SUPPLEMENTAL FIGURES

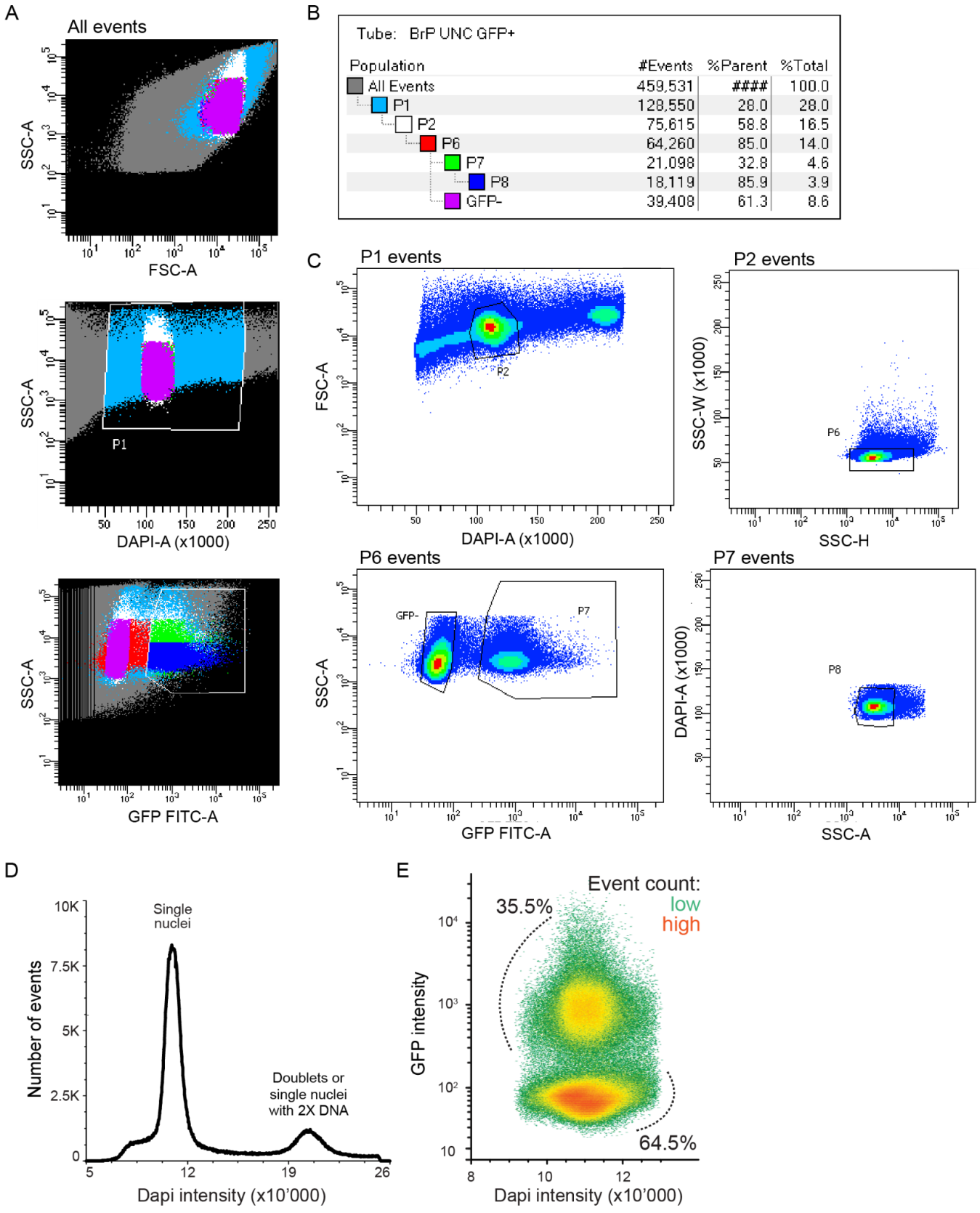
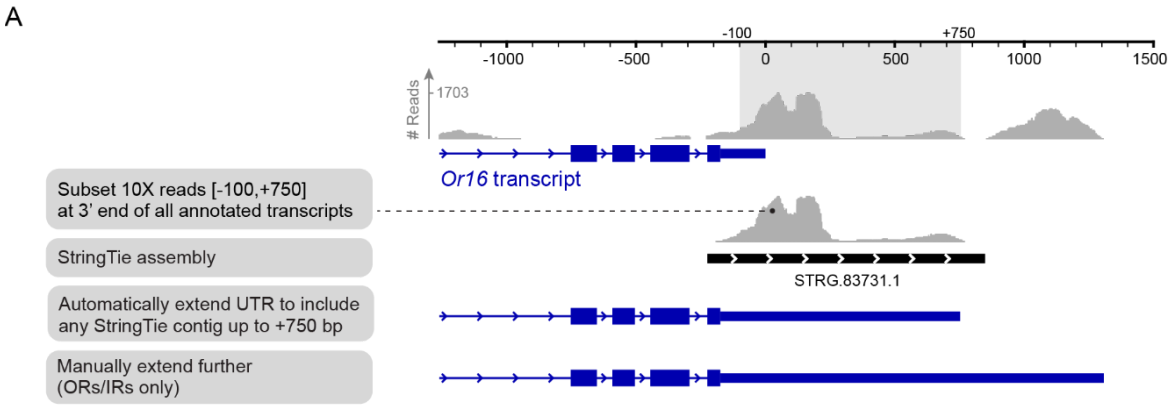
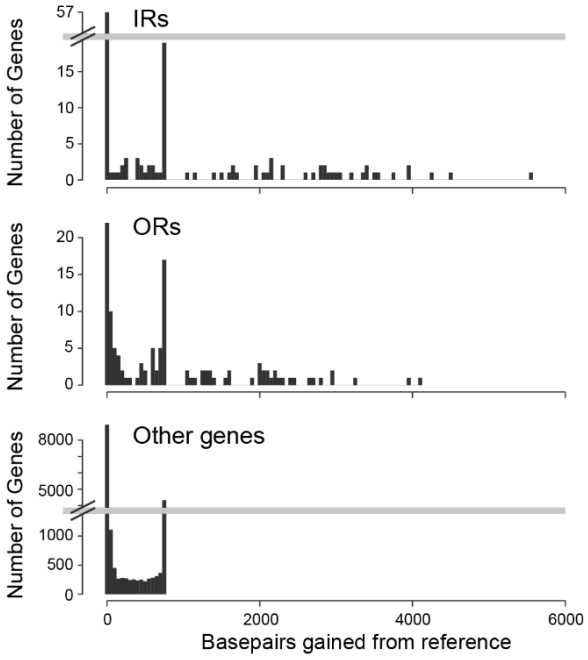
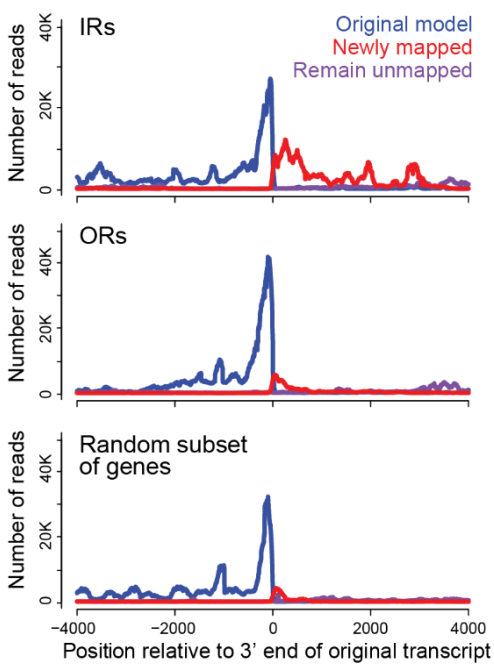


Figure S1. Fluorescent activated cell sorting of GFP-labeled nuclei from female antennal neurons. (**A**) Merged FACS data from 3 independent suspensions of antennal nuclei colored according to the populations shown in (B). (**B**) Table showing the number of events in subsets defined by sequential gates P1-P8 during one FACS run. P8 represents the final subset of GFP+ single nuclei used for 10X library prep. (**C**) Density plots showing sequential gating parameters for one FACS run. (**D**) Distribution of DAPI intensity across all P1 events. Events in the larger 1X peak were retained through gate P2 while those in the smaller 2X peak were discarded. (**E**) Density plot showing all P6 events according to DNA content (DAPI) and GFP intensity. Events in the upper GFP+ peak were retained through gate P7 while those in the lower GFP- peak were discarded. The retained GFP+ neurons represent ~35% of all antennal nuclei. FSC, forward scatter; SSC, side scatter; GFP-FITC, GFP intensity; DAPI, DNA marker intensity. -A, -W, -H correspond to average, width, height components.







C

	Total genes	Automatically extended	Manually extended
ORs	117	89 (76%)	38 (32%)
IRs	135	66 (49%)	40 (30%)
Other genes	18070	10544 (58%)	0

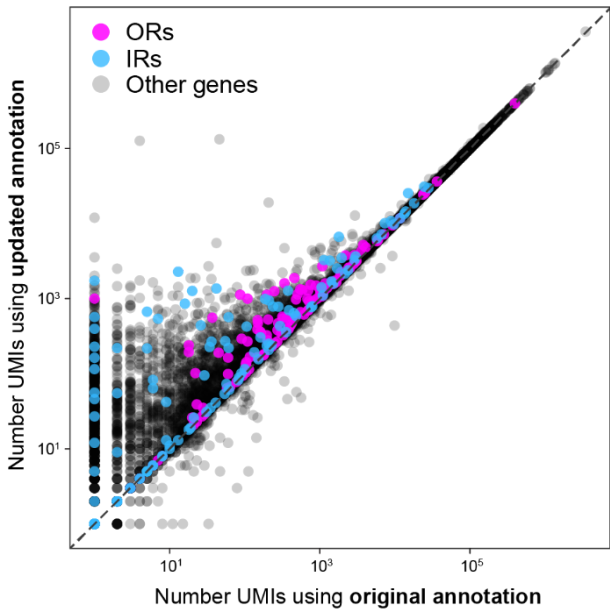
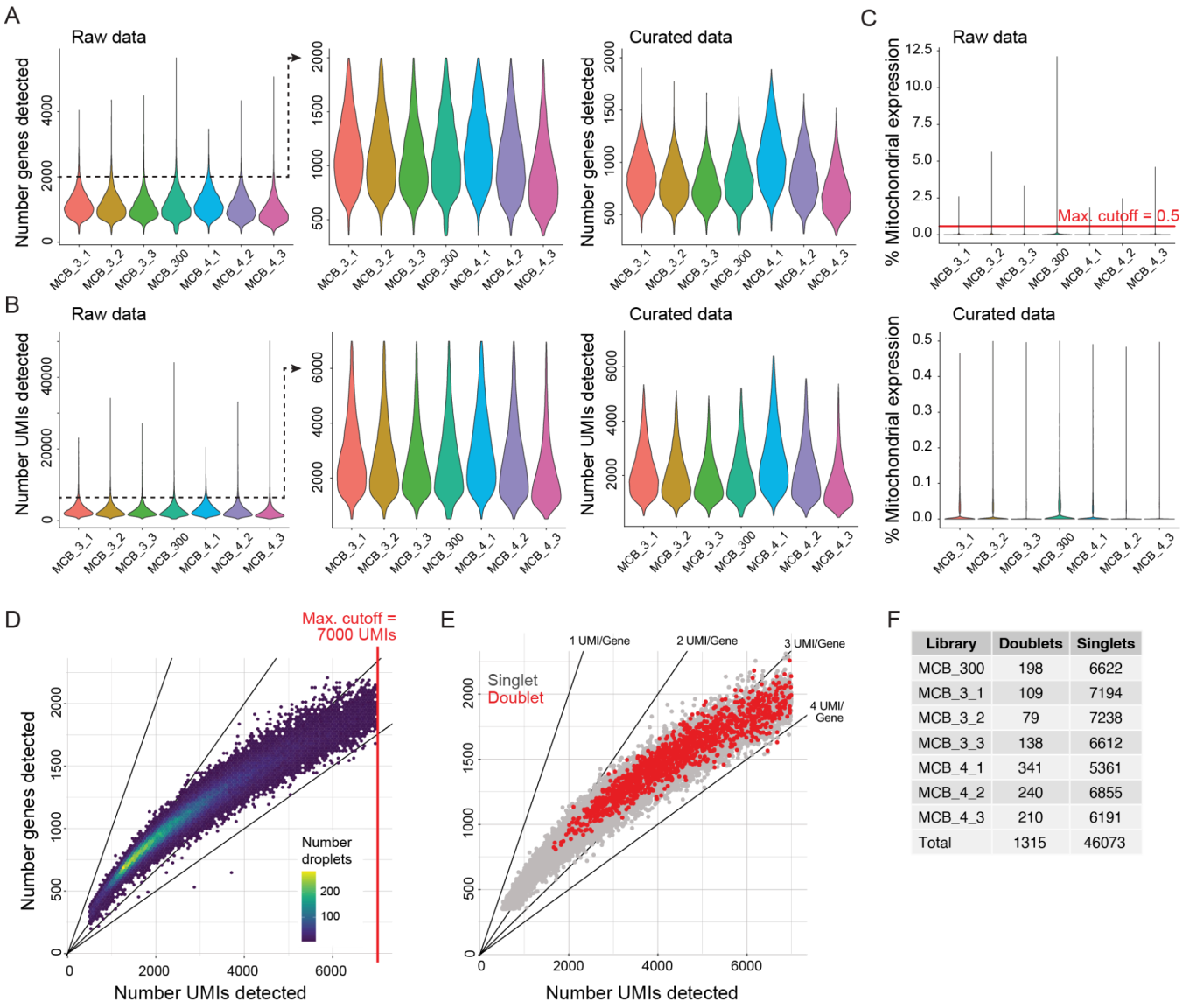


Figure S2. Reannotation or 3'UTRs in the AaegL5 genome. (**A**) Schematic of reannotation process for an example odorant receptor. Top trace shows pile up of 10X reads with base pair coordinates given in reference to the end of the originally annotated transcript. (**B**) Number of base pairs added to the 3' UTRs of ORs, IRs, and other genes. Many genes in all three categories were extended by exactly 750 bps as this was the extension cutoff used in the automated pipeline. (**C**) Number and fraction of genes that received extensions. (**D**) Distribution of summed 10X reads across the end of annotated ORs, IRs, or a random subset of genes. Colors highlight reads that were assigned to genes using the original annotation (blue), reads that were newly assigned using the updated annotation (red), and reads that were not assigned to the focal genes. (**E**) Scatterplot comparing the total number of UMIs assigned to given genes using the original and updated annotations.



F

Library	Doublets	Singlets
MCB_300	198	6622
MCB_3_1	109	7194
MCB_3_2	79	7238
MCB_3_3	138	6612
MCB_4_1	341	5361
MCB_4_2	240	6855
MCB_4_3	210	6191
Total	1315	46073

Figure S3. Preprocessing of droplets in 7 replicate snRNAseq libraries. (**A**–**C**) Distribution of genes detected (A), UMIs detected (B), and percent mtDNA transcripts (C) in each library before (raw) and after (curated) removal of ambient RNA contamination using the program SoupX[60]. Droplets with <350 genes, >7000 UMIs, or >0.5% mtDNA expression were discarded. (**D**) Density of droplets in which specific combinations of UMIs and genes were detected. (**E**) As in (D) but highlighting putative doublets detected using the program Solo[62]. (**F**) Summary of doublets detected in each library.

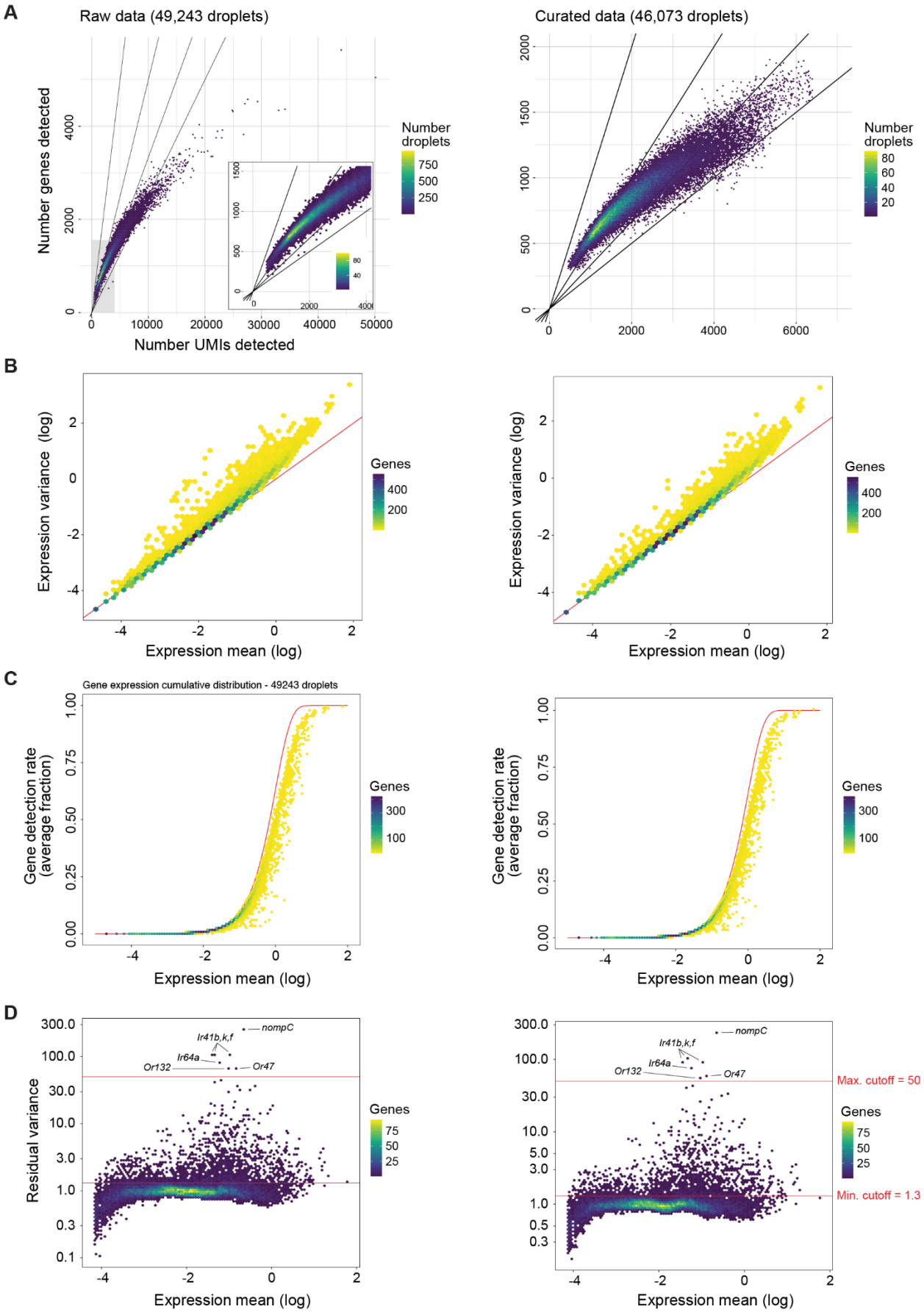


Figure S4. **Gene expression variability and normalization in raw and processed droplets from the full dataset.** (**A**) Average UMI per gene per droplet. (**B**) Linear modeling of gene expression dispersion. (**C**) Poisson modeling of gene expression deviation. Red line depicts the predicted cumulative distribution. (**D**) Residual expression variance by mean expression across all genes. The highlighted clip.range and rv.th threshold values were used for sctransform v2 normalization. In all panels, plots on the left and right show raw and processed data, respectively.

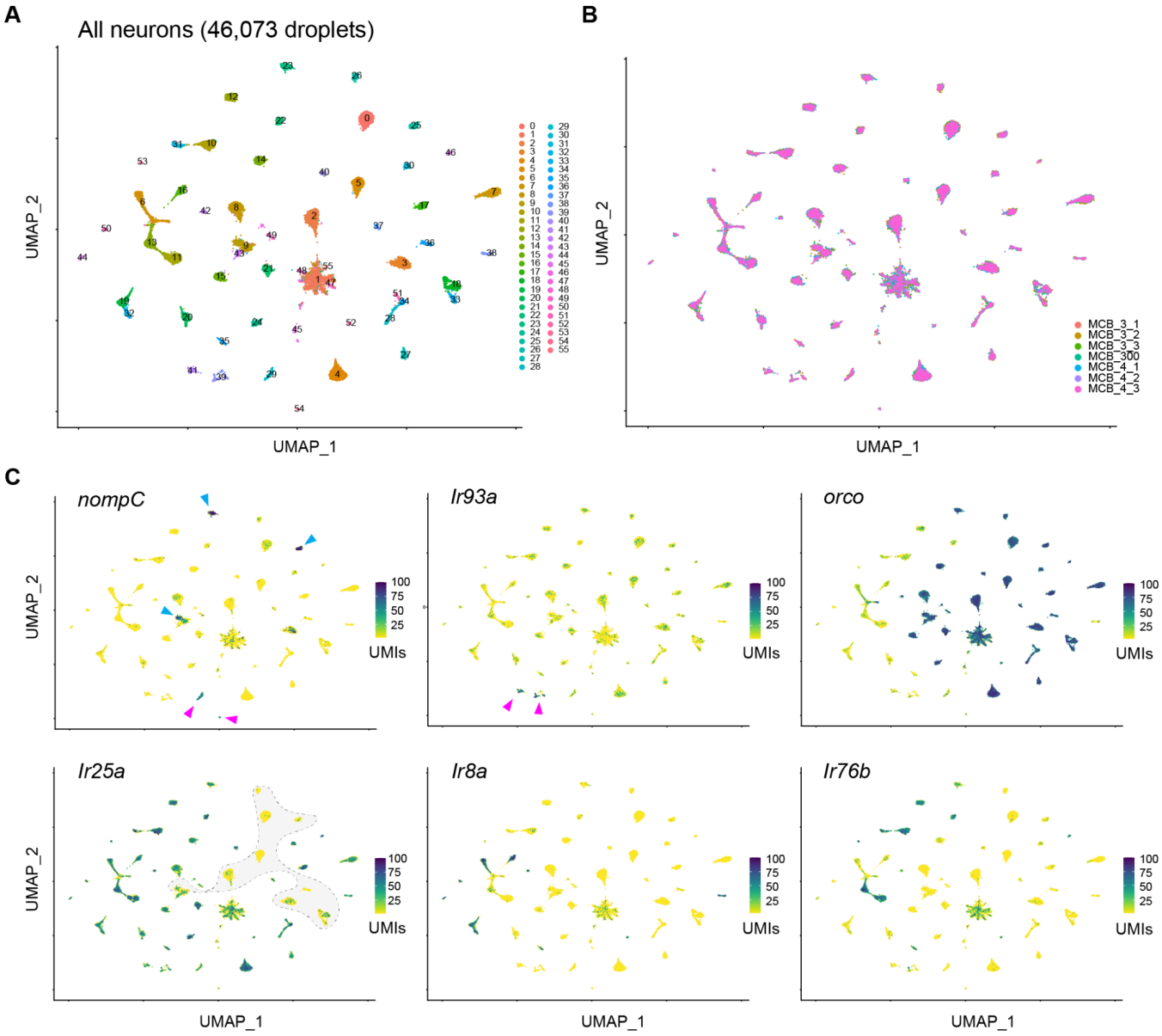


Figure S5. Details of UMAP clustering and marker gene expression in the all-neuron analysis. UMAP clustering for the full data set (as in Fig. 1D) highlighting clusters (**A**), 10X libraries (**B**), and expression of key marker genes (**C**). In (A), clusters 1, 47, 48, and 55 were low complexity 'junk' clusters located in the center of the plot, while cluster 45 was clearly distinct but could not be associated with a specific function. In (C), pink arrowheads indicate putative mechanosensory neuron clusters (*nompC* panel), and thermo/hygrosensory neuron clusters (*Ir93a* panel). The mechanosensory channel *nompC* is also highly expressed in several *orco*+ olfactory sensory neuron clusters (blue arrowheads). The biological significance of this expression is unknown.

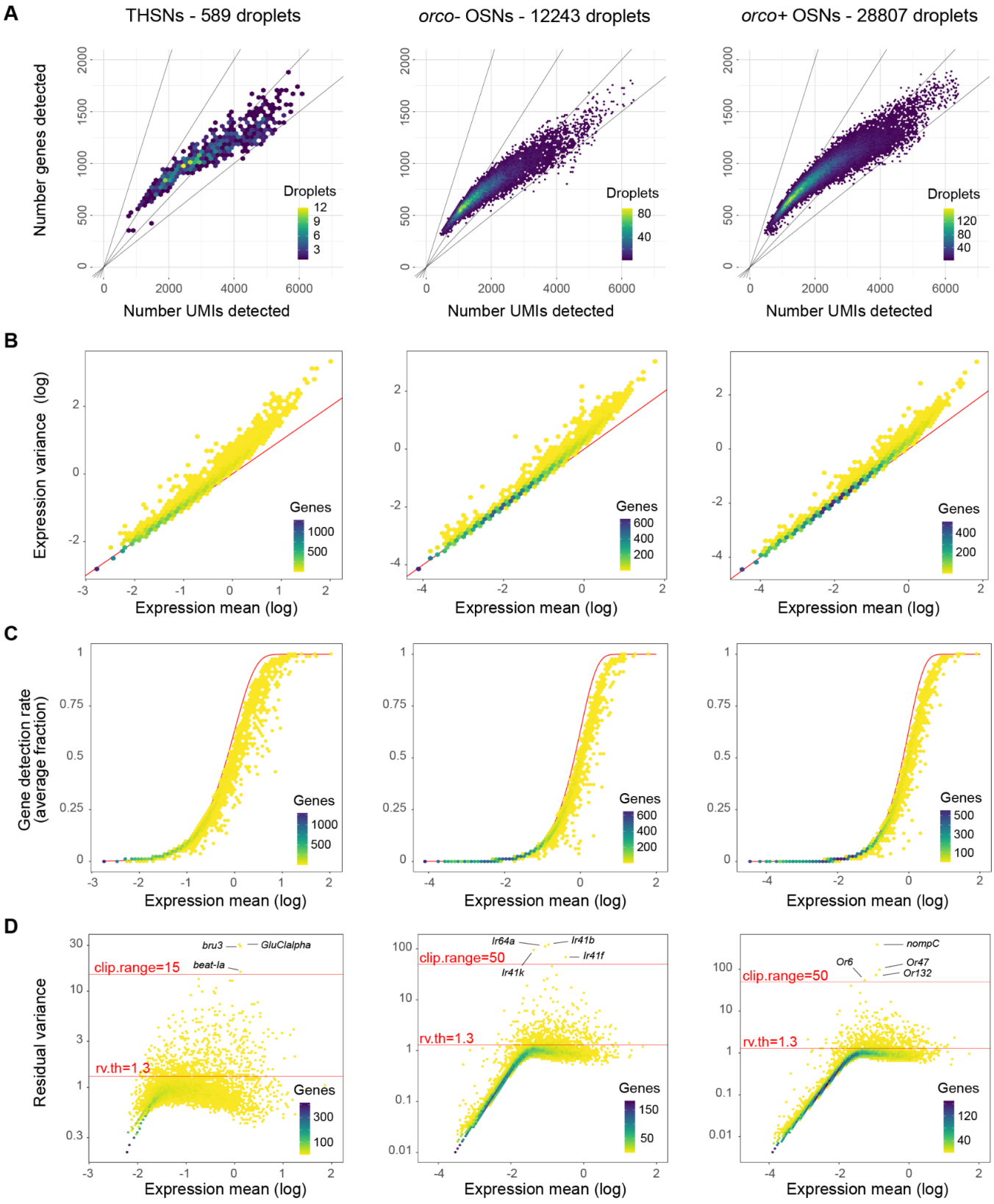


Figure S6. **Gene expression variability and normalization of subsetted sensory neurons.** (**A**) Average UMI per gene per droplet. (**B**) Linear modeling of gene expression dispersion. (**C**) Poisson modeling of gene expression deviation. Red line depicts the predicted cumulative distribution. (**D**) Residual expression variance by mean expression across all genes. The highlighted clip.range and rv.th threshold values were used for

sctransform v2 normalization. In all panels, plots on the left and right show raw and processed data, respectively. Left, middle, and right plots represent thermo/hygrosensory neurons, *orco*- OSNs, and *orco*+ OSNs, respectively.

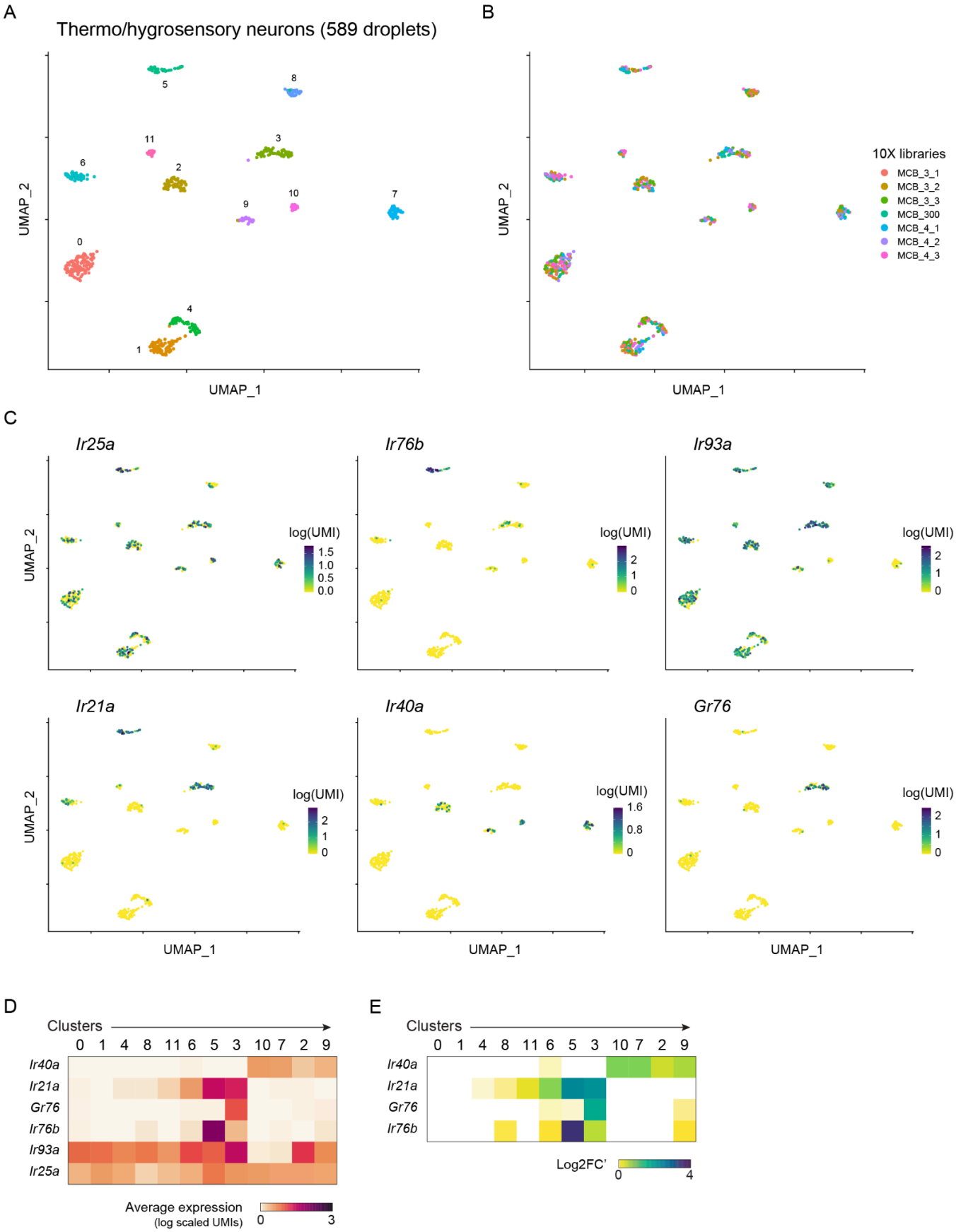


Figure S7. Details of UMAP clustering and receptor expression in putative thermo- and hygrosensory neuron subtypes. (**A**–**C**) UMAPs highlight nuclei assigned to different clusters (A), derived from different 10X libraries (B), or expressing different levels of key chemosensory receptors and coreceptors (C). (**D**–**E**) Heatmaps showing average expression (D) and differential expression (E) of chemosensory receptors across all clusters.

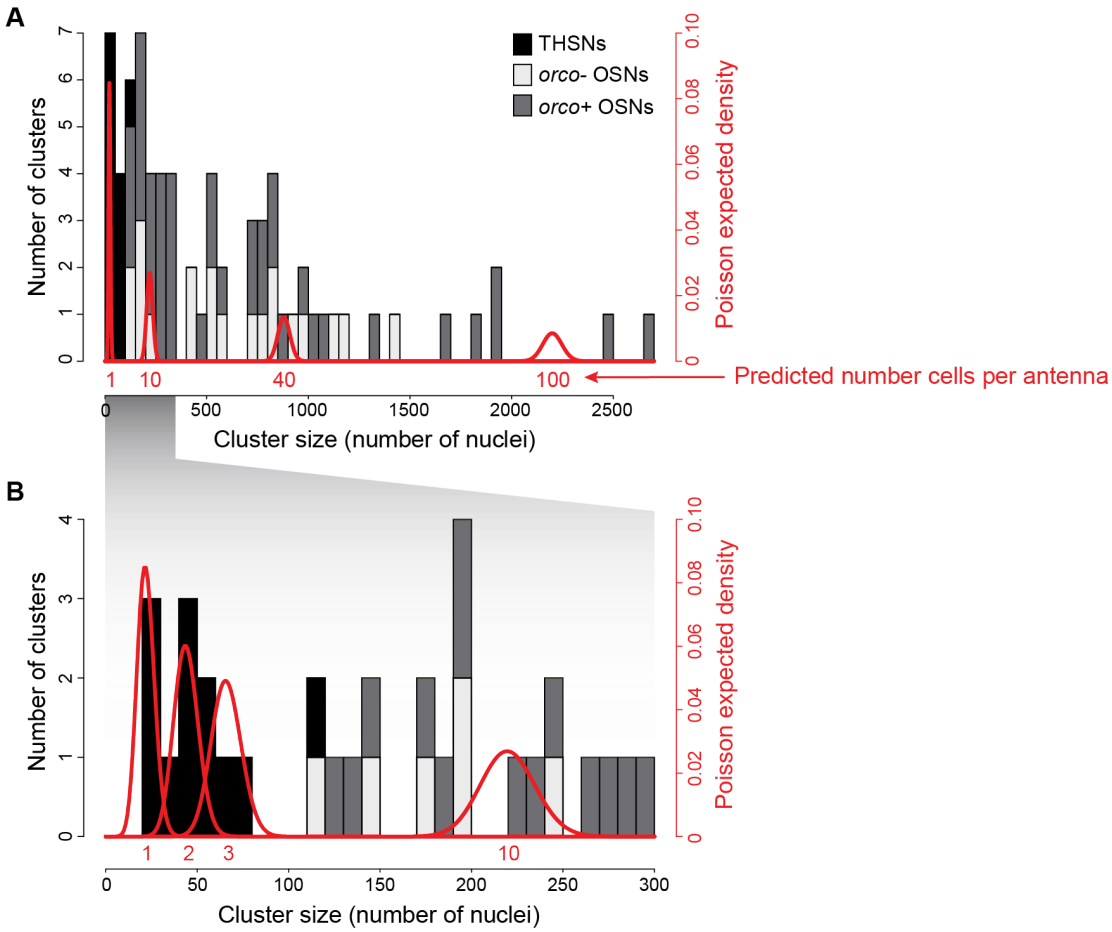


Figure S8. Inferred number of cells on a single female antenna belonging to identified sensory neuron subtypes. (**A**) Histogram showing size distribution of sensory neuron clusters (grey) overlaid by the expected densities for cell types comprising 1, 10, 40, or 100 neurons per antenna (red). Poisson densities were estimated under the assumption that the 46,073 sequenced nuclei were randomly drawn with replacement from the ~2000 neurons present on a female antenna[19]. (**B**) Zoom of (A), highlighting tiny populations of thermo/hygrosensory neurons (black) that we infer to comprise just 1-3 neurons per antenna.

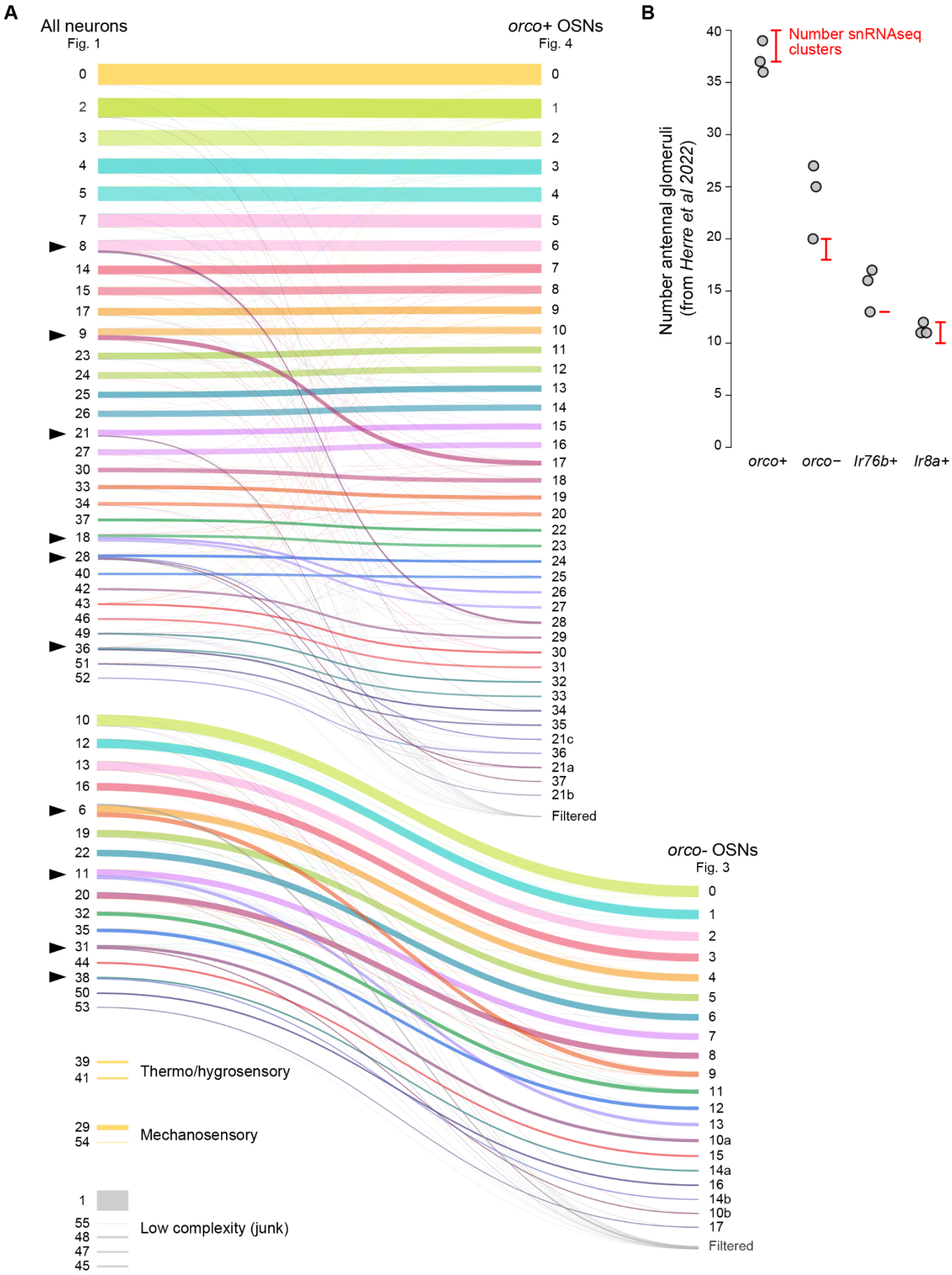


Figure S9. Correspondence between clusters in full and subsetted analyses with comparison to glomerulus counts. (**A**) Sankey plot showing how clusters in the original all neuron analysis (left; Fig. 1D) correspond to those in subsetted *orco+* OSN analysis (top right; Fig. 4) and *orco-* OSN analysis (bottom right; Fig. 3). Colored lines represent groups of nuclei, with line thickness proportional to the size of the group. Black arrowheads mark OSN clusters that were split in the subsetted analyses. (**B**) Number of *orco+*, *orco-*, *Ir76b+*, and *Ir8a+* glomeruli identified in three female brains by *Herre et al 2022* in the LVP strain of *Ae. aegypti* (grey dots)[12] compared to the number of corresponding clusters from our analysis (red). Glomerulus counts exclude those targeted by palp neurons. snRNAseq cluster counts (red) range from the number that express a unique complement of ligand-specific receptors to the total number. Note that *Ir76b+* and *Ir8a+* cluster numbers include those identified among *orco-* OSNs (Fig. 3) plus the 2-3 identified among *orco+* OSNs (Fig. 4).

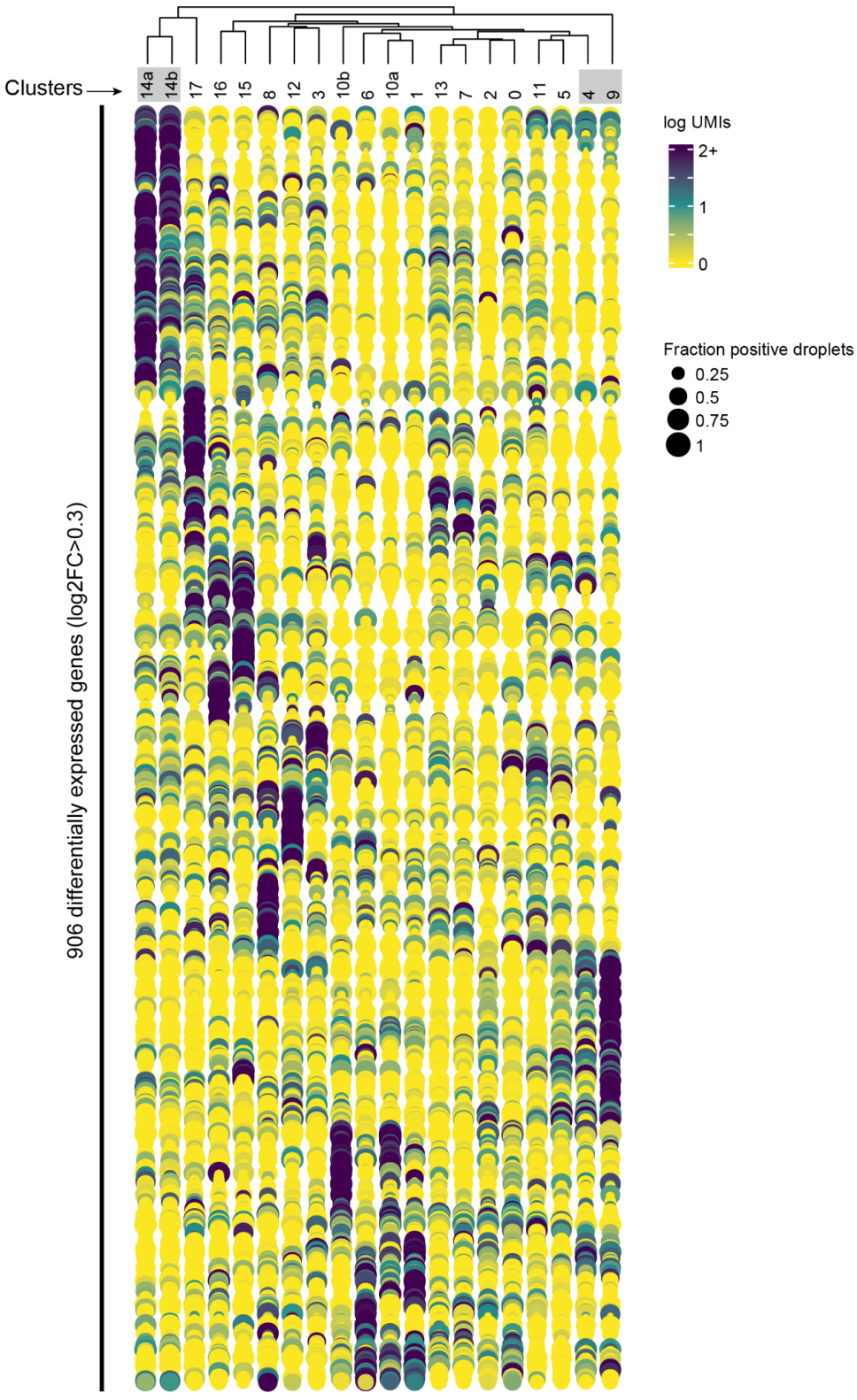


Figure S10. Dotplot of differential expression across 20 *orco*- OSN clusters. Grey boxes around cluster IDs highlight pairs that express the same complement of receptors and were merged in the main analysis (Fig. 3B).

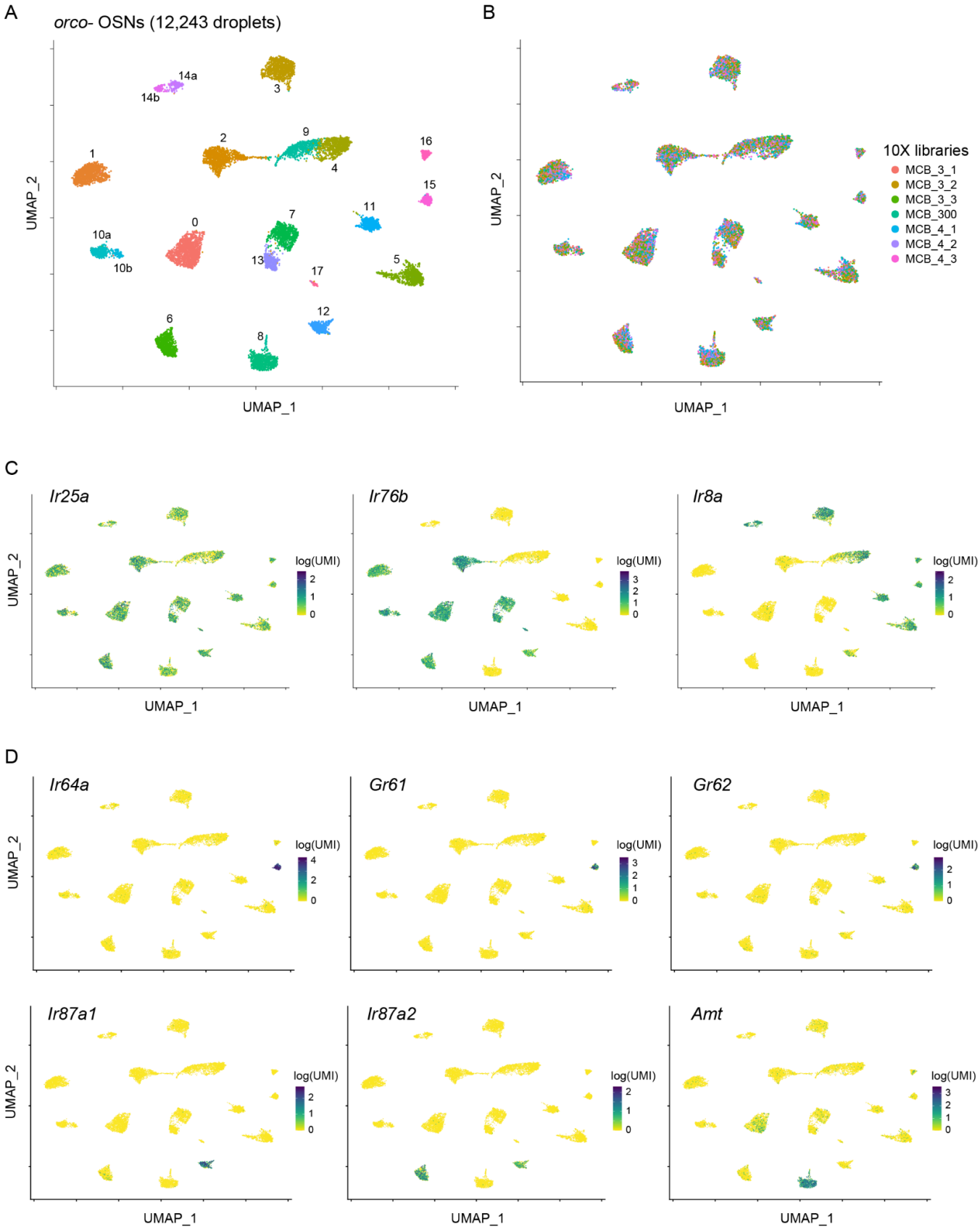


Figure S11. **Details of UMAP clustering of *orco*- OSN subtypes.** (**A**–**D**) UMAPs highlight nuclei assigned to different clusters (A), derived from different 10X libraries (B), and expressing different levels of key chemosensory co-receptors (C) or example ligand-specific receptors (D).

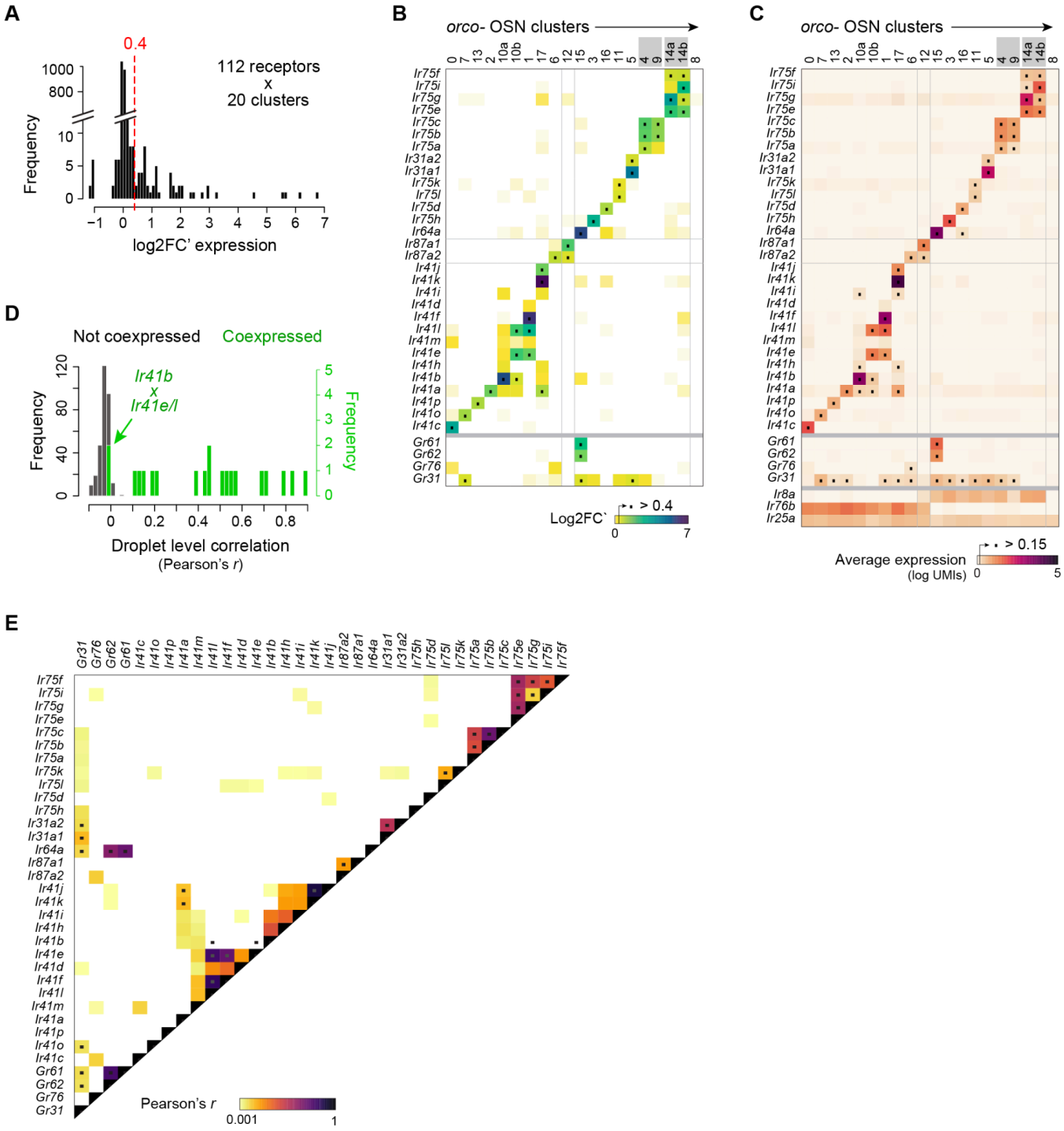


Figure S12. Summary of receptor expression across *orco*- OSN clusters. (**A**) Distribution of log2FC' values used to identify a cutoff (dashed red line) for expression calls. All 112 receptors that were detected in one or more droplets were analyzed across all 20 clusters. (**B**) Heatmap showing differential expression across clusters. Plot includes all ligand-specific *IRs* (n=30), *ORs* (n=0), and *GRs* (n=4) with log2FC' above 0.15 in any cluster, but black dots mark cases where log2FC' exceeded the 0.4 cutoff used to call expression. Grey boxes around cluster IDs highlight pairs that express the same complement of receptors and were merged in the main analysis (Fig. 3). (**C**) Same as (B) but showing absolute expression, with black dots marking an alternative average log-scaled expression > 0.15 cutoff. (**D**) Distribution of droplet-level correlations (Pearson's *r*) for pairs of receptors that were (red) or were not (grey) called as coexpressed based on the log2FC' cutoff. (**E**) Pairwise droplet-level correlations for the receptors shown in (B–C). Black dots mark pairs called as coexpressed. Note that 'coexpressed' receptors

showed elevated correlations in all but two cases. The exceptions (*Ir41b*~*Ir41e* and *Ir41b*~*Ir41l*) reflect a clustering artifact wherein a few *Ir41b*+ droplets from cluster 10a were erroneously lumped with *Ir41e*+/*Ir41l*+ droplets in cluster 10b.

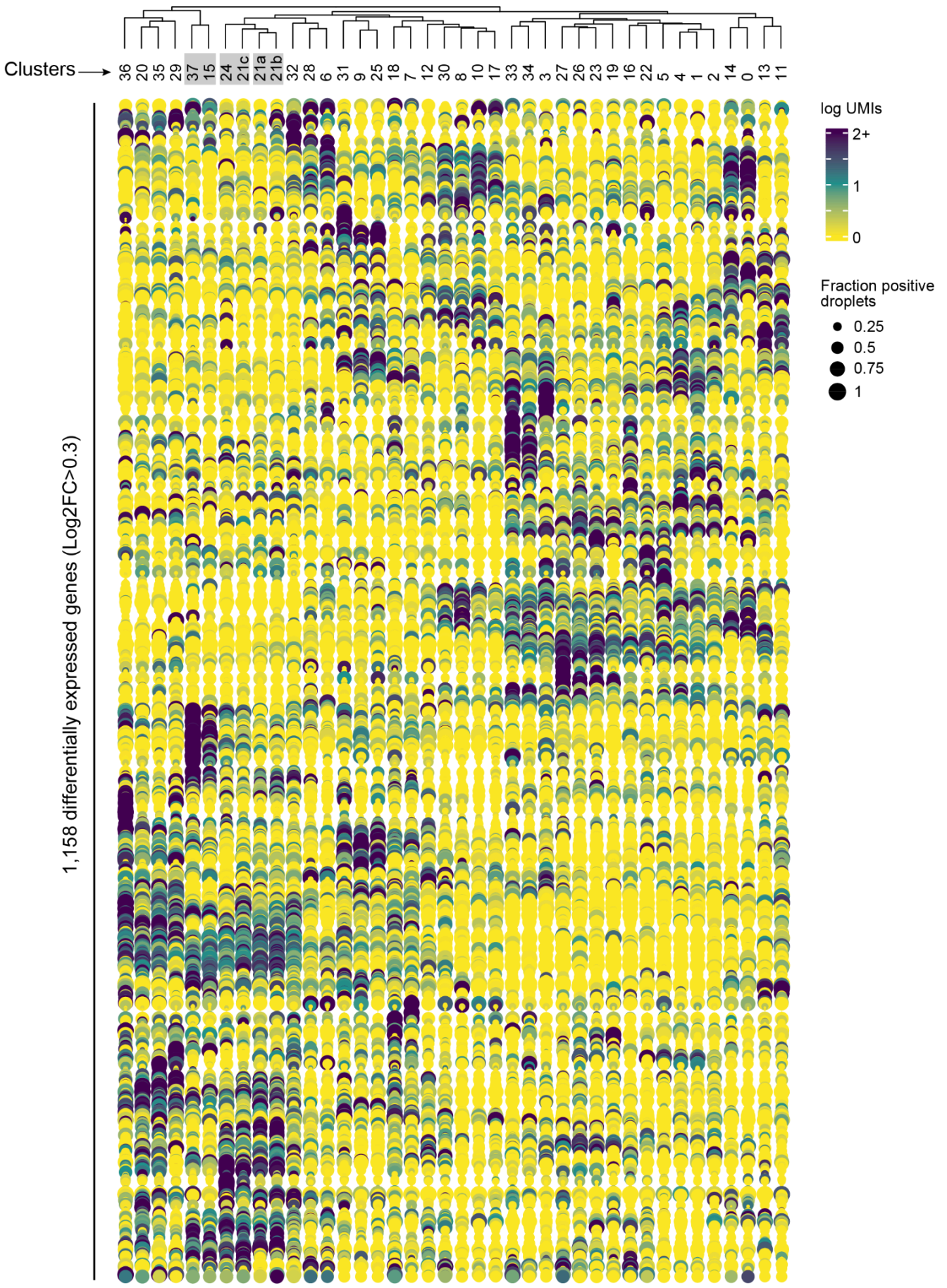


Figure S13. Dotplot of differential expression across 40 *orco*+ OSN clusters. Grey boxes around cluster IDs highlight pairs that express the same complement of receptors and were merged in the main analysis (Fig. 4B).

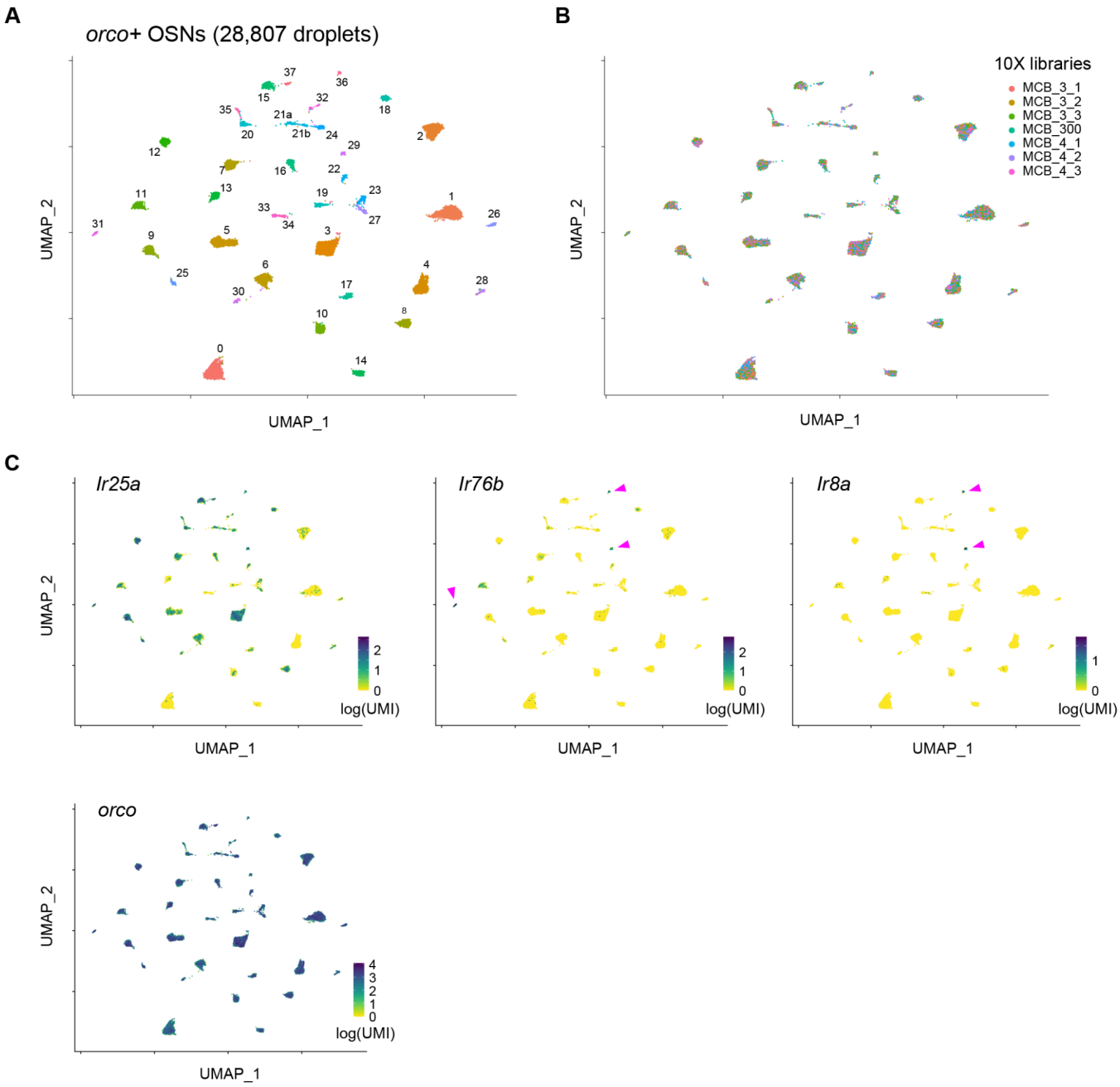


Figure S14. Details of UMAP clustering and co-receptor expression in *orco*+ OSN subtypes. Plots highlight nuclei assigned to different clusters (**A**), derived from different 10X libraries (**B**), and expressing different levels of key chemosensory co-receptors (**C**). Pink arrow in (C) mark the few clusters showing significant *Ir76b* (n=3) and *Ir8a* (n=2) expression.

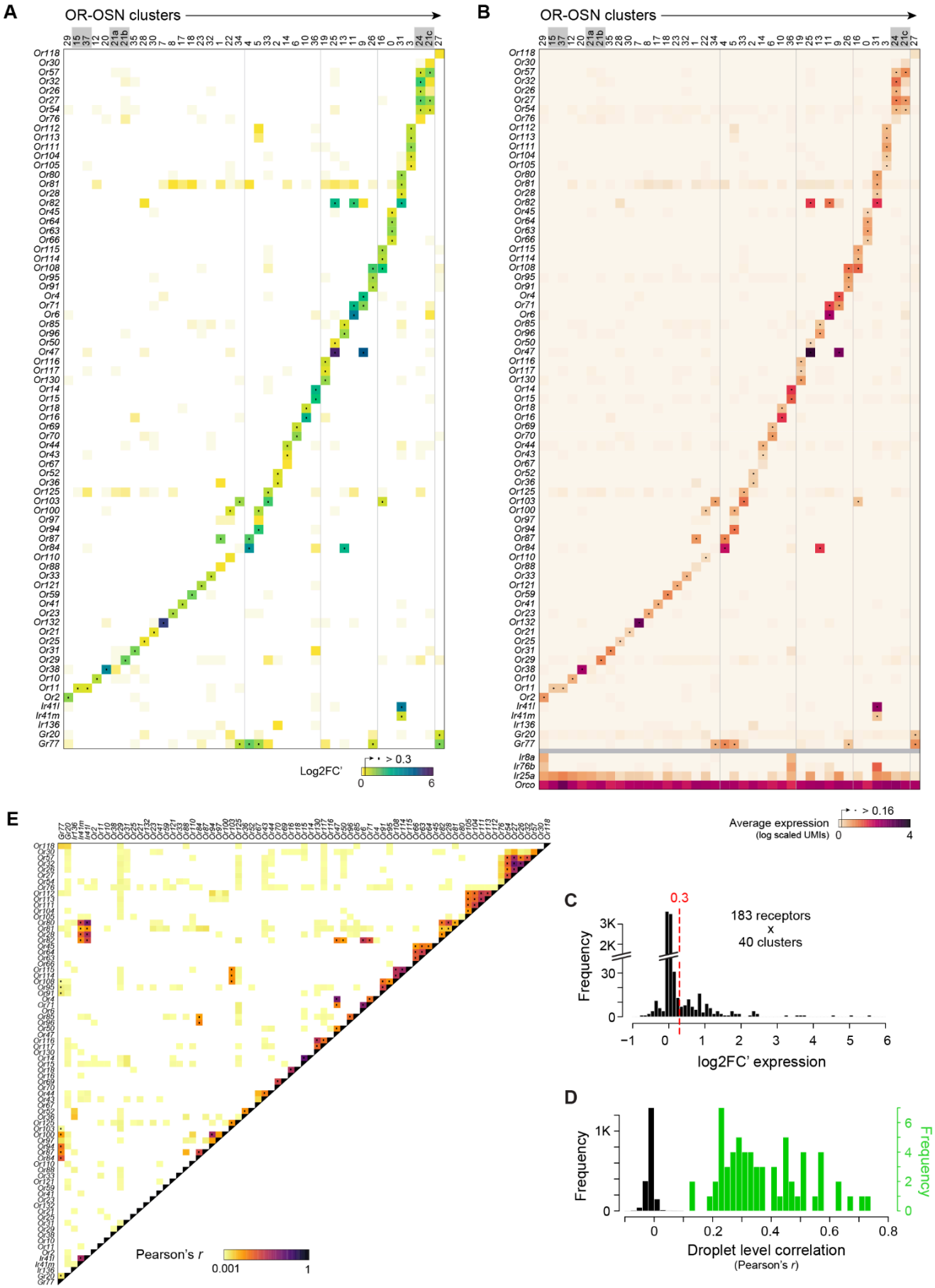


Figure S15. Summary of receptor expression across *orco*+ OSN clusters. (**A**) Heatmap showing differential expression across clusters. Plot includes all ligand-specific receptors (*ORs*, *IRs*, *GRs*) with log2FC' above 0.15 in any cluster, but black dots mark cases where log2FC' exceeded the 0.3 cutoff (see C). Grey boxes around cluster IDs highlight pairs that express the same or similar complement of receptors and were merged in the main analysis (Fig. 4). (**B**) Same as (A) but showing absolute expression, with black dots marking an alternative average log-scaled expression > 0.15 cutoff.(**C**) Distribution of log2FC' values used to identify the 0.3 cutoff (dashed red line) for expression calls. All 183 receptors that were detected in at least one or more droplets were analyzed across all 37 clusters. (**D**) Distribution of droplet-level correlations (Pearson's *r*) for pairs of ORs that were (green) or were not (grey) coexpressed according to the log2FC' threshold. Note the lack of overlap between the two distributions, indicating that all pairs of ORs we consider coexpressed were found not only in the same clusters, but also in the same droplets more often than expected by chance. (**E**) Pairwise droplet-level correlations for the set of receptors shown in (A–B). Black dots mark coexpressed pairs.

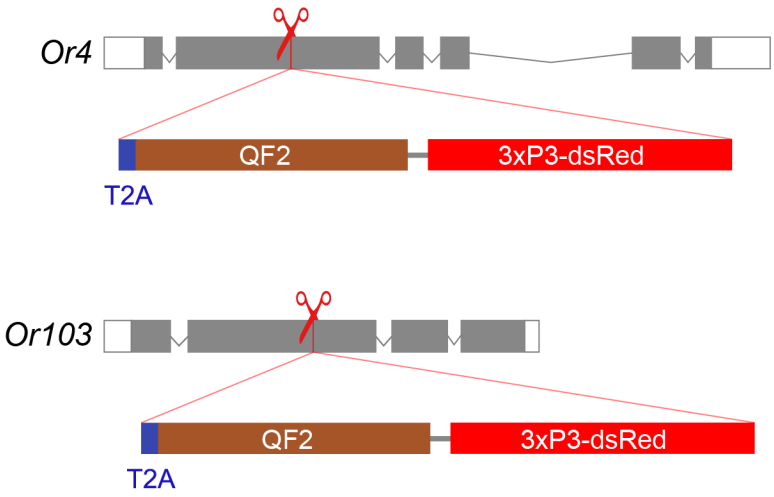


Figure S16. Schematic of *Or4* and *Or103* knock-in constructs. CRISPR-mediated homologous recombination was used to insert an in-frame T2A-QF2 element into the coding sequences of *Or4* and *Or103*. T2A is a ribosomal skipping sequence[66], and QF2 is a fungal transcription factor[47]. For both loci, the insertion is expected to disrupt the coding sequence of the native receptor and result in the dual translation of the partial receptor coding sequence and QF2 from a single mRNA transcript[67].