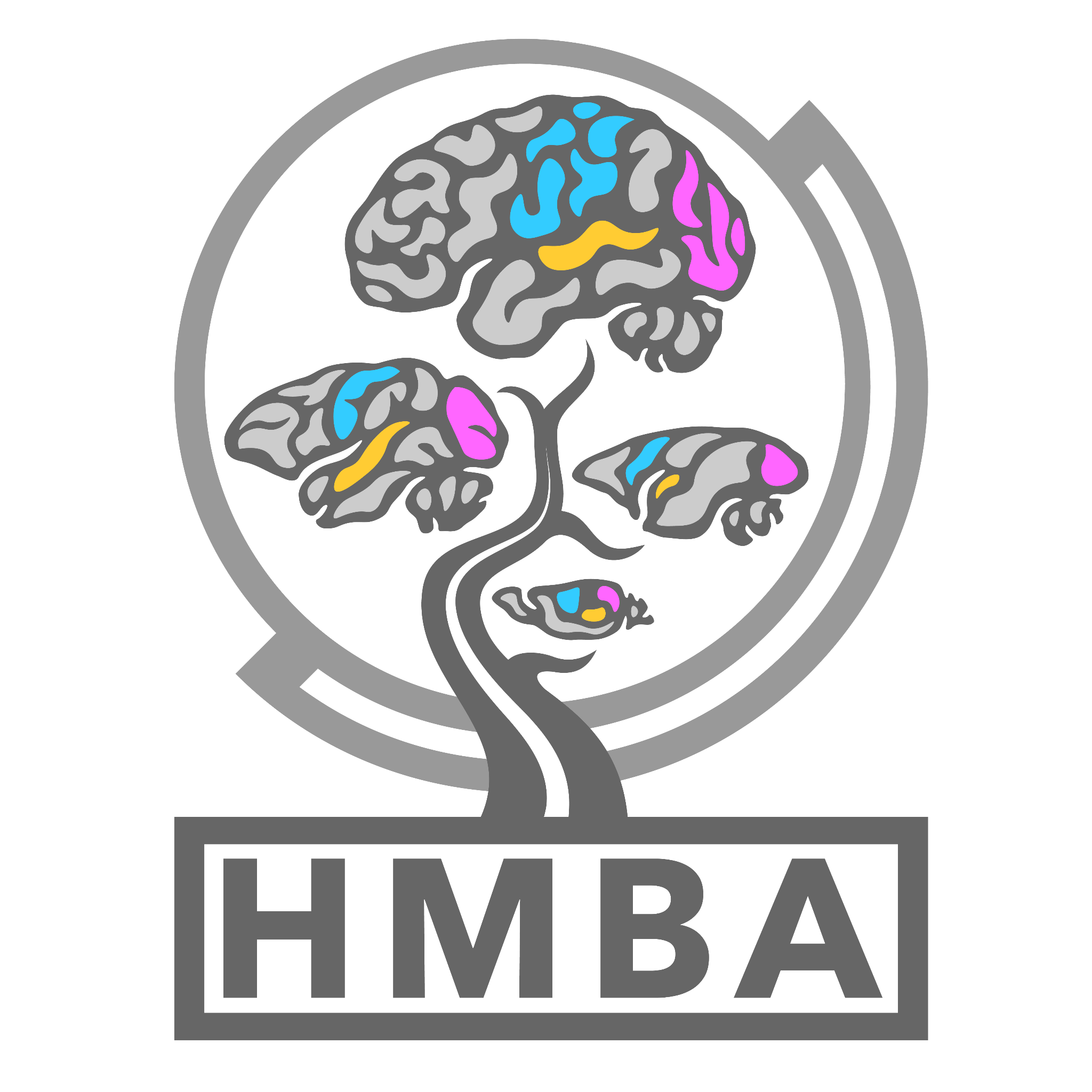
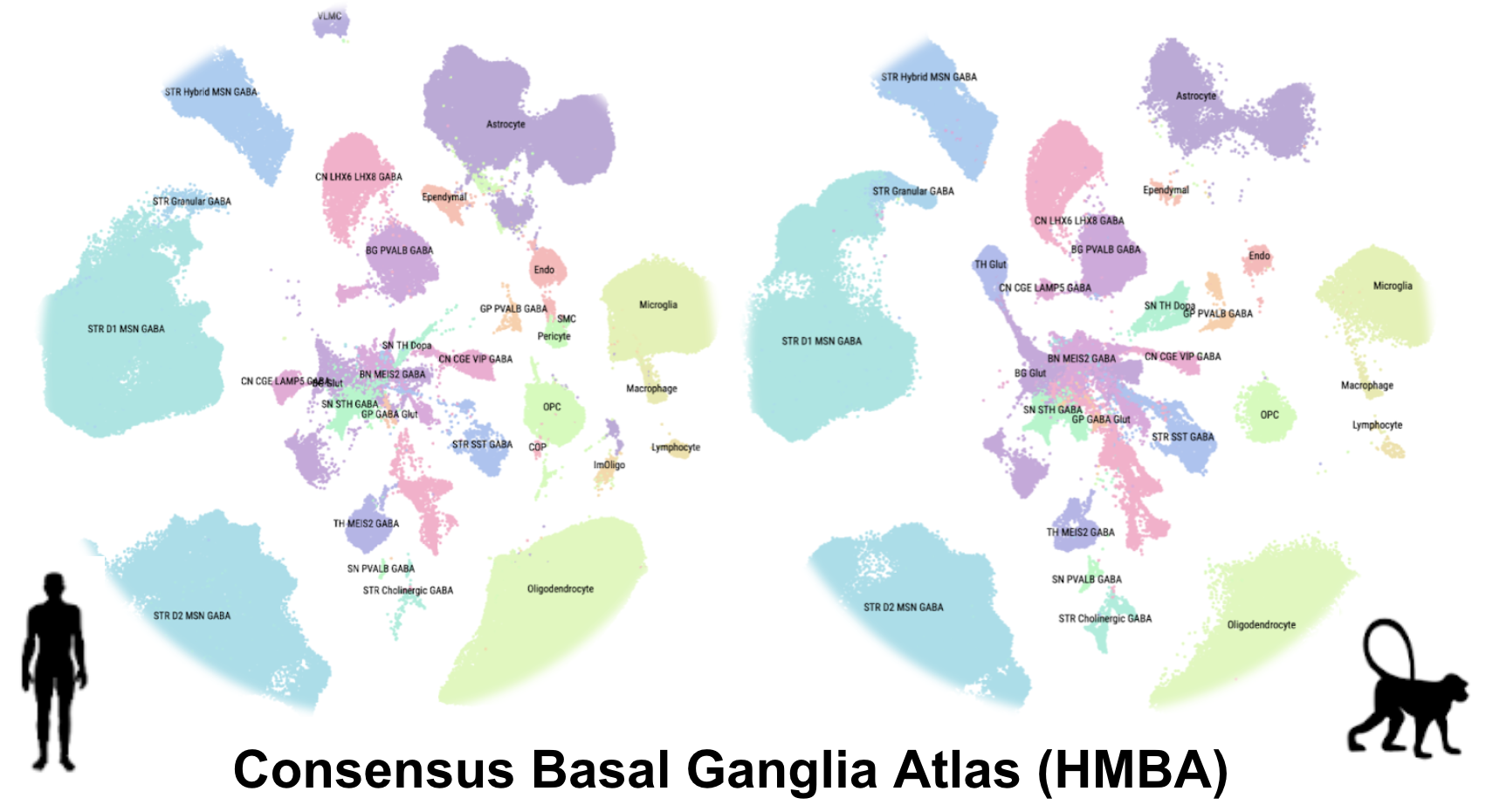
**ARCHIVED - Please edit this [Word version](https://alleninstitute.sharepoint.com/:w:/r/sites/EvoGenTeam-HMBABasalGanglia/_layouts/15/Doc.aspx?sourcedoc=%7B916D2823-8B3B-4C66-A910-A2F853427C50%7D&file=BG%20Taxonomy%20brain-map%20landing%20page.docx&nav=eyJjIjo4NzQ0NTE2Mjl9&action=default&mobileredirect=true)**

HMBA Cross-species Atlas of the Basal Ganglia

[Introduction:](#_dpx4eysyhyzj)

[Understanding the Basal Ganglia](#_zctvu9d0mmvz)

[Human and Mammalian Brain Atlas (HMBA):](#_6r22dagz57v)

[Consensus Cell Type Nomenclature](#_c2dwpxv86pu0)

[Sample and Cell Type Metadata Collection](#_onrl8rd5wu08)

[Data and Tool availability:](#_fboqvxhhtkdn)

[Taxonomies: single nucleus RNA-seq (10X Multiome)](#_4tuxtr3qkn85)

[Data and Cell Type Mapping:](#_nlrd3brjees8)

## Introduction:

### Understanding the Basal Ganglia of

The basal ganglia (BG) are a system of interconnected brain structures that play a crucial role in motor control, learning, behavior, and emotion. With approximately 200 million neurons in the human basal ganglia alone, these structures are involved in a wide range of neurological processes and are implicated in numerous disorders affecting human health, including Parkinson’s disease, Huntington’s disease, and substance abuse disorders. To further understand the complexity of the basal ganglia, researchers have historically classified its neurons into various types based on their cytoarchitecture, connectivity, molecular profile, and functional properties. However, recent advancements in high-throughput transcriptomic profiling have revolutionized our ability to systematically categorize these cell types within species, while the maturation of machine learning technologies have enabled the integration of these taxonomies across species. Through the integration of these advancements, we are now poised to present a unified taxonomy of the mammalian basal ganglia, providing unprecedented insights into the conserved and divergent features of this system across humans, macaques, marmosets, and mice.

## Human and Mammalian Brain Atlas (HMBA)**:**

### Consensus Cell Type Nomenclature

In an effort to unify the field’s understanding of basal ganglia cell types across species, the HMBA has initiated a comprehensive nomenclature system that synthesizes pre-existing literature describing cell type nomenclature with new cross-species transcriptomic data. This initiative integrates HMBA single-nucleus RNA sequencing (snRNA-seq) data from human, macaque, marmoset, and previously published mouse basal ganglia, with the goal of generating a **consensus cell type taxonomy** that can be widely adopted by the scientific community. By focusing on conserved marker genes and shared molecular profiles to supplement established names from the broader community, we have developed a standardized naming system that captures the evolutionary relationships and functional distinctions among basal ganglia cell types. The HMBA consensus basal ganglia taxonomy is designed to streamline communication, foster collaboration, and facilitate the development of novel research tools targeting specific cell types across multiple species.

### Sample and Cell Type Metadata Collection

To support the HMBA consensus cell type taxonomy, we have compiled extensive metadata providing detailed information associated with each identified cell type. This metadata includes information on gene expression patterns and marker genes from snRNA-seq as well as synonymous names from existing literature, including [Yao et al. 2023](https://doi.org/10.1038/s41586-023-06812-z) and [Siletti et al 2023](https://doi.org/10.1126/science.add7046). The metadata fields will be organized under an [aligned taxonomy data format](https://github.com/AllenInstitute/scrattch.taxonomy/tree/main/schema) enabling integration with pre-existing tools such as CellxGene and Taxonomy Development Tools (TDT). This comprehensive metadata repository is available to researchers, offering a valuable resource for exploring the functional diversity and regulatory mechanisms of basal ganglia neurons.

## **Data and Tool availability:**

### Taxonomies: single nucleus RNA-seq (10X Multiome)

Our consensus basal ganglia cell type taxonomy is the result of [iterative clustering](https://github.com/AllenInstitute/transcriptomic_clustering) and cross-species integration of transcriptomic data. The taxonomy encompasses neurons from key structures within the basal ganglia, including the caudate (Ca), putamen (Pu), nucleus accumbens (NAc), the external and internal segments of the globus pallidus (GPe, GPi), subthalamic nucleus (STN), and substantia nigra (SN). By combining data from multiple primate and rodent species, we have developed a consensus taxonomy that highlights both conserved and species-specific cell types. We validate our taxonomy through marker gene expression analysis, comparison with previously published taxonomies, and self-projection, ensuring the accuracy and robustness of each level in the taxonomic hierarchy.

### Data and Cell Type Mapping:

Here we provide an initial version of the HMBA basal ganglia consensus taxonomy annotated onto both human and macaque 10X multiome (RNA-seq) profiling. Within each AIT .h5ad object you will find components necessary for analysis, visualization and cell type mapping using various algorithms via [scrattch.mapping](https://github.com/AllenInstitute/scrattch.mapping) or [cell\_type\_mapper](https://github.com/AllenInstitute/cell_type_mapper).

The AIT .h5ad files are organized according to a [schema](https://github.com/AllenInstitute/scrattch.taxonomy/tree/main/schema):

* **Raw counts:** anndata.raw.X
* **logTPM:** anndata.X
* **Donor corrected latent space (scVI):** anndata.obsm.scVI
* **Donor corrected umap:** anndata.obsm.X\_umap

| **Species** | **File Format** | **Size** | **Data Download** | **Annotation Sheet** |
| --- | --- | --- | --- | --- |
| Human | AIT (.h5ad) | **34GB** | [**AWS S3 URL**](https://released-taxonomies-802451596237-us-west-2.s3.us-west-2.amazonaws.com/HMBA/BasalGanglia/HMBA_Human_BG_082024_AIT.h5ad) | [**Google Sheet**](https://docs.google.com/spreadsheets/d/1P0AnChVqlkCDawg_-ggTY0PIjHvxDyfAmx_p9H4S6PQ/edit?usp=sharing) |
| Macaque | AIT (.h5ad) | **25GB** | [**AWS S3 URL**](https://released-taxonomies-802451596237-us-west-2.s3.us-west-2.amazonaws.com/HMBA/BasalGanglia/HMBA_Macaque_BG_082024_AIT.h5ad) | [**Google Sheet**](https://docs.google.com/spreadsheets/d/1eh_fetTw6rQNmd53n2W6CbmrjgBTJsShqFtewL7VliQ/edit?usp=sharing) |

Additional details about building and organizing Allen Institute Taxonomy (AIT) files can be found here: [scrattch.taxonomy](https://github.com/AllenInstitute/scrattch.taxonomy).

For tutorials showcasing how to use AIT files easily with basic and MapMyCells cell type mapping algorithms please visit our [tutorial](https://github.com/AllenInstitute/scrattch.mapping/blob/main/examples/mapping_BasalGanglia.md) on [scrattch.mapping](https://github.com/AllenInstitute/scrattch.mapping).