Available online at www.sciencedirect.com

SCIENCE DIRECT®

Gene 340 (2004) 303–312

GENE
AN INTERNATIONAL JOURNAL ON GENES AND GENOMES

www.elsevier.com/locate/gene

Rapid turnover and species-specificity of vomeronasal pheromone receptor genes in mice and rats☆

Wendy E. Grus, Jianzhi Zhang*

Department of Ecology and Evolutionary Biology, University of Michigan, 3003 Natural Science Building, 830 North University Avenue, Ann Arbor, MI 48109, USA

Received 5 May 2004; received in revised form 11 June 2004; accepted 29 July 2004
Available online 11 September 2004
Received by S. Yokoyama

Abstract

Pheromones are used by individuals of the same species to elicit behavioral or physiological changes, and they are perceived primarily by the vomeronasal organ (VNO) in terrestrial vertebrates. VNO pheromone receptors are encoded by the *V1r* and *V2r* gene superfamilies in mammals. A comparison of the *V1r* and *V2r* repertoires between closely related species can provide significant insights into the evolutionary genetic mechanisms responsible for species-specific pheromone communications. A total of 137 putatively functional *V1r* genes of 12 families were previously identified from the mouse genome. We report the identification of 95 putatively functional *V1r* genes from the draft rat genome sequence. These genes map primarily to four blocks in two chromosomes. The rat *V1r* genes can be phylogenetically grouped into 10 families, which are shared with mouse, and 2 new families, which are rat-specific. Even in many shared families, gene numbers differ between the two species, apparently due to frequent gene duplication and pseudogenization after the separation of the two species. Molecular dating suggests that most of the rat V1r families emerged before or during the radiation of mammalian orders, but many duplications within families occurred as recently as in the past 10 million years (MY). Our results show that the evolution of the *V1r* repertoire is characterized by exceptionally fast gene turnover via gains and losses of individual genes, suggesting rapid and substantial changes in pheromone communication between species.

Keywords: *V1r*; Gene duplication; Rodents; Gene sorting; Pheromone

1. Introduction

Pheromones are used by individuals of the same species to elicit behavioral or physiological changes, such as male–male aggression, puberty, estrus, and induction of mating, and are perceived primarily by the vomeronasal organ (VNO; reviewed in Keverne, 1999; Dulac and Torello, 2003). Two superfamilies of vomeronasal pheromone receptors, V1r and V2r, are known in mammals and they differ in expression location and gene structure (Dulac and Axel, 1995; Herrada and Dulac, 1997; Matsunami and Buck, 1997; Ryba and Tirindelli, 1997; reviewed in Dulac and Torello, 2003). While both types of receptors are seven-transmembrane (TM) G-protein coupled receptors, V1rs are characterized by an intronless coding region, while V2rs are characterized by a long highly variable N-terminal domain. V1rs are expressed in Gαi2 neurons and V2rs are expressed in Gα0 neurons (reviewed in Dulac and Torello, 2003). Targeted deletion of some *V1r* genes in mice show altered aggression and sexual behaviors (Del Punta et al., 2002). Additionally, a third vomeronasal receptor superfamily, V3r, has been described (Pantages and Dulac, 2000); however, V3rs were later found to be a family of V1rs (Rodriguez et al., 2002). Because they lack introns, *V1r* genes are more accessible

Abbreviations: MY, million years; VNO, vomeronasal organ; ORF, open reading frame; TM, transmembrane; EC, extracellular; IC, intracellular.

☆ Genbank accession numbers for the rat *V1r* genes reported in this paper: AY649001–AY649095.

* Corresponding author. Tel.: +1 734 763 0527; fax: +1 734 763 0544.
E-mail address: jianzhi@umich.edu (J. Zhang).

0378-1119/$ - see front matter © 2004 Elsevier B.V. All rights reserved.
doi:10.1016/j.gene.2004.07.037

than *V2r* genes to bioinformatic as well as experimental studies. For example, the entire mouse *V1r* repertoire, including 137 putatively functional members of 12 families has been described (Rodriguez et al., 2002), while the mouse *V2r* repertoire has yet to be described.

The identification of *V1r* and *V2r* genes has opened the door to studies on the molecular mechanisms and origins of species-specific pheromone communications. V1rs were originally discovered and described in rat (Dulac and Axel, 1995). The absence of highly conserved regions in V1r prohibits the design of degenerate primers that can amplify a large number of genes across wide taxonomic scale (Giorgi and Rouquier, 2002). Therefore, the V1r superfamily was not extensively described in any species until the availability of the human and mouse genome sequences (Lane et al., 2002; Rodriguez et al., 2002; Rodriguez and Mombaerts, 2002). The comparison of V1rs between human and mouse is not informative because the two species are

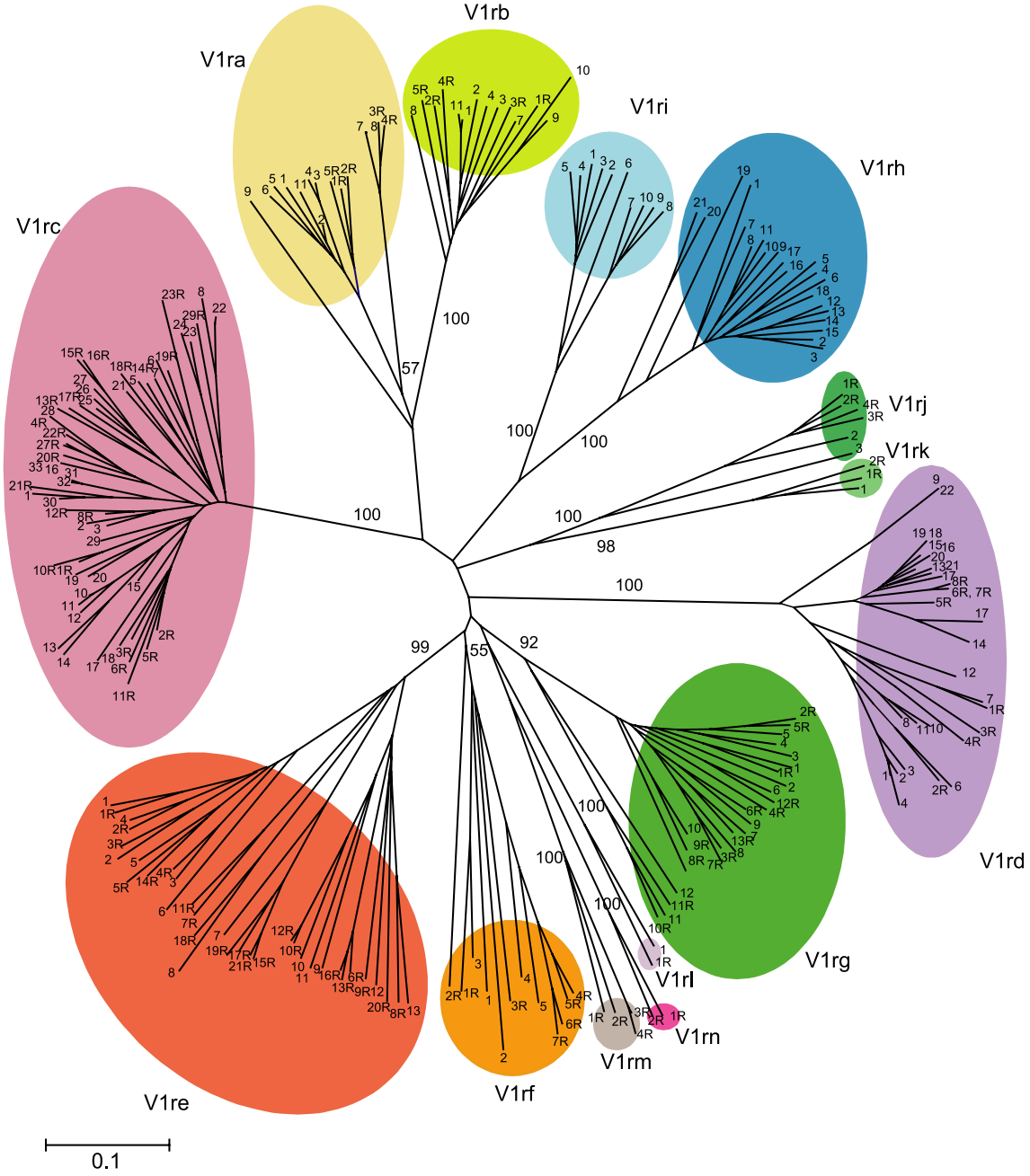

Fig. 1. Unrooted phylogenetic tree of 95 putatively functional *V1r* genes in rat and 137 in mouse. Rat genes are identified with the suffix R. Families V1ra–V1rl were previously defined by Rodriguez et al. (2002) and families V1rm and V1rn are newly identified here. Bootstrap percentages supporting the monophyly of each family are given, except for V1ra, which is not monophyletic due to one sequence (V1ra9). The tree was reconstructed using the neighbor-joining method with Poisson-corrected protein distances. The scale bar shows 0.1 amino acid substitution per site. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article).

distantly related and because humans lack VNO sensitivity due to loss of important components of the VNO pheromone transduction pathway (Zhang and Webb, 2003; Liman and Innan, 2003). Zhang et al. (2004) compared the *V1r* repertoires of two genome assemblies generated from different strains of inbred mice. While this comparison is useful for identifying polymorphisms within species, it does not address interspecific differences which are the hallmark of pheromone communication. The availability of the draft rat genome sequence (Rat Genome Sequencing Consortium, 2004) provides the opportunity to compare for the first time pheromone receptor repertoires of relatively closely related species, as mouse and rat diverged only 18±6 million years (MY) ago (Rat Genome Sequencing Consortium, 2004).

Divergence in the *V1r* repertoires between species can occur in three ways: functional divergence of orthologs, loss or gain of family members, and loss or gain of entire families. Two recent studies compared a small number of *V1r* genes between mouse and rat (Lane et al., 2004; Emes et al., 2004) but did not provide a general picture on the evolution of the V1r repertoire. We compare the entire *V1r* repertoires in an attempt to examine which of the three processes dominate the divergence of *V1r* genes between species.

2. Materials and methods

2.1. Database searches

TBLASTN searches for rat *V1r* genes were done on the rat genome sequence available in the National Center for Biotechnology Information (NCBI) with the rat genome build 2 version 1 (http://www.ncbi.nlm.nih.gov/genome/guide/rat/index.html). The previously described 137 mouse *V1r* genes were used as query sequences. *V1r* pseudogenes were identified by premature stop codons or incomplete sequence across the 13 protein domains (7 transmembrane, 3 extracellular (EC), and 3 intracellular (IC)) of the 15 domains. The N-terminal extracellular and C-terminal intracellular domains were not considered in our criterion because they are highly variable in sequence length. The physical locations of the rat genes were determined by mapping the TBLASTN results to chromosomal contigs. The rat genes were named by their family memberships (e.g., V1re5) and the numbers after the family designation are randomly chosen.

2.2. Sequence alignment and phylogenetic analysis

Gene sequences were aligned according to the protein sequence alignment made by Clustal_X (Thompson et al., 1997), with manual adjustment. Phylogenetic trees were constructed using the neighbor-joining method (Saitou and Nei, 1987) with 1000 bootstrap replications (Felsenstein, 1985). Synonymous (d_S) and nonsynonymous (d_N) substitution rates were computed by the modified Nei–Gojobori method (Zhang et al., 1998). MEGA2 (Kumar et al., 2001) was used for these evolutionary analyses.

2.3. Identifying putative rat–mouse orthologs

To find orthologous *V1r* genes, we identified clades in the phylogenetic tree that contain a single rat gene and a single mouse gene and then calculated the d_S between the putative orthologs. Those with estimated d_S within the range of 0.063–0.317 were accepted as orthologs. This range covers 95% of over 11,000 orthologous mouse–rat gene pairs (Rat Genome Sequencing Consortium, 2004). In the present study, 19 putative orthologous pairs were identified via the phylogenetic analysis and 18 of them passed the above d_S criterion. Our procedure is conservative in the sense that it is unlikely to generate false orthologous pairs but may miss some real pairs.

2.4. Dating duplication events

To determine the dates of duplication events within the rat V1r families, we calculated the pairwise d_S for paralogous genes within families. The average synonymous substitution rate in rodents is about $0.19/36=5.3\times10^{-9}$ per site per year, estimated using the mean d_S of 0.19 and the divergence time of 18±6 million years between the mouse and rat (Rat Genome Sequencing Consortium, 2004).

To determine the time of duplication between gene families, we calculated the average d_S between families and used the above calibration. We assumed that the root of the tree in Fig. 1 is on the stem branch leading to V1rc, as suggested by Rodriguez et al. (2002) when taste receptor genes were used to root the mouse V1r tree.

3. Results

3.1. Composition of the rat V1r gene superfamily

From the database searches with mouse V1r queries, we identified from the rat genome sequence 116 *V1r* genes. Additional searches with human *V1r*-like genes (Rodriguez and Mombaerts, 2002) did not yield any additional sequences. Because the rat genome is over 90% complete (Rat Genome Sequencing Consortium, 2004), the *V1r* repertoire we describe here is probably over 90% complete as well. We examined the conceptually translated protein sequences encoded by the retrieved DNA sequences and found that 95 of them contain open reading frames (ORFs) that cover 13 internal domains. These 95 ORFs are regarded as putatively functional genes. It is possible that some of them will in fact be nonfunctional due to mutations in either protein-coding or regulatory regions. The proportion of putatively functional *V1r* genes appears higher in rats (95/116=82%) than in mice (53%; Rodriguez et al., 2002; Zhang

et al., 2004), although this may be caused by the bias of TBLASTN against pseudogene detection.

A neighbor-joining tree of all putatively functional *V1r* genes from the mouse and rat was reconstructed using protein sequences (Fig. 1). The 137 mouse *V1r* genes were previously classified into 12 families based on phylogenetic analysis as well as the criterion that genes with >40% protein sequence identity belong to the same family (Rodriguez et al., 2002). These 12 families were recovered in our tree of mouse and rat genes, and most (10/12) of them have high bootstrap support (>90%). In 10 (a, b, c, d, e, f, g, j, k, and l) of the 12 families, rat genes are also found, and these are the shared V1r families between the two species. Two mouse families (h and i) do not contain any rat genes and are mouse-specific families (Table 1). In addition, two new families (m and n; defined by the same criterion as used in mouse) contain only rat genes and are referred to as rat-specific families. Because the rat genome sequence was obtained from an inbred line (Rat Genome Sequencing Consortium, 2004) and the *V1r* genes were identified by their chromosomal location, it is unlikely that genes with high sequence identity are only allelic variants and not separate genes. In fact, two of the rat *V1rd* genes (*V1rd6* and *V1rd7*) are identical in DNA sequence, but are located 1 Mb apart on chromosome 1. Our evolutionary analysis provides further support that *V3r* (*V1rd*) genes (Pantages and Dulac, 2000) are not a separate vomeronasal receptor superfamily (Rodriguez et al., 2002) but are a family of *V1r* genes (Fig. 1). In fact, only eight *V1rd* genes are present in the rat genome.

3.2. Chromosomal organization of V1r genes

The 95 putatively functional rat *V1r* genes map to 8 locations in chromosomes 1, 4, and 7, although most of them are located in four major blocks in chromosomes 1 and 4 (Fig. 2). It is interesting to note that members of the same family tend to be located in the same chromosomal region, suggesting tandem gene duplication as the primary mechanism for family expansion.

Using the mouse–rat homologous chromosome map (http://www.genboree.org), we found syntenic regions between the four major blocks of *V1r* genes in rat and four blocks in mouse (Fig. 2). Both blocks on rat chromosome 4 map to mouse chromosome 6. The two blocks on rat chromosome 1 are split mainly between mouse chromosomes 7 and 17. Twenty-eight putatively functional mouse *V1r* genes identified by Rodriguez et al. (2002) have yet to be mapped.

From the phylogeny shown in Fig. 1 and estimates of synonymous distances, we identified 18 orthologous *V1r* gene pairs between mouse and rat. These pairs are shown on the chromosomes as well (Fig. 2). One *V1rg* pair (rat *V1rg9* and mouse *V1rg10*) has a mouse gene in chromosome 5, not part of the large block in *V1rg* family in mouse (Fig. 2). However, a comparative study of a different mouse strain did not identify any *V1r* genes on mouse chromosome 5 (Zhang et al., 2004), indicating that this is a recent translocation within the species *Mus musculus*.

3.3. Comparison with mouse V1r genes

The most notable difference between the *V1r* repertories of the mouse and rat is the presence of two mouse-specific and two rat-specific families. The mouse-specific V1rh and V1ri families contain 23% of all mouse *V1r* genes. We determined that the absence of rat *V1rh* and *V1ri* genes is not due to lack of sequencing in this region, as we can find rat orthologs of non-*V1r* genes that are located in the mouse V1rh and V1ri chromosomal regions and a *V1ri* pseudogene was found in this region in rat. The rat-specific V1rm and V1rn families contain 6% of all rat *V1r* genes. No functional genes or pseudogenes that belong to these two families were found in mouse.

Fig. 1 shows that in addition to the species-specific gene families, many shared families have different numbers of genes in the two species. V1rc is the richest family in both species. However, V1re contains 21 rat genes but only 13 mouse genes. On the other hand, V1rd has 8 rat but 22 mouse genes. These differences show that lineage-specific contraction and/or expansion of certain V1r families must have occurred after the mouse–rat separation. In some shared V1r families (e.g., V1ra), not a single one-to-one ortholgous pair is found between the two species. This implies that every gene in this family has been subject to either duplication or pseudogenization since the mouse–rat split.

3.4. d_N/d_S ratio for orthologous genes and paralogous genes

To examine whether positive selection has been operating during the divergence of orthologous *V1r* genes between species, we computed the number of synonymous substitu-

Table 1
V1r families in rat and mouse

Family	No. of functional genes		Percentage identity (%, amino acid)	
	Rat	Mouse	Rat	Mouse
A	5	10	57–92	54–99
B	5	9	72–84	70–98
C	23	31	74–96	60–99
D	8	22	54–100	61–93
E	21	13	43–97	44–77
F	7	5	43–93	45–75
G	13	12	48–91	53–88
H	0	21	–	61–93
I	0	10	–	65–91
J	4	2	84–90	58
K	2	1	84	–
L	1	1	–	–
M	4	0	70–97	–
N	2	0	81	–
Total	95	137		

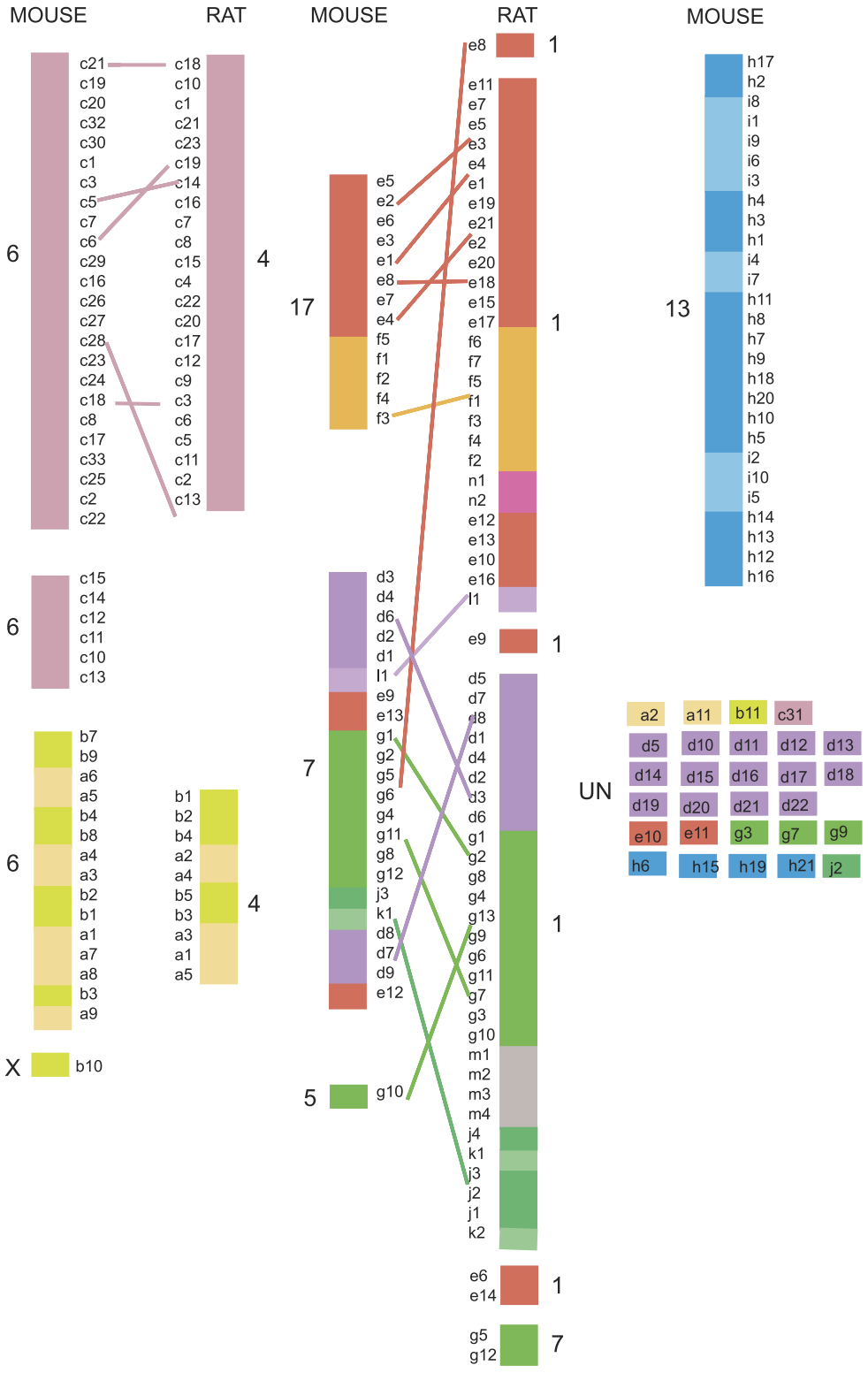

Fig. 2. Chromosomal locations of putatively functional rat and mouse *V1r* genes. Lines between mouse and rat genes indicate orthologous relationships. Unmapped mouse genes are shown as "UN". Colors designate family membership as in Fig. 1. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article).

tion per synonymous site (d_S) and the number of non-synonymous substitution per nonsynonymous site (d_N) for the 18 pairs of orthologous genes between mouse and rat. A d_N/d_S ratio that significantly exceeds 1 provides the most convincing evidence for positive selection, whereas $d_N/d_S<1$ shows overall purifying selection on the gene, although

positive selection at a small number of sites cannot be excluded.

The d_S values for the 18 *V1r* gene pairs range from 0.14 to 0.26 with a mean of 0.19, identical to the mean d_S from over 11,000 orthologous gene pairs between the mouse and rat (Rat Genome Sequencing Consortium, 2004). All 18 d_S values fall within the middle 67% of the d_S distribution determined from this large set of orthologous genes (Rat Genome Sequencing Consortium, 2004), suggesting that our determination of orthology was correct.

The overall d_N/d_S ratio across the entire gene sequence is less than 1 for each of the 18 pairs of orthologous genes (Fig. 3A), with an average d_N/d_S of 0.46. This value is in the 2% upper tail of the distribution of d_N/d_S derived from the over 11,000 gene comparisons aforementioned (mean=0.11; Rat Genome Sequencing Consortium, 2004). As mean d_S for the 11,000 genes and the mean d_S for the *V1r* genes are virtually identical, the difference in the ratio comes from a much higher d_N in the *V1r* genes than in most other genes. This may be due to relatively weak purifying selection on *V1r* genes or the presence of positive selection at some sites. However, even when we examined d_N/d_S for separate domains, all show $d_N/d_S<1$ (Fig. 3B).

We also compared members of the same V1r families in rats to test whether positive selection may be responsible for divergence of V1r duplicates within families. The overall d_N/d_S ratio averages 0.549. The d_N/d_S ratios from the rat-specific families (Fig. 4A) and the mouse-specific families (Fig. 4B) are similar to the d_N/d_S ratios from the families shared between both species (Fig. 4C, D). We did not attempt to estimate d_N/d_S ratios between genes belonging to different V1r families because of the high sequence divergence and expected low reliability of estimation caused by multiple hits.

3.5. Dates of V1r gene duplications

It is important to know the approximate dates of gene duplication events that gave rise to the *V1r* genes. Because d_S increases over time relatively constantly, we estimated d_S values between paralogous *V1r* genes of rats. A calibration of d_S=0.19 corresponding to a gene age of 18 MY was used and dates of the duplication events within and between rat

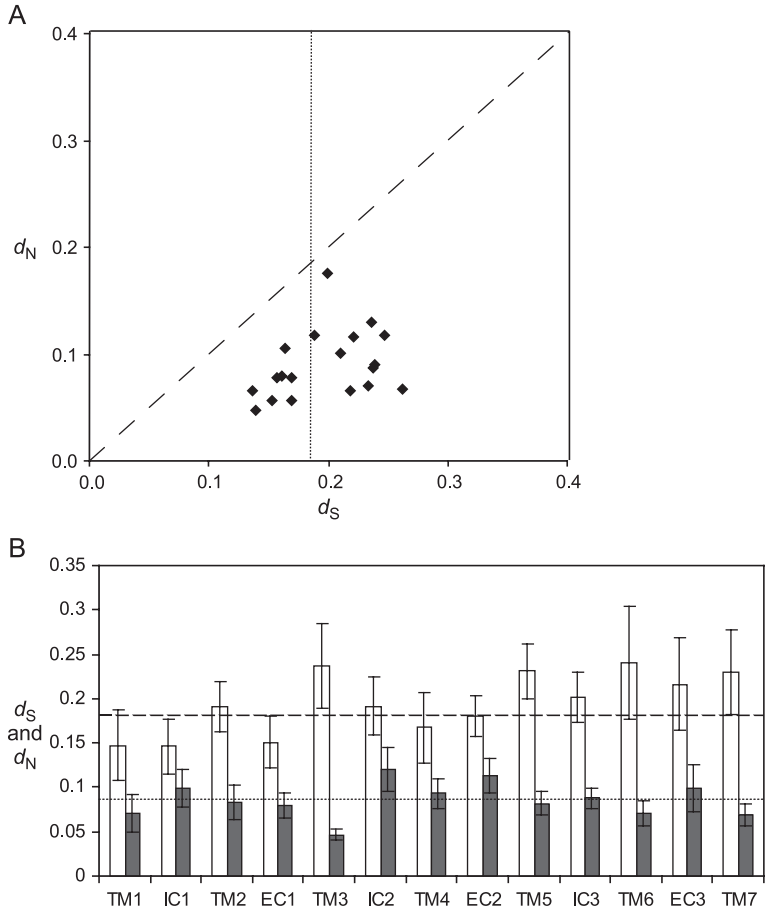

Fig. 3. Synonymous (d_S) and nonsynonymous (d_N) distances between 18 orthologous *V1r* gene pairs of mouse and rat. (A) Overall d_S and d_N across the entire sequences. The dotted vertical line indicates d_S=0.19, which is the mean d_S for over 11,000 mouse–rat orthologous gene pairs analyzed by Rat Genome Sequencing Consortium (2004). The diagonal line shows $d_N=d_S$. (B) The average d_N (black bars) and d_S (white bars) among the 18 orthologous pairs for each of the 13 domains. TM, transmembrane domain; EC, extracellular domain; IC, intracellular domain. The N-terminal extracellular domain and C-terminal intracellular domain are not studied due to high variation in sequence length. Error bars show one standard error of the mean. The dashed line is the average d_S for the entire gene, whereas the dotted line is the average d_N for the entire gene.

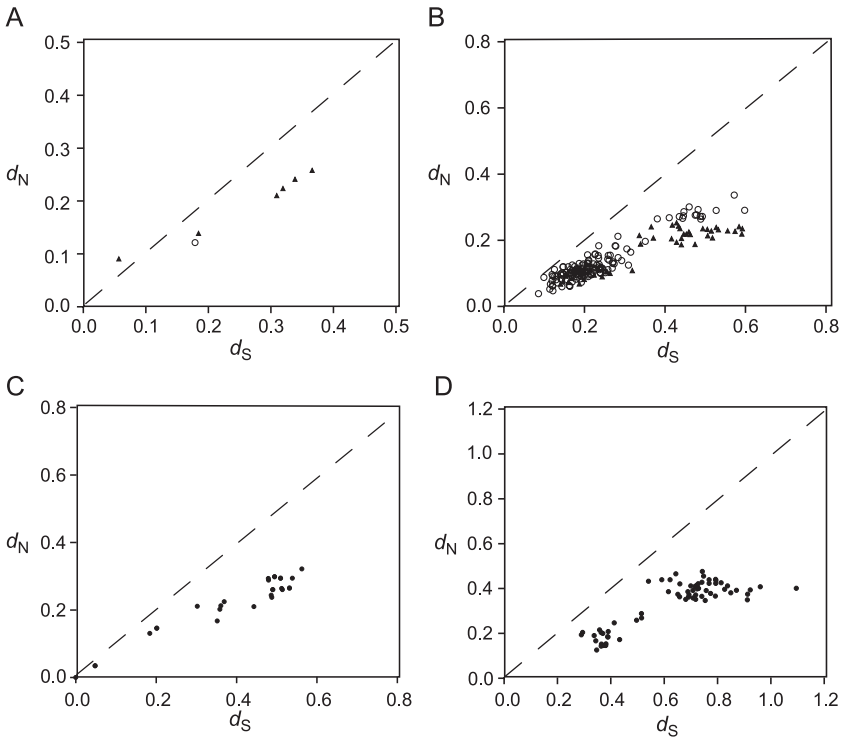

Fig. 4. Synonymous (d_S) and nonsynonymous (d_N) distances between paralogous *V1r* genes within families. The diagonal lines show d_N=d_S. (A) d_S and d_N for rat–specific families. (▲) and (○) represent V1rm and V1rn, respectively. (B) d_S and d_N for mouse-specific families. (▲) and (○) represent V1rh and V1ri, respectively. (C) d_S and d_N between paralogous rat genes in V1re, a typical *V1r* family shared between mouse and rat. (D) d_S and d_N between paralogous mouse genes in V1re.

V1r families were obtained (Fig. 5A, C). Although many duplication events within gene families occurred after the mouse–rat split, about half of them also took place before this split. Additionally, one-fourth of all the duplications that occurred within families are estimated to have taken place in the last 10 MY, including very recent duplications within V1rd, V1re, V1rf, and V1rm. It appears that the number of duplications that led to stably retained genes increased in the recent past, as indicated by a negative correlation (r^2=0.75, P<0.003) between the number of duplications and evolutionary age (Fig. 5B). However, this could be due to the short lifespan for *V1r* genes, such that ancient duplicates are more likely to have been lost than young ones, rendering underestimation of duplication numbers in the ancient past.

Molecular dating suggests that all of the rat V1r families were present at the time of split between rat and mouse (Fig. 5C). Even the youngest rat V1r family, the one descended from the common ancestor of V1ra and V1rb, emerged ~62 MY ago. These estimates are conservative because large d_S values are likely to be underestimated due to the difficulty of correcting multiple hits at synonymous sites.

4. Discussion

In this study, we identified 95 putatively functional *V1r* genes from the rat genome sequence, supporting the prediction of ~100 distinct genes in the original description of the rat *V1r* gene superfamily (Dulac and Axel, 1995). This number is about two-thirds of that in the mouse. Wild mice and rats have different social/reproductive structures (Abbott, 2004). Mice live in groups with one highly aggressive alpha male monopolizing the females, whereas rats are promiscuous and less aggressive. It is possible that the observed difference in the number of their *V1r* genes has biological significance.

Our study localized the majority of the rat *V1r* genes to two blocks on chromosome 4 and two blocks on chromosome 1, which show synteny to four chromosomal regions in mouse. While the families making up each region are the same, there are few orthologous gene pairs within them. The Rat Genome Sequencing Project Consortium (2004) estimated that 86–94% of rat genes have one-to-one mouse orthologs. However, only 19% of the rat *V1r* genes have one-to-one mouse orthologs, indicative of rapid gene turnover. Indeed, many V1r families differ in size between the two species and each species has two families that are absent in the other species. All these observations show frequent gains and losses of *V1r* genes during evolution. Such a pattern of gene family evolution has often been observed in host–defense genes such as the Major-Histocompatibility-Complex, immunoglobulin, and eosinophil-associated RNase genes (Cadavid et al., 1997; Sitnikova and Su, 1998; Zhang et al., 2000). While the rapid turnover of host–defense genes is presumably a response to ever-changing pathogens that infect hosts, that of *V1r* genes is likely due to rapid change in

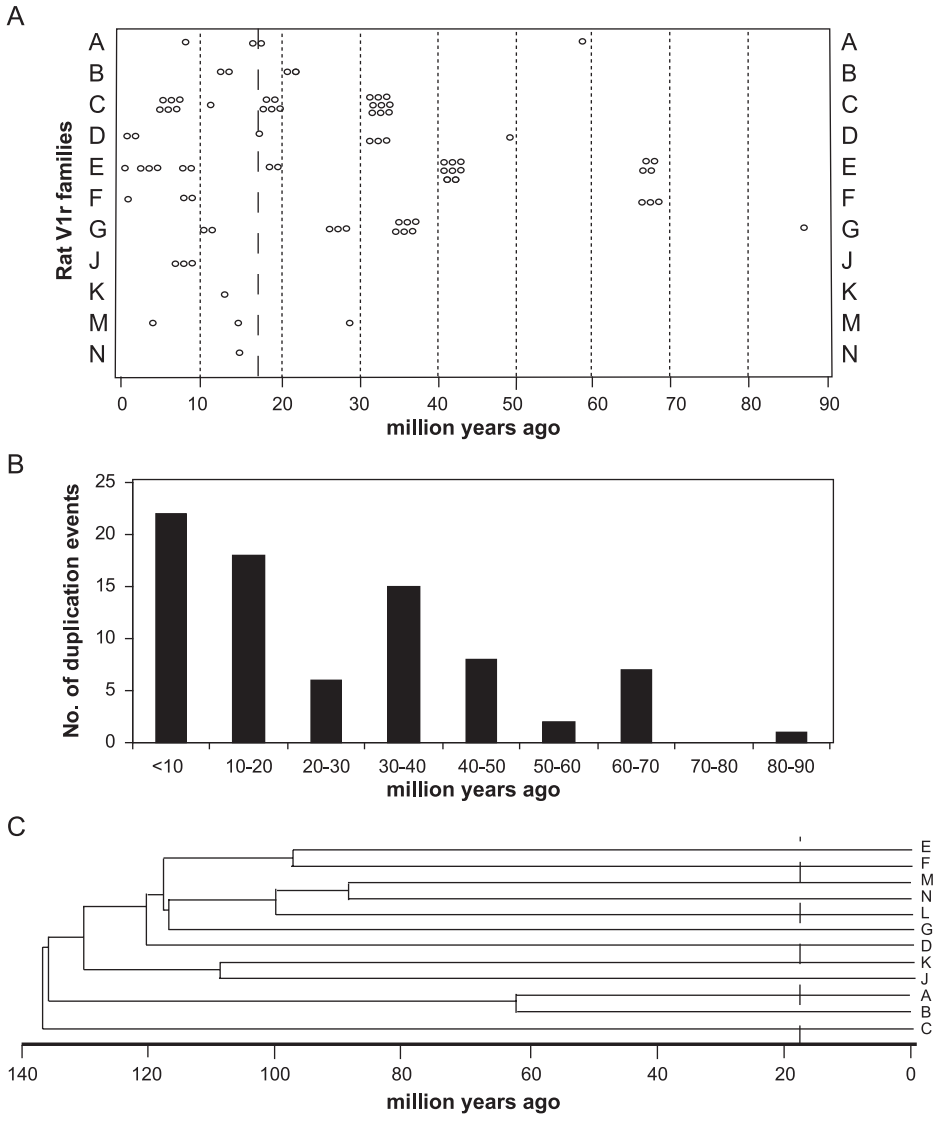

Fig. 5. Estimates of the dates of duplication events in the evolution of the rat *V1r* superfamily. (A) Dates for duplications within families. (○) Duplication event. The vertical dashed line represents the time when mouse and rat were separated. The vertical dotted lines divide the figure into 10-MY intervals. (B) The distribution of the number of duplication events in 10-MY intervals. (C) The dates of the duplication events leading to the 12 families found in rat. The dashed line indicates the time of rat–mouse divergence.

pheromone cues during evolution, which may be under intense sexual and/or natural selection.

The gene family tree in Fig. 1 gives the impression that the *V1r* repertoire enlarges over evolutionary time. This interpretation of the tree could be wrong. It is quite possible that the *V1r* repertoire in the common ancestor of mice and rats was as complex as those of today's mice and rats. Since new genes gained in evolution show up in the tree whereas old genes that have been lost no longer appear, the tree looks as if there are more genes now than in the past. In other words, one should not interpret the evolutionary pattern of *V1r* genes as gene family expansion (Lane et al., 2002), unless there is other evidence. A more appropriate term is "gene sorting", which describes unequal loss, retention, and amplification of ancestral genes between species (Zhang et al., 2000) and is consistent with the birth-and-death model of gene family evolution (Nei et al., 1997). For example, molecular dating suggests that the rat-specific V1rm and V1rn families were present in the common ancestor of the mouse and rat. Thus, the absence of these two families in the mouse must be due to lineage-specific loss that occurred in the mouse lineage since it was separated from the rat lineage. Similarly, the absence of the mouse-specific V1rh and V1ri families in rat must be due to the lineage-specific loss in rats. In fact, a *V1ri* pseudogene is still present in the rat genome.

In addition to the gains and losses of entire V1r families and gains and losses of *V1r* genes within families, divergence between mouse and rat *V1r* repertoires may also occur by functional divergence of orthologous genes. At present, there is no direct evidence showing functional differences between mouse and rat orthologous *V1r* genes.

If positive selection is found to act in the divergence of orthologous genes, one may infer that there is functional divergence. This is because positive selection occurs only if there is a functional change that results in increased organismal fitness. In our analysis, we did not find evidence for higher d_N over d_S for any of the 18 orthologous gene pairs compared. However, the average d_N/d_S for the V1r genes is significantly higher than that for other mouse/rat genes, indicating that V1rs are subject to either weak purifying selection or positive selection that may act at only a few sites of the protein. In this regard, it is interesting to note that a few earlier studies have suggested action of positive selection on V1rs. Emes et al. (2004) analyzed 22 mouse and rat *V1r* genes by a likelihood method (Yang et al., 2000) and identified positive selection acting on 14 codons. Lane et al. (2004) used the same statistical method in an analysis of 14 mouse and rat *V1r* genes and identified positive selection at 5–10% of codons. Because most of their sequences were paralogous, it is not possible to infer whether there has been positive selection between orthologs. Furthermore, the statistical method they used has been shown to be unreliable because it often detects positive selection when there is none (Suzuki and Nei, 2001, 2002, 2004). Recently, Zhang et al. (2004) conducted a comprehensive comparison of *V1r* genes detected from two mouse genome assemblies. Because the two assemblies were derived from different inbred strains of mice, the orthologous differences observed reflect intraspecific polymorphisms. These authors found a d_N/d_S ratio of 1.13 at the polymorphic sites. Although this was presented as evidence for positive selection, we think that it should be interpreted with care for three reasons. First, it has not been shown that the d_N/d_S ratio is significantly higher than 1, thus, the observation does not reject the neutral hypothesis. In fact, a relatively high d_N/d_S ratio within species could result from ineffective purifying selection if there has been recent population shrinking. Second, the observed between-strain genetic differences were likely present in mouse populations before breeding, as the history of managed breeding is short. However, because of intense artificial selection and/or extensive genetic drift during breeding, the differences between the two inbred strains may not accurately represent the polymorphic pattern in natural mouse populations. Third, the evolutionary patterns of *V1r* genes may be different at the intra- and interspecific levels because pheromones are supposed to mediate species-specific communications. Thus, even if positive selection is acting at the intraspecific level, it may not act at the interspecific level. Taken together, although it is quite likely that orthologous *V1r* genes between species have been under positive selection, no conclusive evidence is available.

Previous studies of limited numbers of mouse and rat V1rs suggested that many *V1r* gene duplications occurred at the time of split between mouse and rat. Our comprehensive analysis of all rat V1rs do not show a particularly high number of duplication events around that time. Instead, the number of successful duplications appears to be higher in recent times than in the past (Fig. 5). This observation, if coupled with a constant number of gene losses per unit time, would suggest that the *V1r* repertoire is expanding. It will be interesting to estimate the gene loss rate using pseudogene data as well as functional gene data from additional species that are closely related to the mouse and rat. Our mapping data suggest that gains of *V1r* genes often occur by tandem gene duplication. Lane et al. (2004) found that these duplications appear to be mediated by repetitive elements in the genome.

Our molecular dating of the *V1r* gene families, although approximate, provides two interesting results. First, the V1r families of the mouse and rat were established early in mammalian evolution. The majority of the families observed in mouse and rat were likely present in the most recent common ancestor of extant rodents, which existed about 75 MY ago (Springer et al., 2003). The second observation is that many V1r families emerged between 90 and 140 MY ago. This is the time shortly before and during the radiation of placental mammals (80–110 MY ago; Springer et al., 2003). Although entirely speculative, it is tempting to hypothesize that the rapid diversification of V1r families enabled the development of advanced and intricate pheromone-mediated communications, which facilitated increased speciation and mammalian radiation. Characterizing *V1r* and *V2r* genes from more placental mammals as well as from marsupials would help test this idea.

Acknowledgments

We thank Soochin Cho and David Webb for valuable comments. This work was supported by a startup fund from the University of Michigan and National Institutes of Health grant GM67030 to J.Z.

References

Abbott, A., 2004. Laboratory animals: the Renaissance rat. Nature 428, 464–466.

Cadavid, L.F., Shufflebotham, C., Ruiz, F.J., Yeager, M., Hughes, A.L., Watkins, D.I., 1997. Evolutionary instability of the major histocompatibility complex class I loci in New World primates. Proc. Natl. Acad. Sci. U. S. A. 94, 14536–14541.

Del Punta, K., Leinders-Zufall, T., Rodriguez, I., Jukam, D., Wysocki, C.J., Ogawa, S., Zufall, F., Mombaerts, P., 2002. Deficient pheromone responses in mice lacking a cluster of vomeronasal receptor genes. Nature 419, 70–74.

Dulac, C., Axel, R., 1995. A novel family of genes encoding putative pheromone receptors in mammals. Cell 83, 195–206.

Dulac, C., Torello, A.T., 2003. Molecular detection of pheromone signals in mammals, from genes to behaviour. Nat. Rev., Neurosci. 4, 551–562.

Emes, R.D., Beatson, S.A., Ponting, C.P., Goodstadt, L., 2004. Evolution and comparative genomics of odorant- and pheromone-associated genes in rodents. Genome Res. 14, 591–602.

Felsenstein, J., 1985. Confidence limits on phylogenies: an approach using the bootstrap. Evolution 39, 783–791.

Giorgi, D., Rouquier, S., 2002. Identification of V1R-like putative pheromone receptor sequences in non-human primates. Characterization of *V1R* pseudogenes in marmoset, a primate species that possesses an intact vomeronasal organ. Chem. Senses 27, 529–537.

Herrada, G., Dulac, C., 1997. A novel family of putative pheromone receptors in mammals with a topographically organized and sexually dimorphic distribution. Cell 90, 763–773.

Keverne, E.B., 1999. The vomeronasal organ. Science 286, 716–720.

Kumar, S., Tamura, K., Jakobsen, I.B., Nei, M., 2001. MEGA2: molecular evolutionary genetics analysis software. Bioinformatics 17, 1244–1245.

Lane, R.P., Cutforth, T., Axel, R., Hood, L., Trask, B.J., 2002. Sequence analysis of mouse vomeronasal receptor gene clusters reveals common promoter motifs and a history of recent expansion. Proc. Natl. Acad. Sci. U. S. A. 99, 291–296.

Lane, R.P., Young, J., Newman, T., Trask, B.J., 2004. Species specificity in rodent pheromone receptor repertoires. Genome Res. 14, 603–608.

Liman, E.R., Innan, H., 2003. Relaxed selective pressure on an essential component of pheromone transduction in primate evolution. Proc. Natl. Acad. Sci. U. S. A. 100, 3328–3332.

Matsunami, H., Buck, L.B., 1997. A multigene family encoding a diverse array of putative pheromone receptors in mammals. Cell 90, 775–784.

Nei, M., Gu, X., Sitnikova, T., 1997. Evolution by the birth-and-death process in multigene families of the vertebrate immune system. Proc. Natl. Acad. Sci. U. S. A. 94, 799–806.

Pantages, E., Dulac, C., 2000. A novel family of candidate pheromone receptors in mammals. Neuron 28, 835–845.

Rat Genome Sequencing Consortium, 2004. Genome sequence of the Brown Norway rat yields insights into mammalian evolution. Nature 428, 493–521.

Rodriguez, I., Mombaerts, P., 2002. Novel human vomeronasal receptor-like genes reveal species-specific families. Curr. Biol. 12, R409–R411.

Rodriguez, I., Del Punta, K., Rothman, A., Ishii, T., Mombaerts, P., 2002. Multiple new and isolated families within the mouse superfamily of V1r vomeronasal receptors. Nat. Neurosci. 5, 134–140.

Ryba, N.J., Tirindelli, R., 1997. A new multigene family of putative pheromone receptors. Neuron 19, 371–379.

Saitou, N., Nei, M., 1987. The neighbor-joining method: a new method for reconstructing phylogenetic trees. Mol. Biol. Evol. 4, 406–425.

Sitnikova, T., Su, C., 1998. Coevolution of immunoglobulin heavy- and light-chain variable-region gene families. Mol. Biol. Evol. 15, 617–625.

Springer, M.S., Murphy, W.J., Eizirik, E., O'Brien, S.J., 2003. Placental mammal diversification and the Cretaceous–Tertiary boundary. Proc. Natl. Acad. Sci. U. S. A. 100, 1056–1061.

Suzuki, Y., Nei, M., 2001. Reliabilities of parsimony-based and likelihood-based methods for detecting positive selection at single amino acid sites. Mol. Biol. Evol. 18, 2179–2185.

Suzuki, Y., Nei, M., 2002. Simulation study of the reliability and robustness of the statistical methods for detecting positive selection at single amino acid sites. Mol. Biol. Evol. 19, 1865–1869.

Suzuki, Y., Nei, M., 2004. False-positive selection identified by ML-based methods, examples from the *Sig1* gene of the diatom *Thalassiosira weissflogii* and the tax gene of a human T-cell lymphotropic virus. Mol. Biol. Evol. 21, 914–921.

Thompson, J.D., Gibson, T.J., Plewniak, F., Jeanmougin, F., Higgins, D.G., 1997. The CLUSTAL_X windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. Nucleic Acids Res. 25, 4876–4882.

Yang, Z., Nielsen, R., Goldman, N., Pedersen, A.M., 2000. Codon-substitution models for heterogeneous selection pressure at amino acid sites. Genetics 155, 431–449.

Zhang, J., Webb, D.M., 2003. Evolutionary deterioration of the vomeronasal pheromone transduction pathway in catarrhine primates. Proc. Natl. Acad. Sci. U. S. A. 100, 8337–8341.

Zhang, J., Rosenberg, H.F., Nei, M., 1998. Positive Darwinian selection after gene duplication in primate ribonuclease genes. Proc. Natl. Acad. Sci. U. S. A. 95, 3708–3713.

Zhang, J., Dyer, K.D., Rosenberg, H.F., 2000. Evolution of the rodent eosinophil-associated RNase gene family by rapid gene sorting and positive selection. Proc. Natl. Acad. Sci. U. S. A. 97, 4701–4706.

Zhang, X., Rodriguez, I., Mombaerts, P., Firestein, S., 2004. Odorant and vomeronasal receptor genes in two mouse genome assemblies. Genomics 83, 802–811.