

## Rapid Diversification of RNase A Superfamily Ribonucleases from the Bullfrog, *Rana catesbeiana*

Helene F. Rosenberg,<sup>1</sup> Jianzhi Zhang,<sup>1</sup> You-Di Liao,<sup>2</sup> Kimberly D. Dyer<sup>1</sup>

<sup>1</sup>Laboratory of Host Defenses, National Institute of Allergy and Infectious Diseases, National Institutes of Health, Bethesda, Maryland

<sup>2</sup>Institute of Biomedical Sciences, Academia Sinica, Taipei 115, Taiwan

Received: 27 November 2000 / Accepted: 16 February 2001

**Abstract.** We present sequences of five novel RNase A superfamily ribonuclease genes of the bullfrog, *Rana catesbeiana*. All five genes encode ribonucleases that are similar to Onconase, a cytotoxic ribonuclease isolated from oocytes of *R. pipiens*. With amino acid sequence data from 14 ribonucleases from three *Rana* species (*R. catesbeiana*, *R. japonica*, and *R. pipiens*), we have constructed bootstrap-supported phylogenetic trees that reorganize these ribonucleases into five distinct lineages—the pancreatic ribonucleases (RNases 1), the eosinophil-associated ribonucleases (RNases 2, 3, and 6), the ribonucleases 4, the angiogenins (RNases 5) and the *Rana* ribonucleases—with the *Rana* ribonucleases no more closely related to the angiogenins than they are to any of the other ribonuclease lineages shown. Further phylogenetic analysis suggests the division of the *Rana* ribonucleases into two subclusters (A and B), with positive (Darwinian) selection ( $d_N/d_S > 1.0$ ) and an elevated rate of radical nonsynonymous substitution ( $d_R$ ) contributing to the rapid diversification of ribonucleases within each cluster. This pattern of evolution—rapid diversification via positive selection among sequences of a multigene cluster—bears striking resemblance to what we have described for the eosinophil-associated ribonuclease genes of the rodent *Mus musculus*, a finding that may have implications with respect to the physiologic function of this unique family of proteins.

**Key words:** Ribonucleases — Molecular evolution — Amphibians

### Introduction

The RNase A superfamily is an unusual group of secretory proteins whose members share elements of structure and enzymatic activity while diverging extensively, ostensibly to promote divergent and unique biologic activities (D'Alessio & Riordan 1997; Beintema and Kleinedam 1998). The mammalian pancreatic ribonucleases (RNases 1, including bovine pancreatic RNase A) are the prototype of this family which also includes the eosinophil-derived neurotoxin (EDN/RNases 2) eosinophil cationic protein (ECP/RNases 3), ribonucleases 4 and 6, and the angiogenins (RNases 5).

Among the few nonmammalian members are the RNase A ribonucleases identified in several species of the bullfrog, genus *Rana* (Youle and D'Alessio 1997; Irie et al. 1998). Among the earliest of this group to be identified was Onconase, a protein isolated from oocytes of the leopard frog, *Rana pipiens*, on the basis of its cytostatic, antineoplastic activity (Ardelt et al. 1991). Although the cysteine structure and N-terminus of onconase (rp-onc) differ from those of the mammalian ribonucleases, the catalytic histidines and lysine and enzymatic activity are maintained. Three additional proteins identified in oocytes and liver from the bullfrogs *Rana catesbeiana* (rc-RNase and rc-L1) and *Rana japonica* (rj-lec from oocytes), were introduced as members of this new group of ribonucleases (Titani et al. 1987; Nitta et

Correspondence to: H.F. Rosenberg, MD, Building 10, room 11N104, LHD/NIAID/NIH, 9000 Rockville Pike, Bethesda, MD 20892; tel: 301 402 1545; fax: 301 402 4369; e-mail: hr2k@nih.gov

al. 1989; Kamiya et al. 1990). Liao et al. (Huang et al. 1998) were the first to isolate a cDNA encoding a *Rana* ribonuclease (rc-RNase), and found mRNA expression in the liver rather than in oocytes where the protein is ultimately localized. Most recently, Liao et al. (2000) reported five novel cDNA sequences encoding *R. catesbeiana* ribonucleases (rc-01 encoding rc-L1, and rc-02, rc-03, rc-04, and rc-06) found in oocyte yolk granule extracts and liver and Chen et al. (2000) described a gender-specific cytotoxic ribonuclease (rp-rap) from *R. pipiens*.

In this work, we report the sequences of yet another five *Rana* ribonucleases (rc-203, rc-204, rc-208, rc-212, and rc-218), which we derived directly via polymerase chain reaction (PCR) using *R. catesbeiana* genomic DNA as a template. We have included all 14 known *Rana* ribonucleases in this evaluation. Our phylogenetic analysis suggests the reorganization of RNase A ribonucleases, with the *Rana* ribonucleases representing one of five distinct lineages. Our analysis also supports the existence of two subclusters within the *Rana* ribonucleases, with positive (Darwinian) selection promoting diversification within each subcluster. Overall, we find that the evolutionary pattern of these ribonucleases is remarkably similar to that of the mouse eosinophil-associated ribonuclease (mEAR) gene cluster, a finding that may have important implications with respect to function.

## Materials and Methods

**Genomic DNA.** Genomic DNA from the bullfrog *Rana catesbeiana* was obtained from lysates of cells of the CCL-41 tongue fibroblast cell line (American Type Culture Collection, Manassas VA). Lysates were prepared by homogenization of  $10^4$  cells in 0.5 ml phosphate-buffered saline.

**Isolation of Coding Sequences by Polymerase Chain Reaction (PCR).** The PCR reactions were as described previously (Singhania et al. 1999; Zhang et al. 2000). Primers were constructed from the coding sequence reported for rc-RNase (Huang et al. 1998) as follows: 5' to 3': 5'-AGAATGTGTGCAAAAATCT-3' (nucleotides 241–258, AF039104) and 3' to 5': 5'-TGCTGGTCAAACACTACAGTATG-3' (nucleotides 683–663). Each reaction included 5  $\mu$ l 10X magnesium-containing reaction buffer (Boehringer Mannheim, Indianapolis, IN), 1  $\mu$ l 10 mM dNTPs, 2.5  $\mu$ l of each 20  $\mu$ M stock of primer, 5  $\mu$ l cell lysate, 1  $\mu$ l Taq polymerase (1 unit/ $\mu$ l, Boehringer Mannheim) with distilled water to a 50  $\mu$ l volume. Reactions took place in a Perkin Elmer 9600 thermocycler with the following cycling conditions: 30 sec at 95°C, 30 sec at 55°C, 60 sec at 72°C for 35 cycles. Multiple products present in the resulting amplicons were subcloned en masse into the PCR II TA cloning vector (Invitrogen, San Diego, CA) and multiple clones were sequenced as described (Zhang et al. 2000). To avoid confounding results from PCR-induced mutation (estimated at 1 nucleotide in 1000 to 1 in 10,000 under these conditions), each sequence was isolated at least twice from independent cycling reactions, and only sequences with three or more independent changes in encoded amino acid sequence were considered in the analysis.

**Evolutionary, Statistical, and Sequence Analysis.** Pairwise distances, tree construction, and bootstrap analyses were performed via

the MEGA2b2 analysis software (Kumar et al. 1993; available online at <http://www.megasoftware.net>). Conservative and radical nonsynonymous substitution was determined by the method of Zhang (2000). Sequence alignments, sequence similarity, and determination of isoelectric points were performed with the assistance of the GAP, PRETTY, and PEPTIDESORT algorithms of the Wisconsin Genetics Computer Group program available online at the National Institutes of Health. Averages and standard deviations were determined via Microsoft Excel.

## Results

### *Identification of Novel Ribonuclease Genes from the Bullfrog, Rana catesbeiana*

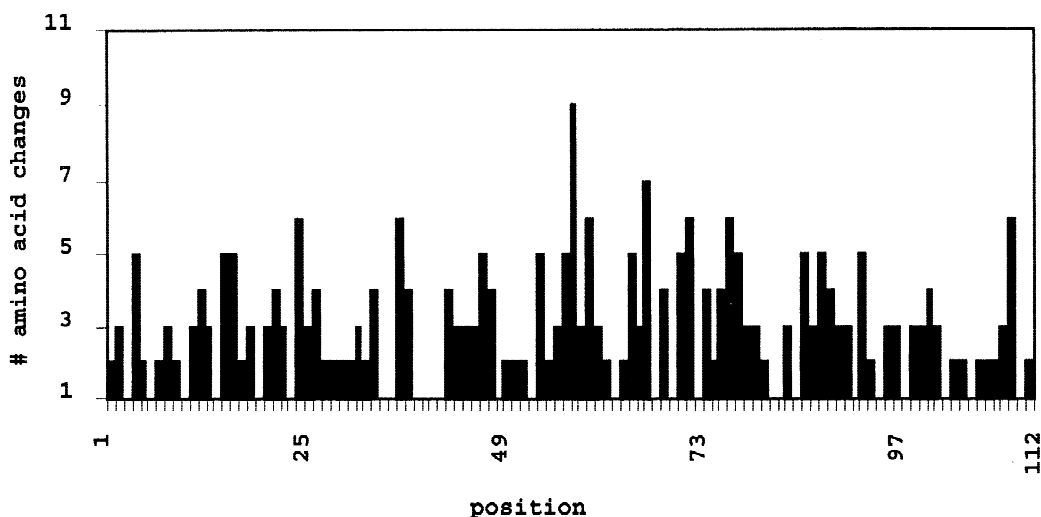
Shown in Fig. 1A is an alignment of the amino acid sequences of the 14 ribonucleases identified in three *Rana* species (*R. catesbeiana*, *R. japonica*, and *R. pipiens*). The novel sequences are those of the 200 series: rc-203, rc-204, rc-208, rc-212, and rc-218. Several of these sequences were also identified by PCR amplification but not reported in the studies performed by Liao et al. (2000). The amino terminal “Q” of these novel sequences has been inferred based on similarity to the amino acid sequence of rc-RNase (rc-RN), whose terminus was determined directly via protein sequencing (Titani et al. 1987); the 22 amino acid signal sequences encoded by the full length cDNAs are not shown. Four of the five novel sequences include eight cysteines (Fig. 1A, at arrow), two histidines and one lysine (at  $\alpha$ 's) spaced in a manner consistent with those of other *Rana* ribonucleases; rc-204, although otherwise a complete sequence, is missing the fourth cysteine. The five novel sequences range in length from 108 to 111 amino acids, and are thus more similar in length to rc-RNase, rc-L1, and rjlec (110 to 111 amino acids) than they are to rc-02, rc-03, rc-04, rc-06, rp-onc, or rp-rap which are all somewhat smaller (104–106 amino acids). The isoelectric points of the five novel sequences range from 8.42 to 9.51, and are indistinguishable from those of the previously studied *Rana* sequences.

**Sequence Variability.** Although there is strict retention of the structural cysteines and catalytic histidines and lysine among the *Rana* ribonucleases (save for rc-204, as above), there is significant variability in other regions. Figure 1B documents the number of different amino acids at a given site, with “1” indicating conserved sequence at the given position in all 14 *Rana* ribonucleases. The average variability across the 112 positions is  $2.9 \pm 1.6$  different amino acids per site. The maximum variability is 9, found at position 57. There are only 26 sites of complete sequence conservation, which include seven of the eight canonical cysteines and the three catalytic residues noted above. There is complete conservation of the RNase A superfamily signature motif CK-NTF (amino acids 34–39, numbered as per Fig. 1A)

A	α	▼	▼α	▼	70		
rc203	QTWAKFQQKH	IPSTSSINCN	TIMDNNIYIV	GGQCKKVNTF	IISSATTVKA	ICNGV-TNSN	VLSPTRFQLD
rc204	.D.PT.....	.....D..	....KD....	R.....	..Y.....	..T...L...	...T.....N
rc208	.N..T.....	.TN.....	.....	.....G.....	.....	..T...I.M.	...T.....N
rc212	.N..T.....	.TN.....S	N..N.SL...	.....	.A.....G	..S....DKK	...S.K....
rc218	.N..T.EQ.	.TN.....S	N..N.SL...	.....	.A.....G	..S....DKK	...S.K....
rc-02	.N.ET..K..	LTD.RDVK.D	AE.KKALF--	--D..QK...	.YARPGR.Q.	L.KNIIVSK.	...TDE.Y.S
rc-03	.D.ET..K..	LTD.KKVK.D	VE.AKALF--	--D...T...	.YALPGR.K.	L.KNIRD.TD	...RDA.L.P
rc-04	.D..T.KK..	LTD.WDVD.D	NL.PTSLF--	--D..DK...	.Y.LPGP...	L.R..IFSAD	...NSE.Y.A
rc-06	.D.DT..K..	LTD.KKVK.D	VE.KKALF--	--D...T...	.FARPPR.Q.	L.KNIKN.T	...RDV.Y.P
rc-RN	.N..T.....	.IN.PI.....	.....	.....R.....	.....	..T...I.M.	...T.....N
rc-L1	.N....KE..	.R.....D..	....KA....	..K..ER...	....EDN...	..S..SPDRK	E..T.S.K.N
rjlec	.N....E..	..N..N....	....KS....	....ER....	.....	..S.AS..R.	...T.R.Q.N
rponc	ED.LT..K..	.TN.RDVD.D	N..ST.LF--	--H..DK...	.Y.RPEP...	..K.IIASK.	..TTSE.Y.S
rprap	.D.LT..K..	LTN.RDVD..	N..ST.LF--	--H..DK...	.Y.RPEP...	..K.IIASK.	..TTSE.Y.S

	▼	▼	▼	▼	α	▼	#aa	pI	%aa seq sim
rc203	TCTRTSITPR	PCPYSSKKET	NKICVKCENQ	LPVHFAGIGKC-			110	9.51	(100)
rc204	--...F..S.	.....T...	.....E	Y.....			108	8.78	84
rc208	.....	.....RT.N	.Y.....	Y.....	.....R.P		111	8.63	87
rc212	I...IF....	.....RT..	.Y.....	Y.....	.....Q.P		111	8.85	79
rc218	I...IF....	.....RT..	.Y.....	Y.....	.....Q.P		111	8.42	79
rc-02	D.N.IKL---	..H.KL..SS	.T..IT..K	....VAVEE.P			105	8.61	46
rc-03	Q.D.IKL---	..H.KL..S	.T..IT.V..	...I...VS.P			106	8.62	50
rc-04	E.NVKPRK--	..K.KL..SS	.R..IR..HE	.....V.I.P			105	8.23	52
rc-06	Q.N.KKL---	..H.RLDGS.	.T..LT.MKE	..I...V...P			105	9.98	49
rc-RN	.....	.....RT..	.Y.....	Y.....	.....R.P		111	8.86	87
rc-L1	..I.D.....	....HPSPDN	.....K.	.....V.....			110	8.84	71
rjlec	..I..ATA..	....N.RT..	.V.....	.....R.....			110	9.60	80
rponc	D.NV.R---	..K.KL..S	..F..T...	A....V.V.S..			104	8.42	58
rprap	D.NV.S---	..K.KL..S	.TF..T...	A....V.V.H.-			104	8.62	57

B.



**Fig. 1** (a) Alignment of the amino acid sequences encoded by RNase A superfamily genes from the bullfrog, *Rana catesbeiana*. Novel sequences are rc-203, rc-204, rc-208, rc-212, and rc-218 (GenBank accession numbers AF351207, AF351208, AF351209, AF351210, AF351211, respectively). Other *R. catesbeiana* (rc) sequences were reported previously (accession numbers AF039104, AF242553–AF242556). Rj-lec is amino acid sequence derived from protein purified from oocytes of the bullfrog *R. japonica* (JX0120), and rp-onc (A39035) and rp-rap (AF165133), from *R. pipiens*. The overhead symbol “α” denotes catalytic histidines or lysine and the symbol “▼” marks the conserved cysteines. Listed at the right are the number of amino acids in each sequence (as shown, without signal sequence); amino terminus inferred from purified protein sequence of rc-L1, JX0085), the calculated isoelectric points (pI), and the percentage similarity of each to the first sequence, rc-203. (b) Sequence variability. Analysis of the 14 *Rana* ribonuclease sequences, indicating the number of different amino acids found at each position within the total 112 sites (one = conserved throughout). The average variability is 2.9 ± 1.6 different amino acids per site.

which includes cysteine-34, the catalytic lysine-35, and threonine-38 of the nucleotide base B<sub>1</sub> binding site. Also conserved is the B<sub>1</sub> binding site element phenylalanine-105, whereas histidine-104 and glutamate-98 of the B<sub>2</sub> binding site are somewhat (but not completely) conserved (Irie et al. 1998).

#### *Evolutionary Dendrogram of Rana Ribonucleases.*

The relationships among the 14 *Rana* ribonucleases are depicted in the neighbor-joining tree in Fig. 2A. There is strong bootstrap support for the division of the *Rana* ribonucleases into two distinct subclusters. Subcluster A includes all the longer (110–111 amino acid) sequences: the five novel sequences (rc-203, rc-204, rc-208, rc-212, and rc-218), the previously identified rc-L1 and rc-RNase (rc-RN), and the single sequence isolated from the species *R. japonica* (rj-lec). Subcluster B includes the shorter (105–106 amino acid) sequences: rc-02, rc-03, rc-04, rc-06, rp-onc, and rp-rap. The relative positions of the ribonucleases within the subclusters have somewhat weaker bootstrap support. Although this has not been investigated exhaustively via direct hybridization or PCR amplification techniques, it is interesting to note that there is no 1:1 pairing among ribonucleases between the *Rana* species. A similar pattern was observed among the mouse ribonucleases of the mEAR cluster, shown by Zhang et al. (2000) to have evolved via a process of rapid birth-death and gene sorting.

*Synonymous and Nonsynonymous Substitution.* We have calculated nonsynonymous ( $d_N$ ) and synonymous ( $d_S$ ) substitutions per respective site for pairs of *Rana* ribonuclease genes both within and between subclusters A and B (Fig. 2B). The ratio of nonsynonymous to synonymous substitutions between subclusters was unremarkable, with all values falling well below 1.0, and an average  $d_N/d_S$  of  $0.63 \pm 0.08$  for the 35 pairs evaluated. In contrast, the values for  $d_N/d_S$  calculated for pairs within subclusters were quite remarkable, and stand in strong support for positive (Darwinian) selection as contributing to their diversification. The average  $d_N/d_S$  for the 31 pairs within clusters (21 A  $\times$  A and 10 B  $\times$  B) is  $1.7 \pm 0.70$ , with the highest value,  $d_N/d_S = 3.99$ , obtained from the comparison of nucleotide sequences encoding subcluster A ribonucleases rc-203 and rc-RN. Values of  $d_N/d_S > 1.0$  were also observed among genes of the mEAR cluster (Singhania et al. 1999), a point that will be considered in greater detail in the Discussion.

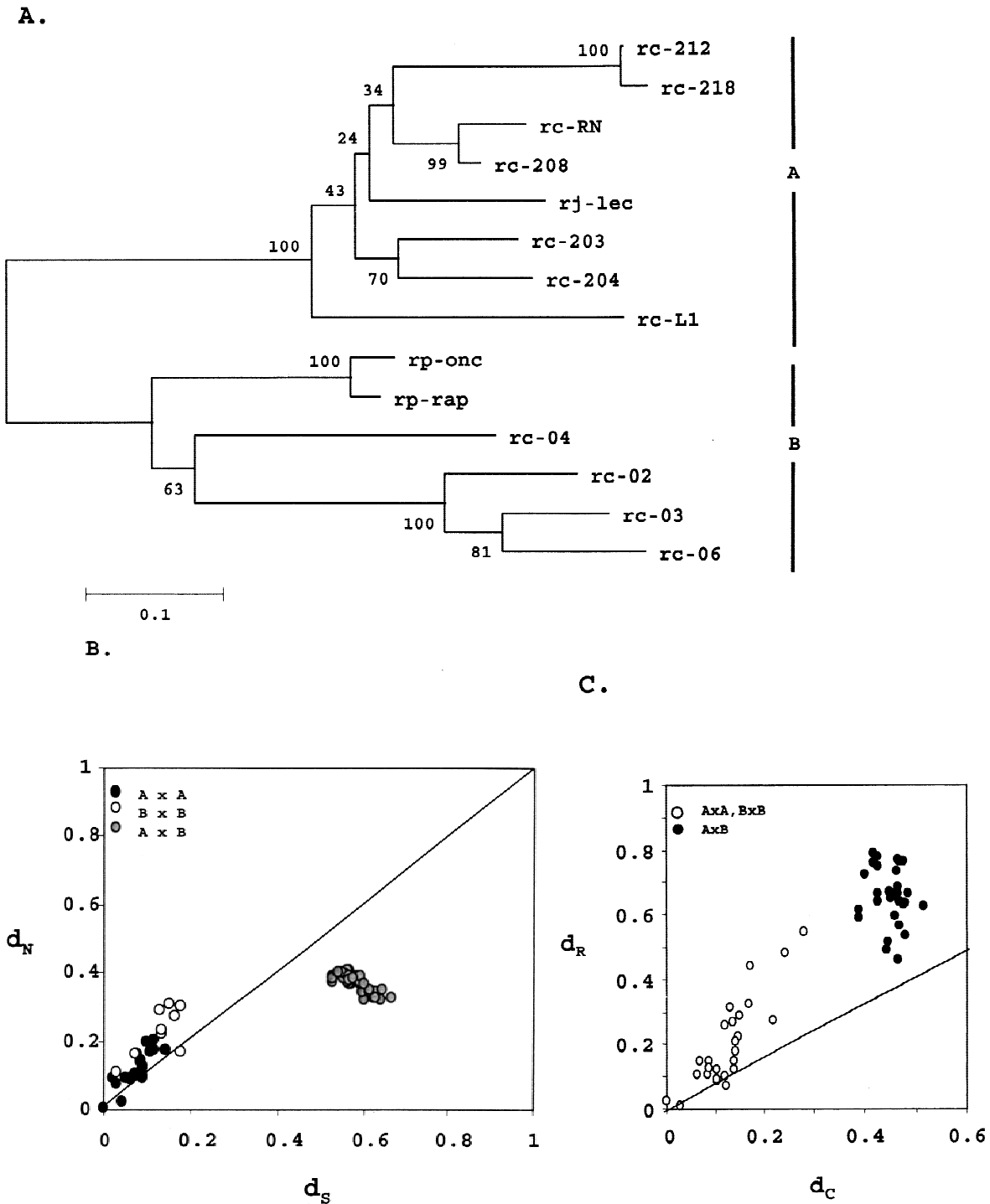
A comparison of radical (charge) to conservative nonsynonymous substitutions per site among pairs of *Rana* ribonucleases are shown in Fig. 2C. The average ratio  $d_R/d_C$  within clusters is  $1.47 \pm 0.56$ , and between clusters,  $1.49 \pm 0.24$ . The ratios both within and between *Rana* ribonuclease subclusters are significantly higher than the average  $d_R/d_C$  calculated by Zhang (2000) for 47 paired sequences from three mammalian species, at 0.73

$\pm 0.24$  ( $P < 0.00001$ ). It is not certain how well this ratio applies directly to comparisons among nonmammalian gene sequences, but it is interesting that the rate of radical nonsynonymous substitution is high, given that the isoelectric points of the various *Rana* ribonucleases do not differ markedly from one another.

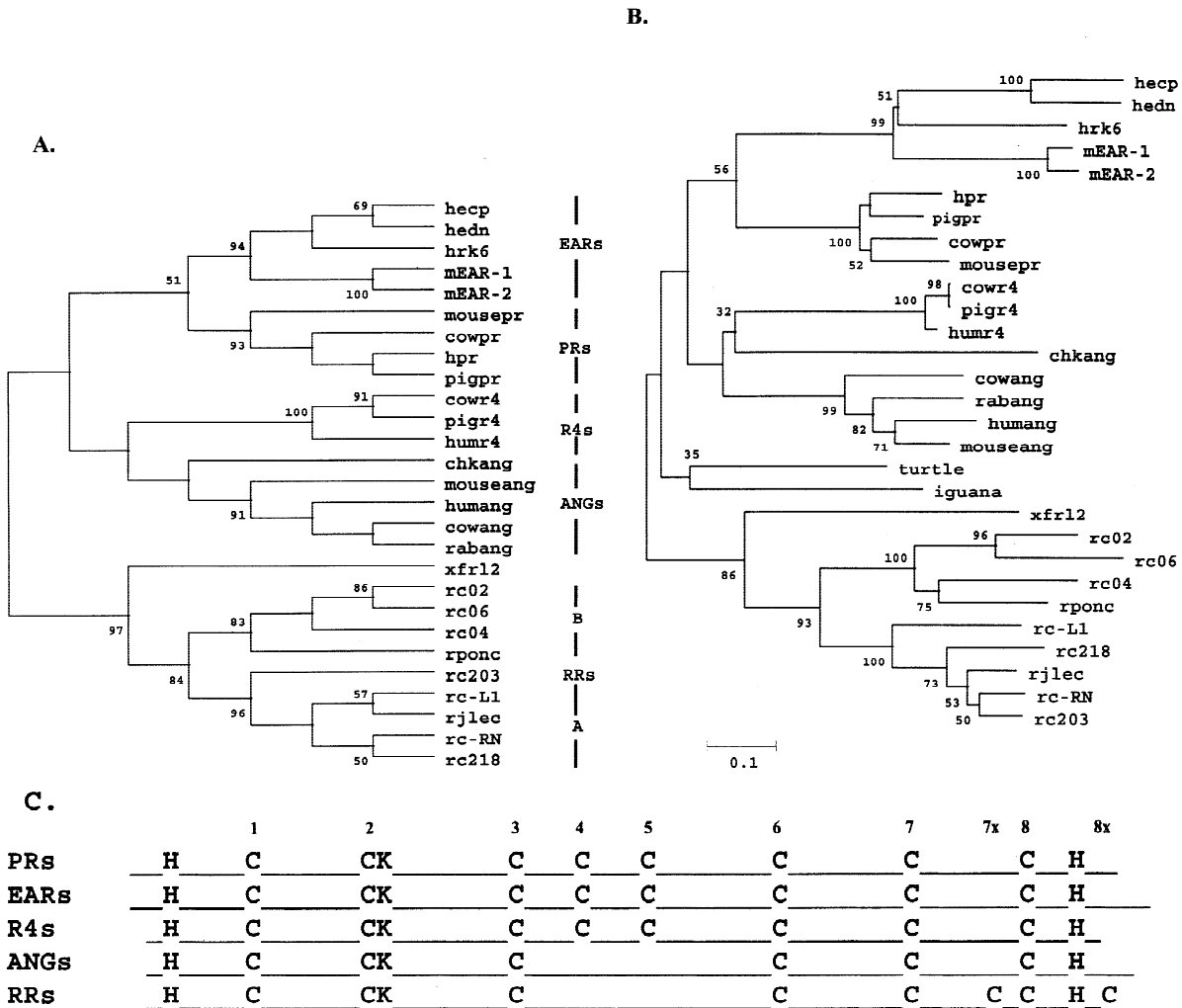
As there are no mammalian genes that are clearly orthologous to the *Rana* ribonucleases, we have used the following simplification to estimate the age of the *Rana* ribonuclease subclusters: given the average  $d_S = 0.46$  synonymous substitutions per site determined by Makalowski and Boguski (1998) for orthologous mouse/human gene pairs and an estimated divergence time 100 million years for these two species (Kumar and Hedges, 1998), the average  $d_S = 0.58 \pm 0.04$  synonymous substitutions per site calculated from the comparison of all subcluster A to B pairs of *Rana* ribonucleases translates into an estimated time of divergence of these two clusters at  $0.58 \times 100/0.46 = 126$ , or  $\sim 130$  million years.

#### *Evolutionary Dendrogram of Major RNase A Ribonuclease Lineages.*

A maximum-parsimony (MP) tree linking *Rana* ribonucleases to the other major RNase A superfamily lineages is shown in Fig. 3A. A minimum-evolution (ME) tree which also includes two reptile ribonucleases (from turtle and iguana) is shown in Fig. 3B. The MP tree demonstrates strong bootstrap support for dividing these RNase A ribonucleases into five distinct lineages: the pancreatic ribonucleases (PRs, also known as RNase 1); the eosinophil-associated ribonucleases (EARs) including the human eosinophil-derived neurotoxin (hEDN, RNase 2) and eosinophil cationic protein (hECP, RNase 3), their mouse orthologs mEAR-1 and mEAR-2, and human ribonuclease k6 (hrk6, RNase 6); the ribonuclease 4s (R4s, RNase 4), the angiogenins (ANGs, RNase 5), and the *Rana* ribonucleases (RRs). There also is strong bootstrap support (84%) for grouping the single ribonuclease sequence from the African clawed frog *Xenopus laevis* (xflr12) (Kinoshita et al. 1995) together with the *Rana* ribonucleases. The division of the true *Rana* ribonucleases into subclusters A and B is likewise maintained in this larger analysis. Interestingly, despite similarities in cysteine structure (Fig. 3C), the *Rana* ribonucleases are no more closely related to the angiogenins than they are to any of the other distinct RNase A lineages. The minimum evolution (ME) tree, which also includes two reptile ribonuclease sequences (from snapping turtle and iguana, respectively), also supports the concepts of *Rana* ribonucleases as a distinct lineage, of the grouping of the *Xenopus* and *Rana* sequences together, and of the division of the *Rana* sequences into subclusters A and B. Interestingly, the ME tree groups the sequence considered to be chicken (*Gallus gallus*) angiogenin with the ribonuclease 4 cluster, although without strong bootstrap support. Similarly, the reptile ribonucleases appear to form yet another distinct cluster. Isolation of additional amphibian and reptile ri-



**Fig. 2** (a) Phylogenetic tree of the *Rana* ribonucleases. Neighbor-joining tree (Poisson-corrected pairwise distances) constructed with encoded (all rc sequences and rp-rap) and experimentally determined (for rj-lec and rp-0nc) amino acid sequences, with accession numbers as listed in the legend to Figure 1A. Bootstrap percentages (500 replications) are shown at the interior branches. On the basis of these findings, the *Rana* ribonucleases have been divided into subclusters A and B as indicated (b) Nonsynonymous ( $d_N$ ) and synonymous ( $d_S$ ) substitutions. Pairwise distances (uncorrected) calculated for *Rana* ribonuclease nucleotide sequence pairs by the method of Nei and Gojobori (1986) within subcluster A (A  $\times$  A, filled circles,  $n = 21$ ), within subcluster B (B  $\times$  B, open circles,  $n = 10$ ), and between subclusters A and B (A  $\times$  B, shaded circles,  $n = 35$ ). Bisecting line represents  $d_N/d_S = 1.0$ . Evaluation was performed with nucleotide sequence encoding amino acid sequences shown in Figure 1A, omitting rj-lec and rp-0nc for which there are no nucleotide sequences available. (c) Radical ( $d_R$ ) vs. conservative ( $d_C$ ) nonsynonymous substitutions. Calculation of  $d_R$  (charge substitution only) and  $d_C$  were as per the method of Zhang (2000) for pairs of sequences both within (open circles) or between (filled circles) subclusters A and B. Bisecting line represents  $d_R/d_C = 0.73$ , which is the experimentally determined average  $d_R/d_C$  for orthologous genes from three mammalian species.



**Fig. 3** (a) Maximum-parsimony (MP) and (b) minimum evolution (ME) phylogenetic trees representing the major RNase A superfamily lineages. The five RNase A lineages are demarcated in (A) include pancreatic ribonucleases (PRs), eosinophil associated ribonucleases (EARS), RNase 4s (R4s), angiogenins (ANGs), and *Rana* ribonucleases (RRs). Bootstrap percentages (100 replications for MP, 500 replications for ME) are shown at the interior branches. Accession numbers for *Rana* sequences are listed in the Legend to Figure 1A; others include human ECP (hecp, X15161); human EDN (hedn, M24157); human RNase k6 (hrk6, U64998); mEAR-1 (U72032); mEAR-2 (U72031); human PR (hpr, A27235); pig PR (A91391); cow PR (S00897); mouse PR (S22598); cow R4 (JX0115); pig R4 (A53180); human R4 (humr4, 152489); chicken ANG (chkgang, X61193); cow ANG (A32474); rabbit ANG (rabang, B43825); human ANG (humang, A90498); mouse ANG (A35932); turtle RNase (A91155); iguana RNase (S41111); xenopus FGF receptor ligand ribonuclease (xfrl2, AF159166). (c) Relative positions of the cysteines and catalytic lysine and histidines in RNase A ribonucleases. Shown are representative structures of the five lineages of RNase A ribonucleases defined in (a). The overhead numbering indicates the eight traditional positions for cysteines as defined in the first four RNase A lineages. The angiogenins (ANGs) and *Rana* ribonucleases (RRs) have no cysteines at positions 4 and 5; the *Rana* ribonucleases have two additional cysteines, at positions 7x (distal to cysteine #7) and 8x (distal to cysteine #8) that form a synapomorphic disulfide bond (Leland et al. 2000). The catalytic histidines (H) and lysine (K) are as shown.

bonuclease sequences will ultimately clarify these relationships.

## Discussion

In this work, we present sequences of five novel RNase A superfamily ribonucleases of the bullfrog, *Rana catesbeiana*. Adding these ribonuclease sequences to those previously characterized in this and in other *Rana* spe-

cies, we present a bootstrap-supported phylogenetic tree defining two *Rana* ribonuclease subclusters (A and B). Pairwise comparisons within clusters suggest that diversification is promoted by positive selection, which is similar to what we observed previously among genes of the mouse eosinophil-associated ribonuclease (mEAR) cluster (Singhania et al. 1999, Zhang et al. 2000). We also present a bootstrap-supported tree suggesting reorganization of the RNase A superfamily, with the *Rana* ribonucleases emerging as one of five distinct lineages.

Since the elucidation of the complete amino acid sequence of the *Rana* ribonuclease, Onconase, the relationship of this unusual protein to those of the more familiar RNase A superfamily ribonucleases has been in question. In the absence of a significant number of related sequences, Onconase and similar proteins from *Rana catesbeiana* and *Rana japonica* were initially included in a group that included the angiogenins (Irie et al. 1998), based on the fact that both the *Rana* sequences and the angiogenins were missing cysteines at the canonical 4<sup>th</sup> and 5<sup>th</sup> positions (see Fig. 3B). With the sequences presented here and with those presented previously, we have been able to construct a bootstrap-supported phylogenetic tree demonstrating that the *Rana* ribonucleases stand apart as a unique lineage, with no closer relationship to the angiogenins than to any of the other mammalian RNase A ribonucleases. Although ribonucleases of this type (i.e., ribonucleases that include cysteines 7× and 8×) have been found only in *Rana* species, our estimate of the origin of the two subclusters at ~130 million years suggests that one or more ribonucleases of this class may be found in other members of the family *Ranidae* if not among others of the class *Lissamphibia*. Identification and further analysis of ribonucleases from the genus *Xenopus* may provide some clarification on this point.

Despite their distinction as a separate lineage of RNase A ribonucleases, the pattern of evolution observed among the *Rana* ribonucleases has much in common with the mouse eosinophil-associated ribonuclease (EAR) gene cluster (Singhania et al. 1999; Zhang et al. 2000). Similar to what was observed among the EAR genes of *Mus musculus*, the *Rana* subclusters are characterized by elevated ratios of nonsynonymous to synonymous substitution, with  $d_N/d_S > 1.0$  for all but two of the 31 pairs of genes within the subclusters. Positive selection ( $d_N/d_S > 1.0$ ) was also observed among the primate eosinophil cationic protein genes, (Zhang et al. 1998) which, together with provocative functional data, has been interpreted as support for a role for the eosinophil-associated ribonucleases in antiviral host defense (Domachowske et al. 1998; Rosenberg and Domachowske 1999). Although it is tempting to consider the possibility of a similar physiologic function for the *Rana* ribonucleases, this interpretation would be decidedly premature, and it is difficult to envision this sort of role for the proteins that are localized in the oocyte. However, it is worthwhile to note that several of the *Rana* ribonucleases were initially isolated as carbohydrate-binding proteins (Titani et al. 1987; Kamiya et al. 1990); we can consider the possibility that these “bifunctional” ribonuclease-lectins function in some way similarly to the combined efforts of the eosinophil-associated ribonucleases and the eosinophil-specific galectin, the Charcot-Leyden crystal protein (Leonidas et al. 1995; Dyer and Rosenberg 1996).

With respect to function in general, the *Rana* ribonucleases have been characterized as cytotoxins, but only in assays targeting mammalian tumor cell lines (Youle and D’Alessio 1997; Irie et al. 1998). Though the pharmaceutical implications of this activity have yet to be fully realized, the physiologic implications are much less clear, as this toxicity may not be telling us much about the way in which *Rana* ribonucleases interact with cells and proteins present in *Rana* (as opposed to mammalian) species. Overall, a more thorough understanding of the localization, cellular distribution, gender- and species-limited expression, and activity of all the *Rana* ribonucleases will provide us with a larger basis by which we can interpret these intriguing evolutionary findings.

*Acknowledgments.* We thank Dr. Harry L. Malech and Dr. John I. Gallin for their ongoing support of work in progress in our laboratory.

## References

- Ardelt A, Mikulski SM, Shogen K (1991) Amino acid sequence of an anti-tumor protein from *Rana pipiens* oocytes and early embryos. *J Biol Chem* 266:245–251
- Beintema JJ, Kleineidam RG (1998) The ribonuclease A superfamily: general discussion. *Cell Mol Life Sci* 54:825–832
- Chen S, Le SY, Newton DL, Maizel JV, Rybak SM (2000) A gender-specific mRNA encoding a cytotoxic ribonuclease contains a 3' UTR of unusual length and structure. *Nucl Acids Res* 28:2375–2382
- D’Alessio G, Riordan JF (1997) Ribonucleases: structures and functions. San Diego, CA: Academic Press
- Domachowske JB, Dyer KD, Bonville CA, Rosenberg HF (1998) Recombinant human eosinophil-derived neurotoxin functions as an effective antiviral agent against respiratory syncytial virus. *J Infect Dis* 177:1458–1464
- Dyer KD, Rosenberg HF (1996) Eosinophil Charcot-Leyden crystal protein binds to beta-galactoside sugars. *Life Sci* 58:2073–2082
- Huang H-C, Wang S-C, Leu Y-J, Lu S-C, Liao Y-D (1998) The *Rana catesbeiana rcr* gene encoding a cytotoxic ribonuclease. *J Biol Chem* 273:6395–6401
- Irie M, Nitta K, Nonaka T (1998) Biochemistry of frog ribonucleases. *Cell Mol Life Sci* 54:775–784
- Kamiya Y, Oyama F, Oyama R, Sakakibara F, Nitta K, Kawauchi H, Takayanagi Y, Titani K (1990) Amino acid sequence of a lectin from Japanese frog (*Rana japonica*) eggs. *J Biochem* 108:139–143
- Kinoshita N, Minshull J, Kirschner MW (1995) The identification of two novel ligands of the FGF receptor by a yeast screening method and their activity in *Xenopus* development. *Cell* 83:621–630
- Kumar S, Tamura K, Nei M (1993) Molecular evolutionary genetics analysis, version 1.02. Institute of Molecular Evolutionary Genetics, Pennsylvania State University, State College, PA
- Kumar S, Hedges SB (1998) A molecular timescale for vertebrate evolution. *Nature* 392:917–920
- Leonidas DD, Elbert BL, Zhou Z, Leffler H, Ackerman SJ, Acharya KR (1995) Crystal structure of human Charcot-Leyden crystal protein, an eosinophil lysophospholipase, identifies it as a new member of the carbohydrate-binding family of galectins. *Structure* 3:1279–1393
- Liao Y-D, Huang H-C, Leu Y-J, Wei C-W, Tang P-C, Wang S-C (2000) Purification and cloning of cytotoxic ribonucleases from *Rana catesbeiana*. *Nucl Acids Res* 28:4097–4104

- Makalowski W, Boguski MS (1998) Evolutionary parameters of the transcribed mammalian genome: an analysis of 2,820 orthologous rodent and human sequences. *Proc Natl Acad Sci USA* 95:9407–9413
- Nei M, Gojobori T (1986) Simple methods for estimating the numbers of synonymous and nonsynonymous nucleotide substitutions. *Mol Biol Evol* 3:418–426
- Nitta R, Katayama N, Okabe Y, Iwama M, Watanabe H, Abe Y, Okazaki T, Ohgi K, Irie M (1989) Primary structure of a ribonuclease from bullfrog (*Rana catesbeiana*) liver. *J Biochem* 106:729–735
- Rosenberg HF, Domachowske JB (1999) Eosinophils, ribonucleases and host defense: solving the puzzle. *Immunol Res* 20:261–274
- 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
- Titani K, Takio K, Kuwada M, Nitta K, Sakakibara F, Kawauchi H (1987) Amino acid sequence of a sialic acid-binding lectin from frog (*Rana catesbeiana*) eggs. *Biochemistry* 26:2189–2194
- Youle RJ, D'Alessio G (1997) Anti-tumor ribonucleases. In: D'Alessio G, Riordan JF (eds) *Ribonucleases: structures and functions*. Academic Press, San Diego, CA, pp 491–515
- 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 (2000) Rates of conservative and radical nonsynonymous nucleotide substitutions in mammalian nuclear genes. *J Mol Evol* 50:56–68
- 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* 97:4701–4706