Dramatic variation of the vomeronasal pheromone receptor gene repertoire among five orders of placental and marsupial mammals

Wendy E. Grus*†, Peng Shi*†‡, Ya-ping Zhang‡§, and Jianzhi Zhang*¶

*Department of Ecology and Evolutionary Biology, University of Michigan, Ann Arbor, MI 48109; ‡Laboratory of Cellular and Molecular Evolution, Kunming Institute of Zoology, Chinese Academy of Sciences, Kunming 650223, China; and §Laboratory for Conservation and Utilization of Bio-resources, Yunnan University, Kunming 650091, China

Communicated by Masatoshi Nei, Pennsylvania State University, University Park, PA, February 25, 2005 (received for review January 15, 2005)

Pheromones are chemicals emitted and sensed by conspecifics to elicit social and sexual responses and are perceived in terrestrial vertebrates primarily by the vomeronasal organ (VNO). Pheromone receptors in the mammalian VNO are encoded by the V1R and V2R gene superfamilies. The V1R superfamily contains 187 and 102 putatively functional genes in the mouse and rat, respectively. To investigate whether this large repertoire size is typical among mammals with functional VNOs, we here describe the V1R repertoires of dog, cow, and opossum based on their draft genome sequences. The dog and cow have only 8 and 32 intact V1R genes, respectively. Thus, the intact V1R repertoire size varies by at least 23-fold among placental mammals with functional VNOs. To our knowledge, this size ratio represents the greatest among-species variation in gene family size of all mammalian gene families. Phylogenetic analysis of placental V1R genes suggests multiple losses of ancestral genes in carnivores and artiodactyls and gains of many new genes by gene duplication in rodents, manifesting massive gene births and deaths. We also identify 49 intact opossum V1R genes and discover independent expansions of the repertoire in placentals and marsupials. We further show a concordance between the V1R repertoire size and the complexity of VNO morphology, suggesting that the latter could indicate the sophistication of pheromone communications within species. In sum, our results demonstrate tremendous diversity and rapid evolution of mammalian V1R gene inventories and caution the generalization of VNO biology from rodents to all mammals.

evolution | dog | cow | opossum | rodents

Pheromones provide conspecific chemical communications that elicit sexual and social changes in behavior and physiology (1). For instance, pheromonal peptides found in mouse urine can mediate individual recognition, induce early puberty, or block pregnancy (1), and pheromones can control behaviors such as maternal aggression (2). In mammals, pheromones are primarily sensed by the vomeronasal organ (VNO), which resides on the bottom of the nasal cavity and is anatomically and physiologically separated from the main olfactory system that detects odorants (1). In rodents, two superfamilies of seven-transmembrane G protein-coupled receptors, V1Rs and V2Rs, serve as pheromone receptors (3). Individual V1R genes are composed of ≈1,000 nucleotides without introns and are expressed in vomeronasal sensory neurons whose cell bodies are located in the apical part of the VNO epithelium (4). These cells also express a G protein subunit named $G\alpha_{i2}$ (3). In contrast, the multiexon V2R genes are characterized by a long, highly variable N-terminal domain and are coexpressed with $G\alpha_O$ in sensory neurons whose cell bodies are basally located (5–7). Both receptor superfamilies were originally identified in rats, and each of them was estimated to contain ≈100 genes (4–7). Because of the simpler gene structure, the complete repertoire of the V1R superfamily has been described in mice, rats, and humans (8–12). The superfamily has 187 and 102 functional genes in the mouse and rat, respectively (8, 12). In humans, ≈200 V1R sequences have been identified, although only four of them have intact ORFs in the majority of individuals (9, 13). The massive V1R pseudogenization observed in the human genome started shortly before the separation of hominoids (i.e., humans and apes) from Old World monkeys, apparently because of the reduced importance of vomeronasal pheromone communications, and the four human V1R ORFs are likely relics of this ongoing pseudogenization process (13).

The morphological complexity of the VNO as well as the complexity of VNO-mediated pheromone communications varies substantially among different mammals (14). These variations led to the hypothesis that the complexity of the V1R and V2R repertoires may also vary greatly among species. However, this hypothesis has not been rigorously tested, because the absence of highly conserved regions in V1Rs makes it difficult to design degenerate primers to amplify a large number of genes across a wide taxonomic scale (15, 16). Although a few limited studies in selected mammals suggested possible among-species variation in the numbers of V1R genes and pseudogenes (15, 17), the extent of this variation and the exact numbers of genes and pseudogenes in these species are unknown. Furthermore, the long-term evolutionary pattern of V1R genes in mammals and other vertebrates remains largely uncharacterized. With the recent availability of the draft genome sequences of dog, cow, and opossum, we now describe the complete V1R repertoires of this diverse array of species. We report a variation in the size of the V1R repertoire among five orders of placental and marsupial mammals that is surprisingly large and unprecedented in any other mammalian gene family.

Materials and Methods

Database Searches. TBLASTN searches for V1R genes were conducted on the dog (*Canis familiaris*), cow (*Bos taurus*), and opossum (*Monodelphis domestica*) genome sequences. The 7.6× coverage dog genome sequence (www.ncbi.nlm.nih.gov/genome/guide/dog) and 3.3× coverage cow genome sequence (www.ncbi.nlm.nih.gov/genome/guide/cow) are available in the National Center for Biotechnology Information. The 7.2× coverage opossum genome sequence is available at ENSEMBL (http://pre.ensembl.org/Monodelphis_domestica). Mouse and rat V1R genes from refs. 8, 10, and 12 were used as query sequences. Putative V1Rs were identified with an E value cutoff of 10^{-5}. They were then used as queries to BLAST the NR database of GenBank. A putative V1R gene was considered to be real if its best hit was a previously known V1R. Use of human V1Rs from ref. 9 as query sequences did not yield additional V1Rs.

Abbreviations: MY, million years; VNO, vomeronasal organ.

†W.E.G. and P.S. contributed equally to this work.

¶To whom correspondence should be addressed. E-mail: jianzhi@umich.edu.

V1Rs were considered pseudogenes if they contained premature stop codons or were incomplete across the 13 internal domains (7 transmembrane, 3 extracellular, and 3 intracellular). The database searches were independently conducted, yielding identical results. The sequences of newly identified intact V1Rs from the dog, cow, and opossum are listed in Data Set 1, which is published as supporting information on the PNAS web site.

Sequence Alignment and Phylogenetic Analysis. Gene sequences were aligned per protein sequence alignment by CLUSTALX (18) with manual adjustment. Phylogenetic trees were reconstructed by using the neighbor-joining method (19) with protein-Poisson distances (20) and were evaluated by 1,000 bootstrap replications (21). We also used protein p-distance (20) and found that the branching patterns with high bootstrap support in the Poisson-distance tree remained unchanged in the *p*-distance tree. After convention (10), we defined gene families by a minimum of 40% amino acid identity among all family members and confirmed the monophyly of gene families by phylogenetic analysis. We conceptually translated each pseudogene sequence according to its alignment with all functional V1R genes (mouse, rat, and dog or cow), and determined the phylogenetic position of each pseudogene by making a neighbor-joining tree with the protein sequence of the pseudogene and those of all functional genes. The distribution of LINE repetitive elements was determined by the REAPEATMASKER program (www.repeatmasker.org). For easy comparison, we used the same criterion as used in ref. 22 to calculate the L1 density. More specifically, we estimated the L1 density in the intergenic regions for tandem-linked V1R genes and that in 10,000 nucleotides upstream and downstream of the coding regions for nontandem-linked V1R genes.

Results

Dramatic Variation in V1R Repertoire Size in Placental Mammals. Placental mammals can be classified into four superorders (23). Rodents and primates belong to the superorder Euarchontoglires, to which all previous V1R evolutionary analyses have been restricted. Here, we investigate the V1R repertoire in the dog and cow, which are members of the superorder Laurasiatheria, the sister clade to Euarchontoglires. From the dog genome sequence, we identified eight complete V1R ORFs, which are presumably functional genes. In addition, 22 V1R pseudogenes were detected. Because the dog genome sequence has a high (7.6×) coverage, it is likely that the majority of, if not all, dog V1R genes have been found. Similarly, we detected 32 functional V1R genes and 41 pseudogenes from the cow genome sequence. However, because the cow sequence has a relatively low coverage (3.3×), it is possible that a few additional genes and pseudogenes may be discovered when a more complete genome sequence becomes available. Even with this limitation, our observations clearly show that the V1R repertoires in cow and dog are substantially smaller than those in mouse and rat, which contain 187 and 102 putatively functional genes, respectively (12). Even when humans are disregarded, there is still a 23-fold variation in the size of functional V1R repertoire among rodents, artiodactyls, and carnivores (Fig. 1). The proportion of putatively functional V1R genes is lower in dogs (27%) than in rodents (49%, refs. 10 and 11), but the proportion in cows (44%) is similar to that in rodents.

Phylogenetic Relationships of Placental V1R Genes. To understand the evolutionary relationships of the V1R genes from the four placental orders (rodents, primates, artiodactyls, and carnivores), we reconstructed a protein neighbor-joining tree with all putatively functional V1Rs from the mouse, rat, human, dog, and cow (Fig. 2). We also included four human V1Rs with complete ORFs in the phylogenetic analysis, although these ORFs are likely relics of an ongoing pseudogenization process (13). A fifth human V1R gene, *hV1RL3*, was not included because a previous study found this gene to be nearly fixed with a nonfunctional allele in human populations (allele frequency >98%) (13). Putatively functional nonhuman primate V1R sequences from ref. 24 were not included because only partial sequences were available.

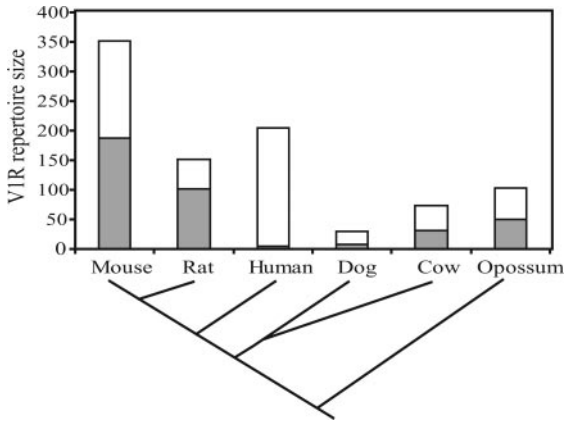

Fig. 1. Variation in the numbers of intact V1R genes (filled bars) and pseudogenes (open bars) among mice, rats, humans, dogs, cows, and opossums with the phylogenetic relationships among the species shown by the tree. Mice have 187 functional V1Rs and 168 pseudogenes (11, 12). Rats have 102 functional V1Rs and at least 50 pseudogenes (ref. 12 and P.S., unpublished data). Humans have 4 V1R ORFs and ≈200 pseudogenes (9, 13). Cows have 32 functional V1Rs and 41 pseudogenes (this study). Dogs have 8 functional V1Rs and 22 pseudogenes (this study). Opossums have 49 functional V1Rs and 53 pseudogenes (this study). The number of pseudogenes in rats is likely an underestimate because there has been no comprehensive study of rat V1R pseudogenes.

The dog and cow V1R genes cluster within the previously described rodent V1R superfamily (8). The mouse and rat V1R genes were previously classified into 14 families (V1RA to V1RN) in ref. 8. A new family (V1RO) is found when the seven recently described rat V1R genes (12) are added (see Fig. 2). Because the families were originally described in rodents, some of the families could be rodent-specific, meaning that their origins postdated the origin of rodents (8, 23). To classify other mammalian V1Rs, we clustered the families into family groups. The family groups contain families, described in rodents, that split around the time of the most recent common ancestor of rodents, cow, and dog. The 15 rodent families may be grouped into nine family groups based on phylogeny (Fig. 2 and Table 1). Based on the rodent synonymous substitution rate, we previously estimated that the V1R family groups appeared before 95 MY ago (8). This finding suggests that the most recent common ancestor of primates, rodents, artiodactyls, and carnivores should have the V1R family groups that are currently observed in rodents, because this ancestor lived ≈95 MY ago (23, 25). Note that this estimate for the age of the V1R family groups is conservative because it is likely that the synonymous substitution rate has been enhanced in rodents compared with that in other mammals (26, 27). We found that all nine family groups contain putatively functional mouse genes and eight of nine (except V1RH/I) contain functional rat genes (Fig. 2 and Table 1). However, only five of the nine family groups include functional dog genes and five include functional cow genes (Fig. 2), suggesting that cows and dogs lost many ancestral V1R genes. For instance, the family groups V1RC and V1RG each contain >10 mouse genes and 10 rat genes, but neither contains functional dog or cow V1Rs. Additionally, there are no functional dog V1RD or V1RE genes, and there are no functional cow V1RH/I genes. Surprisingly, the second-largest V1R gene family

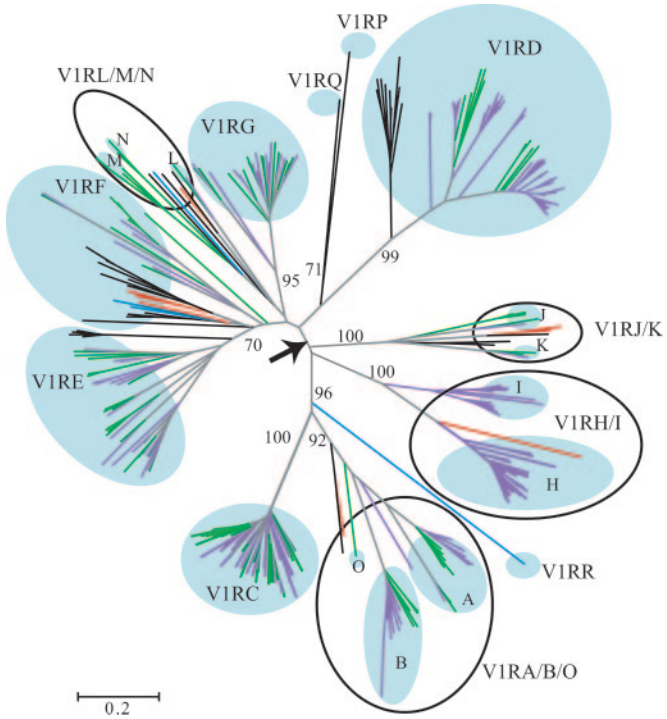

Fig. 2. Phylogeny of intact V1R genes: 8 in dogs, 32 in cows, 4 in humans, 102 in rats, and 187 in mice. Shaded regions group the genes into 18 placental mammalian V1R families previously described in refs. 8, 10, and 12 or described here (see *Results*), with the family names indicated. Black circles mark family groups that contain more than one family as shown in Table 1. Dog branches are in red, cow branches are in black, human branches are in blue, mouse branches are in purple, and rat branches are in green. Bootstrap percentages supporting the family groups are shown if >50. The tree was reconstructed by using the neighbor-joining method with Poisson-corrected protein distances. The arrow points to where the tree is rooted with putative V1Rs of the frog *Xenopus tropicalis* (W.E.G. and J.Z., unpublished data). (Scale bar: 0.2 amino acid substitutions per site.)

in cows (V1RD) contains no functional dog genes. We also identified two cow-specific families, V1RP and V1RQ, each containing a single gene.

Interestingly, several family groups that do not contain functional cow or dog genes possess their pseudogenes. We were able to classify 17 of the 22 dog V1R pseudogenes and 35 of 41 cow V1R pseudogenes by family group (Table 1; see also Fig. 5, which is published as supporting information on the PNAS web site). The remaining pseudogenes have degenerated too much to be included. Family groups V1RC and V1RD each contain two dog pseudogenes, but no functional dog genes. When both functional genes and pseudogenes are considered, only two family groups in dog (V1RE and V1RG) and two in cow (V1RC and V1RG) have been completely lost if the unclassified pseudogenes do not belong to these family groups. These observations provide further evidence that the small sizes of the dog and the cow V1R repertoires can be partially explained by loss of ancestral genes.

Table 1. V1R gene family groups in five placental mammals

Family group	Mouse*	Rat*	Dog†	Cow†	Human‡
A/B/O	19	15	1 (2)	1 (1)	0
C	32	23	0 (2)	0 (0)	0
D	56	8	0 (2)	9 (7)	0
E	16	22	0 (0)	4 (2)	0
F	5	8	3 (5)	10 (19)	2
G	21	13	0 (0)	0 (0)	0
H/I	35	0	1 (3)	0 (1)	0
J/K	2	6	2 (2)	3 (1)	0
L/M/N	1	7	1 (1)	3 (4)	1
P	0	0	0 (0)	1 (0)	0
Q	0	0	0 (0)	1 (0)	0
R	0	0	0 (0)	0 (0)	1
Total	187	102	8 (17)	32 (35)	4

*Refs. 8 and 12.
†This study; numbers in parentheses are pseudogenes that could be classified into family groups. The remaining pseudogenes were too degenerated to determine their family groups.
‡Refs. 9 and 13.

The second factor causing the repertoire-size variation among species is the lineage-specific expansion of families as occurred most prominently in rodents. The phylogenetic analysis shows that most of the dog and cow V1Rs diverged from their closest rodent homologs before the expansions of rodent V1R families (Fig. 2). This divergence pattern is consistent with our previous estimate that the earliest duplication events within rodent V1R families took place ≈88 MY ago (8), postdating the separation of rodents, carnivores, and artiodactyls ≈95 MY ago (23, 25). Family expansions were virtually absent in the dog but were evident in three V1R families of cow. Only two dog families, V1RF and V1RJ/K, contain more than one V1R gene. The other three dog V1Rs are single genes in V1RL/M/N, V1RH, and V1RA/B/O, respectively. Six of the 22 dog pseudogenes were part of the two larger dog V1R families. In contrast, three cow V1R families, V1RD, V1RE, and V1RF, exhibited the duplications characteristic of the rodent V1R gene families. The remaining cow V1Rs are in V1RA/B/O, V1RJ/K, V1RL/M/N, V1RP, and V1RQ. Although some of these families contain more than one cow gene (Table 1), the multiple genes are not the product of species-specific duplication events. Twenty-eight of the 41 cow pseudogenes were part of the three largest cow families. The dog-cow V1RF clade was similar to the rodent gene families with species-specific duplications (Fig. 2).

Based solely on comparison with the mouse, Rodriguez and Mombaerts (9) reported that human V1R genes do not belong to the reported V1R families. Our analysis showed that three human V1Rs can be classified into V1RF and V1RL/M/N, and the fourth (hV1RL5) forms a new human-specific family (V1RR). Giorgi and Rouquier (15) identified several V1R sequences from the chimpanzee, gorilla, and orangutan. We found that these sequences cluster closely with the human sequences, although we did not present them in the phylogeny of Fig. 2 because they are not from complete V1R repertoires. Thus, when all four placental orders are considered, there are 12 V1R family groups, 10 of which have moderate to high bootstrap support (Fig. 2). In the case of V1RF, we maintained the family by at least 40% amino acid identity among all genes. Fig. 2 also shows V1RF as paraphyletic; however, the bootstrap values for the deep branches defining V1RF are low, indicating that it could be monophyletic.

Opossum V1R Repertoire. We also identified 49 putatively functional V1R genes and 53 pseudogenes from the opossum genome sequence. Because the opossum genome sequence has a high coverage (7.2×), we expect that almost all opossum V1R genes have been detected. We reconstructed a phylogenetic tree of the 49 opossum V1Rs with all functional V1Rs of the mouse, rat, dog, and cow (Fig. 3). The phylogeny shows that the opossum genes can be classified into eight opossum-specific families (oV1RA to oV1RH). The families range in size from a single gene (oV1RH) to 15 genes (oV1RA and oV1RC), with variable levels of bootstrap support. The tree shows that the placental and marsupial genes do not form two separate monophyletic groups, suggesting that more than one V1R gene was present in the

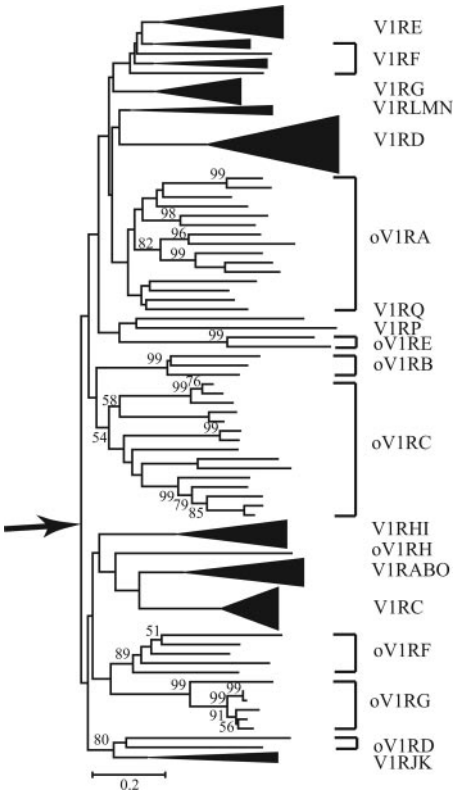

Fig. 3. Phylogeny of intact V1R genes from the opossum, mouse, rat, dog, and cow. V1R families of placental mammals (V1RA-V1RQ) have been collapsed for better illustration, but all 49 V1Rs of the opossum are shown. Opossum genes are divided into eight families (oV1RA-oV1RH). The tree was reconstructed by using the neighbor-joining method with Poisson-corrected protein distances. Bootstrap percentages are shown if >50. The arrow points to where the tree is rooted with *X. tropicalis* V1Rs. (Scale bar: 0.2 amino acid substitutions per site.)

common ancestor of placentals and marsupials. However, it is difficult to estimate the number of V1R genes in the common ancestor because of the low resolution of deep nodes in the tree. Nevertheless, the presence of many well supported opossum-specific and placental-specific gene clusters in the tree provides strong phylogenetic evidence that V1R families expanded independently in marsupials and placentals.

Discussion

In this study, we described the vomeronasal pheromone receptor V1R gene superfamily from the dog, cow, and opossum, extended the study of the superfamily outside of rodents and primates, and revealed extremely high variation in the sizes of functional V1R repertoires among mammals. The sizes of the dog and cow V1R repertoires are vastly smaller than those of rodents and primates. We found only 8 putatively functional V1Rs in dogs and 32 in cows compared with 102 in rat and 187 in mouse (8, 10–12). In humans, the functional V1R repertoire is also small with only four ORFs in most individuals (9, 13). However, the entire human V1R repertoire, including both functional genes and pseudogenes, is quite large, with ≈200 members. In opossum, we identify 49 putatively functional V1Rs, intermediate among what we identify here for dog and cow and what had been previously identified for rodents.

Even when humans who have lost functional VNOs are disregarded, the size of the functional V1R repertoire varies by >23-fold among all mammals or among placental mammals (Table 2). Several gene families, particularly those involved in sensory, immune, and reproductive functions, are known to vary substantially in size among mammalian species (27). For instance, the number of functional olfactory receptor genes in rat is approximately four times that in humans (27, 33). The putatively anti-parasitic eosinophil-associated RNase gene family is 6–17 times larger in rodents than in New World monkeys (29, 30, 37). The human X-linked testis-expressed homeobox genes *OTEX* and *PEPP2* have a total of 15 orthologous genes in the mouse genome, because of multiple gene duplications that postdated the primate-rodent separation (32). The human genome contains >200 Ig heavy chain variable region (VH) genes and ≈80 of them are functional (38). Rabbit also contains >100 VH genes, but only one of them is predominantly used, resulting in very few functional genes (39). The exact number of functional rabbit VH genes is yet to be determined, although at least five have been identified (39). Thus, the number of functional VH genes may vary up to 16-fold among different mammals. Table 2 lists additional gene families known to have wide variations of family size among mammalian species. However, to our knowledge, the size variation in V1R repertoire among mammals exceeds that in any other mammalian gene family. This high variation might be in part because V1Rs are involved in both sensory and reproductive functions. Our phylogenetic analysis indicates that the dramatic size difference in the V1R repertoires of placental mammals is due to two molecular evolutionary mechanisms. First, some ancestral gene families that are still present in rodents have been lost in dogs and cows. Second,

Table 2. Size variation of some gene families among mammalian species

Gene family	Function	Size variation, ratio	Smallest size (organism)	Biggest size (organism)	Sources
V1R	Pheromone receptor	23.4	8 (dog)*	187 (mouse)	12, this study
Morpheus	Nuclear pore complex interacting protein	21	1 (Old World monkeys)	21 (chimpanzee)	28
EAR	Antiparasitic RNases	17	1 (New World monkeys)	17 (ricefield mouse)	29, 30
Ly49	Immunity	17	1 (baboon)	17 (rat)	31
OTEX/PEPP2	Reproduction-related	7.5	2 (human)	15 (mouse)	32
Granzyme	Mast cell chymases	7	4 (human)	28 (rat)	27
KIR	Immunity	7	2 (mouse)	14 (human, macaque)	33, 34
OR	Olfactory receptor	3.7	388 (human)	1,430 (rat)	27, 35
Keratin-associated protein	Epithelial cell function	3.3	3 (human)	10 (mouse)	27
Reverse transcriptase	Polymerase	2.6	25 (mouse)	65 (human)	36

Only functional genes are considered. Species with at least one functional gene in the gene family are compared. Gene families with a size ratio >2 are presented.

*Human V1Rs are likely relics of an ongoing pseudogenization process (see ref. 13). Therefore, the dog has the smallest functional repertoire.

species-specific duplication events characteristic of rodent V1R families were less frequent in cows and dogs. Thus, massive gene deaths and births (40, 41) in different lineages explain the observed size variation.

Is it possible that the smaller V1R repertoire in dogs and cows indicates that the VNO is not functional in these organisms? Pseudogenization of vomeronasal genes and loss of VNO function happened in catarrhine primates (i.e., humans, apes, and Old World monkeys), presumably after the acquisition of full trichromatic vision (13, 42, 43). It is possible that stereoscopic vision in both primates and carnivores (44) compensates for reduced pheromone communication. Complete loss of VNO function, however, is unlikely to be responsible for the small V1R repertoires of dogs and cows, because TRP2, the ion channel necessary for VNO pheromone signal transduction (45, 46), is apparently functional in cows (47), and we were able to identify a complete ORF for dog TRP2 from the genome sequence. Furthermore, there have been reports of bovine pheromones that induce estrus, which is likely mediated by the VNO (48, 49). Is it possible that the V1R repertoires have shrunk during the domestications of dog and cow because of either artificial selection or genetic drift? We think it is unlikely because the domestication started no earlier than 15,000 years ago for dogs and 10,000 years ago for cows (50, 51). Even if a small number of functional genes have become nonfunctional during domestication, their relics should remain readily detectable as pseudogenes. We thus believe that the sizes of V1R repertoires in dogs and cows should be very close to those in their wild ancestors. Another possibility is that the small V1R repertoire could be compensated by a large V2R repertoire. This explanation also seems unlikely, because all V2Rs we identified from dogs and cows were pseudogenes (data not shown). Furthermore, all V2R genes identified from the human and goat genomes are pseudogenes (17). In fact, no functional V2Rs have been reported in nonrodent mammals. If the estimated size of $\approx$100 genes in the rodent V2R repertoire (6) is accurate, the phylogenetic surveys suggest an even more dramatic variation in the V2R repertoire among placental mammals. In this respect, it is interesting to note that a recent study found two types of vomeronasal systems in mammals, with rodents and opossums having both $G\alpha_{i2}$- and $G\alpha_O$-expressing vomeronasal sensory neurons and all other species examined (goats, dogs, horses, musk shrews, and marmosets) having only $G\alpha_{i2}$-expressing vomeronasal sensory neurons (ref. 52, but also see ref. 53). Because V1Rs are expressed in $G\alpha_{i2}$-positive neurons and V2Rs are expressed in $G\alpha_O$-positive neurons (16), it is possible that functional V2Rs exist only in rodents and opossums among mammals. Indeed, our preliminary search confirms the presence of V2R ORFs in the opossum genome (P.S. and J.Z., unpublished results).

It should be noted that not all VNO-mediated functions in rodents are VNO-mediated in other mammals (53), and it is possible that during the evolution of rodents some olfactory cues became detectable by the VNO. In fact, a goat V1R gene is known to be expressed in the olfactory epithelium (17). Before we can determine the effects of gene loss in V1R superfamily evolution, we must understand the functional difference between family groups. Two studies have investigated the functions of V1Rs, one in a strain of mice mutant for *V1Rb2* (54) and one in a strain of mice that had 16 genes of the families V1RA and V1RB knocked out (2). Both studies showed that the mutations resulted in physiological or behavioral changes in the mice. But, no study has compared the phenotypic effects of eliminating different V1R families, and it remains unknown whether each family is necessary for a specific function. Such information would allow us to reason why certain families are lost in some species.

In addition to gene loss, our analysis clearly demonstrated V1R family expansions in some placental mammals. However, it is unknown what factors have promoted the dramatic expansions in rodents but have only allowed limited expansion in cows and virtually no expansion in dogs. Lane and colleagues (22, 55) suggested that the V1R gene duplications in rodents were mediated by L1 repetitive elements. These elements, which are known to have been active in rodents around the time of the mouse-rat divergence, densely populate the regions of the mouse and rat genomes where V1Rs are located. L1 density differs between species, so it would be interesting to see how V1R repertoire size correlates with L1 density. After the procedure of Lane *et al.* (22), we found that 21% of the DNA sequences in the genomic regions harboring dog V1R genes are L1 elements, lower than the corresponding density (40%) in the mouse V1R loci (22) or the average density (25%) in the mouse genome (56). The L1 density of the genomic regions containing cow V1R genes is also 21%. Thus, the low duplicability of both dog and cow V1R genes might be in part due to the low density of L1 elements in the genomic regions. L1 density in cow V1R genomic regions is similar to that in dogs, but cows have four times as many V1Rs as dogs have. Thus, L1 elements might not play as great a role in V1R duplications as originally thought, or the role of L1 elements in V1R duplication might be limited to rodents.

The phylogenetic reconstruction of placental V1Rs with the opossum V1Rs (Fig. 3) suggests that there was more than one V1R gene in the common ancestor of marsupials and placentals, and the mammalian V1R families are then at least 170 MY old (23). Many V1R families expanded in placentals and marsupials independently. Sequencing another marsupial genome will significantly broaden our understanding of V1R evolution in marsupials. Because both primates and rodents have >100 V1R genes (or pseudogenes in the case of humans) and because primates and rodents are more closely related to each other than either of them is to carnivores or artiodactyls (23, 25), one might infer that after the separation of the common ancestor of cows and dogs from the common ancestor of rodents and primates, there was a dramatic expansion of the V1R repertoires. Thus, the large size of the V1R superfamily as observed in rodents might be restricted to organisms derived from the common ancestor of primates and rodents, including the five orders of Rodentia, Lagomorpha (e.g., rabbit), Dermoptera (e.g., flying lemur), Scandentia (e.g., tree shrew), and Primates (25). However, our previous molecular dating indicates that rodent V1R families expanded after the primate-rodent split (8), suggesting independent expansions in rodents and primates. [In the future, the independent expansions in primates and rodents could be tested by examining the phylogenetic positions of human pseudogenes when their sequences (ref. 9) become available.] This independence would further imply that some of the aforementioned mammalian orders might not contain expanded V1R families. Even if expansion is characteristic of these five orders, it is also possible that a functional V1R repertoire subsequently shrank after expansion, as in catarrhine primates (13). These considerations suggest that mouse and rat may be atypical mammals in terms of their pheromone receptor genes and pheromone sensitivities. Of course, independent expansions would also imply great differences in V1R receptors and pheromone sensitivities. Thus, one should be cautious in applying to other mammals the V1R and pheromone-related knowledge learned from the model organisms of mouse and rat. Furthermore, although a large variation in the V1R repertoire is described here, the species we have examined (human, mouse, rat, dog, cow, and opossum) represent only five of $\approx$24 orders of placental and marsupial mammals. A more thorough investigation of theV1R repertoires in other orders will give a better picture of the variation and evolution of V1Rs in mammals.

Interestingly, the small repertoire of vomeronasal pheromone receptor genes that we report in the dog may not be entirely unexpected and may be common to carnivores in general.

Although rodents have a complex VNO with a thick layer of vomeronasal sensory epithelium (VNSE), where vomeronasal receptors are found, carnivores have a different type of VNO with a much thinner VNSE (14, 52). Furthermore, the VNO of the ferret *Mustela putorius*, another carnivore, is rudimentary in size and development and does not change with respect to season, which is in contrast to the seasonal variation observed in other ferret organs involved in sexual reproduction and behavior and in rodent VNOs (57). Additionally, Woodley and Baum (58) found that the main olfactory system, but not the VNO, is necessary for mate identification in the ferret. The findings in dogs and ferrets suggest a limited role of the VNO, which may predict a small V1R repertoire in all carnivores. Among artiodactyls, cow is the only species with a fully sequenced genome. However, a few studies have focused on the VNO of the goat, which belongs to the same family as the cow (17, 52), and a genomic Southern analysis suggested that the goat V1R repertoire is also significantly smaller than that in rodents (17). However, even between mouse and rat, there is a nearly 2-fold difference in the number of functional V1R genes, indicating that ordinal generalizations of V1R repertoire sizes from one species should be drawn cautiously.

As mentioned, the complexity of VNO morphology varies among vertebrates. Takami (14) classified VNOs into five morphological types from the most complex (A) to the simplest (E). Type A is found in ophidian species such as crotaline, garter, and water snakes. Type B is found in rodents and lagomorphs; although marsupials such as opossum also have VNOs similar to type B. Type C is found in ungulates (e.g., horses, cows, and sheep), carnivores (e.g., dogs, cats, and ferrets), prosimian primates, and New World monkeys. Type D is found in amphibians, and type E is found in human fetal VNO. Thus, there is a correlation between the morphological complexity of the VNO and the number of intact V1R genes (Fig. 4), providing genomic support for the notion that the morphological complexity of the VNO may be used as a proxy for the sophistication of pheromone communications within species, as has been assumed by many anatomists (15, 17).

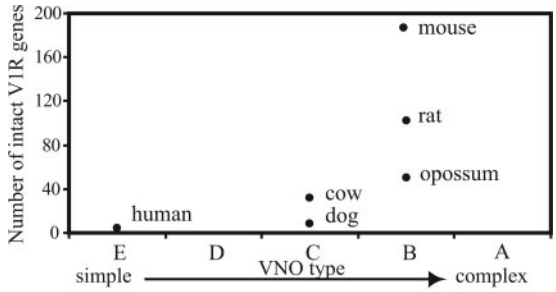

Fig. 4. Correlation between the morphological complexity of VNO and the number of intact V1R genes in the genome. Morphological complexity follows ref. 14.

Note Added in Proof. Recently, Young *et al.* (59) reported the same eight dog V1R genes.

We thank Soochin Cho, Masatoshi Nei, Ondrej Podlaha, David Webb, and three anonymous reviewers for valuable comments. This work was supported in part by National Institutes of Health Grant GM67030 (to J.Z.) and by grants from the National Natural Science Foundation of China (30021004, 30430110), the Natural Science Foundation of Yunnan, and the Chinese Academy of Sciences to Y.-p.Z.

1. Keverne, E. B. (1999) *Science* **286,** 716–720.
2. Del Punta, K., Leinders-Zufall, T., Rodriguez, I., Jukam, D., Wysocki, C. J., Ogawa, S., Zufall, F. & Mombaerts, P. (2002) *Nature* **419,** 70–74.
3. Dulac, C. & Torello, A. T. (2003) *Nat. Rev. Neurosci.* **4,** 551–562.
4. Dulac, C. & Axel, R. (1995) *Cell* **83,** 195–206.
5. Ryba, N. J. & Tirindelli, R. (1997) *Neuron* **19,** 371–379.
6. Herrada, G. & Dulac, C. (1997) *Cell* **90,** 763–773.
7. Matsunami, H. & Buck, L. B. (1997) *Cell* **90,** 775–784.
8. Grus, W. E. & Zhang, J. (2004) *Gene* **340,** 303–312.
9. Rodriguez, I. & Mombaerts, P. (2002) *Curr. Biol.* **12,** R409–R411.
10. Rodriguez, I., Del Punta, K., Rothman, A., Ishii, T. & Mombaerts, P. (2002) *Nat. Neurosci.* **5,** 134–140.
11. Zhang, X., Rodriguez, I., Mombaerts, P. & Firestein, S. (2004) *Genomics* **83,** 802–811.
12. Shi, P., Bielawski, J., Yang, H. & Zhang, Y. P. (2005) *J. Mol. Evol.*, in press.
13. Zhang, J. & Webb, D. M. (2003) *Proc. Natl. Acad. Sci. USA* **100,** 8337–8341.
14. Takami, S. (2002) *Microsc. Res. Tech.* **58,** 228–250.
15. Giorgi, D. & Rouquier, S. (2002) *Chem. Senses* **27,** 529–537.
16. Mombaerts, P. (2004) *Nat. Rev. Neurosci.* **5,** 263–278.
17. Wakabayashi, Y., Mori, Y., Ichikawa, M., Yazaki, K. & Hagino-Yamagishi, K. (2002) *Chem. Senses* **27,** 207–213.
18. Thompson, J. D., Gibson, T. J., Plewniak, F., Jeanmougin, F. & Higgins, D. G. (1997) *Nucleic Acids Res.* **25,** 4876–4882.
19. Saitou, N. & Nei, M. (1987) *Mol. Biol. Evol.* **4,** 406–425.
20. Nei, M. & Kumar, S. (2000) *Molecular Evolution and Phylogenetics* (Oxford Univ. Press, New York).
21. Felsenstein, J. (1985) *Evolution* **39,** 783–791.
22. Lane, R. P., Cutforth, T., Axel, R., Hood, L. & Trask, B. J. (2002) *Proc. Natl. Acad. Sci. USA* **99,** 291–296.
23. Murphy, W. J., Pevzner, P. A. & O'Brien, S. J. (2004) *Trends Genet.* **20,** 631–639.
24. Mundy, N. I. & Cook, S. (2003) *Mol. Biol. Evol.* **20,** 1805–1810.
25. Springer, M. S., Murphy, W. J., Eizirik, E. & O'Brien, S. J. (2003) *Proc. Natl. Acad. Sci. USA* **100,** 1056–1061.
26. Li, W.-H. (1997) *Molecular Evolution* (Sinauer, Sunderland, MA).
27. Rat Genome Sequencing Consortium (2004) *Nature* **428,** 493–521.
28. Johnson, M. E., Viggiano, L., Bailey, J. A., Abdul-Rauf, M., Goodwin, G., Rocchi, M. & Eichler, E. E. (2001) *Nature* **413,** 514–519.
29. Zhang, J., Dyer, K. D. & Rosenberg, H. F. (2000) *Proc. Natl. Acad. Sci. USA* **97,** 4701–4706.
30. Cho, S., Beintema, J. J. & Zhang, J. (2005) *Genomics* **85,** 208–220.
31. Hao, L. & Nei, M. (2004) *Immunogenetics* **56,** 343–354.
32. Wang, X. & Zhang, J. (2004) *Genetics* **167,** 879–888.
33. Hughes, A. L. (2002) *Mol. Phylogenet. Evol.* **25,** 330–340.
34. Hao, L. & Nei, M. (February 17, 2005) *Gene*, 10.1016/j.gene.2004.12.012.
35. Niimura, Y. & Nei, M. (2003) *Proc. Natl. Acad. Sci. USA* **100,** 12235–12240.
36. Mouse Genome Sequencing Consortium. (2002) *Nature* **420,** 520–562.
37. Rosenberg, H. F., Dyer, K. D., Tiffany, H. L. & Gonzalez, M. (1995) *Nat. Genet.* **10,** 219–223.
38. Lefranc, M. P. (2001) *Exp. Clin. Immunogenet.* **18,** 100–116.
39. Su, C. & Nei, M. (1999) *Proc. Natl. Acad. Sci. USA* **96,** 9710–9715.
40. Nei, M. & Hughes, A. L. (1992) in *Proceedings of the 11th Histocompatibility Workshop and Conference*, eds. Tsuji, K., Aizawa, M. & Sasazuki, T. (Oxford Univ. Press, Oxford), pp. 27–38.
41. Hughes, A. L. & Nei, M. (1989) *Mol. Biol. Evol.* **6,** 559–579.
42. Liman, E. R. & Innan, H. (2003) *Proc. Natl. Acad. Sci. USA* **100,** 3328–3332.
43. Webb, D. M., Cortes-Ortiz, L. & Zhang, J. (2004) *Mol. Biol. Evol.* **21,** 697–704.
44. Kral, K. (2003) *Behav. Processes* **64,** 1–12.
45. Stowers, L., Holy, T. E., Meister, M., Dulac, C. & Koentges, G. (2002) *Science* **295,** 1493–1500.
46. Leypold, B. G., Yu, C. R., Leinders-Zufall, T., Kim, M. M., Zufall, F. & Axel, R. (2002) *Proc. Natl. Acad. Sci. USA* **99,** 6376–6381.
47. Wissenbach, U., Schroth, G., Philipp, S. & Flockerzi, V. (1998) *FEBS Lett.* **429,** 61–66.
48. Rekwot, P. I., Ogwu, D., Oyedipe, E. O. & Sekoni, V. O. (2001) *Anim. Reprod. Sci.* **65,** 157–170.
49. Lane, A. J. & Wathes, D. C. (1998) *J. Dairy Sci.* **81,** 2145–2150.
50. Savolainen, P., Zhang, Y. P., Luo, J., Lundeberg, J. & Leitner, T. (2002) *Science* **298,** 1610–1613.
51. Bradley, D. G., MacHugh, D. E., Cunningham, P. & Loftus, R. T. (1996) *Proc. Natl. Acad. Sci. USA* **93,** 5131–5135.
52. Takigami, S., Mori, Y., Tanioka, Y. & Ichikawa, M. (2004) *Chem. Senses* **29,** 301–310.
53. Dennis, J. C., Allgier, J. G., Desouza, L. S., Eward, W. C. & Morrison, E. E. (2003) *J. Anat.* **203,** 329–338.
54. Boschat, C., Pelofi, C., Randin, O., Roppolo, D., Luscher, C., Broillet, M. C. & Rodriguez, I. (2002) *Nat. Neurosci.* **5,** 1261–1262.
55. Lane, R. P., Young, J., Newman, T. & Trask, B. J. (2004) *Genome Res.* **14,** 603–608.
56. Smit, A. F. (1999) *Curr. Opin. Genet. Dev.* **9,** 657–663.
57. Weiler, E., Apfelbach, R. & Farbman, A. I. (1999) *Chem. Senses* **24,** 127–136.
58. Woodley, S. K. & Baum, M. J. (2004) *Eur. J. Neurosci.* **20,** 1025–1032.
59. Young, J. M., Kambere, M., Trask, B. J. & Lane, R. P. (2005) *Genome Res.* **15,** 231–240.