An Integrated Single-Cell and Epigenomic Resource for Comparative Analysis of the Basal Ganglia

Wenjin Zhang[1,#], Wubin Ding[2,#], Kai Li[3,#], Lei Chang[3], Amit Klein[2], Cindy Tatiana Báez-Becerra[4], Jonathan A. Rink[4], Anna Bartlett[2], Huaming Chen[2], Natalie Schenker[2], Nelson Johansen[5], Tyler Mollenkopf[5], Yuanyuan Fu[5], Xie Yang[3], Shane Liu[1], Chanrung Seng[1], Benpeng Miao[1], Tianjie Liu[1], Quan Zhu[6], Rebecca D. Hodge[5], Trygve E. Bakken[5], Ed S. Lein[5], Michael Hawrylycz[5], Xiangmin Xu[7,8,9,10,11], M. Margarita Behrens[4], Bing Ren[3,12], Joseph R. Ecker[2,13], Ting Wang[1,14, *], Daofeng Li[1,*]

[1] Department of Genetics, The Edison Family Center for Genome Sciences and Systems Biology, Washington University School of Medicine, St. Louis, MO 63110, USA
[2] Genomic Analysis Laboratory, The Salk Institute for Biological Studies, La Jolla, CA 92037, USA
[3] New York Genome Center, New York, NY, USA
[4] Computational Neurobiology Laboratory, The Salk Institute for Biological Studies, La Jolla, CA 92037, USA
[5] Allen Institute for Brain Science, Seattle, WA, 98109, USA
[6] Center for Epigenomics, Department of Cellular and Molecular Medicine, University of California, San Diego; La Jolla, CA, USA
[7] Department of Anatomy and Neurobiology, School of Medicine, University of California, Irvine, Irvine, CA 92697, USA
[8] Department of Microbiology and Molecular Genetics, School of Medicine, University of California, Irvine, Irvine, CA 92697, USA
[9] Center for Neural Circuit Mapping, University of California, Irvine, Irvine, CA 92697, USA
[10] Department of Biomedical Engineering, University of California, Irvine, Irvine, CA 92697, USA
[11] Department of Computer Science, University of California, Irvine, Irvine, CA 92697, USA
[12] Department of Genetics and Development, Systems Biology, Biochemistry and Molecular Biophysics, Columbia University Irving Medical Center, New York, NY, USA
[13] Howard Hughes Medical Institute, The Salk Institute for Biological Studies, La Jolla, CA 92037, USA
[14] McDonnell Genome Institute, Washington University School of Medicine, St. Louis, MO 63108, USA
[#] Equal contribution
* Correspondence: twang@wustl.edu (T.W.), dli23@wustl.edu (D.L.)

Abstract

The basal ganglia are central to adaptive behaviors, integrating motor control, cognition, emotion, and reward processing. An improved understanding of basal ganglia molecular organization is critical for elucidating the neural mechanisms underlying movement disorders, cognitive impairments, and neuropsychiatric conditions. To support systematic investigation of basal ganglia biology, we developed an accessible, web-based platform for integrated exploration of single-cell and multi-omics datasets. The platform hosts comprehensive cell-type-resolved datasets encompassing chromatin conformation, DNA methylation, histone modifications, chromatin accessibility, and gene expression, available at https://basalganglia.epigenomes.net/. These data were generated using state-of-the-art single-cell technologies, including snm3C-seq, Paired-Tag, and 10X Multiome, enabling joint analyses of transcriptional and epigenetic regulation at high resolution and scale. In addition to genome browser-based visualization, the platform provides interactive tools for single-cell data exploration, including UMAP embeddings, gene expression and epigenetic distribution plots, and gene set-based heatmaps, allowing users to examine regulatory heterogeneity across cell types, regions, and

modalities. The web portal features an intuitive interface supporting hierarchical cell-type navigation, assay-based filtering, gene- and region-centric queries, and customizable, synchronized visualization panels implemented through the WashU Epigenome Browser. Visualization states can be saved and shared to facilitate reproducible and collaborative research. Furthermore, the platform supports comparative epigenomic analyses across human, macaque, marmoset, and mouse species, leveraging cross-species datasets generated by the BICAN (NIH BRAIN Initiative Cell Atlas Network) consortium. Together, this resource offers a comprehensive and integrated framework for single-cell and comparative epigenomic analysis of the basal ganglia, enabling accelerated discovery of the regulatory mechanisms underlying basal ganglia function and disease.

Introduction

To address the complexity of the human brain, the NIH BRAIN Initiative has established coordinated, large-scale efforts to map brain cell types and their molecular features systematically. The Brain Initiative Cell Atlas Network (BICAN) is a central component of this strategy, aiming to generate a comprehensive, multimodal reference atlas of human and model-organism brain cell types. BICAN integrates transcriptomic, epigenomic, anatomical, and functional measurements across multiple brain regions and species, emphasizing standardized data generation, open data sharing, and cross-consortium interoperability. These principles are intended to enable broad reuse of data resources and to support community-driven discovery (1).

The basal ganglia (BG) serve as a central hub for integrating inputs from the cortex, thalamus, and brainstem, coordinating these signals to guide adaptive motor, cognitive, and affective behaviors. Disruptions within these circuits are associated with a wide range of neurological and psychiatric conditions (2-4). Beyond their translational relevance, the basal ganglia were prioritized by the BICAN consortium due to their conserved structure and functional organization across species, making them an ideal system for cross-species comparative studies and for linking molecular and cellular findings to evolutionary and functional principles.

Advances in single-cell and single-nucleus technologies have been instrumental in enabling these efforts (5). Single-cell RNA sequencing (scRNA-seq) (6) has revealed extensive transcriptional diversity within neuronal and non-neuronal populations, redefining cell-type classification in the brain (7). In parallel, single-cell ATAC-seq has enabled the systematic mapping of chromatin accessibility landscapes, providing insights into cis-regulatory elements and transcription factor networks that underlie cell identity and state (8). Together, these approaches have become foundational tools for constructing cell atlases and for linking molecular profiles to brain function and disease.

More recently, multimodal epigenomic technologies have expanded the scope of single-cell profiling. Single-nucleus methylome and chromatin conformation sequencing

(snm3C-seq) enables joint measurement of DNA methylation and three-dimensional genome organization within individual nuclei, providing an integrated view of epigenetic regulation and higher-order chromatin architecture (9). Paired-Tag technology enables simultaneous profiling of histone modifications and gene expression within the same cells, directly linking chromatin state to transcriptional output (10,11). Such multimodal measurements are essential for understanding how epigenetic mechanisms shape cell-type-specific gene regulation and how these processes vary across brain regions and species.

Epigenetic regulation plays a critical role in basal ganglia development, neuronal subtype specification, and functional specialization. Cell-type- and region-specific patterns of chromatin accessibility, DNA methylation, histone modifications, and chromatin interactions contribute to regulatory programs that are not fully captured by transcriptomic data alone. Integrating these layers is therefore essential for constructing a comprehensive molecular framework of basal ganglia biology and for interpreting disease-associated genetic variation (12-14).

While large-scale efforts such as BICAN have generated extensive multi-omics datasets, effective data exploration and integration remain significant challenges for the broader research community. Web-based platforms that provide standardized access, intuitive navigation, and interactive visualization are critical for maximizing the utility of these resources. By enabling users to explore cell types, genomic features, and regulatory modalities within a unified framework and by supporting comparative analyses across species, such portals facilitate hypothesis generation, cross-study integration, and collaborative research. These capabilities are particularly valuable for advancing a systems-level understanding of basal ganglia function and dysfunction.

Results

Integrated multi-omics basal ganglia resource and data overview

To enable systematic exploration of basal ganglia molecular architecture, we assembled and integrated a comprehensive collection of single-cell and single-nucleus multi-omics datasets spanning major basal ganglia regions based on the BICAN cross-species consensus basal ganglia taxonomy (15). The resource includes transcriptomic, chromatin accessibility, DNA methylation, histone modification, and chromatin conformation data generated using 10X Multiome (scRNA-seq and scATAC-seq), snm3C-seq, and Paired-Tag technologies. These datasets capture both neuronal and non-neuronal cell populations and provide complementary views of gene regulation across multiple epigenetic layers.

All datasets were uniformly processed using standardized pipelines to ensure consistency across assays, regions, and species. Cell-type annotations were harmonized across

modalities using established marker genes and cross-modality integration strategies, enabling coherent comparisons between transcriptomic and epigenomic profiles. Data are organized by consensus cell taxonomy either on subclass (Supplemental Table 1) or group level (Supplemental Table 2), assay modality, and species, supporting both targeted queries and broad exploratory analyses.

To facilitate accessibility and reuse, the integrated datasets are hosted within a web-based portal that provides structured navigation across cell types and assays, as well as gene- and genomic region-centric entry points. Users can interactively explore data through dynamically generated tables and genome browser views, enabling rapid inspection of regulatory features, including chromatin accessibility, DNA methylation, histone modifications, and chromatin interactions, within defined cell populations. The portal provides a genome browser anchored workflow that guides users from cell-type selection to integrated multi-omics visualization (Figure 1). Users first define cell populations using a hierarchical taxonomy, then select assay modalities of interest, after which the corresponding datasets are dynamically rendered in an interactive genome browser environment. This workflow enables coordinated inspection of chromatin accessibility, histone modifications, DNA methylation, transcriptional activity, and regulatory annotations within a unified genomic context, facilitating both exploratory and hypothesis-driven analyses.

Figure 1: Genome browser anchored visualization workflow of the portal. (A) Data exploration begins with cell-type selection using a hierarchical taxonomy presented in the first tab of the home page, allowing selection at both subclass and group levels. (B) In the second tab, users select assay modalities for visualization. All available assays are preselected by default, with search and filtering tools provided to customize the selection. (C) Based on the selected cell types and assays, the portal renders an interactive genome browser view displaying the corresponding datasets. Standard genome browser operations, including zooming, panning, and gene- or coordinate-based searches, support detailed examination of genomic and epigenomic features. ATAC-seq, histone modification, and RNA-seq tracks are displayed using CPM (counts per million) normalization, consistent across all figures. Activity-by-contact (ABC) annotations and interaction links are derived from Wu et al. (submitted). DNA methylation data are shown in methylC format (16), with blue bars representing CpG methylation levels (range 0-1) and gray lines indicating read density.

The platform supports coordinated visualization across multiple datasets by enabling synchronized genomic views, thereby allowing direct comparison of regulatory landscapes across cell types, regions, and modalities. Visualization states can be configured, saved, and shared, facilitating reproducibility and collaborative analysis. In addition, the resource incorporates cross-species datasets generated within the BICAN consortium, enabling comparative analyses between human and non-human primate (NHP) basal ganglia cell types. It supports cross-species comparative epigenomic

analyses by enabling synchronized visualization of homologous genomic regions across multiple species. Using the *CADM1* locus as an illustrative example, users can simultaneously examine chromatin accessibility and regulatory signatures from human, mouse, macaque, and marmoset basal ganglia datasets (Figure 2). *CADM1* encodes a synaptic cell adhesion molecule that plays important roles in neuronal connectivity and synapse formation and has been implicated in neurodevelopmental and neuropsychiatric disorders, making it a biologically relevant locus for comparative analysis (17,18). Alignment of epigenomic tracks across species is achieved using genomeAlign tracks (19) derived from pairwise whole-genome sequence alignments, enabling direct comparison of regulatory features despite differences in genome organization. At the *CADM1* locus, this approach reveals multiple regions of conserved chromatin accessibility that are shared across all four species, consistent with evolutionarily conserved enhancer elements. The ability to identify such conserved regulatory signatures provides insight into core gene regulatory mechanisms underlying basal ganglia function and facilitates prioritization of functionally important non-coding elements. This comparative visualization capability allows users to assess both conserved and species-specific regulatory features within a single interactive framework, supporting evolutionary and translational analyses of basal ganglia gene regulation.

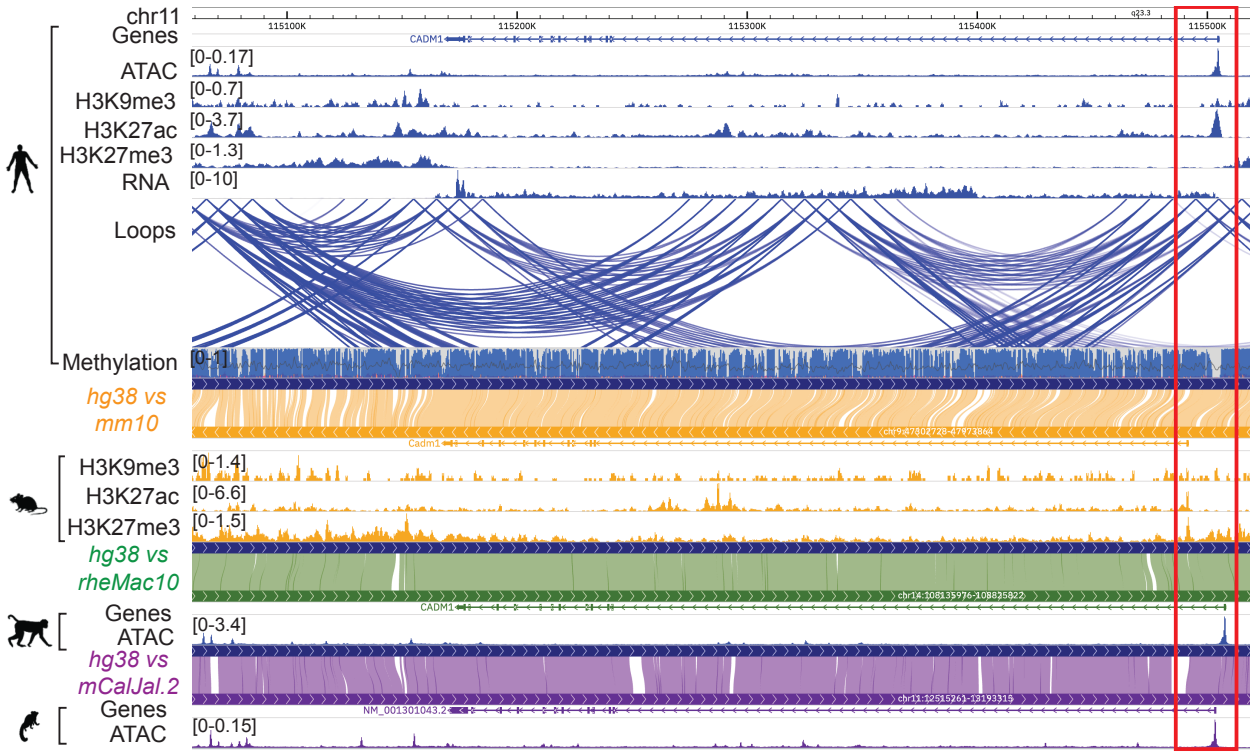


Figure 2: Cross-species comparative epigenomic visualization at the *CADM1* locus. Genome browser view showing epigenomic datasets from human, mouse, macaque, and marmoset aligned at the *CADM1* gene locus. Tracks are color-coded by species: human (blue), mouse (orange), macaque (green), and marmoset (purple). Conserved open chromatin and putative enhancer signatures across species are highlighted, with

red boxes indicating regions exhibiting shared regulatory activity in all four species. Cross-species alignment is enabled by genomeAlign tracks, generated from pairwise whole-genome alignments between each species, allowing homologous regulatory elements to be visualized in a unified coordinate framework.

Together, this integrated data framework establishes a unified entry point for basal ganglia multi-omics data and provides the foundation for downstream analyses of cell-type-specific regulatory mechanisms and disease-relevant genomic features.

Web portal architecture and interactive visualization framework

To support efficient exploration of complex, multimodal datasets, the web-based portal was developed with a modular architecture optimized for scalability, interoperability, and user-driven analysis. The platform integrates curated multi-omics datasets through a structured backend that links cell-type annotations, assay metadata, and genomic coordinates, enabling rapid retrieval and coordinated visualization across modalities. The architecture is designed to accommodate continued data expansion as additional basal ganglia regions, cell types, and assays become available through ongoing consortium and community efforts. The complete list of available data is stored in a TSV (Tab-Separated Values) file that can be updated for future expansion (Supplementary File 1).

The user interface emphasizes intuitive navigation and flexible data access. Users can browse datasets by cell taxonomy, assay modality, or species, or initiate exploration through gene- or genomic region-centric queries. Interactive tables enable filtering, sorting, and expanding metadata, providing transparent access to sample composition, experimental parameters, and cell-type definitions. These entry points are designed to support both hypothesis-driven queries and exploratory analyses.

One central feature of the portal is its integration with the WashU Epigenome Browser (20,21), which enables high-resolution, genome-based visualization of multiple epigenomic and transcriptomic tracks. Users can generate customized visualization panels that combine data across assays, cell types, and regions. Visualization settings, including track selection, scaling, and synchronization state, can be saved and shared with collaborators, supporting reproducibility and collaborative workflows. Comparative epigenomics functionality further enables side-by-side visualization of homologous genomic regions across human and NHPs, leveraging cross-species datasets generated within the BICAN consortium.

In addition to genome browser-based visualization, the portal provides integrated tools for exploring single-cell and single-nucleus analysis results through commonly used summary and dimensionality reduction plots (22). These visualization modules allow users to examine molecular variation across cells, regions, and cell types without requiring external computational workflows. Users can flexibly select datasets, assay modalities, and features of interest, including predefined metadata variables, individual genes, or gene lists, enabling both targeted and exploratory analyses.

As an example, a UMAP visualization of snm3C-seq data can be used to examine cell-type-specific DNA methylation patterns at genes of interest. When coloring UMAP embeddings by normalized CHH methylation (Wu et al., Submitted) levels at the *CADM1* locus, distinct patterns emerge across basal ganglia cell populations, reflecting epigenetic regulation that is not directly apparent from transcriptomic data alone (Figure 3A). The ability to display two UMAP plots side by side further facilitates direct comparison across genes, modalities, or experimental conditions (Supplementary Figure 1). The portal also supports region-level comparisons of gene expression using summary plots. Visualization of *CADM1* expression across basal ganglia regions using Paired-Tag RNA data highlights regional variability in transcriptional output, complementing the epigenomic patterns observed in methylation data (Figure 3B). Such coordinated views enable users to relate epigenetic regulation to transcriptional outcomes within a unified interface. Finally, gene set-based analyses are supported through heatmap visualizations that summarize epigenomic signals across cell classes. For a user-defined gene list, normalized CHH methylation levels can be displayed and grouped by subclass-level taxonomy, enabling identification of cell-type-specific regulatory signatures at scale (Figure 3C). Together, these single-cell visualization tools extend the portal beyond genomic coordinate-centric views and provide a flexible framework for interrogating regulatory and transcriptional heterogeneity in the basal ganglia.

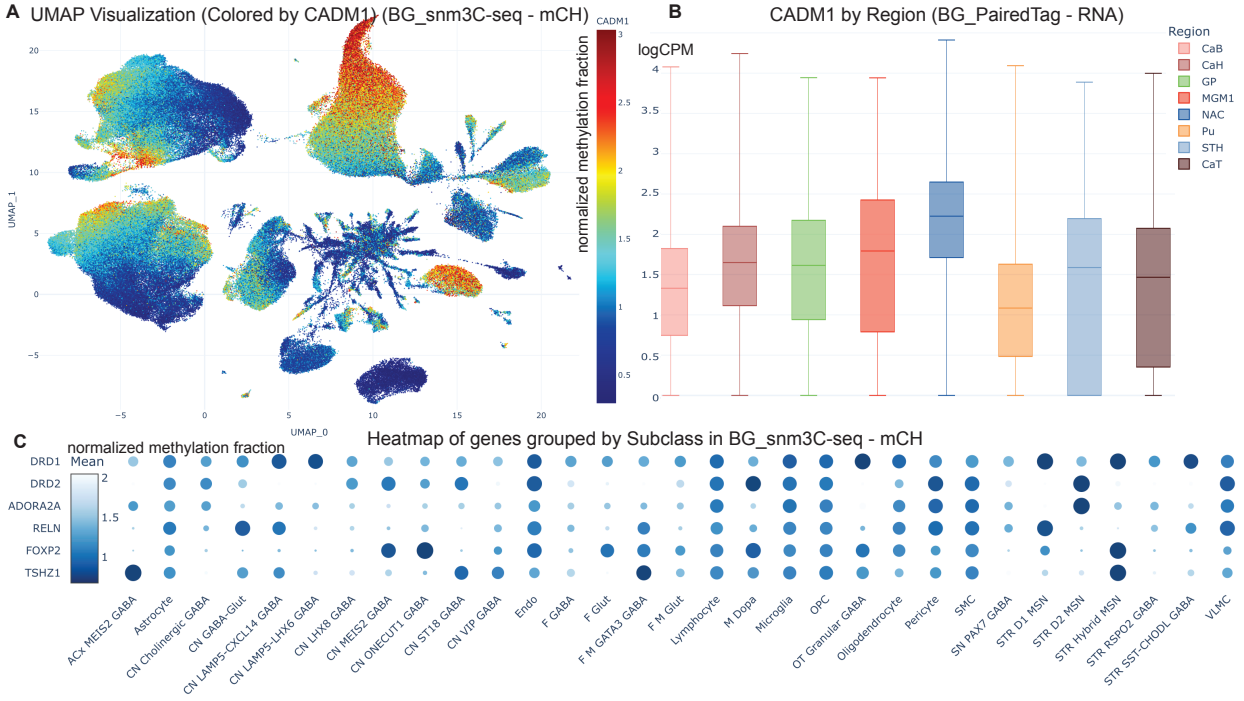


Figure 3: Visualization of single-cell analysis results within the portal. The portal supports interactive visualization of single-cell and single-nucleus analysis results using dimensionality reduction, distribution, and summary plots. Users first select a dataset and assay modality, then generate plots based on predefined variables, individual genes, or gene sets. (A) UMAP projection of single nuclei from snm3C-seq data, colored by CHH DNA methylation levels associated with the *CADM1* gene. The UMAP interface also

supports side-by-side plotting for comparative visualization. (B) Boxplot showing *CADM1* gene expression distributions across basal ganglia regions using RNA expression data from Paired-Tag. (CaB: body of caudate, CaH: head of caudate, GP: globus pallidus, MGM1: gray matter of midbrain, NAC: nucleus accumbens, Pu: putamen, STH: subthalamic nucleus, CaT: tail of caudate, Supplementary file 2 generated by Allen Institute contains other abbreviations used in the portal) (C) Heatmap displaying average CHH methylation levels for a user-defined gene list, grouped by subclass-level cell taxonomy. Rows correspond to genes and columns correspond to cell subclasses.

Example biological insights enabled by integrated basal ganglia multi-omics analyses

To illustrate the utility of the resource for biological discovery, we present representative examples demonstrating how integrated multi-omics visualization across cell types can reveal regulatory features of the basal ganglia. These examples are intended to highlight analytical workflows enabled by the platform rather than to provide a comprehensive biological interpretation.

Cell-type-specific enhancer signature

To illustrate the portal's ability to resolve cell-type-specific regulatory elements, we examined open chromatin accessibility at the *DRD2* locus using 10x Multiome ATAC-seq data (Figure 4). This analysis reveals an enhancer verified by AAV toolbox assay (23) that is selectively accessible in cerebral nuclei (CN) Lateral Ganglionic Eminence (LGE)-derived GABAergic neurons (24), including medium spiny neurons (25), but not in non-neuronal cell populations. The cell-type-restricted accessibility pattern suggests that this enhancer may contribute to the selective regulation of *DRD2* expression in basal ganglia projection neurons. Such examples demonstrate how integrated single-cell chromatin accessibility visualization enables the identification of cell-type-specific candidate regulatory elements, providing mechanistic insight into the transcriptional programs underlying basal ganglia circuit specialization.

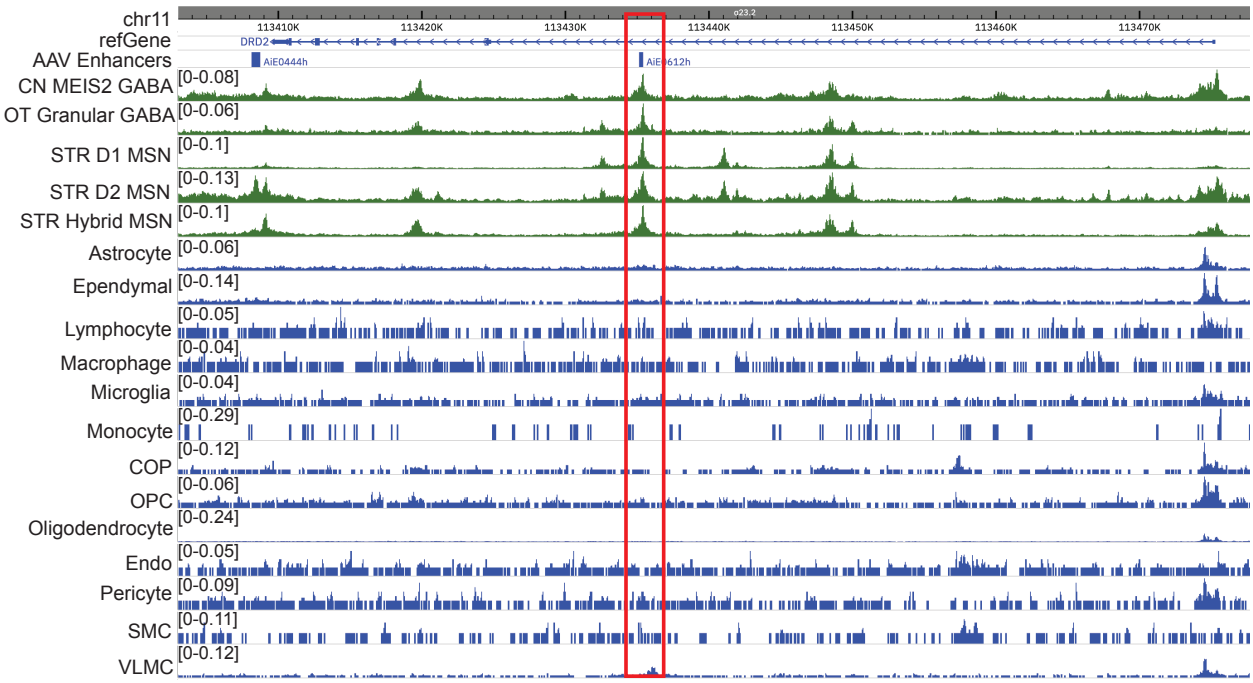


Figure 4: Cell-type-specific open chromatin signatures at a basal ganglia enhancer. Genome browser visualization of open chromatin profiles derived from 10x Multiome ATAC-seq data at the *DRD2* locus on chromosome 11. An enhancer exhibits strong accessibility selectively in CN LGE-derived GABAergic neurons, while remaining inactive in non-neuronal cell types. From top to bottom, tracks display the genomic coordinate ruler, gene annotations including *DRD2*, and annotated AAV enhancer elements shown as blocks beneath the gene model. All tracks represent ATAC-seq signal (CPM), with green tracks corresponding to CN LGE-derived GABAergic neurons and blue tracks representing non-neuronal cells. The red box highlights an enhancer region with enriched chromatin accessibility specifically in medium spiny neurons (MSNs) and two additional GABAergic neuron subtypes.

Integration of chromatin conformation with epigenomic and transcriptional data

To demonstrate how integrated multi-omics data can reveal cell-type-specific regulatory mechanisms, we examined chromatin interactions and transcriptional activity across neuronal and non-neuronal populations using combined snm3C-seq, ATAC-seq, and Paired-Tag datasets (Figure 5). The portal enables simultaneous visualization of chromatin accessibility, histone modifications, three-dimensional chromatin contacts, and gene expression, providing a comprehensive view of regulatory architecture within defined cell types. At the *BCL11B* locus, one of the MSN marker genes (26), striatal hybrid medium spiny neurons exhibit increased ABC (Activity-By-Contact) (27,28) predicted enhancer-promoter interactions (Wu et al., Submitted) accompanied by elevated RNA expression relative to microglia. This illustrates how activity-by-contact modeling, when integrated with chromatin conformation and epigenomic data, can identify putative regulatory links underlying cell-type-specific transcriptional programs. By enabling direct contrast of regulatory interactions and expression across neuronal and non-neuronal cell types, the portal facilitates mechanistic investigation of gene regulation in the basal

ganglia and supports the prioritization of candidate regulatory elements for functional follow-up studies. This also underscores the importance of higher-order chromatin structure in shaping basal ganglia gene expression programs and demonstrates the added value of incorporating chromatin conformation data into integrative analyses.

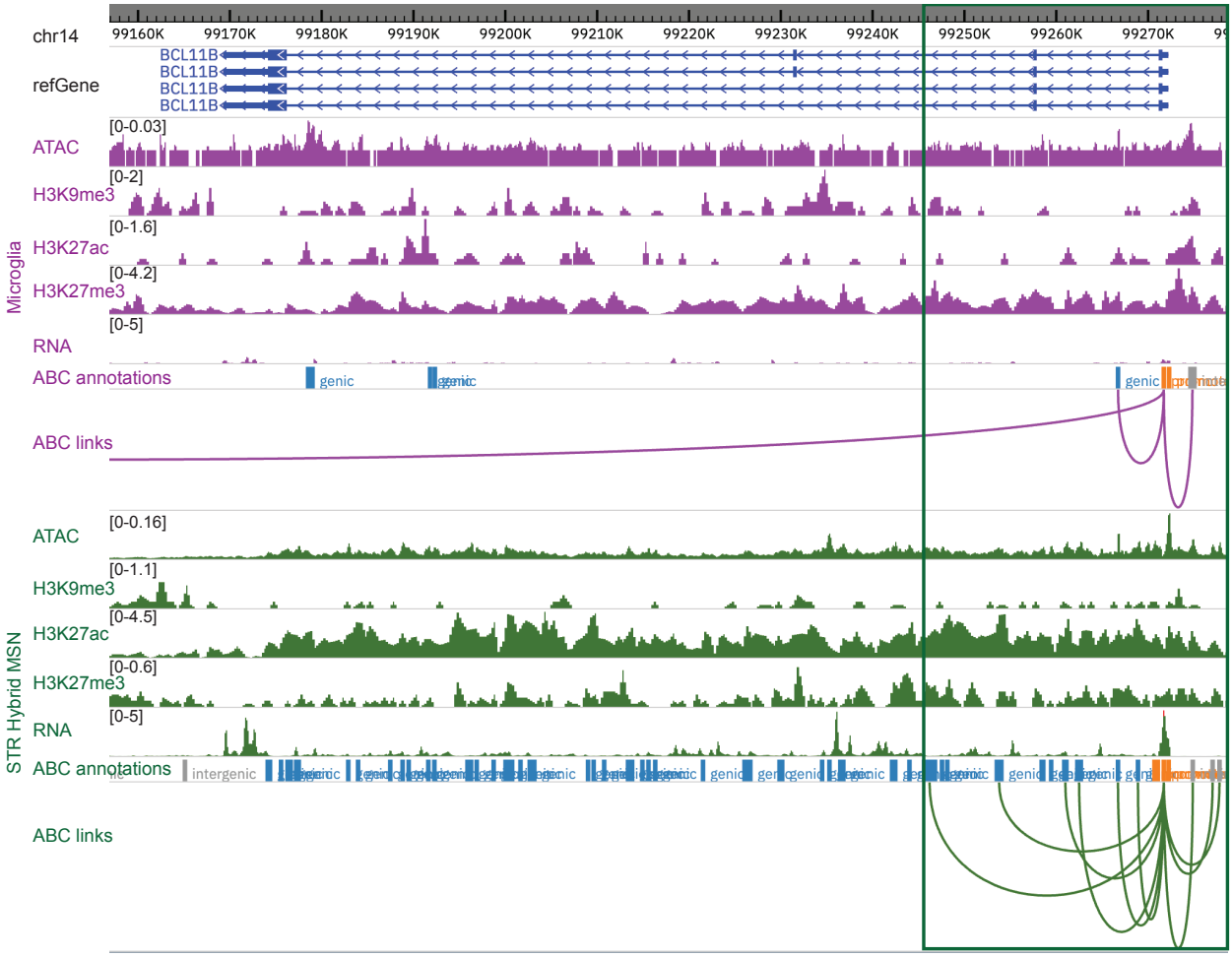


Figure 5: Cell-type-specific chromatin interactions associated with transcriptional differences in neuronal and non-neuronal populations. Genome browser view illustrates coordinated chromatin interaction and transcriptional profiles across the MSN marker gene *BCL11B* locus. Tracks include genomic coordinates, refGene annotations, 10x Multiome ATAC-seq signal, histone modification profiles and RNA expression measured by Paired-Tag, and activity-by-contact (ABC) regulatory links inferred from integrated snm3C-seq and ATAC-seq data. Purple tracks represent microglia, while green tracks represent striatal hybrid medium spiny neurons (STR Hybrid MSNs). The view highlights the *BCL11B* locus, where increased ABC interactions and higher RNA expression are observed in MSNs.

Cross-species conservation and divergence of regulatory elements

We observed that using the portal's comparative epigenomics functionality, users can examine conserved regulatory features between different species (Figure 2). We can also examine the divergence, here using striatal ventral region (STRv) D1 medium spiny neurons (MSNs) as an example, aligned human and mouse datasets reveal distinct

regulatory architectures at the *PLXND1* locus (Figure 6). While mouse D1 MSNs exhibit a pronounced enhancer signature characterized by H3K27ac enrichment accompanied by detectable gene expression, these features are substantially diminished or absent in the corresponding human cell population. Such analyses underscore the importance of human-specific epigenomic data for interpreting regulatory mechanisms in the basal ganglia and caution against direct extrapolation from model organisms without comparative validation.

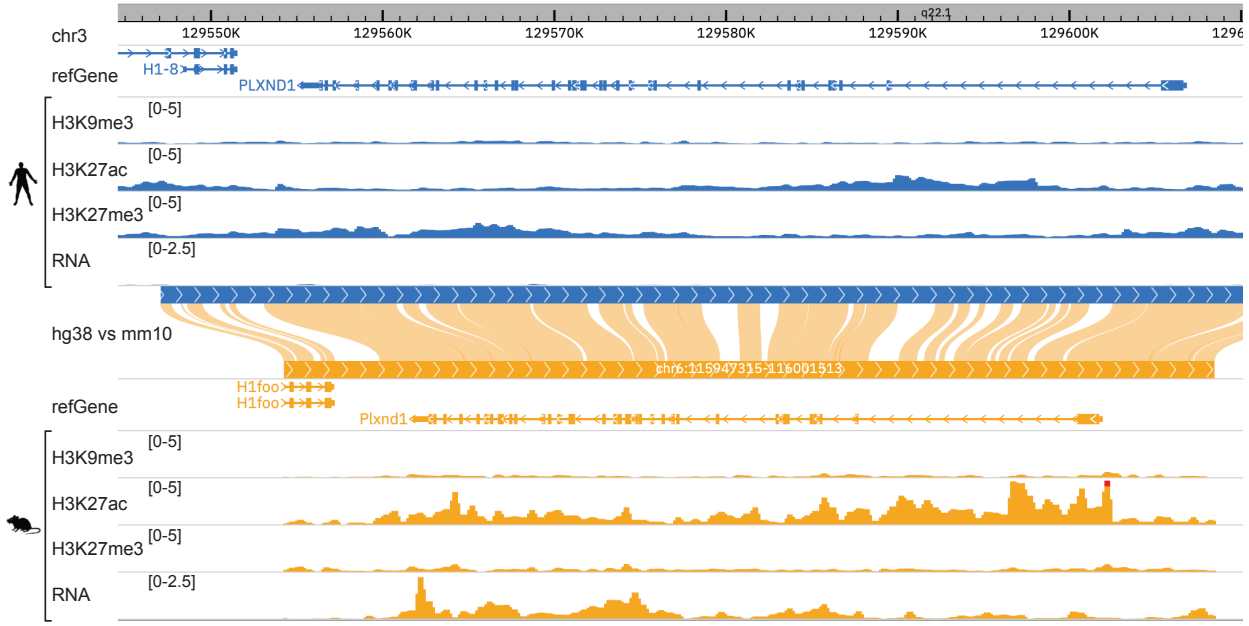


Figure 6: Species-specific epigenomic regulation at the PLXND1 locus in striatal D1 medium spiny neurons. Genome browser view showing aligned human and mouse epigenomic datasets from striatal ventral region (STRv) D1 medium spiny neurons (MSNs) at the *PLXND1* gene locus. Comparative visualization reveals a mouse-specific enhancer signature marked by H3K27ac enrichment and corresponding RNA expression, whereas these regulatory and transcriptional signals are absent or markedly reduced in the human data. This example highlights species-specific differences in epigenomic regulation within homologous basal ganglia cell types.

Discussion

In this study, we present an integrated, web-based resource for exploring the multi-omics regulatory landscape of the basal ganglia at single-cell resolution. By unifying transcriptomic, chromatin accessibility, DNA methylation, histone modification, and chromatin conformation data across cell types, regions, and species, this platform addresses a critical need for accessible and interoperable tools to study the molecular architecture of basal ganglia circuits.

The example analyses presented here illustrate how coordinated visualization of multiple regulatory layers can provide mechanistic insights that are difficult to obtain from individual modalities alone. Cell-type-specific patterns of chromatin accessibility, DNA

methylation, and histone modifications observed at key basal ganglia genes reinforce the concept that epigenetic regulation is a primary driver of neuronal identity and functional specialization. Integration of these epigenomic signals with gene expression data enables users to link regulatory features to transcriptional outcomes directly, facilitating hypothesis generation about gene regulation in defined neuronal populations.

Beyond genome browser-centric views, the platform's single-cell analysis visualizations further enhance interpretation of regulatory heterogeneity. Dimensionality reduction, distribution, and gene set-based summary plots enable users to examine variation in transcriptional and epigenetic features across individual cells, cell types, and brain regions. For example, UMAP visualizations colored by gene-specific methylation or expression values reveal cell-type-specific regulatory patterns that are not readily apparent from aggregated tracks alone, while boxplots and heatmaps summarize region- and subclass-level differences across genes or gene sets. These single-cell views complement coordinate-based visualization by providing a cell-centric perspective on regulatory diversity within the basal ganglia.

Importantly, incorporation of chromatin conformation data from snm3C-seq adds a structural dimension to these analyses, enabling examination of long-range regulatory interactions in a cell-type-specific manner. Combined with single-cell methylation and expression analyses, these data highlight how three-dimensional genome organization, epigenetic state, and transcriptional output jointly shape basal ganglia gene regulation. This integrated framework is particularly relevant for interpreting non-coding regulatory elements and disease-associated variants that act by disrupting regulatory architecture rather than through protein-coding sequences.

Cross-species comparative analyses further underscore the value of this resource. While many regulatory features are conserved between human and mouse basal ganglia cell types, species-specific differences in non-coding regulatory elements are readily apparent. The ability to examine these features both at genomic loci and through single-cell summaries reinforces the importance of human-specific epigenomic data for interpreting genetic risk loci and cautions against over-generalization from model organisms alone. Coordinated visualization of homologous regions across species enables systematic investigation of conserved and divergent regulatory mechanisms.

Beyond biological insights, this work emphasizes the importance of data accessibility and usability. The web portal's modular architecture, interactive visualization framework, and integration with the WashU Epigenome Browser lower technical barriers for researchers with diverse computational backgrounds. Features such as multi-species track views, interactive single-cell plots, and shareable visualization states support reproducible, collaborative exploration and align with the open-science principles of the NIH BRAIN Initiative and BICAN.

As single-cell and multimodal profiling technologies continue to evolve, this resource provides a flexible foundation for incorporating additional datasets, brain regions, and species. Future expansions may include additional data from NHP species and extended single-cell analytical capabilities. Collectively, this platform serves as a community-oriented foundation for advancing our understanding of basal ganglia biology and its role in neurological and psychiatric disease.

Conclusion

In summary, we present an open, interactive platform that integrates single-cell transcriptomic and epigenomic data to enable systematic exploration of the basal ganglia at cellular and regulatory resolution. In addition to genome-centric views, the platform provides dedicated single-cell analysis visualizations that capture regulatory and transcriptional heterogeneity across cell types, regions, and gene sets. By combining complementary modalities with intuitive visualization and comparative tools, this resource lowers barriers to data access and supports integrative investigations of gene regulation, genome organization, and cell-type specialization. As part of BICAN infrastructure studies, this platform serves as a foundational reference for basal ganglia epigenome research and a scalable framework for future expansion to advance our understanding of brain function and disease.

Acknowledgements

We are deeply grateful to the scientists at the Allen Institute who developed the consensus basal ganglia cell taxonomy, and the anonymous donors who contributed the brain samples.

Funding

This work was supported by the NIH BRAIN Initiative Cell Atlas Network (BICAN) fund to the Center for Multiomic Human Brain Cell Atlas (UM1MH130994).

Declaration of interests

J.R.E. is a scientific adviser for Zymo Research Inc. and Ionis Pharmaceuticals. B.R. is a co-founder of Epigenome Technologies and has equity in Arima Genomics Inc. The remaining authors declare no competing interests.

References

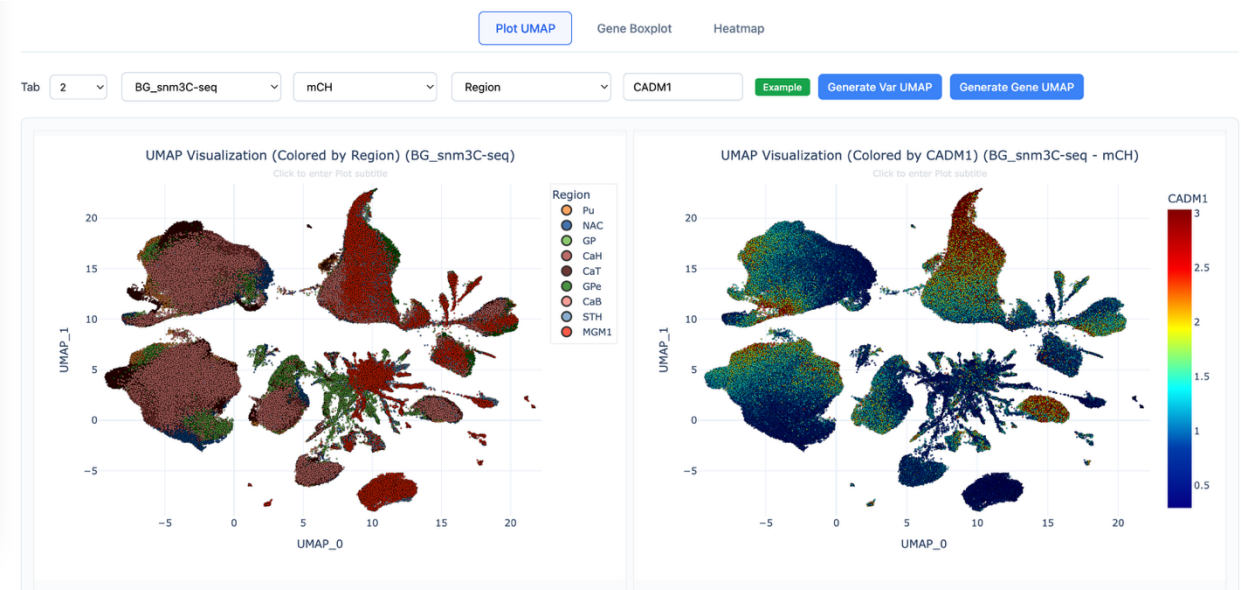
1. Nowakowski, T.J., Nano, P.R., Matho, K.S., Chen, X., Corrigan, E.K., Ding, W., Gao, Y., Heffel, M., Jayakumar, J., Kaplan, H.S. *et al.* (2025) The new frontier in understanding human and mammalian brain development. *Nature*, **647**, 51-59.
2. Graybiel, A.M. (2008) Habits, rituals, and the evaluative brain. *Annu Rev Neurosci*, **31**, 359-387.

3. Haber, S.N. (2016) Corticostriatal circuitry. *Dialogues Clin Neurosci*, **18**, 7-21.
4. Redgrave, P., Prescott, T.J. and Gurney, K. (1999) The basal ganglia: a vertebrate solution to the selection problem? *Neuroscience*, **89**, 1009-1023.
5. Zeng, H. and Sanes, J.R. (2017) Neuronal cell-type classification: challenges, opportunities and the path forward. *Nat Rev Neurosci*, **18**, 530-546.
6. Tang, F., Barbacioru, C., Wang, Y., Nordman, E., Lee, C., Xu, N., Wang, X., Bodeau, J., Tuch, B.B., Siddiqui, A. *et al.* (2009) mRNA-Seq whole-transcriptome analysis of a single cell. *Nat Methods*, **6**, 377-382.
7. Yao, Z., van Velthoven, C.T.J., Kunst, M., Zhang, M., McMillen, D., Lee, C., Jung, W., Goldy, J., Abdelhak, A., Aitken, M. *et al.* (2023) A high-resolution transcriptomic and spatial atlas of cell types in the whole mouse brain. *Nature*, **624**, 317-332.
8. Kim, S.S., Truong, B., Jagadeesh, K., Dey, K.K., Shen, A.Z., Raychaudhuri, S., Kellis, M. and Price, A.L. (2024) Leveraging single-cell ATAC-seq and RNA-seq to identify disease-critical fetal and adult brain cell types. *Nat Commun*, **15**, 563.
9. Liu, H., Zeng, Q., Zhou, J., Bartlett, A., Wang, B.A., Berube, P., Tian, W., Kenworthy, M., Altshul, J., Nery, J.R. *et al.* (2023) Single-cell DNA methylome and 3D multi-omic atlas of the adult mouse brain. *Nature*, **624**, 366-377.
10. Xie, Y., Zhu, C., Wang, Z., Tastemel, M., Chang, L., Li, Y.E. and Ren, B. (2023) Droplet-based single-cell joint profiling of histone modifications and transcriptomes. *Nat Struct Mol Biol*, **30**, 1428-1433.
11. Zhu, C., Zhang, Y., Li, Y.E., Lucero, J., Behrens, M.M. and Ren, B. (2021) Joint profiling of histone modifications and transcriptome in single cells from mouse brain. *Nat Methods*, **18**, 283-292.
12. Ayata, P., Badimon, A., Strasburger, H.J., Duff, M.K., Montgomery, S.E., Loh, Y.E., Ebert, A., Pimenova, A.A., Ramirez, B.R., Chan, A.T. *et al.* (2018) Epigenetic regulation of brain region-specific microglia clearance activity. *Nat Neurosci*, **21**, 1049-1060.
13. Brule, B., Alcala-Vida, R., Penaud, N., Scuto, J., Mounier, C., Seguin, J., Khodaverdian, S.V., Cosquer, B., Birmele, E., Le Gras, S. *et al.* (2025) Accelerated epigenetic aging in Huntington's disease involves polycomb repressive complex 1. *Nat Commun*, **16**, 1550.
14. Zhou, J., Zhang, Z., Wu, M., Liu, H., Pang, Y., Bartlett, A., Peng, Z., Ding, W., Rivkin, A., Lagos, W.N. *et al.* (2023) Brain-wide correspondence of neuronal epigenomics and distant projections. *Nature*, **624**, 355-365.
15. Johansen, N.J., Fu, Y., Schmitz, M., Dubuc, A., Kempynck, N., Wirthlin, M., Garcia, A.D., Hewitt, M., Turner, M.A., Seeman, S.C. *et al.* (2025) Cross-species consensus atlas of the primate basal ganglia. *bioRxiv*.
16. Zhou, X., Li, D., Lowdon, R.F., Costello, J.F. and Wang, T. (2014) methylC Track: visual integration of single-base resolution DNA methylation data on the WashU EpiGenome Browser. *Bioinformatics*, **30**, 2206-2207.
17. Nguyen, H.D. (2025) Elucidate biomarkers and the molecular pathways associated with genetic variants that contribute to the etiology of Parkinson's disease. *Acta Neurol Belg*, **125**, 1621-1641.
18. Wu, X., Azizan, E.A.B., Goodchild, E., Garg, S., Hagiyama, M., Cabrera, C.P., Fernandes-Rosa, F.L., Boulkroun, S., Kuan, J.L., Tiang, Z. *et al.* (2023) Somatic

mutations of CADM1 in aldosterone-producing adenomas and gap junction-dependent regulation of aldosterone production. *Nat Genet*, **55**, 1009-1021.

19. Zhuo, X., Hsu, S., Purushotham, D., Kuntala, P.K., Harrison, J.K., Du, A.Y., Chen, S., Li, D. and Wang, T. (2023) Comparing genomic and epigenomic features across species using the WashU Comparative Epigenome Browser. *Genome Res*, **33**, 824-835.
20. Li, D., Purushotham, D., Harrison, J.K., Hsu, S., Zhuo, X., Fan, C., Liu, S., Xu, V., Chen, S., Xu, J. *et al*. (2022) WashU Epigenome Browser update 2022. *Nucleic Acids Res*, **50**, W774-W781.
21. Seng, C., Liu, S., Zhang, W., Zhuo, X., Li, D. and Wang, T. (2025) WashU Epigenome Browser update 2025. *Nucleic Acids Res*, **53**, W554-W561.
22. McInnes, L.A., John%ASaul, Nathaniel%AGroßberger, Lukas%BJournal Name: Journal of Open Source Software, 3, J.V. and 29, J.I. (2018) UMAP: Uniform Manifold Approximation and Projection. *Journal Name: Journal of Open Source Software; Journal Volume: 3; Journal Issue: 29*, Medium: X.
23. Hunker, A.C., Wirthlin, M.E., Gill, G., Johansen, N.J., Hooper, M., Omstead, V., Vargas, S., Lerma, M.N., Taskin, N., Weed, N. *et al*. (2025) Enhancer AAV toolbox for accessing and perturbing striatal cell types and circuits. *Neuron*, **113**, 1507-1524 e1517.
24. van Velthoven, C.T.J., Gao, Y., Kunst, M., Lee, C., McMillen, D., Chakka, A.B., Casper, T., Clark, M., Chakrabarty, R., Daniel, S. *et al*. (2025) Transcriptomic and spatial organization of telencephalic GABAergic neurons. *Nature*, **647**, 143-156.
25. Berackey, B.T., Tan, Z., Wu, G., Das, S.C., Li, R., Esser, B., Ye, Q., Nafisi, M., Park, S.S., Mendieta, P.A.S. *et al*. (2025) Multiscale Spatial Transcriptomic Atlas of Human Basal Ganglia Cell-Type and Cellular Community Organization. *bioRxiv*.
26. He, J., Kleyman, M., Chen, J., Alikaya, A., Rothenhoefer, K.M., Ozturk, B.E., Wirthlin, M., Bostan, A.C., Fish, K., Byrne, L.C. *et al*. (2021) Transcriptional and anatomical diversity of medium spiny neurons in the primate striatum. *Curr Biol*, **31**, 5473-5486 e5476.
27. Fulco, C.P., Nasser, J., Jones, T.R., Munson, G., Bergman, D.T., Subramanian, V., Grossman, S.R., Anyoha, R., Doughty, B.R., Patwardhan, T.A. *et al*. (2019) Activity-by-contact model of enhancer-promoter regulation from thousands of CRISPR perturbations. *Nat Genet*, **51**, 1664-1669.
28. Gschwind, A.R., Mualim, K.S., Karbalayghareh, A., Sheth, M.U., Dey, K.K., Jagoda, E., Nurtdinov, R.N., Xi, W., Tan, A.S., Jones, H. *et al*. (2023) An encyclopedia of enhancer-gene regulatory interactions in the human genome. *bioRxiv*.

Supplementary Figures



Supplementary Figure1: Multiple UMAP plots in the single cell analysis visualization module allows easier side-by-side comparison.