeply Integrated human Single-Cell Omics (DISCO), a database for spatially resolved transcriptomes (SpatialDB), a comprehensive repository of spatial transcriptomics (CROST), NCBI Sequence Read Archive (NCBI-SRA), ArrayExpress Database (ArrayExpress), and 10x Genomics.

Author contributions: Qiuyan Guo (Data curation [equal], Methodology [equal], Validation [equal], Writing—original draft [equal], Writing—review & editing [equal]), Qian Liu (Data curation [equal], Methodology [equal], Validation [equal], Writing—original draft [equal], Writing—review & editing [equal]), Mengyu Xin (Data curation [equal], Methodology [equal], Validation [equal], Writing—original draft [equal], Writing—review & editing [equal]), Houxing Li (Data curation [equal], Methodology [equal], Validation [equal], Writing—original draft [equal], Writing—review & editing [equal]), Jiatong Li (Methodology [supporting]), Yifan Dai (Methodology [supporting]), Rui Sun (Validation [supporting]), Yujie Zhang (Validation [supporting]), Yuting He (Data curation [supporting]), Borui Xu (Resources [equal]), Xinxin Shan (Resources [equal]), Yue Gao (Software [lead]), Shangwei Ning (Methodology [equal], Writing—original draft [equal], Writing—review & editing [equal]), Hui Zhi (Methodology [equal], Writing—original

draft [equal], Writing—review & editing [equal]), Peng Wang (Conceptualization [lead], Supervision [lead], Funding acquisition [lead], Methodology [equal], Project administration [lead], Writing—original draft [equal], Writing—review & editing [equal]).

Supplementary data

Supplementary data is available at NAR online.

Conflict of interest

None declared.

Funding

The National Natural Science Foundation of China (82573599, 82573333, and 82373408); the Heilongjiang Provincial Natural Science Foundation (YQ2024H013); the Distinguished Young Science Foundation of the First Affiliated Hospital of Harbin Medical University (2024JQ20). Funding to pay the Open Access publication charges for this article was provided by National Natural Science Foundation of China (82373408).

Data availability

All the data can be downloaded from <http://bio-bigdata.hrbmu.edu.cn/LnCeVar> or <http://www.bio-bigdata.net/LnCeVar>.

References

- Human Genome Structural Variation WorkingGroup, Eichler EE, Nickerson DA *et al.*. Completing the map of human genetic variation. *Nature* 2007;447:161–5. <https://www.nature.com/articles/447161a>
- Kirk BW, Feinsod M, Favis R *et al.* Single nucleotide polymorphism seeking long term association with complex disease. *Nucleic Acids Res* 2002;30:3295–311. <https://doi.org/10.1093/nar/gkf466>
- Zhu Y, Spitz MR, Lei L *et al.* A single nucleotide polymorphism in the matrix metalloproteinase-1 promoter enhances lung cancer susceptibility. *Cancer Res* 2001;61:7825–9.
- Alexander RP, Fang G, Rozowsky J *et al.* Annotating non-coding regions of the genome. *Nat Rev Genet* 2010;11:559–71. <https://doi.org/10.1038/nrg2814>
- Li MJ, Zhang J, Liang Q *et al.* Exploring genetic associations with ceRNA regulation in the human genome. *Nucleic Acids Res* 2017;45:5653–65. <https://doi.org/10.1093/nar/gkx331>
- Wang P, Ning S, Zhang Y *et al.* Identification of lncRNA-associated competing triplets reveals global patterns and prognostic markers for cancer. *Nucleic Acids Res* 2015;43:3478–89. <https://doi.org/10.1093/nar/gkv233>
- Wang J, Zou Y, Du B *et al.* SNP-mediated lncRNA–ENTPD3–AS1 upregulation suppresses renal cell carcinoma via miR-155/HIF-1 α signaling. *Cell Death Dis* 2021;12:672. <https://doi.org/10.1038/s41419-021-03958-4>
- Feng T, Feng N, Zhu T *et al.* A SNP-mediated lncRNA (LOC146880) and microRNA (miR-539-5p) interaction and its potential impact on the NSCLC risk. *J Exp Clin Cancer Res* 2020;39:157. <https://doi.org/10.1186/s13046-020-01652-5>
- Dai H, Chu X, Liang Q *et al.* Genome-wide association and functional interrogation identified a variant at 3p26.1 modulating ovarian cancer survival among Chinese women. *Cell Discov* 2021;7:121. <https://doi.org/10.1038/s41421-021-00342-6>
- Wang P, Li X, Gao Y *et al.* LnCeVar: a comprehensive database of genomic variations that disturb ceRNA network regulation. *Nucleic Acids Res* 2020;48:D111–7. <https://doi.org/10.1093/nar/gkz887>
- Vandereyken K, Sifrim A, Thienpont B *et al.* Methods and applications for single-cell and spatial multi-omics. *Nat Rev Genet* 2023;24:494–515. <https://doi.org/10.1038/s41576-023-00580-2>
- Tian L, Chen F, Macosko EZ. The expanding vistas of spatial transcriptomics. *Nat Biotechnol* 2023;41:773–82. <https://doi.org/10.1038/s41587-022-01448-2>
- Zhang K, Deng R, Gao H *et al.* Lighting up single-nucleotide variation *in situ* in single cells and tissues. *Chem Soc Rev* 2020;49:1932–54. <https://doi.org/10.1039/C9CS00438F>
- Zhang T, Jia H, Song T *et al.* De novo identification of expressed cancer somatic mutations from single-cell RNA sequencing data. *Genome Med* 2023;15:115. <https://doi.org/10.1186/s13073-023-01269-1>
- Wilson GW, Derouet M, Darling GE *et al.* scSNV: accurate dscRNA-seq SNV co-expression analysis using duplicate tag collapsing. *Genome Biol* 2021;22:144. <https://doi.org/10.1186/s13059-021-02364-5>
- Muyas F, Sauer CM, Valle-Inclan JE *et al.* De novo detection of somatic mutations in high-throughput single-cell profiling data sets. *Nat Biotechnol* 2024;42:758–67. <https://doi.org/10.1038/s41587-023-01863-z>
- Yang C, Liu Y, Lv C *et al.* CanCellVar: a database for single-cell variants map in human cancer. *Am J Hum Genet* 2024;111:1420–30. <https://doi.org/10.1016/j.ajhg.2024.05.014>
- Yang C, Liu Y, Wang X *et al.* stSNV: a comprehensive resource of SNVs in spatial transcriptome. *Nucleic Acids Res* 2025;53:D1224–34. <https://doi.org/10.1093/nar/gkae945>
- Barrett T, Wilhite SE, Ledoux P *et al.* NCBI GEO: archive for functional genomics data sets—update. *Nucleic Acids Res* 2013;41:D991–5. <https://doi.org/10.1093/nar/gks1193>
- Sun D, Wang J, Han Y *et al.* TISCH: a comprehensive web resource enabling interactive single-cell transcriptome visualization of tumor microenvironment. *Nucleic Acids Res* 2021;49:D1420–30. <https://doi.org/10.1093/nar/gkaa1020>
- Li M, Zhang X, Ang KS *et al.* DISCO: a database of deeply integrated human single-cell omics data. *Nucleic Acids Res* 2022;50:D596–602. <https://doi.org/10.1093/nar/gkab1020>
- Fan Z, Chen R, Chen X. SpatialDB: a database for spatially resolved transcriptomes. *Nucleic Acids Res* 2020;48:D233–7. <https://doi.org/10.1093/nar/gkz934>
- Wang G, Wu S, Xiong Z *et al.* CROST: a comprehensive repository of spatial transcriptomics. *Nucleic Acids Res* 2024;52:D882–90. <https://doi.org/10.1093/nar/gkad782>
- Sarkans U, Fullgrabe A, Ali A *et al.* From ArrayExpress to BioStudies. *Nucleic Acids Res* 2021;49:D1502–6. <https://doi.org/10.1093/nar/gkaa1062>
- Qi Y, Xu R, Song C *et al.* A comprehensive database of exosome molecular biomarkers and disease–gene associations. *Sci Data* 2024;11:210. <https://doi.org/10.1038/s41597-024-03015-7>
- Satija R, Farrell JA, Gennert D *et al.* Spatial reconstruction of single-cell gene expression data. *Nat Biotechnol* 2015;33:495–502. <https://doi.org/10.1038/nbt.3192>
- Qiu X, Mao Q, Tang Y *et al.* Reversed graph embedding resolves complex single-cell trajectories. *Nat Methods* 2017;14:979–82. <https://doi.org/10.1038/nmeth.4402>
- Cao J, Spielmann M, Qiu X *et al.* The single-cell transcriptional landscape of mammalian organogenesis. *Nature* 2019;566:496–502. <https://doi.org/10.1038/s41586-019-0969-x>
- Xu R, He D, Sun R *et al.* CNV-mediated dysregulation of the ceRNA network mechanism revealed heterogeneity in diffuse and intestinal gastric cancers. *J Transl Med* 2025;23:308. <https://doi.org/10.1186/s12967-025-06222-x>
- Guo Q, Liu Q, He D *et al.* LnCeCell 2.0: an updated resource for lncRNA-associated ceRNA networks and web tools based on single-cell and spatial transcriptomics sequencing data. *Nucleic*

- Acids Res* 2024;53:D107–15.
<https://doi.org/10.1093/nar/gkae947>
31. Wang P, Guo Q, Hao Y *et al.* LnCeCell: a comprehensive database of predicted lncRNA-associated ceRNA networks at single-cell resolution. *Nucleic Acids Res* 2021;49:D125–33.
<https://doi.org/10.1093/nar/gkaa1017>
 32. Wang P, Li X, Gao Y *et al.* LncACTdb 2.0: an updated database of experimentally supported ceRNA interactions curated from low- and high-throughput experiments. *Nucleic Acids Res* 2019;47:D121–7. <https://doi.org/10.1093/nar/gky1144>
 33. Wang P, Guo Q, Qi Y *et al.* LncACTdb 3.0: an updated database of experimentally supported ceRNA interactions and personalized networks contributing to precision medicine. *Nucleic Acids Res* 2022;50:D183–9. <https://doi.org/10.1093/nar/gkab1092>
 34. Li JH, Liu S, Zhou H *et al.* starBase v2.0: decoding miRNA–ceRNA, miRNA–ncRNA and protein–RNA interaction networks from large-scale CLIP-Seq data. *Nucleic Acids Res* 2014;42:D92–7. <https://doi.org/10.1093/nar/gkt1248>
 35. Dai H, Li L, Zeng T *et al.* Cell-specific network constructed by single-cell RNA sequencing data. *Nucleic Acids Res* 2019;47:e62. <https://doi.org/10.1093/nar/gkz172>
 36. Gao Y, Shang S, Guo S *et al.* Lnc2Cancer 3.0: an updated resource for experimentally supported lncRNA/circRNA cancer associations and web tools based on RNA-seq and scRNA-seq data. *Nucleic Acids Res* 2021;49:D1251–8.
<https://doi.org/10.1093/nar/gkaa1006>
 37. Gene Ontology Consortium. The Gene Ontology resource: enriching a GOLD mine. *Nucleic Acids Res* 2021;49:D325–34.
<https://doi.org/10.1093/nar/gkaa1113>
 38. Castanza AS, Recla JM, Eby D *et al.* Extending support for mouse data in the Molecular Signatures Database (MSigDB). *Nat Methods* 2023;20:1619–20.
<https://doi.org/10.1038/s41592-023-02014-7>
 39. Yuan H, Yan M, Zhang G *et al.* CancerSEA: a cancer single-cell state atlas. *Nucleic Acids Res* 2019;47:D900–8.
<https://doi.org/10.1093/nar/gky939>
 40. Hanahan D, Weinberg RA. Hallmarks of cancer: the next generation. *Cell* 2011;144:646–74.
<https://doi.org/10.1016/j.cell.2011.02.013>
 41. Hutter C, Zenklusen JC. The Cancer Genome Atlas: Creating Lasting Value beyond Its Data. *Cell* 2018;173:283–5.
<https://doi.org/10.1016/j.cell.2018.03.042>
 42. Liu Y, Liu G. Targeting NEAT1 affects the sensitivity to PARPi in serous ovarian cancer by regulating the homologous recombination repair pathway. *J Cancer* 2024;15:1397–413.
<https://doi.org/10.7150/jca.91896>
 43. He X, Zhang P. Serine/arginine-rich splicing factor 3 (SRSF3) regulates homologous recombination-mediated DNA repair. *Mol Cancer* 2015;14:158. <https://doi.org/10.1186/s12943-015-0422-1>