Supplemental Experimental Procedures

Mice:

GCD-IRES-tauGFP and *OMP-IRES-GFP* mice were obtained from the Jackson Laboratory (strain B6;129P2-Gucy2d^{tm2Mom}/MomJ, stock number 006704 and strain B6;129P2-Omp^{tm3Mom}/MomJ, stock number 006667, respectively). *OR174-9-IRES-tauGFP* mice were obtained from the Axel Laboratory (Sosulski et al., 2011). *Emx1-IRES-Cre* and *GCaMP3* mice were obtained from the Jackson Laboratory (strain B6.129S2-Emx1$^{tm1(cre)Krj}$/J, stock number 005628 and strain B6;129S-Gt(ROSA)26Sor$^{tm38(CAG-GCaMP3)Hze}$/J and stock number 014538, respectively). Unless otherwise noted, all experiments on wild type mice were performed on 6-8 week old C57/BL6 male mice (Jackson Laboratory). All mouse husbandry and experiments were performed following institutional and federal guidelines and approved by Harvard Medical School's Institutional Animal Care and Use Committee.

Epithelial single cell dissociation:

Mice were euthanized with a lethal dose of Xylazine (~50 mg/mouse, Lloyd). Using 1X phosphate-buffered saline (PBS, VWR) to keep the epithelium moist during dissection, the head of the mouse was severed and the olfactory epithelium was rapidly dissected and placed in PBS. The epithelium was then transferred to a round-bottomed glass dish containing 1 mL of papain solution (one vial of Papain, (Worthington) dissolved in 5 mL of Earle's Balanced Salt Solution (Worthington) and then equilibrated 10 minutes at 37 °C, 5% CO_2) and 100 µL of DNAse solution (one vial of DNAse (Worthington) dissolved in 1 mL of Earle's Balanced Salt Solution (Worthington)). Bone was removed from the epithelium under a dissecting microscope (Leica MZ75) and the resultant tissue was placed in a 5 mL Falcon tube (Becton Dickinson) with an additional 1 mL of Papain solution and 100 µL of DNA solution and rocked gently for 30 minutes at 37 °C. The tissue was then gently triturated with a 5 mL stripette 10-15 times and the non-dissociated pieces of tissue were allowed to sediment by gravity for approximately 3 minutes. The supernatant was transferred to a fresh 5 mL Falcon tube and centrifuged 5 minutes at 300 x g. The supernatant was decanted, the cells were washed twice with 5 mL DMEM, high glucose, minus glutamine (Life Technologies)/10% heat inactivated Fetal Bovine Serum (FBS, Life Technologies) and resuspended in 1 mL DMEM/10% FBS to use for experimentation.

FACS:

Olfactory epithelial cells were dissociated as described above. To label hematopoetic cells, cell suspensions were incubated with PE-Cy5-conjugated rat anti-mouse CD45 antibody (1:2000, BD Pharmingen) for one hour at room temperature and then were washed twice with 1 mL DMEM/10% FBS. The cells were then incubated with Hoescht to allow the differential labeling of dead cells, before being placed on ice until they were sorted. Cells were sorted using either an Astrios EQ (Beckman Coulter, Fort Collins, CO) or a FACSAriaII (Becton Dickinson, San Jose, CA). Following the exclusion of dead cells and CD45 positive immune cells, GFP positive cells were then sorted to >99% purity by setting gates based on the fluorescent profile of cells derived from wild type, non-fluorescent mice. Cells were sorted directly into Trizol (Invitrogen) prior to RNA extraction (see below).

RNA isolation:

Fluorescent cells were sorted directly into 750 µLTrizol (Invitrogen) and samples were then vortexed to homogeneity before being placed on ice for the duration of the FACS session. All subsequent RNA work was performed in a RNA-dedicated PCR workstation (Air Clean Systems) that was UV-irradiated for 20 minutes prior to use and then was subsequently decontaminated with RNAseZap wipes (Ambion). Samples were passed through a 1 mL insulin syringe (Westnet) four times and the samples were allowed to sit for 5 minutes at room temperature. 200 µL of chloroform (Sigma) was then added to each sample, and the tubes were shaken vigorously by hand for 15 seconds and then placed at room temperature for 3 minutes. Samples were centrifuged at 12,000 x g for 15 minutes at 4 °C. The aqueous phase was removed to a nuclease free microcentrifuge tube (Ambion) and 10 µg of RNAse-free glycogen (Life Technologies) was added as a carrier along with 500 µL of 100% 2-propanol (Sigma). The samples were inverted several times to homogenize and incubated for 10 minutes at room temperature. The samples were then centrifuged at 12,000 x g for 10 minutes at 4 °C and the supernatant was removed with a pipette. The RNA pellet was washed with 1 mL of 75% ethanol and then centrifuged for five minutes at 4 °C at 7,500 x

g. The supernatant was pipetted off, the pellet was allowed to air dry for five minutes and was resuspended in 5-10 µL of nuclease free water (Ambion) and stored at -80 °C until further use.

cDNA library generation:

cDNA libraries were generated from isolated RNA using a modified version of the Smart-Seq2 protocol that was optimized for reliability and sensitivity using "spike" RNAs (Ambion) (Picelli et al., 2014). Briefly, 0.6 µL of RNA was added to 8 µL lysis buffer (1.6 mM dNTPs, 2.5 µM oligo(dT)-VN primer (5'-AAGCAGTGGTATCAACGCAGAGTACT$_{30}$VN-3', where 'N' is any base and 'V' is A, C, or G), 0.2 µL RNAse inhibitor (Ambion), 0.1% TX-100 v/v) in a nuclease free PCR tube and the samples were incubated in a Master Cycler ProS thermocycler (Eppendorf) for 3 minutes at 72 °C before being held at 4 °C for > 1 minute. Next, 11.4 µL of reverse transcriptase mix (10 µL Superscript II (Invitrogen), 5 µL RNAse inhibitor (Ambion), 40 µL 5X first strand buffer (Invitrogen), 10 µL 0.1M DTT (Invitrogen), 8 µL 5 M Betaine, 37.8 µL nuclease free water (Ambion), 1.2 µL 1M MgCl2, 2 µL of TSO (5'-AAGCAGTGGTATCAACGCAGAGTACATrGrG+G-3', which contains two riboguanisines (rG) and one LNA-modified guanisine (+G) to facilitate template switching) was added to each sample and mixed by pipetting. The samples were incubated in the thermocycler for 90 minutes at 42 °C and then were subjected to 10 cycles of 2 minutes at 50 °C followed by 2 minutes at 42 °C, one 15 minute incubation at 70 °C, and a final hold at 4 °C lasting at least 1 minute. Next the samples were amplified by adding 30 µL of PCR reaction mix (300 µL Kappa HiFi HotStart ReadyMix (2X KAPA Biosystems), 6 µL ISPCR primer (5'-AAGCAGTGGTATCAACGCAGAGT-3'), and 54 µL nuclease free water). The reactions were mixed by pipetting and incubated at 98 °C for 4 minutes followed by 24 cycles of 98 °C for 20 seconds, 67 °C for 15 seconds, and 72 °C for 6 minutes. A final incubation was performed at 72 °C for 5 minutes before the samples were placed at 4 °C for > 1 minute. The resultant cDNA was isolated using an Agencourt Ampure XP kit (Beckman Coulter) following the manufacturer's instructions and was subsequently stored at -80 °C until further use.

RNA sequencing:

cDNA libraries were sheared to a size of approximately 250 bases using a Covaris S2. The sheared cDNA was run on a Wafergen Apollo machine using the Kapa Genomic Library Construction Kit (Kapa Biosystems, KK8234). Ten cycles of PCR were run after the samples were removed from the robot and the reactions were cleaned using magnetic beads. Quality control was performed using an Agilent 2200 Tape Station with a D1000 High Sensitivity Tape with ladder provided. The morphology and overall concentration of the samples were assessed and those samples that passed a concentration cutoff were subjected to qPCR to more accurately determine concentration. qPCR was run with SYBR green using the KAPA SYBR FAST Universal 2X qPCR Master Mix reagent (Kapa Biosystems, KK4602) and primers complementary to the P5 and P7 regions of the adapter sequences. Serial dilutions of PhiX were used to generate a standard curve, which was, in turn, used to determine the concentration of cDNA in each sample prior to sequencing. Samples were sequenced on an Illumina 2500 in rapid mode on a single lane of SR50.

Alignment and differential expression:

RNA-sequencing reads were processed using the RNA-seq pipeline implemented in version 0.8.3a-9483413 of the bcbio-nextgen analysis project. Briefly, poor quality bases with PHRED scores less than five (Macmanes, 2014), contaminant adapter sequences, and polyA tails were trimmed from the ends of reads with cutadapt version 1.4.2, discarding reads shorter than twenty bases. A STAR (Dobin et al., 2013) index was created from a combination of the *Mus musculus* version 10 (mm10) build of the mouse genome and the Ensembl release 75 gene annotation. Trimmed reads were aligned to the STAR index, discarding reads with ten or more multiple matches to the genome. Quality metrics including mapping percentage, rRNA contamination, average coverage across the length of the genes, read quality, adapter contamination and others were calculated using a combination of FastQC, RNA-SeQC (DeLuca et al., 2012), and custom functions from bcbio-nextgen and bcbio.rnaseq (available upon request). Read mapping to genes were counted using featureCounts (Liao et al., 2014) version 1.4.4, excluding reads mapping multiple times to the genome and reads that could not be uniquely assigned to a gene. Counts were normalized and differential expression between cell types was called at the level of the gene using DESeq2

(Love et al., 2014) version 1.6.3. To identify gene families that were expressed in GC-D cells, the data set was filtered to identify transmembrane protein-encoding genes that were at least ten-fold enriched in GC-D cells with an adjusted p-value of < 0.05 after correcting for multiple comparisons using the methods described by Benjamini and colleagues (Benjamini et al., 2001). Despite using these stringent filtering criteria, with very large gene families one might still observe the expression of an occasional family member by chance and thus the data were filtered further to only consider gene families in which at least three members were ten-fold enriched with an adjusted p-value of < 0.05. From this list of gene families, the *Ms4a* genes were selected for further study as they exhibited significant molecular diversity to encode chemoreceptors (see Figure 2) and had relatively uncharacterized function.

Nanostring:

10,000 GFP positive cells from *Gucy2d-IRES-GFP* or *Omp-GFP* mice were sorted into Trizol and the RNA was isolated as described above. Three biological replicate RNA samples were hybridized to Nanostring probes using nCounter Elements reagents according to the manufacturer's specifications. The protocol was modified to perform the hybridization step at 67 °C for 48 hours to maximize the detection of low abundance transcripts. RNA molecules that hybridized to probe were captured and quantified using an automated Nanostring prep station following the manufacturer's instructions. The resultant data were analyzed using nSolver software. Briefly, the average number of detected molecules for six internal negative control probes (whose complementary sequences are not present in the mouse genome) was used to calculate a rate of non-specific hybridization. After subtracting the amount of binding resulting from non-specific interactions, the number of molecules of each RNA transcript found in GC-D samples and OMP samples was compared using Student's t-test. See Supplemental Table 2 for probe sequences used in these experiments.

Multi species alignment of Mus musculus MS4A proteins:

FASTA format sequences of the indicated *Mus musculus* MS4A proteins were downloaded from the NCBI protein database and aligned using the PRALINE sequence alignment program on the Centre for Integrative Bioinformatics VU website using the default settings. Amino acid conservation across family members was scored using the PRALINE default settings where the least conserved amino acids were given a 0 score and the most conserved amino acids were assigned a 10 (Simossis and Heringa, 2005). TOPCONS was used to determine the predicted topology of the MS4A family member that was used on the top line of the alignment. All topographical representations were generated using the Protter program and manually entering the topographical orientation of the MS4A protein as predicted by TOPCONS.

Phylogenetic and selection analyses with MS4A *genes:*

MS4A sequences were retrieved from both Ensembl and NCBI databases and imported into Geneious v8 (Biomatters Ltd). We chose 37 representative taxa from all the major mammalian lineages (see Supplemental Table 1 for full list). When a gene had more than one predicted isoform, the sequence that contained the longest open-reading frame was selected. Coding DNA sequences were translated, aligned with MAFFT v7.017 (Katoh and Standley, 2013a, b) using the E-INS-i algorithm, the BLOSUM80 scoring matrix, and a gap-opening penalty of 1. Sequences were then back-translated into codons. For phylogenetic reconstruction of the multigene family tree, the OpenMPI version of MrBayes v3.2.1 (Ronquist et al., 2012) and the GTR+I+G model as determined by jModelTest 2.1.7 (Darriba et al., 2012) were used. The final dataset consisted of 411 sequences and 447 characters corresponding to sites present in at least 75 percent of the aligned sequences. For individual gene tree reconstructions and evolutionary analyses, sequence subsets were extracted based on their group membership as predicted based on the multigene family tree. Sequences were then realigned corresponding to each subset as above, the resulting alignments were trimmed to remove positions that contained gaps in the majority of sequences. The phylogenetic reconstruction was carried out using the OpenMPI version of RAxML v8 (Stamatakis, 2014).

To identify branches under episodic positive selection, the random-effects likelihood branch-site method (BS-REL) (Kosakovsky Pond et al., 2011) was used as implemented in the HyPhy package (Pond et al., 2005). The branch-site models allow the nonsynonymous to synonymous substitution rate ratio ω (d_N/d_S) to vary both among amino acid sites in the protein and across branches on the tree to detect positive selection affecting specific sites along particular lineages (Anisimova and Yang, 2007). We identified evidence for site-specific positive selection in MS4A homologs using the codeml program in the PAML

v4.8 software package (Anisimova and Yang, 2007). In brief, we compared different site models, in which the evolutionary rate ω is allowed to vary among sites. Here, we focused on comparison of model pairs: M1a (neutral; codon values of ω fitted into two discrete site classes between 0 and 1) versus M2a (positive selection; similar to M1a but with one additional class allowing ω>1); M7 (neutral; values of ω following a beta distribution with ω=1 maximum) versus M8 (positive selection; similar to M7 but with one additional class allowing ω>1); and M8a (neutral; similar to M7 but with one fixed class with ω=1) versus M8. Multiple starting values of ω were used, and either the F3x4 or F61 model of codon frequencies. To evaluate whether the models allowing positive selection provided a significantly better fit to the data, likelihood ratio tests were used. Notably, the M1a-M2a comparison is more stringent and can lack power to detect signatures of diversifying selection compared to the M7-M8 models, which impose less constraints on the distribution of ω. Finally, the M8a vs. M8 comparison can be used to contrast the potential role of reduced purifying selection (or relaxation) versus positive selection. When the null model is rejected, the empirical Bayes procedure was implemented under model M8 to identify sites under positive selection (posterior probability ≥ 0.90). To identify sites that have experienced purifying selection (posterior probability ≥ 0.90), the Fast Unconstrained Bayesian AppRoximation (FUBAR) (Murrell et al., 2013) was used as implemented in the HyPhy package. Consensus topology predictions were made using a standalone version of TOPCONS2.0 (Tsirigos et al., 2015). All computational analyses were run on the Odyssey cluster supported by the FAS Division of Science, Research Computing Group at Harvard University.

Plasmids:

pCI-MOR9-1 was a gift from Hiroaki Matsunami (Addgene plasmid # 22331) (Saito et al., 2009). pGP-CMV-GCaMP6s was a gift from Douglas Kim (Addgene plasmid # 40753) (Chen et al., 2013). The GNAI15 expression plasmid was kindly provided by Steve Liberles. DNA sequences encoding mCherry-MS4A were cloned into the tetracycline inducible mammalian expression plasmid, pcDNA5-FRT-TO using standard molecular biology methods.

Heterologous expression of MS4A proteins:

Flp-In-T-Rex HEK293 cells (Invitrogen R78007) were maintained in complete media (DMEM, high glucose, no glutamine (Life Technologies) supplemented with 10% tetracycline free Fetal Bovine Serum (Clontech), penicillin, streptomycin, and glutamine (Life Technologies)) with 5% CO2 in a 37 °C humidified tissue culture incubator (NuAire) on 10 cm tissue culture plates (BD #430167). Approximately four hours prior to transfection, cells were washed once with 10 mL plain DMEM and then incubated with 2 mL of Trypsin-EDTA solution (ATCC) for approximately 3 minutes at 37 °C. 10 mL of complete media was then added to the plate and the cells were triturated 5-10 times vigorously to generate single cell suspensions. After centrifugation at 300 x g for 5 minutes, the supernatant was aspirated and the cells were resuspended in complete media before they were plated on Round German Glass 15 mm coverslips (Bellco Biotechnology) in 12 well plates, which had been incubated for at least 24 hours with 0.02 mg/mL poly-d-lysine hydrobomide (Sigma) before being washed twice with ddH_20. After four hours, cells were transfected with calcium phosphate. For each coverslip, a 50 µL reaction mix consisting of 250 mM $CaCl_2$, approximately 2.5 µg GCaMP6s encoding plasmid, and 1 µg MS4A encoding plasmid. This mix was homogenized by pipetting 4 times. To this reaction mix, 50 µL of 2X HeBs (274 mM NaCl, 10 mM KCl, 1.4 mM $Na_2PO_4.7H_20$, 15 mM D-glucose, 42 mM Hepes (free acid)) pH 7.04-7.10 was added in a swirling motion from the bottom of the tube and bubbled briefly with air. This mixture was incubated for five minutes at room temperature, pipetted once to mix, and added to the well in a drop-wise manner. The 12 well plate was then shaken 5 times along each major axis before being placed in the incubator. Approximately one hour later tetracycline (Sigma #T7760) was added to a final concentration of 1 µg /mL and cells were allowed to express MS4A proteins overnight.

MS4A6C surface expression:

HEK293 cells were transfected with a plasmid encoding GCaMP6s and either a control plasmid or a plasmid encoding a mCherry-MS4A6C fusion protein as described above. 16 hours following transfection, 1 µg of purified anti-MS4A6C antibody was added to the cells. The cells were incubated with antibody for 1 hour at room temperature in the dark, and then washed 3X with complete media. Cells were subsequently fixed with 4% paraformaldehyde/1X PBS for ten minutes and then washed three times with 1X PBS. Cells were re-blocked with PBS/0.3% Triton X-100/5% donkey serum for 30 minutes and then incubated with Alexa633 goat anti-rabbit secondary antibody (1:300) for 45 minutes. The fixed and stained

cells were washed three times with block, and the coverslips were mounted on slides using VECTASHIELD Mounting Medicum with DAPI (Vector laboratories).

Odors:

Steroids were purchased from Steraloids (Newport, Rhode Island). Additional compounds were purchased from Sigma and were obtained at the highest purity possible.

In vitro *odor screen:*

HEK293 cells were co-transfected as described above with a plasmid encoding GCaMP6s and either a plasmid encoding one of the MS4A proteins or plasmids encoding the odorant receptor MOR9-1 and the G protein GNAI15, (which couples the exogenous GPCR to intracellular calcium stores, see Ukhanov et al., 2014). Coverslips were washed with Ringers solution and then secured in a perfusion/imaging chamber (Warner Instruments) using High Vacuum Grease, Dow Corning (VWR). Ringers solution was constantly perfused over the cells at a rate of ~10 mL/minute using an 8-channel valve-controlled gravity-driven perfusion system (Warner Instruments) and images were acquired using the MetaFluor software package using an Olympus IX83 microscope, a Sutter Lambda DG4 Light System with a Xenon arc lamp, and an Andor Neo 5.5 sCMOS camera. A 20X NA 0.45 air objective (Olympus) was used to acquire images at ~ 0.67 Hz at 3 x 3 pixel binning to reduce the exposure times required to obtain images thereby limiting photobleaching and phototoxicity. Each imaging epoch consisted of 900 seconds. For the first 300 seconds images were acquired but not saved as a significant amount of GCaMP6s signal decay occurred within this window. Subsequently the experiment consisted of 150 seconds of Ringers, valve switch and 75 seconds of Ringers (from a separate reservoir to mimic odor stimulation), valve switch and 150 seconds of Ringers, valve switch and 75 seconds of odorant, followed by a final valve switch and 150 seconds of Ringers. Odorants were delivered as mixtures of chemicals diluted in Ringers solution such that each individual constituent was present at 10 μM final concentration. Each coverslip was exposed to a single imaging epoch to control for response adaptation. To assess the time required to clear dead volumes and to estimate mixing delays, 100 nM RhodamineB dye was flowed instead of odor and dye saturation kinetics were determined. In figures 3A and S4B stimulus bars begin when odor is estimated to reach 90% saturation.

To identify monomolecular compounds that activate HEK293 cells expressing individual MS4A family members, several odorant mixes that elicited statistically significant responses in cells expressing a particular MS4A protein were selected for deconvolution analysis. The final concentration of each odorant in the deconvolution analysis was 50 μM except for the polyunsaturated fatty acid constituents, which were delivered at 10 μM, as at higher concentrations some of these molecules non-specifically activated HEK293 cells.

Analysis of in vitro *screen data:*

Analysis was performed using fully automated custom scripts in iPython that employed NumPy, SciPy, Pymorph, Pandas, and Mahotas source code packages. Each fluorescent image of the time series was locally smoothed and downsampled 3-fold in both dimensions to facilitate subsequent processing. The downsampled image stack was segmented into cell areas using a watershed transform, which were then used to extract the fluorescence values of individual cell areas over time. The resulting traces were baseline-corrected using a wide (100 frames) moving median filter to remove long-timescale drifts such as fluorescence decay due to photobleaching and smoothed using a local (9 frames) moving average filter. A total of seven coverslips from the mixture screen and five coverslips from the deconvolution screen were removed from the analysis for failing to meet quality control.

Processed traces were normalized by calculating an average fluorescence value across the period prior to odor stimulation and then dividing the entire trace by this value (thereby generating a dF/F trace). Responses were identified as fluorescent peaks within a twenty frame window centered at the time point at which maximal odor concentration occurred (determined empirically using RhodamineB dye) that were five standard deviations above the noise observed in the baseline period. The spontaneous activity response rate (determined as responses observed during twenty frames of the Ringers only period) was subtracted from the odor-evoked response rate; this rate did not change based upon the specific position of the window chosen in the trace (as long as it did not overlap with odor delivery). The proportion of cells co-expressing a chemoreceptor and GCaMP6s that responded to a given stimulus was compared by Z-test to the

proportion of cells expressing only GCaMP6s that responded to the same stimulus. Only response tallies with a P-value of < 0.05 following FDR correction for multiple comparisons are depicted.

In vitro *functional experiments with focal stimulus delivery pencil:*

To characterize ligand-induced response kinetics with greater temporal precision than in the bulk calcium assay, HEK293 cells were transfected with a plasmid encoding the fast version of GCaMP6 (GCaMP6f) and plasmids encoding either an MS4A or MOR9-1/GNAI15. The cells were perfused with Ringer's solution and odorants were delivered in liquid phase with a six channel gravity-controlled perfusion manifold (Warner Instruments) through a custom-made "stimulus pencil" made of quartz tubing focally positioned relative to the field of view. Images were acquired as described for the *in vitro* odor screen with an imaging rate of 2 Hz. The delay between valve switch and odor delivery was determining using RhodamineB dye as in the bulk calcium assay.

Dose response curves and EC50 calculations:

Dose response curves were determined as in (Mainland 2015). HEK293 cells co-expressing GCaMP6f and MOR9-1/GNAI15 or an MS4A were odor stimulated with a focal stimulus delivery pencil as described earlier. Each field of view was stimulated using the stimulus pencil with odorant spanning six log orders of odorant (from 10 nM to 1 mM) starting with the lowest concentration. Fluorescent traces were extracted and normalized as described above for the *in vitro* odor screen. Because of the tighter stimulus control afforded by this configuration, cells were considered to have responded if they exhibited fluorescent peaks greater than 4 standard deviations above baseline within a 15 second window centered around the time of maximal odor responses. A delay of 150 seconds was included between stimulus presentations. The cumulative fraction of cells that had responded by each odor concentration was determined for both receptor-expressing and control cells. To account for non-specific activation that might occur at higher concentrations of odorant the fraction of control cells (expressing GCaMP only) that was activated was subtracted from the fraction of receptor-expressing cells and sigmoidal dose response curves were then fit to the receptor data. Between four and twenty coverslips for both control and receptor-expressing cells were imaged to construct each dose response curve and each data point represents the mean +/- SEM of the relative cumulative fraction of cells that responded to that odor concentration.

Analysis of calcium transient temporal dynamics:

Analysis was performed by applying fully automated custom iPython scripts to extracted fluorescence traces. Response onset, half-rise time and peak time were identified using a peak and trough finding package in SciPy (SciPy.signalfind_peaks_cwt) (Du et al., 2006). Only cells that had response peaks greater than 4 standard deviations above baseline were included for analysis. Two-tailed Student's T-tests were run on the distributions of data and were Bonferroni corrected for multiple comparisons.

Requirement of extracellular calcium for MS4A ligand responses:

Experiments were performed similarly to those described using the stimulus pencil delivery system. Cells were either perfused with Ringers supplemented with 1 mM calcium chloride (plus calcium) or 1 mM EGTA (minus calcium) to chelate calcium. The percent of cells that responded was determined as described above.

RNAscope Fluorescent In Situ *Hybridization:*

Fluorescent *in situ* hybridization was performed on dissociated olfactory epithelial cells adhered to glass coverslips. Nasal epithelia from 8-12 week old C57/BL6 male mice, OR174-9-IRES-tauGFP mice, or GC-D-IRES-tauGFP mice were dissociated with the same method used for FACS, except that they were incubated in papain + DNAse I for 90 minutes. The final cell pellet from a single mouse epithelium was resuspended in 900 µL of DMEM + 10% FBS, and 75 µL was placed on each of twelve 12 mm #1 glass coverslips (Bellco Biotechnology #1943-10012) that had been coated with 30 µL of 5 mg/mL poly-D-lysine (Sigma Aldrich #P6407), dried, and washed three times with deionized water. Cells were allowed to settle on coverslips for 30 minutes at room temperature, rinsed once with 1X PBS, and fixed with 4% paraformaldehyde (PFA)/1X PBS (diluted from 20% PFA, Electron Microscopy Sciences #15713) for 30 minutes at room temperature. Coverslips were washed three times with 1X PBS and passed through a dehydration series of 50%, 70%, and 100% ethanol (v/v in water) for 5 minutes each. The last wash was replaced with fresh 100% ethanol and cells were stored at -20 °C overnight.

On the second day, coverslips were stained for specific mRNA targets by the RNAscope protocol (Advanced Cell Diagnostics, RNAscope Multiplex Fluorescent Assay For Fresh Frozen Tissues User Manual rev. 20121003). Several modifications to the RNAscope protocol were made to combine FISH for low abundance targets with immunohistochemistry: after rehydration, coverslips were treated with protease ("Pretreat 4") diluted 1:30 in 1X PBS; RNAscope target probes, which are supplied at 50X concentration, were diluted to final 2X concentration in blank probe solution; hybridization lasted for 4 hrs at 40 °C in the HybEZ oven (Advanced Cell Diagnostics #310010); and for experiments using type C1 and C2 probes (see below) the fluorescent reagent Amp4-altB was used, whereas for experiments using C1 and C3 probes we used Amp4-altC (Advanced Cell Diagnostics #320850).

To combine FISH with immunostaining, at the end of the RNAscope protocol, instead of being counterstained and mounted, the coverslips were rinsed in 1X PBS twice and blocked in 1X DAKO Serum-Free Protein Block (DAKO X0909) in PBS for 20 minutes at room temperature. Coverslips were then incubated with primary antibody (rabbit anti-Car2 or chicken anti-GFP as below) in blocking solution for 1 hour at room temperature, washed three times for 5 minutes in PBS, and stained with secondary antibody in PBS for 30 minutes at room temperature. After three final washes in PBS the coverslips were counterstained with ProLong Diamond Antifade Mountant + DAPI (Life Technologies #P36961), fixed to slides with nail polish, and imaged by epifluorescent or confocal microscopy as described below.

To quantify RNAscope signal, coverslips were viewed under epifluorescence with a 63X/NA1.4 oil immersion objective lens (Zeiss Plan-Apochromat 63X/1.40 OIL DIC 440762-9904-000) and fluorescent puncta visible by eye were counted for each immunopositive cell. In initial experiments puncta in random immunonegative cells were also quantified and the background levels for MS4A probes were found to be ~ 1 punctum/5-10 cells (data not shown). In approximately 10% of experiments much higher background (3+ puncta/cell) was observed, possibly due to sample drying. These coverslips were discarded from the analysis. For the remainder of the analysis puncta were quantified only in identifiable cells (i.e., GC-D/CAR2 or OR174-9 expressing cells.)

The target probes used in this assay are custom reagents designed by Advanced Cell Diagnostics and are currently available in the ACD catalog. Mouse target probes used are as follows: Ms4a1-c1 #318671, Ms4a2-c1 #438171, Ms4a3-c1 #438181, Ms4a4A-c1 #427391, Ms4a4B-c1 #314611, Ms4a4C-c1 #426371, Ms4a4D-c1 #427401, Ms4a5-c1 #318641, Ms4a6B-c1 #313801, Ms4a6C-c1 #314581, Ms4a6D-c1 #314591, Ms4a7-c1 #314601, Ms4a8A-c1 #426361, Ms4a10-c1 #438151, Ms4a15-c1 #427381, Olfr173-c1 #313771, Olfr151-c1 #431161, Olfr66-c1 #431171, Ms4a6C-c3 314581-C3, Ms4a4B-c3 #314611-C3, Car2-c2 #313781-C2, Gucy2d-c2 #425451-C2, Pde2a-c3 #426381-C3.

MS4A Antisera:

Peptides derived from mouse MS4A protein sequences were synthesized by Covance, Inc. (Denver, PA) as follows: 6C-pep: CKQSKELSLIEHDYYQ; 6D-pep: SQNSKNKSSVSSESLC; 7-pep: HKREKTGHTYEKEDD; 4B-pep: HQGTNVPGNVYKNHPC; 8A-pep: TAKSWEPEQERLTWC; 5-pep: TTQEYQTTELTATAYNC. These peptides were coupled by terminal cysteine residues to KLH and used to raise an immunoglobulin response in rabbits and guinea pig (6C-pep only). Sera collected by Covance, Inc. from rabbits and guinea pig were tested for ability to stain 293T cells expressing MS4A protein and GC-D cells (guinea pig anti-6C and rabbit anti-6C antibodies yielded the same pattern of staining *in vitro* and *in vivo*); the best lots of serum were purified by passing over protein A/sepharose (Life Technologies #10-1042) and eluted with 100 mM Glycine, pH 2.7 to collect immunoglobulin. A subset of these (anti-6C, 6D, and 7) were further affinity purified by passing over resin cysteine-coupled to the respective antigenic peptides (Thermo Scientific #44999) and fractions eluted with 100 mM glycine, pH 2.7 were tested for greatest specificity on 293T cells. We also raised antisera against GC-D protein with the peptide (Ac-QRIRTDGKGRRLAC), which was purified by passing over protein A/sepharose and a peptide affinity column as above. For peptide competition experiments, antibody concentration was estimated by reacting with protein assay dye reagent (Bio-Rad #5000006) and measuring 595 nm absorbance with a Cary 60 UV-Vis spectrophotometer (Agilent Technologies). Sample concentrations were estimated by fitting to a line generated with IgG standards; most purified antibodies were between 1.0 and 5.0 mg/mL. Peptides were resuspended in 1X TBS at 1 mg/mL and mixed with diluted antibodies in block at a 10 mg peptide: 1 mg antibody ratio; as the peptides are ~15 amino acid residues and IgGs are ~1500 residues total, it was estimated that this gives ~1000-fold molar excess of peptide. Similar results were obtained with lower peptide concentration (data not shown).

Primary antibodies/concentrations used were as follows: goat anti-Pde2a (1:50, Santa Cruz Biotechnology sc-17227), chicken anti-GFP (1:1000, Abcam #ab13970), rabbit anti-AC3 (1:100, Santa Cruz Biotechnology sc-588), rabbit anti-GC-D (serum 7444, affinity fraction 2, 1:1000), rabbit anti-Car2 (1:500, Abcam #191343), goat anti-CD20 (1:50, Santa Cruz Biotechnology #sc-7735), rabbit anti-MS4A4B (serum 7512, PAS fraction, 1:2000), rabbit anti-MS4A6C (serum 7277, affinity fraction 3, 1:500), anti-MS4A6D (serum 7284, affinity fraction 3, 1:500), anti-MS4A7 (serum 7286, affinity fraction 2, 1:500), anti-MS4A8A (serum 7508, PAS fraction, 1:2000), anti-MS4A5 (serum 7503, PAS fraction, 1:2000), guinea pig anti-MS4A6C (serum 7630, PAS fraction, 1:1000), anti-S6 phosphoSerine240/244 (1:200, Cell Signaling Technologies #22155).

Secondary antibodies/concentrations used were as follows: donkey anti-rabbit-Cy3 (1:300, Jackson Immunoresearch, #711-167-003), donkey anti-rabbit-CF633 (1:300, Biotium #20125), donkey anti-goat-Alexa633 (1:300, Invitrogen #A21082), donkey anti-goat-Alexa488 (1:300, Jackson Immunoresearch #705-546-147), donkey anti-guinea pig-Alexa647 (1:300, Jackson Immunoresearch #706-606-148), donkey anti-chicken-Alexa488 (1:300, Jackson Immunoresearch #703-545-155).

Cultured Cell Preparation for Immunostaining:

293T cells adhered to 12 mm coverslips and transfected with mCherry-MS4A expression plasmids (see above) were fixed 24-48 hours post-transfection with 4% PFA/1X PBS (Electron Microscopy Sciences) for 10 minutes at room temperature. Coverslips were washed three times with 1X PBS and stored at 4 °C in 1X PBS until immunostaining as described below.

Tissue Preparation for Immunostaining:

All preparation of nasal epithelia and olfactory bulbs was performed as follows, except for assays involving phospho-S6: animals were euthanized as above and olfactory epithelia were dissected out from the skull with olfactory bulbs attached and fixed overnight in 4% PFA/1X PBS at RT. After washing 5 minutes with 1X PBS three times, noses were decalcified overnight in 0.45M EDTA/1X PBS at 4 °C. Noses were washed once more with 1X PBS, then sunk in 20% sucrose for 3 hrs at 4 °C, and finally embedded in Tissue Freezing Medium (VWR #15146-025). 15 micron cryosections were cut onto Superfrost Plus glass slides (VWR #48311703) and stored at -80 °C until staining. For experiments involving anti-phosphoS6 staining, nasal epithelia were instead fixed overnight in 4% PFA/1X PBS at 4 °C.

Immunostaining:

For experiments without anti-phosphoS6, cryosections on slides were removed from the freezer, air dried for 10 minutes, and antigen retrieved by immersion for 10 minutes in a 98 °C solution of 10 mM Sodium Citrate, 0.1% Tween-20, pH 6.5. Slides were rinsed in 1X Tris-buffered Saline (TBS: 50 mM Tris-Cl, 150 mM NaCl, pH 7.5) at room temperature and blocked in 5% Normal Donkey Serum (Jackson Immunoresearch #017-000-121)/0.1% Triton-X100/1XTBS for 30 minutes at room temperature. All immunostaining was done at 4 °C overnight with antisera diluted in blocking solution.

On the following day, slides were washed three times for 10 minutes in 1X TBS/0.1% TritonX100 at room temperature, and then incubated in secondary antibody solution (antibodies listed above in blocking solution) for 45 minutes at room temperature. Finally, slides were washed three times for 10 minutes in 1X TBS/0.1% TritonX100, counterstained with Vectashield + DAPI (VWR #101098-044), and coverslipped before confocal microscopic imaging.

For experiments involving anti-phosphoS6, staining was performed as above with the following changes: there was no antigen retrieval step before blocking; blocking solution was 1X PBS/0.1% TritonX100/3% Normal Donkey Serum/3% BSA; all washes were done in 1X PBS/0.1% TritonX100; and the primary antibody solution included 50 nM of a phosphopeptide ("3P") derived from the S6 sequence (Knight et al., 2012). This peptide, which contains phosphorylated Serine235/236/240, is meant to compete away nonspecific antibody binding to S6 protein not phosphorylated at its neuronal activity-dependent sites.

Finally, assaying *Ms4a6d-IRES-GFP*-infected noses required costaining with two rabbit primary antisera, one to MS4A6D and the other to phosphorylated S6. To prevent cross-reactivity with secondary antibodies, phospho-S6 staining was performed first as described above; after application of fluorescent secondary antibody, remaining sites on the rabbit primary antibody were blocked by incubation with 5 µg /mL unconjugated donkey anti-rabbit Fab fragments (Jackson Immunoresearch) in washing solution at

room temperature for 1 hour. After washing out unbound Fab fragments three times with wash solution, tissue sections were stained with anti-MS4A6D overnight as described above and detected with a CF633-conjugated secondary antibody (Biotium). No fluorescence to indicate cross-reactivity was observed, and control experiments without anti-MS4A6D antiserum showed that the later secondary antibody did not interact with the Fab-blocked phospho-S6 primary antibody.

Explant calcium imaging:

Adult offspring of *Emx1-IRES-Cre* and *ROSA-GCaMP3* mice were sacrificed and dissected to expose the lateral olfactory epithelium, adjacent to the olfactory bulb. The epithelium and adjacent skull (including the adjacent olfactory bulb) was embedded in a custom-made perfusion chamber with 5% low-melt agarose (Invitrogen) made from modified Ringer's solution (115 mM NaCl, 5 mM KCl, 2 mM $CaCl_2$, 2 mM $MgCl_2$, 25 mM $NaHCO_3$, 10 mM HEPES, pH 7.4). Small slits were cut anterior to the cul-de-sac regions of exposed olfactory epithelium to allow fluid flow, and for the remainder of the experiment the tissue was superfused with carboxygenated modified Ringer's solution (95% O_2, 5% CO_2). Odorants were delivered to the tissue in liquid phase *via* an 8-to-1 perfusion manifold controlled by solenoid valves (Warner Instruments) through a custom-made "stimulus pencil" positioned near the olfactory epithelium. Each odorant or mixture was diluted in 0.1% DMSO in modified Ringer's solution to a concentration of 100 μM.

Cul-de-sac regions of the tissue containing GCaMP3+ cells were imaged through the epithelial cartilage by standard multiphoton microscopy (Prairie Technologies) with a Ti-sapphire laser (Coherent) at 965 nm through a 25X/NA1.05 water-immersion objective lens (Olympus XL Plan N). 512x512 fluorescent images were acquired at 1 Hz, with each odor trial consisting of 20 seconds of Ringers, 10 seconds of odorant, and finally 50 more seconds of Ringers. After imaging, a high resolution image of the field of view was acquired to aid cell identification. As was done for the *in vitro* assays, the delay time for odor delivery was assessed using RhodamineB dye; these experiments revealed that the maximal lumenal concentration of dye at equilibrium was approximately 10 percent of the initial dye concentration (as assessed via fluorescence intensity). Image registration was accomplished using custom code, and publicly available Python image processing and OpenCV packages. Time-series images were aligned using a feature-based approach that is robust to regional fluorescence intensity fluctuations over the course of the experiment. Images (for alignment purposes) were first contrast enhanced to enable feature detection, and then all frames from a single experiment were registered to a manually chosen target frame in a pairwise manner. For each frame-target pair, positional features, typically corresponding to cell bodies and blood vessels, were then automatically identified using Harris corner detection. Corresponding features were then used to obtain an optimal homography (i.e., projective transformation) between frame and target, and this transform was then applied to the raw images.

Cells from aligned movies of the raw images were identified using a semi-automated approach. First, the centroids of putative cells (i.e., rounded and convex GCaMP3+ objects, sometimes with nuclear exclusion of GCaMP3) in a time-series average projection were manually specified. Cell masks were then generated using a combination of morphological filtering and region-growing. Each cell mask was further refined based on co-fluctuations in the fluorescence of pixels in the cell's vicinity. Independently covarying groups of pixels were first identified using nonnegative matrix factorization. Pixels associated with the cell of interest were then assigned to the current mask; pixels that correspond to adjacent cells were excluded. In the resultant binary mask, each connected component therefore corresponded to a cell; fluorescence time-courses for each cell were then obtained by averaging the pixels in each connected component on a frame-by-frame basis.

Putative necklace cells were defined as cells within the cul-de-sacs that responded to carbon disulfide with at least a 25% increase in average fluorescence and did not respond to DMSO alone. A cell was then categorized as responding to an odor stimulus if its average fluorescence increased at least 25% relative to the previous 10 frames and reached a peak after odor presentation, but not before. For display, plotted traces were smoothed by convolution with a 3-frame rectangular window. For generating the top panels of Figure 6A, the fluorescent signal following vehicle stimulation was subtracted from the fluorescent signal following odorant stimulation, and the resultant image was heat-mapped and overlaid onto a reference image using a custom Python script. For the purposes of data representation these heatmaps are shown as raw change in fluorescence; 100% on this fluorescence scale corresponds to the full dynamic range of image intensity in this acquisition.

Confocal Microscopy and Image Processing:

Slides were imaged under a 63X/NA1.4 oil immersion planar apochromatic objective lens (Zeiss). Digital images were acquired on a Zeiss LSM 510 Meta confocal microscope (Harvard Neurodiscovery Imaging Center). Cyanine and Alexa fluorophores were excited in sequence with an argon laser (488 nm line), a HeNe laser (543 nm), and a second HeNe laser (633 nm). DAPI was imaged with a tuneable Coherent Chameleon Laser at 740 nm in two-photon excitation mode. Emission was detected with standard dichroic mirrors and filter sets. Using Imaris 8.1 software (Bitplane Inc.), multi-channel z-stacks were maximum-intensity-projected into two dimensions and passed through a median filter to remove debris much smaller than structures being assayed.

Adenoviral Infection:

Six week old male mice (Jackson) were anesthetized with 100 mg/kg of Ketamine, 10 mg/kg of Xylazine, and 3 mg/kg of Acepromazine. Infusions of 3~6 x 10^8 FFU of human adenovirus serotype 5 encoding Ms4a6c-IRES-hrGFP or Ms4a6d-IRES-hrGFP in the right nostril of mice were performed as previously described (Holtmaat et al., 1996). Mice were subjected to the odor exposure 6-8 days after injections.

Odor Exposure for Phospho-S6 Immunostaining and Quantification of Positive Cells:

8-10 week old C57/BL6 male mice (Jackson Laboratory) were group housed overnight on a reverse light-dark cycle with 3-5 mice/cage. On the day of the experiment, single mice were habituated in fresh cages for 3 hours in the dark. The odor stimulus was then introduced into each cage as 100 μL of neat odorant blotted onto Whatman paper in 10 cm petri dishes with slits cut into the lid. After 4 hours of odor exposure, animals were sacrificed and their nasal epithelia were dissected and fixed as described above.

To quantify phospho-S6 immunopositive cells, 10-15 images of Pde2a-positive cell enriched cul-de-sacs were acquired with a confocal microscope (see above) for each slide of odor-exposed olfactory epithelial sections. Laser and acquisition parameters were held constant across each experiment. pS6 immunopositive cells were counted manually from the resulting images; a cell was called positive if it showed smooth pS6 fluorescent signal filling up the soma at levels visibly above adjacent tissue and Pde2a-negative cells. Counts from each image were summed to give an estimate of the proportion of activated necklace cells for each animal. In initial experiments with each odorant, the person imaging and counting cells was blind to the identity of the odorant. Because only sulfated steroids have a vehicle control (DMSO alone) these data were analyzed separately from the data in Figure7A.

The same method was used to quantify MS4A6C-ires-GFP and MS4A6D-ires-GFP virus-infected cells, except that GFP+ cells were sparse and therefore called positive for pS6 and MS4A6C (guinea pig antibody) by finding a GFP+ cell, imaging it under laser excitation, and recording the cell as positive or negative manually. 20 or more GFP+ cells were counted for 3-5 epithelia per odorant, and images of representative cells were acquired and processed as described above.

Table S1, List of taxa used for the phylogenetic reconstructions, Related to Figure 2 and Experimental Procedures

Group	Order	Family	Species	Common name
Monotremata		Ornithorhynchidae	*Ornithorhynchus anatinus*	platypus
Marsupalia		Dasyuridae	*Sarcophilus harrisii*	Tasmanian devil
Placentalia	Soricomorpha	Soricidae	*Sorex araneus*	common shrew
	Carnivora	Mustelidae	*Mustela putorius furo*	ferret
		Ursidae	*Ailuropoda melanoleuca*	giant panda
		Canidae	*Canis lupus familiaris*	dog
	Chiroptera	Vespertilionidae	*Myotis lucifugus*	little brown bat
	Perissodactyla	Equidae	*Equus ferus caballus*	horse
	Artiodactyla	Suidae	*Sus scrofa domesticus*	pig
		Bovidae	*Bos taurus*	cow
			Ovis aries	sheep
	Cetacea	Delphinidae	*Tursiops truncatus*	bottlenose dolphin
		Lipotidae	*Lipotes vexillifer*	baiji
	Macroscelidea	Macroscelididae	*Elephantulus edwardii*	elephant shrew
	Afrosoricida	Tenrecidae	*Echinops telfairi*	lesser hedgehog tenrec
	Hyracoidea	Procaviidae	*Procavia capensis*	rock hyrax
	Cingulata	Dasypodidae	*Dasypus novemcinctus*	nine-banded armadillo
	Primates	Galagidae	*Otolemur garnetti*	northern greater galago
		Callitrichidae	*Callithrix jacchus*	common marmoset
		Hominidae	*Gorilla gorilla*	gorilla
			Pan troglodytes	chimpanzee
			Homo sapiens	human
	Scandentia	Tupaiidae	*Tupaia chinensis*	tree shrew
	Lagomorpha	Ochotonidae	*Ochotona princeps*	American pika
		Leporidae	*Oryctolagus cuniculus*	rabbit
	Rodentia	Heteromyidae	*Dipodomys ordii*	kangaroo rat
		Dipodidae	*Jaculus jaculus*	jerboa
		Muridae	*Mus musculus*	mouse
			Rattus norvegicus	rat
		Cricetidae	*Cricetulus griseus*	Chinese hamster
			Mesocricetus auratus	golden hamster
			Microtus ochrogaster	prairie vole
		Sciuridae	*Spermophilus tridecemlineatus*	thirteen-lined ground squirrel
		Bathyergidae	*Heterocephalus glaber*	naked mole-rat
		Chinchillidae	*Chinchilla lanigera*	chinchilla
		Octodontidae	*Octodon degus*	degu
		Caviidae	*Cavia porcellus*	guinea pig

Table S2, List of probe sequences used in Nanostring experiments, Related to Figure 1 and Experimental Procedures

Ms4a1	GCAACCTGCTCCAAAAGTGAACCTCAAAAGGACATCTTCACTGGTGGGCCCCACA CAAAGCTTCTTCATGAGGGAATCAAAGGCTTTGGGGGCTGTCCAA
Ms4a2	ACAGAAAATAGGAGCAGAGCAGATCTTGCTCTCCCAAATCCACAAGAATCCTCCA GTGCACCTGACATTGAACTCTTGGAAGCATCTCCTGCCAAAGCAG
Ms4a3	CCAGGCTTTCAAGGGTTGCCAATCTTCACCGTCACCTGATGTCTGCATTTCCCTGG GTTCCTCATCAGATGGCCTGGTGTCTTTAATGCTGATTCTCACC
Ms4a4a	AACCCAAAATCCTTGGGATTGTGCAGATTGTAATCGCCATCATGAACCTCAGCAT AGGAATTATGATGATAATTGCCACTGTGTCGACCGGTGAAATACC
Ms4a4b	CCTAGGATATTAACACTTCATTGCACTGGCTTTTGAGGTGAATATTAGATTTACTG TAAGTATGTAAGTCAAGCACTTATTAGGTCAACAACACTTCAAC
Ms4a4c	TGGCAAATCTATCTTCTGAACCACTCATTTCTGTGGTCTTAATGGCTCCAATTTGG GGACCAATAATGTTCATTGTCTCAGGATCCCTGTCAATTGCAGC
Ms4a4d	ACAACTGGCACTACCATCGTGGTGAAAACCCAGCTCAAGCATACCCACAAATAGA GTCCCACATCGAAACTCCACCACATTACTCAAGGATACTGTTTCT
Ms4a5	TGAATTTACTTAGTGCTCTGGGAGCAGCAGCTGGAATCATTCTCCTCATATTTGGC TTCCTTCTAGATGGGGAATTCATCTGTGGCTATTCTCCAGATGG
Ms4a6b	AAACAAAACTAAATACCACAAAAACAAATGGAACTATACCGCAGAAGATATGTC TTCATGATAATGCAGAAATTCCAACCATCACAGGGTAGCAATGCTT
Ms4a6c	CATGATTCCACAGGTAGTGACCAATGAGACCATCACAACGATTTCACCAAATGGA ATCAACTTTCCCCAAAAAGACGAGTCCCAGCCTACCCAACAGAGG
Ms4a6d	AGTTTGGCTGCTTTAGAGCCTGCCTTGCAGCAATGTAAGCTGGCTTTCACACAACT AGACACAACCCAAGATGCTTATCATTTCTTTAGCCCTGAGCCAT
Ms4a7	GCCTCCAATGTAGCAAGCTCTGTTGTTGCCGTCATTGGCCTCTTCCTCTTCACCTAT TGTCTGATAGCCCTGGGGAGTGCTTTCCCACACTGTAACTCAG
Ms4a8a	TGTCACTACAACAATCCAGGTGTGGTCATTCCAAATGTCTATGCAGCAAACCCAG TGGTCATCCCAGAACCACCAAACCCAATACCAAGTTATTCCGAAG
Ms4a10	CCTAAGACCTCTCTGAAGGTTCTCTGTGTGATAGCCAACGTTATCAGCTTGTTCTG CGCACTGGCCGGCTTCTTTGTCATTGCCAAGGACCTCTTCCTGG
Ms4a13	TTTCATGGCTGCTAACACCTGATGTAGGTGCCCATGAGATTCCCATATAACAAGGC ACACCTCATGCATTTTGTGCAAAAGGAAATTCACAACAAGGTGA
Ms4a15	GTGGGAAATCTTGGCTTCGCAGAGGTTTCGGAGGTTTGTCTTCAAGATCATTAAGC ACGGAGAACTCAGAATGTTCCAGAATAGACTGGCATTTCAGAGG
Ms4a18	GAATTCATCCTCACCTGCATAGCCTCACATTTTGGATGCCAGGCTGTCTGCTGCGC CCATTTTCAGAACATGACAATGTTCCCAACCATATTTGGTGGCA
Pde1c	GCTGTAATCGATGCATTGAAGGATGTGGATACGTGGTCCTTCGATGTCTTTTCCCT CAATGAGGCCAGTGGAGATCATGCACTGAAGTTCATTTTCTATG
Adcy3	CAACAACGGCGGCATCGAGTGTCTACGCTTCCTCAATGAGATCATCTCTGATTTTG ACTCTCTCCTGGACAATCCCAAATTCCGGGTCATCACCAAGATC
Cnga2	GCTTGTGGATAATGGAGATCATGTGGGTTGAATTTCTAAGAGCGTGACCTCCTAA GTCTCACAAGGAATCAGAGAATAGCTAAATTGTCCTTCCTGAGGC
Actb	CAGGTCATCACTATTGGCAACGAGCGGTTCCGATGCCCTGAGGCTCTTTTCCAGCC TTCCTTCTTGGGTATGGAATCCTGTGGCATCCATGAAACTACAT
Gapdh	AGGTTGTCTCCTGCGACTTCAACAGCAACTCCCACTCTTCCACCTTCGATGCCGGG GCTGGCATTGCTCTCAATGACAACTTTGTCAAGCTCATTTCCTG
Pde2a	CCACTAGCTTCTCTTCTGTTTTGTTCCCTATGTGTCGTGGGTGGGGGAGGGGGCCA CCTGCCTTACCTACTCTGAGTTGCCTTTAGAGAGATGCATTTTT
Car2	TGCCCAGCATGACCCTGCCCTACAGCCTCTGCTCATATCTTATGATAAAGCTGCGT CCAAGAGCATTGTCAACAACGGCCACTCCTTTAACGTTGAGTTT
Golf	ATCGAAGACTATTTCCCGGAGTATGCCAATTATACTGTCCCTGAAGATGCAACAC CAGATGCGGGAGAAGATCCCAAAGTTACAAGAGCAAAGTTCTTTA
Emx1	CAGGCAAGCGACGTTCCCCAGGACGGGCTGCTTTTGCACGGGCCCTTCGCACGCA AGCCCAAGCGGATTCGCACAGCCTTCTCGCCCTCGCAGCTGCTGC

Table S3. List of odorants used for functional imaging experiments, Related to Figure 3 and Experimental Procedures

Alcohols	1-butanol, 2,5-dimethylphenol, eugenol, guaicol, 1-hexanol, isoeugenol, 1-nonanol, 1-octanol, 2-phenylethanol, thymol
Ketones	acetylanilone, acetophenone, 2-butanone, cyclohexanone, 3-decanone, dodecanolactone, 4-heptanone, 2-octanone, 2-pentanone, vanillin
Sulfurs	2,4,5-trimethyl thiazole, TMT, thiophene, tetrahydrothiophene
Acids	formic acid, hexanoic acid, heptanoic acid, ocantoic acid, tiglic acid, valeric acid, isovaleric acid
Esters	allyl cinnamate, amyl acetate, benzyl acetate, cycohexyl aceatate, ethyl benzoate, ethyl propionate, ethyl valerate, ethyl tiglate, piperidine, propyl butyrate
Aldehydes	p-anise aldehyde, butyl formate, butyraldehyde, benzaldehyde, cinnamaldehyde, ethyl formate, heptanal, octanal, propionaldehyde, heptaldehdye
Nitrogenous	2,5-dimethylpyrazine, 2,6-dimethylpyrazine, 2,3-dimethylpyrazine, indole, nicotine, pyrrolidine, pyridine, quinoline
Steroids	4-Androsten-17alpha-ol-3-one sulphate, 5-Androsten-3Beta 17Beta-diol disulphate, 1,3,5(10)-Estratrien-3 17Beta-diol disulphate, 1,3,5(10)-Estratrien-3 17alpha-diol 3-sulphate, 5alpha-pregnen-3alpha-ol-20-one sulphate, 5beta-pregnen-3beta-ol-20-one sulphate, 4-pregnan-11beta 21-diol-3 20-dione 21-sulphate, 4-pregnen-21-ol-3 20-ione glucosiduronate, 1,3,5(10)-Estratrien-3 17Beta-diol 3-sulphate, 4-pregnen-11beta 17,21-triol3 20-dione 21-sulphate
PUFAs	arachidonic acid, docosohexanoic acid, linoleic acid, linolenic acid, nervonic acid, oleic acid, petroselenic acid
Saturated fatty acids	decanoic acid, docosanoic acid, dodecanoic acid, eicosanoic acid, hexanoic acid, myristic acid, octadecanoic acid, octanoic acid, palmitic acid
Terpenes	R-carvone, 1,4-cineole, citral, cintronellal, R-fenchone, E-beta farnesene, geraniol, alpha-ionone, linalool, +-menthone, gamma-terpinene, 1,3-minus-verbenone

Supplemental References

Anisimova, M., and Yang, Z. (2007). Multiple Hypothesis Testing to Detect Lineages under Positive Selection that Affects Only a Few Sites. Molecular Biology and Evolution *24*, 1219-1228.

Benjamini, Y., Drai, D., Elmer, G., Kafkafi, N., and Golani, I. (2001). Controlling the false discovery rate in behavior genetics research. Behav Brain Res *125*, 279-284.

Chen, T.W., Wardill, T.J., Sun, Y., Pulver, S.R., Renninger, S.L., Baohan, A., Schreiter, E.R., Kerr, R.A., Orger, M.B., Jayaraman, V.*, et al.* (2013). Ultrasensitive fluorescent proteins for imaging neuronal activity. Nature *499*, 295-300.

Darriba, D., Taboada, G.L., Doallo, R., and Posada, D. (2012). jModelTest 2: more models, new heuristics and parallel computing. Nat Meth *9*, 772-772.

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

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

Holtmaat, A.J., Hermens, W.T., Oestreicher, A.B., Gispen, W.H., Kaplitt, M.G., and Verhaagen, J. (1996). Efficient adenoviral vector-directed expression of a foreign gene to neurons and sustentacular cells in the mouse olfactory neuroepithelium. Brain Res Mol Brain Res *41*, 148-156.

Katoh, K., and Standley, D.M. (2013a). MAFFT multiple sequence alignment software version 7: improvements in performance and usability. Mol Biol Evol *30*, 772-780.

Katoh, K., and Standley, D.M. (2013b). MAFFT Multiple Sequence Alignment Software Version 7: Improvements in Performance and Usability. Molecular Biology and Evolution *30*, 772-780.

Knight, Z.A., Tan, K., Birsoy, K., Schmidt, S., Garrison, J.L., Wysocki, R.W., Emiliano, A., Ekstrand, M.I., and Friedman, J.M. (2012). Molecular profiling of activated neurons by phosphorylated ribosome capture. Cell *151*, 1126-1137.

Kosakovsky Pond, S.L., Murrell, B., Fourment, M., Frost, S.D.W., Delport, W., and Scheffler, K. (2011). A Random Effects Branch-Site Model for Detecting Episodic Diversifying Selection. Molecular Biology and Evolution *28*, 3033-3043.

Liao, Y., Smyth, G.K., and Shi, W. (2014). featureCounts: an efficient general purpose program for assigning sequence reads to genomic features. Bioinformatics *30*, 923-930.

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

Macmanes, M.D. (2014). On the optimal trimming of high-throughput mRNA sequence data. Front Genet *5*, 13.

Murrell, B., Moola, S., Mabona, A., Weighill, T., Sheward, D., Kosakovsky Pond, S.L., and Scheffler, K. (2013). FUBAR: A Fast, Unconstrained Bayesian AppRoximation for Inferring Selection. Molecular Biology and Evolution *30*, 1196-1205.

Picelli, S., Faridani, O.R., Bjorklund, A.K., Winberg, G., Sagasser, S., and Sandberg, R. (2014). Full-length RNA-seq from single cells using Smart-seq2. Nat Protoc *9*, 171-181.

Pond, S.L.K., Frost, S.D.W., and Muse, S.V. (2005). HyPhy: hypothesis testing using phylogenies. Bioinformatics *21*, 676-679.

Ronquist, F., Teslenko, M., van der Mark, P., Ayres, D.L., Darling, A., Höhna, S., Larget, B., Liu, L., Suchard, M.A., and Huelsenbeck, J.P. (2012). MrBayes 3.2: Efficient Bayesian Phylogenetic Inference and Model Choice across a Large Model Space. Systematic Biology.

Simossis, V.A., and Heringa, J. (2005). PRALINE: a multiple sequence alignment toolbox that integrates homology-extended and secondary structure information. Nucleic Acids Res *33*, W289-294.

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

Tsirigos, K.D., Peters, C., Shu, N., Käll, L., and Elofsson, A. (2015). The TOPCONS web server for consensus prediction of membrane protein topology and signal peptides. Nucleic Acids Research *43*, W401-W407.