nature genetics

Parallel adaptive origins of digestive RNases in Asian and African leaf monkeys

Jianzhi Zhang

Similar morphological or physiological changes occurring in multiple evolutionary lineages are not uncommon. Such parallel changes are believed to be adaptive, because a complex character is unlikely to originate more than once by chance. However, the occurrence of adaptive parallel amino acid substitutions is debated[1–3]. Here I propose four requirements for establishing adaptive parallel evolution at the protein sequence level and use these criteria to demonstrate such a case. I report that the gene encoding pancreatic ribonuclease was duplicated independently in Asian and African leaf-eating monkeys. Statistical analyses of DNA sequences, functional assays of reconstructed ancestral proteins and site-directed mutagenesis show that the new genes acquired enhanced digestive efficiencies through parallel amino acid replacements driven by darwinian selection. They also lost a non-digestive function independently, under a relaxed selective constraint. These results demonstrate that despite the overall stochasticity, even molecular evolution has a certain degree of repeatability and predictability under the pressures of natural selection.

I propose that four requirements should be fulfilled to demonstrate adaptive parallel evolution at the protein sequence level. First, similar changes in protein function occur in independent evolutionary lineages. Second, parallel amino acid substitutions are observed in these proteins. Third, the parallel substitutions are not attributable to chance alone and therefore must have been driven by a common selective pressure. Fourth, the parallel substitutions are responsible for the parallel functional changes. Although more than a dozen cases of molecular parallel evolution have been reported[1,4–7], almost all of them satisfy only the first and/or second requirements. Below I describe the parallel evolution of colobine pancreatic ribonucleases (RNases), which satisfies all four criteria.

Colobines are a subfamily of Old World monkeys and are separated into Asian and African clades, which diverged from each other ~13 million years (Myr) ago[8]. Colobines are different from most other primates in that their primary diet consists of leaves rather than insects or fruits[9]. Because mammals do not possess enzymes to degrade cellulose, a main component of leaves, colobines adopt a ruminant-like alimentary system[9]. That is, they use symbiotic bacteria in their foregut to ferment leaves and then recover nutrients by digesting the bacteria[9]. Because rapidly growing bacteria have the highest ratio of RNA nitrogen to total nitrogen of all cells, a higher-than-usual amount of RNase is required in foregut-fermenters to quickly degrade RNA so that the nitrogen can be recycled efficiently[10]. Indeed, ruminants and colobines have considerably higher concentrations of pancreatic RNase than other mammals do[10,11]. In addition, it is known that the pH in the small intestine, where this enzyme acts, is significantly lower in colobines (6–7) than in humans (7.4–8), probably owing to the large amounts of fatty acids produced during fermentation[9,12]. It is thus interesting to examine how the gene encoding pancreatic RNase has responded to various selective pressures during colobine evolution.

Pancreatic RNase is encoded by the *RNASE1* gene in humans. *RNASE1* is a single-copy gene in all main lineages of non-colobine primates, but it was duplicated in the douc langur (*Pygathrix nemaeus*), an Asian colobine, generating its sister gene *RNASE1B*[13]. Molecular dating suggests that the duplication postdated the separation of Asian and African colobines[13], predicting that *RNASE1B* is limited to Asian colobines. To verify this prediction, I investigated *RNASE1* in an individual from an African colobine species, the guereza (*Colobus guereza*). I amplified, cloned and sequenced a continuous genomic region of ~2 kb covering the single coding exon of *RNASE1* as well as its upstream and downstream flanking sequences (**Supplementary Fig. 1** online). Unexpectedly, the amplified DNA was a mixture of three genomic regions and contained distinct sequences from three genes hereby named *RNASE1*, *RNASE1β* and *RNASE1γ*, respectively (**Fig. 1**). A phylogenetic analysis clearly indicates independent duplications of *RNASE1* in Asian and African colobines (see Methods and **Fig. 2**). Molecular dating suggests that the two gene duplication events in guereza occurred 6.7 (95% confidence interval 4.5–9.4) and 7.8 (95% confidence interval 5.4–10.6) Myr ago, respectively, confirming that they are limited to African colobines.

The independently generated douc langur *RNASE1B* and guereza *RNASE1β* and *RNASE1γ* show exceptionally similar patterns of sequence evolution. First, in both species, these new genes evolved at a significantly faster pace at the protein sequence level than their

Department of Ecology and Evolutionary Biology, University of Michigan, 1075 Natural Science Building, 830 North University Avenue, Ann Arbor, Michigan 48109, USA. Correspondence should be addressed to J.Z. (jianzhi@umich.edu).

Received 5 December 2005; accepted 27 April 2006; published online 11 June 2006; doi:10.1038/ng1812

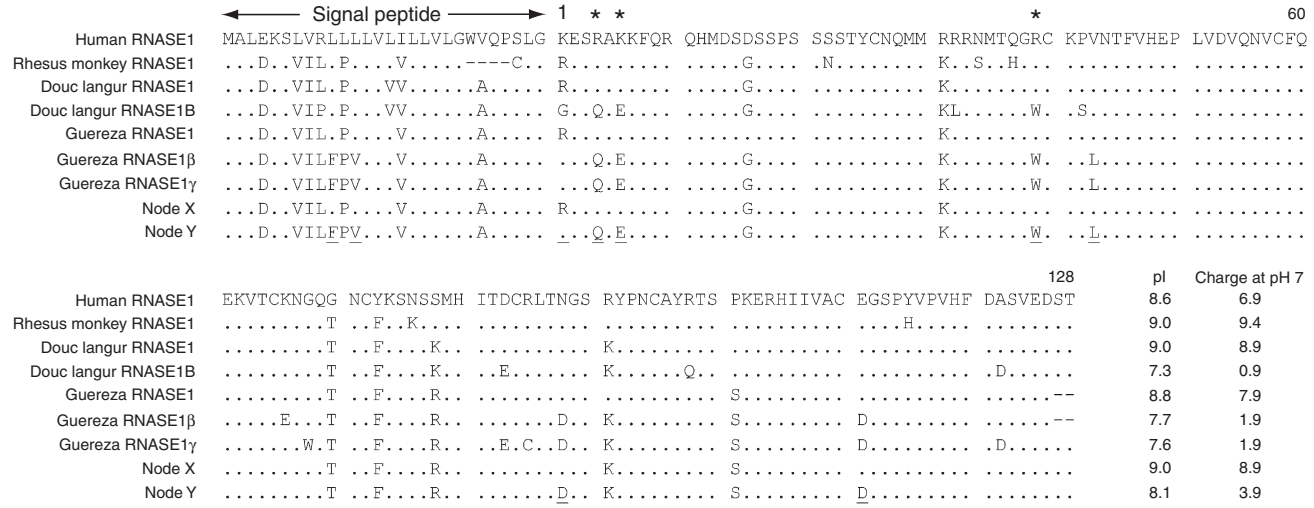

Figure 1 Sequence alignment of pancreatic RNases. I determined the guereza sequences in this paper and inferred the ancestral sequences at node X and Y (see **Fig. 2**) using a Bayesian method. All other sequences were obtained from GenBank. Amino acids identical to human RNASE1 are indicated by dots and alignment gaps are indicated by dashes. The three parallel amino acid substitutions occurring in douc langur RNASE1B and guereza RNASE1β and RNASE1γ are shown by asterisks. Amino acid differences between the ancestral proteins of node X and Y are underlined. The mature peptide starts from position 1, as indicated. pI, isoelectric point of the mature peptide.

paralogs did (**Fig. 2**). For example, there have been ten amino acid substitutions in the douc langur RNASE1B since the duplication, but zero substitutions in RNASE1 ($P = 0.002$). In guereza, since the first duplication (node X in **Fig. 2**), RNASE1β and RNASE1γ have had 10 and 13 amino acid substitutions, respectively, whereas RNASE1 has had only 2 substitutions ($P = 0.02$ and 0.005, respectively). Although the two duplication events in guereza (X and Y in **Fig. 2**) were relatively close in time (~1.1 Myr apart), the number (nine) of amino acid substitutions between the two events was large, indicating that the accelerated evolution occurred primarily in the common ancestor of RNASE1β and RNASE1γ. Second, the accelerated evolution did not occur in the noncoding regions of douc langur *RNASE1B* ($P > 0.1$), guereza *RNASE1β* ($P > 0.5$) or guereza *RNASE1γ* ($P > 0.4$), suggesting that a difference in natural selection, not mutation rate, was responsible for the accelerated protein evolution of the duplicated genes. Third, since duplication, there have been 0.031 nonsynonymous substitutions per nonsynonymous site in douc langur *RNASE1B*, significantly greater than the number of substitutions (0.0077) per synonymous or noncoding site ($P < 0.002$, Fisher's exact test). Similarly, between guereza nodes X and Y, there were 0.0281 nonsynonymous substitutions per nonsynonymous site, significantly greater than that (0.0037) per synonymous or noncoding site ($P < 0.0002$). Because the synonymous and noncoding sites have evolved approximately neutrally in the present case (see Methods), these comparisons strongly suggest that in both douc langur and guereza, the rapid nonsynonymous substitutions in the new genes were driven by positive darwinian selection. Fourth, the amino acid substitutions were nonrandom. In douc langur RNASE1B, nine amino acid substitutions occurred in the mature peptide and seven of them involved charge changes. Notably, all seven increased the negative charge of the protein. Similarly, between nodes X and Y in guereza, seven amino acid substitutions occurred in the mature peptide. Four of them involved charge changes, and all four increased the negative charge of the protein. The chance probability of a substitution pattern that is at least as skewed as observed is 0.00026 for the douc langur and 0.026 for the guereza (see Methods). Because of the charge-altering substitutions, both the net charge at pH 7 and the isoelectric point are considerably lower for douc langur RNASE1B than for RNASE1 (**Fig. 1**). Similar reductions are observed between the ancestral proteins at nodes X and Y in guereza, although to a smaller extent (**Fig. 1**).

These common features in sequence evolution prompt the hypothesis that similar functional changes have occurred independently in

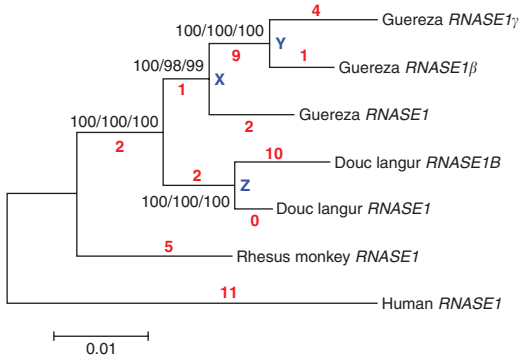

Figure 2 Phylogenetic relationships among colobine pancreatic RNase genes. I reconstructed the tree using the 2-kb DNA sequences, including the RNase coding region and flanking noncoding regions (see Methods), with a total of 1,954 nucleotide sites after the removal of alignment gaps. The maximum-likelihood, neighboring-joining and maximum-parsimony methods all support the same tree topology, with the bootstrap percentages from the three methods shown by black numbers at interior nodes. The branches are drawn to scale, showing the number of nucleotide substitutions per site. I estimated these from the neighbor-joining tree based on the Kimura's two-parameter distance. The red numbers show the inferred total numbers of amino acid substitutions in signal and mature peptides on each branch of the tree. The three nodes where gene duplication took place are indicated by blue X, Y and Z, respectively. See Methods for details of the phylogenetic analysis. Guereza is an African colobine, douc langur is an Asian colobine and rhesus monkey is a cercopithecine.

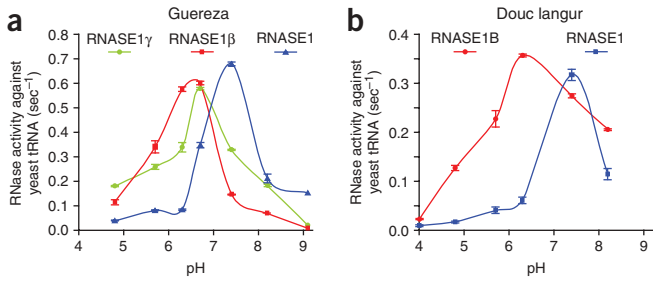

Figure 3 RNase activities against yeast tRNA at different pHs. (**a**) Recombinant proteins of guereza RNases. (**b**) Recombinant proteins of douc langur RNases. Panel **b** is adapted from ref. 13. The mean enzyme activities from three replicates and the associated 95% confidence intervals (bars) are given.

the duplicated RNase genes of douc langur and guereza. Because RNAs, the substrates of RNases, are charged molecules, the charge-altering substitutions in RNases could potentially affect enzyme-substrate interaction, thus influencing catalytic activity. To test this hypothesis, I prepared recombinant proteins from the guereza *RNASE1*, *RNASE1β* and *RNASE1γ* genes and examined their ribonucleolytic activities at different pHs. I found that the optimal pH of RNASE1 was 7.4, but those of RNASE1β and RNASE1γ were both 6.7 (**Fig. 3a**). There were no obvious differences among the enzymes in the catalytic activity at optimal pH. These findings are similar to those in douc langur, where the optimal pH of RNASE1B (6.3) is lower than that of RNASE1 (7.4) (**Fig. 3b**). Because the optimal pH (7.4) of douc langur and guereza RNASE1 is identical to that of human and rhesus monkey RNASE1 (ref. 13), I conclude that parallel decreases of optimal pH have occurred in the douc langur RNASE1B and guereza RNASE1β and RNASE1γ. Notably, the lowered optimal pHs (6.3–6.7) of these RNases match the reduced pH (6–7) of colobine small intestine, suggesting that the positive selection detected in both species arose from the demand for a high ribonucleolytic activity in an acidified microenvironment.

Three parallel amino acid substitutions occurred in douc langur RNASE1B and in the common ancestor of guereza RNASE1β and RNASE1γ (**Fig. 4a**), all increasing the negative charge of the protein.

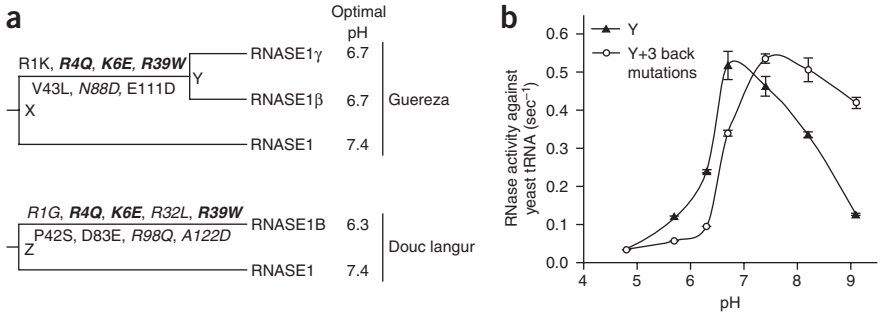

Figure 4 Parallel amino acid substitutions and parallel functional changes in guereza and douc langur pancreatic RNases. (**a**) Parallel amino acid substitutions. The substitutions occurring in the mature peptide are given, with the parallel substitutions shown in boldface and charge-altering substitutions in italics. R1K means that amino acid R (arginine) is replaced by K (lysine) at position 1 of the mature peptide. (**b**) The parallel substitutions are necessary for the decrease of optimal pH in guereza RNASE1β and RNASE1γ. Protein "Y" refers to the ancestral protein at node Y. "Y+3 back mutations" refers to the construct where mutations Q4R, E6K and W39R are introduced to protein "Y". The mean enzyme activities from three replicates and the associated 95% confidence intervals (bars) are given.

Two of them also represent relatively uncommon replacements in most proteins[14]. The chance probability of having three or more parallel amino acid substitutions in the two lineages is between 0.0001 and 0.0026 (see Methods), strongly suggesting that chance alone cannot explain the occurrence of the parallel substitutions. Rather, these substitutions were probably driven by the common selective pressure. To test whether these three substitutions were responsible for the reduction of the optimal pH of colobine pancreatic RNases, I reconstructed ancestral proteins[15]. I first inferred the amino acid sequence of the ancestral protein at node Y using a Bayesian method (**Fig. 4a**). I next reconstructed the ancestral gene by site-directed mutagenesis, expressed it in bacteria and purified the recombinant protein. The optimal pH of this ancestral protein was 6.7 (**Fig. 4b**), identical to those of guereza RNASE1β and RNASE1γ. I then mutated the ancestral gene to reverse the three aforementioned parallel substitutions. The newly engineered protein had an optimal pH of 7.4 (**Fig. 4b**). Thus, the three backward mutations restored the original optimal pH of the enzyme, indicating that the three forward mutations were necessary for the decrease of the optimal pH from 7.4 to 6.7 in the evolution of guereza RNASE1β and RNASE1γ. There are also several non-parallel charge-altering substitutions in Asian and African colobines (**Fig. 4a**) and our preliminary result indicates that they may also contribute to the decrease of the optimal pH.

Why has the optimal pH of RNASE1 been maintained at 7.4 in both douc langur and guereza? Human *RNASE1* is expressed in other tissues besides pancreas[16], suggesting that the gene has other functions in addition to digestion. Indeed, human RNASE1 can degrade double-stranded (ds) RNA, which is not commonly found in the diet. Although the physiological role of this enzymatic activity (EA_{dsRNA}) is unclear, *in vitro* studies suggest that it is involved in antiviral defenses[17]. I found that whereas human, rhesus monkey, douc langur and guereza RNASE1 proteins were all able to degrade dsRNA, this activity was virtually absent in douc langur RNASE1B, guereza RNASE1β and guereza RNASE1γ (**Fig. 5**). Both structural modeling and site-directed mutagenesis showed the importance of the positively charged noncatalytic residues to the EA_{dsRNA} of RNASE1 (refs. 13,18). Consistent with this model, the three parallel amino acid substitutions, all removing positive charges, greatly reduced EA_{dsRNA} in douc langur RNASE1B[13]. Thus, the adaptive substitutions that lowered the optimal pH for degrading dietary RNA abolished the dsRNA degradation activity. Although unproven, it is generally believed that foregut fermentation and leaf-eating emerged in the common ancestor of all colobines. Fossil evidence suggests that these changes occurred at least 10 Myr ago[8], predating the duplications of *RNASE1*. This implies that the shift in optimal pH of the pancreatic RNases was not necessary for the changes in diet and digestive physiology of colobines. Rather, the latter changes provided a selective pressure for more efficient digestive RNases in acidified environments, while gene duplication offered raw genetic materials that enabled this functional improvement. Because of the constraint imposed by the dsRNA degradation activity, the adaptive evolution of the pancreatic RNases would probably have been impossible without gene duplication.

The pancreatic RNase gene was also duplicated twice in the common ancestor of

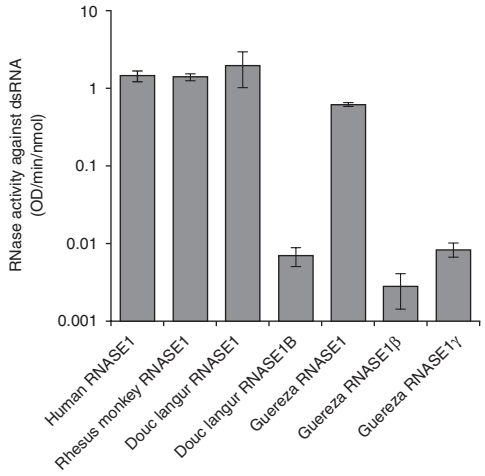

Figure 5 Loss of the RNase activity against dsRNA in douc langur RNASE1B and guereza RNASE1β and RNASE1γ. The mean activities from three replicates and the associated 95% confidence intervals (bars) are given. The original data for human, rhesus monkey and douc langur RNases are from ref. 13.

ruminants[3]. These duplicated RNases have some functional changes similar to those in the duplicated colobine RNases[3], but the parallel functional changes between ruminant and colobine RNases are due to divergent rather than parallel amino acid substitutions[3]. More surprisingly, site-directed mutagenesis shows that some of the ruminant-specific substitutions would have had different functional effects if they had taken place in colobine RNases[3]. This comparison suggests that parallel adaptive substitutions have a relatively narrow window of time to occur during evolution, because when the genetic background (including both intra- and inter-molecular interactions) is altered, even the same substitutions may have different functional consequences. I predict that convergent substitutions, which result in the occurrence of same amino acids beginning from different starting amino acids, are even rarer than parallel substitutions[1,2].

By evolutionary analysis of DNA sequences, functional assays of ancestral proteins and site-directed mutagenesis, I have shown that the pancreatic RNase gene was duplicated in Asian and African colobines independently and that the duplicated genes subsequently experienced parallel functional changes by means of parallel amino acid replacements driven by a common selective pressure. All four requirements for demonstrating adaptive parallel evolution of protein sequences are thus fulfilled. Although molecular evolution is undoubtedly largely stochastic, my results suggest that it is also repeatable and predictable to a certain degree in the presence of similar selective pressures, echoing laboratory observations of large numbers of parallel amino acid substitutions among replicates of *in vitro* evolutionary experiments conducted in viruses[19]. The establishment of the four criteria allows rigorous testing of hypotheses of adaptive parallel evolution at the protein sequence level, which will help assess the relative contributions of chance and necessity in biological evolution[20,21].

METHODS

Amplification and sequencing of guereza pancreatic RNase genes. Using polymerase chain reaction (PCR), I amplified a ~2 kb genomic region (**Supplementary Fig. 1**) covering the single coding exon of the pancreatic RNase genes as well as upstream and downstream noncoding regions from the genomic DNA of a guereza monkey. The primer sequences were from ref. 13. PCR was conducted with high-fidelity polymerase under conditions recommended by the manufacturer (Invitrogen). The products were cloned into pCR4Blunt-TOPO vector (Invitrogen) and sequenced from both directions using the dideoxy chain termination method with an automated sequencer. Multiple colonies were sequenced and three distinct sequences were identified. To exclude the possibility that these sequence variations were generated by PCR errors, I conducted a second PCR-sequencing experiment. The same three sequences were recovered. Although it remains possible that the three sequences were from two loci, the high level of sequence divergence among them (1.8–2.7% for the 2 kb region) suggests the existence of three separate genes, because the allelic difference in the nuclear genome of Old World monkeys is typically an order of magnitude lower[22]. Furthermore, the divergence among the three guereza sequences is greater than that (1.4%) between douc langur *RNASE1* and *RNASE1B* for the same region.

Evolutionary analysis. I retrieved from GenBank sequences of the aforementioned 2 kb genomic region from human (*Homo sapiens*), rhesus monkey (*Macaca mulatta*) and douc langur (both *RNASE1* and *RNASE1B*). I aligned these sequences, along with the three sequences from the guereza, using CLUSTALW. Because the species involved are closely related, the alignment was straightforward and reliable. I then reconstructed the phylogenetic tree of the 7 DNA sequences, using the 1,954 nucleotide sites remaining after the removal of alignment gaps. I used the neighbor-joining method, as implemented in the MEGA3 program. Various nucleotide distances, including Jukes-Cantor, Kimura's two-parameter, Tajima-Nei, Tamura-Nei and Tamura distances, were used; all gave the same tree topology. The same topology was also found when I reconstructed the tree using the maximum-likelihood and maximum-parsimony methods implemented in PAUP*. In the likelihood analysis, I used the general reversible model with a gamma distribution and a fraction of invariable sites that describes the among-site rate variation. Bootstrap percentages derived from 2,000 replicates were used to evaluate the repeatability of the reconstructed trees. Insertions and deletions, in addition to the nucleotide substitutions, also support independent duplications of *RNASE1* in Asian and African colobines (data not shown).

Gene conversion, which homogenizes DNA sequences within a species, might have confounded this analysis. Therefore, I used Sawyer's method[23], as implemented in the computer program GENECONV 1.81, to examine the 2-kb region. No significant signals of gene conversion were found. Thus, the phylogenetic tree is unlikely to have been distorted by this process.

Based on the inferred phylogenetic tree, I reconstructed the ancestral nucleotide and protein sequences at all interior nodes of the tree using the distance-based Bayesian method[24]. Because of the low level of sequence divergence, the inference was highly reliable, reflected by >0.995 mean posterior probabilities. At node Y, for which the ancestral protein was experimentally reconstructed, all sites have posterior probabilities >0.99. I also confirmed these results with the likelihood-based Bayesian method[25]. Tajima's test of molecular clock[26] was applied for comparing the substitution rates of two sequences when a third sequence was used as an outgroup. To date the duplication events in the guereza, I used the noncoding region (~ 1.5 kb), because the molecular clock hypothesis could not be rejected for this region between any pair of the three guereza pancreatic RNase genes. The calibration used was the divergence between cercopithecines (for example, rhesus monkey) and colobines at 15 Myr ago, determined from fossils[8]. I used the bootstrap method to obtain the 95% confidence intervals of the time estimates. The transition/transversion mutational bias was estimated from the noncoding regions of the guereza and rhesus monkey sequences to be 2.1. The potential numbers of synonymous (S) and nonsynonymous (N) sites of a sequence as well as the observed substitutions (s and n) at these sites between two sequences were computed following Zhang *et al.*[27]. I estimated that for colobine pancreatic RNase genes $S = 129.3$ and $N = 320.7$. Fisher's exact test was used to compare the rates of nucleotide substitutions at different types of sites[28]. The noncoding and synonymous sites examined in this study evolve approximately neutrally, because the average divergence at these sites between human and Old World monkeys[13] was similar to the reported genomic average divergence between the human and rhesus monkey at various introns and pseudogenes[29].

Using the program HON-NEW (ref. 30), I estimated the probability (p) that a random nonsynonymous mutation involves a charge change to be 0.377 for the RNase genes. Given M nonsynonymous substitutions, I calculated the probability (P) that at least m of them involve charge changes, all in one direction, using the binomial distribution with the formula

$$P = \sum_{i=m}^{M} \left(\frac{1}{2}\right)^{i-1} \binom{M}{i} p^i (1-p)^{M-i}$$

Isoelectric points (pI) and net charges of mature peptides were computed by the web program Protein Calculator. I used the computer program CAPE to test whether the number of observed parallel amino acid substitutions was significantly higher than expected by chance, given the total numbers of amino acid replacements in the two evolutionary lineages under investigation[2]. Two amino acid substitution models (Poisson and JTT) that were implemented in CAPE were used. To be conservative, I present here the probabilities computed under the JTT model, which were higher than those computed under the Poisson model. When all sites in the mature peptide were assumed to be variable, the chance probability (CP) of observing three or more parallel amino acid substitutions in the tree branch linking node X and Y and that linking node Z and RNASE1B was 0.000071. However, I found that among 20 primate pancreatic RNase sequences that are currently available, there were only 42 sites in the mature peptide that are variable. Based on the conservative assumption that amino acid substitutions can only occur in these 42 sites, I estimated CP to be 0.0026. The true CP is thus between 0.000071 and 0.0026.

Recombinant protein preparation and functional assays. The experimental procedures follow ref. 13. I subcloned the guereza *RNASE1*, *RNASE1β* and *RNASE1γ* genes into the bacterial expression vector pFLAG-CTS (Kodak) and verified them 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 (Sigma), but does not affect the RNase activity[13]. I used the QuikChange site-directed mutagenesis kit (Stratagene) to mutate *RNASE1β* to generate the ancestral sequence at node Y. The inferred ancestral protein sequence at node Y differs from guereza RNASE1β by only one amino acid: node Y has K (lysine) at position 66, whereas RNASE1β has E (glutamic acid). Three more mutations (Q4R, E6K and W39R) were introduced to the ancestral sequence of node Y to generate the construct "Y+3 back mutations". All the mutations were confirmed by DNA sequencing. I isolated, purified and quantified recombinant proteins as previously described[13]. I measured the RNase activity of the recombinant proteins against yeast tRNA at different pHs (40 mM sodium phosphate buffer with pH 4.8–9.1) at 25 °C. Purified RNase (1–10 pmol) was added to 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. The catalytic activity of RNase was computed as nmol of RNA digested per second per nmol of RNase. I measured the RNase activity (EA_{dsRNA}) against dsRNA (poly(U)•poly(A) annealed from poly(U) and poly(A), Pharmacia) 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 by ultraviolet absorbance at 260 nm (ref. 13). EA_{dsRNA} was found to be highest at pH 7 for all three guereza RNases. The results at pH 7 were thus reported for all constructs. Three replicates of experiments were conducted at each condition examined and the means and their 95% confidence intervals were computed.

GenBank accession numbers. The DNA sequences reported in this paper have been submitted to GenBank (DQ516063–DQ516065).

URLs. CLUSTALW, http://bioweb.pasteur.fr/seqanal/interfaces/clustalw-simple.html; MEGA3, http://www.megasoftware.net/mega.html; PAUP*, http://paup.csit.fsu.edu/; Protein Calculator, http://www.scripps.edu/~cdputnam/protcalc.html.

Note: Supplementary information is available on the Nature Genetics website.

ACKNOWLEDGMENTS

I thank M. Bakewell, S. Cho, W. Grus, A. Rooney and D. Webb for comments. This work was supported by the US National Institutes of Health.

COMPETING INTERESTS STATEMENT

The author declares that he has no competing financial interests.

Published online at http://www.nature.com/naturegenetics
Reprints and permissions information is available online at http://npg.nature.com/reprintsandpermissions/

1. Doolittle, R.F. Convergent evolution: the need to be explicit. *Trends Biochem. Sci.* **19**, 15–18 (1994).
2. Zhang, J. & Kumar, S. Detection of convergent and parallel evolution at the amino acid sequence level. *Mol. Biol. Evol.* **14**, 527–536 (1997).
3. Zhang, J. Parallel functional changes in the digestive RNases of ruminants and colobines by divergent amino acid substitutions. *Mol. Biol. Evol.* **20**, 1310–1317 (2003).
4. Stewart, C.B., Schilling, J.W. & Wilson, A.C. Adaptive evolution in the stomach lysozymes of foregut fermenters. *Nature* **330**, 401–404 (1987).
5. Yokoyama, R. & Yokoyama, S. Convergent evolution of the red- and green-like visual pigment genes in fish, *Astyanax fasciatus*, and human. *Proc. Natl. Acad. Sci. USA* **87**, 9315–9318 (1990).
6. Rodriguez-Trelles, F., Tarrio, R. & Ayala, F.J. Convergent neofunctionalization by positive Darwinian selection after ancient recurrent duplications of the xanthine dehydrogenase gene. *Proc. Natl. Acad. Sci. USA* **100**, 13413–13417 (2003).
7. Mundy, N.I. A window on the genetics of evolution: MC1R and plumage colouration in birds. *Proc. Biol. Sci.* **272**, 1633–1640 (2005).
8. Delsen, E. Evolutionary history of the colobine monkeys in paleoenvironmental perspective. in *Colobine Monkeys: Their Ecology, Behaviour and Evolution* (eds. Davies, A.G. & Oates, J.F.) 11–44 (Cambridge Univ. Press, Cambridge, UK, 1994).
9. Kay, R.N.B. & Davies, A.G. Digestive physiology. in *Colobine Monkeys: Their Ecology, Behaviour and Evolution* (eds. Davies, A.G. & Oates, J.F.) 229–249 (Cambridge Univ. Press, Cambridge, UK, 1994).
10. Barnard, E.A. Biological function of pancreatic ribonuclease. *Nature* **221**, 340–344 (1969).
11. 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).
12. Guyton, A.C. & Hall, J.E. *Textbook of Medical Physiology* (Saunders, London, 1996).
13. Zhang, J., Zhang, Y.P. & Rosenberg, H.F. Adaptive evolution of a duplicated pancreatic ribonuclease gene in a leaf-eating monkey. *Nat. Genet.* **30**, 411–415 (2002).
14. Tang, H., Wyckoff, G.J., Lu, J. & Wu, C.I. A universal evolutionary index for amino acid changes. *Mol. Biol. Evol.* **21**, 1548–1556 (2004).
15. Jermann, T.M., Opitz, J.G., Stackhouse, J. & Benner, S.A. Reconstructing the evolutionary history of the artiodactyl ribonuclease superfamily. *Nature* **374**, 57–59 (1995).
16. Futami, J. *et al.* Tissue-specific expression of pancreatic-type RNases and RNase inhibitor in humans. *DNA Cell Biol.* **16**, 413–419 (1997).
17. Sorrentino, S. & Libonati, M. Human pancreatic-type and nonpancreatic-type ribonucleases: a direct side-by-side comparison of their catalytic properties. *Arch. Biochem. Biophys.* **312**, 340–348 (1994).
18. Sorrentino, S., Naddeo, M., Russo, A. & D'Alessio, G. Degradation of double-stranded RNA by human pancreatic ribonuclease: crucial role of noncatalytic basic amino acid residues. *Biochemistry* **42**, 10182–10190 (2003).
19. Bull, J.J. *et al.* Exceptional convergent evolution in a virus. *Genetics* **147**, 1497–1507 (1997).
20. Carroll, S.B. Chance and necessity: the evolution of morphological complexity and diversity. *Nature* **409**, 1102–1109 (2001).
21. Monod, J. *Chance and Necessity* (Vintage, New York, 1972).
22. Magness, C.L. *et al.* Analysis of the *Macaca mulatta* transcriptome and the sequence divergence between *Macaca* and human. *Genome Biol.* **6**, R60 (2005).
23. Sawyer, S. Statistical tests for detecting gene conversion. *Mol. Biol. Evol.* **6**, 526–538 (1989).
24. Zhang, J. & Nei, M. Accuracies of ancestral amino acid sequences inferred by the parsimony, likelihood, and distance methods. *J. Mol. Evol.* **44** (Suppl.), S139–S146 (1997).
25. Yang, Z., Kumar, S. & Nei, M. A new method of inference of ancestral nucleotide and amino acid sequences. *Genetics* **141**, 1641–1650 (1995).
26. Tajima, F. Simple methods for testing the molecular evolutionary clock hypothesis. *Genetics* **135**, 599–607 (1993).
27. 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).
28. Zhang, J., Kumar, S. & Nei, M. Small-sample tests of episodic adaptive evolution: a case study of primate lysozymes. *Mol. Biol. Evol.* **14**, 1335–1338 (1997).
29. Li, W. *Molecular Evolution* (Sinauer, Sunderland, Massachusetts, USA, 1997).
30. Zhang, J. Rates of conservative and radical nonsynonymous nucleotide substitutions in mammalian nuclear genes. *J. Mol. Evol.* **50**, 56–68 (2000).