Evolution of Duplicate Genes in a Tetraploid Animal, *Xenopus laevis*[1]

Marianne K. Hughes and Austin L. Hughes

Department of Biology and Institute of Molecular Evolutionary Genetics, The Pennsylvania State University

To understand the evolution of duplicate genes, we compared rates of nucleotide substitution between 17 pairs of nonallelic duplicated genes in the tetraploid frog *Xenopus laevis* with rates between the orthologous loci of human and rodent. For all duplicated *X. laevis* genes, the number of synonymous substitutions per site (d_S) was greater than the number of nonsynonymous substitutions per site (d_N), indicating that these genes are subject to purifying selection. There was also a significant positive correlation ($r = 0.915$) between d_N for the *X. laevis* genes and d_N for the mammalian genes, suggesting that, at the amino acid level, the *X. laevis* genes and the mammalian genes are under similar constraints. Results of relative-rate tests showed nearly equal rates of nonsynonymous substitution in each copy of the *X. laevis* genes; apparently there are similar constraints on both copies. No correlation was found between d_S for the *X. laevis* genes and d_S for the mammalian genes. There was a significant positive correlation both between members of pairs of duplicated *X. laevis* genes ($r = 0.951$) and between human and rodent orthologues ($r = 0.854$) with respect to third-position G + C content but no such relationship between the *X. laevis* genes and either of their mammalian orthologues. The results indicate that both copies of a duplicate gene can be subject to purifying selection and thus support the hypothesis of selection against all genotypes containing a null allele at either of two duplicate loci.

Introduction

Ohno (1970) emphasized the evolutionary importance of gene duplication: the presence of one functional locus permits the duplicated copy both to accumulate mutations that would otherwise have been eliminated by selection and, eventually, to acquire new function. Kimura and Ohta (1974) concluded that "gene duplication must always precede the evolution of a gene having new function" (p. 2850). However, the more likely outcome of a duplication event is that one copy becomes nonfunctional, as deleterious mutations accumulate that eventually silence the redundant gene (Nei 1969; Nei and Roychoudhury 1973).

Polyploid organisms offer a model for understanding the evolutionary effects of gene duplication (Ferris and Whitt 1979). In vertebrates, the presence of duplicated genes on different chromosomes suggests that polyploidization has been a feature of early chordate evolution (Ohno et al. 1968). Isozyme studies of genes duplicated by polyploidization have centered mainly on tetraploid fishes (reviewed in Li 1980). The results of these studies indicate that the rate of silencing has been surprisingly slow. For example, in the catostomids, which are thought to have arisen from a tetraploid-

1. Key words: *Xenopus laevis,* paralogy, nuclear genes, rates of substitution.

Address for correspondence and reprints: Austin L. Hughes, Department of Biology, Mueller Laboratory, The Pennsylvania State University, University Park, Pennsylvania 16802.

Mol. Biol. Evol. 10(6):1360–1369. 1993.

0737-4038/93/1006-0016$02.00

ization 50 Mya (Uyeno and Smith 1972), the average loss of duplicate-gene function has been only 53% (Ferris and Whitt 1977*a*). Further, Li (1980) has pointed out that, in every enzyme system studied, loss of one gene has occurred in at least one species, so single-locus expression is adequate for normal function. Several hypotheses have been advanced to account for low rates of silencing, as follows: (1) Bailey et al. (1978) suggested that the large effective population sizes characteristic of salmonids and catostomids permit selection against double null homozygotes to affect gene frequencies. (2) Takahata and Maruyama (1979) argued that, in addition to strong selection against double null homozygotes, there could be further selection against all genotypes containing null alleles. (3) Li (1980) observed that, prior to the reestablishment of disomic segregation, gene loss can only occur via chromosome loss. A long period of tetrasomy following tetraploidization, as may be true for salmonids, thus slows the rate of gene silencing. Furthermore, once a divergence in regulation of expression has occurred, the two genes are not redundant (Li 1980, 1982).

To understand the evolution of duplicate genes at the DNA level, we analyzed available sequences duplicated by tetraploidization in the frog *Xenopus laevis*. Cytogenetic study of frogs of the genus *Xenopus* (family Pipidae) has identified species having $2n$ = 20, 36, 40, 72, and 108 chromosomes, suggesting that polyploidy has occurred frequently in this group (Kobel and Du Pasquier 1986). *Xenopus laevis*, in which $2n$ = 36, is thought to have arisen from a tetraploidization event ~30 Mya. The time of this event is based on immunological distance between *X. laevis* and the related nonpolyploid *X. tropicalis* ($2n$ = 20) (Bisbee et al. 1977). The arrangement of *X. laevis* chromosomes as distinct pairs indicates that disomy is complete (Tymowska and Kobel 1972). Electrophoretic studies of enzymes and blood proteins (summarized in Graf and Kobel 1991) suggest that both copies of duplicated pairs are expressed at approximately one-half of all loci. For a number of these genes, DNA sequences of both copies are available. Thus an analysis of the evolutionary consequences of gene duplication is possible.

We analyzed sequences of 17 duplicated *X. laevis* genes by comparing rates of synonymous and nonsynonymous nucleotide substitution between the two copies with corresponding rates between orthologous genes of human and rodent. We also compared rates of nonsynonymous substitution within each pair of duplicated *X. laevis* genes. If one copy of a duplicated *X. laevis* gene is free to accumulate mutations without constraint, the divergence between the two should be greater than expected relative to the divergence between the two mammalian orthologues. Further, freedom from constraint on one copy should be evident in unequal rates of substitution between the two.

Methods

DNA Sequences Analyzed

Table 1 lists DNA sequences used in the analysis, with source and number of codons compared for each gene. Our analysis does not include some of the duplicated *X. laevis* genes given by Graf and Kobel (1991). No sequence was available for one, or both, of the mammalian orthologues of type I cytokeratin, homeobox 6 and 7, src proto-oncogene, and the ribosomal proteins. Preliminary construction of phylogenetic trees for basic fibroblast growth factor and myoD (data not shown) suggested that the two *X. laevis* genes were the result of an ancient duplication, prior to the tetraploidization event 30 Mya. Preliminary analysis of the two *X. laevis* cardiac actin genes, suggested by Mohun et al. (1986) to be products of the tetraploidization, showed a synonymous nucleotide substitution rate (1.2 ± 0.7 substitutions/100 sites) only one-

Table 1
Sources of Sequences Used in Analysis

Gene and Species	Accession No.	No. of Codons Compared
Actin, skeletal:[a,b]		
Xenopus	X03470 and X12525	376
Mouse	M12347	
Human	J00068	
Albumin:		
Xenopus	M18350 and M21442	594
Rat	J00698	
Human	V00494	
Calmodulin:[a,c]		
Xenopus	K01944 and K01945	149
Rat	M19312, M17069, M16659, and X13933	
Human	M19311, J04046, M27319	
Enkephalin A:		
Xenopus	X00852 and X00853	214
Rat	Y07503	
Human	J00122	
Furin:		
Xenopus	M80471 and M80472	591
Mouse	X54056	
Human	X17094	
Homeobox 2/2.3:[a]		
Xenopus	X06592[d]	214
Mouse	X06762	
Human	M16937	
Insulin:[a]		
Xenopus	M24442 and M24443	100
Mouse	X04724 and X04725	
Human	V00565	
Integrin, β1 subunit:[a]		
Xenopus	M20140 and M20180	796
Mouse	Y00769	
Human	X07979	
N-CAM:[a]		
Xenopus	M25696 and M76710	729
Mouse	Y00051, X07244 and X06328	
Human	X16841	
Oncogenes:		
C-ets-1:[a,e]		
Xenopus	X52691 and X52692	267
Mouse	X53953	
Human	J04101	
C-ets-2:[a]		
Xenopus	M86183, M86184 and X52635	464
Mouse	J04103	
Human	J04102	
C-myc:[a,b]		
Xenopus	M14806 and M14807	393
Mouse	L00038	
Human	V00568	
POMC:[a,c]		
Xenopus	X03843 and X03844	229
Mouse	J00610	
Human	K02406	

Table 1 (*Continued*)

Gene and Species	Accession No.	No. of Codons Compared
Thyroid hormone receptor α:		
Xenopus	M35343 and M35344	367
Rat	M18028	
Human	M24748	
Thyroid hormone receptor β:[a]		
Xenopus	M35359 and M35361	367
Rat	J03933	
Human	X04707	
Thyrotropin-releasing hormone:[a,c]		
Xenopus	M34699[f]	184
Rat	M12138	
Human	M63580	
Vimentin:[a]		
Xenopus	X16843 and X16844	456
Mouse	M26251	
Human	X56134	

[a] Both copies expressed.
[b] Expressed at different developmental stages.
[c] Both copies expressed equally.
[d] A second sequence having no accession no. was taken from the work of Fritz et al. (1989).
[e] Expressed in different tissues.
[f] A second sequence having no accession no. was taken from the work of Bulant et al. (1992).

fifteenth as large as the average for duplicated *X. laevis* genes (see below). This suggests that the cardiac actin gene duplication occurred much more recently than 30 Mya. The six pairs of globin gene sequences were not included in the analysis because of the difficulty in establishing accurate orthologies between amphibians and mammals. Vitellogenin, an egg-yolk precursor protein, has no mammalian orthologue (Wallace 1978). However, examination of the GenBank and EMBL sequence databases yielded six additional duplicated *X. laevis* genes not listed by Graf and Kobel (1991) for which human and rodent orthologues were available and phylogenetic tree construction supported the hypothesis of a recent duplication.

All the *X. laevis* genes in our analysis were determined to be nonallelic by the authors who reported the sequences. Whether both genes were expressed, and to what extent, was not reported in every case (table 1). Also, four nonallelic rat calmodulin sequences and three nonallelic human calmodulin sequences were available. These could be assigned to three sets of orthologues, for which d_S and d_N were calculated separately (see below).

Statistical Methods

Each set of sequences was first aligned at the amino acid level by the CLUSTAL V program (Higgins et al. 1992) and corrected by eye where necessary. (Alignments are available on request.) Where the alignment postulated a gap in any sequence, the corresponding codons were removed from all sequences so that a comparable data set was obtained in each case. Numbers of synonymous nucleotide substitutions per synonymous site (d_S) and nonsynonymous substitutions per nonsynonymous site (d_N) were estimated by Nei and Gojobori's (1986) method I. Relative-rate tests (Wu and Li 1985) were conducted by separately comparing d_N between each *X. laevis* sequence and its human orthologue.

Results

Table 2 shows mean d_S and $d_N/100$ sites in comparisons between human and rodent and between duplicated *Xenopus laevis* genes. For all genes, d_S exceeded d_N in both comparisons, as would be expected for genes under purifying selection (Nei 1987, p. 33). Further, both d_S and d_N between the duplicated *X. laevis* genes were in most cases lower than d_S and d_N between the mammalian genes, consistent with a tetraploidization more recent than the divergence of rodents and primates ~80–100 Mya (Li et al. 1990). The average ratio of d_S between the mammalian genes to d_S between the *X. laevis* genes was 2.93; the average ratio of d_N between the mammalian genes to d_N between the *X. laevis* genes was 2.28 (table 2).

To determine whether the duplicated *X. laevis* genes were subject to constraints similar to those on the mammalian genes, we computed correlations for (1) d_S between the *X. laevis* genes and d_S between the mammalian genes (fig. 1A) and (2) d_N between the *X. laevis* genes and d_N between the mammalian genes (fig. 1B). There was no statistically significant correlation between the two values of d_S ($r = 0.185$; fig. 1A). We used the Tajima-Nei correction (Tajima and Nei 1984), which makes no assumptions about the base composition of sequences being compared, to compute K_4, the number of nucleotide substitutions at fourfold degenerate sites (Wolfe et al. 1989). As with d_S, there was no significant relationship between the two values of K_4 ($r = 0.099$). There was, however, a significant positive correlation between the two values

Table 2
Mean d_S and $d_N/100$ Sites (± SE) in Comparisons between Human and Rodent Genes and between Two Genes of *Xenopus laevis*

Gene	Human vs. Rodent		Xenopus A vs. Xenopus B	
	Mean d_S ± SE	Mean d_N ± SE	Mean d_S ± SE	Mean d_N ± SE
Actin, skeletal	53.9 ± 6.1	0.2 ± 0.2	15.8 ± 2.6	0.0 ± 0.0
Albumin	111.3 ± 10.8	15.9 ± 1.2	16.2 ± 2.3	5.9 ± 0.7
Calmodulin:				
Rat a/b vs. human a	66.8 ± 12.6	0.0 ± 0.0		
Rat c vs. human b	52.3 ± 10.1	0.0 ± 0.0	31.6 ± 7.1	0.0 ± 0.0
Rat d vs. human c	43.8 ± 8.8	0.7 ± 0.4		
Enkephalin A	70.1 ± 10.9	9.0 ± 1.4	14.2 ± 3.5	2.3 ± 0.7
Furin	54.1 ± 4.8	1.9 ± 0.4	18.8 ± 2.3	1.5 ± 0.3
Homeobox 2/2.3	23.0 ± 4.3	3.6 ± 0.9	14.5 ± 3.4	2.4 ± 0.7
Insulin	59.6 ± 11.8	11.1 ± 2.2	12.0 ± 4.5	3.8 ± 1.3
Integrin, β1 subunit	62.6 ± 5.0	4.2 ± 0.5	30.7 ± 2.9	1.0 ± 0.2
N-CAM	54.8 ± 4.5	4.7 ± 0.6	14.7 ± 1.9	3.6 ± 0.5
Oncogenes:				
C-ets-1	32.1 ± 5.0	1.3 ± 0.5	21.5 ± 3.8	0.5 ± 0.3
C-ets-2	81.8 ± 8.5	4.1 ± 0.6	32.8 ± 3.9	2.8 ± 0.5
C-myc	48.6 ± 5.6	2.9 ± 0.6	18.9 ± 3.0	2.7 ± 0.6
POMC	48.6 ± 7.3	11.0 ± 1.5	26.5 ± 4.9	3.6 ± 0.8
Thyroid hormone receptor α	18.7 ± 3.0	0.4 ± 0.2	11.8 ± 0.2	0.8 ± 0.3
Thyroid hormone receptor β	54.1 ± 6.3	1.0 ± 0.3	14.7 ± 2.6	0.8 ± 0.3
Thyrotropin-releasing hormone	96.0 ± 15.7	24.2 ± 2.7	17.2 ± 4.1	7.9 ± 1.4
Vimetin	40.5 ± 4.4	1.2 ± 0.3	14.2 ± 2.3	2.8 ± 0.5
Overall mean	56.1 ± 5.7	5.7 ± 1.6	19.2 ± 1.7	2.5 ± 0.5

of d_N ($r = 0.915$, $P < 0.001$; fig. 1B). At the amino acid level, the duplicated *X. laevis* genes are apparently subject to the same constraints that lower d_N in mammals.

Mouchiroud et al. (1988) found a strong correlation between third-position G + C content in orthologous human and rodent genes, indicating that there is conservation of G + C content in orthologous genes in the two mammalian orders. The average third-position G + C content of the *X. laevis* genes in our analysis was 52.6 ± 2.8%, that of the rodent genes was 66.6 ± 2.9%, and that of the human genes was 65.2 ± 3.9%. Both the rodent ($t = 2.62$, $P < 0.01$) and the human ($t = 3.47$, $P < 0.001$) genes were significantly higher in third-position G + C than were the *X. laevis* genes. We found a significant positive correlation between third-position G + C content of the human genes and third-position G + C content of their rodent orthologues ($r = 0.854$, $P < 0.001$). We also found a significant positive correlation between third-position G + C content within pairs of duplicated *X. laevis* genes ($r = 0.951$, $P < 0.001$). However, we found no statistically detectable relationship between third-position G + C content in either of the mammalian genes and that of the *X. laevis* genes (data not shown). Thus the conservation of third-position G + C content of orthologous genes does not extend to groups as distant from mammals as amphibians.

Because the *X. laevis* tetraploidization was relatively recent, low d_N between the duplicated genes could result from purifying selection on one copy only, with the other copy free to accumulate mutations. To examine this hypothesis, we performed relative-rate tests (Wu and Li 1985) comparing each *X. laevis* gene separately to its human orthologue (the reference sequence). The results are shown in table 3. In every comparison, neither of the two copies evolved significantly faster than the other at nonsynonymous sites.

Graf and Kobel (1991) summarize data showing that at 13 of 24 loci examined, duplicate *X. laevis* genes were detectable by electrophoresis. Since electrophoretic detectability requires a net charge difference, not all duplicate genes will be detected by this method. To obtain an estimate of the proportion of duplicate genes not detectable by electrophoresis, we compared protein products of the genes in our analysis with respect to net charge. In 5 of 17 cases, the charges of the two proteins encoded by the duplicate genes were identical, and thus the expression of both copies would not have been detected by electrophoresis. We used 54.2% (13 of 24) as an estimate of the proportion of all duplicate loci at which both copies are expressed and both copies are detectable by electrophoresis. We used 29.4% (5 of 17) as an estimate of the proportion of duplicate loci with both copies expressed at which the two copies are not detectable by electrophoresis. On the basis of these values, we estimated that 76.8% of duplicated *X. laevis* loci have not been silenced.

Discussion

For the duplicated *Xenopus laevis* genes, $d_S > d_N$ in every comparison (table 2) is evidence of purifying selection (Nei 1987, p. 33). Furthermore, the significant positive correlation of d_N between the *X. laevis* genes with d_N between the mammalian genes (fig. 1B), as well as the nearly equal rates of nonsynonymous substitution in the two *X. laevis* genes (table 3), indicates that both copies are under constraints similar to those on the mammalian genes. One explanation of these results is that regulatory divergence between the *X. laevis* genes has occurred in some cases (Li 1980). If this were true, neither gene copy would be redundant, since the two genes would be expressed in different tissues. In four of the duplicated genes in our analysis, the two copies are expressed in different tissues or at different developmental times (table 1). The two

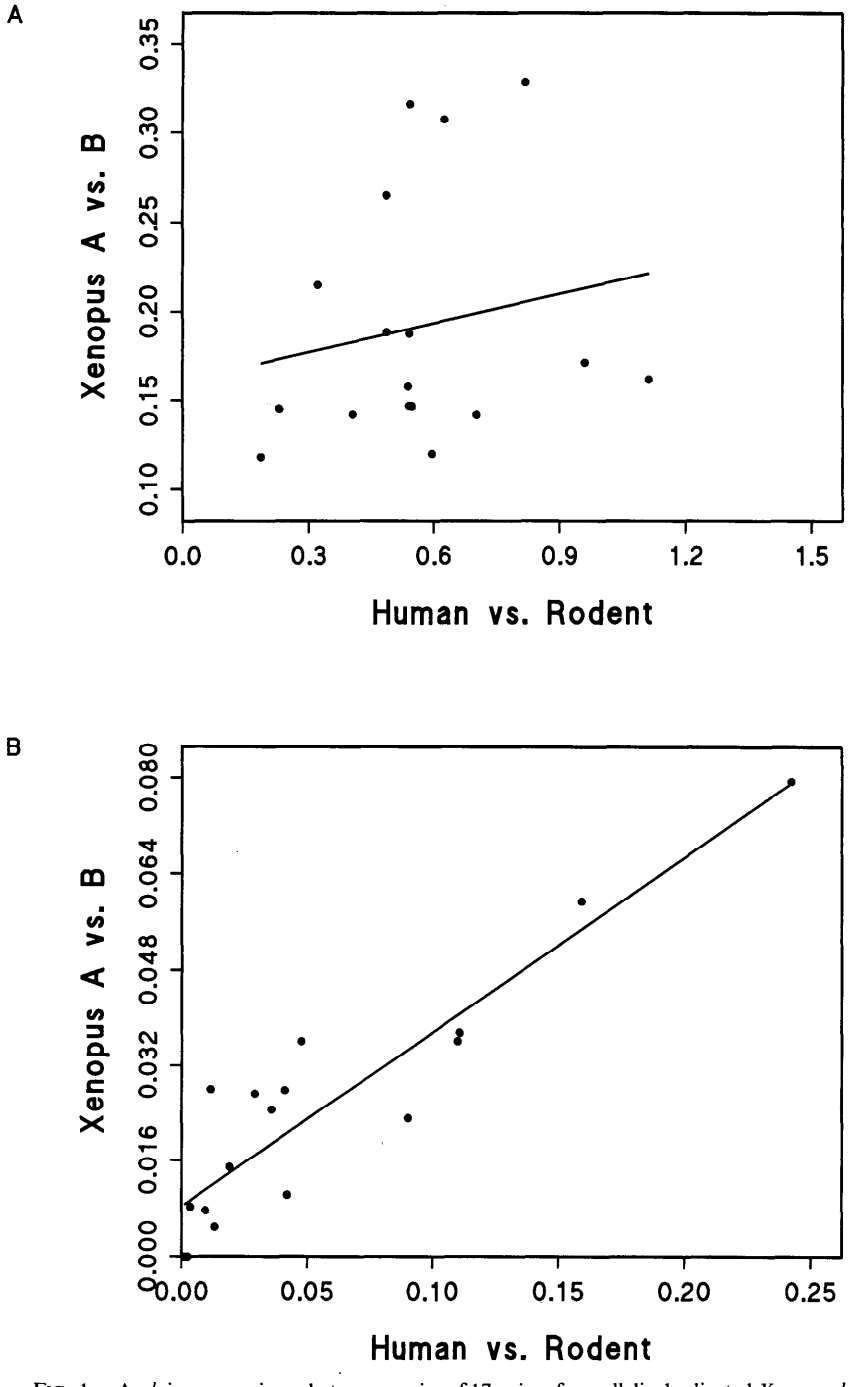

FIG. 1.—A, d_S in comparisons between copies of 17 pairs of nonallelic duplicated *Xenopus laevis* genes, as a function of d_S between orthologous genes from human and rodent. The equation for the regression line is $Y = 0.161 + 0.054\,X\,(r = 0.185)$. B, d_N in comparisons between copies of 17 pairs of nonallelic duplicated *X. laevis* genes, as a function of d_N between orthologous genes from human and rodent. The equation for the regression line is $Y = 0.008 + 0.292\,X\,(r = 0.915)$.

Table 3
Mean d_N/100 Sites (± SE) between Each of Two Duplicated Sequences (A and B) in *Xenopus laevis* and an Orthologous Human Sequence, with Results of Relative-Rate Tests

	Mean d_N ± SE		
Gene	*Xenopus* A vs. Human	*Xenopus* B vs. Human	Difference
Actin, skeletal	2.2 ± 0.5	2.2 ± 0.5	0.0 ± 0.1
Albumin	58.8 ± 2.9	59.1 ± 2.9	−0.3 ± 1.2
Calmodulin	0.7 ± 0.3	0.7 ± 0.3	0.0 ± 0.0
Enkephalin A	24.7 ± 2.5	23.2 ± 2.4	1.5 ± 0.9
Furin	14.6 ± 1.1	14.7 ± 1.1	−0.2 ± 0.4
Homeobox 2/2.3	19.7 ± 2.2	20.0 ± 2.2	−0.3 ± 0.9
Insulin	35.6 ± 4.8	35.1 ± 4.8	0.5 ± 0.2
Integrin, β1 subunit	12.2 ± 0.9	12.3 ± 0.9	−0.1 ± 0.2
N-CAM	19.5 ± 1.2	19.9 ± 1.2	−0.3 ± 0.6
Oncogenes:			
C-ets-1	5.0 ± 0.9	5.0 ± 0.9	0.1 ± 0.3
C-ets-2	21.7 ± 1.6	21.5 ± 1.6	0.2 ± 0.6
C-myc	24.8 ± 1.9	25.0 ± 1.9	−0.2 ± 0.7
POMC	32.1 ± 2.9	32.7 ± 2.9	−0.6 ± 1.2
Thyroid hormone receptor α	4.8 ± 0.8	5.4 ± 0.8	−0.6 ± 0.3
Thyroid hormone receptor β	4.4 ± 0.7	4.5 ± 0.8	−0.1 ± 0.3
Thyrotropin-releasing hormone	66.4 ± 5.8	70.1 ± 6.1	−3.7 ± 2.8
Vimentin	14.4 ± 1.3	15.4 ± 1.3	−1.0 ± 0.6

copies of thyrotropin-releasing hormone, for example, are expressed in skin and brain, respectively (Bulant et al. 1992). All of the duplicated *X. laevis* genes that are not known to be expressed differentially and for which sequences were available have protein products that are involved in interactions with other proteins, either directly or indirectly in biochemical pathways (table 1). In such cases, selection may act to reduce the amount of variation between the duplicated proteins, because a mutation altering the structure of such a protein, rather than being simply a null mutant, may have a deleterious effect on the organism by interfering in protein-protein interactions.

Thus, our evidence of constraint on both copies of duplicated *X. laevis* genes supports Takahata and Maruyama's (1979) hypothesis of selection against all genotypes containing null alleles. By contrast, Bailey et al.'s (1978) hypothesis that loss of duplicate genes is slowed by a large effective population size seems unlikely, since effective population size of *X. laevis* is probably much smaller than that of the salmonids. Furthermore, the evidence of constraint on both copies of a duplicated gene does not support the theory that, after duplication, one gene copy is redundant and therefore free to accumulate mutations (Kimura and Ohta 1974).

It is possible, however, that sometimes selection may act to favor silencing of such genes. For example, silencing of one copy of the major histocompatibility complex (MHC) in *X. laevis* may have been favored by natural selection (Du Pasquier et al. 1977). Immunologists have argued that, because self-reactive T-cell clones are eliminated in development, selection may set a limit on the number of functional MHC loci (Howard 1987).

The ratios of mean d_S and d_N for the mammalian genes to those for the *X. laevis* genes were close to those predicted under the hypothesis of a molecular clock (table 2). Li et al. (1990) estimated 80–100 Mya for the rodent-primate divergence. When

80 Mya is used, the d_S ratio for the genes in our analysis gives a time for the *X. laevis* tetraploidization of 27 Mya, and the d_N ratio gives a time of 35 Mya. When 100 Mya is used, the d_S and d_N ratios give times of 34 Mya and 44 Mya, respectively. These estimates are reasonably close to the 30 Mya estimate of Bisbee et al. (1977), based on immunological distances. Because more sites are involved, d_N is expected to be a better estimator of divergence times than d_S. If this is true here, Bisbee et al.'s calculated time of tetraploidization may be a slight underestimate.

The tetraploid loaches, in which the estimated time of tetraploidization (15–40 Mya) is similar to that of *X. laevis,* have electrophoretically detectable duplication at only ~25% of loci examined (Ferris and Whitt 1977*b*). However, as Ferris and Whitt point out, it is possible that expressed duplicate genes from a relatively recent tetraploidization have not yet diverged enough to be detected by electrophoresis. Direct analysis of DNA in our study has shown that the level of underestimation of duplicate-gene expression can be ≥29%. If our estimation is correct, it suggests that the level of silencing in *X. laevis* is considerably lower than that in the loaches. If so, the most parsimonious explanation would be that the tetraploidization event in the loaches occurred earlier than previously estimated.

Acknowledgments

This research was supported by grants from the National Institutes of Health to A. L. H. and Masatoshi Nei and from the National Science Foundation to Masatoshi Nei.

LITERATURE CITED

BAILEY, G. S., R. T. M. POULTER, and P. A. STOCKWELL. 1978. Gene duplication in tetraploid fish: model for gene silencing at unlinked duplicate loci. Proc. Natl. Acad. Sci. USA **75:** 5575–5579.

BISBEE, C. A., M. A. BAKER, A. C. WILSON, H.-A. IRANDOKHT, and M. FISCHBERG. 1977. Albumin phylogeny for clawed frogs (*Xenopus*). Science **195:**785–787.

BULANT, M., K. RICHTER, K. KUCHLER, and G. KREIL. 1992. A cDNA from brain of *Xenopus laevis* coding for a new precursor of thyrotropin-releasing hormone. FEBS Lett. **296:**292–296.

DU PASQUIER, L., V. C. MIGGIANO, H. R. KOBEL, and M. FISCHBERG. 1977. The genetic control of histocompatibility reactions in natural and laboratory-made polyploid individuals of the clawed toad *Xenopus.* Immunogenetics **5:**129–141.

FERRIS, S. D., and G. S. WHITT. 1977*a.* Loss of duplicate gene expression after polyploidisation. Nature **265:**258–260.

———. 1977*b.* Duplicate gene expression in diploid and tetraploid loaches (Cypriniformes, Cobitidae). Biochem. Genet. **15:**1097–1113.

———. 1979. Evolution of the differential regulation of duplicate genes after polyploidization. J. Mol. Evol. **12:**267–317.

FRITZ, A. F., K. W. Y. CHO, C. V. E. WRIGHT, B. G. JEGALIAN, and E. M. DE ROBERTIS. 1989. Duplicated homeobox genes in *Xenopus.* Dev. Biol. **131:**584–588.

GRAF, J.-D., and H. R. KOBEL. 1991. Genetics. Pp. 19–34 *in* B. K. KAY and H. B. PENG, eds. Methods in cell biology. Vol. 36, *Xenopus laevis:* practical uses in cell and molecular biology. Academic Press, New York.

HIGGINS, D. G., A. J. BLEASBY, and R. FUCHS. 1992. Clustal V: improved software for multiple sequence alignment. Comput. Appl. Biosci. **8:**189–191.

HOWARD, J. C. 1987. MHC organization in the rat: evolutionary considerations. Pp. 397–427 *in* G. KELSOE and D. H. SCHULZE, eds. Evolution and vertebrate immunity. University of Texas Press, Austin.

KIMURA, M., and T. OHTA. 1974. On some principles governing molecular evolution. Proc. Natl. Acad. Sci. USA **71**:2848–2852.

KOBEL, H. R., and L. DU PASQUIER. 1986. Genetics of polyploid *Xenopus*. Trends Genet. **2**: 310–315.

LI, W.-H. 1980. Rate of gene silencing at duplicate loci: a theoretical study and interpretation of data from tetraploid fishes. Genetics **95**:237–258.

———. 1982. Evolutionary change of duplicate genes. Isozymes **6**:55–92.

LI, W.-H., M. GOUY, P. M. SHARP, C. O'HUIGIN, and Y. W. YANG. 1990. Molecular phylogeny of Rodentia, Lagomorpha, Primates, Artiodactyla, and Carnivora and molecular clocks. Proc. Natl. Acad. Sci. USA **87**:6703–6707.

MOHUN, T. J., N. GARRETT, and J. B. GORDON. 1986. Upstream sequences required for tissue-specific activation of the cardiac actin gene in *Xenopus laevis* embryos. EMBO J. **5**:3185–3193.

MOUCHIROUD, D., C. GAUTIER, and G. BERNARDI. 1988. The compositional distribution of coding sequences and DNA molecules in humans and murids. J. Mol. Evol. **27**:311–320.

NEI, M. 1969. Gene duplication and nucleotide substitution in evolution. Nature **221**:40–42.

———. 1987. Molecular evolutionary genetics. Columbia University Press, New York.

NEI, M., and T. GOJOBORI. 1986. Simple methods for estimating the numbers of synonymous and nonsynonymous nucleotide substitutions. Mol. Biol. Evol. **3**:418–426

NEI, M., and A. K. ROYCHOUDHURY. 1973. Probability of fixation of nonfunctional genes at duplicate loci. Am. Nat. **107**:362–372.

OHNO, S. 1970. Evolution by gene duplication. Springer, New York.

OHNO, S., U. WOLF, and N. B. ATKIN. 1968. Evolution from fish to mammals by gene duplication. Hereditas **59**:169–187.

TAJIMA, F., and M. NEI. 1984. Estimation of evolutionary distance between nucleotide sequences. Mol. Biol. Evol. **1**:269–285.

TAKAHATA, N., and T. MARUYAMA. 1979. Polymorphism and loss of duplicate gene expression: a theoretical study with application to tetraploid fish. Proc. Natl. Acad. Sci. USA **76**:4521–4525.

TYMOWSKA, J., and H. R. KOBEL. 1972. Karyotype analysis of *Xenopus muelleri* (Peters) and *Xenopus laevis* (Daudin) (Pipidae). Cytogenetics **11**:270–278.

UYENO, T., and G. SMITH. 1972. Tetraploid origin of the karyotype of catostomid fishes. Science **175**:644–646.

WALLACE, R. A. 1978. Oocyte growth in nonmammalian vertebrates. Pp. 469–502 *in* R. E. JONES, ed. The vertebrate ovary: comparative biology and evolution. Plenum, New York.

WOLFE, K. H., P. M. SHARP, and W.-H. LI. 1989. Mutation rates vary among regions of the mammalian genome. Nature **337**:283–285.

WU, C.-I., and W.-H. LI. 1985. Evidence for higher rates of nucleotide substitution in rodents than in man. Proc. Natl. Acad. Sci. USA **82**:1741–1745.

PAUL M. SHARP, reviewing editor

Received March 15, 1993; revision received June 22, 1993

Accepted June 22, 1993