nature neuroscience

The missense of smell: functional variability in the human odorant receptor repertoire

Joel D Mainland[1–3], Andreas Keller[4], Yun R Li[2,6], Ting Zhou[2], Casey Trimmer[1], Lindsey L Snyder[1], Andrew H Moberly[1,3], Kaylin A Adipietro[2], Wen Ling L Liu[2], Hanyi Zhuang[2,6], Senmiao Zhan[2], Somin S Lee[2,6], Abigail Lin[2] & Hiroaki Matsunami[2,5]

Humans have ~400 intact odorant receptors, but each individual has a unique set of genetic variations that lead to variation in olfactory perception. We used a heterologous assay to determine how often genetic polymorphisms in odorant receptors alter receptor function. We identified agonists for 18 odorant receptors and found that 63% of the odorant receptors we examined had polymorphisms that altered *in vitro* function. On average, two individuals have functional differences at over 30% of their odorant receptor alleles. To show that these *in vitro* results are relevant to olfactory perception, we verified that variations in *OR10G4* genotype explain over 15% of the observed variation in perceived intensity and over 10% of the observed variation in perceived valence for the high-affinity *in vitro* agonist guaiacol but do not explain phenotype variation for the lower-affinity agonists vanillin and ethyl vanillin.

The human genome contains ~800 odorant receptor genes that have been shown to exhibit high genetic variability[1–3]. In addition, humans exhibit considerable variation in the perception of odorants[4,5], and variation in an odorant receptor predicts perception in four cases: loss of function in *OR11H7P*, *OR2J3*, *OR5A1* and *OR7D4* leads to elevated detection thresholds for the respective agonists isovaleric acid[6], *cis*-3-hexen-1-ol[7], β-ionone[8] and androstenone[9]. These results suggest that although the olfactory system uses a combinatorial code in which responses of multiple receptor types lead to recognition of a given odorant, the response of a single receptor can have a large influence on the perception of an odorant.

Understanding the role of a single receptor requires functional data for receptor-odorant pairs. Matching mammalian odorant receptors to ligands has seen limited success, and the picture is even worse when considering human odorant receptors; ligands have been published for only 22 of the ~400 intact human odorant receptors[6,8–17]. This lack of data is a critical bottleneck in the field; matching ligands to odorant receptors is essential for understanding the olfactory system at all levels and for building viable models of olfaction.

Using a high-throughput system for functional testing of odorant receptors[18], we can now study the role of missense single-nucleotide polymorphisms (SNPs) in the function of odorant receptors. Here we identify ligands for several orphan odorant receptors, determine the prevalence and functional consequences of missense mutations in odorant receptors, and measure the effect of these functional changes on human olfactory perception.

RESULTS

High-throughput screening of human odorant receptors

To identify agonists for a variety of odorant receptors, we cloned a library of 511 human odorant receptor genes for a high-throughput heterologous screen. These clones represent 394 (94%) of the 418 intact odorant receptor genes, and 428,793 (47%) of their 912,912 intact odorant receptor alleles present in the 1000 Genomes Project. Some odorant receptors were represented by multiple nonsynonymous alleles in the screen.

We screened the odorant receptor library with a panel of 73 odorants that have been used in previous psychophysical testing[9,19] and used a cyclic adenosine monophosphate (cAMP)-mediated luciferase assay to measure receptor activity[20] (**Supplementary Fig. 1**). In the primary screen, we stimulated at an odorant concentration of 100 μM. We selected 1,572 odorant-receptor pairs from this primary screen for a secondary screen in which we tested each odorant receptor against a no-odor control as well as 1 μM, 10 μM and 100 μM concentrations of the odorant in triplicate. For 425 odorant-receptor pairs, exposure to at least one concentration of the odorant resulted in significantly higher activation than the no-odor control (*t*-test, $P < 0.05$, uncorrected for multiple comparisons). These odorant-receptor pairs included 190 clones representing 160 unique odorant receptors.

We then constructed dose-response curves for at least one putative agonist of 160 odorant receptors. 27 odorant receptors showed a significant response to at least one agonist (extra sums-of-squares *F* test against vector control, $P < 0.05$ divided by the number of

[1]Monell Chemical Senses Center, Philadelphia, Pennsylvania, USA. [2]Department of Molecular Genetics and Microbiology, Duke University Medical Center, Research Drive, Durham North Carolina, USA. [3]Department of Neuroscience, University of Pennsylvania School of Medicine, Philadelphia, Pennsylvania, USA. [4]Laboratory of Neurogenetics and Behavior, The Rockefeller University, New York, New York, USA. [5]Department of Neurobiology and Duke Institute for Brain Sciences, Duke University Medical Center, Research Drive, Durham North Carolina, USA. [6]Present addresses: School of Medicine, University of Pennsylvania, Philadelphia, Pennsylvania, USA (Y.R.L.), Department of Pathophysiology, Key Laboratory of Cell Differentiation and Apoptosis of National Ministry of Education, Shanghai Jiao Tong University School of Medicine, Shanghai, China (H.Z.) and Department of Ecology and Evolutionary Biology, Osborn Memorial Labs, Yale University, New Haven, Connecticut, USA (S.S.L.). Correspondence should be addressed to J.D.M. (joel.mainland@gmail.com).

Received 6 August; accepted 5 November; published online 8 December 2013; doi:10.1038/nn.3598

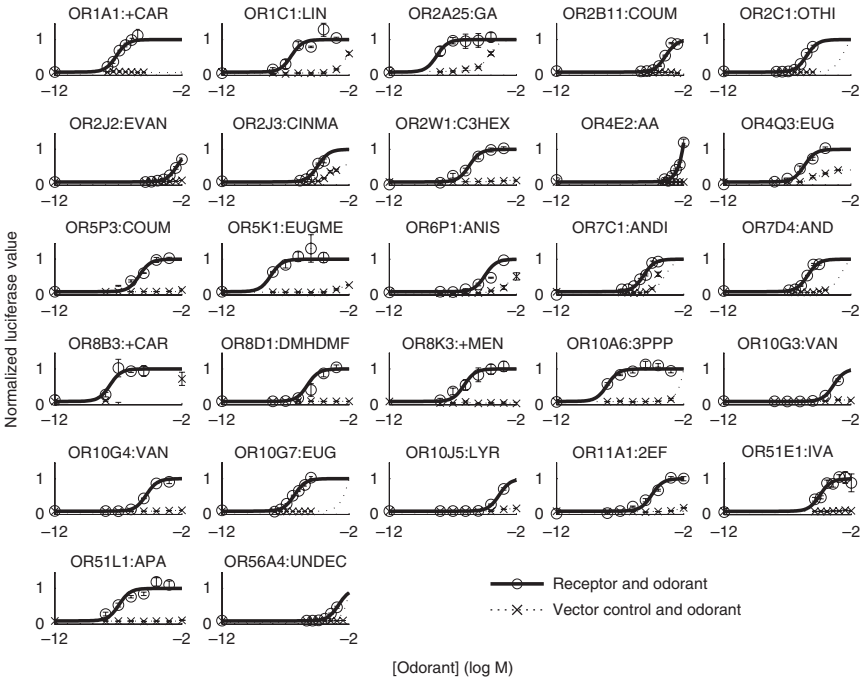


Figure 1 Dose-response curves of the receptor encoded by the most common functional allele for 27 receptors. Responses of cells transfected with either a plasmid encoding the indicated odorant receptor or an empty vector to the indicated odorants. Error bars, s.e.m. over three replicates. Odor abbreviations are defined in **Supplementary Table 1**.

receptors tested), including nine that have previously been shown to respond to at least one agonist[9,16,17] (**Fig. 1**). For the other 18 odorant receptors we identified new agonists. This nearly doubles the total number of published human odorant receptors with known agonists, bringing the total to 40 (refs. 6,8–17). The receptors identified by this method are spread throughout 9 of the 13 gene families of odorant receptors[21] (**Fig. 2**), suggesting that our assay is useful for examining ligand-receptor interactions across a wide variety of odorant receptors.

Genetic variation in odorant receptors

We identified agonists for seven odorant receptors that segregate between intact and disrupted forms (**Table 1**), bringing the total number of segregating pseudogenes with known agonists to eight[6]. Combined with psychophysics data for a genotyped population, these odorant receptor-agonist pairs can be used to probe the role of a single odorant receptor in olfactory perception.

In addition to segregating pseudogenes and missense variation in conserved amino acid residues, a segregating missense variation that alters nonconserved amino acid residues of odorant receptors can also account for a portion of the variance in odor perception[7–9]. How many of the odorant receptors with intact open reading frames have functionally different variants, adding to the already considerable amount of variation in the human odorant receptor repertoire? We found a median of five alleles with an allele frequency greater than 1% across 418 odorant receptors in the 1000 Genomes Project data. 18 odorant receptors had only one allele with an allele frequency over 1% across the 2,184 haplotypes. In contrast, *OR51A2* had 19 different variants with an allele frequency over 1%. The odorant receptors for which we identified agonists did not exhibit a significantly different number of polymorphisms than odorant receptors without identified agonists (median alleles = 5 for both sets, two-sided Mann-Whitney *U* test, $Z = 0.77$, $P = 0.44$).

To test how variability in amino acid sequence affects odorant receptor activation by odorants, we targeted odorant receptors with at least one known agonist and cloned alleles from pooled genomic DNA with the goal of representing the majority of protein-coding alleles seen in the 1000 Genomes Project data. For 16 odorant receptors we successfully cloned 51 alleles, which represented an average of 27,118 (77%) of the receptors' 34,944 alleles present in the 1000 Genomes Project data. One mechanism through which genetic polymorphisms could influence receptor function is by altering cell-surface expression. We assessed the cell-surface expression of odorant receptor variants encoded by these 51 cloned alleles using live-cell immunostaining with an antibody against the N-terminal Rho tag followed by fluorescence-activated cell sorting (FACS). Relative cell-surface expression among each set of variants did not correlate with either relative potency (Spearman $\rho = 0.04$, $P = 0.82$, **Supplementary Fig. 2a**) or relative efficacy (Spearman $\rho = 0.13$, $P = 0.45$, **Supplementary Fig. 2b**) of the variants in the functional assay. Although a complete lack of cell-surface expression eliminated receptor responses to known agonists, high surface expression did not reliably confer additional sensitivity.

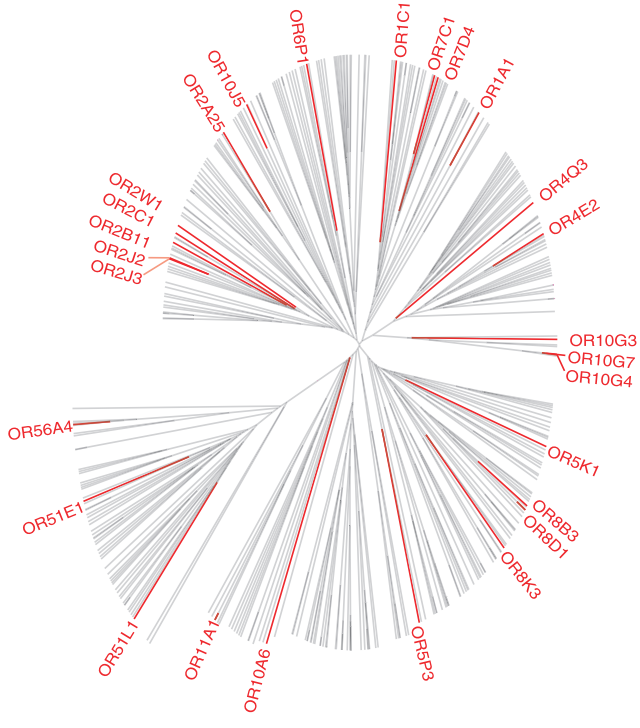


Figure 2 Unrooted tree based on similarity of amino acid properties. 27 odorant receptors with agonists are highlighted in red and represent 9 of the 13 odorant receptor gene families. Grantham's amino acid property scales were used to quantify receptor similarity[30], and distances were calculated using the unweighted pair group method with arithmetic mean (UPGMA).

Table 1 Seven segregating pseudogenes with agonists

Odorant receptor name	Pseudogene allele percentage (%)	Result	Agonist
OR2B11	43	8-amino-acid protein	Cinnamaldehyde
OR4E2	30	MAYDRY domain	Amyl acetate
OR8K3	24	MAYDRY domain	(+)-menthol
OR10A6	22	PMLNPLIY domain	3-phenyl propyl propionate
OR2C1	4	272 amino acid protein	Octanethiol
OR4Q3	1.50	159 amino acid protein	Eugenol
OR10G7	1.40	191 amino acid protein	Eugenol

The frequency of disrupted alleles for the corresponding odorant receptor genes as found in the 1000 Genomes Project[22] is listed. In cases where the variant allele alters a highly conserved domain in the protein, the conserved amino acid that varies is underlined.

A small amount of cell-surface expression was sufficient to confer functional responses. In summary, FACS did not provide enough resolution to determine whether functional variation was due to defects in cell-surface expression.

Functional consequences of genetic variation

We screened odorant receptor variants encoded by 46 of the alleles used in the FACS analysis against 55 odorants chosen quantitatively to span the physicochemical space[17] (**Supplementary Fig. 3**). Across odorants, the absolute magnitudes of response varied, but the relative responses of variants remained consistent (**Fig. 3a**,**b** and **Supplementary Fig. 4**). In other words, if a variant was hypersensitive to one agonist, that variant tended to be hypersensitive to all agonists. We found no case of a genetic change that resulted in a change in odor tuning (**Supplementary Fig. 4**), but we chose our odorant library to span odorant space, and this library was therefore not ideal for identifying more subtle changes.

We then examined how the variant responses compared across a range of concentrations by constructing a dose-response curve from 10 nM to 10 mM (**Fig. 3c** and **Supplementary Fig. 5**). Here we included the 15 odorant receptors tested against all 55 odorants as well as 12 additional odorant receptors. We typically used only a single agonist, as our results from using a broad set of odorants suggested that the differences between variants using one odorant were highly correlated to differences between variants using different odorants. We fit the data to a sigmoid curve and compared the variants using an extra sums-of-squares F test. We classified a pair of variants as hyper- or hypofunctional if one variant in the pair had both a lower potency (half maximal effective concentration; EC50) and a lower efficacy (maximum value). Comparing one variant to all other variants of the same odorant receptor from the 1000 Genomes Project data revealed that 11% of the variants were hyperfunctional, 68% were indistinguishable and 6.8% were hypofunctional. 7.9% of the variants were encoded by pseudogenes, and for 5.5% of the variants, potency and efficacy did not change concordantly, so we could not clearly classify them as hypo- or hyperfunctional (**Fig. 4a**). 63% (17/27) of the odorant receptors we examined had polymorphisms that altered *in vitro* function. Residues that are polymorphic across variants with measured function are shown in **Figure 4b**. There was no obvious pattern to the amino acids that change function; they were found throughout the protein. The odds that a residue altered function in our assay did not correlate with evolutionary conservation (genomic evolutionary rate profiling (GERP) score, regression analysis, $r = -0.04$, $P = 0.83$), predictions from the SIFT tool (sorting intolerant from tolerant; regression analysis, $r = 0.05$, $P = 0.80$) or predictions from the PolyPhen tool (polymorphism phenotyping; regression analysis, $r = -0.05$, $P = 0.81$).

To quantify functional differences across the 1000 Genomes Project population, we assigned *in vitro* results to each participant according to their allele type. We had *in vitro* results for 46,561 (79%) of the 58,968 alleles (27 odorant receptors × 1,092 subjects × 2 alleles). When we conservatively classified all pairwise comparisons, including those involving untested alleles, as functionally identical, we saw an average of 16 functional differences in dose response out of 54 possible functional differences (27 odorant receptors tested in dose response × 2 alleles; **Fig. 5a**). When we classified all pairwise comparisons, including an untested allele, as functionally different, we saw an average of 22 functional differences in dose response out of 54 possible functional differences. These results were consistent

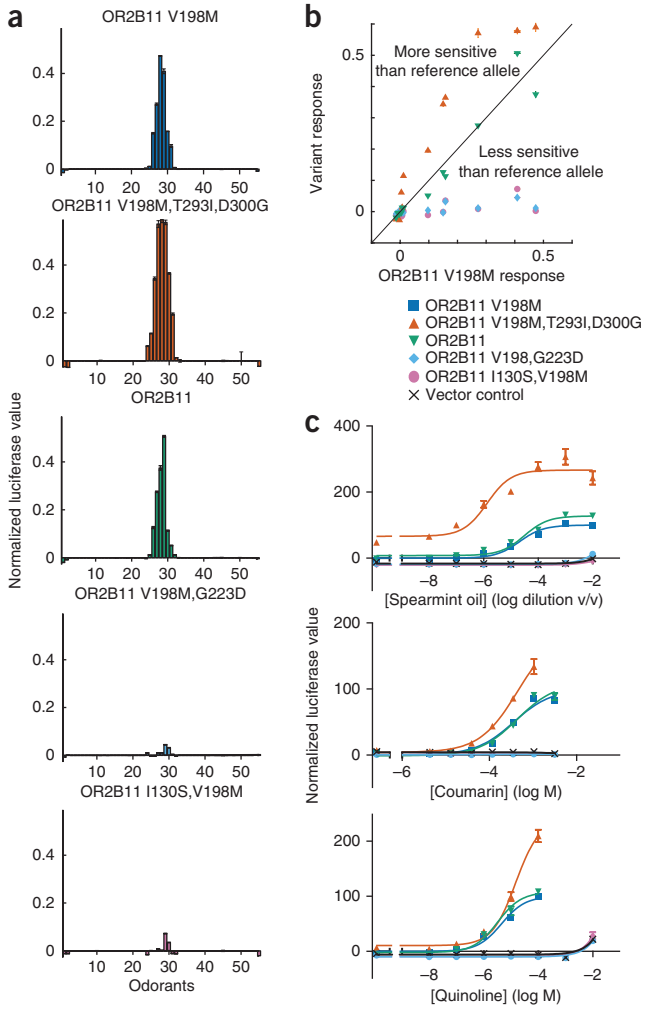


Figure 3 Functional testing of odorant receptor variants. (**a**) Sensitivity-ordered tuning curves for five variants of OR2B11 tested against 55 representative odorants at 100 μM. If a given odorant did not significantly activate any of the variant receptors above the no-odor control (two-tailed t-test, $\alpha = 0.05/55$), that odorant's response was set to zero across all variants. Odorants were ordered along the x axis according to the response they elicited from the receptor encoded by the *OR2B11* reference allele (see **Supplementary Fig. 3** for odor names). Error bars, s.e.m. over three replicates. (**b**) Responses of the receptors encoded by the four variant alleles to the 55 representative odorants at 100 μM, plotted against the responses of the receptor encoded by the *OR2B11* reference allele. Black line, unit slope. (**c**) Dose-response curves for the receptors encoded by *OR2B11* alleles for three indicated odorants. Y axis represents the luciferase value normalized to the value for the receptor encoded by the reference allele. Error bars, s.e.m. over three replicates.

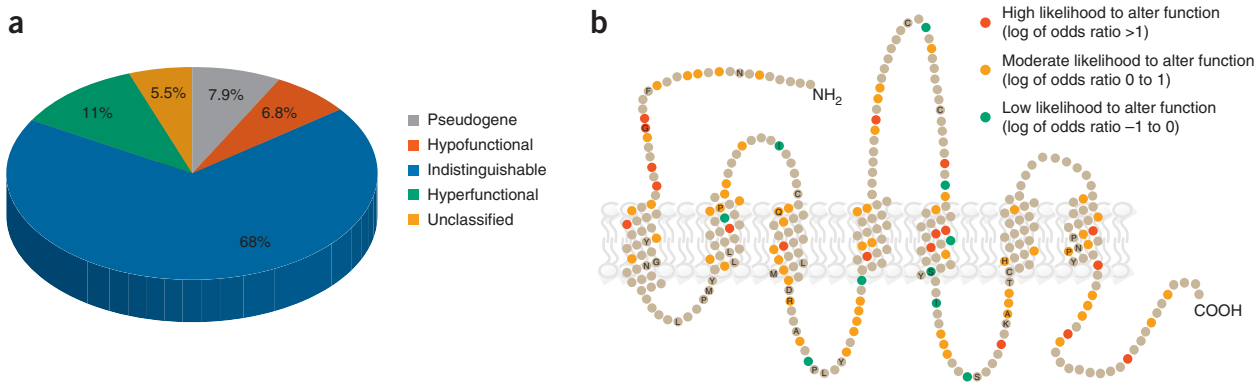


Figure 4 Summary of functional variation. (**a**) Functional differences among 27 odorant receptors of 1,092 participants in the 1000 Genomes Project. Note that pseudogenes account for a small portion of the variability relative to missense variations. (**b**) Snake plot of a typical odorant receptor showing residues where SNPs alter the function of the receptor. Amino acid residues that did not vary between any of the receptors encoded by minor alleles and the receptor encoded by their reference allele are shown in gray. The remaining residues are colored according to the odds that they alter function given our current dose-response data. Amino acid positions conserved in at least 90% of the receptors are labeled with their single-letter amino acid code.

if we excluded the 500 related participants. In other words, two individuals differed functionally at over 30% (16/54) of their odorant receptor alleles. Pairs in which both participants were of Asian ancestry (Han Chinese in Beijing (CHB), Southern Han Chinese (CHS) and Japanese in Tokyo (JPT)) were more functionally similar than pairs in which neither participant had Asian ancestry

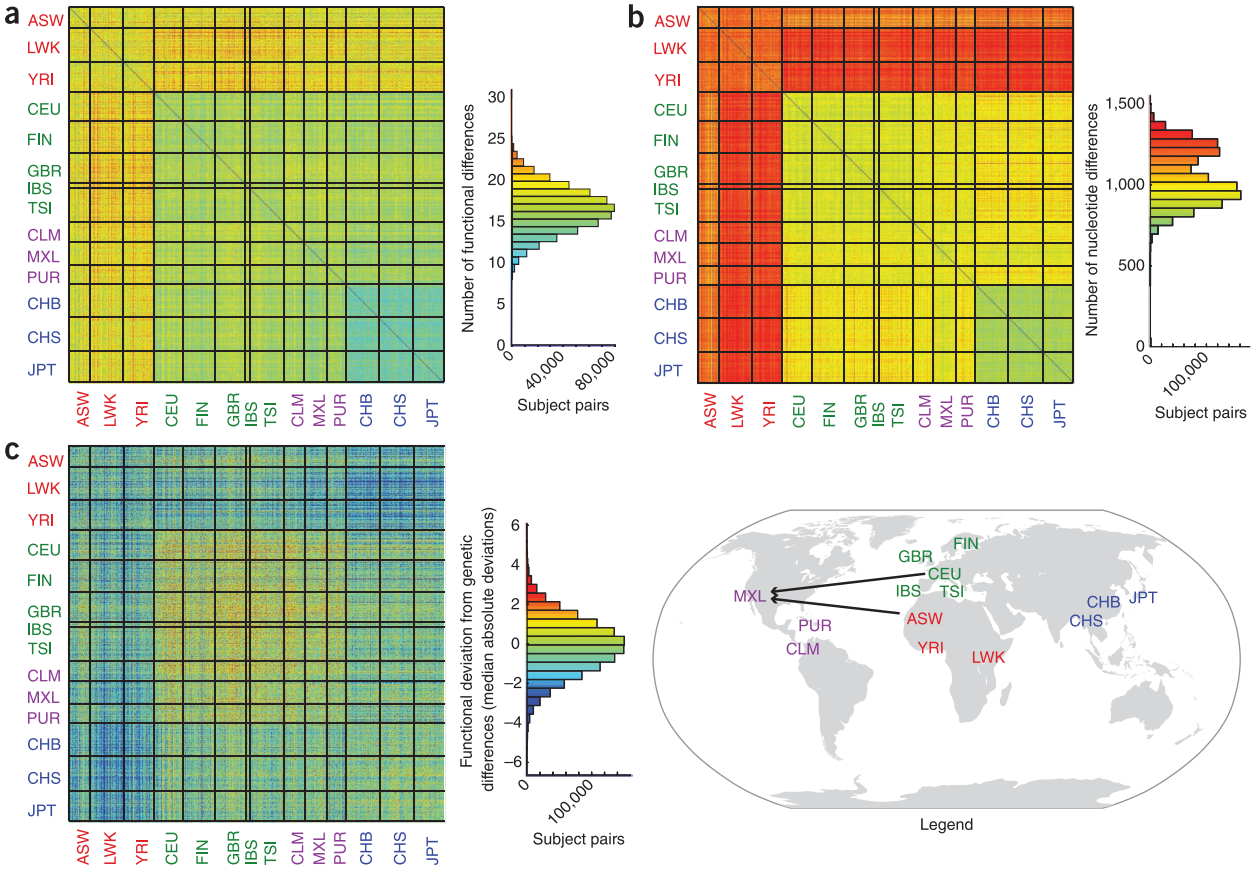


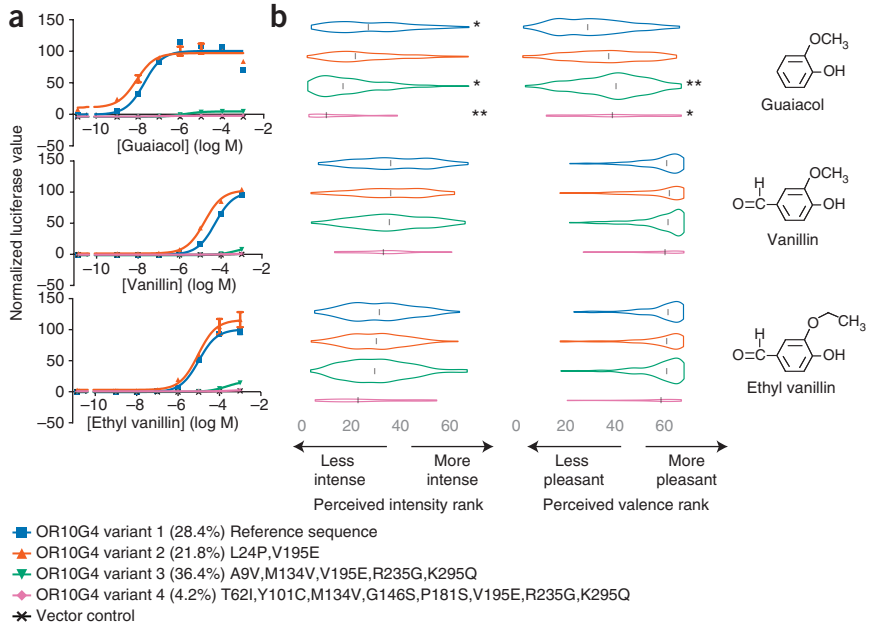
Figure 5 Functional differences between participants. (**a**–**c**) The number of functional differences (**a**), nucleotide differences (**b**), and *z*-scored functional differences minus *z*-scored nucleotide differences (**c**) among 27 odorant receptors of 1,092 participants from the 1000 Genomes Project. The colors of the squares represent the number of differences between participants, with color values indicated in the histograms. Participant populations are labeled on the axes and separated by black grid lines. The legend displays ethnic groups from **a**–**c** at the place of geographic origin; arrows point to the location of sample collection. ASW, African ancestry in Southwest USA; CEU, Utah residents with Northern and Western European ancestry from the CEPH collection; CHB, Han Chinese in Beijing; CHS, Southern Han Chinese; CLM, Colombian in Medellin, Colombia; FIN, Finnish; GBR, British individuals from England and Scotland; IBS, Iberian populations in Spain; JPT, Japanese in Tokyo; LWK, Luhya in Webuya, Kenya; MXL, Mexican ancestry in Los Angeles, California; PUR, Puerto Rican; TSI, Tuscanians in Italy; YRI, Yoruba in Ibadan, Nigeria.

(median Asian = 13; median non-Asian = 17; two-sided Mann-Whitney *U* test, $z = 127$, $P < 0.0001$). Pairs in which both participants were of African ancestry (African Ancestry in Southwest United States (ASW), Luhya in Webuya, Kenya (LWK), and Yoruba in Ibadan, Nigeria (YRI)), were more functionally different than pairs in which neither participant was of African ancestry (median African = 16; median non-African = 15; two-sided Mann-Whitney *U* test, $z = 29$, $P < 0.0001$)[22], in line with those populations having a greater genetic diversity (**Fig. 5b**). However, when taking genetic diversity into account, pairs in which both participants were of African ancestry (ASW, LWK and YRI) were more functionally similar than pairs in which neither participant was of African ancestry (median African = −0.83; median non-African = 0.36; two-sided Mann-Whitney *U* test, $z = 149$, $P < 0.0001$; **Fig. 5c**). This shows that, although there is greater genetic variability among individuals of African descent, much of this diversity did not translate into functional differences relative to other groups.

Perceptual consequences of genetic variation

We have so far shown that genetic changes are widespread in the human population and that these genetic changes result in widespread *in vitro* functional changes. We next determined whether the observed *in vitro* functional changes lead to the predicted consequences in perception. We selected an odorant receptor encoded by *OR10G4* for additional testing because we had genomic DNA of subjects that had been tested for their perception of three agonists of the OR10G4 receptor[19] and because functional and nonfunctional *OR10G4* alleles were common in the 1000 Genomes Project data[22]. We obtained *OR10G4* sequences for 308 of the 391 participants who had rated their perceived intensity and valence for guaiacol, vanillin and ethyl vanillin. We then examined the effect of each *OR10G4* allele on the perceptual phenotypes (**Fig. 6**).

There were four *OR10G4* alleles with minor allele frequency (MAF) >4% in the participant population: the reference allele (allele 1), which encodes ALTYMGPVRK, and variant alleles 2–4 that encode products that differ from the one encoded by the reference allele by two (APTYMGPERK), five (VLTYVGPEGQ) or eight (ALICVSSEGQ) amino acids, respectively. We challenged the odorant receptor variants encoded by these four alleles with guaiacol using a heterologous luciferase assay and fit the resulting data to a sigmoidal curve. The odorant receptor encoded by allele 2 was more sensitive to guaiacol than the reference odorant receptor encoded by allele 1, but the effect was small (log(EC50 for allele 1) = −7.4, log(EC50 for allele 2) = −7.7, extra sum of squares *F* test, $F_{3,42} = 6.38$, $P < 0.002$). The odorant receptor encoded by allele 3 had a much lower affinity to the three odorants than the reference odorant



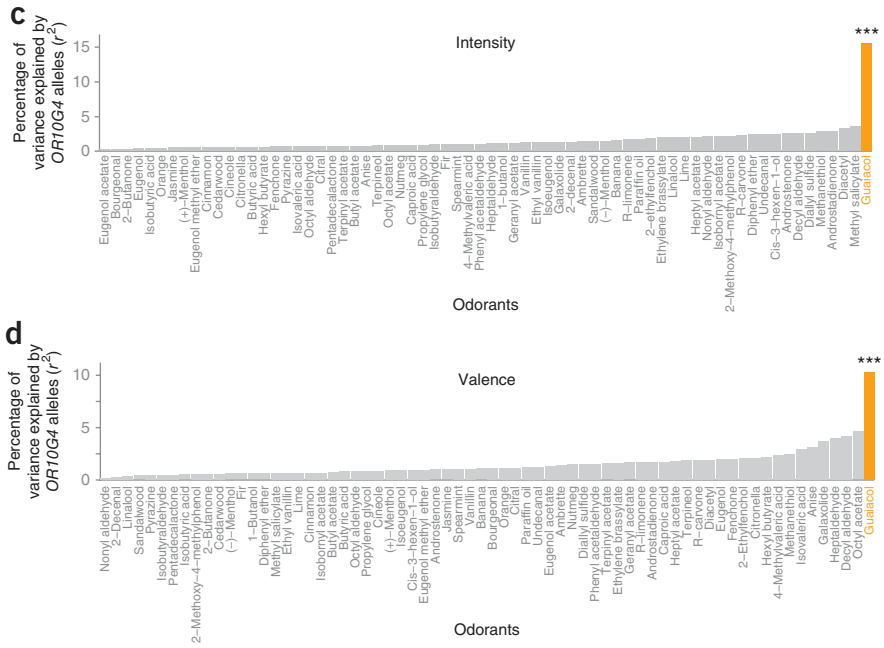


Figure 6 Effects of genetic variation in *OR10G4* on perceived intensity and valence. (**a**) Concentration response curves of receptors encoded by *OR10G4* alleles with a frequency greater than 4% in the participant population. Error bars, s.e.m. of three replicates. *Y*-axis values are normalized to the baseline response of the reference allele. (**b**) Perceived intensity and valence rank for three *in vitro OR10G4* agonists by receptor variants of *OR10G4*. Each participant is represented twice: once for the maternal and once for the paternal allele. The width of each violin is proportional to the number of participants assigning a given rank. The black line inside the violin denotes the median rank. The amino acid changes are relative to the hg19 reference sequence. The frequency listed is the allele frequency in the 308 participants. All unlisted alleles occurred with a frequency lower than 4%. $^*P < 0.05$, $^{**}P < 0.01$ in the regression model, *F* test (only shown for regression models that were overall significantly different from a constant model). (**c**,**d**) Percentage of perceptual variance (r^2) in intensity (**c**) and valence (**d**) ranking explained by *OR10G4* allele types. Each odor was analyzed using the multiple linear regression model outlined in Results. $^{***}P < 0.001$ after FDR correction, *F* test. For all other odorants, $P > 0.05$ after FDR correction.

receptor but still resulted in significant responses (log(EC50) = −5.5, sum of squares test against reference, $F_{3,42} = 459$, $P < 0.001$; extra sum of squares F test against vector control, $F_{3,42} = 149$, $P < 0.001$). There was no difference in response between cells transfected with allele 4 and cells transfected with vector only (extra sum of squares F test against vector control, $F_{3,42} = 2.2$, $P = 0.11$; **Fig. 6a**). We generated odorant receptors with each of the SNPs in a reference background and found that no single SNP accounted for the functional impairment in the odorant receptors encoded by alleles 3 and 4, suggesting that multiple residues interact to cause the decrease in affinity (**Supplementary Fig. 6**).

Using multiple regression analysis, we tested whether *OR10G4* allele type significantly predicted participants' perception of the three *in vitro* agonists. We regressed the predictors, allele counts (0, 1 or 2) for the four alleles with frequency >4% in the participant population, against the odor rating rank. *OR10G4* allele type predicted 15.4% of the variance in perceived intensity of guaiacol (regression analysis, $r^2 = 0.165$, adjusted $r^2 = 0.154$, compared to constant model, $F_{4,303} = 15.0$, $P < 0.001$ after false discovery rate (FDR) correction). The model estimated that subjects with none of the major alleles would rank the intensity of guaiacol 24th relative to the other tested odors. Participants rated guaiacol 2.1 ranks more intense for each copy of the reference allele 1 they possessed ($\beta = 2.10$, $P < 0.04$), and 2.4 and 4.3 ranks less intense for each copy of alleles 3 and 4 they possessed ($\beta = -2.39$, $P < 0.02$ and $\beta = -4.34$, $P < 0.005$), respectively. Participants' intensity ratings were not significantly associated with possession of allele 2 ($\beta = 1.01$, $P = 0.32$).

In addition to intensity, *OR10G4* allele type predicted 10.3% of the variance in perceived valence of guaiacol ($r^2 = 0.115$, adjusted $r^2 = 0.103$, compared to the constant model, $F_{4,303} = 9.85$, $P < 0.001$ after FDR correction). The model estimated that subjects with none of the major alleles would rank the valence of guaiacol 29th relative to the other tested odors. Participants rated guaiacol 3.3 and 3.7 ranks more pleasant for each copy of alleles 3 and 4 ($\beta = 3.33$, $P < 0.002$ and $\beta = 3.71$, $P < 0.03$), respectively, but participants' valence ratings were not significantly associated with possession of alleles 1 and 2 ($\beta = -0.69$, $P = 0.52$; $\beta = 1.88$, $P = 0.08$), respectively.

In contrast to guaiacol, neither perceived intensity nor valence of vanillin and ethyl vanillin were predicted by *OR10G4* allele type (vanillin intensity compared to the constant model, $F_{4,303} = 0.95$, uncorrected $P = 0.44$; ethyl vanillin intensity compared to the constant model, $F_{4,303} = 0.95$, uncorrected $P = 0.44$; vanillin valence compared to the constant model, $F_{4,303} = 0.84$, uncorrected $P = 0.50$; and ethyl vanillin valence compared to the constant model, $F_{4,303} = 0.50$, uncorrected $P = 0.74$). As additional controls, the 308 participants were also psychophysically tested for their intensity and valence perception of 63 odors that are not known to be agonists of the *OR10G4* receptor as well as two solvents. Of the 68 compounds, only guaiacol intensity and valence were significantly correlated with *OR10G4* allele type (**Fig. 6c**,**d**).

DISCUSSION

We identified 27 odorant receptors with known agonists that have functionally different alleles that segregate in the human population and demonstrated that this segregation is relevant to human perception of odorants. This nearly doubles the number of human odorant receptors with a known agonist and, to our knowledge, is the first investigation of the functional role of genetic variation in a large set of odorant receptors. Pairing odorants and odorant receptors, and verifying the functional consequences of segregating polymorphisms *in vitro* allowed us to address previously inaccessible questions regarding how activation of an individual odorant receptor alters olfactory perception. This promises to be a rich future field of study, as we do not currently know how the odorant receptor array codes for odor threshold, intensity or character. Understanding how the functional alteration of an odorant receptor affects the neural code is a crucial first step in a model of olfactory perception.

Each pair of individuals had, on average, differences in 16–22 of a possible 54 alleles (27 odorant receptor genes with dose-response data × 2 alleles per subject). If we extrapolate to the ~400 intact odorant receptors, we expect each pair of individuals to differ at 237–326 of the 800 alleles. This suggests that odor detection at the peripheral level is highly variable. Variation at the peripheral level leads to variability in odor perception across individuals in several cases; in addition to the OR10G4-guaiacol association we demonstrated here, four olfactory perceptual phenotypes have previously been linked to a single odorant receptor genes[6–9] and five additional olfactory phenotypes have been linked to regions of the genome containing genes encoding more than one receptor[23–25]. Each individual, therefore, has a highly personalized set of olfactory receptors that affects his or her perception of odors.

We focused only on SNPs in the coding regions of the odorant receptor genes because of the lack of an efficient assay for testing the effects of noncoding polymorphisms on expression. That said, there is considerable variation in noncoding regions, which can lead to altered gene transcription[26] and even changes in sensory perception[27]. Similarly, we did not examine copy-number variation, which is widespread in human odorant receptors[28,29]. Thus, our data underestimate the potential extent of variation in each individual's expressed odorant receptor repertoire.

Our study did not find any evidence suggesting SNPs that alter *in vitro* function are restricted to a particular domain of the receptor, deviate from neutral evolution or are predicted by two popular computational alogrithms. Note, however, that our study was not designed to carefully detect changes resulting from a particular SNP; because we did not generate every possible combination of SNPs for the majority of odorant receptors, SNP-specific alterations may be confounded by linkage in the tested alleles.

Although we found that OR10G4 has at least three *in vitro* agonists, the *OR10G4* allele type only predicted perceived intensity and valence for guaiacol. The dose-response curves shown in **Figure 6a** reveal that guaiacol is a more potent agonist than either vanillin or ethyl vanillin. Although more data are needed, one possible interpretation is that the intensity and valence of odorants that only weakly activate a receptor will not be altered by functional variation in the receptor. Indeed, this is similar to the association between *OR7D4* genotype and androstenone[9]. In that case, the receptors encoded by both of the major alleles respond to androstenone *in vitro*, but one variant is much less potent than the other. As was the case for *OR7D4*, participants with the allele that encoded a receptor with lower *in vitro* affinity to the ligand found the odor to be less intense and more pleasant. This suggests that not all functional variation *in vitro* will lead to perceptual variation, but the exact rules determining how much of this variation is compensated for at later stages of processing will require further investigation.

Variation in *OR10G4* genotype explains 15.4% of the variance in perception of guaiacol intensity, which is lower than the 39% of variation in perception of androstenone intensity explained by variation in *OR7D4* genotype. The reason for this lower explanatory value is unclear. One possibility is that more odorant receptors have a role in the perception of guaiacol than in the perception of androstenone, therefore reducing the influence of a single odorant receptor on the percept. Another is

that confounding variables, such as culture and genetic background, may have differential effects on the two phenotypes.

The role of a single odorant receptor in olfactory perception is currently unknown, in part because of the large search space for both odorants and odorant receptors, and the redundant nature of the combinatorial code for odorant identity. By assigning ligands to odorant receptors, measuring the functional consequences of segregating polymorphisms *in vitro* and linking *in vitro* function to human behavior, these data provide a solid platform from which to probe the effects of a single odorant receptor on olfactory perception.

METHODS

Methods and any associated references are available in the online version of the paper.

Note: Any Supplementary Information and Source Data files are available in the online version of the paper.

ACKNOWLEDGMENTS

This work was supported by R01 DC005782, R01 DC012095, R03 DC011373, R01 DC013339, T32 DC000014 and a National Research Service Award postdoctoral fellowship F32 DC008932 to J.D.M. A portion of the work was performed using the Monell Chemosensory Receptor Signaling Core and Genotyping and DNA/RNA Analysis Core, which are supported, in part, by funding from the US National Institutes of Health NIDCD Core Grant P30 DC011735. A portion of the work was supported by the Defense Advanced Research Project Agency RealNose Project. Collection of psychophysical data was supported by grant # UL1 TR000043 from the Clinical and Translational Science Award program at the National Center for Advancing Translational Sciences. The FACS analysis was performed using the Duke Cancer Institute Flow Cytometry Core. We thank D. Marchuk for sharing equipment, L.B. Vosshall for supervising the collection of psychophysical data and DNA samples by A.K. in her laboratory, and R. Molday (University of British Columbia Centre for Macular Research) for 4D2 anti-rhodopsin antibody.

AUTHOR CONTRIBUTIONS

J.D.M. and H.M. conceived and designed the project. J.D.M., C.T., A.H.M., L.L.S., S.Z., W.L.L.L., T.Z., Y.R.L., H.Z., S.S.L., A.L. and K.A.A. performed research. A.K. collected the psychophysical data and provided DNA samples. J.D.M. carried out the analysis and wrote the paper with help from all authors. H.M. supervised the project.

COMPETING FINANCIAL INTERESTS

The authors declare no competing financial interests.

1. Menashe, I., Man, O., Lancet, D. & Gilad, Y. Different noses for different people. *Nat. Genet.* **34**, 143–144 (2003).
2. Hasin-Brumshtein, Y., Lancet, D. & Olender, T. Human olfaction: from genomic variation to phenotypic diversity. *Trends Genet.* **25**, 178–184 (2009).
3. Olender, T. *et al.* Personal receptor repertoires: olfaction as a model. *BMC Genomics* **13**, 414 (2012).
4. Ayabe-Kanamura, S. *et al.* Differences in perception of everyday odors: a Japanese-German cross-cultural study. *Chem. Senses* **23**, 31–38 (1998).
5. Amoore, J.E. Specific anosmia and the concept of primary odors. *Chem. Senses* **2**, 267–281 (1977).
6. Menashe, I. *et al.* Genetic elucidation of human hyperosmia to isovaleric acid. *PLoS Biol.* **5**, e284 (2007).
7. McRae, J.F. *et al.* Genetic variation in the odorant receptor OR2J3 is associated with the ability to detect the "grassy" smelling odor, cis-3-hexen-1-ol. *Chem. Senses* **37**, 585–593 (2012).
8. Jaeger, S.R. *et al.* A Mendelian trait for olfactory sensitivity affects odor experience and food selection. *Curr. Biol.* **23**, 1601–1605 (2013).
9. Keller, A., Zhuang, H., Chi, Q., Vosshall, L.B. & Matsunami, H. Genetic variation in a human odorant receptor alters odour perception. *Nature* **449**, 468–472 (2007).
10. Wetzel, C.H. *et al.* Specificity and sensitivity of a human olfactory receptor functionally expressed in human embryonic kidney 293 cells and *Xenopus laevis* oocytes. *J. Neurosci.* **19**, 7426–7433 (1999).
11. Spehr, M. *et al.* Identification of a testicular odorant receptor mediating human sperm chemotaxis. *Science* **299**, 2054–2058 (2003).
12. Sanz, G., Schlegel, C., Pernollet, J.C. & Briand, L. Comparison of odorant specificity of two human olfactory receptors from different phylogenetic classes and evidence for antagonism. *Chem. Senses* **30**, 69–80 (2005).
13. Matarazzo, V. *et al.* Functional characterization of two human olfactory receptors expressed in the baculovirus Sf9 insect cell system. *Chem. Senses* **30**, 195–207 (2005).
14. Jacquier, V., Pick, H. & Vogel, H. Characterization of an extended receptive ligand repertoire of the human olfactory receptor OR17–40 comprising structurally related compounds. *J. Neurochem.* **97**, 537–544 (2006).
15. Neuhaus, E.M., Mashukova, A., Zhang, W., Barbour, J. & Hatt, H. A specific heat shock protein enhances the expression of mammalian olfactory receptor proteins. *Chem. Senses* **31**, 445–452 (2006).
16. Schmiedeberg, K. *et al.* Structural determinants of odorant recognition by the human olfactory receptors OR1A1 and OR1A2. *J. Struct. Biol.* **159**, 400–412 (2007).
17. Saito, H., Chi, Q., Zhuang, H., Matsunami, H. & Mainland, J.D. Odor coding by a Mammalian receptor repertoire. *Sci. Signal.* **2**, ra9 (2009).
18. 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).
19. Keller, A., Hempstead, M., Gomez, I.A., Gilbert, A.N. & Vosshall, L.B. An olfactory demography of a diverse metropolitan population. *BMC Neurosci.* **13**, 122 (2012).
20. Zhuang, H. & Matsunami, H. Evaluating cell-surface expression and measuring activation of mammalian odorant receptors in heterologous cells. *Nat. Protoc.* **3**, 1402–1413 (2008).
21. Hayden, S. *et al.* Ecological adaptation determines functional mammalian olfactory subgenomes. *Genome Res.* **20**, 1–9 (2010).
22. The 1000 Genomes Project Consortium. A map of human genome variation from population-scale sequencing. *Nature* **467**, 1061–1073 (2010).
23. McRae, J.F. *et al.* Identification of regions associated with variation in sensitivity to food-related odors in the human genome. *Curr. Biol.* **23**, 1596–1600 (2013).
24. Eriksson, N. *et al.* Web-based, participant-driven studies yield novel genetic associations for common traits. *PLoS Genet.* **6**, e1000993 (2010).
25. Eriksson, N. *et al.* A genetic variant near olfactory receptor genes influences cilantro preference. *arXiv* 1209.2096 (2012).
26. Zhang, X. *et al.* Characterizing the expression of the human olfactory receptor gene family using a novel DNA microarray. *Genome Biol.* **8**, R86 (2007).
27. Fushan, A.A., Simons, C.T., Slack, J.P., Manichaikul, A. & Drayna, D. Allelic polymorphism within the TAS1R3 promoter is associated with human taste sensitivity to sucrose. *Curr. Biol.* **19**, 1288–1293 (2009).
28. Waszak, S.M. *et al.* Systematic inference of copy-number genotypes from personal genome sequencing data reveals extensive olfactory receptor gene content diversity. *PLoS Comput. Biol.* **6**, e1000988 (2010).
29. Nozawa, M., Kawahara, Y. & Nei, M. Genomic drift and copy number variation of sensory receptor genes in humans. *Proc. Natl. Acad. Sci. USA* **104**, 20421–20426 (2007).
30. Grantham, R. Amino acid difference formula to help explain protein evolution. *Science* **185**, 862–864 (1974).

ONLINE METHODS

Cloning. Odorant receptor open reading frames were amplified from the genomic DNA of 20 participants from the International Hapmap Consortium using Phusion polymerase and subcloned into pCI expression vectors (Promega) containing the sequence encoding the first 20 residues of human rhodopsin (Rho tag). The sequences of the cloned receptors were verified by sequencing (3100 Genetic Analyzer, Applied Biosystems).

Fluorescence-activated cell sorting. We conducted FACS analysis on all tested clones for the 17 odorant receptors where we had more than one clone (**Supplementary Fig. 5**). Hana3A cells were maintained in minimal essential medium (Sigma) containing 10% FBS (Sigma), 500 μg/ml penicillin-streptomycin (Invitrogen) and 6 μg/ml amphotericin B (Sigma). Cells were seeded in 35-mm dishes (Falcon) and grown overnight at 37 °C and 5% CO_2. The following day, each dish was transfected with plasmids using 4 μl Lipofectamine 2000 (Invitrogen), 1,200 ng Rho-tagged odorant receptor plasmid, 300 ng human receptor-transporting protein 1 short (RTPIS) plasmid and 20 ng of EGFP plasmid to control for transfection efficiency. 24 h after transfection, cells were washed with PBS and detached from the dishes using Cellstripper (Cellgro). Primary incubation was carried out at 4 °C using mouse monoclonal anti-rhodopsin 4D2 (ref. 31; gift from R. Molday) diluted 1:50 in PBS containing 2% FBS and 15 mM NaN_3 for 30 min. Cells were washed in PBS containing 2% (vol/vol) FBS and 15 mM NaN_3, followed by secondary incubation with Phycoerythrin (PE)-conjugated donkey anti-mouse antibody (Jackson Immunologicals) diluted 1:100 in PBS containing 2% FBS and 15 mM NaN_3 for 30 min, covered with aluminum foil. Cells were washed and resuspended in PBS containing 2% FBS and 15 mM NaN_3 containing 1:500 dilution of 7-aminoactinomycin D (7AAD; 1 mg/ml; Calbiochem), a fluorescent, cell-impermeant DNA-binding agent that selectively stains dead cells. Sorting of fluorescent cells was conducted using a BD FACSCanto (BD Biosciences). Cells that were EGFP-negative and/or 7AAD-positive were removed from subsequent analysis. Cell-surface expression was quantified as PE fluorescence intensity. Data collection and analysis were not randomized.

Luciferase assay. The Dual-Glo Luciferase Assay System (Promega) was used to measure receptor reponses as previously described[20]. Hana3A cells were transfected with 5 ng/well of RTP1S plasmid[32], 5 ng/well of pRL-SV40, 10 ng/well of luciferase driven by a cyclic AMP response element, 2.5 ng/well of M3 (ref. 33) and 5 ng/well of plasmids encoding odorant receptors. 1 M odorant stocks were diluted in DMSO. 24 h after transfection, transfection medium was removed and replaced with the appropriate concentration of odor diluted from the 1 M stocks in CD293 (Gibco). Four hours after odor stimulation, luminescence was measured using a Polarstar Optima plate reader (BMG). All luminescence values were divided by the *Renilla* luciferase activity to control for transfection efficiency in a given well. Data were analyzed with Microsoft Excel, GraphPad Prism 4 and Matlab (MathWorks).

1000 Genomes Project data. Allele frequency in the human population was derived from the May 2011 phased release of the 1000 Genomes Project public data (ftp://ftp-trace.ncbi.nih.gov/1000genomes/ftp/release/20110521/)[22]. Variant calls were obtained from the public repository in vcf format using tabix[34]. A custom-written Matlab script was used to translate the vcf file into 2,184 full-length phased alleles (two alleles for each of the 1,092 participants in the public database).

Human odorant receptor genotyping. All DNA samples were approved by the Rockefeller University Institutional Review Board. All subjects gave informed consent to participate and were financially compensated for their time and effort. Venous blood (8.5 ml) was collected from participants and genomic DNA was prepared with the Qiagen PAXgene blood DNA kit. For sequencing, human genomic DNAs were amplified with HotStar Taq (Qiagen) with primers upstream (5′-ACCTGGTTGATGCAGTTTCC-3′) and downstream (5′-AAACCTATTGATGAGAAATGAGTCAA-3′) of the OR10G4 open reading frame. The PCR products were then purified using Sephacryl S-400 (GE Healthcare) and sequenced with a 3100 or 3730 Genetic Analyzer (ABI Biosystems).

Procedures for olfactory psychophysics. All psychophysical data were obtained from ref. 19 and approved by the Rockefeller University Institutional Review Board. All subjects gave informed consent to participate and were financially compensated for their time and effort. Exclusion criteria for subjects were: allergies to odors or fragrances, a history of nasal illness, upper respiratory infection, seasonal allergy, prior endoscopic surgery on the nose, preexisting 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. Of the 308 subjects (138 male), 133 were Caucasian, 29 were Asian and 77 were African-American. The median age was 35 years, with a range of 19 to 66. In brief, participants rated the intensity and valence of 66 odorants on a 7-point scale. The intensity scale was labeled with 1 as "extremely weak" and 7 as "extremely strong." The valence scale was labeled with 1 as "extremely unpleasant" and 7 as "extremely pleasant." Stimuli were presented in jars. For a detailed description of the psychophysical methods, see ref. 9. Three of these odorants, ethyl vanillin, vanillin and guaiacol, are *in vitro* agonists to *OR10G4*. We examined the ratings of the higher of two tested concentrations. Ethyl vanillin and vanillin were presented at a 1/200 dilution in propylene glycol; guaiacol was presented at a 1/1,000,000 dilution in paraffin oil. Our data collection and analysis was blind to genotype, as all sequencing was conducted after phenotyping of the human subjects was complete. Data collection and analysis were not randomized.

Statistical analysis. *Screening procedure*. We stimulated the entire odorant receptor library with 73 odorants used in previous psychophysical testing[9]. We applied the odorants at 100 μM (except for androstenone and androstadienone, which were both applied at 10 μM) and ranked odorant-receptor pairs by their activity above the no-odor condition. We selected the top 5% of odorant-receptor pairs from this primary screen; some receptors were very promiscuous, so we tested only the top ten ligands for a given receptor. We then performed a secondary screen in which each odorant receptor was tested against a no-odor control as well as 1 μM, 10 μM and 100 μM of agonists identified in the primary screen. Each comparison was performed in triplicate, where each measure was collected from separate wells, but each well contains cells from the same parent plate of cells. Statistical significance was assessed by two-sided *t*-test comparing the three wells stimulated with odor with the three wells stimulated with medium alone. As this was a screening procedure, the data distribution was assumed to be normal, but this was not formally tested. In addition, the tests were uncorrected for multiple comparisons. We then constructed dose-response curves using concentrations ranging from 10 nM to 10 mM for the odor-receptor pairs that were significantly different from baseline in the secondary screen (*t*-test, $P < 0.05$, uncorrected for multiple comparisons). Each odorant receptor–odorant dose was tested in triplicate, where each measure was collected from separate wells, but each well contained cells from the same parent plate of cells, and a vector-only control was included for each odorant. We fit the data to a sigmoidal curve. We counted an odorant as an agonist if the 95% confidence intervals of the top and bottom parameters did not overlap, the s.d. of the fitted log(EC50) was less than 1 log unit and the extra sums-of-squares test confirmed that the odorant activated the receptor significantly more than the control, which was transfected with an empty vector (extra sums-of-squares *F* test against the vector control, $P < 0.05$). Data collection and analysis were not randomized.

Screening 55 odorants. To choose 55 odorants that quantitatively span chemical space, we generated 20 physicochemical descriptors that predict 62% of the variance in mammalian odorant receptor responses[17] for 2,715 commonly used odorants. We then divided the 2,715 odorants into 55 clusters using *k*-means clustering. For each cluster, we selected the odorant closest to the centroid of the cluster among odorants that are previously shown to activate at least one odorant receptor. If no such agonist was present in the cluster, we selected the odorant closest to the centroid of the cluster to maximize structural diversity. Each odorant was screened against each receptor variant at 100 μM in triplicate where each measure was collected from separate wells but each well contains cells from the same parent plate of cells. We performed an analysis of variance (ANOVA) on the responses from the clones of each odorant receptor. We used 15 odorant receptors where we had more than one allele cloned with an allele frequency greater than 1% in the 1,092 participants and the cloned alleles represented a large percentage of the 2,184 alleles. For 13 odorant receptors, the cloned alleles represented more than 85% of the 2,184 alleles. For *OR2B11*, the cloned alleles represented 37.5% of the alleles, and for *OR10G4*, the cloned alleles represented 29.5% of the alleles. Data collection and analysis were not randomized.

Dose-response curves. We tested odorant receptors with odorants ranging in concentration from 10 nM to 10 mM. All numerical results are reported as

doi:10.1038/nn.3598

mean ± s.e.m. and represent data from a minimum of three replicates, where each measure was collected from separate wells, but each well contains cells from the same parent plate of cells. We fit the resulting data with a three-parameter logistic model. We counted an odorant as an agonist if the 95% confidence intervals of the top and bottom parameters did not overlap, the s.d. of the fitted log EC50 was less than 1 log unit and the extra sums-of-squares test confirmed that the odorant activated the receptor significantly more than was the case for the vector-only control (extra sums-of-squares F test against the vector control, $P < 0.05$).

For each pair of alleles, we determined whether one model fit the data from both alleles better than two separate models using the extra sums-of-squares test. A pair of alleles was classified as hyper- or hypofunctional if one allele in the pair had both a higher EC50 (lower efficacy) and a lower potency (dynamic range or top-bottom). A pair of alleles was designated as 'unclassified' if the potency and efficacy showed discordant changes (i.e., one allele was more sensitive but had a lower efficacy).

To compare each pair of individuals, we took the four alleles from a single odorant receptor and removed any pairs of alleles that were indistinguishable according to the above criteria. Each remaining pair was counted as one functional difference. These values were summed across odorant receptors, with a maximum of 48 possible functional differences per pair of participants. Data collection and analysis were not randomized.

Odds that a SNP alters function. We aligned the nucleotide sequences of the odorant receptor variants to a multiple-sequence alignment of 1,425 intact mouse and human odorant receptors. For each SNP, we calculated the ratio of the odds that a functional change (as defined above, relative to the most common functional variant) occurred in an allele with a nonsynonymous amino acid to the odds that a functional change occurred in an allele with a synonymous amino acid. We used SNPnexus[35] (Ensembl 63 build) to generate GERP, SIFT and PolyPhen scores.

Multiple linear regression model. Multiple regression analysis was used to test whether the number of *OR10G4* alleles significantly predicted participants' perception of the three *in vitro* agonists. To determine the minimum sample size for this analysis, we performed a Monte-Carlo simulation using the data from ref. 9. We ranked each subject's ratings of the odorants to control for differences in general olfactory acuity and usage for the rating scale across subjects. The predictors were allele counts (0, 1 or 2) for the four alleles with MAF > 4% in the participant population. Data collection and analysis were not randomized.

31. Laird, D.W. & Molday, R.S. Evidence against the role of rhodopsin in rod outer segment binding to RPE cells. *Invest. Ophthalmol. Vis. Sci.* **29**, 419–428 (1988).
32. Zhuang, H. & Matsunami, H. Synergism of accessory factors in functional expression of mammalian odorant receptors. *J. Biol. Chem.* **282**, 15284–15293 (2007).
33. Li, Y.R. & Matsunami, H. Activation state of the m3 muscarinic acetylcholine receptor modulates mammalian odorant receptor signaling. *Sci. Signal.* **4**, ra1 (2011).
34. Li, H. Tabix: fast retrieval of sequence features from generic TAB-delimited files. *Bioinformatics* **27**, 718–719 (2011).
35. Chelala, C., Khan, A. & Lemoine, N.R. SNPnexus: a web database for functional annotation of newly discovered and public domain single nucleotide polymorphisms. *Bioinformatics* **25**, 655–661 (2009).