Adaptive evolution of a duplicated pancreatic ribonuclease gene in a leaf-eating monkey

Jianzhi Zhang[1,2], Ya-ping Zhang[3] & Helene F. Rosenberg[1]

Published online: 4 March 2002, DOI: 10.1038/ng852

© 2002 Nature Publishing Group http://genetics.nature.com

Although the complete genome sequences of over 50 representative species have revealed the many duplicated genes in all three domains of life[1–4], the roles of gene duplication in organismal adaptation and biodiversity are poorly understood. In addition, the evolutionary forces behind the functional divergence of duplicated genes are often unknown, leading to disagreement on the relative importance of positive Darwinian selection versus relaxation of functional constraints in this process[5–10]. The methodology of earlier studies relied largely on DNA sequence analysis but lacked functional assays of duplicated genes, frequently generating contentious results[11,12]. Here we use both computational and experimental approaches to address these questions in a study of the pancreatic ribonuclease gene (*RNASE1*) and its duplicate gene (*RNASE1B*) in a leaf-eating colobine monkey, douc langur. We show that RNASE1B has evolved rapidly under positive selection for enhanced ribonucleolytic activity in an altered microenvironment, a response to increased demands for the enzyme for digesting bacterial RNA. At the same time, the ability to degrade double-stranded RNA, a non-digestive activity characteristic of primate RNASE1, has been lost in RNASE1B, indicating functional specialization and relaxation of purifying selection. Our findings demonstrate the contribution of gene duplication to organismal adaptation and show the power of combining sequence analysis and functional assays in delineating the molecular basis of adaptive evolution.

A subfamily of Old World monkeys, colobines are unique primates that use leaves rather than fruits and insects as their primary food source; these leaves are then fermented by symbiotic bacteria in the foregut[13]. Similar to ruminants, colobines recover nutrients by breaking and digesting the bacteria with various enzymes, including pancreatic ribonuclease (RNASE1), which is secreted from the pancreas and transported into the small intestine to degrade RNA[14,15]. Earlier studies revealed a substantially greater amount of ribonuclease (RNase) in the pancreas of foregut fermenting mammals (colobines and ruminants) than in other mammals[14,15]. This is believed to be related to the fact that rapidly growing bacteria have the highest ratio of RNA-nitrogen to total nitrogen of all cells, and high concentrations of RNase are needed to break down bacterial RNA so that nitrogen can be recycled efficiently[14].

Using a screening method based on PCR and sequencing, we detected one *RNASE1* gene in each of the 15 non-colobine primates examined, including 5 hominoids, 5 Old World monkeys, 4 New World monkeys and 1 prosimian. We determined the DNA sequences of these *RNASE1* genes; the deduced protein sequences are shown in Fig. 1*a*. The phylogenetic tree of the *RNASE1* sequences (Fig. 2*a*) is consistent with the known species relationships[16] at all nodes, with greater than 55% bootstrap support, suggesting that the *RNASE1* genes are orthologous. By contrast, two *RNASE1* genes were found in the Asian colobine, douc langur (*Pygathrix nemaeus*). Phylogenetic analysis (Fig. 2*a*) suggests that these two genes were generated by recent duplication postdating the separation of colobines from other Old World monkeys (cercopithecines). The branch lengths of the gene tree indicate that the nucleotide sequence of one daughter gene (*RNASE1*) has not changed since duplication, whereas that of the other gene (*RNASE1B*) has accumulated many substitutions (Fig. 2*a*). Beintema[15] previously purified an RNase from the pancreas of another Asian colobine, hanuman langur (*Presbytis entellus*), and obtained the mature peptide sequence for this protein. Our phylogenetic analysis of these protein sequences shows that the hanuman langur pancreatic RNase clusters with douc langur RNASE1B with 99% bootstrap support (Fig. 2*b*). This result implies an orthologous relationship between these two proteins, which suggests that the douc langur *RNASE1B* is also expressed in the pancreas.

We determined the structures of *RNASE1* of human, rhesus monkey and douc langur and that of douc langur *RNASE1B* by sequencing genomic regions flanking the coding sequences; we found no variation in gene structure (Fig. 1*b*). The entire RNASE1 or RNASE1B protein is encoded by exon 2, which is separated from an upstream noncoding exon by an intron of 703–706 nt. The presence of a homologous intron (98.9% sequence identity) in *RNASE1* and *RNASE1B* suggests that gene duplication was probably due to unequal crossing-over rather than to retroposition, which usually generates intronless duplicates.

To trace the evolutionary history of *RNASE1B*, we inferred the gene sequence of the most recent common ancestor of douc langur *RNASE1* and *RNASE1B*. As the sequences involved are closely related, the parsimony ancestral inference[17] was unambiguous at all sites and the distance-based Bayesian method[18] gave the same inference with nearly 100% probability, indicating high reliability of the ancestral inference. The coding region of the inferred ancestral sequence is identical to that of present-day *RNASE1* of douc langur, in agreement with the zero branch length of the douc langur *RNASE1* lineage (Fig. 2*a*). Thus, the 12 nucleotide differences between the coding regions of douc langur *RNASE1* and *RNASE1B* all occurred in the *RNASE1B* lineage (Fig. 3). We

[1]*Laboratory of Host Defenses, National Institute of Allergy and Infectious Diseases, National Institutes of Health, Bethesda, Maryland, USA.* [2]*Departments of Ecology and Evolutionary Biology and Molecular, Cellular and Developmental Biology, University of Michigan, 3003 Natural Sciences Building, 830 North University Avenue, Ann Arbor, Michigan 48109, USA.* [3]*Kunming Institute of Zoology, Chinese Academy of Sciences, Kunming, Yunnan, China. Correspondence should be addressed to J.Z. (e-mail: jianzhi@umich.edu).*

tested the molecular-clock hypothesis (that is, equal rates of nucleotide substitution) for the two genes of douc langur using rhesus monkey *RNASE1* as an outgroup (Fig. 3); this molecular-clock hypothesis is rejected ($P<0.001$)[19]. When we divided the nucleotide substitutions into synonymous and nonsynonymous (amino acid–altering) substitutions (Fig. 3), we found that the synonymous substitutions passed the clock test ($P>0.10$), whereas the nonsynonymous substitutions did not ($P<0.005$). This suggests that the rate difference between the two genes is due to a difference in natural selection rather than in mutation rate. Consistent with this result, the clock hypothesis cannot be rejected for the noncoding region of roughly 1,500 nt ($P>0.1$), which is presumably free from selection (see below). In addition, the molecular-clock hypothesis for the noncoding region cannot be rejected between rhesus monkey *RNASE1* and douc langur *RNASE1* (or *RNASE1B*) when human *RNASE1* is used as an outgroup ($P>0.2$). These results allowed us to use the noncoding regions to date the gene duplication event. Using the fossil record of a divergence time of 15 million years (Myr) between colobines and cercopithecines[20], we estimated that the duplication of *RNASE1* to *RNASE1B* occurred 4.2 Myr ago, with a 95% bootstrap confidence interval of 2.4–6.4 Myr ago.

To explore the evolutionary forces driving the accelerated evolution of *RNASE1B*, we compared the number of nucleotide substitutions per site at nonsynonymous sites in *RNASE1B* since its origin through gene duplication, and the corresponding number at synonymous and noncoding sites. We found that the number of substitutions per nonsynonymous site (0.0310) is significantly greater than that per synonymous and noncoding sites (0.0077; $P<0.002$, Fisher's exact test). Synonymous and noncoding sites are generally not considered to be subject to purifying selection. In the present case, the percent nucleotide difference between humans and Old World monkeys at synonymous and noncoding sites of the *RNASE1* (or *RNASE1B*) locus is 6.45 ± 0.61, which is similar to the reported average percentage difference (7.1)[21] between orthologous sequences of humans and Old World monkeys at various pseudogenes and introns ($P>0.20$, *t*-test). Taken together, these analyses suggest that the synonymous and noncoding sites at the *RNASE1B* locus are not subject to selective constraints and that the accelerated evolution of the coding sequence of *RNASE1B* is due to positive Darwinian selection. To investigate the nature of the amino-acid substitutions favored by selection, we divided nonsynonymous substitutions into two groups: those altering the amino-acid charge (radical substitutions) and those

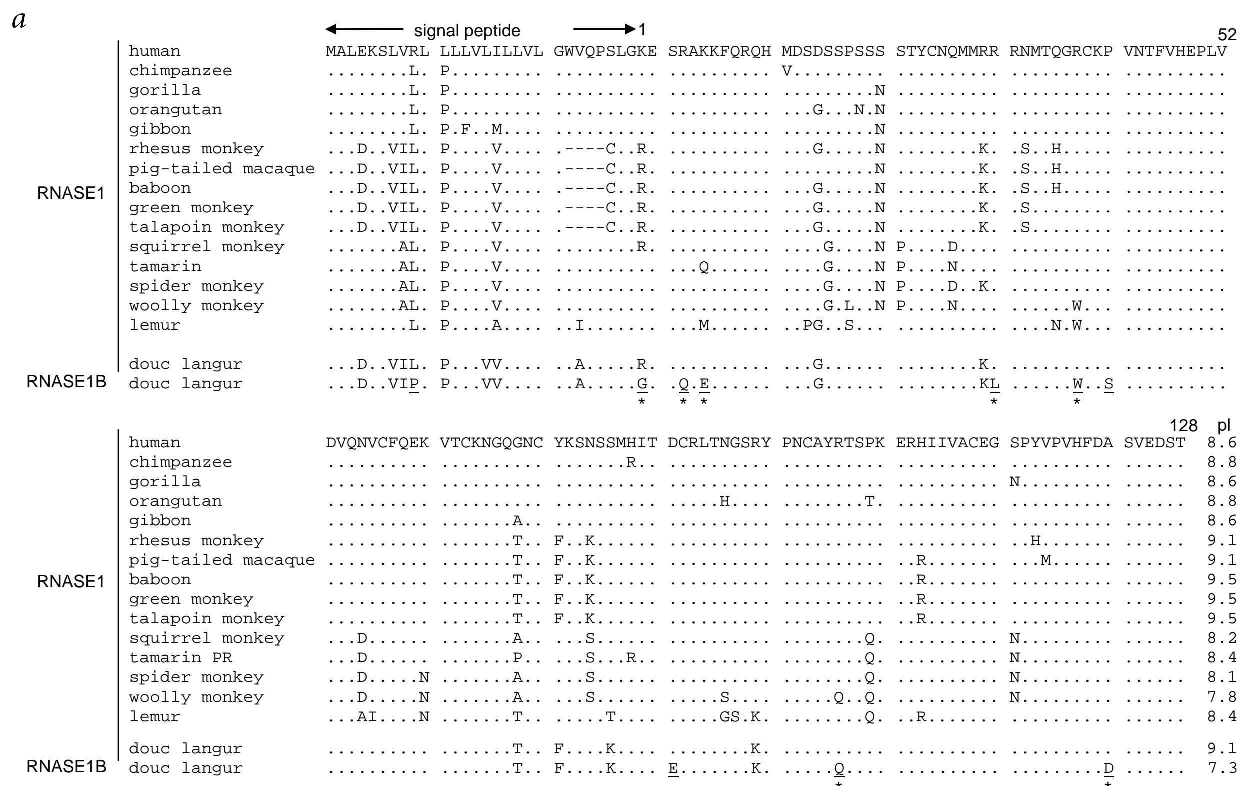


Fig. 1 Protein sequences and genomic structures of *RNASE1* and *RNASE1B* of primates. ***a***, Protein sequence alignment of RNASE1 and RNASE1B. Amino acid substitutions that occurred in RNASE1B since its origin by duplication are underlined, with those involving changes in charge indicated by an asterisk. pI, isoelectric point of mature peptides. ***b***, The conserved structure of *RNASE1* and *RNASE1B*. The structures of douc langur *RNASE1* and *RNASE1B* were determined by homology to that of human *RNASE1*, which was determined by comparing the cDNA and genomic sequences. Compared with douc langur *RNASE1*, there is a 1-nt insertion in the intron of *RNASE1B*. We found no other insertions or deletions between them in the sequenced regions shown here, although there are a total of 28 nucleotide substitutions.

b

DNA sequences surveyed

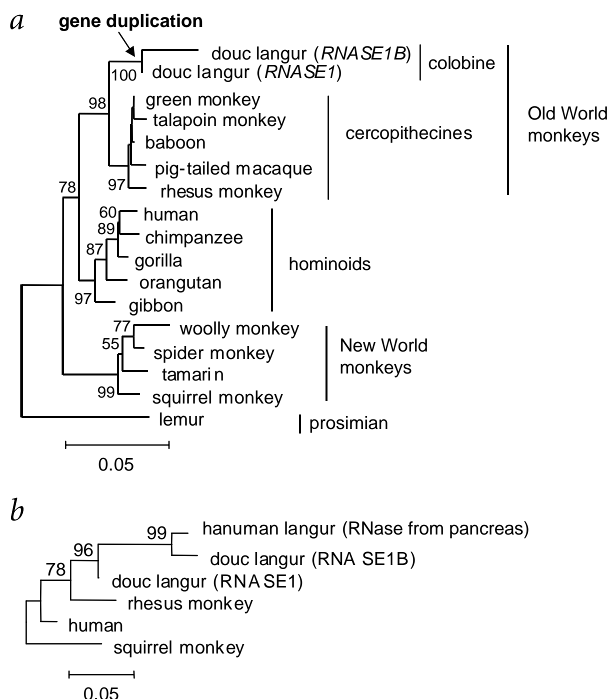
471 bp

protein coding region

		exon 1	intron	exon 2	
human *RNASE1*	408 bp	35 bp	703 bp	634 bp	191 bp
douc langur *RNASE1*	441 bp	35 bp	705 bp	634 bp	191 bp
douc langur *RNASE1B*	441 bp	35 bp	706 bp	634 bp	191 bp

© 2002 Nature Publishing Group http://genetics.nature.com

Fig. 2 Phylogenetic relationships among *RNASE1* and *RNASE1B* of primates. ***a***, The gene tree of *RNASE1* and *RNASE1B*. Kimura's two-parameter distances are used. Virtually identical trees are obtained when Tajima-Nei, Tamura-Nei or Tamura-Nei-γ distances (S. Kumar *et al.*, MEGA2, Arizona State University) are used. The differences only occur at some low-bootstrap (<50%) nodes of the tree shown here. ***b***, Phylogenetic relationship of the purified RNase from hanuman langur and douc langur RNASE1B. Poisson distances of the amino acid sequences of the mature peptides are used. Bootstrap percentages higher than 50 are shown on tree branches. Branch lengths are drawn to scale, indicating the number of nucleotide or amino acid substitutions per site.



that leave charge unaltered (conservative substitutions). Earlier studies showed that, for most mammalian genes, the rate of radical substitution is lower than that of conservative substitution, owing to stronger purifying selection on radical substitution[22]. In *RNASE1B*, however, the opposite is found. The number of radical substitutions per site since duplication (0.067) is significantly greater than that (0.012) of conservative substitutions per site ($P<0.02$; Fisher's exact test). There are nine amino-acid substitutions in the mature peptide of RNASE1B, and seven of them involve charge changes. Unexpectedly, all seven charge-altering substitutions increase the negative charge of the protein (Fig. 1*a*). Apparently, the amino-acid substitutions are nonrandom ($P<0.016$, randomization test), with negatively charged residues being selectively favored. Notably, the rate of radical substitution is not statistically different from the conservative rate when amino acid polarity or volume[22] is considered ($P>0.15$).

The charge-altering substitutions reduced the net charge of RNASE1B from 8.8 to 0.8 (at pH 7) and the isoelectric point from 9.1 to 7.3 (Fig. 1*a*). Because RNA is negatively charged, the net charge of RNase influences its interaction with the substrate and its catalytic performance[23]. We therefore hypothesized that the charge-altering substitutions may have changed the optimal pH of RNASE1B in catalyzing the digestion of RNA. To test this hypothesis, we prepared recombinant proteins from douc langur *RNASE1B* as well as the *RNASE1* genes of human, rhesus monkey and douc langur, and examined their ribonucleolytic activities at different pH levels in a standard RNase assay against yeast tRNA. We determined that the optimal pH for human RNASE1 is 7.4, a value that is within the pH range (7.4–8.0) measured in the small intestine of humans[24,25]. The same optimal pH was observed for RNASE1 of rhesus monkey and douc langur (Fig. 4*a*). Probably because of foregut fermentation and related changes in digestive physiology, the pH in the small intestine of colobine monkeys shifts to 6–7 (ref. 13). Notably, the optimal pH for douc langur RNASE1B was found to be 6.3 (Fig. 4*a*). At pH 6.3, RNASE1B is about six times as active as RNASE1 in digesting RNA, and the difference in their activities is statistically significant ($P<0.001$, *t*-test). These results suggest that the rapid amino acid substitutions in RNASE1B were driven by selection for enhanced RNase activity at the relatively low pH environment of the colobine small intestine.

	nucleotide substitutions coding	(syn, nonsyn)	noncoding	amino-acid changes sig pep	mat pep
douc langur *RNASE1B*	12**	(2, 10**)	11	1	9**
douc langur *RNASE1*	0	(0, 0)	5	0	0
rhesus monkey *RNASE1*	14	(4, 10)	49	2	7

Fig. 3 Tests of the molecular clock hypothesis for *RNASE1* and *RNASE1B* of douc langur. Rhesus monkey *RNASE1* is used as an outgroup. The numbers of substitutions on each of the three branches of the tree are determined by comparing the present-day sequences with the ancestral sequence at the interior node of the tree. Significance level of the Tajima's test: *, 5% ; **, 0.5%. syn, synonymous; nonsyn, nonsynonymous; sig pep, signal peptide; mat pep, mature peptide.

Sequence conservation of douc langur RNASE1 after gene duplication and its unchanged optimal catalytic pH at 7.4 suggest that this protein acts in non-digestive processes. Of note, human RNASE1 is found in many other tissues besides the pancreas[26] and has enzyme activity (EA_{dsRNA}) in degrading double-stranded (ds) RNA, although the physiological relevance of this catalytic activity is unclear[23]. We found similar EA_{dsRNA} among RNASE1 of human, rhesus monkey and douc langur (Fig. 4*b*), with that of douc langur RNASE1B reduced to approximately 0.3% (Fig. 4*b*). As one interpretation, RNASE1B can afford to loose EA_{dsRNA} function because the paralogous RNASE1 retains it; it is likely that some of the adaptive charge-altering substitutions in RNASE1B are detrimental to EA_{dsRNA}. To determine which of the nine amino-acid substitutions in RNASE1B are responsible for loss of EA_{dsRNA}, we used site-directed mutagenesis to create mutant forms of douc langur RNASE1, each with one substitution. We found that eight of the nine substitutions reduce EA_{dsRNA} substantially ($P<0.005$, two-tail *t*-test), whereas the other (R4Q, Fig. 4*b*) has a mild and marginally significant effect ($P=0.069$, two-tail *t*-test and $P=0.035$, one-tail test). The detrimental effect of these substitutions on EA_{dsRNA} might also be predicated from the fact that seven of the nine substitutions that occurred in RNASE1B are not found in any of the 16 primate RNASE1 proteins examined, and that five of the substitutions occurred in positions that are otherwise invariant (Fig. 1*a*). Two of the most influential substitutions are Arg→Leu at position 32 and Asp→Glu at position 83, each reducing EA_{dsRNA} to approximately 3%. Both Arg32 and Asp83 are invariant among primate RNASE1 proteins, suggesting that they are essential for RNASE1 function and that mutations at these sites have been subject to strong purifying selection. It should also be noted that each of the nine RNASE1 single-substitution constructs has a significantly higher EA_{dsRNA} than that of RNASE1B ($P<0.005$), suggesting that it is not a single substitution, but a collective effect of multiple substitutions, that has dramatically reduced the EA_{dsRNA} of RNASE1B. Future analyses of mutant RNASE1 proteins with multiple substitutions may uncover possible interactions among these amino-acid changes.

© 2002 Nature Publishing Group http://genetics.nature.com

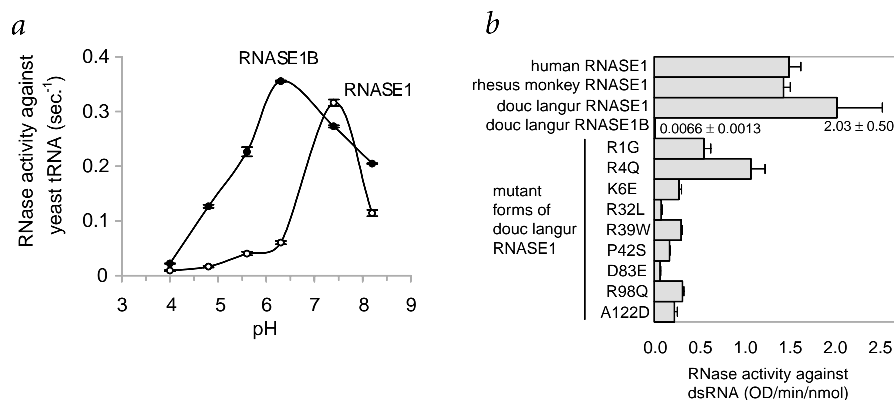


Fig. 4 Enzyme activities of recombinant RNASE1B, RNASE1 and mutant forms of RNASE1. ***a***, RNase activity against yeast tRNA at different pH levels. ***b***, RNase activity against dsRNA. Mutant forms of douc langur RNASE1 are indicated by formula *XyZ*, in which amino acid *X* is replaced by *Z* at position *y* of the mature peptide. Error bars indicate 1 s.e.m.

© 2002 Nature Publishing Group http://genetics.nature.com

Using statistical analysis of nucleotide substitutions and biochemical assays of recombinant proteins, we have described the adaptive evolution of the duplicated douc langur *RNASE1B* in response to increased demands for RNase in an altered microenvironment of the enzyme. The origin and functional changes of RNASE1B probably made the digestive system of these leaf-eating monkeys more efficient. Taken together, our results provide evidence of the important contribution of gene duplication to adaptation of organisms to their environments. It has been debated whether positive selection or relaxation of purifying selection drives functional divergence of duplicated genes[5–10]. In the present case, had the functional constraints for EA_{dsRNA} not been relaxed, the mutations affecting the catalytic optimal pH for RNASE1B could not have been accepted because, as we have shown, they result in a loss of EA_{dsRNA}. On the other hand, without positive selection, it is unlikely that the net charge of RNASE1B would have undergone such a dramatic change in a short period of evolutionary time before the gene was deactivated by random nonsense mutations. Functional relaxation clearly made these otherwise deleterious mutations acceptable, and positive selection further enhanced the fixation probability of the mutations. In short, the two evolutionary forces had complementary roles in the functional divergence of RNASE1B from RNASE1. Our observation that EA_{dsRNA} is retained in RNASE1, with the digestive role transferred to RNASE1B, supports the proposal that gene duplication provides the opportunity for daughter genes to achieve functional specialization[27,28]. Fossil records show the emergence of leaf-eating and foregut fermentation in colobines no later than 10 Myr ago[20], predating the origin of *RNASE1B*. This suggests that changes in diet and digestive physiology in colobines provided the selective forces for the evolution of a more effective digestive RNase, whereas gene duplication provided the raw genetic material. We also note the temporal proximity of the gene duplication and the radiation of Asian colobines about 3.5 Myr ago[20], which, together with the presence of RNASE1B in at least two genera of Asian colobines (*Pygathrix* and *Presbytis*), suggests the possibility of a causal link between these events.

Methods

Isolation of *RNASE1* and *RNASE1B*. We amplified the coding region of *RNASE1* and *RNASE1B* from the genomic DNA of one individual each of human (*Homo sapiens*), chimpanzee (*Pan troglodytes*), gorilla (*Gorilla gorilla*), orangutan (*Pongo pygmaeus*), gibbon (*Hylobates leucogenys*), douc langur (*Pygathrix nemaeus*), rhesus monkey (*Macaca mulatta*), pig-tailed macaque (*Macaca nemestrina*), baboon (*Papio hamadryas*), green monkey (*Cercopithecus aethiops*), talapoin monkey (*Miopithecus talapoin*), squirrel monkey (*Saimiri sciureus*), tamarin (*Saguinus oedipus*), spider monkey (*Ateles geoffroyi*), woolly monkey (*Lagothrix lagotricha*) and lemur (*Lemur catta*), with primers PR5 and PR3. We carried out PCR with high-fidelity Taq, under conditions recommended by the manufacturer (Life Technology), cloned the products into pCR4Blunt-TOPO vector (Invitrogen) and sequenced from both directions using the dideoxy chain termination method with the Perkin-Elmer 377 automatic sequencer. We sequenced several colonies for each species and found no sequence variation within species, except for douc langur, for which we identified two distinct sequences. Although possible, it is unlikely that the two sequences of douc langur are derived from two alleles rather than two genes, because of their unusually high divergence (7.8% at the protein sequence level). If they were allelic sequences, overdominant selection would have to be considered to explain the existence of this trans-specific polymorphism (Fig. 2*b*). In addition, our preliminary study from another Asian colobine (*Presbytis francoisi*) identified at least three distinct RNASE1 sequences in one individual (data not shown), providing definite evidence of *RNASE1* gene duplication in colobines. We also ruled out the possibility that *RNASE1B* exists in non-colobine primates but was not detected because it has not diverged in sequence from *RNASE1* (homoduplication). If *RNASE1-RNASE1B* duplication had occurred before the separation of colobines from other Old World monkeys, the age of *RNASE1B* would be at least 15 Myr (ref. 20), which converts to a nucleotide difference of 4.5% ($2 \times 15 \times 10^6 \times 1.5 \times 10^{-9}$) between the duplicates in noncoding regions, given that the nucleotide mutation rate in higher primates is about 1.5×10^{-9} per site per year[21]. Thus, the expected number of nucleotide differences between *RNASE1* and *RNASE1B* should be 69 ($1{,}533 \times 4.5\%$) in the 1,533 bp of noncoding regions we sequenced, and our experiment would have easily detected two sequences with this level of divergence if they indeed existed in a non-colobine primate such as the rhesus monkey. The noncoding regions of *RNASE1* and *RNASE1B* were amplified with primers 263 and 264 using a Platinum TaqPCRx system (Life Technology) under conditions recommended by the manufacturer, and the products were cloned into pCR4-TOPO of Invitrogen and sequenced. PCR primers are available upon request.

Evolutionary analysis. Phylogenetic trees were reconstructed by the MEGA2 program (S. Kumar *et al.*, Arizona State Univ.) using the neighbor-joining method with 1,000 bootstrap replications. We used PHYLIP v. 3.57c (J. Felsenstein, Univ. of Washington) to confirm the MEGA2 results. Ancestral gene sequences were inferred by the parsimony[17] and distance-based Bayesian methods[18]. The transition/transversion mutational bias[29] was estimated from the noncoding region to be 4.37. We computed the potential numbers of noncoding (I), synonymous (S), nonsynonymous (N), conservative nonsynonymous (C) and radical nonsynonymous (R) sites of a sequence as well as the observed substitutions (i, s, n, c, r), at these sites, between two sequences[9,22,30]. For the common ancestral gene of douc langur *RNASE1* and *RNASE1B*, I=1538, S=144.8, N=323.2, C=166.4 (mature peptide) and R=103.9 (mature peptide), and for the *RNASE1B* lineage since gene duplication, i=11, s=2, n=10, c=2, and r=7. We used Fisher's exact test to compare the rates of substitutions at different types of sites[31]. We tested the molecular clock hypothesis using Tajima's method[19]. The duplication event was dated using the noncoding DNA sequences of douc langur *RNASE1*, *RNASE1B*, and rhesus monkey *RNASE1*, and the bootstrap method was used to obtain the 95% confidence interval of the time estimate. We computed isoelectric points (pI) as well as the net charges of mature peptides with the Wisconsin GCG program.

Recombinant proteins and their enzymatic activities. Human, rhesus monkey, and douc langur *RNASE1* and douc langur *RNASE1B* were subcloned into the bacterial expression vector pFLAG CTS (Kodak) and verified by sequencing. The vector adds the octapeptide DYKDDDDK (FLAG) to the recombinant protein, which facilitates its purification and detection with M2 anti-FLAG monoclonal antibody but does not affect the RNase activity[32]. We used the QuikChange site-directed mutagenesis kit (Stratagene) to mutate douc langur *RNASE1* and confirmed the mutations by sequencing. Recombinant proteins were isolated, purified and quantified as described[32]. The RNase activity of the recombinant proteins in digesting yeast tRNA was measured at different pHs (40 mM sodium acetate buffer with pH=4.0–5.6 and 40 mM sodium phosphate buffer with pH=6.3–8.2) at 25 °C. We added purified RNase (0.1–1.0 pmol) into 0.8 ml of the aforementioned buffer with 1.42 nmol tRNA. The reaction was stopped by 0.5 ml of 20 mM lanthanum nitrate with 3% perchloric acid, and insoluble tRNA was removed by centrifugation. The amount of solubilized RNA was determined from ultraviolet absorbance at 260 nm. We computed the catalytic activity of the RNase as the nmol of RNA digested per second per nmol of RNase[32]. The RNase activity (EA_{dsRNA}) against dsRNA (poly(U)•poly(A) combined from poly(U) and poly(A); Pharmacia) was measured at 25 °C in 1 ml buffer of 0.15 M sodium chloride and 0.015 M sodium citrate (pH 6.3–8.4) with 5 ng substrate and 10–100 pmol RNase, and was determined from ultraviolet absorbance at 260 nm (ref. 33). EA_{dsRNA} for douc langur RNASE1 and RNASE1B were both found to be highest at pH 7; the results at pH 7 are thus reported for all constructs. We carried out at least three replications of experiments at each condition examined and computed the means and their standard errors.

GenBank accession numbers. Human *RNASE1* cDNA, W84323; Human *RNASE1* genomic sequence, AL133371. The DNA sequences reported in this paper have been submitted to GenBank (AF449628–46).

Acknowledgments

We thank K. Dyer for technical assistance and J. Beintema, A. Rooney and three anonymous referees for their comments on early versions of the manuscript. This work was supported in part by a start-up fund and a Rackham grant from the University of Michigan (to J.Z.) and grants from the Natural Science Foundation of China and Chinese Academy of Sciences (to Y.P.Z.).

Received 20 September 2001; accepted 23 January 2002.

1. Tomb, J.F. *et al*. The complete genome sequence of the gastric pathogen *Helicobacter pylori*. *Nature* **388**, 539–547 (1997).
2. Makarova, K.S. *et al*. Comparative genomics of the Archaea (Euryarchaeota): evolution of conserved protein families, the stable core, and the variable shell. *Genome Res.* **9**, 608–628 (1999).
3. Rubin, G.M. *et al.* Comparative genomics of the eukaryotes. *Science* **287**, 2204–2215 (2000).
4. Gaudieri, S., Kulski, J.K., Dawkins, R.L. & Gojobori, T. Different evolutionary histories in two subgenomic regions of the major histocompatibility complex. *Genome Res.* **9**, 541–549 (1999).
5. Ohno, S. Ancient linkage groups and frozen accidents. *Nature* **244**, 259–262 (1973).
6. Goodman, M., Moore, G.W. & Matsuda, G. Darwinian evolution in the genealogy of haemoglobin. *Nature* **253**, 603–608 (1975).
7. Li, W.H. & Gojobori, T. Rapid evolution of goat and sheep globin genes following gene duplication. *Mol. Biol. Evol.* **1**, 94–108 (1983).
8. Kimura, M. *The Neutral Theory of Molecular Evolution.* (Cambridge University Press, Cambridge, 1983).
9. Zhang, J., Rosenberg, H.F. & Nei, M. Positive Darwinian selection after gene duplication in primate ribonuclease genes. *Proc. Natl Acad. Sci. USA* **95**, 3708–3713 (1998).
10. Hughes, A.L. *Adaptive Evolution of Genes and Genomes* (Oxford University Press, New York, 1999).
11. Brookfield, J.F. & Sharp, P.M. Neutralism and selectionism face up to DNA data. *Trends Genet.* **10**, 109–111 (1994).
12. Charlesworth, B., Charlesworth, D., Nurminsky, D.I. & Hartl, D.L. How was the Sdic gene fixed? *Nature* **400**, 519–520 (1999).
13. Kay, R.N.B. & Davies, A.G. in *Colobinem Monkeys: Their Ecology, Behaviour and Evolution* (eds Davies, A.G. & Oates, J.F.) 229–250 (Cambridge University Press, Cambridge, 1994).
14. Barnard, E.A. Biological function of pancreatic ribonuclease. *Nature* **221**, 340–344 (1969).
15. Beintema, J.J. The primary structure of langur (*Presbytis entellus*) pancreatic ribonuclease: adaptive features in digestive enzymes in mammals. *Mol. Biol. Evol.* **7**, 470–477 (1990).
16. Goodman, M. *et al*. Toward a phylogenetic classification of primates based on DNA evidence complemented by fossil evidence. *Mol. Phylogenet. Evol.* **9**, 585–598 (1998).
17. Fitch, W.M. Toward defining the course of evolution: minimum change for a specific tree topology. *Syst. Zool.* **20**, 406–416 (1971).
18. Zhang, J. & Nei, M. Accuracies of ancestral amino acid sequences inferred by parsimony, likelihood, and distance methods. *J. Mol. Evol.* **44**, S139–S146 (1997).
19. Tajima, F. Simple methods for testing the molecular evolutionary clock hypothesis. *Genetics* **135**, 599–607 (1993).
20. Delsen, E. in *Colobine Monkeys: Their Ecology, Behaviour and Evolution* (eds Davies, A.G. & Oates, J.F.) 11–44 (Cambridge University Press, Cambridge, 1994).
21. Li, W.H. *Molecular Evolution* 221–224 (Sinauer Associates, Sunderland, Massachusetts, 1997).
22. Zhang, J. Rates of conservative and radical nonsynonymous nucleotide substitutions in mammalian nuclear genes. *J. Mol. Evol.* **50**, 56–68 (2000).
23. Sorrentino, S. & Libonati, M. Structure-function relationships in human ribonucleases: main distinctive features of the major RNase types. *FEBS Lett.* **404**, 1–5 (1997).
24. Guyton, A.C. & Hall, J.E. *Textbook of Medical Physiology* 9th edn (Saunders, London, 1996).
25. Code, C.F. *Handbook of Physiology* Vol. III (American Physiological Association, Washington DC, 1968).
26. Futami, J. *et al.* Tissue-specific expression of pancreatic-type RNases and RNase inhibitor in humans. *DNA Cell Biol.* **16**, 413–419 (1997).
27. Hughes, A.L. The evolution of functionally novel proteins after gene duplication. *Proc. R. Soc. Lond. B Biol. Sci*. **256**, 119–124 (1994).
28. Lynch, M. & Force, A. The probability of duplicate gene preservation by subfunctionalization. *Genetics* **154**, 459–473 (2000).
29. Kimura, M. A simple method for estimating evolutionary rates of base substitutions through comparative studies of nucleotide sequences. *J. Mol. Evol.* **16**, 111–120 (1980).
30. Rooney, A.P. & Zhang, J. Rapid evolution of a primate sperm protein: relaxation of functional constraint or positive Darwinian selection? *Mol. Biol. Evol.* **16**, 706–710 (1999).
31. Zhang, J., Kumar, S. & Nei, M. Small-sample tests of episodic adaptive evolution: a case study of primate lysozyme genes. *Mol. Biol. Evol.* **14**, 1335–1338 (1997).
32. Rosenberg, H.F. & Dyer, K.D. Eosinophil cationic protein and eosinophil-derived neurotoxin. Evolution of novel function in a primate ribonuclease gene family. *J. Biol. Chem.* **270**, 21539–21544 (1995).
33. Libonati, M. & Floridi, A. Breakdown of double-stranded RNA by bull semen ribonuclease. *Eur. J. Biochem.* **8**, 81–87 (1969).

© 2002 Nature Publishing Group http://genetics.nature.com