Comparative Analysis of AAV Serotypes for Transduction of Olfactory Sensory Neurons

Benjamin D.W. Belfort*[1,2,3,4,5], Johnathan D. Jia*[3,5,6], Alexandra Garza^[1,2,3,5], Anthony M. Insalaco^[2,3], JP McGinnis[3,7], Brandon T. Pekarek[1,2,3], Joshua Ortiz-Guzman[3], Burak Tepe[2,3], Hu Chen[3,5,6], aSCENT-PD Investigators, Zhandong Liu[3,5,6], Benjamin R. Arenkiel°[2,3,5,8]

1 Genetics and Genomics Graduate Program, Baylor College of Medicine, Houston, Texas 77030, USA.
2 Department of Molecular and Human Genetics, Baylor College of Medicine, Houston, Texas 77030, USA.
3 Jan and Dan Duncan Neurological Research Institute, Texas Children's Hospital, Houston, Texas 77030, USA
4 Medical Scientist Training Program, Baylor College of Medicine, Houston, Texas 77030, USA.
5 Aligning Science Across Parkinson's (ASAP) Collaborative Research Network, Chevy Chase, MD 20815.
6 Department of Pediatrics, Section of Neurology, Baylor College of Medicine, Houston, Texas 77030, USA
7 Department of Neurosurgery, Baylor College of Medicine, Houston, Texas 77030, USA
8 Department of Neuroscience, Baylor College of Medicine, Houston, Texas 77030, USA.
*,^ These authors contributed equally

° Corresponding author: Benjamin R. Arenkiel

Email: Arenkiel@bcm.edu

Author Contributions: BDWB and BRA- project conceptualization; JJ and BDWB- bioinformatic analyses; BDWB, AG, AMI- performed research; JPM, JOG- provided essential resources; BDWB, JJ, AG, AMI- performed analyses; BDWB, JJ- wrote the paper; BP, BT, HC- provided feedback and technical support.

Collaborators: aSCENT-PD Investigators: Benjamin Arenkiel, Zhandong Liu, Brit Mollenhauer, Josef Penninger, Maxime Rousseaux, Armen Saghatelyan, Natalina Salmaso, Christine Stadelmann, Michael G Schlossmacher, Julianna J Tomlinson, John M Woulfe

Competing Interest Statement: Disclose any competing interests here.

Classification: Neuroscience

Keywords: AAV, gene therapy, olfaction, olfactory sensory neurons

This file includes:

Main Text
Figures 1 to 2
Table 1

Abstract

Olfactory sensory neurons within the nasal epithelium detect volatile odorants and relay odor information to the central nervous system. Unlike other sensory inputs, olfactory sensory neurons interface with the external environment and project their axons directly into the central nervous system. The use of adeno-associated viruses to target these neurons has garnered interest for applications in gene therapy, probing olfactory sensory neuron biology, and modeling disease. To date, there is no consensus on the optimal AAV serotype for efficient and selective transduction of olfactory sensory neurons *in vivo*. Here we utilized serial confocal imaging and single-nucleus RNA sequencing to evaluate the efficacy of 11 different AAV serotypes in transducing murine olfactory sensory neurons via non-invasive nasal inoculation. Our results reveal that AAV1, while highly effective, exhibited broad tropism, whereas AAV-DJ/8 showed the greatest specificity for olfactory sensory neurons.

Main Text

Introduction

The mammalian olfactory epithelium (OE) is a specialized epithelial tissue within the nose that plays a critical role in the sense of smell. Olfactory sensory neurons (OSNs) are specialized chemosensory neurons that reside in the OE and detect airborne chemicals through G-protein coupled odorant receptors. OSNs project their axons directly into the olfactory bulb (OB), creating a single-neuron chain between the external environment and the central nervous system (CNS). This conduit into the CNS provides unique opportunities for probing OSN biology, administering targeted gene therapies, and modeling disease. The use of AAVs has been explored for these purposes [1], yet information is lacking on the optimal serotype for transducing OSNs. In the present study, we leveraged confocal imaging and single-nucleus RNA sequencing (snRNAseq) to identify optimal AAV serotypes for *in vivo* transduction of murine OSNs.

Results

To test the efficacy of AAV transduction across commonly used serotypes, we packaged an identical construct (rAAV-EF1a-TdTomato-WPRE-PolyA) into 11 AAV different capsid serotypes with reported OE and neuronal transduction[1]. Serotypes included AAV1, AAV2, AAV5, AAV7, AAV8, AAV9, AAV-DJ/8 [2,3] AAV-PhP.eB [4], AAV-PhP.S [4], AAV-rh10 [5,6], and AAV-SCH9 [7]. Each AAV serotype was introduced via non-invasive nasal inoculation (NINI) [8] into three male mice per serotype, for a total of 33 experimental animals. After a four-week expression period, the OB was processed to quantify colocalized olfactory marker protein (OMP) and TdTomato expression within OSN axon terminals as a readout for the number of transduced OSNs [Fig. 1A]. By first focusing on the reporter expression in the glomerular layer of the OB, we avoided potentially confounding TdTomato signals from other transduced accessory cell types within the OE [Fig. 1B].

For quantification, we performed confocal image analysis across OSN terminal fields within the glomerular layer via immunostaining, measuring levels of tdTomato reporter expression alongside OMP. Importantly, we found no difference in total OMP area between samples from each serotype [Fig. 1C], suggesting that transduction itself did not compromise overall OSN viability or OE integrity. We next quantified the amount of TdTomato surface area (normalized to OMP surface area) per section, and found that AAV1, AAV7, AAV-DJ/8, and AAV-rh10 displayed the greatest levels of expression [Fig. 1D] in OSN terminals that project to the olfactory bulb.

Following this primary reporter expression screen, we next performed snRNAseq to more accurately quantify OSN transduction efficiency of the top 4 candidates in the different OE cell types. Towards this, we designed 4 new AAV expression constructs, all identical aside from a short barcode sequence ("ID") that identifies the corresponding AAV serotype upon transduction and snRNA sequencing. The 4 AAVs were mixed at equal titer and introduced into wildtype mice via NINI. After 4 weeks of *in vivo* expression, OE was harvested for snRNAseq [Fig. 2A]. Data were processed following standard scRNAseq guidelines [9]. Using 20,081 high-quality nuclei, we generated an annotated UMAP from known cell type-specific markers (see SI appendix) to quantify the distribution of cell types [Fig. 2B-D] and identify those

that were differentially transduced by the four candidate AAV serotypes. We confirmed cell type lineages by comparing terminal fate probabilities from Palantir pseudotime analysis, with Horizontal Basal Cells (HBCs) designated as the starting point and mature OSN (mOSN), mature sustentacular (mSus), microvillar (MVC), and Bowman's gland (BowG) cells as terminal fates [10] [Fig. E-F]. A 2-sided t-test ($\alpha = 0.05$) and chi-squared goodness of fit indicated that the terminal fate probabilities between GBCs and iSus cells were different for mOSN and mSus lineages [Fig. 2G]. Finally, we quantified the number and type of cells that were effectively transduced for each AAV serotype [Fig. 2H]. In total, we identified 382 cells that harbored AAV barcodes. While mOSN transduction efficiency was highest for AAV1 [Fig. 2I], AAV-DJ/8 showed the greatest specificity for mOSNs [Fig. 2J]. Further examination of AAV1 expression showed broad tropism, with immature sustentacular (iSus) and airway ciliated cells (ACCs) contributing to the highest normalized cell counts [Fig. 2K].

Discussion

In the present study, we leveraged confocal imaging and snRNAseq to identify optimal AAV serotypes for *in vivo* OSN transduction via NINI. Despite variability in our imaging dataset, attributable to unpredictable flow of fluid through the nasal turbinates during NINI, AAV1, AAV7, AAV-DJ/8, and AAV-rh10 displayed the greatest efficacy in their ability to transduce OSNs compared to other AAV serotypes we tested [Fig. 1D]. To determine which serotype was most efficient at transducing OSNs in a more quantitative, higher-resolution, and cell type-specific manner, we employed snRNA sequencing from isolated OE. Though this is the first study to employ snRNAseq of the OE, we found that clustering and cell type distribution were similar to other single-cell sequencing studies previously performed on the OE [11,12, 13] [Fig. 2B-D]. We found that cell subsets comprised known lineages of interest, namely HBCs, globose basal cells (GBCs), immature OSNs (iOSNs), mOSNs, iSus, mSus, MVCs, and BowGs, and determined the corresponding pseudotemporal lineage patterns [Fig. 2E]. Through this, we substantiated cell type identities by comparing their terminal fate probabilities. Of interest, the terminal fate probability distribution of HBCs properly reflected multiple cell types, indicating their pluripotent ability to divide into all lineages of interest, but favoring mOSNs, mSus, and MVCs [Fig. 2F]. By comparing the terminal fates of iSus and GBC, we further validated the identified OSN and sustentacular cell lineages. Having confirmed the cellular composition of the transduced tissue, we next quantified the number of AAV+ cells within each cell type cluster [Fig. 2H]. Of the serotypes tested, AAV1 showed the highest transduction efficiency of OSNs [Fig. 2I], but AAV-DJ/8, AAV7, and AAV-rh10 all showed greater OSN specificity, with AAV-DJ/8 showing the greatest specificity compared to AAV1 [Fig. 2J]. From these data, we also found that AAV1 showed the broadest tropism in the OE; transducing cell types such as HBCs, GBCs, ACCs, and particularly iSus, all with greater efficiency than OSNs [Fig. 2K]. Alongside a census of the cell types that comprise the mouse olfactory epithelium, together these findings also present useful implications for future research, including OSN-targeted gene therapy, probing OSN biology, modeling diseases where the OE and OB are loci of interest (e.g., Parkinson's disease and Alzheimer's disease), and informing the development of future AAV capsids for targeting OSNs.

Materials and Methods

Key Resources Table

Reagent type (species) or resource	Designation	Source or reference	Identifiers	Additional information
Strain, strain background (*Mus musculus, male*)	C57BL/6J	The Jackson Laboratory	RRID:IMSR_JAX:00066	8-9 weeks old for single serotype inoculations, 10 weeks old for combined serotype inoculations

Chemical compound, drug	OptiPrep	Millipore Sigma	D1556-250ML	Also known as iodixanol
Chemical compound, drug	OptiSeal 16x67 mm tubes	Beckman Coulter	Cat# 362181	Loaded with the iodixanol gradient
Other	NVT 65 Near-Vertical Rotor	Beckman Coulter	Cat# 362755	Centrifuged at 60,000 rpm for 90 minutes
Other	Amicon Ultra-15 Centrifugal Filter	Millipore Sigma	UFC910024	For OptiPrep removal and AAV concentration
Commercial assay, kit	qPCR AAV Titer Kit	Applied Biological Materials	Cat# G931	For viral titer
Other	AAV production	This paper	dx.doi.org/10.17504/protocols.io.81wgbzwj3gpk/v1	For production of the 11 serotypes described
Recombinant DNA reagent	rAAV-Ef1α-TdTomato-WPRE-polyA	This paper		Packaged into the 11 AAV serotypes described
Recombinant DNA reagent	rAAV-Ef1α-EGFP-WPRE-ID-PolyA	This paper		AAV1 ID: CGACGCCTTGTGGATTTTCGTTTTA AAV7 ID: CGACGATTTTCGTTTTACCTTGTGG AAV-DJ/8 ID: CGACGCGTTTTACCTTGTGGATTTTA AVrh10 ID: CGACGGATTTTCGTTTTACCTTGTG
Chemical compound	16% PFA diluted in PBS	Electron Microscopy Sciences	Cat# 15710	Diluted to 4% in 30 mL PBS
Chemical compound	Tissue-Plus O.C.T. Compound	Fisher Scientific	Cat# 23-730-571	

Other	CM1860 Cryostat Microtome	Leica Biosystems		40 um sections
Antibody	anti-OMP (Goat, polyclonal)	FUJIFILM Wako	Cat# 544-10001-WAKO	Immunofluorescence (1:20,000)
Antibody	anti-Goat Alexa Fluor 488 (Donkey, polyclonal)	Invitrogen	Cat# A-11055	Immunofluorescence (1:1000)
Chemical compound, drug	Hoechst	Thermo Fisher Scientific	Cat# 62249	Immunofluorescence (1:1000)
Chemical compound, drug	Fluoromount-G	SouthernBiotech	Cat# 0100-01	Tissue mounting media
Other	TCS SP8 confocal Microscope	Leica Biosystems		Equipped with 20x/.75 objective
Software, algorithm	LAS X	Leica Biosystems	RRID:SCR_013673	Imaging software for the Confocal Microscope
Software, algorithm	Imaris v10.0	Bitplane	RRID:SCR_007370	Analysis of OB sections
Other	Image analysis	This paper	dx.doi.org/10.17504/protocols.io.n2bvjn1rxgk5/v1	For quantifying TdTomato expression in OB axon terminals
Chemical compound, drug	Nuclei Extraction Buffer	Miltenyi Biotec	Cat# 130-128-024	Lysis buffer
Chemical compound, drug	Protector RNAse Inhibitor	Millipore Sigma	Cat# 3335402001	
Other	gentleMACS C tubes	Miltenyi Biotec	Cat# 130-093-237	
Other	gentleMACS Octo Dissociator	Miltenyi Biotec	Cat# 130-096-427	
Other	MACS SmartStrainers 70um	Miltenyi Biotec	Cat# 130-098-462	

Other	MACS SmartStrainers 30um	Miltenyi Biotec	Cat# 130-098-458	
Commercial assay, kit	Chromium Single Cell 3' Reagent Kits User Guide (v3.1 Chemistry)	10X Genomics	PN-1000121	Single-cell library preparation
Software, algorithm	STAR v2.7.11b	Thomas Gingeras Lab	RRID:SCR_004463	Align paired-end snRNAseq FASTQ files
Software, algorithm	STARSolo	Alexander Dobin Lab	RRID:SCR_021542	Align paired-end snRNAseq FASTQ files
Ssoftware, algorithm	GRCm39 vM34	GENCODE	RRID:SCR_014966	Custom FASTA and GTF file creation
Software, algorithm	CellBender v0.3.0	Broad Institute	https://doi.org/10.1038/s41592-023-01943-7	Empty droplet filtering
Software, algorithm	Python v3.10.14	The Python Software Foundation	RRID:SCR_008394	
Software, algorithm	Scanpy v1.10.1	Fabian Theis Lab	RRID:SCR_018139, https://github.com/scverse/scanpy	Processing and QC of single-cell data
Software, algorithm	Solo (scvi-tools v1.0.4)	Nir Yosef Lab	https://github.com/scverse/scvi-tools	Doublet removal
Software, algorithm	rapids single-cell v0.10.5	Nir Yosef Lab	https://rapids-singlecell.readthedocs.io	GPU-acceleration of Scanpy processing steps
Software, algorithm	pyMDE v0.1.18	Minimum-distortion embedding, 2021	https://pymde.org	MDE generation
Software, algorithm	Palantir v1.3.3	Dana Pe'er Lab	https://github.com/dpeerlab/Palantir	Pseduotime determinations
Software, algorithm	RNA analysis code	This paper	https://doi.org/10.5281/zenodo.13620762	Generated in house

Animals

All mice used in this study were male C57BL/6J (Jax :000664) and were used in compliance with Baylor College of Medicine IACUC. For single serotype inoculations, mice were 8 to 9 weeks old, and for combined serotype inoculations, mice were 10 weeks old.

Plasmid Constructs and Nasal Inoculation

For all nasal inoculations, mice were briefly anesthetized with isoflurane. While under anesthesia, the mice were nasally inoculated in 5 ul doses for a total of 40 ul (alternating between nostrils, with 20 ul of virus per nostril).

For inoculations of individual serotypes, rAAV-Ef1α-TdTomato-WPRE-polyA was packaged in serotype AAV1, AAV2, AAV5, AAV7, AAV8, AAV9, AAV-DJ/8, AAV-PHP.eB, AAV-PHP.S, AAVrh10, or AAV-SCH9 and were normalized to a final concentration of $4.85e^{11}$ vg/mL. 3 mice were inoculated with each individual serotype, for a total of 33 mice.
For mixed inoculations, unique variants of an rAAV-Ef1α-EGFP-WPRE-ID-PolyA construct were packaged in AAV1, AAV7, AAV-DJ/8, and AAVrh10, with unique barcodes (ID) corresponding to each serotype. The AAVs were combined in equal concentrations (a total of $6.10e^{9}$ vg per serotype). The IDs are listed below:
AAV1 ID: CGACGCCTTGTGGATTTTCGTTTTA
AAV7 ID: CGACGATTTTCGTTTTACCTTGTGG
AAV-DJ/8 ID: CGACGCGTTTTACCTTGTGGATTTT
AAVrh10 ID: CGACGGATTTTCGTTTTACCTTGTG

Tissue Processing for Immunofluorescence
Four weeks after nasal inoculation with individual AAV serotypes, the 33 mice were deeply anesthetized with isoflurane and transcardially perfused with 10mL of ice-cold PBS followed by 10mL of 4% paraformaldehyde (16% PFA diluted in PBS; Electron Microscopy Sciences 15710). The olfactory bulbs (OBs) were then harvested, fixed in 4% PFA at 4°C overnight, and moved to 30% sucrose (36 hours) for cryoprotection. OBs were frozen in OCT (Fisher HealthCare cat# 23-730-571) and sectioned at 40um on a cryostat (Leica CM1860), with every third section collected for analysis. Tissue sections were blocked in 5% Donkey Serum (1X PBS, 0.1% Triton) for one hour at room temperature, washed three times in PBST (0.1% Triton) and incubated with anti-OMP (1:20,000, FUJIFILM Wako Chemicals U.S.A. Corporation cat# 544-10001-WAKO) antibody at 4°C overnight. Sections were then washed three times with PBST (0.1% Triton), incubated with secondary antibody (anti-goat, Alexa Fluor 488 Invitrogen cat# A-11055) for 1 hour at room temperature, washed three times with PBST, incubated with Hoechst (1:1000, Thermo Fisher cat# 62249) for 15 minutes at room temperature, and finally washed one more time with PBST before being covered with Fluoromount-G (Southern Biotech cat# 0100-01) and sealed for microscopy.

Microscopy and Image analysis
OB images were taken using a Leica TCS SP8 confocal microscope equipped with a 20x/.75 objective and Leica LAS X software (RRID:SCR_013673, https://www.leica-microsystems.com/products/microscope-software/details/product/leica-las-x-ls). To obtain the entire volume of each OB section, Z-stacks of ~30-40um per sample were collected at 1024x1024 resolution with 4x line averaging and 2x frame averaging. All images were stitched within the LAS X software. OB images were analyzed using Imaris software (v10.0, RRID:SCR_007370, http://www.bitplane.com/imaris/imaris). For each image, a region of interest was drawn over the glomerular layer (excluding the olfactory nerve layer) and was used to generate a mask. Volumetric surfaces were then generated following absolute intensity thresholding to account for per-sample background fluorescence signal. To quantify transduction efficiency of each serotype, the total area (um^2) of TdTomato signal and OMP signal within the mask were quantified. The ratio of TdTomato signal to OMP signal was then calculated. An average of 23.3 OB sections were imaged for each of the 33 mice, for a total of 760 confocal images.
To blind the researchers to which serotype was being analyzed, image files were automatically renamed and randomly sorted into separate folders using a program generated in house. The files were converted back into their identifiable names after all Imaris-based quantification was completed.
For a more comprehensive guide on how to analyze images in this manner, please refer to the following protocols.io DOI: dx.doi.org/10.17504/protocols.io.n2bvjn1rxgk5/v1

To access the complete imaging dataset, please refer to “Comparative Analysis of AAV Serotypes for Transduction of Olfactory Sensory Neurons”, accession number S-BIAD1370 (DOI: 10.6019/S-BIAD1370), on BioImage Archive.

Tissue harvest and Nuclei Isolation

Four weeks after nasal inoculation with AAV1, AAV7, AAV-DJ/8, and AAVrh10 (mixed together in equal proportions, final concentration of 1.22e11 vg/ml for each serotype), 4 male C57BL/6J mice were deeply anesthetized and transcardially perfused with ice-cold PBS. The OE was then rapidly collected. Immediately following dissection, samples were cut into small pieces and processed using GentleMACS nuclei isolation protocol (Nuclei Extraction Buffer [Miltenyi Biotec, cat# 130-128-024], Protector RNAse Inhibitor [Millipore Sigma, cat# 3335402001], GentleMACS C tubes [Miltenyi Biotec, cat# 130-093-237], GentleMACS Octo Dissociator [Miltenyi Biotec, cat# 130-096-427], MACS SmartStrainers 70um [Miltenyi Biotec, cat# 130-098-462], MACS SmartStrainers 30um [Miltenyi Biotec, cat# 130-098-458]). In brief, samples were placed in 2mL of Miltenyi Nuclei Isolation Buffer and Protector RNAse Inhibitor in GentleMACS C tubes. Samples then underwent the preprogrammed “nuclei isolation” program in a GentleMACS Octo Dissociator. Immediately after dissociation, samples were strained through a 70um MACS SmartStrainer and collected in a 15ml tube and centrifuged at 500 x g for 5 minutes at 4°C. The supernatant was extracted and discarded, and the resulting pellet resuspended in 1mL of ice-cold PBS. Resuspended samples were then run through a 30um MACS SmartStrainer. Upon visual inspection of nuclei following isolation, it was determined that debris levels were low enough for samples to proceed immediately to library preparation.

snRNAseq Bioinformatic Analysis

Alignment and Quantification of Transcripts

4 sets of paired-end snRNAseq FASTQ files, each corresponding with a technical replicate, were aligned and quantified using STAR v2.7.11b (RRID:SCR_004463, http://code.google.com/p/rna-star/) and STARSolo (RRID:SCR_021542, https://github.com/alexdobin/STAR/blob/master/docs/STARsolo.md)[14]. For STARSolo to generate the spliced/unspliced count matrix, we used the velocyto flag available in the pipeline. All downstream steps were performed using the spliced count matrix unless specified otherwise. To include the AAV serotypes in the alignment, we generated a custom reference FASTA and GTF file using the GENCODE (RRID:SCR_014966, https://www.gencodegenes.org) GRCm39 vM34 release[15].

Ambient RNA Correction

After alignment and quantification, raw spliced, unspliced, and ambiguous count matrices were generated for each sample. To reduce the effects of ambient RNA contamination and filter out empty droplets, we applied CellBender v0.3.0 (https://github.com/broadinstitute/CellBender) to each replicate’s corresponding spliced/mature RNA count matrix[16]. The resulting corrected matrix was transformed into a Scanpy object for further analysis in Python v3.10.14 (RRID:SCR_008394, https://www.python.org).

Processing, Quality Control, and Integration of snRNAseq data

Each replicate was processed and filtered according to the standard processing guidelines from Scanpy v1.10.1 (RRID:SCR_018139, https://github.com/theislab/scanpy)[17]. Initial thresholds of minimum 10 cells and 200 genes were set to remove any empty droplets missed by CellBender. Gene/feature initial cutoffs were left more lenient to avoid any possibility of filtering out the AAV genes in case low transduction and expression in the dataset. We set feature/gene and UMI barcode cutoffs to remove dead, low-quality cells, and doublets/multiplets. The cutoffs were set as: 1500 for number of features, 2000 for total counts, and 1% for percent mitochondrial counts. Any cells with gene, count, and mitochondrial count

percentages above these thresholds were removed from the dataset. After filtering, samples 1, 2, 3, and 4 had 5810, 6065, 5480, and 5820 cells, respectively.

Doublet removal was performed using the raw count data using Solo from scvi-tools v1.0.4 (https://github.com/scverse/scvi-tools)[18,19]. Filtering, including doublet removal, was performed on each individual sample and after integration of the count matrices. For data integration and batch effect removal, we used the single-cell Variational Inference (scVI) model from scvi-tools. Briefly, the model learns a nonlinear mapping between the latent space and the parameters of a zero-inflated negative binomial distribution used to generate gene expression counts. Batch correction is performed by including batch annotations as an input to the decoder network, allowing the model to learn batch-specific effects that can be removed when sampling from the latent space. Finally, the standard Scanpy processing pipeline including library log-normalization, selection of highly variable genes (N=3500), dimensional reduction with PCA and UMAP, and clustering was applied to the integrated data. We did not use scVI to regress out covariates due to the likelihood of over-correction in the batch-effect removal step. The resulting cell x gene matrix was 20081 x 3500 (24415 raw). When possible, computational speeds were accelerated using an NVIDIA A30 GPU and the rapids single-cell library v0.10.5 (https://rapids-singlecell.readthedocs.io), a library for GPU-acceleration of Scanpy processing steps[20].

Classification of AAV+ Cells

For identification of AAV+ cells, we set a threshold based off the log-normalized expression. Any cell i is considered positive for AAV gene j if

$$X^{log}_{ij} > 10^{-5} \text{ where } i \in \{0,1,2,\ldots,N\} \text{ and } j \in \{0,1,2,3\}$$

Where N is the total number of cells minus one and j corresponds to each of the 4 AAV serotypes. By stratifying according to the cell types (Celltype Annotation), we were able to find the cell-specific frequency distribution of all AAV serotypes in the snRNAseq dataset.

Cell Type Annotation

Using the latent representation generated by scVI after integration, a uniform manifold approximation projection (UMAP) was generated. Clusters were generated using the Leiden algorithm at multiple resolutions to capture different size populations and cell type specific markers were used to confirm the identities. Using canonical markers derived from the literature, we identified and annotated each cluster according to the cell type that it belonged to (see table below). For clusters that were seemingly distinct but expressed the same cell type clusters, we included a separate metadata column to annotate the clusters with an additional identifier (i.e., mSus and mSus2).

Table 1. Cell Types and Cell Type Markers

Cell Type	Markers
Airway ciliated cell (ACC)	Foxj1, Cdhr3, Cdhr4
Bowman's gland (BowG)	Muc5b, Bpifa1, Aqp5, Sox9, Bpifb9b, Bpifb9a
Globose basal cell (GBC)	Sox11, Kit, Neurod1, Neurog1, Hes6, Ascl1
Goblet cell (Gob)	Agr2, Ltf, Spdef

Horizontal basal cell (HBC)	Krt5, Krt14, Krt17, Trp63
Immune cells (Immune)	Cd19, Cd74
Lymphatic epithelium (LE)	Cc21a
Microvillar cell (MVC)	Ascl3, Cftr, Hepacam2, Lcn11
Mesenchymal cell (Mes)	Alx1, Tfap2b, Carmn, Col1a1, Col1a2, Pdgfra, Vim, Runx2, Pdgfrb
Mature olfactory ensheathing cell (mOEC)	Alx3, Ptprz1, Ptn, Plp1, Mmd2
Immature olfactory sensory neuron (iOSN)	Gap43, Neurod1, Sox11, Dcx, Lhx2, Gng8, Tubb3, Stmn1, Stmn2
Mature olfactory sensory neuron (mOSN)	Ano2, Olfm1, Chga, Kcnk10, Cnga2, S100a5, Gnai
Respiratory epithelium (RE)	Bpifb4
Solitary chemosensory neuron (SCC)	Trpm5
Immature sustentacular cell (iSus)	Bpifa1, Mybl1, Lgr5
Mature sustentacular cell (mSus)	Muc2, Cyp2g1, SLc2a3, Ackr3
Undefined secretory cell (UnSec)	Tmsb4x, Scgb1c1, S100a5,
Vascular endothelial cell (VE)	Vwf, Flt1

Pseudotime and Terminal Fate Probability

We selected only relevant lineages of interest for downstream analysis and confirmation, specifically OSN, Sus, MVC, and BowG lineages. The relevant cell types include HBC, GBC, iOSN, mOSN, iSus, mSus, MVC, and BowG. After subsetting, the nearest neighbor graph was re-calculated again to find the relationships between the remaining clusters. To improve the visualization of our dataset, we used the minimum distortion embedding (MDE) technique available in pyMDE v0.1.18 (https://pymde.org) to generate a new embedding space[21].

To confirm the identities of our cell type annotations, we determined the pseudotime relationships between clusters using Palantir[22] v1.3.3 (https://github.com/dpeerlab/Palantir). We specified HBC as the root cell while defining mOSN, mSus, MVC, and BowG as the terminal points for Palantir. We then examined the Palantir-generated terminal fate probabilities of HBC, GBC, and iSus to confirm the annotated cluster identities. The terminal fate probabilities were evaluated with a 2-sided t-test for each individual lineage and a chi-squared goodness of fit (GOF) to evaluate the total terminal fates with $\alpha=0.05$.

Resource Availability

All code used in Bioinformatic Analysis is publicly available on Github at https://github.com/LiuzLab/Mouse-AAV-OSN. Raw count matrices and the final processed AnnData

object are available with the Github code on Zenodo (https://doi.org/10.5281/zenodo.13620762). Raw FASTQ files are available on the CRN Cloud.
All images collected for this manuscript are publicly available through BioImage Archive (accession S-BIAD1370, DOI: 10.6019/S-BIAD1370)
CC BY: Comparative Analysis of AAV Serotypes for Transduction of Olfactory Sensory Neurons © 2024 by Benjamin Belfort is licensed under CC BY 4.0. To view a copy of this license, visit https://creativecommons.org/licenses/by/4.0/

For extended materials and methods regarding AAV production and snRNAseq library preparation and sample submission, please refer to the supplemental information.

Acknowledgments
This research was funded in whole or in part by Aligning Science Across Parkinson's [10000400] through the Michael J. Fox Foundation for Parkinson's Research (MJFF). For the purpose of open access, the author has applied a CC BY public copyright license to all Author Accepted Manuscripts arising from this submission. Special thanks to the Texas Children's Hospital Neuroconnectivity Core for assistance with viral packaging and Baylor College of Medicine's Single Cell Genomics Core for assistance with snRNAseq library preparation.

References

1. Gadenstaetter AJ, Schmutzler L, Grimm D, Landegger LD. Intranasal application of adeno-associated viruses: a systematic review. Transl Res. 2022 Oct;248:87-110. doi: 10.1016/j.trsl.2022.05.002. Epub 2022 May 18. PMID: 35597541.
2. Grimm D, Lee JS, Wang L, Desai T, Akache B, Storm TA, Kay MA. In vitro and in vivo gene therapy vector evolution via multispecies interbreeding and retargeting of adeno-associated viruses. J Virol. 2008 Jun;82(12):5887-911. doi: 10.1128/JVI.00254-08. Epub 2008 Apr 9. PMID: 18400866; PMCID: PMC2395137.
3. Hammond SL, Leek AN, Richman EH, Tjalkens RB. Cellular selectivity of AAV serotypes for gene delivery in neurons and astrocytes by neonatal intracerebroventricular injection. PLoS One. 2017 Dec 15;12(12):e0188830. doi: 10.1371/journal.pone.0188830. PMID: 29244806; PMCID: PMC5731760.
4. Chan KY, Jang MJ, Yoo BB, Greenbaum A, Ravi N, Wu WL, Sánchez-Guardado L, Lois C, Mazmanian SK, Deverman BE, Gradinaru V. Engineered AAVs for efficient noninvasive gene delivery to the central and peripheral nervous systems. Nat Neurosci. 2017 Aug;20(8):1172-1179. doi: 10.1038/nn.4593. Epub 2017 Jun 26. PMID: 28671695; PMCID: PMC5529245.
5. Cearley CN, Wolfe JH. Transduction characteristics of adeno-associated virus vectors expressing cap serotypes 7, 8, 9, and Rh10 in the mouse brain. Mol Ther. 2006 Mar;13(3):528-37. doi: 10.1016/j.ymthe.2005.11.015. Epub 2006 Jan 18. PMID: 16413228.
6. Gao GP, Alvira MR, Wang L, Calcedo R, Johnston J, Wilson JM. Novel adeno-associated viruses from rhesus monkeys as vectors for human gene therapy. Proc Natl Acad Sci U S A. 2002 Sep 3;99(18):11854-9. doi: 10.1073/pnas.182412299. Epub 2002 Aug 21. PMID: 12192090; PMCID: PMC129358.
7. Ojala DS, Sun S, Santiago-Ortiz JL, Shapiro MG, Romero PA, Schaffer DV. In Vivo Selection of a Computationally Designed SCHEMA AAV Library Yields a Novel Variant for Infection of Adult Neural Stem Cells in the SVZ. Mol Ther. 2018 Jan 3;26(1):304-319. doi: 10.1016/j.ymthe.2017.09.006. Epub 2017 Sep 8. PMID: 28988711; PMCID: PMC5762983.
8. Santry LA, Ingrao JC, Yu DL, de Jong JG, van Lieshout LP, Wood GA, Wootton SK. AAV vector distribution in the mouse respiratory tract following four different methods of administration. BMC Biotechnol. 2017 May 15;17(1):43. doi: 10.1186/s12896-017-0365-2. PMID: 28506256; PMCID: PMC5433059.
9. Wolf FA, Angerer P, Theis FJ. SCANPY: large-scale single-cell gene expression data analysis. Genome Biol. 2018 Feb 6;19(1):15. doi: 10.1186/s13059-017-1382-0. PMID: 29409532; PMCID: PMC5802054.
10. Setty M, Kiseliovas V, Levine J, Gayoso A, Mazutis L, Pe'er D. Characterization of cell fate probabilities in single-cell data with Palantir. Nat Biotechnol. 2019 Apr;37(4):451-460. doi: 10.1038/s41587-019-0068-4. Epub 2019 Mar 21. Erratum in: Nat Biotechnol. 2019 Oct;37(10):1237. doi: 10.1038/s41587-019-0282-0. PMID: 30899105; PMCID: PMC7549125.
11. Brann DH, Tsukahara T, Weinreb C, Lipovsek M, Van den Berge K, Gong B, Chance R, Macaulay IC, Chou HJ, Fletcher RB, Das D, Street K, de Bezieux HR, Choi YG, Risso D, Dudoit S, Purdom E, Mill J, Hachem RA, Matsunami H, Logan DW, Goldstein BJ, Grubb MS, Ngai J, Datta SR. Non-neuronal expression of SARS-CoV-2 entry genes in the olfactory system suggests mechanisms underlying COVID-19-associated anosmia. Sci Adv. 2020 Jul 31;6(31):eabc5801. doi: 10.1126/sciadv.abc5801. Epub 2020 Jul 24. PMID: 32937591; PMCID: PMC10715684.
12. Fletcher RB, Das D, Gadye L, Street KN, Baudhuin A, Wagner A, Cole MB, Flores Q, Choi YG, Yosef N, Purdom E, Dudoit S, Risso D, Ngai J. Deconstructing Olfactory Stem Cell Trajectories at Single-Cell Resolution. Cell Stem Cell. 2017 Jun 1;20(6):817-830.e8. doi: 10.1016/j.stem.2017.04.003. Epub 2017 May 11. PMID: 28506465; PMCID: PMC5484588.
13. Li W, Wu T, Zhu K, Ba G, Liu J, Zhou P, Li S, Wang L, Liu H, Ren W, Yu H, Yu Y. A single-cell transcriptomic census of mammalian olfactory epithelium aging. Dev Cell. 2024 Aug 17:S1534-5807(24)00482-9. doi: 10.1016/j.devcel.2024.07.020. Epub ahead of print. PMID: 39173624.
14. Kaminow B, Yunusov D, Dobin A. STARsolo: accurate, fast and versatile mapping/quantification of single-cell and single-nucleus RNA-seq data. doi:10.1101/2021.05.05.442755
15. Frankish A, Carbonell-Sala S, Diekhans M, et al. GENCODE: reference annotation for the human and mouse genomes in 2023. *Nucleic Acids Res*. 2023;51(D1):D942-D949. doi:10.1093/nar/gkac1071

16. Fleming SJ, Chaffin MD, Arduini A, et al. Unsupervised removal of systematic background noise from droplet-based single-cell experiments using CellBender. *Nat Methods*. 2023;20(9):1323-1335. doi:10.1038/s41592-023-01943-7
17. Wolf FA, Angerer P, Theis FJ. SCANPY: Large-scale single-cell gene expression data analysis. *Genome Biol*. 2018;19(1). doi:10.1186/s13059-017-1382-0
18. Bernstein NJ, Fong NL, Lam I, et al. Solo: Doublet Identification in Single-Cell RNA-Seq via Semi-Supervised Deep Learning. *Cell Syst*. 2020;11(1):95-101.e5. doi:10.1016/j.cels.2020.05.010
19. Gayoso A, Lopez R, Xing G, et al. A Python library for probabilistic analysis of single-cell omics data. *Nat Biotechnol*. 2022;40(2):163-166. doi:10.1038/s41587-021-01206-w
20. Avantika L, Corey NJ, Rajesh I, et al. (2023). NVIDIA-Genomics-Research/rapids-single-cell-examples: v2022.12.0 (v2022.12.0). *Zenodo*. doi:10.5281/zenodo.7566335
21. Agrawal A, Ali A, Boyd S. Minimum-Distortion Embedding. *Foundations and Trends in Machine Learning*. 2021;14(3):211-378. doi:10.1561/2200000090
22. Setty M, Kiseliovas V, Levine J, et al. Characterization of cell fate probabilities in single-cell data with Palantir. *Nat Biotechnol*. 2019;37(4):451-460. doi:10.1038/s41587-019-0068-4

Figures

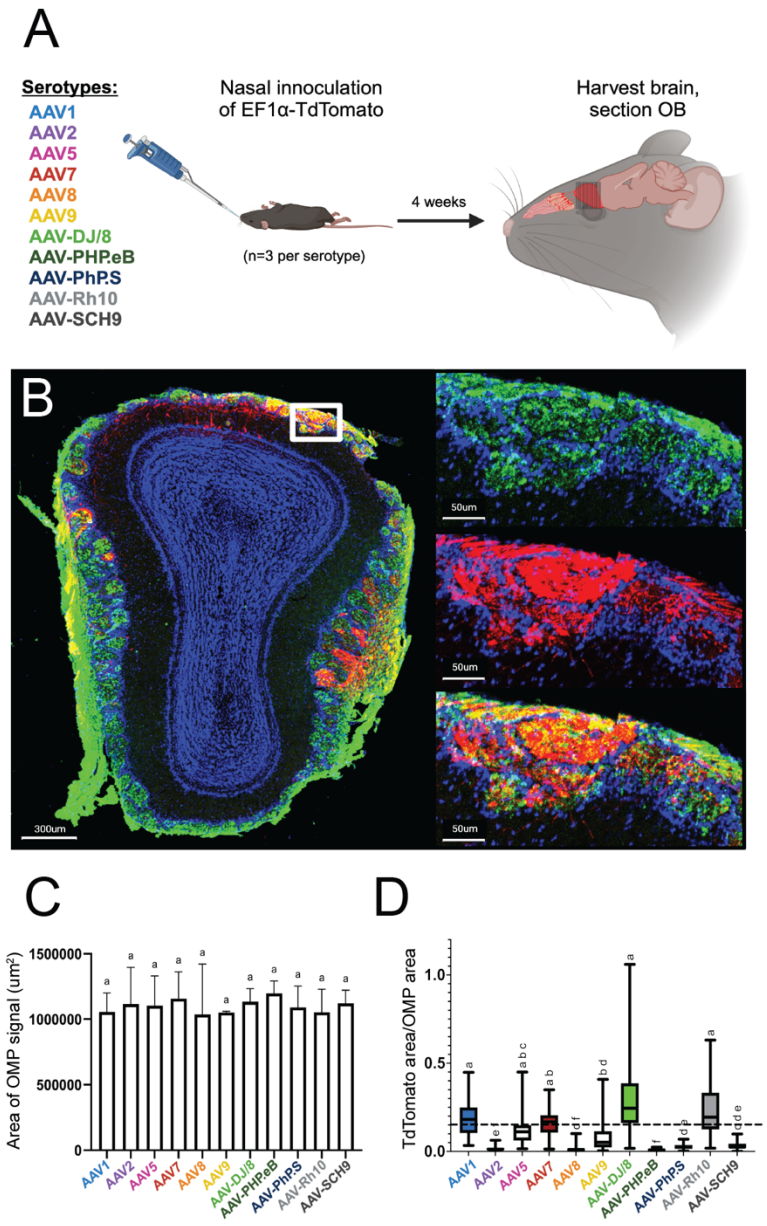

Figure 1. Confocal imaging identifies AAV1, AAV7, AAV-DJ/8, and AAV-rh10 as efficient AAV serotypes for OSN transduction. A) Overview of experimental paradigm. Identical plasmid constructs were packaged in 11 unique AAV serotypes, and each serotype was individually introduced via NINI. B) Representative confocal image of OB cross-section from an AAV1 inoculated mouse. Upper right: close-up of OSNs forming the structures of the glomerular and olfactory nerve layers of the OB, with OMP (green) channel isolated. Middle right: close-up of the same region, with TdTomato (red) channel isolated. Bottom right: overlay of OMP and TdTomato channels. Blue= Hoechst. C) Total area of OMP signal per OB for each serotype. D) Quantification of the percentage of TdTomato/OMP in the glomerular layer per OB section. Error

bars represent the minimum and maximum values. Statistical differences were assessed using one-way ANOVA, the results of which are represented by compact letter display.

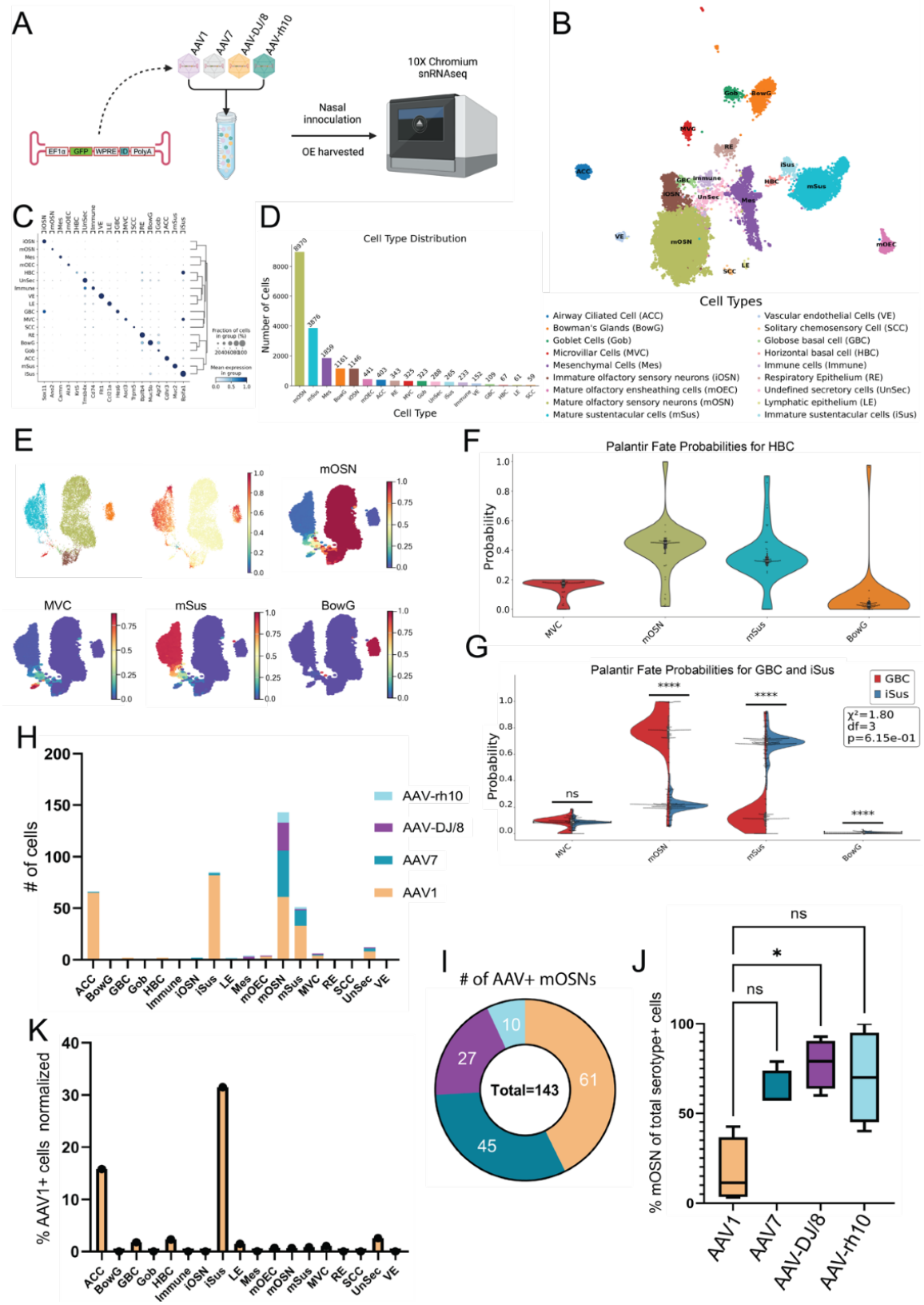

Figure 2. Single nucleus transcriptomics shows AAV1 has the greatest transduction rate while AAV-DJ/8 has the greatest specificity for mOSNs. A) Experimental workflow. B) UMAP

of snRNAseq dataset. C) Dot plot showing the cell-type specific markers used for annotations. The x-axis is the marker, and the y-axis is the cell type. D) Cell type distribution with total cell counts. The x-axis is the cell type with each bar's color reflecting the corresponding color on the UMAP, and the y-axis is the absolute count of the respective cell types. E) Minimum distortion embedding (MDE) with our specific lineages of interest including mOSN, mSus, MVC, and BowG lineages. The total OE cell type-specific pseudotime is shown above and individual lineage pseudotimes shown below, with red indicating late pseudotime and dark blue indicating early pseudotime. F) Violin-box plot of terminal fate probability of HBCs calculated from the Palantir pseudotime. G) Terminal fate probabilities of GBCs and iSus were compared using a 2-sided t-test ($\alpha = 0.05$) and chi-squared goodness of fit for each possible terminal fate. The stars represent the level of significance. If $p < 1e^{-4}$ there are 4 stars, $< 1e^{-3}$ are 3 stars, 2 stars for < 0.01, 1 star for < 0.05, and ns for not significant. Chi-squared GOF was used to evaluate overall differences between terminal fate probabilities of each cluster. H) Total AAV+ cell counts per cell type. I) Total AAV+ mOSN cell counts. J) Percent of total serotype positive cells which are mOSNs. * $p < .05$, Kruskal-Wallis test. K) Percent AAV1+ cell counts of each cell type cluster total cell count.