J Mol Evol (2002) 54:102–106
DOI: 10.1007/s00239-001-0022-0

JOURNAL OF MOLECULAR EVOLUTION

Age and Detection of Retroprocessed Pseudogenes in Murine Rodents

Felix Friedberg, Allen R. Rhoads

Department of Biochemistry and Molecular Biology, Howard University, College of Medicine, Washington, DC 20059, USA

Received: 23 May 2001 / Accepted: 31 May 2001

Abstract. Retroprocessed pseudogenes, calmodulin II (ψ_1, ψ_2, and ψ_3 CALMII), ψ α-tubulin, π-glutathione *S*-transferase (ψ π-GST) from rat, lactic acid dehydrogenase (ψ LDH) from mouse, and heat shock protein 60 chaperonin (ψ HSP60) from Chinese hamster, were examined for their presence in these species by polymerase chain reaction (PCR). Pseudogenes of these murine rodents were detected by PCR only in those species in which the genes were originally identified, suggesting that the selected pseudogene of one species arose too recently to be detected in the genomes of the other rodent species. The calculated ages of the rodent pseudogenes ranged from 1.7 Myr (ψ α-tubulin) to 7.5 Myr (ψ_3 CALMII) when employing a homologous functional gene of the taxon as a reference in the relative rate test with the mouse or rat as the outgroup. Given the high rate of divergence of the genes of rodents relative to other species, selection of an outgroup with similar mutation rates seems warranted. To justify further the conclusion that the selected pseudogenes were indeed retroprocessed after these three taxa diverged, the presence of the pseudogenes in the genome of different rat species was examined. The existence of ψ_3 CALMII and ψ α-tubulin pseudogenes of *Rattus norvegicus* among species belonging to *Rattus sensu stricto* is evidence for the common ancestry of this group.

Key words: Retroprocessed pseudogenes — Murine rodents — Age — Calmodulin II — α-Tubulin

Correspondence to: Felix Friedberg; *email:* ffriedberg@howard.edu

Introduction

Recently, we examined retroprocessed pseudogenes in the genome of certain members of the primate order by comparing the data calculated utilizing the molecular clock theory with the actual detection of these pseudogenes by molecular probe analysis employing the polymerase chain reaction (PCR). It was found that the order of evolution as conventionally accepted for these primates was in good agreement when the calculated ages were compared to the observed detection of the genes by PCR analysis (Friedberg and Rhoads 2000). We suggested that such studies, by providing species-specific markers, might ultimately help to ascertain minimum divergence times for related taxa, i.e., the presence or absence of a given pseudogene might provide justification for the postulated branching of phylogenetic trees. Here, we report the results obtained when the same methodology of pseudogene neutral rate analysis was applied to the genomes of the mouse, rat, and hamster as well as to selected taxa of *Rattus sensu stricto* (Verneau et al. 1998).

Materials and Methods

Rodent genomic DNA was prepared from the livers of *Rattus norvegicus, Mus musculus,* and *Mesocricetus auratus* by a standard protocol involving osmotic shock, proteinase K digestion, phenol extraction followed by precipitation, and washing with ethanol (Sambrook et al. 1989). The DNA was judged pure if the A_{260}/A_{280} ratio was >1.75. The DNA for the various members of *Rattus sensu stricto* and *Rattus sensu lato* was generously provided by Anthony Furano and Francois Catezflis and was from the collection of the Institut des Sciences de L'Evolution (Montpellier, France). Detection of pseudogenes in the DNA from the various members of Rodentia was performed by PCR

Table 1. PCR primers used to amplify unique regions of rodent pseudogenes

Gene name, accession no. (reference)[a]	Forward primer (5′ → 3′)	Reverse primer (5′ → 3′)	Expected product size (bp)
Rat ψ_1 CALMII, X13813 (Nojima 1989)	TCCTTAGTCCTAGTGGAGCT	CCTCATCTGCTGCACTAATAG	387
Rat ψ_2 CALMII, M17068 (Nojima et al. 1987)	ACCACTCTCTTCTCTAGCATC	CTCCTTGGTTGTTACTGTCTG	816
Rat ψ_3 CALMII, X14264 (Nojima 1989)	GTGACTGAGGCTCGAATTAG	CTCTGAAGAGTTTAGCTCACC	549
Rat ψ α-tubulin,[b] J00799	AGAGATCTGGCGCCCTCTG	CTGTCCAAGGTCAATGATC	561
(Lemischka and Sharp 1982)	CAGCTCGCTCGACAATTTC		605
Rat ψ π-GST,[b] J02690 (Okuda et al. 1987)	GTGGGAATATGGATACGACC	ATCCACTAGGGCGGCCTTCC	295
	GGAAGGCCGCCCTAGTGGA	CCTAGGCCTACTGGGCAAGC	541
Mouse ψ LDH, M31035 (Fukasawa et al. 1987)	GGTCGAGAGCATAATGAACC	GAACAATCCTCTCTTCTCTC	748
Hamster ψ HSP60, M34562 (Venner et al. 1990)	CTCACTCTCCGCCACCAGCC	CTATGAGCAGTAATGGGCTC	750

[a] The accession numbers for the functional rat and mouse genes used for comparison were M19312 and M27844 for CALMII, V01227 and M13441 for α-tubulin, X01964 and Y00309 for LDH-A, X02904 and X53451 for π-GST, and X53585 and X533584 for HSP60, respectively. The functional hamster and mouse genes for HSP60 were M22383 and X53584, respectively.

[b] Multiple sets of primers were employed for the detection of this pseudogene. Both pairs yielded the same results as listed in Table 3.

using the primers listed in Table 1. In all instances a functional gene [a 3′-untranslated region of calmodulin III (CALMIII)] and a PCR reaction mixture excluding template were employed as controls. Selective primers (19- to 21-mer) were designed using unique regions of the pseudogenes compared to bona fide genes as indicated in Fig. 1 and synthesized by Bioserve Biotechnologies (Laurel, MD). The T_m of these primers ranged from 56 to 64°C, and the amplified regions ranged from 295 to 816 bp in length. PCR was performed in 1× PCR buffer (PCR Master Mix; Qiagen Inc.). Amplification was performed in a Gene-Amp PCR System 2400 (Perkin–Elmer), with an initial denaturation at 94°C for 5 min, followed by 30 cycles of denaturation at 94°C for 30 s, annealing at 62°C for 30 s, and extension at 72°C for 1 min, with a final extension at 72°C for 5 min. Two PCR reactions were run in tandem to enhance the sensitivity of detection and reduce template requirements. The first reaction utilized 10 μl of primary reaction mixture followed by the addition of fresh primers and PCR Master Mix to give a secondary reaction volume of 30 μl. Low-stringency PCR conditions were employed using the same initial denaturation followed by 35 cycles of denaturation for 30 s, annealing at 54°C for 30 s, and extension at 72°C for 1 min, with a final extension at 72°C for 5 min, only to confirm negative results obtained with tandem PCR. Electrophoresis was performed on 1.5% agarose in 40 m*M* Tris–HCl, 20 m*M* sodium acetate, and 1 m*M* EDTA, pH 7.8, and 0.0001% ethidium bromide. For further product identification, bands were extracted in TE buffer and a second PCR was performed, followed by removal of primers and nucleotides using a PCR Select-II column (5 Prime → 3 Prime, Inc., Boulder, CO) and dye terminator cycle sequencing of the isolated DNA (Veritas Inc., Gaithersburg, MD). Multiple sequence alignment of homologous coding and sequenced regions was performed by CLUSTAL analysis (Thompson et al. 1994) and by the Pileup program of the Genetic Computer Group (GCG), Madison, WI. The pairwise number of synonymous and nonsynonymous substitutions per site was determined by the Diverge program (GCG) using the Pileup sequence alignment output. This estimate of the number of synonymous and nonsynonymous substititutions is based on the method of Li (1993) as modified by Pamilo and Bianchi (1993) using Kimura's (1980) two-parameter method. Gaps or indels were not considered in the determination of the substitutions (Li et al. 1981).

Results

The evolutionary distances calculated from the corrected rates of substitution of the selected pseudogenes and the corresponding functional genes from the mouse, rat, and Chinese hamster are listed in Table 2. Utilizing these rates, the ages of divergence were calculated for the pseudogenes; they ranged from 1.7 to 7.5 Myr and were well within the estimated divergence time of 11 Myr between rat and mouse (Catzeflis et al. 1987) or of 17 Myr between rat and hamster estimated by DNA hybridization studies (Catzeflis et al. 1993). When examined by PCR, all of the selected pseudogenes were detectable only in the species where the pseudogene was first identified, i.e., the rat pseudogene was present only in DNA from rat, and not in DNA from mouse or hamster (Table 3). This finding is in agreement with the estimated ages of the pseudogenes that were calculated based on comparisons with homologous functional genes in rodents (Table 2).

When the presence of rat ψ_3 CALMII and ψ α-tubulin was probed by PCR analysis of DNA isolated from various members of *Rattus sensu stricto,* the results depicted in Table 4 were obtained. These two pseudogenes tested positive for the examined *Rattus sensu stricto* species with the exceptions of *Sundamys muelleri* and *Berylmys bowersi.*

Discussion

When examining the distribution of pseudogenes in primates, we indicated a potential caveat to calculating the age of divergence of pseudogenes using bona fide genes selected from the human and the rat as a frame of reference in estimating the rate of nucleotide substitution for the functional genes. The molecular clock is based on the assumption that the same rates of mutation apply across ages as well as lineages (Zuckerkandl and Pauling 1965). Catzeflis et al. (1987) suggested that the rate of muroid rodent DNA evolution is about 10 times higher than that of hominoid primates, and Li et al. (1987), using statistical analysis, estimated the rate to be 4–8 times higher in

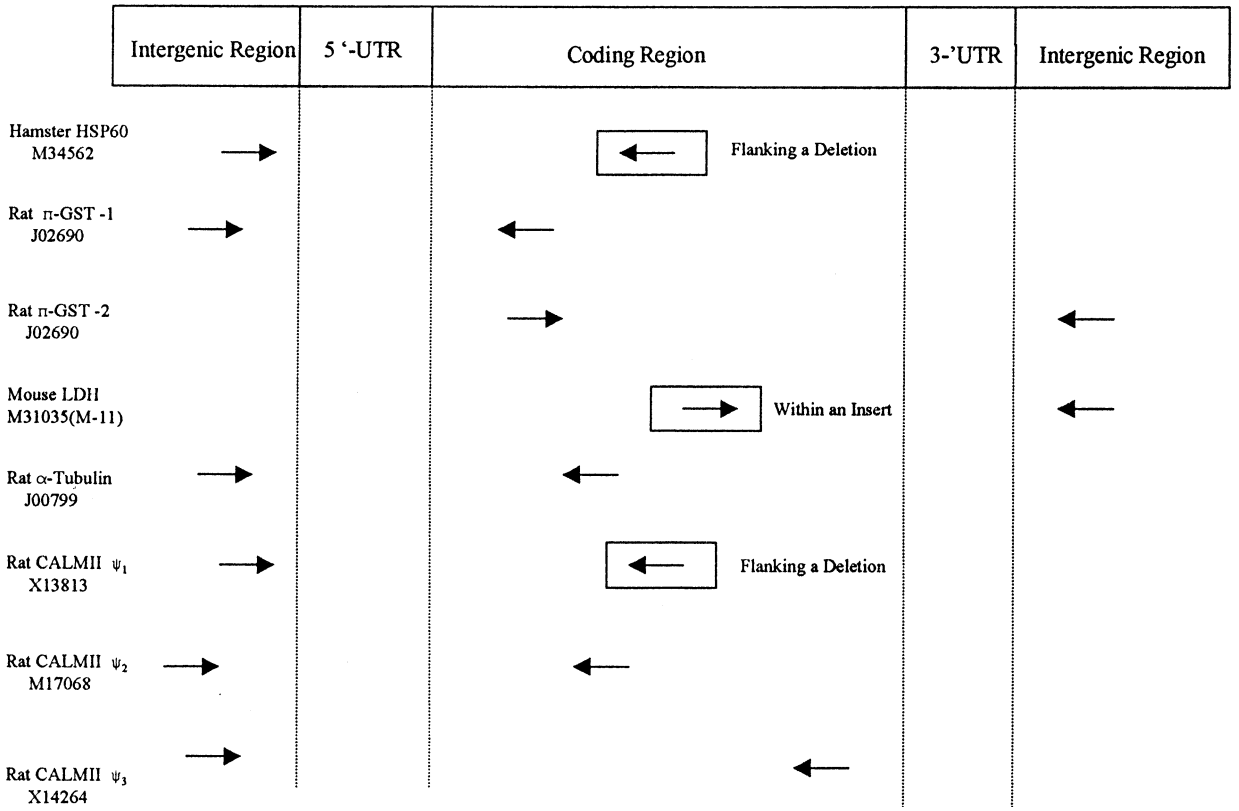

Fig. 1. Primer design—rodent pseudogenes.

rodents than in primates. Certainly, no global molecular clock seems to exist in mammals, and evolutionary time appears to be relative within different lineages. For example, the gene that encodes the copper/zinc superoxide dismutase (SOD) has a variable rate of evolution depending on the group of organisms measured (Ayala 1997; Li and Tarimura 1987). This caveat is strengthened when the data obtained in this study for rodents are considered (Table 2). In all instances (rat, mouse, and hamster) PCR data were positive only in the species wherein the original pseudogene had been shown to exist, but the pseudogenes were not detected in the other two species of rodent. For the three CALMII-derived pseudogenes of rat with calculated ages of 2.2 to 7.5 Myr based on the mouse as an outgroup, no homologous PCR products could be detected in the genome of the mouse or the hamster. These ages were well within the generally accepted divergence times for these rodents (11–17 Myr). By employing the mouse as the outgroup with each of the seven selected pseudogenes, divergence times ranging from 1.7 to 7.5 Myr were calculated. In the case of α-tubulin, where the sequences of functional genes were available for the mouse, the rat, and the hamster, the calculated divergence times for rat ψ α-tubulin based on these rat/mouse and rat/hamster comparisons were in good agreement, 1.7 and 2.1 Myr, respectively. This suggests that an acceptable reference for calculating the ages of pseudogenes in rodents would be obtained by the comparison of functional genes from other rodents when applying the relative rate test of Sarich and Wilson (1973) (i.e., the rate within murine rodents is more consistent). This study supports the molecular clock theory when applied to lineages having nearly the same mutational rate. While the selected retroprocessed pseudogenes cannot be employed to verify the ages of divergence of distantly related rodents because of their relatively recent origin, they allow a minimum estimate of such ages of divergence and serve as species-specific markers.

Neutrally evolving pseudogenes may additionally permit the establishment of phylogenetic relationships within a genus (Tables 3 and 4). The results demonstrating that the ψ_3 CALMII pseudogene (7.5 Myr) exists in most of the rat species examined in this study are in agreement with the postulated emergence of the common ancestor of the *Rattus sensu stricto* lineage at not less than 7.5 Myr (Verneau et al. 1998). The appearance of the youngest pseudogene (ψ α-tubulin, 1.7 Myr) would predate the divergence of the examined members of the *Rattus sensu stricto* lineage. From the sequence data, unique sequence regions not shared by the functional gene confirmed that we identified the claimed pseudogenes. The exceptionally low percentage of substitutions among these selected taxa indicates a very close genetic relationship and precludes a reliable phylogenetic analysis based on these data. In contrast to the conclusions of Verneau et al. (1998), however, we suggest that *B. bowersi* and *S. muelleri* belong to the *Rattus sensu lato* rather

Table 2. Rat and mouse pseudogene divergence[a]

Comparison gene	Versus mouse functional gene		Versus rat functional gene		T_n (Myr)
	K_s	K_a	K_s	K_a	
Rat CALMII	23.0	0.0	—	—	—
Rat ψ_1 CALMII	16.4	11.6	17.1	11.4	5.7
Rat ψ_2 CALMII	23.8	6.48	9.88	6.35	2.2
Rat ψ_3 CALMII	18.3	9.39	26.5	9.40	7.5
Rat α-tubulin	13.9	1.56	—	—	—
Rat ψ α-tubulin	14.9	2.83	5.33	1.29	1.7[b]
Mouse LDH	—	—	25.4	1.43	—
Mouse ψ LDH	11.4	4.28	25.8	5.68	2.4
Rat π-GST	16.7	4.10	—	—	—
Rat ψ π-GST	15.4	6.05	12.9	4.3	4.7
Hamster HSP60	19.7	0.53	16.8	0.46	—
Hamster ψ HSP60	19.6	8.88	18.8	8.84	3.5[c]

[a] T_n (time since retroposition) values were calculated utilizing the molecular clock as suggested by Li (1997). K_s and K_a are the number of synonymous substitutions per 100 synonymous sites and the number of nonsynonymous substitutions per 100 nonsynonymous sites, respectively. The actual number of nucleotide substitutions was estimated by the two-parameter method of Kimura (1980). Mouse/rat and rat/hamster reference divergence times of 11 and 17 Myr, respectively, were used to calculate the T_n. The age of the rat pseudogenes was calculated using mouse as a reference for the divergence of the functional gene.

[b] An age of 2.1 Myr for rat ψ α-tubulin was obtained using Chinese hamster instead of mouse as a reference.

[c] The K_s between hamster HSP60 and hamster ψ HSP60 was 9.00. An age of 3.9 Myr for hamster ψ HSP60 was obtained using rat instead of mouse as a reference.

Table 3. PCR detection of pseudogenes in different rodents

Gene	Species		
	Rat	Mouse, 11 Myr[a]	Hamster, 17 Myr
Rat α-tubulin ψ, 1.7 Myr[b]	+[c]	–[d]	–
Rat CLAMII ψ_2, 2.2 Myr	+	–	–
Mouse LDH ψ, 2.4 Myr	–	+	–
Hamster HSP60 ψ, 3.1 Myr	–	–	+
Rat π-GST ψ, 4.7 Myr	+	–	–
Rat CALMII ψ_1, 5.7 Myr	+	–	–
Rat CALMII ψ_3, 7.5 Myr	+	–	–

[a] Divergence times for mouse and hamster are taken from Catzeflis et al. (1993).

[b] Calculated divergence times from Table 2.

[c] A plus sign indicates detection of a single intense band of the appropriately sized fragment upon electrophoresis of the PCR products.

[d] A minus sign indicates no detection of any product under either PCR condition.

than the *Rattus sensu stricto* group since they tested negative for this pseudogene. Negative results may be caused by many factors, but *B. bowersi* and *S. muelleri* gave additional confirmatory negative results when tested for the existence of the ψ_2 CALMII pseudogene,

Table 4. Sequence identification and substitutions of rat pseudogenes[a]

Species	Substitutions/100 residues	
	Rat ψ_3 CALMII (7.5 Myr)	Rat ψ α-tubulin (1.7 Myr)
Rattus exulans	3.1	3.1[b]
Rattus moluccarius	2.4	3.0
Rattus satarae	1.4	6.0[b]
Rattus fuscipes	2.6	ND[c]
Niviventer fulvenscens	0.2	ND
Sundamys muelleri	–[d]	–
Berylmys bowersi	–	–

[a] Isolated PCR products were sequenced using the same primer set in both the 5′ and the 3′ directions except where indicated utilizing primers listed in Table 1. The sequences were aligned and compared with the *R. norvegicus* pseudogene to determine substitutions/100 residues. Sequenced regions ranged from 414 to 495 nucleotides in length.

[b] The sequence of the PCR product was determined in only the 5′ or 3′ direction.

[c] Not determined due to either a mixed or a low signal for sequencing reaction. Additional sequencing primers may be required.

[d] A minus sign indicates no detection of any product under either PCR condition.

with a calculated retropositioning time of 2.2 Myr (data not shown). Members of the *Rattus senso stricto* group would be expected to test positive for the three CALMII pseudogenes given their rather recent divergence from the common ancestor.

Acknowledgment. This work was partially supported by Grant MCB930015P from the Pittsburgh Supercomputing Center, Pittsburgh, Pennsylvania.

References

Ayala FJ (1997) Vagaries of the molecular clock. Proc Natl Acad Sci USA 94:7776–7783

Catzeflis FM, Sheldon FH, Alquist JE, Sibley CG (1987) DNA-DNA hybridization evidence of the rapid rate of muroid rodent DNA evolution. Mol Biol Evol 4:242–253

Catzeflis FM, Dickerman AW, Michaux J, Kirsch JAW (1993) In: Szalay FS, Novacek MJ, McKenna MC (eds) Mammalian phylogeny (placentals). Springer Verlag, New York, pp 159–172

Friedberg F, Rhoads AR (2000) Calculation and verification of the ages of retroprocessed pseudogenes. Mol Phylogenet Evol 16:127–130

Fukasawa KM, Tanimura M, Sakai I, Sharief FS, Chung F-Z, Li SS-L (1987) Molecular nature of spontaneous mutations in mouse lactate dehydrogenase-A processed pseudogenes. Genetics 115:177–184

Kimura M (1980) A simple method for estimating evolutionary rates of base substitutions through comparative studies of nucleotide sequences. J Mol Evol 16:111–120

Lemischka I, Sharp PA (1982) The sequences of an expressed rat α-tubulin gene and a pseudogene with an inserted repetitive element. Nature 300:330–335

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

Li W-H (1997) Molecular evolution. Sinauer, Sunderland, MA

Li W-H, Tanimura M (1987) The molecular clock runs more slowly in man than in apes and monkeys. Nature 326:93–96

Li W-H, Gojobori T, Nei M (1981) Pseudogenes as a paradigm of neutral evolution. Nature 292:237–239

Li W-H, Tanimura M, Sharp PM (1987) An evaluation of the molecular clock hypothesis using mammalian DNA sequences. J Mol Evol 25:330–342

Nojima H (1989) Structural organization of multiple rat calmodulin genes. J Mol Biol 208:269–282

Nojima H, Kishi K, Sokabe H (1987) Multiple calmodulin mRNA species are derived from two distinct genes. Mol Cell Biol 7:1873–1880

Okuda A, Sakai M, Muramatsu M (1987) The structure of the rat glutathione S-transferase P gene and related pseudogenes. J Biol Chem 262:3858–3863

Pamilo P, Bianchi NO (1993) Evolution of the Zfx and Zfy genes: Rates and interdependence between the genes. Mol Biol Evol 10: 271–281

Sambrook J, Fritsch EF, Maniatis T (1989) Molecular cloning—A laboratory manual. Cold Spring Harbor Laboratory, Cold Spring Harbor, NY

Sarich VM, Wilson AC (1973) Generation time and genomic evolution in primates. Science 179:1144–1147

Thompson JD, Higgins DG, Gibson TJ (1994) CLUSTALW: Improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. Nucleic Acids Res 22:4673–4680

Venner TJ, Singh B, Gupta RS (1990) Nucleotide sequences and novel structural features of human and Chinese hamster hsp60 (chaperonin) gene families. DNA Cell Biol 9:545–552

Verneau O, Catzeflis F, Furano AV (1998) Determining and dating recent rodent speciation events by using L1 (LINE-1) retrotransposons. Proc Natl Acad Sci USA 95:11284–11289

Zuckerkandl E, Pauling L (1965) Evolutionary divergence and convergence in proteins. In: V Bryson, HJ Vogel (eds) Evolving genes and proteins. Academic Press, New York, pp 97–166