



Contents lists available at ScienceDirect

Molecular Phylogenetics and Evolution

journal homepage: www.elsevier.com/locate/ympev

Letter to the Editor

Phylogenetic evidence for parallel adaptive origins of digestive RNases in Asian and African leaf monkeys: A response to Xu et al. (2009)

Identification of parallel amino acid substitutions accompanying parallel phenotypic evolution is of considerable interest to molecular evolutionists because such parallel substitutions are likely to be adaptive and functionally important (Stewart et al., 1987; Yokoyama and Yokoyama, 1990; Zhang and Kumar, 1997). In 2006, I reported that the gene encoding pancreatic ribonuclease (RNase1) was duplicated independently in Asian and African colobine monkeys (Zhang, 2006). Statistical analyses of DNA sequences, functional assays of reconstructed ancestral proteins, and site-directed mutagenesis showed that the new genes acquired enhanced digestive efficiencies through three parallel amino acid replacements driven by positive selection. They also lost a non-digestive function independently, under a relaxed selective constraint. In a recent Short Communication, Xu and colleagues suggested that the independent duplications that I reported were actually only one duplication event and that the adaptive parallel substitutions I described were no longer existent or were explainable by hypermutations at CpG sites (Xu et al., 2009). Xu et al.'s claims were not supported by available evidence. Below, I provide a detailed response and offer additional phylogenetic evidence for parallel adaptive evolution of digestive RNases in Asian and African colobines.

In my original analysis (Zhang, 2006), the RNase gene tree was reconstructed using both the coding region and flanking noncoding regions, totaling 1954 nucleotides. The tree was reconstructed with maximum-likelihood, neighbor-joining, and maximum-parsimony methods, and all methods supported the same tree with high bootstrap values (>98% for every node by every method). In the reanalysis, however, Xu et al. reconstructed the gene tree using only ~450 nucleotides of the coding region. Their tree (Fig. 2 in Xu et al. (2009)) is less reliable than my tree (Fig. 2 in Zhang (2006)) for five reasons.

First, it is well-known that molecular phylogenies are generally more reliable when they are based on many nucleotides than based on few nucleotides. Xu et al.'s tree was based on a much smaller number of nucleotide sites than was my tree. As a result, the bootstrap values in their tree are low, as they admitted (Xu et al., 2009). For example, the clade of all RNase1B sequences has a bootstrap value of only 75% in their Fig. 2A and 61% in their Fig. 2B. The statistical support for their tree topology is not significantly greater than for mine even when only the coding region is considered. Second, the topology of their tree is not consistent with their hypothesis of one duplication event in the common ancestor of colobines. If their hypothesis were correct, one should observe that (i) all RNase1 sequences of colobine monkeys form a monophyletic group and that (ii) the phylogenetic relationships among colobine RNase1 genes are identical to the relationships

among colobine RNase1B genes. Neither of these predictions was observed in their Fig. 2, indicating that their hypothesis is incorrect or the coding sequences are simply too short to generate reliable trees. Their tree topology cannot be explained by any simple evolutionary scenario; multiple gene duplications and losses would have to be invoked. Third, they claimed that their hypothesis of a single gene duplication event is more parsimonious than my hypothesis of multiple independent duplication events. However, multiple duplication events must have happened because some colobine monkeys have three or even four RNase genes (Schienman et al., 2006; Zhang, 2006). Use of the parsimony principle for gene duplication is apparently inappropriate here. Fourth, they claimed that my use of noncoding sequences in tree-making could result in a wrong tree, possibly due to gene conversion, but they did not provide direct evidence for gene conversion. In my original analysis, I specifically tested the possibility of gene conversion by a statistical method (Sawyer, 1989) but found no such evidence (Zhang, 2006, p. 822, right column, paragraph 3). So, in short, Xu et al. used partial data to generate an unreliable tree and argued that it is better than the full-data-based tree, without providing explicit evidence for the inferiority of the full data to the partial data. Finally and most importantly, use of the coding sequences alone as in Xu et al. (2009) could result in a wrong tree when there are parallel amino acid substitutions. In fact, some earlier studies used misleading gene trees as evidence for the presence of parallel substitutions (Stewart et al., 1987; Swanson et al., 1991). To illustrate this point more clearly, I conducted additional phylogenetic analysis of the RNase sequences. I use sequences published in (Zhang, 2006) because of the availability of sufficient noncoding sequences for comparison with coding sequences. When noncoding and coding sequences are combined, the tree indicates independent duplication events (Fig. S1A). When only noncoding sequences are used, the same tree is obtained (Fig. S1B). When only coding sequences are used, the tree topology changes so that the duplicated genes of douc langur (*Pygathrix nemaeus*) and guereza (*Colobus guereza*) are clustered (Fig. S1C). However, with a bootstrap value of 58–69%, this clustering is not statistically significant (Fig. S1C). Furthermore, the tree cannot be easily interpreted as one duplication event because the RNase1 sequences of the two colobines are not clustered as would be expected if only one duplication event took place. The tree of Fig. S1C is likely wrong due to the three parallel substitutions in the duplicated RNase genes of douc langur and guereza. Site-directed mutagenesis and functional assays directly demonstrated that the three parallel substitutions affected the enzyme activities of the duplicated RNases in such a way that the enzymes are now more adaptive to their microenvironment in the colobine small intestine (Zhang, 2006). When I remove the three codons where the parallel substitutions occurred, the duplicated genes of the two colobines are no longer clustered (Fig. S1D). Because the short coding sequences are not expected to provide reliable trees, it is not surprising that the tree of

Fig. S1D is still not identical to that of Fig. S1A. Taken together, multiple lines of evidence show that Xu et al.'s tree in their Fig. 2 is at least unreliable and most likely incorrect due to the use of fewer nucleotides as well as the influence of parallel adaptive substitutions.

If Xu et al.'s tree is unreliable or incorrect, as demonstrated above, most of their further analyses collapse. However, I do want to respond to the observation of substitutions at CpG sites. Of the three parallel nonsynonymous substitutions I reported for the duplicated RNase genes of douc langur and guereza, two occurred at CpG sites (nucleotide change G95A that resulted in amino acid change R4Q; nucleotide change C199T that resulted in amino acid change R39W). Can we explain these parallel substitutions by mutation and drift without invoking positive selection? The answer is "No". The point mutation rate in Old World monkeys is about 1×10^{-9} per site per year (Yi et al., 2002). For CpG sites, the rate may be 10 times high (Robertson and Wolfe, 2000). Thus, the neutral substitution rate, which is identical to the neutral mutation rate (Kimura, 1983), can be as high as 1×10^{-8} per site per year at CpG sites. In other words, one should expect to see one substitution per 100 million years (MY) at a CpG site. However, the three parallel substitutions I reported in guereza all occurred in a short window of ~ 1.1 MY (between node X and Y in Fig. 2 of Zhang (2006)). One may argue that the estimate of 1.1 MY may not be accurate. But even if it were 10 MY, it is still obvious that mutation and drift alone would not be sufficient to drive the fixations of the three parallel changes, one of which was not even at a CpG site. By contrast, positive selection can increase the fixation probability drastically. The probability of fixation is approximately $2s$ for a beneficial allele with a fitness advantage of s , but is $1/(2N)$ for a neutral allele (Kimura, 1983), where N is the effective population size. Their ratio is $4Ns$. If we assume $s = 0.01$ and $N = 5 \times 10^4$ for the monkeys studied here, the ratio becomes $4Ns = 2000$. That is, even a 1% selective advantage can increase the rate of substitution by 2000-fold, much more effective than a 10-fold increase of mutation rate at CpG sites.

Together, the available evidence supports parallel adaptive evolution of digestive RNases in Asian and African colobines. It would be of significant interest to study the pancreatic RNase genes and other digestive enzyme genes in additional colobines to under-

stand the detailed molecular evolutionary mechanisms underlying the adaptation of colobines to leaf-eating.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.ympev.2009.07.003.

References

- Kimura, M., 1983. The Neutral Theory of Molecular Evolution. Cambridge University Press, Cambridge.
- Robertson, K.D., Wolfe, A.P., 2000. DNA methylation in health and disease. *Nat. Rev. Genet.* 1, 11–19.
- Sawyer, S., 1989. Statistical tests for detecting gene conversion. *Mol. Biol. Evol.* 6, 526–538.
- Schienman, J.E., Holt, R.A., Auerbach, M.R., Stewart, C.B., 2006. Duplication and divergence of 2 distinct pancreatic ribonuclease genes in leaf-eating African and Asian colobine monkeys. *Mol. Biol. Evol.* 23, 1465–1479.
- Stewart, C.B., Schilling, J.W., Wilson, A.C., 1987. Adaptive evolution in the stomach lysozymes of foregut fermenters. *Nature* 330, 401–404.
- Swanson, K.W., Irwin, D.M., Wilson, A.C., 1991. Stomach lysozyme gene of the langur monkey: tests for convergence and positive selection. *J. Mol. Evol.* 33, 418–425.
- Xu, L., Su, Z., Gu, Z., Gu, X., 2009. Evolution of RNases in leaf monkeys: being parallel gene duplications or parallel gene conversions is a problem of molecular phylogeny. *Mol. Phylogenet. Evol.* 50, 397–400.
- Yi, S., Ellsworth, D.L., Li, W.H., 2002. Slow molecular clocks in Old World monkeys, apes, and humans. *Mol. Biol. Evol.* 19, 2191–2198.
- Yokoyama, R., Yokoyama, S., 1990. 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.
- Zhang, J., 2006. Parallel adaptive origins of digestive RNases in Asian and African leaf monkeys. *Nat. Genet.* 38, 819–823.
- Zhang, J., Kumar, S., 1997. Detection of convergent and parallel evolution at the amino acid sequence level. *Mol. Biol. Evol.* 14, 527–536.

Jianzhi Zhang*

Department of Ecology and Evolutionary Biology,
University of Michigan, 1075 Natural Science Building,
830 North University Avenue, Ann Arbor, MI 48109, USA

* Fax: +1 734 763 0544.

E-mail address: jianzhi@umich.edu

Available online 10 July 2009