Pseudogenization of the tumor-growth promoter angiogenin in a leaf-eating monkey

Jianzhi Zhang\textsuperscript{a,\textdagger}, Ya-ping Zhang\textsuperscript{b}

\textsuperscript{a}Department of Ecology and Evolutionary Biology, University of Michigan, 3003 Natural Sciences Building, 830 North University Avenue, Ann Arbor, MI 48109, USA

\textsuperscript{b}Laboratory of Molecular Evolution and Genome Diversity, Kunming Institute of Zoology, Chinese Academy of Sciences, Kunming 650223, China

Received 16 December 2002; received in revised form 23 January 2003; accepted 18 February 2003

Received by S. Yokoyama

Abstract

Physiological functions of human genes may be studied by gene-knockout experiments in model organisms such as the mouse. This strategy relies on the existence of one-to-one gene orthology between the human and mouse. When lineage-specific gene duplication occurs and paralogous genes share a certain degree of functional redundancy, knockout mice may not provide accurate functional information on human genes. Angiogenin is a small protein that stimulates blood-vessel growth and promotes tumor development. Humans and related primates only have one angiogenin gene, while mice have three paralogous genes. This makes it difficult to generate angiogenin-knockout mice and even more difficult to interpret the genotype-phenotype relation from such animals should they be generated. We here show that in the douc langur (\textit{Pygathrix nemaeus}), an Asian leaf-eating colobine monkey, the single-copy angiogenin gene has a one-nucleotide deletion in the sixth codon of the mature peptide, generating a premature stop codon. This nucleotide deletion is found in five unrelated individuals sequenced, and therefore is likely to have been fixed in the species. Five colobine species that are closely related to the douc langur have intact angiogenin genes, suggesting that the pseudogenization event was recent and unique to the douc langur lineage. This natural knockout experiment suggests that primate angiogenin is dispensable even in the wild. Further physiological studies of douc langurs may offer additional information on the role of this cancer-related gene in normal physiology of primates, including humans. Our findings also provide a strong case for the importance of evolutionary analysis in biomedical studies of gene functions.

\textsuperscript{\textdagger}2003 Elsevier Science B.V. All rights reserved.

Keywords: Pseudogene; Colobine; Primate; Ribonuclease; Douc langur

1. Introduction

Angiogenin (ANG) was originally isolated from human tumor cell-conditioned medium based on its ability to stimulate formation of new blood vessels (angiogenesis) (Fett et al., 1985). It was subsequently found to belong to the RNase A superfamily by sequence homology (Kurachi et al., 1985; Strydom et al., 1985) and was named RNase 5 (reviewed in Strydom, 1998). RNase A superfamily has 8 members in humans (Zhang et al., 2002a) and they exhibit drastically different physiological functions, albeit with the common ribonuclease activity (D’Alession and Riordan, 1996; Beintema and Kleineidam, 1998). ANG is the only member of the superfamily known to be angiogenic. Although the ribonuclease activity of ANG is about a million times lower than that of pancreatic RNase, the prototype of the superfamily, the ribonuclease activity is required for the angiogenic activity (Shapiro et al., 1986). The molecular mechanism of the angiogenic activity has not been fully elucidated, though a working model has been proposed (Vallee and Riordan, 1997). ANG has elevated expression in various tumors, including breast, pancreatic, gastric, colorectal, and urothelial cancers, and the angiogenic activity is related to cancer progression (Shimoyama et al., 1996, 1999; Montero et al., 1998; Miyake et al., 1999; Shimoyama and Kaminishi, 2000). Not surprisingly, ANG antagonists have the ability to inhibit cancer growth in vivo (Olson et al., 1994, 1995; Piccoli et al., 1998; Kao et al., 2003).
2. Materials and methods

2.1. DNA sequencing

The ANG gene has only one coding exon, which was amplified from genomic DNAs of six species belonging to the subfamily Colobinae of Old World monkeys by PCR with primers ANG-5 (5'-GTGGTGGAAAGATGTT-GATGCCC-3') and ANG-3 (5'-AGCACCTGAC-CAGGGCCCCTGTTA-3'). The first four codons of the ANG gene were encoded in the ANG-5 primer. The six species are the douc langur (Pygathrix nemaeus), Tonkin snub-nosed monkey (Rhinopithecus avunculus), Biet's snub-nosed monkey (Rhinopithecus bieti), golden snub-nosed monkey (Rhinopithecus roxellanae), white-side-burned black leaf monkey (Trachypithecus francoisi), and guereza (Colobus guereza). One individual per species was examined except for the douc langur, for which five animals were investigated. The PCR product was purified and sequenced from both directions as described in Zhang and Rosenberg (2000). In some cases, the PCR products were also cloned into pCR4TOPO cloning vector (Invitrogen, San Diego, CA). Multiple colonies were sequenced in order to find sequence variations within individuals as a result of either polymorphism or gene duplication.

2.2. Evolutionary analysis

ANG sequences of other primates were obtained from Zhang and Rosenberg (2002). Phylogenetic trees were reconstructed by the parsimony and neighbor-joining methods (Saitou and Nei, 1987) with 2000 bootstrap replications (Felsenstein, 1985). The MEGA2 program (Kumar et al., 2000) was used for the phylogenetic analysis. Tajima's test (Tajima, 1993) as implemented in MEGA2 was used for the test of the molecular clock hypothesis. The distance-based Bayesian method (Zhang and Nei, 1997) was used for inferring ancestral gene sequences from present-day sequences. The modified Nei–Gojobori method (Zhang et al., 1998) was used for estimating synonymous and nonsynonymous substitutions. Test of the difference in the ratio of nonsynonymous to synonymous substitutions among evolutionary lineages is performed by the Fisher's exact test as described in Zhang and Nei (1997).

3. Results

3.1. ANG pseudogene of the douc langur

We previously demonstrated that the ANG gene evolves rapidly under diversifying positive selection from a study of 11 noncolobine primates (Zhang and Rosenberg, 2002). In an expanded evolutionary survey of the gene among different primates, we inadvertently found in the douc langur (P. nemaeus) a one-nucleotide deletion in the ANG coding region. The deletion occurs at the second position of the sixth codon of the mature peptide, and shifts the open reading frame (Fig. 1). This results in the occurrence of a premature stop codon only four codons downstream of the deletion. We do not know if the douc langur ANG gene is expressed and translated. Should it be translated, the mature peptide would have only 9 amino acids, substantially shorter than the normal length of 123 amino acids. Because the ANG gene only has one coding exon, it is impossible to have an alternative splicing that may alter the reading frame. Clearly, this deletion-containing gene is nonfunctional and is a pseudogene.

We ruled out the possibility that the deletion was due to a PCR error by confirming the sequence from multiple PCR experiments. We also cloned the PCR products and sequenced ten colonies, all showing the same sequence. These results indicate that at least for the individual examined, the ANG gene is a pseudogene. In order to understand whether the pseudogene has been fixed in the
douc langur species, we examined four additional unrelated individuals. Our results show that among the five individuals, there are two allelic sequences (A and B) of the ANG gene and both sequences contain the same deletion. Thus, the pseudogene is likely to be fixed in the douc langur species, although a larger sample is needed to prove that. The two alleles identified have only one synonymous nucleotide difference (Fig. 1). Individuals #1, #3, and #5 have the genotype of AA, while #2 and #4 have the genotype of AB. The allele frequencies of A and B are therefore 0.8 and 0.2, respectively. The Hardy-Weinberg equilibrium cannot be rejected ($\chi^2 = 0.31, P > 0.5$). The presence of both AA and AB individuals strongly suggests that the A and B sequences were from two alleles rather than two paralogous genes.

3.2. ANG genes of other colobines

To understand when the pseudogenization of the ANG gene occurred, we examined five other colobine species. These include one African colobine from the genus Colobus and four Asian colobines from the genera Rhinopithecus and Trachypithecus. We found that all five species have intact ANG genes through sequencing of at least five colonies for each species. Fig. 2 shows the neighbor-joining tree of the colobine ANG genes, with the human and macaque genes used as outgroups. This tree is consistent with the current understanding of colobine systematics (Zhang and Ryder, 1998). That is, *Rhinopithecus* is closely related to and is one of the sister genera of the genus *Pygathrix*, to which the douc langur belongs (Delson, 1994; Zhang and Ryder, 1998), and *Colobus* is most distantly related to *Pygathrix* (Delson, 1994). Since all the species examined except the douc langur has an intact ANG gene, the pseudogenization event likely occurred in the douc langur lineage after its separation from other colobines (Fig. 2). The fossil record suggests that *Pygathrix* and *Rhinopithecus* separated in early Pleistocene about 2 million years (MY) ago (Delson,
1994). The limited molecular data, however, suggests a divergence of about 5 MY (Zhang and Ryder, 1998). Thus, it is safe to say that the pseudogenization event occurred within the past 5 MY. The tree in Fig. 2 also shows that the two ANG alleles of the douc langur form a statistically supported monophyletic group in exclusion of other sequences, suggesting that the two alleles originated recently within the douc langur lineage.

3.3. Variation of the rate of ANG gene evolution

It is interesting to note that in the tree of Fig. 2, the branch leading to *Pygathrix* (douc langur) is quite long, in comparison to those leading to other Asian colobines. In fact, using the African colobine *C. guereza* as an outgroup, we can reject the molecular clock hypothesis for the comparison between the sister lineages of *Pygathrix* and *Rhinopithecus* (Table 1). The rate difference between *Pygathrix* and *Trachypithecus*, however, is not statistically significant. The enhanced substitution rate in *Pygathrix*, in comparison to its sister lineage, may reflect the reduced functional constraints in the ANG pseudogene. However, this is not necessarily so, because ANG genes of noncolobine primates were previously found to evolve rapidly under positive Darwinian selection (Zhang and Rosenberg, 2002) and pseudogenization of ANG should decrease, not increase, the rate of nonsynonymous substitutions. In order to examine the substitution rate of ANG in details, we inferred the ancestral ANG gene sequences for all nodes in the tree of Fig. 2, and counted the numbers of synonymous (*s*) and nonsynonymous (*n*) substitutions for each branch of the tree (Fig. 3). Because the species are closely related, the average posterior probabilities of the inferred ancestral states are higher than 99% for all nodes and the estimated numbers of substitutions are expected to be reliable. We found that the *n/s* ratio for the set of branches that link functional ANG sequences of colobines (thick branches in Fig. 3) is 12/4 = 3, which is not significantly (*P* > 0.2) different from 4.1, a value pre-

---

**Table 1**

Tajima’s test of the molecular clock hypothesis

<table>
<thead>
<tr>
<th>Sequence 1</th>
<th>Sequence 2</th>
<th>Probability</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>Pygathrix nemaeus-a</em></td>
<td><em>Rhinopithecus bieti</em></td>
<td>0.014</td>
</tr>
<tr>
<td><em>Pygathrix nemaeus-b</em></td>
<td><em>Rhinopithecus bieti</em></td>
<td>0.008</td>
</tr>
<tr>
<td><em>Pygathrix nemaeus-a</em></td>
<td><em>Rhinopithecus avunculus</em></td>
<td>0.014</td>
</tr>
<tr>
<td><em>Pygathrix nemaeus-b</em></td>
<td><em>Rhinopithecus avunculus</em></td>
<td>0.008</td>
</tr>
<tr>
<td><em>Pygathrix nemaeus-a</em></td>
<td><em>Rhinopithecus roxellanae</em></td>
<td>0.059</td>
</tr>
<tr>
<td><em>Pygathrix nemaeus-b</em></td>
<td><em>Rhinopithecus roxellanae</em></td>
<td>0.034</td>
</tr>
<tr>
<td><em>Pygathrix nemaeus-a</em></td>
<td><em>Trachypithecus francoisi</em></td>
<td>0.317</td>
</tr>
<tr>
<td><em>Pygathrix nemaeus-b</em></td>
<td><em>Trachypithecus francoisi</em></td>
<td>0.206</td>
</tr>
</tbody>
</table>

*Colobus guereza* is used as the outgroup. In all cases, sequence 1 has a higher rate of substitution than sequence 2. The statistical significance of the rate difference is determined by the chi-square test.
viously obtained from 11 noncolobine primates (Zhang and Rosenberg, 2002). The branch leading to douc langur ANG has an \( n/s \) ratio of \( 6/1 = 6 \). But because the numbers of substitutions are small, this ratio is not significantly different from either 3 or 4.1. Thus, because the colobine species are so closely related and the pseudogenization event is so recent that a statistical comparison in the rate ratio of nonsynonymous to synonymous substitutions is not powerful enough to reveal the details of possible changes in selective pressure on the pseudogene. Nevertheless, compared to the rhesus monkey, the functional ANG genes of colobines seem to evolve more slowly. When the human sequence is used as the outgroup, this rate difference is statistically significant (\( P < 0.01 \); Tajima’s test) for the comparison between the rhesus money and any of the four Asian colobines that have intact ANG genes. The cause of this rate difference is unclear. However, it is not uncommon that the selective pressure on a gene changes over relatively short evolutionary time, as exemplified in the male-sex determination gene \( SRY \) of primates (Wang et al., 2002).

### 4. Discussion

In this study, we showed that the ANG gene of the douc langur \( P. nemaeus \) was inactivated by a one-nucleotide deletion in its coding region. We believe that this pseudogene is the sole ANG gene in the douc langur, because our extensive PCR-cloning-sequencing experiments resulted in only one gene. Furthermore, a search in the human genome sequence identified only one ANG gene (Zhang and Rosenberg, 2002; Zhang et al., 2002a) and previous phylogenetic analyses of the ANG sequences from 11 noncolobine primates (four hominoids, four Old World monkeys, and three New World monkeys) revealed no signs of gene duplication (Zhang and Rosenberg, 2002). If the douc langur had a duplicated copy of the ANG gene that originated after colobines diverged from other Old World monkeys \( \sim 15 \) MY ago (Delson, 1994), it would be quite unlikely that we could amplify the highly divergent human and New World monkey ANG genes, which separated over 40 MY ago (Kumar and Hedges, 1998), but could not amplify a gene that diverged within the past 15 MY ago. The fact that only the pseudogene was detected strongly suggests the nonexistence of functional ANG genes in the douc langur.

Our evolutionary analysis of the pseudogenization event indicated that it occurred after the separation of genus \( Pygathrix \) from its closely related sister genus \( Rhinopithecus \) within the last 5 MY. \( Pygathrix nemaeus \) is the only species recognized in the genus \( Pygathrix \), although the subspecies \( P. nigripes \) is considered by some authors to be a separate species (Nowak, 1999). Unfortunately, we have no access to genomic DNAs of \( P. nigripes \) and cannot test the presence of the ANG pseudogene in this organism. In any case, it is quite likely that the douc langur is the only extant species whose progenitors experienced this specific one-nucleotide deletion that inactivated the ANG gene. Douc langurs are found to the east of the Mekong River in central and southern Vietnam, eastern Cambodia, and central Laos (Nowak, 1999). They inhabit tropical rainforests from sea level to 2000 m. The species is classified as endangered by the IUCN (the World Conservation Union).

While the pseudogenization of the ANG gene presumably has a deleterious effect on the organism as the intact ANG gene is present in all other mammals examined, that douc langurs likely do not possess functional ANG strongly suggests that the gene is dispensable even in the wild. It is plausible that the population size of douc langurs has been small for a long time, which may have made it possible for nonlethal alleles to be fixed by genetic drift. Although one may also imagine that the fixation of the ANG pseudogene allele was driven by positive selection because of ANG’s adverse role in promoting tumor growth, there is no explicit supporting evidence for this hypothesis, as all other mammals retain functional ANG genes. Furthermore, ANG’s role in tumor growth is mainly a post-reproductive event in the mammalian life cycle. It should be mentioned that mice are known to have ANG pseudogenes in addition to three copies of functional ANG genes (Brown et al., 1995). In that case, however, the fixations of the pseudogenes were likely a neutral process because of the presence of the functional genes.

The generation of mice without the ANG genes is currently under way (G. Hu, personal communication). It would be very interesting to compare the phenotype of the knockout mice with that of the douc langur, as they represent different genetic systems with regard to the ANG gene number. To our knowledge, no specific phenotypes of douc langurs may be attributable to the inactivation of the ANG gene. But this may simply because not many physiological studies have been conducted with this species. For example, no information on the rate of tumor occurrence is known in douc langurs, which will certainly be worth knowing, given ANG’s role in tumor-growth promotion. Due to the evolutionary proximity of humans and monkeys and their common feature in having only one ANG gene, douc langurs likely yield more accurate functional information on human ANG than the knockout mice do. Furthermore, because the douc langur ANG gene was inactivated within the past 5 MY, it is unexpected that new genetic changes could have occurred to substitute the ANG function. In other words, douc langur can be regarded as an ANG-knockout animal for studying human ANG function. Captive douc langurs are available in several zoos (e.g. the San Diego Zoo) and noninvasive physiological studies seem possible.

We previously found in a study of 11 primates that ANG evolves rapidly by diversifying positive selection (Zhang and Rosenberg, 2002). Although the selective agent was not identified, we conjectured that the rapid substitutions might be related to ANG-binding proteins. In particular, ANG/
RNase inhibitor (RI) was suggested to be a likely target of selection because nearly 40% of the RI-binding sites in ANG are variable among the 11 primates. The identification of the ANG pseudogene in the douc langur provides a rare opportunity to examine the potential coevolution between ANG and RI.

As mentioned, ANG is the only known member of the primate RNase A superfamily that is angiogenic. At the protein sequence level, the closest relative to ANG is RNase 4 (Zhang et al., 2002a), and they share only 43% sequence identity. It is thus very unlikely that the RNase 4 gene could substitute ANG in the douc langur. However, it is interesting to note that the pancreatic RNase gene of the douc langur was duplicated about 4 MY ago and the duplicated gene has shifted its function to become specialized in digesting bacterial RNA (Zhang et al., 2002b). Although ANG and pancreatic RNase both belong to the RNase A gene superfamily, the inactivation of the former gene and duplication of the latter are presumably unrelated, as ANG and pancreatic RNase have very different functions and they only share about 36% sequence identity. Nevertheless, these evolutionary events demonstrate the high plasticity of the RNase A gene superfamily (Zhang et al., 2002a, 2003) and support the birth-and-death model of gene family evolution (Nei et al., 1997).

Acknowledgements

We thank Jaap Beintema, David Webb, and two anonymous reviewers for their comments on an early version of the manuscript. This work was supported by a startup fund from the University of Michigan to J.Z. and research grants from the National Institutes of Health (R01-GM67030) to J.Z., Natural Science Foundation of China to Y.Y.P.Z., and Chinese Academy of Sciences (KSCX2-1-05) to J.Z., Natural Science Foundation of China to Y.P.Z., and Chinese Academy of Sciences (KSCX2-1-05) to J.Z., Natural Science Foundation of China to Y.P.Z., and Chinese Academy of Sciences (KSCX2-1-05) to J.Z., Natural Science Foundation of China to Y.P.Z., and Chinese Academy of Sciences (KSCX2-1-05) to J.Z., Natural Science Foundation of China to Y.P.Z., and Chinese Academy of Sciences (KSCX2-1-05) to J.Z., Natural Science Foundation of China to Y.P.Z., and Chinese Academy of Sciences (KSCX2-1-05) to J.Z., Natural Science Foundation of China to Y.P.Z., and Chinese Academy of Sciences (KSCX2-1-05) to J.Z., Natural Science Foundation of China to Y.P.Z., and Chinese Academy of Sciences (KSCX2-1-05) to J.Z., Natural Science Foundation of China to Y.P.Z., and Chinese Academy of Sciences (KSCX2-1-05) to J.Z., Natural Science Foundation of China to Y.P.Z., and Chinese Academy of Sciences (KSCX2-1-05) to J.Z., Natural Science Foundation of China to Y.P.Z., and Chinese Academy of Sciences (KSCX2-1-05) to J.Z., Natural Science Foundation of China to Y.P.Z., and Chinese Academy of Sciences (KSCX2-1-05) to J.Z., Natural Science Foundation of China to Y.P.Z., and Chinese Academy of Sciences (KSCX2-1-05) to J.Z., Natural Science Foundation of China to Y.P.Z., and Chinese Academy of Sciences (KSCX2-1-05) to J.Z., Natural Science Foundation of China to Y.P.Z., and Chinese Academy of Sciences (KSCX2-1-05) to J.Z., Natural Science Foundation of China to Y.P.Z., and Chinese Academy of Sciences (KSCX2-1-05) to J.Z.,. This work was supported by a version of the manuscript. This work was supported by a version of the manuscript. This work was supported by a version of the manuscript.

References


Kumar, S., Tamura, K., Jakobsen, I., Nei, M., 2000. MEGA2, Molecular Evolutionary Genetic Analysis, Arizona State University, Tempe, AZ.


