Molecular characterization of projection neuron subtypes in the mouse olfactory bulb

Sara Zeppilli[1,2], Tobias Ackels[3,4,10], Robin Attey[1,10], Nell Klimpert[1], Kimberly D. Ritola[5], Stefan Boeing[6,7], Anton Crombach[8,9], Andreas T. Schaefer[3,4*], Alexander Fleischmann[1,2*]

[1]Department of Neuroscience, Division of Biology and Medicine, Brown University, Providence, USA

[2]Center for Interdisciplinary Research in Biology (CIRB), Collège de France, and CNRS UMR 7241 and INSERM U1050, Paris, France

[3]The Francis Crick Institute, Neurophysiology of Behaviour Laboratory, London, UK

[4]Department of Neuroscience, Physiology & Pharmacology, University College London, UK

[5]Janelia Research Campus, Howard Hughes Medical Institute, Ashburn, USA

[6]The Francis Crick Institute, Bioinformatics and Biostatistics, London, UK

[7]The Francis Crick Institute, Scientific Computing - Digital Development Team, London, UK

[8]Inria Antenne Lyon La Doua, Villeurbanne, France

[9]Université de Lyon, INSA-Lyon, LIRIS, UMR 5205, Villeurbanne, France

[10]These authors contributed equally

[*]Corresponding authors: andreas.schaefer@crick.ac.uk, alexander_fleischmann@brown.edu

Abstract

Projection neurons (PNs) in the mammalian olfactory bulb (OB) receive direct input from the nose and project to diverse cortical and subcortical areas. Morphological and physiological studies have highlighted functional heterogeneity, yet no molecular markers have been described that delineate PN subtypes. Here, we used viral injections into olfactory cortex and fluorescent nucleus sorting to enrich PNs for high-throughput single nucleus and bulk RNA deep sequencing. Transcriptome analysis and RNA *in situ* hybridization identified three mitral and five tufted cell populations with characteristic transcription factor network topology and cell adhesion and excitability-related gene expression. Finally, by integrating bulk and snRNA-seq data we propose that different mitral cell populations selectively project to different regions of olfactory cortex. Together, we have identified potential molecular and gene regulatory mechanisms underlying PN diversity and provide new molecular entry points into studying the diverse functional roles of mitral and tufted cell subtypes.

41 **Introduction**

42 The mammalian olfactory system is unique among sensory systems in that it bypasses the
43 thalamus: olfactory receptor neurons (ORNs) in the nose project to the olfactory bulb (OB), a
44 forebrain structure containing – in the mouse – approximately 500,000 neurons per
45 hemisphere (Parrish-Aungst et al., 2007). There, they synapse onto various interneurons and
46 projection neurons. The latter directly project to a variety of cortical structures, including the
47 anterior olfactory nucleus, piriform cortex, cortical amygdala and the lateral entorhinal cortex
48 (Ghosh et al., 2011; Haberly and Price, 1977; Miyamichi et al., 2011; Sosulski et al., 2011).
49 This places OB projection neurons at a pivotal position to distribute processed olfactory
50 information broadly across the brain.

51

52 Each ORN in the mouse expresses only one of approximately 1000 olfactory receptor genes
53 (Buck and Axel, 1991; Niimura, 2012; Zhang and Firestein, 2002). ORNs expressing the same
54 receptor project axons onto defined spherical structures, glomeruli (Mori and Sakano, 2011),
55 containing a variety of neuropil including the apical dendrites of 10-50 projection neurons
56 (Bartel et al., 2015; Schwarz et al., 2018). Historically, OB projection neurons have been
57 divided into mitral and tufted cells (MCs, TCs), largely based on their soma location and
58 dendritic and axonal projection pattern (**Figure 1-figure supplement 1**) (Haberly and Price,
59 1977; Imamura et al., 2020; Mori et al., 1983; Orona et al., 1984): MC somata are located
60 predominantly in a thin layer with their dendrites covering the deeper part of the OB external
61 plexiform layer. Their axons project to a wide range of structures including posterior piriform
62 cortex. TC axons, on the other hand, are restricted to more anterior forebrain structures and
63 their cell bodies are distributed across the external plexiform layer, with dendrites largely
64 restricted to superficial layers. Within the TC population, several subdivisions have been made
65 into deep, middle, superficial and external TCs, largely based on soma position. MCs on the
66 other hand are often morphologically described as a largely homogeneous population.
67 However, branching patterns of lateral dendrites as well as soma size and apical dendrite
68 length might allow for further subdivision (Mouradian and Scott, 1988; Orona et al., 1984;

Schwarz et al., 2018). Moreover, projection patterns might differ based on soma position along the dorsomedial-ventrolateral axis of the OB (Inokuchi et al., 2017).

In parallel to this morphological diversity, numerous studies have described physiological heterogeneity both as a result of differential inputs from granule cells onto TCs and MCs (Christie et al., 2001; Ezeh et al., 1993; Geramita et al., 2016; Phillips et al., 2012) as well as intrinsic excitability and possibly glomerular wiring (Burton and Urban, 2014; Gire et al., 2019). Consequently, TCs respond more readily, with higher peak firing rates, and to lower odor concentration *in vivo* (Griff et al., 2008; Kikuta et al., 2013; Nagayama et al., 2014), and earlier in the respiration cycle compared to MCs (Ackels et al., 2020; Fukunaga et al., 2012; Igarashi et al., 2012; Jordan et al., 2018; Phillips et al., 2012).

Within the TC and MC populations, biophysical heterogeneity has been more difficult to tie to specific cell types or subtypes. MCs show diversity in biophysical properties that is thought to aid efficient encoding of stimulus-specific information and is, at least in part, experience-dependent (Angelo et al., 2012; Burton et al., 2012; Padmanabhan and Urban, 2010; Tripathy et al., 2013). Both *in vivo* and *in vitro* recordings suggest that a subset of MCs show regular firing, whilst others show ‘stuttering’ behavior characterized by irregular action potential clusters (Angelo et al., 2012; Balu et al., 2004; Bathellier et al., 2008; Buonviso et al., 2003; Carey and Wachowiak, 2011; Desmaisons et al., 1999; Fadool et al., 2011; Friedman and Strowbridge, 2000; Margrie and Schaefer, 2003; Padmanabhan and Urban, 2010; Schaefer et al., 2006). While TCs are heterogeneous, with for example external TCs displaying prominent rhythmic bursting, driving the glomerular circuitry into long-lasting depolarizations *in vitro* (De Saint Jan et al., 2009; Gire and Schoppa, 2009; Gire et al., 2019; Najac et al., 2011), a systematic assessment of biophysical variety is lacking so far. Moreover, differential centrifugal input from cortical and subcortical structures might further amplify this overall heterogeneity both between MCs and TCs as well as potentially within those different classes

96 (Boyd et al., 2012; Kapoor et al., 2016; Markopoulos et al., 2012; Niedworok et al., 2012;
97 Otazu et al., 2015).
98
99 Thus, anatomical projection patterns, *in vivo* odor responses, and intrinsic properties are
100 known to show substantial variability across different projection neurons. Systematic
101 investigation of different projection neurons, however, has been hampered by a scarcity of
102 specific molecular tools. Interneuron diversity, on the other hand, in general has received
103 considerable attention with numerous studies including in the OB (Parrish-Aungst et al., 2007;
104 Tavakoli et al., 2018) aiming to provide a systematic assessment of morphology, physiology,
105 chemotype and the basis for genetic targeting of distinct types of interneurons. For OB
106 projection neurons, however, only little information about chemotypes is available at this point:
107 *Pcdh1* and *Tbx21* (Haddad et al., 2013; Nagai et al., 2005) have been shown to be selectively
108 expressed in a subset of OB projection neurons. CCK distinguishes a subset of TCs
109 (superficial TCs, (Liu and Shipley, 1994; Seroogy et al., 1985; Short and Wachowiak, 2019;
110 Sun et al., 2020). Vasopressin expressing cells might constitute a further subset of superficial
111 TCs (Lukas et al., 2019), and recently the *Lbhd2* gene has been used to obtain more specific
112 genetic access to MCs (Koldaeva et al., 2020). Heterogeneous expression of both the GABAa
113 receptor as well as voltage-gated potassium channel subunits have been observed
114 (Padmanabhan and Urban, 2010; Panzanelli et al., 2005), albeit not linked to specific cell
115 types. Expression of axon guidance molecules such as *Nrp2* might further allow subdivision
116 of projection neurons across the OB (Inokuchi et al., 2017).
117
118 Hence, while some molecular markers can be used to define specific subsets of projection
119 neurons, this description is far from complete. A comprehensive molecular definition of
120 projection neuron types would help to classify and collate existing biophysical, morphological
121 and physiological data and delineate the distinct output streams of the OB. Moreover, it would
122 provide a platform upon which further focused experimental approaches could be tied.
123

124 Single cell or single nucleus RNA sequencing has been used effectively to map out cell type
125 across a variety of brain areas (Macosko et al., 2015; Zeisel et al., 2018), including inhibitory
126 interneurons in the mouse OB (Tepe et al., 2018). As M/TCs constitute only ~10% of all OB
127 neurons, we decided to enrich for projection neurons for single nucleus (sn)RNA-seq. We then
128 combined snRNA-seq with bulk RNA deep seq for OB neurons projecting to different cortical
129 areas, thereby allowing us to disentangle different projection neuron classes by target area.
130 We found that indeed both MCs and TCs fall into several, separable types, defined by
131 expression of both common and overlapping gene regulatory networks. This work will
132 therefore provide a molecular entry point into disentangling the diversity of OB projection
133 neurons and defining the functional roles of different MC/TC types.
134

135 **Results**

136

137 **Single nucleus RNA sequencing of olfactory bulb projection neurons distinguishes**
138 **mitral and tufted cell types**

139 To characterize the molecular diversity of OB projection neurons, we devised two
140 complementary experimental strategies. We used viral targeting and Fluorescence-Activated
141 Nuclei Sorting (FANS) to enrich for OB projection neurons, and we then characterized their
142 transcriptomes using single nucleus RNA sequencing (snRNA-seq) and bulk RNA deep
143 sequencing **(Figure 1A, Figure 1-figure supplement 2).**

144

145 First, we injected a retrogradely transported Adeno-Associated Virus expressing nuclear GFP
146 (rAAV-retro-CAG-H2B-GFP (Tervo et al., 2016)) into multiple sites along the antero-posterior
147 axis of the olfactory cortex (**Figure 1A).** Histological analysis revealed that virus injections
148 resulted in GFP expression in a heterogeneous population of OB cells labelling cells in the
149 mitral cell, external plexiform, glomerular and granule layers **(Figure 1B)**. Sparse GFP
150 expression in putative periglomerular and granule cells may have resulted from viral infection
151 of migrating immature neurons from the rostral migratory stream or from diffusion of the virus
152 from the injection site.

153

154 We dissected the olfactory bulbs of three injected mice, generated three independent
155 replicates of single nuclei, enriched for GFP expression using FANS (**Figure 1-figure**
156 **supplement 2)**, and performed snRNA-seq using 10X Genomics technology **(Figure 1A).** We
157 performed a detailed quality check of the individual replicates, then combined nuclei for
158 downstream analyses **(Figure 2-figure supplement 1)**. We analyzed a total of 31,703 nuclei,
159 grouped in 22 clusters that were annotated *post hoc* based on the expression of established
160 marker genes for excitatory and inhibitory neurons and glial cell populations **(Figure 2A-C).**
161 We initially used the combinatorial expression patterns of glutamatergic markers and

162 previously characterized M/T cell markers (*Vglut1, Vglut2, Vglut3, Tbx21, Pcdh21, Thy1)* to
163 identify putative OB projection neurons, comprising 23.66% (n=7504) of all nuclei.
164 Next, we further subclustered the selected profiles, resulting in a total of 10 molecularly distinct
165 subpopulations. To assign preliminary labels to each of these cell types, we used marker
166 genes previously employed in functional or single cell RNA-seq studies (Nagayama et al.,
167 2014; Tepe et al., 2018). We also used available RNA *in situ* hybridization data from the Allen
168 Institute for Brain Science to corroborate our preliminary assignments (**Figure 3-figure**
169 **supplement 1 and 2**). Our analysis revealed eight molecularly distinct clusters of putative
170 projection neurons and two clusters of putative periglomerular cells **(Figure 2D and E)**. Among
171 projection neurons, we identified three clusters of MCs (M1, M2, M3) and five clusters of
172 middle and external TCs (T1, ET1, ET2, ET3, ET4).
173
174 **smFISH validates mitral and tufted cell types**
175 To identify genes selectively expressed in OB projection neurons, we used the R-package
176 glmGamPoi (Ahlmann-Eltze and Huber, 2020). We calculated the combined average raw
177 expression of the top differentially expressed genes for each cell type and found it to be highly
178 specific for each cluster (**Figure 3A and C)**. We then selected a few specific marker genes
179 **(Figure 3B and D)** to validate projection neuron type identity by combining single molecule
180 Fluorescent *In Situ* Hybridization (smFISH) with GFP staining upon rAAV-retro-CAG-H2B-
181 GFP injection in olfactory cortex.
182
183 We first investigated MC type-specific gene expression. Differential expression (DE) analysis
184 identified the voltage-gated potassium channel *Kcng1*, the transcriptional regulator LIM
185 homeobox 5 (*Lhx5)* and the serine-rich transmembrane domain 1 (*Sertm1)* as putative M1-
186 specific marker genes. Two-color smFISH revealed selective co-localization of *Kcng1* and
187 *Lhx5* transcripts within the same subpopulation of cells in the MC layer (**Figure 3-figure**
188 **supplement 1E)**. Furthermore, *Kcng1, Lhx5* and *Sertm1* expression was consistently
189 observed in neurons expressing GFP **(Figure 3H, Figure 3-figure supplement 1B-D)**. Next,

190 DE analysis identified the mechanosensory ion channel *Piezo2*, the transcription cofactor
191 vestigial like family member 2 (*Vgll2)* and the zinc finger protein 114 (*Zfp114)* as putative M2-
192 specific markers. Two-color smFISH revealed extensive co-localization of *Piezo2* and *Vgll2*
193 transcripts within the same subpopulation of cells in the mitral cell layer **(Figure 3-figure**
194 **supplement 1J),** and co-localization of M2-specific marker genes with GFP **(Figure 3I, Figure**
195 **3-figure supplement 1F and G).** Finally, we identified the calcium-dependent secretion
196 activator 2 (*Cadps2)*, calcitonin (*Calca)* and follistatin (*Fst)* as putative M3-specific markers.
197 smFISH revealed selective and extensive co-localization of *Cadps2* with *Calca* or with *Fst*
198 transcripts within the same subpopulation of cells in the MC layer **(Figure 3G and G', Figure**
199 **3-figure supplement 1N).** Furthermore, *Cadps2, Calca* and *Fst* expression was consistently
200 observed in neurons expressing GFP **(Figure 3G and H, Figure 3-figure supplement 1K-**
201 **M).** Importantly, two-color smFISH revealed that type-specific M1, M2 and M3 markers were
202 expressed in largely non-overlapping populations of MCs: M1-specific *Kcng1* and *Lhx5*
203 transcripts did not co-localize with M2-specific *Vgll2* and *Piezo2* transcripts (**Figure 3-figure**
204 **supplement 1O and P**); M1-specific *Kcng1* transcripts did not co-localize with M3-specific
205 *Cadps2* and *Calca* transcripts **(Figure 3H and H', Figure 3-figure supplement 1Q)**; M2-
206 specific *Vgll2* transcripts did not co-localize with M3-specific *Cadps2* transcripts **(Figure 3I**
207 **and I').** Together, the selective co-localization of type-specific genes in non-overlapping
208 populations in the mitral cell layer validates these three types as accurate and meaningful
209 groupings of MCs, and their co-localization with GFP validates their identity as projection
210 neurons.
211
212 DE analysis for TC type-specific genes identified the transcription factor BarH-like homeobox
213 2 (*Barhl2)*, the gamma-sarcoglycan *Sgcg*, the vitamin D receptor (*Vdr)*, and the olfactory
214 receptors *Olfr110* and *Olfr111* as putative T1 markers. Two-color smFISH revealed extensive
215 co-localization of *Barhl2* and *Sgcg* or *Olfr110/Olfr111* transcripts within the same
216 subpopulation of cells in the external plexiform layer, indicative of middle tufted cells **(Figure**

217 **3J and J', Figure 3-figure supplement 2B)**. Furthermore, *Barhl2* and *Sgcg* expression was
218 observed in neurons expressing GFP **(Figure 3J).**
219 The coagulation factor C homolog (*Coch)* and the Wnt family member 5b (*Wnt5b)* were
220 identified as putative ET1 markers. smFISH confirmed the expression of *Coch* and *Wnt5b* in
221 a subpopulation of cells in the external plexiform and glomerular layers **(Figure 3K, figure**
222 **supplement Figure 2D)**. Moreover, *Coch* expression was observed in neurons expressing
223 GFP **(Figure 3K)**. The LIM homeobox 1 (*Lhx1)* and the early B-cell factor 3 (*Ebf3)* were
224 identified as putative ET2 markers. Two-color smFISH revealed *Lhx1* and *Ebf3* co-expression
225 in a sparse subpopulation of cells at the boundary between the external plexiform and
226 glomerular layers **(Figure 3-figure supplement 2F)**, indicative of external TCs. Finally, we
227 identified the lymphocyte antigen 6 family member 6GE (*Ly6g6e)* and the transcription factor
228 Forkhead box O1 (*Foxo1)* as putative ET4 markers. smFISH revealed selective expression of
229 *Ly6g6e* and *Foxo1* in a subpopulation of cells in the glomerular layer **(Figure 3-figure**
230 **supplement 2H),** and co-localization of *Ly6g6e* with GFP validates their identity as external
231 TCs **(Figure 3L).** Importantly, two-color smFISH revealed that type-specific tufted cell markers
232 were expressed in largely non-overlapping populations of cells: ET1-specific *Coch* transcripts
233 did not co-localize with T1-specific *Barhl2* transcript or with ET4-specific *Ly6g6e* transcripts
234 **(Figure 3K and K', 3L and L').** Furthermore, ET4-specific *Foxo1* and *Ly6g6e* transcripts did
235 not co-localize with T1-specific *Barhl2* transcript **(Figure 3-figure supplement 2I)**, with ET2-
236 specific *Ebf3* transcript **(Figure 3-figure supplement 2J)** or with ET1-specific *Wnt5b*
237 transcript **(Figure 3-figure supplement 2K)**. Overall, the selective co-localization of type-
238 specific genes, their location within the olfactory bulb, their non-overlapping nature, and their
239 co-localization with GFP validates these five types of middle and external TCs as accurate
240 and meaningful classifications.
241
242 **Inferring gene regulatory networks for projection neurons**
243 The differential gene expression patterns revealed by transcriptome analysis are determined
244 by the concerted action of transcription factors (TFs). We therefore set out to characterize cell

245 types by their TF activity, and we used independent information about TF binding sites to
246 group genes by TF interactions. These gene regulatory networks are biologically meaningful
247 and more robust against technical artifacts than the expression of individual genes, providing
248 a complementary set of axes along which to cluster MCs and TCs. Ultimately, gene regulatory
249 network analysis can yield more detail for classifying cell types and for understanding the
250 molecular mechanisms that underlie their transcriptional differences.
251
252 To infer the regulatory networks of each type of projection neuron, we used the Single-Cell
253 Regulatory Network Inference and Clustering pipeline (SCENIC, (Aibar et al., 2017)). SCENIC
254 is a three-step computational protocol based around regulons. A regulon is a TF and its
255 (predicted) target genes **(Figure 4A)**. In brief, SCENIC consists of co-expression analysis,
256 followed by TF binding motif enrichment analysis, and finally evaluation of a regulon's activity.
257 The results are a list of regulons and a matrix of all the single cells with their regulon activity
258 scores (RAS, essentially an Area-Under-the-Curve metric, see Methods for details).
259
260 **Clustering on regulon activity corroborates molecular groupings of mitral and tufted**
261 **cell types and allows further subdivision**
262 We applied SCENIC to MCs and TCs (6472 cells), computing 64 regulons with a range of 8
263 to 724 target genes (median = 41). These regulons greatly reduce the dimensionality of the
264 data from >30,000 genes to 64 regulons, defining a new biologically meaningful space in which
265 to analyze relationships between cells. Using the regulon activity matrix, we performed a
266 Leiden clustering (Traag et al., 2019; Wolf et al., 2018) in UMAP space on the putative
267 projection neurons and confirmed that doing so recapitulated the cell types we defined based
268 on transcriptome analysis **(Figure 4-figure supplement 1).** We observed that the two
269 approaches produced similar clusters, with only minor differences in drawing the boundaries
270 between clusters **(Figure 4-figure supplement 1)**. The strong overlap between these
271 classification methods validates the cell types we characterize as meaningful divisions within
272 the data.

273 In addition, constraining genome-wide transcriptome data by TF-target gene interactions
274 shows that the transcriptome-defined MC and TC types are heterogenous clusters in terms of
275 regulon activity **(Figure 4D)**. This suggests the existence of subtypes within mitral and tufted
276 cell types. To investigate these further, we defined more fine-grained groupings within cell
277 types. We used hierarchical clustering analysis to further subdivide cell types, finding 5
278 subtypes of M1, 2 subtypes each of M2 and M3, 5 subtypes each of T1 and ET1, and 3
279 subtypes of ET4. We confirmed that these subtypes were well-recovered with a Leiden
280 clustering on regulon activity **(Figure 4-figure supplement 2)**.
281
282 **Combinations of regulon modules characterize mitral and tufted cell subtypes**
283 TFs activity is thought to be organized into coordinated network modules that determine
284 cellular phenotypes (Alexander et al., 2009; Irons and Monk, 2007; Suo et al., 2018). To
285 characterize how TFs are organized into such modules in OB projection neurons, we searched
286 for combinatorial patterns of regulon activity. We used the Connection Specificity Index (CSI)
287 to this end, which is an association index known to be suited for the detection of modules
288 (Fuxman Bass et al., 2013; Suo et al., 2018). By computing the CSI on the basis of pairwise
289 comparisons of regulon activity patterns across cells, we found that the 68 regulons grouped
290 into 7 modules (mod1-7) **(Figure 4B and C)**. We confirmed these modules by looking at the
291 activity of individual regulons in each cell and visually verifying that individual regulons act
292 together as the identified modules **(Figure 4D)**.
293
294 Interestingly, regulon and module activity were not uniform within cell types. Rather, regulon
295 activity suggested distinct subtypes of each MC or TC type **(Figure 4D)**, corroborating that
296 using biologically relevant information to reduce dimensionality facilitates more fine-grained
297 classification. To further investigate these subtypes, we used the modules to describe how the
298 combined regulatory logic of distinct TFs contributes to the diversity of MC and TC subtypes.
299 We asked if combinations of modules could uniquely describe the subtypes. To do so, we
300 calculated the average module activity score per cell subtype. Next, we performed a

301 hierarchical clustering (correlation distance, complete linkage) on the subtypes **(Figure 4E)**
302 and mapped average module activity on the UMAP space defined by transcriptome analysis
303 **(Figure 4F)**. We found that, when grouped by module activity, subtypes do not group by type;
304 rather, subtypes of different cell types share similar module activity **(Figure 4E)**, providing
305 another complementary axis for grouping cells. Interestingly, module activity also forms
306 gradients along the UMAP plot, representing gradual transitions between subtypes **(Figure**
307 **4F)**. Therefore, while subtypes were identified most easily using modules of regulons, this
308 subtype structure is also apparent in the full transcriptome space, further validating the
309 divisions between subtypes as accurate and meaningful. Moreover, it makes explicit that a
310 continuous activity gradient of TFs is transformed in a non-linear manner into distinct
311 transcriptome differences between mitral and tufted cell types.
312

313 **Regulon-based transcription factor networks reveal overlapping features of cell type**
314 **identity**

315 While TFs regulate large numbers of target genes, central to cell identity are the interactions
316 between TFs: TFs can regulate their own expression as well as the expression of other TFs,
317 generating a TF network thought to be a core determinant of cell type identity (Arendt et al.,
318 2016; Becskei et al., 2001; Thieffry, 2007). We thus looked at TF-TF interactions to visualize
319 the TF network topology that defines MCs and TCs. We specifically asked whether MC and
320 TC classes share common TF network features, or whether, as suggested by the analysis of
321 genome-wide transcriptome and regulon analysis **(Figures 2-4)**, MC and TC subtypes are
322 defined by specific yet overlapping TF network features.
323 As a regulon is defined by a TF and a set of target genes, we constructed a (directed) network
324 of TFs by taking from each regulon's target genes only the TFs that have regulons themselves
325 **(Figure 5A)**. Overall, we found one large set of interconnected TFs, three small components
326 of five TFs or fewer, and twelve isolated TFs. 38 of the 64 TFs (59%) show possible self-
327 activating regulation, and several others form mutually-activating pairs (e.g. *Mxi1* and *Phf8*).
328 The network is dominated by three hub genes, two of which may self-activate: *Pbx3* (activates

329 19 TFs), *Bmyc* (activates 10 TFs) and *Bclaf1* (activates 7 TFs). Their central position in this
330 network suggests a role as hub regulators of MC and TC identity.
331 As anticipated from our module analysis, we find features that are shared across certain types
332 of MCs and TCs rather than MC- or TC-specific network features **(Figure 5B and C)**. For
333 instance, we find that M1 and ET2 are characterized by the hub *Pbx3* and *Dlx1*, which are
334 both self-activating and together form a positive feedback loop. Yet both cell types are
335 distinguishable by additional cycles: M1 has *Dlx2* "in the loop" with *Pbx3* and *Dlx1*, whilst ET2
336 has a 5-cycle involving *Pbx3*, *Dlx1*, *Dlx5*, *Sox2* and *Egr1* (i.e. a superset of the shared cycle
337 with M1). For M2 and T1, we observe that they are mostly characterized by peripheral nodes
338 of the network, that in most cases do not regulate each other, even if many are grouped
339 together in module 1 **(Figure 4C)**. And ET1 and ET4 are characterized by the target-gene-rich
340 3-cycle of *Ppargc1a* and the estrogen receptors *Esrra*, *Esrrg* (for ET4 the estrogen receptors
341 are in the top 20 of specific regulons), with ET4 also featuring the cycle between *Junb* and
342 *Fos*. Interestingly, these two cycles are connected, for instance via *Esrrg* and *Junb*.
343
344 Taken together, the analysis of TF regulatory networks suggests that individual MC and TC
345 types share key TF network features, which might point towards common physiology or
346 connectivity features. The differentially active TF network hubs and loops provide starting
347 points for future investigation of the functional differences between the MC and TC types
348 described here. Thus, the modules and the network serve as complementary approaches for
349 studying cell type identity, with modules suited to classifying cells into types and subtypes and
350 network analysis suited to investigating their functional differences.
351
352 **Simulating single nucleus gene expression from bulk RNA deep sequencing**
353 TCs preferentially target anterior olfactory regions, including the Anterior Olfactory Nucleus
354 (AON), while MCs target anterior and posterior olfactory areas (Imamura et al., 2020).
355 Therefore, we asked whether the genetic diversity between MC and TC types could provide

356 information about their projection targets. To investigate this question, we again injected rAAV-
357 retro-CAG-H2B-GFP into olfactory cortex, albeit now *either* into the AON *or* the posterior
358 Piriform Cortex (pPCx) **(Figure 1A).** For each injection site and in three replicates, we then
359 enriched GFP-expressing OB nuclei, and prepared RNA for bulk RNA deep sequencing
360 **(Figure 1A).**
361
362 Bulk RNA sequencing represents molecular information from a variety of different cell types.
363 Given that a substantial fraction of isolated nuclei in our experiments was comprised of granule
364 and periglomerular cells, in addition to projection neurons, we devised a novel computational
365 approach to recapitulate the constituent cell types from bulk RNA sequencing data by
366 simulating single nucleus expression profiles. Previous methods simulated the transcriptome
367 of a single cell based on the overall distribution of gene expression levels in the bulk RNA
368 sequencing data, producing many nuclei that were similar to each other and to the original
369 bulk expression profile (Konstantinides et al., 2018). This worked well for clean bulk RNA-seq
370 datasets with only one cell type, but for our mixed datasets, the simulated nuclei resembled
371 an unrealistic average of the constituent cell types. Therefore, to capture the diversity
372 contained within our bulk RNA-seq datasets, we used regulons as the unit of analysis to create
373 simulated nuclei with more biologically realistic transcriptomes (**Figure 6A and B**).
374
375 We first compared simulated and snRNA-seq nuclei by using principal component analysis to
376 project both into a shared low-dimensional space. We used these principal components as
377 input to a UMAP projection to visually inspect relationships between simulated and snRNA-
378 seq nuclei (**Figure 6B**). Consistent with histology and snRNA-seq analyses (**Figures 1 and**
379 **2**), simulations from both AON-injected and pPCx-injected bulk RNA-seq datasets contained
380 cells types other than projection neurons, and this contamination was more pronounced in
381 simulations from AON-injected bulk RNA-seq datasets. The dispersion of simulated nuclei
382 throughout this space indicated that simulations successfully recapitulated the diversity of cell
383 types in the bulk RNA-seq datasets, as each cell class from the combined snRNA-seq data

384 had some simulated nuclei in its vicinity (**Figure 6B**). To account for this contamination and
385 filter for only putative simulated projection neurons, we trained linear discriminant analysis
386 (LDA) classifiers to predict whether snRNA-seq nuclei were projection neurons based on the
387 30 top principal components that defined the shared low-dimensional space. These classifiers
388 accurately and consistently classified projection neurons, with a mean Jaccard index (a
389 measure of similarity between predicted and true labels) of 97.3% and a standard deviation of
390 0.17% over 1000 classifiers. We then applied this classifier to the simulated nuclei, designating
391 those simulations predicted to be projection neurons by all 1000 classifications as putative
392 simulated projection neurons (**Figure 6B**).
393
394 To directly compare these simulated projection neurons to MCs characterized through snRNA-
395 seq, we next used principal component analysis to define a shared low-dimensional space for
396 MCs and simulated projection neurons only. To investigate potential differences in projection
397 target between MC types, we trained 1000 LDA classifiers to predict the projection target of
398 simulated projection neurons based on the 30 top principal components that defined the
399 shared low-dimensional space (mean Jaccard index: 85.1%; standard deviation: 1.1%). We
400 then used these classifiers to predict the projection targets of snRNA-seq MCs. Interestingly,
401 we consistently found different predicted targets for the molecularly-defined MC types. 98.4%
402 of M1 MCs were classified as AON-projecting by at least 90% of classifiers, suggesting that
403 M1 cells preferentially project to anterior targets (**Figure 6C**). In contrast, 71.4% of M2 MCs
404 were consistently classified as pPCx-projecting, suggesting that M2 cells preferentially project
405 to posterior targets (**Figure 6C**). M3 MCs segregated into distinct populations: 48.2% of M3
406 MCs were consistently classified as AON-projecting, while 42.9% were consistently classified
407 as pPCx-projecting (**Figure 6C**). These findings suggest that the molecular subcategorization
408 of MCs may delineate differences in connectivity. For M1 and M2, these labels were defined
409 by gene expression and regulon activity, and projection target specificity. However, in the case
410 of M3, our results suggest that the molecular category contains neurons with distinct projection

411 patterns (**Figure 6D)**. These results suggest that neuronal connectivity provides an
412 independent axis along which to investigate cell type identity.
413
414

415 **Discussion**

416 Morphological differences between OB mitral and tufted cells have been described since the
417 time of Cajal (The Croonian lecture, 1894). Electrophysiological and functional imaging
418 experiments *in vivo* and *in vitro*, developmental studies as well as anatomical reconstructions
419 from light and electron microscopy studies have further highlighted the heterogeneity of OB
420 projection neurons (Christie et al., 2001; Ezeh et al., 1993; Fukunaga et al., 2012; Geramita
421 et al., 2016; Haberly and Price, 1977; Kawasawa et al., 2016; Mori et al., 1983; Phillips et al.,
422 2012). We here provide the first detailed molecular profiling of projection neurons of the mouse
423 olfactory bulb and delineate the types and subtypes of mitral and tufted cells together with
424 their key gene regulatory networks.

425 We have performed single nucleus and bulk RNA deep sequencing to characterize the
426 molecular diversity of mouse OB projection neurons. We identified, based on transcriptome
427 and RNA *in situ* analysis, three distinct types of MCs and five distinct types of TCs. We then
428 used comprehensive gene regulatory network analysis to reveal candidate gene regulatory
429 mechanisms that underlie cell type diversity. Finally, we describe a novel computational
430 approach for integrating single nucleus and bulk RNA sequencing data, and we use this
431 approach to propose that different MC types selectively project to anterior versus posterior
432 regions of olfactory cortex. Our analyses identified potential molecular determinants of cell
433 type-specific functional properties and projection target connectivity and provide a
434 comprehensive resource for investigating odor processing and olfactory circuit function and
435 evolution.

436

437 **The molecular diversity of olfactory bulb projection neurons**

438 Given that the vast majority of OB neurons are interneurons, notably granule cells and
439 juxtaglomerular neurons, we devised a retrograde viral targeting strategy to substantially
440 enrich for OB projection neurons. This allowed us to analyze the transcriptomes of over 7500
441 putative projection neurons that could in turn be grouped into 8 molecularly distinct mitral and
442 tufted cell types. We validated neuronal cell identity using smFISH with multiple type-specific

443 marker genes, and we determined the localization of identified neuronal types within the mitral
444 cell, external plexiform and glomerular layers of the olfactory bulb. Finally, we used retrograde
445 viral tracing to confirm that the molecularly distinct neuronal types we describe indeed project
446 to the olfactory cortex. Based on our analysis we define three molecularly distinct types of
447 MCs and five distinct types of TCs.
448 The neuronal cell types we have characterized here likely represent the major categories of
449 OB projection neurons. More extensive sampling might reveal additional rare cell types, and
450 more fine-grained clustering could further subdivide subtypes of neurons. However, our
451 samples contained 7,504 putative projection neurons compared to current estimates of 10,000
452 – 30,000 projection neurons overall per OB (Nagayama et al., 2014; Richard et al., 2010).
453 Furthermore, we sorted nuclei rather than whole cells, which is thought to more accurately
454 reflect relative cell type abundance (Habib et al., 2017; Lake et al., 2017). Altogether we are
455 therefore confident that our analysis captures the key biologically relevant types of projection
456 neurons. Independent from the number of molecularly distinct neuronal cell types, the gene
457 expression profiles we have described here provide critical new tools for refining projection
458 neuron cell type identities by aligning a cell's molecular features with its functional properties.
459 Previous experiments have highlighted the heterogeneity of odor responses of MCs and TCs
460 (Balu et al., 2004; Bathellier et al., 2008; Carey and Wachowiak, 2011; Desmaisons et al.,
461 1999; Friedman and Strowbridge, 2000; Schaefer et al., 2006). We propose that this functional
462 diversity can be explained, at least in part, by the molecular diversity of OB projection neurons,
463 a model that can now be tested experimentally.
464

465 **Specificity of mitral and tufted cell projections**

466 A critical feature of neuronal cell type identity is their projection target specificity. Earlier
467 studies have shown that TCs project to anterior regions of the olfactory cortex only, while MCs
468 project to both anterior and posterior olfactory cortex (Imamura et al., 2020; Nagayama et al.,
469 2010; Scott et al., 1980). Furthermore, and in contrast to the organization of neuronal
470 projections to sensory cortexes for vision, hearing, or touch, projections from the OB to the

471 piriform cortex lack apparent topographical organization (Ghosh et al., 2011; Miyamichi et al.,
472 2011; Sosulski et al., 2011).
473 We have analyzed, using bulk RNA deep sequencing, gene expression profiles of cells
474 projecting to anterior versus posterior olfactory cortex. Using simulations based on gene
475 regulatory network analysis, we have then mapped the bulk RNA sequencing data onto MC
476 types defined by single nucleus RNA sequencing. Interestingly, our analysis suggests that
477 cells of the M1 MC type preferentially target the anterior olfactory cortex, while M2 cells
478 preferentially target the posterior olfactory cortex. Furthermore, M3 cells, while clustering as a
479 single cell type based on their transcriptomes, project to either anterior or posterior cortical
480 targets. These results are consistent with the model that gene expression and connectivity
481 provide complementary axes for cell type classification (Kim et al., 2020). Importantly, our
482 results provide critical new molecular markers for investigating the anatomical organization of
483 connectivity from the OB to cortex.
484

485 **Molecular mechanisms underlying olfactory bulb projection neuron diversity**

486 Gene regulatory network analysis can reveal the transcriptional programs that determine the
487 functional properties of neuronal subtypes. Here, we describe cell type-specific modules of
488 gene regulation, defined by the interactions of transcription factors and their target genes. One
489 intriguing result of this analysis is that cell subtypes do not fall into clearly delineated MC and
490 TC classes. For example, module activity of M1d and M3b MC subtypes is more similar to
491 those of ET2 TCs than other M1 and M2 MC subtypes. We obtained similar results from
492 analyzing the TF-TF network that is thought to be closely linked to maintenance of cell identity.
493 This demonstrated that while some MC subtypes indeed share key TFs with each other,
494 regulon activity and network features in MC subtypes and TC subtypes are highly overlapping.
495 For example, the most prominent hub, *Pbx3,* is strongly active in M1 and ET2 but not any of
496 the other MC or TC subtypes. Generally, projection neurons are characterized by a variety of
497 hubs and cycles in the TF-TF network that are used by both MC and TC subtypes in a
498 combinatorial manner. The prominent hub and cycle-related genes we have identified here,

499 may act as candidate master regulators of neuronal function, which can be targeted for
500 experimental validation. Together, these results suggest that subtypes of MCs and TCs may
501 share important functional properties, possibly blurring at the transcriptional level the
502 traditional division into tufted and mitral cells as the two major classes of OB projection
503 neurons. Moreover, the gradients of module activity that we observed over the MC and TC
504 subtypes theoretically provide a mechanism for generating multiple distinct cellular
505 phenotypes, similar to how morphogenetic gradients allow for spatial patterning and cell
506 differentiation (Wolpert, 1969). Through non-linear regulatory interactions gradual differences
507 at the transcriptomic level can be translated into selective expression of functional genes.
508 For example, we found that the *Kcng1* gene was selectively expressed in M1 MCs. The *Kcng1*
509 gene encodes for a voltage-gated potassium channel, which forms heterotetrameric channels
510 with the ubiquitously expressed delayed rectifying Kv2.1 potassium channel (indeed also
511 expressed in the M1 cluster) and modifies the kinetics of channel activation and deactivation
512 (Kramer et al., 1998). Other voltage-gated potassium channels exhibiting prominent
513 differential expression levels in MC and TC subtypes include *Kcnd3, Kcng1, Kcnh5, Kcnq3,*
514 *Kcnj2* and 6, and *Hcn1* (for details see accompanying website link in Materials and Methods).
515 These channels represent intriguing candidates for controlling the differential excitability of
516 different MC and TC types.
517 We also found that a large number of cell adhesion and axon guidance genes known to control
518 the formation and maintenance of neuronal connectivity were differentially expressed in OB
519 projection neuron types. Examples include members of the cadherin superfamily of cell
520 adhesion glycoproteins (*Cdh6, 7, 8c, 9, 13,* and *20*), and components of the
521 Semaphorin/Neuropilin/Plexin complexes, including *Nrp2, Plexna3, Sema3e,* and *Sema5b.*
522 Semaphorin/Neuropilin/Plexin complexes are known to play critical roles in the development
523 and maintenance of neuronal connections, including in OB MCs (Inokuchi et al., 2017; Saha
524 et al., 2007). Heterogeneity in these cell adhesion and guidance genes might inform
525 subdivisions in projection neurons across the OB, in particular along the dorsomedial-
526 ventrolateral axis.

Our data set provides an important resource for studying the evolution of olfactory neural circuits across species. Adaptation to distinct olfactory environments, and the critical role of olfaction in survival and reproduction has shaped the evolution of the repertoire of odorant receptors and olfactory sensory neurons (Bargmann, 2006; Niimura, 2012). However, little is known about how evolving sensory inputs from the nose are accommodated at the level of the OB and its connections to the olfactory cortex. A detailed molecular description of mouse OB projection neurons provides a first step towards understanding the evolution of olfactory sensory processing across species.

538 **Materials and methods**

539

540 **Experimental model and subject details**

541 Male and female C57Bl/6 mice (6- to 8-week-old) were used in this study and obtained by in-
542 house breeding. All animal protocols were approved by the Ethics Committee of the board of
543 the Francis Crick Institute and the United Kingdom Home Office under the Animals (Scientific
544 Procedures) Act 1986, as well as Brown University's Institutional Animal Care and Use
545 Committee followed by the guidelines provided by the National Institutes of Health.

546

547 **Stereotaxic injections and histology**

548 Mice were anaesthetized using isoflurane and prepared for aseptic surgery in a stereotaxic
549 frame (David Kopf Instruments). A retrogradely transported Adeno Associated Virus (rAAV-
550 retro-CAG-H2B-GFP, (Tervo et al., 2016)) was injected stereotaxically into multiple sites of
551 piriform cortex (PCx) and anterior olfactory nucleus (AON). The following coordinates, based
552 on the Paxinos and Franklin Mouse Brain Atlas, were used: Coordinates (AP / ML / DV) in mm
553 for PCx injections: (1) -0.63 / -4.05 / -4.10, (2) -0.8 / -4.00 / -4.10. For AON injections: (1) 2.8
554 / 1.25 / 2.26 and 2.6, (2) 2.68 / 1.25 / 2.3 and 2.75, (3) 2.34 / 0.7 / 3.5. Using a
555 micromanipulator, a pulled glass micropipette was slowly lowered into the brain and left in
556 place for 30 seconds before the virus was dispensed from the micropipette using a Nanoject
557 injector (Drummond Scientific) at a rate of 46 nl/min (0.3 µl for PCx and 0.2 µl for AON per
558 injection site). The micropipette was left in place for an additional 5 min before being slowly
559 withdrawn to minimize diffusion along the injection tract. Craniotomies were covered with
560 silicone sealant (WPI) and the skin was sutured. Mice were provided with 5 mg/kg Carprofen
561 in their drinking water for 2 days following surgery.

562

563 Histology was used to validate viral targeting of olfactory bulb projection neurons. Mice were
564 deeply anaesthetized with 2.5% of 250mg/kg Avertin and transcardially perfused with 10 ml of
565 ice-cold phosphate-buffered saline (PBS) followed by 10 ml of 4% paraformaldehyde (PFA).

566 Brains were dissected and post-fixed for 5 h in 4 % PFA at 4 °C. Coronal sections (100 µm
567 thick) were prepared using a vibrating-blade Leica VT100S Vibratome. Sections were rinsed
568 in PBS and incubated in PBS / 0.1% Triton X-100 and Neurotrace counterstain (1:1000,
569 ThermoFisher) at 4 °C overnight, then mounted on SuperFrost Premium microscope slides
570 (Fisher, cat# 12-544-7) in Fluorescent Vectashield Mounting Medium (Vector). Images were
571 acquired at 20X using a Nikon A1R-HD confocal microscope.
572
573 **Single nuclei isolation, FANS and RNA extraction**
574 To isolate GFP-labeled nuclei, 9 individual replicates were used (for bulk RNA deep
575 sequencing: 3 replicates of AON-injected mice and 3 replicates of PCx-injected mice; for single
576 nuclei RNA sequencing: 3 replicates of AON and PCx-injected mice). Mice were deeply
577 anaesthetized with an overdose of ketamine/xylazine and transcardially perfused with ice-cold
578 PBS. Both hemispheres of the olfactory bulb were dissected, and the hemisphere ipsilateral
579 to the injection site was carefully separated from the contralateral hemisphere. Both
580 hemispheres were minced separately and placed into two different tubes. To dissociate single
581 nuclei, Nuclei PURE Prep was used according to the manufacturer instructions (Sigma, cat#
582 NUC201-1KT) with some modifications. The minced tissue was gently homogenized in 2.75
583 ml Nuclei PURE Lysis Buffer and 27.5 µl 10% Triton X-100 using an ice-cold dounce and
584 pestle, and filtered two times through a 40 µm cell strainer on ice. After centrifuging at 500
585 rpm for 5 min at 4 °C, the supernatant was aspirated and gently resuspended in 500 µl of cold
586 buffer (1x of cold Hanks' Balanced Salt Solution HBSS, 1% nuclease-free BSA, 22.5 µl of
587 RNasin Plus (Promega N2611) and 1/2000 DRAQ5).
588 Fluorescence-activated nuclei sorting of single nuclei was performed using a BD FACSAria™
589 III Cell Sorter with a 70 µm nozzle at a sheath pressure of 70 psi. Precision mode (yield mask
590 set to 16, purity mask set to 16 and phase mask set to 0) was used for stringent sorting. For
591 single nucleus RNA sequencing, GFP+ nuclei were sorted into a 1.5 ml centrifuge tube. For
592 bulk RNA deep sequencing, GFP+ nuclei were sorted into 100 µl Trizol and 1.43 µl of RNA

593 carrier, and total RNA was extracted using the Arcturus PicoPure RNA Isolation Kit
594 (ThermoFisher, cat# KIT0204).
595
596 **Single nucleus RNA sequencing**
597 Libraries were prepared using the Next Single Cell / Low Input RNA Library Prep Kit (New
598 England Biolabs). The quality and quantity of the final libraries were assessed with the
599 TapeStation D5000 Assay (Agilent Technologies) before sequencing with an Illumina HiSeq
600 4000 platform using the 10X kit version Chromium Single Cell 3' v3. RNA concentrations were
601 measured as: 14.4, 23.3, 7.9 ng/µl (n = 3 animals).
602
603 **Bulk RNA deep sequencing**
604 Libraries were prepared using the Next Single Cell / Low Input RNA Library Prep Kit (New
605 England Biolabs). The quality and quantity of the final libraries was assessed with the
606 TapeStation D1000 Assay (Agilent Technologies) before sequencing with an Illumina HiSeq
607 4000 platform. RNA concentrations were measured as: AON injections (n = 3 animals), 1.495,
608 1.682, 1.881 ng/µl and RNA integrity numbers (RIN) 8.3, 8.7, 9.0; PCx injections (n = 3
609 animals), 0.257, 0.165, 0.133 ng/µl; RIN = 8.0, 10.0, 7.8 for each replicate, respectively.
610
611 **Single nucleus RNA sequencing analysis**
612 Raw sequencing datasets were processed using the Cell Ranger pipeline (10x Genomics).
613 Count tables were loaded into R (version 3.6, https://www.r-project.org) and further processed
614 using the Seurat 3 R-package (Butler et al., 2018).
615 For each of the three replicates, we removed all nuclei with fewer than 500 distinct genes
616 detected or with more than 5% of unique molecular identifiers stemming from mitochondrial
617 genes. After quality control, we merged the replicates and retained a total of 31,703 nuclei
618 (median of 2,300 genes per nucleus; for each replicate median genes per nucleus: R1=2,266;
619 R2=2419; R3=2,322). Principal component analysis (PCA) was then performed on
620 significantly variable genes and the first 30 principal components were selected as input for

621 clustering and UMAP, based on manual inspection of a principal component variance plot (PC
622 elbow plot). Clustering was performed using the default method (Louvain) from the Seurat
623 package, with the resolution parameter of the FindClusters function set to 0.3.
624 Subclustering of projection neurons was carried out by selecting clusters M1, M2/M3, T1, ET1
625 and ET2 from the initial single-nuclei analysis based on the combinatorial expression patterns
626 of glutamatergic and previously characterized mitral/tufted cell markers (*Tbx21, Pcdh21, Thy1,*
627 *Vglut1, Vglut2 and Vglut3).* Subclustered nuclei were subjected to a new clustering with the
628 Seurat resolution parameter of the FindClusters function set to 0.3.
629 Differential gene expression analysis on single-nuclei data was performed using the
630 glmGamPoi R-package (Ahlmann-Eltze and Huber, 2020). Gene set enrichment analysis
631 (GSEA) on the resulting log-fold changes was performed as described in (Subramanian et al.,
632 2005).
633
634 **Network inference**
635 Gene regulatory networks were inferred using the pySCENIC pipeline (Single-Cell rEgulatory
636 Network InferenCe, (Aibar et al., 2017)) and visualized using Jupyter notebooks and
637 Cytoscape (Shannon et al., 2003). pySCENIC is a three-step approach: (1) predict TF-target
638 gene pairs using Arboreto; (2) filter TF-target gene associations for false positives using TF
639 binding site enrichment in a window of 5kb around a target's Transcription Start Site (TSS)
640 and group TFs with their target genes into so-called regulons; (3) calculate the activity of
641 regulons in each cell in terms of the Area Under the recovery Curve (AUC). Step 1 depends
642 on a stochastic search algorithm and is therefore performed n = 100 times. Only TFs that are
643 found >80 times and with TF-target gene interactions that occur >80 times are considered. To
644 avoid technical issues in the analysis, regulons with fewer than 8 target genes are removed
645 from the final list. Subsequent analysis in Step 3 involves a stochastic downsampling to speed
646 up computation, hence we verified that the chosen sample size was sufficient for accurate
647 AUC approximations. We calculated n = 25 AUC matrices and confirmed that they contained

648 few zeros and the variance of each matrix entry (i.e. approximated regulon activity in a given
649 cell) was low.
650
651 **Bulk RNA deep sequencing analysis**
652 The 'Trim Galore!' utility version 0.4.2 was used to remove sequencing adaptors and to quality
653 trim individual reads with the q-parameter set to 20. Sequencing reads were then aligned to
654 the mouse genome and transcriptome (Ensembl GRCm38 release-89) using RSEM version
655 1.3.0 (Li and Dewey, 2011) in conjunction with the STAR aligner version 2.5.2 (Dobin et al.,
656 2013) . Sequencing quality of individual samples was assessed using FASTQC version 0.11.5
657 and RNA-SeQC version 1.1.8 (DeLuca et al., 2012). Differential gene expression was
658 determined using the R-bioconductor package DESeq2 version 1.24.0 (Love et al., 2014).
659 Gene set enrichment analysis (GSEA) was conducted as described in (Subramanian et al.,
660 2005).
661
662 **Integration of single nucleus and bulk RNA deep sequencing data**
663 Nuclei were simulated from each bulk RNA-seq replicate using a weighted random sampling
664 of regulons with replacement. A regulon's relative weight corresponded to the prevalence of
665 its transcription factor in the given bulk RNA-seq sample. Each time a regulon was selected,
666 the counts for its transcription factor and all target genes increased by one. The number of
667 regulons expressed in each simulated nucleus was randomly selected from a list of how many
668 unique transcription factors each snRNA-seq nucleus expressed (normalized expression > 2).
669 Simulated nuclei were treated as raw count matrices and integrated with snRNA-seq nuclei
670 using the SCT package in R (Hafemeister and Satija, 2019). To filter simulated nuclei, we
671 trained 1000 linear discriminant analysis (LDA) classifiers with the python package scikit-learn
672 (Pedregosa et al., 2011). For each classifier, snRNA-seq nuclei were split into test and train
673 datasets, with 75% of nuclei used for training and the other 25% used for test. Each classifier
674 was trained to predict whether a nucleus was a projection neuron (whether it was selected for
675 subclustering in the initial Seurat analysis) based on values for the 30 top principal

676 components from the SCT integration. Each classifier was applied to the remaining snRNA-
677 seq nuclei for testing, and accuracy was evaluated using the Jaccard index calculated by
678 scikit-learn (Pedregosa et al., 2011). The classifiers were then applied to the simulated nuclei.
679 Simulated nuclei predicted to be projection neurons by all 1000 classifiers were designated as
680 putative simulated projection neurons and selected for further analysis. Similarly, these
681 putative simulated projection neurons were integrated with snRNA-seq mitral cells using SCT.
682 1000 LDA classifiers were trained to classify simulated nuclei as AON-projecting or PCx-
683 projecting based on values for the 30 top principal components from the SCT integration. Each
684 classifier was trained on 75% of the simulated projection neurons and tested on the other 25%,
685 with accuracy evaluated using the Jaccard index. Each classifier was then applied to snRNA-
686 seq mitral cells.
687
688 **smFISH in tracing experiments**
689 Experiments were performed according to the manufacturer's instructions, using the
690 RNAscope Fluorescent Multiplex kit (Advanced Cell Diagnostics (ACD)) for fresh frozen
691 tissue. Briefly, a total of 6 mice were injected into the AON and PCx with the rAAVretro-CAG-
692 H2B-GFP. After 15 days post-injection, mice were deeply anaesthetized with 2.5% of
693 250mg/kg Avertin and transcardially perfused with 10 ml of ice-cold phosphate-buffered saline
694 (PBS). The brains were dissected out from the skull, immediately embedded in Tissue Plus
695 O.C.T compound (Fisher Healthcare) and snap frozen in a bath of 2-methylbutane on dry ice.
696 Brains were cryo-sectioned coronally at 20 µm thickness, mounted on Fisherbrand™
697 Superfrost™ Plus microscope slides (Fisher Scientific) and stored at -80˚C until use. *In situ*
698 probes against the following mouse genes were ordered from ACD and multiplexed in the
699 same permutations across sections: *Foxo1* (#485761-C2 and 485761), *Kcng1* (#514181-C2),
700 *Lxh1* (#488581), *Sertm1* (#505401-C2), *Ebf3*(#576871-C3), *Sgcg* (#488051-C3), *Cadps2*
701 (#529361-C3 and 529361), *Coch* (#530911-C3), *Ly6g6e* (#506391-C2), *Wnt5b* (#405051), *Fst*
702 (#454331), *Barhl2* (#492331-C2), *Vdr* (524511-C3), *Gfp* (#409011, #409011-C2 and #409011-
703 C3), *Piezo2* (#500501), *Olfr110/111* (#590641), *Calca* (#578771), *Lhx5* (#885621-C3), and

704 *Vgll2* (#885631-C2). Following smFISH, high resolution images of a single z-plane were
705 obtained using a 60x oil immersion objective on an Olympus FV3000 confocal microscope
706 and a 40x oil immersion objective on a Nikon A1R-HD confocal microscope.
707
708 **Data and code availability**
709 Raw single nucleus RNA and bulk RNA deep sequencing data have been deposited in Gene
710 Expression Omnibus (GEO) under the accession numbers GSE162654 and GSE162655
711 respectively. The R and Python analysis scripts developed for this paper are available at the
712 GitLab links https://gitlab.com/fleischmann-lab/molecular-characterization-of-projection-
713 neuron-subtypes-in-the-mouse-olfactory-bulb and https://gitlab.inria.fr/acrombac/projection-
714 neurons-mouse-olfactory-bulb. Extensive computational tools for additional in-depth
715 exploration of our data sets are available through our website:
716 https://biologic.crick.ac.uk/OB_projection_neurons.
717
718

Acknowledgements

We thank the Crick Advanced Sequencing Facility, especially Robert Goldstone and Amelia Edwards for their excellent support. We thank Debipriya Das and Ana Agua-Doce from the Crick Flow Cytometry Facility for technical assistance, and the Crick animal facility for animal care. We thank the members of the Crick Digital Development Team, particularly Amy Strange, Luke Nightingale, Jude Pinnock and Marc Pollitt for excellent technical support. We thank the Harvard Neurobiology Imaging Facility for consultation and RNAscope services that supported this work. This facility is supported in part by the Neural Imaging Center as part of an NINDS P30 Core Center grant #NS072030. We thank Keeley Baker, Gilad Barnea, Bob Datta, Kevin Franks and Paul Greer for critical comments on the manuscript. Work in the ATS lab was supported by the Francis Crick Institute, which receives its core funding from Cancer Research UK (FC001153), the UK Medical Research Council (FC001153), and the Wellcome Trust (FC001153); a Wellcome Trust Investigator grant to ATS (110174/Z/15/Z), and a DFG postdoctoral fellowship to TA. Work is the AF lab was supported by grants from the NIH (1U19NS112953-01, 1R01DC017437-03) and the Robert J and Nancy D Carney Institute for Brain Science.

Competing interests

The authors declare that no competing interests exist.

742 **References**

743 Ackels, T., Jordan, R., Schaefer, A.T., and Fukunaga, I. (2020). Respiration-Locking of
744 Olfactory Receptor and Projection Neurons in the Mouse Olfactory Bulb and Its Modulation
745 by Brain State. Front. Cell. Neurosci. *14*, 220.

746 Ahlmann-Eltze, C., and Huber, W. (2020). glmGamPoi: Fitting Gamma-Poisson Generalized
747 Linear Models on Single Cell Count Data (Bioinformatics).

748 Aibar, S., González-Blas, C.B., Moerman, T., Huynh-Thu, V.A., Imrichova, H., Hulselmans, G.,
749 Rambow, F., Marine, J.-C., Geurts, P., Aerts, J., et al. (2017). SCENIC: single-cell regulatory
750 network inference and clustering. Nat. Methods *14*, 1083–1086.

751 Alexander, R.P., Kim, P.M., Emonet, T., and Gerstein, M.B. (2009). Understanding modularity
752 in molecular networks requires dynamics. Sci. Signal. *2*, pe44.

753 Angelo, K., Rancz, E.A., Pimentel, D., Hundahl, C., Hannibal, J., Fleischmann, A., Pichler, B.,
754 and Margrie, T.W. (2012). A biophysical signature of network affiliation and sensory
755 processing in mitral cells. Nature *488*, 375–378.

756 Arendt, D., Musser, J.M., Baker, C.V.H., Bergman, A., Cepko, C., Erwin, D.H., Pavlicev, M.,
757 Schlosser, G., Widder, S., Laubichler, M.D., et al. (2016). The origin and evolution of cell
758 types. Nat. Rev. Genet. *17*, 744–757.

759 Balu, R., Larimer, P., and Strowbridge, B.W. (2004). Phasic stimuli evoke precisely timed
760 spikes in intermittently discharging mitral cells. J. Neurophysiol. *92*, 743–753.

761 Bargmann, C.I. (2006). Comparative chemosensation from receptors to ecology. Nature *444*,
762 295–301.

763 Bartel, D.L., Rela, L., Hsieh, L., and Greer, C.A. (2015). Dendrodendritic synapses in the
764 mouse olfactory bulb external plexiform layer. J. Comp. Neurol. *523*, 1145–1161.

765 Bathellier, B., Buhl, D.L., Accolla, R., and Carleton, A. (2008). Dynamic ensemble odor coding
766 in the mammalian olfactory bulb: sensory information at different timescales. Neuron *57*,
767 586–598.

768 Becskei, A., Séraphin, B., and Serrano, L. (2001). Positive feedback in eukaryotic gene
769 networks: cell differentiation by graded to binary response conversion. EMBO J. *20*, 2528–
770 2535.

771 Boyd, A.M., Sturgill, J.F., Poo, C., and Isaacson, J.S. (2012). Cortical feedback control of
772 olfactory bulb circuits. Neuron *76*, 1161–1174.

773 Buck, L., and Axel, R. (1991). A novel multigene family may encode odorant receptors: a
774 molecular basis for odor recognition. Cell *65*, 175–187.

775 Buonviso, N., Amat, C., Litaudon, P., Roux, S., Royet, J.-P., Farget, V., and Sicard, G. (2003).
776 Rhythm sequence through the olfactory bulb layers during the time window of a respiratory
777 cycle. Eur. J. Neurosci. *17*, 1811–1819.

778 Burton, S.D., and Urban, N.N. (2014). Greater excitability and firing irregularity of tufted cells
779 underlies distinct afferent-evoked activity of olfactory bulb mitral and tufted cells. J. Physiol.
780 *592*, 2097–2118.

781 Burton, S.D., Ermentrout, G.B., and Urban, N.N. (2012). Intrinsic heterogeneity in oscillatory
782 dynamics limits correlation-induced neural synchronization. J. Neurophysiol. *108*, 2115–
783 2133.

784 Butler, A., Hoffman, P., Smibert, P., Papalexi, E., and Satija, R. (2018). Integrating single-cell
785 transcriptomic data across different conditions, technologies, and species. Nat. Biotechnol.
786 *36*, 411–420.

787 Carey, R.M., and Wachowiak, M. (2011). Effect of sniffing on the temporal structure of
788 mitral/tufted cell output from the olfactory bulb. J. Neurosci. Off. J. Soc. Neurosci. *31*,
789 10615–10626.

790 Christie, J.M., Schoppa, N.E., and Westbrook, G.L. (2001). Tufted cell dendrodendritic
791 inhibition in the olfactory bulb is dependent on NMDA receptor activity. J. Neurophysiol. *85*,
792 169–173.

793 De Saint Jan, D., Hirnet, D., Westbrook, G.L., and Charpak, S. (2009). External tufted cells
794 drive the output of olfactory bulb glomeruli. J. Neurosci. Off. J. Soc. Neurosci. *29*, 2043–
795 2052.

796 DeLuca, D.S., Levin, J.Z., Sivachenko, A., Fennell, T., Nazaire, M.-D., Williams, C., Reich, M.,
797 Winckler, W., and Getz, G. (2012). RNA-SeQC: RNA-seq metrics for quality control and
798 process optimization. Bioinforma. Oxf. Engl. *28*, 1530–1532.

799 Desmaisons, D., Vincent, J.D., and Lledo, P.M. (1999). Control of action potential timing by
800 intrinsic subthreshold oscillations in olfactory bulb output neurons. J. Neurosci. Off. J. Soc.
801 Neurosci. *19*, 10727–10737.

802 Dobin, A., Davis, C.A., Schlesinger, F., Drenkow, J., Zaleski, C., Jha, S., Batut, P., Chaisson, M.,
803 and Gingeras, T.R. (2013). STAR: ultrafast universal RNA-seq aligner. Bioinforma. Oxf. Engl.
804 *29*, 15–21.

805 Ezeh, P.I., Wellis, D.P., and Scott, J.W. (1993). Organization of inhibition in the rat olfactory
806 bulb external plexiform layer. J. Neurophysiol. *70*, 263–274.

807 Fadool, D.A., Tucker, K., and Pedarzani, P. (2011). Mitral cells of the olfactory bulb perform
808 metabolic sensing and are disrupted by obesity at the level of the Kv1.3 ion channel. PloS
809 One *6*, e24921.

810 Friedman, D., and Strowbridge, B.W. (2000). Functional role of NMDA autoreceptors in
811 olfactory mitral cells. J. Neurophysiol. *84*, 39–50.

812 Fukunaga, I., Berning, M., Kollo, M., Schmaltz, A., and Schaefer, A.T. (2012). Two distinct
813 channels of olfactory bulb output. Neuron *75*, 320–329.

814 Fuxman Bass, J.I., Diallo, A., Nelson, J., Soto, J.M., Myers, C.L., and Walhout, A.J.M. (2013).
815 Using networks to measure similarity between genes: association index selection. Nat.
816 Methods *10*, 1169–1176.

817 Geramita, M.A., Burton, S.D., and Urban, N.N. (2016). Distinct lateral inhibitory circuits drive
818 parallel processing of sensory information in the mammalian olfactory bulb. ELife *5*.

819 Ghosh, S., Larson, S.D., Hefzi, H., Marnoy, Z., Cutforth, T., Dokka, K., and Baldwin, K.K.
820 (2011). Sensory maps in the olfactory cortex defined by long-range viral tracing of single
821 neurons. Nature *472*, 217–220.

822 Gire, D.H., and Schoppa, N.E. (2009). Control of on/off glomerular signaling by a local
823 GABAergic microcircuit in the olfactory bulb. J. Neurosci. Off. J. Soc. Neurosci. *29*, 13454–
824 13464.

825 Gire, D.H., Zak, J.D., Bourne, J.N., Goodson, N.B., and Schoppa, N.E. (2019). Balancing
826 Extrasynaptic Excitation and Synaptic Inhibition within Olfactory Bulb Glomeruli. ENeuro *6*.

827 Griff, E.R., Mafhouz, M., and Chaput, M.A. (2008). Comparison of identified mitral and
828 tufted cells in freely breathing rats: II. Odor-evoked responses. Chem. Senses *33*, 793–802.

829 Haberly, L.B., and Price, J.L. (1977). The axonal projection patterns of the mitral and tufted
830 cells of the olfactory bulb in the rat. Brain Res. *129*, 152–157.

831 Habib, N., Avraham-Davidi, I., Basu, A., Burks, T., Shekhar, K., Hofree, M., Choudhury, S.R.,
832 Aguet, F., Gelfand, E., Ardlie, K., et al. (2017). Massively parallel single-nucleus RNA-seq with
833 DroNc-seq. Nat. Methods *14*, 955–958.

834 Haddad, R., Lanjuin, A., Madisen, L., Zeng, H., Murthy, V.N., and Uchida, N. (2013). Olfactory
835 cortical neurons read out a relative time code in the olfactory bulb. Nat. Neurosci. *16*, 949–
836 957.

837 Hafemeister, C., and Satija, R. (2019). Normalization and variance stabilization of single-cell
838 RNA-seq data using regularized negative binomial regression. Genome Biol. *20*, 296.

839 Igarashi, K.M., Ieki, N., An, M., Yamaguchi, Y., Nagayama, S., Kobayakawa, K., Kobayakawa,
840 R., Tanifuji, M., Sakano, H., Chen, W.R., et al. (2012). Parallel mitral and tufted cell pathways
841 route distinct odor information to different targets in the olfactory cortex. J. Neurosci. Off. J.
842 Soc. Neurosci. *32*, 7970–7985.

843 Imamura, F., Ito, A., and LaFever, B.J. (2020). Subpopulations of Projection Neurons in the
844 Olfactory Bulb. Front. Neural Circuits *14*, 561822.

845 Inokuchi, K., Imamura, F., Takeuchi, H., Kim, R., Okuno, H., Nishizumi, H., Bito, H., Kikusui, T.,
846 and Sakano, H. (2017). Nrp2 is sufficient to instruct circuit formation of mitral-cells to
847 mediate odour-induced attractive social responses. Nat. Commun. *8*, 15977.

848 Irons, D.J., and Monk, N.A. (2007). Identifying dynamical modules from genetic regulatory
849 systems: applications to the segment polarity network. BMC Bioinformatics *8*, 413.

850 Jordan, R., Fukunaga, I., Kollo, M., and Schaefer, A.T. (2018). Active Sampling State
851 Dynamically Enhances Olfactory Bulb Odor Representation. Neuron *98*, 1214-1228.e5.

852 Kapoor, V., Provost, A.C., Agarwal, P., and Murthy, V.N. (2016). Activation of raphe nuclei
853 triggers rapid and distinct effects on parallel olfactory bulb output channels. Nat. Neurosci.
854 *19*, 271–282.

855 Kawasawa, Y.I., Salzberg, A.C., Li, M., Sestan, N., Greer, C.A., and Imamura, F. (2016). RNA-
856 seq analysis of developing olfactory bulb projection neurons. Mol. Cell. Neurosci. *74*, 78–86.

857 Kikuta, S., Fletcher, M.L., Homma, R., Yamasoba, T., and Nagayama, S. (2013). Odorant
858 response properties of individual neurons in an olfactory glomerular module. Neuron *77*,
859 1122–1135.

860 Kim, E.J., Zhang, Z., Huang, L., Ito-Cole, T., Jacobs, M.W., Juavinett, A.L., Senturk, G., Hu, M.,
861 Ku, M., Ecker, J.R., et al. (2020). Extraction of Distinct Neuronal Cell Types from within a
862 Genetically Continuous Population. Neuron *107*, 274-282.e6.

863 Koldaeva, A., Zhang, C., Huang, Y.-P., Reinert, J., Mizuno, S., Sugiyama, F., Takahashi, S.,
864 Soliman, T., Matsunami, H., and Fukunaga, I. (2020). Transcriptome analysis for the
865 development of cell-type specific labeling to study olfactory circuits (Neuroscience).

866 Konstantinides, N., Kapuralin, K., Fadil, C., Barboza, L., Satija, R., and Desplan, C. (2018).
867 Phenotypic Convergence: Distinct Transcription Factors Regulate Common Terminal
868 Features. Cell *174*, 622-635.e13.

869 Kramer, J.W., Post, M.A., Brown, A.M., and Kirsch, G.E. (1998). Modulation of potassium
870 channel gating by coexpression of Kv2.1 with regulatory Kv5.1 or Kv6.1 alpha-subunits. Am.
871 J. Physiol. *274*, C1501-1510.

872 Lake, B.B., Codeluppi, S., Yung, Y.C., Gao, D., Chun, J., Kharchenko, P.V., Linnarsson, S., and
873 Zhang, K. (2017). A comparative strategy for single-nucleus and single-cell transcriptomes
874 confirms accuracy in predicted cell-type expression from nuclear RNA. Sci. Rep. *7*, 6031.

875 Li, B., and Dewey, C.N. (2011). RSEM: accurate transcript quantification from RNA-Seq data
876 with or without a reference genome. BMC Bioinformatics *12*, 323.

877 Liu, W.L., and Shipley, M.T. (1994). Intrabulbar associational system in the rat olfactory bulb
878 comprises cholecystokinin-containing tufted cells that synapse onto the dendrites of
879 GABAergic granule cells. J. Comp. Neurol. *346*, 541–558.

880 Love, M.I., Huber, W., and Anders, S. (2014). Moderated estimation of fold change and
881 dispersion for RNA-seq data with DESeq2. Genome Biol. *15*, 550.

882 Lukas, M., Suyama, H., and Egger, V. (2019). Vasopressin Cells in the Rodent Olfactory Bulb
883 Resemble Non-Bursting Superficial Tufted Cells and Are Primarily Inhibited upon Olfactory
884 Nerve Stimulation. ENeuro *6*.

885 Macosko, E.Z., Basu, A., Satija, R., Nemesh, J., Shekhar, K., Goldman, M., Tirosh, I., Bialas,
886 A.R., Kamitaki, N., Martersteck, E.M., et al. (2015). Highly Parallel Genome-wide Expression
887 Profiling of Individual Cells Using Nanoliter Droplets. Cell *161*, 1202–1214.

888 Margrie, T.W., and Schaefer, A.T. (2003). Theta oscillation coupled spike latencies yield
889 computational vigour in a mammalian sensory system. J. Physiol. *546*, 363–374.

890 Markopoulos, F., Rokni, D., Gire, D.H., and Murthy, V.N. (2012). Functional properties of
891 cortical feedback projections to the olfactory bulb. Neuron *76*, 1175–1188.

892 Miyamichi, K., Amat, F., Moussavi, F., Wang, C., Wickersham, I., Wall, N.R., Taniguchi, H.,
893 Tasic, B., Huang, Z.J., He, Z., et al. (2011). Cortical representations of olfactory input by
894 trans-synaptic tracing. Nature *472*, 191–196.

895 Mori, K., and Sakano, H. (2011). How is the olfactory map formed and interpreted in the
896 mammalian brain? Annu. Rev. Neurosci. *34*, 467–499.

897 Mori, K., Kishi, K., and Ojima, H. (1983). Distribution of dendrites of mitral, displaced mitral,
898 tufted, and granule cells in the rabbit olfactory bulb. J. Comp. Neurol. *219*, 339–355.

899 Mouradian, L.E., and Scott, J.W. (1988). Cytochrome oxidase staining marks dendritic zones
900 of the rat olfactory bulb external plexiform layer. J. Comp. Neurol. *271*, 507–518.

901 Nagai, Y., Sano, H., and Yokoi, M. (2005). Transgenic expression of Cre recombinase in
902 mitral/tufted cells of the olfactory bulb. Genes. N. Y. N 2000 *43*, 12–16.

903 Nagayama, S., Enerva, A., Fletcher, M.L., Masurkar, A.V., Igarashi, K.M., Mori, K., and Chen,
904 W.R. (2010). Differential axonal projection of mitral and tufted cells in the mouse main
905 olfactory system. Front. Neural Circuits *4*.

906 Nagayama, S., Homma, R., and Imamura, F. (2014). Neuronal organization of olfactory bulb
907 circuits. Front. Neural Circuits *8*, 98.

908 Najac, M., De Saint Jan, D., Reguero, L., Grandes, P., and Charpak, S. (2011). Monosynaptic
909 and polysynaptic feed-forward inputs to mitral cells from olfactory sensory neurons. J.
910 Neurosci. Off. J. Soc. Neurosci. *31*, 8722–8729.

911 Niedworok, C.J., Schwarz, I., Ledderose, J., Giese, G., Conzelmann, K.-K., and Schwarz, M.K.
912 (2012). Charting monosynaptic connectivity maps by two-color light-sheet fluorescence
913 microscopy. Cell Rep. *2*, 1375–1386.

914 Niimura, Y. (2012). Olfactory receptor multigene family in vertebrates: from the viewpoint
915 of evolutionary genomics. Curr. Genomics *13*, 103–114.

916 Orona, E., Rainer, E.C., and Scott, J.W. (1984). Dendritic and axonal organization of mitral
917 and tufted cells in the rat olfactory bulb. J. Comp. Neurol. *226*, 346–356.

918 Otazu, G.H., Chae, H., Davis, M.B., and Albeanu, D.F. (2015). Cortical Feedback Decorrelates
919 Olfactory Bulb Output in Awake Mice. Neuron *86*, 1461–1477.

920 Padmanabhan, K., and Urban, N.N. (2010). Intrinsic biophysical diversity decorrelates
921 neuronal firing while increasing information content. Nat. Neurosci. *13*, 1276–1282.

922 Panzanelli, P., Perazzini, A.-Z., Fritschy, J.-M., and Sassoè-Pognetto, M. (2005).
923 Heterogeneity of gamma-aminobutyric acid type A receptors in mitral and tufted cells of the
924 rat main olfactory bulb. J. Comp. Neurol. *484*, 121–131.

925 Parrish-Aungst, S., Shipley, M.T., Erdelyi, F., Szabo, G., and Puche, A.C. (2007). Quantitative
926 analysis of neuronal diversity in the mouse olfactory bulb. J. Comp. Neurol. *501*, 825–836.

927 Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M.,
928 Prettenhofer, P., Weiss, R., Dubourg, V., et al. (2011). Scikit-learn: Machine Learning in
929 Python. J. Mach. Learn. Res. *12*, 2825–2830.

930 Phillips, M.E., Sachdev, R.N.S., Willhite, D.C., and Shepherd, G.M. (2012). Respiration drives
931 network activity and modulates synaptic and circuit processing of lateral inhibition in the
932 olfactory bulb. J. Neurosci. Off. J. Soc. Neurosci. *32*, 85–98.

933 Richard, M.B., Taylor, S.R., and Greer, C.A. (2010). Age-induced disruption of selective
934 olfactory bulb synaptic circuits. Proc. Natl. Acad. Sci. U. S. A. *107*, 15613–15618.

935 Saha, B., Hari, P., Huilgol, D., and Tole, S. (2007). Dual role for LIM-homeodomain gene Lhx2
936 in the formation of the lateral olfactory tract. J. Neurosci. Off. J. Soc. Neurosci. *27*, 2290–
937 2297.

938 Schaefer, A.T., Angelo, K., Spors, H., and Margrie, T.W. (2006). Neuronal oscillations enhance
939 stimulus discrimination by ensuring action potential precision. PLoS Biol. *4*, e163.

940 Schwarz, D., Kollo, M., Bosch, C., Feinauer, C., Whiteley, I., Margrie, T.W., Cutforth, T., and
941 Schaefer, A.T. (2018). Architecture of a mammalian glomerular domain revealed by novel
942 volume electroporation using nanoengineered microelectrodes. Nat. Commun. *9*, 183.

943 Scott, J.W., McBride, R.L., and Schneider, S.P. (1980). The organization of projections from
944 the olfactory bulb to the piriform cortex and olfactory tubercle in the rat. J. Comp. Neurol.
945 *194*, 519–534.

946 Seroogy, K.B., Brecha, N., and Gall, C. (1985). Distribution of cholecystokinin-like
947 immunoreactivity in the rat main olfactory bulb. J. Comp. Neurol. *239*, 373–383.

948 Shannon, P., Markiel, A., Ozier, O., Baliga, N.S., Wang, J.T., Ramage, D., Amin, N.,
949 Schwikowski, B., and Ideker, T. (2003). Cytoscape: a software environment for integrated
950 models of biomolecular interaction networks. Genome Res. *13*, 2498–2504.

951 Short, S.M., and Wachowiak, M. (2019). Temporal Dynamics of Inhalation-Linked Activity
952 across Defined Subpopulations of Mouse Olfactory Bulb Neurons Imaged In Vivo. ENeuro *6*.

953 Sosulski, D.L., Bloom, M.L., Cutforth, T., Axel, R., and Datta, S.R. (2011). Distinct
954 representations of olfactory information in different cortical centres. Nature *472*, 213–216.

955 Subramanian, A., Tamayo, P., Mootha, V.K., Mukherjee, S., Ebert, B.L., Gillette, M.A.,
956 Paulovich, A., Pomeroy, S.L., Golub, T.R., Lander, E.S., et al. (2005). Gene set enrichment
957 analysis: a knowledge-based approach for interpreting genome-wide expression profiles.
958 Proc. Natl. Acad. Sci. U. S. A. *102*, 15545–15550.

959 Sun, X., Liu, X., Starr, E.R., and Liu, S. (2020). CCKergic Tufted Cells Differentially Drive Two
960 Anatomically Segregated Inhibitory Circuits in the Mouse Olfactory Bulb. J. Neurosci. Off. J.
961 Soc. Neurosci. *40*, 6189–6206.

962 Suo, S., Zhu, Q., Saadatpour, A., Fei, L., Guo, G., and Yuan, G.-C. (2018). Revealing the Critical
963 Regulators of Cell Identity in the Mouse Cell Atlas. Cell Rep. *25*, 1436-1445.e3.

964 Tavakoli, A., Schmaltz, A., Schwarz, D., Margrie, T.W., Schaefer, A.T., and Kollo, M. (2018).
965 Quantitative Association of Anatomical and Functional Classes of Olfactory Bulb Neurons. J.
966 Neurosci. Off. J. Soc. Neurosci. *38*, 7204–7220.

967 Tepe, B., Hill, M.C., Pekarek, B.T., Hunt, P.J., Martin, T.J., Martin, J.F., and Arenkiel, B.R.
968 (2018). Single-Cell RNA-Seq of Mouse Olfactory Bulb Reveals Cellular Heterogeneity and
969 Activity-Dependent Molecular Census of Adult-Born Neurons. Cell Rep. *25*, 2689-2703.e3.

970 Tervo, D.G.R., Hwang, B.-Y., Viswanathan, S., Gaj, T., Lavzin, M., Ritola, K.D., Lindo, S.,
971 Michael, S., Kuleshova, E., Ojala, D., et al. (2016). A Designer AAV Variant Permits Efficient
972 Retrograde Access to Projection Neurons. Neuron *92*, 372–382.

973 Thieffry, D. (2007). Dynamical roles of biological regulatory circuits. Brief. Bioinform. *8*, 220–
974 225.

975 Traag, V.A., Waltman, L., and van Eck, N.J. (2019). From Louvain to Leiden: guaranteeing
976 well-connected communities. Sci. Rep. *9*, 5233.

977 Tripathy, S.J., Padmanabhan, K., Gerkin, R.C., and Urban, N.N. (2013). Intermediate intrinsic
978 diversity enhances neural population coding. Proc. Natl. Acad. Sci. U. S. A. *110*, 8248–8253.

979 Wolf, F.A., Angerer, P., and Theis, F.J. (2018). SCANPY: large-scale single-cell gene expression
980 data analysis. Genome Biol. *19*, 15.

981 Wolpert, L. (1969). Positional information and the spatial pattern of cellular differentiation.
982 J. Theor. Biol. *25*, 1–47.

983 Zeisel, A., Hochgerner, H., Lönnerberg, P., Johnsson, A., Memic, F., van der Zwan, J., Häring,
984 M., Braun, E., Borm, L.E., La Manno, G., et al. (2018). Molecular Architecture of the Mouse
985 Nervous System. Cell *174*, 999-1014.e22.

986 Zhang, X., and Firestein, S. (2002). The olfactory receptor gene superfamily of the mouse.
987 Nat. Neurosci. *5*, 124–133.

988 (1894). The Croonian lecture.—La fine structure des centres nerveux. Proc. R. Soc. Lond. *55*,
989 444–468.

990

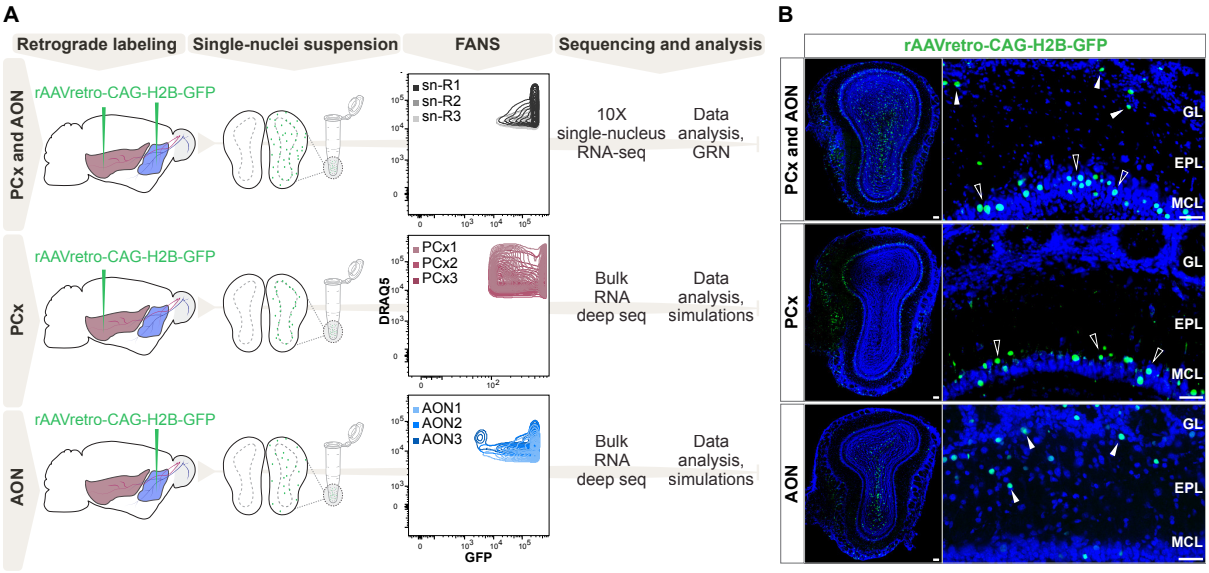


Figure 1: Comprehensive molecular profiling of olfactory bulb projection neurons.
(A) Schematic representation of experimental design. Top: after injection of rAAVretro-CAG-H2B-GFP into PCx and AON, single nuclei were dissociated from 3 mice (single nuclei (sn) R1,2,3: replicates 1,2,3) and sorted using Fluorescence-activated Nuclei Sorting (FANS). The population of nuclei is selected based on GFP and DRAQ5 (far-red fluorescent DNA dye). See **Figure 1-figure supplement 2** for detailed FANS plots. Sorted nuclei were sequenced using 10X single-nucleus RNA-seq. Middle and bottom: after injection of rAAVretro-CAG-H2B-GFP into PCx (middle) or AON (bottom), single nuclei were dissociated from 3 mice for each injection site and sorted using FANS (as described above and **Figure 1-figure supplement 2**). RNA extracted from sorted nuclei was sequenced using bulk RNA deep sequencing. PCx: Piriform Cortex; AON: Anterior Olfactory Nucleus; R: replicate; GRN: Gene Regulatory Network.
(B) Representative coronal sections and high magnification images showing GFP expression (in green) in the olfactory bulb after injection of rAAVretro-CAG-H2B-GFP into PCx and AON (top), PCx only (middle), and AON only (bottom). Injection of the virus into PCx and AON resulted in GFP-expressing nuclei located in the mitral cell (empty arrowheads), external plexiform, glomerular (white arrowheads), and granule cell layers; injection into PCx resulted in GFP-expressing nuclei located in the mitral cell layer (empty arrowheads); injection into AON resulted in GFP-expressing nuclei located in the external plexiform and glomerular layers (white arrowheads) and granule cell layers. GL: glomerular layer; EPL: external plexiform layer; MCL: mitral cell layer; GCL: granule cell layer. Neurotrace counterstain in blue. Scale bars, 100µm and 50µm (high magnification).

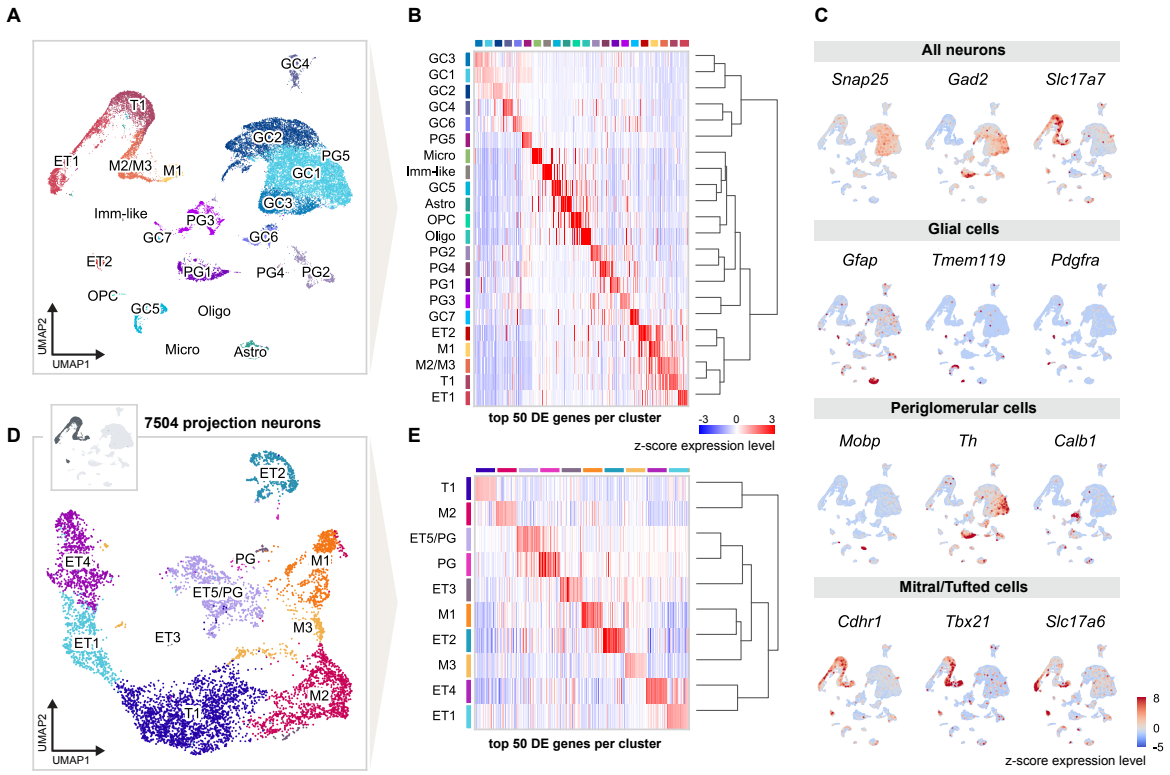


Figure 2: Single nucleus RNA sequencing distinguishes distinct cell types and molecular signatures of OB projection neurons.

(A) UMAP representation of gene expression profiles of 31,703 single nuclei combined from all replicates (1, 2, 3) of mice injected into both AON and PCx, grouped into 22 clusters color-coded by cell type membership (GC: granule cell, PG: periglomerular cell, OPC: oligodendrocyte precursor cell, Micro: microglia, Astro: astrocyte, Oligo: oligodendrocyte, ET: external tufted cell, M: mitral cell, T: tufted cell, Imm-like: Immature-like cell). See **Figure 2-figure supplement 1** for detailed quality check of each replicate.

(B) Matrixplot showing the z-score expression level of the top 50 differentially expressed (DE) genes for each cell population ordered by hierarchical relationships between distinct clusters. Each column represents the average expression level of a gene in a given cluster, color-coded by the UMAP cluster membership **(**from **A)**. The dendrogram depicts the hierarchical relationships and is computed from the PCA representation of the data using Pearson correlation as distance measure and link by complete linkage.

(C) UMAP representations of known marker genes for main cell populations (*Snap25*: neurons; *Gad2*: GABAergic neurons; *Slc17a7*: glutamatergic neurons; *Gfap, Tmem119, Pdgfra*, *Mobp*: glial cells; *Th*, *Calb1*: periglomerular neurons; *Cdhr1, Tbx21, Slc17a6*: mitral/tufted cells). Nuclei are color-coded by the z-score expression level of each transcript.

(D) UMAP representation of subclustering from initial clusters M1, M2/M3, T1, ET1 and ET2 (cluster IDs from **Figure A**), selected for the expression of known excitatory and mitral/tufted cell markers **(**shown in **C)**, showing 7,504 putative projection neurons grouped into 10 distinct types.

(E) Same matrixplot as described in **B** showing the z-score expression level of the top 50 DE genes for each projection neuron type ordered by hierarchical relationships and color-coded by the UMAP subcluster membership **(**from **D).**

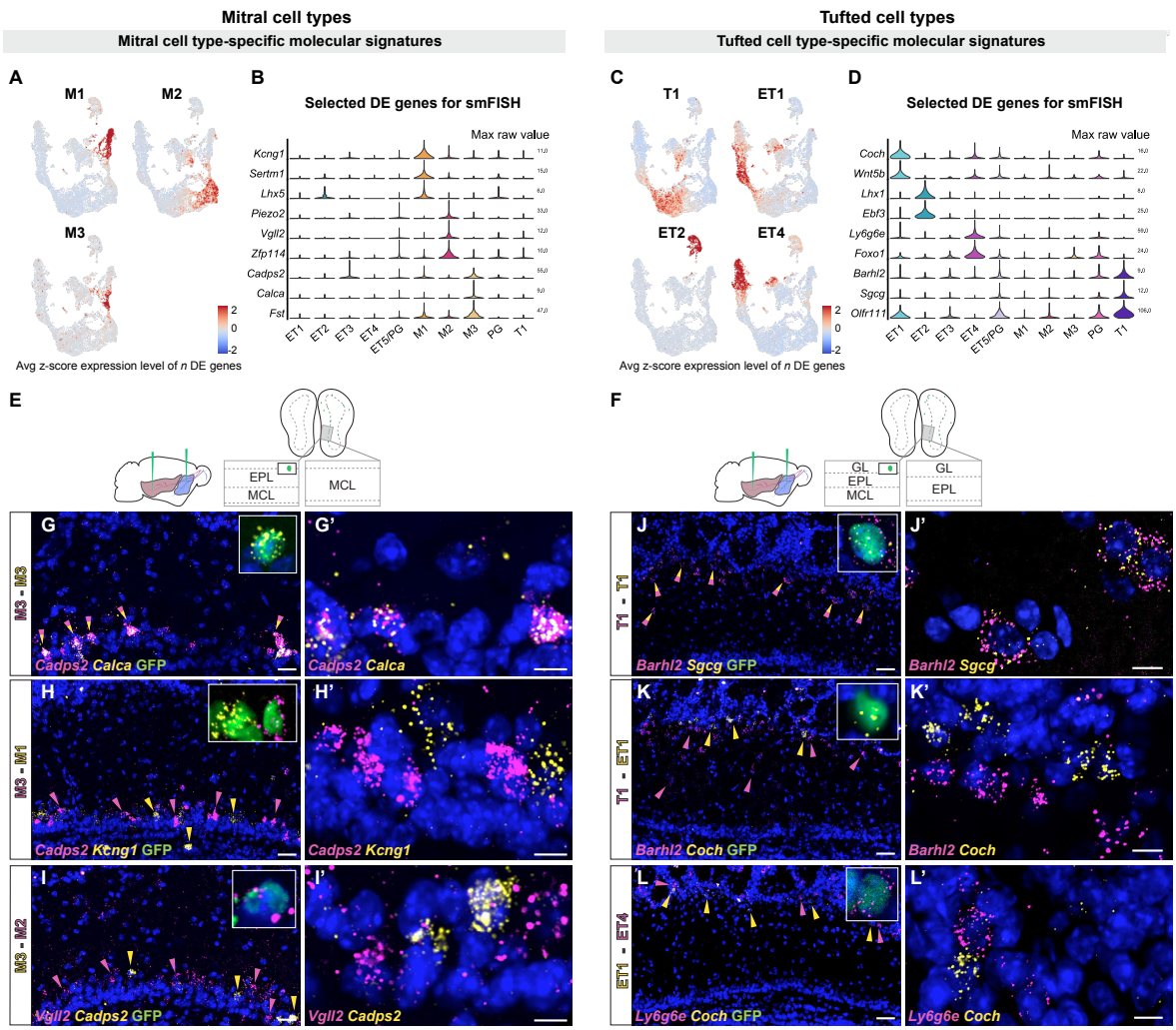


Figure 3: Histological validation of molecularly distinct mitral and tufted cell types.
(A) Combined average (avg) z-score expression level of the top *n* differentially expressed (DE) genes for each mitral cell type (M1 *n*=14, M2 *n*=11, M3 *n*=10), overlaid on the projection neuron UMAP space (**Figure 2D**). DE genes were selected if their log fold change was greater than 4 (see Methods for details). M1-specific genes: *Kcng1, Lhx1, Sertm1, Gabra2, Doc2b, Cntn6, Olfr1259, Nrp2, C1ql1, Ebf1, Baiap3, Adgrl2, Dsc2, Chrna5;* M2-specific genes*: Piezo2, Vgll2, Zfp114, Nts, Ros1, Samsn1, Grid2, Smpx, Itga4, Itga9, Sema6d;* M3-specific genes*: Cadps2, Calca, Fst, Ets1, Ednra, Cdkn1c, Mustn1, Smoc2, Cnr1, Ccno.*
(B) Violin plots showing maximum raw expression value of selected mitral cell type-specific DE genes across mitral and tufted cell clusters for further validation with smFISH.
(C) Combined average (avg) z-score expression level of the top *n* DE genes for each tufted cell type (T1 *n*=9, ET1 *n*=7, ET2 *n*=9, ET4 *n*=6), overlaid on the projection neuron UMAP space (**Figure 2D**). T1-specific genes: *Barhl2, Sgcg, Vdr, Olfr111, Olfr110, Cacna1g, Fam84b, Kcna10, Tspan10;* ET1-specific genes*: Coch, Wnt5b, Rorb, Chst9, Tpbgl, Clcf1, Rxfp1;* ET2-specific genes*: Lhx1, Ebf3, Trp73, Edn1, Ebf2, Nr2f2, Uncx, Psrc1, Dsp;* ET4-specific genes*: Ly6g6e, Foxo1, Siah3, Galnt12, Itga8, Ets2, Grik4.*
(D) Violin plots showing maximum raw expression value of selected tufted cell type-specific DE genes across mitral and tufted cell clusters for further validation with smFISH.
(E, F) Schematic representations of the smFISH images for validating projection neuron type-

specific selected marker genes upon rAAVretro-CAG-H1B-GFP injection into PCx and AON. The schemes depict the laminar location visualized in the histological images from a coronal section of the ipsilateral hemisphere to the injection site. EPL: external plexiform layer; MCL: mitral cell layer; GL: glomerular layer.

(G - I) smFISH showing combinatorial expression of mitral cell type-specific marker genes for M1, M2 and M3 cells in the mitral cell layer. High magnifications (top right) show co-labeling of viral GFP with the *in situ* mRNA probe. **(G and G')** The M3 markers *Cadps2* and *Calca* are co-expressed in subpopulations of cells in the mitral cell layer, indicated by the yellow/magenta arrowheads. **(H and H')** The M3 marker *Cadps2* and M1 marker *Kcng1* are mutually exclusive in subpopulations of cells in the mitral cell layer, indicated by the magenta and yellow arrowheads respectively. **(I and I')** The M3 marker *Cadps2* and M2 marker *Vgll2* are mutually exclusive in subpopulations of cells in the mitral cell layer, indicated by the yellow and magenta arrowheads respectively. For additional histological analysis see **Figure 3-figure supplement Figure 1.**

(J - L) smFISH images showing combinatorial expression patterns of tufted cell type-specific marker genes for validating T1, ET1, ET2 and ET4 clusters as distinct projection neuron types in the external plexiform and glomerular layers. High magnifications (top right) show co-labeling of viral GFP with the *in situ* mRNA probe. As described for the mitral cell types, yellow or magenta arrowheads show mutually exclusive patterns (**K, K':** T1-ET1 and **L, L':** ET1-ET4), and yellow/magenta arrowheads show co-expression patterns (**J, J':** T1-T1). For additional histological analysis see **Figure 3-figure supplement 1.** DAPI counterstain in blue. Scale bars, 50μm and 10μm (high magnifications).

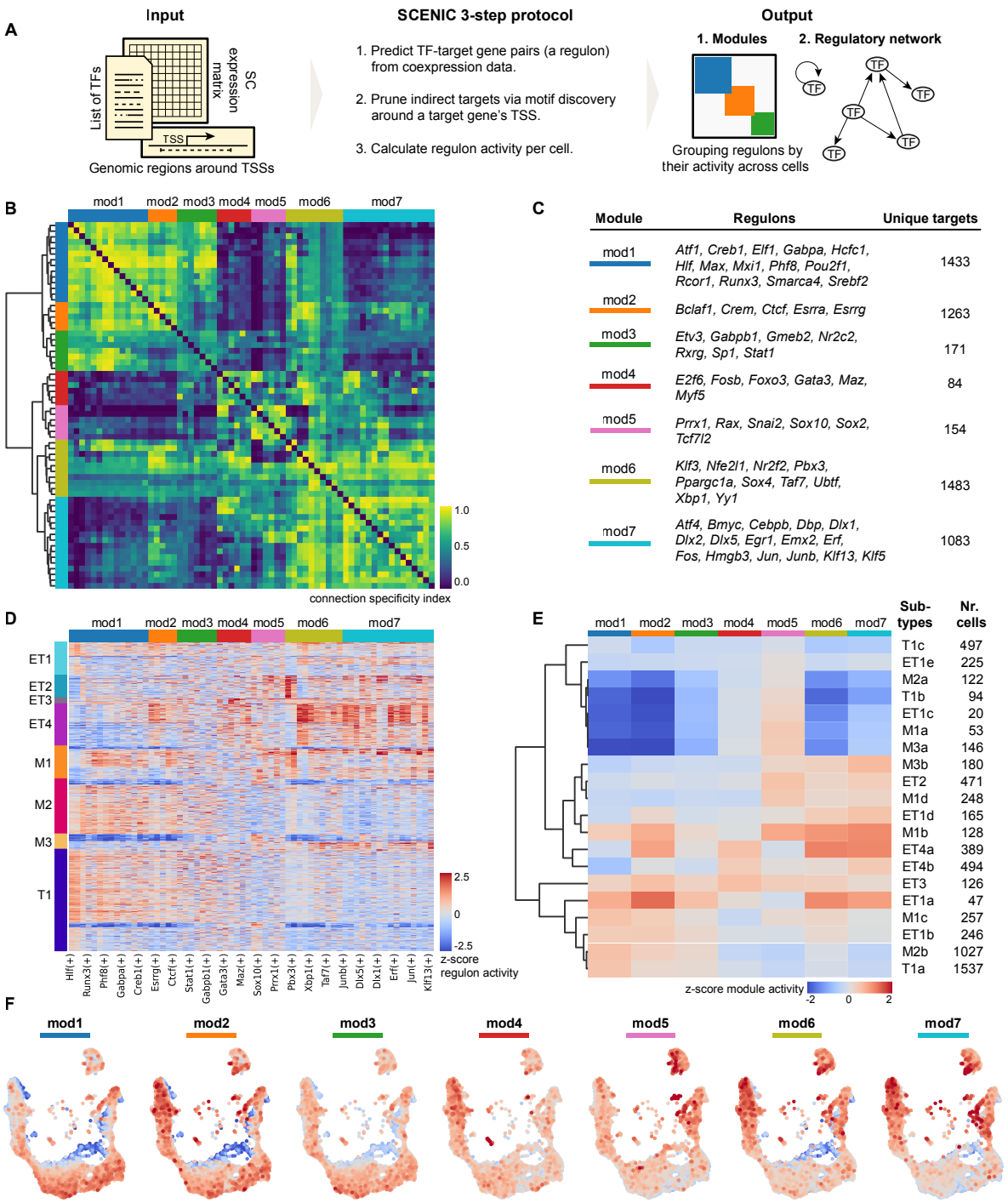


Figure 4: Mitral and tufted cell-specific regulons combine into modules.
(A) Schematic representation of the network analysis pipeline, including the required input, the SCENIC protocol, and the output in the form of regulon modules and a regulatory network **(Figure 5)**.
(B) Hierarchical clustering of regulons using the Connection Specificity Index (CSI) as a distance measure results in 7 modules (Ward linkage). The CSI of two regulons is based on Pearson correlation coefficients (PCC): for each PCC between regulon A and B, the CSI is the fraction of regulons that have a PCC with A and with B lower than PCC (AB). Prominent cross-module interactions are observed for mod1-2-3 and mod6-7. Moreover, module 6 showed interactions with all other modules.

1095 **(C)** Table listing the modules, their regulons, and the number of unique target genes in each
1096 module.
1097 **(D)** Projection neuron cell types defined by transcriptome analysis **(Figure 2)** and subtypes
1098 (rows) defined by regulon activity (columns). Rows were ordered by cell type and within a cell
1099 type by hierarchical clustering (Euclidean distance, Ward linkage). Columns clustered as in
1100 panel **B**.
1101 **(E)** Module activity per cell subtype. Module activity is calculated as the average activity of its
1102 regulons for a given cell subtype. Each cell subtype may be defined by a combination of active
1103 and inactive modules. For example, T1a (bottom row) is defined by high activity in modules 1,
1104 2, and 3.
1105 **(F)** Module activity mapped on the transcriptome UMAP space. Color range as in panel **E**.
1106

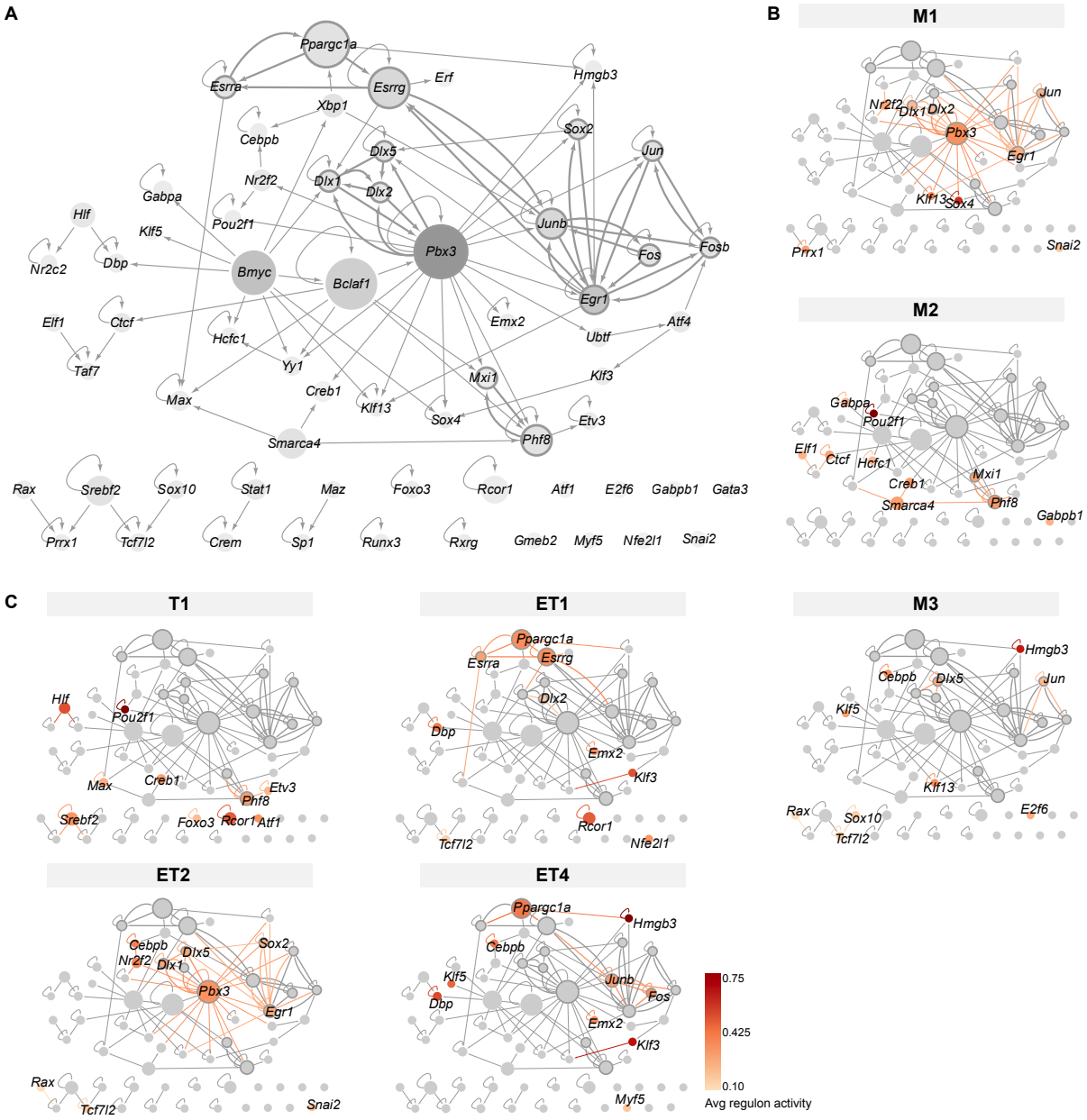


Figure 5: Transcription factor network derived from regulons.
(A) Overview of mitral and tufted cell transcription factor (TF) networks, with node size scaled by the number of target genes and nodes colored with different shades of gray based on outdegree (number of outgoing edges). Thick borders and edges denote cycles of 2 or 3 regulons. The three main hubs are: *Pbx3* (outdegree 19, target genes 688), *Bmyc* (outdegree 10, target genes 578) and *Bclaf* (outdegree 7, target genes 724).
(B) Mitral cell types (M1, M2, M3) with standardized regulon activity for the top 10 most specific regulons mapped onto the corresponding TF nodes (compare **Figure 4D, E**).
(C) Tufted cell types (T1, ET1, ET2, ET4) with standardized regulon activity for the top 10 most specific regulons. We omitted ET3 as it only had a few cells.

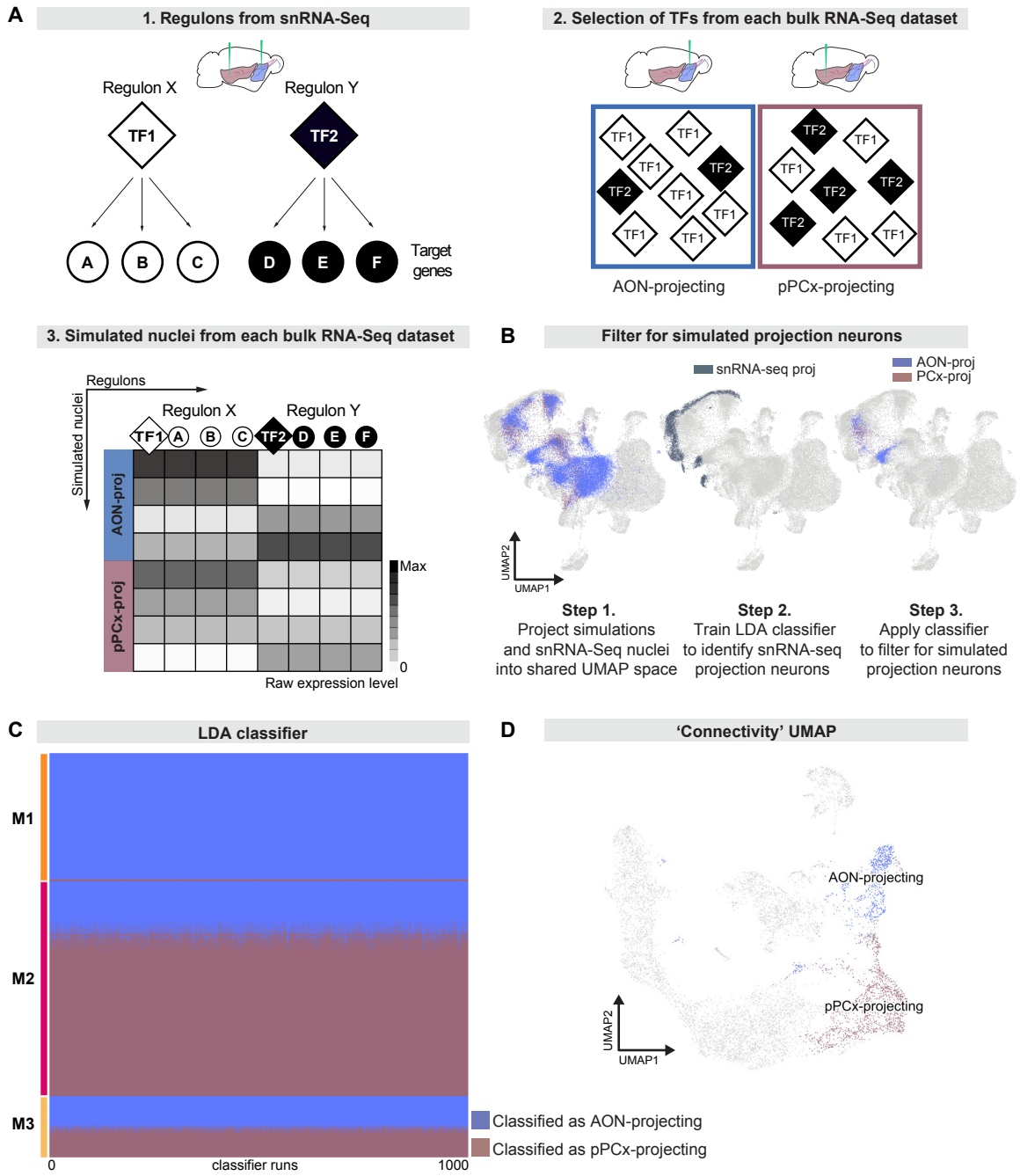
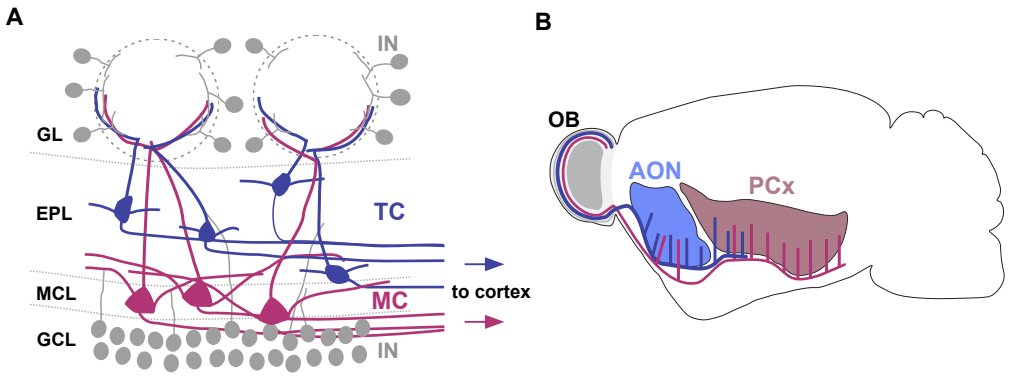


Figure 6: Simulations from bulk RNA deep sequencing data suggest mitral types have distinct projection targets.

(A) Schematic representation of strategy to integrate bulk RNA-seq and single nucleus RNA-seq data. **(A1)** Simulations use regulons inferred from snRNA-seq data. A regulon consists of a transcription factor and all target genes that are activated by that transcription factor. **(A2)** When simulating nuclei from bulk RNA-seq data, the probability of selecting a given regulon is determined by the abundance of its transcription factor in the bulk RNA-seq dataset. **(A3)** Nuclei are simulated from each bulk dataset through random sampling of regulons with replacement. This method maintains broad differences between datasets while accounting for heterogeneity within each dataset.

(B) Simulated nuclei and snRNA-seq nuclei projected into a shared UMAP representation.

1131 Step 1: Blue indicates all simulated nuclei from AON-projecting bulk RNA-seq and purple
1132 indicates all simulated nuclei from pPCx-projecting bulk RNA-seq. Step 2: Darker color
1133 indicates snRNA-seq projection neurons. Step 3: Blue indicates AON-projecting simulated
1134 projection neurons and purple indicates pPCx-projecting simulated projection neurons.
1135 **(C)** Linear Discriminant Analysis (LDA) classifiers were trained on simulated projection
1136 neurons, then used to predict the projection target of snRNA-seq mitral cells to investigate
1137 projection targets of snRNA-seq derived types. Each row represents one mitral cell. Each
1138 column represents one of 1000 LDA classifiers. Blue indicates that the mitral cell was
1139 classified as AON-projecting, and purple indicates that the mitral cell was classified as pPCx-
1140 projecting. Within each mitral cell type, cells are sorted vertically by their predicted projection
1141 target.
1142 **(D)** UMAP representation color-coded by predicted projection target. Cells in blue were
1143 predicted to be AON-projecting by all 1000 classifiers. Cells in purple were predicted to be
1144 pPCx-projecting by all 1000 classifiers.
1145



1146
1147 **Figure 1-figure supplement 1: Schematic representation of olfactory bulb cell types and**
1148 **their cortical projection targets**
1149 **(A)** Schematic representation of cell types and their distribution within the olfactory bulb (IN:
1150 interneuron, TC: tufted cell, MC: mitral cell, GL: glomerular layer, EPL: external plexiform layer,
1151 MCL: mitral cell layer, GCL: granule cell layer). Tufted and mitral cells project their axons to
1152 downstream cortical regions.
1153 **(B)** Schematic representation of the distinct axonal projection targets for mitral and tufted cells
1154 in the olfactory cortex (AON: anterior olfactory nucleus, PCx: piriform cortex). Tufted cells
1155 project primarily to AON and to the anterior part of PCx, whereas mitral cell axons
1156 predominantly target the posterior portion of PCx.
1157

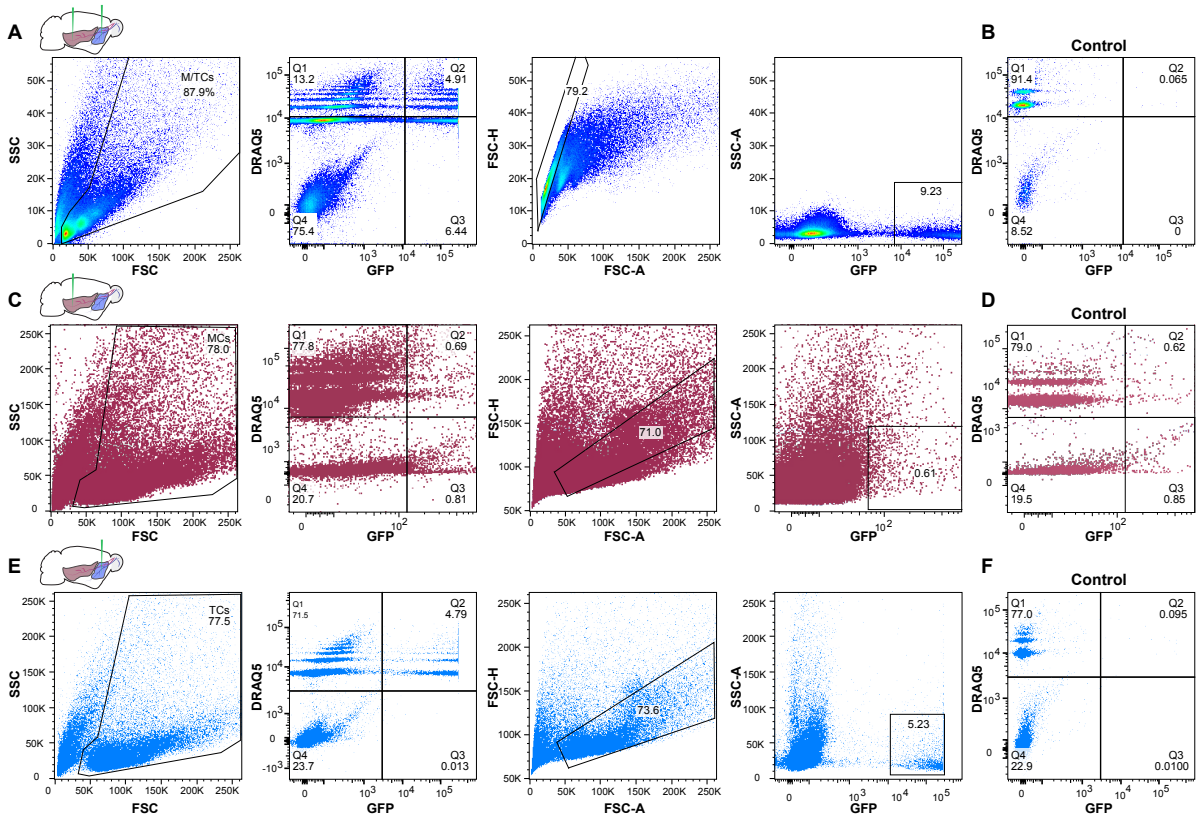


Figure 1-figure supplement 2: Enrichment of GFP-expressing nuclei using Fluorescence-Activated Nuclei Sorting (FANS).

(A) Representative FANS data of GFP-expressing nuclei after injection of rAAVretro-CAG-H2B-GFP into AON and PCx to label OB projection neurons. From left to right: gating strategy for enrichment of GFP-expressing nuclei. Nuclei are identified based on size and granularity in the FSC (forward scatter) versus SSC (side scatter) plot. Next, the population of nuclei restricted to Quadrant 2 is selected that is both GFP-expressing and DRAQ5-positive (far-red fluorescent DNA dye). To exclude doublets, the population of nuclei around the diagonal in the FSC-A (forward scatter area) versus FSC-H (forward scatter height) plot is selected. Lastly, GFP-expressing nuclei filtered through the preceding steps are plotted against the SSC-A (side scatter area).

(B) FANS results for control specimen for AON and PCx (olfactory bulb contralateral to the injection hemisphere from the same animal) showing the absence of GFP-expressing nuclei (same gating strategy as shown in **(A)**).

(C) Representative FANS data after injection of rAAVretro-CAG-H2B-GFP into PCx to enrich for mitral cell nuclei.

(**D)** FANS results for control specimen for PCx injections showing the absence of GFP-expressing nuclei (same gating strategy as shown in **(C)**).

(E) Same as in **(C)** but after injection of rAAVretro-CAG-H2B-GFP into AON to enrich for tufted cell nuclei.

(F) FANS results for control specimen for AON (olfactory bulb contralateral to the injection hemisphere from the same animal) showing the absence of GFP-expressing nuclei (same gating strategy as shown in **(E)**).

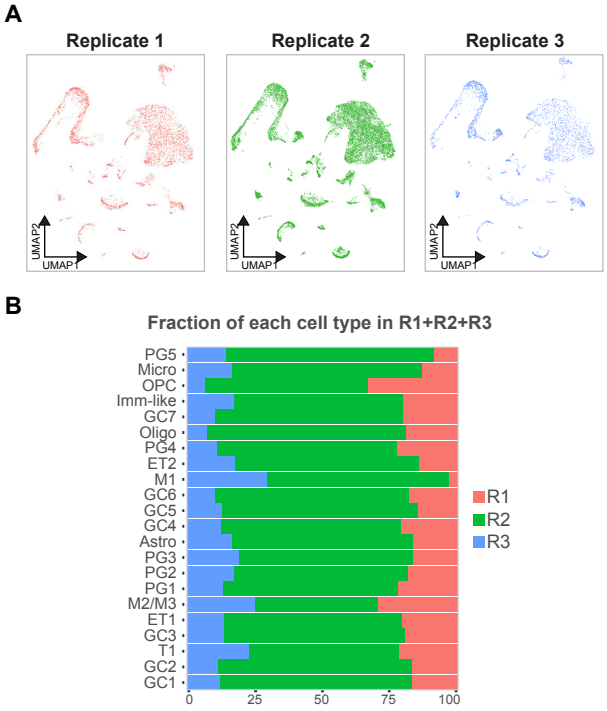


Figure 2-figure supplement 1: Quality check of individual replicates of sn-RNA-seq shows the reliability of the data and the replicability of each cell type.
(A) Depiction of nuclei for each replicate (in red R1, in green R2, in blue R3) embedded in the UMAP space showing that replicates are very similar to each other and can be combined for downstream analyses.
(B) Barchart showing the fraction of each cluster when combined R1, 2 and 3 color-coded by the replicate membership **(A),** indicating that each cell type is consistently represented in each replicate.

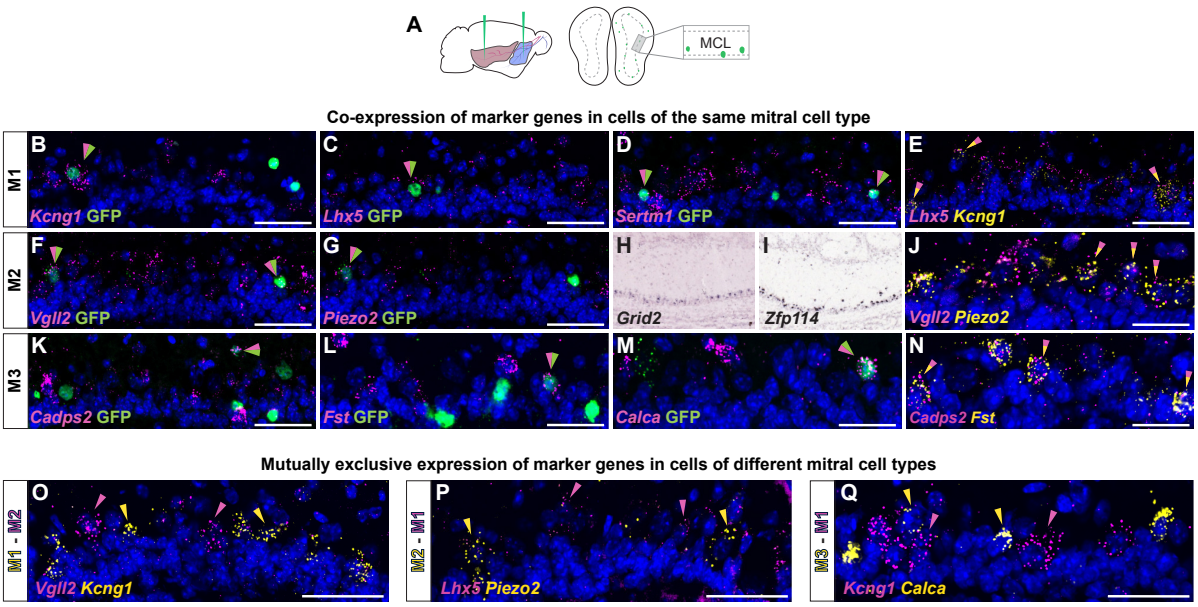


Figure 3-figure supplement 1: Histological analysis of DE genes for distinct mitral cell types.

(A) Schematic representation of the smFISH images for validating selected mitral cell type-specific marker genes upon rAAVretro-CAG-H2B-GFP injection into PCx and AON. The scheme depicts the laminar location visualized in the histological images from a coronal section of the ipsilateral hemisphere to the injection site. MCL: mitral cell layer.

(B - D) smFISH showing co-expression of M1-specific marker genes with viral GFP. Labeled nuclei are indicated by the magenta/green arrowheads.

(E) smFISH showing co-expression of two M1-specific marker genes. Labeled nuclei are indicated by the yellow/magenta arrowheads.

(J) smFISH showing co-expression of two M2-specific marker genes. Labeled nuclei are indicated by the yellow/magenta arrowheads.

(N) smFISH showing co-expression of two M3-specific marker genes. Labeled nuclei are indicated by the yellow/magenta arrowheads.

(F, G) smFISH showing co-expression of M2-specific marker genes with viral GFP. Labeled nuclei are indicated by the magenta/green arrowheads.

(H, I) *In situ* hybridization images from the Allen Brain Atlas showing additional M2-specific DE genes.

(K - M) smFISH showing co-expression of M3-specific marker genes with viral GFP. Labeled nuclei are indicated by the magenta/green arrowheads.

(O-P-Q) smFISH images showing mutually exclusive expression of mitral cell type-specific marker genes for M1, M2, and M3. Yellow and magenta arrowheads show mutually exclusive expression of M1 versus M2 **(O-P),** and M1 versus M3 **(Q)** marker genes.

DAPI counterstain in blue. Scale bars, 50µm.

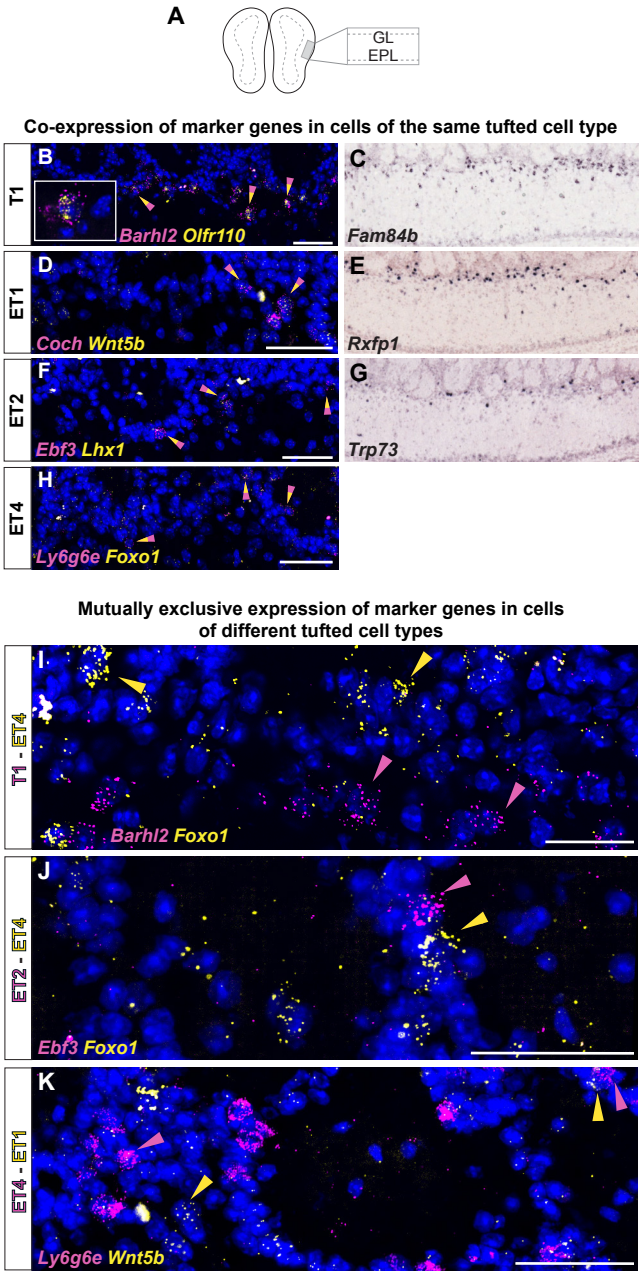


Figure 3-figure supplement 2: Histological analysis of DE genes for distinct tufted cell types.

(A) Schematic of the smFISH images for validating selected tufted cell type-specific marker genes upon rAAVretro-CAG-H2B-GFP injection into PCx and AON. The scheme depicts the laminar location visualized in the histological images from a coronal section of the ipsilateral hemisphere to the injection site. GL=glomerular layer; EPL=external plexiform layer.

(B) smFISH showing co-expression of two T1-specific marker genes. Labeled nuclei are indicated by the yellow/magenta arrowheads. High magnification (left) shows clear co-labeling of the two mRNA probes *Barhl2* and *Olfr110*.

(C) *In situ* hybridization images from the Allen Brain Atlas showing one additional T1-specific DE gene.

(D) smFISH showing co-expression of two ET1-specific marker genes. Labeled nuclei are indicated by the yellow/magenta arrowheads.

(E) *In situ* hybridization images from the Allen Brain Atlas showing additional ET1-specific DE genes.
(F) smFISH showing co-expression of two ET2-specific marker genes. Labeled nuclei are indicated by the yellow/magenta arrowheads.
(G) *In situ* hybridization images from the Allen Brain Atlas showing additional ET2-specific DE genes.
(H) smFISH showing co-expression of two ET4-specific marker genes. Labeled nuclei are indicated by the yellow/magenta arrowheads.
(I-J-K) smFISH images showing mutually exclusive expression of tufted cell type-specific marker genes for T1, ET1, ET2 and ET4. Yellow and magenta arrowheads show mutually exclusive expression of T1 versus ET4 **(I)**, ET2 versusET4 **(J)**, and ET1 versus ET4 **(K)** marker genes.
DAPI counterstain in blue. Scale bars, 50µm.

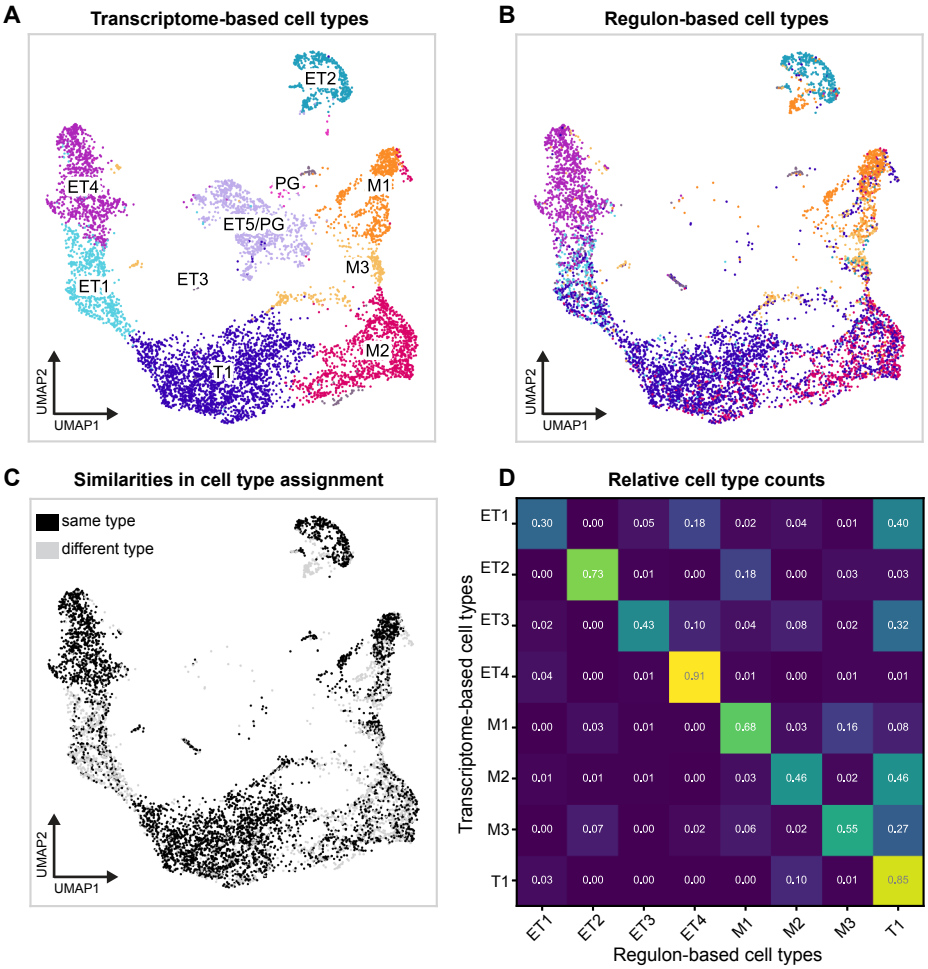


Figure 4-figure supplement 1: Regulon-based clustering and transcriptome-based clustering provide complementary axes for cell type identification.
(A) Leiden clustering of projection neurons in a UMAP space computed from regulon activity, visualized on the UMAP coordinates computed from the whole transcriptome. The major cell type clusters are consistent between the two approaches, even if boundaries are more diffuse and have shifted. Regulon-based clusters were computed through a fine-grained clustering of cells on their regulon activity score. We computed a knn graph (nr of neighbors=10) and then clustered cells with the Leiden algorithm (resolution = 5.5), resulting in 86 clusters. Regulon-based cell types were named for the most abundant transcriptome-based cell type present in the given cluster. This reduced the 86 clusters to 8. Note that the Leiden algorithm is an improved version of the Louvain clustering algorithm (Traag 2018, arxiv).
(B) UMAP of cells assigned the same type (black) or a different type (grey) by both methods.
(C) Overlap between transcriptome-based and regulon-based cell type assignment. Each entry represents the fraction of cells in a given transcriptome-based cell type that were assigned to a given regulon-based cell type. Cell type assignments were mainly consistent, with changes mostly observed in the expansion of T1 under regulon-based clustering to include many cells from ET1, ET3, M2 and M3.

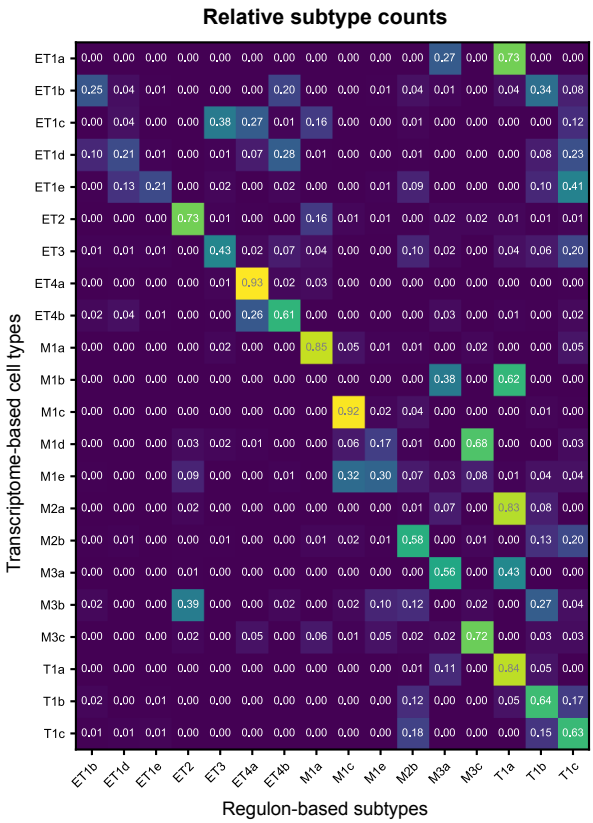


Figure 4-figure supplement 2: Overlap between transcriptome-based and regulon-based cell subtype assignment.
Subtypes were assigned using the method described in **Figure 4-figure supplement 1A.** Each entry represents the fraction of cells in a given transcriptome-based cell subtype that were assigned to a given regulon-based cell subtype. Differences are mainly observed in the expansion of T1a, M3c and ET2 in regulon-based assignment.