# Letter to the Editor

## Erratic Evolution of SRY in Higher Primates

*Xiaoxia Wang,* Jianzhi Zhang,† and Ya-ping Zhang**

*Laboratory of Molecular Evolution and Genome Diversity, Kunming Institute of Zoology, Chinese Academy of Sciences; and†Departments of Ecology and Evolutionary Biology and Molecular, Cellular, and Developmental Biology, University of Michigan

*SRY,* a Y-chromosomal gene that is pivotal in initiating the development of testis and a determinant of male sex in mammals (Goodfellow and Lovell-Badge 1993), has been subject to intense evolutionary study in the past decade (Foster et al. 1992; Tucker and Lundrigan 1993; Whitfield, Lovell-Badge, and Goodfellow 1993; Pamilo and O'Neill 1997; Nagai 2001). SRY protein is a transcription factor containing the conserved DNA-binding HMG domain (~78 amino acids) in the middle of the sequence and highly variable N- and C-terminal sequences (Whitfield, Lovell-Badge, and Goodfellow 1993). A higher rate of nonsynonymous nucleotide substitution than that of synonymous substitution in the terminal regions of SRY has been reported in primates and rodents, with suggestions that these regions may be subject to positive Darwinian selection (Tucker and Lundrigan 1993; Whitfield, Lovell-Badge, and Goodfellow 1993). Here, we sequence the SRY genes of nine Old World (OW) monkeys and show that the evolution of the SRY has been significantly decelerated in these species. Thus, the SRY shows a complex pattern of erratic evolution among different groups of primates, raising an intriguing possibility of varied selective pressures on this fundamentally important gene in evolution.

Using PCR and DNA sequencing, we determined the SRY gene sequences of nine OW monkeys (see fig. 1). The SRY sequences of six hominoids, one OW monkey, and one New World (NW) monkey were available in the GenBank at the time of this study. Thus, a total of 17 sequences are analyzed here (fig. 1). Because the HMG domain of SRY is conserved throughout mammalian evolution, we only analyze the N- and C-terminal regions, which have 120 codons after the indels are removed. We first computed the numbers of synonymous ($d_S$) and nonsynonymous ($d_N$) nucleotide substitutions between pairs of the six hominoid sequences, using the modified Nei-Gojobori method (Zhang, Rosenberg, and Nei 1998). We found that $d_N$ is greater than $d_S$ for 13 of the 15 pairwise comparisons (fig. 2), suggesting possible actions of positive selection. We then conducted the same analysis for the 10 OW monkeys and found that $d_N$ is smaller than $d_S$ for most of these comparisons (fig. 2). Only among closely related OW monkeys did we find greater $d_N$ than $d_S$. In such cases, however, the numbers of substitutions are too few to draw any solid conclusion of the possible action of positive selection. Regardless, the contrasting patterns of high $d_N/d_S$ ratios in hominoids and low ratios in OW monkeys are clear (fig. 2). Use of Li's (1993) method gave similar results.

To examine whether the $d_N/d_S$ ratio has been reduced in OW monkeys or enhanced in hominoids during evolution and to rigorously test the difference in $d_N/d_S$ among groups of organisms, we used a phylogeny-based analysis following Zhang, Kumar, and Nei (1997). The phylogenetic relationships of the main groups (families and subfamilies) of primates have been well established (e.g., Goodman et al. 1998), although the detailed relationships within groups are still uncertain (e.g., Zhang and Ryder 1998). We assume that the phylogeny of the 17 species analyzed here is as shown in figure 1. Use of trees slightly different from the one shown here gave essentially identical results. We used the distance-based Bayesian method (Zhang and Nei 1997) to infer the ancestral SRY sequences at all interior nodes of the tree. Because the species involved are closely related, the inference was highly reliable, with average posterior probabilities over 99% for all nodes. We then counted the numbers of synonymous ($s$) and nonsynonymous ($n$) substitutions on each branch of the tree (fig. 1). For the 10 branches that link the six hominoids, the sum of $n$ and $s$ are 27 and 7, respectively. These numbers become 18 and 15, respectively, for the group of 10 OW monkeys (fig. 1). Therefore, the $n/s$ ratio is about 3.2 times higher in hominoids ($27/7 = 3.86$) than in OW monkeys ($18/15 = 1.2$), and this difference is statistically significant ($P = 0.028$, Fisher's exact test), confirming the observation from the pairwise sequence comparisons (fig. 2). The branch linking the NW monkey and the common ancestor of hominids and OW monkeys (fig. 1) may be used as a reference to test whether the $n/s$ ratio has been increased in hominoids or decreased in OW monkeys. This reference branch shows a high $n/s$ ratio of $42.5/13.5 = 3.15$. Fisher's test indicates that the difference in $n/s$ is not significant between this branch and those of hominoids ($P = 0.43$) but is significant between this branch and those of OW monkeys ($P = 0.036$). These results suggest that the $n/s$ ratio remains unchanged in hominoids but has been lowered in OW monkeys. To further examine whether the decreased $n/s$ ratio in OW monkeys is caused by a reduction in the nonsynonymous substitution rate or an increase in the synonymous rate, we compared the average number of

Key words: SRY, evolution, hominoids, Old World monkeys, natural selection.

Address for correspondence and reprints: Ya-ping Zhang, Laboratory of Molecular Evolution and Genome Diversity, Kunming Institute of Zoology, Chinese Academy of Sciences, Kunming 650223, People's Republic of China. E-mail: zhangyp@public.km.yn.cn.

Address for correspondence and reprints: Jianzhi Zhang, Departments of Ecology and Evolutionary Biology and Molecular, Cellular, and Developmental Biology, University of Michigan, 3003 Natural Science Building, 830 North University Avenue, Ann Arbor, Michigan 48109. E-mail: jianzhi@umich.edu.

*Mol. Biol. Evol.* 19(4):582–584. 2002
 ISSN: 0737-4038

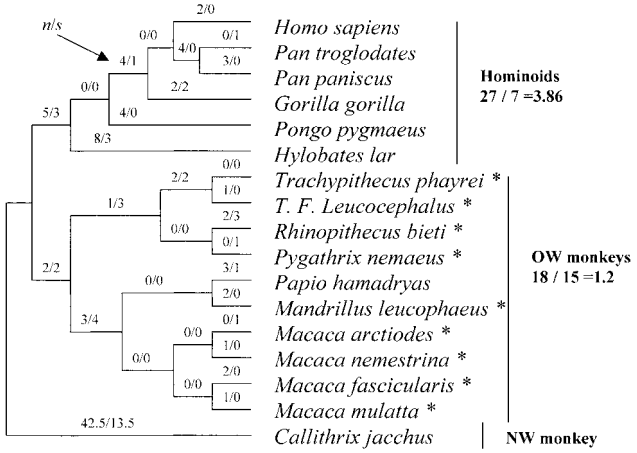


FIG. 1.—Evolution of non-HMG regions of the SRY gene in higher primates. The numbers of nonsynonymous (*n*) and synonymous (*s*) substitutions occurred in each tree branch are shown above the branches. The sums of *n* and *s* for a group of species are shown below the group names. The newly determined sequences are marked with * after the species names. The estimated numbers of potential nonsynonymous (*N*) and synonymous (*S*) sites are 252 and 108, respectively. Use of alternative trees of primates gave essentially identical results. Two pairs of primers, SRYlf2 (5′-GATCAGCAGGGCAAGTAGTC-3′) and SRYhm (5′-TGTGCCTCCTGGAAGAATGG-3′), SRYlm2 (5′-AGATGGCTC-TAGAGAATCCC-3′), and SRYhf2 (5′-TTGTAGCCAATGTTACCC-GA-3′), were used for PCR amplification and sequencing. PCR conditions were as follows: 95°C for 2 min, followed by 39 cycles of 94°C for 50 s, 44–62°C for 1 min, and 72°C for 1 min, followed by 73°C for 10 min. Automatic DNA sequencing was conducted on ABI 377 DNA sequencer. GenBank accession numbers for the new sequences are AF454965 to AF454973.

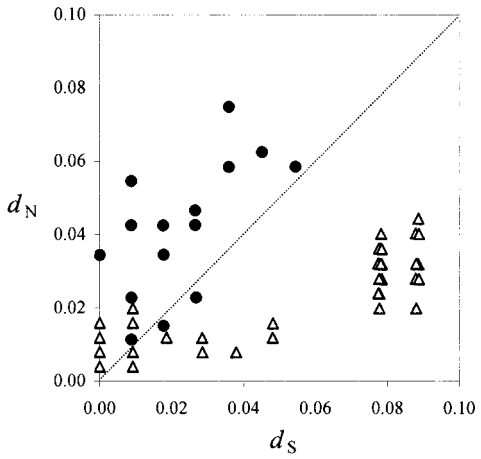


FIG. 2.—Pairwise synonymous and nonsynonymous distances of SRY genes among six hominoids (dots) and among 10 OW monkeys (triangles). Only non-HMG regions are considered.

synonymous substitutions in hominoids and OW monkeys since their separation. We find that there were on average 5.0 and 6.7 synonymous substitutions, respectively, per lineage, in hominoids and OW monkeys. These numbers become 11.7 and 5.8 for nonsynonymous substitutions in the two groups, respectively. This analysis suggests that the main reason for the decreased *n*/*s* ratio in OW monkeys is deceleration of nonsynonymous substitutions. There are reports (Li and Tanimura 1987) of reduced mutation rate in hominoids (hominoid slowdown hypothesis). Our observation of a slightly lower rate of synonymous substitution in hominoids than in OW moneys appears consistent with this hypothesis, although the observed difference is so small that it may simply because of stochastic errors.

With the estimates of the potential numbers of synonymous (*S*) and nonsynonymous (*N*) sites in SRY sequences, one may test the hypothesis of neutral evolution (Zhang, Kumar, and Nei 1997). In the present case, we estimated that $S = 108$ and $N = 252$. Thus, for hominoids, the *n*/*s* ratio (3.86) is not significantly greater than *N*/*S* (252/108 = 2.33) ($P = 0.14$, Fisher's test). For the reference branch, *n*/*s* (3.15) and *N*/*S* are not significantly different either ($P = 0.21$). For OW monkeys, however, *n*/*s* (1.2) is significantly lower than *N*/*S* ($P = 0.04$), suggesting the action of purifying selection.

To sum up, our analysis revealed reduced rate of nonsynonymous substitution and action of purifying selection in the terminal regions of SRY in OW monkeys. Although confirming earlier results of high nonsynonymous substitution rates of SRY in hominoids (Whitfield, Lovell-Badge, and Goodfellow 1993), we were not able to reject the neutral evolution hypothesis. As such, whether the rapid evolution of hominoid SRY is caused by positive selection or relaxed functional constraints remains an open question. Assuming that positive selection did occur in SRY, Hurst (1994) proposed that the selection is associated with a promiscuous mating system and predicated that there will be more nonsynonymous substitutes in polygamous than monogamous taxa. Our data, however, are not consistent with his prediction, as polygamous primates (such as langurs, macaques, and chimpanzees) do not show higher $d_N/d_S$ ratio than the monogamous gibbon. The SRY gene is located in the nonrecombining region of the Y chromosome, which means that it may be subject to certain evolutionary forces that are nonexistent or relatively unimportant in the case of autosomal genes. These evolutionary forces and phenomena include the Hill-Roberson effect, Muller's ratchet, genetic hitchhiking, background selection, male-driven evolution, hemizygosity, and reduced effective population size (Tucker and Lundrigan 1995; Charlesworth and Charlesworth 2000). However, the relative contributions of these confounding effects on the rate of nucleotide substitution are still unclear, and it is unknown whether natural selection is more likely or less likely to be detected on Y-linked genes. To understand the unexpected evolutionary pattern of SRY, we turn to the structural and functional data of the protein. It is interesting to note that all clinical mutations in human SRY resulting in phenotypic sex reversal are found in the HMG domain, except for one case of a nonsense mutation in the C-terminal region (Hawkins et al. 1992; Goodfellow and Lovell-Badge 1993; Tajima et al. 1994; Werner et al. 1995). This suggests that the non-HMG regions are under relatively relaxed selective pressures. Computational analysis on the variation of substitution rate among sites also led to this conclusion (Zhang and Gu 1998). These notions, however, do not exclude the possibility that the non-HMG regions may be under weak positive or purifying selec-

tion. In fact, the action of purifying selection is detected in OW monkeys. The variation of $d_N/d_S$ ratio among primate lineages also suggests to us that these non-HMG regions are not functionless. Rather, our findings suggest that they may play varied roles in SRY functioning among different lineages. Structure-function analysis of several primate SRY proteins may thus shed light on the functions of these non-HMG regions and help resolve the evolutionary mysteries of this fundamentally important protein.

## Acknowledgments

We thank Dr. Priscilla Tucker for valuable comments. This work was supported by Natural Science Foundation of China and Chinese Academy of Sciences grants (KSCX2-1-05) to Y-p.Z. and a start-up fund of the University of Michigan to J.Z.

LITERATURE CITED

CHARLESWORTH, B., and D. CHARLESWORTH. 2000. The degeneration of Y chromosomes. Philos. Trans. R. Soc. Lond. **355**:1563–1572.

FOSTER, J. W., F. E. BRENNAN, G. K. HAMPIKIAN, P. N. GOODFELLOW, A. H. SINCLAIR, R. LOVELL-BADGE, L. SELWOOD, M. B. RENFREE, D. W. COOPER, and J. A. M. GRAVES. 1992. Evolution of sex determination and the Y chromosome: SRY-related sequences in marsupials. Nature **359**:531–532.

GOODFELLOW, P. N., and R. LOVELL-BADGE. 1993. SRY and sex determination in mammals. Annu. Rev. Genet. **27**:71–92.

GOODMAN, M., C. A. PORTER, J. CZELUSNIAK, S. L. PAGE, H. SCHNEIDER, J. SHOSHANI, G. GUNNELL, and C. P. GROVES. 1998. Toward a phylogenetic classification of primates based on DNA evidence complemented by fossil evidence. Mol. Phylogenet. Evol. **9**:585–598.

HAWKINS, J. R., A. TAYLOR, P. BERTA, J. LEVILLIERS, B. VAN DER AUWERA, and P. N. GOODFELLOW. 1992. Mutational analysis of SRY: nonsense and missense mutations in XY sex reversal. Hum. Genet. **88**:471–474.

HURST, L. D. 1994. Embryonic growth and the evolution of the mammalian Y chromosome. I. The Y as an attractor for selfish growth factors. Heredity **73**:223–232.

LI, W.-H. 1993. Unbiased estimation of the rates of synonymous and nonsynonymous substitution. J. Mol. Evol. **36**: 96–99.

LI, W.-H., and M. TANIMURA. 1987. The molecular clock runs more slowly in man than in apes and monkeys. Nature **326**: 93–96.

NAGAI, K. 2001. Molecular evolution of Sry and Sox gene. Gene **270**:161–169.

PAMILO, P., and R. J. O'NEILL. 1997. Evolution of the SRY genes. Mol. Biol. Evol. **14**:49–55.

TAJIMA, T., J. NAKAE, N. SHINOHARA, and K. FUJIEDA. 1994. A novel mutation localized in the 3′ non-HMG box region of the SRY gee in 46X,Y gonasal dysgenesis. Hum. Mol. Genet. **3**:1187–1189.

TUCKER, P. K., and B. L. LUNDRIGAN. 1993. Rapid evolution of the sex determining locus in Old World mice and rats. Nature **364**:715–717.

———. 1995. The nature of gene evolution on the mammalian Y chromosome: lessons from Sry. Philos. Trans. R. Soc. Lond. **350**:221–227.

WERNER, M. H., J. R. HUTH, A. M. GRONENBORN, and G. M. CLORE. 1995. Molecular basis of human 46X, Y sex reversal revealed from the three-dimensional solution structure of the human SRY-DNA complex. Cell **81**:705–714.

WHITFIELD, L. S., R. LOVELL-BADGE, and P. N. GOODFELLOW. 1993. Rapid sequence evolution of the mammalian sex-determining gene SRY. Nature **364**:713–715.

ZHANG, J., and X. GU. 1998. Correlation between the substitution rate and rate variation among sites in protein evolution. Genetics **149**:1615–1625.

ZHANG, J., S. KUMAR, and M. NEI. 1997. Small-sample tests of episodic adaptive evolution: a case study of primate lysozymes. Mol. Biol. Evol. **14**:1335–1338.

ZHANG, J., and M. NEI. 1997. Accuracies of ancestral amino acid sequences inferred by the parsimony, likelihood, and distance methods. J. Mol. Evol. **44**(Suppl. 1):S139–S146.

ZHANG, J., H. F. ROSENBERG, and M. NEI. 1998. Positive Darwinian selection after gene duplication in primate ribonuclease genes. Proc. Natl. Acad. Sci. USA **95**:3708–3713.

ZHANG, Y. P., and O. A. RYDER. 1998. Mitochondrial cytochrome *b* gene sequence of old world monkey: with special reference on evolution of Asian colobines. Primates **39**:39–49.

RODNEY HONEYCUTT, reviewing editor

Accepted December 18, 2001