Different noses for different people

Idan Menashe, Orna Man, Doron Lancet & Yoav Gilad

Of more than 1,000 human olfactory receptor genes, more than half seem to be pseudogenes. We investigated whether the most recent of these disruptions might still segregate with the intact form by genotyping 51 candidate genes in 189 ethnically diverse humans. The results show an unprecedented prevalence of segregating pseudogenes, identifying one of the most pronounced cases of functional population diversity in the human genome.

The olfactory receptor repertoire is the largest gene superfamily in mammals, including 1,000–1,400 coding regions distributed in clusters over most chromosomes[1–3]. In the mouse, olfactory receptor pseudogenes comprise 20% of the gene range; in humans, a fraction roughly three times larger seems to be inactivated[1,3]. This extreme diminution of the functional olfactory receptor repertoire is a relatively recent genomic process[4] and is probably ongoing. Accordingly, we conjectured that a substantial fraction of the human olfactory receptors might segregate between an intact and pseudogene form. Indeed, isolated cases of olfactory segregating pseudogenes have been reported[5–7]. Yet, these could underlie only a small part of the reported widespread phenotype variation[8,9]. We therefore launched a whole-genome search for single-nucleotide polymorphisms (SNPs) that exchange between the intact and pseudogenic forms in olfactory receptor loci.

We used two strategies to screen for such polymorphisms. First, we focused on olfactory receptor pseudogenes that have only one open reading frame disruption[1]. We sequenced 50 of these in a chimpanzee, owing to the notion that if the pseudogene state is not shared between the two higher apes, it may be recent and thus have a better chance of generating a human polymorphism. This identified 33 olfactory receptor loci for further scrutiny. Second, we searched the Celera human SNP database for variations with potential to affect protein integrity. In this realm, we included 9 cases of in-frame stop codons and 9 cases in which a SNP represented a missense change in a highly conserved amino acid and, hence, was probably functionally important[10].

We genotyped the total of 51 olfactory receptor loci in 189 individuals from several ethnic origins. We confirmed that 26 of these loci segregate in our sample, either between the disrupted and intact

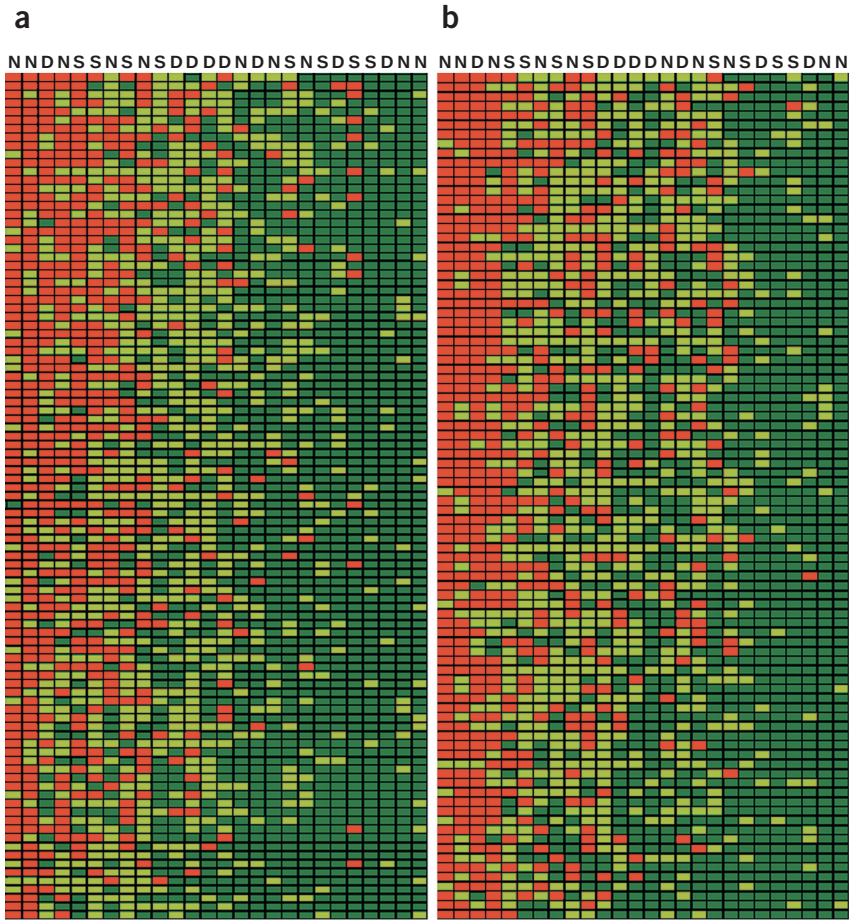

Figure 1 The observed individual olfactory receptor genotypes in African American (**a**) and non-African (**b**) individuals. Red indicates homozygous olfactory receptor disruption; dark green denotes homozygously intact olfactory receptor; light green represents heterozygotes. Individuals (rows) and olfactory receptor loci (columns) are ordered according to their disruption level. Disruption type is indicated above: N, nonsense mutation; D, deletion or insertion; S, missense substitution at a highly conserved site. The 178 putative phenotypes seen in the 189 individuals based on recessive inheritance are in accordance with a computed expectation value assuming Hardy–Weinberg equilibrium and based on the allele frequencies present.

Department of Molecular Genetics and the Crown Human Genome Center, The Weizmann Institute of Science, Rehovot 76100, Israel. Correspondence should be addressed to D.L. (doron.lancet@weizmann.ac.il).

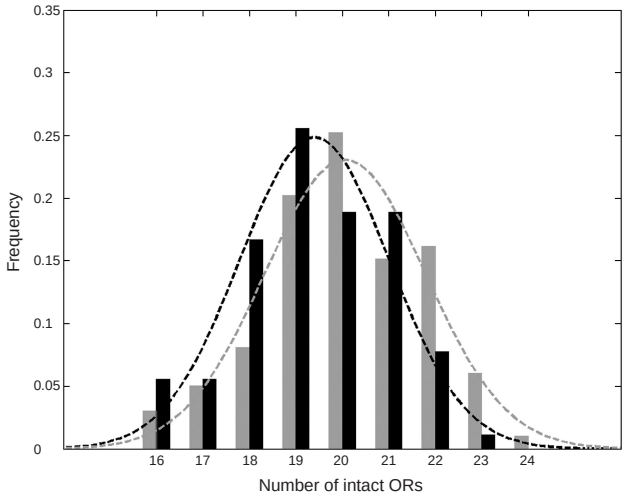

Figure 2 Phenotype distribution for different ethnic groups. The population frequency of the counts of deduced functional loci in African American (light bars) and non-African (dark bars) individuals. Olfactory receptor loci were considered intact if the individual carried at least one copy of the intact allele. The broken line curves are Gaussian fits with $\mu = 20.06$ and $\sigma = 1.73$ for African American individuals and $\mu = 19.40$ and $\sigma = 1.61$ for non-African individuals. The statistical significance of the difference in the number of intact olfactory receptors (ORs) between African American and non-African individuals was assessed by a Mann–Whitney U test ($P = 0.0022$) or by a χ^2 test ($P = 0.014$). A comparable P value was obtained if we did not assume dominance of the functional allele and used the three different genotypes at each locus (Mann–Whitney U test, $P = 0.0001$; χ^2 test, $P = 0.015$).

alleles (18 loci) or between the conserved and modified forms (8 loci; **Fig. 1**). Consequently, each of the individuals examined had a unique genotypic pattern. The resulting phenotypic effect may be evaluated on the basis of the suggestion that olfactory dysfunction is a recessive trait[9,11]. We observed great diversity: 178 different functional genomes among the 189 individuals studied. Such a high level of documented interindividual variability in a gene family is unprecedented, except in the case of the major histocompatibility complex. The latter includes the most polymorphic loci in humans; two of the primary genes, *HLA-B* and *HLA-DRB1*, have ~200 allelic variants each, and every individual carries a unique haplotype signature[12]. In the case of olfactory receptor genes, each locus is only biallelic, but the number of functionally variable loci is much larger.

Notably, non-African individuals had significantly fewer functional olfactory receptors than did African American individuals (**Fig. 2**). This result substantiates our previous reports[6,7] suggesting that different evolutionary pressures may have shaped the chemosensory repertoire in different human populations.

We extrapolated the number of segregating olfactory receptor pseudogenes in the entire human genome to be at least 60, of which 48 are expected to have a minor allele frequency above 1% (**Supplementary Note** and **Supplementary Table 1** online). These numbers are in rough agreement with the reported count of different modes of specific anosmias, human odorant-specific sensory deficits[8]. Thus, the genotypic disparity that we observed might underlie at least some of the reported human phenotypic variation. Future association studies will help substantiate the detailed relationships between individual olfactory receptor disruptions and defined cases of odorant-specific olfactory threshold variability.

Most genotype–phenotype association studies are based either on rare gene disruptions (that is, mutations underlying monogenic traits) or on combinations of frequent variants involving missense rather than nonsense DNA alterations. The widespread occurrence of segregating olfactory pseudogenes that we report is rather unusual, and only a few analogous cases have been described[13].

It is interesting to ask whether this functional polymorphism is unique to olfactory receptors, stemming from their gene superfamily undergoing a recent rapid decline in the number of functional genes[4]. An intriguing alternative is that this phenomenon is more widespread and affects other gene families with high expansion rate and low selective pressure due to partial functional redundancy. A potential target for a relevant genome-wide search could be the ~20,000 pseudogenes estimated to be present in the human genome[14]. Although many of these are processed pseudogenes or may never have been functional, it is possible that an appreciable number of pseudogenes, which have arisen by recent mutations, will eventually be shown to constitute segregating null alleles. These would comprise a hitherto unexplored domain of human genotypic heterogeneity.

Note: Supplementary information is available on the Nature Genetics website.

ACKNOWLEDGMENTS

We thank M. Przeworski for helpful discussions. D.L. holds the Ralph and Lois Silver Chair in Human Genomics. This work was supported by the Crown Human Genome Center at the Weizmann Institute of Science and by an Israel Ministry of Science grant to the National Laboratory for Genome Infrastructure.

COMPETING INTERESTS STATEMENT

The authors declare that they have no competing financial interests.

Received 15 January; accepted 28 March 2003
Published online 5 May 2003; doi:10.1038/ng1160

1. Glusman, G., Yanai, I., Rubin, I. & Lancet, D. *Genome Res.* **11**, 685–702 (2001).
2. Zhang, X. & Firestein, S. *Nat. Neurosci.* **5**, 124–133 (2002).
3. Young, J.M. *et al. Hum. Mol. Genet.* **11**, 535–546 (2002).
4. Gilad, Y., Man, O., Paabo, S. & Lancet, D. *Proc. Natl. Acad. Sci. USA* **100**, 3324–3327 (2003).
5. Ehlers, A. *et al. Genome Res.* **10**, 1968–1978 (2000).
6. Menashe, I., Man, O., Lancet, D. & Gilad, Y. *Hum. Mol. Genet.* **11**, 1381–1390 (2002).
7. Gilad, Y. & Lancet, D. *Mol. Biol. Evol.* **20**, 307–314 (2003).
8. Amoore, J.E. *Ann. NY Acad. Sci.* **237**, 137–143 (1974).
9. Whissell-Buechy, D. & Amoore, J.E. *Nature* **242**, 271–273 (1973).
10. Schoneberg, T., Schulz, A. & Gudermann, T. *Rev. Physiol. Biochem. Pharmacol.* **144**, 143–227 (2002).
11. Wysocki, C.J. & Beauchamp, G.K. *Proc. Natl. Acad. Sci. USA* **81**, 4899–4902 (1984).
12. Yeager, M. & Hughes, A.L. *Immunol. Rev.* **167**, 45–58 (1999).
13. Barron, K.S. & Robinson, M.A. *Hum. Immunol.* **40**, 17–19 (1994).
14. Harrison, P.M. *et al. Genome Res.* **12**, 272–280 (2002).