

Name: Dorzhey Matkhanov, M.S.

Email: dmatkhanov@unmc.edu

Adaptive cross-validation single-cell annotation enables cross-species cell-type transfer and neuroimmune responses in HIV/SIV models

Accurate cell-type annotation is a prerequisite for genetic and epigenetic analyses in complex neuroimmune single-cell datasets, yet performance is highly sensitive to reference composition, disease context, and tissue- or species-specific transcriptomic structure. We developed a reproducible, explainable, and adaptable pipeline for automated annotation with uncertainty quantification and Cell Ontology-based verification and applied it to Single Cell Opioid Responses in the Context of HIV (SCORCH) Consortium human brain data and complementary rhesus macaque datasets relevant to addiction and HIV comorbidity. The framework integrates ensemble label transfer (popV) with consensus-derived confidence scores, flexible reference selection matched to query tissue, species, and health condition, and an orthogonal auditing layer using Cello, a hierarchical Cell Ontology classifier designed to avoid marker-gene circularity.

Reference construction combined atlas-scale brain datasets, a microglia-focused atlas, and manually annotated samples to ensure representation of HIV-associated neuroimmune states; gene identifier harmonization and standardized minimal QC enabled consistent processing across sources. To obtain defensible performance estimates and demonstrate relaunch stability, we implemented 5-fold cross-validation across reference sources, yielding highly consistent metrics and confusion-heatmap structure across reruns and across clinically relevant strata, including HIV-focused subsets and encephalitis-relevant sample groups. For downstream neuroimmune analyses supporting epigenetics study design, we used the pipeline to extract a high-recall CNS macrophage set, refined it to high-specificity microglia using Cello-guided clustering, and performed conservative differential expression and microglial state scoring as a foundation for cell-type resolved molecular readouts.

We further established training and inference workflows for cross-species transfer, demonstrating strong out-of-the-box human-to-monkey and monkey-to-human performance with expected trade-offs when incorporating other-species labels, enabling comparative studies in model systems. Finally, we applied the framework to compare SIV-associated transcriptional perturbations between spleen myeloid cells and brain macrophage populations under saline versus morphine exposure, quantifying concordant and discordant gene-level responses that motivate follow-up integration with epigenetic modalities. Collectively, this work provides a practical annotation-to-inference blueprint that can support collaboration on cell-type-resolved genetic and epigenetic mechanisms of opioid exposure and HIV/SIV-related neuroimmune comorbidities.

Authors:

Submitter: Dorzhey Matkhanov, Bioinformatics Research Analyst, University of Nebraska Medical Center, Omaha

Howard Fox, PhD, M.D., University of Nebraska Medical Center

Niu Meng, PhD, University of Nebraska Medical Center