Cell Reports, Volume *38*

Supplemental information

A 3D transcriptomics atlas of the mouse nose sheds light on the anatomical logic of smell

Mayra L. Ruiz Tejada Segura, Eman Abou Moussa, Elisa Garabello, Thiago S. Nakahara, Melanie Makhlouf, Lisa S. Mathew, Li Wang, Filippo Valle, Susie S.Y. Huang, Joel D. Mainland, Michele Caselle, Matteo Osella, Stephan Lorenz, Johannes Reisert, Darren W. Logan, Bettina Malnic, Antonio Scialdone, and Luis R. Saraiva

SUPPLEMENTAL FIGURES

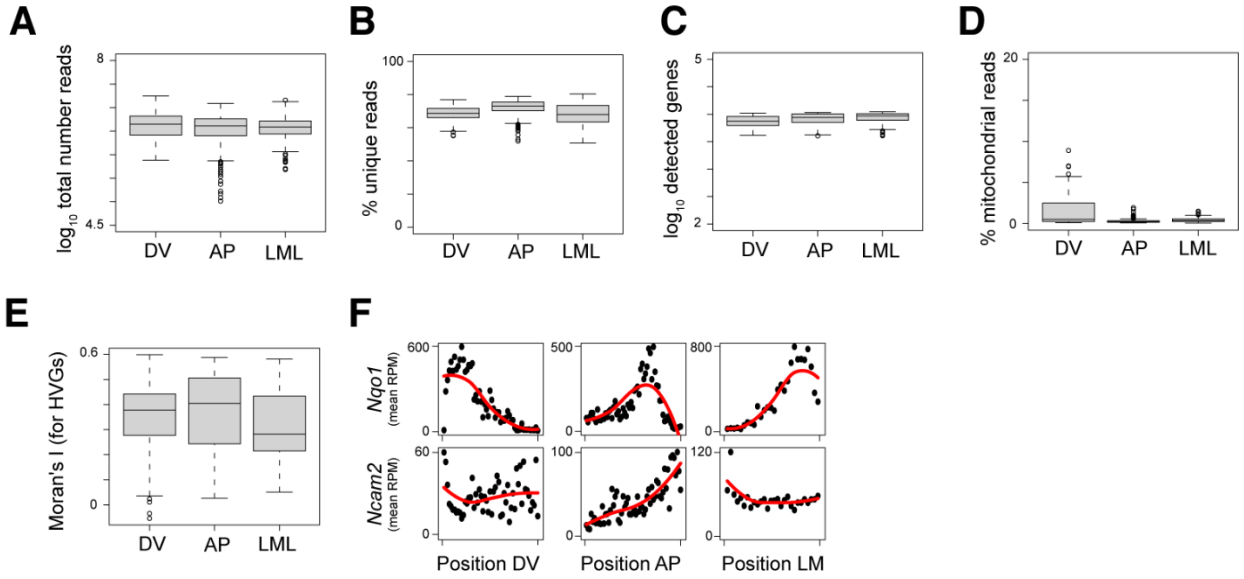

Figure S1. TOMO-seq data QC, Related to Figure 1 and Table S1. (A) Boxplots showing the distributions of the $\log_{10}$ total number of reads per sample in each axis (DV = dorsal-ventral; AP = anterior – posterior; LML = lateral-mid-lateral). (B) Boxplots of percentage of uniquely mapped reads per sample per axis. (C) Boxplots of distributions of $\log_{10}$ detected genes per sample per axis. (D) Boxplots of percentage of mitochondrial reads per sample per axis. (E) Boxplots showing the distribution of the Moran's I statistics calculated for the top 100 Highly Variable Genes per axis. P-values are computed for each gene and then combined with the Simes' method. The combined p-values are $< 2.2 \times 10^{-16}$ for all axes. (F) Normalized expression of canonical OM spatial marker genes along the three axes. Red line showing fits with local polynomial models.

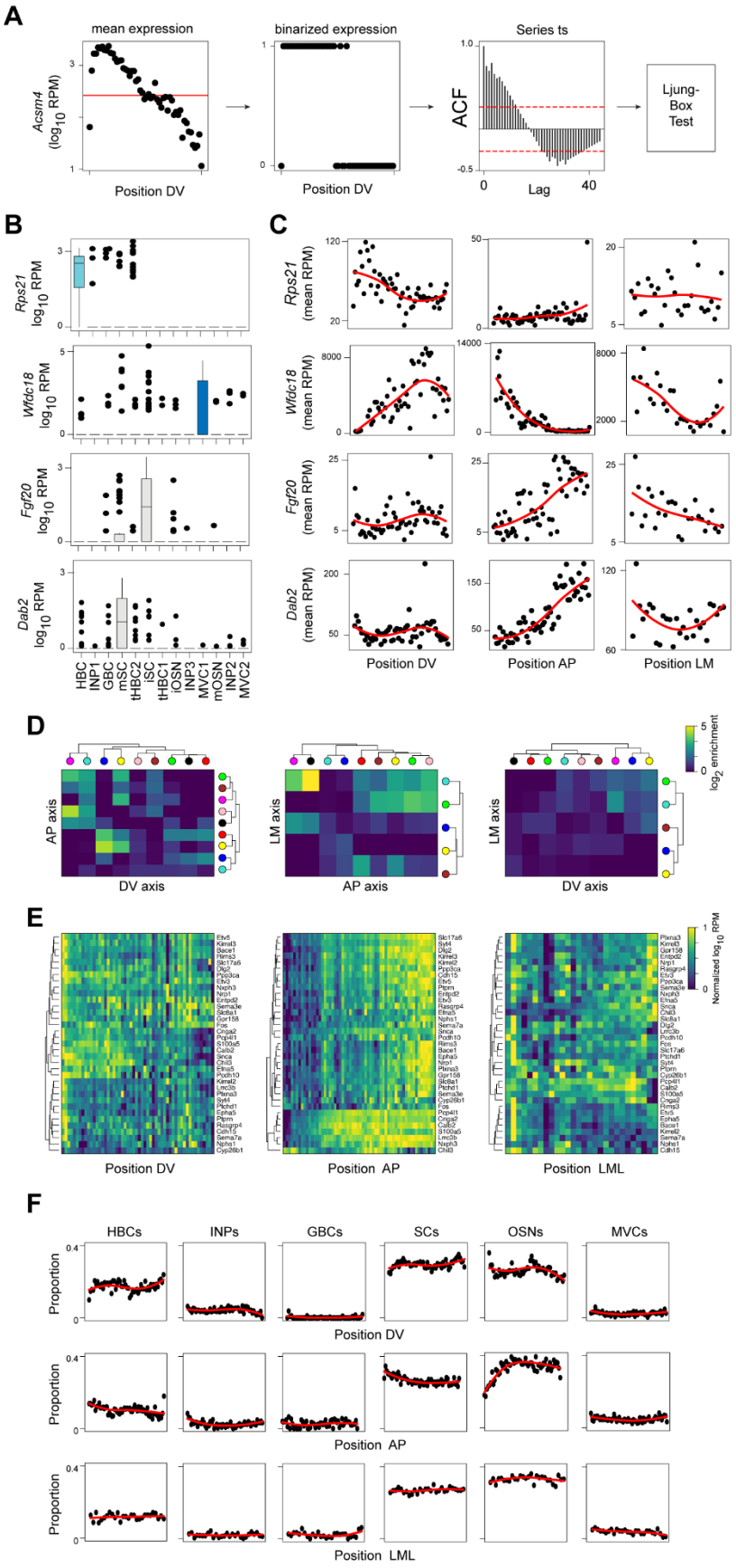
A
mean expression
binarized expression
Series ts
Acsm4 (log10 RPM)
Position DV
Position DV
ACF
Lag
Ljung-Box Test
B
Rps21 log10 RPM
Wfdc18 log10 RPM
Fgf20 log10 RPM
Dab2 log10 RPM
HBC
INP1
GBC
mSC
tHBC2
iSC
tHBC1
iOSN
INP3
MVC1
mOSN
INP2
MVC2
C
Rps21 (mean RPM)
Wfdc18 (mean RPM)
Fgf20 (mean RPM)
Dab2 (mean RPM)
Position DV
Position AP
Position LM
D
AP axis
DV axis
LM axis
AP axis
LM axis
DV axis
log2 enrichment
E
Position DV
Position AP
Position LML
Normalized log10 RPM
F
HBCs
INPs
GBCs
SCs
OSNs
MVCs
Proportion
Position DV
Position AP
Position LML

Figure S2. Spatial differential expression analysis, **Related to Figure 2 and Table S2.** (A) Schematics of strategy to find spatially differentially expressed genes; as an example, data for *Acsm4* along the dorsal-ventral (DV) axis is shown: Gene expression was binarized according to whether the expression per slice was higher or lower than the median expression (red horizontal line). Then, we computed the autocorrelation function for different values of the lags, and we applied the Ljung-Box test to verify whether the autocorrelation values are significantly higher than zero. (B) Box plots of example genes' expression (log_{10} reads-per-million, RPMs) distributions in different cell types. None of these genes is expressed in mOSNs (INP = Immediate Neuronal Precursors; GBC = Globose Basal Cells; mOSNs = mature Olfactory sensory neurons; iOSNs = immature Olfactory Sensory Neurons; MVC = Microvillous Cells; iSC = Immature Sustentacular Cells; mSC = Mature Sustentacular Cells; HBCs = Horizontal Basal Cells). (C) Spatial gene expression trends along each axis of the example genes shown in panel B. (D) Heatmap showing the log_2 enrichment for the intersection between different gene clusters (indicated by colored circles) across pairs of axes, after excluding *Olfr* genes. (E) Heatmaps showing normalized mean expression of the neuronal activity marker genes listed in Table S2 from (Wang et al., 2017) along the three axes. (F) We used cell type deconvolution analysis to estimate the cell type composition per section along the three axes. The red line marks the fit with local polynomial models.

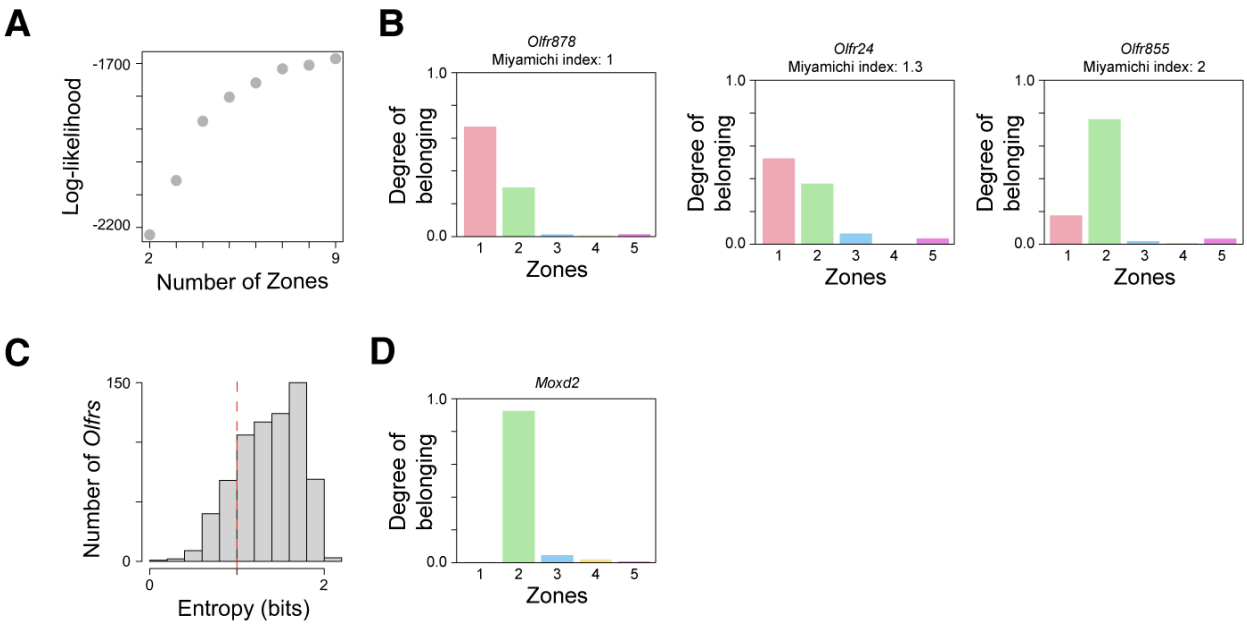

Figure S3. *Olfr* genes 3D zones, Related to Figure 3. (A) Log-likelihood values for fits with LDA models as a function of the number of zones. (B) Bar plot showing the degrees of belonging of *Olfr* genes with overlapping spatial patterns (Miyamichi indexes of 1, 1.3 and 2 respectively). (C) Distribution of entropy values of our 689 spatially differentially expressed *Olfrs*. The *Olfrs* with entropy values less than 1 bit (vertical red line) can be considered to fit mostly in one zone. (D) Bar plot showing the degrees of belonging of *Moxd2*.

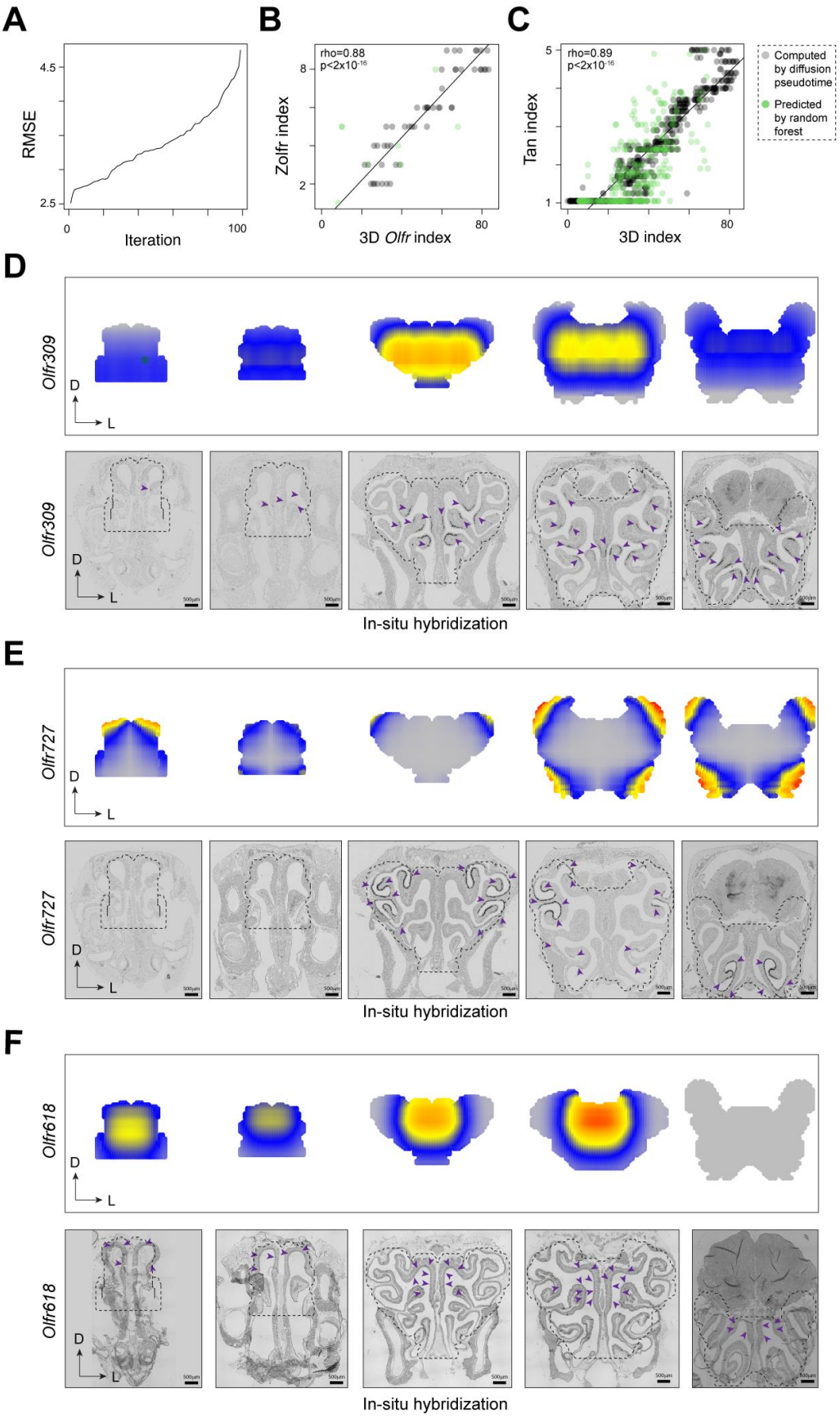
A
RMSE
4.5
2.5
0
100
Iteration
B
rho=0.88
p<2x10-16
Zolfr index
8
2
0
80
3D Olfr index
C
rho=0.89
p<2x10-16
Tan index
5
1
0
80
3D index
Computed by diffusion pseudotime
Predicted by random forest
D
Olfr309
D
L
Olfr309
In-situ hybridization
E
Olfr727
Olfr727
In-situ hybridization
F
Olfr618
Olfr618
In-situ hybridization

Figure S4. *Olfr* 3D index prediction, Related to Figure 4 and Tables S3 and S4. (A) Root mean square error (RMSE) per iteration of the cross-validation test for the Random Forest model used to predict 3D indexes. (B) Scatter plot illustrating the comparison of our 3D indexes versus the "Zolfr indexes" defined by (Zapiec and Mombaerts, 2020) from ISH data. For this comparison, these zones were numbered from 1 to 9 from the most dorsal to the most ventral. Black circles indicate *Olfrs* detected in our dataset; green circles are *Olfrs* for which indexes were predicted with Random Forest. The correlation coefficients computed separately on these two sets of *Olfrs* are respectively rho=0.92, p-value<$2x10^{-16}$ and rho=0.44, p-value>0.05. (C) Scatter plot showing the correlation of our 3D indexes with the "Tan Indexes" estimated by (Tan and Xie, 2018), who performed RNA-seq on 12 samples at different positions along the dorsal-ventral axis of the OM and estimated indexes using as reference the ~80 *Olfrs* analyzed in (Miyamichi et al., 2005) via ISH. Black circles indicate *Olfrs* detected in our dataset; green circles are *Olfrs* for which indexes were predicted with Random Forest. The correlation coefficients computed separately on these two sets of *Olfrs* are respectively rho=0.95, p-value<$2x10^{-16}$, and rho=0.68, p-value < $2x10^{-16}$.(D-F) In-situ hybridization experiment validating the predicted 3D spatial expression patterns for *Olfr309* (D), *Olfr727* (E), and *Olfr618* (F). Note that *Olfr618* is expressed in Zone 1, consistent with its predicted spatial expression pattern and calculated 3D index of 7.42 (Figure 4 N, O). Purple arrowheads indicate the location of ISH labeled cells. The dotted outline indicates the borders of the OM dissected and used in the RNA-seq experiments and for the construction of the 3D model.

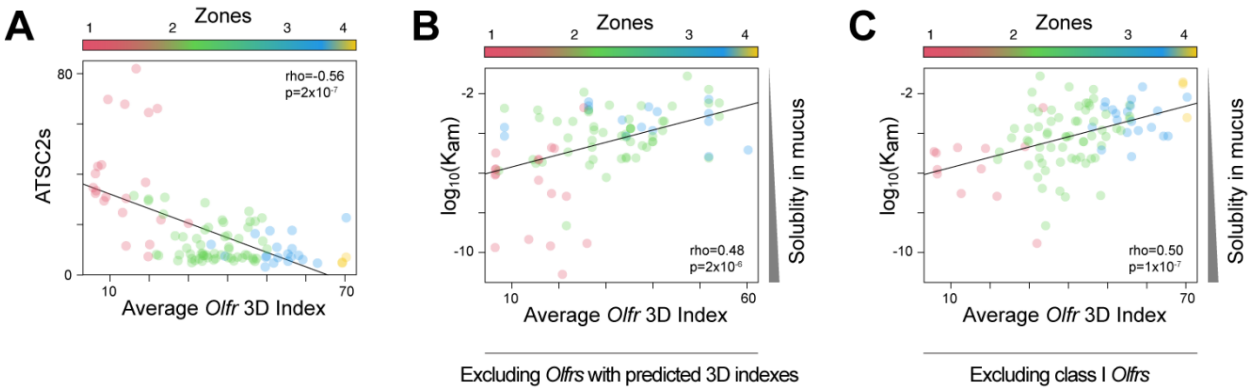

Figure S5. Physiological role of the zones, Related to Figure 6 and Table S6. Scatter plot illustrating the correlation between ATSC2s of the odorants and the average 3D indexes of their cognate Olfrs. Only odorants for which we know at least two cognate Olfrs (110) were used here. Odorants are colored according to the zone they belong to (defined as the zone with the highest average degree of belonging computed over all cognate receptors). (B) Scatter plot illustrating the correlation between air/mucus partition coefficients of the odorants and the average 3D indexes of their cognate Olfrs. Only odorants which are detected by Olfrs present in our TOMO-seq dataset (87) were used here. (C) Scatter plot illustrating the correlation between air/mucus partition coefficients of the odorants and the average 3D indexes of their cognate Olfrs. Only odorants which are detected by Class II Olfrs (101) were used here.