Update on the olfactory receptor (OR) gene superfamily

Tsviya Olender,1 Doron Lancet1 and Daniel W. Nebert2*

1The Crown Human Genome Center, Department of Molecular Genetics, The Weizmann Institute of Science, Rehovot 76100, Israel
2Department of Environmental Health and Center for Environmental Genetics (CEG), University Cincinnati Medical Center, Cincinnati, OH 45267–0056, USA
*Correspondence to: Tel: +1 513 821 4664; Fax: +1 513 558 0925; E-mail: dan.nebert@uc.edu

Date received (in revised form): 5th May, 2008

Abstract

The olfactory receptor gene (OR) superfamily is the largest in the human genome. The superfamily contains 390 putatively functional genes and 465 pseudogenes arranged into 18 gene families and 300 subfamilies. Even members within the same subfamily are often located on different chromosomes. OR genes are located on all autosomes except chromosome 20, plus the X chromosome but not the Y chromosome. The gene:pseudogene ratio is lowest in human, higher in chimpanzee and highest in rat and mouse — most likely reflecting the greater need of olfaction for survival in the rodent than in the human. The OR genes undergo allelic exclusion, each sensory neurone expressing usually only one odourant receptor allele; the mechanism by which this phenomenon is regulated is not yet understood. The nomenclature system (based on evolutionary divergence of genes into families and subfamilies of the OR gene superfamily) has been designed similarly to that originally used for the CYP gene superfamily.

Keywords: classification of gene families and subfamilies, OR gene superfamily, CYP gene superfamily, nasal olfactory neurone, olfaction, olfactory receptor gene superfamily, allelic exclusion, opossum genome, platypus genome

Introduction

Before 1980, the names of genes and classification of their encoded proteins were highly variable and non-systematic — especially to anyone slightly outside a particular field or to a new graduate student entering the field. Professor Margaret Oakley Dayhoff was a pioneer in attempting to create order out of chaos in the naming of genes and gene families by means of computerised protein alignments.1 She was widely recognised as the founder in this new field of gene/protein classification, before her untimely death in 1983.

Cytochrome P450 (CYP) genes are conveniently arranged into families and subfamilies based on the percentage amino acid sequence identity.2–7 Enzymes that share approximately ≥40 per cent identity are assigned to a particular family designated by an Arabic numeral, whereas those sharing approximately ≥55 per cent identity are grouped into a particular subfamily designated by a letter. For example, the sterol 27-hydroxylase enzyme and the 25-hydroxy-vitamin D3 1α-hydroxylase enzyme are both assigned to the CYP27 family because they share >40 per cent sequence identity. Furthermore, the sterol 27-hydroxylase is assigned to the CYP27 ‘A’ subfamily and the 25-hydroxy-vitamin D3 1α-hydroxylase to the CYP27 ‘B’ subfamily because their protein sequences are <55 per cent identical. If an additional enzyme were to be discovered that shared >55 per cent identity with the sterol 27-hydroxylase, then it would be named CYP27A2. If an additional enzyme were to be discovered that
shared \(<55\) per cent but \(>40\) per cent identity with the sterol 27-hydroxylase as well as the 25-hydroxy-vitamin D\(_3\) 1\(\alpha\)-hydroxylase, then it would be named CYP27C1. The development and application of this delightfully logical system of nomenclature to the genes of many animals, plants and bacteria\(^8\) has eliminated the confusion that often had plagued the naming of gene families and superfamilies. Subsequently, this ‘divergent evolution’ nomenclature system was adopted for several hundred other gene families and superfamilies — including the olfactory receptor superfamily.

**Background and history**

Vertebrate olfactory receptor (OR) genes represent a category of G-protein-coupled receptors (GPCRs) that contain seven transmembrane \(\alpha\)-helical domains and function in the reception of innumerable odour molecules in the environment.\(^9\) The OR gene superfamily is the largest in vertebrate genomes.\(^10\)–\(^13\) The genomic architecture of mammalian OR gene clusters shows an ancient evolutionary origin, preceding the marsupial–eutherian split; species-specific evolution has further shaped the different OR gene families, by means of both gains and losses of complete clusters, as well as expansion and contraction of existing clusters.\(^11\)

This dynamic flexibility is also reflected among individual humans; examining 51 candidate OR genes on DNA chips in 189 ethnically diverse subjects, a striking amount of population diversity was found.\(^14\) Segregating pseudogenes (SPGs) are genes that segregate in populations between intact genes and pseudogenes — due to a disruptive single nucleotide polymorphism (SNP). A range of 16–24 functional OR genes was found, just in this study alone, indicating that the OR gene superfamily is among the most pronounced examples of functional population diversity in the human genome.\(^14\) Copy number variations (CNVs), another type of polymorphism, are also highly prevalent among human OR genes.\(^15,16\) All these genomic events are evidence of a relatively recent process, whereby the extreme diminution of a functional repertoire in humans has occurred — a process which is presumably still ongoing.

For most mammalian species, the ability to detect millions of different odourants is critical to their survival. Based on recent OR gene mining data in the platypus, opossum, cow and dog genomes — compared with that in the rat, mouse, macaque and human genomes\(^13\) — we are now certain that there has been a substantial expansion of the OR gene superfamily since the mammalian radiation \(\sim100\) million years ago.

The evolutionary change in the number of OR genes in insects is not nearly as extensive as that in mammals. *Drosophila melanogaster* has a relatively

| Family No. of | No. of functional | No. of pseudogenes |
|---------------|------------------|-------------------|
| 1             | 21               | 28                | 11               |
| 2             | 41               | 68                | 45               |
| 3             | 3                | 4                 | 2                |
| 4             | 21               | 51                | 78               |
| 5             | 49               | 47                | 65               |
| 6             | 21               | 30                | 21               |
| 7             | 9                | 11                | 100              |
| 8             | 18               | 24                | 24               |
| 9             | 12               | 9                 | 14               |
| 10            | 29               | 35                | 30               |
| 11            | 11               | 8                 | 15               |
| 12            | 1                | 2                 | 1                |
| 13            | 11               | 12                | 10               |
| 14            | 6                | 6                 | 1                |
| 51            | 21               | 23                | 21               |
| 52            | 22               | 26                | 23               |
| 55            | 1                | 0                 | 1                |
| 56            | 2                | 6                 | 3                |
| Totals        | 299              | 390               | 465              |

**Table 1. Summary of the olfactory receptor (OR) gene superfamily (18 families)**
Table 2. Summary of OR genes in families 1 to 4

| Subfamily | Genes | Pseudo | Subfamily | Genes | Pseudo | Subfamily | Genes | Pseudo |
|-----------|-------|--------|-----------|-------|--------|-----------|-------|--------|
| OR1A      | 2     | 0      | OR2I      | 0     | 1      | OR3A      | 4     | 0      |
| OR1B      | 1     | 0      | OR2J      | 2     | 2      | OR3B      | 0     | 1      |
| OR1C      | 1     | 0      | OR2K      | 1     | 0      | OR3D      | 0     | 1      |
| OR1D      | 3     | 1      | OR2L      | 5     | 3      |           |       |        |
| OR1E      | 2     | 1      | OR2M      | 5     | 1      | OR4A      | 4     | 27     |
| OR1F      | 2     | 1      | OR2N      | 0     | 1      | OR4B      | 1     | 1      |
| OR1G      | 1     | 0      | OR2P      | 0     | 1      | OR4C      | 10    | 11     |
| OR1H      | 0     | 1      | OR2Q      | 0     | 1      | OR4D      | 7     | 3      |
| OR1I      | 1     | 0      | OR2R      | 0     | 1      | OR4E      | 1     | 1      |
| OR1J      | 3     | 0      | OR2S      | 1     | 1      | OR4F      | 9     | 7      |
| OR1K      | 1     | 0      | OR2T      | 16    | 1      | OR4G      | 0     | 6      |
| OR1L      | 5     | 0      | OR2U      | 0     | 2      | OR4H      | 0     | 2      |
| OR1M      | 1     | 1      | OR2V      | 2     | 0      | OR4K      | 8     | 7      |
| OR1N      | 2     | 0      | OR2W      | 3     | 3      | OR4L      | 1     | 0      |
| OR1P      | 0     | 1      | OR2X      | 0     | 1      | OR4M      | 2     | 0      |
| OR1Q      | 1     | 0      | OR2Y      | 1     | 0      | OR4N      | 3     | 2      |
| OR1R      | 0     | 1      | OR2Z      | 1     | 0      | OR4P      | 1     | 1      |
| OR1S      | 2     | 0      | OR2AD     | 0     | 1      | OR4Q      | 1     | 2      |
| OR1X      | 0     | 2      | OR2AE     | 1     | 0      | OR4R      | 0     | 3      |
| OR1AA     | 0     | 1      | OR2AF     | 0     | 1      | OR4S      | 2     | 0      |
| OR1AB     | 0     | 1      | OR2AG     | 2     | 0      | OR4T      | 0     | 1      |
|           |       |        |           |       |        |           |       |        |
| OR2A      | 9     | 6      | OR2AJ     | 1     | 0      | OR4W      | 0     | 1      |
| OR2B      | 4     | 3      | OR2AK     | 1     | 0      | OR4X      | 2     | 2      |
| OR2C      | 2     | 0      | OR2AL     | 0     | 1      |           |       |        |
| OR2D      | 2     | 0      | OR2AM     | 0     | 1      |           |       |        |
| OR2E      | 0     | 1      | OR2AO     | 0     | 1      |           |       |        |
| OR2F      | 2     | 0      | OR2AP     | 1     | 0      |           |       |        |
| OR2G      | 3     | 1      | OR2AS     | 0     | 2      |           |       |        |
| OR2H      | 2     | 2      | OR2AT     | 1     | 2      |           |       |        |

Family OR1 genes are located on chromosomes 9, 17, 19, 11, 16, 5, 1, 6 and X.
Family OR2 genes are located on chromosomes 1, 6, 7, 11, 5, 9, 12, 16, 19 and X.
Family OR3 genes are located on chromosomes 17, 1 and X.
Family OR4 genes are located on chromosomes 11, 14, 15, 1, 19, 17, 18, 21, 6, 4, 5, 8 and X.
small receptor repertoire of 62 odourant receptors. A comparison of 12 Drosophila species, encompassing ~60 million years of divergence, shows that the number of functional OR genes has remained fairly stable. Caenorhabditis elegans has a highly developed chemosensory system, which enables it to detect a wide variety of volatile (olfactory) and water-soluble (gustatory) cues associated with food, danger or other animals; between 500 and 1,000 different GPCRs are expressed in chemosensory neurones, and these may be supplemented by alternative sensory pathways as well. The vertebrate OR gene repertoire has thus evolved from a subset of ancestral genes in the fly and worm.

There appear to be three important periods in the evolution of the vertebrate olfactory system, as evidenced by comparative genomics: (1) the adaptation to land in amphibian ancestors; (2) the decline of olfaction in primates; and (3) the delineation of putative pheromone receptors concurrent with rodent speciation. The gene: pseudogene ratio is lowest in human, higher in chimpanzee and highest in rat and mouse. This most likely reflects the necessity of olfaction for survival — more so in the rodent than in the human.

Whereas the chicken, platypus and primate genomes carry <400 functional OR genes, the opossum and rodent genomes, not surprisingly, contain between 1,000 and 1,210 functional OR genes. Curiously, however, it is difficult to explain why the cow genome, with 970 functional OR genes, shows more than the dog genome, with ~811 functional OR genes, when dogs are considered to have such a keen sense of olfaction. Thus, the number of OR genes in a species does not appear to be directly related to the environmental ‘requirement’ or to lifestyle.

**Current bioinformatics about the OR gene superfamily**

The OR gene superfamily comprises 18 gene families and 300 subfamilies (Table 1). Presently, there are 390 putatively functional (protein-coding) OR genes and 465 OR pseudogenes located in multiple clusters of varying sizes scattered throughout all autosomes except chromosome (Chr) 20, and on the X but not the Y chromosome. The members of each subfamily have been placed therein because of divergent evolution, as described above. These subfamilies differ from CYP subfamilies, in that individual members within one subfamily are often located on two or more different chromosomes. The OR2T (Table 2) subfamily contains 16 functional genes — more than in any other subfamily. Evolutionary divergence of each of the 18 gene families is illustrated in Figure 1.

**Figure 1.** A phylogenetic analysis of one representative from each family of the human OR gene repertoire. In this tree, one can see the following: (1) a general guideline for how the different families relate to one another (although this is very general, and the branching is not always this well defined); (2) the numbers near each branch denote the OR family number; (3) each pie chart size is scaled to represent the number of the OR genes inside that family (black = functional genes, grey = pseudogenes, yellow = segregating pseudogenes [SPGs]). SPGs are genes that segregate in populations between intact genes and pseudogenes — due to a disruptive SNP. This disruptive mutation can introduce a stop codon, or alter a highly conserved amino acid that is important for proper function of the protein. In Tables 1–5, the SPGs are counted as ‘functional genes’ or ‘pseudogenes’, according to the Human Genome Project public version. Additional information can be found at the HORDE database (http://bioportal.weizmann.ac.il/HORDE/).
| Subfamily | Genes | Pseudo | Subfamily | Genes | Pseudo | Subfamily | Genes | Pseudo |
|-----------|-------|--------|-----------|-------|--------|-----------|-------|--------|
| OR5A      | 2     | 0      | OR5BC     | 0     | 1      | OR7A      | 3     | 8      |
| OR5B      | 5     | 4      | OR5BD     | 0     | 1      | OR7C      | 2     | 0      |
| OR5C      | 1     | 0      | OR5BE     | 0     | 1      | OR7D      | 2     | 1      |
| OR5D      | 4     | 4      | OR5BH     | 0     | 1      | OR7E      | 1     | 85     |
| OR5E      | 0     | 1      | OR5BJ     | 0     | 1      | OR7G      | 3     | 1      |
| OR5F      | 1     | 1      | OR5BK     | 0     | 1      | OR7H      | 0     | 2      |
| OR5G      | 0     | 4      | OR5BL     | 0     | 1      | OR7K      | 0     | 1      |
| OR5H      | 5     | 5      | OR5BM     | 0     | 1      | OR7L      | 0     | 1      |
| OR5I      | 1     | 0      | OR5BN     | 0     | 2      | OR7M      | 0     | 1      |
| OR5J      | 1     | 2      | OR5BP     | 0     | 1      |           |       |        |
| OR5K      | 4     | 0      | OR5BQ     | 0     | 1      | OR8A      | 1     | 2      |
| OR5L      | 2     | 0      | OR5BR     | 0     | 1      | OR8B      | 5     | 6      |
| OR5M      | 6     | 8      | OR5BS     | 0     | 1      | OR8C      | 0     | 1      |
| OR5P      | 2     | 2      | OR5BT     | 0     | 1      | OR8D      | 3     | 0      |
| OR5R      | 1     | 0      |           |       |        | OR8F      | 0     | 1      |
| OR5S      | 0     | 1      | OR6A      | 1     | 0      | OR8G      | 2     | 3      |
| OR5T      | 3     | 0      | OR6B      | 3     | 0      | OR8H      | 3     | 0      |
| OR5V      | 1     | 0      | OR6C      | 11    | 8      | OR8I      | 1     | 2      |
| OR5W      | 1     | 1      | OR6D      | 0     | 1      | OR8J      | 2     | 1      |
| OR5AC     | 1     | 2      | OR6E      | 0     | 1      | OR8K      | 3     | 2      |
| OR5AH     | 0     | 1      | OR6F      | 1     | 0      | OR8L      | 0     | 1      |
| OR5AK     | 1     | 3      | OR6J      | 1     | 0      | OR8Q      | 0     | 1      |
| OR5AL     | 0     | 2      | OR6K      | 3     | 3      | OR8R      | 0     | 1      |
| OR5AM     | 0     | 1      | OR6L      | 0     | 2      | OR8S      | 1     | 0      |
| OR5AN     | 1     | 1      | OR6M      | 1     | 2      | OR8T      | 0     | 1      |
| OR5AO     | 0     | 1      | OR6N      | 2     | 0      | OR8U      | 3     | 0      |
| OR5AP     | 1     | 1      | OR6P      | 1     | 0      | OR8V      | 0     | 1      |
| OR5AQ     | 0     | 1      | OR6Q      | 1     | 0      | OR8X      | 0     | 1      |
| OR5AR     | 1     | 0      | OR6R      | 0     | 2      |           |       |        |
| OR5AS     | 1     | 0      | OR6S      | 1     | 0      |           |       |        |
| OR5AU     | 1     | 0      | OR6T      | 1     | 0      |           |       |        |

Continued
Table 3. Continued

| Subfamily | Genes | Pseudo | Subfamily | Genes | Pseudo | Subfamily | Genes | Pseudo |
|-----------|-------|--------|-----------|-------|--------|-----------|-------|--------|
| OR5AW     | 0     | 1      | OR6U      | 0     | 1      | OR6V      | 1     | 0      |
| OR5AZ     | 0     | 1      | OR6W      | 0     | 1      | OR6X      | 1     | 0      |
| OR5BA     | 0     | 1      | OR6Y      | 1     | 0      |
| OR5BB     | 0     | 1      |           |       |        |           |       |        |

Family OR5 genes are located on chromosomes 11, 3, 6, X, 14, 19, 2, 4 and 9.
Family OR6 genes are located on chromosomes 12, 1, 11, 14, 10, 7, 2 and 8.
Family OR7 genes are located on chromosomes 19, 11, 3, 8, 4, 13, 2, 12, 7, 10, 14, 9, 21, 5 and X.
Family OR8 genes are located only on chromosomes 11 and 12.

Table 4. Summary of OR genes in families 9 to 13

| Subfamily | Genes | Pseudo | Subfamily | Genes | Pseudo | Subfamily | Genes | Pseudo |
|-----------|-------|--------|-----------|-------|--------|-----------|-------|--------|
| OR9A      | 2     | 2      | OR10A     | 6     | 0      | OR11A     | 1     | 0      |
| OR9G      | 3     | 2      | OR10B     | 0     | 1      | OR11B     | 1     | 1      |
| OR9H      | 0     | 1      | OR10C     | 1     | 0      | OR11H     | 5     | 4      |
| OR9I      | 1     | 2      | OR10D     | 0     | 4      | OR11I     | 0     | 1      |
| OR9K      | 1     | 1      | OR10E     | 7     | 2      | OR11J     | 0     | 3      |
| OR9L      | 0     | 1      | OR10F     | 5     | 0      | OR11K     | 0     | 2      |
| OR9M      | 0     | 1      | OR10G     | 3     | 6      | OR11L     | 1     | 0      |
| OR9N      | 0     | 1      | OR10H     | 2     | 0      | OR11M     | 0     | 1      |
| OR9P      | 0     | 1      | OR10I     | 0     | 1      | OR11N     | 0     | 1      |
| OR9Q      | 2     | 0      | OR10J     | 1     | 0      | OR11P     | 0     | 1      |
| OR9R      | 0     | 1      | OR10K     | 1     | 1      | OR11Q     | 0     | 1      |
| OR9S      | 0     | 1      | OR10L     | 1     | 2      |           |       |        |
| OR10S     | 1     | 0      | OR12D     | 1     | 2      |
| OR10T     | 1     | 1      |
| OR10U     | 0     | 1      | OR13A     | 1     | 0      |
| OR10V     | 1     | 2      | OR13B     | 6     | 3      |
| OR10W     | 1     | 0      | OR13D     | 1     | 2      |
| OR10X     | 1     | 0      | OR13E     | 0     | 1      |
| OR10Y     | 0     | 1      | OR13F     | 1     | 0      |
| OR10Z     | 1     | 0      | OR13G     | 1     | 0      |
| OR10AA    | 0     | 1      | OR13H     | 1     | 0      |
| OR10AB    | 0     | 1      | OR13I     | 0     | 1      |

Continued
Table 4. Continued

| Subfamily | Genes | Pseudo | Subfamily | Genes | Pseudo | Subfamily | Genes | Pseudo |
|-----------|-------|--------|-----------|-------|--------|-----------|-------|--------|
| OR10AC    | 0     | 1      | OR13J     | 1     | 0      | OR13K     | 0     | 1      |
| OR10AD    | 1     | 0      | OR13K     | 0     | 1      | OR13Z     | 0     | 2      |
| OR10AE    | 0     | 2      | OR13Z     | 0     | 2      | OR10AF    | 0     | 1      |
| OR10AG    | 1     | 0      | OR14A     | 2     | 0      | OR10AH    | 0     | 1      |
| OR10AH    | 0     | 1      | OR14C     | 1     | 0      | OR10AK    | 0     | 1      |
| OR10AK    | 0     | 1      | OR14I     | 1     | 0      | OR14J     | 1     | 0      |
|           |       |        | OR14K     | 1     | 0      | OR14L     | 0     | 1      |

Family OR9 genes are located on chromosomes 11, 7, 12, 1 and 2.
Family OR10 genes are located on chromosomes 11, 1, 19, 12, 14, 7 and 6.
Family OR11 genes are located on chromosomes 14, 15, 1, X, 6, 12 and 22.
Family OR12 genes are located only on chromosome 6.
Family OR13 genes are located on chromosomes 9, 1, X and 10.
Family OR14 genes are located only on chromosome 1.

Table 5. Summary of OR genes in families 51, 52, 55 and 56

| Subfamily | Genes | Pseudo | Subfamily | Genes | Pseudo | Subfamily | Genes | Pseudo |
|-----------|-------|--------|-----------|-------|--------|-----------|-------|--------|
| OR51A     | 3     | 7      | OR52A     | 3     | 0      | OR55B     | 0     | 1      |
| OR51B     | 4     | 2      | OR52B     | 3     | 3      | OR56A     | 4     | 1      |
| OR51C     | 0     | 2      | OR52D     | 1     | 0      | OR56A     | 4     | 1      |
| OR51D     | 1     | 0      | OR52E     | 5     | 3      | OR56B     | 2     | 2      |
| OR51E     | 2     | 0      | OR52H     | 1     | 1      |           |       |        |
| OR51F     | 2     | 3      | OR52I     | 2     | 0      |           |       |        |
| OR51G     | 2     | 0      | OR52J     | 1     | 2      |           |       |        |
| OR51H     | 0     | 2      | OR52K     | 2     | 1      |           |       |        |
| OR51I     | 2     | 0      | OR52L     | 1     | 1      |           |       |        |
| OR51J     | 1     | 0      | OR52M     | 1     | 1      |           |       |        |
| OR51K     | 0     | 1      | OR52N     | 4     | 1      |           |       |        |
| OR51L     | 1     | 0      | OR52P     | 0     | 2      |           |       |        |
| OR51M     | 1     | 0      | OR52Q     | 0     | 1      |           |       |        |
| OR51N     | 0     | 1      | OR52R     | 1     | 0      |           |       |        |
| OR51P     | 0     | 1      | OR52S     | 0     | 1      |           |       |        |
| OR51Q     | 1     | 0      | OR52T     | 0     | 1      |           |       |        |

Continued
Note that, in many instances, some subfamilies contain only a single gene or only a single pseudogene (Tables 2–5). In fact, the OR7E subfamily has only one functional gene, and all the other 85 members are pseudogenes (Table 3). The OR7E subfamily is the largest subfamily in the human OR gene repertoire, and probably has expanded in the human genome through a series of segmental gene duplication events. The newly described human OR14 gene family (Table 4) was realised after analysis of the platypus and opossum OR gene repertoires. This analysis revealed that six human OR functional genes and one OR pseudogene (which previously had been classified within the OR5 family) are actually derived from a distinct platypus OR gene family. The evolutionary divergence of the OR14 gene family is shown in Figure 2.

The ‘shotgun’ splattering of OR genes throughout the human genome must have happened before speciation of Homo sapiens and the development of its 22 autosomes plus the X and Y chromosomes; this can be inferred from the high conservation of the OR genes’ genomic organisation among marsupial and eutherian mammals, and the phylogenetic analysis of the platypus OR gene repertoire — by comparison with that in mammals. In contrast to this OR gene arrangement would be the establishment of the CYP gene subfamilies, which arose as syntenic clusters of members within a single chromosomal segment. This finding suggests that gene duplication events within CYP subfamilies occurred after mammalian speciation and development of the autosomes and sex chromosomes.

The two largest OR gene clusters are located on Chr 11, with 38 functional genes (51 per cent of total) on 11q (Cluster 11@5.0) and 44 functional genes (45 per cent) on 11p (Cluster 11@55.6). These genes are predominantly in OR families 51, 52, 55 and 56 (Table 5). Immersed within these two clusters are dozens of other non-OR-related genes. This intrusion of other non-OR-related genes can also be seen in all other OR gene clusters throughout the genome.

**Future directions: Additional subsets of sensory reception genes and identification of ligands**

A recently appreciated discovery in olfaction is the unique specialisation of sensory neurones, such that each individual sensory neurone is stochastically chosen to express usually only one odourant receptor allele. This mechanism of ‘allelic exclusion’, by which mutually exclusive expression of odourant receptor genes is regulated, remains unclear at present.

The vomeronasal-1 receptor genes (VN1R) also encode GPCRs and, while they encode odourant receptors, they are evolutionarily distinct from the very large OR gene superfamily. There are five VN1R genes and nine VN1R pseudogenes. The VN1R1, VN1R2 and VN1R4 genes and VN1R6P pseudogene are located at Chr 19q13.42; the VN1R10, VN1R11, VN1R12, VN1R13 and VN1R14P pseudogenes are located on Chr 6p21; VN1R7 and VN1R8P are on Chr 21p11.2;
Figure 2. A phylogenetic analysis of platypus, opossum and human OR genes for the new family 14 only. Opossum = black for intact genes, grey for pseudogenes. Platypus = red for intact genes, pink for pseudogenes. Human = blue. OR14 is an expansion of three ancestral OR gene subfamilies (A, B and C); the expansion, in both platypus and opossum, took place after speciation, whereas only one branch shows an orthologous relationship between platypus and human (marked with *). The tree was generated with MEGA4, using the nearest-neighbour-joining algorithm, and distances with the Poisson correction model. Bootstrap units are also indicated.34 This tree is grounded with the OR51E1 gene. Only genes with no more than two frame disruptions were considered in the analysis.
VN1R3 is alone on Chr 16p11.2; VN1R5 is alone on Chr 1q44; VN1R9P is alone on Chr 22.28

At the present time, information about the ligands for mammalian OR genes is very limited. The smell of lemons (limonene), the perception of a floral or woody smell (acetophenone)29 and the ability to smell isovaleric acid30 have been mapped of sugars, bitter substrates, carbon dioxide and neurones and are known to mediate the perception of the arista, peripheral proprioceptive neurones in the legs, neurones in the larval and adult brain, and oenocytes.31,32 Among these same lines, we and others have observed several OR genes being significantly up- or downregulated in the liver or kidney of knockout mouse lines — that is, in tissues not normally known to be involved in olfaction. It is therefore tempting to speculate that the receptors encoded by OR genes, as well as by Gr genes, might participate in the roles of detecting endogenous ligands.

Acknowledgments

We thank our colleagues for valuable discussions and a critical reading of this manuscript. The writing of this article was funded, in part, by NIH grant P30 ES06096 (D.W.N.).

References

1. Dayhoff, M.O., Barker, W.C. and Hunt, L.T. (1983), ‘Establishing homologies in protein sequences’, Methods Enzymol. Vol. 91, pp. 524–545.
2. Nebert, D.W., Adesnik, M., Coon, M.J. et al. (1987), ‘The P450 gene superfamily. Recommended nomenclature’, DNA Vol. 6, pp. 1–11.
3. Nebert, D.W. and Gonzalez, EJ (1987), ‘P450 genes: Structure, evolution, and regulation’, Ann. Rev. Biochem. Vol. 56, pp. 945–993.
4. Nebert, D.W., Nelson, D.R., Adesnik, M. et al. (1989), ‘The P450 superfamily: Updated listing of all genes and recommended nomenclature for the chromosomal loci’, DNA Vol. 8, pp. 1–13.
5. Nebert, D.W., Nelson, D.R., Coon, M.J. et al. (1991), ‘The P450 superfamily: Update on new sequences, gene mapping, and recommended nomenclature’, DNA Cell Biol. Vol. 10, pp. 1–14.
6. Nelson, D.R., Kamataki, T., Wixman, D.J. et al. (1993), ‘The P450 superfamily: Update on new sequences, gene mapping, accession numbers, early trivial names of enzymes, and nomenclature’, DNA Cell Biol. Vol. 12, pp. 1–51.
7. Nelson, D.R., Koymans, L., Kamataki, T. et al. (1996), ‘P450 superfamily: Update on new sequences, gene mapping, accession numbers and nomenclature’, Pharmacogenet. In press.
8. http://drnelson.utmem.edu/CytochromeP450.html.
9. Buck, L. and Axel, R. (1991), ‘A novel multigene family may encode odorant receptors: A molecular basis for odor recognition’, Cell Vol. 65, pp. 175–187.
10. Gaillard, I., Rouquier, S. and Giorgi, D. (2004), ‘Olfactory receptors’, Cell Mol. Life Sci. Vol. 61, pp. 456–469.
11. Aloni, R., Olender, T. and Lancet, D. (2006), ‘Ancient genomic architecture for mammalian olfactory receptor clusters’, Genome Biol. Vol. 7, p. R88.
12. Henion, T.R. and Schwarting, G.A. (2007), ‘Patternning the developing and regenerating olfactory system’, J. Cell Physiol. Vol. 210, pp. 290–297.
13. Niimura, Y. and Nei, M. (2007), ‘Evolutionary dynamics of olfactory receptor genes in mammalian evolution’, PLoS ONE Vol. 2, p. e708.
14. Menashe, I., Man, O., Lancet, D. and Gilad, Y. (2003), ‘Different noses for different people’, Nat. Genet. Vol. 34, pp. 143–144.
15. Nguyen, D.Q., Webber, C. and Ponting, C.P. (2006), ‘Bias of selection on human copy-number variants’, PLoS Genet. Vol. 2, p. e20.
16. Hasin, Y., Olender, T., Khen, M. et al. (2008), ‘High resolution copy-number variation map reflects human olfactory receptor diversity and evolution’, PLoS Genet. in press.
17. Smith, D.P. (2007), ‘Olfactory receptor gene family in Drosophila melanogaster’, Pfleger’s Arch. Vol. 454, pp. 749–758.
18. Nozawa, M. and Nei, M. (2007), ‘Evolutionary dynamics of olfactory receptor genes in Drosophila species’, Proc. Natl. Acad. Sci. USA Vol. 104, pp. 7122–7127.
19. Bargmann, C.I. (2006), ‘Chemosensation in C. elegans’, Wormbook, October 25, pp. 1–29.
20. Kambere, M.B. and Lane, R.P. (2007), ‘Co-regulation of a large and rapidly evolving repertoire of odorant receptor genes’, BMC Nucosci. Vol. 8 (Suppl. 3), p. S2.
21. http://bioprortal.weizmann.ac.il/HORDE/.
22. Safran, M., Chalifa-Caspi, V., Shmueli, O. et al. (2003), ‘Human gene-centric databases at the Weizmann Institute of Science: GeneCards, UDB, CroW 21 and HORDE’, Nucleic Acids Res. Vol. 31, pp. 142–146.
23. Olender, T., Feldmesser, E., Arotat, T. et al. (2004), ‘The olfactory receptor universe — From whole genome analysis to structure and evolution’, Genet. Mol. Res. Vol. 3, pp. 545–553.
24. Yue, Y. and Haaf, T. (2006), ‘7E olfactory receptor gene clusters and evolutionary chromosome rearrangements’, Cytogenet. Genome Res. Vol. 112, pp. 6–10.
25. Warren, W.C., Hillier, L.W., Marshall-Graves, J.A. et al. (2008), ‘Genome analysis of the platypus reveals unique signatures of evolution’, Nature Vol. 453, pp. 175–183.
26. Serizawa, S., Miyamichi, K., Nakatani, H. et al. (2007), ‘Negative feedback regulation ensures the one receptor-one olfactory neuron rule in mouse’, Science Vol. 320, pp. 2088–2094.
27. Rodriguez, I. (2007), ‘Odorant and pheromone receptor gene regulation in vertebrates’, Curr. Opin. Genet. Dev. Vol. 17, pp. 465–470.
28. http://www.genenames.org/.
29. Malnic, B., Godfrey, P.A. and Buck, L.B. (2004), ‘The human olfactory receptor gene family’, Proc. Natl. Acad. Sci. USA Vol. 101, pp. 2584–2589.
30. Griff, I.C. and Reed, R.R. (1995), ‘The genetic basis for specific anosmia to isovaleric acid in the mouse’, Cell Vol. 83, pp. 407–414.
31. Menashe, I., Abaffy, T., Hasin, Y. et al. (2007), ‘Genetic elucidation of human hyperosmia to isovaleric acid’, PLoS Biol. Vol. 5, p. e284.
32. Keller, A., Zhuang, H., Chi, Q. et al. (2007), ‘Genetic variation in a human odorant receptor alters odour perception’, Nature Vol. 449, pp. 468–472.
33. Thorne, N. and Amrein, H. (2006), ‘Atypical expression of Drosophila gustatory receptor genes in sensory and central neurons’, J. Comp. Neurol. Vol. 506, pp. 548–568.
34. Tamura, K., Dudley, J., Nei, M. and Kumar, S. (2007), ‘MEGA4: Molecular evolutionary genetics analysis (MEGA) software version 4.0’, Mol. Biol. Evol. Vol. 24, pp. 1596–1599.