The Microevolution of *V1r* Vomeronasal Receptor Genes in Mice

Seong Hwan Park[†,1,2], Ondrej Podlaha[1], Wendy E. Grus[1], and Jianzhi Zhang[*,1]

[1]Department of Ecology and Evolutionary Biology, University of Michigan, Ann Arbor, Michigan

[2]Department of Legal Medicine, College of Medicine, Korea University, Seoul, Korea

*Corresponding author: E-mail: jianzhi@umich.edu.

†Present address: Department of Pathology, Chucheon Sacred Heart Hospital, College of Medicine, Hallym University, Chuncheon, Korea

Accepted: 21 April 2011

Abstract

Vomeronasal sensitivity is important for detecting intraspecific pheromonal cues as well as environmental odorants and is involved in mating, social interaction, and other daily activities of many vertebrates. Two large families of seven-transmembrane G-protein–coupled receptors, V1rs and V2rs, bind to various ligands to initiate vomeronasal signal transduction. Although the macroevolution of *V1r* and *V2r* genes has been well characterized throughout vertebrates, especially mammals, little is known about their microevolutionary patterns, which hampers a clear understanding of the evolutionary forces behind the rapid evolutionary turnover of *V1r* and *V2r* genes and the great diversity in receptor repertoire across species. Furthermore, the role of divergent vomeronasal perception in enhancing premating isolation and maintaining species identity has not been evaluated. Here we sequenced 44 *V1r* genes and 25 presumably neutral noncoding regions in 14 wild-caught mice belonging to *Mus musculus* and *M. domesticus*, two closely related species with strong yet incomplete reproductive isolation. We found that nucleotide changes in *V1r*s are generally under weak purifying selection and that only ~5% of *V1r*s may have been subject to positive selection that promotes nonsynonymous substitutions. Consistent with the low functional constraints on *V1r*s, 18 of the 44 *V1r*s have null alleles segregating in one or both species. Together, our results demonstrate that, despite occasional actions of positive selection, the evolution of *V1r*s is in a large part shaped by purifying selection and random drift. These findings have broad implications for understanding the driving forces of rapid gene turnovers that are often observed in the evolution of large gene families.

Key words: *Mus musculus*, *Mus domesticus*, vomeronasal receptor, pheromone detection, V1r, evolution.

Introduction

Olfaction plays a critical role in the daily life of vertebrates, such as prey detection, predator avoidance, mating, and territoriality (Mombaerts 1999). Two distinct nasal olfactory systems exist in most terrestrial vertebrates: the main olfactory system (MOS) and the vomeronasal system (VNS) (Dulac and Torello 2003; Grus and Zhang 2006). Although partially overlapping in function, the MOS appears to be mainly responsible for recognizing environmental odorants, whereas the VNS primarily detects pheromones, which constitute a poorly defined class of chemicals that are emitted and sensed by individuals of the same species to elicit sexual/social behaviors and physiological changes (Restrepo et al. 2004; Spehr et al. 2006). The MOS and the VNS are anatomically and neurologically separated; they use different receptors and have distinct signal transduction pathways (Dulac and Torello 2003; Grus and Zhang 2006). The VNS is of particular interest to evolutionists because of its high diversity in complexity among species (Grus et al. 2005; Young et al. 2005, 2010; Shi and Zhang 2007; Grus and Zhang 2008). Although the morphological components of the VNS are believed to first emerge in the common ancestor of tetrapods, its genetic components have been inferred to exist in the common ancestor of all vertebrates (Grus and Zhang 2006, 2009). Among tetrapods, the VNS varies from completely absent in birds, catarrhine primates (humans, apes, and Old World monkeys), most bats, and many cetaceans to rudimentary in amphibians to highly complex in murids, opossums, and the platypus (Zhang and Webb 2003; Grus et al. 2005, 2007; Grus and Zhang 2006; Shi and Zhang 2007, 2009; Zhao et al. 2011).

Vomeronasal sensitivity is mediated by two families of G-protein–coupled receptors known as V1rs and V2rs (Mombaerts 2004). In the genome of the laboratory mouse,

there are about 190 putatively functional *V1r* and 70 *V2r* genes (Shi et al. 2005; Yang et al. 2005). Note that despite the availability of genome sequences, the gene numbers are only approximate due to among-strain variations and/or incomplete genomic sequencing (Zhang et al. 2004). The *V1r* and *V2r* gene repertoires, especially the former, have been examined in many mammalian genomes (Rodriguez and Mombaerts 2002; Rodriguez et al. 2002; Grus and Zhang 2004, 2008; Grus et al. 2005, 2007; Young et al. 2005, 2010; Shi and Zhang 2007; Young and Trask 2007). It was reported that the among-species size variation in *V1r* and *V2r* gene repertoires is among the highest of all mammalian gene families (Grus et al. 2005, 2007; Yang et al. 2005). This variation is not random, at least in the case of *V1r*s, because a clear positive correlation exists between the morphological complexity of the VNS and the number of putatively functional *V1r* genes (Grus et al. 2005, 2007). An evolutionary hallmark of *V1r*s and *V2r*s is the exceptionally rapid gene turnover that results in lineage-specific receptors. For example, between 187 mouse and 106 rat *V1r*s examined, only 18 are one-to-one orthologous (Grus and Zhang 2004, 2008), in sharp contrast to the genome-wide estimate that 86–94% of rat genes have one-to-one mouse orthologs (Gibbs et al. 2004). Despite the striking macroevolutionary diversity of *V1r*s and *V2r*s, the evolutionary forces acting on these genes are unclear due to the lack of knowledge about the population genetic dynamics of *V1r* and *V2r* genes. Specifically, it would be interesting to test whether *V1r*s and *V2r*s evolve by divergent selective pressures in sibling species because pheromones are by definition species specific (Brennan and Keverne 2004).

In this work, we study the microevolution of vomeronasal receptor genes in two closely related mouse species, *Mus musculus* (abbreviated as *Mm*) and *Mus domesticus* (*Md*). *Mm* is distributed from Eastern Europe to Japan, across Russia and northern China, whereas *Md* is common in Western Europe, Africa, and the near-East and was transported by humans to the Americas and Australia (Guenet and Bonhomme 2003). *Mm* and *Md* diverged within the last 500 thousand years (Salcedo et al. 2007). The two species form a narrow zone of hybridization through Central Europe that extends from the Jutland Peninsula to the Bulgarian coast of the Black Sea (Sage et al. 1993; Tucker 2007). Mice from the center of the hybrid zone have higher parasite loads than those from the edges of the hybrid zone (Sage et al. 1986; Moulia et al. 1991, 1993), indicative of reduced fitness due to hybrid inviability. Additional laboratory studies have documented reproductive incompatibility between *Mm* and *Md*. There is clear evidence of hybrid male sterility between the two species (Forejt and Ivanyi 1974; Forejt 1996; Alibert et al. 1997; Storchova et al. 2004; Britton-Davidian et al. 2005; Trachtulec et al. 2005; Vyskocilova et al. 2005). There is also evidence for limited female sterility in some crosses but not others (Forejt and Ivanyi 1974; Britton-Davidian et al. 2005). Evidence for partial premating isolation is also ample (Laukaitis et al. 1997; Smadja and Ganem 2002; Smadja et al. 2004). Although some authors regard *Mm* and *Md* as two subspecies of the species *M. musculus* (Tucker 2007), for simplicity, we treat them as two species that are in an early stage of divergence with a low degree of gene flow.

There are several reasons why we chose to study *Mm* and *Md*. First, the laboratory mouse, a mosaic of *Mm*, *Md*, and *Mus castaneus* (Frazer et al. 2007; Yang et al. 2007), is a model organism for studying vomeronasal sensitivity. A substantial amount of genetic, neurological, and behavioral data related to vomeronasal sensitivity is available for the laboratory mouse, allowing a more accurate interpretation of the evolutionary and population genetic data that we collect. Second, the genome sequence of the laboratory mouse is known, making the experimental design much easier. Third, mice represent those vertebrates with a relatively high level of vomeronasal sensitivity (Takami 2002; Grus et al. 2005). Thus, their vomeronasal sensitivities may be more important in determining organismal fitness and under stronger natural selection.

Avoiding the hybrid zone, we trapped seven wild *Mm* and seven wild *Md* in Czech Republic and France, respectively. Our present study focuses on *V1r*s because they have only one coding exon, making DNA amplification and sequencing much easier. Here, we report the microevolution of 44 *V1r*s and 25 presumably neutral noncoding regions in these 14 mice.

Materials and Methods

Seven *M. musculus* and seven *M. domesticus* individuals were collected from Czech Republic and France, respectively. Although the mice from each species were sampled from restricted geographic areas (supplementary table 1, Supplementary Material online), it should not affect our results because mice have little geographic differentiation (Salcedo et al. 2007). The identity of the mice was confirmed by sequencing a 683-nt segment of the mitochondrial cytochrome c oxidase subunit I (COXI) gene that is commonly used as a barcode for identifying animal species.

The liver genomic DNAs of the mice were extracted using the PUREGENE genomic DNA purification kit (Gentra Systems, Minneapolis, MN), following the manufacturer's instruction. Gene-specific primers for amplifying 44 *V1r* genes were designed according to the *Mus musculus* reference sequence from GenBank (supplementary table 2, Supplementary Material online). The protein-coding region of each *V1r* gene studied has 870–1,104 nt, which were completely amplified in our experiments. Polymerase chain reactions (PCRs) were performed with GoTaq DNA Polymerase (Promega Corp, Madison, WI) under conditions recommended by the manufacturer. PCR products were examined on 1.5% agarose gel. Samples showing duplicated

electropherograms due to insertions/deletions were cloned with TOPO PCR cloning kit (Invitrogen, Carlsbad, CA) and sequenced with universal T7 and M13 primers using the Sanger method on an automatic DNA sequencer. Otherwise, the PCR products were enzymatically processed using calf intestinal phosphatase and exonuclease I (Exo I) (New England Biolabs, Ipswich, MA) before being sequenced bidirectionally with the gene-specific primers. Sequencher (GeneCodes, Ann Arbor, MI) and MEGA4 (Tamura et al. 2007) were used to edit and align the sequences. Twenty-five presumably neutral noncoding regions (supplementary table 3, Supplementary Material online), most with ~1,000 nt, were also amplified and directly sequenced in the same 14 mice. Watterson's θ, nucleotide diversity π, Tajima's *D*, and Fu and Li's *D** were computed using DnaSP (Librado and Rozas 2009). Tajima's test (Tajima 1989) and Fu and Li's test (Fu and Li 1993) were conducted by 10,000 coalescent simulations in DnaSP. Hudson–Kreitman–Aguade (HKA) test (Hudson et al. 1987) was conducted using a program written by J. Hey (http://life-sci.rutgers.edu/~heylab/). The sequences reported in this article have been submitted to GenBank (accession numbers JF782602–JF783819, JF782044–JF782601, and JF783820–JF783959).

Results

Intraspecific Variations of *V1r*s and Noncoding Regions

Based on previous studies (Grus and Zhang 2004; Shi et al. 2005), there are at least 188 putatively functional *V1r* genes in the mouse genome (fig. 1). We carefully selected 44 of them for an in-depth study in 14 mice. These 44 genes were chosen to represent major lineages of mouse *V1r*s, to include genes with (14) and without (30) rat one-to-one orthologs, and to allow gene-specific amplification and sequencing (fig. 1). For comparison, we also sequenced 25 presumably neutral noncoding regions in these 14 mice. Five of these 25 noncoding regions were from a previous study (Baines and Harr 2007), and the sequenced segments are either in introns or in intergenic regions that are >5 kb upstream of coding regions. The remaining 20 noncoding regions were randomly picked from the genome, with the criteria that the regions are at least 200 kb away from any known gene. All sequenced noncoding regions are on autosomes, so are the sequenced *V1r* genes. The average length of the noncoding regions sequenced (936 nt) is similar to the average length of the *V1r*s sequenced (934 nt).

The basic population genetic parameters of individual *V1r*s and noncoding regions are presented in table 1 and table 2, respectively. Consistent with other population genetic studies (Salcedo et al. 2007), we found nucleotide diversity per site (π) at the 25 noncoding regions to be higher in *Md* (0.0021) than in *Mm* (0.0013), although the difference is not statistically significant ($P = 0.15$, two-tailed paired *t*-test; table 2). However, the opposite is found for *V1r*s, although the difference is again not significant ($P = 0.09$, two-tailed paired *t*-test; table 1). Compared with the noncoding regions, *V1r*s show an overall higher π in *Mm* ($P = 0.24$, two-tailed Mann–Whitney test) but a lower π in *Md* ($P = 0.014$, two-tailed Mann–Whitney test) (tables 1 and 2). In neither species is there a significant difference in π between *V1r*s with one-to-one rat orthologs and those without such orthologs ($P > 0.2$, two-tailed Mann–Whitney test; table 1).

We applied Tajima's test of neutrality (Tajima 1989) to each of the *V1r*s (table 1) and noncoding regions (table 2). Note that the null hypothesis in Tajima's test is the Wright–Fisher model of strict neutrality. Thus, rejection of the null hypothesis may indicate one or more of the following: purifying selection, positive selection, and demographic changes. For both *V1r*s and noncoding regions, several loci show significantly negative or positive Tajima's *D*. For example, in *Mm*, five *V1r*s and two noncoding regions show significantly positive *D* (nominal $P < 0.05$), whereas five *V1r*s and one noncoding region show significantly negative *D*. In *Md*, zero *V1r* and two noncoding regions show significantly positive *D*, whereas two *V1r*s and one noncoding region show significantly negative *D*. In neither species is there a significant difference between *V1r*s and noncoding regions in the fraction of loci with significantly positive or negative *D* ($P > 0.1$ in all cases, χ^2 test). We also compared the frequency distribution of Tajima's *D* between *V1r*s and noncoding regions (fig. 2a and 2b) but found no significant differences ($P = 0.44$ for *Mm* and 0.12 for *Md*, Kolmogorov–Smirnov test). We found similar results from comparing the distribution of Fu and Li's *D** between *V1r*s and noncoding regions ($P = 0.74$ for *Mm* and 0.30 for *Md*, Kolmogorov–Smirnov test, fig. 2c and 2d).

Interspecific Divergences of *V1r*s and Noncoding Regions

The mean number of nucleotide difference per site between *Mm* and *Md* is 0.00442 for the *V1r*s (table 1) and 0.00804 for the noncoding regions (table 2). The mean nonsynonymous nucleotide difference per nonsynonymous site (d_N) of the 44 *V1r*s divided by the mean synonymous difference per synonymous site (d_S) of the same set of genes is 0.315. Seven *V1r*s have $d_N/d_S > 1$, but none of them significantly exceeds 1 by Fisher's exact test (Zhang et al. 1997). The fraction of *V1r*s with a d_N/d_S ratio below 1 is significantly greater than 50% ($P = 3 \times 10^{-6}$, binomial test). These results indicate that the evolutionary divergence of *V1r*s is overall governed by purifying selection.

Combining the polymorphism and divergence data, we conducted several McDonald–Kreitman tests (McDonald and Kreitman 1991) by varying the consideration of the polymorphic data from one or both species (table 3). In all cases, nonsynonymous/synonymous ratio is lower for

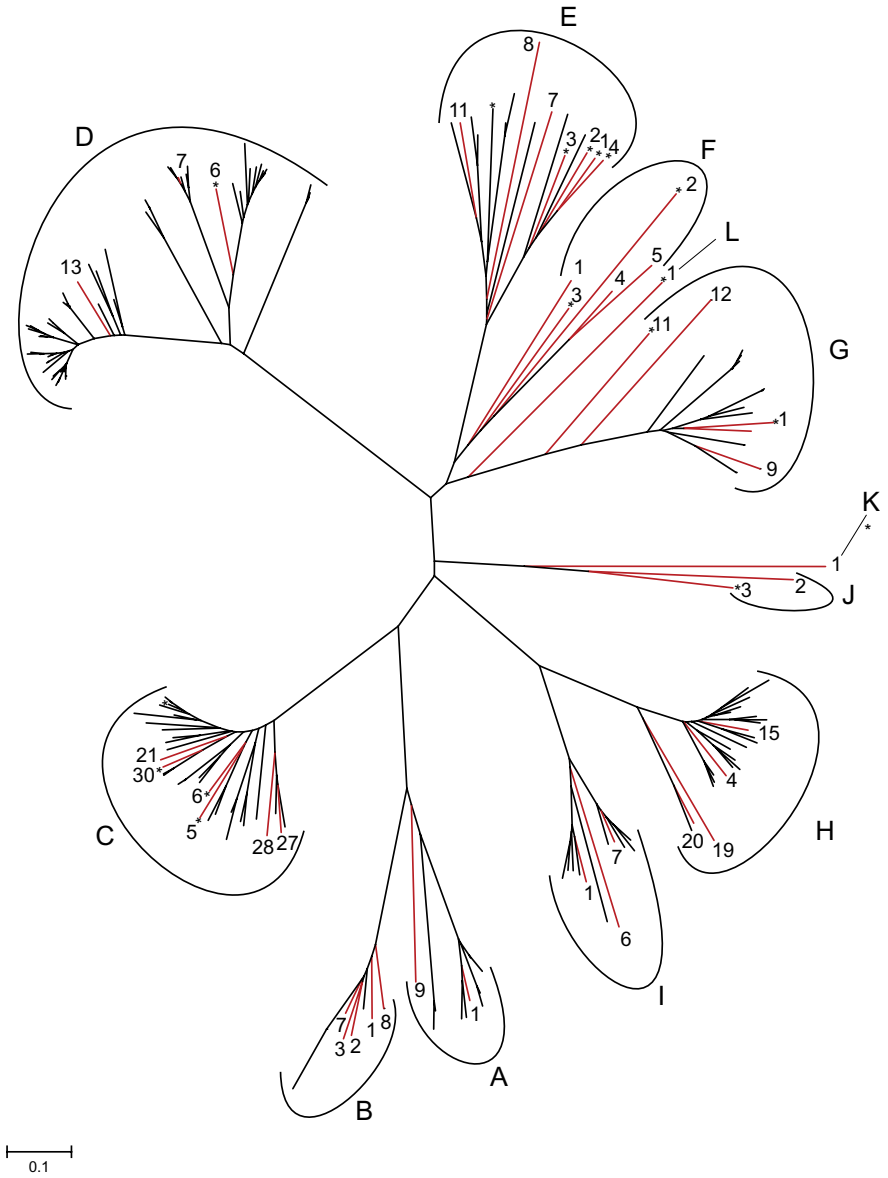

Fig. 1.—An unrooted tree of 188 putatively functional mouse V1rs (Shi et al. 2005). Red branches denote the 44 V1rs surveyed in this study. Branches denoted with * have putatively functional rat V1r orthologs (Grus and Zhang 2008). Gene family names are from Rodriguez et al. (2002). The tree was reconstructed using the Neighbor-Joining method (Saitou and Nei 1987) with Poisson-corrected protein distances. The scale bar shows 0.1 amino acid substitution per site.

divergence than for polymorphism, although the differences are not statistically significant (table 3). These findings suggest that the evolution of mouse *V1rs* is largely neutral, with the presence of only weak purifying selection that hampers the fixation of some nonsynonymous changes. For instance, the nonsynonymous/synonymous ratio is 102/56 = 1.82 for intraspecific polymorphisms in *Md* but 32/31 = 1.03 for interspecific divergences ($P = 0.068$, Fisher's exact test). Consistent with the above interpretation, we found the nonsynonymous/synonymous ratios for polymorphisms and divergences to be more similar to each other when only derived alleles with frequencies equal to or greater than 2/14 are considered for polymorphisms. For instance, the nonsynonymous/synonymous ratio now becomes 62/44 = 1.41 for polymorphisms in *Md*, closer to the ratio of 1.03 for interspecific divergences ($P = 0.76$, Fisher's exact test).

Because of the relatively small numbers of synonymous polymorphisms and substitutions in our *V1r* data, we augmented this dataset with the 25 noncoding regions to enhance the statistical power of the McDonald–Kreitman test. That is, we lumped synonymous and noncoding changes and compared them with nonsynonymous changes (table

Table 1

Intra- and Interspecific Sequence Variations of 44 Mouse *V1r* Genes

V1r genes	Length (nt)	Within *M. musculus*					Within *M. domesticus*					Between species				
		S[1]	θ ($\times 10^4$)[2]	π ($\times 10^4$)[3]	Tajima's D[4]	Fu and Li's D*[4]	S[1]	θ ($\times 10^4$)[2]	π ($\times 10^4$)[3]	Tajima's D[4]	Fu and Li's D*[4]	K[5]	d ($\times 10^4$)[6]	d_N ($\times 10^4$)[7]	d_S ($\times 10^4$)[8]	d_N/d_S
A1	990	4	12.7	16.0	0.85	1.16	5	15.9	13.2	−0.58	−0.95	0	31.0	16.0	81.9	0.195
A9	948	1	3.4	3.9	0.32	0.72	2	6.6	10.3	1.51	0.94	12	137.1	77.8	345.7	0.225
B1	933	10	33.8	55.7	2.53	1.42	6	20.2	13.0	−1.29	−1.88	2	78.1	57.3	152.2	0.376
B2	933	7	23.6	39.6	2.50	1.33	6	20.2	13.3	−1.23	−1.88	2	65.1	52.3	111.0	0.471
B3	933	2	6.7	7.5	0.32	0.94	6	20.2	24.1	0.70	0.65	0	23.7	21.0	33.4	0.628
B7	933	1	3.4	5.7	1.43	0.72	3	10.1	12.5	0.72	0.02	1	29.1	37.9	0.0	NA
B8	930	1	3.4	1.5	−1.16	−1.40	2	6.8	4.4	−0.96	−0.45	11	121.4	101.2	197.3	0.513
C5[9]	903	0	0.0	0.0	0.00	0.00	2	7.0	6.9	−0.01	0.94	11	125.8	77.6	296.9	0.261
C6	900	4	14.0	20.1	1.45	1.16	8	28.0	21.2	−0.90	−1.16	5	89.7	81.2	121.5	0.669
C21	894	0	0.0	0.0	0.00	0.00	1	3.5	3.0	−0.34	0.72	0	9.6	0.0	41.6	0.000
C27	927	18	61.1	85.9	1.69	1.54	3	10.2	11.1	0.29	0.02	0	87.1	53.1	209.0	0.254
C28[10]	894	7	27.7	41.8	2.16	1.38	4	14.1	6.4	−1.80	−2.27	0	34.5	21.5	81.0	0.265
C30	900	1	3.5	2.9	−0.34	0.72	5	17.5	16.6	−0.17	1.23	1	31.0	2.1	222.1	0.009
D6	999	3	9.4	6.7	−0.89	0.02	2	5.1	6.3	−0.53	−0.45	0	20.7	4.7	75.1	0.062
D7	951	1	3.3	1.5	−1.16	−1.40	4	13.2	13.6	0.10	0.30	7	83.4	54.3	182.2	0.298
D13	915	1	3.4	5.4	1.21	0.72	6	20.6	9.4	−1.96	−2.52	0	10.7	5.1	29.9	0.170
E1	921	7	23.9	24.1	0.03	1.33	1	3.4	5.4	1.21	0.72	0	28.7	31.5	20.4	1.542
E2	897	1	3.5	5.9	1.43	0.72	1	3.5	4.0	0.32	0.72	2	31.1	25.9	49.7	0.521
E3	1,002	5	15.7	8.3	−1.62	−1.68	0	0.0	0.0	0.00	0.00	1	14.3	2.8	54.4	0.051
E4	921	4	13.7	6.2	−1.80	−2.27	2	6.8	8.2	0.56	0.94	0	36.4	28.3	64.8	0.437
E7	936	2	6.7	7.5	0.32	0.94	0	0.0	0.0	0.00	0.00	1	19.8	2.0	81.2	0.024
E8	951	3	9.9	4.5	−1.67	−2.09	4	13.2	18.4	1.28	1.16	0	15.0	9.8	33.2	0.295
E11	933	3	13.5	8.7	−1.16	−0.55	5	16.9	20.1	0.68	0.51	0	22.4	20.4	29.6	0.689
F1	990	4	12.8	9.4	−0.88	−1.41	1	3.2	1.4	−1.16	−1.40	0	6.5	7.6	3.2	2.362
F2	987	9	28.7	15.5	−2.00	−2.36	2	6.4	7.1	0.32	0.94	0	11.8	2.8	45.5	0.062
F3[11]	870	16	100.3	122.6	2.26	2.26	1	3.6	1.6	−1.16	−1.40	0	92.8	76.4	151.4	0.505
F4	1,005	5	15.6	8.3	−1.62	−1.68	1	3.1	4.4	0.84	0.72	0	26.0	33.6	2.9	11.502
F5	960	7	23.0	16.8	−1.00	0.21	0	0.0	0.0	0.00	0.00	0	46.1	24.6	119.2	0.206
G1	921	3	10.2	10.1	−0.03	0.02	21	71.7	60.4	−0.66	1.12	0	40.9	32.3	70.1	0.461
G6	918	19	65.1	100.9	2.29	1.07	6	20.6	20.6	0.01	0.02	0	75.8	40.8	253.4	0.176
G9	921	7	23.9	35.0	1.71	0.77	2	6.8	5.7	−0.44	0.94	0	42.7	27.8	94.2	0.295
G11	957	2	6.6	6.1	−0.20	−0.45	2	6.6	6.1	−0.44	0.94	0	11.2	12.6	6.5	1.941
G12	915	8	27.5	13.8	−1.88	−2.17	4	13.7	8.9	−1.16	−0.55	0	74.2	46.7	162.8	0.287
H4	1,104	1	2.8	4.0	0.84	0.72	4	11.4	6.3	−1.48	−1.41	4	48.5	16.8	157.3	0.107
H15	897	2	7.0	7.5	0.18	−0.45	11	38.6	50.1	1.18	1.06	0	49.4	32.5	105.2	0.309
H19	906	3	10.4	16.4	1.75	1.07	3	10.4	12.6	0.65	1.07	0	28.5	19.9	57.0	0.349
H20	897	9	31.6	40.7	1.11	1.40	3	10.5	6.1	−1.28	−1.04	0	43.0	21.9	114.9	0.190
I1	903	0	0.0	0.0	0.00	0.00	4	13.9	11.0	−0.70	−0.55	0	5.9	7.3	3.4	2.142
I6	948[12]	2	6.6	5.6	−0.44	0.94	3	10.0	10.0	0.01	0.02	3	40.7	51.4	9.3	5.551
I7	909	5	17.3	22.3	0.99	1.23	2	6.9	5.8	−0.44	0.94	0	44.0	35.2	75.5	0.467
J2	951	5	16.6	8.8	−1.62	−1.68	3	9.9	9.8	−0.03	0.02	0	29.3	4.9	72.8	0.067
J3	927	6	20.4	13.4	−1.23	−1.88	0	0.0	0.0	0.00	0.00	0	45.5	27.3	106.5	0.256
K1	888	3	10.6	4.8	−1.67	−2.09	2	7.1	8.9	0.70	0.94	0	18.3	7.4	55.6	0.132
L1	897	5	17.5	10.7	−1.36	−0.95	5	17.5	17.3	−0.05	0.51	0	15.9	11.5	30.9	0.371

Table 1
Continued

V1r genes	Length (nt)	Within *M. musculus*					Within *M. domesticus*					Between species				
		S[1]	θ ($\times 10^4$)[2]	π ($\times 10^4$)[3]	Tajima's D[4]	Fu and Li's D*[4]	S[1]	θ ($\times 10^4$)[2]	π ($\times 10^4$)[3]	Tajima's D[4]	Fu and Li's D*[4]	K[5]	d ($\times 10^4$)[6]	d_N ($\times 10^4$)[7]	d_S ($\times 10^4$)[8]	d_N/d_S
Mean 1[13]	916	3.8	13.0	18.1	0.08	0.00	3.9	13.5	12.0	−0.18	0.17	1.3	40.0	29.8	76.5	0.389
Mean 2[14]	943	5.1	17.1	19.3	0.09	0.00	3.4	11.4	10.9	−0.17	−0.07	1.5	46.1	30.3	104.7	0.289
Mean[15]	934	4.7	15.8	18.9	0.08	0.00	3.6	12.1	11.3	−0.18	0.00	1.4	44.2	30.1	95.7	0.315

NOTE.—[1]Number of polymorphic sites.
[2] Watterson's polymorphism per site.
[3] Nucleotide diversity per site.
[4] Values significantly different from 0 at the 5% level are underlined. Significance is determined by 10,000 coalescent simulations.
[5] Number of fixed nucleotide differences between species.
[6] Mean number of nucleotide differences between species.
[7] Mean number of nonsynonymous differences per nonsynonymous site between species.
[8] Mean number of synonymous differences per synonymous site between species.
[9] Shaded genes have one-to-one rat orthologs, whereas unshaded genes do not have one-to-one rat orthologs.
[10] Two *M. musculus* samples could not be amplified.
[11] Five *M. musculus* samples could not be amplified.
[12] All *M. domesticus* sequences have an amino acid insertion and are three bases longer than those of *M. musculus*.
[13] Average for the mouse genes that have one-to-one rat orthologs.
[14] Average for the mouse genes that have no one-to-one rat orthologs.
[15] Average for all genes.

3). The results are consistent with those from the comparison between synonymous and nonsynonymous changes, but the difference between polymorphism and divergence becomes statistically more significant (table 3). For example, the nonsynonymous/(noncoding+synonymous) ratio is 0.510 for intraspecific polymorphisms in *Md*, significantly higher than that (0.227) for interspecific divergences ($P < 0.001$, Fisher's exact test). Together, the various McDonald–Kreitman tests demonstrate the overall action of purifying selection hampering the spread and fixation of nonsynonymous changes in *V1r*s. We did not perform McDonald–Kreitman tests for individual *V1r*s because of the low numbers of synonymous and nonsynonymous changes in each *V1r* and the consequent low statistical power.

We examined d/θ for each *V1r* in each species, where d is the average nucleotide difference per site between an *Mm* allele and an *Md* allele and θ is Watterson's estimate of polymorphism per site in a species (table 4). Because some sequences have no polymorphic sites, we used the actual number of polymorphic site plus 1 in calculating θ for each *V1r* gene or noncoding region. In *Mm*, the mean d divided by mean θ is 2.59 for *V1r*s, whereas the corresponding ratio is 7.30 for the noncoding regions. In *Md*, the ratios are 3.63 and 4.20 for *V1r*s and noncoding regions, respectively. Thus, overall, *V1r*s have lower divergence-to-polymorphism ratios than noncoding regions, indicative of purifying selection on *V1r*s. When each gene is examined separately by the HKA test (Hudson et al. 1987), however, three genes (A9, B8, and C5) show significantly greater d/θ than the 25 noncoding regions in both *Mm* and *Md* and three additional genes (F3, F5, and J3) show significantly greater d/θ than noncoding regions only in *Md* (table 4). Because 44 tests were conducted in each species, some of the significant cases (on average 2.2 cases per species) may be artifacts of multiple testing. After examining the fixed differences between *Mm* and *Md*, we believe that F3, F5, and J3 are probably false positives because they lack any fixed differences, whereas A9, B8, and C5 are likely to be true positives because they each contain at least five fixed nonsynonymous differences between the two species, and the statistical significance of the HKA test is high ($P < 0.0025$ for each gene in each species). Even after the conservative Bonferroni correction, A9 remains significant in both species and C5 remains significant in *Mm*. Thus, it is likely that a small fraction of *V1r*s has been subject to positive selection in the divergence of *Mm* and *Md*.

Abundant Segregating Null Alleles of *V1r*s

We observed a large number of *V1r* genes that have segregating null alleles in either one or both mouse species based on the occurrences of single nucleotide polymorphisms and/or insertions/deletions (indels) that introduce premature stop codons. For two *V1r* genes (C28 and F3), amplification was unsuccessful in some but not all mouse individuals even

Table 2

Intra- and Interspecific Sequence Variations of 25 Noncoding Regions in Mice

				Within *M. musculus*					Within *M. domesticus*					Between species	
Region name[1]	Chromos. no.	Nucleotide position[2]	length (nt)[3]	*S*[4]	θ ($\times 10^4$)[5]	π ($\times 10^4$)[6]	Tajima's D[7]	Fu and Li's D*[7]	*S*[4]	θ ($\times 10^4$)[5]	π ($\times 10^4$)[6]	Tajima's D[7]	Fu and Li's D*[7]	*K*[8]	*d* ($\times 10^4$)[9]
032	3	61421835–61422742	908	1	3.5	1.6	−1.16	−1.40	10	34.6	26.4	−0.93	0.58	2	85.7
con1	5	4750295–4751185	885	4	14.2	22.4	1.89	1.16	5	17.8	23.6	1.14	1.23	0	44.7
con2	5	4177067–4178013	948	6	19.9	26.9	1.26	1.29	4	13.3	11.1	−0.53	1.16	0	40.0
062	6	13370671–13371601	930	1	3.6	5.7	1.38	0.75	5	16.9	17.7	0.17	0.51	9	122.8
063	6	13335122–13336060	940	0	0.0	0.0	0.00	0.00	7	23.5	32.5	1.41	0.77	7	93.5
071	7	34198101–34199051	951	6	19.8	14.9	−0.89	−0.61	8	26.5	29.5	0.43	1.37	9	150.0
072	7	1485242–1486253	1,012	2	6.2	5.2	−0.44	0.94	1	3.1	4.9	1.21	0.72	11	117.9
073	7	14736902–14737811	902	15	55.1	90.7	2.78	1.51	7	24.2	30.1	0.90	0.21	4	148.9
074	7	13161241–13162059	818	1	3.8	1.7	−1.16	−1.40	10	38.4	65.4	2.73	1.42	1	71.6
con3	7	12877447–12878480	1,034	3	9.1	12.4	0.07	0.30	4	12.2	10.5	−0.44	−0.55	1	26.2
102	10	54958829–54959659	812	0	0.0	0.0	0.00	0.00	1	3.8	1.7	−1.16	−1.40	9	111.7
131	13	4938296–4939215	920	2	6.8	10.5	1.46	0.94	3	10.3	13.3	0.90	0.02	9	117.2
con4	13	5177266–5178204	939	3	10.0	14.4	1.32	1.07	6	20.1	28.8	1.55	1.29	4	81.4
133	13	14942772–14943791	1,021	1	3.1	5.3	1.51	0.72	7	21.6	18.1	−0.60	1.33	4	60.9
con5	13	10340883–10341849	984	3	9.6	8.3	−0.42	−1.04	8	25.5	21.9	−0.53	0.86	6	86.4
con6	14	12999764–13000706	948	5	16.6	15.1	−0.32	0.51	2	6.6	4.3	−0.96	−0.45	1	39.9
con7	14	55603471–55604414	944	3	10.5	5.3	−1.63	−1.95	17	56.6	29.3	−1.99	−2.45	1	45.9
con8	17	1418876–1419796	921	5	17.1	13.2	−0.78	−0.95	5	17.1	13.6	−0.70	−0.22	2	42.7
173	17	27883797–27884681	885	2	7.1	6.6	−0.20	−0.45	1	3.6	5.1	0.84	0.72	15	181.6
con9	17	64305167–64306026	860	10	36.6	30.2	−0.68	−1.52	3	11.0	6.4	−1.28	−1.04	0	64.8
GGH	4	20166795–201667294	499	1	5.0	3.5	−0.34	0.72	4	18.3	19.3	0.00	0.00	0	2.9
MELK	4	2162811–2163729	918	0	0.0	0.0	0.00	0.00	14	50.6	75.0	2.06	1.20	5	116.2
NKD1	8	91403263–91404440	1,178	8	16.9	14.9	−0.87	−0.22	11	21.1	12.8	0.83	0.02	0	26.7
PUM1	4	130022004–130023187	1,183	4	8.8	2.2	0.00	0.00	5	10.0	3.9	−1.16	−1.40	0	7.9
SFRP1	8	2480594–2481312	1,071	1	3.0	5.0	1.43	0.72	12	30.1	24.3	−0.75	−0.68	10	122.0
Mean			936.4	3.5	11.7	12.6	0.17	0.04	160	20.7	21.2	0.13	0.21	4	80.4

NOTE.—[1]Most noncoding regions do not have standard names, and the names listed are the idendification numbers used in our laboratory.

[2] Nucleotide positions in the mouse genome sequence of National Center for Biotechnology Information Build 37.

[3] Length in *M. musculus*. Length in *M. domesticus* may be slightly different due to indels.

[4] Number of polymorphic sites.

[5] Watterson's polymorphism per site.

[6] Nucleotide diversity per site.

[7] Values significantly different from 0 at the 5% level are underlined. Significance is determined by 10,000 coalescent simulations.

[8] Number of fixed nucleotide differences between species.

[9] Mean number of nucleotide differences between species.

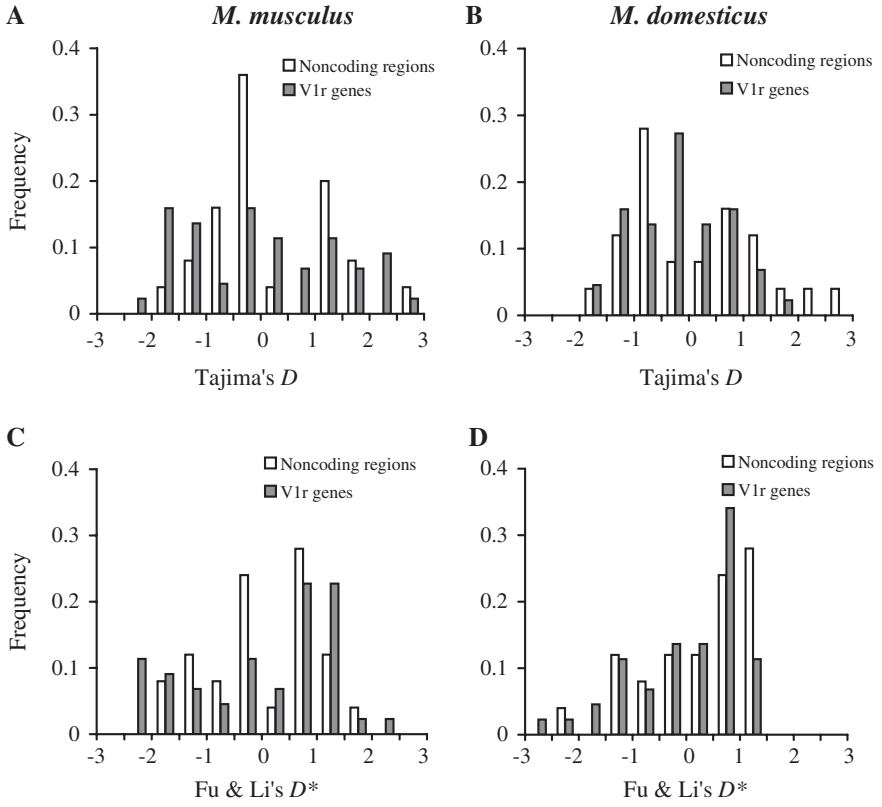

Fig. 2.—Frequency distributions of Tajima's *D* and Fu and Li's *D** among 44 *V1r*s and 25 noncoding regions in mouse. (A) Tajima's *D* in *M. musculus*; (B) Tajima's *D* in *M. domesticus*; (C) Fu and Li's *D** in *M. musculus*; and (D) Fu and Li's *D** in *M. domesticus*.

after extensive experimentation with multiple primer sets including those within coding regions, suggesting that the two genes may have been partially or entirely deleted in these individuals. We thus regard these cases as null alleles as well. In total, 14 *V1r*s harbor null alleles in *Mm* and 7 in *Md* (table 5). Given these numbers, we should expect $14 \times 7/44 = 2.23$ *V1r*s to have segregating null alleles in both species if *V1r* pseudogenization in the two species is independent. Consistent with this expectation, two *V1r*s harbor segregating null alleles in both species, and the pseudogenization events were independent because the null alleles in the two species were generated by different open reading frame (ORF)–disrupting mutations (table 5). In addition to the prevalence of pseudogenized *V1r*s, the frequencies of the null alleles are not particularly low (table 5), especially in *Mm*, suggesting the lack of strong selection preventing the null alleles from spreading through the populations. This finding is consistent with

Table 3
Numbers of Synonymous and Nonsynonymous Sequence Variations in *V1r*s

				Nonsynonymous/ synonymous		Nonsynonymous/ (noncoding + synonymous)	
	Nonsynonymous	Synonymous	Noncoding[1]	Ratio	*P* value[2]	Ratio	*P* value[2]
Polymorphisms in *M. musculus*	111	97	80	1.144	0.774	0.627	<0.001
Polymorphisms in *M. domesticus*	102	56	144	1.821	0.068	0.510	<0.001
High-frequency[3] polymorphisms in *M. musculus*	78	77	65	1.013	1.000	0.549	<0.001
High-frequency[3] polymorphisms in *M. domesticus*	62	44	104	1.409	0.330	0.419	0.01
Fixed differences between the two species	32	31	110	1.032		0.227	

Note.—[1]Variations in the 25 noncoding regions.
[2] *P* value is from Fisher's exact test in comparison with fixed differences.
[3] Frequency of the derived allele is equal to or greater than 2 of 14.

Table 4
Comparison Between *V1r*s and the 25 Noncoding Regions by the HKA Test

	M. musculus		*M. domesticus*	
***V1r* genes**	***d*/θ**[1]	***P* value**[2]	***d*/θ**[1]	***P* value**[2]
A1	1.95	2.6E-03	1.63	3.5E-01
A9	20.67	1.5E-04	13.78	1.4E-04
B1	2.11	5.0E-04	3.31	9.1E-01
B2	2.41	4.4E-03	2.76	8.5E-01
B3	2.35	1.0E-01	1.01	1.5E-02
B7	4.32	8.5E-01	2.16	8.7E-01
B8	17.95	1.6E-03	11.96	2.4E-03
C5	36.12	3.2E-10	12.04	2.1E-03
C6	5.13	3.3E-01	2.85	8.3E-01
C21	2.73	9.5E-01	1.36	9.9E-01
C27	1.35	2.9E-07	6.42	5.9E-01
C28	1.09	1.2E-07	1.96	6.9E-01
C30	4.44	7.7E-01	1.48	2.4E-01
D6	1.65	1.7E-03	2.20	9.6E-01
D7	12.61	7.0E-02	5.04	8.6E-01
D13	1.56	1.8E-01	0.45	1.7E-05
E1	1.05	5.9E-08	4.20	9.7E-01
E2	4.43	7.7E-01	4.43	9.5E-01
E3	0.76	4.5E-09	4.54	4.0E-01
E4	2.14	4.5E-03	3.56	9.8E-01
E7	1.97	4.2E-02	5.91	1.6E-01
E8	1.14	3.4E-05	0.91	2.8E-02
E11	1.66	1.9E-03	1.11	4.9E-02
F1	0.41	2.2E-11	1.02	9.5E-01
F2	0.37	1.7E-14	1.24	6.4E-01
F3	0.87	7.2E-10	12.83	1.5E-03
F4	1.38	1.8E-05	4.15	9.7E-01
F5	1.76	1.4E-03	14.08	4.2E-06
G1	2.99	1.2E-01	0.54	9.3E-06
G6	1.11	1.1E-08	3.16	9.0E-01
G9	1.56	2.5E-05	4.16	9.6E-01
G11	1.14	6.4E-04	1.14	5.0E-01
G12	2.40	3.2E-03	4.32	9.2E-01
H4	8.52	4.2E-01	3.41	9.4E-01
H15	4.69	5.2E-01	1.17	1.6E-02
H19	2.05	1.2E-02	2.05	8.4E-01
H20	1.23	3.2E-07	3.07	9.6E-01
I1	1.70	1.0E+00	0.34	3.6E-05
I6	4.09	4.7E-01	3.07	9.6E-01
I7	2.12	2.8E-03	4.24	9.6E-01
J2	1.48	2.2E-05	2.21	8.8E-01
J3	1.91	5.4E-04	13.40	1.8E-04
K1	1.29	1.4E-04	1.72	8.9E-01
L1	0.76	4.5E-09	0.76	3.0E-03
All *V1r* genes	2.59	1.5E-96	3.63	6.9E-05
25 noncoding regions	7.30		4.20	

NOTE.—[1]Interspecific divergence per site divided by Watterson's polymorphism per site. The actual number of polymorphc site plus 1 was used in calculating Watterson's polymorphism per site. Values of d/θ that are significantly greater than expected from the 25 noncoding regions are underlined.

[2] *P* values are from chi-squares tests.

an overall low purifying selection acting on *V1r*s and provides a microevolutionary explanation for the rapid gene turnover observed at the macroevolutionary time scale.

Table 5
Mouse *V1r*s With Segregating Null Alleles[1]

***V1r*s**	***M. musculus***	***M. domesticus***
A9	3 (deletion)	
B1	8 (deletion)	
B2		1 (deletion)
C6	4 (SNP)	
C27	5 (SNP) + 1 (deletion)	
C28	4 (no amplification)	1 (SNP)
E1	2 (deletion)	
F1	1 (deletion) + 2 (deletion)	
F3	10 (no amplification) + 2 (SNP)	
F5	1 (deletion)	
H4		1 (insertion)
H15		6 (SNP)
I6	2 (insertion)	1 (insertion) + 1 (SNP)
I7	2 (deletion)	1 (insertion)
J2	2 (SNP)	
J3	1 (SNP)	
K1	1 (SNP) + 2 (deletion)	
L1		1 (SNP)

NOTE.—Numbers in the table are the numbers of null alleles (of 14 per species) for *V1r*s that harbor null alleles in the species. The type of null mutations is also indicated.

Discussion

In this study, we characterized the intra- and interspecific sequence variations of 44 *V1r* genes and 25 noncoding regions in two closely related *Mus* species. Both intraspecific polymorphisms and interspecific divergences are generally reduced in *V1r*s compared with the noncoding regions, suggesting that the overall force in *V1r* evolution is purifying selection. The strength of the purifying selection, however, is relatively weak. This is reflected by a ratio of approximately 0.32 between the mean nonsynonymous substitution rate and mean synonymous substitution rate of the 44 *V1r*s. A similar ratio (0.34) is obtained when only fixed differences between *Mm* and *Md* are considered. In comparison, this ratio is on average ~0.11 when all 11,503 mouse–rat orthologous genes are compared. The Rat Genome Sequencing Project Consortium reported that 15% of mouse–rat orthologous genes have a d_N/d_S ratio greater than 0.28, but we found that 57% of *V1r*s belong to this category, the difference being highly significant ($P < 10^{-6}$, binomial test). The overall weak purifying selection is also reflected by the high fraction of *V1r* loci that have segregating null alleles in either one or both *Mus* species examined. It is likely that a sizable proportion of *V1r* genes are not functionally constrained in each mouse species, consistent with the previous observation of virtually neutral variations of *V1r* gene copy number within and between species (Nozawa et al. 2007; Zhang 2007). In other words, the seemingly rapid *V1r* gene turnover is at least in part caused by neutral genomic drift (Nozawa et al. 2007). We did not find any *V1r* that has been duplicated in one of the two *Mus* species since their

separation, but this is attributable to our intentional avoidance of studying *V1r*s with closely related paralogs to ease gene-specific amplification. In fact, in our preliminary study, one gene (E13) appeared to be duplicated in some individuals, but it was removed from the subsequent study due to the difficulty in designing gene-specific primers. Our findings of weak purifying selection in the microevolution of *V1r*s is generally consistent with a recent interspecific comparison of a subset of *V1r*s between the laboratory mouse and *M. spretus* (Kurzweil et al. 2009), which diverged from each other much earlier than the separation between *Mm* and *Md*. Kurzweil et al. sequenced a genomic segment of *M. spretus* that harbors two subfamilies of *V1r*s, including 18 genes. They observed that 1 of 11 genes in subfamily a and 2 of 7 genes in subfamily b have been lost in *M. spretus*, whereas 2 genes in subfamily a are becoming pseudogenes in *M. musculus*. They also identified two orthologous pairs that have been subject to positive selection. However, they did not analyze population genetic dynamics of *V1r*s because no intraspecific polymorphism data were collected. Our data thus complement theirs in providing necessary information for inferring the microevolutionary forces acting on *V1r*s.

It should be noted that, across mammals, there is a strong positive correlation between the morphological (and presumably physiological) complexity of the VNS in a species and the number of intact *V1r* genes the species has (Grus et al. 2005, 2007). Furthermore, during the evolutionary transition of vertebrates from water to land, there was a ~50-fold increase in the ratio of the number of V1rs, which likely bind to airborne ligands, to that of V2rs, which likely bind to water-soluble ligands (Shi and Zhang 2007). Thus, it is likely that the evolution of the *V1r* repertoire is also subject to positive selection. Indeed, using the HKA test, we detected positive selection for nucleotide substitutions in ~5% of the mouse *V1r*s surveyed after controlling for multiple testing. Extrapolating this result to all *V1r*s suggests that there are ~10 *V1r*s that have adaptive differences between the two *Mus* species compared. Although we have no knowledge about the number of pheromonal differences between the two species, it is not unlikely that they differ by no more than a dozen pheromones. Recently, Karn et al. suggested that a particular *V1r* gene, *Vmn1r67* (or E10 in our nomenclature), experienced adaptive divergences among *Mus* species and that it might be responsible for detecting the androgen-binding protein, a species-specific cue for species recognition (Karn et al. 2010). Although this gene is not included in our 44 *V1r*s, positive selection on this gene would not be inconsistent with our estimate of ~5% of positively selected *V1r*s in the divergence between *Mm* and *Md*. In the future, it will be interesting to confirm their finding in the wild mice used here. An earlier paper compared two assemblies of mouse genome sequences, which were acquired from different mouse inbred lines, and reported an overall d_N/d_S ratio between the two assemblies to be 1.13 for *V1r*s (Zhang et al. 2004). The authors interpreted this finding as evidence for positive selection acting on most *V1r*s without actually testing whether the d_N/d_S ratio is significantly greater than 1, which is required for establishing positive selection. In our study, we found the mean d_N/d_S ratio for polymorphisms to be ~0.5 when all *V1r* polymorphisms in the 14 mice sequenced here are considered. It remains to be seen whether this disparity is due to the difference between the genes we sampled and the rest of *V1r*s, sequencing errors in draft genome sequences, or the differences between the wild mice and inbred laboratory mice. Furthermore, it will be important to reexamine the polymorphism and divergence of mouse *V1r*s at sites important for binding to ligands when such sites are identified.

It is important to note that the sample size (seven mice per species) is relatively small in our study, making population genetic tests of positive selection less powerful. Nonetheless, multiple observations from our data are consistent with one another in supporting the conclusion that weak purifying selection is the predominant force in mouse *V1r* evolution. Furthermore, the McDonald–Kreitman test also strongly rejects the strict neutrality in support of purifying selection rather than positive selection. Thus, it is unlikely that the observed paucity of positive selection in *V1r*s is an artifact of our small sample. We focused our analysis exclusively on coding regions rather than on regulatory regions because all *V1r*s are expressed in the VNS and are not expected to have important evolutionary changes in gene regulation. A previous analysis of the promoter regions of *V1r*s supports this view (Stewart and Lane 2007). One limitation of our study is that we focused exclusively on point mutations and small indels, whereasthe macroevolution of *V1r*s is also known to be characterized by gene duplication, deletion, and possibly gene conversion. In the future, it would be especially interesting to examine the dynamics of copy number variations for *V1r*s as has been analyzed for odorant receptor genes (Nozawa et al. 2007; Zhang 2007; Hasin et al. 2008; Young et al. 2008).

Evolutionary changes of large gene families appear to contribute disproportionately to genomic evolution because several of the largest gene families in eukaryotic genomes evolve rapidly (Shiu et al. 2004; Nei and Rooney 2005; Niimura 2009; Shi and Zhang 2009). It will be interesting to study whether our findings on the microevolution of *V1r*s extend to other large gene families.

Supplementary Material

Supplementary tables S1–S3 are available at *Genome Biology and Evolution* online (http://www.oxfordjournals.org/our_journals/gbe/).

Acknowledgments

We thank John Baines and Bettina Harr at University of Cologne for the primer sequences used for amplifying some

noncoding regions. We thank Soochin Cho, Nathan Pearson, and two anonymous reviewers for valuable comments. This work was supported by a research grant from the U.S. National Institutes of Health. S.H.P. was in part supported by a visiting fellowship from Korea University.

Literature Cited

Alibert P, FelClair F, Manolakou K, BrittonDavidian J, Auffray JC. 1997. Developmental stability, fitness, and trait size in laboratory hybrids between European subspecies of the house mouse. Evolution. 51:1284–1295.

Baines JF, Harr B. 2007. Reduced X-linked diversity in derived populations of house mice. Genetics. 175:1911–1921.

Brennan PA, Keverne EB. 2004. Something in the air? New insights into mammalian pheromones. Curr Biol. 14:R81–R89.

Britton-Davidian J, Fel-Clair F, Lopez J, Alibert P, Boursot P. 2005. Postzygotic isolation between the two European subspecies of the house mouse: estimates from fertility patterns in wild and laboratory-bred hybrids. Biol J Linn Soc Lond. 84:379–393.

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

Forejt J. 1996. Hybrid sterility in the mouse. Trends Genet. 12:412–417.

Forejt J, Ivanyi P. 1974. Genetic studies on male sterility of hybrids between laboratory and wild mice (Mus musculus L.). Genet Res. 24:189–206.

Frazer KA, et al. 2007. A sequence-based variation map of 8.27 million SNPs in inbred mouse strains. Nature. 448:1050–1053.

Fu YX, Li WH. 1993. Statistical tests of neutrality of mutations. Genetics 133:693–709.

Gibbs RA, et al. 2004. Genome sequence of the Brown Norway rat yields insights into mammalian evolution. Nature. 428:493–521.

Grus WE, Shi P, Zhang J. 2007. Largest vertebrate vomeronasal type 1 receptor gene repertoire in the semiaquatic platypus. Mol Biol Evol. 24:2153–2157.

Grus WE, Shi P, Zhang YP, Zhang J. 2005. Dramatic variation of the vomeronasal pheromone receptor gene repertoire among five orders of placental and marsupial mammals. Proc Natl Acad Sci U S A. 102:5767–5772.

Grus WE, Zhang J. 2004. Rapid turnover and species-specificity of vomeronasal pheromone receptor genes in mice and rats. Gene. 340:303–312.

Grus WE, Zhang J. 2006. Origin and evolution of the vertebrate vomeronasal system viewed through system-specific genes. BioEssays. 28:709–718.

Grus WE, Zhang J. 2008. Distinct evolutionary patterns between chemoreceptors of 2 vertebrate olfactory systems and the differential tuning hypothesis. Mol Biol Evol. 25:1593–1601.

Grus WE, Zhang J. 2009. Origin of the genetic components of the vomeronasal system in the common ancestor of all extant vertebrates. Mol Biol Evol. 26:407–419.

Guenet JL, Bonhomme F. 2003. Wild mice: an ever-increasing contribution to a popular mammalian model. Trends Genet 19:24–31.

Hasin Y, et al. 2008. High-resolution copy-number variation map reflects human olfactory receptor diversity and evolution. PLoS Genet. 4:e1000249.

Hudson RR, Kreitman M, Aguade M. 1987. A test of neutral molecular evolution based on nucleotide data. Genetics. 116:153–159.

Karn RC, Young JM, Laukaitis CM. 2010. A candidate subspecies discrimination system involving a vomeronasal receptor gene with different alleles fixed in M. m. domesticus and M. m. musculus. PLoS One. 5:e12638.

Kurzweil VC, Getman M, Green ED, Lane RP. 2009. Dynamic evolution of V1R putative pheromone receptors between Mus musculus and Mus spretus. BMC Genomics. 10:74.

Laukaitis CM, Critser ES, Karn RC. 1997. Salivary androgen-binding protein (ABP) mediates sexual isolation in Mus musculus. Evolution. 51:2000–2005.

Librado P, Rozas J. 2009. DnaSP v5: a software for comprehensive analysis of DNA polymorphism data. Bioinformatics. 25:1451–1452.

McDonald JH, Kreitman M. 1991. Adaptive protein evolution at the Adh locus in Drosophila. Nature. 351:652–654.

Mombaerts P. 1999. Molecular biology of odorant receptors in vertebrates. Annu Rev Neurosci. 22:487–509.

Mombaerts P. 2004. Genes and ligands for odorant, vomeronasal and taste receptors. Nat Rev Neurosci. 5:263–278.

Moulia C, et al. 1991. Wormy mice in a hybrid zone—a genetic-control of susceptibility to parasite infection. J Evol Biol. 4:679–687.

Moulia C, Le Brun N, Dallas J, Orth A, Renaud F. 1993. Experimental evidence of genetic determinism in high susceptibility to intestinal pinworm infection in mice: a hybrid zone model. Parasitology. 106(Pt 4):387–393.

Nei M, Rooney AP. 2005. Concerted and birth-and-death evolution of multigene families. Annu Rev Genet. 39:121–152.

Niimura Y. 2009. On the origin and evolution of vertebrate olfactory receptor genes: comparative genome analysis among 23 chordate species. Genome Biol Evol. 1:34–44.

Nozawa M, Kawahara Y, Nei M. 2007. Genomic drift and copy number variation of sensory receptor genes in humans. Proc Natl Acad Sci U S A. 104:20421–20426.

Restrepo D, Arellano J, Oliva AM, Schaefer ML, Lin W. 2004. Emerging views on the distinct but related roles of the main and accessory olfactory systems in responsiveness to chemosensory signals in mice. Horm Behav. 46:247–256.

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.

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

Sage R, Atchley W, Capanna E. 1993. House mice as models in systematic biology. Syst Biol. 42:523–561.

Sage RD, Heyneman D, Lim KC, Wilson AC. 1986. Wormy mice in a hybrid zone. Nature. 324:60–63.

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

Salcedo T, Geraldes A, Nachman MW. 2007. Nucleotide variation in wild and inbred mice. Genetics. 177:2277–2291.

Shi P, Bielawski JP, Yang H, Zhang YP. 2005. Adaptive diversification of vomeronasal receptor 1 genes in rodents. J Mol Evol. 60:566–576.

Shi P, Zhang J. 2007. Comparative genomic analysis identifies an evolutionary shift of vomeronasal receptor gene repertoires in the vertebrate transition from water to land. Genome Res. 17:166–174.

Shi P, Zhang J. 2009. Extraordinary diversity of chemosensory receptor gene repertoires among vertebrates. In: Meyerhof W, Korsching S, editors. Chemosensory systems in mammals, fishes, and insects. Berlin (Germany): Springer. pp. 1–23.

Shiu SH, Karlowski WM, Pan R, Tzeng YH, Mayer KF, Li WH. 2004. Comparative analysis of the receptor-like kinase family in Arabidopsis and rice. Plant Cell. 16:1220–1234.

Smadja C, Catalan J, Ganem G. 2004. Strong premating divergence in a unimodal hybrid zone between two subspecies of the house mouse. J Evol Biol. 17:165–176.

Smadja C, Ganem G. 2002. Subspecies recognition in the house mouse: a study of two populations from the border of a hybrid zone. Behav Ecol. 13:312–320.

Spehr M, Spehr J, Ukhanov K, Kelliher KR, Leinders-Zufall T, Zufall F. 2006. Parallel processing of social signals by the mammalian main and accessory olfactory systems. Cell Mol Life Sci. 63:1476–1484.

Stewart R, Lane RP. 2007. V1R promoters are well conserved and exhibit common putative regulatory motifs. BMC Genomics. 8:253.

Storchova R, Gregorova S, Buckiova D, Kyselova V, Divina P, Forejt J. 2004. Genetic analysis of X-linked hybrid sterility in the house mouse. Mamm Genome. 15:515–524.

Tajima F. 1989. Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. Genetics. 123:585–595.

Takami S. 2002. Recent progress in the neurobiology of the vomeronasal organ. Microsc Res Tech. 58:228–250.

Tamura K, Dudley J, Nei M, Kumar S. 2007. MEGA4: Molecular Evolutionary Genetics Analysis (MEGA) software version 4.0. Mol Biol Evol. 24:1596–1599.

Trachtulec Z, Mihola O, Vlcek C, Himmelbauer H, Paces V, Forejt J. 2005. Positional cloning of the Hybrid sterility 1 gene: fine genetic mapping and evaluation of two candidate genes. Biol J Linn Soc. 84:637–641.

Tucker PK. 2007. Systematics of the genus *Mus*. In: Fox J, et al., editors. Mouse in Biomedical Research, 2nd ed. Boston: Elsevier Press. p. 13–23.

Vyskocilova M, Trachtulec Z, Forejtv J, Pialek J. 2005. Does geography matter in hybrid sterility in house mice? Biol J Linn Soc Lond. 84:663–674.

Yang H, Bell TA, Churchill GA, Pardo-Manuel de Villena F. 2007. On the subspecific origin of the laboratory mouse. Nat Genet. 39:1100–1107.

Yang H, Shi P, Zhang YP, Zhang J. 2005. Composition and evolution of the V2r vomeronasal receptor gene repertoire in mice and rats. Genomics. 86:306–315.

Young JM, Endicott RM, Parghi SS, Walker M, Kidd JM, Trask BJ. 2008. Extensive copy-number variation of the human olfactory receptor gene family. Am J Hum Genet. 83:228–242.

Young JM, Kambere M, Trask BJ, Lane RP. 2005. Divergent V1R repertoires in five species: amplification in rodents, decimation in primates, and a surprisingly small repertoire in dogs. Genome Res. 15:231–240.

Young JM, Massa HF, Hsu L, Trask BJ. 2010. Extreme variability among mammalian V1R gene families. Genome Res. 20:10–18.

Young JM, Trask BJ. 2007. V2R gene families degenerated in primates, dog and cow, but expanded in opossum. Trends Genet. 23:212–215.

Zhang J. 2007. The drifting human genome. Proc Natl Acad Sci U S A. 104:20147–20148.

Zhang J, Kumar S, Nei M. 1997. Small-sample tests of episodic adaptive evolution: a case study of primate lysozymes. Mol Biol Evol. 14:1335–1338.

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

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

Zhao H, Xu D, Zhang S, Zhang J. 2011. Widespread losses of vomeronasal signal transduction in bats. Mol Biol Evol. 28:7–12.

Associate editor: Bill Martin