Genes and smells

Andreas Keller (The Rockefeller University, New York, USA)

The recently sequenced genomes of several vertebrate species revealed large differences between the odorant receptor (OR) repertoires of even closely related species. These differences are a consequence of frequent gene duplications and subsequent mutations of duplicated genes. The same mechanisms are at work within a species, and different human genomes therefore show an unparalleled variability in the number and sequence of OR genes. Recent research has started to explore the consequences of this genetic variability for how we perceive odours.

In 1991, Linda Buck and Richard Axel discovered that, in vertebrates, odours are sensed by members of a family of seven-transmembrane G-protein-coupled receptors, the vertebrate-like ORs. Since this discovery of the molecular substrate of the sense of smell, our understanding of this most mysterious of our senses has changed dramatically. In this article, I review recent discoveries about the evolution and function of OR genes.

Odorant receptors in the genome

The first vertebrate-like OR gene evolved from a rhodopsin-like gene in the chordate lineage after the echinoderm lineage (which includes the sea urchins) split off around 550 million years ago. As a consequence, sea urchins have no OR genes, whereas most chordates and all vertebrate species that have been studied so far have at least some OR genes[1].

It is remarkable how much the number of OR genes varies between species (Figure 1). There are only two OR genes in the shark genome[1], but elephants have 4364 ORs[2]. Huge differences in the size of the OR family are also found in more closely related species. Among mammals, the armadillo has 3146 ORs, whereas the dolphin has only 26. Within the primates, the number of ORs ranges from 344 in the tarsier to 977 in the mouse lemur[2]. Humans have about 800 OR genes.

A second remarkable feature of the OR gene family is that many of its members are pseudogenes. Pseudogenes are DNA sequences that show many features of genes, yet they are not translated into functional proteins due to genetic alterations such as stop codons. In humans, about half of the OR genes are pseudogenes. Some researchers concluded from this that in humans olfactory abilities are rapidly degrading and looked for explanations of why this might be the case. One suggestion was that the acquisition of trichromatic colour vision reduced adaptive pressure on the OR genes[3].

Key words: *genomic drift, odorant receptor, odour perception, OR7D4 gene, pseudogene*

Figure 1. The OR gene family in mammals. Vastly different numbers of intact OR genes and OR pseudogenes are found in the genomes of 32 mammals. Data from Hayden et al.[2].

It is now clear that the high percentage of OR pseudogenes is not a specific feature of the human genome. Most of the 50 mammalian species studied by Hayden et al.[2] have a higher percentage of pseudogenes than humans. The percentage of OR pseudogenes ranges in mammals from 9% in the mouse to 85% in the dolphin (Figure 1)[2].

Evolution of ORs

What caused the large differences in the size of the OR gene family between species and the high percentage of OR pseudogenes in most species? It has been suggested that OR gene repertoires are shaped by a process that was termed 'genomic drift', a random process of gene duplications and deletions that is analogous to genetic drift, the stochastic change of allele frequencies in a population[4]. This process was also termed 'birth-and-death evolution' because of the large number of events that are equivalent to the 'birth' and 'death' of genes. In the human genome, for example, 212 intact OR genes were lost through mutation and 57 gained through duplication since the most recent common ancestor of the primates 48 million years ago[5] (Figure 2). Some 22 of the ORs that were gained in the human lineage were gained since we split from the chimpanzee lineage and are therefore uniquely human ORs that may sense odours of special relevance for our species[5].

Genomic drift shapes the OR family through random processes. If there is no adaptive advantage or disadvantage in having two copies of an OR instead of one, then OR genes can duplicate at random and duplicated OR genes can mutate at random. This explains the variability in the size of the OR family (through duplication) and the high percentage of OR pseudogenes (through mutation). The challenge is to find evidence for adaptive selection among the remnants of these chance events. The clearest evidence for adaptive processes shaping the OR family comes from comparing aquatic and terrestrial animals. The OR repertoires of fish are characteristically different from those of terrestrial vertebrates, and the repertoire of the frog represents an intermediate state[1]. Furthermore, the re-adaptation of terrestrial vertebrates to an aquatic environment in dolphins and sea snakes is accompanied by dramatic changes in OR genes. Hayden et al.[2] analysed more than 50 000 OR genes from 50 mammalian species and found that certain types of OR genes are more common in aquatic species and others are more common in terrestrial species.

It is interesting that Young et al.[6] found the same variability in gene number and percentage of pseudogenes for another family of genes encoding

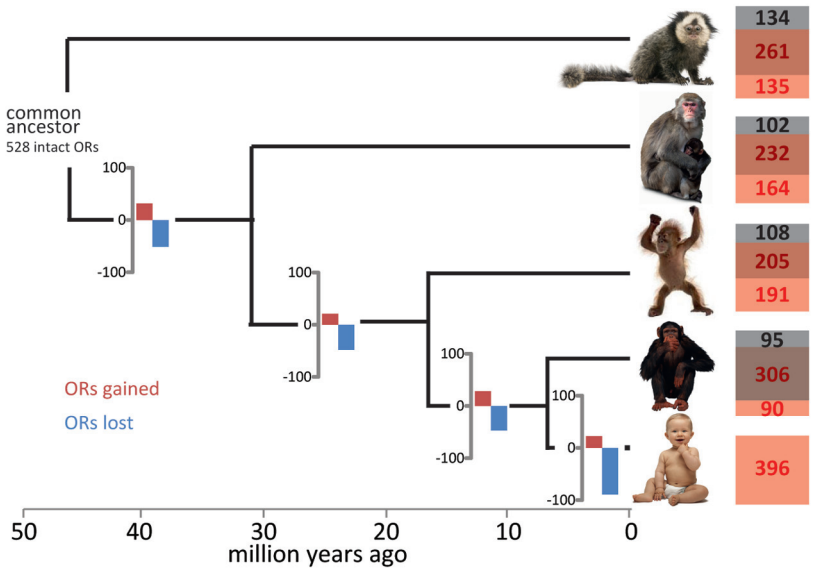

Figure 2. Evolution of the OR gene family in primates. The repertoire of intact OR genes changed rapidly during primate evolution. The intact ORs that are gained through duplication and that are lost through deletion or pseudogenization in each step are shown in the inserted charts in the phylogenetic tree. On the right, the overlap between the human intact OR repertoire (red) and the intact OR repertoire of the other species (grey) is shown. Data from Matsui et al.[5].

receptors involved in odour sensing, the V1R family. The 37 mammalian genomes that they studied range from 24 V1R genes in the little brown bat to 1405 in the platypus. The percentage of V1R genes that are pseudogenized varies from 17% in the mouse lemur to 100% in two bat species, and the macaque and the dolphin, which have no functional V1R genes. This shows that the evolutionary processes that shaped the OR gene repertoires are not specific for this family, but a common feature of the evolution of gene families.

Variability in human ORs

The duplications and mutations that shaped the OR repertoires of different species are genetic processes that occur in individuals. If genomic drift plays an important role in OR evolution, one would therefore expect to find large variability not only between species, but also within a species. So far, the only vertebrate species in which a large number of individuals has been sequenced is humans and it was indeed found that the human OR genes are extraordinarily variable.

At the genomic level, the number of OR genes varies dramatically from one person to another because of duplications and deletions. Waszak et al.[7] investigated 150 human genomes for copy number variations in OR genes. They found copy number variations for 307 OR genes. The person with the fewest OR genes in their study smells odours with 77 receptors less than the person with the most OR genes[7].

Vast variability between human genomes is also seen at the sequence level within OR genes. In the most extreme cases, a mutation can turn an intact OR into a pseudogene. There are over 60 human ORs known that are segregating pseudogenes, which means that some individuals carry a pseudogenized variant, whereas others carry an intact variant of the gene[8]. The completion of the 1000 Genomes Project (www.1000genomes.org) will undoubtedly show that an even larger percentage of human ORs are segregating pseudogenes.

Some mutations do not result in pseudogenization, but merely change the

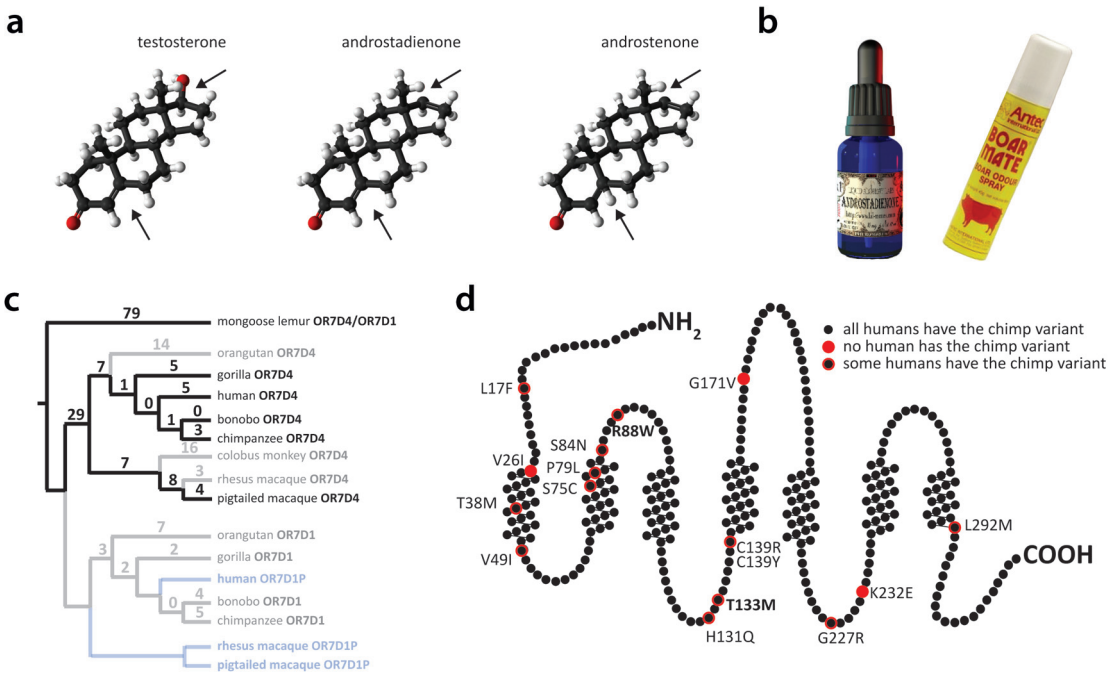

Figure 3. OR7D4 structure and function. (**a**) OR7D4 is specifically activated by odorous steroids such as androstenone and androstadienone which are structurally similar to the sex hormone testosterone. (**b**) Androstadienone is the ingredient in commercially available products that are meant to attract women (on the left) and to induce mating stance in female pigs (on the right). (**c**) The evolution of OR7D4 in the primate lineage involved a duplication of an ancestral OR gene into *OR7D4* and *OR7D1*. Pseudogenes are shown in blue. Variants that were determined to code for non-functional proteins are shown in grey. The numbers represent the changes in amino acid residues at each step. Data from Zhuang et al.[15]. (**d**) OR7D4 is a 312-amino-acid-long seven-transmembrane protein.

structure of the OR protein, thereby potentially altering the odour-sensitivity. Mutations that alter a single amino acid residue in a protein are known as missense single nucleotide polymorphisms (SNPs). Missense SNPs are very common in OR genes. An average of about 10 missense SNPs per OR gene are known, and many of these variants occur with high frequency (NCBI dbSNP, BUILD 132).

What do these huge differences between two individuals' OR repertoires mean for how they perceive their odorous environment? There is considerable variability in odour perception. Some people are exquisitely sensitive to odours that others can only smell at very high concentrations or not at all[9]. The odour excreted in the urine after eating asparagus, for example, can be detected by most people, but not by everyone. Furthermore, differences in perceived odour pleasantness, odour quality and similarity between two odours have been described. It is tempting to speculate that at least part of the variability of the sense of smell has to do with the fact that we all smell odours with very different repertoires of receptors.

A human OR for steroids

To study the connection between an OR gene and odour perception, one needs to know which odour the OR is detecting. Of the over 50 000 OR genes that have been identified to date, it is only known for about 100 to which odour the OR is sensitive[10]. The human OR7D4, for example, responds specifically to odorous steroids that are similar in their structure to testosterone[11] (Figure 3a). These odorous steroids, such as androstenone and androstadienone, are found in human excretions, including sweat, urine, saliva and semen. They occur at higher concentrations in males than in females[12]. The same is true in pigs, where odorous steroids are produced and secreted in large amounts by boars. Oestrous female pigs are attracted to these odours and they are more likely to get into the rigid mating posture when exposed to them[13]. Some researchers have suggested that odorous steroids act as pheromones in humans too[14] (Figure 3b).

The *OR7D4* gene is only found in mammals. There are no *OR7D4* orthologues in non-mammalian vertebrates such as chicken and zebrafish. Within

mammals, several primates, as well as dogs, cows, mice and rats, have an *OR7D4* orthologue (www.genecards.org). The gene was probably lost in some mammalian lineages, and no orthologue has so far been found in pigs or in new world monkeys. The closest paralogue of *OR7D4* in the human genome is the pseudogenized *OR7D1* that is probably the result of an *OR7D4* duplication after the divergence of the new world monkeys[15] (Figure 3c). The chimpanzee OR7D4 orthologue has three amino acid residues that have not yet been found in humans (Figure 3d).

In terms of variability between individuals, *OR7D4* is probably representative of human ORs. No copy number variation has been found for the gene among 150 human genomes. In total, 14 missense SNPs and six synonymous SNPs are known (NCBI dbSNP, BUILD 132). Freely combining the 14 missense SNPs would result in over 15 000 different OR7D4 proteins. Five of these variants were found in a study of 391 individuals[11]. The two most frequent variants (in this population) are OR7D4 RT and OR7D4 WM which differ in two linked amino acid residues[11] (Figure 3d).

In vitro experiments revealed that OR7D4 RT, but not OR7D4 WM, responds to the odorous steroids[11]. Consequently, individuals with the non-functional variant were less sensitive to both androstenone and androstadienone than individuals carrying the functional variant. The perceived pleasantness and what the odours smelled like were also dependent on what OR7D4 variant the individual carried[11].

Because odorous steroids are an important part of human body odour, differences in the *OR7D4* gene may have behavioural consequences beyond the psychophysical tests in a smell laboratory. In a suggestive study, it was shown that the perceived pleasantness of the smell of an individual's heterosexual partner was significantly associated with a SNP in OR7D4[16]. Because the SNP under investigation in this study is synonymous and does not affect the protein, the causal relationship between the OR7D4 SNP and this phenotype will require further investigation.

Perspective

Owing to the evolutionary processes that shape the OR gene family, there is large variability not only between species, but also between members of the same species. As a consequence, each person smells their odorous environment with a different set of receptors. Studying odour-guided behaviours in humans therefore offers the unique opportunity to directly study the effects of individual genes on human behaviour. Studies of the first gene for which this has been made possible, *OR7D4*, are already starting to produce interesting results and several hundred more ORs are waiting to have their role in human behaviours and social interactions uncovered. ■

I am supported by a Branco Weiss fellowship from the Society in Science Foundation and by a Young Investigator Award from NARSAD. I thank Jessica Keiser, Joel Mainland and Leslie Vosshall for comments on earlier versions of this article.

Andreas Keller received his PhD from the University of Würzburg in Germany where he studied visually guided behaviours in Drosophila melanogaster *in the laboratory of Martin Heisenberg. He then joined the laboratory of Leslie B. Vosshall at the Rockefeller University in New York where he studies the genetic basis of odour perception in humans. email: kellera@rockefeller.edu*

References

1. Niimura, Y. (2009) Genome Biol. Evol. **1**, 34–44
2. Hayden, S., Bekaert, M., Crider, T.A., Mariani, S., Murphy, W.J. and Teeling, E.C. (2010) Genome Res. **20**, 1–9
3. Gilad, Y., Wiebe, V., Przeworski, M., Lancet, D. and Paabo, S. (2004) PLoS Biol. **2**, E5
4. Nei, M., Niimura, Y. and Nozawa, M. (2008) Nat. Rev. Genet. **9**, 951–963
5. Matsui, A., Go, Y. and Niimura, Y. (2010) Mol. Biol. Evol. **27**, 1192–1200
6. Young, J.M., Massa, H.F., Hsu, L. and Trask, B.J. (2010) Genome Res. **20**, 10–18
7. Waszak, S.M., Hasin, Y., Zichner, T. et al. (2010) PLoS Comput. Biol. **6**, e1000988
8. Hasin–Brumshtein, Y., Lancet, D. and Olender, T. (2009) Trends Genet. **25**, 178–184
9. Amoore, J.E. (1977) Chem. Senses Flavour **2**, 267–281
10. Zhuang, H., Chien, M.S. and Matsunami, H. (2009) Proc. Natl. Acad. Sci. U.S.A. **106**, 21247–21251
11. Keller, A., Zhuang, H., Chi, Q., Vosshall, L.B. and Matsunami, H. (2007) Nature **449**, 468–472
12. Gower, D.B. and Ruparelia, B.A. (1993) J. Endocrinol. **137**, 167–187
13. Dorries, K.M., Adkins-Regan, E. and Halpern, B.P. (1997) Behav. Brain Sci. **49**, 53–62
14. Pause, B. (2004) Physiol. Behav. **83**, 21–29
15. Saito, H., Chi, Q., Zhuang, H., Matsunami, H. and Mainland, J.D. (2009) Sci. Signaling **2**, ra9
16. Sookoian, S., Burgueño, A., Gianotti, T.F., Marillet, G. and Pirola, C.J. (2011) Biol. Psychol. **86**, 153–157