Published in final edited form as:

Mamm Genome. 2009 ; 20(11-12): 749–757. doi:10.1007/s00335-009-9215-4.

RNase 1 genes from the Family Sciuridae define a novel rodent ribonuclease cluster

Steven J. Siegel^{a,c}, Caroline M. Percopo^{a,c}, Kimberly D. Dyer^a, Wei Zhao^{a,d}, V. Louise Roth^b, John M. Mercer^b, and Helene F. Rosenberg^{a,e}

^aLaboratory of Allergic Diseases, National Institute of Allergy and Infectious Diseases, National Institutes of Health, Bethesda, Maryland ^bBiology Department, Duke University, Durham, North Carolina

Abstract

The RNase A ribonucleases are complex group of functionally diverse secretory proteins with conserved enzymatic activity. We have identified novel RNase 1 genes from four species of squirrel (order Rodentia, family Sciuridae). Squirrel RNase 1 genes encode typical RNase A ribonucleases, each with eight cysteines, a conserved CKXXNTF signature motif, and a canonical His¹²-Lys⁴¹-His¹¹⁹ catalytic triad. Two alleles encode *Callosciurus prevostii* RNase 1, which include a Ser¹⁸↔Pro, analogous to the sequence polymorphisms found among the RNase 1 duplications in the genome of *Rattus exulans*. Interestingly, although the squirrel RNase 1 genes are closely related to one another (77 to 95% amino acid sequence identity), the cluster as a whole is distinct and divergent from the clusters including RNase 1 genes from other rodent species. We examined the specific sites at which Sciuridae RNase 1s diverge from Muridae / Cricetidae RNase 1s, and determined that the divergent sites are located on the external surface, with complete sparing of the catalytic crevice. The full significance of these findings awaits a more complete understanding of biological role of mammalian RNase 1s.

Keywords

ribonucleases; rodent; diversity

Introduction

RNase 1 is the most thoroughly characterized lineage of the RNase A ribonucleases, which is an extensive family of secretory enzymes characterized in vertebrate species (Beintema and Kleineidam, 1994; Dubois *et al.*, 1999; Dyer and Rosenberg 2006; Rosenberg 2008; Marshall *et al.*, 2008; Cho *et al.*, 2005; Cho and Zhang, 2006; Yu and Zhang 2006). Bovine RNase A, the RNase 1 ortholog initially isolated from pancreatic tissue, has served as a prototype “master protein” and was the source of many of the earliest discoveries on primary structure, crystallographic ultrastructure, protein folding and catalytic mechanism (Hirs *et al.*, 1956; Anfinsen, 1957; Aqvist and Anfinsen, 1959; Scheraga and Rupley, 1962). However, despite

^ePlease address correspondence to: Helene F. Rosenberg, M.D., Section Chief, LAD, NIAID, NIH, Building 10, Room 11C215, 9000 Rockville Pike, Bethesda, MD 20892, tel: 301 420 1545; hrosenberg@niaid.nih.gov.

^cS. J. Siegel and C. M. Percopo contributed equally to this work.

^dCurrent address: Biochemical Engineering College, Beijing Union University, 18th Fatou Xili, Chaoyang District, Beijing PR China

Sequence data from this article have been deposited with the DDBJ/EMBL/GenBank Data Libraries under Accession Nos. FJ659109 – FJ659113

the complex biochemical characterization, our understanding of the biology of RNase 1 remains rudimentary. Given its abundance in bovine pancreas, localization as a secretory protein, and overall catalytic efficiency, RNase A was presumed to function as a digestive enzyme, serving to promote catabolism of ingested ribonucleic acid polymers. However, more recent studies have shown that RNase 1 expression is substantially less prominent in pancreatic tissue of non-ruminant mammals (Beintema *et al.*, 1973), and is not limited to the gastrointestinal tract (Futami *et al.* 1997). Similarly, the RNase 1 lineage was considered to be relatively stable from an evolutionary perspective, at least compared to extensive divergence and duplication reported for the eosinophil ribonucleases (RNase 2 / RNase 3; Rosenberg *et al.*, 1995; Zhang *et al.*, 2000). However, recently Dubois and colleagues (2002) identified multiple RNase 1 duplications in the genomes of rat species (family Muridae), suggesting that RNase 1 genes, at least among rodents, may be responding to more complex evolutionary constraints than may have been previously appreciated.

As part of a larger exploration of the RNase 1 lineage in rodent species, we have isolated and characterized full coding sequences from genomic DNAs from four distinct species of squirrels (order Rodentia, family Sciuridae). Squirrels are widely distributed, morphologically primitive rodents that have diversified in concert with major changes in geography and climate (Mercer and Roth, 2003); an analysis of nuclear genes *c-myc* and *RAG-1* has confirmed five major squirrel lineages (Steppan *et al.*, 2004). In this work, we explore the sequences of the Sciuridae RNase 1s, note the similarities and differences from their counterparts among species from rodent families, and evaluate the divergent sites within the context of the tertiary protein structure.

Materials and Methods

Source material

The *Callosciurus notatus* fibroblast cell line NZP-46 was obtained from American Type Culture Collection (CRL-1926, ATCC, Manassas, VA). Genomic DNA was isolated from cells by lysis in 10 mM Tris-HCl, pH 8.0, with 100 mM NaCl, 25 mM EDTA, 0.5% SDS and 0.1 mg/mL proteinase K at 56°C overnight followed by phenol: chloroform extraction. All other genomic DNAs were prepared from animal tissue as described (Mercer and Roth, 2003).

PCR amplification of an initial RNase 1 fragment from *C. notatus* genomic DNA

Primers were designed from the 5' and 3' ends of an RNase 1 fragment originally reported for *Oryctolagus cuniculus* (Dubois *et al.*, 2003; [Table 1]). The 338 bp product resulting from amplification using these primers and *C. notatus* genomic DNA as a template (standard conditions, 35 cycles) was cloned into PCR2.1, sequenced (ABI 3730xl Genetic Analyzer) and the information used to design internal unidirectional primers in order to isolate the complete *C. notatus* RNase 1 gene and to identify additional squirrel RNase 1s, as described in the following sections. Primers used to clone other *C. notatus* genes in order to confirm the identity of this cell line are also listed in Table 1.

Unidirectional PCR / universal genome walk

Full length squirrel RNase 1 sequences were generated using the Genome Walker Universal kit (BD Biosciences / Clontech, Mountainview, CA). Two nested 5'-directed and two nested 3'-directed primers were designed based on sequence from the original *C. notatus* RNase 1 fragment [Figure 1]. Manufacturer's protocols were followed, save for the use of 0.25 µg of genomic DNA to generate each linker-based library rather than the recommended 2.5 µg. Sequences were amplified using Advantage 2 PCR enzyme (Clontech, Mountain View, CA), cycling parameters including a hold 95°C for 2 minutes, followed by 25 cycles of 95°C for 20 seconds, 50°C for 20 seconds and 72°C for 30 seconds, followed by a 5 minute hold at 72°C.

Amplification products to be evaluated were cloned into the PCR2.1 sequencing vector. Full length sequence information for each RNase 1 was initially determined from overlapping clones from the genome walk libraries, and then full length sequences were re-isolated *de novo* from fresh samples of unmanipulated genomic DNA, utilizing primers designed from sequence situated outside the designated open reading frames. All final clones were isolated from at least two independent amplification reactions and sequenced completely in two directions.

Genomic Southern blot

Genomic Southern blot was performed by Lofstrand Labs Limited (Gaithersburg, MD) with genomic DNA and unlabelled cDNA probes supplied by our group. Genomic DNA was isolated as described above and subjected to restriction digestion and gel electrophoresis (1 µg / lane), acid / base denaturation, neutralization, and transfer to nylon membrane by standard methods. DNA was UV linked to the membrane, and the membrane subjected to prehybridization (68°C, 3 hrs in 6X SSC, 5X Denhardt's solution, 0.5% SDS) followed by hybridization for 25 hrs at 68°C with random primed probes. Filters were washed at low stringency in 2X SSC with 0.01% SDS at 68°C with three buffer changes over 60 minutes and exposed to film for 2.5 days with an intensifier screen at -80°C.

Sequences, structures and evolutionary analysis

Primary sequence analysis was determined with the assistance of Sequencher analysis software (GeneCodes, Ann Arbor, MI). Amino acid sequence similarities were determined via algorithms within the Wisconsin Genetics Computer Group software package available online at the National Institutes of Health. Isoelectric points of encoded proteins were determined by EMBOSS at <http://isoelectric.ovh.org/>. Amino acid and nucleotide sequence alignments were generated via ClustalW (<http://clustalw.genome.jp/>). Phylogenetic analysis was conducted using MEGA version 4.0 (Tamura *et al.*, 2007). The ribbon diagram of human RNase 1 (pdb 2k11) was from the RCSB protein data bank (<http://www.rcsb.org/pdb/home/home.do>), and was manipulated via the Java-based MBT protein workshop algorithms. Nucleotide and amino acid sequences of all RNase 1s other than those identified in this study were obtained via direct accession number searches and via BLAST searches at the NCBI / National Library of Medicine portal at <http://www.ncbi.nlm.nih.gov/sites/entrez>.

Results

Isolation of complete RNase 1 coding sequences

The strategy used to isolate complete open reading frames of RNase 1 genes from squirrel genomic DNAs is shown in Figure 1. The NZP-46 *C. notatus* fibroblast cell line was used as an initial source of genomic DNA. We confirmed this cell line as derived from *C. notatus* by cloning and sequencing a ~400 bp fragment of the 16S ribosomal RNA gene reported for this species (GenBank acc no. AY227453). As shown in Figure 1A, we generated a primary 338 bp fragment encoding *C. notatus* RNase 1 from genomic DNA isolated from the NZP-46 cell line using 24 bp 5' and 3' primers derived from sequence originally reported to encode RNase 1 from the European rabbit (*O. cuniculus*, see Table 1 for these and all primer sequences). The published rabbit sequence (listed as Genbank AJ535681 in Dubois *et al.*, 2003) was remarkably similar to squirrel sequences isolated in this report; this sequence file been changed to AJ535681.1, now denoting RNase 1 from *Sciurus vulgaris*. We utilized sequence from this primary *C. notatus* fragment to design nested primer pairs (3A, 3B, 5A, 5B) for two rounds of unidirectional PCR amplification reactions with restriction fragmented, linked libraries created from multiple sources of squirrel genomic DNA as templates (Figure 1B, and as described in Methods). Amplification products were sequenced, overlapping contigs were aligned, and the complete open reading frames were ultimately re-isolated from unmanipulated

genomic DNA using primers derived from genomic sequence located 5' to the start ATG and 3' to the stop codon [Figure 1C]. Each gene sequence reported was isolated from multiple independent amplification trials.

Properties of amino acid sequences encoded by squirrel RNase 1 genes

Single intronless open reading frames were identified in three of the four sources of genomic DNA except for *C. prevostii*, in which two highly homologous RNase 1 sequences were identified. The squirrel sequences encode proteins that are typical of the RNase 1 lineage [Figure 2]. All encoded RNase 1s have a 28 amino acid hydrophobic signal sequence preceding an amino terminus that was inferred by homology to human and mouse RNase 1s (Mizuta *et al.*, 1990; Lenstra and Beintema, 1979). All polypeptides maintain the His-Lys-His catalytic triad, eight appropriately spaced cysteines, and the RNase A family signature CKXXNTF motif. The polypeptides are 128 amino acids in length, and have isoelectric points ranging from 8.40 to 8.58.

RNase 1s of *C. prevostii*

As noted above, two RNase 1 sequences were identified in the DNA sample from *C. prevostii*; each unique sequence was identified in multiple amplification trials. We performed Southern blot analysis, which revealed single bands in each of three restriction-digested *C. notatus* DNA samples probed with radiolabelled *C. notatus* RNase 1, and likewise a single band in the lane containing restriction-digested *C. prevostii* DNA probed with radiolabelled *C. prevostii* RNase 1 [Figure 3]. These findings suggest that the two *C. prevostii* RNase 1 sequences are alleles rather than gene duplications. Of the four divergent sites within this allelic pair, two are in the signal sequence, and two are within the mature protein, at positions Ser¹⁸↔Pro and Arg⁵⁵↔Gln [Figure 2]. Interestingly, Ser¹⁸↔Pro was identified in the gene duplication of the Polynesian rat (*Rattus exulans*), both the comparison of RNase 1α ↔ RNase 1β, and also RNase 1α ↔ RNase 1δ (Dubois *et al.*, 2002).

Phylogenetic relationships among mammalian RNase 1s

An unrooted neighbor-joining tree displaying relationships among mammalian RNase 1 DNA sequences is shown in Figures 4A. The Sciuridae RNase 1s form a cluster that is distinct from those of other mammalian species, and that is divergent from the cluster that includes RNase 1s from Muridae and Cricetidae species. Sciridae RNase 1s in relation to all rodent sequences, including those of the rodent infraorder Hystricognathi (guinea pig, capybara, cuis), for which DNA sequence is unavailable, are shown in the tree in Figure 4B. Average distances within and between clusters were determined [Table 2]. Clearly, the rodent Sciuridae RNase 1s are no more similar to rodent Muridae / Cricetidae RNase 1s than they are to any of the other phylogenetically more distant mammalian species. Interestingly, the same is true when this type of analysis is applied to the data shown in Figure 4B; the Sciuridae RNase 1 sequences are no more similar to those of the Hystricognathi RNase 1s than they are to the Muridae RNase 1s (data not shown).

We identified 18 sites that demarcate specific divergence between the squirrel RNases and those in the Muridae cluster [Table 3]. There are no linear regions or domains with a specific propensity for divergence. However, when the divergent sites are mapped onto the three dimensional ribbon diagram of the homologous human RNase 1 (pdb #2k11; Figure 5), they localize universally on the exterior surface of the protein. The segments that interact with the incoming ribonucleotide, those demarcated by and encompassing the His¹² - Lys⁴¹ - His¹¹⁹ catalytic triad, are remarkably conserved.

Discussion

In this work, we isolated and characterized four novel RNase 1 genes from squirrel genomic DNAs via a two-tiered unidirectional PCR approach. We used a variant of this method in a previous study in which we described the isolation of a unique RNase A ribonuclease from the amphibian species *Iguana iguana* (Nitto *et al.*, 2005). This approach is ideally suited for isolation of novel RNase A ribonucleases, which have relatively small, intronless open reading frames with characterized sequence motifs and specific regions of predictable cross-species homologies. Here, we chose to re-isolate the complete open-reading frames *de novo* from unmanipulated genomic DNA once the complete sequences were assembled from genomic fragments. In this way, we were able to assess the sequence of each gene directly, and we were able to eliminate the possibility that our isolated sequences contained unrelated fragments that became ligated to one another during library assembly.

Among the interesting findings, we observed that rodent RNase 1s from squirrels and from other rodent families emerged as distinct clusters. The constraints leading to divergence among the rodent RNase 1s remain obscure for several reasons. First, despite the historical focus on bovine pancreatic RNase A, the full biological role of RNase 1 in non-ruminant species remains obscure. Futami and colleagues (1997) have demonstrated that RNase 1s are expressed widely in tissues outside the gastrointestinal tract, and Beintema and colleagues (1973) have demonstrated that there is little to no actual enzymatic activity in the pancreas other than in mammals with ruminant digestive systems, and even less in pancreatic tissue of squirrels than in those of other rodents, all findings that lead one to question the role of this protein as purely a digestive enzyme. Yet few other hypotheses have emerged in its place.

Although extensive gene duplication has been described for the rodent RNase 2/3s and the *Rana* ribonucleases (Rosenberg *et al.*, 2001), there were only a few duplications characterized among RNase 1 genes until quite recently; the best known of these are those among ruminants, the bovine brain and bovine seminal RNases (D'Alessio *et al.*, 1991). More recently, Zhang and colleagues (2002, 2006) have described duplications of RNase 1s of African and Asian langurs, and Dubois and colleagues (2002) have identified extensive RNase 1 duplication among several rat species. Duplication of RNase 1 in the Polynesian rat (*Rattus exulans*) includes a nucleotide polymorphism resulting in the conversion of Pro¹⁸↔Ser, similar to that observed between *C. prevostii* RNase 1 alleles (Ser¹⁸↔Pro). Interestingly, Ser¹⁸ is a component of a crucial hydrogen bond in the native structure of bovine RNase A and provides stability for the amino terminal segment of the protein (Richards and Wycoff, 1971); conversion to proline likely eliminates this bond, which could release the amino terminal segment from this constrained position, permitting novel interactions to occur.

Despite our inability to come to any universal conclusions regarding function, information obtained from cross-species evaluations provides us with some clues regarding functionally crucial vs. flexible sites. We evaluated individual sites and found that, despite the divergence calculated for the rodent Sciuridae vs. rodent Muridae / Cricetidae sequences, the essential nature of the RNase 1 proteins remained unchanged. In other words, none of the RNase A family signature features were lost, there was no prominent change in net charge, nor were there any systematic alterations in hydrophilicity / hydrophobicity resulting from the changes at the 18 individual divergent sites. However, when these sites were examined in three-dimensional perspective, it was clear that mutational change was all focused at the external surface of the protein. Not only were the catalytic amino acids left intact (His¹², Lys⁴¹, His¹¹⁹), but the entire catalytic crevice remained free of any mutational change (Figure 5). Furthermore, only one of the 18 total sites (Q69) was among those recognized as a contact site for interaction with intracellular ribonuclease inhibitor [RI; (Kobe and Deisenhofer, 1995)].

This finding suggests that maintaining strong interactions with RI is likewise an important evolutionary constraint.

In summary, we have identified four novel RNase 1 genes from the rodent family Sciuridae. The sequences encode proteins with properties typical of the RNase A ribonuclease family, including the signal sequence, His-Lys-His catalytic triad and eight canonically spaced cysteines. The most interesting feature of the squirrel sequences is their nature of their divergence from the rodent Muridae/Cricetidae RNase 1 counterparts. Three-dimensional projection indicates that the divergent sites are located on the external surface of RNase 1, while the catalytic crevice remains unchanged. The significance of this finding awaits a clearer understanding of the full spectrum of activities of the mammalian RNase 1 lineage.

Acknowledgments

The work presented in this manuscript was supported by NIAID DIR funding to HFR and National Science Foundation #DEB 9726855 to JMM and VLR.

References

- Anfinsen CB Jr. Structural basis of ribonuclease activity. *Fed Proc* 1957;16:783–791. [PubMed: 13480361]
- Aqvist SE, Anfinsen CB. The isolation and characterization of ribonucleases from sheep pancreas. *J Biol Chem* 1959;234:1112–1117. [PubMed: 13654330]
- Beintema JJ, Scheffer AJ, van Dijk H, Welling GW, Zwiers H. Pancreatic ribonuclease distribution and comparisons in mammals. *Nat New Biol* 1973;241:76–78. [PubMed: 4512404]
- Beintema JJ, Kleineidam RG. The ribonuclease A superfamily: general discussion. *Cell Mol Life Sci* 1994;54:825–832. [PubMed: 9760991]
- Cho S, Beintema JJ, Zhang J. The ribonuclease A superfamily of mammals and birds: identifying new members and tracing evolutionary histories. *Genomics* 2005;85:208–220. [PubMed: 15676279]
- Cho S, Zhang J. Ancient expansion of the ribonuclease A superfamily revealed by genomic analysis of placental and marsupial mammals. *Gene* 2006;373:116–125. [PubMed: 16530354]
- D'Alessio G, Di Donato A, Parente A, Piccoli R. Seminal RNase: a unique member of the ribonuclease superfamily. *Trends Biochem Sci* 1991;16:104–106. [PubMed: 2057997]
- Dubois JY, Catzeflis FM, Beintema JJ. The phylogenetic position of "Acomyinae" (Rodentia, Mammalia) as sister group of a Murinae + Gerbillinae clade: evidence from the nuclear ribonuclease gene. *Mol Phylogenet Evol* 1999;13:181–192. [PubMed: 10508551]
- Dubois JY, Jekel PA, Mulder PP, Bussink AP, Catzeflis FM, Carsana A, Beintema JJ. Pancreatic-type ribonuclease 1 gene duplications in rat species. *J Mol Evol* 2002;55:522–533. [PubMed: 12399926]
- Dubois JY, Ursing BM, Kolkman JA, Beintema JJ. Molecular evolution of mammalian ribonucleases I. *Molec Phylog & Evol* 2003;27:453–463.
- Dyer KD, Rosenberg HF. The RNase A superfamily: generation of diversity and innate host defense. *Mol Diversity* 2006;10:585–597.
- Futami J, Tsushima Y, Murato Y, Tada H, Sasaki J, Seno M, Yamada H. Tissue-specific expression of pancreatic-type RNases and RNase inhibitor in humans. *DNA Cell Biol* 1997;16:413–419. [PubMed: 9150428]
- Hirs CH, Moore S, Stein WH. Studies on structure of ribonuclease. *Fed Proc* 1956;15:840–848. [PubMed: 13356878]
- Kobe B, Deisenhofer J. A structural basis of the interactions between leucine-rich repeats and protein ligands. *Nature* 1995;374:183–186. [PubMed: 7877692]
- Lenstra JA, Beintema JJ. The amino acid sequence of mouse pancreatic ribonuclease. Extremely rapid evolutionary rates of the myomorph rodent ribonucleases. *Eur J Biochem* 1979;98:399–408. [PubMed: 556267]
- Marshall GR, Feng JA, Kuster DJ. Back to the future: ribonuclease A. *Biopolymers* 2008;90:259–277. [PubMed: 17868092]

- Mercer JM, Roth VL. The effects of Cenozoic global change on squirrel phylogeny. *Science* 2003;299:1568–1572. [PubMed: 12595609]
- Mizuta K, Awazu S, Yasuda T, Kishi K. Purification and characterization of three ribonucleases from human kidney: comparison with urine ribonucleases. *Arch Biochem Biophys* 1990;281:144–151. [PubMed: 2383019]
- Nitto T, Lin C, Dyer KD, Wagner RA, Rosenberg HF. Characterization of a ribonuclease gene and encoded protein from the reptile, Iguana iguana. *Gene* 2005;352:36–44. [PubMed: 15893436]
- Richards RM, Wyckoff HW. Bovine pancreatic ribonuclease. *Enzymes* 1971;4:647–652.
- Rosenberg HF, Dyer KD, Tiffany HL, Gonzalez M. Rapid evolution of a unique family of primate ribonuclease genes. *Nat Genet* 1995;10:219–223. [PubMed: 7663519]
- Rosenberg HF, Zhang J, Liao YD, Dyer KD. Rapid diversification of RNase A superfamily ribonucleases from the bullfrog, *Rana catesbeiana*. *J Mol Evol* 2001;53:31–38. [PubMed: 11683320]
- Rosenberg HF. RNase A ribonucleases and host defense: an evolving story. *J Leukoc Biol* 2008;83:1079–1087. [PubMed: 18211964]
- Scheraga HA, Rupley JA. Structure and function of ribonuclease. *Adv Enzymol Relat Subj Biochem* 1962;24:161–261. [PubMed: 14498258]
- Steppan SJ, Storz BL, Hoffmann RS. Nuclear DNA phylogeny of the squirrels (Mammalia: Rodentia) and the evolution of arboreality from c-myc and RAG1. *Mol. Phylogenetics and Evol* 2004;30:703–719.
- Tamura K, Dudley J, Nei M, Kumar S. MEGA4: Molecular Evolutionary Genetics Analysis (MEGA) software version 4.0. *Mol Biol Evol* 2007;24:1596–1599. [PubMed: 17488738]
- Yu L, Zhang YP. The unusual adaptive expansion of pancreatic ribonuclease genes in carnivore. *Mol Biol Evol* 2006;23:2326–2335. [PubMed: 16950759]
- Zhang J, Dyer DK, Rosenberg HF. Evolution of the rodent eosinophil-associated RNase gene family by rapid gene sorting and positive selection. *Proc Natl Acad Sci U S A* 2000;97:4701–4708. [PubMed: 10758160]
- Zhang J, Zhang YP, Rosenberg HF. Adaptive evolution of a duplicated pancreatic ribonuclease gene in a leaf-eating monkey. *Nat Genet* 2002;30:411–415. [PubMed: 11925567]
- Zhang J. Parallel adaptive origins of digestive RNases in Asian and African leaf monkeys. *Nat Genet* 2006;38:819–823. [PubMed: 16767103]

Figure 1. Strategy used to isolate complete RNase 1 coding sequences from genomic DNAs from the Family Sciuridae

Tricolored squirrel a	MALEKPLALF PLLVLVLLLL EWVSPTVGRE SAAKKFQQRH MDSSDSFSSS PTYCNEMMKR RNMTKGHC	VP	VNTFVHEPLV
Tricolored squirrel b	M.N....	P....	
Plantain squirrel	V.Q....		
Eastern grey squirrel	S.V.V.P.L.K.N....	A.SGI.R.S..	
Sthern flying squirrel	-.S.V.F....V.PS.L.K. TS.H....	NG.SRV.R.TR...Q	S..
Tricolored squirrel a	DVRAVCTQEN VTCKNGQTN	C YQSKSSMHIT DCRLTRGSKY PDCRYQTTVK	
Tricolored squirrel b	..Q.....		
Plantain squirrel	N.Q.....		
Eastern grey squirrel	..Q.....	R.N....	..N.A.K.SQ.
Sthern flying squirrel	..Q.....	..M.....	..T.NL...E...N.A.K..Q.
Tricolored squirrel a	EKHIVACEG NPSVPVFDA SVEGST	pI 8.58	%Identity - Genus/species C. prevostii GenBank # FJ659109
Tricolored squirrel b		8.40	97 C. prevostii FJ659110
Plantain squirrel	.RY.I.....V..	8.58	95 C. notatus FJ659111
Eastern grey squirrel	.R..I.....Y.....E..	8.40	84 S. carolinensis FJ659112
Sthern flying squirrel	.R..I.....Y.....E..	8.40	77 G. volans FJ659113

Figure 2. Amino acid sequences encoded by squirrel RNase 1 genes
All encoded polypeptides include signal sequence (italicized), eight appropriately spaced cysteines (blue), the His (H) -Lys (K)- His (H) catalytic triad (red), and the RNase A family CKXXNTF signature motif (within black box). The amino termini of the secreted form of RNase 1 were inferred via homology with the human and bovine RNase 1 and are indicated by the blue arrow; dots within the sequences indicate identity with the reference sequence (tricolored squirrel RNase 1a). Calculated isoelectric points, % identity with the reference sequence, genus-species designations and Genbank accession numbers are as shown. The *C. prevostii* sequences shown are RNase 1a and RNase 1b alleles.

Figure 3. Genomic Southern blot

Genomic DNA from squirrels *C. notatus* (*C. not*) and *C. prevostii* (*C. pr*) subjected to restriction digestion with Eco RI (E), Hind III (H), or Pst I (P) and probed with radiolabelled *C. notatus* RNase 1 or *C. prevostii* RNase 1a, respectively. Molecular size markers (kB) as shown; arrow indicates single hybridizing band detected in *C. prevostii* genomic DNA.

Figure 4A

Figure 4. Neighbor-joining trees documenting relationships among mammalian RNase 1s
(A) Squirrel (Sciuridae) RNase 1s (highlighted in blue) are shown in relation to RNase 1s of primates, carnivores, ruminants, aquatic mammals, and other rodents. Tree is constructed via algorithms in MEGA 4.0 (Tamura *et al.*, 2007) using coding sequences derived from cDNA or genomic DNA, including signal sequences where available, aligned via ClustalW. Model for tree construction: maximum composite likelihood, complete deletions, substitutions: transitions and transversions. Bootstrap values based on 1000 replications; those > 50 are shown. GenBank accession numbers are included after the common names; genus and species designations are available within Genbank files. **(B)** Squirrel RNase 1 encoded polypeptide sequences (highlighted in blue) are shown in relation to RNase 1 polypeptide sequences (both

those inferred from DNA sequence and those sequenced directly from tissue samples) from various rodent species. Tree is constructed using MEGA 4.0 (Tamura *et al.*, 2007) as noted above, with complete deletions, bootstrap values (1000 replications) as noted in A. Genus and species designations are available in Protein Database (PDB) files, accession numbers listed after common names.

Figure 5. Three-dimensional localization of divergent sites

Localization of divergent sites inferred via homology with human RNase 1 (ribbon plot, PDB ID# 2k11, amino terminus is position 1). Shown are the universally conserved catalytic components (His¹², Lys⁴¹, His¹¹⁹, positions indicated as blue dots), and the 18 sites at which there is near or complete amino acid sequence conservation among Sciuridae RNase 1s and corresponding divergence among the Muridae/Cricetidae sequences (indicated as red dots; see Table 2). The arrow denotes the catalytic crevice.

Table 1

Primers used to generate PCR amplification products

Name	Description	Sequence
16-1	5'→3' from Cn 16S RNA	5'-CGA GGG TTA ATC TGT CTC TTA CTT-3'
16-2	3'→5' from Cn 16S RNA	5'- TAG CGG TTG CAC CAT TTG GAT GTC-3'
R1	5'→3' from rabbit RNase 1	5'- AAT GAG TCC CCA GCC ATG AAG TTC-3'
R2	3'→5' from rabbit RNase 1	5'- CAT AGA TGG GTT CCC CTC ACA AGC-3'
5A	3'→5' primary genewalk	5' - TTA AGT GGA CAC CTC CAC CGA AGC ATC GGA - 3'
5B	3'→5' nested genewalk	5' - CTC ACA GGC CAC GAT GAT CTA CCT CTC TTT - 3'
3A	5'→3' primary genewalk	5' - CAG CAC ATG GAC TCG AGC GAT TCC TT - 3'
3B	5'→3' nested genewalk	5' - AGC AGC CCC ACC TAC TGC AAC GAA AT-3'
Cn-S1A	5'→3' completion primary	5'-GTG CTC TTC CTT TCC CTT TTC-3'
Cn-S1B	5'→3' completion nested	5'-GTT AGG CTC TCT CAG GCA CTG AG-3'
Cn-S2A	3'→5' completion primary	5'-GGT CAA GAA AGG AAG AAG GGA-3'
Cn-S2B	3'→5' completion nested	5'-GTC GGA ACC ATG AGG TGG AG-3'
Cp-S1A	5'→3' completion primary and nested	5'-TTA GGC TCT CTC AGG CAC TGA GGC C-3'
Cp-S2A	3'→5' completion primary	5'-GAA GAA GGG AGA GGC AAG CAT-3'
Cp-S2B	3'→5' completion nested	5'-GAA CCA TGA GGT GGA GAG GA-3'
Gv-S1A	5'→3' completion primary	5' TTA AGC ATG CTC GAG CGG -3'
Gv-S1B	5'→3' completion nested	5' GGC CGC CAG TGT GCT GGA A-3'
Gv-S2A	3'→5' completion primary	5' GGA CAA GGA AGG AAG AAG GG-3'
Gv-S2B	3'→5' completion nested	5' GCA CGC GTG AGA ACC ATG AGG T-3'
Sc-S1A	5'→3' completion primary	5'- TGA GGT CTG GCT CTT TCT CCT TTT CCA G-3'
Sc-S1B	5'→3' completion nested	5'-TTA GGC TCT CTC AAG CAG TGA GGC CAC C-3'
Sc-S2A	3'→5' completion primary	5' TTA GAA GGG AGA GGC AAG AAC CAT GAG GT-3'
Sc-S2B	3'→5' completion nested	5'-AAG GAG AGG GAC ATC CCA CTG CTC TGT-3'

Cn, *C. notatus*; Cp, *C. prevostii*; Gv, *G. volans*; Sc, *S. carolinensis*.

Table 2

Estimates of evolutionary divergence over sequence pairs within and between clusters

Avg. Distances within clusters (sites $\pm$ SE)	
Primate RNase 1s (15 genes)	14.5 $\pm$ 1.9
Rodent RNase 1s (Muridae / Cricetidae; 25 genes)	52.1 $\pm$ 3.5
Rodent RNase 1s (Sciuridae; 4 genes)	27.5 $\pm$ 3.7
Carnivore RNase 1s (11 genes)	39.2 $\pm$ 3.2
Ruminant / Aquatic mammal RNase 1s (9 genes)	40.1 $\pm$ 3.4
Avg. Distances between clusters (sites $\pm$ SE)	
Sciuridae vs. Primate	48.6 $\pm$ 5.1
Sciuridae vs. Muridae / Cricetidae	75.0 $\pm$ 5.4
Sciuridae vs. Carnivores	65.8 $\pm$ 5.4
Sciuridae vs. Ruminant / Aquatic mammals	59.8 $\pm$ 5.0
Primate vs. Muridae / Cricetidae	75.4 $\pm$ 5.4
Primate vs. Carnivores	61.9 $\pm$ 5.4
Primate vs. Ruminant / Aquatic mammals	60.2 $\pm$ 5.1
Muridae / Cricetidae vs. Carnivores	80.6 $\pm$ 5.3
Muridae / Cricetidae vs. Ruminant / Aquatic mammals	74.9 $\pm$ 5.1
Carnivores vs. Ruminant / Aquatic mammals	74.4 $\pm$ 4.4

Shown are the number of divergent sites per sequences ($\pm$ standard error, calculated via 1000 bootstrap replications (Tamura *et al.*, 2007)). Sequences within each cluster are as shown in Figure 4A. Gaps were eliminated via complete deletion. 64 gene sequences, 317 sites per gene sequence were included in the analysis.

Table 3

Sites at which RNase 1s of Sciuridae diverge from RNase 1s of Muridae/Cricetidae.

Position relative to amino terminus	Amino acid(s) shared among <i>Sciuridae</i> at this position	Amino acid(s) shared among <i>Muridae</i> / <i>Cricetidae</i> at this position
28	E	Q
39	H	Y, Q, S
52	V	T, Q, A, E, D
59	T	S
69 [*]	Q	N, R, K
70 [*]	T	S, Q, R, N
74	Q	K
78	N, S	A, T
87 [*]	T	K
88	R	D, G
89	G	N, S
96	A, R	D, N, E
102	K	L, H, Q, S, Y
103	E	Q
125	E	-
126	G	-
127	S	-
128	T	-

* Amino acid sequence of the white-footed mouse (*Peromyscus leucopus*) aligns with Sciuridae at these sites only.