

Supplementary Methods

Cloning of human odorant receptors

423 human odorant receptors were cloned with sequence information from The Olfactory Receptor Database (<http://senselab.med.yale.edu/senselab/ORDB/default.asp>). Of these, 335 were predicted to encode functional receptors, 45 were predicted to encode pseudogenes, 29 were putative variant pairs of the same genes, and 14 were duplicates. We adopted the nomenclature proposed by Doron Lancet¹. OR7D4 and the six intact odorant receptor genes in the OR7D4 gene cluster (OR1M1, OR7G2, OR7G1, OR7G3, OR7D2, and OR7E24) were used for functional analyses. SNPs in these odorant receptors were identified from the NCBI dbSNP database (<http://www.ncbi.nlm.nih.gov/projects/SNP>) or through genotyping. OR7D4 single nucleotide variants were generated by cloning the reference sequence from a subject or by inducing polymorphic SNPs by site-directed mutagenesis using overlap extension PCR. Single nucleotide and frameshift variants for the six intact odorant receptors in the same gene cluster as OR7D4 were generated by cloning the respective genes from the genomic DNA of each subject. The chimpanzee OR7D4 orthologue was amplified from chimpanzee genomic DNA (Coriell Cell Repositories).

Odorant receptors that contain the first 20 amino acids of human rhodopsin tag² in pCI (Promega) were expressed in the Hana3A cell line along with a short form of *mRTP1* called *RTP1S*, (M37 to the C-terminal end), which enhances functional expression of the odorant receptors³. For experiments with untagged odorant receptors, OR7D4 RT and S84N variants without the Rho tag were cloned into pCI.

Cell culture, immunocytochemistry, and flow cytometry

Cells were grown in a 37°C incubator containing 5% CO₂. For immunocytochemistry, cells were seeded in a 35mm dish (Falcon) containing a piece of cover glass coated with poly-D-lysine (Sigma) 24 hrs prior to transfection in Minimum Essential Medium containing 10% FBS (M10). Lipofectamine2000 (Invitrogen) was used for transfection of plasmid DNA. For permeabilized staining, blue fluorescent protein (BFP) was cotransfected as a control for transfection efficiency. 24 hrs post-transfection, cells were fixed in 4% paraformaldehyde for 15 min and permeabilized with methanol at 4°C. Cells were blocked in 5% skim milk diluted in PBS and incubated in 5% skim milk/PBS containing mouse monoclonal anti-rhodopsin antibody, 4D2, at room temperature for 45 min. Cells were then washed in PBS followed by incubation with Cy3-conjugated donkey anti-mouse IgG (Jackson Immunologicals) at room temperature for 30 min. For FACS analysis, Hana3A cells were seeded in 35mm dishes. At the time of transfection, green fluorescent protein (GFP) expression vector was co-transfected to monitor the transfection efficiency. About 24 hrs post-transfection, cells were incubated with PBS containing 4D2, 15mM NaN₃, and 2% FBS and then washed and incubated with phycoerythrin (PE)-conjugated donkey anti-mouse IgG (Jackson Immunologicals). 7-amino-actinomycin D (7-AAD; Calbiochem), a fluorescent, cell-impermeant DNA binding agent, was added before flow cytometry to eliminate dead cells from analysis⁴. 7-AAD selectively stains dead cells. The intensity of PE signal among the GFP-positive population was measured and plotted.

Western blot analysis

Hana3A cells in 35mm dishes were transfected with Rho-tagged receptor variants and RTP1S using Lipofectamine2000 (Invitrogen). GFP expression vector was cotransfected as a control. 24 hrs post-transfection, cells were lysed with sample loading buffer (20mM Tris [pH 7.5], 2mM EDTA, 5% SDS, 20% glycerol, 0.002% BPB, 0.25M DTT) and sonicated. SDS-PAGE and Western blot analysis were performed according to Mini-Protein 2 Cell (Bio-Rad) protocol. Membranes were incubated with 4D2 and subsequently with donkey anti-mouse HRP (Jackson Immunologicals). ECL (GE HealthCare Biosciences AB; Uppsala, Sweden) was used for detecting proteins on membranes. After exposure, the membrane was incubated with stripping buffer (25mM glycine-HCl [pH2], 1% SDS, 25mM glycine, 0.036N HCl, 1% SDS) for 30 min at room temperature and then with rabbit anti-GFP and subsequently with donkey anti-rabbit HRP.

Luciferase assay and data analysis

Dual-Glo™ Luciferase Assay System (Promega) was used to measure the firefly and *Renilla* luciferase activities as previously described⁵. CRE-luciferase (Stratagene) was used to measure receptor activation. *Renilla* luciferase driven by a constitutively active SV40 promoter (pRL-SV40; Promega) was used as an internal control for cell viability and transfection efficiency. Hana3A cells were plated on poly-D-lysine-coated 96-well plates (BioCoat; Becton Dickinson). Plasmid DNA of the receptor variants and RTP1S was transfected using Lipofectamine2000. For each 96-well plate, 1μg of CRE-Luc, 1μg of pRL-SV40, 5μg total of odorant receptors, and 1μg RTP1S were transfected.

In addition to these constructs, for untagged receptors, 0.25µg of mouse Ric8b and 0.25µg of mouse G_olf were transfected per plate⁶. About 24 hrs post-transfection, the medium was replaced with CD293 chemically defined medium (Gibco) and then the cells were incubated for 30 min at 37°C and 5% CO₂. The medium was then replaced with odorant solution diluted in CD293 and the cells were incubated for 4 hrs at 37°C and 5% CO₂. The odours used in this study and the concentrations used in luciferase assays can be found in Supplementary Tables S4 and S5. We followed manufacturer's protocols for measuring luciferase and *Renilla* luciferase activities. Luminescence was measured using Wallac Victor 1420 plate reader (Perkin-Elmer). Normalized luciferase activity was calculated by the formula $[\text{luc(N)} - \text{luc(lowest)}] / [\text{luc(highest)} - \text{luc(lowest)}]$, where luc(N) = luminescence of firefly luciferase divided by luminescence of *Renilla* luciferase of a certain well; luc(lowest) = lowest luminescence of firefly luciferase divided by luminescence of *Renilla* luciferase of a plate or a set of plates; and luc(highest) = highest luminescence of firefly luciferase divided by luminescence of *Renilla* luciferase of a plate or a set of plates. Data was analyzed with Microsoft Excel and GraphPad Prism 4.

Human odorant receptor genotyping and sequencing

Venous blood (8.5 ml) was collected from all subjects and genomic DNA was prepared with the Qiagen PAXgene blood DNA kit.

Polymorphisms in OR7D4 were assayed by both sequencing and allele-specific PCR. In allele-specific PCR, an OR7D4 RT and an OR7D4 WM PCR were performed on each genomic DNA sample, each using a pair of internal primers containing the residues of interest. The RT forward primer is specific for R88 (5'-CTAGTGAGCATCCAGGCAC-3') and the reverse primer is specific for T133 (5'-

CAGGGGTTCATGATGACCG-3'). The WM forward primer contains W88 (5'-CTAGTGAGCATCCAGGCAT-3') and the reverse primer contains M133 (5'-CAGGGGTTCATGATGACCA-3'). The PCR was done using HotStar Taq (Qiagen). Cycling protocol was: 95°C, 15 min; 30 cycles of 95°C, 15 sec; 66°C, 30s; 72°C, 1 min; and then 72°C, 10 min. Fifty percent of each reaction was analyzed on a 1% agarose gel (Research Products International Corp.). For sequencing, human genomic DNAs were amplified with HotStar Taq (Qiagen) with primers upstream (5'-AAGTGATGACAAGCTGAGCTGC-3') and downstream (5'-CCACAACATTTGCCTTAGGGGTA-3') of the OR7D4 open reading frame. The PCR products were then either gel-purified using MiniElute Gel Extraction Kit (Qiagen) or Sephadex™-purified (GE HealthCare) and sequenced with 3100 or 3730 Genetic Analyzer (ABI Biosystems) or by GeneWiz (New Brunswick, NJ). All samples were sequenced in addition to allele-specific PCR.

Polymorphisms in the six other odorant receptors that are in the same cluster as OR7D4 were assayed by sequencing. Human genomic DNAs were amplified with HotStar Taq (Qiagen) with primers upstream (5'-AACCATCTCCCTGTCATTC-3' for OR1M1, 5'-GTTTATCAGCAAGAAGTCTG-3' for OR7G2, 5'-CCCGCAGTCTAGAAAACAC-3' for OR7G1, 5'-CTTATAACTGGTTTGGTTTTG-3' for OR7G3, 5'-TGCCTGGCTAATGACCTC-3' for OR7D2, and 5'-GAGGGTGTATAATCCTATGTG-3' for OR7E24) and downstream (5'-TTAGCCCAAGACTCCCAG-3' for OR1M1, 5'-GTTTATCAGCAAGAAGTCTG-3' for OR7G2, 5'-AAAAAATCCAGGTGTGGTG-3' for OR7G1, 5'-ATTACTTCTTTCTCCCTGAC-3' for OR7G3, 5'-CCATTGGTGCTCACAAAAC-3'

for OR7D2, and 5'-GAATGAAAGCCATCAAGCAAC-3' for OR7E24) of the odorant receptor open reading frame. The PCR products were purified using Sephadryl™ S-400 (GE HealthCare) and sequenced using the same upstream and downstream primer pairs. Internal primers or alternative downstream primers were used for sequencing in the case of OR7G1 (downstream 5'-CAAGGAATCCTGTTATGATG-3'), OR7G3 (internal 5'-GGTACAATGTCATCATGAAC-3') and OR7E24 (internal 5'-CAGAGGCGTGGGTTTCATG-3' and 5'-TGGACATGCAAACCTCACAG-3').

Linkage disequilibrium of the SNPs was analyzed using JLIN⁷.

Human subjects

Subjects for this study were recruited from the greater New York City area for the Rockefeller University Smell Study under Rockefeller University IRB approved protocol #LVO-0539. All subjects gave informed consent to participate and were financially compensated for their time and effort. Exclusion criteria for subjects were: allergies to odours or fragrances, a history of nasal illness, upper respiratory infection, seasonal allergy, prior endoscopic surgery on the nose, pre-existing medical condition that has caused a reduced sense of smell such as head injury, cancer therapy, radiation to head and neck, or alcoholism. Pregnant women and children under 18 were excluded from this study. To control for inter-test variation, all subjects completed the same protocol on two different visits that were four or more days apart.

The intensity and valence ratings and the assignment of descriptors to odours shown in Figures 3 and 4 were measured in sessions between March 2005 and May 2006. The genotype of the subjects was unknown during this period of the study. The

thresholds shown in Figure 3 were measured between August and December 2006. At this time the genotypes of the subjects were known, but the subjects and the test administrators were blind to subject genotype information. We invited 100/242 subjects with the RT/RT genotype and all subjects with other genotypes back for thresholding to androstenone and androstadienone. Not all invited subjects participated in androstenone and androstadienone thresholding. All subjects participating in this thresholding previously participated in the intensity and valence ratings and the assigning of descriptors to odours. We obtained evaluable data from 412 subjects (~77% of enrolled subjects), who had to meet the minimum criteria of qualifying for the study, completing two study sessions, and providing a blood sample of adequate size for DNA isolation.

Procedures for olfactory psychophysics

All testing was performed in a well-ventilated room in the Outpatient Unit of the Rockefeller University Hospital. On the first visit, vital signs were collected from each subject and an 8.5 ml venous blood sample was collected. All women of child-bearing age took a urine pregnancy test to confirm that they were not pregnant. The average subject took 2.5 hours for the sessions that included the rating and descriptor data shown here and 30 minutes for the sessions in which the thresholds were determined. During the first visit only, we collected data on the demographics and habits of the subjects in a computer-administered questionnaire. Smell tests were self-administered and computerized using custom-written applications in Filemaker Pro and Microsoft Access. A screenshot from the computerized intensity and valence rating is shown in Supplementary Figure S4. The computer application for the intensity and valence rating

had a built-in mandatory 15 seconds inter-stimulus interval. However, it took most subjects longer to move from one stimulus to the next so this was rarely enforced. This application as well as the application in which odour descriptors were assigned to odours was written in Filemaker Pro. The application for thresholding was written in Microsoft Access. To prevent sampling errors, all odour vials used in this study were barcoded. Subjects scanned each odour vial containing the stimulus before sampling the stimulus and were only permitted to proceed if the correct vial was scanned.

Screening for subjects with general hyposmia and malingerers

To exclude malingerers and subjects with general hyposmia, a battery of olfactory tests that did not include the two odours under study (androstenone and androstadienone) was performed. In detail, thresholds for vanillin, pentadecalactone, and isovaleric acid were obtained following the procedure described below, and intensity ratings for 66 odours at two concentrations (“high” and “low”) and for the solvents were performed (see Supplementary Table S5). The subjects’ performance in the three thresholding experiments was ranked, along with ranking three performance indicators derived from the rating data:

1. The percentage of odors for which the “low” concentration was rated higher than the respective solvent.
2. The percentage of odors for which the “high” concentration was rated higher than the respective solvent.
3. The percentage of odors for which the “high” concentration was rated higher than the “low” concentration.

The average of the six ranks representing three indicators of performance in thresholding and three indicators of performance in intensity ranking was calculated among all 412 subjects and the 21 subjects (5%) with the lowest olfactory acuity were excluded from evaluation in this study. 13 of these 21 subjects were RT/RT, 4 RT/WM, and 4 RT/P79L. Therefore, the 391 subjects evaluated for this study were as follows:

OR7D4 genotype	# Subjects
RT/RT	242
RT/WM	96
RT/P79L	26
WM/WM	10
RT/S84N	7
WM/P79L	4
RT/D52G	2
WM/S84N	2
WM/L162P	1
S84N/P79L	1
TOTAL	391

Stimuli for olfactory psychophysics

All odours (Supplementary Table S4) were presented as one ml of an odour dilution in either propylene glycol or paraffin oil in 20 ml amber glass vials. Odour vials were used for 40 sessions and then replaced by a new set. Master stocks of each odour were established at the beginning of the study to avoid variability in odour concentrations. The concentrations of odours used in the intensity and valence rating are shown in Supplementary Table S5. The perceived quality of odours used in this study is listed in Supplementary Table S6.

For the descriptors task, the following odours and concentrations were used:

- propylene glycol: pure
- pentadecalactone: 1/500

- vanillin: 1/200
- androstenone: 1/10,000

The following odours were used for thresholding:

- androstenone
- androstadienone

These were tested at an initial concentration of binary dilution 23 (1/8,388,608) in propylene glycol and moved from there in binary steps. The range of dilutions tested was from binary dilution 27 (1:134,217,728) to binary dilution 6 (1:64).

Olfactory ratings

The intensity and valence of 66 odours were rated at two different concentrations (high and low) and the intensity and valence of the two solvents (paraffin oil and propylene glycol) was rated three times but for consistency only the first two ratings were evaluated. Prior to these ratings, six stimuli that represented the spectrum of intensity and valence of the stimuli used in the study were presented to allow the subjects to calibrate their usage of the scale. The subjects were unaware that the first six stimuli served this purpose. After the rating of the 66 odours at two concentrations and the rating of the solvents, 15 stimuli that were presented earlier in the experiment were repeated to test for the effect of adaptation and olfactory fatigue on the ratings. The subjects were not aware that the last 15 stimuli served this purpose. These 15 stimuli and the six first stimuli are not included in the analysis presented here. Odour stimuli were presented in the same order in all sessions to avoid bias caused by adaptation and olfactory fatigue and to make the measurements between different sessions as comparable as possible.

The 15 control stimuli that were presented twice during a session were as follows:

1. guaiacol (high)
2. octyl acetate (high)
3. undecanal (high)
4. paraffin oil
5. heptyl acetate (low)
6. hexyl butyrate (low)
7. butyric acid (high)
8. hexyl butyrate (high)
9. decyl aldehyde (high)
10. 2-decenal (low)
11. cis-3-hexenal (low)
12. nonyl aldehyde (high)
13. 2-methoxy-4-methylphenol (low)
14. decyl aldehyde (low)
15. propylene glycol

Although there was variability between the first and second presentation of these stimuli, there was no indication of a systematic difference between the intensity rating at the beginning and end of the session. Eight of the 15 stimuli were rated on average as more intense at the end of the session, whereas seven were rated as less intense.

A seven point scale was used to rate intensity and valence with these choices:

INTENSITY RATING

- Extremely Weak

- Very Weak
- Slightly Weak
- Neither Weak nor Strong
- Slightly Strong
- Very Strong
- Extremely Strong

VALENCE RATING

- Extremely Unpleasant
- Very Unpleasant
- Slightly Unpleasant
- Neither Unpleasant nor Pleasant
- Slightly Pleasant
- Very Pleasant
- Extremely Pleasant

In addition there was a button on the screen labeled "I can't smell anything" and a button labeled "Don't Know". If the "Don't Know" button was pressed, no rating was recorded. If the "I can't smell anything" button was pressed, a 0 was recorded for the intensity rating and no rating was recorded for the valence rating.

Prior to the study (in November and December 2004) the concentrations used for each odorant were determined in intensity matching experiments in which subjects rated the intensity of stimuli. Odours were considered "low" intensity when the intensity rating was within one standard deviation of the rating for a 1:10,000 dilution of 1-butanol. Odours were considered "high" intensity when the intensity rating was within one

standard deviation of a 1:1,000 dilution of 1-butanol. For ethylene brassylate, eugenol methyl ether, (-)-menthol, (+)-menthol, and vanillin the pure odour or the saturated dilution was rated less intense than the criteria for "high" intensity and these odours were therefore presented at the highest possible concentration. Androstenone and androstadienone could not be intensity matched in any meaningful way because of the high variability in the responses across subjects. The intensity matching protocol was approved by the Rockefeller University Institutional Review Board. All subjects gave their informed consent to participate and were financially compensated for their time and effort. Ten subjects participated in the intensity matching and six visits for each subject were necessary to match all stimuli. The subjects were aware of the purpose of the intensity matching and were instructed to focus on the intensity of the stimulus and disregard the valence. The concentrations of stimuli that resulted from the intensity matching and were subsequently used for the intensity and valence rating are shown in Supplementary Table S5.

Assigning descriptors to odours

Subjects assessed the quality of six odours using an odour profiling method that has shown to produce stable profiles of odorants⁸. Subjects were asked to rate a list of 146 odour descriptors on a scale from 0 ("descriptor does not at all describe my perception of the odour") to 5 ("descriptor perfectly describes my perception of the odour"). The 146 descriptors used in odor profiling were:

COMMON: sweet, fragrant, perfumery, floral, cologne, aromatic, musky, incense, bitter, stale, sweaty, light, heavy, cool/cooling, warm

FOUL: fermented/rotten fruit, sickening, rancid, putrid/foul/decayed, dead animal, mouse-like

FOODS: buttery (fresh), caramel, chocolate, molasses, honey, peanut butter, soupy, beer, cheesy, eggs (fresh), raisins, popcorn, fried chicken, bakery/fresh bread, coffee

MEATS: meat seasoning, animal, fish, kippery/smoked fish, blood/raw meat, meat/cooked good, oily/fatty

FRUITS: cherry/berry, strawberry, peach, pear, pineapple, grapefruit, grape juice, apple, cantaloupe, orange, lemon, banana, coconut, fruity/citrus, fruity/other

VEGETABLES: fresh vegetables, garlic/onion, mushroom, raw cucumber, raw potato, bean, green pepper, sauerkraut, celery, cooked vegetables

SPICES: almond, cinnamon, vanilla, anise/licorice, clove, maple/syrup, dill, caraway, minty/peppermint, nut/walnut, eucalyptus, malt, yeast, black pepper, tea leaves, spicy

BODY: dirty linen, sour milk, sewer, fecal/manure, urine, cat urine, seminal/like sperm

MATERIALS: dry/powdery, chalky, cork, cardboard, wet paper, wet wool/wet dog, rubbery/new, tar, leather, rope, metallic, burnt/smoky, burnt paper, burnt candle, burnt rubber, burnt milk, creosote, sooty, fresh tobacco smoke, stale tobacco smoke

CHEMICALS: sharp/pungent/acid, sour/acid/vinegar, ammonia, camphor, gasoline/solvent, alcohol, kerosene, household gas, chemical, turpentine/pine oil, varnish, paint, sulphidic, soapy, medicinal, disinfectant/carbolic, ether/anaesthetic, cleaning fluid/carbona, mothballs, nail polish remover

OUTDOORS: hay, grainy, herbal/cut grass, crushed weed, crushed grass, woody/resinous, bark/birch, musty/earthy, moldy, cedarwood, oakwood/cognac, rose, geranium leaves, violets, lavender, laurel leaves

Odour profiling typically took five minutes per odorant and was performed as a computer-controlled self-test in which the subject's responses were directly recorded. The default setting for each descriptor was set to 0, such that subjects recorded values from 1-5 for only those descriptors that pertained to their perception of a given odour. We provided large display posters listing all 146 odour descriptors so that subjects could study these before beginning this part of the test. Of the six odours, the first (spearmint) familiarized subject with the procedure and the descriptors. The subjects were not aware that the first odour served this purpose. Data obtained with this odour were not included in the analysis. For a second odour (methanethiol), we failed to produce reproducible intensities of the odour due to the high volatility of methanethiol, and these data were

therefore also not evaluated. The descriptors used to describe the other four odours (vanillin, pentadecalactone, androstenone, and propylene glycol) were used in this study. We evaluated the descriptors that were used in more than 10% of the sessions for a given odor. This represented 23/146 descriptors for vanillin (Fig. 4d-e), and 21/146 descriptors for androstenone (Fig. 4f-g). Of these 44 descriptors, only the five shown in Figure 4e and Figure 4g showed statistically significant differences between the genotypes. In Figure 4e and Figure 4g we plot the percentage with which a given descriptor was used by subjects of a given genotype. The descriptors used in more than 10% of the sessions to describe pentadecalactone (19/146) and propylene glycol (11/146) were not significantly affected by OR7D4 genotype.

Thresholding

Detection thresholds (Fig. 3d-e) were determined using the “Single Staircase Threshold Detection Method”^{9, 10}. This method produces accurate data on the threshold concentration of a given odour and is easy to administer and for the subjects to perform. We tested the thresholds for each subject on two occasions at least four days apart to control for inter-trial variability in olfactory performance. The median difference between the thresholds determined on the two days was 1.2 and 3.8 binary dilution steps for androstenone and androstadienone, respectively.

A custom-built, computer-controlled, self-administered thresholding test was completed by each subject. Odour vials had barcode labels and the procedure was carried out at a computer equipped with a bar code scanner. Subjects were instructed to sniff two vials, one containing the solvent, the other a dilution of the odorant. Subjects were asked

to scan the vial with the stronger odour. Depending on the answer, the procedure was repeated at an adjusted concentration. The total time to determine the threshold varied between subjects, but was typically between 15 and 25 minutes per odorant. We started the thresholding procedure at binary dilution 27 for the steroidal odours. If the subject failed to identify the right vial, the computer prompted the subject to move to a higher concentration in binary dilution steps. This continued until the subject chose the correct vial at one concentration five times in a row. At this point, the direction of the change in concentration was reversed and a lower concentration was tested. After this first reversal, the direction of the change in concentration was reversed whenever the subject made an incorrect choice when concentrations were decreased or two correct choices when concentrations were increased. The experiment continued until the seventh reversal. The thresholds reported represent the average of the last four reversals. The data in Figure 3 show the distribution of the thresholds for different genotypes in histograms. The data are displayed as the number of subjects of a given genotype having a given detection threshold. Each subject's threshold is the average of two replicates of the experiments on two visits four or more days apart. An example of a threshold procedure is shown in Supplementary Figure S5.

Statistical Analysis

Intensity and valence ratings: The intensity and valence ratings of the 68 stimuli (66 odours and 2 solvents) were ranked. Intensity ratings were ranked so that the rank “1” was assigned to the strongest odor and higher ranks to weaker odors. Valence ratings were ranked so that the rank “1” was assigned to the most pleasant odor and higher ranks

to less pleasant odours. Both intensity and valence ratings were analyzed with a Mann-Whitney-U test. To correct for multiple comparisons with the ratings, a Bonferroni correction was performed on the intensity and pleasantness ratings setting the significance threshold at 0.05. Since $n=68$ stimuli (66 odours and 2 solvents) were used, the significance thresholds are: 0.000735 for “significant” (*); 0.000147 for “highly significant” (**); 1.47E-05 for “extremely significant” (***).

Odour thresholding: The thresholds of different groups were compared with a Mann-Whitney-U test. The significance threshold (*) was 0.05.

Odour profiling: Differences in the usage of descriptors of odour quality was assessed with a chi-square test to compare the p-value from a contingency table with “descriptor used” and “descriptor not used” as the two outcomes. To correct for multiple descriptors in the usage of descriptors of odour quality, a Bonferroni correction was performed on the odour profiling data using 0.05 as the threshold for significance. The significance thresholds differ from odour to odour since they are dependent on the number of odours used by more than 10% of the subjects because only these descriptors were used in the statistical analysis. For androstenone, 21 descriptors were used frequently enough to be included in the analysis and therefore the significance thresholds are: 0.002381 for “significant” (*); 0.000476 for “highly significant” (**); 4.76E-05 for “extremely significant” (***).

Variance analysis of phenotype/genotype effect: The effect of the OR7D4 genotype on ratings of the odorous steroids was determined using non-parametric LOWESS (Locally weighted regression and smoothing scatter plots) regression¹¹. The R^2 values were 0.31 for the androstadienone intensity and 0.30 for the androstenone

intensity. For the valence, the values were 0.16 and 0.10, respectively. For this analysis, the average ratings of the two concentrations of the respective odour were the explanatory variables. When the average rating of both concentrations of both steroid odours was the explanatory variable, the R^2 values were 0.39 for the intensity rating and 0.19 for the valence rating ($p < 0.0001$ in all cases). Gender, race, as well as the average solvent rating, the average rating of low concentration of odour stimuli, and the average rating of high concentration of odour stimuli were used as covariates. These last three covariates were necessary to correct for the individual differences in how subjects used the rating scales.

References Cited

1. Glusman, G., Yanai, I., Rubin, I. & Lancet, D. The complete human olfactory subgenome. *Genome Res* 11, 685-702 (2001).
2. Krautwurst, D., Yau, K. W. & Reed, R. R. Identification of ligands for olfactory receptors by functional expression of a receptor library. *Cell* 95, 917-26 (1998).
3. Zhuang, H. & Matsunami, H. Synergism of accessory factors in functional expression of mammalian odorant receptors. *J Biol Chem* 282, 15284-93 (2007).
4. Philpott, N. J. et al. The use of 7-amino actinomycin D in identifying apoptosis: simplicity of use and broad spectrum of application compared with other techniques. *Blood* 87, 2244-51 (1996).
5. Saito, H., Kubota, M., Roberts, R. W., Chi, Q. & Matsunami, H. RTP family members induce functional expression of mammalian odorant receptors. *Cell* 119, 679-691 (2004).
6. Von Dannecker, L. E., Mercadante, A. F. & Malnic, B. Ric-8B promotes functional expression of odorant receptors. *Proc Natl Acad Sci U S A* 103, 9310-4 (2006).
7. Carter, K. W., McCaskie, P. A. & Palmer, L. J. Genome-wide linkage and association mapping of disease genes with the GAW14 simulated datasets. *BMC Genet* 6 Suppl 1, S41 (2005).
8. Dravnieks, A. Odor quality: semantically generated multidimensional profiles are stable. *Science* 218, 799-801 (1982).
9. Doty, R. L., McKeown, D. A., Lee, W. W. & Shaman, P. A study of the test-retest reliability of ten olfactory tests. *Chem Senses* 20, 645-56 (1995).
10. Doty, R. L. & Laing, D. G. in *Handbook of Olfaction and Gustation* (ed. Doty, R. L.) pp. 203-228 (Marcel Dekker, Inc., New York, 2003).
11. Cleveland, W. S. Robust locally weighted regression and smoothing scatterplots. *J. Am. Stat. Assoc.* 74, 829-836 (1979).

Legends to Supplementary Figures

Figure S1| Responses of two untagged OR7D4 alleles to androstenone and androstadienone

Dose-response curves measured in a luciferase assay for untagged RT and S84N variants of OR7D4 stimulated with androstenone and androstadienone (0, 1, 2, 4, 8, 16, and 32 μ M) are shown. Similar to the Rho-tagged counterparts, the S84N variant shows stronger responses compared to RT. Y-axis denotes normalized luciferase response \pm S.E.M. (n=3).

Figure S2| Linkage disequilibrium (LD) plot for OR7D4 gene cluster and responses of the major variants of the neighbouring odorant receptors to androstadienone

a, Linkage disequilibrium (LD) as measured by R^2 for non-synonymous SNPs or a frameshift (fs) is shown in the odorant receptor cluster including OR7D4. All SNPs detected in the 24 selected subjects regardless of allele frequency are shown. **b-c**, The major variants of six odorant receptors that are in the same odorant receptor gene cluster as OR7D4 on chromosome 19 (7D2, 1M1, 7G1, 7G2, 7G3, 7E24) were assessed for responses to 0.3 μ M, 3 μ M, and 30 μ M androstenone (**b**) androstadienone (**c**) in a luciferase assay. The responses of these six odorant receptors are compared to the responses of two OR7D4 alleles (RT, left; WM, second to left). Y-axis denotes normalized luciferase response \pm S.E.M. (n=4).

Figure S3| Distribution of variant OR7D4 proteins in permeabilized Hana3 cells

Permeabilized cell immunofluorescence of Hana3A cells expressing OR7D4 RT, WM, P79L, and S84N. Cells were stained with anti-rhodopsin antibody. Scale bar = 50 μ m.

Figure S4| Screenshot from the computerized intensity and valence rating

Shown here is the screen the subjects saw after scanning in vial #2 in the intensity and valence rating portion of the smell test. Subjects were instructed to click the "I Can't Smell Anything" button if they could not perceive any odour. If subjects were able to perceive the odour, they were asked to rate the strength and valence of the odour by clicking on the appropriate words. After both selections were made, subjects were prompted to scan vial #3 on the next screen.

Figure S5| Example of the threshold detection procedure

Data from one subject performing the thresholding procedure (see Supplementary Methods) are shown. In the representation above, when the subject scanned the vial containing the odour, this is denoted with a red "1". When the subject scanned the vial containing the solvent this is denoted with a red "0". The thresholds reported in the paper are the average of the last four reversals. In this example, the threshold was computed to be 13.25 (red horizontal line), since the last four reversals were at dilutions 12, 14, 13, and 14.

Figure S6| RT/RT, RT/WM, WM/WM intensity rating

The data used to calculate the values shown in Figure 3b are shown. The histograms in the two uppermost rows show all the data used to calculate Figure 3b. N=242 for RT/RT, N=96 for RT/WM, and N=10 for WM/WM. Since each subject was tested twice, the number of data points represented in the histograms is twice the number of the subjects. In Figure 3b and in the other figures showing results from ratings (Fig. 3a and c, Fig. 4a-c), the ratings for the two concentrations of an odour are pooled. Note that the high concentration of androstenone (left in lower row) is rated "extremely weak" by fewer than five percent of the RT/RT subjects, but by a much higher percentage of RT/WM, and WM/WM subjects. For "extremely strong", this situation is reversed. In the middle row, histograms showing the distribution of ranks of the 68 stimuli are shown, with the data represented in bins of 10 rankings (blue lines), with 1 being the odour ranked as most intense and 66 the odour ranked as least intense. In the bottom row, these rank-histograms are converted to median, 1.quartile, 3.quartile, lower and upper limit (before correction: * $p < 0.00073$, ** $p < 0.00014$, *** $p < 1.47E-05$; after correction: * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$).

Figure S7| Ratings divided by genotype and gender

The difference in intensity (a) and valence (b) ranking of RT/WM and RT/RT subjects is shown divided by gender (male RT/RT: N=111; male RT/WM: N=41; female RT/RT: N=131; female RT/WM: N=55). Significance was assessed with a Mann-Whitney-U test with a Bonferroni correction for the 66 odours and two solvents tested (before correction:

* $p < 0.00073$, ** $p < 0.00014$, *** $p < 1.47 \times 10^{-5}$; after correction: * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$).

Figure S8| Ratings divided by genotype and race

The difference in intensity (a) and valence (b) ranking of RT/WM and RT/RT subjects is shown divided by race for Caucasian and African-American subjects (African-American RT/RT: $N=55$; African-American RT/WM: $N=9$; Caucasian RT/RT: $N=107$; Caucasian RT/WM: $N=62$). Significance was assessed with a Mann-Whitney-U test with a Bonferroni correction for the 66 odours and two solvents tested (before correction: * $p < 0.00073$, ** $p < 0.00014$, *** $p < 1.47 \times 10^{-5}$; after correction: * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$).

Figure S9| Thresholds of WM/WM subjects

The data for WM/WM subjects ($n=7$) who were tested for their androstenone (a) and androstadienone (b) thresholds are shown. This is compared to the median of the RT/RT subjects ($n=46$) and the median of the WM/P79L ($n=2$) subjects, who carry two non-functional alleles of OR7D4.

Figure S10| RT/RT, RT/WM, WM/WM valence rating

The data used to calculate the values shown in Figure 4b are shown. The histograms in the two uppermost rows show all the data used to calculate Figure 4b. $N=242$ for RT/RT, $N=96$ for RT/WM, and $N=10$ for WM/WM. Since each subject was tested twice, the number of data points represented in the histograms is twice the number of the subjects.

The ratings for the two concentrations of an odour are pooled. Note that at both concentrations of androstenone and androstadienone "extremely unpleasant" was used more frequently by RT/RT subjects than by RT/WM subjects. None of the WM/WM subjects rated any of these four stimuli as "extremely unpleasant". In the middle row, histograms showing the distribution of ranks of the 68 stimuli are shown, with the data represented in bins of 10 rankings (blue lines), with 1 being the odour ranked as most pleasant and 66 the odour ranked most unpleasant. In the bottom row these rank-histograms are converted to median, 1.quartile, 3.quartile, lower and upper limit (before correction: * $p < 0.00073$, ** $p < 0.00014$, *** $p < 1.47E-05$; after correction: * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$).

	OR name	Normalized response
1	OR1A1	0.130
2	OR1A2	0.012
3	OR1B1	0.042
4	OR1C1	0.016
5	OR1D2	0.042
6	OR1D4	-0.006
7	OR1D5	-0.026
8	OR1E2	0.049
9	OR1E5	0.029
10	OR1F13	-0.011
11	OR1G1	0.005
12	OR1I1	-0.001
13	OR1J2	0.066
14	OR1J4	0.008
15	OR1K1	0.004
16	OR1L1	0.078
17	OR1L3	0.068
18	OR1L4	0.044
19	OR1L8	0.039
20	OR1M1	0.017
21	OR1N1	0.000
22	OR1N2	0.025
23	OR1Q1	0.027
24	OR1S1	0.017
25	OR1S2	0.014
26	OR2A2	0.018
27	OR2A5	0.083
28	OR2A6	0.038
29	OR2A7	0.024
30	OR2A10	0.008
31	OR2A12	0.016
32	OR2A25	0.215
33	OR2AE1	0.018
34	OR2AG2	0.023
35	OR2AJ1	0.034
36	OR2AK2	0.031
37	OR2B3	0.045
38	OR2B6	0.079
39	OR2B9	0.005

40	OR2B11	-0.004
41	OR2C3	0.029
42	OR2C4	-0.001
43	OR2D2	0.008
44	OR2D3	-0.015
45	OR2F1	0.052
46	OR2F2	0.037
47	OR2G2	0.013
48	OR2G3	0.003
49	OR2G6	-0.004
50	OR2H1	0.013
51	OR2H2	0.041
52	OR2H3	0.023
53	OR2J2	0.255
54	OR2J3	0.042
55	OR2K2	0.014
56	OR2L13	0.032
57	OR2L2	0.055
58	OR2L8	0.031
59	OR2M2	0.017
60	OR2M3	0.126
61	OR2M4	0.018
62	OR2M5	0.027
63	OR2S2	-0.058
64	OR2T1	0.002
65	OR2T4	0.014
66	OR2T5	0.032
67	OR2T6	0.049
68	OR2T7	0.006
69	OR2T10	0.041
70	OR2T11	0.032
71	OR2T12	0.042
72	OR2T29	0.051
73	OR2T34	0.031
74	OR2T35	0.044
75	OR2V3	0.067
76	OR2W1	-0.342
77	OR2W3	0.014
78	OR2W5	0.026
79	OR2Y1	0.029
80	OR2Z1	-0.003

81	OR3A1	-0.002
82	OR3A2	-0.036
83	OR4A4	0.022
84	OR4A5	0.015
85	OR4A8	0.065
86	OR4A15	0.027
87	OR4B1	0.007
88	OR4C3	0.018
89	OR4C5	0.001
90	OR4C6	0.033
91	OR4C11	0.012
92	OR4C12	0.032
93	OR4C15	0.015
94	OR4C16	0.010
95	OR4C46	0.062
96	OR4D1	0.021
97	OR4D2	0.002
98	OR4D5	0.021
99	OR4D6	0.012
100	OR4D9	0.022
101	OR4D10	0.015
102	OR4D11	0.032
103	OR4E2	-0.103
104	OR4F12	0.042
105	OR4F14	0.053
106	OR4F15	0.027
107	OR4F16	0.019
108	OR4F17	0.015
109	OR4F21	-0.007
110	OR4K1	-0.005
111	OR4K2	0.004
112	OR4K3	0.019
113	OR4K5	0.010
114	OR4K13	0.019
115	OR4K14	0.024
116	OR4K17	0.007
117	OR4L1	0.037
118	OR4M2	0.002
119	OR4N2	0.020
120	OR4N4	0.002
121	OR4N5	0.026

122	OR4P4	0.032
123	OR4Q3	0.001
124	OR4S1	-0.009
125	OR4S2	0.039
126	OR4X1	0.055
127	OR4X2	0.026
128	OR5A1	0.029
129	OR5A2	0.046
130	OR5AC2	0.003
131	OR5AK2	-0.041
132	OR5AK2	0.026
133	OR5AK3	0.017
134	OR5AN1	0.182
135	OR5AP2	0.020
136	OR5AR1	0.042
137	OR5AT1	0.061
138	OR5AU1	-0.024
139	OR5AU1	0.026
140	OR5AV1	0.019
141	OR5AY1	0.022
142	OR5B2	0.031
143	OR5B13	0.030
144	OR5B16	0.015
145	OR5B17	0.022
146	OR5B21	0.022
147	OR5BF1	0.048
148	OR5C1	0.007
149	OR5D13	0.016
150	OR5D14	0.018
151	OR5D18	0.031
152	OR5F1	0.021
153	OR5H1	0.010
154	OR5H2	0.015
155	OR5H6	0.019
156	OR5H14	-0.018
157	OR5H15	0.031
158	OR5I1	0.027
159	OR5J2	-0.009
160	OR5K1	-0.085
161	OR5K2	0.020
162	OR5L1	0.005

163	OR5L2	-0.003
164	OR5M1	0.034
165	OR5M9	0.014
166	OR5M10	0.036
167	OR5M11	0.025
168	OR5P2	0.005
169	OR5P3	0.003
170	OR5R1	0.011
171	OR5T1	0.040
172	OR5T2	0.029
173	OR5T3	0.040
174	OR5U1	-0.005
175	OR5V1	0.018
176	OR6A2	0.041
177	OR6B1	0.046
178	OR6B2	-0.008
179	OR6B3	0.045
180	OR6C1	0.035
181	OR6C2	0.048
182	OR6C3	0.022
183	OR6C4	0.035
184	OR6C6	0.017
185	OR6C65	0.028
186	OR6C68	0.062
187	OR6C70	0.060
188	OR6C74	0.030
189	OR6C76	0.008
190	OR6F1	0.050
191	OR6K2	0.032
192	OR6K3	0.013
193	OR6K6	-0.013
194	OR6M1	0.009
195	OR6N1	0.023
196	OR6N2	0.024
197	OR6P1	0.026
198	OR6Q1	0.010
199	OR6S1	-0.002
200	OR6T1	0.023
201	OR6X1	0.009
202	OR6Y1	0.054
203	OR7A5	-0.003

204	OR7A10	0.015
205	OR7A17	-0.007
206	OR7C1	0.233
207	OR7C2	-0.013
208	OR7D4	1.000
209	OR7E24	-0.025
210	OR7G1	-0.012
211	OR7G2	0.024
212	OR7G3	0.013
213	OR8A1	-0.002
214	OR8B2	0.026
215	OR8B4	0.022
216	OR8B8	0.011
217	OR8B12	-0.005
218	OR8D1	0.012
219	OR8D2	0.016
220	OR8D4	0.047
221	OR8G1	0.058
222	OR8G5	-0.001
223	OR8H1	0.018
224	OR8H2	0.016
225	OR8H3	0.023
226	OR8I2	0.025
227	OR8J1	-0.005
228	OR8J3	0.007
229	OR8K1	0.021
230	OR8K3	0.006
231	OR8K5	0.035
232	OR8S1	0.005
233	OR8U1	0.041
234	OR8U8	0.002
235	OR8U9	0.023
236	OR9A2	0.054
237	OR9A4	0.058
238	OR9G1	0.039
239	OR9G4	0.020
240	OR9G5	0.051
241	OR9G9	0.017
242	OR9I1	0.003
243	OR9K2	0.018
244	OR9Q1	0.008

245	OR9Q2	-0.031
246	OR10A2	0.026
247	OR10A3	0.011
248	OR10A4	0.010
249	OR10A5	0.009
250	OR10A6	0.070
251	OR10A7	0.004
252	OR10AG1	0.019
253	OR10C1	0.034
254	OR10G3	0.027
255	OR10G6	0.033
256	OR10G7	0.060
257	OR10G8	-0.001
258	OR10G9	0.007
259	OR10H1	0.012
260	OR10H2	-0.052
261	OR10H4	-0.013
262	OR10H5	0.024
263	OR10J1	0.067
264	OR10J3	0.038
265	OR10J5	0.021
266	OR10K1	0.075
267	OR10K2	0.049
268	OR10P1	0.003
269	OR10Q1	0.018
270	OR10R2	0.008
271	OR10T2	0.056
272	OR10W1	0.019
273	OR10X1	0.019
274	OR10Z1	0.064
275	OR11A1	0.044
276	OR11G2	0.051
277	OR11H1	0.016
278	OR11H4	0.008
279	OR11H6	0.042
280	OR11L1	0.073
281	OR12D2	0.021
282	OR12D3	0.022
283	OR13A1	-0.008
284	OR13C3	0.005
285	OR13C4	0.051

286	OR13C5	0.047
287	OR13C8	0.035
288	OR13C9	0.052
289	OR13D1	0.021
290	OR13F1	0.028
291	OR13G1	0.023
292	OR13H1	0.031
293	OR13J1	0.066
294	OR51A4	0.029
295	OR51D1	0.024
296	OR51E1	-0.051
297	OR51E2	-0.459
298	OR51F1	0.017
299	OR51F2	-0.005
300	OR51G1	0.039
301	OR51G2	0.019
302	OR51H1	0.008
303	OR51I1	-0.002
304	OR51I2	0.019
305	OR51L1	-0.098
306	OR51M1	0.026
307	OR51Q1	0.045
308	OR51T1	-0.004
309	OR52A1	0.019
310	OR52B2	0.008
311	OR52B6	-0.002
312	OR52D1	0.030
313	OR52E1	0.024
314	OR52E2	-0.001
315	OR52E4	0.015
316	OR52E5	0.062
317	OR52E6	0.009
318	OR52E8	0.020
319	OR52H1	0.010
320	OR52I2	0.020
321	OR52J3	0.031
322	OR52K1	0.010
323	OR52K2	0.022
324	OR52L1	0.026
325	OR52L2	0.016
326	OR52M1	0.020

327	OR52N1	0.002
328	OR52N4	0.038
329	OR52N5	0.028
330	OR52P1	0.010
331	OR52W1	0.028
332	OR56A1	0.028
333	OR56A6	0.063
334	OR56B1	0.021
335	OR56B4	0.018
	no OR	0.021
	no OR SD (N=8)	0.023

Supplementary Table S2| Single nucleotide polymorphisms in OR7D4 and their distribution among 391 subjects

SNP	dbSNP ID	Chromosomal Position	Allele	AA change	Codon	Protein region	Allele frequency	Racial self-identification***					
								African-American	Asian	Caucasian	Do not wish to specify	Native American	Other**
refseq							615/782 0.786	24%	8%	45%	3%	1%	19%
1		9186359	T/C	D52G	2	IC1	2/782 0.003	100%	0%	0%	0%	0%	0%
2	rs5020281	9186290	G/C	S75C	2	TM2	0/782 0						
3		9186278	G/A	P79L	2	TM2	31/782 0.040	81%	0%	0%	6%	0%	13%
4	rs5020280	9186263	C/T	S84N	2	EC1	10/782 0.013	50%	0%	10%	0%	0%	40%
5		9186252	G/A	R88W	1	EC1	123/782 0.157	13%	9%	63%	4%	1%	11%
6	rs5020279	9186121	G/C	H131Q	3	IC2	0/782 0						
7	rs5020278	9186116	G/A	T133M	2	IC2	123/782 0.157	13%	9%	63%	4%	1%	11%
8	rs5020277	9186106	C/T	M136I	3	IC2	0/782 0						
9	rs5020276	9186099	A/G	C139R	1	IC2	0/782 0						
10	rs5020275	9186098	C/T	C139Y	2	IC2	0/782 0						
11		9186029	A/G	L162P	2	TM4	1/782 0.001	0%	0%	100%	0%	0%	0%
12		9185678	G/T	A279D	2	TM7	0/782 0*						
13	rs4564704	9185640	G/T	L292M	1	TM7	0/782 0						

*One individual with this SNP was found but was not used for psychophysical analysis.

**Of the 70 subjects who chose "Other" as a race category, 74% self-identified as Hispanic, 17% as mixed race, and 7% as African.

***Percentages are rounded off to the nearest whole number.

Supplementary Table S3| Single nucleotide polymorphisms in other odorant receptors in the OR7D4 locus and their distribution among 24 selected subjects

Gene name	AA change	Allele frequency	
OR1M1	A91T	1/48	0.021
	V167I	1/48	0.021
	S274F	1/48	0.021
OR7G2	V284A	24/48	0.500
	F302V	23/48	0.479
OR7G1	Y35C	1/48	0.021
	V83A	12/48	0.250
	W141C	19/48	0.396
	A156V	1/48	0.021
	F248I	1/48	0.021
	Y252C	11/48	0.229
OR7G3	M29V	13/48	0.271
	S64F	1/48	0.021
	R122*	1/48	0.021
	R220*	1/48	0.021
	V247I	1/48	0.021
	S310fs**	11/48	0.229
OR7D2	T197M	1/48	0.021
	V247I	1/48	0.021
OR7E24	I64V	3/48	0.063
	P242S	22/48	0.458

* nonsense mutation

** fs: frameshift caused by a 5-bp insertion at nt929

Supplementary Table S4| Odours used in this study

ODOUR			
NAME	ALTERNATIVE NAME	CAS #	MW
(-)-menthol	l-menthol	2216-51-5	156
(+)-menthol	d-menthol	15356-60-2	156
1-butanol		71-36-3	74
2-butanone		78-93-3	72
2-decenal		3913-71-1	154
2-ethylfenchol		18368-91-7	182
2-methoxy-4-methylphenol	creosol	93-51-6	138
4-methylvaleric acid	isocaproic acid	646-07-1	116
ambrette		8015-62-1	
androstadienone		794-58-9	270
androstenone		18339-16-7	272
anise		8007-70-3	
banana			
bourgeonal		18127-01-0	190
butyl acetate		123-86-4	116
butyric acid	butanoic acid	107-92-6	88
caproic acid	hexanoic acid	142-62-1	116
cedarwood		68990-83-0	
cineole	eucalyptol	470-82-6	154
cinnamon		8015-91-6	
cis-3-hexen-1-ol		928-96-1	100
citral		5392-40-5	152
citronella		8000-29-1	
decyl aldehyde	decanal	112-31-2	156

Supplementary Table S4-cont'd

ODOUR			
NAME	ALTERNATIVE NAME	CAS #	MW
diacetyl	2,3-butanedione	431-03-8	86
diallyl sulfide		592-88-1	114
diphenyl ether	diphenyl oxide	101-84-8	170
ethyl vanillin	3-ethoxy-4-hydroxybenzaldehyde	121-32-4	166
ethylene brassylate	Musk T	105-95-3	270
eugenol		97-53-0	164
eugenol acetate		93-28-7	206
eugenol methyl ether		93-15-2	178
fenchone	L-alpha-fenchone	7787-20-4	152
fir		8002-09-3	
galaxolide	musk galaxolide	1222-05-5	258
geranyl acetate		105-87-3	196
guaiacol		90-05-1	124
heptaldehyde	heptanal	111-71-7	114
heptyl acetate		112-06-1	158
hexyl butyrate		2639-63-6	172
isobornyl acetate		125-12-2	196
isobutyraldehyde		78-84-2	72
isobutyric acid		79-31-2	88
isoeugenol		97-54-1	164
isovaleric acid		503-74-2	102
jasmine		8022-96-6	
lime		8008-26-2	
linalool	(+/-) linalool	78-70-6	154

Supplementary Table S4-cont'd

ODOUR			
NAME	ALTERNATIVE NAME	CAS #	MW
methanethiol	methyl mercaptan	5188-07-8	70
methyl salicylate		119-36-8	152
nonyl aldehyde	nonanal	124-19-6	142
nutmeg		8008-45-5	
octyl acetate		112-14-1	172
octyl aldehyde	octanal	124-13-0	128
orange		8008-57-9	
pentadecalactone	Thibetolide	106-02-5	240
phenyl acetaldehyde	phenyl ethanal	122-78-1	120
pyrazine		290-37-9	80
r-carvone	L-carvone	6485-40-1	150
r-limonene	D-limonene	5989-27-5	136
sandalwood		8006-87-9	
spearmint		8008-79-5	
terpineol	alpha-terpineol	98-55-5	154
terpinyl acetate		80-26-2	196
undecanal		112-44-7	170
vanillin		121-33-5	152
paraffin oil		8012-95-1	
propylene glycol	dipropylene glycol	57-55-6	

Supplementary Table S5| Dilutions of odours used in this study

ODOUR	PSYCHOPHYSICS			LUCIFERASE ASSAY	
NAME	LOW CONCENTRATION	HIGH CONCENTRATION	SOLVENT	CONCENTRATION	SOLVENT
(-)-menthol	1/400	1/40	propylene glycol	1/213,310 (=30μM)	CD293
(+)-menthol	1/400	1/40	propylene glycol	1/213,310 (=30μM)	CD293
1-butanol	1/10,000	1/1,000	paraffin oil	1/449,707 (=30μM)	CD293
2-butanone	1/10,000	1/5,000	paraffin oil	1/462,279 (=30μM)	CD293
2-decenal	1/10,000	1/1,000	paraffin oil	1/216,097 (=30μM)	CD293
2-ethylfenchol	1/100,000	1/5,000	paraffin oil	1/182,844 (=30μM)	CD293
2-methoxy-4-methylphenol	1/1,000,000	1/100,000	paraffin oil	1/241,256 (=30μM)	CD293
4-methylvaleric acid	1/5,000,000	1/10,000	paraffin oil	1/286,962 (=30μM)	CD293
ambrette	1/1,000,000	1/1,000	paraffin oil	1/30,000	CD293
androstadienone	1/100,000	1/1,000	propylene glycol	1/123,457 (=30μM)	CD293
androstenone	1/100,000	1/1,000	propylene glycol	1/122,356 (=30μM)	CD293
anise	1/50,000	1/5,000	paraffin oil	1/30,000	CD293
banana	1/250,000	1/10,000	paraffin oil	1/30,000	CD293
bourgeonal	1/2,000	1/200	paraffin oil	1/175,176 (=30μM)	CD293
butyl acetate	1/1,000,000	1/1,000	paraffin oil	1/286,962 (=30μM)	CD293
butyric acid	1/1,000,000	1/250,000	paraffin oil	1/378,332 (=30μM)	CD293
caproic acid	1/1,000,000	1/250,000	paraffin oil	1/286,962 (=30μM)	CD293
cedarwood	1/5,000	1/2,000	paraffin oil	1/30,000	CD293
cineole	1/100,000	1/1,000	paraffin oil	1/216,097 (=30μM)	CD293
cinnamon	1/50,000	1/10,000	paraffin oil	1/30,000	CD293
cis-3-hexen-1-ol	1/250,000	1/100,000	paraffin oil	1/332,800 (=30μM)	CD293
citral	1/50,000	1/5,000	paraffin oil	1/218,959 (=30μM)	CD293
citronella	1/250,000	1/10,000	paraffin oil	1/30,000	CD293
decyl aldehyde	1/25,000	1/5,000	paraffin oil	1/213,310 (=30μM)	CD293

Supplementary Table S5-cont'd

ODOUR		PSYCHOPHYSICS		LUCIFERASE ASSAY	
NAME	LOW CONCENTRATION	HIGH CONCENTRATION	SOLVENT	CONCENTRATION	SOLVENT
diacetyl	1/10,000,000	1/10,000	paraffin oil	1/387,191 (=30µM)	CD293
diallyl sulfide	1/2,000,000	1/100,000	paraffin oil	1/291,873 (=30µM)	CD293
diphenyl ether	1/500	1/200	paraffin oil	1/195,836 (=30µM)	CD293
ethyl vanillin	1/1,000	1/200	propylene glycol	1/200,590 (=30µM)	CD293
ethylene brassylate	1/500	1/100	paraffin oil	1/123,289 (=30µM)	CD293
eugenol	1/25,000	1/1,000	paraffin oil	1/203,000 (=30µM)	CD293
eugenol acetate	1/1,000,000	1/100	paraffin oil	1/161,623 (=30µM)	CD293
eugenol methyl ether	1/500	1/10	paraffin oil	1/187,024 (=30µM)	CD293
fenchone	1/25,000	1/1,000	paraffin oil	1/218,959 (=30µM)	CD293
fir	1/100,000	1/10,000	paraffin oil	1/30,000	CD293
galaxolide	1/10	1/1,000	paraffin oil	1/128,998 (=30µM)	CD293
geranyl acetate	1/10,000	1/200	paraffin oil	1/169,818 (=30µM)	CD293
guaiacol	1/50,000,000	1/1,000,000	paraffin oil	1/268,516 (=30µM)	CD293
heptaldehyde	1/10,000,000	1/25,000	paraffin oil	1/291,919 (=30µM)	CD293
heptyl acetate	1/25,000	1/2,500	paraffin oil	1/210,650 (=30µM)	CD293
hexyl butyrate	1/1,000	1/100	paraffin oil	1/193,498 (=30µM)	CD293
isobornyl acetate	1/2,000,000	1/100	paraffin oil	1/169,818 (=30µM)	CD293
isobutyraldehyde	1/100,000	1/1,000	paraffin oil	1/462,279 (=30µM)	CD293
isobutyric acid	1/10,000	1/1,000	paraffin oil	1/378,332 (=30µM)	CD293
isoeugenol	1/25,000	1/2,000	paraffin oil	1/203,000 (=30µM)	CD293
isovaleric acid	1/2,000,000	1/20,000	paraffin oil	1/326,372 (=30µM)	CD293
jasmine	1/200,000	1/1,000	paraffin oil	1/30,000	CD293
lime	1/2,000,000	1/5,000	paraffin oil	1/30,000	CD293
linalool	1/100,000	1/100	paraffin oil	1/216,097 (=30µM)	CD293

Supplementary Table S5-cont'd

ODOUR		PSYCHOPHYSICS		LUCIFERASE ASSAY	
NAME	LOW CONCENTRATION	HIGH CONCENTRATION	SOLVENT	CONCENTRATION	SOLVENT
methanethiol	1/50,000,000	1/10,000,000	water	1/475,617 (=30μM)	CD293
methyl salicylate	1/25,000	1/1,000	paraffin oil	1/219,083 (=30μM)	CD293
nonyl aldehyde	1/100,000	1/5,000	paraffin oil	1/234,345 (=30μM)	CD293
nutmeg	1/25,000	1/1,000	paraffin oil	1/30,000	CD293
octyl acetate	1/100	1/200	paraffin oil	1/193,498 (=30μM)	CD293
octyl aldehyde	1/250,000	1/25,000	paraffin oil	1/259,982 (=30μM)	CD293
orange	1/2,500	1/100	paraffin oil	1/30,000	CD293
pentadecalactone	1/2,000	1/500	propylene glycol	1/138,666 (=30μM)	CD293
phenyl acetaldehyde	1/8,000,000	1/2,000,000	paraffin oil	1/277,430 (=30μM)	CD293
pyrazine	1/500	1/10	propylene glycol	1/416,204 (=30μM)	CD293
r-carvone	1/100,000	1/1,000	paraffin oil	1/221,897 (=30μM)	CD293
r-limonene	1/250	1/10	paraffin oil	1/244,673 (=30μM)	CD293
sandalwood	1/10,000	1/1,000	paraffin oil	1/30,000	CD293
spearmint	1/250,000	1/5,000	paraffin oil	1/30,000	CD293
terpineol	1/10,000	1/100	paraffin oil	1/216,097 (=30μM)	CD293
terpinyl acetate	1/1,000	1/500	paraffin oil	1/169,818 (=30μM)	CD293
undecanal	1/10,000	1/1,000	paraffin oil	1/195,740 (=30μM)	CD293
vanillin	1/1,000	1/200	propylene glycol	1/195,740 (=30μM)	CD293
paraffin oil					CD293
propylene glycol					CD293

Supplementary Table S6| Perceived quality of odours used in this study

ODOUR		
NAME	QUALITY	DESCRIPTORS
(-)-menthol	Refreshing, light, diffusive odor with a sweet pungency. Characteristic resemblance to main odor of Peppermint, and a cooling effect upon the mucous membrane...	cool (cooling), minty (peppermint), medicinal
(+)-menthol	Odor very similar to that of (L-)Menthol, but somewhat more woody, not quite as sweet, and the cooling sensation is not perceptible at the same low concentration as it is in (L-)Menthol.	
1-butanol	Mild "fusel"-like odor, more volatile and more choking than fusel oil itself. Although somewhat winery in character, its odor is really nondescript, rather "chemical".	
2-butanone	Ethereal, slightly nauseating odor, not exactly pleasant.	
2-decenal	Very powerful, waxy-Orange-like, sweet-"aldehydic" odor of fair tenacity and high diffusive power.	
2-ethylfenchol		
2-methoxy-4-methylphenol	Sweet-spicy, phenolic-leathery odor with distinctly Vanilla-like undertones, balsamic-warm sweetness...	
4-methylvaleric acid	Unpleasant sour and penetrating odor. Less fatty-sweat-like than Caproic acid, more pungent.	
ambrette	(ambrette seed oil:) ...rich, sweet, floral-musky, distinctly wine-like or brandy-like odor...	
androstadienone		
androstenone	Reported as having a penetrating urine-like odor.	
anise	A very common description is that of "licorice odor".	
banana		
bourgeonal		
butyl acetate	Very diffusive, ethereal-fruity, pungent odor, reminiscent of many kinds of (ripe and over-ripe) fruit... The odor is often described as resembling that of Pear, Banana, Strawberry, etc...	
butyric acid	Powerful, penetrating, diffusive odor, reminiscent of rancid butter.	sickening, putrid (foul, decayed), rancid
caproic acid	Heavy, acrid-acid, fatty-rancid odor, often described as "sweat-like".	sour (vinegar), sickening, sweaty
cedarwood	(cedarwood oil, Virginia:) ... at first oily-woody and almost sweet, mild and pleasant, somewhat balsamic and typical of cedarwood (lumber).	(Cedrone S:) cedarwood, wood (resinous), fragrant
cineole	Fresh, diffusive, camphoraceous-cool odor of poor tenacity.	camphor, medicinal, eucaliptus
cinnamon	(cinnamon bark oil:) ...extremely powerful, diffusive, warm-spicy, sweet and tenacious odor.	(cinnamon bark oil:) cinnamon, spicy, fragrant
cis-3-hexen-1-ol	Powerful and intensely green, grassy odor.	
citral	Widely used as a powerful Lemon-fragrance chemical...	(geranial:) lemon, fruity (citrus), fragrant
citronella	(citronella oil, Ceylon:) The odor is very peculiar, warm-woody and yet fresh, grassy and somewhat reminiscent of wet leaves.	
decyl aldehyde	Penetrating and very powerful, sweet-waxy, Orange-peel-like odor. In extreme dilution refreshing, Citrus-peel-like.	

Supplementary Table S6-cont'd

ODOUR		
NAME	QUALITY	DESCRIPTORS
diacetyl	Very powerful and diffusive, pungent, buttery odor. Chlorine-Quinone-like in high concentration, oily-buttery in extreme dilution.	
diallyl sulfide	Powerful, penetrating Garlic-Radish-like odor of poor tenacity.	
diphenyl ether	Harsh floral-green, metallic Geranium type odor.	aromatic, fragrant, chemical
ethyl vanillin	Intensely sweet odor, warm, slightly floral and with some resemblance to Vanilla in its creamy sweetness.	
ethylene brassylate	Sweet-musky, somewhat oily odor of outstanding tenacity.	
eugenol	Powerful, warm-spicy, rather dry and almost sharp odor, drier and harder than that of Clove bud oil, less peppery-woody than that of Clove leaf oil.	clove, spicy, fragrant
eugenol acetate	Fruity-balsamic, warm and faintly spicy odor, somewhat clove-like, but with rosy floral notes and vanilla-like sweetness.	
eugenol methyl ether	Peculiar musty-tealike, warm and mildly spicy, slightly earthy, tenacious odor. The warm notes are mostly herbaceous... there is an overall resemblance to notes of Ginger and Tea.	
fenchone	Warm-camphoraceous, powerful and diffusive, basically sweet odor.	
fir	(fir needle oil, Siberian:) Its odor is refreshingly balsamic, slightly fatty or oily with a powerful pine-forest odor, and a peculiar fruity-balsamic undertone.	
galaxolide	Sweet and musky odor of good tenacity,... it has a tendency of appearing somewhat woody and not sufficiently sweet.	fragrant, perfumery, floral
geranyl acetate	Sweet, fruity-floral, rosy, somewhat green and remotely Lavender-like odor of moderate tenacity.	
guaiacol	Powerful smoke-like, somewhat medicinal odor, sweeter than the prototype phenolic odor.	burnt (smoky), medicinal, woody (resinous)
heptaldehyde	Very powerful and diffusive oily-fatty, "rancid" odor. Penetrating and pungent at high concentration, almost fruity, "fermented-fruit"-like in extreme dilution.	oily (fatty), woody (resinous), sickening
heptyl acetate	Fruity, fatty-green and slightly floral odor with pleasant, leafy undertones. The delicate, leafy notes resemble those found in Apple-rose or Hiprose. ...earthy-green or mushroom-like note.	
hexyl butyrate	Powerful, fruity, heavy odor, reminiscent not of one particular fruit, but of a "melange" of unripe fruits.	
isobornyl acetate	Mild oily-piney, balsamic-camphoraceous odor, reminiscent of Spruce Needles or certain Pine Needles.	
isobutyraldehyde	Extremely diffusive, penetrating odor, pungent and - undiluted - unpleasant, sour, repulsive. In extreme dilution it becomes almost pleasant, fruity, Banana-like, "overripe fruit-like".	
isobutyric acid	Powerful, diffusive sour (acid) odor, slightly less repulsive, and also less buttery than the n-Butyric acid. In extreme dilution the odor becomes almost pleasant, fruity.	
isoeugenol	Mild and sweet, deep-floral, very tenacious odor with great warmth and resemblance to Carnation, Sweet Williams or Wallflower.	
isovaleric acid	Very diffusive, acid-acrid, in moderate dilution cheesy, unpleasant odor of poor tenacity...	
jasmine	(jasmin absolute from concrete:) ... intensely floral, warm, rich and highly diffusive odor with a peculiar waxy-herbaceous, oily-fruity and tea-like undertone.	
lime	(lime oil, distilled:) ... sharp, fresh, terpene-like, somewhat perfumery-fruity citrus-type odor.	
linalool	Light and refreshing, floral-woody odor with a faintly citrusy note.	(l-linalool:) fragrant, light, aromatic

Supplementary Table S6-cont'd

ODOUR		
NAME	QUALITY	DESCRIPTORS
methanethiol	Odor of rotten cabbage, very diffusive and objectionable.	
methyl salicylate	Pungent-sweet, fruity-rooty odor with burning sensation at high concentration... the ester will obviously have very different odor/flavor descriptions in different countries.	minty (peppermint), fragrant, aromatic
nonyl aldehyde	Very powerful and diffusive fatty-floral, waxy odor of moderate tenacity. In proper dilution, the fatty notes become more pleasant, floral-waxy, more rosy and sweet, fresh as Neroli.	
nutmeg	(nutmeg oil:) ... a light, fresh, warm-spicy and aromatic odor, a distinctly terpeny topnote and a rich, sweet-spicy, warm bodynote.	
octyl acetate	Fruity, slightly fatty, waxy-floral odor with a discretely green, Apple-like or almost "weedy" note, while the undertones are mildly woody. Moderate to poor tenacity.	
octyl aldehyde	Powerful, and in undiluted state harsh-fatty, penetrating odor. In extreme dilution sweet, Orange-like, slightly fatty, Honey-like and of moderate to poor tenacity.	
orange	(orange oil, bitter:) The odor is very peculiar, fresh and yet "bitter" in the sense of "dry", but with a rich and lasting, seet undertone.	
pentadecalactone	Delicately animal, musky and sweet, extremely tenacious odor of outstanding uniformity.	
phenyl acetaldehyde	Peculiar oily-green, earthy-rooty, but mild and sweet odor of good tenacity.	
pyrazine	Pungent, sweet odor, in dilution floral with remote resemblance to Heliotrope. Very diffusive, poor tenacity.	
r-carvone	Warm-herbaceous, breadlike, penetrating and diffusive odor, somewhat spicy, in extreme dilution also floral, overall reminiscent of Spearmint oil (rectified).	(carvone:) minty (peppermint), fragrant, cool (cooling)
r-limonene	Fresh, light and sweet citrusy odor with strong resemblance to Orange peel oil	(limonene:) fruity (citrus), lemon, orange
sandalwood	(sandalwood oil, East Indian:) ... extremely soft, sweet-woody and almost animal-balsamic odor, presenting little or no particular topnote...	
spearmint	(spearmint oil:) ... very warm, slightly green-herbaceous odor, penetrating and powerful...	minty (peppermint), sweet, cool (cooling)
terpineol	...delicately floral and sweet of Lilac type...	fragrant, aromatic, disinfectant (carbolic)
terpinyl acetate	Mildly herbaceous, sweet and refreshing odor of spicy Bergamot-Lavender type, with variations into piney notes according to quality of the ester.	
undecanal	Pleasant waxy-floral, refreshing odor with a discrete fruity overtone and moderate tenacity.	
vanillin	Intensely sweet and very tenacious creamy Vanilla-like odor... A common remark from laymen, smelling highly diluted Vanillin is "Chocolate"...	vanilla, sweet, chocolate
paraffin oil		
propylene glycol	Odorless when absolutely pure.	woody (resinous), light, sour (vinegar)

Odour quality is taken from Arctander, S. (1969) *Perfume and Flavor Chemicals (Aroma Chemicals)* (Arctander, Montclair, N.J.) or Arctander, S. (1961) *Perfume and Flavor Materials of Natural Origin* (Arctander, Montclair, N.J.). The odour descriptors are from Dravnieks, A. (1985) *Atlas of Odor Character Profiles* (ASTM, Philadelphia, P.A.). The three most frequently used descriptors out of 146 are listed.

Figure S1

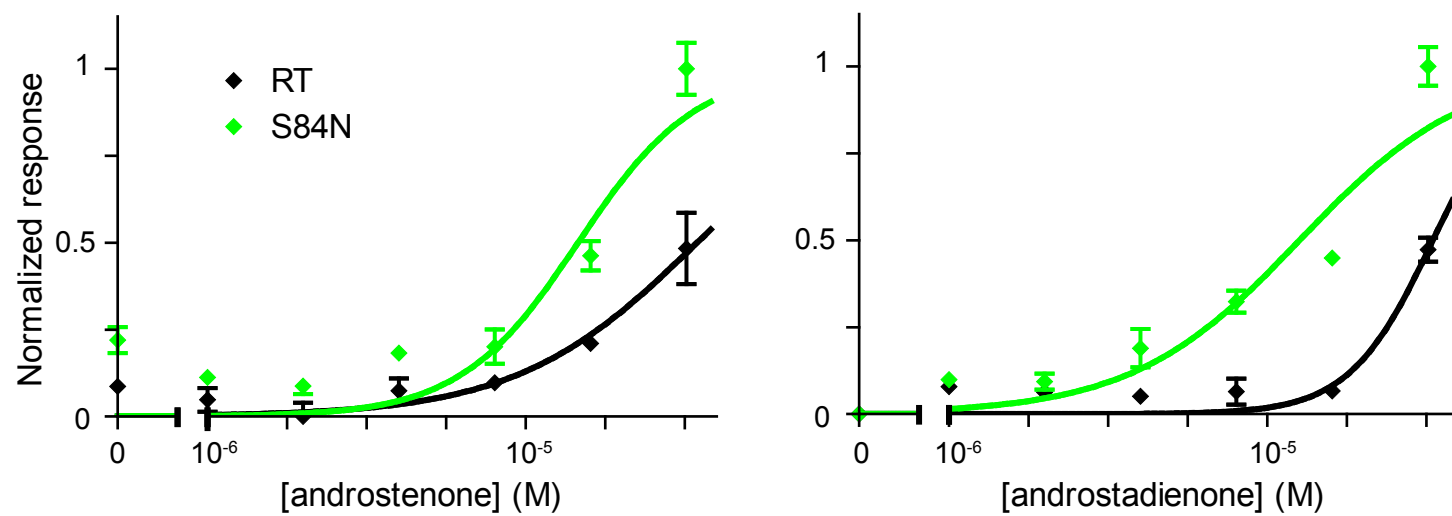
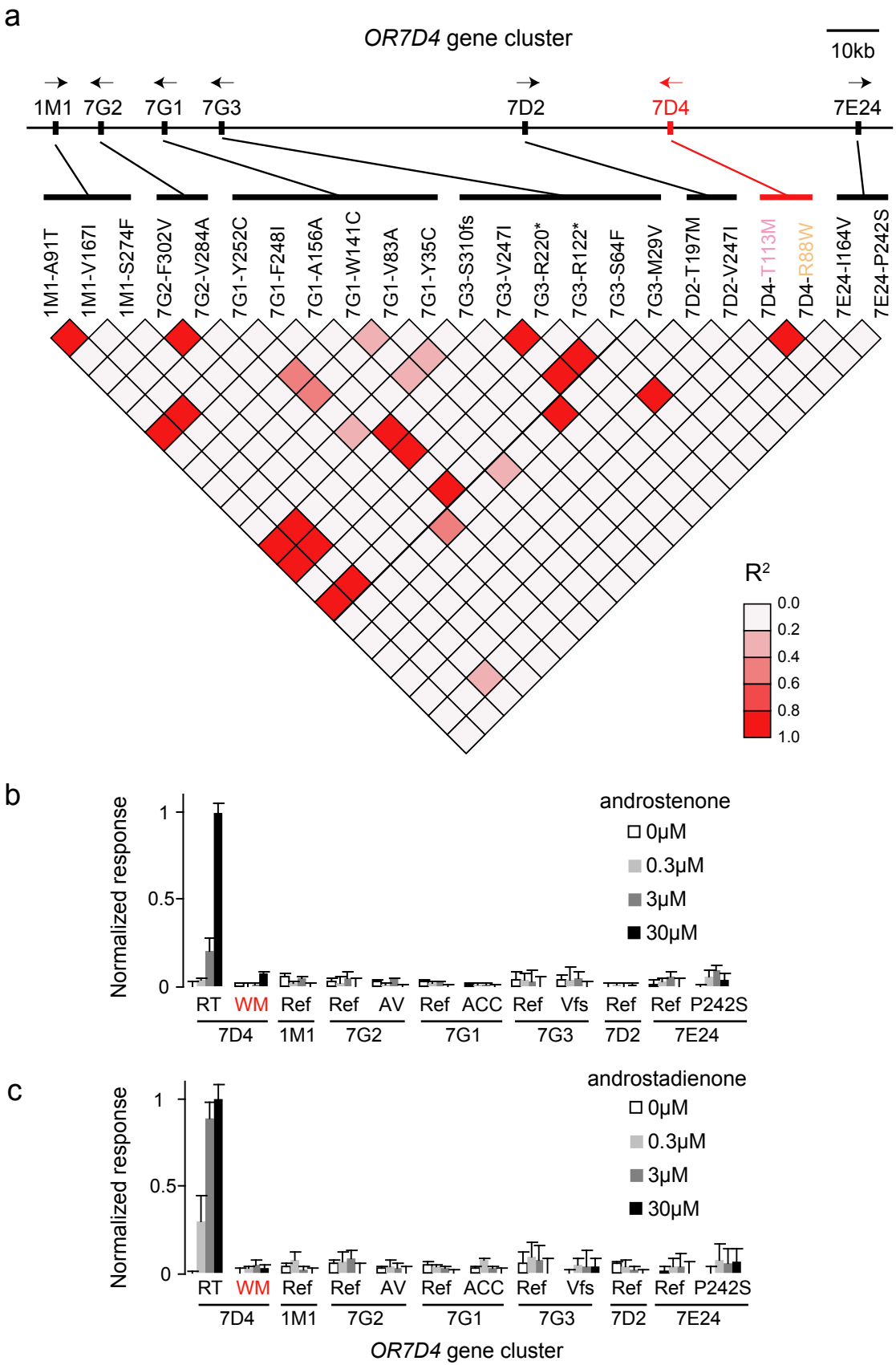
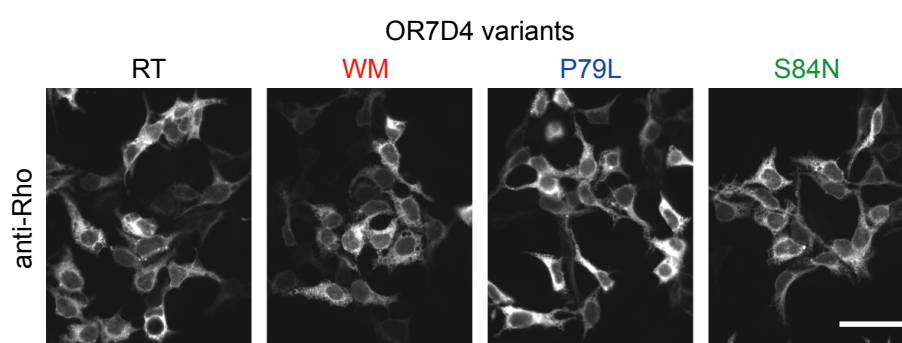
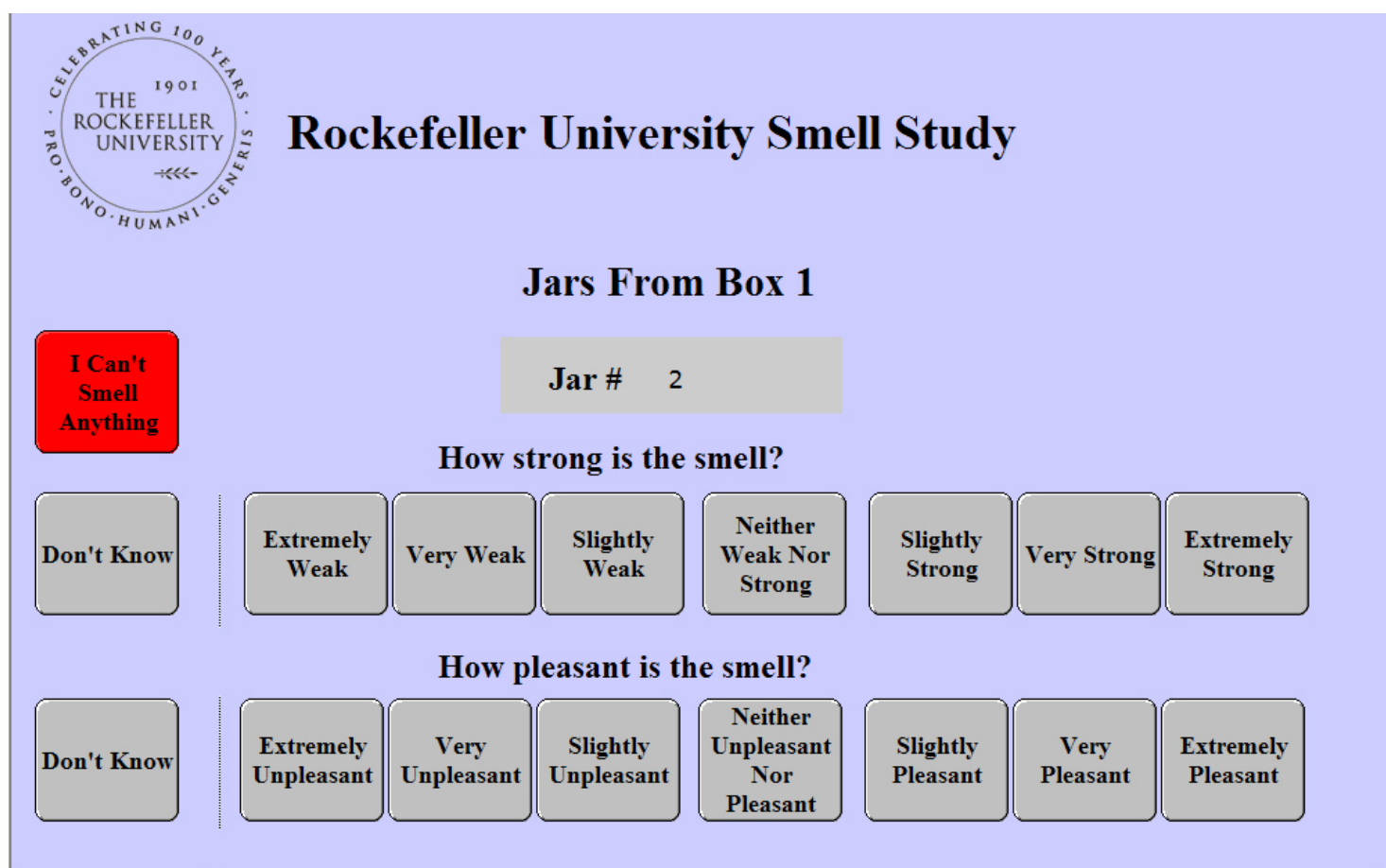


Figure S2







The image shows a screenshot of a web-based smell study interface. At the top left is the Rockefeller University logo, which is a circular seal with the text 'CELEBRATING 100 YEARS' at the top, '1901' in the center, 'THE ROCKEFELLER UNIVERSITY' below that, and 'PRO·BONO·HUMANI·GENE·RIS' at the bottom. To the right of the logo is the title 'Rockefeller University Smell Study' in a large, bold, black serif font. Below the title is the section header 'Jars From Box 1' in a bold, black serif font. Underneath this is a grey rectangular box containing the text 'Jar # 2'. To the left of the 'Jar # 2' box is a red rectangular button with the text 'I Can't Smell Anything' in white. Below the 'Jar # 2' box are two rows of response options. The first row is titled 'How strong is the smell?' and contains eight grey buttons with the following text from left to right: 'Don't Know', 'Extremely Weak', 'Very Weak', 'Slightly Weak', 'Neither Weak Nor Strong', 'Slightly Strong', 'Very Strong', and 'Extremely Strong'. The second row is titled 'How pleasant is the smell?' and contains eight grey buttons with the following text from left to right: 'Don't Know', 'Extremely Unpleasant', 'Very Unpleasant', 'Slightly Unpleasant', 'Neither Unpleasant Nor Pleasant', 'Slightly Pleasant', 'Very Pleasant', and 'Extremely Pleasant'. A vertical dotted line is positioned between the 'Don't Know' button and the first response button in each row.

Rockefeller University Smell Study

Jars From Box 1

I Can't Smell Anything

Jar # 2

How strong is the smell?

Don't Know | Extremely Weak | Very Weak | Slightly Weak | Neither Weak Nor Strong | Slightly Strong | Very Strong | Extremely Strong

How pleasant is the smell?

Don't Know | Extremely Unpleasant | Very Unpleasant | Slightly Unpleasant | Neither Unpleasant Nor Pleasant | Slightly Pleasant | Very Pleasant | Extremely Pleasant

Figure S5

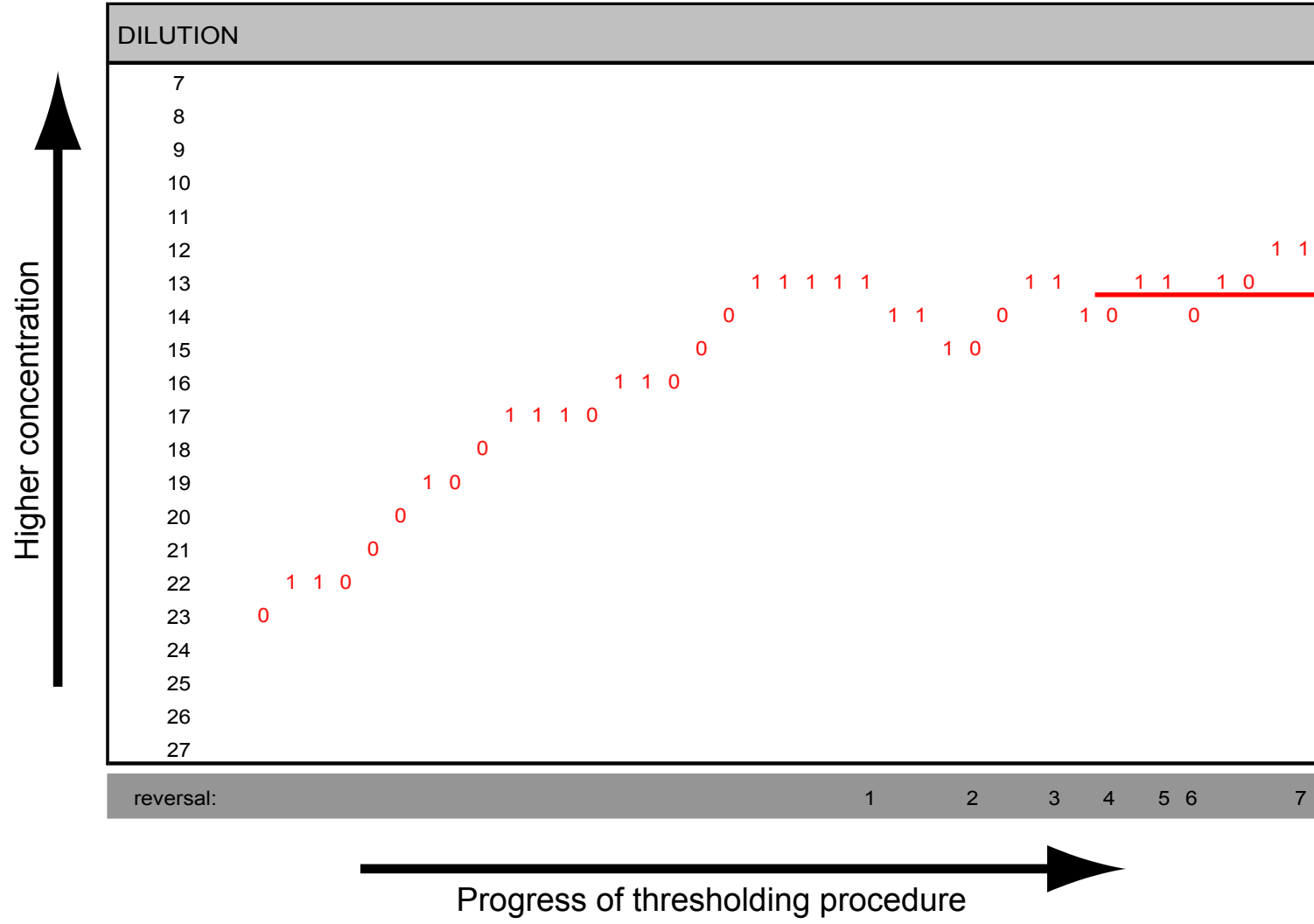


Figure S6

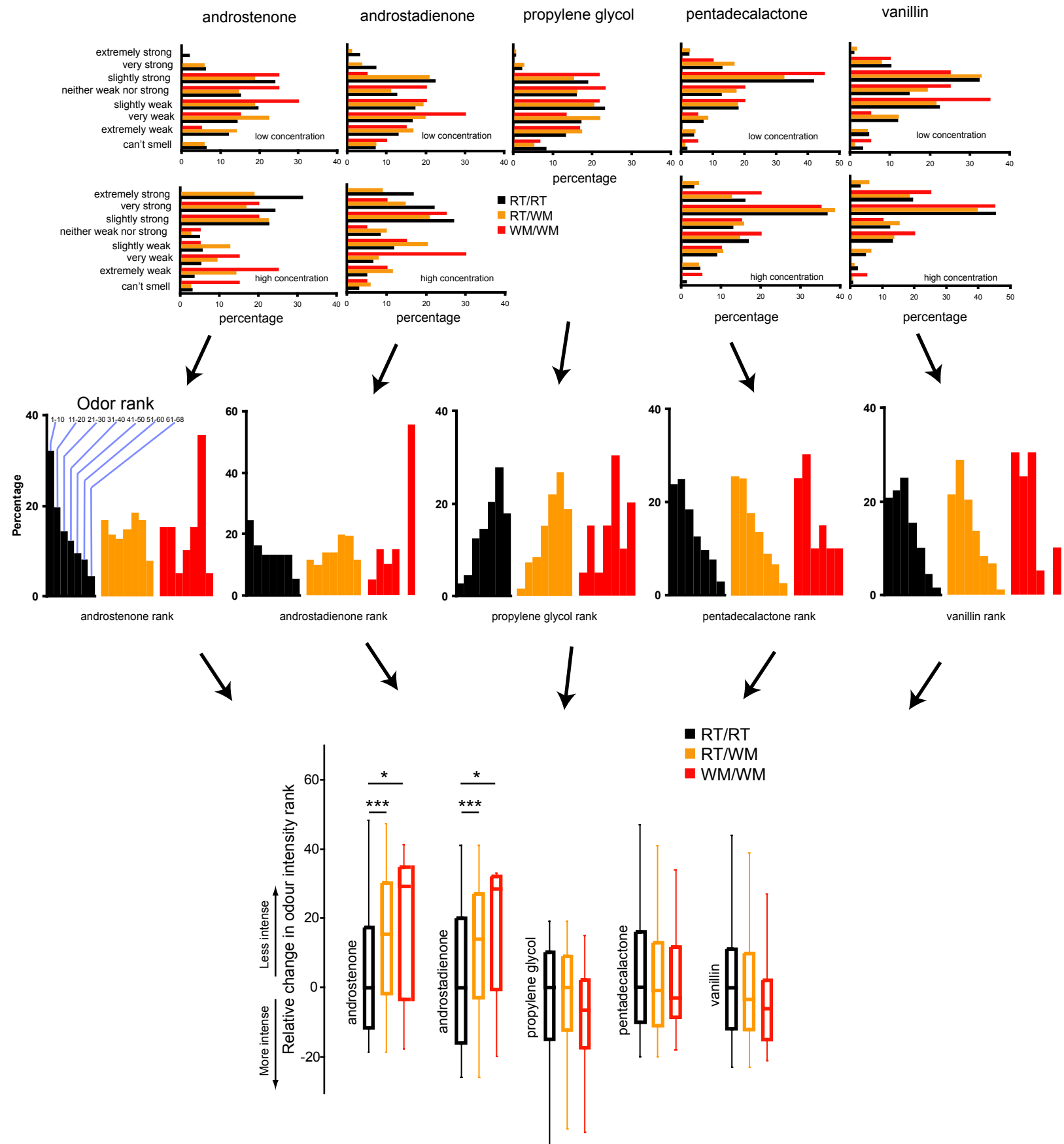


Figure S7

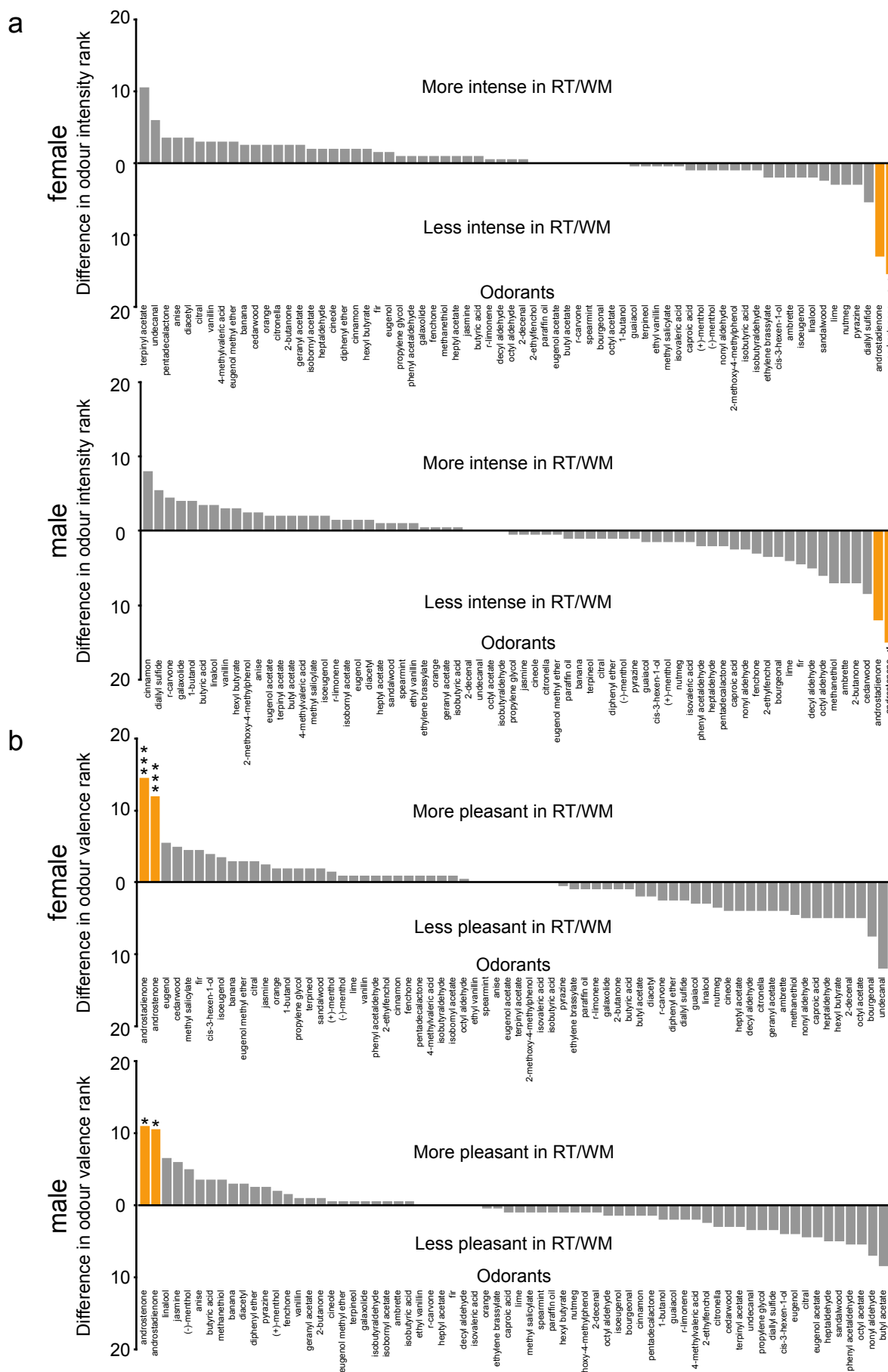


Figure S8

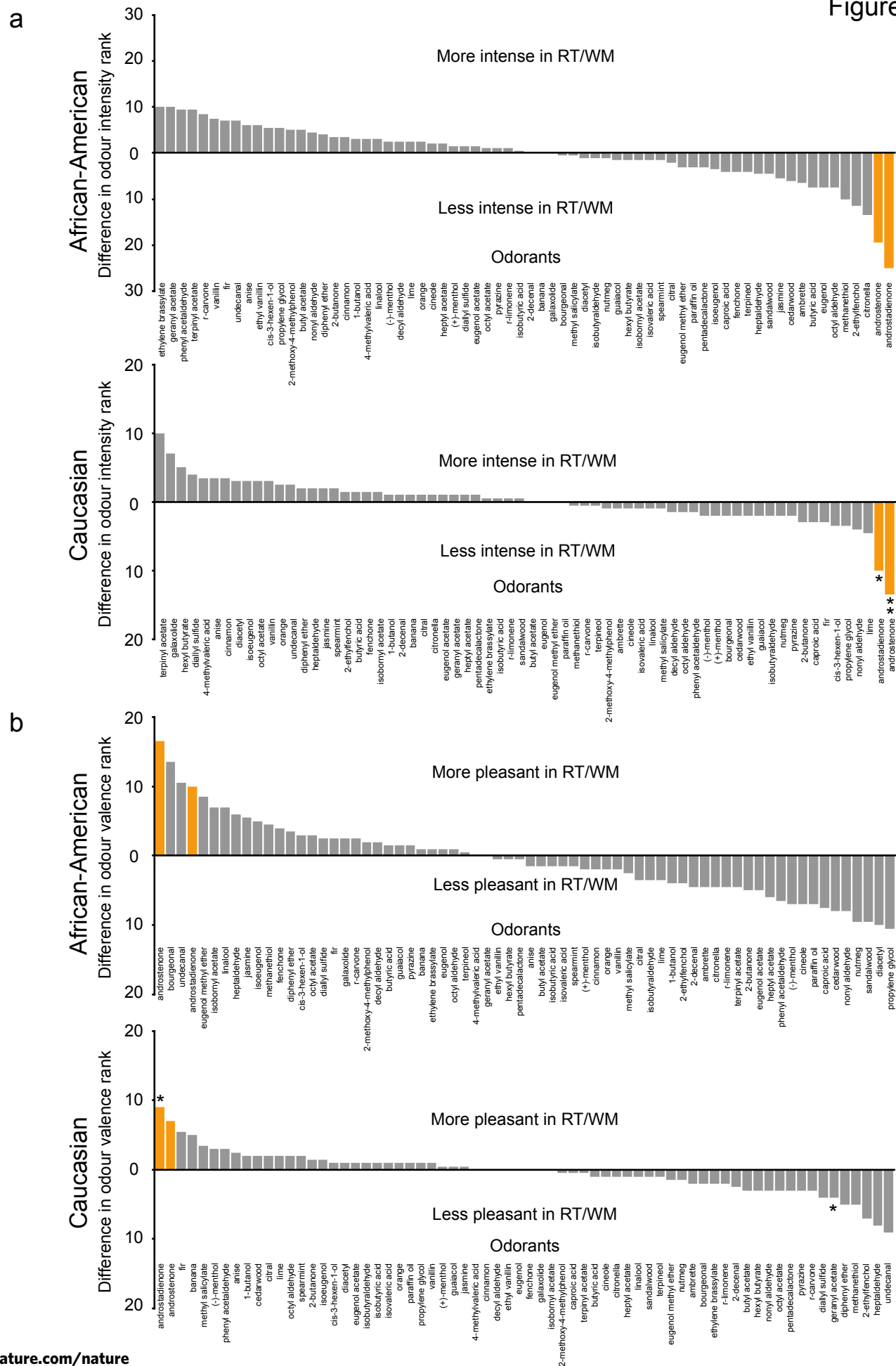


Figure S9

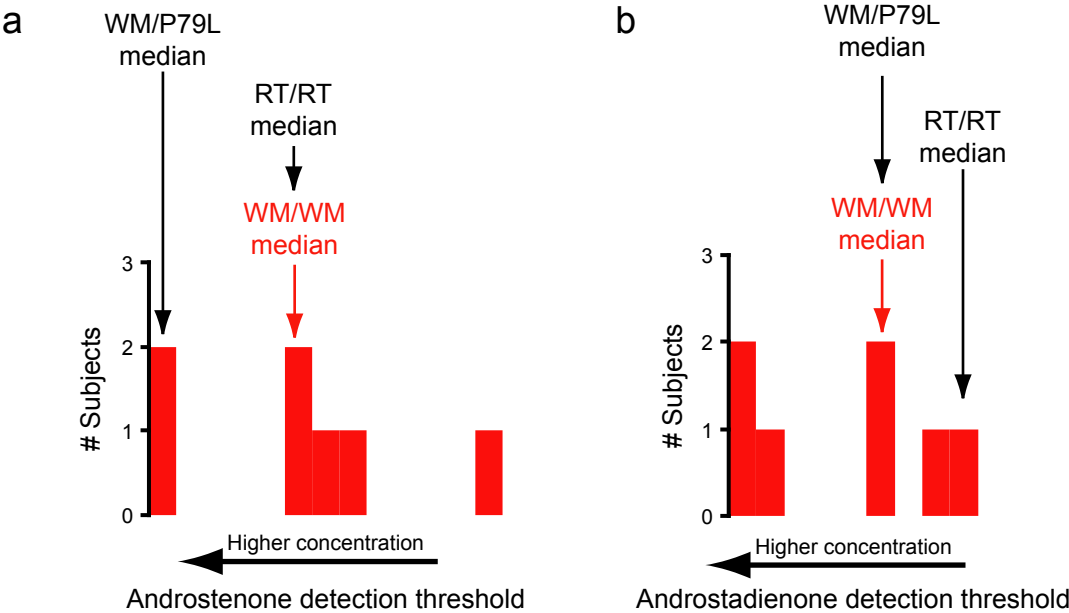


Figure S10

