Downloaded from genome.cshlp.org on January 1, 2012 - Published by Cold Spring Harbor Laboratory Press



Ribonuclease k6: Chromosomal Mapping and Divergent Rates of Evolution within the RNase A Gene?Superfamily

Madeleine S. Deming, Kimberly D. Dyer, Alan T. Bankier, et al.

Genome Res. 1998 8: 599-607
Access the most recent version at doi:10.1101/gr.8.6.599

References This article cites 32 articles, 17 of which can be accessed free at:
http://genome.cshlp.org/content/8/6/599.full.html#ref-list-1

Article cited in:
http://genome.cshlp.org/content/8/6/599.full.html#related-urls

Email alerting service Receive free email alerts when new articles cite this article - sign up in the box at the top right corner of the article or **click here**

To subscribe to *Genome Research* go to:
http://genome.cshlp.org/subscriptions

Cold Spring Harbor Laboratory Press

Downloaded from genome.cshlp.org on January 1, 2012 - Published by Cold Spring Harbor Laboratory Press

RESEARCH

Ribonuclease k6: Chromosomal Mapping and Divergent Rates of Evolution within the RNase A Gene Superfamily

Madeleine S. Deming,[1] Kimberly D. Dyer,[1] Alan T. Bankier,[2] Michael B. Piper,[2] Paul H. Dear,[2] and Helene F. Rosenberg[1,3]

[1]Laboratory of Host Defenses, National Institute of Allergy and Infectious Diseases, National Institutes of Health, Bethesda, Maryland 20892 USA; [2]Medical Research Council Laboratory of Molecular Biology, Protein and Nucleic Acid Chemistry Division, Cambridge CB2 2QH, UK

We have localized the gene encoding human RNase k6 to within ~120 kb on the long (q) arm of chromosome 14 by HAPPY mapping. With this information, the relative positions of the six human RNase A ribonucleases that have been mapped to this locus can be inferred. To further our understanding of the individual lineages comprising the RNase A superfamily, we have isolated and characterized 10 novel genes orthologous to that encoding human RNase k6 from Great Ape, Old World, and New World monkey genomes. Each gene encodes a complete ORF with no less than 86% amino acid sequence identity to human RNase k6 with the eight cysteines and catalytic histidines (H^{15} and H^{123}) and lysine (K^{38}) typically observed among members of the RNase A superfamily. Interesting trends include an unusually low number of synonymous substitutions (K_s) observed among the New World monkey RNase k6 genes. When considering nonsilent mutations, RNase k6 is a relatively stable lineage, with a nonsynonymous substitution rate of 0.40×10^{-9} nonsynonymous substitutions/nonsynonymous site/year (ns/ns/yr). These results stand in contrast to those determined for the primate orthologs of the two closely related ribonucleases, the eosinophil-derived neurotoxin (EDN) and eosinophil cationic protein (ECP), which have incorporated nonsilent mutations at very rapid rates (1.9×10^{-9} and 2.0×10^{-9} ns/ns/yr, respectively). The uneventful trends observed for RNase k6 serve to spotlight the unique nature of EDN and ECP and the unusual evolutionary constraints to which these two ribonuclease genes must be responding.

[The sequence data described in this paper have been submitted to the GenBank data library under accession nos. AF037081–AF037090.]

Ribonuclease k6 (RNase k6) is a recently identified member of the RNase A superfamily, a group of otherwise divergent proteins defined by specific elements of sequence homology to the prototype, Ribonuclease A. In previous work we have shown that human RNase k6 is encoded by a single-copy gene, with mRNA transcripts expressed in numerous somatic tissues (Rosenberg and Dyer 1996). Within this superfamily, RNase k6 is most closely related to the two eosinophil ribonucleases, the eosinophil-derived neurotoxin (EDN/RNase 2) and eosinophil cationic protein (ECP/RNase 3), although mRNA encoding RNase k6 was detected in peripheral blood neutrophils and monocytes, but not eosinophils. From an enzymatic perspective, recombinant RNase k6 was found to be a relatively efficient ribonuclease, with catalytic activity only ~40-fold lower than that of recombinant EDN (Rosenberg and Dyer 1995a, 1996); in contrast, the catalytic actvity of recombinant ECP was found to be ~2000-fold lower than that of recombinant EDN (Rosenberg and Dyer 1997).

The RNase A superfamily has been the subject of numerous studies that have elucidated patterns of evolution at the molecular level (Beintema et al. 1986, 1988; Benner and Allemann 1989; Fitch and Beintema 1990; Confalone et al. 1995; Jermann et al. 1995; Rosenberg et al. 1995; Trabesinger-Ruef et al. 1996). At present, this superfamily includes at least four distinct sequence lineages (Riordan and D'Alessio 1997) and two species-limited clusters (Palmieri et al. 1985; Watanabe et al. 1988; Sasso et al. 1991; Batten et al. 1997; Larson et al. 1997), as well as the two most rapidly evolving functional coding sequences known among primates (Rosenberg et al. 1995). The basis for this remarkable expenditure of evolutionary energy remains uncertain.

As part of our ongoing interest in the evolution

[3]Corresponding author.
E-MAIL hr2k@nih.gov; FAX (301) 402-4369.

Downloaded from genome.cshlp.org on January 1, 2012 - Published by Cold Spring Harbor Laboratory Press

of the RNase A gene superfamily, we have focused our attention on RNase k6; we describe here the isolation and characterization of ten nonhuman primate orthologs of the human RNase k6 gene. In contrast to the rapid rates observed for both EDN and ECP (Rosenberg et al. 1995), we have found that nonsilent mutations have been incorporated into genes of the RNase k6 lineage at a relatively conservative pace. In addition, we have determined the chromosomal localization of human RNase k6 by HAPPY mapping (Dear and Cook 1989, 1993; Walter et al. 1993; Dear 1997; Dear et al. 1997), and have used this information to position RNase k6 with respect to other members of this family along the long (q) arm of human chromosome 14.

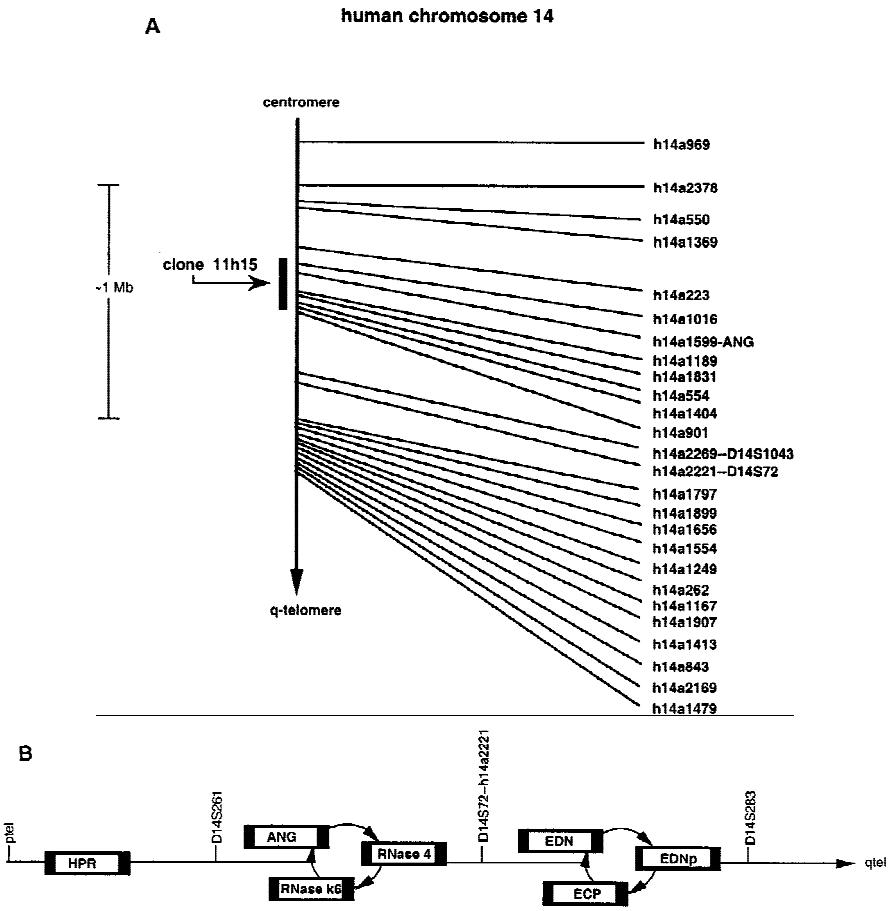


Figure 1 (*A*) Chromosomal localization of human RNase k6 by HAPPY mapping. Shown is a region of the HAPPY map of human chromosome 14 (Dear and Cook 1989, 1993; Walter et al. 1993; Dear 1997; Dear et al. 1997). The markers, STSs listed at the *right* are encoded by inter-*Alu* segments of genomic DNA; those located within one PAC distance of a previously defined marker (D14S1043, D14S72) are also noted. Clone 11h15 from the de Jong RPCI1 library (Ioannou and de Jong 1996) includes the gene encoding human RNase k6 and the five HAPPY map markers h14a1016, h14a1599, h14a1189, h14a1831, and h14a554, within a span that corresponds to the the q11 region of chromosome 14. Marker h14a1599 is found within the gene encoding angiogenin (marked ANG). The GenBank accession nos. for the STSs are as follows: h14a1016, G35789; h14a1599, G36057; h14a1189, G35876; h14a1831, G36118; h14a554, G35567. (*B*) Relative positions of the six human RNase A family members on chromosome 14. Assembling the information presented in *A* with that available from the combined maps of the NCBI. (http://www.ncbi.nlm.nih.gov), the most likely linear arrangement of these ribonucleases can be inferred; those listed in circular fashion cannot be placed in a definitive order with respect to one another with the information available. (HPR) Human pancreatic ribonuclease/RNase 1; (ANG) angiogenin/RNase 5; (RNase 4) ribonuclease 4; (RNase k6) ribonuclease k6; (EDN) eosinophil-derived neurotoxin/RNase 2; (EDNp) eosinophil-derived neurotoxin pseudogene; (ECP) eosinophil cationic protein/RNase 3. The diagrams are not drawn to scale.

RESULTS

Chromosomal Localization of RNase k6 by HAPPY Mapping

A HAPPY map of human chromosome 14, including 1001 novel sequence tagged sites (STSs), has been reported recently by Dear and colleagues (Dear and Cook 1989, 1993; Walter et al. 1993; Dear 1997; Dear et al. 1997). The novel STSs were generated by PCR amplification of genomic sequences located between *Alu* sites. These inter-*Alu* sites were mapped to ~100-kb resolution by determining their cosegregation frequency in radiation-sheared, size-selected fragments of genomic DNA. The plasmid artificial clone (PAC) 11h15, isolated from the RPCI1 human PAC library (Ioannou and de Jong 1996) was identified as containing the RNase k6 coding sequence. This clone was also found to include five specific HAPPY markers (Fig. 1A) that span ~100–120 kb within the q11 region of human chromosome 14. One of these markers, h14a1599, is

Downloaded from genome.cshlp.org on January 1, 2012 - Published by Cold Spring Harbor Laboratory Press

found within the gene encoding angiogenin, another member of the RNase A gene superfamily. Taken together, these results indicate that angiogenin and RNase k6 lie within a PAC distance (~100 kb) of one another.

With this information, together with information provided from a composite of the OMIM, RH Consortium, and Whitehead maps (http://www.ncbi.nlm.nih.gov), the relative positions of these ribonucleases with respect to one another can be inferred (Fig. 1B).

Isolation and Characterization of Primate Orthologs of RNase k6

By use of genomic DNA templates from various primate species, ~400-bp amplification products were generated with primers complementary to the human RNase k6 sequence (Rosenberg and Dyer 1996). All genomic fragments isolated encode ORFs with no less than 86% amino acid sequence identity to human RNase k6 (Table 1). As anticipated from the accepted phylogeny of these primate species, sequences within one family (e.g., Great Apes, Old World monkeys, or New World monkeys) were more similar to one another than to those of another superfamily, and increasing divergence was observed with increasing time elapsed since species divergence (Sibley and Ahlquist 1984). Also, as anticipated, the amino acid sequence of the nonprimate ortholog, RNase k2, originally isolated by Irie and colleagues (Niwata et al. 1985; Irie et al. 1988) from the bovine species *B. taurus*, differed significantly from all the primate isolates shown. Interestingly, several of the conservative amino acid changes maintained constancy within family groupings. Specifically, residue 1 was identified as tryptophan (W) in humans, Great Apes and Old World monkeys, and found to be isoleucine (I) in all three New World monkey sequences evaluated. Similar patterns were seen for V^{61} and F^{101} (Fig. 2A).

An alignment of the encoded amino acid sequences of both human and nonhuman primate RNase k6 orthologs are shown in Figure 2A. Each sequence contains the eight appropriately spaced cysteines that characterize most members of this superfamily, as well as the histidine and lysine residues analogous to those found in the catalytic site of RNase A. Similarly, all primate orthologs include the CKXXNTF sequence bracketing the catalytic ly-

Table 1. Percent Nucleotide and Amino Acid Sequence Identities: Ribonuclease k6 Genes

		Great Apes			Old World monkeys				New World monkeys		
	H. sapiens	*P. troglodytes*	*G. gorilla*	*P. pygmaeus*	*M. talapoin*	*M. milatta*	*C. aethiops*	*P. hamadryas*	*A. trivirgatus*	*S. sciureus*	*S. oedipus*
H. sapiens		*99*	*99*	*98*	*95*	*96*	*96*	*96*	*95*	*93*	*94*
P. troglodytes	98		*99*	*98*	*95*	*97*	*97*	*97*	*95*	*93*	*94*
G. gorilla	98	98		*98*	*96*	*97*	*96*	*97*	*95*	*94*	*94*
P. pygmaeus	98	98	98		*94*	*96*	*95*	*96*	*94*	*92*	*93*
M. talapoin	94	94	94	93		*98*	*98*	*90*	*93*	*91*	*91*
M. mulatta	95	95	95	95	97		*98*	*99*	*95*	*93*	*94*
C. aethiops	95	95	95	94	96	97		*98*	*94*	*93*	*93*
P. hamadryas	95	95	95	95	97	98	98		*95*	*94*	*94*
A. trivirgatus	92	92	92	91	91	92	91	92		*96*	*97*
S. sciureus	87	87	87	86	87	87	86	88	94		*96*
S. oedipus	89	89	89	88	89	92	90	91	96	91	
B. taurus	69	69	69	71	68	70	68	70	69	68	68

Species names indicated here by binomial nomenclature are listed by common name in the legend to Fig. 1. Nucleotide sequence identities are shown in italics; amino acid sequence identities, in plain text. Signal sequences were not included in the amino acid sequence analysis. Sequence identities were determined by the GAP algorithm of the Wisconsin Genetics Computer Group program on-line at the National Institutes of Health.

Downloaded from genome.cshlp.org on January 1, 2012 - Published by Cold Spring Harbor Laboratory Press

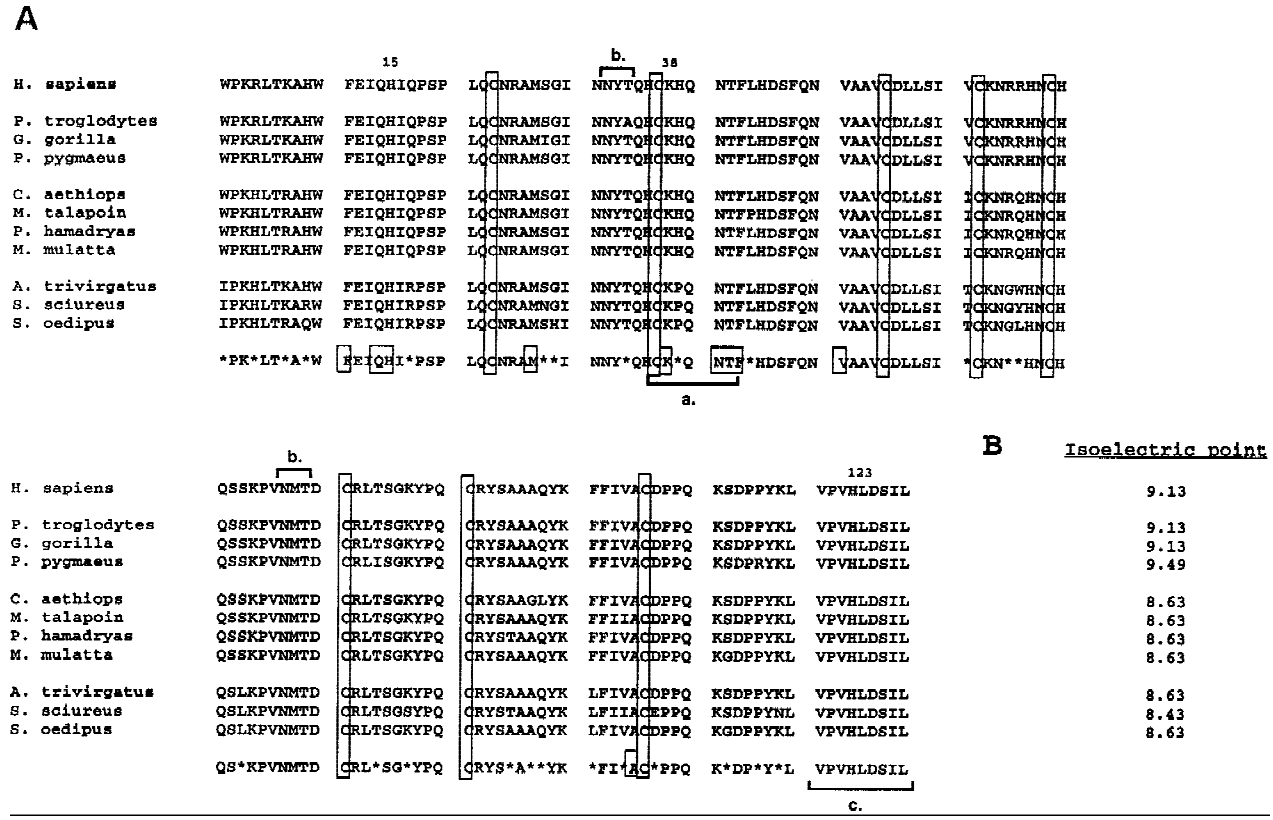


Figure 2 (*A*) Alignment of the primate orthologs of RNase k6. The DNA sequences encoding the polypeptides shown were generated by PCR from genomic DNAs as described. Each DNA sequence encodes an ORF that includes the eight canonical cysteines (shaded boxes), H^{15}, H^{123}, and K^{38} corresponding to residues found in the active site of RNase A (numbered), and the CKXXNTF motif (bracket *a*) common to all members of this superfamily. The amino acid sequence shown in the lowermost line denotes residues at which the eleven orthologs are identical (lettered) or divergent (*); residues included in boxes are common to all RNase A superfamily members (Riordan and D'Alessro 1997). Glycosylation sites initially identified in the human sequence are indicated by the brackets marked *b.* Each DNA sequence also encodes a 23 amino acid signal sequence (not shown); the residues demarcated by bracket *c* were included within the 3′ primer sequence. Common names for each genus/species include *H. sapiens,* human (GenBank accession no. U64998); *P. troglodytes,* chimpanzee (AF037081); *G. gorilla,* gorilla (AF037088); *P. pygmaeus,* orang-utan (AF037082); *C. aethiops,* African green monkey (AF037090); *M. talapoin,* talapoin (AF037087); *P. hamadryas,* baboon (AF037083); *M. mulatta,* rhesus monkey (AF037089); *A. trivirgatus,* owl monkey (AF037084); *S. sciureus,* squirrel monkey (AF037085); *S. oedipus,* cottontop tamarin (AF037086). The DNA sequence of human RNase k6 has been corrected from that originally reported (Rosenberg and Dyer 1996), resulting in the conversion of R^{53} to A. (*B*) pIs of the RNase k6 orthologs were calculated by the PEPTIDESORT algorithm of the Wisconsin Genetics Computer Group Program on-line at the National Institutes of Health.

sine (K^{38}), a motif found to be invariant among the divergent members of this superfamily (Rosenberg and Dyer 1996; Riordan and D'Alessio 1997). Of the two glycosylation sites initially identified in the human RNase k6 sequence, the first shows a limited degree of divergence, whereas the second is conserved throughout. Each sequence also includes a 23-residue signal sequence (not shown); subsequent analyses were limited to the sequence of the mature protein. The amino termini of the primate sequences were identified by comparison with the amino terminus of the bovine RNase k2 protein as determined by peptide sequencing (Niwata et al. 1985; Irie et al. 1988).

Isoelectric points (pI) calculated for each primate sequence are listed in Figure 2B. The pIs range from a low of 8.43 (for the New World monkey, *S. scuireus*) to a high of 9.49 (for the Great Ape, *P. pygmaeus*), which are all within the range typically seen for ribonucleases of this superfamily. Interestingly, the trend of increased pI in higher primates is analogous to that observed among primate or-

Downloaded from genome.cshlp.org on January 1, 2012 - Published by Cold Spring Harbor Laboratory Press

thologs of both EDN/RNase 2 and ECP/RNase 3 (Rosenberg and Dyer 1995a).

Molecular Evolution of RNase k6; Quantitative Analysis

Values for nonsynonymous substitution per nonsynonymous site (K_a) and synonymous substitution per synonymous site (K_s), comparing each ortholog with the human sequence were calculated as described by Li (1993) and are listed in Table 2. As anticipated, the values for K_a increase with increasing time elapsed since species divergence. Interestingly, the values calculated for K_s among the three New World monkeys do not diverge to the extent one might expect, resulting in elevated K_a/K_s ratios for all three New World monkey sequences.

Values obtained for both Ka and Ks for the human-New World monkey coding sequence pairs were compared with 13 similar coding sequence pairs reported previously (Rosenberg et al. 1995). Similar to results reported earlier (Graur 1985; Li et al. 1985; Ticher and Graur 1989; Wolfe and Sharp 1993; Mouchiroud et al. 1995), we observed a distinct correlation between K_a and K_s for the pairs evaluated (Fig. 3; $r^2 = 0.4$). Whereas all three points for RNase k6 pairs fell below the line, indicating relatively lower K_s values even at the given lower K_a, the degree of divergence was not statistically significant.

Table 2. Evolutionary Analysis of Primate RNase k6

Species	K_a	K_s	K_a/K_s
Great Apes			
P. troglodytes	0.005	0.017	0.29
G. gorilla	0.003	0.025	0.12
p. pygmaeus	0.009	0.069	0.13
Old World monkeys			
C. aethiops	0.015	0.080	0.19
M. talapoin	0.021	0.080	0.26
P. hamadryas	0.013	0.070	0.19
M. mulatta	0.013	0.052	0.25
New World monkeys			
A. trivirgatus	0.032	0.061	0.52
S. sciureus	0.064	0.098	0.65
S. oedipus	0.050	0.063	0.79

Nonsynonymous substitutions per nonsynonymous site (K_a) and synonymous substitutions per synonymous site (K_s) ratios of each nonhuman primate vs. human were calculated as described (Li 1993). Common names for each primate are listed in the legend to Fig. 1.

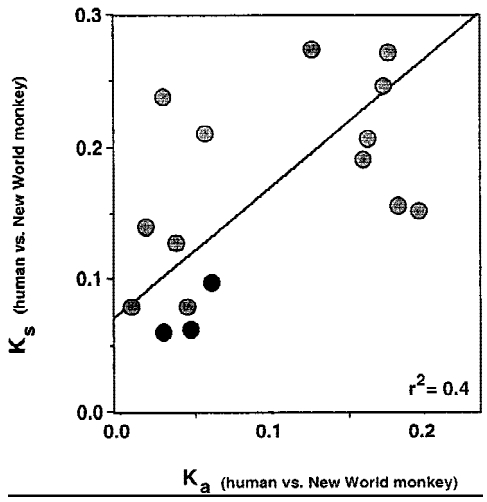


Figure 3 Synonymous substitution per synonymous site (K_s) vs. nonsynonymous substitution per nonsynonymous site (K_a) for 16 human–New World monkey coding sequence pairs. The values calculated for RNase k6 pairs are indicated with black circles. The shaded circles represent values calculated for 13 coding sequence pairs originally reported by Rosenberg et al. (1995) including ECP, EDN, interleukin-3, CD59, protamine P2, protamine P1, SRY, insulin a-chain, interferon-γ; complement C4a, insulin b-chain, c-*myc* proto-oncongene, and EGF-like growth factor. (See Rosenberg et al. 1995 for Genbank accession nos. for these sequence pairs).

Human and New World monkey RNase k6 sequences were subjected to relative-rate testing with a bovine RNase k6 ortholog as the outgroup (Table 3). In all three instances, positive values are obtained when the K_s calculated for the New World monkey-bovine pair (K_{ac}) is subtracted from that calculated for the human-bovine pair (K_{ab}), suggesting diverging patterns of synonymous substitution (see Discussion).

In Figure 4, the values for K_a are plotted against years of divergent evolution, by use of dates of species divergence determined by Sibley and Ahlquist (1984). The slope of this line, representing the rate of incorporation of nonsilent mutations, was calculated to be 0.40×10^{-9} nonsynonymous substitutions/nonsynonymous site/year (ns/ns/yr). This value can be compared with those calculated in the series reported by Rosenberg et al. (1995), which includes 63 human/non-human primate pairs from 26 independent functional coding sequences. The rate determined for RNase k6 is most similar to those reported for interferon-gamma (0.48×10^{-9} ns/ns/yr) and apolipoprotein A1 (0.42×10^{-9} ns/ns/yr), and is only slightly higher than the series average (0.37×10^{-9} ns/ns/yr).

Downloaded from genome.cshlp.org on January 1, 2012 - Published by Cold Spring Harbor Laboratory Press

Table 3. Relative Rate Analysis

Species a	Species b	K_{ac}	K_{bc}	$K_{ac}-K_{bc}$
H. sapiens		0.51 ± 0.019		
	A. trivirgatus		0.33 ± 0.011	+0.18
	S. sciureus		0.50 ± 0.013	+0.01
	S. oedipus		0.43 ± 0.010	+0.08

Differences in the ratios of synonymous substitutions per site (K_s) between human (species a) and New World monkey (species b) using the ribonuclease k6 gene from *B. taurus* (Genbank accession no. U64997) as an outgroup (species c). Values for synonymous substitution per site between the primate and bovine sequences K_{ac} and K_{bc} ± variance) were calculated according to Li (1993).

In Table 4, the rate determined for RNase k6 is compared with those determined for RNase 2 (EDN) and RNase 3 (ECP), two members of the RNase A superfamily previously shown to be incorporating nonsilent mutations more rapidly than any other functional sequences studied among primates (Rosenberg et al. 1995). Despite the fact that RNase k6 is a close relative of both EDN and ECP, the rates of sequence divergence differ dramatically (see Discussion).

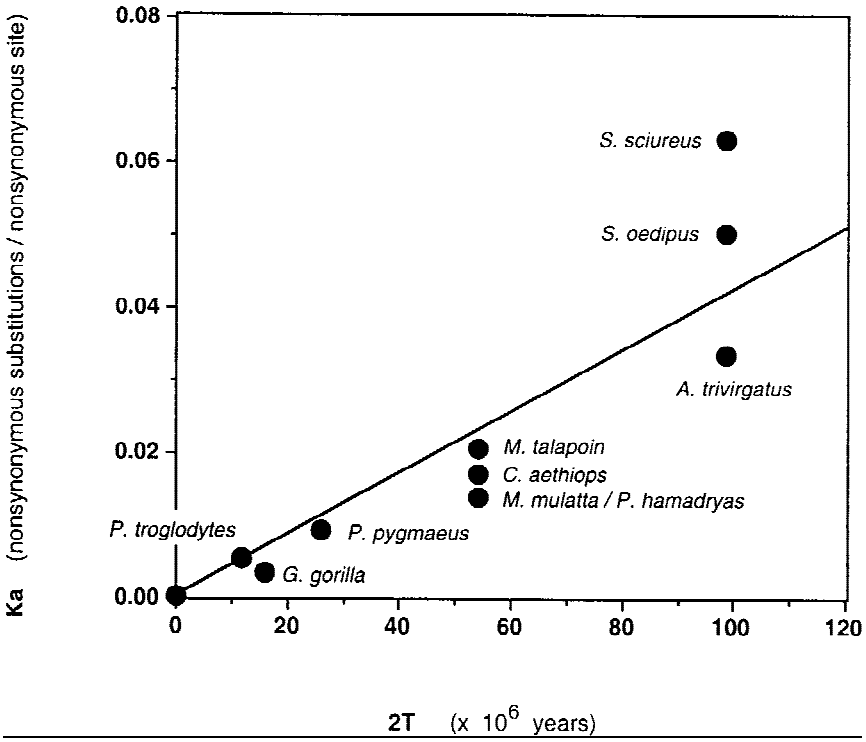


Figure 4 Nonsynonymous substitution per nonsynonymous site (K_a) vs. years of divergent evolution. The values for K_a are listed in Table 2, and the years of divergent evolution (2T) are as per Sibley and Ahlquist (1984). Common name for each genus/species are as listed in the legend to Fig. 2. Both K_a and 2T for the Old World monkeys *M. mulatta* and *P. hamadryas* are identical and are thus represented by a single point. The slope of the line representing a best fit through these data points is 0.40×10^{-9} ns/ns/yr, with a correlation coefficient (r^2) of 0.93.

DISCUSSION

As part of this work, we have isolated and characterized 10 primate orthologs of human RNase k6, a member of the RNase A superfamily (Rosenberg and Dyer 1996). The sequences encode complete ORFs, all with the appropriately spaced cysteines and catalytic histidines and lysine identifying them as members of the RNase A superfamily. The amino acid sequence similarities among these isolates also place them within the RNase k6 sequence lineage. The degree of amino acid sequence divergence among the various human/nonhuman primate RNase k6 pairs is consistent with the accepted phylogeny of these primate species. Interestingly, although the genes encoding primate orthologs of the related RNases EDN and ECP have incorporated nonsilent mutations at very rapid rates (1.9×10^{-9} and 2.0×10^{-9} ns/ns/yr, respectively), the genes encoding primate orthologs RNase k6 have remained relatively stable (0.40×10^{-9} ns/ns/yr). The fact that RNase k6 does not have a dynamic evolutionary history underscores the unique nature of the EDN/ECP lineage, and the distinct evolutionary constraints to which these two genes must be responding. These constraints have been shown to promote the development of increased toxicity and enhanced ribonuclease activity of ECP and EDN, respectively (Rosenberg and Dyer 1995a, 1997), features of particular significance to the roles played by these proteins in eosinophil-mediated host defense (Spry 1988; Makino and Fukuda 1993; Domachowske and Rosenberg 1997; Domachowske et al. 1998).

Downloaded from genome.cshlp.org on January 1, 2012 - Published by Cold Spring Harbor Laboratory Press

Table 4. Rates of Nonsynonymous and Synonymous Substitution of Three Distinct Ribonucleases of the RNase A Gene Superfamily

Coding sequence	K_a vs. 2T	r^2	K_s vs. 2T	r^2
RNase k6	0.40	0.81	1.1	0.54
EDN (RNase 2)	1.9	0.98	2.9	0.80
ECP (RNase 3)	2.0	0.92	1.7	0.97

The rates of ns/ns/yr (K_a vs. 2T) and synonymous substitution/synonymous site/year (K_s vs. 2T) for RNase k6 were calculated as described (Li 1993), with years of divergent evolution (2T) from Sibley and Ahlquist (1984). Values for K_a and K_s for EDN (RNase 2) and ECP (RNase 3) were taken from Rosenberg et al. (1995).

Whereas the rate of nonsynonymous substitution in the RNase k6 lineage is relatively conservative, several interesting features of specific primate orthologs merit comment. It is interesting to note that the New World monkey sequences (*A. trivirgatus*, *S. scuireus*, and *S. oedipus*) are more highly divergent from one another (Table 1; Fig. 2). An assessment of the relative ribonuclease activities of these proteins in recombinant form may provide some insight into the functional significance of this observation, specifically whether divergence is actively promoted or simply tolerated under the given set of evolutionary constraints. Although the K_a/K_s ratios for the New World monkey/human pairs (Table 2) are relatively high, suggesting that divergence is promoted, the unusually low values for K_s render this conclusion somewhat suspect. Additionally, it is not clear whether this degree of divergence is an idiosyncratic feature of the New World monkeys, or whether this pattern persists among more distant species; evaluation of prosimian and nonprimate mammal sequences may serve to clarify this issue.

The values calculated for synonymous substitutions per synonymous site (K_s) among the New World monkey/human RNase k6 pairs appear to be unusually low, ranging from ~30%–50% below the average (0.18 synonymous substitutions per synonymous site (ss/ss) calculated for a series of 13 independent human-New World monkey coding sequence pairs (Rosenberg et al. 1995). Although seemingly independent variables, several groups have reported correlations between Ks and Ka for interspecies coding pairs (Graur 1985; Li et al. 1985; Ticher and Graur 1989; Wolfe and Sharp 1993; Mouchiroud et al. 1995). To examine our results in this context, we compared K_a and K_s for each of the aforementioned human–New World monkey coding sequence pairs, as shown in Figure 3. Whereas all three K_s values calculated for RNase k6 pairs fall below the best-fit line, and thus are low even in the setting of low K_a, the difference does not reach statistical significance. To investigate this phenomenon further, we performed a relative rate test, using the sequence of bovine RNase k6 (Rosenberg and Dyer 1996) as an outgroup (Table 3). Interestingly, the results of relative rate analysis suggest that the K_s values for New World monkeys are unusually low, as the synonymous substitution probability calculated for the human-bovine RNase k6 coding pair (K_{ab}) exceeds that calculated for all three New World monkey-bovine pairs (K_{bc}), even before generation time is taken into account. The explanation for this finding is not readily apparent. Although synonymous substitutions are generally perceived as silent, mutations in specific sequences may introduce subtle alterations in gene structure and/or stability that are not considered in this type of analysis.

The calculated pIs for each of the RNase k6 coding sequence display a range typical for members of the RNase A superfamily, including the trend upward in higher versus lower primates analogous to that observed for the related RNases EDN and ECP (Rosenberg and Dyer 1995a). Increased cationicity has been associated with enhanced cytoxicity in the EDN/ECP lineage; whether this finding is replicated in the RNase k6 lineage remains to be evaluated.

Finally, we have used the HAPPY mapping technique (Dear and Cook 1989, 1993; Walter et al. 1993; Dear 1997; Dear et al. 1997) to localize the gene encoding RNase k6 to within ~120 kb within the q11 region of human chromosome 14. All six human members of the RNase A superfamily have been localized to chromosome 14 (Kurachi et al. 1985; Hamann et al. 1990; Rosenberg and Dyer 1995b), a configuration most likely to be the result of multiple gene duplication events. Although the evolutionary history of these RNase A family loci cannot be traced with assurance, the duplication that gave rise to distinct genes encoding EDN and ECP has been traced to the divergence of Old World from New World monkeys (Rosenberg et al. 1995). As there may be as yet undiscovered members of the RNase A superfamily, a more thorough evaluation of this region of human chromosome 14 is warranted.

METHODS

Localization of RNase k6 by HAPPY Mapping

The HAPPY map of human chromosome 14, recently described by Dear and colleagues (Dear and Cook 1989, 1993;

Downloaded from genome.cshlp.org on January 1, 2012 - Published by Cold Spring Harbor Laboratory Press

Walter et al. 1993; Dear 1997; Dear et al. 1997) includes 1001 STSs, most mapped to within 100-kb resolution on the basis of cosegregation in vitro after controlled irradiation and size selection of human genomic DNA. By use of RNase k6-specific oligonucleotide primers, 5′-CAGCCAAGTCCTCTCCAATGCAACAGGG-3′ (nucleotides 118–145, GenBank U64998) and 5′-CTGGGCAGCAGCACTATAGCGGCACTGG-3′ (nucleotides 363–336), the PAC clone 11h15 was isolated from the de Jong RPCI1 human PAC library (Ioannou and de Jong 1996), which included both the gene encoding RNase k6 and five specific HAPPY map markers (Fig. 1).

Isolation of Primate RNase k6 Orthologs by PCR

Oligonucleotide primers complementary to the human sequence included 5′-ATGGTGCTATGCTTTCCTCTTCTTTTACTG-3′ (nucleotides 1–30, GenBank U64998) and 5′-TTAGAGAATACTATCTAAGTGTACAGGAAC-3′ (nucleotides 453–424). Genomic DNAs from chimpanzee (*Pan troglodytes*), orangutan (*Pongo pygmaeus*), cottontop tamarin (*Saguinus oedipus*), owl monkey (*Aotus trivirgatus*), green monkey (*Cercopithecus aethiops*) and rhesus monkey (*Macaca mulatta*) were purchased from BIOS Labs, New Haven, CT. Genomic DNAs were also isolated from cell lines purchased from American Type Tissue Collection, Rockville, MD, including ROK, gorilla (*Gorilla gorilla*); DPSO 14474, squirrel monkey (*Saimiri sciureus*); FRhK-4, rhesus monkey (*M. mulatta*); MT1.k, talapoin (*Miopithecus talapoin*); and 26CD-1, baboon (*Papio hamadryas*). Concentrations and cycling temperatures of the individual PCR reactions were as described previously (Rosenberg et al. 1995; Batten et al. 1997). Amplification products of ~400 bp were isolated and subcloned into the PCR II TA cloning vector (Invitrogen, San Diego, CA) and subjected to automated sequencing analysis (ABI Prism, Applied Biosystems, Foster City, CA). Sequence analysis and comparisons were performed with the assistance of Sequencher (Gene Codes Corporation, Ann Arbor, MI) and ABI Prism software (Perkin Elmer, Foster City, CA) and algorithms within the Wisconsin Genetics Computer Group program available on-line at the National Institutes of Health.

Determination of Rates of Synonymous and Nonsynonymous Substitution

Sequences of the primate RNase k6 orthologs were compared with the human RNase k6 sequence, and values of synonymous (K_s) and nonsynonymous (K_a) substitution probability were calculated as described (Li 1993). Values for K_a and K_s were plotted against years of divergent evolution (Sibley and Ahlquist 1984) and slopes of the best-fitting lines determined by least squares method with correlation coefficients (r^2) included.

ACKNOWLEDGMENTS

We thank Harry L. Malech and John I. Gallin for their ongoing support of the work in our laboratory.

The publication costs of this article were defrayed in part by payment of page charges. This article must therefore be hereby marked "advertisement" in accordance with 18 USC section 1734 solely to indicate this fact.

REFERENCES

Batten, D., K.D. Dyer, J.B. Domachowske, and H.F. Rosenberg. 1997. Molecular cloning of four novel murine ribonuclease genes: Unusual expansion within the Ribonuclease A gene family. *Nucleic Acids Res.* **25:** 4235–4239.

Beintema, J.J., W.M. Fitch, and A. Carsana. 1986. Molecular evolution of pancreatic-type ribonucleases. *Mol. Biol. Evol.* **3:** 262–275.

Beintema, J.J., C. Schuller, M. Irie, and A. Carsana. 1988. Molecular evolution of the ribonuclease superfamily. *Prog. Biophys. Mol. Biol.* **51:** 165–192.

Benner, S.A. and R.K. Allemann. 1989. The return of pancreatic ribonucleases. *Trends Biochem. Sci.* **14:** 396–397.

Confalone, E., J.J. Beintema, M.P. Sasso, A. Carsana, M. Palmieri, M.T. Vento, and A. Furia. 1995. Molecular evolution of genes encoding ribonucleases in ruminant species. *J. Mol. Evol.* **41:** 850–858.

Dear, P.H. 1997. In *Genome mapping: a practical approach* (ed. P.H. Dear), pp. 95–123. IRL Press, Oxford, UK.

Dear, P.H. and P.R. Cook. 1989. Happy mapping: A proposal for linkage mapping the human genome. *Nucleic Acids Res.* **17:** 6795–6807.

———. 1993. Happy mapping: Linkage mapping using a physical analogue of meiosis. *Nucleic Acids Res.* **21:** 13–20.

Dear, P.H., A.T. Bankier, and M.B. Piper. 1997. A high resolution metric HAPPY map of human chromosome 14. *Genomics* **48:** 232–241.

Domachowske, J.B. and H.F. Rosenberg. 1997. Eosinophils inhibit retroviral transduction of human target cells by a ribonuclease-dependent mechanism. *J. Leukoc. Biol.* **62:** 363–368.

Domachowske, J.B., K.D. Dyer, C.A. Bonville, and H.F. Rosenberg. 1998. Recombinant human eosinophil-derived neurotoxin/RNase 2 functions as an effective antiviral agent against respiratory syncytial virus. *J. Infect. Dis.* **177:** (in press).

Fitch, W.M. and J.J. Beintema. 1990. Correcting parsimonious trees for unseen nucleotide substitutions: The effect of dense branching as exemplified by ribonuclease. *Mol. Biol. Evol.* **7:** 438–443.

Graur, D. 1985. Amino acid composition and the evolutionary rates of protein-coding genes. *J. Mol. Evol.* **22:** 53–62.

Hamann, K.J., R.M. Ten, D.A. Loegering, R.B. Jenkins, M.T. Heise, C.R. Schad, L.R. Pease, G.J. Gleich, and R.L. Barker. 1990. Structure and chromosome localizaton of the human eosinophil-derived neurotoxin and eosinophil cationic protein genes: Evidence for intronless coding sequences in the ribonuclease gene family. *Genomics* **7:** 535–546.

Downloaded from genome.cshlp.org on January 1, 2012 - Published by Cold Spring Harbor Laboratory Press

Ioannou, P.A. and P.J. de Jong. 1996. Construction of bacterial artificial chromosome libraries using the modified P1 (PAC) system. In *Current protocols in human genetics* (ed. N.C. Dracopoli), vol. 5, pp. 1–24.

Irie, M., R. Nitta, K. Ohgi, Y. Niwata, H. Watanabe, M. Iwama, J.J. Beintema, A. Sanda, and Y. Takizawa. 1988. Primary structure of a non-secretory ribonuclease from bovine kidney. *J. Biochem.* **104**: 289–296.

Jermann, T.M., J.G. Opitz, J. Stackhouse, and S.A. Benner. 1995. Reconstructing the evolutionary history of the artiodactyl ribonuclease superfamily. *Nature* **374**: 57–59.

Kurachi, K., E.W. Davie, D.J. Strydom, J.F. Riordan, and B.L. Vallee. 1985. Sequence of the cDNA and gene for angiogenin, a human angiogenesis factor. *Biochemistry* **24**: 5494–5499.

Larson, K.A., E.V. Olson, B.J. Madden, G.J. Gleich, N.A. Lee, and J.J. Lee. 1997. Two highly homologous ribonuclease genes expressed in mouse eosinophils identify a larger subgroup of the mammalian ribonuclease superfamily. *Proc. Natl. Acad. Sci.* **93**: 12370–12375.

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

Li, W.-H., C.-I. Wu, and C.-C. Luo. 1985. A new method for estimating synonymous and nonsynonymous rates of nucleotide substitution considering the relative likelihood of nucleotide and codon changes. *Mol. Biol. Evol.* **2**: 150–174.

Makino, S. and T. Fukuda. 1993. *Eosinophils: biological and clinical aspects.* CRC Press, Boca Raton, FL.

Mouchiroud, D., C. Gautier, and G. Bernardi. 1995. Frequencies of synonymous substitutions in mammals are gene-specific and correlated with frequencies of nonsynonymous substitutions. *J. Mol. Evol.* **40**: 107–113.

Niwata, Y., K. Ohgi, A. Sanda, A. Takizawa, and M. Irie. 1985. Purification and properties of bovine kidney ribonucleases. *J. Biochem.* **97**: 923–934.

Palmieri, M., A. Carsana, A. Furia, and M. Libonati. 1985. Sequence analysis of a cloned cDNA for bovine seminal ribonuclease. *Eur. J. Biochem.* **152**: 275–277.

Riordan, J.F. and G. D'Alessio. 1997. *Ribonucleases: structure and function.* Academic Press, New York, NY.

Rosenberg, H.F. and K.D. Dyer. 1995a. Eosinophil cationic protein and eosinophil-derived neurotoxin: Evolution of novel function in a primate ribonuclease gene family. *J. Biol. Chem.* **270**: 21539–21544.

———. 1995b. Human ribonuclease 4 (RNase 4): Coding sequence, chromosomal localization, and identification of two distinct transcripts in human somatic tissues. *Nucleic Acids Res.* **23**: 4290–4295.

———. 1996. Molecular cloning and characterization of a novel human ribonuclease (RNase k6): Increasing diverstiy in the enlarging ribonuclease gene family. *Nucleic Acids Res.* **24**: 3507–3513.

———. 1997. Diversity among the primate eosinophil-derived neurotoxin genes: A specific carboxy-terminal sequence is necessary for enhanced ribonuclease activity. *Nucleic Acids Res.* **25**: 3532–3536.

Rosenberg, H.F., K.D. Dyer, H.L. Tiffany, and M. Gonzalez. 1995. Rapid evolution of a unique family of primate ribonuclease genes. *Nature Genet.* **10**: 219–223.

Sasso, M.P., A. Carsana, E. Confalone, C. Cosi, S. Sorrentino, M. Viola, M. Palmieri, E. Russo, and A. Furia. 1991. Molecular cloning of the gene encoding bovine brain ribonuclease and its expression in different regions of the brain. *Nucleic Acids Res.* **19**: 6469–6474.

Sibley, C.G. and J.E. Ahlquist. 1984. The phylogeny of the hominoid primates as indicated by DNA-DNA hybridization. *J. Mol. Evol.* **20**: 2–15.

Spry, C.J.F. 1988. *Eosinophils: A comprehensive review and guide to the scientific and medical literature.* Oxford University Press, Oxford, UK.

Trabesinger-Ruef, N., T. Jermann, T. Zankel, B. Durrant, G. Frank, and S.A. Benner. 1996. Pseudogenes in ribonuclease evolution: A source of new biomacromolecular function? *FEBS Lett.* **382**: 319–322.

Ticher, A. and D. Graur. 1989. Nucleic acid composition, codon usage, and the rate of synonymous substitution in protein-coding genes. *J. Mol. Evol.* **28**: 286–298.

Walter, G., I.M. Tomlinson, G.P. Cook, G. Winter, T.H. Rabbitts, and P.H. Dear. 1993. HAPPY mapping of a YAC reveals alternative haplotypes in the human immunoglobulin VH locus. *Nucleic Acids Res.* **21**: 4524–4529.

Watanabe, H., H. Katoh, M. Ishii, Y. Komoda, A. Sanda, Y. Takizawa, K. Ohgi, and M. Irie. 1988. Primary structure of a ribonuclease from bovine brain. *J. Biochem.* **104**: 939–945.

Wolfe, K.H. and P.M. Sharp. 1993. Mammalian gene evolution: Nucleotide sequence divergence between mouse and rat. *J. Mol. Evol.* **37**: 441–456.

Received January 7, 1998; accepted in revised form April 22, 1998.