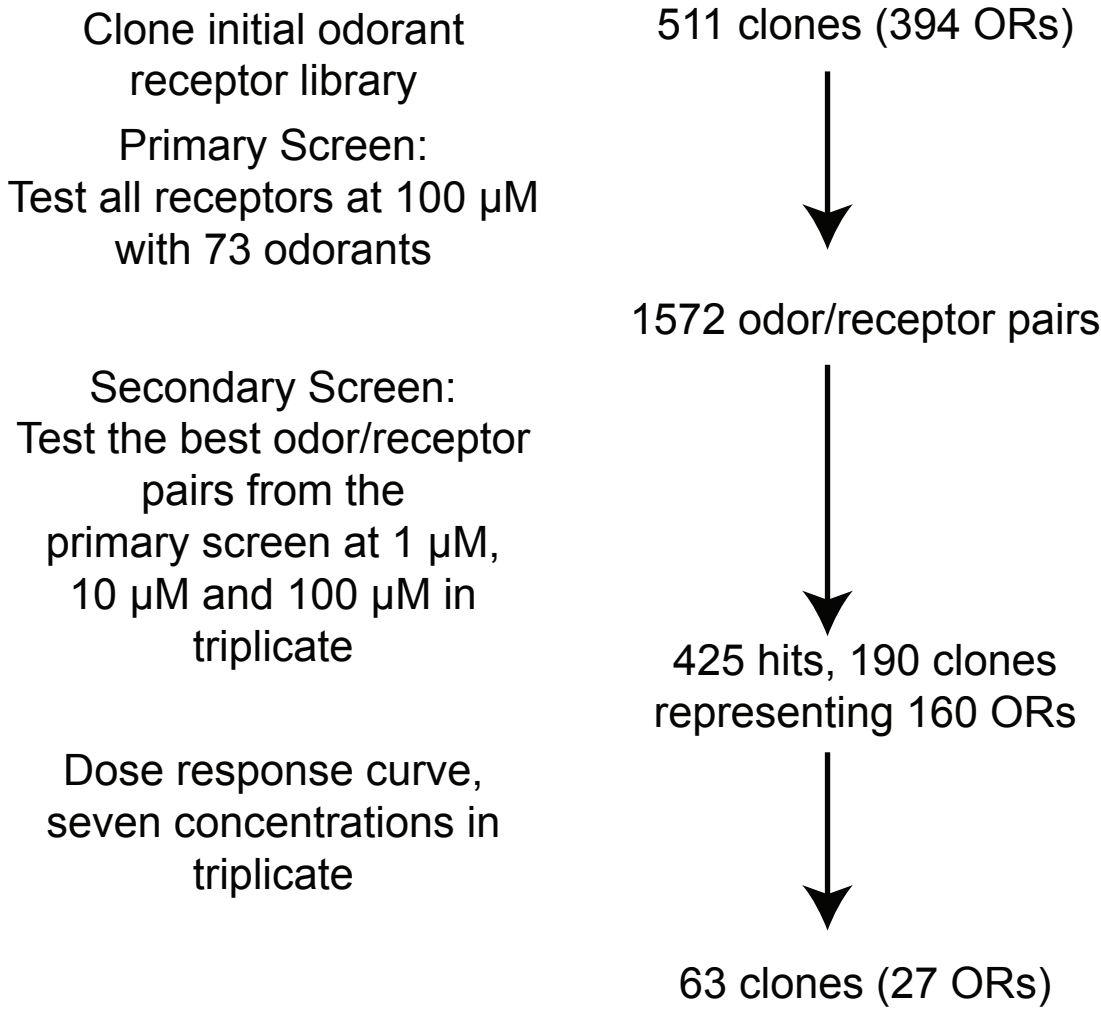
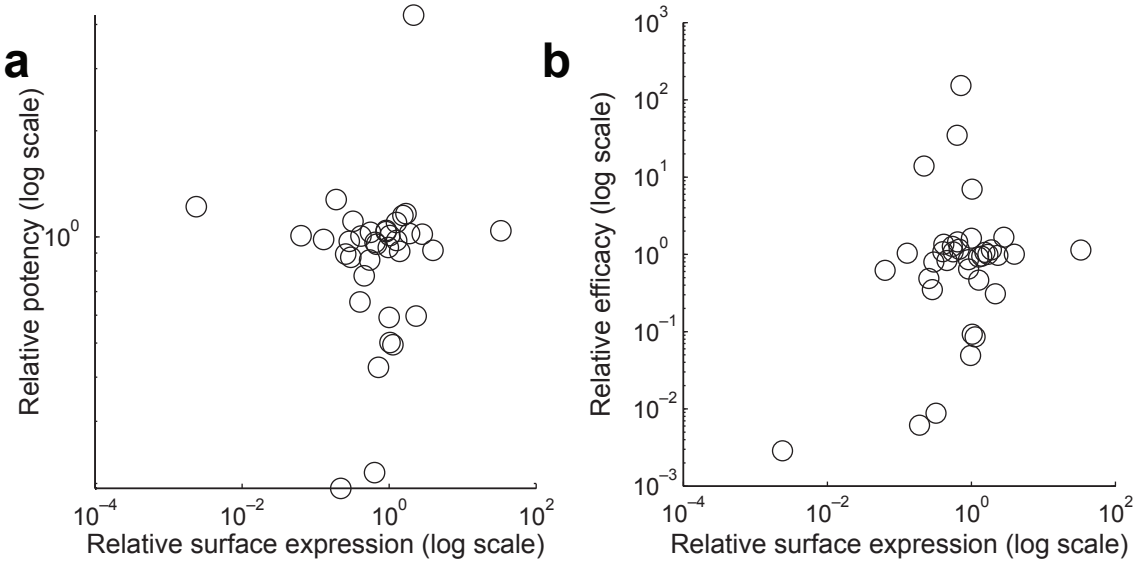
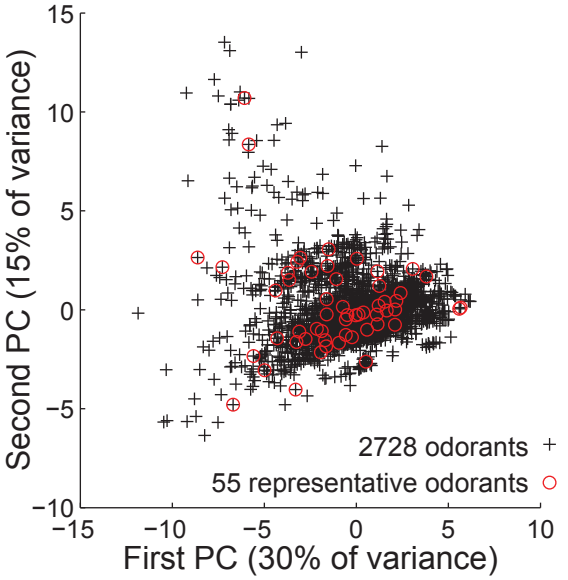
Screening procedure



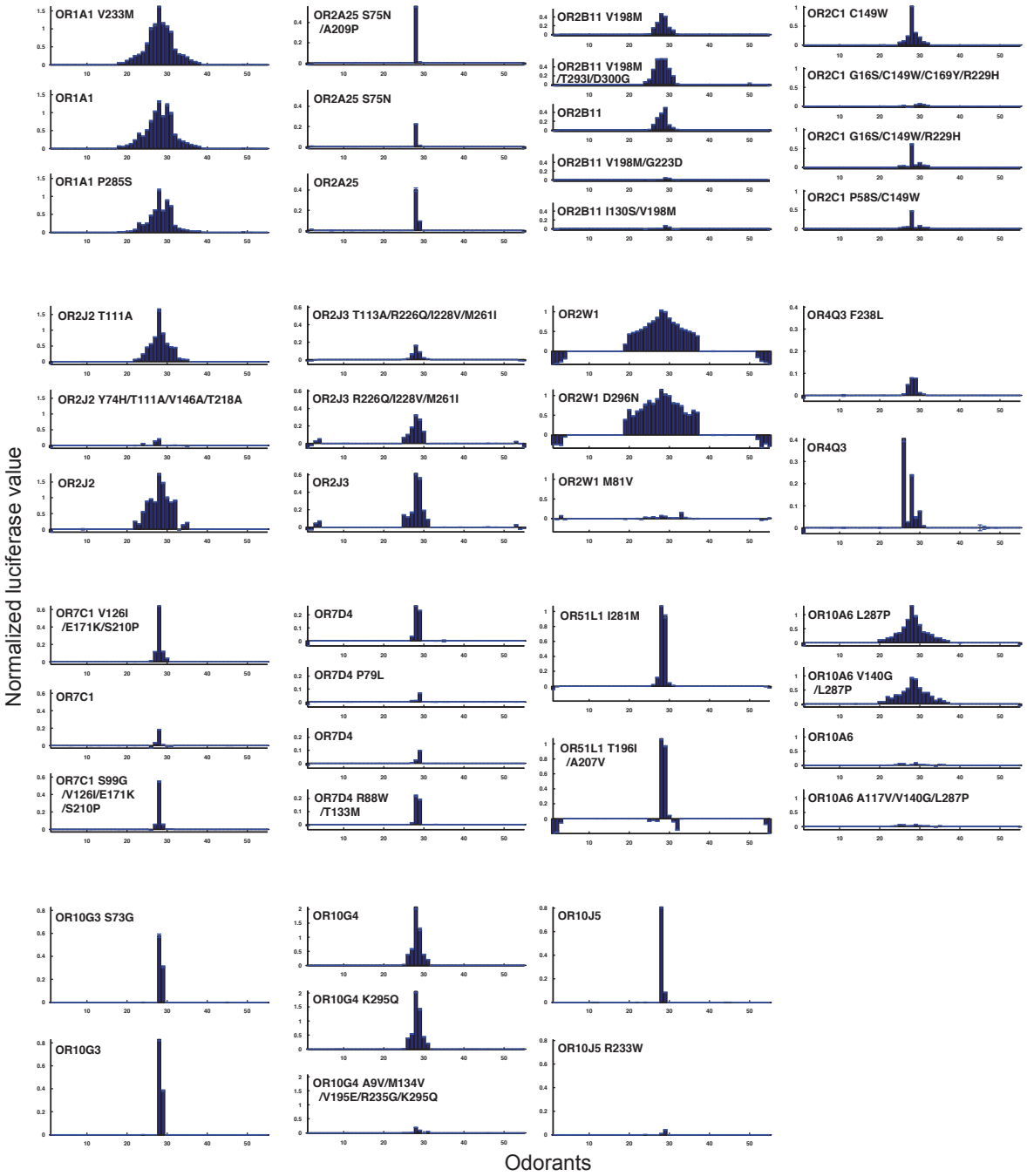
Supplementary Figure 1 Outline of the screening procedure.



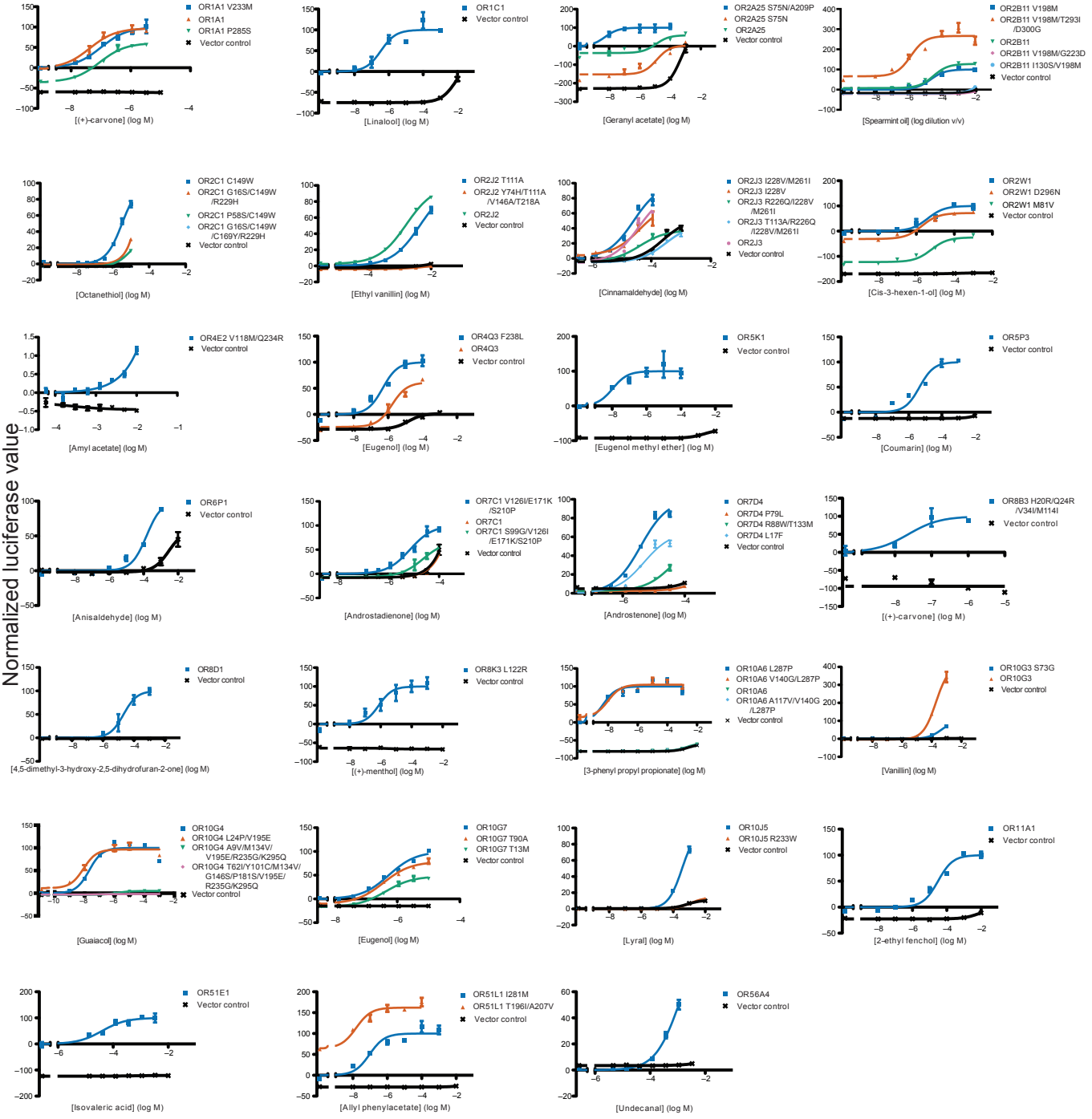
Supplementary Figure 2. Relative surface expression, measured by a fluorescence-activated cell sorter, does not correlate with relative potency (a) (Spearman rho=-0.01, n=38, p=0.95) or relative efficacy (b) (Spearman rho=0.16, n=38, p=0.33). For each variant we compared the change in mean surface expression (PE fluorescence intensity) to the change in the receptor potency (EC50 from the best-fit sigmoidal model) or receptor efficacy (top-bottom from the best-fit sigmoidal model) relative to the most common functional variant.



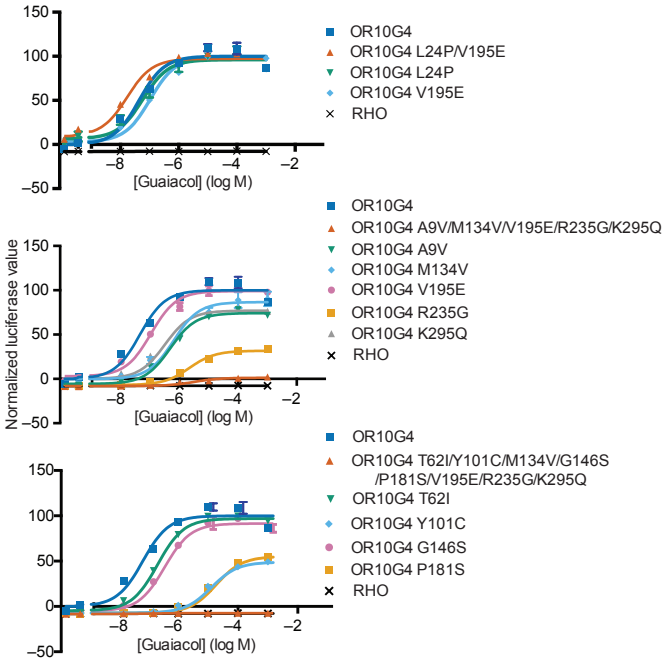
Supplementary Figure 3. We calculated 20 chemical descriptors, previously shown to explain more than 62% of the variance in functional responses in a heterologous system, for 2728 odorants 1. For display purposes, the odorants are projected onto a 2D space made of the first and second principal components. Black crosses represent all 2728 odorants, red circles represent the 55 odorants chosen to span olfactory space.



Supplementary Figure 4. Sensitivity-ordered tuning curves for odorant receptors. The y-axis represents the luciferase response normalized by Renilla Luciferase. The 55 odorants are displayed along the x-axis according to the strength of response they elicited from the reference allele of each receptor (defined as the most frequent allele that shows activity significantly higher than the vector control). If a given odorant did not significantly activate any of the variant receptors above the no-odor control (2-tailed t-test, $\alpha=0.05/55$), that odorant's response was set to zero across all variants. The odors eliciting the strongest response are placed at the center of the distribution. The order of odorants is the same across all variants of a given receptor, but is different across receptors. Error bars represent standard error.



Supplementary Figure 5. Dose response curves for odorant receptors. Y-axis values are normalized to the reference allele (shown in dark blue), here defined as the most frequent allele in the 1000 genomes population that shows activity significantly higher than the vector control. Error bars represent standard error.



Supplementary Figure 6 Functional characterization of OR10G4 polymorphisms. Y-axis values are normalized to the reference allele (shown in dark blue), here defined as the most frequent allele in the 1000 genomes population that shows activity significantly higher than the vector control. Error bars represent standard error.

CAS number	SMILES	Common name	Abbreviation
2244-16-8	CC(=C)[C@H]1CC=C(C(=O)C1)C	(+)-carvone	+CAR
15356-60-2	CC1CCC(C(C1)O)C(C)C	(+)-menthol	+MEN
121-14-2	[O-][N+](=O)c1ccc(c(c1)[N+](=O)[O-])C	2,4-DNT	24DNT
821-55-6	CCCCCCCC(=O)C	2-nonanone	29ONE
18368-91-7	CCC1(O)C2(C)CCC(C1(C)C)C2	2-ethylfenchol	2EF
15679-19-3	CCOc1nccs1	2-ethoxythiazole	2ET
3149-28-8	COc1cnccn1	2-Methoxypyrazine	2MPYR
122-74-7	CCC(=O)OCCCc1ccccc1	3-phenyl propyl propionate	3PPP
1076-38-6	Oc1cc(=O)c2c(o1)cccc2	4-Hydroxycoumarin	4HC
628-63-7	CCCCCOC(=O)C	N-amyl acetate	AA
18339-16-7	O=C1CC[C@]2([C@H](C1)CC[C@@H]1[C@@H]2CC[C@]2([C@H]1C	androstenone	AND
4075-07-4	O=C1CCC2(C(=C1)CCC1C2CCC2(C1CC=C2)C)C	androstadienone	ANDI
123-11-5	COc1ccc(cc1)C=O	anisaldehyde	ANIS
1797-74-6	C=CCOC(=O)Cc1ccccc1	allyl phenyl acetate	APA
7756-96-9	CCCCOC(=O)c1ccccc1N	Butyl anthranilate	BA
7492-70-8	CCCCOC(=O)C(OC(=O)CCC)C	butyl butyryl lactate	BBL
23726-91-2	C/C=C/C(=O)C1=C(C)CCCC1(C)C	beta-damascone	BDAM
24330-52-7	SCC[C@H](S)C	1,3-butane dithiol	BDT
71-43-2	c1ccccc1	benzene	BEN
592-84-7	CCCCOC=O	butyl formate	BF
928-96-1	OCCC=CCC	cis-3-hexen-1-ol	C3HEX
14371-10-9	O=C/C=C/c1ccccc1	Cinnamaldehyde	CINMA
4437-20-1	S(Cc1ccco1)SCc1ccco1	coffee difuran	COFFDF
91-64-5	O=c1ccc2c(o1)cccc2	Coumarin	COUM
112-31-2	CCCCCCCCCC=O	decyl aldehyde	DA
334-48-5	CCCCCCCCCC(=O)O	decanoic acid	DECAC
10022-28-3	CCCCCCCC(OC)OC	1,1-dimethoxy-octane	DM8
28664-35-9	CC1OC(=O)C(=C1C)O	caramel furanone	DMHDMF
75-18-3	CSC	dimethyl sulfide	DMS
105-95-3	O=C1CCCCCCCCCCCCC(=O)OCCO1	ethylene brassylate	EB
97-53-0	C=CCc1ccc(c(c1)OC)O	eugenol	EUG
93-15-2	C=CCc1ccc(c(c1)OC)OC	eugenol methyl ether	EUGME
121-32-4	CCOc1cc(C=O)ccc1O	ethyl vanillin	EVAN
105-87-3	C/C(=C\COC(=O)C)/CCC=C(C)C	geranyl acetate	GA
90-05-1	COC1=CC=CC=C1O	guaiacol	GUAIA
78-81-9	NCC(C)C	isobutyl amine	IA
513-44-0	SCC(C)C	isobutyl mercaptan	IM
2035-99-6	CCCCCCCC(=O)OCCC(C)C	isoamyl octanoate	IO
503-74-2	CC(CC(=O)O)C	isovaleric acid	IVA
5328-37-0	OCC(C(C(C=O)O)O)O	laevo-arabinose	LA
78-70-6	C=CC(CCC=C(C)C)(O)C	linalool	LIN
31906-04-4	O=CC1CCC(=CC1)CCCC(O)(C)C	Lyral	LYR
57500-00-2	CSSCc1ccco1	methyl furfuryl disulfide	MFUR
554-12-1	CCC(=O)OC	Methyl propionate	MP
2345-24-6	C/C(=C/COC(=O)C(C)C)/CCC=C(C)C	neryl isobutyrate	NI
2306-88-9	CCCCCCCCOC(=O)CCCCCCC	octyl octanoate	OO
111-88-6	CCCCCCCCS	octanethiol	OTHI
122-78-1	O=CCc1ccccc1	phenyl acetaldehyde	PA
4286-15-1	CCC(c1ccccc1)C(=O)O	(S)-(+)-2-Phenylbutyric acid	PBA
290-37-9	n1ccncc1	Pyrazine	PYR
91-22-5	c1ccc2c(c1)nccc2	quinoline	QUIN
5989-27-5	CC1=CC[C@@H](CC1)C(=C)C	r-limonene	RLIM
15707-24-1	CCc1nccnc1CC	shoyu pyrazine	SP
4023-65-8	OC(=O)C/C(=C\C(=O)O)/C(=O)O	trans-Aconitic acid	TAA
98-55-5	CC1=CC[C@@H](CC1)C(O)(C)C	terpineol	TERP
80-26-2	CC(=O)OC(C1CCC(=CC1)C)(C)C	terpinyl acetate	TERPA
68-11-1	OC(=O)CS	thioglycolic acid	THIOA
60633-24-1	CC1SC(C(=N1)C)C	TMT	TMT
112-44-7	CCCCCCCCCCC=O	undecanal	UNDEC
121-33-5	COc1cc(C=O)ccc1O	vanillin	VAN

Supplementary Table 1 Odorants used in the study. Odorants are listed with their Chemical Abstract Service (CAS) registry number, simplified molecular-input line-entry system (SMILES) identifier, common name, and abbreviation used in the figures.