Olfactory Receptor Genes: Evolution

Yoshihito Niimura, *Department of Applied Biological Chemistry, Graduate School of Agricultural and Life Sciences, The University of Tokyo, Tokyo, Japan and ERATO Touhara Chemosensory Signal Project, JST, The University of Tokyo, Tokyo, Japan*

Advanced article

Article Contents

Online posting date: 15th August 2014

Many mammal genomes have approximately 1000 genes encoding olfactory receptors (ORs), and OR genes constitute the largest multigene family in mammals. Comparisons among the OR gene repertoires in a broad range of species demonstrates that gene duplication and pseudogenization cause frequent gene gain and loss in this family, causing drastic evolutionary changes in the number of genes depending on species' ecological niches and other sensory modalities. For example, higher primates are equipped with a well-developed visual system, and they have a reduced OR gene repertoires relative to mammals with lesser visual systems. Additionally, aquatic and terrestrial vertebrates retain different sets of OR genes, and these sets reflect the capacity to detect water-soluble and airborne odorants, respectively. The origin of vertebrate OR genes can be traced back to the common ancestor of chordates, but insects and nematodes each use a distinct family of genes to encode chemoreceptors; therefore, multiple distinct chemoreceptor gene families emerged independently during animal evolution.

Introduction

Among the five senses, olfaction, the sense of smell, may seem to be the least important for humans. However, the sense of smell is essential to our humanity – emotionally, physically, sexually and socially (Herz, 2007). Loss of olfaction severely affects a person's quality of life. For many animal species, olfaction is of the great importance to survival and fitness. Olfactory signals are used to find food, identify mates and offspring, recognise territories and avoid danger. Moreover, some animal species have a much more refined and yet broader olfactory system than humans.

When a molecule of β-phenylethyl alcohol enters your nose, your brain interprets it as a rose-like fragrance. But why does that molecule have the scent of roses? Actually, the relationship between odour molecules and the perceived odours is enigmatic. Certainly, there are many cases in which molecules with an identical or similar functional group are perceived as similar odours. For example, carboxylic esters usually exhibit pleasant fruity odours. However, molecules with similar structures can be perceived as different odours (**Figure 1a** and **b**); conversely, molecules that are completely different structurally can be perceived as similar odours (**Figure 1c** and **d**). Therefore, the relationships between the structure of odorant molecules (stimulus) and odours (perception) are complicated. Still, no existing general rules can be used to reliably predict a perceived odour based on a given molecular structure.

The olfactory system contrasts sharply with the colour vision system. Humans can normally see light (electromagnetic waves) with a wavelength between approximately 380 and 780 nm. As the wavelength of the light (stimulus) changes gradually, the colour (perception) also changes continuously from blue to red along the visible spectrum. Light is detected by visual pigments in photoreceptor cells in the retinas of eyes. Each visual pigment comprises a protein named opsin and a chromophore named retinal. Most humans (except for colour-blind people) have three opsin genes in their genome; therefore, they produce three different types of visual pigments. Each type is activated by a specific range of wavelengths. Each of the three ranges corresponds to red, green or blue. Consequently, most humans are trichromats and have a colour vision system by which all perceivable colours can be reproduced by an appropriate mixture of three primary colours, red, green and blue. **See also**: Visual Pigment Genes: Evolution

eLS subject area: Evolution & Diversity of Life

How to cite:
Niimura, Yoshihito (August 2014) Olfactory Receptor Genes: Evolution. In: eLS. John Wiley & Sons, Ltd: Chichester.
DOI: 10.1002/9780470015902.a0020789.pub2

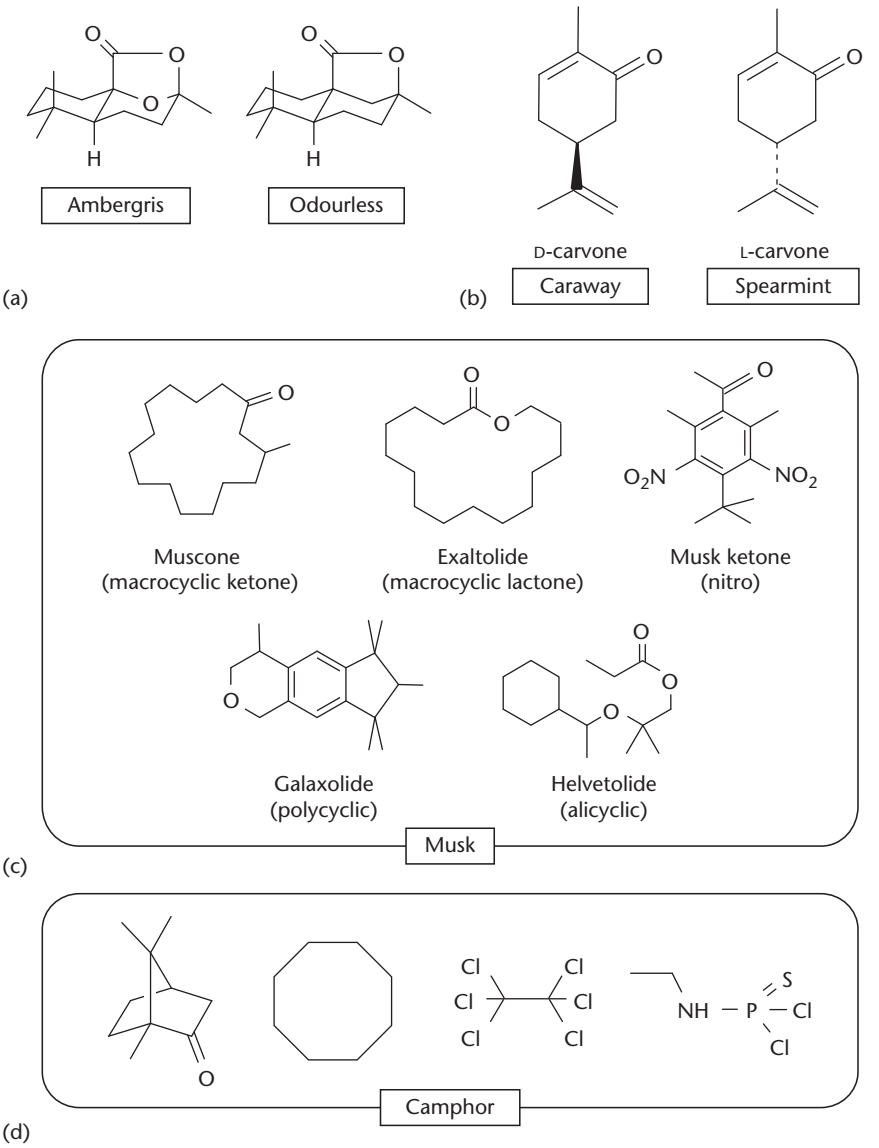


Figure 1 Complexity of structure–odour relationships (Rossiter, 1996). (a) The molecule on the left conveys a strong ambergris odour, but molecule on the right is odourless because it lacks just one of the oxygen atoms. (b) Enantiomers that have different odours. (c) Five very different molecular structures all have musky odours. (d) Four very different molecular structures all have camphoraceous odours, though there is no one functional group that is common to each of them.

Odour molecules in the environment are detected by olfactory receptors (ORs) that are expressed in the olfactory epithelium of the nasal cavity. Surprisingly, most mammals have as many as approximately 1000 OR genes. Mammalian genomes generally encode 20 000–25 000 genes; therefore, 4–5% of a typical mammalian proteome is dedicated to odour detection. OR genes constitute the largest multigene family in mammals. The human genome contains approximately 400 (see below) OR genes; this number is yet much larger than the number of opsin genes. The presence of a large OR gene repertoire in part explains why odour perception is so complicated. Olfaction, unlike vision, does not involve a small number of 'primary odour' which can generate any perceivable odours by their appropriate mixture.

OR genes were first identified in rats by Buck and Axel (1991). They discovered a huge multigene family that encodes G-protein coupled receptors (GPCRs) of which expression is restricted to the olfactory epithelium. Their discovery opened the door for the molecular studies of chemical senses, and they were awarded the Nobel Prize in Physiology or Medicine in 2004. Subsequently, genes with considerable homology to rat OR genes were found in the olfactory epithelium of channel catfish; therefore, other vertebrates also use chemoreceptors similar to mammalian ORs (Ngai *et al.*, 1993). Studies in subsequent decades

revealed that various types of non-OR genes are also involved in chemosensation, including pheromone and taste detection. Currently, seven different multigene families are known to be involved in vertebrate chemosensation: ORs, vomeronasal receptors type 1 and type 2 (V1Rs and V2Rs), trace amine-associated receptors (TAARs), formyl peptide receptors (FPRs) and taste receptors type 1 and type 2 (T1Rs and T2Rs) (Niimura, 2012a). The OR gene family is by far the largest of these families. Chemoreceptor genes were also identified in insect, nematode and other invertebrate genomes. **See also**: Chemosensory Systems; Comparative Genomics of the Major Chemosensory Gene Families in Arthropods; Genetics of Taste Perception; Mammalian Pheromones

In this article, the author reflects on OR gene evolution from the perspective of comparative genomics. The author mainly focuses on vertebrate OR genes, and henceforth 'OR' refers to vertebrate ORs unless otherwise noted.

OR Genes and Proteins

ORs are GPCRs, each containing seven α-helical transmembrane (TM) regions. GPCR genes can be classified into five or six groups based on sequence similarities; OR genes belong to the largest of these groups, the rhodopsin-like GPCR superfamily. This superfamily includes other genes that encode receptors for neurotransmitters, peptide hormones, chemokines, lipids, nucleotides, etc. (Fredriksson *et al*., 2003). The opsin genes involved in colour perception are also members of this superfamily; therefore, OR and opsin genes are distant relatives of each other. Each OR is on average approximately 310 amino acids long, and has several OR-specific motifs; for example, 'MAYDRYVAIC' motifs are located at the junction of each third TM region and the adjacent downstream intracellular loop (Niimura, 2012b). Mammalian ORs can be definitively classified into two groups, Class I or Class II, based on amino acid sequence similarities (see **Figure 5a**).

OR genes generally do not have any introns in the coding regions. This intronless gene structure is widely observed among GPCR genes. However, the number of exons in the 5′-untranslated region often varies among OR genes; moreover, these noncoding exons can be alternatively spliced to generate multiple messenger ribonucleic acid (mRNA) isoforms (Young *et al*., 2003). The biological significance of the presence of multiple isoforms is unknown. **See also**: G Protein-coupled Receptors; Human Intronless Genes and their Associated Diseases

OR genes are mainly expressed in sensory neurons of the olfactory epithelium. It is generally thought that each olfactory neuron expresses only a single functional OR gene among approximately 1000 genes in a monoallelic manner. This 'one neuron–one receptor rule' is thought to be necessary for discrimination among many different odorants, such that only a subset of olfactory neurons responds to a given odorant. Moreover, axons of olfactory neurons that express the same type of OR converge onto a specific target glomerulus in an olfactory bulb (Mori and Sakano, 2011). This phenomenon is called 'one receptor–one glomerulus rule'.

Notably, OR gene expression is not completely restricted to the olfactory epithelium. Parmentier *et al*. (1992) discovered OR gene expression in mammalian testis, and later, it was demonstrated that these testicular ORs mediated sperm chemotaxis (Spehr *et al*., 2003). Additionally, some OR genes are expressed in various other non-olfactory tissues, including brain, tongue, prostate, placenta, gut and kidney (Flegel *et al*., 2013). However, the function of such non-olfactory OR expression is unknown in most cases.

OR-odorant Relationship

It is generally thought that the relationships between ORs and odorants are not one-to-one, but multiple-to-multiple; one OR recognises multiple odorants, and one odorant is recognised by multiple ORs. Therefore, different odorants are represented as different combinations of activated ORs. Such a combinatorial coding scheme involving approximately 1000 ORs could allow discrimination among an almost unlimited number of odorants. Actually, a recent study provided evidence that humans can discriminate among more than one trillion olfactory stimuli, and indicated that the human olfactory system far outperforms the other senses with regard to the number of physically different stimuli that are discernible (Bushdid *et al*., 2014). With a one receptor–one glomerulus rule and a combinatorial coding scheme, a signal from each odorant can be converted into a topographical map of multiple glomeruli activated with varying magnitudes (Mori and Sakano, 2011).

However, the relationships between ORs and odorants are still largely unknown. To date, ligands have been identified for fewer than 100 mammalian ORs, and none have been identified for non-mammalian ORs. Both *in vivo* and *in vitro* approaches have been used to decode OR-odorant relationships. An *in vivo* approach involves visualising glomeruli activated by a given odour via optical imaging methods; OR gene expression in the sensory neurons projecting to the activated glomeruli is then analysed via single-cell reverse transcriptase polymerase chain reaction (RT-PCR). For example, Shirasu *et al*. (2014) identified a mouse OR (MOR215–1) that is specifically activated by muscone, a natural component of musk. Interestingly, the muscone-responsive glomeruli are not activated by polycyclic or other kinds of musks (see **Figure 1c**).

An *in vitro* approach involves expression of a target OR and additional signal transduction proteins (e.g. a G-protein) in mammalian cultured cells or *Xenopus* oocytes; downstream signals induced by a given odour can then be observed. Using this approach, Saito *et al*. (2009) performed high-throughput screening of 464 human and mouse ORs against 93 diverse odorants. However, they succeeded to identify ligands for only 52 mouse and 10 human ORs. Their results showed that the combinatorial

coding scheme is indeed correct; furthermore, they found some ORs are 'generalists' that are broadly tuned to a variety of structurally related ligands, whereas others are 'specialists' that are narrowly tuned and specific to a limited number of ligands. They demonstrated that Class I and Class II ORs tend to bind hydrophilic and hydrophobic ligands, respectively. However, there are currently no simple methods for predicting ligand-OR pairs from the sequence of a given OR, and a larger number of OR-odorant relationships must be examined to decipher the 'odour code'.

Bioinformatic Analysis of OR Genes

Because of advances in sequencing technologies, whole genome sequences from diverse organisms have be determined and made available via the Internet. **Figure 2** summarises the numbers of OR genes identified from the whole genome sequences of 30 chordate species by using bioinformatic methods. The numbers of OR genes are highly variable among different species. The fraction of OR pseudogenes is generally high (20–60%), and these fractions also vary considerably among species.

Each identified OR gene was classified into one of three categories: 'intact gene', 'truncated gene' or 'pseudogene' (**Figure 2**). An intact gene was defined as an intact coding sequence from the initiation codon to the stop codon that lacked any deletions in well-conserved regions. In contrast, a pseudogene was defined as a sequence that contained a nonsense mutation, frameshift, deletion within well-conserved regions or some combination thereof. A truncated gene was defined as a partial intact gene sequence located at a contig end. When a quality of the genome sequence is

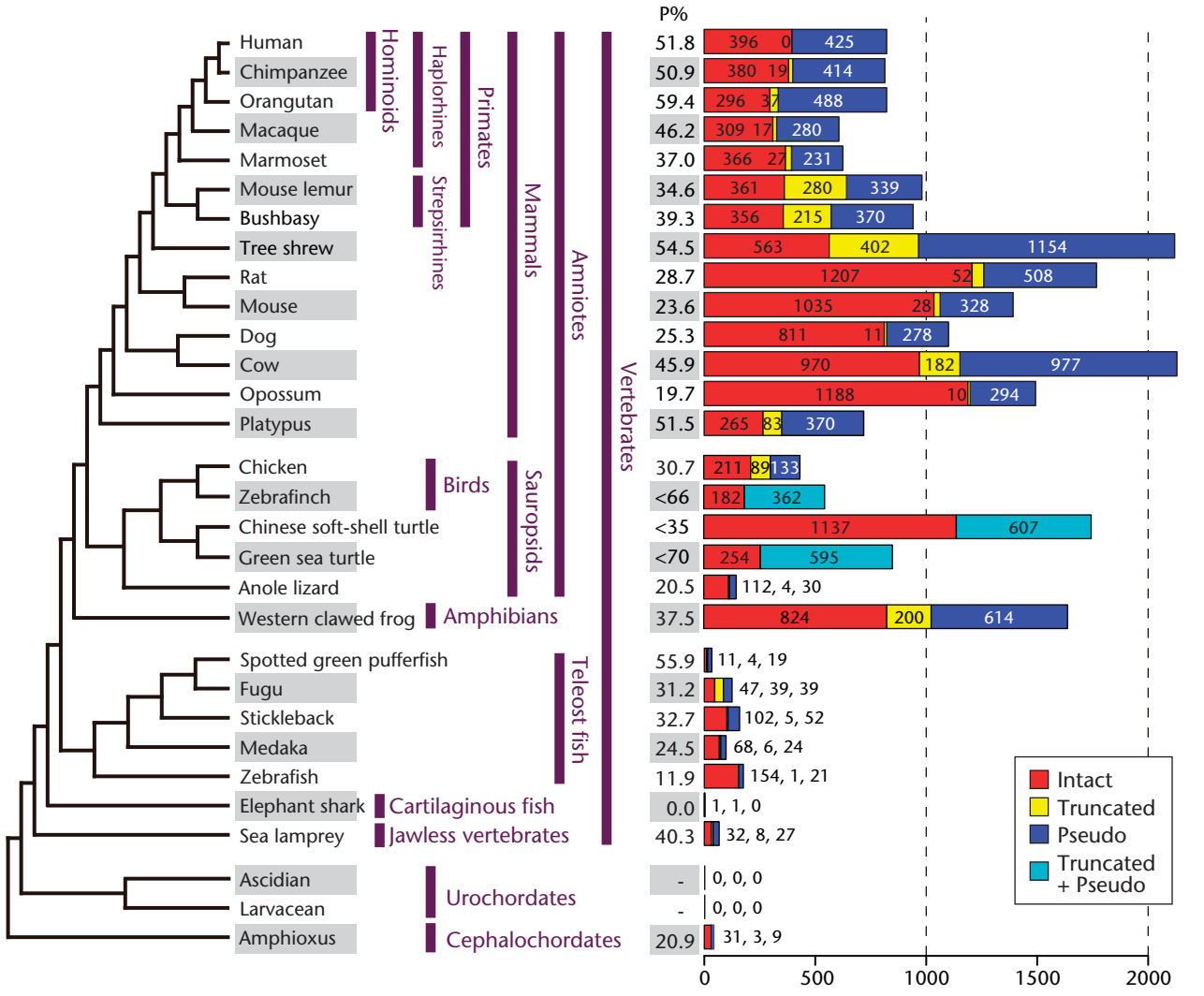


Figure 2 Number of OR genes identified from the whole genome sequences of 30 chordate species. Each number in the bar graphs indicates the number of intact genes (red), truncated genes (yellow) or pseudogenes (dark blue). For zebra finch and two turtle species, no distinction was made between truncated genes and pseudogenes. 'P%' indicates the fraction of pseudogenes. Data from Niimura and Nei (2007), Niimura (2009), Matsui *et al.* (2010) and Wang *et al.* (2013).

improved, a truncated gene will be classified into either an intact gene or a pseudogene. With low-coverage sequence information, the fraction of truncated genes in a genome tends to be high (e.g. mouse lemur, bushbaby or tree shrew) because contig lengths are relatively short due to incomplete assembly. Note that intact genes are potentially functional, but usually there is no experimental verification.

OR Genes in Humans

Genomic clusters

There are approximately 820 OR genes in the human genome (Niimura and Nei, 2003; Matsui *et al*., 2010). Among them, approximately 400 are intact genes, and more than a half are pseudogenes. Human OR genes reside in genomic clusters, and each chromosome, except chromosome 20 and the Y chromosome, encodes OR genes (**Figure 3a**). Chromosome 11 contains >40% of all human OR genes. Totally there are >30 genomic clusters each of which contains five or more OR genes. All Class I genes are found in a single cluster on chromosome 11. The largest human OR gene cluster contains approximately 100 Class II genes (intact genes or pseudogenes) and occupies an approximately 2Mb genomic region on chromosome 11.

OR genes in close proximity to each other within a cluster tend to be evolutionarily closely related (**Figure 3b**). This observation indicates that repeated tandem gene duplications have increased the number of OR genes (**Figure 3c**). However, the relationship between the evolutionary relatedness and the chromosomal positions is not always straightforward. A single OR gene cluster can contain evolutionarily distantly related genes, and evolutionarily closely related genes can reside in different clusters or on different chromosomes. These observations can be explained by assuming that several chromosomal rearrangements have occurred at regions containing OR gene clusters and that genes in different clusters were shuffled during evolution.

Some OR gene clusters are involved in human diseases – reciprocal translocation between two OR gene clusters, one on chromosome 4 and another on 8, causes Wolf–Hirschhorn syndrome (Niimura and Nei, 2003). Patients with this disease have a craniofacial phenotype described as a 'Greek warrior helmet' appearance (wide-set eyes, a broad or beaked nose, low-set malformed ears, and a small head), cognitive impairment and growth retardation. However, neither OR gene cluster involved in this disease-associated translocation contain any intact OR genes; therefore, the disease is not due to aberration of OR genes. Among the approximately 420 OR pseudogenes in humans, approximately 80 have highly similar DNA sequences, and have apparently all arisen from a single functional gene, *OR7E24*. These pseudogenes are collectively called the 7E (or H*) pseudogenes (Newman and Trask, 2003; Niimura and Nei, 2005a), and the two clusters involved in the Wolf–Hirschhorn syndrome contain only 7E pseudogenes.

Polymorphism and the diversity of odour perception

OR gene loci exhibit remarkably high between-individual diversity, and are among the most diverse regions of the human genome. Olender *et al*. (2012) used data from the 1000 Genome Project to investigate the diversity of OR gene repertoires among individuals. They identified 244 segregating OR pseudogenes, for which both intact and pseudogene forms are present in the population. They also found 63 OR loci exhibiting deletion copy number variation (CNV); such loci are present in some individuals, but not in others. In all, 66% of the approximately 400 human intact OR loci are affected by nonfunctional single nucleotide polymorphisms (SNPs), insertion–deletions (indels) and/or CNVs. Therefore, each individual has a unique set of functional OR genes.

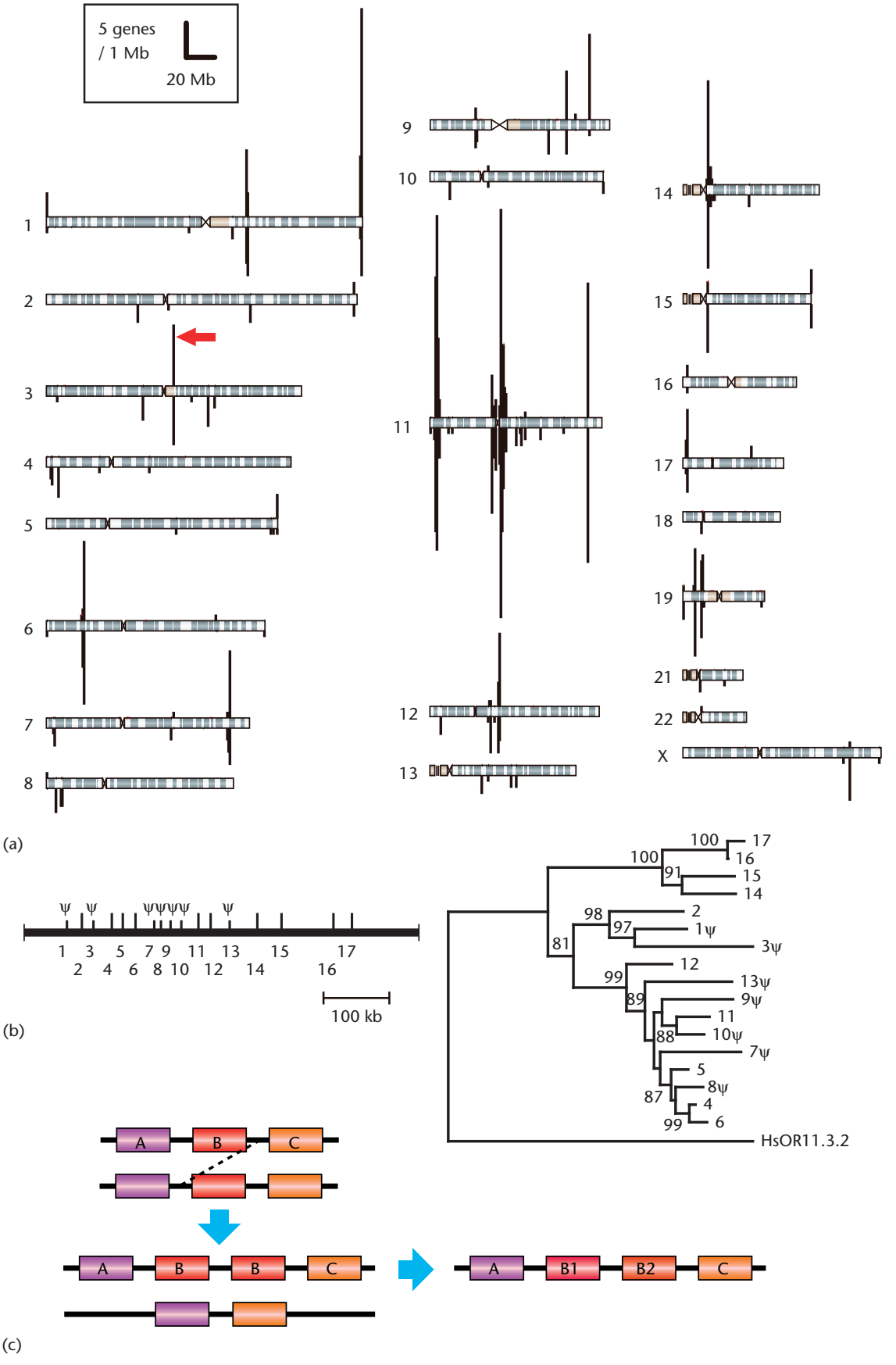
Olfactory perception differs considerably among individuals. Specific anosmia refers to individuals who lack the ability to perceive a particular odour, though they generally have a good sense of smell. For example, 7% of tested subjects exhibit specific anosmia to the macrocyclic musk Exaltolide (Whissell-Buechy and Amoore, 1973), and 9% exhibit anosmia to the polycyclic musk Galaxolide (Baydar *et al*., 1993; see **Figure 1c**).

Androstenone, a pig pheromone, is also subject to specific anosmia. People exhibit three different types of perception of this molecule: offensive (sweaty or urinous), pleasant (sweet or floral) and odourless. Keller *et al*. (2007) revealed that perception of androstenone is associated with the SNPs in *OR7D4*, an OR gene. They also showed that the OR7D4 protein is activated by androstenone *in vivo*. Variants of this locus include two non-synonymous SNPs linked to each other, R88W and T133M, and subjects having a RT/WM or WM/WM genotype were less sensitive and felt less unpleasant to androstenone than did RT/RT subjects.

Several other studies have also demonstrated associations between odour perception and SNPs in OR gene loci: *OR11H7P* is associated with isovaleric acid (sweaty) perception (Menashe *et al*., 2007), *OR2J3* with *cis*-3-hexen-1-ol (grassy) perception (McRae *et al*., 2012), *OR5A1* with β-ionone (floral) perception (Jaeger *et al*., 2013) and *OR10G4* with guaiacol (smoky) perception (Mainland *et al*., 2014).

OR Genes in Primates

Higher primates generally have much smaller numbers (300–400) of intact OR genes than do most other mammals (approximately 1000) (**Figure 2**). This observation is thought to reflect that higher primates heavily rely on vision instead of olfaction, and that primate olfaction has degenerated. Matsui *et al*. (2010) demonstrated that the



common ancestor of hominoids (humans and apes), Old World monkeys (OWMs), and New World monkeys (NWMs) had approximately 550 functional OR genes, and each species has lost >200 OR genes during evolution (**Figure 4a**). Notably, the number of intact OR genes in humans is similar to that of chimpanzees, and it is even larger than that of orangutans or macaques (Go and Niimura, 2008; Matsui *et al*., 2010). This observation may mean that our olfactory ability is not particularly worse than that of other higher primates.

Which factors have caused the shrinkage of OR gene repertoires during the primate evolution? Loss of olfactory capacity in primate lineages might be related to the acquisition of well-developed colour vision. As mentioned above, most humans have trichromatic vision mediated by three opsin genes. However, trichromacy is not the norm among mammalian species. Hominoids and OWMs are trichromatic, but most other mammalian taxa have dichromatic vision mediated by two opsin genes; such dichromatic vision is called colour-blindness in humans. In the common ancestor of hominoids and OWMs, duplication of an opsin gene on the X chromosome resulted in two divergent and functionally distinct opsin genes that separately mediate red and green vision; this gene duplication and divergence resulted in trichromacy in hominoids and OWMs (**Figure 4a**). Colour vision systems in NWMs are complicated. There is a single X-linked opsin gene locus that is usually polymorphic; therefore, heterozygous females are trichromatic, whereas homozygous females and all males are dichromatic.

To determine whether colour vision and olfaction are evolutionarily linked, Gilad *et al*. (2004) investigated the fractions of OR pseudogenes from 19 primate species by examining 100 randomly chosen OR gene sequences. Their results showed that the fractions of OR pseudogenes in hominoid and OWM species are significantly higher than those in NWM or other mammalian species. Based on these observations, they proposed the 'colour vision priority hypothesis', specifically that OR genes were lost concomitantly with the acquisition of complete trichromatic vision. However, analyses using deep-coverage genomes (Matsui *et al*., 2010) indicated that there are no significant differences between hominoids/OWMs and NWMs with regard to numbers of intact OR genes. Moreover, results (**Figure 4a**) indicate that gradual OR gene loss occurred repeatedly in every lineage leading from the NWM/OWM/hominoid common ancestor to humans and that one rapid, large-scale OR gene loss event did not occur near the branch-point that separated OWMs/hominoids from NWMs and at which trichromatic vision emerged. Therefore, the colour vision priority hypothesis was not supported by these findings.

Based on morphology of nostrils, the order Primates can be divided into two suborders: (1) strepsirrhines, which means 'twisted nose' and includes lemurs and lorises, and (2) haplorhines, which means 'simple nose' and includes tarsiers, NWMs, OWMs and hominoids. This classification is supported by molecular studies. Strepsirrhines and haplorhines are characterised by the presence or absence of the rhinarium, respectively. The rhinarium is a moist and hairless surface at the tip of the nose, and is used to detect the directional origin of odorants. Many mammalian species, including cats and dogs, have rhinarium.

Generally, haplorhines have a smaller olfactory epithelium based on relative size than strepsirrhines (Barton, 2006). Moreover, most strepsirrhines are nocturnal, whereas most haplorhines are diurnal. Therefore, haplorhines are apparently less dependent on olfaction than strepsirrhines. To determine which factor or factors led to the shrinkage of OR gene repertoires during primate evolution, a wide variety of primate species that inhabit a wide range of ecological niches must be examined. **See also**: Primates (Lemurs, Lorises, Tarsiers, Monkeys and Apes); Visual Pigment Genes: Evolution

OR Genes in Mammals

Mammals are extremely diverse in size, shape and habitat use. Mammals occupy all habitats: terrestrial, fossorial, arboreal, volant and aquatic. Their feeding habitats are also highly diversified. Insectivores, herbivores, carnivores and omnivores are found among mammals; some feed on fish, others leaves, yet other on grains or seeds; some are even ant specialist. Therefore, OR gene repertoires are predictably highly variable among mammals and reflect the ecological diversity of mammals (Hayden *et al*., 2010).

We previously estimated the numbers of OR gene gains or losses in mammalian lineages based on a mammalian phylogenetic tree (Niimura and Nei, 2007). The results (**Figure 4b**) showed that (1) gene expansion occurred in the placental mammal lineage after it diverged from the monotreme and from marsupial lineages and that (2) hundreds of gains and losses of OR genes have occurred in an order-specific manner. The latter finding suggests that, although the numbers of functional OR genes in several mammalian species are similar (approximately 1000), these OR gene repertoires are often highly diverged from one

Figure 3 OR genes in the human genome. (a) Vertical bars above and below the chromosomes represent locations of intact OR genes and OR pseudogenes, respectively. The height of each bar indicates the number of OR genes existing in a nonoverlapping 1-Mb window. (b) The OR gene cluster indicated by the red arrow in (a). The diagram (left) represents an expanded view of a 0.6-Mb region on chromosome 3. 'Ψ' represents a pseudogene. All genes are encoded on the same strand. The neighbor-joining phylogenetic tree (right) for the genes contained in this 0.6-Mb cluster indicates that neighbouring genes within the cluster tend to be more closely related to each other than to more distantly located genes within the cluster. For example, genes 16 and 17 are more closely related to each other than they are to the other genes. *HsOR11.3.2* was used as the outgroup in the phylogenetic tree. Bootstrap values greater than 80% are shown. (a) and (b) were modified from Nei *et al.* (2008). © Nature Publishing Group. (c) Schematic representation of tandem gene duplication. Unequal crossing-over generates a new gene copy ('B') adjacent to the original gene. Subsequently, independent accumulation of mutations causes the sequences of the duplicates to diverge and potentially to acquire distinctive functions ('B1' and 'B2').

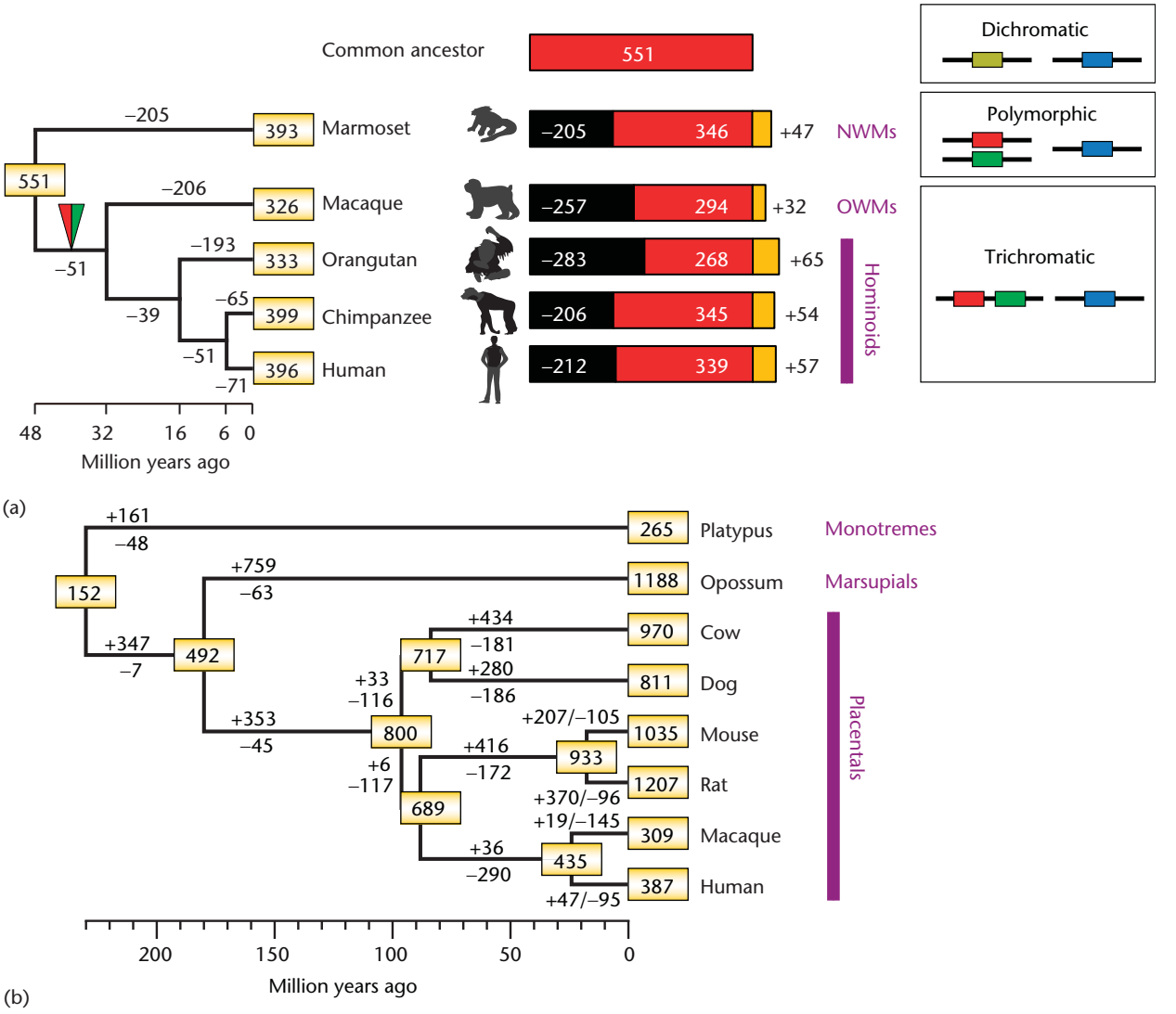


Figure 4 Evolutionary dynamics of OR genes in primates (a) and in mammals (b). (a) OR gene losses during primate evolution. The common ancestor of the five primate species is estimated to have had 551 functional OR genes. The number above each branch indicates the number of the ancestral OR genes that were lost in that lineage. For example, humans have lost 212 of the 551 putative ancestral OR genes, but 57 gene duplications apparently occurred; therefore, humans currently have 396 functional OR genes. An arrowhead with red and green represents the branch at which the duplication of red/green opsin genes occurred. Colour vision system in each species is also shown (right). X-linked red/green opsin genes and an autosomal blue opsin gene are represented schematically. Modified with permission from Matsui *et al.* (2010). © Oxford University Press. (b) Gains and losses of OR genes during mammalian evolution. The number in each box indicates the number of functional OR genes in the indicated extant species or ancestral species. The numbers with a plus or minus sign indicate estimated numbers of gene gains and losses, respectively, along each branch. Modified from Niimura and Nei (2007). © PLoS One.

another. Therefore, the spectrum of detectable odorants might be quite different among different mammalian species.

This kind of dynamic gene gain and loss in a multigene family is called 'birth-and-death evolution'. In this model, new genes are created by gene duplication, and some duplicated genes are maintained in the genome for a long time, whereas others are deleted or become nonfunctional through deleterious mutations. This model was first proposed to explain the evolutionary pattern of the major histocompatibility complex (MHC) genes involved in the immune system (Nei and Rooney, 2005). It is now known that most multigene families are subject to birth-and-death evolution to some extent, but OR genes provide one of the most extreme examples of this mode of evolution.

In addition to higher primates, platypuses also have a small OR gene repertoire (**Figure 2**; Niimura and Nei, 2007). Platypuses are semiaquatic egg-laying mammals endemic to Australia. The platypus bill is a sensor; it houses electroreceptors and mechanoreceptors that can detect weak electric fields generated by prey (e.g. freshwater shrimp) in the mud at the bottom of streams. Therefore, platypuses

can find prey with their eyes, ears and nostrils closed. These faculties are reminiscent of those of toothed whales (e.g. dolphins), which completely lack an olfactory system and have developed an echolocation system to adapt to a fully aquatic life. In fact, the fractions of OR pseudogenes in toothed whale genomes are reportedly very high (Hayden *et al*., 2010). Therefore, different sensory modalities do seem to affect one another. The OR gene repertoire present in each organism's genome is thought to reflect its ecological niche and the extent of reliance on olfaction.

However, it is unclear which aspect of olfactory ability is reflected in the number of OR genes in the genome. Dogs are supposed to have a very keen sense of smell, but they do not have particularly a large repertoire of OR genes (**Figure 2**). This observation may be explained by the hypothesis that the number of OR genes in a given species is correlated with the number of odorants it can discriminate among, whereas the sensitivity to a specific odorant may be determined by an absolute amount of expressed ORs. Carnivores may not need to distinguish among many different types of odours, but they may be very sensitive to the odours that they can discern. **See also**: Cetacea (Whales, Porpoises and Dolphins); Mammalia; Monotremata

OR Genes in Vertebrates

Fish, like mammals, use olfactory cues to find food, avoid danger and identify conspecific individuals. Olfactory information is also used to recognise places within an organism's environment. Salmon have a remarkable homing ability; specifically, they return to the river where they were spawned, and this behaviour depends on olfaction. Salmon imprint to place-specific odours during a sensitive developmental period, and adults use the odorant memory to return to their natal streams. Fish can detect mainly four groups of water-soluble molecules as odorants: amino acids, gonadal steroids, bile acids and prostaglandins. These odorants are nonvolatile chemicals; therefore, humans cannot smell them.

As shown in **Figure 2**, teleost fish have much smaller numbers of OR genes than mammals. However, OR gene repertoires among fish species are more diverse than those in mammals (**Figure 5a**). Extensive phylogenetic analyses showed that each OR gene in jawed vertebrates can be classified into one of seven groups, designated α–η (Niimura and Nei, 2005b; Niimura, 2009). For mammalian OR genes, group γ corresponds to Class II, and groups α and β correspond to Class I. The numbers of intact OR genes belonging to each group varies among taxa (**Figure 5b**). The distribution of genes exhibits an intriguing pattern: groups α and γ are well represented in tetrapods (amphibians, reptiles, birds and mammals), but are absent from all fish (with one exception in zebrafish). Conversely, groups δ, ε, ζ and η are found in teleost fish and amphibians, but amniotes (reptiles, birds and mammals) completely lack these groups. Therefore, groups α and γ genes are considered to be terrestrial-type genes, and groups δ, ε, ζ and η are regarded as aquatic-type genes. Interestingly, only amphibians have both types.

These observations indicate that terrestrial-type genes function in detection of volatile odorants, and aquatic-type genes function in detection of water-soluble odorants. Group β genes are exceptional because they were present both in aquatic and terrestrial vertebrates. Therefore, group β genes may function to detect odorants that are both volatile and water-soluble (e.g. alcohol). For example, β-phenylethyl alcohol conveys a rose-like fragrance, but this molecule can also be detected by fish at a low concentration (Niimura, 2009).

The author recently analysed whole genome sequences of two turtle species – the Chinese soft-shell turtle and the green sea turtle (Wang *et al*., 2013). Although both are aquatic, we did not find any aquatic-type OR genes. This observation is not surprising, given the phylogenetic position of turtles. Molecular studies show that turtles are more closely related to birds than are lizards. Notably, reptiles are a paraphyletic group, not a monophyletic group. Sauropsid is the clade that comprises reptiles and birds (see **Figure 2**). During the process of terrestrial adaptation, the common ancestor of amniotes (sauropsids and mammals) apparently lost all aquatic-type OR genes; therefore, although some turtles have secondarily adapted to the aquatic life, they lack any aquatic-type OR genes. Interestingly, however, we found that the fractions of group α OR genes are high (46–62%) in both turtle species; this preponderance of group α genes was not characteristic of the other sauropsids examined ($<5\%$; **Figure 5b**). Moreover, phylogenetic analysis showed that the group α genes are greatly expanded in the turtle lineage (Wang *et al*., 2013). Because group α (Class I) genes tend to detect hydrophilic volatile odorants (see Section 'OR-odorant relationship'), lineage-specific expansion of the group α OR genes in turtles may be related with adaptation to aquatic life. **See also**: Reptilia (Reptiles)

OR Genes in Invertebrates

Vertebrates belong to the phylum Chordata. Chordates include two more invertebrate subphyla, cephalochordates (including amphioxus or lancelet) and urochordates (tunicates). Amphioxi have fish-like appearance, but they lack any distinctive sensory apparatus corresponding to the eyes, ears or nose. Thus, amphioxi are also called 'acraniates', meaning headless animals. Nevertheless, we found >30 vertebrate-like OR genes when analysing the whole genome sequence of the Florida lancelet, *Branchiostoma floridae* (Niimura, 2009). Amphioxus OR genes have diverged from vertebrate OR genes in amino acid sequence and form an amphioxus-specific clade; nevertheless, they are clearly distinguishable from other non-OR GPCRs (**Figure 5a**). The olfactory system of amphioxus is not well understood, and the function of these genes remains unclear.

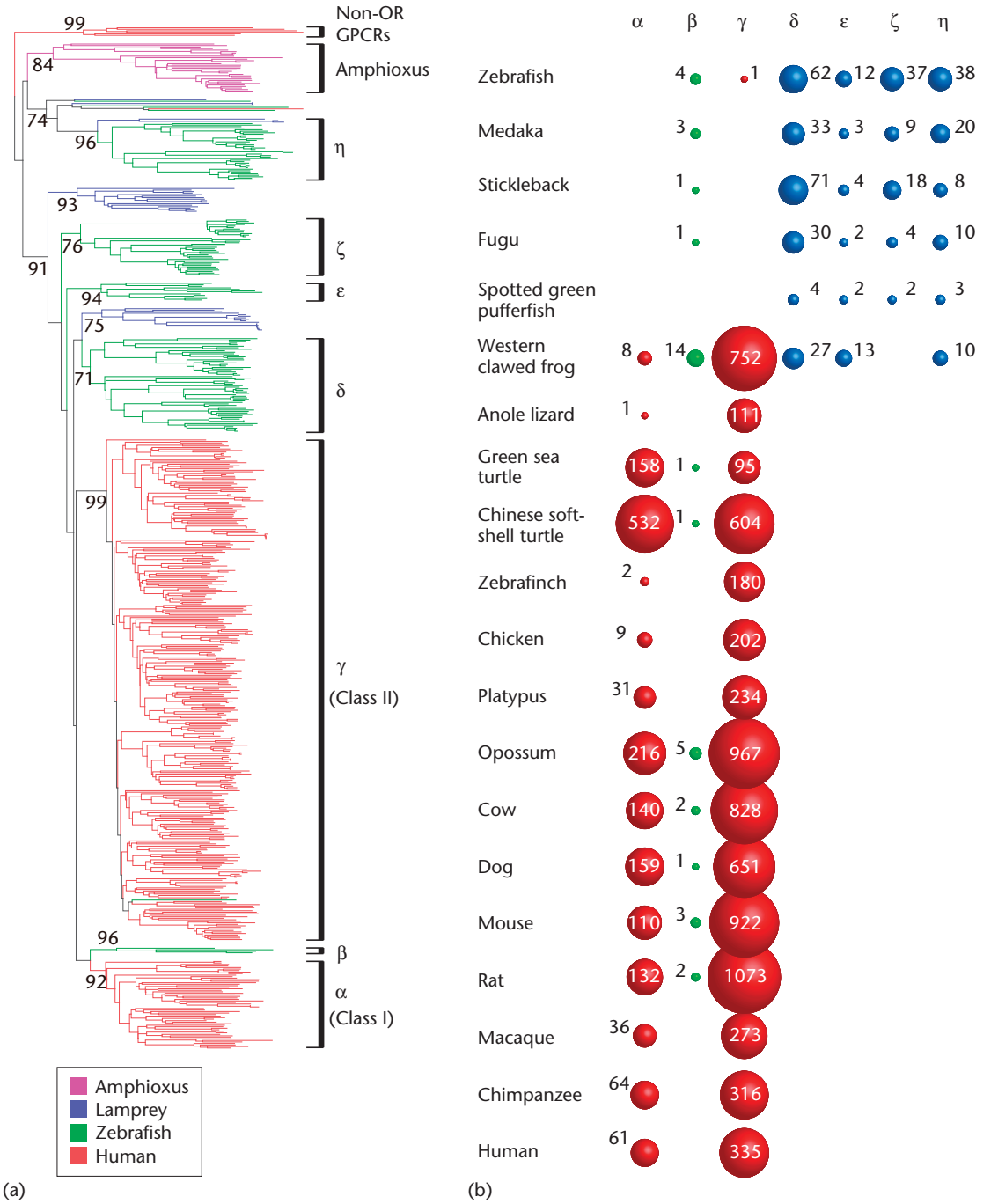


Figure 5 Evolution of OR genes in vertebrates. (a) Neighbour-joining phylogenetic tree constructed by using all intact OR genes identified from amphioxus, lamprey, zebrafish, and human. Several non-OR GPCR genes were used as the outgroup. Bootstrap values are shown for major clades. Modified from Niimura (2009). © Oxford University Press. (b) Number of intact OR genes in each gene group for each species. The volume of a sphere is proportional to the number of genes in the group. Terrestrial-type and aquatic-type OR genes are represented by red and blue, respectively. Data from Niimura (2009) and Wang *et al.* (2013).

No vertebrate-like OR genes were found in genome sequences of three urochordate species, the ascidians *Ciona intestinalis* and *Ciona savignyi* and the larvacean *Oikopleura dioica* (Niimura, 2009). Although the morphology of ascidians are highly diverged from those of vertebrates, molecular phylogenomic studies revealed that urochordates, and not cephalochordates, are the sister group of vertebrates. Because amphioxus, the most basal chordates, retains vertebrate-like OR genes, the origin of vertebrate-like OR genes can be traced back to the common ancestor

of all chordates. Therefore, the absence of vertebrate-like OR genes in the urochordate genomes indicates that all OR genes were lost in the urochordate lineage. Ascidians are sessile filter-feeders, whereas larvaceans have a floating planktonic lifestyle. Reflecting their inactive lifestyles, the nervous systems of urochordates are highly reduced, and sensory organs are poorly developed. Nevertheless, the possibility that other families of genes function as chemoreceptors in urochordates cannot be excluded. **See also**: Analysis of the Amphioxus Genome

Chemoreceptor genes were also identified in other invertebrates including insects, nematodes, echinoderms and mollusks. Among these groups of genes, insect chemoreceptor genes are the most thoroughly studied. Insect chemoreceptors are classified into two evolutionarily related gene families, insect ORs and gustatory receptors (GRs). Insect OR/GRs are seven-TM proteins, as are vertebrate ORs. However, the membrane topology of the insect chemoreceptor is inverted relative to that of vertebrate ORs. Insect OR/GRs and vertebrate ORs do not share any amino acid sequence similarities; therefore, they have independent evolutionary origins. Insect ORs are not GPCRs; they are odorant-gated ion channels that assemble into functional heterodimers (Sato *et al*., 2008; Wicher *et al*., 2008).

There are 62 functional insect OR genes and 73 functional GR genes in the fruit fly genome. Bioinformatic analyses of many insect species showed that the numbers of putatively functional insect OR/GR genes, like those of vertebrate OR genes, vary among species and these numbers range from 265 ORs and 220 GRs in the red flour beetle (*Tribolium* Genome Sequencing Consortium, 2008) to 10 ORs and 6 GRs in human body lice (Kirkness *et al*., 2010). GR genes were also identified in the genome sequence of the water flea *Daphia pulex*, an aquatic crustacean arthropod, but OR genes are completely absent from this genome (Peñalva-Arana *et al*., 2009). Therefore, insect-like OR may be limited only to insects, whereas GRs may have a more ancient origin. **See also**: Comparative Genomics of the Major Chemosensory Gene Families in Arthropods

The nematode *Caenorhabditis elegans* is a small roundworm comprising only approximately 1000 somatic cells with a simple nervous system of only 302 neurons. Nevertheless, *C. elegans* has a surprisingly large number of chemoreceptor genes. There are as many as approximately 1300 functional chemoreceptor genes and approximately 400 pseudogenes in the *C. elegans* genome; therefore, chemoreceptors account for approximately 8.5% of the entire *C. elegans* proteome (Thomas and Robertson, 2008). The *C. elegans* chemoreceptor genes encode GPCRs with seven-TM regions. They are more diverse than vertebrate ORs and are classified into 19 subfamilies. Among these subfamilies, only the 'srw' subfamily shows sequence similarities to vertebrate ORs; the other 18 subfamilies are nematode-specific. Unlike vertebrates and insects, nematodes lack vision and hearing; this lack of other sensory modalities may explain the high genetic investment in chemosensation by this tiny animal.

Putative chemoreceptor genes were also identified in the sea urchin *Strongylocentrotus purpuratus* and the marine mollusk *Aplysia californica*. Raible *et al*. (2006) extensively examined rhodopsin-like GPCR genes from the sea urchin genome. They found that some GPCR gene families are greatly expanded within the sea urchin lineage and that the member genes are prominently expressed in the pedicellariae and tube feet of adult sea urchins. Cummins *et al*. (2009) discovered *A. californica* chemoreceptor genes. They identified novel families of rhodopsin-like GPCR genes expressed in the rhinophore and oral tentacles; in all, 90 chemoreceptor genes were found in the low-coverage (2x) *A. californica* genome.

Chemosensory systems are thought to be present in essentially all motile organisms. However, chordates, insects, nematodes, echinoderms and mollusks use evolutionarily independent gene families to encode chemosensory receptor. Apparently, genes that encode chemosensory receptors have emerged independently many times during animal evolution.

Acknowledgements

This work was supported in part by Grant-in-Aid for Young Scientists (B) (JSPS KAKENHI Grant Number 23770271) and ERATO Touhara Chemosensory Signal Project from JST, Japan.

References

Barton RA (2006) Olfactory evolution and behavioral ecology in primates. *American Journal of Primatology* **68**: 545–558.

Baydar A, Petrzilka M and Schott M (1993) Olfactory thresholds for androstenone and Galaxolide: sensitivity, insensitivity and specific anosmia. *Chemical Senses* **18**: 661–668.

Buck L and Axel R (1991) A novel multigene family may encode odorant receptors: a molecular basis for odor recognition. *Cell* **65**: 175–187.

Bushdid C, Magnasco MO, Vosshall LB and Keller A (2014) Humans can discriminate more than 1 trillion olfactory stimuli. *Science* **343**: 1370–1372.

Cummins SF, Erpenbeck D, Zou Z *et al*. (2009) Candidate chemoreceptor subfamilies differentially expressed in the chemosensory organs of the mollusc Aplysia. *BMC Biology* **7**: 28.

Flegel C, Manteniotis S, Osthold S, Hatt H and Gisselmann G (2013) Expression profile of ectopic olfactory receptors determined by deep sequencing. *PLoS One* **8**: e55368.

Fredriksson R, Lagerström MC, Lundin LG and Schiöth HB (2003) The G-protein-coupled receptors in the human genome form five main families. Phylogenetic analysis, paralogon groups, and fingerprints. *Molecular Pharmacology* **63**: 1256–1272.

Gilad Y, Przeworski M and Lancet D (2004) Loss of olfactory receptor genes coincides with the acquisition of full trichromatic vision in primates. *PLoS Biology* **2**: E5.

Go Y and Niimura Y (2008) Similar numbers but different repertoires of olfactory receptor genes in humans and chimpanzees. *Molecular Biology and Evolution* **25**: 1897–1907.

Hayden S, Bekaert M, Crider TA *et al*. (2010) Ecological adaptation determines functional mammalian olfactory subgenomes. *Genome Research* **20**: 1–9.

Herz R (2007) *The Scent of Desire: Discovering Our Enigmatic Sense of Smell*. New York, NY: HarperCollins Publishers.

Jaeger SR, McRae JF, Bava CM *et al*. (2013) A mendelian trait for olfactory sensitivity affects odor experience and food selection. *Current Biology* **23**: 1601–1605.

Keller A, Zhuang H, Chi Q, Vosshall LB and Matsunami H (2007) Genetic variation in a human odorant receptor alters odour perception. *Nature* **449**: 468–472.

Kirkness EF, Haas BJ, Sun W *et al*. (2010) Genome sequences of the human body louse and its primary endosymbiont provide insights into the permanent parasitic lifestyle. *Proceedings of the National Academy of Sciences of the USA* **107**:12168–12173.

Mainland JD, Keller A, Li YR *et al*. (2014) The missense of smell: functional variability in the human odorant receptor repertoire. *Nature Neuroscience* **17**: 114–120.

Matsui A, Go Y and Niimura Y (2010) Degeneration of olfactory receptor gene repertories in primates: no direct link to full trichromatic vision. *Molecular Biology and Evolution* **27**: 1192–1200.

McRae JF, Mainland JD, Jaeger SR *et al*. (2012) Genetic variation in the odorant receptor OR2J3 is associated with the ability to detect the ''grassy'' smelling odor, cis-3-hexen-1-ol. *Chemical Senses* **37**: 585–593.

Menashe I, Abaffy T, Hasin Y *et al*. (2007) Genetic elucidation of human hyperosmia to isovaleric acid. *PLoS Biology* **5**: e284.

Mori K and Sakano H (2011) How is the olfactory map formed and interpreted in the mammalian brain? *Annual Review of Neuroscience* **34**: 467–499.

Nei M and Rooney AP (2005) Concerted and birth-and-death evolution of multigene families. *Annual Review of Genetics* **39**:121–152.

Nei M, Niimura Y and Nozawa M (2008) The evolution of animal chemosensory receptor gene repertoires: roles of chance and necessity. *Nature Reviews Genetics* **9**: 951–963.

Newman T and Trask BJ (2003) Complex evolution of 7E olfactory receptor genes in segmental duplications. *Genome Research* **13**: 781–793.

Ngai J, Dowling MM, Buck L, Axel R and Chess A (1993) The family of genes encoding odorant receptors in the channel catfish. *Cell* **72**: 657–666.

Niimura Y and Nei M (2003) Evolution of olfactory receptor genes in the human genome. *Proceedings of the National Academy of Sciences of the USA* **100**: 12235–12240.

Niimura Y and Nei M (2005a) Comparative evolutionary analysis of olfactory receptor gene clusters between humans and mice. *Gene* **346**: 13–21.

Niimura Y and Nei M (2005b) Evolutionary dynamics of olfactory receptor genes in fishes and tetrapods. *Proceedings of the National Academy of Sciences of the USA* **102**: 6039–6044.

Niimura Y and Nei M (2007) Extensive gains and losses of olfactory receptor genes in mammalian evolution. *PLoS One* **2**: e708.

Niimura Y (2009) On the origin and evolution of vertebrate olfactory receptor genes: comparative genome analysis among 23 chordate species. *Genome Biology and Evolution* **1**: 34–44.

Niimura Y (2012a) Evolution of chemosensory receptor genes in primates and other mammals. In: Hirai H, Imai G, Go Y (eds) *Post-Genome Biology of Primates*, pp. 43–62. Tokyo: Springer.

Niimura Y (2012b) Olfactory receptor multigene family in vertebrates: from the viewpoint of evolutionary genomics. *Current Genomics* **13**: 103–114.

Olender T, Waszak SM, Viavant M *et al*. (2012) Personal receptor repertoires: olfaction as a model. *BMC Genomics* **13**: 414.

Parmentier M, Libert F, Schurmans S *et al*. (1992) Expression of members of the putative olfactory receptor gene family in mammalian germ cells. *Nature* **355**: 453–455.

Peñalva-Arana DC, Lynch M and Robertson HM (2009) The chemoreceptor genes of the waterflea *Daphnia pulex*: many Grs but no Ors. *BMC Evolutionary Biology* **9**: 79.

Raible F, Tessmar-Raible K, Arboleda E *et al*. (2006) Opsins and clusters of sensory G-protein-coupled receptors in the sea urchin genome. *Developmental Biology* **300**: 461–475.

Rossiter KJ (1996) Structure-odor relationships. *Chemical Reviews* **96**: 3201–3240.

Saito H, Chi Q, Zhuang H, Matsunami H and Mainland JD (2009) Odor coding by a mammalian receptor repertoire. *Science Signaling* **2**: ra9.

Sato K, Pellegrino M, Nakagawa T *et al*. (2008) Insect olfactory receptors are heteromeric ligand-gated ion channels. *Nature* **452**: 1002–1006.

Shirasu M, Yoshikawa K, Takai Y *et al*. (2014) Olfactory receptor and neural pathway responsible for highly selective sensing of musk odors. *Neuron* **81**: 165–178.

Spehr M, Gisselmann G, Poplawski A *et al*. (2003) Identification of a testicular odorant receptor mediating human sperm chemotaxis. *Science* **299**: 2054–2058.

Thomas JH and Robertson HM (2008) The *Caenorhabditis* chemoreceptor gene families. *BMC Biology* **6**: 42.

Tribolium Genome Sequencing Consortium (2008) The genome of the model beetle and pest *Tribolium castaneum*. *Nature* **452**: 949–955.

Wang Z, Pascual-Anaya J, Zadissa A *et al*. (2013) The draft genomes of soft-shell turtle and green sea turtle yield insights into the development and evolution of the turtle-specific body plan. *Nature Genetics* **45**: 701–706.

Wicher D, Schäfer R, Bauernfeind R *et al*. (2008) *Drosophila* odorant receptors are both ligand-gated and cyclic-nucleotide-activated cation channels. *Nature* **452**: 1007–1111.

Whissell-Buechy D and Amoore JE (1973) Odour-blindness to musk: simple recessive inheritance. *Nature* **242**: 271–273.

Young JM, Shykind BM, Lane RP *et al*. (2003) Odorant receptor expressed sequence tags demonstrate olfactory expression of over 400 genes, extensive alternate splicing and unequal expression levels. *Genome Biology* **4**: R71.

Further Reading

Bargmann CI (2006) Comparative chemosensation from receptors to ecology. *Nature* **444**: 295–301.

Crasto CJ (ed.) (2013) *Olfactory Receptors*. New York, NY: Humana Press.

Wyatt TD (2003) *Pheromones and Animal Behavior: Chemical Signals and Signatures*. Cambridge: Cambridge University Press.