Zebrafish Ribonucleases Are Bactericidal: Implications for the Origin of the Vertebrate RNase A Superfamily

Soochin Cho and Jianzhi Zhang

Department of Ecology and Evolutionary Biology, University of Michigan

Understanding the evolutionary origin of the ribonuclease (RNase) A superfamily is of great interest because the superfamily is the sole vertebrate-specific enzyme family known to date. Although mammalian RNases have a diverse array of biochemical and physiological functions, the original function of the superfamily at its birth is enigmatic. Such information may be obtained by studying basal lineages of the vertebrate phylogeny and is necessary for discerning how and why this superfamily originated. Here, we clone and characterize 3 RNase genes from the zebrafish, the most basal vertebrate examined for RNases. We report 1) that all the 3 zebrafish RNases are ribonucleolytically active, with one of them having an RNase activity comparable to that of bovine RNase A, the prototype of the superfamily; 2) that 2 zebrafish RNases have prominent expressions in adult liver and gut, whereas the 3rd is expressed in adult eye and heart; and 3) that all 3 RNases have antibacterial activities in vitro. These results, together with the presence of antibacterial and/or antiviral activities in multiple distantly related mammalian RNases, strongly suggest that the superfamily started as a host-defense mechanism in vertebrate evolution.

Introduction

The ribonuclease (RNase) A superfamily is the sole vertebrate-specific enzyme family known to date (Lander et al. 2001; Zhang, Dyer, Rosenberg 2002). The prototype of the superfamily, bovine pancreatic RNase (also known as RNase A), has a long history in biochemistry as it has served as an excellent model for studying protein folding (Anfinsen 1973; Udgaonkar and Baldwin 1988; Klefhaber et al. 1995; Sambashivan et al. 2005). The superfamily has also been a model in molecular phylogenetic and evolutionary studies (Beintema et al. 1977; Jermann et al. 1995; Rosenberg et al. 1995; Zhang et al. 1998, 2000; Kleineidam et al. 1999; Zhang and Rosenberg 2002a; Zhang J, Zhang YP, Rosenberg 2002; Yu and Zhang 2006; Zhang 2006, 2007). In humans, the superfamily comprises 8 canonical members (Zhang, Dyer, Rosenberg 2002), including the pancreatic ribonuclease (RNase 1), eosinophil-derived neurotoxin (or RNase 2), eosinophil cationic protein (or RNase 3), RNase 4, angiogenin (ANG or RNase 5), RNase 6 (or k6), RNase 7, and RNase 8. More recent studies expanded the superfamily to include 5 additional members, RNases 9–13 (Penttinen et al. 2003; Castella, Benedetti, et al. 2004; Castella, Fouchecourt, et al. 2004; Devor et al. 2004; Cho et al. 2005). RNases have a wide variety of functions, including degradation of dietary RNAs (Barnard 1969), stimulation of blood vessel formation (Fett et al. 1985), tumor growth suppression (Ardelt et al. 1991), innate immunity (Dyer and Rosenberg 2006), and involvement in motor neuron proliferation (Greenway et al. 2006) and male reproduction (Zhu et al. 2007). Notwithstanding the great functional diversity, members of the superfamily have a number of common features. First, as secreted proteins, they all have a signal sequence at the N-terminus. Second, the entire coding region is contained in a single exon. Third, canonical RNases (i.e., RNases 1–8) possess 6–8 conserved cysteines forming disulfide bridges. Fourth, the catalytic triad and several other signature motifs are well conserved among the canonical RNases (Beintema and Kleineidam 1998).

Except frogs (Rosenberg et al. 2001), nonmammalian species possess relatively few RNases, which have only 6 conserved cysteines much like mammalian RNase 5 (Strydom 1998). Phylogenetic analysis of RNase genes identified from numerous completely sequenced genomes suggested a substantial expansion of the RNase superfamily in early mammalian evolution prior to the placental–marsupial divergence, presumably from an RNase 5–like ancestral gene (Cho et al. 2005; Cho and Zhang 2006). Among the 8 mammalian canonical RNases, RNase 2, 3, 5, and 7 have antibacterial and/or antiviral activities (Lehrer et al. 1989; Domachowske et al. 1998; Harder and Schroder 2002; Hooper et al. 2003; Zhang et al. 2003). Whether RNase 8 is bactericidal is controversial as different labs obtained different results (Zhang, Dyer, Rosenberg 2002; Rudolph et al. 2006). At any rate, the presence of antibacterial/antiviral activities in diverse lineages of the mammalian RNase superfamily suggests that host defense might be the original physiological role of the superfamily at its birth (Cho et al. 2005; Dyer and Rosenberg 2006). Testing this hypothesis requires studies of nonmammalian vertebrates, especially basal vertebrates such as fish, to examine whether their RNases also have host-defense functions. Recently, Nitto et al. (2006) reported their detailed characterization of 2 chicken RNases and showed that one of them has bactericidal activities. RNases have also been identified in frogs (Huang et al. 1998; Irie et al. 1998; Rosenberg et al. 2001), turtles (Beintema et al. 1985; Katekaew et al. 2006), and iguana (Zhao et al. 1994; Nitto et al. 2005) with their ribonucleolytic activities characterized. However, it has not been extensively tested, compared with the mammalian RNases, as whether these nonmammalian RNases have any host-defense functions. In 2005, we reported our preliminary identification of RNase genes in the then incomplete zebrafish genome sequence (Cho et al. 2005). Here, we characterize 3 RNase genes that we identified from the now much improved zebrafish genome sequence. We show that all 3 fish RNases are bactericidal, strongly supporting the hypothesis that the RNase A superfamily started as a host-defense mechanism in early vertebrate evolution.

Key words: RNase, angiogenin, *Danio rerio*, antibacterial, host defense, molecular evolution.

E-mail: jianzhi@umich.edu.

Mol. Biol. Evol. 24(5):1259–1268. 2007
doi:10.1093/molbev/msm047
Advance Access publication March 7, 2007

Table 1
Three Zebrafish RNase Genes Identified in This Study

Name	Chromosome	ORF Location	Direction	ORF Length[a]	MW (kDa)[b]	pI[c]
Dr-RNase1	14	77167683–77168129	+	447 bp	14.68	8.95
Dr-RNase2	22	6683218–6683667	+	447 bp	14.76	9.18
Dr-RNase3	22	6647326–6647730	+	405 bp	12.82	8.82

[a] Excluding the stop codon.
[b] Molecular weight of the mature protein.
[c] Isoelectric point of the mature protein.
NOTE.—ORF, open reading frame.

Materials and Methods

In Silico Identification of Fish RNase Genes

For convenience, in this paper "RNase" is used to refer to a member of the RNase A superfamily. To identify all RNase genes in the zebrafish (*Danio rerio*) genome, we performed TBlastN searches (*E* value cutoff = 10^{-10}) in the latest zebrafish genome assembly (Zv6) available at the Ensembl server (http://www.ensembl.org/Danio_rerio/index.html) and the UCSC (University of California, Santa Cruz) Genome Browser (http://genome.ucsc.edu/) using all known human RNases (RNases 1–13) as queries (Cho et al. 2005). The genome sequence has 6.5× to 7× coverage with >95% of the sequences placed on chromosomes. Our search yielded 3 sequences (*Dr*-RNase 1–3) that show significant similarities to the query sequences and contain key features of RNases. In addition, we performed TBlastN searches against the Atlantic salmon (*Salmo salar*) and rainbow trout (*Oncorhynchus mykiss*) expressed sequence tag (EST) databases available at the cGRASP (http://web.uvic.ca/cbr/grasp/) and the medaka (*Oryzias latipes*) genome sequence (6.7× coverage) available at NIG (National Institute of Genetics) DNA Sequencing Center (http://dolphin.lab.nig.ac.jp/medaka/). We also conducted thorough searches against the genome sequences of stickleback (*Gasterosteus aculeatus*, BROAD S1, 11× coverage), fugu (*Takifugu rubripes*, FUGU 4.0, 8.7x coverage), and Tetraodon (*Tetraodon nigroviridis*, TETRAODON 7, 6x coverage) that are available from the Ensembl server (http://www.ensembl.org/).

Zebrafish Samples and Nucleic Acid Purification

We used the wild-type AB zebrafish strain for all our tissue samples in this study. We purified genomic DNAs from the muscle tissue of the sample fish using the PUREGENE genomic DNA purification kit (Gentra Systems, Minneapolis, MN), following the manufacturer's instruction. The final concentration of the purified genomic DNA was adjusted to ~50 ng/μl. Total RNA was purified using TRIZOL reagent (Invitrogen, Carlsbad, CA) from whole-fish bodies at 6 different developmental stages (0–5 days old) and from 9 different adult tissues (eye, heart, brain, liver, gut, testis, ovary, skin, and muscle). Except for the reproductive organs, several fishes of mixed sexes were used. We note that we were not able to completely isolate the gut tissue from other attached soft organs such as pancreas due to technical difficulties.

Rapid Amplification of cDNA Ends and Reverse Transcription–Polymerase Chain Reaction

5′ and 3′ cDNA ends were determined using the FirstChoice RLM-rapid amplification of cDNA ends (RACE) Kit (Ambion Inc., Austin, TX) following manufacturer's instruction. The gene-specific primers used for RACE experiments are listed in supplementary table 1 (Supplementary Material online). Amplified cDNA products were cloned into pCR4-TOPO vector (Invitrogen) and at least 3 independent clones were sequenced for verification. For reverse transcription–polymerase chain reaction (RT-PCR), we generated 1st-strand cDNA using RETROscript Kit (Ambion) and then amplified each gene using gene-specific primers (supplementary table 1, Supplementary Material online). For each gene, the forward and reverse gene-specific primers anneal to 2 separate exons so that cDNA amplifications can be distinguished from genomic contaminant amplifications by product size differences. We used Amplitaq Gold DNA polymerase (Applied Biosystems Inc., Foster City, CA) for all RT-PCRs, and the number of cycles was 35 for all reactions. Amplified products were analyzed using 10% polyacrylamide gels (Invitrogen).

Cloning, Bacterial Expression, and Purification of the Zebrafish RNase Genes

The signal peptides of the 3 zebrafish RNases were predicted using the SignalP 3.0 server (Bendtsen et al. 2004). The mature peptide portion of each gene was amplified using PCR, cloned into the pFLAG-CTS vector of *Escherichia coli* (Sigma, St. Louis, MO), and verified by sequencing. The vector produces a recombinant protein with an octapeptide tag DYKDDDDK (FLAG) attached to its C-terminus, which facilitates its purification and detection with M2 anti-FLAG monoclonal antibody (Sigma) (Zhang et al. 2003). Recombinant proteins were isolated, purified, and quantified as previously described (Zhang et al. 2003).

RNase Assay

The RNase activities of the recombinant proteins against a standard yeast tRNA substrate were measured in 40 mM sodium phosphate buffer (pH 7.4) at 25 °C. Purified RNase was added into 0.8 ml of the above buffer with 1.42 nmol tRNA. The reaction was stopped by 0.5 ml of 20 mM lanthanum nitrate with 3% perchloric acid, and

```
                                                    *                 ↓
RNase A        MALKSLVLLS LLVLVLLLVR VQPSLGKETA AAKFERQHMD SSTSAASSSN YCNQMMK-SR  59
Human RNase5   .VMGLG...L VF..G.G.TP PTLAQDNS-R YTH.LT..Y. AKP-QGRDDR ..ESI.R-R.
Dr-RNase1      .GIHRCTAVV ..L.CAS.ST YGQPAEIRRR YEH.LT..VY G----GITEQ T.DRV.R-Q.
Dr-RNase2      .EILQSAVIF ...FSFSFTV KV.DNES--P YE..L...V. P----DM.VQ K..SEIS-K.
Dr-RNase3      .KTRQSFIIL ...ICAS.AV NSQ------S YND.K.K.LA P---.GMKED D.TTLIVTE.

                     ↓*                  ↓       ↓       ↓           ↓
RNase A        NLT----KDR CKPVNTFVHE SLADVQAVCS QKNVACKNGQ TNCYQSYSTM SITDCRETGS  115
Human RNase5   G..-----SP ..DI...I.G NKRSIK.I.E NE-NGNPHRE -.LRI.K.SF QV.T.KLH.G
Dr-RNase1      GI.RFPTGND ..E....IQA NGNH.RT..T GGGTRQTENR -DL.M.NNQF TVIT.TLRSG
Dr-RNase2      KI.-AKAGND ..K....IQA NKR..N...G NAGNRVVDT- -.LTK.NQPF PVVT.QLKSG
Dr-RNase3      KIK---E.NQ ..KI...IL. TEDKIKG..N TPATDG..HK ------GTGF TVIN.TKI--

                   ↓                   ↓       *
RNase A        SKYPNCAYKT TQANKHIIVA CEGNPYVPVH FDASV-----  --  140
Human RNase5   .PW.P.Q.RA .AGFRNVV.. ..NG--L... L.Q.IFRRP- --
Dr-RNase1      ERH...R.RG KESSRK.V.. ..SE--W.T. YEKG.IV--- --
Dr-RNase2      ERR.H.Q.RG RSSTRY.VLR .DKG--W... Y.EGIIDVNS SG
Dr-RNase3      ENIID.K.NG VKRTTD..LT ..NR--L... YGR.----NN S-
```

FIG. 1.—Protein sequence alignment of the zebrafish RNases (*Dr*-RNases 1–3) with RNase A (bovine pancreatic RNase) and human RNase 5 (angiogenin). Dashes indicate alignment gaps and dots represent the same amino acids as in the 1st sequence. The 8 structural cysteines are indicated by arrows, whereas the 3 catalytic residues are labeled by asterisks. The conserved CKXXNTF motif is highlighted in a box. The first amino acid residue of the mature peptide is underlined. For zebrafish RNases, we used the SignalP 3.0 server to predict the signal peptide cleavage sites. Numbers indicate residue numbers in RNase A. GenBank accession numbers for the 3 zebrafish RNase genes are EF382669–71.

insoluble tRNA was removed by centrifugation. The amount of solubilized tRNA was determined by ultraviolet absorbance at 260 nm. The catalytic activity of the RNase was determined as the pmol of RNA digested per second per pmol of RNase (Zhang et al. 2003). In the experiments, we used 54.2 pmol *Dr*-RNase 1, 30.6 pmol *Dr*-RNase 2, and 53.9 pmol *Dr*-RNase 3. For comparison, we also used 456 pmol of commercially available (Sigma) RNase A and 286 pmol of human RNase 5 provided by Zhang and Rosenberg (2002b). The results shown are normalized for 1 pmol of each RNase. The average values from 3 experiments and their 95% confidence intervals are presented.

Antibacterial Assay

We used pathogenic strains of bacteria *E. coli* (ATCC number 11303), *Pseudomonas aeruginosa* (27853), and *Staphylococcus aureus* (27217) and conducted the bactericidal assay following Zhang et al. (2003). The bacteria were grown overnight and diluted 1:1000 in 10 mM sodium phosphate buffer (pH 7.5). Two microliters (40,000 colony-forming units [CFUs]) of bacteria were incubated with varying concentrations of recombinant RNases for 8 h at 37 °C. Serial dilutions of each protein–bacteria incubation were prepared and plated, and CFUs remaining after each treatment were determined. For the positive control, we mixed kanamycin solution with bacteria at a final concentration of 10 mg/ml. The negative control was identically treated, but with proteins from a sham isolation (with vector only).

Phylogenetic Analyses

Protein sequences were aligned by ClustalX (Thompson et al. 1997) with manual adjustments. MEGA3 (Kumar et al. 2004) was used for evolutionary analyses. Phylogenetic trees were reconstructed using the Neighbor-Joining method (Saitou and Nei 1987) with 2,000 bootstrap replications. We used the complete deletion option and the Poisson-corrected protein distance for tree making.

Results

Three RNase Genes Are Identified from the Zebrafish Genome Sequence

We performed TBlastN and BLAT searches against the current zebrafish genome assembly (Zv6) using all human RNases as queries and identified 3 sequences that are homologous to the query sequences (table 1). All 3 sequences have known features of canonical RNases (fig. 1). First, the entire coding region of 405–447 nt is contained in a single exon. Second, a signal peptide of ~20 amino acids is detected computationally at the N-terminus of the protein, suggesting that it is a secreted protein. Third, each protein has the "CKXXNTF" signature motif and the catalytic triad, as well as conserved cysteine residues. All canonical mammalian RNases except RNase 5 (angiogenin) have 8 cysteines forming 4 disulfide bridges, but mammalian RNase 5 and all nonmammalian RNases identified to date have only 6 cysteines (with the exception of frog RNases that have an additional disulfide bond at another position [Rosenberg et al. 2001]). Consistent with this pattern, all the 3 zebrafish RNases have 6 cysteines as in RNase 5 (fig. 1). We named the sequences *Dr*-RNase 1, *Dr*-RNase 2, and *Dr*-RNase3 according to the order in which they were discovered.

Dr-RNase 1 and *Dr*-RNase 2 are more similar to each other (42% amino acid sequence identity) than either is to *Dr*-RNase 3 (28% and 30%, respectively). They are also identical in protein length (149 amino acids), longer than *Dr*-RNase 3 by 14 amino acids. However, *Dr*-RNase 1 and *Dr*-RNase 2 are located in different chromosomes, whereas *Dr*-RNase 2 and *Dr*-RNase 3 are located in the same chromosome with an interval of just ~35 kb

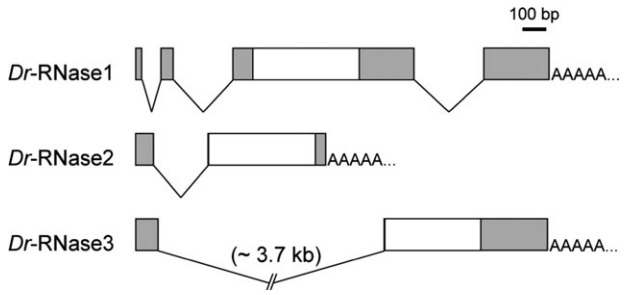

FIG. 2.—Structures of zebrafish RNase genes. Boxes represent exons and lines indicate introns. Open reading frames are shown by open boxes, whereas untranslated regions are shown by gray boxes. The symbol "AAAAA..." indicates the polyadenylated tail of messenger RNA. The intron of *Dr*-RNase 3 is not drawn to scale.

(table 1). These observations suggest that the 1st tandem gene duplication produced *Dr*-RNase 3 and *Dr*-RNase 2 and the subsequent gene duplication generated *Dr*-RNase 1 on a different chromosome.

Dr-RNase 1 Has an Unusual Gene Structure

We determined the entire mRNA sequences of the zebrafish RNase genes by 5′ and 3′ RACE using heart's and liver's total mRNAs, and compared the mRNA sequences with the genomic sequences to determine exon–intron boundaries for each gene. The resulting gene structures are drawn to scale in figure 2. Almost invariably, vertebrate RNase genes have 2 exons: one noncoding exon followed by another exon containing the entire open reading frame. This is the case for *Dr*-RNase 2 and *Dr*-RNase 3, but *Dr*-RNase 1 has 4 exons with its open reading frame residing in the 3rd exon. It is also noteworthy that the single intron of *Dr*-RNase 3 is much longer (~3.7 kb) than the introns in the other 2 genes. We found that a large section of this intron (nucleotides 800–2150) is homologous to mermaid short interspersed repetitive elements. Thus, the insertion of the transposable element is likely the reason why this intron is unusually large.

Zebrafish RNases Are Expressed in Limited Numbers of Adult Tissues

To determine when the zebrafish RNases are expressed during development, we performed RT-PCR analyses using total mRNA samples purified from various developmental stages including fertilized eggs (day 0) and 1- to 5-day-old hatchlings. No strong expression was detected in any of these early stages but a weak expression of *Dr*-RNase 3 was detected in 5-day-old fish (fig. 3). We did the same RT-PCR tests using mRNAs purified from 9 different adult tissues including eye, heart, brain, liver, gut, testis, ovary, skin, and skeletal muscle (fig. 3). Interestingly, *Dr*-RNase 1 and *Dr*-RNase 2 show a similar expression pattern consistent with their sequence similarity: strong expressions in the liver and gut and a weak signal in the heart. In addition, a weak expression of *Dr*-RNase 1 is detected in the testis. Different from the other 2 genes, *Dr*-RNase 3 is expressed in the eye and heart with much lower intensity. From these results, we conclude that zebrafish RNases function in late developmental stages and in adults.

Zebrafish RNases Have Ribonucleolytic Activities

We purified recombinant proteins of zebrafish RNases as described in Materials and Methods and determined their ribonucleolytic activities using the standard assay against yeast tRNA substrates. All the 3 zebrafish RNases have ribonucleolytic activities (fig. 4). Among them, *Dr*-RNase 3 has the highest activity, comparable to that of bovine pancreatic ribonuclease (RNase A). *Dr*-RNase 1 and *Dr*-RNase 2 have lower activities, but greater than that of human RNase 5 (angiogenin).

Zebrafish RNases Are Bactericidal

As mentioned, many members of the RNase A superfamily have antibacterial and/or antiviral activities. We examined whether zebrafish RNases are also bactericidal, using 2 gram-negative and 1 gram-positive pathogenic bacterial strains as targets. At maximum doses, all the 3 zebrafish RNases have potent activities against the gram-negative bacteria *E. coli* and *P. aeruginosa* but have only mild effects against the gram-positive bacteria *S. aureus* (fig. 5*A*). We also confirmed that their antibacterial activities against *E. coli* are dose dependent (fig. 5B). The LD_{50} value, the dose that kills 50% of bacteria, is estimated to be 1.0–1.5 μM for all 3 zebrafish RNases, which is comparable to that of human RNase 7 (Zhang et al. 2003). As expected, no live bacteria were detected in positive controls.

Evolution of Fish RNase Genes

To gain a broader picture of fish RNase evolution, we searched the Atlantic salmon (*S. salar*) and rainbow trout (*O. mykiss*) EST databases and the medaka (*O. latipes*) genome sequence. We identified 2, 2, and 10 RNase sequences from the salmon, trout, and medaka, respectively, and named them *Ss*-RNases 1 and 2 (salmon), *Om*-RNases 1 and 2 (trout), and *Ol*-RNases 1–10 (medaka), according to the order in which they were identified. We also conducted thorough searches in the genome sequences of stickleback (*G. aculeatus*), fugu (*T. rubripes*), and Tetraodon (*T. nigroviridis*) but found no RNase sequences. Because these genome sequences are of high coverage (see Materials and Methods), we believe that the negative result reflects a lack of RNase genes in these 3 fishes.

To determine the phylogenetic positions of fish RNases in the vertebrate RNase superfamily, we constructed a protein Neighbor-Joining tree of fish RNases, previously identified nonmammalian RNases, and all human RNases (fig. 6). The tree leads to several inferences on the evolution of fish RNases and that of the superfamily. First, the RNases of each fish species do not cluster into a species-specific clade, except for the 2 trout RNases that are closely related to each other. For example, the 3 zebrafish RNases do not form

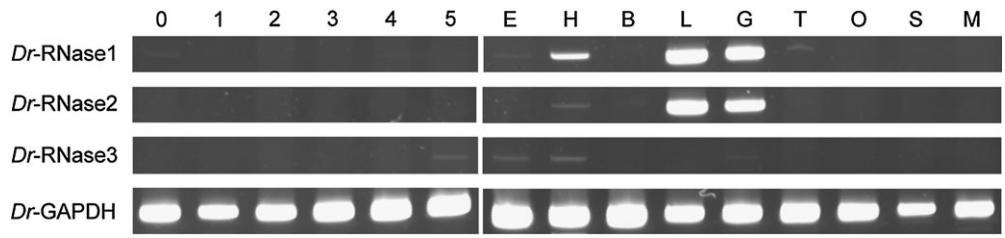

FIG. 3.—Expression patterns of zebrafish RNase genes measured by RT-PCR. cDNAs were made and used for PCR from total RNA samples obtained from fertilized eggs (lane 0) and 1–5 days whole-body hatchlings (lane 1–5). To examine adult expression, cDNAs were purified from 9 different organs: eye (lane E), heart (H), brain (B), liver (L), gut (G), testis (T), ovary (O), skin (S), and skeletal muscle (M). The expression of the housekeeping glyceraldehyde 3-phosphate dehydrogenase gene (*Dr*-GAPDH) was used as a positive control.

1 clade and the 10 medaka RNases form 2 separate clades. These results indicate that the diversification of fish RNases is ancient and that there were more than one RNase gene in the last common ancestor of all the fish species studied here, which lived around 150 MYA (Benton and Donoghue 2007). Later, each fish lineage underwent differential retention and expansion by gene sorting (Zhang et al. 2000) leading to the varied RNase repertoires as seen in figure 6. Second, although all the fishes studied here belong to a monophyletic group (Teleostei or teleost fishes), their RNases do not form a monophyletic group in exclusion of all other vertebrate RNases. This pattern suggests that there were at least 2 RNase genes in the common ancestor of teleosts and tetrapods (amphibians, reptiles, birds, and mammals). Third, the noncanonical RNases (human RNases 9–13 in the tree) form a monophyletic group. If we consider the many differences between canonical and noncanonical RNases and root the canonical RNase part of the tree with noncanonical RNases, fish and frog RNases appear to be the basal canonical RNases. Interestingly, this rooting also suggests that the initial diversification among human RNases 1–8 occurred between RNase 5 and all other RNases in the common ancestor of mammals and birds. The tree showed that the 3 chicken RNases are orthologous to mammalian RNase 5, consistent with the findings of similar functions between some chicken RNases and mammalian RNase 5 in terms of angiogenic and bactericidal activities (Nitto et al. 2006). It should be noted that because vertebrate RNases are quite divergent in sequence but are short in length, many bootstrap values in the tree of figure 6 are not high, particularly for some deep nodes. This makes it difficult to infer RNase evolution with a high certainty.

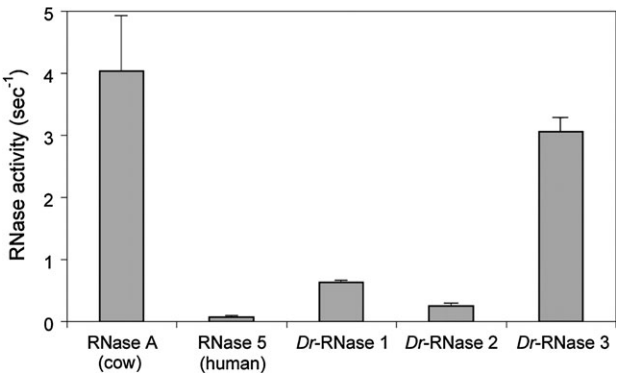

FIG. 4.—Ribonucleolytic activities of zebrafish RNases. Bovine pancreatic ribonuclease (RNase A) and human RNase 5 (*Hs*-ANG) are used for comparison. The unit of the ribonucleolytic activity is pmol of substrate digested per pmol of enzyme per second. Shown here are the average activities from 3 measurements and their 95% confidence intervals.

Discussion

We previously reported our discovery of RNase genes in the then incomplete zebrafish genome sequence (Cho et al. 2005). Here, we report the identification and functional characterization of 3 RNase genes in the now much improved zebrafish genome sequence. The zebrafish genome assembly (Zv6) used here has a high coverage (6.5× to 7×). It is thus likely that we have identified most, if not all, zebrafish RNase genes. We found that all 3 zebrafish RNases have potent bactericidal activities against gram-negative *E. coli* and *P. aeruginosa*, whereas this activity against gram-positive *S. aureus* is much weaker. These results mark the first discovery of bactericidal activities of fish RNases. As mentioned above, multiple divergent canonical RNases of mammals are known to have antibacterial and/or antiviral activities (Dyer and Rosenberg 2006). Because fish are basal vertebrates, our finding of the bactericidal activities in their RNases support the hypothesis that host defense was the initial function of the RNase A superfamily when it originated in vertebrate evolution (Cho et al. 2005).

It has been suggested and experimentally verified in several antibacterial RNases that positively charged amino acid residues are important for the disruption of negatively charged bacterial cell membranes and thus are key to the bactericidal activity (Zhang et al. 1998; Carreras et al. 2003; Zhang et al. 2003; Huang et al. 2006). Indeed, all known bactericidal RNases, including mammalian RNases 3, 5, and 7, as well as chicken RNases, have relatively high isoelectric points (Zhang et al. 2003; Nitto et al. 2006). However, isoelectric points of the zebrafish RNases are not particularly high (table 1), although their total numbers of positively charged amino acids are as high as those of previously known bactericidal RNases (table 2). This suggests that local clusters of positive charges may be more important than global isoelectric points for bactericidal activities. Table 2 shows that some RNases from the salmon, trout, and medaka also have relatively high numbers of

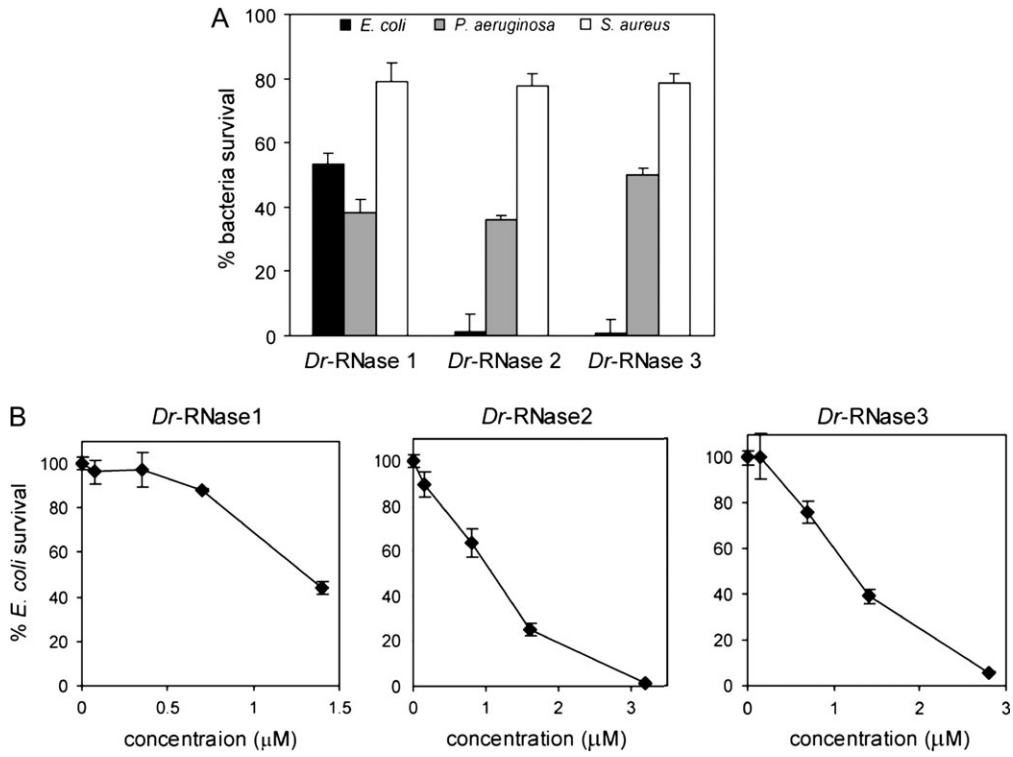

FIG. 5.—Bactericidal activities of zebrafish RNases at (*A*) fixed, maximum doses and at (*B*) variable doses. For (*A*), 1.4 μM (final concentration) of *Dr*-RNase 1, 2.8 μM of *Dr*-RNase 2, and 3.2 μM of *Dr*-RNase 3 were incubated with each bacterial strain and percentage survival compared with negative control (100%, not shown) was obtained for each experiment. The average of 3 measurements, along with its standard error, is shown. For (*B*), various concentrations of zebrafish RNases were made by dilution using negative control at appropriate ratios and were then incubated with *Escherichia coli*. The negative control was made from identically treated sham isolation (with vector only). The average percentage survival value compared with negative control from the 3 experiments is shown with its standard error.

positively charged amino acids, suggesting that these RNases may also be bactericidal.

As mentioned, most nonmammalian RNases have only 6 conserved cysteines, much like mammalian RNase 5 (Strydom 1998). Some of them are also similar to mammalian RNase 5 (Strydom 1998) in having low ribonucleolytic activities (Nitto et al. 2005, 2006). It is thus unsurprising that *Dr*-RNases 1 and 2 have only weak ribonucleolytic activities. However, we show that *Dr*-RNase 3 has a much stronger ribonucleolytic activity, comparable to that of bovine pancreatic RNase (RNase A). This suggests the possibility that a potent ribonucleolytic activity originated quite early in the evolutionary history of canonical RNases. It has been demonstrated previously in human RNase 3 and a chicken RNase that ribonucleolytic activities are not necessary for bactericidal activities (Rosenberg 1995; Nitto et al. 2006). It would be interesting to examine if this is the case in zebrafish RNases.

Our RT-PCR analyses show that none of the 3 zebrafish RNase genes is expressed during early embryonic development. This is in line with the fact that no RNases studied to date is involved in early development or housekeeping functions. It is noteworthy that *Dr*-RNases 1 and 2, which are more similar to each other than either is to *Dr*-RNase 3 in protein sequence, are also similar to each other in their expression pattern, which is different from that of *Dr*-RNase 3. The strong expression of *Dr*-RNases 1 and 2 in the liver and gut is similar to some other RNases such as human RNase 1, 2, 4, 5, and 7 (Futami et al. 1997; Zhang et al. 2003). All but a few of over 100 blood plasma proteins are made in liver, and a great diversity of these plasma proteins serve as defense molecules in the innate immune system (Putnam 1985). Furthermore, it has been reported that mouse Ang4, one of the several duplicates of mouse RNase 5, is produced by Paneth cells in the intestine and secreted out to the gut lumen and has bactericidal activities against intestinal microbes (Hooper et al. 2003). Thus, it is not unreasonable that *Dr*-RNases 1 and 2 produced in the liver and gut participate in innate immunity in a similar way. Alternatively, it is possible that *Dr*-RNases 1 and 2, expressed in the gut, might be involved in regulating the microbiota of the gut rumen by selectively suppressing certain phylotypes of microorganisms. For example, a recent study of reciprocal gut microbiota transplants from zebrafish and mice to germ-free recipients revealed host habitat selection (Rawls et al. 2006). It should be noted that for the preparation of the gut sample, we had to take the intestinal tissue along with other organs attached to it, such as pancreas, due to technical limitations. Thus, it is possible that *Dr*-RNase 1 and 2 are expressed in those tissues rather than or as well as the gut, which would not be surprising because RNase A, the prototype of the RNase superfamily, is expressed in the pancreas. Compared with *Dr*-RNase 1 and 2, we could detect only weak expressions of *Dr*-RNase 3 in the whole 5-day-old fish and in the adult eye and heart. It is, however, possible that this gene is highly expressed under specific

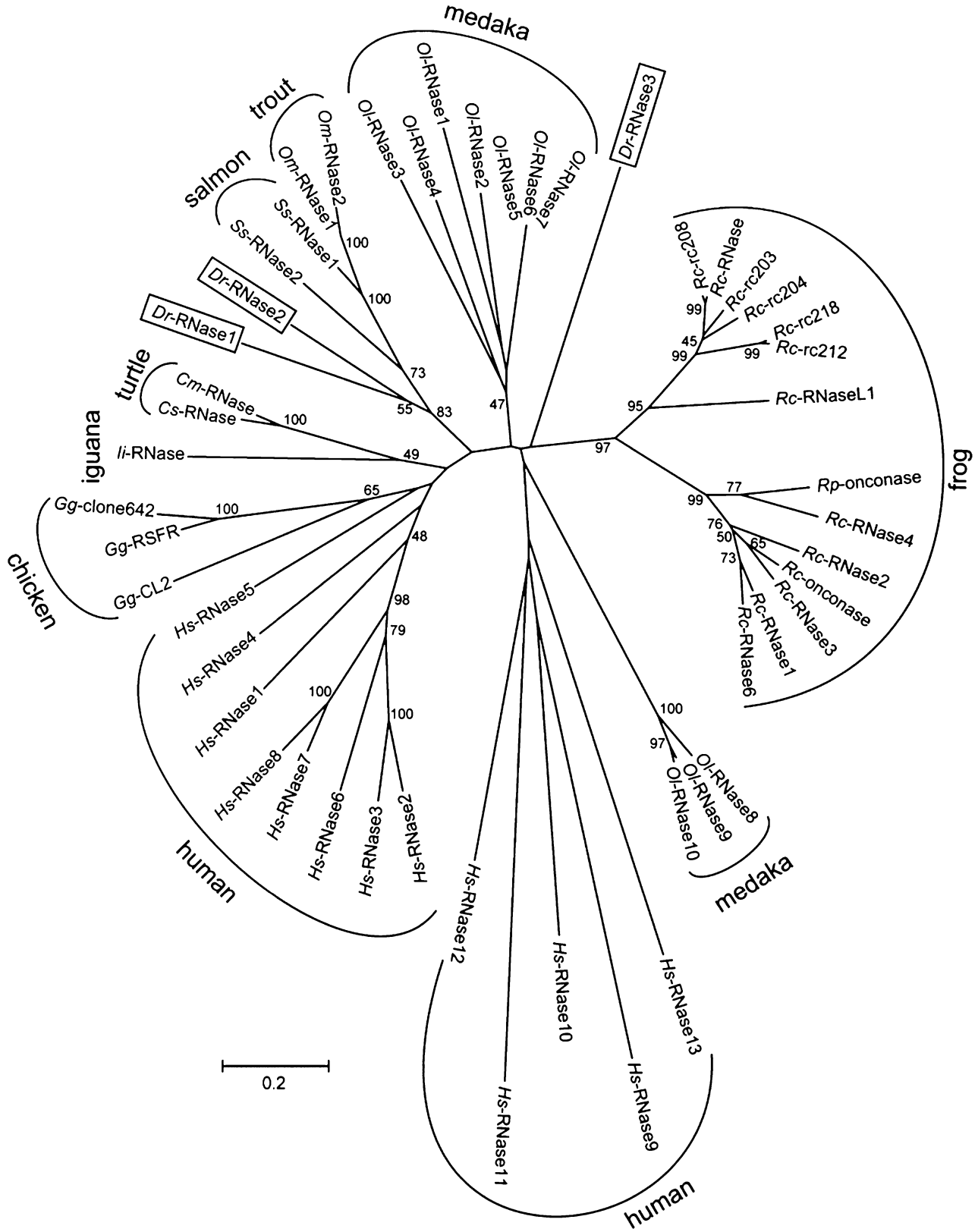

FIG. 6.—Phylogenetic relationships of fish RNases, other nonmammalian RNases, and human RNases. Three zebrafish RNases are boxed for distinction. A total of 76 amino acid sites were used in tree making after the removal of alignment gaps. The Neighbor-Joining method with the protein Poisson-corrected distance was used. Bootstrap percentages (in 2,000 replications) that are greater than 40 are shown on interior branches. Prefixes are used to indicate organisms: *Dr* for zebrafish (*Danio rerio*), *Ss* for Atlantic salmon (*Salmo salar*), *Om* for rainbow trout (*Oncorhynchus mykiss*), *Ol* for medaka (*Oryzias latipes*), *Hs* for human (*Homo sapiens*), *Gg* for chicken (*Gallus gallus*), *Ii* for common green iguana (*Iguana iguana*), *Cm* for green turtle (*Chelonia mydas*), *Cs* for snapping turtle (*Chelydra serpentine*), and *Rc* and *Rp* for 2 bullfrog species (*Rana catesbeiana* and *Rana pipiens*).

Table 2
Number of Basic Amino Acids in All Fish RNases and the Chicken and Human RNases Known to be Bactericidal

Protein	pI	Number of Lysines	Number of Arginines	Total
Dr-RNase1	8.95	4	15	19
Dr-RNase2	9.18	11	10	21
Dr-RNase3	8.82	15	5	20
Ss-RNase1	9.09	7	10	17
Ss-RNase2	8.24	13	2	15
Om-RNase1	8.95	11	8	19
Om-RNase2	8.31	11	5	16
Ol-RNase1	8.01	7	11	18
Ol-RNase2	7.40	9	7	16
Ol-RNase3	8.45	14	7	21
Ol-RNase4	8.57	6	12	18
Ol-RNase5	7.61	14	4	18
Ol-RNase6	8.17	14	4	18
Ol-RNase7	8.17	14	4	18
Ol-RNase8	8.48	12	5	17
Ol-RNase9	8.27	7	6	13
Ol-RNase10	8.24	7	7	14
Gg-RSFR	10.20	4	14	18
Hs-RNase3	10.72	1	19	20
Hs-RNase5	9.65	7	13	20
Hs-RNase7	9.80	18	4	22

NOTE.—The isoelectric point of the mature peptide of each protein was calculated using the Protein Calculator v3.3 (http://www.scripps.edu/~cdputnam/protcalc.html). Prefixes are used to indicate species: *Dr* for zebrafish (*Danio rerio*), *Ss* for Atlantic salmon (*Salmo salar*), *Om* for rainbow trout (*Oncorhynchus mykiss*), *Ol* for medaka (*Oryzias latipes*), *Gg* for chicken (*Gallus gallus*), and *Hs* for human (*Homo sapiens*).

conditions (e.g., pathogenic infection) or tissues that were not examined in our study.

It has been noted that gene sorting, defined as a process leading to differential retention and amplification of ancestral genes in different lineages, is a common feature in the evolution of the RNase A superfamily (Zhang et al. 2000; Cho et al. 2005; Cho and Zhang 2006) and other host-defense gene families (reviewed in Nei and Rooney 2005). Presumably, this evolutionary pattern reflects a high rate of gene turnover that is consistent with the role of the gene family in defending against ever-changing pathogens. It is noteworthy from the tree in figure 6 that the fish RNase genes also show patterns of gene sorting, such as the presence of different gene numbers in different species and the lack of clear orthologs among species. Such evolutionary patterns are consistent with the hypothesis that fish RNases are involved in host defense.

Because teleost fish are the most basal vertebrates in which RNase genes have been identified, a natural question is whether RNase genes can be found in more basal vertebrates such as sharks and lampreys. We could not find any RNase genes in the 1.4× coverage elephant shark (*Callorhinchus milii*) genome sequence (Venkatesh et al. 2006) and the trace files of the incomplete sea lamprey (*Petromyzon marinus*) genome sequence (http://genome.wustl.edu/genome.cgi?GENOME=Petromyzon%20marinus). Furthermore, we were not able to detect RNase genes in the genome sequences of invertebrate deuterostomes, including the amphioxus *Branchiostoma floridae* (v1.0, 8.1× coverage; http://shake.jgi-psf.org/Brafl1/), sea squirt *Ciona intestinalis* (v2.0, 6× coverage; http://genome.jgi-psf.org/Cioin2/Cioin2.home.html), sea squirt *Ciona savignyi* (13× coverage; http://www.broad.mit.edu/annotation/ciona/index.html) and sea urchin *Strongylocentrotus purpuratus* (Build 2.1, 6× coverage; http://www.ncbi.nlm.nih.gov/genome/guide/sea_urchin/). Because these invertebrates are closely related to vertebrates (Delsuc et al. 2006) and because their genome sequences are of high coverage, our negative results confirm that the RNase A gene superfamily is vertebrate specific. Our searches in the shark and lamprey genomes suggest that the superfamily may in fact be specific to bony vertebrates, although this hypothesis requires further scrutiny when more complete genome sequences from the shark and lamprey become available.

While completing our experiments, we became aware of the recent publication on zebrafish RNases of Pizzo et al. (2006). For easy comparison, we summarize our findings and their findings of zebrafish RNases in table 3. They identified 3 zebrafish RNase genes from EST databases and showed that these RNases have ribonucleolytic and angiogenic activities. Interestingly, only 2 of the 3 RNases that they studied correspond to what we discovered and reported here. Their ZF-RNase3 and ZF-RNase2 are identical to our *Dr*-RNase 1 and *Dr*-RNase 2, respectively. We confirmed that the zebrafish genome assembly (Zv6) does not contain

Table 3
Properties of Zebrafish RNases found in this study and in Pizzo et al. (2006)

	Dr-RNase 1 (Zf-RNase-3[a])	*Dr*-RNase 2 (Zf-RNase-2[a])	*Dr*-RNase 3	Zf-RNase-1[a]	Sources
Gene structure	Extra exons	Standard	Standard	Unknown	This study
Expression pattern	Liver, gut, heart, testis	Liver, gut, heart	Eye, heart, 5-day hatchling	Unknown	This study
RNase activity					
Simple assay (sec^{-1})	0.637 ± 0.015	0.254 ± 0.024	3.061 ± 0.112	Unknown	This study
k_{cat}/K_m (M^{-1} sec^{-1})	$6.0 \pm 0.74 \times 10^3$	$6.3 \pm 0.50 \times 10^2$	Unknown	$2.3 \pm 0.25 \times 10^2$	Pizzo et al. (2006)
Bactericidal activities					
Against gram-positive bacteria	Mild	Mild	Mild	Unknown	This study
Against gram-negative bacteria	Potent	Potent	Potent	Unknown	This study
LD_{50} against *E. coli* (μM)	1.3	1.1	1.1	Unknown	This study
Angiogenic activity	No	Yes	Unknown	Yes	Pizzo et al. (2006)

[a] Gene names used by Pizzo et al. (2006).

their ZF-RNase1. It is probable that ZF-RNase1 actually exists in the genome, but is not covered in the genome sequence assembly. Thus, zebrafish most likely has 4 RNase genes.

Supplementary Material

Supplementary table 1 is available at *Molecular Biology and Evolution* online (http://www.mbe.oxfordjournals.org/).

Acknowledgments

We thank Eric Horstick and Shawn Sprague (University of Michigan) and Hiromi Hirata (Nagoya University) for generously providing the zebrafish samples and Jaap Beintema for valuable comments. This work was supported by National Institutes of Health grant GM67030 to J.Z.

Literature Cited

Anfinsen CB. 1973. Principles that govern the folding of protein chains. Science. 181:223–230.

Ardelt W, Mikulski SM, Shogen K. 1991. Amino acid sequence of an anti-tumor protein from Rana pipiens oocytes and early embryos. Homology to pancreatic ribonucleases. J Biol Chem. 266:245–251.

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

Beintema JJ, Broos J, Meulenberg J, Schuller C. 1985. The amino acid sequence of snapping turtle (Chelydra serpentina) ribonuclease. Eur J Biochem. 153:305–312.

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.

Bendtsen JD, Nielsen H, von Heijne G, Brunak S. 2004. Improved prediction of signal peptides: signalP 3.0. J Mol Biol. 340:783–795.

Benton MJ, Donoghue PC. 2007. Paleontological evidence to date the tree of life. Mol Biol Evol. 24:26–53.

Carreras E, Boix E, Rosenberg HF, Cuchillo CM, Nogues MV. 2003. Both aromatic and cationic residues contribute to the membrane-lytic and bactericidal activity of eosinophil cationic protein. Biochemistry. 42:6636–6644.

Castella S, Benedetti H, de Llorens R, Dacheux JL, Dacheux F. 2004. Train A, an RNase A-like protein without RNase activity, is secreted and reabsorbed by the same epididymal cells under testicular control. Biol Reprod. 71:1677–1687.

Castella S, Fouchecourt S, Teixeira-Gomes AP, Vinh J, Belghazi M, Dacheux F, Dacheux JL. 2004. Identification of a member of a new RNase A family specifically secreted by epididymal caput epithelium. Biol Reprod. 70:319–328.

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.

Cho S, Zhang J. 2006. Ancient expansion of the ribonuclease A superfamily revealed by genomic analysis of placental and marsupial mammals. Gene. 373:116–125.

Delsuc F, Brinkmann H, Chourrout D, Philippe H. 2006. Tunicates and not cephalochordates are the closest living relatives of vertebrates. Nature. 439:965–968.

Devor EJ, Moffat-Wilson KA, Galbraith JJ. 2004. LOC 390443 (RNase 9) on chromosome 14q11.2 is related to the RNase A superfamily and contains a unique amino-terminal preprotein-like sequence. Hum Biol. 76:921–935.

Domachowske JB, Dyer KD, Bonville CA, Rosenberg HF. 1998. Recombinant human eosinophil-derived neurotoxin/RNase 2 functions as an effective antiviral agent against respiratory syncytial virus. J Infect Dis. 177:1458–1464.

Dyer KD, Rosenberg HF. 2006. The RNase A superfamily: generation of diversity and innate host defense. Mol Divers. 10:585–597.

Fett JW, Strydom DJ, Lobb RR, Alderman EM, Bethune JL, Riordan JF, Vallee BL. 1985. Isolation and characterization of angiogenin, an angiogenic protein from human carcinoma cells. Biochemistry. 24:5480–5486.

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 in humans. DNA Cell Biol. 16:413–419.

Greenway MJ, Andersen PM, Russ C, et al. (15 co-authors) 2006. ANG mutations segregate with familial and 'sporadic' amyotrophic lateral sclerosis. Nat Genet. 38:411–413.

Harder J, Schroder JM. 2002. RNase 7, a novel innate immune defense antimicrobial protein of healthy human skin. J Biol Chem. 277:46779–46784.

Hooper LV, Stappenbeck TS, Hong CV, Gordon JI. 2003. Angiogenins: a new class of microbicidal proteins involved in innate immunity. Nat Immunol. 4:269–273.

Huang YC, Lin YM, Chang TW, Wu SJ, Lee YS, Chang DT, Chen C, Wu SH, Liao YD. 2006. The flexible and clustered lysine residues of human ribonuclease 7 are critical for membrane permeability and antimicrobial activity. J Biol Chem. 282:4626–4633.

Huang HC, Wang SC, Leu YJ, Lu SC, Liao YD. 1998. The Rana catesbeiana rcr gene encoding a cytotoxic ribonuclease. Tissue distribution, cloning, purification, cytotoxicity, and active residues for RNase activity. J Biol Chem. 273:6395–6401.

Irie M, Nitta K, Nonaka T. 1998. Biochemistry of frog ribonucleases. Cell Mol Life Sci. 54:775–784.

Jermann TM, Opitz JG, Stackhouse J, Benner SA. 1995. Reconstructing the evolutionary history of the artiodactyl ribonuclease superfamily. Nature. 374:57–59.

Katekaew S, Torikata T, Araki T. 2006. The complete amino acid sequence of green turtle (Chelonia mydas) egg white ribonuclease. Protein J. 25:316–327.

Klefhaber T, Labhardt AM, Baldwin RL. 1995. Direct NMR evidence for an intermediate preceding the rate-limiting step in the unfolding of ribonuclease A. Nature. 375:513–515.

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 Bioinform. 5:150–163.

Lander ES, Linton LM, Birren B, et al. (256 co-authors) 2001. Initial sequencing and analysis of the human genome. Nature. 409:860–921.

Lehrer RI, Szklarek D, Barton A, Ganz T, Hamann KJ, Gleich GJ. 1989. Antibacterial properties of eosinophil major basic protein and eosinophil cationic protein. J Immunol. 142:4428–4434.

Nei M, Rooney AP. 2005. Concerted and birth-and-death evolution of multigene families. Annu Rev Genet. 39:121–152.

Nitto T, Dyer KD, Czapiga M, Rosenberg HF. 2006. Evolution and function of leukocyte RNase A ribonucleases of the avian species, Gallus gallus. J Biol Chem. 281:25622–25634.

Nitto T, Lin C, Dyer KD, Wagner RA, Rosenberg HF. 2005. Characterization of a ribonuclease gene and encoded protein from the reptile, Iguana iguana. Gene. 352:36–44.
Penttinen J, Pujianto DA, Sipila P, Huhtaniemi I, Poutanen M. 2003. Discovery in silico and characterization in vitro of novel genes exclusively expressed in the mouse epididymis. Mol Endocrinol. 17:2138–2151.
Pizzo E, Buonanno P, Di Maro A, Ponticelli S, De Falco S, Quarto N, Cubellis MV, D'Alessio G. 2006. Ribonucleases and angiogenins from fish. J Biol Chem. 281:27454–27460.
Putnam FW. 1985. structure, function and genetic control. The plasma proteins. London: Academic Press.
Rawls JF, Mahowald MA, Ley RE, Gordon JI. 2006. Reciprocal gut microbiota transplants from zebrafish and mice to germ-free recipients reveal host habitat selection. Cell. 127:423–433.
Rosenberg HF. 1995. Recombinant human eosinophil cationic protein. Ribonuclease activity is not essential for cytotoxicity. J Biol Chem. 270:7876–7881.
Rosenberg HF, Dyer KD, Tiffany HL, Gonzalez M. 1995. Rapid evolution of a unique family of primate ribonuclease genes. Nat Genet. 10:219–223.
Rosenberg HF, Zhang J, Liao YD, Dyer KD. 2001. Rapid diversification of RNase A superfamily ribonucleases from the bullfrog, Rana catesbeiana. J Mol Evol. 53:31–38.
Rudolph B, Podschun R, Sahly H, Schubert S, Schroder JM, Harder J. 2006. Identification of RNase 8 as a novel human antimicrobial protein. Antimicrob Agents Chemother. 50: 3194–3196.
Saitou N, Nei M. 1987. The neighbor-joining method: a new method for reconstructing phylogenetic trees. Mol Biol Evol. 4:406–425.
Sambashivan S, Liu Y, Sawaya MR, Gingery M, Eisenberg D. 2005. Amyloid-like fibrils of ribonuclease A with three-dimensional domain-swapped and native-like structure. Nature. 437:266–269.
Strydom DJ. 1998. The angiogenins. Cell Mol Life Sci. 54: 811–824.
Thompson JD, Gibson TJ, Plewniak F, Jeanmougin F, Higgins DG. 1997. The CLUSTAL_X windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. Nucleic Acids Res. 25:4876–4882.
Udgaonkar JB, Baldwin RL. 1988. NMR evidence for an early framework intermediate on the folding pathway of ribonuclease A. Nature. 335:694–699.
Venkatesh B, Kirkness EF, Loh YH, et al. (12 co-authors) 2006. Ancient noncoding elements conserved in the human genome. Science. 314–1892.
Yu L, Zhang YP. 2006. The unusual adaptive expansion of pancreatic ribonuclease gene in carnivora. Mol Biol Evol. 23:2326–2335.
Zhang J. 2006. Parallel adaptive origins of digestive RNases in Asian and African leaf monkeys. Nat Genet. 38:819–823.
Zhang J. 2007. Disulfide-bond reshuffling in the evolution of an ape placental ribonuclease. Mol Biol Evol. 24:505–512.
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.
Zhang J, Dyer KD, Rosenberg HF. 2002. RNase 8, a novel RNase A superfamily ribonuclease expressed uniquely in placenta. Nucleic Acids Res. 30:1169–1175.
Zhang J, Dyer KD, Rosenberg HF. 2003. Human RNase 7: a new cationic ribonuclease of the RNase A superfamily. Nucleic Acids Res. 31:602–607.
Zhang J, Rosenberg HF. 2002a. 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. 2002b. Diversifying selection of the tumor-growth promoter angiogenin in primate evolution. Mol Biol Evol. 19:438–445.
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.
Zhao W, Beintema JJ, Hofsteenge J. 1994. The amino acid sequence of iguana (Iguana iguana) pancreatic ribonuclease. Eur J Biochem. 219:641–646.
Zhu CF, Liu Q, Zhang L, Yuan HX, Zhen W, Zhang JS, Chen ZJ, Hall SH, French FS, Zhang YL. 2007. RNase9, an androgen-dependent member of the RNase A family, is specifically expressed in the rat epididymis. Biol Reprod. 76:63–73.

David Irwin, Associated Editor

Accepted March 4, 2007