Nucleic Acids Research, 2026, **54**, D958–D970
https://doi.org/10.1093/nar/gkaf1092
Advance access publication date: 17 November 2025
Database issue

scBrainScope: cross-species multidimensional brain atlas

Siying Qin[1,†], Yuexi Yang[1,†], Haoyu Wang[1,†], Mengrou Li[1,†], Yushan Deng[1], Yangfeng Chen[1], Changxiao Li[1], Lingmin Peng[1], Linlin Xu[1], Xiwen Qiu[1,2], Yuerwei Guan[1], Shikai Wang[1], Yuanqing He[1], Xiangdong Wang[3,*], Yuting Ma[4,*], Li Huang[1,*], Dongsheng Chen[1,*]

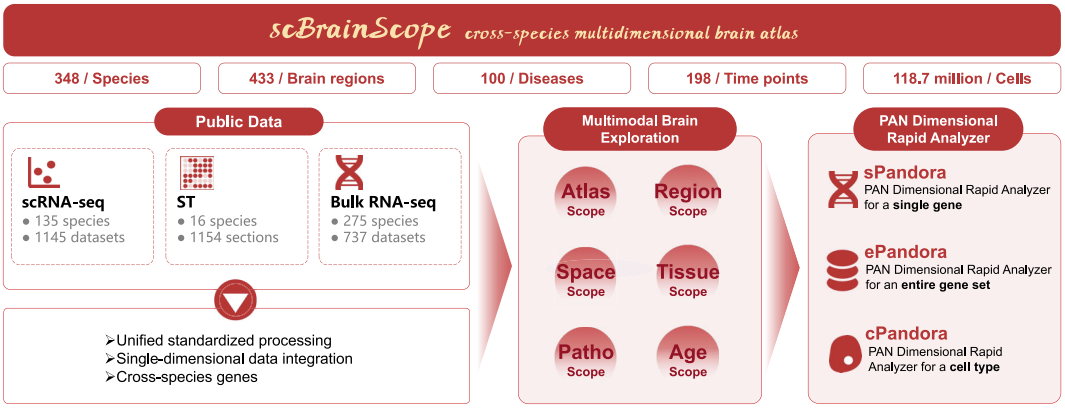
[1]State Key Laboratory of Common Mechanism Research for Major Diseases, Suzhou Institute of Systems Medicine, Chinese Academy of Medical Sciences & Peking Union Medical College, Suzhou 215123, Jiangsu, China
[2]Department of Obstetrics, The Affiliated Suzhou Hospital of Nanjing Medical University, Suzhou Municipal Hospital, Gusu School of Nanjing Medical University, Suzhou 215000, Jiangsu, China
[3]Zhongshan Hospital, Department of Pulmonary and Critical Care Medicine, Institute for Clinical Science, Shanghai Institute of Clinical Bioinformatics, Shanghai 200032, China
[4]National Key Laboratory of Immunity and Inflammation, Suzhou Institute of Systems Medicine, Chinese Academy of Medical Sciences & Peking Union Medical College, Suzhou 215123, Jiangsu, China

*To whom correspondence should be addressed. Email: xdwang@fuccb.com
Correspondence may also be addressed to Yuting Ma. Email: yuting_ma1984@163.com
Correspondence may also be addressed to Li Huang. Email: hl@ism.cams.cn
Correspondence may also be addressed to Dongsheng Chen. Email: cds@ism.pumc.edu.cn
†The first four authors should be regarded as Joint First Authors.

Abstract

As one of the most evolutionarily complex and functionally diverse organs, the brain is characterized by its intricate structure, developmental alterations, and cellular heterogeneity. Although extensive brain atlases exist, there remains an urgent need for an integration platform that brings together gene expression profiles and analysis tools to explore features across species, brain regions, developmental stages, and pathological conditions. Thus, we developed scBrainScope, a detailed and interactive transcriptomic atlas of the brain. Our platform brings together single cell sequencing data from 135 species, covering 433 brain regions, 198 developmental stages, and 100 neurological diseases. In addition, we compiled 737 bulk RNA sequencing datasets from 275 species, along with 1154 spatial datasets of brain, spinal cord, and embryonic tissues. scBrainScope comprises six atlas modules (AtlasScope, RegionScope, TissueScope, SpaceScope, PathoScope, and AgeScope) and three analytical modules (sPandora, ePandora, and cPandora). Together, these tools enable researchers to investigate cell identity, gene programs, and spatial organization at multiple scales and dimensions. scBrainScope is freely available at http://8.142.154.29/scBrainScope or http://www.brainscopes.org, offering an interactive, data-rich resource for neuroscience, evolutionary biology, and translational medicine.

Graphical abstract



Introduction

The vertebrate brain is molecularly, cellularly, and anatomically complex, supporting various functions and behaviors. To elucidate this complexity, comprehensive examination of gene expression, cell types, and development stages across species is essential [1]. In recent years, high-throughput transcriptomic

Received: August 6, 2025. **Revised:** September 18, 2025. **Accepted:** September 23, 2025

technologies such as bulk RNA sequencing (bulk RNA-seq), single-cell RNA sequencing (scRNA-seq), and spatial transcriptomics (ST) have enabled studies of the brain at unprecedented resolution [2–4]. Meanwhile, publicly available brain datasets have grown substantially. However, this rapid expansion of multispecies, multiplatform data presents increasingly challenges for integration, visualization, and analysis.

Several databases have been developed for brain-related single-cell and ST data. MAPbrain provides a multi-omics atlas of the primate brain, covering 21 million brain cells from 45 brain regions across 164 development time points [5]. STomicsDB offers ST data for 17 species [6]. scREAD focuses on Alzheimer's disease-related single-cell sequencing datasets comprising 713 640 cells with an analysis framework for sc/snRNA-seq and ST data [7]. SCAN [8] and TEDD [9] integrate multi-omics data across diverse brain regions and provide tools for interactive visualization and dataset-level exploration.

While these pioneering works have significantly advanced our understanding of brain biology, they remain fragmented in their integration of single-cell, spatial, and bulk transcriptomic datasets at scale and across species, limiting efforts to unravel the evolutionary complexities of the central nervous system. To address this gap, we developed scBrainScope, a comprehensive, integrative platform for multidimensional analysis of brain transcriptomic data. scBrainScope brings together 118.7 million single-cell transcriptomes from 135 species, covering 433 brain regions, 198 developmental stages, 100 brain disorders, and 38 different techniques. It further incorporates 737 bulk RNA-seq datasets from 275 species and 1154 ST sections of brain, spinal cord, and embryonic tissues.

Beyond its multimodal atlas modules, scBrainScope offers a unified analytical toolkit that enables cross-dimensional analysis of genes, gene sets, and cell types across all integrated datasets. This toolkit comprises three core modules, sPandora, ePandora, and cPandora, that together systematically compare transcriptomic patterns across species, brain regions, developmental stages, disease states, and spatial contexts. By integrating data from various methods and a wide range of species (including mammals, birds, fishes, reptiles, amphibians, and invertebrates), scBrainScope facilitates large-scale exploration of the brain's molecular architecture across evolutionary and biological contexts. These analytical modules provide researchers with a powerful and flexible toolkit for studying evolutionary patterns, developmental pathways, disease-related changes, and the spatial organization of brain cell types and gene activity.

Materials and methods

Data collection, library construction, and curation

Data were sourced primarily from publicly available databases like the Allen Brain Atlas [10, 11], Single Cell Portal [12], NCBI/GEO [13], EMBL-EBI [14], and from articles indexed in PubMed. For scRNA-seq, raw FASTQ files were processed with Cell Ranger [15], and preprocessed expression matrices (RDS, H5AD, or Loom formats) were standardized using SCANPY (v1.11.1) [16] and Seurat (v4.3.0) [17]. Metadata were manually curated and harmonized across datasets to include species, tissue origin, brain region, developmental stage, disease condition, sequencing strategy, and available cell-type annotations. For bulk RNA-seq data, species lacking genome references, *de novo* transcriptome assemblies were generated with Trinity [18], followed by coding sequence prediction using TransDecoder [19]. ST data were curated from 1154 tissue sections across multiple species and technologies. Stereo-seq data sets [20] were processed using Stereopy [21], while data from other platforms were converted to H5AD format and spatial coordinates assigned accordingly.

Quality control

Gene and cell-level quality control for scRNA-seq and spatial data followed uniform criteria, including filtering cells with <200 or >6000 detected genes and removing doublets via Scrublet [22]. Bulk RNA-seq data were processed using fastp [23] for quality control.

Orthologous gene identification

For ortholog mapping, BioMart [24] annotations were used when available. Otherwise, genome annotation files, and proteome sequence files were downloaded from the NCBI RefSeq database [25] or other sources. Orthologs were inferred using OrthoFinder [26] or reciprocal best hits via DIAMOND [27]. All genes were mapped to human orthologs to enable cross-species comparisons.

Reference-based cell type annotation

One high-quality single-cell dataset was curated from published literature [28] and manually consolidated into eight major brain cell types: Neurons (Neu), Astrocytes (Ast), Oligodendrocytes (Olg), Oligodendrocyte Precursor Cells (OPC), Microglia (Mic), Smooth Muscle Cells (SMC), Endothelial cells (End), and Pericytes (PC). This curated dataset was then directly used as the training set for CellTypist [29] to predict cell types.

Atlas-level integration of datasets

Single-cell transcriptomics data were grouped along four biological dimensions: species, brain regions, developmental stages, and disease states. Each biological sample (assigned a unique BSID) was downsampled to 1000 cells per dimension. Integrated embeddings were computed using the Harmony algorithm [30] via sc.pp.harmony_integrate, using BSID as the batch variable.

Four-dimensional single-cell transcriptomics modules

scBrainScope integrates four single-cell transcriptomic modules: AtlasScope, RegionScope, PathoScope, and AgeScope. These collectively encompass over 118 million cells from diverse species and conditions. All datasets underwent standardized processing and format harmonization. Users can interrogate single-gene (sPandora) and gene-set (ePandora) expression under different contexts, examine altered cell-type proportions through cPandora, and access differentially expressed gene (DEG) lists in the GeneLists module for downstream enrichment and functional analyses.

Across AtlasScope, RegionScope, PathoScope, and AgeScope, we implemented eight shared analytical functions on the single-cell dataset pages, using a standardized preprocessing pipeline that includes quality control, down-

sampling, dimensionality reduction, and clustering. Gene-set scores for 81 curated sets across 25 categories were computed and visualized in Uniform Manifold Approximation and Projection (UMAP) and violin plots, with a subset of cells displayed interactively via Seurat objects deployed in ShinyCell [31]. HomoGene links a human gene to its ortholog in a target species by querying a precomputed homolog table, and exporting it for downstream analysis. Transcriptional regulation was examined by computing gene-set scores across 100 sampled cells per dataset and correlating them with transcription factors from JASPAR [32, 33], using Pearson correlation with False Discovery Rate (FDR) correction. Dataset quality was evaluated, and differential expression analyses were performed across species, brain regions, developmental stages, and disease conditions. Gene lists were generated using statistical filters (adjusted $P < .01$, logFC > 1), and top markers were visualized using dot plots, heatmaps, stacked violins, and UMAPs. The top five DEGs per cell type were examined across seven representative species using sc.pl.dotplot. Enrichment analyses of DEGs were conducted with the ToppGene API [34, 35]. Motif analysis was performed using HOMER [36], and protein-protein interaction (PPI) networks were generated with the rbioapi R package [37]. Processed datasets, including raw and normalized count matrices and full metadata, were provided in H5AD format for download.

TissueScope module

TissueScope provides a bulk RNA-seq brain atlas spanning 275 vertebrate species, enabling broad evolutionary and tissue-specific transcriptomic analyses. Raw count matrices were filtered, imputed, and normalized using DESeq2 [38]. Transcripts Per Million (TPM) values were derived using gene lengths estimated by Kallisto [39]. For organisms lacking reference genomes, *de novo* transcript assembly was performed using Trinity, followed by redundancy reduction with cd-hit-est and coding sequence prediction via TransDecoder. The platform supports sPandora and ePandora for gene-level and gene-set profiling, GeneLists for identifying class-specific markers, and TimeTree [40] for visualizing phylogenetic relationships.

SpaceScope module

SpaceScope enables ST analysis using standard clustering and gene-scoring approaches. A total of 81 gene sets were scored and visualized via spatial UMAP projections. Cluster-specific marker genes were identified using sc.tl.rank_genes_groups, and density estimation was applied to annotated chips using sc.tl.embedding_density. SpatialDE [41] was employed to detect spatially variable genes. For RGB visualization, gene expression intensities were mapped to red, green, and blue channels based on spot-level data. The ggplot2 [42] in R was utilized for visualization. Each spot was plotted as a point on a scatterplot.

Three pan-dimensional analysis modules

sPandora: gene expression at the single-gene level

sPandora was designed to support cross-dimensional exploration of single-gene expression patterns, enabling users to identify conserved or context-specific features across species, brain regions, developmental stages, and disease conditions. This fine-grained resolution is particularly useful for validating biomarkers, investigating regulatory genes, and generating hypotheses about gene function. It compiles expression profiles of individual genes from AtlasScope, RegionScope, PathoScope, and AgeScope. Results are stored in comma-delimited files and displayed as interactive heatmaps. TissueScope provides bulk profiles, and SpaceScope visualizes ST measurements.

ePandora: gene expression at the entire gene-set level

Single-gene expression can be noisy and variable, especially in single-cell data, making it difficult to extract robust biological signals. In contrast, gene set scoring aggregates the expression of biologically related genes, reducing noise and enhancing interpretability. ePandora computes gene set scores using SCANPY's sc.tl.score_genes function across AtlasScope, RegionScope, PathoScope, and AgeScope. Scores are calculated for each species, region, disease state, age group, and cell type, saved as CSV files, and visualized as heatmaps. By applying standardized scoring methods, it improves interpretability and highlights coordinated biological programs. In SpaceScope, the same function is applied to ST arrays for *in situ* visualization. In TissueScope, gene set expression is estimated by calculating the mean TPM value of all genes within each set.

cPandora: gene expression at the entire cell-type level

cPandora facilitates comparison of major brain cell type proportions across species and conditions, helping users detect shifts in cellular composition and understand underlying biological dynamics. Cell type proportion is a key indicator of biological variation across species, brain regions, developmental stages, and disease conditions, which can be highly variable and context dependent. Changes in the abundance of major cell types (e.g. neurons, microglia, oligodendrocytes) often reflect fundamental biological processes such as neurodevelopment, aging, inflammation, or degeneration. cPandora quantifies the proportions of six brain cell types across species, brain regions, disease states, and development stage in single-cell datasets (AtlasScope, RegionScope, PathoScope, AgeScope). Results are shown as bar charts of relative abundances. In SpaceScope, cell-type signals are inferred by scoring spots against predefined cell-type signatures and displayed graphically.

Results

Overview of scBrainScope database

Currently, scBrainScope comprises 118.7 million single-cell transcriptomes from 135 species, spanning 433 brain regions, 100 brain disorders, and 198 developmental stages. The platform also integrates 737 bulk RNA-seq datasets from 275 species and 1154 spatially resolved transcriptomic sections from brain, spinal cord, and embryonic tissues.

The database is organized into nine functional modules: AtlasScope, RegionScope, TissueScope, SpaceScope, PathoScope, AgeScope, sPandora, ePandora, and cPandora. The first six provide atlas-based views of species, brain regions, tissues, spatial contexts, disease states, and developmental stages. The latter three offer analytical tools: PAN Dimensional Rapid Analyzer for a single gene (sPandora), for a gene set (ePandora), and for a cell type (cPandora), enabling multidimensional exploration of brain function The "Home" page gives

Database	scRNA-Seq					ST		RNA-Seq		Analysis pipeline		
	Cells	Species	Brain regions	Diseases	Time points	Datasets	Species	Datasets	Species	Pan-dimensional celltype analysis	Pan-dimensional geneset analysis	Pan-dimensional gene analysis
scBrainScope	118.7 million	135	433	100	198	1154	16	737	275	Yes	Yes	Yes
MAPbrain	21 million	6	45	NA	164	NA	6	0	0	No	No	No
Toti	1176	2	NA	NA	NA	NA	NA	NA	NA	No	No	No
SCAN	10.6 million	81	NA	100	7	NA	NA	0	0	No	No	No
STOmicsDB	NA	0	NA	NA	NA	218	17	0	0	No	No	No
SpatialRef	NA	0	NA	NA	NA	552	3	0	0	No	No	No
ssREAD	7.3 million	2	10	1	NA	381	2	0	0	No	No	No
SODB	NA	NA	6	NA	NA	1361	12	0	0	No	No	No
Aquila	NA	0	NA	30	NA	107	5	0	0	No	No	No
STellaris	NA	2	1	1	NA	101	2	0	0	No	No	No
SPASCER	NA	4	1	NA	NA	NA	4	0	0	No	No	No
STAB	144047	1	20	NA	11	0	0	0	0	No	No	No
SC2disease	946481	1	NA	25	NA	0	0	0	0	No	No	No
SpatialDB	NA	0	NA	NA	NA	305	5	0	0	No	No	No
scREAD	713640	2	10	1	NA	0	0	0	0	No	No	No

Figure 1. Comprehensive comparison between scBrainScope and existing brain-related transcriptomic databases. scBrainScope brings together 118.7 million single-cell transcriptomes (scRNA-seq) from 135 species, covering 433 brain regions, 100 brain disorders, and 198 developmental stages. It also includes 1154 ST sections from 16 species and 737 bulk RNA-seq datasets from 275 species. scBrainScope's pan-dimensional analytical tools feature specialized modules for cross-species, cross-region, cross-development, and cross-disease analysis at the single-gene (sPandora), gene-set (ePandora), and cell-type (cPandora) levels. Compared to other databases like MAPbrain [5], Toti [43], SCAN [8], STOmicsDB [6], SpatialRef [44], ssREAD [45], SODB [46], Aquila [47], STellaris [48], SPASCER [49], STAB [50], SC2disease [51], SpatialDB [52], and scREAD [7] scBrainScope stands out with broader species coverage, greater data dimensionality, and fully integrated multimodal datasets and analysis tools. "NA" indicates missing data for certain metrics in the listed databases.

an overview and easy access to each module, while the "Help" page offers detailed guidance for navigation and use.

Compared with existing brain-related databases, scBrainScope exhibits superior data coverage across multiple dimensions (Fig. 1). scBrainScope addresses cross-species divergence, cellular heterogeneity, complex brain anatomy, and dynamic spatiotemporal signaling under both healthy and disease situations. It currently covers the largest collection of multi-omics brain datasets and provides an integrated analysis of cell-proportion dynamics, gene-sets score, and gene-expression patterns from the perspectives of species, anatomy, development, and pathology, thereby offering an unprecedented opportunity to investigate brain evolution, development, and disease.

Single-cell transcriptomics modules (atlasscope, regionscope, pathoscope, and agescope)

scBrainScope integrates large-scale scRNA-seq datasets to facilitate high-resolution molecular profiling across diverse biological contexts. Its four single-cell atlas modules (AtlasScope, RegionScope, PathoScope, AgeScope) (Fig. 2A) allow users to explore gene expression, cell-type composition, and dynamic changes across species, brain regions, developmental stages, and disease conditions. Each module offers various analysis options for single-gene queries (sPandora), gene-set assessments (ePandora), and cell-type comparisons (cPandora), as well as access to curated lists of DEGs (GeneLists).

Users can easily explore brain single-cell sequencing data for species, region, disease, or development stage of interest at Single-dataset exploration on every single-cell transcriptomics module by using the filter panel or clicking images (Fig. 2B Step1 and Step2). Once selected, a result table will appear providing essential dataset information (Fig. 2B Step3). The AtlasScope platform exemplifies this functionality. When a species is selected (e.g. zebrafish), the system automatically generates a tabulated overview of available scRNA-seq datasets for that organism. This curated table presents critical dataset metadata, including species name, sample origin, developmental stage, disease status, sequencing technology, and publication information, thereby facilitating efficient dataset selection for downstream analyses. To ensure rigorous data traceability, scBrainScope employs the following standardized metadata fields:

- BSID, represents the unique identifier of the dataset within our database;
- CommonName, the common name of the species from which the data was derived;
- Region, the specific brain region where the data was collected;
- Timing, the developmental stage at the time of data collection;
- Diseases, the healthy condition of the sample (e.g. normal or disease-specific);
- Tech, the sequencing technology used for data generation;
- Journal, the source journal in which the dataset was published;
- PMID: the PubMed ID corresponding to the publication associated with the dataset;
- Year, the publication year of the dataset's reference article.

Clicking "Detail" displays basic dataset information and a download link for the single-cell expression matrix, while "ViewData" opens the scRNA-Seq DataSet page for deeper exploration.

The scRNA-seq DataSet page serves as a comprehensive platform for examining scRNA-seq data (Supplementary Fig. S1). It includes eight analysis parts:

1. DataInfo&Gene: Gene-level expression atlas and metadata overview;

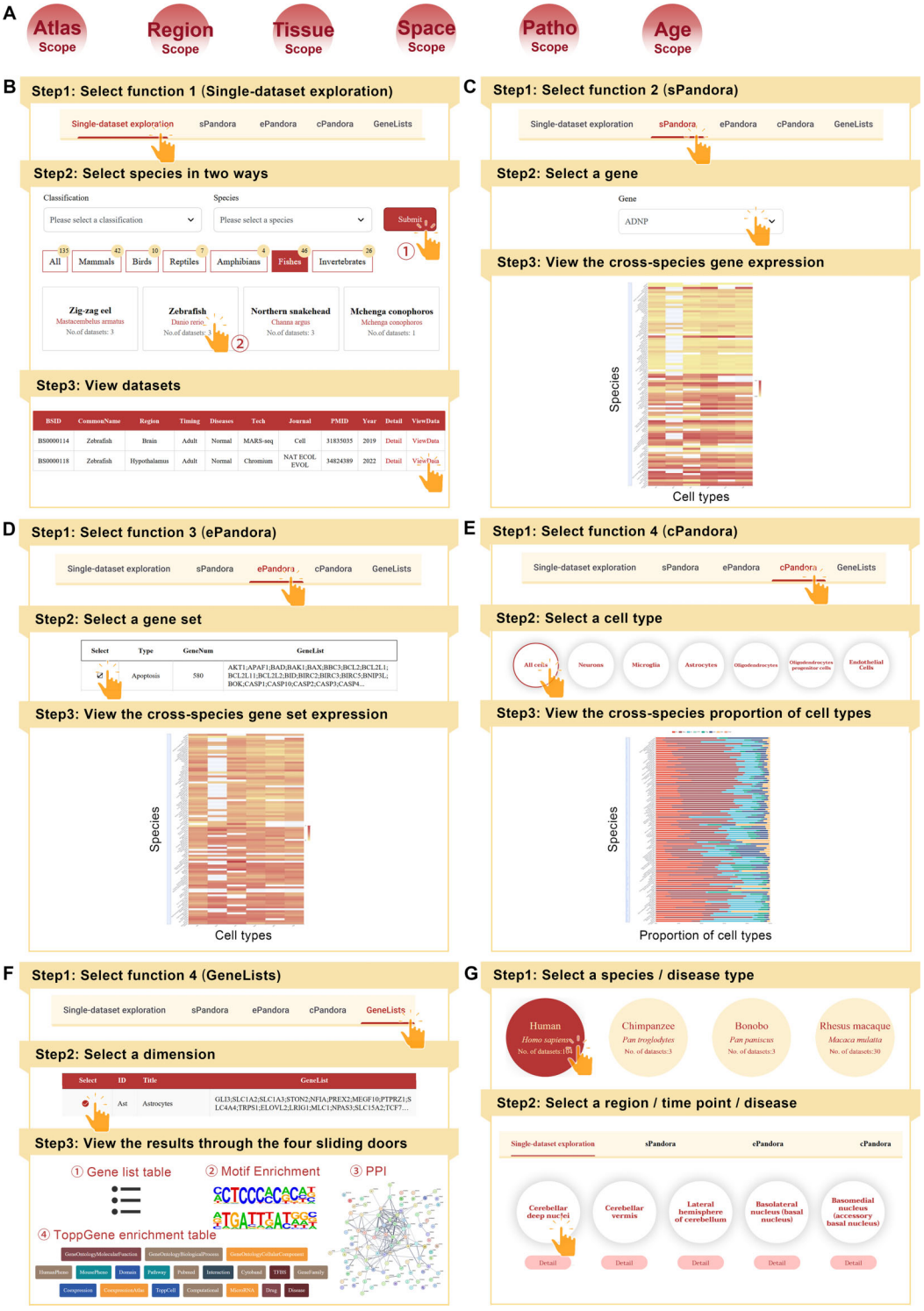


Figure 2. Overview of scBrainScope single-cell transcriptomics modules (AtlasScope, RegionScope, PathoScope, and AgeScope) and user interface. (**A**) Diagrams showing the six transcriptomics atlas modules: AtlasScope, RegionScope, TissueScope, SpaceScope, PathoScope, and AgeScope. (**B**) Taking AtlasScope as example to showcase Single-dataset exploration workflow. Users start by selecting a species either by browsing classification or clicking on the species image [53, 54]. (**C**) Example of sPandora analysis workflow of four single cell transcriptomics modules. Users enter a single gene of interest to check its expression patterns across species and cell types, displayed as a heatmap. (**D**) Example of ePandora analysis workflow of four single cell transcriptomics modules. Users select a gene set [55] from 25 curated categories to explore cross-species expression patterns across cell types. The result is shown by a heatmap. (**E**) Example of cPandora analysis workflow of four single cell transcriptomics modules Users select a cell type to examine its proportion across species, visualization as stacked and conventional bar plots. (**F**) GeneLists function workflow of four single cell transcriptomics modules. Users select a dimension (cell types in AtlasScope module) to access DEG lists, with downstream analysis including enrichment (Motif [36], ToppGene [34, 35]), PPI networks [37, 56]. (**G**) Example of dataset selection in RegionScope, PathoScope, and AgeScope. Unlike AtlasScope, which begins with species-level or disease category-level dataset selection, these modules first require users to choose a broader category (e.g. species or disease category) before refining the selection to a specific brain region, developmental stage, or disease. The panel illustrates this two-step process using representative examples (RegionScope), while the available options extend beyond those shown.

2. ShinyAtlas: ShinyCell-based interactive exploration of datasets;
3. HomoGene: Downloadable orthology tables for cross-species gene conversion;
4. GeneSets&TF: Transcriptome-based association network between curated gene sets and transcription factors (TF);
5. Stats&Anno: Sample-wise statistics and annotation summaries;
6. TimeScaleReview: Cell type & lineage timeline;
7. Enrichment: Enrichment of gene list;
8. Download: Processed expression matrices for all cells and genes.

For example, selecting the BS0000118 dataset (zebrafish, *Danio rerio*) [53] at AtlasScope (Fig. 2B) opens its corresponding scRNA-seq DataSet Page (Supplementary Fig. S1). In Part 1 (DataInfo & Gene), users can investigate gene expression at both the single gene and gene set levels using interactive visualization tools (Supplementary Fig. S1, Part1). The Information panel provides detailed metadata related to BS0000118 [53]. The samples were collected from the hypothalamus of adult zebrafish under normal conditions. They are sequenced using the chromium platform and quantified using the GRCz10 reference genome with homologous gene conversion applied. The dataset contains a total of 64 332 cells. If users wish to explore the data from a gene-level perspective, they can enter the name of a gene of interest in the panel on the right side of the interactive UMAP plot to view its expression pattern. For example, by selecting *SOX2*, a key transcription factor involved in neural stem cell maintenance, users can observe its expression in both Neu (neurons) and OPC populations (Supplementary Fig. S1 Part1). Moreover, users can click on the "GeneSet Exp" tab to examine the scores of 25 curated gene set categories within the dataset. The Part 2, ShinyAtlas integrates the ShinyCell framework. It offers seven interactive modules that enable exploration of data by cell type, gene expression, gene coexpression, metadata, and various visualization methods (Supplementary Fig. S1, Part2). Part 3, HomoGene supports evolutionary analyses by providing downloadable gene homology tables across species. For example, in zebrafish (*Danio rerio*), *camk1d* (calcium/calmodulin-dependent protein kinase I delta) has been identified as one of the genes with characteristic expression in the hypothalamus [57]. If users wish to examine the homology of *camk1d* its human counterpart and its expression in the dataset, they can enter camk1d into the table within Part 3: HomoGene to retrieve its human ortholog, *CAMK1D* (Supplementary Fig. S1, Part3). Using the "CellInfo vs GeneExpr" module in Part 2: ShinyAtlas, users can then visualize the expression pattern of *CAMK1D*, which shows broad expression in Neu cells (Supplementary Fig. S1, Part2). Transcriptional regulation can be further explored through the GeneSets&TF (Supplementary Fig. S1, Part4), which identifies associations between gene sets and transcription factors using correlation-based methods. The stemness-related gene set "Hs_ESC_SOX2_targets_Boyer" [58] comprises human genes that were identified as *SOX2* transcription factor targets in embryonic stem cells through Chromatin Immunoprecipitation (ChIP)-on-chip analysis. By selecting this gene set, users can examine transcription factors that are either positively or negatively correlated with its expression in the dataset. Our analysis shows that two transcription factors, *ID4* and *LITAF*, are significantly positively correlated with the Hs_ESC_SOX2_targets_Boyer gene set (Supplementary Fig. S1, Part4). This observation is consistent with earlier reports demonstrating that *ID4* can drive the formation of glioma stem-like cells, enhance chemoresistance, and maintain stemness properties by activating the *SOX2* pathway [59]. Quality assessment is facilitated by the Stats&Anno module (Supplementary Fig. S1, Part5).

The shared analytical framework comprises four core functional submodules, sPandora, ePandora, cPandora, and GeneLists, designed to support multidimensional transcriptomic analysis. Using the four submodules within AtlasScope as an example, sPandora allows for detailed exploration of single-gene expression across different species and brain cell types. It visualizes gene distribution patterns using heatmaps, providing cross-species comparisons of expression in brain samples (Fig. 2C). For example, when analyzing the gene *ANDP*, the heatmap displays its expression across major brain cell types – astrocytes, endothelial cells, microglia, neurons, oligodendrocytes, and OPCs—for all included species.

ePandora supports cross-species and cross-cell-type analysis of predefined gene sets. This submodule features 25 curated gene set categories and supports coordinated expression profiling. Users by selecting a gene set category (e.g. Death) and choosing a specific gene set (e.g. Apoptosis [55]), ePandora displays a heatmap of expression across all species (Fig. 2D), allowing insights into the evolutionary conservation, divergence, and context-specific regulation.

cPandora offers a comprehensive view of brain cell-type distributions across species (Fig. 2E). It quantifies proportions of six major brain cell types (neurons, microglia, astrocytes, oligodendrocytes, OPCs, and endothelial cells) across datasets. Selecting "All cells" displays their relative abundances as a stacked bar chart, while choosing a specific cell type (e.g. neurons) shows its distribution across species.

GeneLists provides cell-type-specific DEG lists generated within our database (Fig. 2F). Users can select one of six major cell types to retrieve DEGs across species and perform downstream analyses, such as ToppGene enrichment, PPI network construction, and motif enrichment. The ToppGene panel offers enrichment across functional categories such as Gene Ontology [60], Human Phenotype Ontology [61], Mouse Phenotype [62], transcription factor binding sites, pathways, diseases, and co-expression networks. Significant enrichments are displayed with detailed *P*-values and FDR-adjusted q-values, and users can click on each entry for further information. The PPI panel visualizes interactions between the selected DEGs, highlighting network connectivity and potential hub genes. The Motif Enrichment panel identifies significantly enriched transcription factor motifs, illuminating upstream regulatory elements. All results, including expression heatmap, bar charts, DEG lists, enrichment tables, and network visualizations, are downloadable for offline analysis. These four submodules are also embedded in other single-cell transcriptomics modules (RegionScope, PathoScope, and AgeScope). Guidance for the other single-cell transcriptomics modules is provided in Fig. 2G.

SpaceScope for cross-species ST data analysis

SpaceScope is a spatially resolved transcriptomic brain atlas for high-resolution analysis of gene expression using ST

data. The module comprises 1145 spatially resolved sections, covering brain, spinal cord, and embryonic tissues from 16 species, providing a comprehensive resource for studying spatial gene expression patterns across different developmental and physiological states. In the Single-dataset exploration submodule, after selecting species and tissue type (Fig. 3A Step1–3), users are presented with an information table detailing (Fig. 3A Step4): SpaceDataSetID (Unique identifier for the dataset), Common Name & Latin Name, Tissue Type, and Spatial Transcriptomics Chip ID, as well as action buttons (Select, Detail, and ViewData).

Each dataset provides multiple spatial analyses, represented by hexagonal buttons with a legend shown above the information table (Fig. 3A, Step5). Red buttons indicate available analyses, whereas gray buttons denote unsupported options. For example, the "RGB" function visualizes three user-selected genes in red, green, and blue channels, along with a merged image to display co-expression patterns (Fig. 3A, Step5). In the SP121_SDS001 sample from the human dorsolateral prefrontal cortex [63] (Fig. 3A, Step5). Separate images display the expression of *SNAP25* (red), *MOBP* (green), and *PCP4* (blue). The merged image delineates the boundary between gray matter/neurons (*SNAP25*) and white matter/oligodendrocytes (*MOBP*), while *PCP4* expression defines cortical layer 5 (L5), confirming the spatial orientation of the sample. Clicking "Cluster" shows the clustering of spots, markers of each cluster, and expression patterns of selected marker genes. The "SVG" button provides a table containing information of spatially variable genes. "Enrichment", "PPI" and "Motif" enables gene list enrichment, PPI, and motif enrichment for SVGs, respectively. The "scRNA" button links to single-cell atlas of selected datasets. The "CT" button can be used for viewing cell type atlas or brain region atlas. As shown in the image (Fig. 3A, Step 5), it provides a schematic representation of cortical layers L1-L4 in the SP121_SDS001 sample. The spatial orientation indicated by this atlas is consistent with that inferred from expression patterns of *SNAP25*, *MOBP*, and *PCP4* using the "RGB" analysis tool. "ViewData" opens the data visualization page: the Dataset Information Panel provides metadata on the dataset, and the Analysis Selection Panel provides access to different analysis tools by clicking hexagonal buttons.

SpaceScope extends the analytical framework of the single-cell modules to ST. sPandora supports single-gene exploration across spatial chips. It enables visualization of expression across four classes: Species, Development (Devo), Tissue, and Disease (Diz) (Fig. 3B). ePandora provides gene set-level exploration under the same conditions, assessing coordinated activity across spatial datasets (Fig. 3C). cPandora estimates cell-type composition from spatial data and projects inferred proportions onto spatial coordinates (Fig. 3D).

TissueScope for cross-species bulk RNA-seq data analysis

To complement single-cell and ST data, scBrainScope incorporates a large-scale bulk RNA-seq module, TissueScope, designed for cross-species transcriptomic comparisons. This module enables investigation of conserved gene-expression programs and interspecies variation across a broad phylogenetic spectrum. TissueScope integrates 737 bulk RNA-seq datasets from 275 species, allowing users to examine macro-level transcriptomic patterns across evolutionary dimensions. Users can explore bulk RNA-seq sequencing data for species of interest at Single-Dataset exploration (Fig. 4A). After users click on a species image or select a species (Fig. 4A Step1,2), an information table appears showing details like Classification, Common Name & Latin Name, Tissue Source, Journal & PMID Information, Detail, and ViewData (Fig. 4A Step3). Clicking "Detail" reveals in-depth metadata, including the study title, journal, PMID, and additional information. Clicking "ViewData" opens the bulk RNA-seq data for further exploration.

The Bulk RNA-Seq DataSet page serves as a comprehensive platform for exploring bulk RNA sequencing data across species. The page begins with a metadata overview that displays basic dataset information, sequencing quality-control metrics, and categorized gene expression matrices [69]. Predicted coding sequences and translated products are inferred via TransDecoder. In Part 3, sequence homology and functional inference are performed through BLASTP [70] alignment. To facilitate functional exploration, Part 4 showcases expression matrices by gene categories such as TFs, GPCRs, and immune related genes. Part 5 gives users a phylogenetic tree encompassing 275 species for evolutionary analysis. PCA analysis is available at Part 6, which projects samples into a shared expression space, with the plot color-coded by species categories (mammals, birds, reptiles, amphibians, and fishes). Furthermore, Part7 highlights genes with high, conserved expression across seven representative species [71].

TissueScope offers four main analytical functions: sPandora, ePandora, GeneLists, and TimeTree. sPandora enables searching and comparing single gene expression across species (Fig. 4B). ePandora provides gene-set-based comparisons of transcriptomic programs across bulk RNA-seq datasets (Fig. 4C). GeneLists allows exploration of DEG lists for each species classification (Fig. 4D).

sPandora for cross-species single-gene analysis

sPandora supports cross-dimensional single-gene analysis across all scBrainScope modules. Users can query a single gene and visualize its expression. Here we use *ADNP* [72] gene as an example, which has been reported related to autism spectrum disorder, as an example (Fig. 5A Step1). In AtlasScope, users can view *ADNP* expression patterns across six major brain cell types in 135 species (Fig. 5A Step2, Step3 I). RegionScope displays *ADNP* expression across different cell types from brain regions within a selected species (e.g. human) (Fig. 5A Step3 II). Our analysis revealed that *ADNP* is highly expressed in endothelial cells in the dorsolateral prefrontal cortex (DLPFC) of humans. TissueScope provides a broad overview of *ADNP* expression in 275 species using bulk RNA seq datasets (Fig. 5A Step3 III). The result offers insight into single gene evolutionary conservation and potential species-specific function. In SpaceScope, users can view the single gene expression on different ST chips by browsing different groups (Species, Devo, Tissue, Diz). The Detail buttons provide additional metadata for the ST chip such as the common name of the source species and the source tissues (Fig. 5A Step3 IV). PathoScope allows users to explore *ADNP* expression patterns across different cell types under various disease conditions (Fig. 5A Step3 V). The heatmap result indicates that *ADNP* expression is relatively high in multiple cell types under neurodegenerative diseases. AgeScope offers the *ADNP* expression across different cell types at various developmental

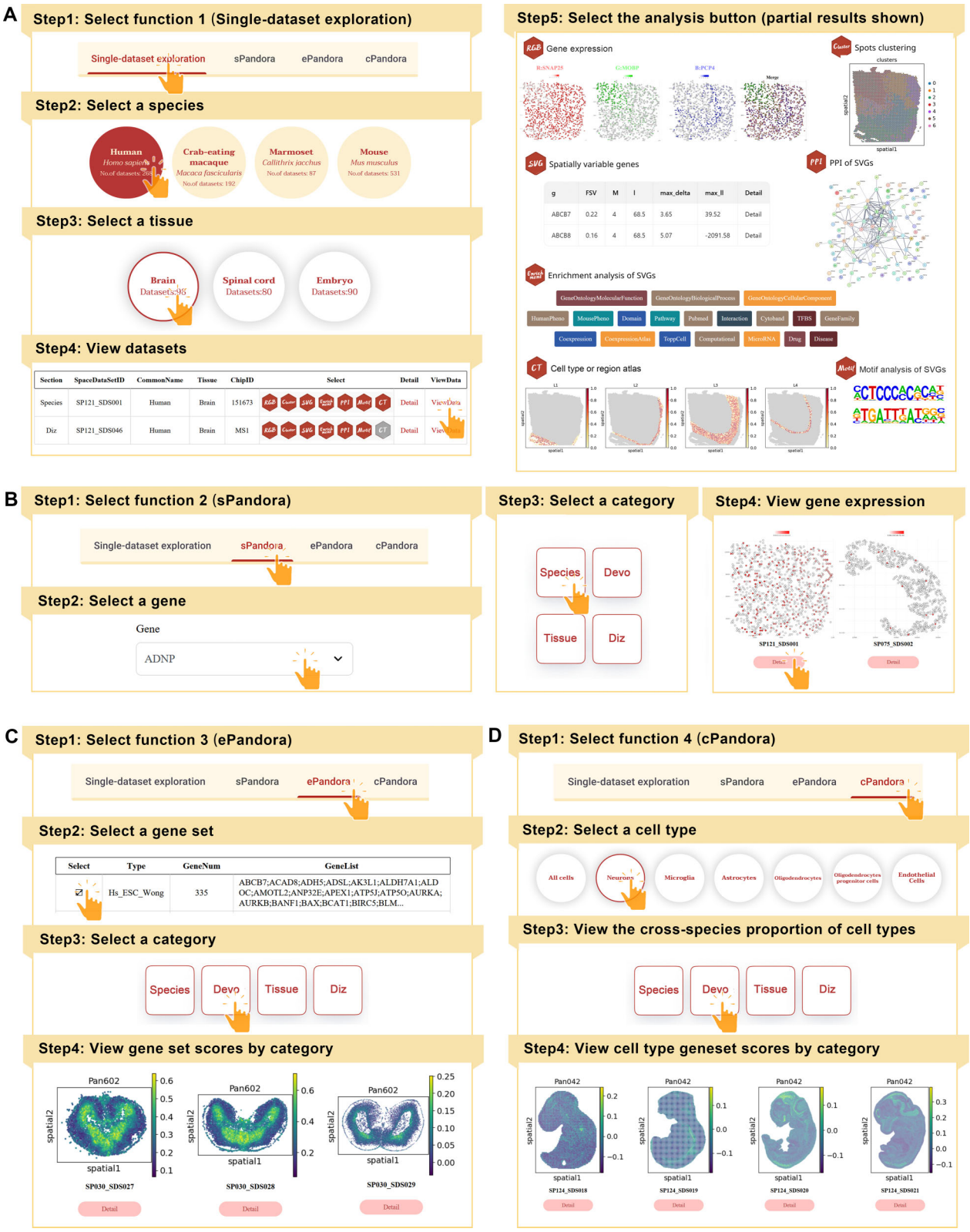


Figure 3. Overview of scBrainScope ST module (SpaceScope) and user interface. (**A**) Single-dataset exploration workflow in SpaceScope. Users start by picking a species and tissue type (such as brain, spinal cord, or embryo), then browse the available datasets and review the detailed metadata. The analysis panel offers multiple functions (e.g. spot visualization, spatial variable gene identification, enrichment analysis and spatial region annotations) [63, 64]. (**B**) sPandora workflow in SpaceScope. Users select a gene of interest and choose a category (Species, Devo, Tissue, Diz) to explore how that gene is expressed spatially across different conditions [63, 65]. (**C**) ePandora workflow in SpaceScope. Users pick predefined gene sets [66] and examine their spatial expression across various categories [67]. (**D**) Users choose a cell type to explore its spatial distribution across different categories [20].

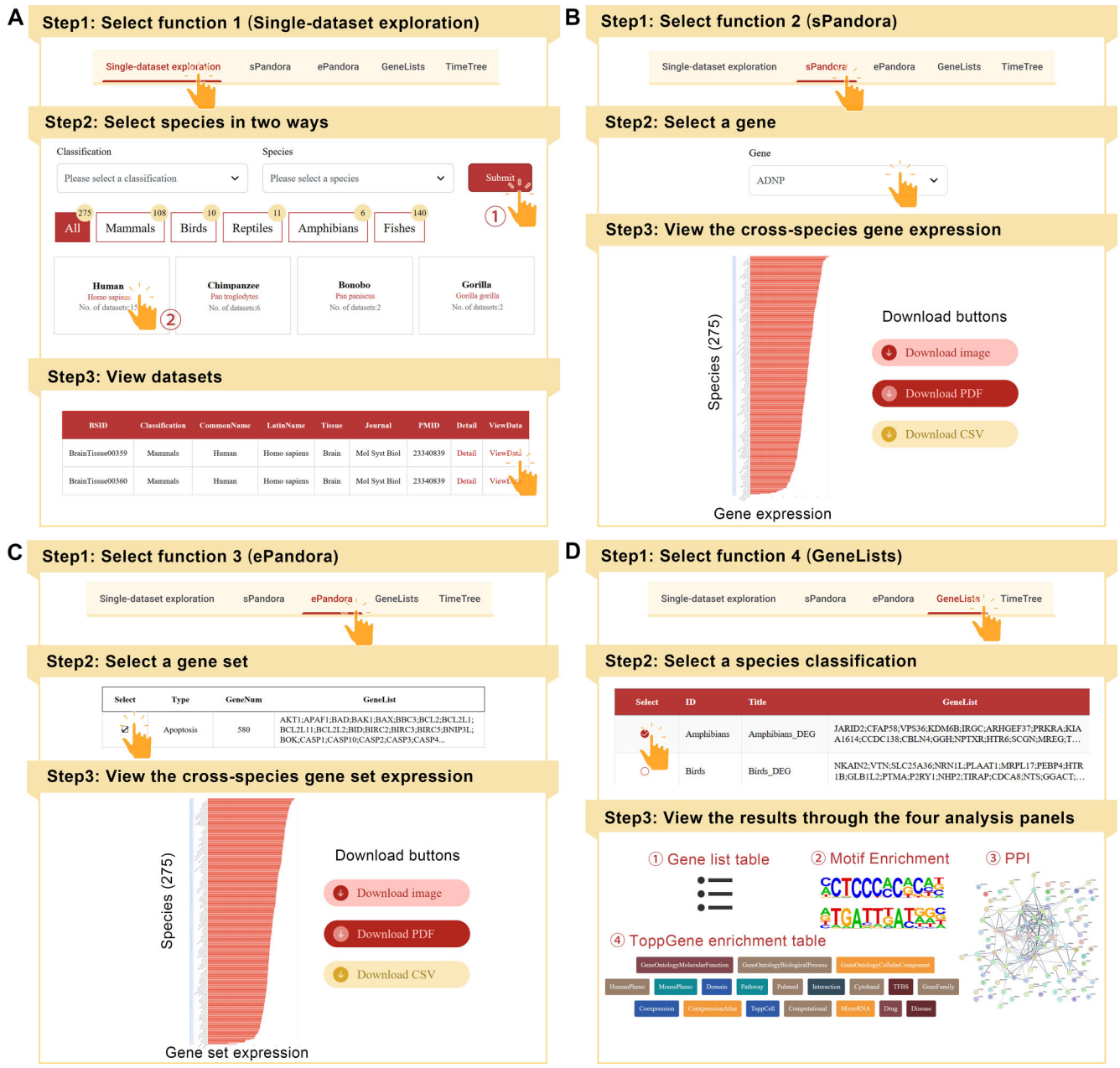


Figure 4. Overview of scBrainScope bulk RNA-seq module (TissueScope) and user interface. (**A**) Single-dataset exploration workflow. Users select species by browsing the classification or clicking a species image to view the available bulk RNA-seq datasets, and detailed metadata [68]. (**B**) sPandora workflow on TissueScope. Users pick a single gene of interest and visualize its expression across 275 species with options to download the result in image, PDF, or CSV format. (**C**) ePandora workflow on TissueScope. Users can check a predefined gene set and examine its expression across species [55]. (**D**) GeneLists workflow on TissueScope. Users choose species classifications to access differently expressed gene lists and run enrichment analysis like pathway enrichment, motif enrichment [36], PPI networks [37, 56], and ToppGene enrichment [34, 35].

stages from a selected species (Fig. 5A Step3 VI). This allows users to track how *ADNP* expression changes over time and to identify stage-specific patterns.

ePandora for cross-species entire gene set analysis

ePandora lets users explore the expression of predefined gene sets from 25 biological categories (Fig. 5B Step1). Users can apply ePandora across all scBrainScope modules to examine selected gene sets (Fig. 5B Step2). After choosing an interested category (e.g. Death) and a specific gene set (*e.g.* Apoptosis [55]), users access expression heatmaps in the following three modules. AtlasScope showcases the "Apoptosis" signatures in six cell types across species (Fig. 5B Step3 I). The result shows that "apoptosis" is comparatively highly expressed in microglial cells, and this pattern appears to be conserved across the species analyzed. Using RegionScope, users can explore the "Apoptosis" expressed across regions of a selected species (e.g. Rhesus macaque) (Fig. 5B Step3 II). TissueScope provides the "Apoptosis" expression in our bulk RNA seq datasets from 275 species (Fig. 5B Step3 III). The result gives a broad overview of "Apoptosis" expression patterns across a wide evolutionary spectrum.

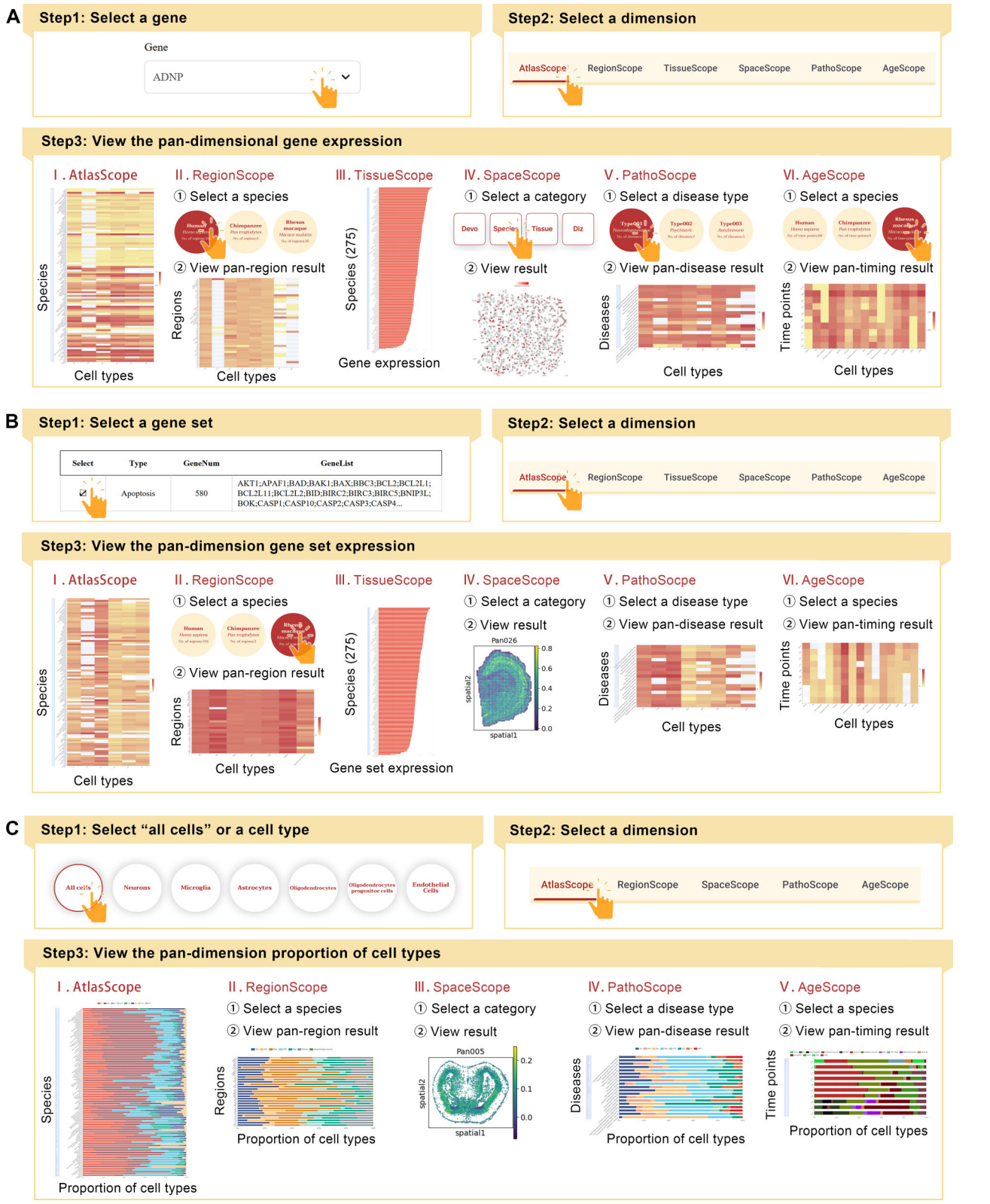


Figure 5. Overview of scBrainScope analysis modules (sPandora, ePandora, cPandora) for pan-dimensional comparison of genes, cell types, and gene sets. (**A**) sPandora workflow: users enter a single-gene query and investigate its expression across all scBrainScope dimensions. The results are provided with heatmaps (AtlasScope, RegionScope, PathoScope, AgeScope), bar plots (TissueScope), or expression plots (SpaceScope) [63]. (**B**) ePandora workflow: users select the gene sets of interest and dimensions to view the result [55]. Users can explore the expression pattern of predefined gene sets across dimensions including species (AtlasScope, TissueScope), brain regions (RegionScope), spatial locations (SpaceScope) [76], disease conditions (PathoScope), and development stages (AgeScope). The results are visualized as heatmaps, bar plots or expression plots. (**C**) cPandora workflow: users select a cell type (e.g. neurons, microglia, astrocytes) or all cell types, and examine its proportion across species (AtlasScope), brain regions (RegionScope), spatial locations (SpaceScope) [67], disease conditions (PathoScope), or developmental stages (AgeScope). The corresponding results are displayed by stacked bar plots or expression plots.

SpaceScope in ePandora divides results into four groups (Species, Devo, Tissue, Diz) (Fig. 5B Step3 IV), allowing users to view "Apoptosis" expression on different ST chips. For example, by selecting Species, users can examine the "Apoptosis" expression pattern on ST chips from different species. These cross-species comparisons may provide valuable clues for studying conserved expression patterns across vertebrate animals. PathoScope explores "Apoptosis" expression pattern across cell types (Fig. 5B Step3 V) within seven disease categories: Neurodegenerative, Psychiatric, AutoImmune, Developmental, Tumors, Infection and Other Diseases. Results show that "Apoptosis" expression is consistently high in astrocytes, endothelial cells, and microglial cells across nearly all analyzed neurodegenerative diseases like Alzheimer's disease and Parkinson's disease. This widespread upregulation suggests that apoptosis may play a critical role in the cellular vulnerability associated with these conditions. The finding aligns with prior evidence that neuronal death underlies the symptoms of many human neurological disorders, including Alzheimer's, Parkinson's, and Huntington's diseases [73, 74]. AgeScope provides the "Apoptosis" expression across different developmental stages of a selected species (e.g. Rhesus macaque [75]) (Fig. 5B Step3 VI). Users can view "Apoptosis" expression in multiple cell types at each stage. By comparing patterns across cell types, users can determine whether certain populations are more vulnerable to apoptosis during stages.

cPandora for cross-species cell type analysis

cPandora allows users to compare the distribution of brain cell types (neurons, microglia, astrocytes, oligodendrocytes, OPCs, and endothelial cells) across modules. Users can select "All cells" or a specific cell type to explore (Fig. 5C, Step1). If "All cells" is selected, AtlasScope (Fig. 5C Step2) major cell-type proportions across all species, providing a global cross-species overview (Fig. 5C Step3 I). RegionScope provides insights into cell type distribution across brain region-specific datasets within selected species, highlighting regionally enriched or depleted cell types (Fig. 5C Step3 II). SpaceScope maps the distributions of cell-type features onto spatial-transcriptomics chips, with data grouped by Species, Devo (developmental stage), Tissue, or Diz (disease) categories (Fig. 5C Step3 III). Users are able to explore the SpaceScope result by grouping the data into four categories: species, development stage (Devo), tissue type (Tissue), and disease condition (Diz). PathoScope presents cell type distributions across brain disease-related datasets, organized into seven categories: neurodegenerative disorders, psychiatric conditions, autoimmune diseases, developmental disorders, tumors, infections, and other disease types. PathoScope can be used to explore how specific diseases influence the abundance of particular cell types (Fig. 5C Step3 IV). The comparative analysis may help highlight conserved cellular response and may pinpoint candidate cell population for therapeutic targets. AgeScope displays the cell type distribution across different developmental stages of a selected species (Fig. 5C Step3 V). Users can utilize AgeScope to identify critical window of cell type expansion or reduction, offering insight into key category events during brain development. Overall, cPandora systematically uncovers both conserved and context-specific trends in cell type dynamics.

Summary and future perspectives

The rapid expansion of transcriptomic data across multiple species, brain regions, development stages, and disease conditions, presents both unprecedented opportunities and significant challenges for neuroscience research. To address these, we developed scBrainScope, an integrative, interactive atlas unifying 118.7 million single-cell profiles, 737 bulk RNA-seq datasets, and 1154 ST datasets across evolutionary, anatomical, developmental, and pathological dimensions. By combining three transcriptomic modalities (single-cell, bulk, and spatial) and providing pan-dimensional analysis pipelines at the single-gene level (sPandora), gene-set level (ePandora), and cell-type level (cPandora), scBrainScope enables comprehensive comparative studies across species, regions, stages, and diseases. In the future, we will expand with high-quality transcriptomic datasets from newly sequenced species and integrate emerging multi-omics modalities including epigenomics, proteomics, and metabolomics.

Acknowledgements

We acknowledge the high-throughput sequencing and high-performance computing platform of Institute of Systems Medicine, Chinese Academy of Medical Sciences/Suzhou Institute of Systems Medicine (ISM) for technical support.

Author contributions: D.C., L.H., Y.M., and X.W. conceived and supervised the project. S.Q., Y.Y., H.W., M.L., Y.D., Y.C., C.L., L.P., L.X., X.Q., Y.G., and S.W. collected the datasets and developed bioinformatic pipelines. S.Q., Y.Y., H.W., M.L., Y.D., Y.C., C.L., L.P., and L.X. performed preprocessing and integrative analyses. D.C., S.Q., Y.Y., M.L., and Y.H. developed the scBrainScope website. D.C., L.H., Y.Y., H.W., and S.Q. wrote the manuscript. All authors reviewed and approved the final version of the manuscript.

Supplementary data

Supplementary data is available at NAR online.

Conflict of interest

None declared.

Funding

The work was supported by Science and Technology Innovation 2030 Major Project (STI2030-Major Projects 2022ZD0205700), National Key Research and Development Program of China (2024YFC3607500), Non communicable Chronic Diseases-National Science and Technology Major Project (2024ZD0519600), National Natural Science Foundation of China under grant 32300560, the Natural Science Foundation of Jiangsu Province (grants no. BK20220279), the CAMS Innovation Fund for Medical Sciences (CIFMS) (2021-I2M-1-061, 2021-I2M-1–074, 2022-I2M-2–004), the Nonprofit Central Research Institute Fund of Chinese Academy of Medical Sciences (2022-RC416-01), the NCTIB Fund for R&D Platform for Cell and Gene Therapy, the Suzhou Municipal Key Laboratory (SZS2022005, SZS2023005), Gusu Innovation and Entrepreneurship Leading Talents Program (ZXL2022475), Beijing Research Ward Excellence Program (BRWEP2024W114060100) and 2025 PHC Cai Yuanpei

project, the Special Research Fund for Central Universities, Peking Union Medical College (3332024089).

Data availability

Datasets are available at http://8.142.154.29/scBrainScope.

References

1. Saunders A, Macosko EZ, Wysoker A *et al.* Molecular diversity and specializations among the cells of the adult mouse brain. *Cell* 2018;174:1015–30. https://doi.org/10.1016/j.cell.2018.07.028
2. Dumitrascu B, Villar S, Mixon DG *et al.* Optimal marker gene selection for cell type discrimination in single cell analyses. *Nat Commun* 2021;12:1186. https://doi.org/10.1038/s41467-021-21453-4
3. Palla G, Fischer DS, Regev A *et al.* Spatial components of molecular tissue biology. *Nat Biotechnol* 2022;40:308–18. https://doi.org/10.1038/s41587-021-01182-1
4. 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
5. Zhuo L, Wang M, Song T *et al.* MAPbrain: a multi-omics atlas of the primate brain. *Nucleic Acids Res* 2025;53:D1055–65. https://doi.org/10.1093/nar/gkae911
6. Xu Z, Wang W, Yang T *et al.* STOmicsDB: a comprehensive database for spatial transcriptomics data sharing, analysis and visualization. *Nucleic Acids Res* 2024;52:D1053–61. https://doi.org/10.1093/nar/gkad933
7. Jiang J, Wang C, Qi R *et al.* scREAD: a single-cell RNA-seq database for Alzheimer's disease. *iScience* 2020;23:101769. https://doi.org/10.1016/j.isci.2020.101769
8. Deng Y, Lu Y, Li M *et al.* SCAN: spatiotemporal cloud atlas for neural cells. *Nucleic Acids Res* 2024;52:D998–1009. https://doi.org/10.1093/nar/gkad895
9. Zhou Z, Tan C, Chau MHK *et al.* TEDD: a database of temporal gene expression patterns during multiple developmental periods in human and model organisms. *Nucleic Acids Res* 2023;51:D1168–78. https://doi.org/10.1093/nar/gkac978
10. Sunkin SM, Ng L, Lau C *et al.* Allen Brain Atlas: an integrated spatio-temporal portal for exploring the central nervous system. *Nucleic Acids Res* 2013;41:D996–1008. https://doi.org/10.1093/nar/gks1042
11. Dataset: Allen Institute for Brain Science . *Cross-species Iso-Seq of mammalian M1*. 2022. https://knowledge.brain-map.org/data/976C22IARSIYLTG9T08/summary (20 August 2024, date last accessed).
12. Tarhan L, Bistline J, Chang J *et al.* Single Cell Portal: an interactive home for single-cell genomics data. bioRxiv, https://doi.org/10.1101/2023.07.13.548886. 17 July 2023, preprint: not peer reviewed.
13. 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
14. Thakur M, Bateman A, Brooksbank C *et al.* EMBL's European Bioinformatics Institute (EMBL-EBI) in 2022. *Nucleic Acids Res* 2023;51:D9–D17. https://doi.org/10.1093/nar/gkac1098
15. Zheng GXY, Terry JM, Belgrader P *et al.* Massively parallel digital transcriptional profiling of single cells. *Nat Commun* 2017;8:14049. https://doi.org/10.1038/ncomms14049
16. Wolf FA, Angerer P, Theis FJ. SCANPY: large-scale single-cell gene expression data analysis. *Genome Biol* 2018;19:15. https://doi.org/10.1186/s13059-017-1382-0
17. Hao Y, Hao S, Andersen-Nissen E *et al.* Integrated analysis of multimodal single-cell data. *Cell* 2021;184:3573–3587.e29. https://doi.org/10.1016/j.cell.2021.04.048
18. Grabherr MG, Haas BJ, Yassour M *et al.* Full-length transcriptome assembly from RNA-Seq data without a reference genome. *Nat Biotechnol* 2011;29:644–52. https://doi.org/10.1038/nbt.1883
19. Haas BJ, Papanicolaou A, Yassour M *et al.* *De novo* transcript sequence reconstruction from RNA-seq using the Trinity platform for reference generation and analysis. *Nat Protoc* 2013;8:1494–512. https://doi.org/10.1038/nprot.2013.084
20. Chen A, Liao S, Cheng M *et al.* Spatiotemporal transcriptomic atlas of mouse organogenesis using DNA nanoball-patterned arrays. *Cell* 2022;185:1777–92. https://doi.org/10.1016/j.cell.2022.04.003
21. Fang S, Xu M, Cao L *et al.* Stereopy: modeling comparative and spatiotemporal cellular heterogeneity via multi-sample spatial transcriptomics. *Nat Commun* 2025;16:3741. https://doi.org/10.1038/s41467-025-58079-9
22. Wolock SL, Lopez R, Klein AM. Scrublet: computational identification of cell doublets in single-cell transcriptomic data. *Cell Syst* 2019;8:281–91. https://doi.org/10.1016/j.cels.2018.11.005
23. Chen S, Zhou Y, Chen Y *et al.* fastp: an ultra-fast all-in-one FASTQ preprocessor. *Bioinformatics* 2018;34:i884–90. https://doi.org/10.1093/bioinformatics/bty560
24. Kasprzyk A. BioMart: driving a paradigm change in biological data management. *Database* 2011;2011:bar049. https://doi.org/10.1093/database/bar049
25. O'Leary NA, Wright MW, Brister JR *et al.* Reference sequence (RefSeq) database at NCBI: current status, taxonomic expansion, and functional annotation. *Nucleic Acids Res* 2016;44:D733–45. https://doi.org/10.1093/nar/gkv1189
26. Emms DM, Kelly S. OrthoFinder: phylogenetic orthology inference for comparative genomics. *Genome Biol* 2019;20:238. https://doi.org/10.1186/s13059-019-1832-y
27. Buchfink B, Xie C, Huson DH. Fast and sensitive protein alignment using DIAMOND. *Nat Methods* 2015;12:59–60. https://doi.org/10.1038/nmeth.3176
28. Siletti K, Hodge R, Mossi Albiach A *et al.* Transcriptomic diversity of cell types across the adult human brain. *Science* 2023;382:eadd7046. https://doi.org/10.1126/science.add7046
29. Domínguez Conde C, Xu C, Jarvis LB *et al.* Cross-tissue immune cell analysis reveals tissue-specific features in humans. *Science* 2022;376:eabl5197. https://doi.org/10.1126/science.abl5197
30. Korsunsky I, Millard N, Fan J *et al.* Fast, sensitive and accurate integration of single-cell data with Harmony. *Nat Methods* 2019;16:1289–96. https://doi.org/10.1038/s41592-019-0619-0
31. Ouyang JF, Kamaraj US, Cao EY *et al.* ShinyCell: simple and sharable visualization of single-cell gene expression data. *Bioinformatics* 2021;37:3374–6. https://doi.org/10.1093/bioinformatics/btab209
32. Rauluseviciute I, Riudavets-Puig R, Blanc-Mathieu R *et al.* JASPAR 2024: 20th anniversary of the open-access database of transcription factor binding profiles. *Nucleic Acids Res* 2024;52:D174–82. https://doi.org/10.1093/nar/gkad1059
33. Castro-Mondragon JA, Riudavets-Puig R, Rauluseviciute I *et al.* JASPAR 2022: the 9th release of the open-access database of transcription factor binding profiles. *Nucleic Acids Res* 2022;50:D165–73. https://doi.org/10.1093/nar/gkab1113
34. Granger B, Berto S. scToppR: a coding-friendly R interface to ToppGene. *Bioinformatics* 2024;40:btae582.https://doi.org/10.1093/bioinformatics/btae582
35. Chen J, Bardes EE, Aronow BJ *et al.* ToppGene Suite for gene list enrichment analysis and candidate gene prioritization. *Nucleic Acids Res* 2009;37:W305–11. https://doi.org/10.1093/nar/gkp427
36. Heinz S, Benner C, Spann N *et al.* Simple combinations of lineage-determining transcription factors prime cis-regulatory elements required for macrophage and B cell identities. *Mol Cell* 2010;38:576–89. https://doi.org/10.1016/j.molcel.2010.05.004
37. Rezwani M, Pourfathollah AA, Noorbakhsh F. rbioapi: user-friendly R interface to biologic web services' API.

Bioinformatics 2022;38:2952–3. https://doi.org/10.1093/bioinformatics/btac172
38. 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
39. Bray NL, Pimentel H, Melsted P *et al.* Near-optimal probabilistic RNA-seq quantification. *Nat Biotechnol* 2016;34:525–7. https://doi.org/10.1038/nbt.3519
40. Kumar S, Suleski M, Craig JM *et al.* TimeTree 5: an expanded resource for species divergence times. *Mol Biol Evol* 2022;39:msac174.https://doi.org/10.1093/molbev/msac174
41. Svensson V, Teichmann SA, Stegle O. SpatialDE: identification of spatially variable genes. *Nat Methods* 2018;15:343–6. https://doi.org/10.1038/nmeth.4636
42. Wickham H. *ggplot2: Elegant Graphics for Data Analysis*. New York: Springer-Verlag, 2016.
43. Chai Y, Zhang R, Jia S *et al.* Toti: an integrated multi-omics database to decipher the epigenetic regulation of gene expression in totipotent stem cells. *Stem Cell Rep* 2025;20:102605. https://doi.org/10.1016/j.stemcr.2025.102605
44. Cui T, Li Y-Y, Li B-L *et al.* SpatialRef: a reference of spatial omics with known spot annotation. *Nucleic Acids Res* 2025;53:D1215–23. https://doi.org/10.1093/nar/gkae892
45. Wang C, Acosta D, McNutt M *et al.* A single-cell and spatial RNA-seq database for Alzheimer's disease (ssREAD). *Nat Commun* 2024;15:4710. https://doi.org/10.1038/s41467-024-49133-z
46. Yuan Z, Pan W, Zhao X *et al.* SODB facilitates comprehensive exploration of spatial omics data. *Nat Methods* 2023;20:387–99. https://doi.org/10.1038/s41592-023-01773-7
47. Zheng Y, Chen Y, Ding X *et al.* Aquila: a spatial omics database and analysis platform. *Nucleic Acids Res* 2023;51:D827–34. https://doi.org/10.1093/nar/gkac874
48. Li X, Xiao C, Qi J *et al.* STellaris: a web server for accurate spatial mapping of single cells based on spatial transcriptomics data. *Nucleic Acids Res* 2023;51:W560–8. https://doi.org/10.1093/nar/gkad419
49. Fan Z, Luo Y, Lu H *et al.* SPASCER: spatial transcriptomics annotation at single-cell resolution. *Nucleic Acids Res* 2023;51:D1138–49. https://doi.org/10.1093/nar/gkac889
50. Song L, Pan S, Zhang Z *et al.* STAB: a spatio-temporal cell atlas of the human brain. *Nucleic Acids Res* 2021;49:D1029–37. https://doi.org/10.1093/nar/gkaa762
51. Zhao T, Lyu S, Lu G *et al.* SC2disease: a manually curated database of single-cell transcriptome for human diseases. *Nucleic Acids Res* 2021;49:D1413–9. https://doi.org/10.1093/nar/gkaa838
52. Fan Z, Chen R, Chen X. SpatialDB: a database for spatially resolved transcriptomes. *Nucleic Acids Res* 2019;48:D233–7.
53. Shafer MER, Sawh AN, Schier AF. Gene family evolution underlies cell-type diversification in the hypothalamus of teleosts. *Nat Ecol Evol* 2022;6:63–76. https://doi.org/10.1038/s41559-021-01580-3
54. Geirsdottir L, David E, Keren-Shaul H *et al.* Cross-species single-cell analysis reveals divergence of the primate microglia program. *Cell* 2019;179:1609–22. https://doi.org/10.1016/j.cell.2019.11.010
55. Zou Y, Xie J, Zheng S *et al.* Leveraging diverse cell-death patterns to predict the prognosis and drug sensitivity of triple-negative breast cancer patients after surgery. *Int J Surg* 2022;107:106936. https://doi.org/10.1016/j.ijsu.2022.106936
56. 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
57. Toro S, Wegner J, Muller M *et al.* Identification of differentially expressed genes in the zebrafish hypothalamic-pituitary axis. *Gene Expr Patterns* 2009;9:200–8. https://doi.org/10.1016/j.gep.2008.12.007
58. Boyer LA, Lee TI, Cole MF *et al.* Core transcriptional regulatory circuitry in human embryonic stem cells. *Cell* 2005;122:947–56. https://doi.org/10.1016/j.cell.2005.08.020
59. Jeon HM, Sohn YW, Oh SY *et al.* ID4 imparts chemoresistance and cancer stemness to glioma cells by derepressing miR-9∗-mediated suppression of SOX2. *Cancer Res* 2011;71:3410–21. https://doi.org/10.1158/0008-5472.CAN-10-3340
60. Consortium TGO, Aleksander SA, Balhoff J *et al.* The Gene Ontology knowledgebase in 2023. *Genetics* 2023;224:iyad031.https://doi.org/10.1093/genetics/iyad031
61. Köhler S, Gargano M, Matentzoglu N *et al.* The Human Phenotype Ontology in 2021. *Nucleic Acids Res* 2021;49:D1207.
62. Groza T, Gomez FL, Mashhadi HH *et al.* The International Mouse Phenotyping Consortium: comprehensive knockout phenotyping underpinning the study of human disease. *Nucleic Acids Res* 2023;51:D1038–45. https://doi.org/10.1093/nar/gkac972
63. Maynard KR, Collado-Torres L, Weber LM *et al.* Transcriptome-scale spatial gene expression in the human dorsolateral prefrontal cortex. *Nat Neurosci* 2021;24:425–36. https://doi.org/10.1038/s41593-020-00787-0
64. Kaufmann M, Schaupp A-L, Sun R *et al.* Identification of early neurodegenerative pathways in progressive multiple sclerosis. *Nat Neurosci* 2022;25:944–55. https://doi.org/10.1038/s41593-022-01097-3
65. Chen A, Sun Y, Lei Y *et al.* Single-cell spatial transcriptome reveals cell-type organization in the macaque cortex. *Cell* 2023;186:3726–43. https://doi.org/10.1016/j.cell.2023.06.009
66. Wong DJ, Liu H, Ridky TW *et al.* Module map of stem cell genes guides creation of epithelial cancer stem cells. *Cell Stem Cell* 2008;2:333–44. https://doi.org/10.1016/j.stem.2008.02.009
67. Wei X, Fu S, Li H *et al.* Single-cell Stereo-seq reveals induced progenitor cells involved in axolotl brain regeneration. *Science* 2022;377:eabp9444. https://doi.org/10.1126/science.abp9444
68. Mazin P, Xiong J, Liu X *et al.* Widespread splicing changes in human brain development and aging. *Mol Syst Biol* 2013;9:633. https://doi.org/10.1038/msb.2012.67
69. Brawand D, Soumillon M, Necsulea A *et al.* The evolution of gene expression levels in mammalian organs. *Nature* 2011;478:343–8. https://doi.org/10.1038/nature10532
70. Altschul SF, Madden TL, Schäffer AA *et al.* Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res* 1997;25:3389–402. https://doi.org/10.1093/nar/25.17.3389
71. Geirsdottir L, David E, Keren-Shaul H *et al.* Cross-species single-cell analysis reveals divergence of the primate microglia program. *Cell* 2019;179:1609–22. https://doi.org/10.1016/j.cell.2019.11.010
72. D'Incal CP, Van Rossem KE, De Man K *et al.* Chromatin remodeler Activity-Dependent Neuroprotective Protein (ADNP) contributes to syndromic autism. *Clin Epigenet* 2023;15:45. https://doi.org/10.1186/s13148-023-01450-8
73. Mattson MP. Apoptosis in neurodegenerative disorders. *Nat Rev Mol Cell Biol* 2000;1:120–30. https://doi.org/10.1038/35040009
74. Moujalled D, Strasser A, Liddell JR. Molecular mechanisms of cell death in neurological diseases. *Cell Death Differ* 2021;28:2029–44. https://doi.org/10.1038/s41418-021-00814-y
75. Micali N, Ma S, Li M *et al.* Molecular programs of regional specification and neural stem cell fate progression in macaque telencephalon. *Science* 2023;382:eadf3786. https://doi.org/10.1126/science.adf3786
76. Liao K, Xiang Y, Huang F *et al.* Spatial and single-nucleus transcriptomics decoding the molecular landscape and cellular organization of avian optic tectum. *iScience* 2024;27:109009. https://doi.org/10.1016/j.isci.2024.109009

Received: August 6, 2025. **Revised:** September 18, 2025. **Accepted:** September 23, 2025