Adaptive Diversification of Bitter Taste Receptor Genes in Mammalian Evolution

Peng Shi,† Jianzhi Zhang,‡ Hui Yang,*† and Ya-ping Zhang*§*

*Laboratory of Molecular Evolution and Genome Diversity, Kunming Institute of Zoology, Chinese Academy of Sciences, Kunming, China; †The Graduate School, Chinese Academy of Sciences, Beijing, China; ‡Department of Ecology and Evolutionary Biology, University of Michigan; and §Laboratory for Conservation and Utilization of Bio-resources, Yunnan University, Kunming, China

The diversity and evolution of bitter taste perception in mammals is not well understood. Recent discoveries of bitter taste receptor (T2R) genes provide an opportunity for a genetic approach to this question. We here report the identification of 10 and 30 putative T2R genes from the draft human and mouse genome sequences, respectively, in addition to the 23 and 6 previously known T2R genes from the two species. A phylogenetic analysis of the T2R genes suggests that they can be classified into three main groups, which are designated A, B, and C. Interestingly, while the one-to-one gene orthology between the human and mouse is common to group B and C genes, group A genes show a pattern of species- or lineage-specific duplication. It is possible that group B and C genes are necessary for detecting bitter tastants common to both humans and mice, whereas group A genes are used for species-specific bitter tastants. The analysis also reveals that phylogenetically closely related T2R genes are close in their chromosomal locations, demonstrating tandem gene duplication as the primary source of new T2Rs. For closely related paralogous genes, a rate of nonsynonymous nucleotide substitution significantly higher than the rate of synonymous substitution was observed in the extracellular regions of T2Rs, which are presumably involved in tastant-binding. This suggests the role of positive selection in the diversification of newly duplicated T2R genes. Because many natural poisonous substances are bitter, we conjecture that the mammalian T2R genes are under diversifying selection for the ability to recognize a diverse array of poisons that the organisms may encounter in exploring new habitats and diets.

Introduction

Mammals can perceive five major tastes: sweet, sour, bitter, salty, and umami (Kinnamon and Cummings 1992; Lindemann 1996; Stewart, Desimone, and Hill 1997; Chaudhari, Landin, and Roper 2000; Lindemann 2000). The ability to distinguish bitter-tasting substances is particularly important, as it enables us and other mammals to avoid potentially deadly environmental toxins (Garcia and Hankins 1975; Glendinning, 1994; Glendinning, Tarre, and Asaoka 1999; Chandrashekar et al. 2000). It has been widely believed that the sensation of bitter taste is initiated by the interaction of tastants with G protein–coupled receptors (GPCRs) in the membrane of taste receptor cells (Wong, Gannon, and Margolskee 1996). Two research groups recently identified putative bitter taste receptor genes from the human and mouse and named them T2R or TRB genes (Matsunami, Montmayeur, and Buck 2000; Adler et al. 2000). These candidate receptors have seven transmembrane domains and conserved amino acid residues that are often seen in GPCRs. Different from the putative sweet taste receptors (T1Rs), which have a large N-terminal domain, the bitter taste receptors possess only a short extracellular N-terminus. T2Rs display 30%~70% sequence identity among themselves. They also have highly conserved sequence motifs in the first, second, third, and seventh transmembrane domains and in the second intracellular loop (Adler et al. 2000). The most divergent parts in T2R sequences are the extracellular regions, which potentially bind tastants (Adler et al. 2000; Gilbertson, Damak, and Margolskee 2000). As in many other GPCR genes, there are no introns breaking the coding sequence of T2R genes. Although only four T2Rs have been functionally characterized and were shown to respond to bitter tastants (Chandrashekar et al. 2000; Bufe et al. 2002), substantial evidence is available for the role of other putative T2Rs in bitter taste perception (Adler et al. 2000). For instance, the T2R genes of humans and mice are organized in clusters in chromosomes, and are genetically linked to loci associated with responses to various bitter compounds. The identified human T2R genes are in chromosomes 12p13, 7q31, and 5p15, which are homologous to mouse chromosomes 6 and 15 (Matsunami, Montmayeur, and Buck 2000; Adler et al. 2000). The putative T2R gene ht2r1 (at 5p15) of humans is linked to genetic loci associated with the response to the bitter substance 6-n-propyl-2-thiouracil (PROP; Reed et al. 1999). The same is true for the T2R genes at human 7q31 (Adler et al. 2000; Matsunami, Montmayeur, and Buck 2000). Notably, the human T2R gene cluster at 12p13 contains six salivary proline-rich protein (PRP) genes (Azen, Lush, and Taylor 1986), which are closely linked to four loci (Soa, Rua, Cyx, and Qui) that are known to influence bitter perception in mice (Lush 1984; Lush 1986; Lush and Holland 1988; Capeless, Whitney, and Azen 1992; Adler et al. 2000).

Much information is available on the electrophysiological, biochemical, genetic, and functional aspects of bitter taste receptors. However, little is known about the evolution of these proteins. In the present study, we report the nearly complete repertoires of human and mouse T2R genes and conduct an evolutionary analysis of these genes.

Key words: evolution, bitter taste receptor, positive selection, gene duplication.

E-mail: zhangyp1@263.net.cn.

Mol. Biol. Evol. 20(5):805–814. 2003
DOI: 10.1093/molbev/msg083

Methods

Data Mining and Evolutionary Analyses

The following T2R genes of the mouse were retrieved from the GenBank: mT2R5 (AF227147), mT2R8

(AF227148), mT2R19 (AF227149), mTRB1 (AF247731), mTRB2 (AF247732), mTRB3 (AF247733), mTRB4 (AF247734), mTRB5 (AF247735). The following human T2R genes were similarly obtained: hT2R1 (AF227129), hT2R3 (AF227130), hT2R4 (AF227131), hT2R5 (AF227132), hT2R7 (AF227133), hT2R8 (AF227134), hT2R9 (AF227135), hT2R10 (AF227136), hT2R13 (AF227137), hT2R14 (AF227138), hT2R16 (AF227139). Additional T2R genes were obtained by screening the human and mouse genome sequences of the February and June 2002 assembly (http://genome.ucsc.edu) and July 2002 assembly (http://www.ensembl.org), respectively, using programs BLASTN or TBLASTN (Altschul et al. 1997). Because both research groups that identified bitter taste receptor genes now agree that these receptors should be called T2Rs (Montmayeur et al. 2001; Nelson et al. 2001), we follow the same nomenclature here.

T2R amino acid sequences were aligned by ClustalW (Thompson, Higgins, and Gibson 1994) with manual adjustments. The nucleotide sequences were then aligned following the amino acid sequence alignment and were used in tree reconstruction. The mouse V1Rd8 and V1Re9 genes, members of the type 1 vomeronasal pheromone receptor (V1R) gene family, were used as outgroups, because among GPCRs, V1R genes are relatively close to T2R genes (our unpublished data). Phylogenetic analysis was conducted using MEGA2 (Kumar et al. 2001). The reliability of the trees obtained was evaluated by the bootstrap method (Felsenstein 1985) with 1,000 replications. Sawyer's (1989) method as implemented in the computer program of Drouin et al. (1999) was used to examine gene conversion among paralogous human and mouse genes. To examine the pattern of nucleotide substitution in different regions of the gene, the numbers of synonymous substitutions per synonymous site (d_S) and the number of nonsynonymous substitutions per nonsynonymous site (d_N) were estimated using modified Nei-Gojobori method (Zhang, Rosenberg, and Nei 1998). Furthermore, we used the likelihood method of Yang et al. (2000) to test positive selection. This method involves two steps. The first is to use the likelihood ratio test (LRT) to test positive selection. This test is performed by comparing a null model with a more general one. The null model does not allow for sites with $\omega(d_N/d_S) > 1$, but the more general one does. Here we used two LRTs. The first LRT compares M0 (null model) with M3 (general model) for variation of ω among sites. M0 assumes one ω for all codons, whereas M3 assumes a discrete distribution of ω with three site classes. The second LRT compares M7 (null model) with M8 (general model) for the presence of sites under positive selection. M7 assumes that ω ratios were distributed among sites by a beta distribution; M8, an extension of M7, adds a discrete ω class to M7. Thus, this model allows for codons with $\omega > 1$. In general, positive selection may be inferred when the ω value estimated under M3 or M8 is greater than 1. The second step is to identify residues under positive selection when the LRT suggests its presence. This step is fulfilled by using the Bayes theorem to calculate the posterior probability that a site has $\omega > 1$. The computation was done by using the PAML software package (Yang et al. 2000). We are aware of a recent report that the likelihood method may be liberal in detecting positive selection (Suzuki and Nei 2001). Therefore, to reach reliable conclusions, we use both the likelihood method and the conventional method in comparing d_S and d_N.

Results

Complete T2R Gene Repertoires of the Human and Mouse

It is relatively easy to identify T2R genes through a computational approach because they are ''intronless.'' Based on the known T2R sequences, we searched the mouse draft genome sequence for new T2R genes, using BlastN or TBlastN (Altschul et al. 1997). Putative T2R genes were determined on the basis of high BLAST E-values and the presence of ~900 bp open reading frames (ORFs). Thirty-six genes were identified, including six that had been reported before. Among these sequences, three ORFs are interrupted by stop codons, and they are regarded as pseudogenes. The putative T2R genes are organized in the genome in three clusters: two major clusters on mouse chromosome 6, which are homologous to human chromosomes 12p13 and 7q31, respectively, and a minor cluster on mouse chromosome 15, which is syntenic to human chromosome 5. In fact, this minor cluster has only one T2R gene. We also searched the draft human genome sequence and found 33 putative T2R genes, including 23 genes that had been reported previously. Of the 10 new genes, two are putatively functional, one in chromosome 12 and the other in chromosome 7. The other 8 newly identified genes are pseudogenes.

Thus, there are 25 and 33 putatively functional T2R genes in the human and mouse, respectively. All of these genes, as well as T2R pseudogenes of the human and mouse, are listed in tables 1 and 2. Figure 1 shows the amino acid sequence alignment of the 58 putatively functional T2Rs, and that alignment confirms the earlier observation that the extracellular regions exhibit the highest sequence variability.

Evolutionary Relationships of Functional T2R Genes from Humans and Mice

To clarify the evolutionary relationships among the T2R genes, a phylogenetic tree of the 58 putatively functional genes from the human and mouse was reconstructed using the Neighbor-Joining method (Saitou and Nei 1987) (fig. 2). The tree shows that the T2R genes may be classified into three major groups, A, B, and C. This grouping, however, is tentative, as the bootstrap supports for the groups are low. But the available genomic information, as presented below, strongly supports this grouping. Group A genes of humans are located on chromosome 12 and are linked to the PRP loci, which are known to influence bitter taste perception. All of the mouse group A genes are located on chromosome 6, which is homologous to human chromosome 12. The human group C genes are located on chromosomes 5 and 7, and are linked to genetic loci associated with the ability to respond to the bitter substance PROP. The mouse group

Table 1
Human T2R Genes

Gene	Start	End	Chromosome	Other Names	References	Accession Numbers
ht2r7	10659234	10658278	Chr12	hTAS2R7; hTRB4	Adler et al. 2000	AF227133
ht2r8	10663644	10662715	Chr12	hTAS2R8; hTRB5	Matsunami et al. 2000	AF227134
ht2r9	10666739	10665801	Chr12	hTAS2R9; hTRB6		AF227135
ht2r10	10682933	10682010	Chr12	hTAS2R10; hTRB2		AF227136
ht2r13	10765962	10765051	Chr12	hTAS2R13; hTRB3		AF227137
ht2r14	10795886	10794918	Chr12	hTAS2R14; hTRB1		AF227138
hps8	10822016	10821246	Chr12		This paper	AY168289
ht2r50	10843524	10842625	Chr12	hTAS2R50	Bufe et al. 2002	AF494235
ht2r49	10854539	10853610	Chr12	hTAS2R49	Bufe et al. 2002	AF494236
ht2r48	10879292	10878336	Chr12	hTAS2R48	Bufe et al. 2002	AF494234
ht2r44	10887999	10887070	Chr12	hTAS2R44	Bufe et al. 2002	AF494228
ht2r47	10929992	10929033	Chr12	hTAS2R47	Bufe et al. 2002	AF494233
hps4	10954513	10955429	Chr12		This paper	AY168285
hps6	10982585	10983481	Chr12		This paper	AY168287
ht2r46	10996519	10995590	Chr12	hTAS2R46	Bufe et al. 2002	AF494227
hps2	11012465	11011707	Chr12		This paper	AY168283
ht2r43	11026452	11025523	Chr12	hTAS2R43	Bufe et al. 2002	AF494237
ht2r55	11220284	11221306	Chr12		This paper	AY161925
hps5	11226841	11227785	Chr12		This paper	AY168286
hps7	11240917	11241816	Chr12		This paper	AY168288
ht2r45	11255241	11254342	Chr12	hTAS2R45	Bufe et al. 2002	AF494226
ht2r16	121113518	121112643	Chr7	hTAS2R16	Adler et al. 2000	AF227139
ht2r3	139734689	139735639	Chr7	hTAS2R3		AF227130
ht2r4	139749019	139749918	Chr7	hTAS2R4		AF227131
hps3	139758344	139759180	Chr7		This paper	AY168284
ht2r5	139760892	139761791	Chr7	hTAS2R5	Adler et al. 2000	AF227132
ht2r38	139943217	139944218	Chr7	hTAS2R38	Bufe et al. 2002	AF494231
ht2r39	141206354	141207370	Chr7	hTAS2R39	Bufe et al. 2002	AF494230
ht2r40	141245014	14124985	Chr7	hTAS2R40	Bufe et al. 2002	AF494229
hps1	141459970	141460596	Chr7		This paper	AY168282
ht2r56	141466388	141467344	Chr7		This paper	AY161926
ht2r41	141500808	141501731	Chr7	hTAS2R41	Bufe et al. 2002	AF494232
ht2r1	9798165	9797266	Chr5	hTAS2R1; hTRB7	Adler et al. 2000; Matsunami et al. 2000	AF227129

NOTE.—ps represents pseudogenes; the starting and ending nucleotide positions are from the human June 2002 assembly (http://genome.ucsc.edu).

C genes are located on chromosomes 15 and 6, homologous to human chromosome 5 and 7. There are only two genes in group B, the putative orthologous pair of hT2R3 and mT2R41. The hT2R3 gene is located on 7p31, and its mouse orthologue is on chromosome 6, which is syntenic to human chromosome 7p31. Our phylogenetic tree shows that, in general, the human and mouse genes do not form two separate clusters. Rather, they intermingle. This suggests that many gene duplication events predated the separation of primates and rodents. For instance, almost every human gene in group C (except ht2r5) has a one-to-one orthologue from the mouse, and vice versa. It is possible that every gene of this group has a conserved function between humans and mice. In contrast, some genes from one species (human or mouse) cluster together to form species-specific clades in group A. For instance, figure 2 shows that 8 human genes (ht2r43, ht2r44, ht2r45, ht2r46, ht2r47, ht2r48, ht2r49, ht2r50) and 8 mouse genes (mt2r54, mt2r55, mt2r57, mt2r59, mt2r60, mt2r62, mt2r63, mt2r64) form two separate clusters. These genes are probably products of duplications after the primate-rodent divergence, and they may have species-specific functions that are distinct from those of group B and C genes. Our unpublished data suggest that species-specific duplications also occurred in other mammals.

Evolutionary Relationships of T2R Genes Within Species

To understand the evolutionary dynamics, we conducted a detailed analysis of the evolutionary relationships of T2R genes within species. The phylogenetic trees of human and mouse T2R genes are given in figures 3 and 4. A comparison between the phylogenetic tree and the chromosome map shows that genes that are phylogenetically closer are also physically closer on chromosomes (figs. 3 and 4). This pattern strongly suggests that new T2R genes were mainly generated by tandem gene duplication.

Our intraspecific gene trees also include pseudogenes. Unexpectedly, the pseudogenes do not evolve much faster than functional genes. One possible explanation is that the T2R pseudogenes may be subject to frequent gene conversion from functional genes. However, gene conversions among paralogous T2R genes are not detected by Sawyer's (1989) test for either the human or the mouse, suggesting that gene conversion may not have played a major role in our case. Another possible explanation is that the pseudogenes lost their functions very recently. This hypothesis may be tested when more primate and rodent species are examined. The third explanation is that functional T2R genes evolve as fast as pseudogenes due to either low functional constraints or positive selection (see below).

Table 2
Mouse T2R Genes

Gene	Start	End	Chromosome	Other Name	References	Accession Numbers
mt2r40	23917353	23918252	Chr6		This paper	AY161895
mt2r41	40689399	40690349	Chr6		This paper	AY161896
mt2r8	40691702	40692595	Chr6	mt2r8	Adler et al. 2000	AF227148
mt2r31	40812719	40813714	Chr6		This paper	AY161919
mt2r34	42338612	42339571	Chr6		This paper	AY161920
mt2r33	42418638	42419597	Chr6		This paper	AY161921
mt2r36	42603784	42604665	Chr6		This paper	AY161922
mt2r38	42608786	42609739	Chr6		This paper	AY161923
mt2r35	42637781	42638707	Chr6		This paper	AY161924
mt2r42	132230474	132231412	Chr6		This paper	AY161897
mt2r43	132259838	132260764	Chr6		This paper	AY161898
mt2r44	132278187	132279113	Chr6		This paper	AY161899
mt2r45	132285064	132285972	Chr6		This paper	AY161890
mt2r5	132286789	132287691	Chr6	mt2r5	Adler et al. 2000	AF227147
mt2r46	132289362	132290291	Chr6		This paper	AY161901
mt2r47	132996398	132997285	Chr6		This paper	AY161902
mt2r48	133039247	133040164	Chr6		This paper	AY161903
mt2r49	133076147	133077079	Chr6		This paper	AY161904
mt2r50	133093825	133094754	Chr6		This paper	AY161905
mt2r51	133101281	133102270	Chr6		This paper	AY161906
mt2r52	133116275	133117192	Chr6		This paper	AY161907
mps3	133125830	133126609	Chr6		This paper	AY168292
mt2r54	133142597	133143589	Chr6		This paper	AY161908
mt2r55	133224557	133225558	Chr6		This paper	AY161909
mt2r56	133232851	133233768	Chr6	mTRB1 (partial)	Matsunami et al. 2000	AF247731;AY161910
mt2r57	133245421	133246422	Chr6		This paper	AY161911
mt2r58	133270830	133271759	Chr6		This paper	AY161912
mt2r59	133287470	133288405	Chr6		This paper	AY161913
mps1	133315371	133316270	Chr6		This paper	AY168290
mt2r60	133332310	133333272	Chr6		This paper	AY161914
mt2r61	133338120	133339052	Chr6		This paper	AY161915
mt2r62	133361196	133362146	Chr6		This paper	AY161916
mps2	133382918	133383828	Chr6		This paper	AY168291
mt2r63	133405016	133405954	Chr6	mTRB2 (partial)	Matsunami et al. 2000	AF247732;AY161917
mt2r64	133423784	133424722	Chr6	mTRB3 (partial)	Matsunami et al. 2000	AF247733; AY161918
mt2r19	32273061	32274065	Chr15	mt2r19	Adler et al. 2000	AF227149

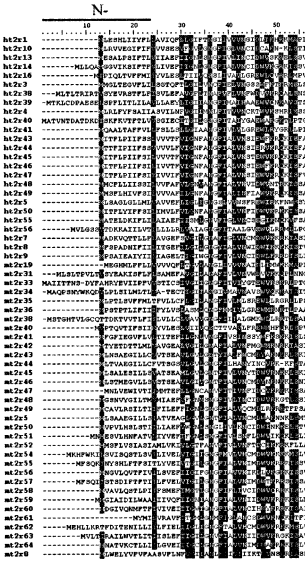
NOTE.—ps represents pseudogenes; the starting and ending nucleotide positions are from the mouse February 2002 assembly (http://genome.ucsc.edu).

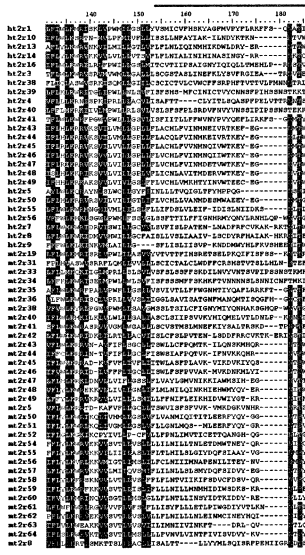
Positive Selection in the Extracellular Regions of T2R Genes

A comparison of the rates of synonymous (silent) and nonsynonymous (amino-acid-replacement) nucleotide substitutions may reveal the evolutionary forces shaping gene evolution (Nei and Kumar 2000). We computed the mean nonsynonymous distances (d_N) in extracellular regions (ER), transmembrane regions (TR), and intracellular regions (IR) of the human and mouse T2R genes and found that d_N is higher in ER than in TR and IR for all comparisons. We also compared the mean synonymous distance (d_S) and mean d_N for the three regions (table 3) and found the mean d_S in all three regions to be similar. Mean d_N is significantly smaller than mean d_S in both TR and IR ($P < 0.01$), but it is slightly greater than d_S in ER ($P > 0.1$). This suggests that while purifying selection dominates the evolution of TR and IR, positive selection may have operated in ER. To reveal the signal of positive selection, we first compared d_S and d_N in ER for the closely related paralogous genes of mouse clusters 1 and 2 (fig. 2). It is interesting that mouse cluster 2 genes all show $d_N < d_S$, but cluster 1 genes show $d_N > d_S$ when d_S is no more than 0.8 (fig. 5*a*). This suggests that mouse clusters 1 and 2 may have been under different selective pressures in addition to their distinct phylogenetic positions (fig. 2). Whereas the function of cluster 2 genes may be more conservative, cluster 1 genes may have been under positive selection and have plasticity in function. In particular, the d_N/d_S ratios for cluster 1 genes decline as d_S increases beyond 0.8, indicating that a saturation effect may exist in nonsynonymous substitutions (Tanaka and Nei 1989). In fact, at this level of divergence, it may be difficult to detect positive selection among duplicated genes simply because positive selection may only occur in a short evolutionary time after gene duplication during the functional shift of the protein, and its effect can be obscured by later substitutions (Zhang, Rosenberg, and Nei 1998). If this is the case, signs of positive selection should be most obvious for very recent duplications. The 8 genes of the human cluster (marked in fig. 2) are phylogenetically closely related. The d_S and d_N values for ER between each

→

FIG. 1.—Alignment of the complete sequences of all putatively functional T2Rs from the human and mouse. Residues highlighted are conserved in at least 75% of the sequences. Predicted extracellular regions, which are also putative ligand-binding regions, are indicated by bars above the sequences.





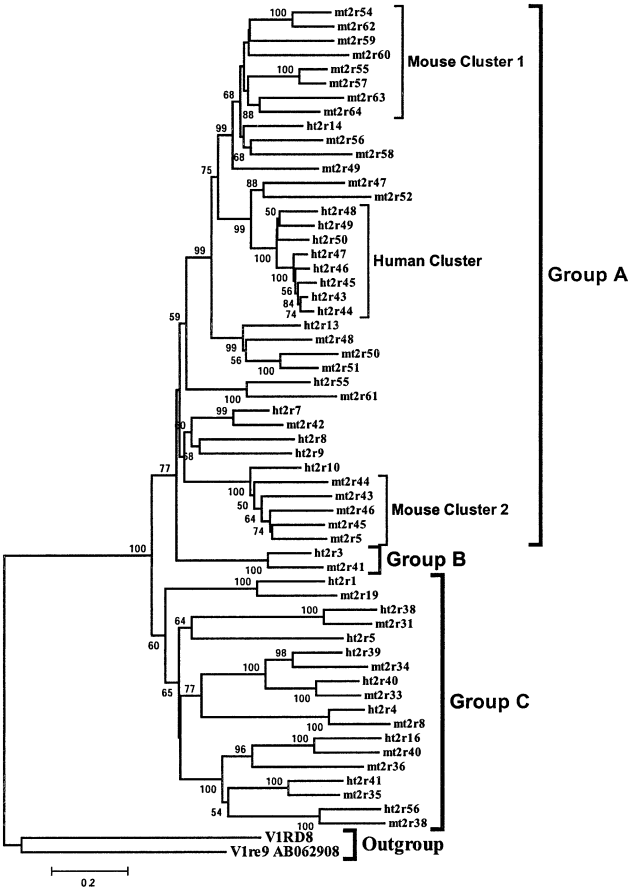


FIG. 2.—Neighbor-Joining tree of 58 putatively functional T2R genes from the human and mouse. After the removal of gaps, a total of 663 nucleotide sites are used in reconstructing the tree. Jukes-Cantor distances are used. Percentage bootstrap values (≥50) are shown on interior branches. The mouse V1Re9 and V1RD8 are used as the outgroup.

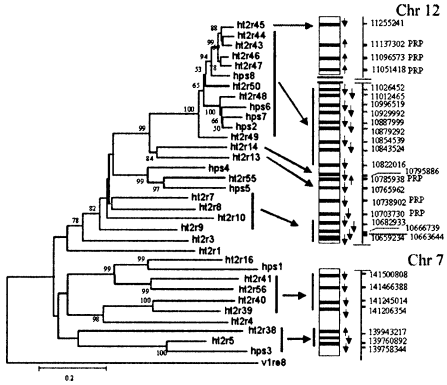


FIG. 3.—Phylogenetically closer T2R genes from humans are also closer in their chromosomal locations. Shown on the left is the phylogenetic tree for 33 T2R genes from humans. A total of 481 nucleotide sites are used in this Neighbor-Joining tree with Jukes-Cantor distances. Percentage bootstrap values (≥50) are shown. On the right are the regions of human chromosomes 12 and 7 that contain T2R genes. PRP refers to salivary proline-rich-protein genes (accession numbers M13058, M13057, S79048, XM_006909, XM_006910, and NM_002723). Arrows indicate the direction of transcription. Arrowheads between the tree and the map indicate that the phylogenetically closely related genes are in proximity in the chromosome. A color figure is available as online Supplementary Material.

pair of the 8 human genes are shown in figure 5*b*. Twenty-four of the 28 pairwise comparisons show $d_N > d_S$. The average d_N (0.232 ± 0.023) is significantly greater than the average d_S (0.169 ± 0.027) ($P < 5\%$, Z test). However, because the pairwise distances are not independent of each other, a phylogeny-based analysis (Zhang, Rosenberg, and Nei 1998) is preferred. In such an analysis, the ancestral sequences at interior nodes of the tree are inferred, and the numbers of synonymous (s) and nonsynonymous (n) substitutions on each tree branch are counted. To make sure that the ancestral sequences are accurately inferred, we used the bottom five sequences of the human cluster in the tree of figure 2, because they are very closely related, and theory predicts that the accuracy of ancestral inference is high with closely related sequences (Zhang and Nei 1997). This analysis shows that the total s and n values for the subtree of these 5 sequences are 11, and 56, respectively (fig. 5*d*). The potential numbers of synonymous (S) and nonsynonymous (N) sites of the sequences are 69.5 and 203.5, respectively. Fisher's test shows that $n/s = 5.1$ is significantly higher than $N/S = 2.9$ ($P = 0.031$). These results provide evidence for the operation of positive selection in the early divergence of paralogous T2R genes after duplication. Note that the above results on positive selection were all from group A genes. For group B and C genes, all except one orthologous human-mouse gene pair show $d_N < d_S$ (fig. 5*c*). For many pairs, the d_N/d_S ratio is lower than 0.5 (fig. 5*c*), which is rarely seen for group A genes (fig. 5*a* and *b*). Although the d_N/d_S ratios for group B and C genes are much lower than 1, they are still relatively high, in comparison to an average mammalian gene, which has a d_N/d_S of about 0.23 (Zhang 2000). This suggests that the extracellular regions of T2Rs are

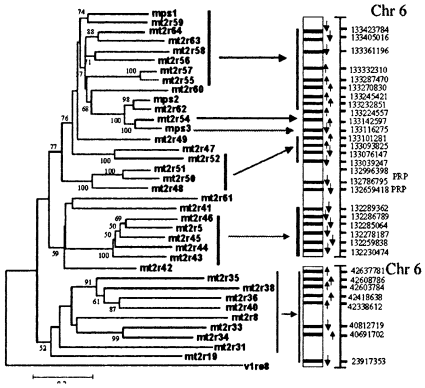


FIG. 4.—Phylogenetically closer T2R genes from the mouse are also closer in their chromosomal locations. Shown on the left is the phylogenetic tree for 36 T2R genes from the mouse. A total of 597 nucleotide sites are used in the Neighbor-Joining tree with Jukes-Cantor distances. On the right are the regions of mouse chromosome 6 that contain the T2R genes. Percentage bootstrap values (≥50) are shown. PRP refers to salivary proline-rich-protein genes (accession numbers BC011176, XM_162813). Arrows indicate the direction of transcription. Small arrows between the tree and the map indicate that phylogenetically closely related genes are in proximity in the chromosome. A color figure is available as online Supplementary Material.

Table 3
Mean Numbers of Nonsynonymous (d_N) and Synonymous (d_S) Substitutions per Site (±SE) in Different Regions of T2Rs

	Mouse			Human		
Region	d_N	d_S	p^a	d_N	d_S	p^a
ER	1.429 ± 0.085	1.426 ± 0.043	NS	1.489 ± 0.080	1.406 ± 0.061	NS
TR	0.613 ± 0.034	1.417 ± 0.040	**	0.563 ± 0.035	1.226 ± 0.057	**
IR	0.637 ± 0.065	1.156 ± 0.065	**	0.602 ± 0.056	1.182 ± 0.068	**
p^b	**	NS		**	NS	

[a] Comparison between d_S and d_N for the same region.
[b] Comparison of d_N between ER and the other two regions combined.
** $P < 0.01$.

generally not very conserved, probably because of the presence of some functionally less important sites.

Positive Selection at Individual Amino Acid Sites of T2R

In the above discussion, we showed that positive selection may have occurred in the ER of closely related T2Rs. To identify which amino acid positions may be under selection, we applied the maximum likelihood method of Yang and coworkers (Nielsen and Yang 1998; Yang et al. 2000). For this analysis, we used the human cluster and two mouse clusters of group A genes as marked in figure 2. Table 4 shows the results. Models M0 and M3 were compared and showed the LRT to be significant in all three clusters examined, suggesting that the selective pressure varies among amino acid sites for each cluster. In addition, the estimated additional ω ratios under M3 are all >1, indicating that positive selection may have operated in these clusters. In the comparison between M7 and M8, M8 fits the data significantly better than M7 in the human cluster and mouse cluster 1 but not in mouse cluster 2, although the estimated ω ratios of all three clusters are >1. The sites with posterior probabilities >95% under M8 are listed in table 4. These sites are similar to those estimated from M3.

We examined the distribution of the inferred positively selected sites (fig. 6). Heterogeneous distribution of positively selected sites was clear from the comparison of the proportion of positively selected sites in ER and that in the rest of T2R. If only the two statistically

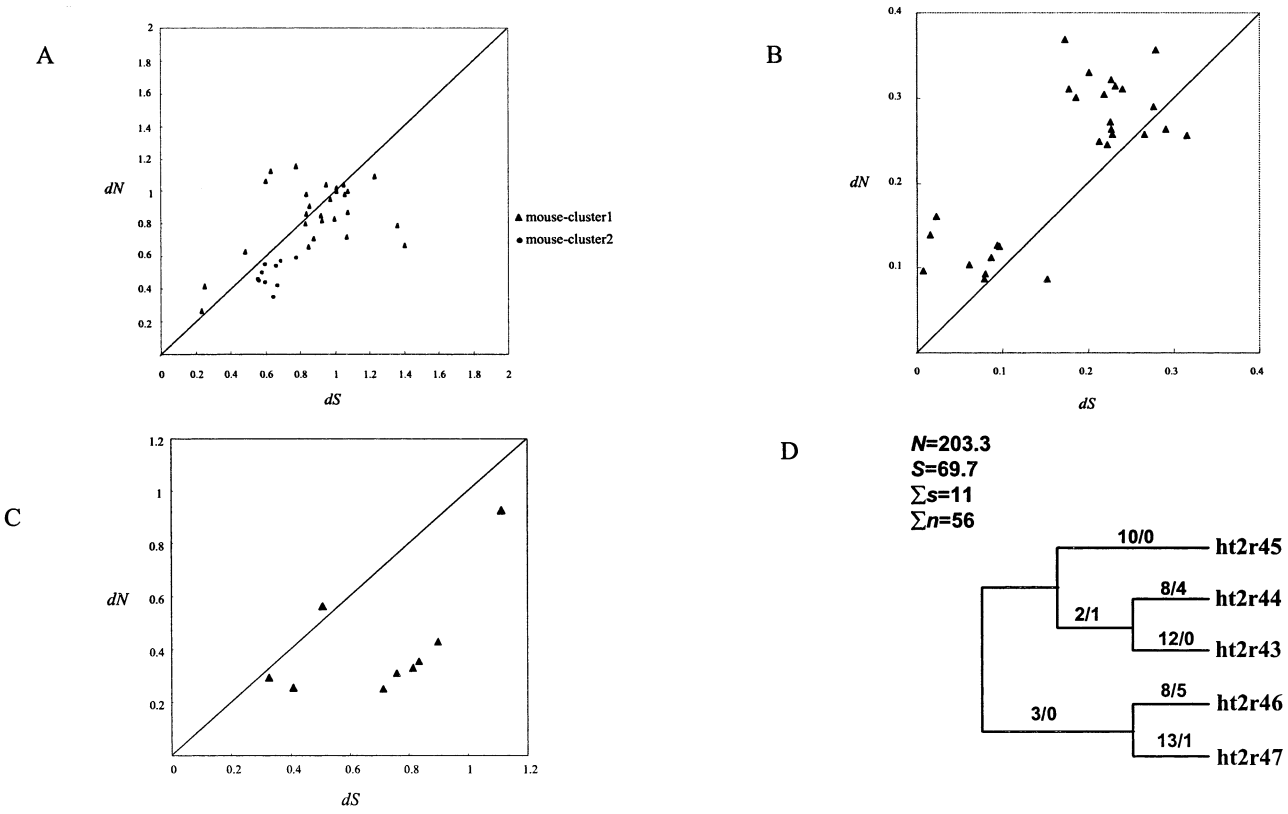


FIG. 5.—Pairwise synonymous (d_S) and nonsynonymous (d_N) nucleotide distances for extracellular regions. *A*, Mouse cluster I and II genes from group A. *B*, Human cluster in group A. *C*, Orthologous human-mouse pairs in group B and C. *D*, Phylogeny-based testing of positive selection for recently duplicated human T2R genes. On each branch is the number of inferred nonsynonymous (*n*) substitutions followed by that of synonymous (*s*) substitutions. *N* and *S* are potential numbers of nonsynonymous and synonymous sites of the sequences, respectively.

Table 4
Likelihood Ratio Tests of Positive Selection for the Three Species-Specific Clusters of Group A T2R Genes

			2Δl			
Clusters	*n*	Lc	M3 vs. M0	M8 vs. M7	Parameters Estimated Under M8	Positively Selected Sites[a]
Human	8	897	38.86**	24.34**	$p1 = 0.189$ $\omega = 3.13$ $P0 = 0.811$ $p = 0.695$ $q = 0.212$	16T 177M 253G 254S 268R
Mouse cluster1	8	903	226.46**	18.74**	$p1 = 0.223$ $\omega = 1.93$ $P0 = 0.777$ $p = 0.736$ $q = 0.741$	6E 169L 259A 292S
Mouse cluster 2	5	891	57.88**	4.04	$p1 = 0.326$ $\omega = 1.31$ $P0 = 0.674$ $p = 1.539$ $q = 2.93$	83L 158Y

NOTE.—*n*: number of sequences in the data set; Lc: number of codons used in the data set after removing alignment gaps.
[a] The sites with posterior probabilities >95% under M8 are listed.
** Significant at 1% level.

significant clusters are considered, the proportion of positively selected sites in ER is 77% for the human cluster and 74% for mouse cluster 1. These numbers are significantly greater than expected under the homogeneous distribution model (chi-square test, $P < 1\%$), as ER only constitute about 30% of the entire protein. This result is consistent with that from pairwise comparisons (table 3).

Discussion

In this work, we searched the mouse and human genome sequences and identified new members of the bitter taste receptor T2R gene family. Together with T2R genes reported earlier, we conducted an evolutionary analysis of all known and putative bitter taste receptor genes of the human and mouse. We found that T2R genes may be classified into three groups. Although most of the group B and C genes show one-to-one orthology between the human and mouse, group A genes exhibit several species-specific gene clusters. This pattern suggests the presence of ''species (lineage)-specific'' and ''species-general'' bitter taste receptors. Species (lineage)-specific T2Rs are receptors that exist in only one species (lineage), with no well-defined orthologous receptors in other species (lineages). Groups of these receptors may have evolved separately in different species to deal with the specific bitter tastants they encounter. Species-general receptors are those receptors potentially common to many mammals. Each of these proteins may be used for detecting one or several distinct bitter compounds that are encountered by all or many species. For instance, human receptor ht2r4 and its mouse orthologue mt2r8 (fig. 2) are activated by denatonium and high concentrations of PROP, but not by other bitter tastants tested (Chandrashekar et al. 2000). The intriguing presence of species-specific and species-general T2Rs awaits further scrutiny from additional mammalian species.

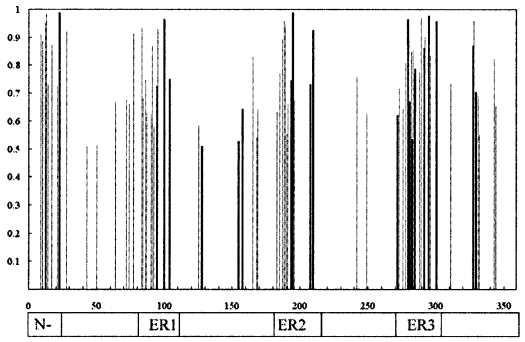


FIG. 6.—Posterior probabilities (>50%) for sites under positive selection. X-axis denotes position in the amino acid alignment. Y-axis denotes posterior probability of sites under positive selection. Sites with black line are those inferred from the human cluster, whereas those with darkish line are from mouse cluster 1. Boxes under the graph denote extracellular regions. A color figure is available as online Supplementary Material.

There is experimental evidence that different T2Rs respond to different bitter tastants (Adler et al. 2000; Chandrashekar et al. 2000; Matsunami, Montmayeur, and Buck 2000). For instance, mouse T2R5 responds to cycloheximide, while human receptor ht2r4 and mouse receptor mt2r8 respond to denatonium and PROP (Chandrashekar et al. 2000). Most recently, it is reported that human T2R16 receptor responds to the bitter tastant β-glucopyranosides (Bufe et al. 2002). Many physiological and neurophysiological studies demonstrated that species-specific bitter perception occurs in rats, pigs, and primates, and indicated that bitter-sensitive taste cells may be responsive to a variety of bitter compounds (Dahl, Erickson, and Simon 1997; Hellekant, Danilova, and Ninomiya 1997; Danilova et al., 1998; Glendinning, Tarre, and Asaoka 1999). With accumulation of such data and molecular characterization of individual T2Rs, it might be possible to understand the common and specific bitter tastants that each species detects.

It has been estimated that the human and mouse genomes contain many T2R pseudogenes (Adler et al. 2000). Indeed, we identified T2R pseudogenes from these species. It appears that the pseudogenes are not necessarily products of recent duplication events (figs. 3 and 4). That is, certain T2Rs had been functional for a long time before being inactivated. We speculate that they became dispensable when the specific bitter tastants no longer existed in the environment that the species occupied and new T2Rs were acquired for detecting newly encountered tastants.

Among the extracellular, transmembrane, and intracellular regions of the T2R molecule, the extracellular regions show the highest sequence variability. Among paralogous human and mouse genes, mainly the ''species (lineage)-specific'' genes of group A, both the comparison of synonymous and nonsynonymous substitution rates and the examination of d_N/d_S at individual amino acid sites

reveal the action of positive selection in the extracellular regions. If the extracellular regions are indeed involved in tastant binding, as predicated by several authors (Adler et al. 2000; Gilbertson, Damak, and Margolskee 2000), our results would suggest that the ability to detect a diverse array of bitter tastants is selectively favored in the evolution of mammals. This is understandable, as many poisonous substances in nature taste bitter (Garcia and Hankins 1975; Glendinning 1994; Glendinning, Tarre, and Asaoka 1999; Chandrashekar et al. 2000), and an organism capable of recognizing a greater number of bitter tastants has a lower probability of ingestion of harmful substances and thus has a higher fitness. In the future, it would be interesting to study the T2R repertoires from additional mammals to test this hypothesis and to search for the molecular basis of adaptation of organisms to their specific environments, such as the unique digestive ribonuclease found in the leafing-eating colobine monkeys (Zhang, Zhang, and Rosenberg 2002).

Supplementary Material

The sequences reported in this paper have been deposited in the GenBank database. Accession numbers: AY161895–AY161926 and AY168282-AY168292.

Acknowledgments

This work was supported by grants from the Chinese Academy of Sciences (KSCX2-1–05), Program for Key International S & T Cooperation Project of the People's Republic of China (2001CB711103), and National Natural Science Foundation of China to Y.P.Z., and by a start-up fund from the University of Michigan to J.Z.

Literature Cited

Adler, E., M. A. Hoon, K. L. Mueller, J. Chandrashekar, N. J. Ryba, and C. S. Zuker. 2000. A novel family of mammalian taste receptors. Cell **100**:693–702.

Altschul, S. F., T. L. Madden, A. A. Schaffer, J. Zhang, Z. Zhang, W. Miller, and D. J. Lipman. 1997. Gapped Blast and PSI-Blast: a new generation of protein database search programs. Nucleic Acids Res. **25**:3389–3402.

Azen, E. A., I. E. Lush, and B. T. Taylor. 1986. Close linkage of mouse genes for salivary proline-rich proteins (Prps) and taste. Trends Genet. **2**:199–200.

Bufe, B., T. Hofmann, D. Krautwurst, J. D. Raguse, and W. Meyerhof. 2002. The human TAS2R16 receptor mediates bitter taste in response to beta-glucopyranosides. Nat. Genet. **32**:397–401.

Capeless, C. G., G. Whitney, and E. A. Azen. 1992. Chromosome mapping of Soa, a gene influencing gustatory sensitivity to sucrose octaacetate in mice. Behav. Genet. **22**:655–663.

Chandrashekar, J., K. L. Mueller, M. A. Hoon, E. Adler, L. Feng, W. Guo, C. S. Zuker, and N. J. Ryba. 2000. T2Rs function as bitter taste receptors. Cell **100**:703–711.

Chaudhari, N., A. M. Landin, and S. D. Roper. 2000. A metabotropic glutamate receptor variant functions as a taste receptor. Nat. Neurosci. **3**:113–119.

Dahl, M., R. P. Erickson, and S. A. Simon. 1997. Neural responses to bitter compounds in rats. Brain Res. **756**:22–34.

Danilova, V., G. Hellekant, T. Roberts, J. M. Tinti, and C. Nofre. 1998. Behavioral and single chorda tympani taste fiber responses in the common marmoset, *Callithrix jacchus jacchus*. Ann. N.Y. Acad. Sci. **855**:160–164.

Drouin, G., F. Prat, M. Ell, and G. D. Clarke. 1999. Detecting and characterizing gene conversions between multigene family members. Mol. Biol. Evol. **16**:1369–1390.

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

Garcia, J., and W. G. Hankins. 1975. The evolution of bitter and the acquisition of toxiphobia. Pp. 39–45 *in* Olfaction and taste. V, Proceedings of the 5th International Symposium. Academic Press, Melbourne, Australia.

Gilbertson, T. A., S. Damak, and R. F. Margolskee. 2000. The molecular physiology of taste transduction. Curr. Opin. Neurobiol. **10**:519–527.

Glendinning, J. I. 1994. Is the bitter rejection response always adaptive? Physiol. Behav. **56**:1217–1227.

Glendinning, J. I., M. Tarre, and K. Asaoka. 1999. Contribution of different bitter-sensitive taste cells to feeding inhibition in a caterpillar (*Manduca sexta*). Behav. Neurosci. **113**: 840–854.

Hellekant, G., V. Danilova, and Y. Ninomiya. 1997. Primate sense of taste: behavioral and single chorda tympani and glossopharyngeal nerve fiber recordings in the rhesus monkey, *Macaca mulatta*. J. Neurophysiol. **77**:978–993.

Kinnamon, S. C., and T. A. Cummings. 1992. Chemosensory transduction mechanisms in taste. Annu. Rev. Physiol. **54**: 715–731.

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

Lindemann, B. 1996. Taste reception. Physiol. Rev. **76**:718–766.

———. 2000. A taste for umami. Nat. Neurosci. **3**:99–100.

Lush, I. E. 1984. The genetics of tasting in mice. III. Quinine. Genet. Res. **44**:151–160.

———. 1986. The genetics of tasting in mice. IV. The acetates of raffinose, galactose and beta-lactose. Genet. Res. **47**:117–123.

Lush, I. E., and G. Holland. 1988. The genetics of tasting in mice. V. Glycine and cycloheximide. Genet. Res. **52**:207–212.

Matsunami, H., J. P. Montmayeur, and L. B. Buck. 2000. A family of candidate taste receptors in human and mouse. Nature **404**:601–604.

Montmayeur, J. P., S. D. Liberles, H. Matsunami, and L. B. Buck. 2001. A candidate taste receptor gene near a sweet taste locus. Nat. Neurosci. **4**:492–498.

Nei, M., and S. Kumar. 2000. Molecular evolution and phylogenetics. Oxford University Press, New York.

Nelson, G., M. A. Hoon, J. Chandrashekar, Y. Zhang, N. J. Ryba, and C. S. Zuker. 2001. Mammalian sweet taste receptors. Cell **106**:381–390.

Nielsen, R., and Z. Yang. 1998. Likelihood models for detecting positively selected amino acid sites and applications to the HIV-1 envelope gene. Genetics **148**:929–936.

Reed, D. R., E. Nanthakumar, M. North, C. Bell, L. M. Bartoshuk, and R. A. Price. 1999. Localization of a gene for bitter-taste perception to human chromosome 5p15. Am. J. Hum. Genet. **64**:1478–1480.

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

Sawyer, S. 1989. Statistical tests for detecting gene conversion. Mol. Biol. Evol. **6**:526–538.

Stewart, R. E., J. A. DeSimone, and D. L. Hill. 1997. New perspectives in a gustatory physiology: transduction, development, and plasticity. Am. J. Physiol. **272**:C1–C26.

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

Tanaka, T., and M. Nei. 1989. Positive Darwinian selection observed at the variable-region genes of immunoglobulins. Mol. Biol. Evol. **6**:447–459.

Thompson, J. D., D. G. Higgins, and T. J. Gibson. 1994. ClustalW: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. Nucleic Acids Res. **22**:4673–4680.

Wong, G. T., K. S. Gannon, and R. F. Margolskee. 1996. Transduction of bitter and sweet taste by gustducin. Nature **381**: 796–800.

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

Zhang, J. 2000. Rates of conservative and radical nonsynonymous nucleotide substitutions in mammalian nuclear genes. J. Mol. Evol. **50**:56–68.

Zhang, J., and M. Nei. 1997. Accuracies of ancestral amino acid sequences inferred by the parsimony, likelihood, and distance methods. J. Mol. Evol. **44**(Suppl.):S139–S146.

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

Zhang, J., Y. P. Zhang, and H. F. Rosenberg. 2002. Adaptive evolution of a duplicated pancreatic ribonuclease gene in a leaf-eating monkey. Nat. Genet. **30**:411–415.

Naruya Saitou, Associate Editor

Accepted January 2, 2003