## MATERIALS AND METHODS

### Mice

All procedures involving mice were approved by the Fred Hutchinson Cancer Research Center Institutional Animal Care and Use Committee. C57BL/6J mice were obtained from The Jackson Laboratory.

### Single cell RNA sequencing (RNA-Seq)

cDNA libraries were prepared from single olfactory epithelium neurons as previously described (*10, 28*). In brief, epithelial tissue was isolated from adult or neonatal animals (P2-P6), dissociated cells were plated on coverslips, and single cells transferred to individual tubes using a microcapillary pipet. Oligo dT-primed cDNAs were prepared from mRNAs in each cell using reverse transcriptase, a poly(A) extension was added to the 3' end of each cDNA using deoxynucleotidyl transferase, and a universal primer then used to amplify the cDNAs. One-third of each cell cDNA mix was used for amplification. For some cells, a duplicate sample was amplified and sequenced starting with a different third of the sample.

To assess the cell stages of neurons used for libraries prior to sequencing, aliquots of single cell libraries were used in PCR reactions with primers for genes expressed at different stages. Primers used were: *Ascl1*, 5' primer, CCACGGTCTTTGCTTCTGTTTTC , 3' primer, GTACGCAGAGGTAATCTCATTACATG, *Neurog1,* 5' primer,

CCCTGAAGACGAGGTGAAAAGTC, 3' primer,

CCAGTGCCTGAATAGCTATGCTAG, *Gap43,* 5' primer,

CTGAACTTTAAGAAATGGCTTTCCAC , 3' primer,

GTTTAAGCCACACTGTTGGACTTG , *Omp,* 5' primer

GCATTTGCTGCTCGCTGGTG, 3' primer, GTGCCACCGTTTTCCTGTCAG.

cDNA libraries were prepared for sequencing using the Illumina TruSeq DNA Sample Prep Kit. Briefly, cDNAs were fragmented to ~300 bp, ligated to adaptors, and PCR amplified with adaptor primers. Samples were subjected to multiplexed sequencing using an Illumina HiSeq 2500 instrument and a paired end, 50 bp read-length sequencing strategy. Image analysis and base calling were performed using Illumina Real Time Analysis v1.13.48 software, followed by 'demultiplexing' of indexed reads and generation of FASTQ files using Illumina CASAVA v1.8.2 software. Adapters were trimmed with the "trim_galore" utility from the cutadapt package (v1.2.1, available from http://www.bioinformatics.babraham.ac.uk/projects/trim_galore/) using the options "--adapter TATAGAATTCGCGGCCGCTCGCGATTTTT --stringency 8 --quality 0 -e 0.15 --length 20 --paired --retain_unpaired".

Sequencing libraries were analyzed with TopHat (*22*) (v2.0.14) and Cufflinks (*23*) (v.2.2.1). Reads were mapped to the mouse genome (mm10, downloaded from the UCSC genome browser) and with the GENCODE vM2 database of genome annotations. TopHat was provided with default options. Because 3' untranslated regions of most *Olfr* transcripts have not been annotated, we sought to refine GENCODE *Olfr* gene structures

by assembling our RNA-Seq reads. We assembled the transcripts recovered from each cell using Cufflinks, which was provided with default options. Assembled transcriptomes for each cell were then merged with the "cuffmerge" utility to create a single transcriptome assembly for the entire experiment. This assembly was compared to GENCODE vM2 to identify and select assembled transcript structures that overlapped with existing gene annotations. To construct a single mouse genome annotation used for downstream analysis, we selected assembled transcripts that overlapped GENCODE *Olfr* genes, along with all GENCODE genes not overlapped by one of our assembled *Olfrs*. This updated GENCODE transcriptome annotation was used for subsequent steps in the analysis. Gene expression profiles were estimated by first running the "cuffquant" tool on the aligned reads for each cell with the "-u" option, which performs additional algorithmic steps designed to better assign ambiguously mapped reads to the correct gene of origin. Per-cell gene expression profiles were subsequently normalized with the "cuffnorm" utility for use in downstream analysis.

Technical quality metrics were analyzed for RNA-Seq libraries using RSeqC version 2.6.2 (*29*) to obtain the number of aligned reads per cell, the percent of reads in each library mapping to exons of mouse genes, (GENCODE M2, mouse genome build mm10), and the number of genes detected at or above 10 fragments per kilobase of transcript per million mapped reads (FPKM).

**Pseudotemporal analysis of the olfactory neuron differentiation trajectory**

Single cell expression data were analyzed with Monocle (v1.2.0) as previously described (*24*). Briefly, Monocle orders differentiating cells according to developmental progress using an unsupervised algorithm and does not require that a user specify a set of marker genes that define progress through differentiation. Each cell can be viewed as a point in a high-dimensional state space, where each gene is a distinct dimension. Monocle aims to reconstruct the "trajectory" the cells travel along through this expression state space as they differentiate. The algorithm first reduces the dimensionality of the log-transformed expression data to two dimensions using Indepenedent Component Analysis. Next, Monocle constructs a minimum spanning tree (MST) on the cells in the reduced-dimensional space. The diameter path of this tree is then used to order the cells. Cells on the diameter path are placed in the order they fall along it. Cells not on the diameter path are placed between cells on it using heuristics designed to minimize the difference in expression between adjacent cells in the final ordering. Once the cells have been placed in order along the trajectory, Monocle assigns the first cell in the ordering a pseudotime value of zero. The second cell receives a pseudotime value equal to the distance between it and the first cell in the reduced space. The third cell is assigned a pseudotime value equal to the pseudotime of the second cell plus the distance between the second and third cell, and so on. Thus, Monocle takes as input a matrix of expression values for the cells and produces as output a measurement of each cell's progress through differentiation derived solely from that data in an entirely unsupervised way.

Monocle is available open-source through the Bioconductor project (https://www.bioconductor.org/). We followed the standard Monocle workflow to analyze differentiating olfactory neurons, as documented in the Monocle vignette.

Although Monocle does not require the user to specify the genes that define progress, it does recommend that genes measured at or below a certain expression threshold be excluded from some steps to ensure reliable ordering. We included all genes with a median expression level of 10 FPKM or higher, after excluding measurements below 0.1. The latter filtering step ensures that the decision to include genes is not simply based on technical "drop-out" artifacts, which are common in current single cell RNA-Seq protocols (*30, 31*). Subsequent steps from the Monocle vignette for reducing dimensionality and ordering cells (via the functions "reduceDimension" and "orderCells") were performed as described with default options.

**Differential expression analysis**

Once it places cells in pseudotime order, Monocle can identify genes that vary over differentiation using statistically robust methods (*24*). To identify genes with pseudotime-dependent changes in expression, we used Monocle's "differentialGeneTest" function on all genes expressed at FPKM >= 0.1.

Single-cell RNA-Seq data are typically log-normally distributed but also highly zero-inflated, probably owing to stochasticity in mRNA recovery and reverse transcription. Although typically used to account for censoring, the Tobit model controls for and

accommodates zero-inflation in FPKM-valued single-cell expression data (*24*). Monocle performs differential analysis by fitting a Tobit-valued vector generalized linear model (VGLM) to the expression data for each gene. This model, specified by the R model formula "expression ~ sm.ns(Pseudotime, df=3)" describes changes in each gene's expression as a function of pseudotime. To regularize the data and improve power, smooth expression changes are smoothed via a natural spline with three degrees of freedom. This model was compared to a reduced model in which pseudotime assignments for each cell were excluded. The two models were compared with a likelihood ratio test to assess statistical significance of pseudotime dependent expression. Fitting and testing was performed using the VGAM package (*32*) which can fit a wide variety of generalized linear and additive models using iteratively reweighted least squares. The p-values from the test for each gene were corrected for multiple testing by Benjamini and Hochberg's method (*33*).

Reproducibility between technical replicates was assessed by mapping reads and quantifying expression for libraries prepared from the same cell as independent samples. The scatterplots in Fig. S4B were prepared by comparing the FPKM values for replicates, after adding pseudo counts, and log-transforming the values.

**Analysis of *Olfr* expression in single neurons**

To ensure an accurate estimate of the number of *Olfr* genes expressed in each individual cell, we curated FPKM values by manual inspection. First, we generated a list of all *Olfr* genes with FPKM >= 1 in each cell. Second, we examined data for each *Olfr* using the

Broad Institute Integrative Genomics Viewer (IGV) (broadinstitute.org/Igv) to visualize the locations of sequenced reads in the gene. *Olfr* transcripts that did not extend over at least a portion of the gene, for example, those showing only one or a few artifactual "read towers", were removed from consideration. Third, we compared *Olfr* transcripts obtained from cDNAs from different cells sequenced in the same Illumina lane. In some cases, it was apparent that an *Olfr* with very high FPKM in one cell sample was present at much lower FPKM in one or more other samples run in the same Illumina lane but not in different Illumina lanes, suggesting that imperfect de-multiplexing ("bleed-through") had created the appearance of low, but detectable *Olfr* expression. *Olfr* FPKM values were manually set to zero in cells in which the measurement was deemed due to a read tower or imperfect bleed-through.

We used a threshold of >10 FPKM for *Olfrs* and other transcripts of interest that were further analyzed, such as olfactory sensory transduction molecules. Examination of *Olfrs* with <10 FPKM, but >1 FPKM using IGV indicated that the vast majority were artifactual.

**RNA *in situ* hybridization (ISH)**

Animals were perfused transcardially with 4% paraformaldehyde (PFA). Nasal tissue was dissected and soaked in 4% PFA for 4 h, in 30% sucrose for 48 h, and then frozen in OCT (Sakura) and cut into 12 μm coronal sections using a cryostat. For animals >3.5 weeks of age, the tissue was decalcified for 3 d in 0.25 M EDTA/2% PFA prior to soaking in sucrose.

Conventional dual fluorescence ISH was performed essentially as described previously (*34, 35*). Using mouse genome sequence data and BLASTN searches available online from NCBI, sequences of approximately 800 bp to 1 kb that were unique to *Olfrs* of interest were identified, PCR-amplified from mouse genomic DNA (Zyagen**)**, and cloned into the pCR4 Topo vector (450030, Life Technologies). Digoxigenin (DIG)-labeled riboprobes were prepared using the DIG RNA Labeling Kit (11277073910, Roche). Dinitrophenyl (DNP)-labeled riboprobes were prepared using a DNP RNA labeling mix containing DNP-11-UTP (NEL555001EA, Perkin Elmer) and NTPs (Roche). Sections were hybridized with probes at 58ºC for 14-16 h and then washed twice for 5 min at 63ºC in 5x SSC followed by twice for 30 min at 63ºC in 0.2x SSC. DIG-labeled probes were detected by incubating sections with horseradish peroxidase (POD)-conjugated sheep anti-DIG antibodies (1:200; 1207733910, Roche) diluted in blocking buffer (1% Blocking reagent, FP1012, Perkin Elmer) for 2h at room temperature (RT). Sections were then washed three times for 5 min at RT in TNT (0.1M Tris-HCl pH7.5, 0.5M NaCl, 0.05 % Tween) buffer, and incubated for 8 min in Biotin-tyramide (1:100; NEL749A001KT, TSA plus Biotin Kit; Perkin Elmer) followed by 20 min in AlexaFluor 568-conjugated Streptavidin (1:2000; S11226, Life Technologies). DNP probes were detected by incubating sections for 2 hr at RT with rabbit anti-DNP-KLH antibodies (1:200; A6430, Life Technologies). Sections were then washed three times for 5 min at RT in TNT buffer, incubated for 2 h at RT with AlexaFluor 488 conjugated donkey anti-rabbit antibodies (1:400; A-21206, Life technologies), and washed. Sections were mounted in

DAPI Fluoromount-G (0100-20, SouthernBiotech) and then analyzed and imaged using AxioImager.Z1 (Zeiss), and LSM 780 NLO confocal (Zeiss), microscopes.

High sensitivity dual ISH was conducted using the QuantiGene ViewRNA Kit (Panomics, Affymetrix) following manufacturer's instructions with minor modifications. Briefly, sections were treated with Proteinase K (5 μg $ml^{-1}$, 03115828001, Roche) for 10 min at RT, washed twice for 5 min in 1X PBS, and then incubated in 4 % PFA for 10 min at RT. After three washes in 1X PBS for 5 min at RT, sections were incubated in 0.1 % Sudan Black B dye (S2380, Sigma; diluted in 70 % ethanol) to eliminate autofluorescence and then hybridized to custom designed QuantiGene ViewRNA probes for *Olfr1507* and *Olfr286* at 40° C overnight. Bound probes were next amplified by incubating for 1h each at 40° C with Affymetrix PreAmplifier and Amplifier molecules. Sections were then incubated for 1h at 40° C with multiple label probes conjugated to Cy3 or AlexaFluor 488. Sections were mounted using Prolong Gold Antifade Mountant (Life Technologies) and then analyzed and imaged using AxioImager.Z1 (Zeiss), and LSM 780 NLO confocal (Zeiss), microscopes.

**Comparison of olfactory epithelium expression patterns of *Olfrs* coexpressed in individual OSNs**

Conventional dual fluorescence ISH was performed essentially as described previously (*35*). Digoxigenin (DIG)-labeled and fluorescein-labeled riboprobes for 2 *Olfrs* identified in the same cell by RNA-Seq were synthesized using the DIG and fluorescein (FLU) RNA Labeling Kit (11685619910, Roche). Sections were hybridized with the two probes

at 58ºC for 14-16 h and then washed twice in 5x SSC for 5 min at 63ºC followed by twice in 0.2x SSC for 30 min at 63ºC. Probes were detected by incubating sections with sheep anti-DIG-alkaline phosphatase antibodies (1:200; 11093274910, Roche) and sheep anti-fluorescein horseradish peroxidase (POD) antibodies (1:200; 11426346910, Roche) for 2h at room temperature. Sections were then washed in TNT buffer (Tris-NaCL-Tween buffer). DIG and FLU labeled probes were detected using the HNPP Fluorescent Detection Set (11758888001, Roche) and TSA Fluorescein kit (NEL741001KT, Perkin Elmer) to produce red and green fluorescent signals, respectively. Sections were mounted in DAPI Fluoromount-G (SouthernBiotech) and then analyzed microscopically for olfactory epithelium spatial zone expression and imaged using the TissueFAXS system (Tissuegnostics).

**Chromosome mapping of coexpressed *Olfrs***

The chromosome locations of individual *Olfrs* were obtained from NCBI online. As previously, *Olfrs* more than 1 Mb apart were considered to be at different *Olfr* chromosomal loci *(4)*.

## SUPPLEMENTAL FIGURE LEGENDS

**Figure S1. Quality metrics for single-cell RNA-Seq libraries used in this study.**

A) Aligned reads per cell. B) Percent of reads in each library mapping to exons of mouse genes (GENCODE M2, mouse genome build mm10). C) Genes detected at or above 10 FPKM. Individual cells are shown in the same order as in Fig. 2 and colored according to cell stage, as indicated.

**Figure S2. Expression of specific genes during OSN development.**

Kinetic diagrams show the expression of known markers of different developmental stages over the developmental progression defined by Monocle (see Fig. 1). Parentheses indicate the groups in which genes were found in the analysis of differentially expressed genes shown in Fig. 1B. Dots indicate individual cells colored according to developmental stage as in Fig. 1C. Black lines indicate loess smoothing (span = 0.75, degree = 2) of log-transformed FPKM values over developmental pseudotime.

**Figure S3. Quality metrics for duplicate libraries prepared from the same cells.**

Quality metrics were determined as in Fig. S1 for duplicate RNA-Seq libraries (1 and 2) (technical replicates) from eight cells. Names of cells and their representative colors are indicated.

**Figure S4**. **Comparisons of technical replicates.**

A. Monocle trajectory and spanning tree, as in Fig. 1A, but with the addition of replicate libraries for 8 cells. Individual cells represented in duplicate (D187-D251) are colored as indicated. Other cells are shown in gray. Cell stages assigned via marker gene analysis are indicated by shape, as indicated.

B. Representative scatterplots are shown of transcript expression levels in log10(FPKM+1) detected in technical replicates (x and y axes) from eight single OSNs.

**Figure S5. A model for *Olfr* gene choice.**

In a "winner-takes-all" model for *Olfr* gene choice, a zone-specific enhancer (red circle) allows multiple *Olfrs* to be expressed at a low level in the early immature OSN. The capture of one or more limiting factors by one *Olfr* permits its high level expression. Low level expression of other *Olfrs* then subsides, owing to the closing of a developmental time window or feedback from the highly expressed *Olfr*.

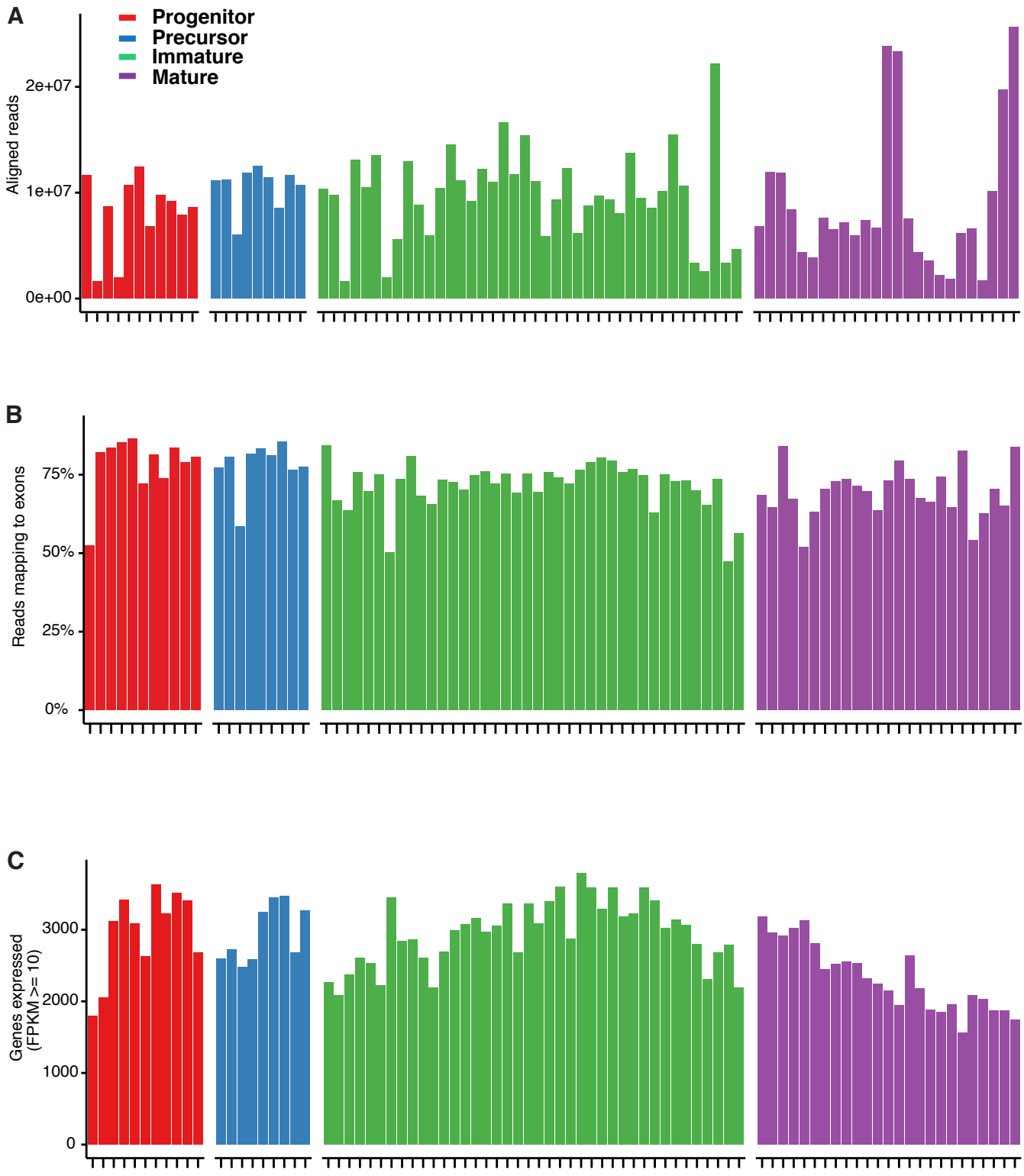
A
Progenitor
Precursor
Immature
Mature
Aligned reads
2e+07
1e+07
0e+00
B
Reads mapping to exons
75%
50%
25%
0%
C
Genes expressed
(FPKM >= 10)
3000
2000
1000
0

**Fig. S1**

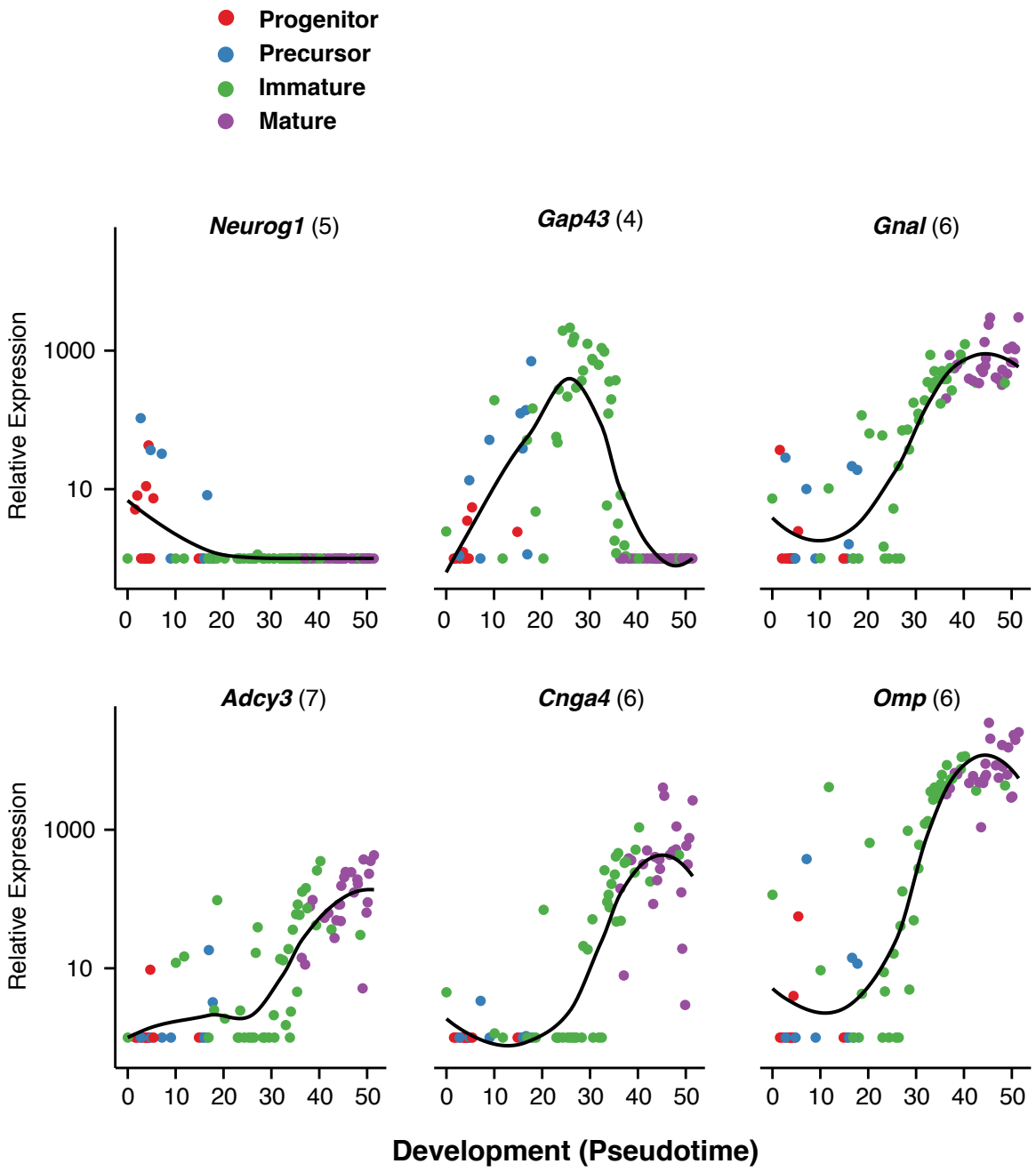
Progenitor
Precursor
Immature
Mature
Neurog1 (5)
Gap43 (4)
Gnal (6)
Adcy3 (7)
Cnga4 (6)
Omp (6)
Relative Expression
1000
10
0
10
20
30
40
50
Development (Pseudotime)

**Fig. S2**

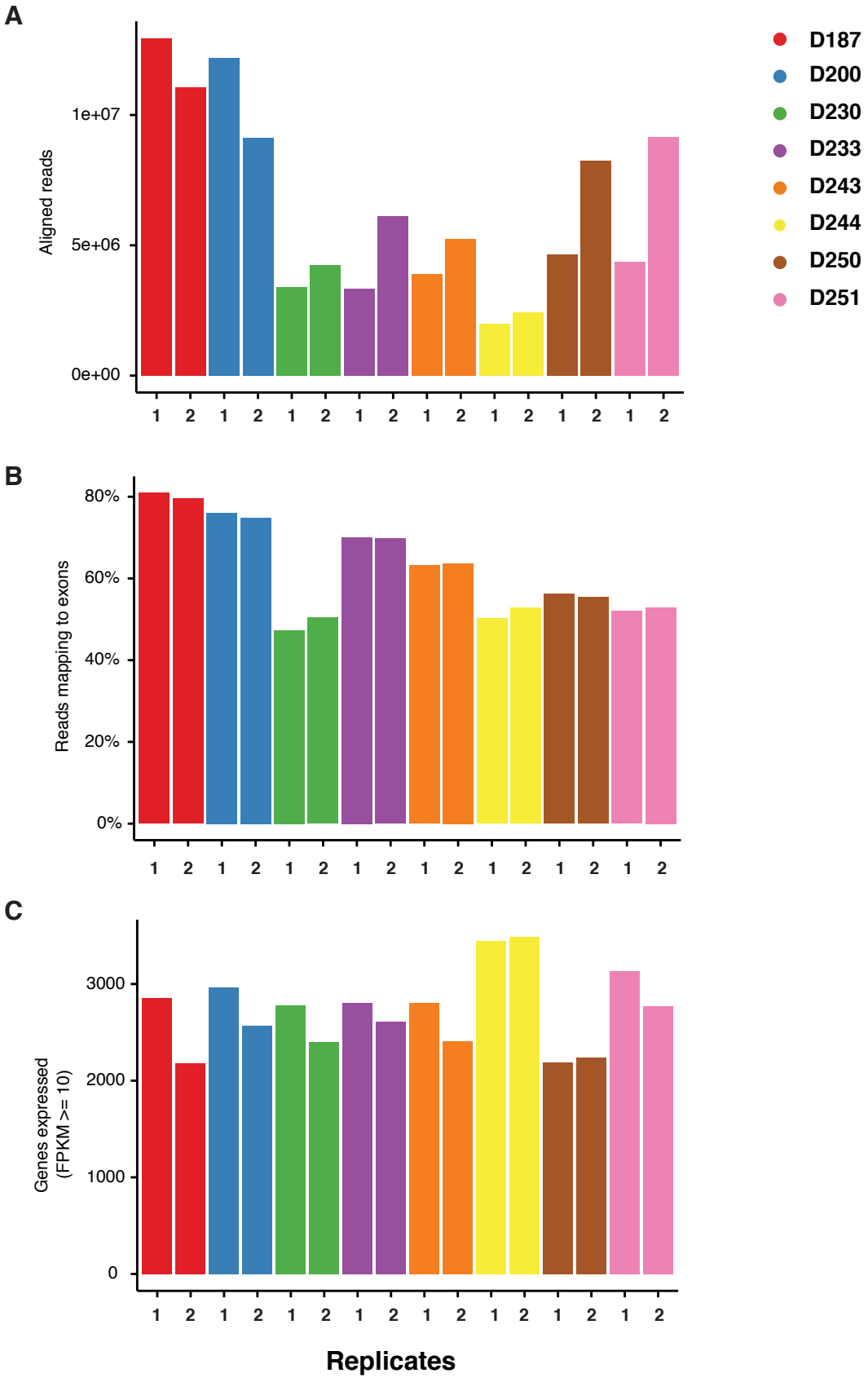
A
Aligned reads
1e+07
5e+06
0e+00
B
Reads mapping to exons
80%
60%
40%
20%
0%
C
Genes expressed
(FPKM >= 10)
3000
2000
1000
0
1 2 1 2 1 2 1 2 1 2 1 2 1 2 1 2
Replicates
D187
D200
D230
D233
D243
D244
D250
D251

**Fig. S3**

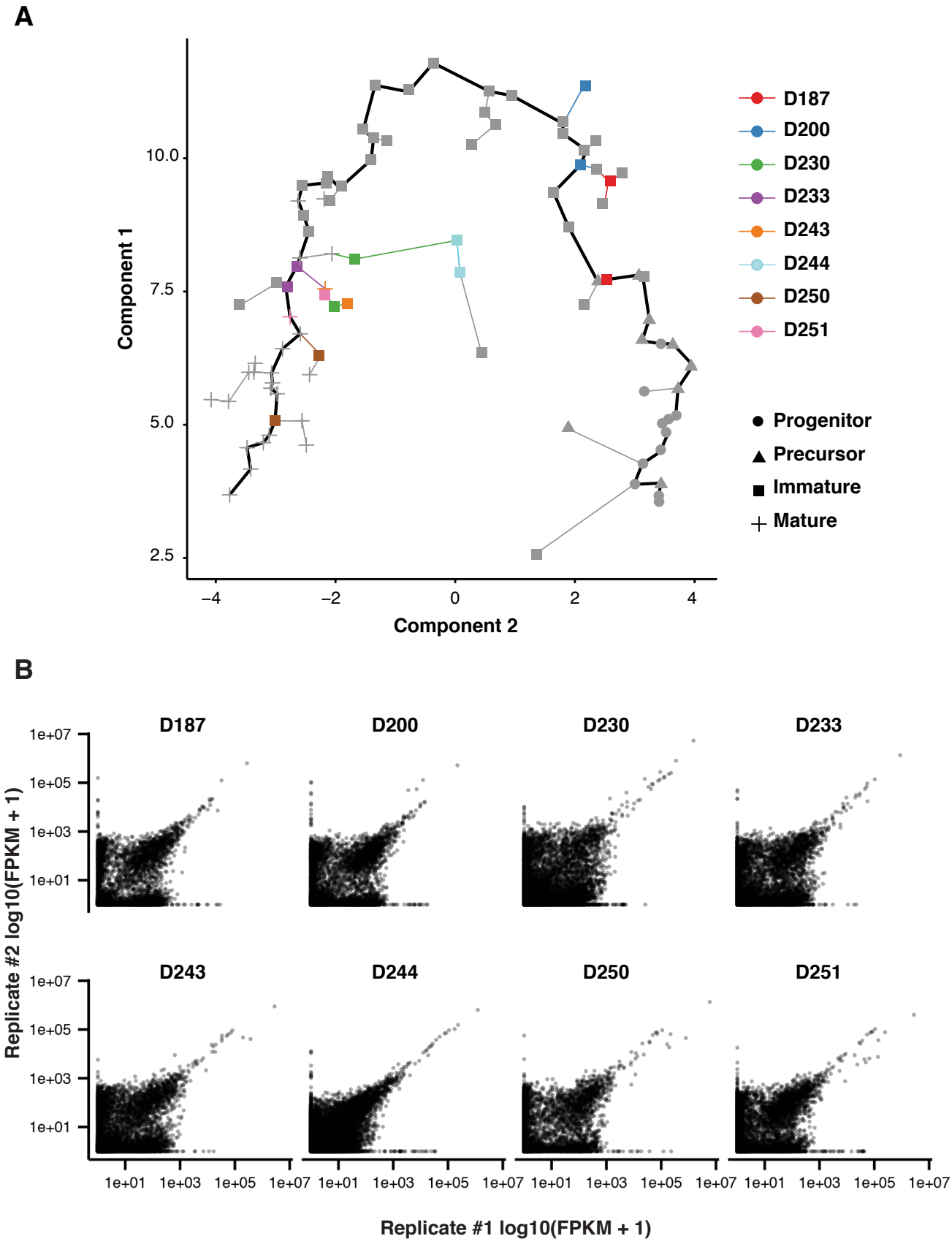
A
Component 1
Component 2
10.0
7.5
5.0
2.5
−4
−2
0
2
4
D187
D200
D230
D233
D243
D244
D250
D251
Progenitor
Precursor
Immature
Mature
B
D187
D200
D230
D233
D243
D244
D250
D251
Replicate #2 log10(FPKM + 1)
Replicate #1 log10(FPKM + 1)
1e+01
1e+03
1e+05
1e+07

Fig. S4

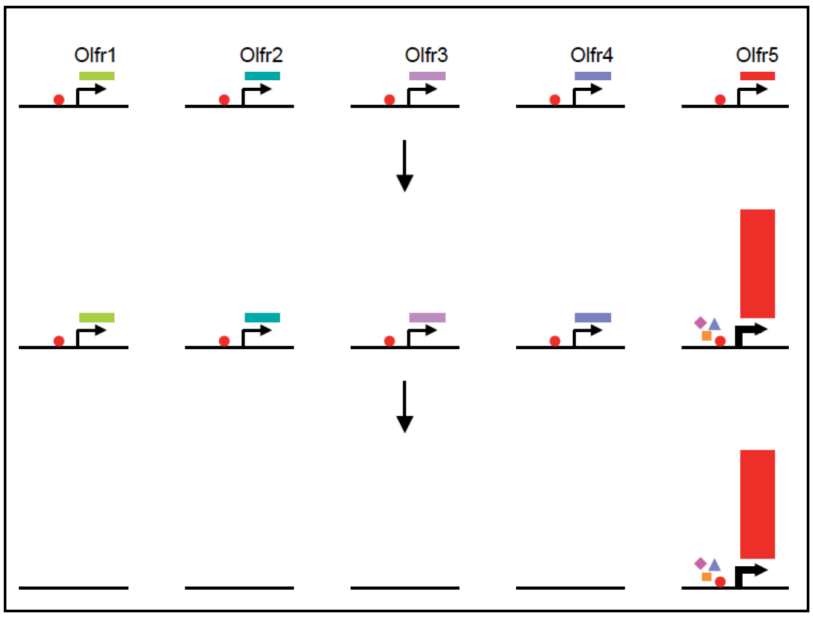


**Fig. S5**

## Expression of transcription and chromatin modifiers during OSN neurogenesis

| | | GO:0045449~regulation of transcription | | | GO:0016568~chromatin modification | | |
|---|---|---|---|---|---|---|---|
| #Cluster | #Total Genes [a] | #Genes [b] | % [c] | list of genes | #Genes [a] | % [b] | list of genes [c] |
| 1 | 253 | 44 | 17.4 | *DPF2, PTOV1, CBX2, ZKSCAN1, GM14420, MXI1, ZKSCAN3, CBFB, FUBP1, RCBTB1, NONO, **KDM1A**, TCERG1, MBTD1, 2210018M11RIK, BCL11B, TARDBP, PSIP1, KDM5B, CHD4, SSBP3, KLF13, RCOR1, TRIM28, CDK9, HMG20B, ZFP148, UBE2B, ZFP575, CCNL2, **FEZF1**, HDAC3, CHMP1A, HIF1A, HDAC2, ILF2, SFPQ, GTF2F2, THRAP3, RFX1, TCEB2, LIME1, USP22, ZFP513* | 15 | 5.9 | *KDM6A, RCOR1, HMG20B, CBX2, UBE2B, RCBTB1, RNF8, **KDM1A**, HDAC3, HDAC2, 2210018M11RIK, H2AFY2, USP22, KDM5B, CHD4* |
| 2 | 186 | 21 | 11.3 | *YEATS4, TSHZ1, BACH2, OTX1, ASXL1, CNOT1, ILF3, TLE1, KEAP1, SUZ12, ZFP579, PCGF3, DRAP1, ZSCAN21, SMARCA5, NHLH2, ZIK1, ZFP821, KDM3B, RBPJ, CHD3* | 5 | 2.7 | *SUZ12, YEATS4, PHF16, SMARCA5, KDM3B* |
| 3 | 204 | 30 | 14.7 | *ELL, HIRA, MED12L, 2610008E11RIK, RGMB, ZFP668, **LHX2**, ZFP90, POU2F1, NPAT, PDE8A, OLIG2, ZFP810, NFX1, SETDB1, ASXL3, ZFP423, **SOX11**, FOXA1, AFF3, ZFP322A, ZFP747, GZF1, CTNNBIP1, MED6, TAF13, BAZ1B, DLX6, **DLX5**, NHLH1* | 3 | 1.5 | *SETDB1, BAZ1B, HIRA* |
| 4 | 291 | 22 | 7.6 | *ZCCHC12, DNMT3A, SSBP2, CTBP2, CRTC1, ZFP30, **EMX2**, TADA1, STAT1, PRPF6, SUV420H2, HDAC5, TARBP2, TCF20, NCOA2, ZSCAN22, SRFBP1, TRAK1, AGRN, MCTS2, MLLT3, HIP1* | 3 | 1.0 | *HDAC5, SUV420H2, DNMT3A* |
| 5 | 467 | 68 | 14.6 | *E2F1, ITGB3BP, ENY2, E2F5, E2F6, ARNT2, **PAX6**, NFKB1, SAP25, FNTB, HEY1, MED26, TOP2A, PITX1, IRAK1, ELP3, CCNK, POGZ, TADA2A, DMRT3, YY1, OTX2, RBBP7, JUNB, MXD3, CDK2, NOC2L, EYA1, TIMELESS, ZFP280C, MED17, **SIX1**, RUVBL2, ZFP260, 3110039M20RIK, HMGB1, HMGB2, **SOX2**, HDGF, NFIX, TRRAP, MEIS2, TSC22D4, ZKSCAN17, ZFP410, RQCD1, TAF9, **RUNX1**, HELLS, SREBF1, SAMD11, SMAD5, RYBP, **NEUROG1**, CENPF, ATAD2, NR4A1, TEAD2, PMF1, GM14418, MSL3, NOTCH1, ATF3, DR1, ZFP7, KDM4B, IRF3, CHAF1A* | 9 | 1.9 | *ENY2, MSL3, EYA1, KDM4B, CHD1, RUVBL2, TRRAP, RBBP7, HELLS* |
| 6 | 538 | 37 | 6.9 | *EFCAB6, TAF9B, CBX4, PDLIM1, MED21, POU1F1, ZFP318, ZBTB38, TSC22D1, APP, CREB3L1, MYCBP, SPIC, NAT14, NFATC1, ZBTB7B, AEBP2, KHDRBS3, FOXJ1, EOMES, ARID3A, RHOX8, PKIA, PPARGC1A, PURB, **ATF5**, TRAF3IP1, PRKCQ, RNF6, GCM1, EAF1, **EBF4**, DPM1, ID4, NFE2L2, RFX2, MAPRE3* | 4 | 0.7 | *AEBP2, CBX4, BNIP3, MAP3K12* |
| 7 | 246 | 17 | 6.9 | *ABLIM2, KAT2B, RARG, EPAS1, RCOR2, NDUFA13, BRIP1, RORB, NEO1, CBX8, ZBTB44, ZFP871, PHF1, NR1D2, CSDE1, TIGD3, HDAC9* | 4 | 1.6 | *KAT2B, PHF1, CBX8, HDAC9* |
| 8 | 562 | 41 | 7.3 | *AKNA, MEF2B, CAMTA2, TSHZ2, HTATIP2, ABCA2, ZEB1, FOXO3, CBX6, MYCBP2, ZFP91, TRIM66, CRY2, NFAT5, CASZ1, ZFP523, GM4924, ETV5, ETV3, ZBTB7A, TESC, ESRRA, IKZF2, CCNC, BMYC, SIRT2, DDIT3, SOD2, PKNOX1, NPTXR, CSRNP2, TRPS1, USP21, KHSRP, BRDT, RFX3, MAPK8IP1, VOPP1, PBX3, KCNH3, MCTS1, KCNH4* | 3 | 0.5 | *NPTXR, USP21, BRDT, CBX6* |
| 9 | 390 | 24 | 6.2 | *4933413G19RIK, NFKBIZ, ZBTB46, RFX4, CCNH, PHF10, MALT1, ZFP777, NR3C1, HDAC11, MAX, DCAF6, TSC22D3, RPS6KA4, ATF7, PER2, KCNH7, ZFP800, HIVEP2, NFE2L1, VGLL4, JMJD1C, NFIC, SERTAD1* | 4 | 1.0 | *BRE, HDAC11, NR3C1, JMJD1C* |
| 10 | 76 | 7 | 9.2 | *MYT1L, NCOA1, CDYL, **EBF3**, **EBF2**, **EBF1**, ELL3* | none | | |
| 11 | 627 | 82 | 13.1 | *EZH2, CBX3, MED24, CNOT2, CTCF, PRDX2, TCEAL8, CBFA2T2, CBX5, MAGED1, ZGPAT, BZW1, SAP30, EPC2, AES, MCM7, SMARCD2, ASH2L, MED29, MED27, PATZ1, MKL2, ZFP503, CAT, CRY1, CCNA2, ZFP422, KHDRBS1, IKBKAP, SATB1, RBBP4, GTF2H4, SF1, ADNP, TLE4, MTA1, MBD3, **HES6**, MCM4, HMGA1, FLNA, MCM5, BAZ1A, EYA2, BTG2, **FOXG1**, RUVBL1, BCLAF1, CARHSP1, NACA, SOX4, HAT1, ZFP451, CHD7, HNRNPD, ZFP219, ACTL6A, TCF4, TCF3, HNRNPAB, ZFP251, HMGA1-RS1, TRP53, POLR3K, SUV39H1, 4930522L14RIK, TEAD1, SIRT6, SAFB2, 2410018L13RIK, MNAT1, ID2, NUP62, ID1, PHB2, SMARCC1, ZFP740, **NEUROD1**, PBX1, ID3, TBL1X, TCF12, PHF6* | 22 | 3.5 | *SATB1, RBBP4, NASP, EZH2, SUV39H1, CBX3, HAT1, CTCF, MBD3, KDM1B, EPC2, EYA2, CHD7, BAZ1A, SMARCE1, SMARCD2, PRMT8, SMARCC1, RTF1, ACTL6A, RUVBL1, TCF3* |

Clusters of differentially expressed genes that covaried during OSN neurogenesis (Fig. 1B) were analyzed for functional annotations using DAVID Bioinformatics Resources 6.7 (david.ncicrf.gov). Genes annotated under "regulation of transcription" (Gene ontology ID: 0045449) and "chromatin modification" (Gene ontology ID: 0016568) are listed here. Genes indicated in red and blue have been implicated in Olfr expression and OSN neurogenesis, respectively (*33, 18*). [a] total genes in cluster, [b] # genes in each category, [c] percent of genes in each category.

## Table S1

## Comparison of *Olfrs* in technical replicates

| Cell | *Olfr* | FPKM in Replicate# 1 | FPKM in Replicate# 2 |
|---|---|---|---|
| D187 (Early Immature) | *Olfr877* | 76 | 95 |
| | *Olfr876* | 172 | 0 |
| | *Olfr231* | 64 | 0 |
| | *Olfr917* | 0 | 314 |
| | *Olfr1157* | 0 | 187 |
| | *Olfr875* | 0 | 157 |
| D200 (Early immature) | *Olfr570* | 305 | 850 |
| | *Olfr1253* | 165 | 57 |
| | *Olfr610* | 320 | 0 |
| | *Olfr518* | 0 | 290 |
| | *Olfr923* | 0 | 90 |
| | *Olfr225* | 0 | 32 |
| D244 (Early Immature) | *Olfr1322* | 373 | 267 |
| | *Olfr1030* | 75 | 0 |
| | *Olfr1223* | 35 | 0 |
| | *Olfr1459* | 0 | 36 |
| | *Olfr1* | 0 | 24 |
| | *Olfr536* | 0 | 15 |
| | *Olfr1382* | 0 | 12 |
| D233 (late immature) | *Olfr325* | 15286 | 20123 |
| | *Olfr1284* | 616 | 1259 |
| | *Olfr725* | 67 | 0 |
| D250 (late immature) | *Olfr730* | 3541 | 6360 |
| | *Olfr1251* | 393 | 0 |
| D230 (Late Immature) | *Olfr1241* | 916 | 1622 |
| | *Olfr810* | 782 | 1440 |
| | *Olfr983* | 245 | 614 |
| | *Olfr821* | 217 | 37 |
| | *Olfr1229* | 115 | 21 |
| | *Olfr481* | 138 | 0 |
| | *Olfr52* | 82 | 0 |
| | *Olfr726* | 0 | 37 |
| | *Olfr180* | 0 | 29 |
| D243 (Mature) | *Olfr728* | 10308 | 3762 |
| | *Olfr378* | 156 | 0 |
| | *Olfr1425* | 62 | 0 |
| D251 (Mature) | *Olfr1281* | 1800 | 1616 |
| | *Olfr777* | 64 | 0 |
| | *Olfr56* | 0 | 97 |
| | *Olfr568* | 0 | 69 |
| | *Olfr1193* | 0 | 17 |

FPKM values of Olfrs detected in replicates.

## Table S2

## Conventional RNA-FISH of *Olfr* coexpression in OSNs

| Animal | *Olfr1507+* | *Olfr1507+* that Z4 Olfr+ [a] | % of *Olfr1507+* dual labeled |
|---|---|---|---|
| P3 # 6 | 1732 | 2 | 0.12 |
| P3 # 8 | 1958 | 3 | 0.15 |
| P3 # 9 | 1740 | 7 | 0.40 |
| P3 # 10 | 2043 | 7 | 0.34 |
| P3 # 11 | 2623 | 4 | 0.15 |
| P3 # 12 | 1689 | 3 | 0.18 |
| **total** | **11785** | **26** | **0.22 ± 0.05** [c] |
| | | | |
| Adult # A2 | 511 | 0 | 0.00 |
| Adult # A3 | 968 | 0 | 0.00 |
| Adult # A7 | 445 | 0 | 0.00 |
| Adult # A9 | 1012 | 0 | 0.00 |
| **total** | **2936** | **0** | **0.00** |

| Animal | *Olfr1507+* | *Olfr1507+* that Z4 Olfr+ [b] | % of *Olfr1507+* dual labeled |
|---|---|---|---|
| P3 # 11 | 1783 | 7 | 0.39 |
| P3 # 20 | 1164 | 6 | 0.52 |
| P3 # 5 | 533 | 3 | 0.56 |
| P3 # 4 | 868 | 3 | 0.35 |
| **total** | **4348** | **19** | **0.45 ± 0.05** |

| Animal | *Olfr211+* | *Olfr211+* that Z3 Olfr | % of *Olfr211+* dual labeled |
|---|---|---|---|
| P3 # 9 | 755 | 1 | 0.13 |
| P3 # 10 | 723 | 2 | 0.28 |
| P3 # 11 | 572 | 4 | 0.70 |
| P3 # 110 | 308 | 0 | 0.00 |
| P3 # 6 | 345 | 0 | 0.00 |
| **total** | **2703** | **7** | **0.22 ± 0.12** |
| | | | |
| Adult # A3 | 662 | 0 | 0.00 |
| Adult # F3 | 639 | 0 | 0.00 |
| Adult # F4 | 693 | 0 | 0.00 |
| Adult # A7 | 562 | 0 | 0.00 |
| Adult # M8 | 567 | 0 | 0.00 |
| **total** | **3123** | **0** | **0.00** |

| Animal | *Olfr743+* | *Olfr743+* that *Olfr1229+* | % of *Olfr743+* dual labeled |
|---|---|---|---|
| Adult # 28 | 834 | 0 | 0.00 |
| Adult # A9 | 2586 | 0 | 0.00 |
| Adult # A7 | 1432 | 0 | 0.00 |
| Adult # 26 | 937 | 0 | 0.00 |
| **total** | **5789** | **0** | **0.00** |

Sections from animals aged P3 or 2-3 months (adult) were analyzed for OSNs double-labeled using probes for *Olfr1507* and a mix of 9 (*a*) or 12 (*b*) zone 4 *Olfrs*, *Olfr211* and a mix of 10 zone 3 *Olfrs*, or *Olfr743* and *Olfr1229*. *c,* percentages are shown mean ± S.E.M..

**Table S3**

## High sensitivity RNA-FISH of *Olfr* coexpression in OSNs

| **Animal** | ***Olfr1507+*** | ***Olfr286+*** | ***Olfr1507+* that *Olfr286+*** | **% of *Olfr1507+* dual labeled** | **% of *Olfr286+* dual labeled** |
|---|---|---|---|---|---|
| P3 # 4 | 1422 | 663 | 5 | 0.35 | 0.75 |
| P3 # 12 | 790 | 415 | 4 | 0.51 | 0.96 |
| P3 # 14 | 618 | 353 | 3 | 0.49 | 0.85 |
| P3 # 15 | 728 | 394 | 1 | 0.14 | 0.25 |
| P3 # 16 | 792 | 557 | 0 | 0.00 | 0.00 |
| P3 # 19 | 140 | 155 | 1 | 0.71 | 0.65 |
| P3 # 35 | 311 | 265 | 2 | 0.64 | 0.75 |
| **total** | **4801** | **2802** | **16** | **0.41 ± 0.09** | **0.60 ± 0.13** |
| | | | | | |
| P21 # 7 | 4278 | 2513 | 5 | 0.12 | 0.20 |
| P21 # 27 | 2268 | 1123 | 3 | 0.13 | 0.27 |
| P21 # 29 | 2188 | 1129 | 1 | 0.05 | 0.09 |
| **total** | **8734** | **4765** | **9** | **0.1 ± 0.02** | **0.18 ± 0.05** |
| | | | | | |
| Adult # 52 | 5185 | 2175 | 0 | 0 | 0 |
| Adult # 53 | 9276 | 2943 | 1 | 0.01 | 0.03 |
| **total** | **14461** | **5118** | **1** | **0.01** | **0.02** |

| **Animal** | ***Olfr1507*-Sense** | ***Olfr286+*** | **% of *Olfr286+* dual labeled** |
|---|---|---|---|
| P3 # 15 | 0 | 429 | 0 |
| P3 # 4 | 0 | 429 | 0 |
| P3 # 16 | 0 | 353 | 0 |
| P21 # 7 | 0 | 1530 | 0 |
| **total** | **0** | **2741** | **0** |
| | | | |
| **Animal** | ***Olfr1507+*** | ***Olfr286*-Sense** | **% of *Olfr1507+* dual labeled** |
| P3 # 15 | 594 | 0 | 0 |
| P3 # 4 | 699 | 0 | 0 |
| P21 # 7 | 1258 | 0 | 0 |
| **total** | **2551** | **0** | **0** |

Sections from animals aged P3, P21, or 2-3 months (adult) were analyzed for OSNs double-labeled using antisense or sense probes for *Olfr1507* and *Olfr286*. Percentages are shown mean ± S.E.M..

**Table S4**

## Expression of immature markers in *Olfr1507+* OSNs at different ages

| **Animal** | ***Olfr1507+*** | ***Olfr1507+* that *Gap43+*** | **% of *Olfr1507+* dual labeled** |
|---|---|---|---|
| P3 # 24 | 669 | 557 | 83.3 |
| P3 # 17 | 654 | 503 | 76.9 |
| **total** | **1323** | **1060** | **80.1 ± 3.2** |
| | | | |
| P21 # 29 | 929 | 185 | 19.9 |
| P21 # 32 | 605 | 115 | 19.0 |
| **total** | **1534** | **300** | **19.5 ± 0.5** |
| | | | |
| Adult # 52 | 949 | 65 | 6.9 |
| Adult # 53 | 1049 | 78 | 7.4 |
| **total** | **1998** | **143** | **7.1 ± 0.3** |

| **Animal** | ***Olfr1507+*** | ***Olfr1507+* that *Gng8+*** | **% of *Olfr1507+* dual labeled** |
|---|---|---|---|
| P3 # 24 | 564 | 349 | 61.9 |
| P3 # 17 | 519 | 330 | 63.6 |
| **total** | **1083** | **679** | **62.8 ± 0.9** |
| | | | |
| P21 # 3 | 847 | 112 | 13.2 |
| P21 # 29 | 904 | 138 | 15.3 |
| **total** | **1751** | **250** | **14.3 ± 1.1** |
| | | | |
| Adult # 52 | 1085 | 21 | 1.9 |
| Adult # 53 | 1026 | 23 | 2.2 |
| **total** | **2111** | **44** | **2.1 ± 0.2** |

Sections from animals aged P3, P21, or 2-3 months (adult) were analyzed by dual RNA-FISH using probes for *Olfr1507* and the immature OSN markers, *Gap43* and *Gng8*. Percentages are shown as mean ± S.E.M..

**Table S5**

## OE expression zones of *Olfrs* detected in single neurons

| Cell | # *Olfrs* | *Olfr* | Zone 1 | Zone 2-3 | Zone 4 | Overlap |
|---|---|---|---|---|---|---|
| D233 | 2 | *Olfr325* | - | - | + | |
| | | *Olfr1284* | - | - | + | +++ |
| D215 | 2 | *Olfr78* | + | - | - | |
| | | *Olfr877* | + | - | - | +++ |
| D200 | 3 | *Olfr570* | + | - | - | |
| | | *Olfr610* | + | - | - | +++ |
| | | *Olfr1253* | - | + | - | + |
| D187 | 3 | *Olfr876* | + | - | - | |
| | | *Olfr877* | + | - | - | ++ |
| | | *Olfr231* | + | - | - | +++ |
| D197 | 4 | *Olfr1104* | + | - | - | |
| | | *Olfr32* | + | - | - | ++ |
| | | *Olfr1537* | + | - | - | ++ |
| | | *Olfr1354* | - | + | - | + |
| D97 | 2 | *Olfr743* | - | - | + | |
| | | *Olfr1229* | - | - | + | ++ |
| D243 | 3 | *Olfr728* | - | - | + | |
| | | *Olfr378* | - | + | - | ++ |
| | | *Olfr1425* | - | + | - | ++ |

Dual RNA-FISH with adult OE sections was used to compare the expression zones of *Olfrs* identified in the same OSN. The first *Olfr* listed for each cell was was compared with each of the others listed for that cell. Rough estimates of the extent of overlap between expression zones of different *Olfrs* are indicated: +, 5-29%; ++, 30-59%, +++, 60-90%.

**Table S6**

## Chromosomal Locations of *Olfr* genes identified in single neurons

| Cell | Olfr | Chr | Megabase | #Olfrs | #Chrs | #Loci |
|---|---|---|---|---|---|---|
| D424 | *Olfr1431* | 19 | 12209568..12210506 | 6 | 4 | 6 |
| | *Olfr323* | 11 | 58625073..58626044, complement | | | |
| | *Olfr113* | 17 | 37574483..37575421, complement | | | |
| | *Olfr209* | 16 | 59361224..59362240, complement | | | |
| | *Olfr43* | 11 | 74206198..74207289, complement | | | |
| | *Olfr1395* | 11 | 49148259..49149212 | | | |
| D215 | *Olfr78* | 7 | 102740721..102759471, complement | 2 | 2 | 2 |
| | *Olfr877* | 9 | 37854820..37855755 | | | |
| D411 | *Olfr467* | 7 | 107814586..107816819 | 10 | 7 | 7 |
| | *Olfr992* | 2 | 85399602..85400531, complement | | | |
| | *Olfr1331* | 4 | 118868783..118869736 | | | |
| | *Olfr967* | 9 | 39750388..39751320 | | | |
| | *Olfr1032* | 2 | 86007778..86008710 | | | |
| | *Olfr1451* | 19 | 12995189..12999920 | | | |
| | *Olfr323* | 11 | 58625073..58626044, complement | | | |
| | *Olfr1500* | 19 | 13827459..13828394, complement | | | |
| | *Olfr284* | 15 | 98340022..98340939, complement | | | |
| | *Olfr923* | 9 | 38827615..38828682 | | | |
| D244 | *Olfr1322* | X | 49885469..49886401, complement | 3 | 2 | 3 |
| | *Olfr1030* | 2 | 85979312..85984798 | | | |
| | *Olfr1223* | 2 | 89144086..89151336, complement | | | |
| D371 | *Olfr661* | 7 | 104688017..104688976 | 2 | 1 | 1 |
| | *Olfr628* | 7 | 103731928..103732878 | | | |
| D332 | *Olfr1222* | 2 | 89124794..89125729, complement | 12 | 7 | 9 |
| | *Olfr788* | 10 | 129472694..129473629 | | | |
| | *Olfr1231* | 2 | 89302649..89303590, complement | | | |
| | *Olfr175-ps1* | 16 | 58823781..58826761, complement | | | |
| | *Olfr786* | 10 | 129436814..129437752 | | | |
| | *Olfr119* | 17 | 37696685..37701720 | | | |
| | *Olfr1357* | 10 | 78611669..78618074, complement | | | |
| | *Olfr1107* | 2 | 87071115..87072223, complement | | | |
| | *Olfr20* | 11 | 73350859..73354699 | | | |
| | *Olfr913* | 9 | 38592803..38595161 | | | |
| | *Olfr112* | 17 | 37563140..37569696, complement | | | |
| | *Olfr686* | 7 | 105203388..105204341, complement | | | |
| D318 | *Olfr1102* | 2 | 87001932..87002994 | 10 | 7 | 9 |
| | *Olfr808* | 10 | 129765483..129768436 | | | |
| | *Olfr745* | 14 | 50642195..50643369 | | | |
| | *Olfr957* | 9 | 39510783..39511718, complement | | | |
| | *Olfr374* | 8 | 72107040..72110509 | | | |
| | *Olfr338* | 2 | 36376778..36377698 | | | |
| | *Olfr1166* | 2 | 88123360..88124993, complement | | | |
| | *Olfr1448* | 19 | 12919363..12920307, complement | | | |
| | *Olfr218* | 1 | 173203358..173204299 | | | |
| | *Olfr341* | 2 | 36479187..36480128, complement | | | |
| D187 | *Olfr876* | 9 | 37803889..37804936 | 3 | 2 | 2 |
| | *Olfr877* | 9 | 37854820..37855755 | | | |
| | *Olfr231* | 1 | 174117091..174118014, complement | | | |
| D373 | *Olfr1314* | 2 | 112091761..112092699, complement | 4 | 3 | 4 |
| | *Olfr796* | 10 | 129607547..129608479, complement | | | |
| | *Olfr27* | 9 | 39128167..39145072 | | | |
| | *Olfr1016* | 2 | 85799339..85800268, complement | | | |
| D197 | *Olfr1104* | 2 | 87021610..87022542, complement | 4 | 3 | 4 |
| | *Olfr1354* | 10 | 78916842..78917936 | | | |
| | *Olfr32* | 2 | 90138081..90268639, complement | | | |
| | *Olfr1537* | 9 | 39237487..39238431, complement | | | |
| D336 | *Olfr187* | 16 | 59035780..59039749, complement | 3 | 2 | 2 |
| | *Olfr186* | 16 | 59026976..59027905, complement | | | |
| | *Olfr1463* | 19 | 13234252..13235184 | | | |
| D200 | *Olfr610* | 7 | 103505997..103506944, complement | 3 | 2 | 2 |
| | *Olfr570* | 7 | 102813310..102901380 | | | |
| | *Olfr1253* | 2 | 89751870..89752826, complement | | | |
| D168 | *Olfr1013* | 2 | 85769803..85770720 | 2 | 2 | 2 |
| | *Olfr131* | 17 | 38082032..38082976, complement | | | |
| D334 | *Olfr285* | 15 | 98312589..98313548, complement | 2 | 2 | 2 |
| | *Olfr531* | 7 | 140400130..140401044, complement | | | |
| D233 | *Olfr325* | 11 | 58580837..58581877 | 3 | 3 | 3 |
| | *Olfr1284* | 2 | 111379002..111379937 | | | |
| | *Olfr725* | 14 | 50034391..50035500, complement | | | |
| D234 | *Olfr56* | 11 | 49050733..49135387 | 2 | 2 | 2 |
| | *Olfr1148* | 2 | 87833041..87833985 | | | |
| D230 | *Olfr810* | 10 | 129790649..129791587, complement | 7 | 4 | 5 |
| | *Olfr1241* | 2 | 89482189..89483133, complement | | | |
| | *Olfr983* | 9 | 40092017..40093558, complement | | | |
| | *Olfr821* | 10 | 130030390..130034560 | | | |
| | *Olfr481* | 7 | 108080796..108081734 | | | |
| | *Olfr1229* | 2 | 89282196..89283131, complement | | | |
| | *Olfr52* | 2 | 86181150..86182109, complement | | | |
| D250 | *Olfr730* | 14 | 50186262..50187218, complement | 2 | 2 | 2 |
| | *Olfr1251* | 2 | 89666928..89667884, complement | | | |
| D243 | *Olfr728* | 14 | 50139702..50140637, complement | 3 | 3 | 3 |
| | *Olfr378* | 11 | 73425037..73428477, complement | | | |
| | *Olfr1425* | 19 | 12073695..12074630, complement | | | |
| D251 | *Olfr1281* | 2 | 111328421..111329338 | 2 | 2 | 2 |
| | *Olfr777* | 10 | 129268386..129270550, complement | | | |
| D232 | *Olfr398* | 11 | 73983662..73984606, complement | 2 | 2 | 2 |
| | *Olfr147* | 9 | 38401711..38403829 | | | |
| D274 | *Olfr1033* | 2 | 86020640..86044811 | 2 | 2 | 2 |
| | *Olfr685* | 7 | 105180361..105181311, complement | | | |
| D97 | *Olfr743* | 14 | 50533414..50534349 | 2 | 2 | 2 |
| | *Olfr1229* | 2 | 89282196..89283131, complement | | | |
| D240 | *Olfr871* | 9 | 20212207..20213353 | 2 | 2 | 2 |
| | *Olfr1140* | 2 | 87746161..87747180 | | | |
| | | | **Mature Neurons** | | | |
| D243 | *Olfr728* | 14 | 50139702..50140637, complement | 3 | 3 | 3 |
| | *Olfr378* | 11 | 73425037..73428477, complement | | | |
| | *Olfr1425* | 19 | 12073695..12074630, complement | | | |
| D251 | *Olfr1281* | 2 | 111328421..111329338 | 2 | 2 | 2 |
| | *Olfr777* | 10 | 129268386..129270550, complement | | | |
| D232 | *Olfr398* | 11 | 73983662..73984606, complement | 2 | 2 | 2 |
| | *Olfr147* | 9 | 38401711..38403829 | | | |
| D274 | *Olfr1033* | 2 | 86020640..86044811 | 2 | 2 | 2 |
| | *Olfr685* | 7 | 105180361..105181311, complement | | | |
| D97 | *Olfr743* | 14 | 50533414..50534349 | 2 | 2 | 2 |
| | *Olfr1229* | 2 | 89282196..89283131, complement | | | |
| D240 | *Olfr871* | 9 | 20212207..20213353 | 2 | 2 | 2 |
| | *Olfr1140* | 2 | 87746161..87747180 | | | |

Chromosome and megabase locations were obtained from NCBI. *Olfrs* <1 megabase apart were assigned to the same locus. All cells shown were classified as immature neurons except those indicated as mature neurons.

## Table S7