SI Appendix for:

Perceptual Convergence of Large Mixtures in Olfaction Implies an Olfactory White

Tali Weiss, Kobi Snitz, Adi Yablonka, Rehan M Khan, Danyel Gafsou, Elad Schneidman, Noam Sobel

Content:

Equated-intensity odorants

All odorants were purchased from Aldrich Chemicals (St. Louis, MO) in the highest available purity. All odorants were diluted with either mineral oil, 1,2-propanediol or deionized distilled water to a point of approximately equally perceived-intensity. This perceived-intensity equation was conducted according to previously published methods. In brief, we identified the odorant with lowest perceived intensity, and first diluted all others to equal perceived intensity as estimated by experienced lab members (1). Next, 24 naïve subjects (10 women) smelled the odorants and rated their intensity. We then further diluted any odorant that was 2 or more standard deviations away from the mean intensity of the series, and repeated the process until we had no outliers. This process is suboptimal, but considering the natural variability in intensity perception, together with naive subjects' bias to identify “a difference”, and

the iterative nature of this procedure, any stricter criteria would generate an endless process.

The odorants used and their final dilutions are listed in Table S1.

Selecting components for mixtures in the similarity experiments

We selected mixtures of odorants out of the initial pool of odorants in a pseudo-random manner, which we applied to two separate databases. The first database consists of 144 single molecules, each having 146 perceptual descriptors as in Dravnieks' Atlas (2, 3), we refer to this database as the "odor perceptual space". The second database contains 1492 odorants commonly used in olfaction research (1). For each of the 1492 odorants we obtained 1438 physicochemical descriptors using Dragon software (Talete, Milan, Italy) (4). The values of the 1438 physicochemical descriptors are on vastly different scales. Thus, to eliminate this source of variance, the values of each property were z-score. We refer to this data as the "odor physiochemical space". Each odorant was represented by a vector whose coordinates describe its 146 perceptual or 1438 physicochemical properties. Our initial pool of 86 odorants for which the perceived intensity equation process was applied is contained in both the perceptual and physicochemical odor space. The selection process of the different mixture sizes from our collection of odorants was as follows. The distance between an odorant x and a mixture G was defined as the minimum Euclidean distance between the odorant vector and the vector of one of the mixture's components: $D(x,G) = min_{g \in G} d(x,g)$ where d(x,g) is the Euclidean distance between the vectors x and g. For example, if the coordinates of the odorants in a mixture $G = \{g : g \in G\}$ of a hypothetical database were ([4,-6,8], [3,-3,10], [-4,0,3]),

then the distance of an odorant x with coordinates [-2,1,1] to that mixture would be 3, because that is the distance to the closest point: [-4,0,3].

Next, we defined a measure T(G) of how well a collection of odorants is spread out over the database of odorants. T(G) is the average distance between the mixture and all of the odorants in the database (144 in the perceptual space and 1492 in the physicochemical space) :

$$T(G) = \frac{\sum D(x,G)}{|X|}$$

Our initial experiment required that we find two disjoint collections of 40 odorants, which are well distributed over the database. From a pool of 80 odorants (picked randomly out of the 86), 10,000 random samples of 40 were selected, and for each sample G, we calculated $T(G)+T(G^C)$, where G^C represents the complementary set of G. We selected the collection G which gave the lowest sum of $T(G)+T(G^C)$ and labeled it 40A and its complement G^C as 40B. Next, we took another 10,000 samples of 30 odorants out of 40A and picked the one sample G such that T(G) was minimal and labeled it 30A. In a similar manner we picked 20A, 10A, and 4A. 1A was picked randomly out of the odorants in 40A. We repeated the procedure to create the subsets 30B, 20B, 10B, 4B, and 1B, out of 40B. The mixtures described above constitute one set that was used to conduct an experiment with a group of ~14 subjects. A total of four different sets of mixtures were prepared using this pseudo-random process, each created for a separate experiment. Two of the sets, which will be referred to as set-1 and set-2, were constructed based on the perceptual space (144 vectors with 146 coefficients). The two other sets, set-3 and set-4, were selected based on the physicochemical space (1492 vectors with 1438 coefficients). In set-4, the largest mixture was composed of 43 components rather than 40, out of the full pool of 86 odorants, and the resulting mixture sizes in this set were 43, 30, 20, 15, 10, 4 and 1.

Pairwise similarity tests

In each experiment, each subject was presented with 36 comparisons (39 in the fourth experiment) of non-overlapping mixtures and was asked to rate their similarity on a visual-analogue scale ranging from 1 to 9, with 9 corresponding to most similar. For assessing the subjects' reliability, we added comparisons of the same mixture (optimally should be rated as "9"). In total, 147 comparisons were between non-overlapping mixtures and additions of 44 comparisons were of mixtures to themselves. The ratings were normalized to a scale of 0 to 100% based on the range of each subject's ratings. Each subject repeated the experiment on two different days to assess test-retest reliability. If the difference between 2 repetitions of the same comparison was greater than 70%, this rating was excluded. This amounted to 109 out of 2070 ratings (~5%). The ratings by subjects whose similarity ratings for identical mixtures were poorer by at least 2 standard deviations from the mean were discarded. This amounted to 3 subjects, retaining in the analysis: 12, 13, 13 and 18 subjects for the four experiments. The average rated similarities were calculated across subjects.

Three-alternative forced-choice discrimination test

Twenty subjects (10 women) were presented with odorant triplets. Two of the jars in each triplet contained the same mixture, and the third was different. Subjects were asked to pick out the jar that contained the dissimilar smell and rate their confidence level on a scale of 1 to 5, with 5 corresponding to "most certain". The subjects were allowed to smell each of the three jars twice. In all four experiments (5 subjects in each), the discrimination was between a 35-component mixture (different versions across experiments), and non-overlapping mixtures of different component sizes which spanned perceptual space (2, 3, 4, 5, 7, 8, 10, 14, 15, 17, 20, 25, and 30-

components). Each subject repeated the experiment on two different days to assess test-retest reliability. On one of the days, the aberrant jar was the 35-component mixture (e.g., one jar with the 35-component mixture, and two jars each containing a 2-component mixture), and on the other day the aberrant jar was the non-35-component mixture (e.g., two jars containing the 35-component mixture, and one jar containing a 2-component mixture) (counterbalanced across days). Correct discrimination scored 100 and incorrect discrimination scored 0. The scores for each mixture component number were averaged across the four experiments (Fig. 2B). The confidence level ratings of subjects who answered correctly in less than 70% of the discrimination trials were discarded, retaining 16 out of 20 subjects for this analysis. The results were normalized to a scale of 0 to 100% (Fig. 2C), with 100% corresponding to "most certain".

Identification experiments

After a three-day acquaintance with 40-component Laurax, 12 subjects performed a four-alternative forced-choice identification task for 23 different novel odorant mixtures of 1-, 4-, 10-, 20-, 30- components (4 versions each), and 40-components (3 versions), all selected to span physicochemical space. 40-component Laurax was also used in order to verify that subjects remember its smell. Each mixture was repeated twice, adding up to 48 trials in this task. Subjects were asked to choose the most appropriate label for each mixture out of four options: one being "Laurax", and the other three were labels provided by a professional perfumer. In the second experiment, after two days of acquaintance with 40-component Laurax, 13 different subjects had to choose the appropriate label for the same mixtures out of five options, four the same as the above, and the fifth option being "other". Choosing "Laurax"

scored 100 and choosing any of the other names scored 0. The number of times "Laurax" was chosen was averaged across four versions of each number of components in a mixture (three versions in the case of 40 component mixture). Two out of the 25 subjects were excluded because they did not identify the very same mixture they smelled in the two previous days as "Laurax". Thus, the number of subjects for further analysis was 11 and 12, respectively.

Delayed match-to-sample

After a two-day acquaintance with 30-component Laurax, 12 subjects were asked to decide if the smell in each jar was Laurax, by pressing "yes" or "no". For analysis, "yes" was scored 100, and "no" was scored 0. In total, each subject had 69 trials: 21 target mixtures of various component sizes (30, 25, 20, 15, 10, 5 and 1; three versions each; 3 repetitions) were selected out of the remaining 56 odorants (i.e., these mixtures shared no components in common with the Laurax mixture). To verify that subjects remember the smell of 30-component Laurax, it was given in two concentrations (high and low); each had 3 repetitions. All subjects correctly identified Laurax (100%), with the exception of one subject who correctly identified it at 66.6%. In the second delayed match-to-sample test, 16 subjects were acquainted with a 25-component Laurax for two days (there were two versions of Laurax, 8 subjects were assigned to each). In the test day, subjects had 35 trials in which they rated whether the odor was Laurax using a VAS ranging between "yes" and "no". 16 mixtures (25 components each) were selected out of the remaining 103 molecules (pool of 144 molecules, Table S1A, S1B): 1. 'Spanned'- nine mixtures whose components were spanned in physicochemical space; 2.'not-spanned'- five mixtures whose components were clustered in this space ; 3. 'not-equated' - two mixtures that had the same

components as the Laurax, however, one single molecule provided 50% of the total mixture solution volume: in the first mixture this molecule was isopropyl alcohol (CAS 67-63-0) and in the second mixture, this molecule was ethyl butyrate (CAS 105-54-4). Each mixture was repeated twice. In addition, the 25-component Laurax was used in order to verify that subjects remember its smell (3 repetitions). After normalization of ratings according to the minimum and maximum values of each subject, the median and the median absolute deviation (MAD) were calculated between subjects. When denoting answers above 75% as "yes" (score 100), and below 25% as "no" (score 0), the average and standard deviation were calculated between subjects. The same paradigm and analysis were used in an additional delayed match-to-sample task when the mixtures were of 15 components (9 subjects were assigned to each version).

In the third delayed match-to-sample test, after a two-day acquaintance with a 60-component Laurax (Supplementary Table 4, 60A), 14 subjects had 33 trials: 15 mixtures of different sizes (1, 4, 10, 20, 30, 40, 50; two version for each, and one 60 component mixture). Those mixtures were selected out of the remaining 84 molecules (pool of 144 molecules, Table S1A, S1B), each mixture was repeated twice. 60-component Laurax was also used in order to verify that subjects remember its smell (3 repetitions). Two subjects were excluded from further analysis, one did not identify the Laurax, and the second did not use the VAS scale. After normalization of the ratings according to the minimum and maximum values of each subject, the median and the median absolute deviation (MAD) were calculated among subjects.

Supplementary Figures and Tables:

Supplementary Figure 1| Rated similarity as a function of number of components in a mixture (mixture components spread in physicochemical space)

A-H. Pearson correlation of the normalized average rated similarities between 1, 4, 10, 15, 20, 30 and 40/43-component mixtures with a nonoverlapping mixture consisting of 43 components (A), 40 components (B), 30 components (three), 20 components (D), 15 components (E), 10 components (F), 4 components (G), and 1 component (H). Each point represents the average rated similarity of one pairwise comparison. These data reflect the comparisons from odorant sets 3 (pink) and 4 (blue) of Supplementary Table 7, in which the mixtures were selected according to their components physicochemical spread. Error bars indicate SE among participants.

Supplementary Figure 2| Rated similarity as a function of number of components in a mixture (mixture components spread in perceptual space)

A-F. Pearson correlation of the normalized average rated similarities between 1, 4, 10, 20, 30 and 40-component mixtures with a nonoverlapping mixture consisting of 40 components (A), 30 components (B), 20 components (C), 10 components (D), 4 components (E), and 1 component (F). Each point represents the average rated similarity of one pairwise comparison. These data reflect the comparisons from odorant sets 1 (green) and 2 (black) of Supplementary Table 7, in which the mixtures were selected according to their components perceptual spread. Error bars indicate SE among participants.

Supplementary Figure 3| Identification after six months

The labels applied to Nerol, Citronellol, Geraniol, β-phenylethyl alcohol (PEA), 'rose' mixture and five versions of Laurax. The averaged response of 10 participants (in percentage) in (A) seven-alternative and (B) four-alternative forced-choice identification test.

Supplementary Figure 4| Natural odors form clusters in olfactory space

The T(G) value of Laurax (average of 30, 40 and 60 component mixtures which were used in the identification and delayed match-to-sample experiments) (blue), 63 components of rose (red)(5) 35 components of coffee (black)(6), and 37 components of apple (green)(7) (we used the components for which we could extract physicochemical properties using Pubchem). We ran 10,000 iterations to generate random mixtures of the appropriate component size (i.e., 30, 40, 60, 63, 35, and 37). The histogram represents all 60,000 T(G) values (z-score). Smaller values of T(G) represent a better spanning of the physicochemical space. As clearly evident, whereas natural odors form an extreme cluster in this physicochemical space, Laurax reflects an average spread in this space.

Supplementary Table 1A: List of odorants and concentrations. On the first column are the ordinal numbers from *Dravnieks' Atlas of Odor Character Profiles* (2, 3). On the sixth column are the odorants phase: L= liquid, S= solid.

	Name	CID Number	CAS Number	v/v % or w(g)/v(ml) %	Phase	Solvent
1	abhexone	61199	698-10-2	0.01	S	water
2	acetophenone	7410	98-86-2	0.15	L	mineral oil
3	ortho-acetyl pyridine	14286	1122-62-9	0.001	L	1,2-propanediol
5	strawberry aldehyde	6501	77-83-8	1	L	1,2-propanediol
7	gamma-nonalactone	7710	104-61-0	0.5	L	mineral oil
9	isoamyl acetate	31276	123-92-2	0.1	L	mineral oil
10	amyl butyrate	10890	540-18-1	1	L	1,2-propanediol
12	iso-pentyl phenyl acetate	7600	102-19-2	0.5	L	mineral oil
13	pentyl valerate	62433	2173-56-0	50	L	mineral oil
16	anisole	7519	100-66-3	0.2	L	1,2-propanediol
18	benzaldehyde	240	100-52-7	0.25	L	mineral oil
20	iso-bornyl acetate	93009	5655-61-8	5	L	mineral oil
21	butanoic acid	264	107-92-6	1	L	water
23	butyl sulfide	11002	544-40-1	0.15	L	mineral oil
27	caryophyllene	5281515	87-44-5	15	L	1,2-propanediol
29	celeriax	6259976	17369-59-4	0.2	L	1,2-propanediol
30	chlorothymol	6982	89-68-9	**100**	S	1,2-propanediol
31	cinnamic aldehyde	307	104-55-2	0.01	L	1,2-propanediol
34	coumarin	323	91-64-5	10	S	mineral oil
36	p-cresol	2879	106-44-5	0.1	S	1,2-propanediol
37	p-cresyl acetate	8797	140-39-6	0.03	L	mineral oil
38	p-cresyl-iso-butyrate	7685	103-93-5	10	L	mineral oil
39	4-methyl anisole	7731	104-93-8	0.15	L	mineral oil
40	cuminic aldehyde	326	122-03-2	1	L	mineral oil
44	cyclohexanol	7966	108-93-0	0.4	L	mineral oil
47	2,4-trans-trans-decadienal	5283349	25152-84-5	0.5	L	mineral oil
49	dibutyl amine	8148	111-92-2	3	L	1,2-propanediol
50	diethyl sulfide	9609	352-93-2	0.01	L	1,2-propanediol
51	dimethyl benzyl carbinyl butyrate	24915	10094-34-5	20	L	1,2-propanediol
52	muguet carbinol	7632	103-05-9	**100**	S	1,2-propanediol
53	2,3-dimethyl pyrazine	22201	5910-89-4	0.2	L	1,2-propanediol
54	2,5-dimethyl pyrazine	31252	123-32-0	0.3	L	mineral oil
56	dimethyl trisulfide	19310	3658-80-8	0.001	L	1,2-propanediol
58	diphenyl oxide	7583	101-84-8	1	S	mineral oil

59	ethyl butyrate	7762	105-54-4	0.01	L	mineral oil
60	ethyl propionate	7749	105-37-3	0.25	L	mineral oil
61	2-ethyl pyrazine	26331	13925-00-3	0.4	L	mineral oil
63	eucalyptol	2758	470-82-6	2	L	mineral oil
64	eugenol	3314	97-53-0	0.3	L	mineral oil
68	furfuryl mercaptan	7363	98-02-2	0.001	L	1,2-propanediol
70	guaiacol	460	90-05-1	0.2	S	mineral oil
71	heptanal	8130	111-71-7	0.04	L	mineral oil
72	1-heptanol	8129	111-70-6	0.4	L	mineral oil
73	hexanal	6184	66-25-1	0.1	L	1,2-propanediol
74	hexanoic acid	8892	142-62-1	2	L	1,2-propanediol
75	1-hexanol	8103	111-27-3	1	L	mineral oil
76	3-hexanol	12178	623-37-0	0.4	L	mineral oil
77	trans-1-Hexanal	5281168	6728-26-3	0.1	L	mineral oil
81	2-phenyl propionaldehyde dimethyl acetal	62336	90-87-9	15	L	mineral oil
82	hydroxy citronellal	7888	107-75-5	50	L	mineral oil
83	Indole	798	120-72-9	0.5	S	1,2-propanediol
89	linalool	443158	126-91-0	2.5	L	mineral oil
93	melonal	61016	106-72-9	1	L	mineral oil
94	l-menthol	16666	2216-51-5	3	S	1,2-propanediol
95	2-methoxy naphthalene	7119	93-04-9	10	S	mineral oil
96	methyl anthranilate	8635	134-20-3	0.35	L	mineral oil
97	methyl acetaldehyde dimethyl acetal	20859	462-95-3	0.15	L	1,2-propanediol
100	para-methyl quinoline	7059	91-62-3	0.25	L	1,2-propanediol
102	methyl salicylate	4133	119-36-8	0.25	L	mineral oil
103	S-(methyl thio) butyrate	62444	2432-51-1	0.03	L	mineral oil
104	musk galaxolide	91497	1222-05-5	5	L	1,2-propanediol
107	nonyl acetate	8918	143-13-5	25	L	mineral oil
109	1-octanol	957	111-87-5	0.75	L	mineral oil
110	1- octen 3- ol	18827	3391-86-4	0.04	L	mineral oil
111	pentanoic acid	7991	109-52-4	0.1	L	mineral oil
112	4-pentenoic acid	61138	591-80-0	0.3	L	mineral oil
113	phenyl acetic acid	999	103-82-2	5	S	mineral oil
114	phenyl acetylene	10821	536-74-3	0.3	L	mineral oil
115	phenyl ethanol	6054	60-12-8	50	L	mineral oil
117	iso-phorone	6544	78-59-1	3	L	mineral oil
118	alpha-pinene	6654	80-56-8	15	L	mineral oil
119	propyl butyrate	7770	105-66-8	1.3	L	mineral oil
121	propyl sulfide	8118	111-47-7	0.1	L	mineral oil
126	skatole	6736	83-34-1	0.005	S	1,2-propanediol
127	α-Terpineol	17100	10482-56-1	50	S	1,2-propanediol
131	thioglycolic acid	31277	123-93-3	10	S	water

132	thiophene	8030	110-02-1	2.5	L	mineral oil
133	thymol	6989	89-83-8	1	S	1,2-propanediol
134	ortho-tolualdehyde	10722	529-20-4	0.1	L	mineral oil
135	toluene	1140	108-88-3	3	L	mineral oil
138	gamma-undecalactone	7714	104-67-6	10	L	mineral oil
139	undecylenic acid	5634	112-38-9	50	S	1,2-propanediol
140	iso-valeraldehyde	11552	590-86-3	0.0003	L	mineral oil
141	iso-valeric acid	10430	503-74-2	0.01	L	mineral oil
142	gamma-valerolactone	7921	108-29-2	0.2	L	1,2-propanediol
143	vanillin	1183	121-33-5	**100**	S	1,2-propanediol

Supplementary Table 1B: List of 58 additional odorants and concentrations.

Name	CID Number	CAS Number	v/v % or w(g)/v(ml) %	Phase	Solvent
acetic acid	176	64-19-7	10	L	water
acetaldehyde	177	75-07-0	5	L	water
acetoin	179	513-86-0	0.1	S	1,2-propanediol
propan-2-one	180	67-64-1	25	L	water
butanal	261	123-72-8	0.002	L	water
octanal	454	124-13-0	0.25	L	mineral oil
2-hydroxypropanoic acid	612	50-21-5	**100**	L	
butane-2,3-dione	650	431-03-8	0.5	L	1,2-propanediol
propan-1-ol	1031	71-23-8	15	L	water
2-oxopropanoic acid	1060	127-17-3	0.8	L	water
methylsulfanylmethane	1068	75-18-3	0.004	L	water
decanoic acid	2969	334-48-5	**100**	S	1,2-propanediol
propan-2-ol	3776	67-63-0	**100**	L	
2-methylpropanal	6561	78-84-2	0.025	L	1,2-propanediol
butan-2-one	6569	78-93-3	1.5	L	1,2-propanediol
methyl acetate	6584	79-50-9	7.5	L	water
2-methylpropanoic acid	6590	79-31-2	1.25	L	1,2-propanediol
methyl benzoate	7150	93-58-3	0.5	L	1,2-propanediol
ethyl benzoate	7165	93-89-0	0.005	L	1,2-propanediol
ethyl 2-hydroxypropanoate	7344	97-64-3	20	L	1,2-propanediol
2-methyl-5-propan-2-ylcyclohexa-1,3-diene	7460	99-83-2	0.5	L	mineral oil
1-methyl-4-propan-2-ylbenzene	7463	99-87-6	0.5	L	1,2-propanediol
2-phenylethyl acetate	7654	103-45-7	**100**	L	
(3S)-3,7-dimethyloct-6-en-1-ol	7793	7540-51-4	1	L	1,2-propanediol
ethyl octanoate	7799	106-32-1	1	L	1,2-propanediol
propane-1-thiol	7848	107-03-9	0.0005	L	1,2-propanediol
ethyl formate	8025	109-94-4	8.33	L	1,2-propanediol

ethyl decanoate	8048	110-38-3	**100**	L	
heptan-2-one	8051	110-43-0	0.33	L	1,2-propanediol
methyl octanoate	8091	111-11-5	3.5	L	1,2-propanediol
undecanal	8186	112-44-7	0.2	L	1,2-propanediol
ethyl acetate	8857	141-78-6	3	L	1,2-propanediol
4-methylpent-3-en-2-one	8858	141-79-7	1	L	1,2-propanediol
butane-2-thiol	10560	513-53-1	0.0005	L	1,2-propanediol
ethyl pentanoate	10882	539-82-2	0.025	L	1,2-propanediol
heptan-2-ol	10976	543-49-7	5	L	1,2-propanediol
methyl propanoate	11124	554-12-1	2.5	L	1,2-propanediol
hexan-3-one	11509	589-38-8	0.5	L	1,2-propanediol
pent-1-en-3-ol	12020	616-25-1	0.3	L	1,2-propanediol
methyl butanoate	12180	623-42-7	0.15	L	1,2-propanediol
(methyldisulfanyl) methane	12232	624-92-0	0.025	L	1,2-propanediol
pentyl acetate	12348	628-63-7	0.75	L	1,2-propanediol
nonan-2-ol	12367	628-99-9	1	L	1,2-propanediol
decan-2-one	12741	693-54-9	**100**	L	
bis(methylsulfanyl)methane	15380	1618-26-4	0.002	L	1,2-propanediol
pentan-2-ol	22386	6032-29-7	4	L	1,2-propanediol
hexyl hexanoate	22873	6378-65-0	26.66	L	1,2-propanediol
4-methoxybenzaldehyde	31244	123-11-5	3.3	L	1,2-propanediol
diethyl butanedioate	31249	123-25-1	**100**	L	
ethyl hexanoate	31265	123-66-0	0.5	L	1,2-propanediol
butyl acetate	31272	123-86-4	1	L	1,2-propanediol
4-ethyl guaiacol	62465	2785-89-9	0.1	L	1,2-propanediol
(R)-(+)-beta-citronellol	101977	1117-61-9	**100**	L	
limonene	440917	5989-27-5	5	L	mineral oil
laevo-beta-pinene	440967	18172-67-3	10	L	1,2-propanediol
geraniol	637566	106-24-1	50	L	1,2-propanediol
thiolane	1127	110-01-1	0.0005	L	1,2-propanediol
trimethyl amine	1146	75-50-3	0.025	L	water

Supplementary Table 2: Three perfumer-provided names for each mixture. The first column: Number of components. The second column: 4- or 5- alternative forced-choice (AFC) identification task.

		Version I	**Version II**	**Version III**	**Version IV**
40	4- AFC	Thyme Vanilla -Cinnamon Coconut	Cherry Candy Vanilla Butter Cinnamon	Vanilla Butter Animalistic Urine	Cherry Candy Vanilla Butter Cinnamon
	5- AFC	Oregano Cinnamon Cherry Candy	Red fruit Vanilla baking Cinnamon Tea	Vanilla baking Cinnamon Tea Butter	Cinnamon Nutty Butter
30	4- AFC	Jasmine Red fruit Vinegar	Beurre noisette Cinnamon stick Vanilla	Bengay Thyme Vanilla Butter	Cherry Candy Feces Urine
	5- AFC	Air freshener Green Apple Cinnamon	Vanilla baking Coffee Nutty	Oregano Cinnamon Butter	Cherry Candy Nutty Vanilla baking
20	4- AFC	Jasmine Lily of the Valleys Water vapor	Patchouli - Vanilla Coconut Butter	Cat feces Jasmine Animalistic	Patchouli - Vanilla Cinnamon Urine/ Feces
	5- AFC	Nutty Vanilla baking Butter	Nutty Seasoned- Sweet Butter	Ink Cat urine Cinnamon	Cinnamon Apple Feces
10	4- AFC	Cat urine Feces Oregano	Talc Cream Baby smell Urine	Oregano Vinegar butter	Naphthalene Jasmine Rose water
	5- AFC	Oregano Urine/ Feces Rotten eggs	Fruity Flowery Butter	peppermint Oregano Feces	Naphthalene Feces/ Cat urine Vinegar
4	4- AFC	Oregano Pizza Cat urine	Feces Cat urine Wild cat/ Street cat	Banana Cat urine Feces	Oregano Pizza Butter Vinegar
	5- AFC	Weedy Seasoned Peppermint	Diluted Tomato paste Fruity Sour	Cider Cherry Candy Banana	Thyme Oregano Pizza Sauce
1	4- AFC	Coconut Bitter almond Butter	Naphthalene Narcissus Jasmine	Banana Pear Nectar Cherry Candy	Rose water Blancmange Sweet almond
	5- AFC	Coconut Fruity Creamy	Flowery naphthalene medicinal	Banana Green Apple Cherry candy	Rose water Pine needles Green Root

Supplementary Table 3: Verbal descriptors of white. The ratings of 146 descriptors by 85 subjects who smelled Laurax in all of the identification and delayed match-to-sample experiments; P.A. is *percentage of applicability* (geometric mean of the *percent usage* and *percentage of the maximum possible score*). In **BOLD** are the descriptors with the highest P.A. values.

	Descriptors	P.A.
1	**Fruity, citrus**	**56.2**
2	Lemon	45.9
3	Grapefruit	40.4
4	Orange	39.5
5	**Fruity, other than citrus**	**53.7**
6	Pineapple	34.3
7	Grape juice	30.7
8	Strawberry	21.1
9	Apple	27.5
10	Pear	28.8
11	Melon	30.7
12	Peach	28.3
13	Banana	19.0
14	**Floral**	**62.1**
15	Rose	43.2
16	Violets	24.8
17	Lavender	39.9
18	Cologne	40.1
19	Musky	45.5
20	**Perfumery**	**63.6**
21	**Fragrant**	**72.9**
22	**Aromatic**	**62.3**
23	Honey	26.4
24	Cherry	24.3
25	Almond	19.4
26	Nail polish remover	37.2
27	Nutty	19.3
28	Spicy	15.0
29	Clove	23.6
30	Cinnamon	24.1
31	Laurel leaves	18.6
32	Tea leaves	23.9
33	Seasoning (for meat)	5.5

	Descriptors	P.A.
74	**Soapy**	**60.7**
75	Leather	13.3
76	Cardboard	10.4
77	Rope	6.7
78	Wet paper	10.0
79	Wet wool, wet dog	8.9
80	Dirty linen	9.2
81	Stale	24.2
82	Musty, earthy, moldy	8.0
83	Raw potato	7.3
84	Mouse-like	6.0
85	Mushroom	10.1
86	Peanut butter	7.0
87	Beany	3.8
88	Eggy (fresh eggs)	7.2
89	Bark, birch bark	25.3
90	Cork	16.2
91	Burnt, smoky	6.4
92	Fresh tobacco smoke	15.2
93	Incense	44.2
94	Coffee	4.8
95	Stale tobacco smoke	17.4
96	Burnt paper	2.9
97	Burnt milk	5.4
98	Burnt rubber	12.3
99	Tar	5.1
100	Creosote	8.2
101	Disinfectant, carbolic	51.3
102	**Medicinal**	**56.1**
103	**Chemical**	**66.1**
104	Bitter	29.3
105	Sharp, pungent, acid	40.2
106	Sour, vinegar	35.1

34	Black pepper	7.1	107	Sauerkraut	14.9
35	Green pepper	5.9	108	Ammonia	20.0
36	Dill	11.4	109	Urine	10.6
37	Caraway	6.7	110	Cat urine	6.7
38	Oak wood, cognac	32.3	111	Fishy	2.9
39	Woody, resinous	31.7	112	Kippery (smoked fish)	2.5
40	Cedarwood	16.8	113	Seminal, sperm-like	7.7
41	Mothballs	28.9	114	New rubber	28.1
42	Minty, peppermint	30.9	115	Sooty	4.3
43	Camphor	31.8	116	Burnt candle	8.5
44	Eucalyptus	39.7	117	Kerosene	9.4
45	Chocolate	7.4	118	Oily, fatty	15.9
46	Vanilla	30.4	119	Buttery, fresh butter	10.5
47	**Sweet**	**56.8**	120	Paint	32.2
48	Maple syrup	22.2	121	Varnish	41.6
49	Caramel	20.0	122	Popcorn	2.8
50	Malty	16.1	123	Fried chicken	2.0
51	Raisins	20.2	124	Meaty (cooked, good)	1.9
52	Molasses	25.6	125	Soupy	7.7
53	Coconut	19.3	126	Cooked vegetables	7.4
54	Anise (licorice)	27.1	127	Rancid	20.6
55	Alcoholic	43.5	128	Sweaty	11.3
56	Etherish, anesthetic	40.1	129	Cheesy	9.3
57	Cleaning fluid	49.8	130	Household gas	12.3
58	Gasoline, solvent	12.5	131	Sulfidic	10.2
59	Turpentine (pine oil)	21.7	132	Garlic, onion	6.0
60	Geranium leaves	33.4	133	Metallic	19.5
61	Celery	8.5	134	Blood, raw meat	7.3
62	Fresh green vegetables	9.0	135	Animal	4.7
63	Crushed weeds	15.2	136	Sewer odor	4.8
64	Crushed grass	14.9	137	Putrid, foul, decayed	17.8
65	Herbal, green, cut grass	18.1	138	Fecal (like manure)	6.2
66	Raw cucumber	10.8	139	Cadaverous (dead animal)	9.8
67	Hay	13.0	140	Sickening	22.1
68	Grainy (as grain)	9.6	141	Dry, powdery	22.3
69	Yeasty	11.8	142	Chalky	15.9
70	Bakery (fresh bread)	5.3	143	Light	48.4
71	Sour milk	15.5	144	Heavy	38.6
72	Fermented (rotten) fruit	23.3	145	Cool, cooling	48.9
73	Beery	13.0	146	Warm	30.6

Supplementary Table 4: recipes for olfactory whites of 60 and 30 components (no overlap between A and B). Odorants listed by CID number.

60 component white (A)	
177	7921
179	8030
180	8103
261	8148
326	8186
460	8635
612	8858
999	10560
1060	10821
1068	10890
1127	10976
1140	12020
1146	12178
2879	12180
3314	12348
3776	12367
5634	17100
6544	18827
6561	22873
6654	24915
6982	26331
7344	31249
7463	31265
7519	61016
7583	62444
7632	62465
7654	93009
7710	637566
7714	5281168
7799	5283349

60 component white (B)	
176	8048
240	8051
264	8118
307	8130
798	8892
1031	10430
2758	10722
2969	10882
4133	11002
6054	11124
6184	11509
6501	11552
6590	12232
6736	12741
6989	14286
7059	15380
7150	16666
7165	19310
7363	22201
7410	22386
7600	31244
7685	31276
7731	62336
7749	62433
7762	91497
7793	440917
7848	440967
7888	443158
7966	5281515
7991	6259976

30 component white (B)
264
307
454
1031
6054
6736
7165
7460
7600
7685
7793
7848
7888
7966
8025
8048
8091
8118
8797
8918
10722
12232
16666
22201
31272
31276
61138
91497
440967
6259976

Supplementary Table 5: The four main components of a rose (8) **which yield the rose odor.** The most prominent labels which were given by independent group of ten people were: rose, narcissus, geranium, lemon and citrus blossoms.

Molecule name	CID Number	Cas Number	% of the rose mixture
β-phenylethyl alcohol	60-12-8	6054	98 %
Geraniol	106-24-1	637566	1.4 %
Citronellol	106-22-9	8842	0.3 %
Nerol	106-25-2	643820	0.3 %

Supplementary Table 6: Five versions of 30-components mixtures. The four molecules' CID numbers of the rose (in red) and the complementary 26 molecules' CID numbers for ideal spanning of the physicochemical space.

Version I	**Version II**	**Version III**	**Version IV**	**Version V**
177	179	179	176	179
264	261	261	180	1060
307	307	307	454	1127
460	460	612	6054	2879
1031	2969	3314	6544	6054
6054	6054	4133	6654	6561
4133	6569	6054	7059	6736
5634	6584	6584	7344	7059
6184	6989	7685	7460	7460
6501	7654	7793	7654	7583
7165	7685	7799	7685	7685
7654	7793	7921	7731	7714
7731	8048	8048	7799	8048
8091	8130	8103	7966	8118
8118	8148	8842	7991	8139
8842	8842	10882	8048	8186
8858	8892	11002	8139	8842
12180	12348	12180	8186	8858
17100	16666	12232	8842	10882
22873	24915	12741	16666	10890
24915	31249	16666	20859	12180
31272	31272	17100	24915	16666
62444	31276	20859	26331	24915
62465	62433	24915	31265	31244
93009	91497	91497	31272	31276

440917	93009	93009	31276	62444
443158	440917	440967	62465	93009
637566	637566	637566	637566	443158
643820	643820	643820	643820	637566
5281515	5281515	5283349	5281168	643820

Supplementary Table 7: The odorants in each mixture.

The odorants selected for each mixture in the similarity experiments. Number of components: # 30; ^ 20; $ 15; * 10; & 4; Bold 1. The odors are identified by numbers according to Supplementary Table 1 (ordinal numbers from *Dravnieks' Atlas*). Set-1 and set-2, were selected based on the perceptual space, set-3 and set-4, were selected based on the physicochemical space.

Set-1	
7	2 #^
9 ^	3
10 #	5 #^*
18 *	12 #
20 #^*&	13
23 #^	16 #
30 #	21 #^*
34 #^*	27 #^
37 #	29 #^
44 #^	31 #^
49 #^*	36 #*
51	38 #^*
52 #^	39 #*&
53 ^	40 #^*
54 #	**47** #^
56 #^	58 #
59 #^*	60 #^&
61	68 #^
63 ^	72 #
64 #^*	81
70 #^*&	82 #^*&
71	93 #

Set-2	
3 #	2
7	5 #^
9 #	13 #
10 ^	18
12 #	20
16 #^	21 #^
23 #	27 #
30 #	29 ^
31 #	37 ^
34 #^*	38 #^
36	40 #
39 &	44 #
47 #	49 #
58 ^	51 #
59 #^*	52 #^
61 #^	53 #^*
64 #	54 #
68 #*	56 #
74 #^&	60 *&
75 #	63 #^
82 #^*&	70 ^*
89 #^	71 #

Set-3	
2 #^*	3 #^
10 #^*&	5
16 *	7 #^
20 #^	**9** #^
23 #	12 #^*
29 ^	13 #*&
34	18 #
36 ^&	21 #
37 #^	27 #^
38 #^*	30 #^
39 #	31 #*
40	47 #^*
44 #&	49
51 #^*	52 #^
54	53 #
58 #	56 #
61 #*	59 ^
70	60 #
72 *	63 #
75 ^	64 #^
77 #^&	68 &
82 #^	71

Set-4	
1 #^$	2 #^$
5 #^*&	3 ^*
9 #	**7** #^$
10 #$	12 #^$*&
16 #	13 #^$&
18 #	27 #^
20 #$	30 #^
21 ^$	31 #$
23 #	34
29 #$	49
36 #^	52 #
37 *	53 #^
38 #	54 #
39 #^	56 #
40 #^*	60 #^$
44	63
47 #	64 #^$*
50	68
51 ^$	71 ^
58 #	73 #
59 ^$	75 ^
61 *	81

73 #*	94
74 #	95 ^
75 #	96 #^
76 #	107 #^
77 #	109 #
83 #*	110
89 #^&	111 #
97 #^	112 #*
100 #^	118 #
103	121 #
104 #	126 #^
113 #^	132 ^*
114 #	133
115 *	134 #
117 #^&	135 #
119 #^	139 ^
138 #	142 ^
141 #^	143 #^*&

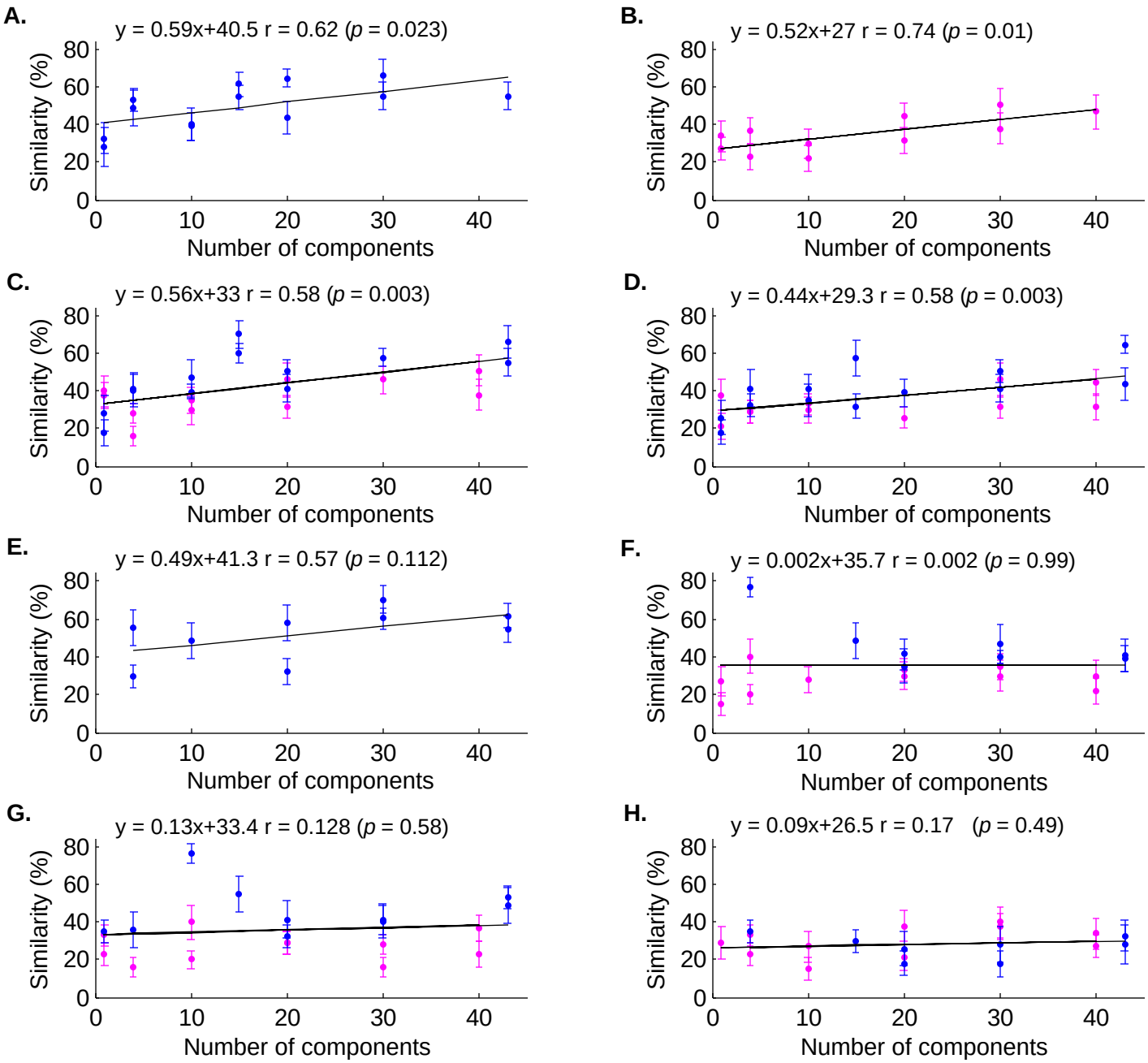
94 #^*	72 #^
96 #^*	73 #
100 #*&	76 #
103	77 ^
104 #^	**81** ^
107 ^	83 #^
109 #^	93 #^*
110 #^	95 #*
111 *	97 #^
119 #	112 ^*&
121 ^	113 #*
126 #*	114 #^*
132 #^*	115 #^*&
133 #^	117 #*&
134 #	118 #
139 #^	135 #^
141 #^	138 #
142	143 #^

83 #	73 #^
93 #	74 #
94 #^*	76
95 #	81
97 #	89 #^*&
110 #^	96 *
111 #^	100 #
114	103
115 #^	104 #^
118 #^	107 #^*
119 #^*	109 #^
121 #	112 #^&
132 #^	113 ^*
135 #	117 #^
138 #	126 #*
139 #^	133 #^
141 #^*	134 #*
143 #	142 #^

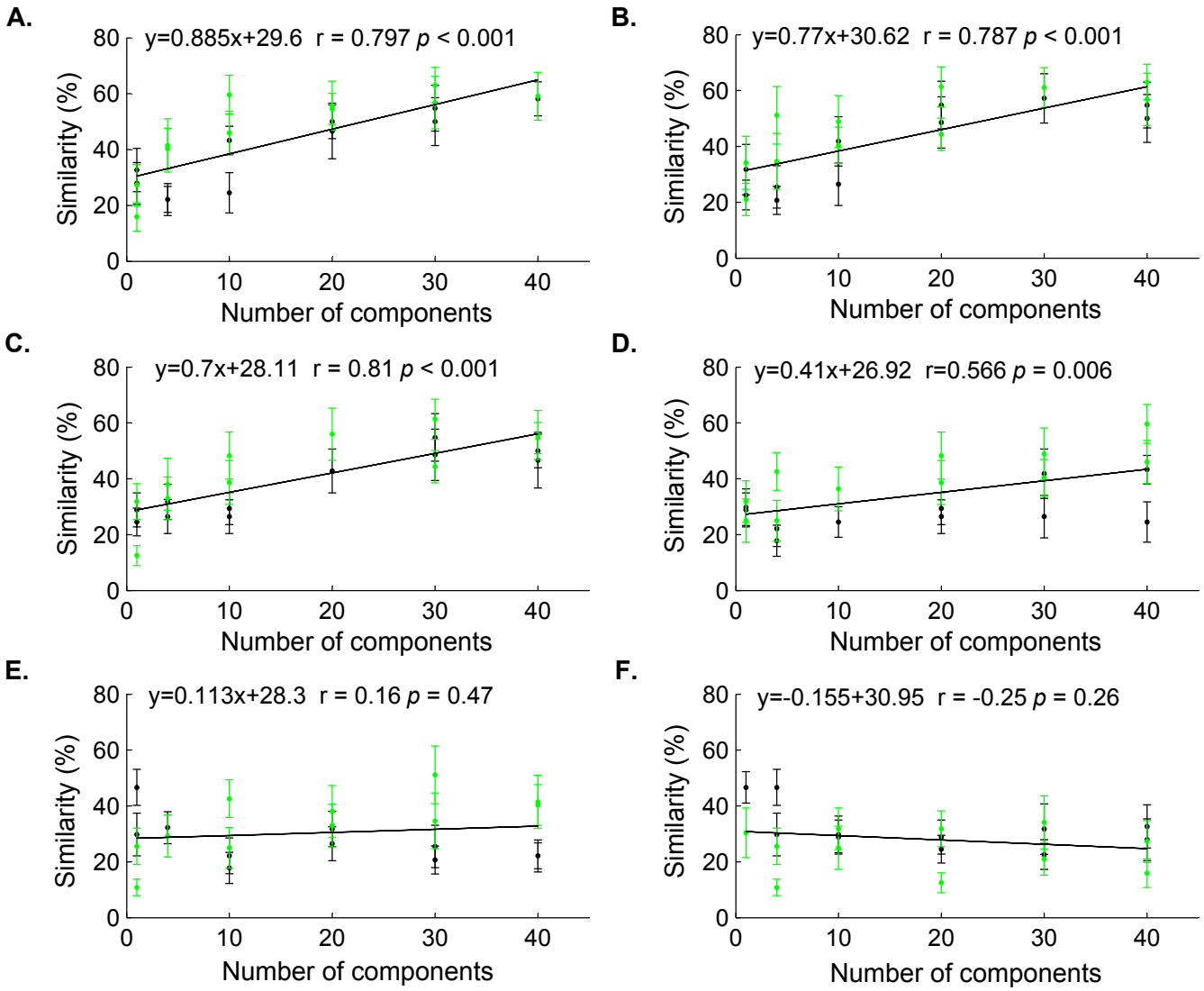
70 #^$&	82 #^
72 #^$*	83
74 #^	93 #$
76 #^*	94 #^$*
77 ^$	95 #
89 #^$	96
97	100
102 #^	104 #^$*
103	110 #*
107 #$	111 #*&
109 #$	115 #$
112 #	118 #^$
113 #^$*&	119 #^*
114 #^	121 #
117 #*&	132 #$
126 *	133 #*&
127 #^$	135
131	138 #
134	139 #^$*
140 #^*	141 #^$
143 #^	142 ^

1. Khan RM*, et al.* (2007) Predicting odor pleasantness from odorant structure: pleasantness as a reflection of the physical world. (Translated from eng) *J Neurosci* 27(37):10015-10023 (in eng).
2. Dravnieks A (1982) Odor quality: semantically generated multidimensional profiles are stable. (Translated from eng) *Science* 218(4574):799-801 (in eng).
3. Dravnieks A ed (1985) *Atlas of odor character profiles* (ASTM Press).
4. Todeschini R & Consonni V eds (2009) *Molecular Descriptors for Chemoinformatics* (Wiley-VCH), Vol 41.
5. Aycı F, Aydınlı M, Bozdemir ÖA, & Tutaş M (2005) Gas chromatographic investigation of rose concrete, absolute and solid residue. *Flavour and Fragrance Journal* 20(5):481-486.
6. Akiyama M*, et al.* (2007) Analysis of the headspace volatiles of freshly brewed arabica coffee using solid-phase microextraction. (Translated from eng) *J Food Sci* 72(7):C388-396 (in eng).
7. Komthong P, Hayakawa S, Katoh T, Igura N, & Shimoda M (2006) Determination of potent odorants in apple by headspace gas dilution analysis. *LWT - Food Science and Technology* 39(5):472-478.
8. Jirovetz L*, et al.* (2005) Solid phase microextraction/gas chromatographic and olfactory analysis of the scent and fixative properties of the essential oil of *Rosa damascena* L. from China. *Flavour and Fragrance Journal* 20(1):7-12.

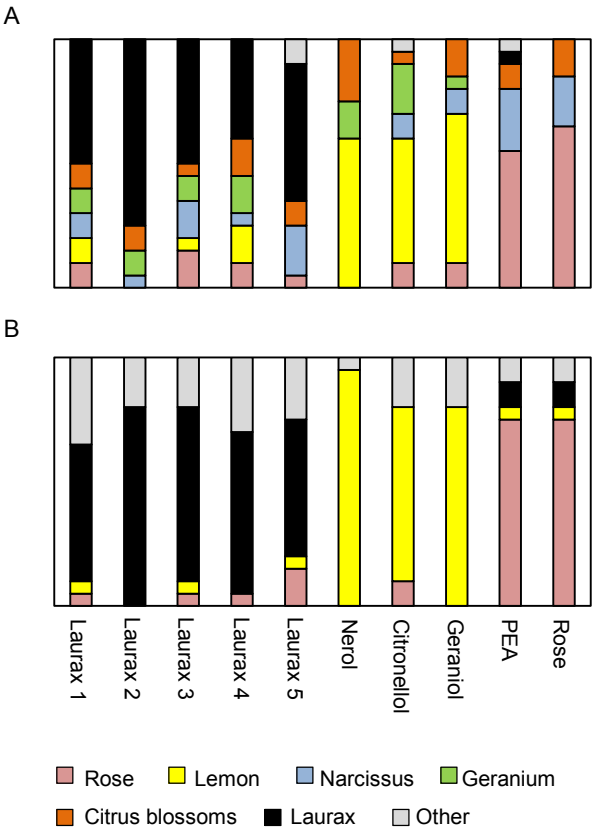
Supp. Figure 1



Supp. Figure 2



Supp. Figure 3



Supp. Figure 4