Abstract

The enzyme enolase [EC 4.2.1.11] is found in all organisms, with vertebrates exhibiting tissue-specific isozymes encoded by three genes: alpha (α), beta (β), and gamma (γ) enolase. Limited taxonomic sampling of enolase has obscured the timing of gene duplication events. To help clarify the evolutionary history of the gene family, cDNAs were sequenced from six taxa representing major lineages of vertebrates: Chiloscyllium punctatum (shark), Amia calva (bowfin), Salmo trutta (trout), Latimeria chalumnae (coelacanth), Lepidosiren paradoxa (South American lungfish), and Neoceratodus forsteri (Australian lungfish). Phylogenetic analysis of all enolase and related gene sequences revealed an early gene duplication event prior to the last common ancestor of living organisms. Several distantly related archaeobacterial sequences were designated as ‘enolase-2’, whereas all other enolase sequences were designated ‘enolase-1’. Two of the three isozymes of enolase-1, α- and β-enolase, were discovered in actinopterygian, sarcopterygian, and chondrichthian fishes. Phylogenetic analysis of vertebrate enolases revealed that the two gene duplications leading to the three isozymes of enolase-1 occurred subsequent to the divergence of living agnathans, near the Proterozoic/Phanerozoic boundary (approximately 550 Mya). Two copies of enolase, designated α1 and α2, were found in the trout and are presumed to be the result of a genome duplication event. © 2000 Elsevier Science B.V. All rights reserved.

Keywords: Archaeabacteria; Cambrian explosion; Gene duplication; Horizontal transfer; Polyploidization; Vertebrate evolution

1. Introduction

Gene and genome duplications are believed to have contributed significantly to the evolution of complex body plans in the metazoans (Holland et al., 1994; Brooke et al., 1998), especially vertebrates (Nadeau and Sankoff, 1997; Stock et al., 1997; Williams and Holland, 1998). The importance of such duplications for animal evolution relies on knowledge of the timing and order of the duplications. For example, if gene duplication occurred before the divergence of deuterostome and protostome animals, it would not be directly related to developmental and morphological events unique to vertebrates (Hughes, 1999), although it could be related through later recruitment (Ono-Koyanagi et al., 2000). Broad taxonomic sampling enhances any evolutionary analysis of gene relationships, and examples of well-studied gene families include hemoglobin (Hardison, 1996; Gorr et al., 1999), lactate dehydrogenase (Crawford et al., 1989; Stock and Whitt, 1992b; Mannen Sanko V, 1997; Stock et al., 1997; Williams and Holland, 1998). Enolase (2-phospho-d-glycerate hydrolase) [EC 4.2.1.11], a dimeric enzyme of the glycolytic pathway responsible for catalyzing the conversion of 2-phosphoglycerate to phosphoenolpyruvate (Day et al., 1993; Brown and Deolittle, 1997), is another conserved gene family that allows for the evolutionary study of gene duplication in animals. It is found in all organisms and is a member of the enolase gene superfamily, which includes carboxyphosphonoenolpyruvate synthase, mandelate racemase, galactonate dehydratase, glutarate dehydratase, muconate-lactonizing enzymes, N-acylamino acid racemase, beta-methylaspartate ammonia-lyase, and o-succinylbenzoate synthase (Babbitt et al., 1996; Pegg and Babbitt, 1999). Vertebrates exhibit three,
more stringently conserved than a of evolution may di
V
bility does exist that the degree of conservation and rate
b
(Wistow et al., 1988). Presently, it is not known if either
Zymomonas mobilis
-t-crystallin in lampreys, other fishes, and reptiles (Wistow et al., 1989). An \( \gamma \)-heterodimer from neuronal tissue also has been reported (Schmechel et al., 1984).

Analyses of amino acid and nucleotide sequences have suggested that the paralogs rapidly arose from a single ancestral gene through a "burst" event early in vertebrate evolution (Day et al., 1993). While numerous sequences are available for enolases, especially vertebrate \( \tau \)-enolases, sequences for vertebrate \( \beta \) and \( \gamma \)-enolases are restricted to chicken, human, rat, and mouse (an amino acid sequence for rabbit \( \beta \)-enolase also is available). Based on intron lengths, Galliongo et al. (1990) proposed that \( \beta \) and \( \gamma \)-enolases are the two most closely related genes. Because of the paucity of sequence data, the timing and order of the gene duplications have remained unresolved.

Several studies have placed the enolase gene duplications between 200 and 300 Mya (Rider and Taylor, 1975; Clark-Rosenberg and Marangos, 1980; Segl et al., 1988). However, Landrey et al. (1978) proposed that a gene duplication was responsible for the six electrophoretically distinct enolase isozymes that they found in Coho salmon. With the likely origin of actinopterygian fishes occurring approximately 450 Mya (Benton, 1997; Kumar and Hedges, 1998), this conclusion regarding Coho salmon should be re-examined. Landrey et al. (1978) suggested that the polyplody of salmonids could be a complicating factor, but, without direct sequence information, it is impossible to determine whether the isozymes found by Landrey et al. (1978) are the result of a gene duplication in the vertebrate lineage, a genome duplication in the salmonid lineage, or both. Furthermore, it remains unclear which vertebrates might exhibit multiple copies of the enolase gene.

Phylogenetic analyses of enolase paralogs could be complicated by the possible effects of gene sharing. Gene sharing occurs when a gene, without duplication, acquires a second function while maintaining its original function. This places two different sets of selective constraints on mutational change within the shared gene (Piatigorsky, 1989; Piatigorsky and Wistow, 1989). The \( \tau \)-enolase gene also is expressed as the lens protein \( \tau \)-crystallin in lampreys, other fishes, and reptiles (Wistow et al., 1988). Presently, it is not known if either \( \beta \) or \( \gamma \)-enolase functions in this manner, but the possibility does exist that the degree of conservation and rate of evolution may differ markedly between the paralogs. Indeed, McAleese et al. (1988) found that \( \gamma \)-enolase is more stringently conserved than \( \tau \)-enolase. More recently, Day et al. (1993) reported that NNE appears to be evolving faster than either NSE or MSE.

To elucidate more clearly the evolutionary history of the enolase gene family, we have generated a phylogeny from all available, complete (or nearly so) enolase amino acid sequences. We have also sequenced \( \alpha \) and \( \beta \)-enolase cDNAs from primitive vertebrates, including chondrichthyian, actinopterygian, and sarcopterygian fishes. The presence of multiple, identifiable forms of enolase in these organisms places a lower limit on the age of the duplications and allows predictions of the presence or absence of the duplicated genes in other vertebrates. Finally, we have addressed the timing of the enolase duplication events and the issue of evolutionary rate variation between the duplicated genes.

2. Materials and methods

The following is a list of amino acid sequences obtained from the public databases (all are enolase-1 unless designated otherwise). Archaeabacteria: Archaeoglobus fulgidus, GenBank Accession No. = AAB90112 (enolase-2); Aeropyrum pernix, BAA31473; Halocarcula marismortui, P29201 (enolase-2); Methanobacterium thermoautotrophicum, AAB84350 (enolase-2); Methanococcus jannaschii, Q60173; Pyrococcus abyssi, CAB30622; CAB49458 (enolase-2); Pyrococcus horikoshi, BAA31009; P71042 (enolase-2). Eubacteria: Aquifex aeolicus, AAC06738; Bacillus subtilis, P37809; Borrelia burgdorferi, O51312; Campylobacter jejuni, CAB73659; Chlamydia muridarum, AAF39672; Chlamydia trachomatis, AAC68189; Chlamydocila pneumoniae, AAF3843; Deinococcus radioodans, AAF12713; Escherichia coli, P03824, P31458 (galactonate dehydratase); Haemophilus influenzae, P43806; Helicobacter pylori, AAD05723; Mycobacterium tuberculosis, P96377; Mycoplasma genitalium, U39723; Mycoplasma pneumoniae, P75189; Neisseria meningitidis, AAF41661; Nitrosomonas europaea, AF061753; Pseudomonas putida, A28700 (mandle-late racemase); P42206 (glutarate dehydratase); Staphylococcus aureus, AF063594; Streptococcus intermedius, BAA38185; Streptococcus thermophilus, AF027167; Synecocystis PCC6803, P77972; Thermostoga maritima, AAD3958; Tropenoma pallidum, P74934; Ureaplasma urealyticum, AAF30591; and Zymomonas mobilis, P33675. Protozoa: Entamoeba histolytica, P51555; Mastigamoeba balamuthi, AAF13454; Plasmodium falciparum, Q72727; Toxoplasma gondii, AAD51128. Fungi: Aspergillus oryzae, Q12560; Candida albicans, P30575; Cladosporium herbarum, X78226; Cunninghamella elegans, Y17298; Neurospora crassa, P42944; Pneumocystis carinii, AF036247; Saccharomyces cerevisiae, P00924; Schizosaccharomyces pombe, P40370. Viridiplantae: Alna glutinosa, Q43221;
Arabidopsis thaliana, P25696; Chlamydomonas reinhardtii, P31683; Lupinus betaea, CAB73428; Lycopersicon esculentum, P26300; Mesembryanthemum crystallinum, Q43130; Oryza sativa, Q42971; Ricinus communis, P42896; and Zea mays, U17973. Invertebrates: Caenorhabditis elegans, Q27527; Drosophila melanogaster, X17034; D. pseudoobscura, AF025805; D. subobscura, AF025806; Homarus gammarus, P56252; Loligo pealei, O06254; Peneaus monodon, AAC78141; Schistosoma japonicum, P53676; and Schistosoma mansoni, U30175. Mammals: Oryctolagus cuniculus, P25704 (β-enolase).

The following amino acid and cDNA nucleotide sequences for vertebrates were obtained from the public databases. γ-enolases. Amphibians: Xenopus laevis, Y00718. Reptiles: Alligator mississippiensis, AF072586; Caiman crocodilus, AF115855; Eumeces ineptus, AF115857; Pelias subsingeri, AF115859; Python regius, AF072589; Sceloporus undulatus, AF072587; Sphenodon punctatus, AF115856; and Trachemys scripta, AF072588. Birds: Anas platyrhynchos, M20749; and Gallus gallus, D37900. Mammals: Bos taurus, M20749; and Rattus norvegicus, M22349; and R. norvegicus, M22349; and R. norvegicus, M20749; and Mus musculus, M20749; and Mus sp., X25237; Rattus norvegicus, X012610; and Mus sp., X25237; and R. norvegicus, M22349. The following cyclostome enolase sequences were included: Eptatretus burgeri, BAA88479; Lamperna reissneri, BAA88482; and Lamperna reissneri, BAA88483 (the partial lamperna sequence, BAA88483, is identical to that of the 97-residue region within BAA88483). CLUSTAL X (Thompson et al., 1997) was used to align amino acid and nucleotide sequences.

The nucleotide sequence alignment was then used to construct three sets of primers for reverse transcription and PCR amplification: non-specific (i.e. primers that did not distinguish between paralogous forms of the enolase gene), γ-enolase specific primers, and β-enolase specific primers. Specific primers were created by placing the 3’ end of the primer on a site that was diagnostic between β- and β-enolase. Sites that were different between α and β sequences yet identical among ≥ 5 (3’) to 3” were:


The alpha specific primers and sequences (5’ to 3’) were:


The beta specific primers and sequences (5’ to 3’) were:


Polyadenylated RNAs [poly(A)⁺ RNA] were isolated from liver tissue of a brown spotted cat shark (Chiloscyllium punctatum; pet trade), South American lungfish (Lepidostomus paradoxus; pet trade), and brown trout (Salmo trutta; domestic). Polyadenylated RNAs were isolated from skeletal muscle tissue from a brown spotted cat shark, South American lungfish, Australian lungfish (Neoceratodus forsteri; locality unknown), coelacanth (Latimeria chalumnae; Virginia Institute of Marine Science 8118, Comoros Islands), and brown trout. Polyadenylated RNAs were isolated from a mixture of liver and smooth muscle tissue from a bowfin (Amia calva).

Isolation of poly(A)⁺ mRNA from all samples except bowfin followed the protocol of the Micro-FastTrack mRNA isolation Kit (Invitrogen) with minor modifications. Initial incubation of the tissue suspension was 30 min. Ethanol precipitates were prepared by adding 10 μl of glycogen (2 mg/ml), 60 μl of 2 M sodium acetate, and 1200 μl of 200 proof ethanol to 400 μl of
elutent containing the poly(A)$^+$ RNA. Isolation of purified as described above, and 1 μl was used as template for a second amplification conducted under the same conditions in 50 μl of solution. The double-stranded product of the second amplification was purified with four passes in a 30,000 MW filter (Millipore). Cycle sequencing reactions were performed by the Nucleic Acids Facility, The Pennsylvania State University, using 3’ dye-labeled deoxynucleotide triphosphates. Nucleotide sequences were generated for the forward and reverse strands of each fragment of each gene on an ABI PRISM 377 DNA Sequencer (Perkin-Elmer).

Nucleotide sequences were added to the alignment and were used to deduce amino acid sequences. Neighbor-joining trees (Saitou and Nei, 1987) were constructed with MEGA (Kumar et al., 1993) using the Kimura 2-parameter transitions + transversions (s+v) and transversions only (v) distances (Kimura, 1980) for nucleotides. For amino acid analyses, a Poisson-corrected distance and a gamma distance ($\alpha = 1.05$; Wang et al., 1999) were used with neighbor-joining. A maximum likelihood analysis was conducted on the restricted vertebrate protein data with MOLPHY (Adachi and Hasegawa, 1996) using the JTT-F option. The statistical significance of groups in the neighbor-joining trees was assessed by the bootstrap probability (Felsenstein, 1985) with 2000 replications (Hedges, 1992).

Fig. 1. Phylogenetic tree (neighbor-joining) of enolase and closely related proteins. Archaeabacterial enolase amino acid sequences, and representative eubacterial and eukaryote sequences, are included along with other genes of the enolase gene superfamily. Letters following taxon labels indicate the three domains: Archaeabacteria (A), Eubacteria (E), and Eukaryota (K). Numbers at nodes are bootstrap P-values ($\geq 50\%$). The tree is rooted by the midpoint of the longest branch.
3. Results

3.1. Global amino acid analyses

It was clear from initial analyses that several archaebacterial enolase sequences were not closely related to the majority of enolase sequences, including other archaebacterial sequences. Therefore, we aligned and compared all archaebacterial sequences with representatives of eu- and eukaryotic enolases, and with several other genes in the enolase gene superfamily (Fig. 1). The well-defined (100% bootstrap confidence) cluster containing the majority of enolase sequences from all three domains is designated here as ‘enolase-1’. The remaining five enolase sequences, all archaebacterial, are more distantly related and form two clusters in the tree. However, the two clusters are not separated by a significant confidence value and may belong to a single gene. Therefore, they are conservatively designated as ‘enolase-2’ (additional sequences and analyses may resolve an additional gene). The fact that the archaebacterial representatives of enolase-1 belong to two kingdoms (Crenarchaeota and Euryarchaeota) and that both genes (enolase-1 and enolase-2) have been sequenced from the same species (Pyrococcus abyssii and P. horikoshii) suggests that gene duplication is involved rather than horizontal transfer, although the latter cannot be ruled out. These highly divergent (and, in some cases, shorter) sequences were not included in later analyses to maximize the number of aligned sites.

The tree generated from amino acid sequences of enolase-1 (Fig. 2) shows monophyly of the vertebrate sequences (99%) and of each enolase-1 paralog: α-enolase (93%), β-enolase (95%), and γ-enolase (99%). However, the branching order of the three genes is not significantly resolved. The remaining animals (invertebrates), except the squid (basal), form a cluster that groups with the vertebrate cluster. Animal enolases are monophyletic (<50%). Plant (<58%) and fungal (50%) enolases also form monophyletic clusters, and the three kingdoms form a monophyletic group (<50%). The four protists are basal on the monophyletic (99%) eukaryote branch (Fig. 2). The four archaeabacteria (95%) and the 24 eubacteria (85%) each form monophyletic groups. The tree was rooted with eubacteria because the eukaryotes appear to be derived by normal vertical evolution from archaeabacteria and not by horizontal transfer from proteobacteria (see Doolittle, 1999). Virtually identical topologies were obtained with either a Poisson-corrected distance or a gamma-corrected (z = 1.05) distance.

3.2. New vertebrate sequences

Enolase-1 cDNA sequences and deduced amino acid sequences were generated for six taxa: C. punctatum (shark) α (GenBank Accession No. AY005159) and β (AY005160), A. calva (bowfin) α (AY005152) and β (AY005153), S. trutta (trot) α (AY005161) and β (AY005162), L. chalumnae (coelacanth) α (AY005154), L. paradoxus (South American lungfish) β (AY005157), and N. forsteri (Australian lungfish) β (AY005158). Partial L. paradoxus α (AY005156) and L. chalumnae β (AY005155) sequences (not used in analyses) were also generated.

For the vertebrate enolase nucleotide alignment, 1083 (625 variable) sites were aligned across all species, beginning at site 85 of the complete D. melanogaster cDNA sequence. The deduced amino acid sequence alignment for the vertebrate enolases began at residue 209 of the complete D. melanogaster sequence and contained 360 (217 variable) sites.

The nucleotide and amino acid sequences for Gallus gallus α-enolase were obtained from GenBank, but these sequences differed qualitatively from other sequences. At numerous positions in the amino acid alignment, this sequence has an amino acid residue that differs from either all other enolases (except Drosophila melanogaster) or all other vertebrate enolases. Furthermore, at site 202, G. gallus β-enolase has a histidine (H) residue, which is exceptional among enolases. At this site, D. melanogaster and all vertebrate α-enolases share a lysine (K) residue, β-enolases share an asparagine (N) residue, and γ-enolases share a serine (S) residue. It is possible that some of these differences represent sequence errors. Also, the nucleotide sequence for R. norvegicus β-enolase from GenBank appeared to contain a sequencing and/or alignment error when placed in the overall enolase alignment. To correct the problem, the adenine at position 471 was removed and replaced by an ambiguity at position 489.

The percentage sequence identity was calculated for all cases in which more than one paralog had been completely sequenced for an individual species. Both nucleotide and amino acid sequences showed a high degree of identity (excluding the S. trutta paralogs: 73–80% for nucleotides and 80–86% for amino acids). The trout paralogs showed an exceptionally high identity for both nucleotides (86%) and amino acids (90%).

3.3. Amino acid analyses

A neighbor-joining tree was constructed using amino acid sequences and a gamma-corrected (z = 1.05) distance (Fig. 3). A virtually identical topology was obtained with a Poisson-corrected distance. The two trout sequences clustered with other α-enolase sequences, and each paralog was found to be monophyletic: α-enolase (72%), β-enolase (<50%), and γ-enolase (99%). The maximum likelihood analysis resulted in a single ML tree (ln L = −5614.7), also supporting the monophyly of each paralog, with the exception of fowl
Fig. 2. Phylogenetic tree of enolase-1 amino acid sequences. A gamma-corrected distance ($\alpha = 1.05$) is used with the neighbor-joining method. The tree is rooted with Eubacteria.

(Gallus) beta, which appeared in a basal position on the tree (see comments above about anomalies of this sequence). However, the relationships of the three paralogues could not be resolved. The hagfish sequence clustered with the lamprey-2 sequence, as was found by Kuraku et al. (1999), which provides additional molecu-
Fig. 3. Phylogenetic tree of vertebrate enolase-1 amino acid sequences. A gamma-corrected distance ($\alpha = 1.05$) is used with the neighbor-joining method. Newly sequenced enolases (this study) are indicated in bold. The tree is rooted with *Drosophila melanogaster*.

3.4. Nucleotide analyses

The emphasis in this study was on amino acid analysis because saturation often occurs in nucleotide analyses of distantly-related sequences. None the less, one advantage of nucleotide data is the increased number of sites for analysis, which may provide a better resolution of some aspects of the phylogeny. As in the amino acid tree (Fig. 3), analysis of the nucleotide sequences showed that each of the three paralogs was monophyletic. The two *S. trutta* sequences clustered together significantly and were renamed *S. trutta* $\alpha_1$ and $\alpha_2$.

4. Discussion

4.1. Time of enolase gene duplications within vertebrates

A single enolase gene gave rise to the three enolase paralogs through two closely spaced duplication events. The discovery of multiple isozymes of enolase in acti-
npterygian, sarcopterygian, and chondrichthian fishes
constrains the timing of the duplication events. These
duplications apparently occurred in the chordate lineage
subsequent to the cyclostome-gnathostome divergence
(564 Mya) and prior to the chondrichthian-osteichthian
divergence (528 Mya) (Kumar and Hedges, 1998;
Hedges, 2001). This would place these duplications
during a relatively short period of evolutionary time
near the Proterozoic/Phanerozoic boundary. A much
more recent timing for the duplications, 200–300 Mya,
was estimated previously (Rider and Taylor, 1975;
Clark-Rosenberg and Marangos, 1980; Segel et al.,
1988).

Although the trichotomy remains unresolved, β- and
γ-enolase are presumed to be most closely related based
on intron lengths (Giallongo et al., 1990). If true, and
considering the presence of α- and β-enolase in chon-
drichthians, all living gnathostomes should possess
the three vertebrate enolase paralogs, regardless of the
branching pattern for these paralogs.

Besides enolase, isozymes of other genes also show a
pattern of gene duplication around the time of the
cyclostome-gnathostome divergence (Ono-Koyanagi
et al., 2000). It has been proposed that this burst in
gene duplication led to functionally similar isozymes,
whereas an earlier period of extensive gene duplication,
early the origin of animals, led to di-
V
V
V
V
V
V
V
V
V
V
V

4.2. Early evolution of enolase

Horizontal gene transfer can explain many of the
complex patterns seen in gene and gene-family relation-
ships involving the three domains of life (Feng et al.,
1997; Doolittle, 1999). However, another explanation
that should be considered in some cases is a simple lack
of phylogenetic resolution caused by the inscrutable
limitation of protein length. None the less, two patterns
commonly observed are (1) a close relationship between
eukaryotes and archaeabacteria and (2) a close relation-
ship between eukaryotes and eubacteria (σ-proteobacteria). The first pattern has been attributed
to vertical transmission, whereas the second pattern has
been attributed to horizontal gene transfer subsequent
to the symbiotic origin of mitochondria from an
σ-proteobacterium (Doolittle, 1999). Although not uni-
versally excepted, this has formed a working hypothesis
for the further study of organismal genomes. The rela-
tionships of the three domains are not statistically
resolved in the global analyses of enolase sequences
(Figs. 1 and 2), but the fact that eubacteria are mono-
phyletic and do not cluster closely with eukaryotes
favors a vertical transmission for enolase-1. In addition,
the finding of two enolase genes (each) in the archaeabac-
teria Pyrococcus abyssi and P. horikoshii supports the
interpretation here of an early gene duplication, prior
to the last common ancestor of living organisms, leading
to enolase-1 and enolase-2. If this interpretation is
correct, enolase-2 was subsequently lost in eubacteria
and in some archaeabacteria. Thus, the polyphyly of
archaeabacteria in previous enolase phylogenies
(Hannaert et al., 2000; Keeling and Palmer, 2000) is
interpreted here as an artefact of rooting.

It has been proposed (Rivera and Lake, 1992),
based on some genes showing vertical transmission,
that eukaryotes are closer to some archaeabacteria
(Crenarchaeota) than to other archaeabacteria
(Euryarchaeota). Only one crenarchaeotan enolase
sequence is available (Pseudoalteromonas haloplanktis
A. pernix), and this clus-
ters closely (95%) with other archaeabacteria (Fig. 2).
However, it has an unusually long branch, and therefore,
additional crenarchaeotan sequences are needed to con-
firm this phylogenetic position.

4.3. Origin of salmonid enolases

Because the two copies of enolase from trout (α
salmonid fish) group with σ-enolases in the amino acid
analysis (Fig. 3) and cluster with each other in nucleo-
tide analyses suggests that they are σ-enolases and the
result of a genome duplication event. Salmonid fishes
are known to have undergone a polyploidization event
with subsequent retention of functional gene duplicates
(Bailey et al., 1978). An additional line of evidence
suggesting that these two trout enolases are the result
of genome duplication is that they have a relatively high
percentage sequence identity. The possibility that they
represent γ-enolase can be eliminated because neuronal
tissue was not used as a source for poly(A) mRNA in
this study. Likewise, β-enolase seems unlikely because
the sequences were obtained from liver tissue where only
σ-enolase is believed to be expressed.

The identification of these trout enolase sequences as
belonging to the σ gene and being the result of a
polyploidization event explains the results obtained by
Landrey et al. (1978). A minimum of three paralogs is necessary for an organism to exhibit six isoforms of a dimeric enzyme (e.g., αα, αβ, ββ, γγ, and γγ for enolase). Even though all gnathostomes have been predicted to possess the three standard enolase paralogs, a βγ heterodimer is unlikely, given the tissue-specific localization of these isoforms. Therefore, the result obtained by Landrey et al. (1978) is best explained by genome duplication.

Acknowledgements

We thank Carla A. Hass, Sadhir Kumar, Laura L. Peling, Stephen W. Scharff, Marcell van Tuinen, Billie J. Swalla, and Daniel Wang for comments and discussion; Maureen Lyons-Weiler for laboratory assistance; Linda R. Maxson for donating some taxa, including Latimeria (courtesy of J.A. Musick, R.W. Chapman, The Explorer’s Club, and Virginia Institute of Marine Science) and Neoceratodus (courtesy of William Bemis); and Chris T. Amemiya for samples. This work was supported by grants from NSF and NASA (S.B.H.).

References


Kumar, S., Tamura, K., Nei, M., 1993. MEGA: Molecular Evolutionary Genetic Analysis Version 1.0. The Pennsylvania State Univer-


