Contrasting Modes of Evolution Between Vertebrate Sweet/Umami Receptor Genes and Bitter Receptor Genes

Peng Shi and Jianzhi Zhang

Department of Ecology and Evolutionary Biology, University of Michigan

Taste reception is fundamental to diet selection in many animals. The genetic basis underlying the evolution and diversity of taste reception, however, is not well understood. Recent discoveries of T1R sweet/umami receptor genes and T2R bitter receptor genes in humans and mice provided an opportunity to address this question. Here, we report the identification of 20 putatively functional T1R genes and 167 T2R genes from the genome sequences of nine vertebrates, including three fishes, one amphibian, one bird, and four mammals. Our comparative genomic analysis shows that orthologous T1R sequences are relatively conserved in evolution and that the T1R gene repertoire remains virtually constant in size across most vertebrates, except for the loss of the T1R2 sweet receptor gene in the sweet-insensitive chicken and the absence of all T1R genes in the tongueless western clawed frog. In contrast, orthologous T2R sequences are more variable, and the T2R repertoire diverges tremendously among species, from only three functional genes in the chicken to 49 in the frog. These evolutionary patterns suggest the relative constancy in the number and type of sweet and umami tastants encountered by various vertebrates or low binding specificities of T1Rs but a large variation in the number and type of bitter compounds detected by different species. Although the rate of gene duplication is much lower in T1Rs than in T2Rs, signals of positive selection are detected during the functional divergences of paralogous T1Rs, as was previously found among paralogous T2Rs. Thus, functional divergence and specialization of taste receptors generally occurred via adaptive evolution.

Introduction

Taste perception varies enormously across different lineages and species of vertebrates. For example, among birds, characteristic response to sweet stimuli is not found in chicken (Halpern 1962; Ganchrow, Steiner, and Bartana 1990), although strong preferences for sweet substances are well documented in hummingbirds and other nectar feeders (Kare and Medway 1959; Duncan 1960; Whittow 2000). In mammals, different trophic groups differ in their selectivity and sensitivity to bitter compounds (Glendinning 1994). Amphibians in the family Pipidae, including *Xenopus*, are tongueless but were reported to use the gustatory sensory receptor cells on the floor of the oral cavity to respond to bitter substances and amino acids (Yoshii et al. 1982). Because the taste perception of a species is intimately related to its diet and environment, an understanding of the variation of taste perception among species, particularly in the context of their environments, can tell us how ecology influences the emergence of biodiversity and adaptation. It would be especially illuminating if the variation and evolution of the genes controlling taste reception can be studied.

Vertebrate taste perception is mediated by taste receptor cells, which usually cluster in taste buds (Roper 1989). Taste receptor cells can detect a wide range of chemicals and trigger transduction cascades, which cause excitation of the nerve fibers (Herness and Gilbertson 1999; Lindemann 2001). Although there are numerous chemicals to which taste receptor cells respond, it is generally thought that there are only five basic tastes: sour, salty, bitter, sweet, and umami (Kinnamon and Margolskee 1996; Lindemann 2001). Among them, the sweet and umami tastes can result in appetitive reactions and generally reflect the identification of nutrients, whereas the bitter taste may cause aversion and thus is a defensive mechanism against ingestion of toxins (Herness and Gilbertson 1999). In mammals, two distinct families of seven-transmembrane G protein–coupled receptors (GPCRs), T1Rs and T2Rs, have been identified as sweet/umami taste receptors and bitter taste receptors, respectively (Hoon et al. 1999; Adler et al. 2000; Chandrashekar et al. 2000; Matsunami, Montmayeur, and Buck 2000; Kitagawa et al. 2001; Max et al. 2001; Montmayeur et al. 2001; Nelson et al. 2001, 2002; Sainz et al. 2001; Zhao et al. 2003; Meyerhof 2005). In humans and mice, the T1R family includes three members: T1R1, T1R2, and T1R3 (Hoon et al. 1999; Kitagawa et al. 2001; Max et al. 2001; Montmayeur et al. 2001; Nelson et al. 2001; Sainz et al. 2001; Zhao et al. 2003), whereas the T2R family comprises ~30 genes (Conte et al. 2003; Shi et al. 2003). T1Rs belong to GPCR C family and are related to V2R vomeronasal pheromone receptors (Pin, Galvez, and Prezeau 2003), whereas T2Rs belong to GPCR A family and are related to V1R vomeronasal receptors (Adler et al. 2000; Matsunami, Montmayeur, and Buck 2000; Shi, Huang, and Zhang 2005). T1R genes contain multiple introns, and the proteins are characterized by a long N-terminal domain that may be involved in ligand binding (Pin, Galvez, and Prezeau 2003). By contrast, T2R genes are intronless, and the proteins have a short N-terminal extracellular domain. Furthermore, T1Rs and T2Rs differ in function and expression. Multiple, if not all, T2Rs are coexpressed in individual cells that also express α-gustducin, a G protein subunit (Adler et al. 2000; Nelson et al. 2001). The coexpression of T2Rs in individual cells results in a common response to a wide array of bitter tastants (Adler et al. 2000; Mueller et al. 2005). By contrast, T1Rs are not coexpressed with α-gustducin. T1R1 and T1R2 are expressed in distinct taste receptor cells, but they are always coexpressed with T1R3. Functional assays showed that the T1R1 + T1R3 heteromeric receptor can detect L-amino acids and monosodium L-glutamate (MSG), which is the taste of umami, whereas the T1R2 + T1R3 heteromeric receptor can respond broadly to sweet tastants (Mombaerts 2004). To date, the evolution of T1R and T2R gene repertoires have been described only in primates and rodents (Conte et al. 2003; Shi et al. 2003;

Key words: taste, receptor, T1R, T2R, sweet, umami, bitter, evolution.

E-mail: jianzhi@umich.edu.

Mol. Biol. Evol. 23(2):292–300. 2006
doi:10.1093/molbev/msj028
Advance Access publication October 5, 2005

© The Author 2005. Published by Oxford University Press on behalf of the Society for Molecular Biology and Evolution. All rights reserved.
For permissions, please e-mail: journals.permissions@oxfordjournals.org

Wang, Thomas, and Zhang 2004; Fischer et al. 2005; Go et al. 2005). Little is known about the evolution of these two families in other vertebrates. In this paper, we identify T1R and T2R genes from nine vertebrate genome sequences and perform comparative evolutionary analyses of these two families. We show that T1Rs and T2Rs exhibit drastically different modes of evolution, with the former being conserved and latter being radical in the changes of gene number and gene sequence across vertebrates.

Materials and Methods

Identification of T1R and T2R Genes

Sequences of previously reported rodent and primate T1R and T2R genes were retrieved from the GenBank. Additional genes belonging to the two gene families were obtained by searching the available vertebrate genome sequences at UCSC (http://genome.cse.ucsc.edu/) and Ensembl (http://www.ensembl.org/), respectively. These genome sequences include three mammals (dog *Canis familiaris*, cow *Bos taurus*, and opossum *Monodelphis domestica*), one bird (chicken *Gallus gallus*), one amphibian (western clawed frog *Xenopus tropicalis*), and three teleost fishes (zebrafish *Danio rerio*, pufferfish *Tetraodon nigroviridis*, and fugu *Takifugu rubripes*). Note that the chimpanzee genome sequence is excluded in this study because it is very close to human sequence and because chimpanzee T2Rs have been characterized (Parry, Erkner, and le Coutre 2004; Wang, Thomas, and Zhang 2004; Fischer et al. 2005; Go et al. 2005). TBlastN (Altschul et al. 1990) was employed to search for T2R genes in the above genome sequences, with the previously known T2R genes used as queries. Because T1R genes contain introns, we employed a more complex search strategy, as described in Yang et al. (2005). Briefly, we first used Blast to identify the genomic locations of putative T1R genes in a genome with the previously reported T1R sequences as queries. Second, the genomic DNA sequences of the putative T1R genes and the known T1R cDNA sequences were used to conduct a cDNA-to-genomic sequence alignment by Spidey (http://www.ncbi.nlm.nih.gov/IEB/Research/Ostell/Spidey/), which provided the exon/intron structures and full-length protein sequences of the putative genes. To ensure the accurate prediction of T1R and T2R genes, the obtained putative protein sequences were examined by the TMHMM method (Sonnhammer, von Heijne, and Krogh 1998) for the presence of seven transmembrane domains. In addition, these putative T1R (or T2R) genes were Blasted against the entire GenBank to ensure that their best hits were known T1Rs (or T2Rs). This step was necessary because T1Rs and T2Rs are known to be closely related to some other GPCR family C receptors (e.g., Ca^{2+}-sensing receptors and V2R vomeronasal receptors) and family A receptors (e.g., V1R vomeronasal receptors), respectively. A T1R or T2R sequence was regarded as a pseudogene if it did not have seven transmembrane domains or its open reading frame (ORF) was disrupted. We only considered those disrupted sequences that were longer than 200 nt as pseudogenes. Sequences shorter than 200 nt were discarded. This stringent criterion led to a low rate of false detection of pseudogenes, although some true pseudogenes might have been missed. The amino acid sequences of the newly identified T1R and T2R genes are provided in Supplementary data set 1 and Supplementary data set 2 (Supplementary Material online), respectively.

Evolutionary Analyses

Deduced amino acid sequences of T1Rs and T2Rs were separately aligned by Clustal_X (Thompson et al. 1997), with manual adjustments. Phylogenetic trees were reconstructed using the neighbor-joining method (Saitou and Nei 1987) with protein Poisson distances (Nei and Kumar 2000). Gap sites in the alignment were not used in the phylogenetic reconstruction (i.e., the complete-deletion option). The reliability of the estimated trees was evaluated by the bootstrap method (Felsenstein 1985) with 1,000 pseudoreplications. MEGA2 (Kumar et al. 2001) was used for the phylogenetic analysis. Positive selection in the divergence of paralogous genes was tested by the recently implemented branch-site test 2, which has been shown to be conservative and robust against violations of various model assumptions (Zhang, Nielsen, and Yang 2005), unlike the previous version of the test (Yang and Nielsen 2002), which has a relatively high rate of false positives (Zhang 2004). The branch-site test 2 compares two models. The alternative model assumes four classes of sites in terms of ω, which is the rate ratio of nonsynonymous substitutions to synonymous substitutions (Nei and Kumar 2000). Site class 0 contains codons that are conserved throughout the tree, with $0 < \omega_0 < 1$ estimated. Site class 1 contains codons that are neutral throughout the tree with $\omega_1 = 1$. Site classes 2a and 2b contain codons that are conserved or neutral on the background branches but become under positive selection on the foreground branches with $\omega_2 > 1$, estimated from the data. The model has four parameters in the ω distribution: p_0, p_1, ω_0, and ω_2. The null model differs from the alternative model in that ω_2 is fixed at 1. Thus, a significant higher likelihood of the alternative model than that of the null model indicates positive selection on the foreground branches.

Results

Vertebrate T1R Genes

To characterize the T1R gene repertoire in vertebrates, we searched seven vertebrate genome sequences, including two mammals (dog and opossum), one bird (chicken), one amphibian (western clawed frog), and three fishes (fugu, pufferfish, and zebrafish), with the previously reported T1R sequences as queries. We did not search T1Rs in the cow genome sequence because its chromosomal assembly was unavailable and the prediction of multiexon genes was unfeasible. There are three types of newly identified genes: functional genes, which are full-length ORFs with intact seven-transmembrane domains; putative functional genes, which are partial sequences due to incomplete genome sequences but contain undisrupted ORFs in the available sequences; and pseudogenes, which are sequences with premature stop codons or lack intact seven-transmembrane domains. In the case of T1Rs, no pseudogenes were detected. Both dog and opossum have three T1R genes, as in the human, mouse, and rat (table 1), but the number of T1R genes varies in some nonmammalian vertebrates.

Table 1
Number of Genes and Pseudogenes in Vertebrate T1R and T2R Gene Families

	T1R		T2R				
Species	Functional	Putatively Functional	Functional	Putatively Functional	Pseudogenes	Total	Proportion of Pseudogenes
Human[a]	3	0	25	0	11	36	31%
Mouse[b]	3	0	35	0	6	41	15%
Rat[c]	3	0	37	0	5	42	12%
Dog	3	0	15	1	5	21	24%
Cow	Unknown	Unknown	12	7	15	34	44%
Opossum	3	0	26	3	5	34	15%
Chicken	2	0	3	0	0	3	0
Frog	0	0	49	3	12	64	19%
Fugu fish	4	1	4	0	0	4	0
Puffer fish	5	1	6	0	0	6	0
Zebrafish	1	0	4	0	0	4	0

[a] T1R gene number is from Liao and Schultz (2003). T2R gene number is from Shi et al. (2003) and Go et al. (2005).
[b] T1R gene number is from Hoon et al. (1999) and Max et al. (2001). T2R gene number is from Shi et al. (2003) and Go et al. (2005).
[c] T1R gene number is from Li et al (2002).

For example, fugu and pufferfish have five and six T1R genes, respectively. However, only one T1R gene was found in the zebrafish, which may be an artifact due to the relatively low quality of its genome sequence. In chicken, T1R1 and T1R3 are present, but T1R2 is missing. The mouse T1R2 gene is located in chromosome 4, flanked by *Aldh4a1* (NM_175438) and *Pax7* (AF25442). In the chicken genome, these two genes are tandem linked on chromosome 21, which is homologous to mouse chromosome 4. Thus, the absence of the T1R2 gene in chicken is likely to be real. No T1R genes were found in the western clawed frog. This result is puzzling because the African clawed frog (*Xenopus laevis*), a species closely related to the western clawed frog, was reported to have high gustatory sensitivity to amino acids (Yoshii et al. 1982), which are detected in mice by T1Rs. There are three possible explanations. First, although the African clawed frog can sense amino acids, the western clawed frog cannot. However, if this were true, we would have identified T1R pseudogenes as the two species are closely related. Second, the T1R genes exist in the frog genome but were not sequenced in the draft genome sequence. Because the T1R genes are unlinked to each other, it is unlikely that all T1R genes would have been missed in genome sequencing. We examined whether the genes flanking the three T1R genes in the mouse can be found in the frog genome. Indeed, all the flanking genes are found. For T1R3, there is no sequencing gap between the two flanking genes in the frog. For T1R1 and T1R2, no sequence gaps that can potentially harbor these two genes are found between the flanking genes. Thus, the loss of the T1R genes in the frog is likely true. Third, there are non-T1R genes in frogs that are responsible for detecting amino acids. It is known in fish that V2R receptors can detect amino acids as odorants (Speca et al. 1999). Frogs have numerous V2R genes (P. Shi and J. Zhang, unpublished data) that may be responsible for detecting amino acids. It would be interesting to test this hypothesis and to examine whether other amphibians retain T1Rs.

To understand the evolutionary relationships among vertebrate T1R genes, we reconstructed a phylogenetic tree of the newly identified functional T1Rs and the previously known T1Rs of the human, mouse, and rat, using the neighbor-joining method with Poisson protein distances (fig. 1). Vertebrate V2R vomeronasal receptors were used to root the tree. The tree shows that T1Rs form three monophyletic clades, supported by 67%, 98%, and 100% bootstrap values, respectively. Each of the three clades contains one of the three T1Rs previously characterized in humans and mice. We therefore named the three clades as T1R1, T1R2, and T1R3 clades, respectively. With the exception of frog, zebrafish, and chicken, every vertebrate examined has at least one gene in each of the three clades, indicating that the divergence among the three clades occurred before the separation of tetrapods and teleosts. The phylogenetic tree also reveals that a few gene duplication events occurred in teleosts. For example, duplication of T1R2 occurred before the divergence of pufferfish and fugu, whereas another duplication took place after the split of pufferfish and fugu. Overall, the number of T1R genes remains more or less constant among the vertebrates examined. It is interesting to note that in all three T1R clades, the human sequences are phylogenetically closer to the dog sequences than to the rodent sequences, contradictory to the current understanding of mammalian phylogeny (Murphy, Pevzner, and O'Brien 2004). However, only for one of the three T1Rs, the human-dog clustering has a high bootstrap support (fig. 1).

Vertebrate T2R Genes

We also investigated T2R genes in the seven genome sequences mentioned above. Because T2R genes are intronless, it is possible to search these genome sequences for which only contig assemblies are available. Thus, we also searched the cow genome sequence. In addition, we searched the rat genome sequence because not all rat T2R genes had been previously described (Wang, Thomas, and Zhang 2004). The number of T2R genes varies considerably between mammals and nonmammalian vertebrates (table 1). A total of three, four, six, and four T2R genes were identified from chicken, fugu, pufferfish, and zebrafish, respectively. By contrast, a total of 64 T2R genes were found

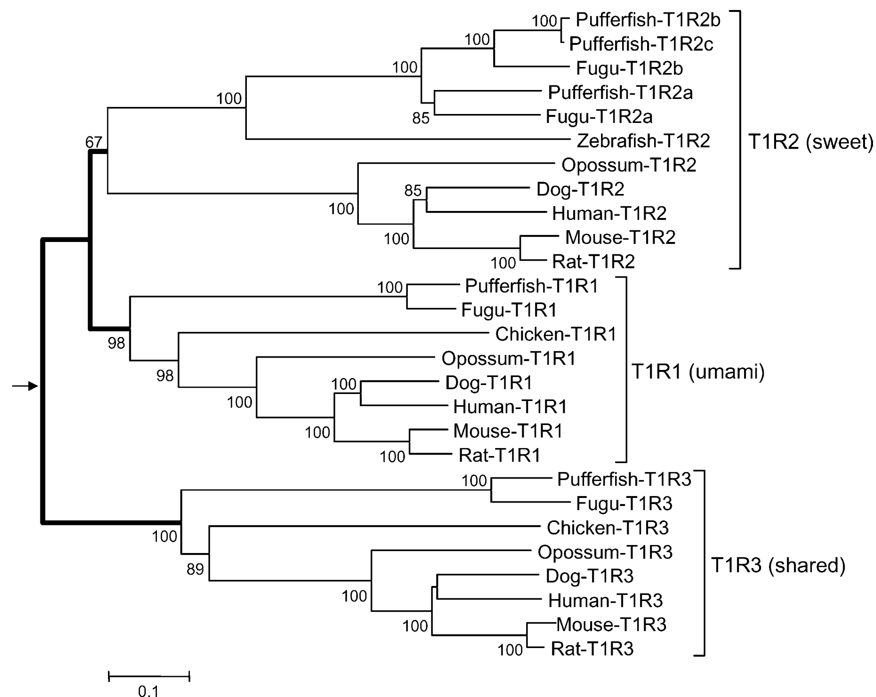

FIG. 1.—Phylogenetic relationships of vertebrate intact T1R genes. The tree is reconstructed by the neighbor-joining method with protein Poisson distances. Bootstrap percentages higher than 50 are shown on interior branches. The arrow points to the tree root, which is determined by using vertebrate V2R genes as outgroups (Yang et al. 2005). The scale bar shows 0.1 amino acid substitution per site. The branches in bold, connecting the three T1R gene clades, are used as foreground branches for testing positive selection (table 2). The T1R1 + T1R3 heteromeric receptor can detect L-amino acids and MSG, which is the taste of umami, whereas the T1R2 + T1R3 heteromeric receptor can respond broadly to sweet tastants.

in the frog, including 49 functional genes, 3 putatively functional genes, and 12 pseudogenes. Mammals also have many T2R genes. For instance, 37 functional genes and 5 pseudogenes were found in rat, which is close to what was previously described in mouse (Conte et al. 2003; Shi et al. 2003; Go et al. 2005). While most mammals have 34–42 T2R genes and pseudogenes, the dog has only 21. When only functional or putatively functional T2Rs of mammals are considered, the among-species variation is even greater, from 12 in dog to 37 in rat. The proportion of pseudogenes in the T2R repertoire is lower in rodents and opossum (12%–15%) than in dog (24%), which is in turn lower than that in the human and cow (31%–44%).

We reconstructed a protein neighbor-joining tree of all newly identified functional T2Rs, together with previously described human and mouse T2Rs (fig. 2; see Supplementary Fig. 1 [Supplementary Material online] for bootstrap results). We excluded the putatively functional genes newly identified because only partial sequences were available. The tree was rooted using vertebrate V1R vomeronasal receptors (Grus et al. 2005; Pfister and Rodriguez 2005). T2R genes from teleost fishes do not form a monophyletic clade, suggesting that the common ancestor of tetrapods and teleosts already had multiple T2Rs (fig. 2). In a previous study, mouse and human T2R genes were classified into three groups: A, B, and C (Shi et al. 2003). Although this grouping was tentative because of the low bootstrap support, it was corroborated by the observation that genes within groups also cluster in chromosomal locations (Shi et al. 2003). We highlighted these previously defined groups in the present tree (fig. 2) and found that the grouping is consistent with the chromosomal distributions of the T2R genes in additional species. For example, group C genes in the dog are located on chromosomes 16, 14, and 34, which are homologous to human chromosome 7 and 5, on which human group C genes are located. All of the dog group A genes are located on chromosome 27, syntenic to human chromosome 12, where human group A genes are found. Lineage-specific gene duplications took place in fishes, frogs, and mammals (fig. 2). In teleosts, four pufferfish T2R genes cluster to form a species-specific clade. In frogs, multiple recent duplications have occurred. In mammals, almost all lineage-specific gene duplications occurred in group A. For example, opossum-specific and cow-specific duplications were found in group A, in addition to the previously described primate- and rodent-specific duplications (Shi et al. 2003). In contrast to the rapidly duplicating group A genes, most group C genes have one-to-one orthologs from the mammalian species considered. Nevertheless, some orthologous groups do not include every species, which may in part be due to the exclusion of putatively functional genes from the tree as the complete sequence information is lacking for these genes. Group B is relatively small. It has one-to-one orthologs in placental mammals but a species-specific clade for the opossum, apparently due to the gene duplications postdating the divergence of placental and marsupial mammals. It is interesting to note that while the chicken and fishes have small T2R repertoires, the frog and mammals have large repertoires. This raised the question of when the T2R repertoire expanded in evolution. One scenario is that the expansion occurred in the common ancestor of tetrapods but later shrank in birds. Another scenario is that the repertoire

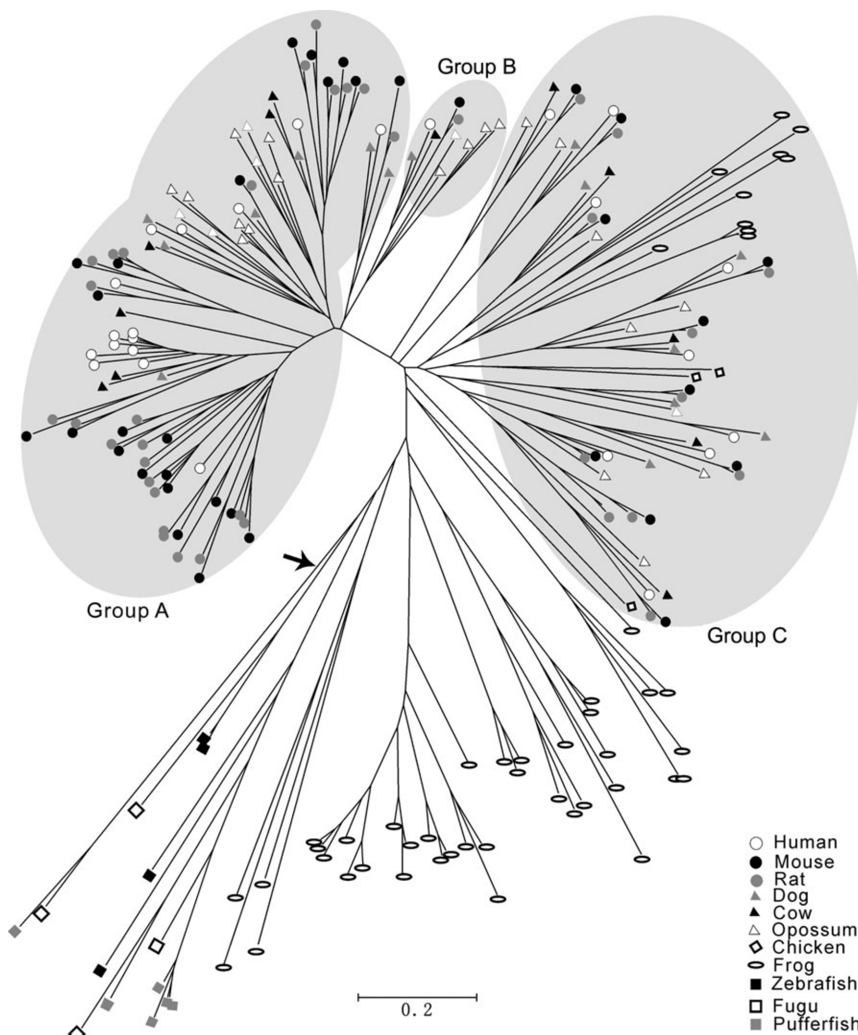

FIG. 2.—Neighbor-joining tree of 216 functional T2R genes from 10 vertebrates. The tree is reconstructed with protein Poisson distances. Arrow indicates the root of the tree, which is determined by using vertebrate V1R genes as outgroups (Grus et al. 2005; Pfister and Rodriguez 2005). Groups A, B, and C were previously defined based on the tree of human and mouse T2R genes (Shi et al. 2003). The same tree with bootstrap information is presented in Supplementary Fig. 1 (Supplementary Material online).

expanded independently in frogs and mammals. Because there were already numerous T2R lineages when frogs and mammals diverged (see group C, fig. 2), it is likely that the repertoire expansion started in their common ancestor. However, recent duplications also suggest that the expansions occurred in the mammals and frogs independently. Thus, a composite scenario with an initial expansion in the common ancestor of tetrapods, followed by additional independent expansions in frogs and mammals, is more likely to be true. According to this scenario, the chicken T2R repertoire may have shrunk to some extent. It should be noted that the tree in figure 2 has low bootstrap values (see Supplementary Fig. 1, Supplementary Material online), and the scenario we describe here requires further confirmation.

Rates of Sequence Divergence in T1Rs and T2Rs

In the above paragraphs, we compared the rate of gene birth/death between T1R and T2R families. It is also interesting to examine the rate of gene sequence divergence between the two families. Figure 3 shows the protein *p*-distances (i.e., proportional differences) between orthologous T1Rs from the human, mouse, rat, and opossum and the corresponding numbers for T2Rs. In addition to the three T1R genes, we were able to examine only six one-to-one orthologs in the T2R family because of frequent lineage-specific gene birth/death. These six genes were identified from the tree in figure 2, with the criteria that each gene has orthologous sequences from the human, mouse, rat, and opossum and the phylogeny of the orthologous sequences is consistent with the species tree. For orthologous pairs between the human and opossum, the *p*-distance is significantly lower for the three T1R genes than the six T2R genes ($P = 0.024$, two-tailed Mann-Whitney U test). The same is true for the human-mouse orthologous pairs ($P = 0.024$, two-tailed Mann-Whitney U test). For the mouse-rat orthologous pairs, T1Rs are again more conserved than T2Rs, although the difference is not significant ($P = 0.262$,

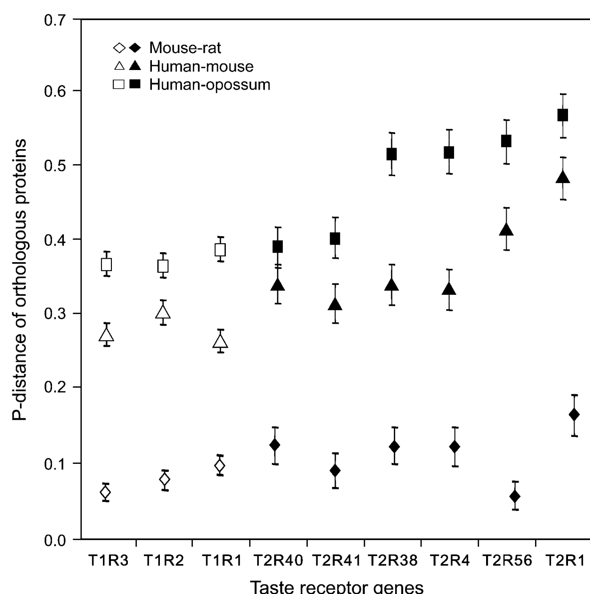

FIG. 3.—Protein *p*-distances of T1R and T2R orthologous genes in mammals. Diamonds indicate *p*-distances of mouse and rat orthologs, triangles indicate *p*-distances of human and mouse orthologs, and squares indicate *p*-distances of human and opossum orthologs. T1R and T2R genes are represented by open and closed symbols, respectively. Error bars show one standard error of *p*.

two-tailed Mann-Whitney *U* test), probably due to the large sampling errors relative to the estimates (fig. 3). Taken together, these comparisons show that T1Rs evolve more slowly than T2Rs at the protein sequence level when the divergences of orthologous genes are concerned.

We next examined the rate of sequence divergence among paralogous genes. Our previous study of paralogous T2Rs showed that they diverged rapidly by positive selection (Shi et al. 2003). Here, we test whether the same is true among the T1R1, T1R2, and T1R3 genes of vertebrates. Specifically, we used a recently improved branch-site likelihood method (Zhang, Nielsen, and Yang 2005) to test positive selection in the three interior branches that connect the three T1R gene clades (fig. 1). We assigned these three branches as foreground branches and all other branches in the tree as background branches. The test result (table 2) showed significant signals of positive selection in the foreground branches. Twenty-two sites were identified to be under positive selection with posterior probabilities >0.95, computed by the conservative Bayes empirical Bayes method (Yang, Wong, and Nielsen 2005; Zhang, Nielsen, and Yang 2005). To see whether the positive selection was present in one, two, or all three of the three foreground branches, we conducted three additional tests. In each test, one of the three branches was assigned as the foreground branch, whereas all other branches in the tree were assigned as background branches. The results showed that each of the three branches that connect the three T1R gene clades experienced positive selection. Thus, although T1R genes do not duplicate as frequently as T2R genes do, the divergences among duplicate genes were similarly governed by positive Darwinian selection in the two families. Furthermore, we found that 18 of the 22 inferred positively selected sites are located in the N-terminal extracellular domain, which is believed to be the ligand-binding domain (Pin, Galvez, and Prezeau 2003). Similar findings of positively selected sites in this domain were previously reported for V2R vomeronasal receptors, which are evolutionarily related to T1Rs (Emes et al. 2004; Yang et al. 2005). It has been shown that the N-terminal domain is necessary for the sweet taste receptor T1R2 to bind sweet substances such as aspartame and neotame (Jiang et al. 2004; Xu et al. 2004). It is possible that the positively selected amino acid substitutions that occurred shortly after the origin of T1R2 were responsible for the emergence of this receptor's specificity toward certain sweet tastants. When the ancestral T1R3 branch was examined, 20 sites were found to be under

Table 2
Detection of Positive Selection in the Divergences of T1R Paralogous Genes

Foreground Branches[a]	$2\Delta l$[b]	Estimates of the Parameters in the Modified Model A[c]	Positively Selected Sites[d]
Three ancestral branches	17.8 ($P = 2 \times 10^{-5}$)	$p_0 = 0.444$, $p_1 = 0.233$, $p_{2a} = 0.212$, $p_{2b} = 0.111$, $\omega_0 = 0.165$, $\omega_2 = 4.05$	58L, **67P**, 77L, 80A, 123S, 165K, **291S**, **369C**, **415H**, **417Q**, **427S**, **486L**, 489W, 490V, **491W**, 497P, 572D, **576T**, *687T*, *715G*, *799R*, ***846A***
Ancestral T1R1 branch	25.48 ($P = 4 \times 10^{-7}$)	$p_0 = 0.566$, $p_1 = 0.304$, $p_{2a} = 0.085$, $p_{2b} = 0.046$, $\omega_0 = 0.168$, $\omega_2 > 999$	None
Ancestral T1R2 branch	27.08 ($P = 2 \times 10^{-7}$)	$p_0 = 0.590$, $p_1 = 0.321$, $p_{2a} = 0.057$, $p_{2b} = 0.031$, $\omega_0 = 0.168$, $\omega_2 > 999$	None
Ancestral T1R3 branch	30.72 ($P = 3 \times 10^{-8}$)	$p_0 = 0.487$, $p_1 = 0.258$, $p_{2a} = 0.167$, $p_{2b} = 0.088$, $\omega_0 = 0.168$, $\omega_2 = 503$	**77L**, **80A**, 165K, 382D, 415H, 417Q, 427S, 486L, **490V**, **491W**, 572D, **576T**, ***687T***, *691Q*, *709L*, *715G*, *745E*, *765S*, *773H*, ***799R***

[a] The foreground branches are either one or all three of the interior branches that connect the three T1R gene clades in figure 1.

[b] Twice the difference between the log likelihood of the alternative model and that of the null model. The modified model A is used as the alternative model. The modified model A with ω_2 fixed at 1 is the null model.

[c] Here, ω values are the nonsynonymous/synonymous rate ratios. p_0 is the proportion of codons that have ω_0 in all branches, p_1 is the proportion of codons that have $\omega_1 = 1$ in all branches, p_{2a} is the proportion of codons that have ω_0 in the background branches but ω_2 in the foreground branches, and p_{2b} is the proportion of codons that have ω_1 in the background branches but ω_2 in the foreground branches. Note that as long as ω_2 significantly exceeds 1 (as indicated by the likelihood ratio test), its exact value has little biological meaning due to the large estimation error.

[d] Sites with the Bayes empirical Bayes posterior probabilities higher than 95% are given (>99% are in bold). Sites in the extracellular N-terminus are shown in roman typeface, while those in the rest of the sequences are italic. It is possible that some of these sites are not subject to positive selection because the posterior probabilities are computed under a set of assumptions that may be inappropriate for T1R genes.

positive selection. Among them, eight are in the non–N-terminal regions. Interestingly, several studies showed that the non–N-terminal regions of T1R3 are required for its recognition of sweetener cyclamate and sweet taste inhibitor lactisole (Jiang et al. 2004, 2005; Xu et al. 2004). Thus, our sequence analysis is consistent with the experimental results and may help further unravel the key residues of each receptor's specificity. However, we caution that although the improved branch-site likelihood method for detecting positive selection was shown to be reliable in computer simulation, the evolutionary patterns in T1R genes could be quite different from the conditions examined in the simulation, and it remains possible that the method makes false detections of positive selection (Zhang, Nielsen, and Yang 2005).

Discussion

In this study, we identified T1R sweet/umami taste receptor genes and T2R bitter taste receptor genes from available vertebrate genome sequences. Our subsequent phylogenetic analysis showed contrasting evolutionary modes between these two gene families, which play vital roles in perceiving nutrients and toxins, respectively. Figure 4 summarizes the variation in gene number across vertebrates for T1Rs and T2Rs. It can be seen that the size of the T1R sweet/umami taste receptor gene family rarely changes across vertebrates. This observation may reflect the necessity of both sweet and umami tastes and the relative constancy in the number and type of these tastants in many vertebrates. Behavioral appetite to sweet and umami stimuli is well documented among vertebrates, including fishes (Hoar and Randall 1971), amphibians (Yoshii et al. 1982; Feder and Burggren 1992), birds (Whittow 2000), and mammals (Lindemann 2001). Consistent with these data, we found that T1R1, T1R2, and T1R3 diverged before the separation of teleost fish and tetrapods. In addition to the western clawed frog, which does not have any T1R genes, loss of a T1R gene was identified only in the chicken genome. Because fish T1R2 genes would be more difficult to detect than chicken T1R2 when mammalian T1R2 genes are used as queries, our observation is unlikely due to methodological problems. Furthermore, the chicken draft genome sequence has 6.6× coverage, and the flanking genes of mouse T1R2 have been identified in the syntenic regions in the chicken genome. Thus, we believe that chicken has lost the T1R2 gene. Interestingly, the electrophysiological study showed that chicken taste buds do not generate impulses when washed with sucrose or saccharine (Duncan 1960; Halpern 1962; Thomson 1964). Moreover, chicken has no preference of sweet stimuli in behavioral reaction tests (Ganchrow, Steiner, and Bartana 1990). Thus, the loss of T1R2 provides the genetic explanation of chicken's indifference to sweet tastants. It would be interesting to investigate when and why chicken lost the sweet taste. Similarly, we found strong evidence that the western clawed frog has lost all T1R genes. It was recently reported that the T1R2 gene is inactivated in cats and related species (Li et al. 2005), explaining why these carnivores cannot sense the sweet taste. Apparently, pseudogenization of T1R2 occurred multiple times in evolution.

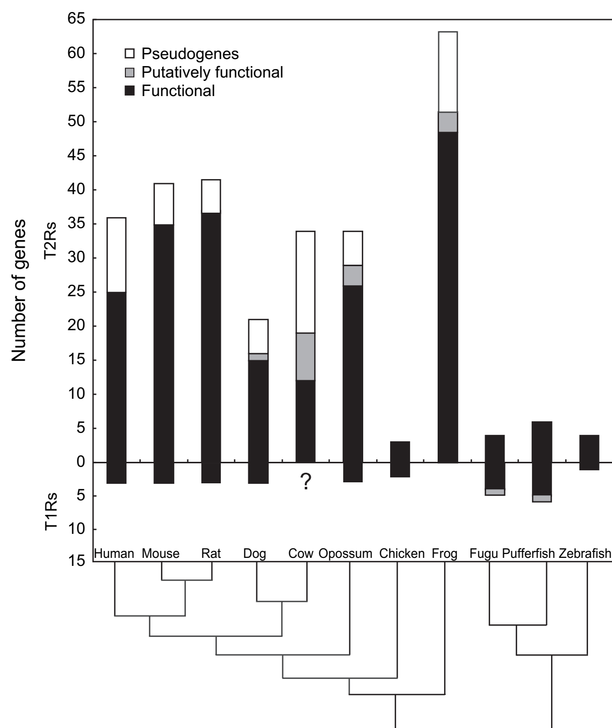

FIG. 4.—Variations of the numbers of functional genes, putative functional genes, and pseudogenes in the T1R and T2R families of 11 vertebrates. The phylogenetic relationships among the species are shown by the tree.

In contrast to T1Rs, the T2R family has remarkable expansions in frogs and mammals. We found only four to six T2R genes in the fishes and three in the chicken, in comparison to 21–42 in mammals and 49 in the frog. Phylogenetic analysis suggests that the T2R repertoire expanded in the common ancestor of tetropods, followed by further expansions in frogs and mammals, with prominent patterns of lineage-specific gene duplication. Interestingly, lineage-specific gene duplications have also been described in other mammalian chemoreception receptor families such as vomeronasal receptors and olfactory receptors (Zhang and Firestein 2002; Grus and Zhang 2004; Grus et al. 2005; Shi et al. 2005; Yang et al. 2005). It is unknown what factors have promoted the dramatic expansions of the T2R gene family in frogs and mammals. It is interesting to note that while T2R-mediated bitter perception is the only bitter transduction known in mammals, some nonmammalian vertebrates have other bitter transduction pathways and thus may be less dependent on T2Rs. For example, quinine, a bitter substance for various vertebrates, can activate a cation channel in bullfrog taste receptor cells (Tsunenari et al. 1996). In catfish, taste responses to L-arginine, a bitter substance, are mediated by nonselective cation channels (Kumazawa, Brand, and Teeter 1998; Bigiani, Ghiaroni, and Fieni 2003). In addition, the expansion of the T2R family may reflect the ecological and dietary diversity and complexity in frogs and mammals, although more studies are needed to test this hypothesis.

The T2R gene repertoire also varies considerably among mammals in terms of gene number and protein

sequences. In humans, T2R genes are under relaxed functional constraints and several have become or are becoming pseudogenes, possibly because humans are less dependent on the bitter taste to avoid ingesting toxic food (Wang, Thomas, and Zhang 2004). Thus, humans have a high proportion of T2R pseudogenes. Interestingly, cows and humans have similar total numbers of T2R genes, but cows have a higher proportion of pseudogenes than humans do. This could be due to the high detoxification capacity of cow's rumen microbes. In other words, detecting poisons in diet is not as important in ruminants as in other mammals. Indeed, ruminants generally are more tolerant to plant poisons than nonruminants (Freeland and Janzen 1974). It appears that the omnivorous mammals have the largest T2R gene repertoire and the lowest proportion of pseudogenes. In nature, poisonous substances generally taste bitter and are unevenly distributed in plant and animal tissues. Food chemistry, toxicology, and feeding ecology studies showed that bitter compounds are more common in plant than in animal tissues (Glendinning 1994). Thus, mammals in different trophic groups have different probabilities of encountering various kinds of bitter compounds. Because of consuming both animal and plant tissues, omnivorous mammals may encounter more toxic compounds than herbivorous mammals do, which in turn encounter more toxins than carnivores do. Indeed, our results showed that dog, a carnivore, has a small number of functional T2R genes. The lower number of functional T2R genes in cow than in dog may be due to the additional factor of rumination. The less complete cow genome sequence may add additional complexities to this comparison. It would be interesting to test these hypotheses when the T2R repertoire is known from more mammalian species.

Supplementary Material

Supplementary data sets 1 and 2 and Supplementary Fig. 1 are available at *Molecular Biology and Evolution* online (http://www.mbe.oxfordjournals.org/).

Supplementary data set 1: Amino acid sequences of newly identified T1R genes.

Supplementary data set 2: Amino acid sequences of newly identified T2R genes.

Supplementary Fig. 1: Neighbor-joining tree of 216 functional T2R genes from 10 vertebrates. The tree is reconstructed with protein Poisson distances. Arrow indicates the root of the tree, which is determined by using vertebrate V1R genes as outgroups (Grus et al. 2005). The teleost fishes, frog, chicken, and mammals are represented by squares, triangles, diamonds, and circles, respectively. The same tree without bootstrap information is presented in figure 2.

Acknowledgments

We thank Soochin Cho, Wendy Grus, Naruya Saitou, Xiaoxia Wang, and an anonymous reviewer for valuable comments. This study was supported by research grants from National Institutes of Health and 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., W. Gish, W. Miller, E. W. Myers, and D. J. Lipman. 1990. Basic local alignment search tool. J. Mol. Biol. **215**:403–410.

Bigiani, A., V. Ghiaroni, and F. Fieni. 2003. Channels as taste receptors in vertebrates. Prog. Biophys. Mol. Biol. **83**: 193–225.

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.

Conte, C., M. Ebeling, A. Marcuz, P. Nef, and P. J. Andres-Barquin. 2003. Evolutionary relationships of the Tas2r receptor gene families in mouse and human. Physiol. Genomics **14**:73–82.

Duncan, C. J. 1960. The sense of taste in birds. Ann. Appl. Biol. **48**:409–414.

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

Feder, M. E., and W. W. Burggren. 1992. Environmental physiology of the amphibians. University of Chicago Press, Chicago, Ill.

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

Fischer, A., Y. Gilad, O. Man, and S. Paabo. 2005. Evolution of bitter taste receptors in humans and apes. Mol. Biol. Evol. **22**:432–436.

Freeland, W. J., and D. H. Janzen. 1974. Strategies in herbivory by mammals: the role of plant secondary compounds. Am. Nat. **108**:889–894.

Ganchrow, J. R., J. E. Steiner, and A. Bartana. 1990. Behavioral reactions to gustatory stimuli in young chicks (*Gallus gallus domesticus*). Dev. Psychobiol. **23**:103–117.

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

Go, Y., Y. Satta, O. Takenaka, and N. Takahata. 2005. Lineage-specific loss of function of bitter taste receptor genes in humans and nonhuman primates. Genetics **170**:313–326.

Grus, W. E., P. Shi, Y. P. Zhang, and J. Zhang. 2005. Dramatic variation of the vomeronasal pheromone receptor gene repertoire among five orders of placental and marsupial mammals. Proc. Natl. Acad. Sci. USA **102**:5767–5772.

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

Halpern, B. P. 1962. Gustatory nerve impulses in the chicken. Am. J. Physiol. **203**:541–544.

Herness, M. S., and T. A. Gilbertson. 1999. Cellular mechanisms of taste transduction. Annu. Rev. Physiol. **61**:873–900.

Hoar, W. S., and D. J. Randall. 1971. Fish physiology. Academic Press, New York.

Hoon, M. A., E. Adler, J. Lindemeier, J. F. Battey, N. J. Ryba, and C. S. Zuker. 1999. Putative mammalian taste receptors: a class of taste-specific GPCRs with distinct topographic selectivity. Cell **96**:541–551.

Jiang, P., M. Cui, B. Zhao, Z. Liu, L. A. Snyder, L. M. Benard, R. Osman, R. F. Margolskee, and M. Max. 2005. Lactisole interacts with the transmembrane domains of human T1R3 to inhibit sweet taste. J. Biol. Chem. **280**:15238–15246.

Jiang, P., Q. Ji, Z. Liu, L. A. Snyder, L. M. Benard, R. F. Margolskee, and M. Max. 2004. The cysteine-rich region of T1R3 determines responses to intensely sweet proteins. J. Biol. Chem. **279**:45068–45075.

Kare, M. R., and W. Medway. 1959. Discrimination between carbohydrates by the fowl. Poult. Sci. **38**:1119–1127.
Kinnamon, S. C., and R. F. Margolskee. 1996. Mechanisms of taste transduction. Curr. Opin. Neurobiol. **6**:506–513.
Kitagawa, M., Y. Kusakabe, H. Miura, Y. Ninomiya, and A. Hino. 2001. Molecular genetic identification of a candidate receptor gene for sweet taste. Biochem. Biophys. Res. Commun. **283**:236–242.
Kumar, S., K. Tamura, I. B. Jakobsen, and M. Nei. 2001. MEGA2: molecular evolutionary genetics analysis software. Bioinformatics **17**:1244–1245.
Kumazawa, T., J. G. Brand, and J. H. Teeter. 1998. Amino acid-activated channels in the catfish taste system. Biophys. J. **75**:2757–2766.
Li, X., W. Li, H. Wang et al. (11 co-authors). 2005. Pseudogenization of a sweet-receptor gene accounts for cats' indifference toward sugar. PLoS Genet. **1**:e3.
Lindemann, B. 2001. Receptors and transduction in taste. Nature **413**:219–225.
Matsunami, H., J. P. Montmayeur, and L. B. Buck. 2000. A family of candidate taste receptors in human and mouse. Nature **404**:601–604.
Max, M., Y. G. Shanker, L. Huang, M. Rong, Z. Liu, F. Campagne, H. Weinstein, S. Damak, and R. F. Margolskee. 2001. Tas1r3, encoding a new candidate taste receptor, is allelic to the sweet responsiveness locus Sac. Nat. Genet. **28**:58–63.
Meyerhof, W. July 20, 2005. Elucidation of mammalian bitter taste. Rev. Physiol. Biochem. Pharmacol. 10.1007/s10254-005-0041-0.
Mombaerts, P. 2004. Genes and ligands for odorant, vomeronasal and taste receptors. Nat. Rev. Neurosci. **5**:263–278.
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.
Mueller, K. L., M. A. Hoon, I. Erlenbach, J. Chandrashekar, C. S. Zuker, and N. J. Ryba. 2005. The receptors and coding logic for bitter taste. Nature **434**:225–229.
Murphy, W. J., P. A. Pevzner, and S. J. O'Brien. 2004. Mammalian phylogenomics comes of age. Trends Genet. **20**:631–639.
Nei, M., and S. Kumar. 2000. Molecular evolution and phylogenetics. Oxford University, New York.
Nelson, G., J. Chandrashekar, M. A. Hoon, L. Feng, G. Zhao, N. J. P. Ryba, and C. S. Zuker. 2002. An amino-acid taste receptor. Nature **416**:199.
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.
Parry, C. M., A. Erkner, and J. le Coutre. 2004. Divergence of T2R chemosensory receptor families in humans, bonobos, and chimpanzees. Proc. Natl. Acad. Sci. USA **101**:14830–14834.
Pfister, P., and I. Rodriguez. 2005. Olfactory expression of a single and highly variable V1r pheromone receptor-like gene in fish species. Proc. Natl. Acad. Sci. USA **102**:5489–5494.
Pin, J. P., T. Galvez, and L. Prezeau. 2003. Evolution, structure, and activation mechanism of family 3/C G-protein-coupled receptors. Pharmacol. Ther. **98**:325–354.
Roper, S. D. 1989. The cell biology of vertebrate taste receptors. Annu. Rev. Neurosci. **12**:329–353.
Sainz, E., J. N. Korley, J. F. Battey, and S. L. Sullivan. 2001. Identification of a novel member of the T1R family of putative taste receptors. J. Neurochem. **77**:896–903.
Saitou, N., and M. Nei. 1987. The neighbor-joining method: a new method for reconstructing phylogenetic trees. Mol. Biol. Evol. **4**:406–425.
Shi, P., J. P. Bielawski, H. Yang, and Y. P. Zhang. 2005. Adaptive diversification of vomeronasal receptor 1 genes in rodents. J. Mol. Evol. **60**:566–576.
Shi, P., J. F. Huang, and Y. P. Zhang. 2005. Bitter and sweet/umami taste receptors with differently evolutionary pathways. Acta Genet. Sin. **32**:346–353.
Shi, P., J. Zhang, H. Yang, and Y. P. Zhang. 2003. Adaptive diversification of bitter taste receptor genes in mammalian evolution. Mol. Biol. Evol. **20**:805–814.
Sonnhammer, E. L., G. von Heijne, and A. Krogh. 1998. A hidden Markov model for predicting transmembrane helices in protein sequences. Proc. Int. Conf. Intell. Syst. Mol. Biol. **6**:175–182.
Speca, D. J., D. M. Lin, P. W. Sorensen, E. Y. Isacoff, J. Ngai, and A. H. Dittman. 1999. Functional identification of a goldfish odorant receptor. Neuron **23**:487–498.
Thompson, J. D., T. J. Gibson, F. Plewniak, F. Jeanmougin, and D. G. Higgins. 1997. The CLUSTAL_X windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. Nucleic Acids Res. **25**:4876–4882.
Thomson, A. L. 1964. A new dictionary of birds. McGraw-Hill, New York.
Tsunenari, T., Y. Hayashi, M. Orita, T. Kurahashi, A. Kaneko, and T. Mori. 1996. A quinine-activated cationic conductance in vertebrate taste receptor cells. J. Gen. Physiol. **108**:515–523.
Wang, X., S. D. Thomas, and J. Zhang. 2004. Relaxation of selective constraint and loss of function in the evolution of human bitter taste receptor genes. Hum. Mol. Genet. **13**:2671–2678.
Whittow, G. C. 2000. Sturkie's avian physiology. Academic Press, San Diego, Calif.
Xu, H., L. Staszewski, H. Tang, E. Adler, M. Zoller, and X. Li. 2004. Different functional roles of T1R subunits in the heteromeric taste receptors. Proc. Natl. Acad. Sci. USA **101**:14258–14263.
Yang, H., P. Shi, Y. P. Zhang, and J. Zhang. 2005. Composition and evolution of the V2r vomeronasal receptor gene repertoire in mice and rats. Genomics **86**:306–315.
Yang, Z., and R. Nielsen. 2002. Codon-substitution models for detecting molecular adaptation at individual sites along specific lineages. Mol. Biol. Evol. **19**:908–917.
Yang, Z., W. S. Wong, and R. Nielsen. 2005. Bayes empirical Bayes inference of amino acid sites under positive selection. Mol. Biol. Evol. **22**:1107–1118.
Yoshii, K., C. Yoshii, Y. Kobatake, and K. Kurihara. 1982. High sensitivity of *Xenopus* gustatory receptors to amino acids and bitter substances. Am. J. Physiol. **243**:R42–R48.
Zhang, J. 2004. Frequent false detection of positive selection by the likelihood method with branch-site models. Mol. Biol. Evol. **21**:1332–1339.
Zhang, J., R. Nielsen, and Z. Yang. August 17, 2005. Evaluation of an improved branch-site likelihood method for detecting positive selection at the molecular level. Mol. Biol. Evol. 10.1093/molbev/msi237.
Zhang, X., and S. Firestein. 2002. The olfactory receptor gene superfamily of the mouse. Nat. Neurosci. **5**:124–133.
Zhao, G. Q., Y. Zhang, M. A. Hoon, J. Chandrashekar, I. Erlenbach, N. J. Ryba, and C. S. Zuker. 2003. The receptors for mammalian sweet and umami taste. Cell **115**:255–266.

Naruya Saitou, Associate Editor

Accepted September 26, 2005