The Unusual Adaptive Expansion of Pancreatic Ribonuclease Gene in Carnivora

Li Yu† and Ya-ping Zhang*†*

*Laboratory for Conservation and Utilization of Bio-resource, Yunnan University, Kunming, China and †Laboratory of Molecular Biology of Domestic Animals, and Cellular and Molecular Evolution, Kunming Institute of Zoology, Kunming, China

Pancreatic ribonuclease (RNASE1) is a digestive enzyme that has been recognized to be one of the most attractive model systems for molecular evolutionary studies. The contribution of RNASE1 gene duplication to the functional adaptation of digestive physiology in foregut-fermenting herbivores, mostly in ruminants, has been well documented. However, no one has ever done a comprehensive study on the carnivores, which are sister to the artiodactyls. Here, we sequenced this gene from 15 species of the superfamily Caniformia in order Carnivora, which all have a relatively simple digestive system and lack the microbial digestion in rumen or cecum typical of most herbivores. In contrast to our initial expectation that only a single RNASE1 gene is present in these carnivores, we observed a "birth (gene duplication)-and-death (gene deactivation)" process for the evolution of RNASE1 genes in all 3 species of Mustelidae family examined here, adding the growing diversity of RNASE1 gene family evolution. In addition, bursts of positive selection have been shown to contribute the enigmatic diversification of these RNASE1 genes in Mustelidae. The finding of the adaptive expansion of RNASE1 in animals without foregut fermentation provides another opportunity for further studies of the structure, function, and evolution of this gene, raising the possibility that new tissue specificity or other functions of RNASE1 genes might have developed in these species.

Introduction

Pancreatic ribonuclease (RNASE1), the prototype of a mammalian RNASE1 superfamily, is a digestive enzyme secreted by the pancreas of almost all vertebrates. It has been and is still being recognized to be one of the most attractive model systems for molecular evolutionary studies. The highest levels of expression are found in ruminants having multicompartmentalized stomachs and species with ruminant-like or cecal digestions, which can be attributed to the need to deal with large quantities of RNA released from the microflora of the stomach or cecum of these herbivores (Barnard 1969; Beintema et al. 1973, 1977; Beintema et al. 1986; Beintema 1990). It is also clear that gene duplication events have occurred independently in some of these species, including ox, sheep (Breukelman et al. 1998, 2001), camel (Kleineidam et al. 1999), rat (Dubois et al. 2002), guinea pigs (Beintema and Neuteboom 1983), and elephants (Dubois et al. 2003). The most compelling example demonstrating the contribution of RNASE1 gene duplication to the functional adaptation of digestive physiology comes from recent evolutionary and functional analyses of the duplicated pancreatic RNASE1 gene in Asian and African leaf-eating colobine monkeys, which have ruminant-type alimentary tract (Zhang et al. 2002; Zhang 2006). Evolving rapidly under positive selection for enhanced ribonucleolytic activity in an altered microenvironment, these duplicated RNASE1 genes (*RNASE1B* in Asia and *RNASE1*β and *RNASE1*γ in African colobines) arose as an adaptive response to the increased demands for the enzyme for digesting bacterial RNA (Zhang et al. 2002; Zhang 2006).

RNASE1 gene has been extensively studied in foregut-fermenting animals, particularly in ruminants; however, no one has ever done a comprehensive study on the carnivores, which are sister to the artiodactyls. The mammalian order Carnivora is characterized by a diverse array of feeding habits, for example, carnivorous weasels, omnivorous bears, and entirely herbivorous pandas (Nowak 1991); however, it has been suggested that they all have a relatively simple digestive system and lack the microbial digestion in rumen or cecum typical of most herbivores mentioned above (Schaller et al. 1989). Pancreatic RNASE1s are expressed at an extremely low level in the pancreas of carnivoran species (Beintema and Lenstra 1982; Beintema et al. 1988). Therefore, although the carnivoran RNASE1s have not yet been characterized by far at the molecular level, one might predict that only a single RNASE1 gene is present in carnivoran species, as were expected of nonruminant artiodactyls and other mammalian species investigated before (Dubois et al. 1999, 2003; Kleineidam et al. 1999). However, our study revealed that such was not the case.

Pancreatic RNASE1 genes from 15 species of the superfamily Caniformia in Carnivora were analyzed. Surprisingly, in all species of Mustelidae family examined, we observed a birth (gene duplication)-and-death (gene deactivation) process for the evolution of RNASE1 genes in contrast to only one pancreatic RNASE1 gene in other carnivores, adding the growing diversity of this gene family evolution. Gene duplication has long been thought as the primary source for evolutionary innovations and functional adaptations (Ohno 1970; Clegg et al. 1997; Force et al. 1999; Zhang 2003). Thus, the characterization of the molecular evolution of RNASE1 genes in Mustelidae family of Carnivora can be of great importance, in view of learning what evolutionary force has possibly shaped the enigmatic diversification of these RNASE1 genes in them and what the implications with respect to biological significance might be. We started the study by phylogenetic reconstruction of all the carnivoran RNASE1 genes and then discussed possible selection pressures responsible for the scenarios of gene duplications and deactivations in Mustelidae species through evolutionary analyses.

Key words: RNASE1, Carnivora, Mustelidae, enzyme activity; positive selection.

E-mail: zhangyp1@263.net.cn, zhangyp@mail.kiz.ac.cn.

Mol. Biol. Evol. 23(12):2326–2335. 2006
doi:10.1093/molbev/msl101
Advance Access publication August 31, 2006

Materials and Methods

DNA Samples and Polymerase Chain Reaction Amplifications

Fifteen species belonging to 6 families of Caniformia of order Carnivora were examined in this study. These

species encompass a broad range of feeding type, from pure meat eaters to frugivores and plant specialist. The 6 families are Procyonidae, Mustelidae, Canidae, Ursidae, Otariidae, and Phocidae. Among these 15 species, the classifications of the 2 pandas have been evolutionary puzzles over the past decades. Here, we assigned the giant panda (*Ailuropoda melanoleuca*) to Ursidae in view of the growing evidence supporting it as the most basal offshoot of that family, whereas the red panda (*Ailurus fulgens*) was kept at an orphan position owing to intense controversy in previous studies. At least 2 species from each of the 6 families except Phocidae are included. Procyonidae were represented by the raccoon (*Procyon lotor*), the coatimundi (*Nasua nasua*); Mustelidae by the marten (*Martes flavigula*), the yellow-bellied weasel (*Mustela kathiab*), and the hog badger (*Arctonyx collaris*); Canidae by the dog (*Canis familiaris*), the red wolf (*Canis rufus*), and the gray wolf (*Canis lupus*); Ursidae by the Asiatic black bear (*Ursus thibetanus*), the spectacled bear (*Tremarctos ornatus*), and the giant panda (*A. melanoleuca*); Otariidae by the sea lion (*Zalophus californianus*) and the fur seal (*Callorhinus ursinus*); and Phocidae by the bearded seal (*Erignathus barbatus*). For each sample, total genomic DNA was isolated from blood or frozen tissues using standard proteinase K, phenol/chloroform extraction (Sambrook et al. 1989).

Pancreatic RNASE1 genes with a length of about 420 bp, which span nearly the entire region coding sequence for the signal and mature peptides, were amplified by polymerase chain reaction (PCR) using primer pair CarRnF (5′-TCCTTGCCGGGAGAATTG-3′) and CarRnR (5′-AAGCGGTGAGGTGGCTAA-3′). The primers were designed based on the conserved regions from comparisons of RNASE1 sequences from available human, rat, mouse, and dog genomes. The optimal conditions adopted in PCR reactions were 95 °C initial denaturation for 5 min, 35 cycles of 94 °C denaturation for 1 min, 50 °C annealing for 1 min, and 72 °C extension for 1 min, followed by a final 72 °C extension for 10 min. For a few species from which amplifications of pancreatic RNASE1 gene using CarRnF and CarRnR are weak or fail, alternative primer pair CarRnF1 (5′-AGGCGGGAGCTGGGGTTCT-3′) and CarRnR1 (5′-AGGCATCAAAGTGGACTGG-3′) was employed in the same PCR condition. The new determined RNASE1 sequences are in GenBank under accession numbers DQ462206–DQ462222 and DQ462226–DQ462230.

Cloning and Sequencing

The amplified PCR products were cloned into PMD18-T vector (Takara, China) and transformed into an ultracompetent *Escherichia coli* cell (Takara, China). Plasmids containing the RNASE1 inserts were extracted using GenElute Plasmid Miniprep Kit (Sigma-Aldrich Co., Shanghai, China). About 20 clones per ligation reaction were sequenced in both directions with an ABI PRISM 3700 DNA sequencer using standard protocols provided by the manufacturer (PE Biosystems, Foster City, CA). Only those sequences with more than 3 mutations in protein sequence and corroborated by at least 2 times of independent amplification and sequencing were used in the analysis.

Phylogenetic Reconstruction and Selection Constraint Analyses

The nearly complete CDS region of 23 RNASE1 genes (2 of them are pseudogenes) corresponding to a 28-residue signal peptide and a 112-residue mature peptide were aligned with ClustalX program (Thompson et al. 1997) and thereafter refined by eye. Only gene regions coding for the mature peptide (336 nt) were analyzed.

Phylogenetic trees were reconstructed using MEGA3 (Kumar et al. 2004) for Neighbor-Joining (NJ) analysis based on Kimura 2-parameters model, as well as PAUP*4.0b8 (Swofford 2001) for maximum likelihood (ML) analysis. In ML analyses, the best-fitting models of sequence evolution were estimated using program Modeltest 3.06 (Posada and Crandall 1998). The reliability of the tree topologies was evaluated using bootstrap support (BS; Felsenstein 1985) with 1,000 replicates for NJ and 100 replicates for ML analysis. The phylogeny was estimated for the data set including the pseudogenes or not.

The molecular clock hypothesis was examined by the relative-rate test of Li and Bousquet (1992) using the program RRTREE (Robinson-Rechavi and Huchon 2000). One advantage of this distance-based method is that it can divide substitution rates into nonsynonymous and synonymous rates for comparisons.

The nonsynonymous to synonymous rate ratio ω (dn/ds) provided an indication of the change of selective pressures. A dn/ds ratio = 1, <1, and >1 will indicate neutral evolution, purifying selection, and positive selection on the protein involved, respectively. The number of synonymous substitutions per synonymous site (ds) and the number of nonsynonymous substitutions per nonsynonymous site (dn) in pairwise comparisons between nucleotide sequences was first estimated in MEGA3 (Kumar et al. 2004) using the modified Nei–Gojobori method (Zhang et al. 1998) with Jukes–Cantor correction. In addition, we adopt several likelihood models with different assumptions of evolution of these RNASE1 genes based on the codon-based likelihood analysis implemented in CODEML program of PAML package (Yang 2000). All models correct the transition/transversion rate and codon-usage biases (F3 × 4). Different starting ω values were also used to avoid the local optima on the likelihood surface (Suzuki and Nei 2001).

In the ML method, site-specific models, which allow for variable selection patterns among amino acid sites, M1a, M2a, M8a, and M8, were used to test for the presence of sites under positive selection and identify them. We construct 2 likelihood ratio tests (LRT), which compare M1a with M2a and M8a with M8. Significant differences between the 2 models were evaluated by calculating twice the log-likelihood difference following a χ^2 distribution, with the number of degrees of freedom equal to the difference in the numbers of free parameters between the 2 models. M2a and M8 models allow for positively selected sites. When these 2 positive-selection models fitted the data significantly better than the corresponding null models (M1a and M8a), the presence of sites with $\omega > 1$ is suggested. The conservative Empirical Bayes approach (Yang et al. 2005) will then be used to calculate the posterior probabilities of a specific codon site and identify those most likely to

be under positive selection. In addition, considering that positive selection may act in very short episodes during the evolution of a protein (Gillespie 1991) and affect only a few sites along a few lineages in the phylogeny, recently developed likelihood models accommodating ω ratios to vary both among lineages of interest and amino acid sites, that is, an improved version of "branch-site" model, were also considered here (Zhang et al. 2005). We use branch-site Model A for stringent test and identification of sites under positive selection along the lineages of interest (Zhang et al. 2005). In addition, the Bonferroni correction was also applied for multiple testing in the analysis.

Results

Carnivoran RNASE1 Sequences

A total of 23 RNASE1 sequences were identified in 15 species of superfamily Caniformia examined. Aligned sequences are 420 bp in length and include the gene regions coding for the complete signal peptide (1–84 bp) and the majority of mature peptide. Only one RNASE1 gene was separately detected in 3 canids, 3 ursids, 2 otariids, 1 phocids, 2 procyonids, and the red panda, whereas in each of 3 mustelids, at least 3 RNASE1 genes and 1 pseudogene are present with the exception of the yellow-bellied weasel, from which no pseudogene has been found. The numbers of amino acid residue differences between 3 paralogous genes in each Mustelidae species were 26 in the hog badger, 22 in the yellow-bellied weasel, and 18 in the marten on average. Thus, they most likely represent distinct loci derived from gene duplications, rather than alleles of a single-gene locus. The sequences of 3 canids were found identical (fig. 1).

Sequence comparison of these carnivoran RNASE1 genes with those published in other mammalian orders indicated that there is 1 amino acid deletion (near residue 24) in the 5′ region of the mature peptide. The same deletion has only been found by far in some representatives of order Edentata (aardvark), Chiroptera (fruit bat), and Rodent (squirrel) (Dubois et al. 2003). Therefore, the amino acid residue N described hereafter in present paper corresponded to residue N + 1 in other RNASE1 studies of most species outside the order Carnivora. Our alignment showed remarkable sequence variations among these carnivoran RNASE1 genes. In all, 50.48% nucleotide sites were variable and 37.86% parsimonious informative. However, those amino acid residues prerequisite for RNASE1 activity and invariant through the whole RNASE1 superfamily are conserved (fig. 1). Of these RNASE1 genes, *Mar*-Ψ sequence in the marten and *Bad*-Ψ in the sable demonstrate typical characteristics of pseudogenes. *Mar*-Ψ contained a premature stop codon at residue 27 in mature peptide region, whereas *Bad*-Ψ contained several frameshifting deletions, including a single deletion at nucleotide 62 in signal peptide region and a single deletion of nucleotide 94 as well as a 6-base deletion of nucleotides 136–141 in mature peptide region, leading to 4 premature stop codons (data not shown). Only gene regions coding for mature peptide (336 nt) were used in our further analysis. The aligned mature peptide sequences contained 181 variable nucleotide sites (53.87%) and 137 informative sites (40.77%).

Phylogenetic Results of Carnivoran RNASE1 Genes

The ML and NJ dendrograms depicting the relationships among 21 functional carnivoran RNASE1s are shown in figure 2. We can find that the 2 trees present an identical topology and similar BSs for branches involving higher-level (i.e., interfamilial) relationships. Although the true phylogeny of caniformian families is not yet well established thus far, a series of recent studies based on both mitochondrial and nuclear DNA (Flynn and Nedbal 1998; Yu et al. 2004; Flynn et al. 2005; Yu and Zhang 2006) have provided a preferred topology, which is convincingly supported by present RNASE1 phylogenetic analysis, with the exception of a minor change in the relationship between Procyonidae, Mustelidae, and the red panda. Without regard to outgroup family Canidae, the bear family Ursidae (including the giant panda) diverged earliest within Caniformia, followed by the pinnipeds, which include families Otariidae and Phocidae. Procyonidae, Mustelidae, and the red panda form the monophyletic Musteloidea sensu stricto clade. Within the clade, Mustelidae RNASE1 genes form a moderate statistical supported group with the red panda as its closest relatives.

Below the level of family, phylogenetic subtree of the family Mustelidae receives special attention as gene duplication and deactivation (i.e., birth-and-death evolution) has occurred in 3 Mustelidae species compared with their non-mustelid homologs. Phylogenetic analysis based on NJ and ML methods both showed that 9 functional Mustelidae RNASE1 sequences did not form species-specific clusters and divided into 2 distinct groups (fig. 2). Group 1 includes 4 RNASE1 sequences, 2 from the marten, *Mar*-1 and *Mar*-3, 1 from the yellow-bellied weasel, *Wea*-1, and 1 from the hog badger, *Bad*-1, although this grouping was not supported by high-bootstrap value. Group 2 consists of the other 5 sequences supported by moderate BS, one from the marten, *Mar*-2, 2 from the yellow-bellied weasel, *Wea*-2/*Wea*-3, and 2 from the hog badger *Bad*-2/*Bad*-3. As shown in figure 2, those paralogous genes in the yellow-bellied weasel, *Wea*-2 and *Wea*-3, and those in the hog badger, *Bad*-2 and *Bad*-3, were consistently closely related in group 2.

ML and NJ analyses gave different results about the internal relationships of the RNASE1 genes within group 1. NJ analysis (fig. 2*B*) revealed sequences from group 1 clustered in the same species relationships as that from group 2, that is, the hog badger diverges first, followed by the yellow-bellied weasel and the marten. The branching order of these 3 species in Mustelidae is in agreement with previous phylogenetic results inferred from the analyses of nuclear and mitochondria genes (Yu and Zhang 2006). However, it was not the case based on ML analysis (fig. 2*A*). Despite this, the respective statistical supports for those inconsistent branches are both rather weak. The inclusion of 2 pseudogenes, *Bad*-Ψ and *Mar*-Ψ, did not alter the 2 tree topologies and the node support levels. NJ analysis groups *Bad*-Ψ with group 2 as the basal branch, whereas ML analysis suggests that it is the earliest diverging branch relative to all Mustelidae RNASE1s, but both placements of *Bad*-Ψ are not convincingly supported (ML and NJ BS $<$ 50%). In contrast, a strong indication for sister grouping of

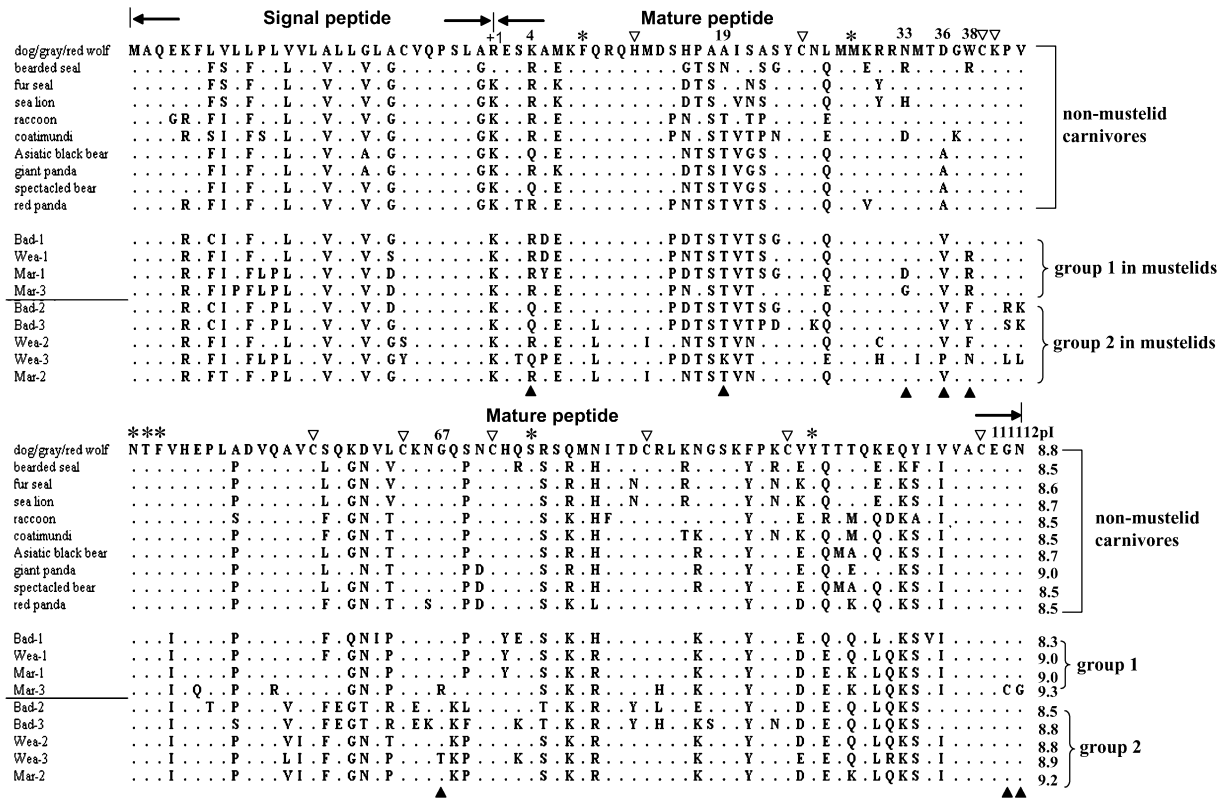


FIG. 1.—Protein sequence alignment of carnivoran RNASE1. Active-site residues and those invariant through the RNASE1 superfamily are marked with open arrows and asterisks, respectively. Eight positively selected residues identified in the study are also indicated. *Wea-1*, *Wea-2*, and *Wea-3* represent 3 RNASE1 genes in *Mustela kathiab* (the yellow-bellied weasel); *Bad-1*, *Bad-2*, and *Bad-3* in *Arctonyx collaris* (the hog badger); and *Mar-1*, *Mar-2*, and *Mar-3* in *Martes flavigula* (the marten).

Mar-Ψ and *Mar*-3 within group 1 has been obtained in both methods (ML and NJ BS = 100%). Until now, the few examples of RNASE1 pseudogenes were only reported in ruminants, elephants, and rats (Breukelman et al. 1998; Dubois et al. 2002, 2003).

From our phylogenetic analysis, we can find that multiple gene duplications and pseudogenizations have occurred during the evolutionary history of these Mustelidae RNASE1s. An initial duplication event emerged before the divergence of 3 Mustelidae species examined and after the divergence of them from the red panda. Then several additional independent duplications produced 3 different RNASE1 genes in each species of Mustelidae. In group 2, 1 duplication event occurred in the yellow-bellied weasel and the hog badger, respectively, whereas in group 1 from NJ analysis, paralogs *Mar*-1 and *Mar*-3 in the marten were produced by 1 duplication event after the marten diverged. However, to explain the gene relationships within group 1 from ML analysis (fig. 2*A*), we have to either invoke at least 2 additional independent gene loss and a duplication event or assume that there are more yet unidentified orthologs present in the yellow-bellied weasel and hog badger. Therefore, NJ result seems more likely not only from its consistent branching order of the 3 Mustelidae species with previous phylogenetic results but also from the parsimony perspective. Anyway, in future research, it is desirable to add more sequence data for clarifying the multifurcations and disagreement in present 2 trees.

Evolutionary Rate and Selective Patterns among Carnivoran Lineages

From the carnivoran phylogeny shown in figure 2*B*, we can see that there are no significant rate fluctuations among RNASE1 genes of those nonmustelid species. In contrast, several short and long branches are evidently intermixed within Mustelidae, reflecting the impression of quite variable rate of evolution among these duplicated RNASE1 genes. The two-cluster relative-rate test, excluding those identical sequences, confirmed that. Using dog RNASE1 sequence as outgroup, we evaluate the relative rates between mustelids and other carnivores. Although all were found to evolve with similar synonymous rates ($P > 0.0011$; after Bonferroni correction; Bonferroni 1936), *Bad*-3 and *Wea*-3 in Mustelidae family evolve significantly more rapidly than others at nonsynonymous sites ($P < 0.0011$). When tests were conducted within Mustelidae using the red panda as outgroup, group 2 evolve significantly faster than group 1 at nonsynonymous sites ($P = 0.0388$), whereas 7 out of 36 nonsynonymous substitution pairwise comparisons rejected the hypothesis of equal rates at the 0.13888% level (after Bonferroni correction; table 1). Therefore, evolutionary-rate changes may have occurred following RNASE1 gene duplications in Mustelidae.

To find out whether the rapid evolution in some Mustelidae RNASE1 gene copies was driven by positive

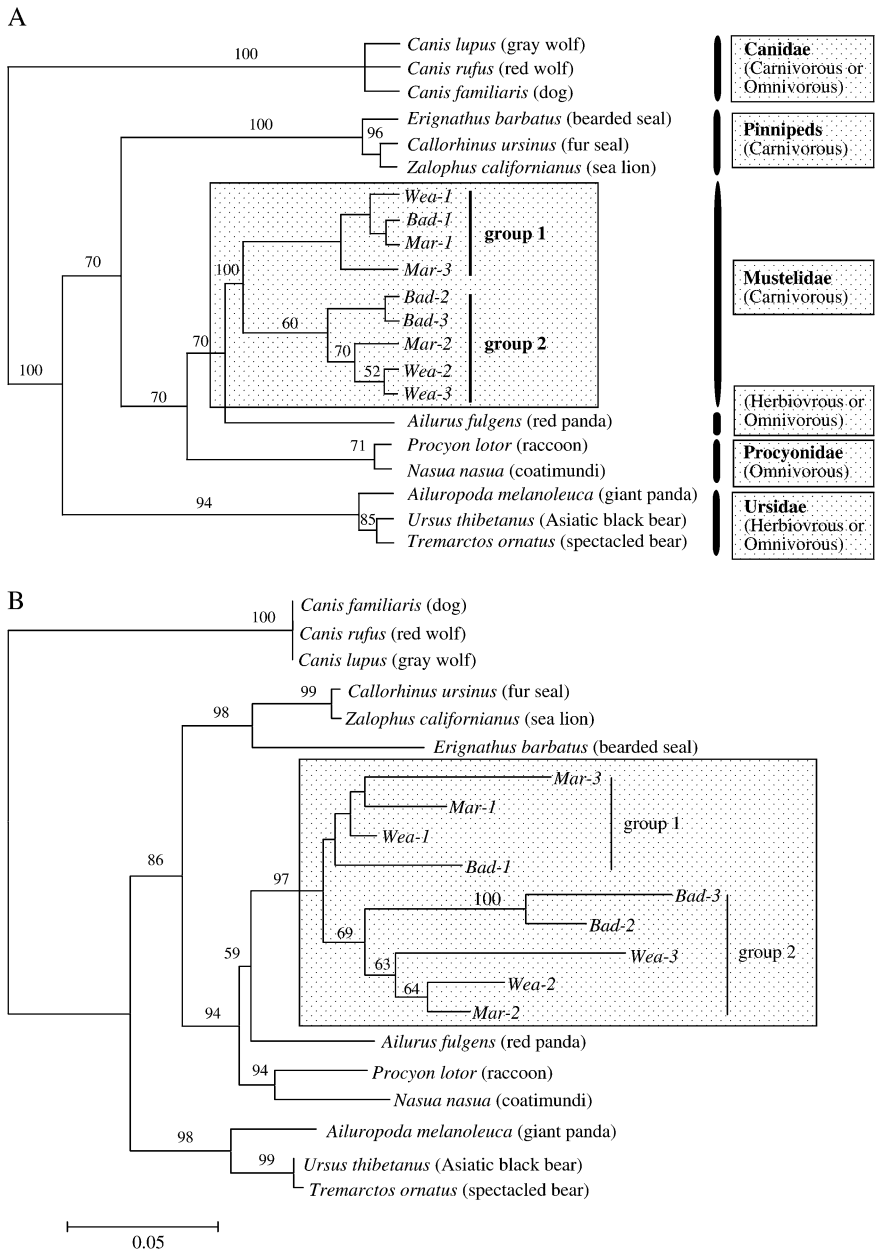


FIG. 2.—Phylogeny of 21 functional RNASE1 genes in Carnivora inferred from (*A*) ML and (*B*) NJ analyses. BS values larger than 50% are presented on the branches. Branch lengths are proportional to the number of nucleotide substitutions per sites for NJ tree.

selection or relaxation of purifying selection resulting from the gene duplication, we first applied the modified Nei–Gojobori method to estimate nonsynonymous (dn) and synonymous (ds) substitution rates for functional pairs of carnivoran RNASE1 genes. The plot of dn against ds is shown in figure 3*A*. In comparisons involving those non-mustelid carnivores, ds is higher than dn (dn/ds < 1), whereas 14 out of 36 comparisons within the mustelids have a ω (dn/ds) > 1 (fig. 3*B*). These observations indicate purifying selection is prevalent throughout carnivoran RNASE1's evolution due to functional constraint but periods of positive selection may have acted on sequences from the mustelids. Therefore, an in-depth examination of the selective agents behind multiple gene duplications in Mustelidae family was needed. The likelihood-based analysis (CODEML program in PAML) of a "small" phylogeny inferred from only 9 functional Mustelidae RNASE1 sequences (fig. 4) served the purpose. Because the likelihood analysis may be sensitive to tree topology used, so those inconsistent nodes produced by the ML and NJ phylogenies (fig. 2), which were also poorly supported from bootstrapping analysis, were collapsed into polytomy in figure 4.

Table 2 shows the results of detecting positive selection from likelihood analyses. Both positive-selection models (M2a and M8) provided a significantly better fit to the data than did the neutral models (M1a and M8a) ($P <$ 0.001) and suggested the presence of positively selected

Table 1
Significantly Heterogeneous Nonsynonymous Rates of Evolution of RNASE1 Genes in Family Mustelidae as Indicated in the 2-Cluster Test of RRTREE

Comparisons between Mustelids and Other Carnivores					
Outgroup	Lineage 1	Lineage 2	dn1	dn2	*P* Value
Dog	Other carnivores	*Bad*-3	0.16519	0.274298	0.000773
Dog	Other carnivores	*Wea-3*	0.16519	0.274465	0.000223
Comparisons between Group 1 and Group 2 within Mustelidae					
Outgroup	Lineage 1	Lineage 2	dn1	dn2	*P* Value
The red panda	Group 1	Group 2	0.0973022	0.12842	0.0388264
Comparisons between All Sequences within Mustelidae					
Outgroup	Lineage 1	Lineage 2	dn1	dn2	*P* Value
The red panda	*Bad-1*	*Bad-3*	0.0924298	0.193321	0.00066161
	Mar-1	*Bad-3*	0.0885905	0.193321	0.00020595
	Mar-2	*Bad-3*	0.0851564	0.193321	6.58E-05
	Wea-1	*Bad-3*	0.0682123	0.193321	6.45E-06
	Wea-2	*Bad-3*	0.0986609	0.193321	0.00051856
	Wea-1	*Bad-2*	0.0682123	0.150043	0.00021178
	Wea-1	*Wea-3*	0.0682123	0.152802	0.00032555

NOTE.—Red wolf and gray wolf in dog family are not included in the test owing to the sequence identity between dog, red wolf, and gray wolf.

residues (table 2). Residue 38 was consistently identified to suffer strong positive selection in M2a and M8 models with 100% posterior probability ($\omega = 11.06159$ and 10.45917, respectively).

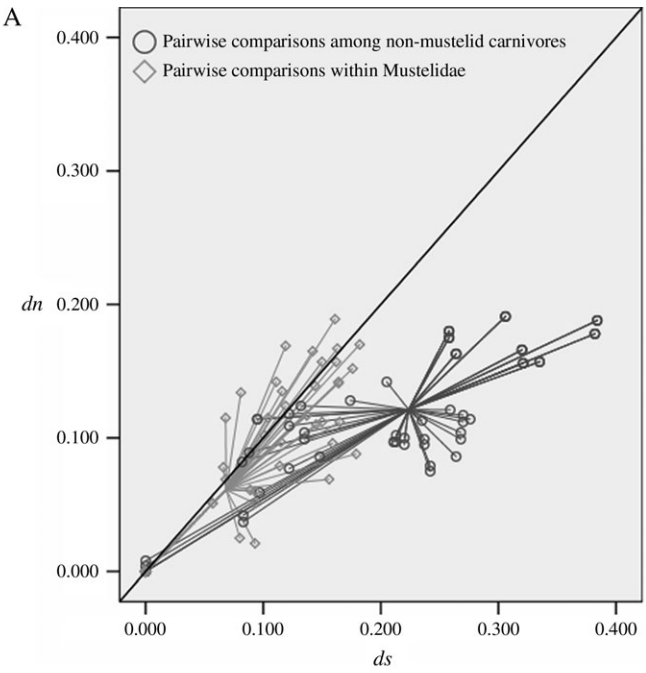
Different from the above tests estimating ω ratio across all codons in the gene, the recently developed branch-site model account for different selective pressure among both amino acid sites and lineages. Moreover, this model is considered powerful to distinguish positive selection from relaxation of purifying selection (Zhang et al. 2005). To examine the possible selective agent behind multiple RNASE1 gene duplications in Mustelidae family, we conduct LRTs based on the branch-site models for those branches resulting from gene duplications (8 such branches in total, a–h as indicated in fig. 4). The analyses suggest there is significant evidence along the ancestral branch to group 2 (branch b; $P = 0.01648$), the terminal branch to *Wea*-3 (branch f; $P = 0.00147$), and that to *Mar*-3 (branch d; $P = 0.00258$). After performing Bonferroni correction for multiple testing, we find that LRT tests are still significant in branches f and d ($P < 0.00625\%$) (table 2). Therefore, positive selection might have indeed been acted in the lineages leading to one of the yellow-bellied weasel copies, *Wea*-3 (branch f), and the marten copies, *Mar*-3 (branch d), during the Mustelidae RNASE1 gene evolution. Residue 38, which has been identified as positively selected codon in site-specific model, was again predicted for branch f in branch-site models with a high posterior probability of >95% (table 2). Beside this, residues 4, 19, 36, and 67 were also detected along this branch, whereas 33, 111, and 112 along branch d.

Discussions

The mammalian RNASE1 superfamily is an unusual group of secretory proteins characterized by the shared elements of structure and enzymatic activity but divergent expression patterns and biological functions, for example, digestion of bacteria RNA (pancreatic RNASE1), angiogenesis (angiogenins), and host defense (eosinophil cationic protein) (D'Alession and Riordan 1996; Beintema et al. 1997; Beintema and Kleineidam 1998; Cho et al. 2005). Increasing cases of demonstrating dynamic evolution of this gene superfamily have also been reported in ruminants, primates, and rodents (Singhania et al. 1999; Zhang et al. 2000, 2002; Zhang 2006).

In the present paper, we analyzed pancreatic RNASE1 gene, the prototype of this gene superfamily, in 15 species of Carnivora, adding the growing diversity of the RNASE1 gene evolution. It was rather a surprise to find that all of the 3 Mustelidae species examined here show evidence of multiple RNASE1 genes, whereas the other carnivores have 1 gene copy. Phylogenetic reconstruction of these sequences places the first gene duplication event, which produced 2 major clades corresponding to group 1 and group 2, in the common ancestor of the Mustelidae carnivores and after the divergence of the red panda (fig. 2). Subsequently, in each of the 2 clades, group 1 and group 2, multiple gene copies and pseudogenes, which did not cluster together according to species in the phylogeny, were produced by additional independent gene duplication and deactivation events. Thus, the overall phylogenetic pattern revealed that the process shaping the short-term evolution of these paralogous RNASE1 lineages in Mustelidae family is a good example of gene birth (duplication)–and-death (gene deactivation), as described in the studies of rat RNASE1 (Dubois et al. 2002), rodent eosinophil-associated RNASE (Zhang et al. 2000), and also of a few other gene families, for example, immunoglobulin (Ota and Nei 1994), the major histocompatibility complex (Piontkivska and Nei 2003), and ubiquitin (Nei et al. 2000). But contrary to the situation observed in the rat RNASE1 genes, we further show that nonsynonymous substitution rate has remarkably accelerated for some gene copies in the Mustelidae following gene duplications, and likelihood analysis confirmed that positive selection has promoted this unusual evolutionary pattern, at

A



B

	Bad-1	*Bad-2*	*Bad-3*	*Mar-1*	*Mar-2*	*Mar-3*	*Wea-1*	*Weal-2*	*Wea-3*
Bad-1									
Bad-2	0.081(0.134)								
Bad-3	0.119(0.169)	0.068(0.069)							
Mar-1	0.093(0.060)	0.165(0.112)	0.176(0.152)						
Mar-2	0.159(0.096)	0.137(0.117)	0.164(0.142)	0.156(0.069)					
Mar-3	0.116(0.135)	0.163(0.167)	0.161(0.189)	0.114(0.079)	0.115(0.097)				
Wea-1	0.057(0.051)	0.101(0.102)	0.111(0.142)	0.080(0.025)	0.093(0.051)	0.066(0.078)			
Wea-2	0.145(0.110)	0.150(0.112)	0.164(0.141)	0.179(0.088)	0.093(0.021)	0.137(0.117)	0.089(0.061)		
Wea-3	0.150(0.157)	0.162(0.157)	0.182(0.170)	0.144(0.139)	0.104(0.115)	0.142(0.165)	0.119(0.124)	0.068(0.115)	

Note: Numbers in the table mean *ds* (*dn*) for pairwise comparisons of 9 RNASE1 genes within Mustelidae.
The comparisons of dn/ds >1 (no test if they are significant >1) were shaded.

FIG. 3.—Nonsynonymous substitutions (*d*n) versus synonymous substitutions (*d*s) for pairwise comparisons of RNASE1 genes in Carnivora. Numbers in the table mean *d*s (*d*n) for pairwise comparisons of 9 RNASE1 genes within Mustelidae. The comparisons of *d*n/*d*s > 1 (no test if they are significant >1) were shaded.

least along the lineages leading to one of the yellow-bellied weasel copies, *Wea*-3, and the marten copies, *Mar*-3 (branch f and branch d in fig. 4). As far as we know, besides the adaptive report in leaf-eating colobine monkeys' RNASE1 gene (Zhang et al. 2002; Zhang 2006), our work is the first case of demonstrating that positive selection has acted on the duplicated RNASE1 genes during the evolution.

We proposed that the excessive nonsynonymous substitutions observed in the Mustelidae species might be associated with novel functional adaptations of RNASE1 genes. Presently, the experimental data concerning the expression of these paralogous RNASE1 genes in the Mustelidae is unavailable, so it is difficult to tell what selective agent has promoted RNASE1's complex evolution in this animal group. The results from calculated isoelectric point (pI) shown in figure 1 indicate that more positive charges may be required for the expansion of RNASE1 gene family in the Mustelidae. In addition, our study provided valuable information on the potentially important adaptive amino acid replacements. A total of 8 such residues were predicated here (4, 19, 33, 36, 38, 67, 111, and 112). Among them, residue 38 was selected in both site-specific and branch-site likelihood models with high posterior probabilities (>95%) in the analysis. Interestingly, residues 4 and 38 have also been considered important in RNASE1 gene evolution from previous studies, given that the residue substitutions that occurred at these 2 positions can both change the enzyme activities of degrading dsRNA (Sorrentino 1998; Zhang et al. 2002; Zhang 2006). Moreover, 7 of these 8 positively selected residues involved either the amino acid charge change or the polarity change (fig. 1). Thus, characterization of these substitutions in functional assays as those in residue 4 and 38 is an important issue especially considering that the changing of the enzyme activity on dsRNA cleavage in RNASE1 may imply potential functional consequences. For example, the duplicated RNASE1

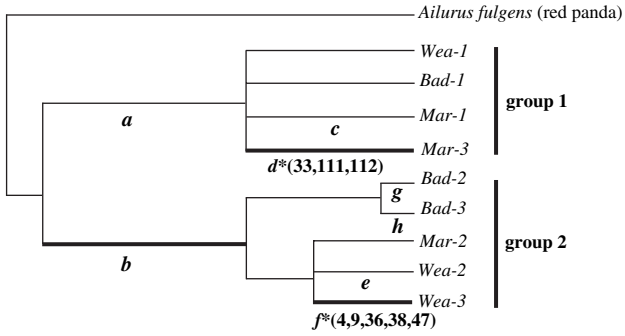


FIG. 4.—Phylogeny of functional RNASE1 genes in Mustelidae used for ML analysis in PAML. The abbreviations for sequence names are the same as in figure 1. Analysis indicated that there is significant evidence along branches b (ancestral branch to group 2), d (terminal branch to *Mar*-3), and f (terminal branch to *Wea*-3) (thick branches in the tree). After correction of multiple testing, LRT test are significant only in branch d (*) and f (*). Eight positively selected residues predicated in total along both branches were also indicated.

genes in douc langur (*RNASE1B*) and guereza (*RNASE1β* and *RNASE1γ*) demonstrate an enhanced role in digestion although at the same time has lost the dsRNA cleavage activity (Zhang et al. 2002; Zhang 2003, 2006).

Interestingly, when we mapped the 8 positive selected residues on the 3-dimensional crystal structure of the molecule, all are found to locate on the surface of the RNASE1 enzyme and extend away from the active site (data not shown). In the review of Golding and Dean (1998), this gene has been taken as an excellent example of adaptive replacements not being solely confined to the active sites. Our analysis of the order Carnivora provides additional evidence in support of their argument. During RNASE1 gene evolution, those replacements that contribute functional adjustment may affect enzyme activities or substrate binding in an indirect (Iijima et al. 1977) or complementary way, as in the case of decarboxylating dehydrogenase, where 2 substitutions outside the active site greatly improve coenzyme specificities by producing subtle conformational changes (Hurley et al. 1996; Golding and Dean 1998). Recently, Zhang and Rosenberg (2002) also found that the enhancement of ribonucleolytic activity of primate eosinophil-derived neurotoxin (*RNASE2*) can be ascribed to the substitutions at 2 complementary sites.

Though the functional changes brought by these 8 positively selected residues cannot be predicted now, they may have influenced the RNASE1 enzyme activity of degrading dsRNA, raising the possibility that new tissue specificity or some more crucial physiologic functions of RNASE1 genes have been developed in Mustelidae family. The carnivoran species sampled in this paper have different diet habits (fig. 2). Yet, no correlation can be apparently found between the diversity in diet and the evolution of RNASE1 genes in Carnivora, in contrast to previous investigations of the same gene in leaf-eating monkeys and ruminants, both being foregut-fermenting mammals. In those studies, RNASE1 duplications and positive selection is justifiable as adaptive responses to digestive physiology. So, the unusual RNASE1 gene evolution observed in Mustelidae family is intriguing. One speculation is that the presence of multiplicity of RNASE1 gene product is an adaptation to improve the efficiency of recovering nutrients from the taken food for the mustelids because mustelids have long been characterized by the highest rates of food passage, that is, the shortest mean retention time of food in the digestive tract, compared with other carnivores (Zharova and

Table 2
Evidence of Adaptive Evolution from Site-Specific and Branch-Site Model Analyses for Mustelidae RNASE1 Genes

Models	ln L[a]	Parameter Estimates[b]	$2\Delta L$[c]	Positively Selected Sites[d]
Site-specific models				
M1a	−1314.504025	$\omega_0 = 0.10899$, $\omega_1 = 1$; $P_0 = 0.59821$, $P_1 = 0.40179$		Not allowed
M2a	−1306.716571	$\omega_0 = 0.10391$, $\omega_1 = 1$, $\chi_2 =$ **11.06159**; $P_0 = 0.56065$, $P_1 = 0.42909$, $P_2 = 0.01026$	(M1a vs. M2a) 15.574908***	***38***
M8a	−1289.722829	$P = 0.92830$, $q = 5.29856$, $P_0 = 0.63971$, $P_1 = 0.36029$; $\omega = 1$		Not allowed
M8	−1306.803379	$P = 0.27796$, $q = 0.32889$, $P_0 = 0.98946$, $P_1 = 0.01054$; $\chi =$ **10.45917**	(M8a vs. M8) 15.182126***	***38***
Branch-site models[e]				
Branch **d**				
Model A	−1308.071604	$\omega_0 = 0.09380$, $\omega_1 = 1$, $\chi_2 =$ **26.22392**; $P_0 = 0.57969$, $P_1 = 0.37805$, $P_2 = 0.04226$		33, ***111***, ***112***
M1a	−1312.611693	$\omega_0 = 0.09108$, $\omega_1 = 1$, $\omega_2 = 1$; $P_0 = 0.48499$, $P_1 = 0.30484$, $P_2 = 0.21017$	(Model A vs. M1a) 9.080178[e]	
Branch **f**				
Model A	−1309.445113	$\omega_0 = 0.12226$, $\omega_1 = 1$, $\chi_2 =$ **28.02641**; $P_0 = 0.57878$, $P_1 = 0.34837$, $P_2 = 0.07285$		4, ***19***, ***36***, ***38***, 67
M1a	−1314.504024	$\omega_0 = 0.10899$, $\omega_1 = 1$, $\omega_2 = 1$; $P_0 = 0.59821$, $P_1 = 0.40179$, $P_2 = 0$	(Model A vs. M1a) 10.117824[e]	

[a] ln L is the log-likelihood scores.
[b] ω estimates in boldface type indicate the presence of positive selection.
[c] LRT to detect adaptive evolution ***$P < 0.001$.
[d] Those codons with posterior probabilities >95% are shown in boldface and italic.
[e] Branches d and f are the lineages leading to *Mar*-3 and *Wea-3*, respectively, in figure 4. After Bonferroni correction for multiple testing, only branches d and f are still significant.

Naumova 2002). An alternative hypothesis is that RNASE1 plays an important role in a function other than digestive activities. A number of previous studies have suggested that RNASE1 expressed in a variety of tissues and blood fluids besides pancreas and served other yet unknown physiological functions, albeit with common nondigestive activity in dsRNA cleavage (Beintema et al. 1988; Futami et al. 1997; Kleineidam et al. 1999; Zhang 2003). This hypothesis can also be indirectly argued for by the report of Dubois et al. (2002), in which, similar to ours, they also surprisingly found that RNASE1 expansion happened in an animal group that does not have foregut fermentation, the genus *Rattus*, whereas one of its gene copies was expressed in spleen tissue and not in pancreas of *Rattus norvegicus*.

Anyway, more information on the structure, function, and evolution of RNASE1 in Carnivora is required for resolving the puzzle why multiplicity of RNASE1 gene should occur in Mustelidae family and not other carnivores. Our study established a necessary foundation for the experimental investigations. It will be interesting to test the expression pattern of these RNASE1 genes and the functional effects of 8 positively selected amino acid substitutions. In addition, other regions of the RNASE1 gene and more sequence data from closely related Mustelidae species are desirable to be included in later studies to unequivocally localize the evolutionary history of multiple RNASE genes in Mustelidae family.

Acknowledgments

We thank Prof Zhang Jianzhi, Dr Shi Peng, and Dr Soochin Cho, Department of Ecology and Evolutionary Biology, University of Michigan, for helpful comments on improving the manuscript. This work was supported by a start-up fund from the University of Yunnan and by grants from National Natural Science Foundation of China (30430110) to Y.P.Z.

Literature Cited

Barnard EA. 1969. Biological function of pancreatic ribonuclease. Nature. 221:340–344.

Beintema JJ. 1990. The primary structure of langur (*Presbytis entellus*) pancreatic ribonuclease: adaptive features in digestive enzymes in mammals. Mol Biol Evol. 7:470–477.

Beintema JJ, Breukelman HJ, Carsana A, Furia A. 1997. Evolution of vertebrate ribonucleases: ribonuclease A superfamily. In: D'Alessio G, Riordan JF, editor. Ribonucleases, structures and functions. San Diego (CA): Academic Press. p. 245–269.

Beintema JJ, Fitch WM, Carsana A. 1986. Molecular evolution of pancreatic- type ribonucleases. Mol Biol Evol. 3:262–275.

Beintema JJ, Gaastra W, Lenstra JA, Welling GW, Fitch WM. 1977. The molecular evolution of pancreatic ribonuclease. J Mol Evol. 10:49–71.

Beintema JJ, Kleineidam RG. 1998. The ribonuclease A superfamily: general discussion. Cell Mol Life Sci. 54:825–832.

Beintema JJ, Lenstra JA. 1982. Evolution of mammalian pancreatic ribonucleases. In: Goodman M, editor. Macromolecular sequences in systematic and evolutionary biology. New York: Plenum. p. 43–73.

Beintema JJ, Neuteboom B. 1983. Origin of the duplicated ribonuclease gene in guinea-pig: comparison of the amino acid sequences with those of two close relatives: capybara and cuis ribonuclease. J Mol Evol. 19:145–152.

Beintema JJ, Scheffer AJ, Van Dijk H, Welling GW, Zwiers H. 1973. Pancreatic ribonuclease: distribution and comparisons in mammals. Nature New Biol. 241:76–78.

Beintema JJ, Schuller C, Irie M, Carsana A. 1988. Molecular evolution of the ribonuclease superfamily. Prog Biophys Mol Biol. 51:165–192.

Bonferroni CE. 1936. Teoria statistica delle classi e calcolo delle probabilità. Pubblicazioni del R Istituto Superiore di Scienze Economiche e Commerciali di Firenze 8:3–62.

Breukelman HJ, Jekel PA, Dubois JY, Mulder PP, Warmels HW, Beintema JJ. 2001. Secretory ribonucleases in the primitive ruminant chevrotain (*Tragulus javanicus*). Eur J Biochem. 268:3890–3897.

Breukelman HJ, van der Munnik N, Kleineidam RG, Furia A, Beintema JJ. 1998. Secretory ribonuclease genes and pseudogenes in true ruminants. Gene. 212:259–268.

Cho S, Beintema JJ, Zhang J. 2005. The ribonuclease A superfamily of mammals and birds: identifying new members and tracing evolutionary histories. Genomics. 85:208–220.

Clegg MT, Cummings MP, Durbin ML. 1997. The evolution of plant nuclear genes. Proc Natl Acad Sci USA. 94:7791–7798.

D'Alession G, Riordan JF. 1996. RNases, structures and functions. New York: Academic Press.

Dubois JY, Catzeflis FM, Beintema JJ. 1999. The phylogenetic position of "Acomyinae" (Rodentia, Mammalia) as sistergroup of a murinae+gerbillinae clade: evidence from the nuclear ribonuclease gene. Mol Phylogenetic Evol. 13: 181–192.

Dubois JY, Jekel PA, Mulder PP, Bussink AP, Catzeflis FM, Carsana A, Beintema JJ. 2002. Pancreatic-type ribonuclease 1 gene duplications in rat species. J Mol Evol. 55:522–533.

Dubois JY, Ursing BM, Kolkman JA, Beintema JJ. 2003. Molecular evolution of mammalian ribonucleases 1. Mol Phylogenet Evol. 27:453–463.

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

Flynn JJ, Finarelli JA, Zehr S, Hsu J, Nedbal MA. 2005. Molecular phylogeny of the Carnivora (Mammalia): assessing the impact of increased sampling on resolving enigmatic relationships. Syst Biol. 54:317–337.

Flynn JJ, Nedbal MA. 1998. Phylogeny of the Carnivora (Mammalia): congruence vs incompatibility among multiple data sets. Mol Phylogenet Evol. 9:414–426.

Force A, Lynch M, Pickett F, Amores A, Yan YL, Postlethwait J. 1999. Preservation of duplicate genes by complementary, degenerative mutations. Genetics. 151:1531–1545.

Futami J, Tsushima Y, Murato Y, Tada H, Sasaki J, Seno M, Yamada H. 1997. Tissue-specific expression of pancreatic-type RNases and RNase inhibitor. DNA and Cell Biol. 16:413–419.

Gillespie JH. 1991 The causes of molecular evolution. Oxford: Oxford University Press.

Golding GB, Dean AM. 1998. The structural basis of molecular adaptation. Mol Biol Evol. 15:355–369.

Hurley JH, Chen R, Dean AM. 1996. Determinants of cofactor specificity in isocitrate dehydrogenase: structure of an engineered $NADP^+ \rightarrow NAD^+$ specificity- reversal mutant. Biochemistry. 35:5670–5678.

Iijima H, Patrzyc H, Bello J. 1977. Modification of amino acids and bovine pancreatic ribonuclease A by kethoxal. Biochim Biophys Acta. 491:305–316.

Kleineidam RG, Pesole G, Breukelman HJ, Beintema JJ, Kastelein RA. 1999. Inclusion of cetaceans within the order Artiodactyla based on phylogenetic analysis of pancreatic ribonuclease genes. J Mol Evol. 48:360–368.

Kumar S, Tamura K, Nei M. 2004. MEGA3: integrated software for molecular evolutionary genetics analysis and sequence alignment. Brief Bioinformatics. 5:150–163.

Li WH, Bousquet J. 1992. Relative-rate test for nucleotide substitutions between two lineages. Mol Biol Evol. 9:1185–1189.

Nei M, Rogozin LB, Piontkivska H. 2000. Purifying selection and birth-and-death evolution in the ubiquitin gene family. Proc Natl Acad Sci USA. 97:10866–10871.

Nowak RM. 1991. Walker's mammals of the world. Baltimore (MD): Johns Hopkins University Press.

Ohno S. 1970. Evolution by gene duplication. Heidelberg (Germany): Springer-Verlag.

Ota T, Nei M. 1994. Divergent evolution and evolution by the birth-and-death process in the immunoglobulin V_H gene family. Mol Biol Evol. 11:469–482.

Piontkivska H, Nei M. 2003. Birth-and-death evolution in primate MHC class I genes: divergence time estimates. Mol Biol Evol. 20:601–609.

Posada D, Crandall KA. 1998. Modeltest: testing the model of DNA substitution. Bioinformatics. 14:817–818.

Robinson-Rechavi M, Huchon D. 2000. RRTree: relative-rate tests between groups of sequences on a phylogenetic tree. Bioinformatics. 16:296–297.

Sambrook E, Fritsch F, Maniatis T. 1989. Molecular cloning. Cold Spring Harbor (NY): Cold Spring Harbor Press.

Schaller GB, Teng QT, Johnson KG, Wang XM, Shen HM, Hu JC. 1989. The feeding ecology of giant pandas and Asiatic black bears in the Tanjiahe Reserve, China. In: Gittleman JL, editor. Carnivore behavior, ecology, and evolution. New York: Cornell University Press. p. 212–241.

Singhania NA, Dyer KD, Zhang J, Deming MS, Bonville CA, Domachowske JB, Rosenberg HF. 1999. Rapid evolution of the ribonuclease A superfamily: adaptive expansion of independent gene clusters in rats and mice. J Mol Evol. 49: 721–728.

Sorrentino S. 1998. Human extracellular ribonuclease: multiplicity, molecular diversity and catalytic properties of the major RNase types. Cell Mol Life Sci. 54:785–794.

Swofford DL. 2001. PAUP*: phylogenetic analysis using parsimony (* and other methods). Version 4.0b8. Sunderland (MA): Sinauer Associates.

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

Thompson JD, Gibson TJ, Plewniak F, Jeanmougin F, Higgins DG. 1997. The ClustalX windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. Nucleic Acids Res. 24:4876–4882.

Yang Z. 2000. Phylogenetic analysis by maximum likelihood (PAML). Version 3.0. London: University College.

Yang Z, Wong WS, Nielsen R. 2005. Bayes empirical Bayes inference of amino acid sites under positive selection. Mol Biol Evol. 22:1107–1118.

Yu L, Li QW, Ryder OA, Zhang YP. 2004. Phylogenetic relationships within mammalian order Carnivora indicated by sequences of two nuclear DNA genes. Mol Phylogenet Evol. 33:694–705.

Yu L, Zhang YP. 2006. Phylogeny of the caniform carnivora: evidence from multiple genes. Genetica. 127:65–79.

Zhang J. 2003. Parallel functional changes in the digestive RNases of ruminants and colobines by divergent amino acid substitutions. Mol Biol Evol. 20:1310–1317.

Zhang J. 2006. Parallel adaptive origins of digestive RNases in Asian and African leaf monkeys. Nat Genet. 38:819–823.

Zhang J, Dyer KD, Rosenberg HF. 2000. Evolution of the rodent eosinophil-associated RNase gene family by rapid gene sorting and positive selection. Proc Natl Acad Sci USA. 97:4701–5706.

Zhang J, Nielsen R, Yang Z. 2005. Evaluation of an improved branch-site likelihood method for detecting positive selection at the molecular level. Mol Biol Evol. 22:2472–2479.

Zhang J, Rosenberg HF. 2002. Complementary advantageous substitutions in the evolution of an antiviral RNase of higher primates. Proc Natl Acad Sci USA. 99:5486–5491.

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

Zhang J, Zhang YP, Rosenberg HF. 2002. Adaptive evolution of a duplicated pancreatic ribonuclease gene in a leaf-eating monkey. Nat Genet. 30:411–415.

Zharova GK, Naumova EI. 2002. Structural and functional adaptations of the sable digestive tract to plant food. Dokl Biol Sci. 382:31–33.

David Irwin, Associate Editor

Accepted August 24, 2006