

Origin of Tetrapods Inferred from Their Mitochondrial DNA Affiliation to Lungfish

Axel Meyer* and Allan C. Wilson

Division of Biochemistry and Molecular Biology, University of California, Berkeley, California 94720, USA

Summary. This paper shows that questions of an unexpected phylogenetic depth can be addressed by the study of mitochondrial DNA (mtDNA) sequences. For decades, it has been unclear whether coelacanth fishes or lungfishes are the closest living relatives of land vertebrates (Tetrapoda). Segments of mtDNA from a lungfish, the coelacanth, and a ray-finned fish were sequenced and compared to the published sequence of a frog mtDNA. A tree based on inferred amino acid replacements, silent transversions, and ribosomal RNA (rRNA) substitutions showed with statistical confidence that the lungfish mtDNA is more closely related to that of the frog than is the mtDNA of the coelacanth. This result appears to rule out the possibility that the coelacanth lineage gave rise to land vertebrates; hence, morphological characters that link the latter two groups are possibly due to convergent evolution or reversals and not to common descent. Besides supporting the theory that land vertebrates arose from an offshoot of the lineage leading to lungfishes, the molecular tree facilitates an evolutionary interpretation of the morphological differences among the living forms. It would appear that the common ancestor of lungfishes and tetrapods already possessed multiple morphological traits preadapting their locomotion, circulation, and respiration for life on land.

Key words: Direct sequencing — Polymerase chain reaction — Cytochrome *b* — 12S rRNA — Coelacanth — Ray-finned fishes — Frog — Evolutionary

tree — Statistical testing — Morphological traits — Conquest of land

Introduction

The recognition that land vertebrates (tetrapods) evolved from lobe-finned fishes was one of the great achievements of comparative morphology (Romer 1966). For more than a century, however, there has been uncertainty (Forey 1988) about which of the two lobe-finned groups, the *Crossopterygii* or lungfishes, are more akin to tetrapods (Fig. 1, see legend). So long as this matter was unsettled, it was impossible to develop a detailed model of how fishes conquered the land.

Fifty years ago, the discovery of a living species representing an otherwise extinct group of coelacanths (Smith 1939, 1953), which comprise a part of the *Crossopterygii*, created hope that this living fossil was the missing link and close to the ancestry of tetrapods. However, even with extant material at hand, comparative morphologists and paleontologists favored different relationships (Fig. 1A) (Wahlert 1968; Løvtrup 1977; Wiley 1979; Rosen et al. 1981; Lagios 1982; Forey 1987; Fritzsche 1987; Northcutt 1987; Schultze 1987). Although the above question is unsolved, it is generally accepted, as shown in Fig. 1A, that ray-finned fishes (*Actinopterygii*) are more distantly related to the tetrapods and are the sister-group of all lobe-finned fishes (Romer 1966; Forey 1988; but see Wahlert 1968). A molecular approach to defining the genealogical relationships among the four lineages (tetrapod,

* Current address: Department of Ecology and Evolution, State University of New York, Stony Brook, New York 11794, USA
Offprint requests to: Axel Meyer

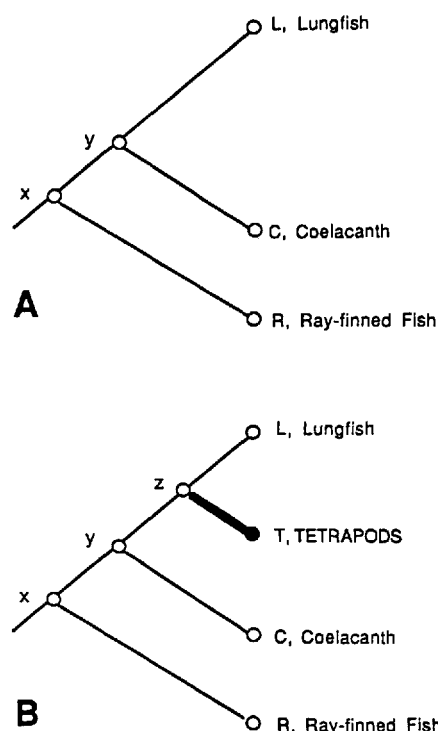


Fig. 1. Phylogenetic relationships of living fishes (Osteichthyes) to one another and to land vertebrates. **A** Tree relating the ray-finned fishes (R) and the two groups of lobe-finned fishes, coelacanths (C) and lungfishes (L). In this tree *x* refers to the common ancestor of all three groups and *y* to the common ancestor of all lobe-finned fishes. In our terminology, *y*L is the lungfish lineage, *y*C is the coelacanth lineage, and *x*R is the ray-finned fish lineage. The question addressed in this paper is whether tetrapods are derived from the *y*L, the *y*C, or the *xy* lineage. Some workers favor the *y*L hypothesis (Wahlert 1968; Rosen et al. 1981), others favor the *y*C hypothesis (Frittsch 1987; Northcutt 1987), whereas still others favor the *xy* hypothesis (Forey 1987); all three hypotheses treat ray-finned fishes as the most distant relatives of tetrapods. **B** Tree relating tetrapods to the fish lineages shown in Fig. 1A. In this tree, *z* refers to the common ancestor of tetrapods and modern lungfish. According to the molecular results reported in this study, tetrapods are an offshoot of the *y*L lineage.

lungfish, coelacanth, and ray-finned fish) is thus in order (Maeda et al. 1984; Hillis and Dixon 1989).

There is growing confidence among molecular evolutionists that when statistical criteria are used to evaluate genealogical trees, biochemical methods can often solve phylogenetic questions definitively (Wilson et al. 1989). The delay in achieving this advance arose in part from technical limitations existing in the era before direct sequencing via the polymerase chain reaction had been invented and also because awareness of the availability of statistical tests was not widespread. For example, no statistical tests were done when two different genes (Morden and Golden 1989; Turner et al. 1989) were used to examine the phylogenetic position of prochlorophytes. Had statistical tests been done on the nucleotide sequences, both genes would have been found to support the same tree, i.e., the tree that links *Fremyella* (a stan-

dard cyanobacterium) more closely than prochlorophytes with green land plants (A.C. Wilson, unpublished). To the best of our knowledge, there are very few cases (for orthologous genes) in which statistically significant conflicts exist among genes as regards genealogical relationships. The exceptions are rare cases of horizontal transfer of genes (e.g., Normand and Bousquet 1989). Therefore, we have used direct sequencing via the polymerase chain reaction to obtain a statistically adequate sample of sequence information with which to answer the question of how tetrapods are related to lobe-finned fishes.

Our use of mitochondrial DNA (mtDNA) for this study may seem surprising because it is known chiefly as a genealogical tool for examining closely related species (Wilson et al. 1985). This ostensible limitation was a consequence of using restriction enzymes for mtDNA comparisons. Since the advent of direct sequencing via the polymerase chain reaction, it has become possible to probe deeper branches in genealogical trees (Kocher et al. 1989; Irwin et al. 1990). Such probing is facilitated by confining attention to base positions that change very slowly (e.g., replacement sites in protein-coding genes and most sites in rRNA genes) and to base changes of the sort that accumulate slowly (i.e., transversions).

Materials and Methods

From the Midas cichlid fish [*Cichlasoma citrinellum*, Actinopterygii (ray-finned fish)], the South American lungfish (*Lepidosiren paradoxa*, Dipnoi), and the coelacanth (*Latimeria chalumnae*, Actinistia) DNA was extracted as described (Kocher et al. 1989) from tissues taken from frozen specimens. [Formalin-fixed tissues of museum specimens of *Latimeria* and the other two extant genera of lungfishes (*Protopterus* and *Neoceratodus*) have been used as well. The sequences obtained from these materials were shorter and supported the conclusions based on the frozen specimens (A. Meyer and A.C. Wilson, unpublished). Partial sequences of the mitochondrially encoded cytochrome oxidase III gene and of the nuclear 18S rRNA were also obtained. They too support the conclusions based on our studies of the genes for cytochrome *b* and 12S rRNA (A. Meyer and A.C. Wilson, unpublished).]

Amplifications and Direct Sequencing. Amplifications were done in 25 μ l of Tris (67 mM, pH 8.8) containing $MgCl_2$ (2 mM), 1 mM of each dNTP, 1 μ M of each primer, template DNA (10–1000 ng), and *Taq* polymerase (1.25 units, Perkin-Elmer-Cetus). The primer sequences for the 12S rRNA are L1091 5'-AAAAAGCTTCAAAGTGGGATTAGATACCCCACTAT-3' and H1478 5'-TGACTGCAGAGGGTGACGGGCGGTGTGT-3' (Kocher et al. 1989); the sequences of the cytochrome *b* primers are L14724 5'-CGAAGCTTGATATGAAAAACCATCGTTG-3' (designed by S. Pääbo) and H15149 5'-AAACTGCAGCCCC-TCAGAATGATATTTGTCCTCA-3' (Kocher et al. 1989). L and H refer to the light and heavy strands, respectively, and the numbers refer to the 3' position of the primers in human mtDNA (Anderson et al. 1981). The methods by which single-stranded

implication is that the sequences compared are not deeply within the multiple-hit zone and therefore may be used for phylogenetic analysis.

Considering the cytochrome and rRNA results together, there are 198 slowly changing positions suitable for phylogenetic analysis. Of these, 33 were phylogenetically informative (Fig. 2). That is, at each such site, the three possible trees relating the four species (see legend of Fig. 1) differ as regards the minimum number of mutations needed to explain the variation at that site.

For 19 of these 33 sites, tree B in Fig. 1 explains the variation with a single mutation per site. By contrast, for only seven of the sites can the sequence variation be explained with one mutation when either of the alternative trees is used. The likelihood that a score of 19:7 could arise by chance in 26 trials is less than 0.015 (Prager and Wilson 1988). Hence tree B in Fig. 1 is significantly better than either of the alternative trees at explaining the sequence variation.

A parsimony analysis (Swofford 1989), with the ray-finned fish as an outgroup, also favored the joining of the lungfish and tetrapod lineages over the other two possibilities when only positions with amino acid replacement changes in cytochrome and the first 240 bp of rRNA were used. Tree B in Fig. 1 required 208 steps at the 154 variable positions (consistency index = 0.952), whereas the two other possible trees (not shown) were 13 and 11 steps longer. A bootstrapping test that, given the use of a large number of phylogenetically informative characters, allows one to attach a probability value to a branching order of a tree (Felsenstein 1985) was conducted using the computer program PAUP (Swofford 1989). The branching order in tree B received support in 99–100 out of 100 trials. Distance trees constructed by the neighbor-joining (Saitou and Nei 1987) and the distance Wagner (Farris 1972) methods using the sums of the rRNA differences and the replacement differences in cytochrome *b* also supported tree B. We did not find a base-compositional bias in the data that could account for the lungfish–tetrapod affiliation. The result of our analysis was unchanged by the inclusion or exclusion of the last 64 bp of the 12S rRNA sequence shown in Fig. 2.

Discussion

This study demonstrates that the phylogenetic information contained in mtDNA can reach further back into the evolutionary history of vertebrates than was suspected and can answer an old evolutionary question with statistical confidence. This evolutionary depth might only be attainable in cold-

blooded vertebrates, as evidence is accumulating that the rate of amino acid replacement in mtDNA is lower in these organisms than in warm-blooded vertebrates (Kocher et al. 1989; Thomas and Beckenbach 1989; Johansen et al. 1990). A lower rate of sequence change reduces the likelihood of parallelism and reversals and thus raises the resolving power of a genealogical analysis (Prager and Wilson 1988). Accordingly, if a mouse sequence is used instead of the frog, tree B still wins, but less decisively.

One implication of our findings is that the coelacanth lineage (γC , Fig. 1) was not the direct source of the tetrapod lineage, whereas the lineage leading to modern lungfishes (γL , Fig. 1) was. Another possible implication emerges from the surprising strength of the lungfish–tetrapod link, namely that their common stem lineage (γZ , Fig. 1) may reach back further in time than the current interpretations of the fossil record suggest. Traditionally, the period of common ancestry represented by the γZ lineage in Fig. 1B tended to be regarded as probably short, in comparison to that represented by the zL or zT lineages, the reason being that lungfishes, crossopterygians, and actinopterygians all appear in the fossil record about 400 million years ago (e.g., Romer 1966).

Morphological Traits Apportioned on the Molecular Tree

With the framework provided by the molecular tree (Fig. 1B), it is possible to gain a better understanding of the morphological transition from water to land, i.e., the loss or gain of morphological traits during the evolution of tetrapods from lobe-finned ancestors and the selective forces that made common ancestors of living lungfishes and tetrapods particularly apt to conquer the land. Thus, the states of 22 morphological traits (from living members of the groups considered) listed in Table 2 can now be allocated to lineages on the molecular tree.

In accordance with the parsimony principle, we postulate that each of the first 14 traits underwent a single change on the stem lineage (γZ , Fig. 1) leading to the common ancestor (z , Fig. 1) of lungfishes and tetrapods. These changes include gain of the abilities to breathe through internal nostrils, to control access to the trachea with a glottis, and thus to feed as a tetrapod does while breathing. Changes in the circulatory system preadapted them to life on land by beginning the separation of their bloodstream into oxygenated blood coming from their lungs and reduced gills and deoxygenated blood coming from the rest of their body. In addition it would appear that their locomotion had become in some respects tetrapod-like, through fusion of their

Table 2. Distribution of 22 morphological traits in extant members of the four taxonomic groups

Trait	State of trait			
	L	T	C	R
Common ancestral (head)				
1 Internal nostrils	+	+	-	-
2 Palate bone fused with neurocranium	+	+	-	-
3 Glottis	+	+	-	-
4 Pharyngobranchial gill arch elements	-	-	+	+
5 Autopalatine bone	-	-	+	+
6 Depressor mandibulae muscle	+	+	-	-
7 Free hyomandibular bone	+	+	-	-
8 Ethmoid commissure sensory canal	-	-	+	+
9 Saccus vasculosus of pituitary gland	-	-	+	+
10 Pars nervosa of pituitary gland	+	+	-	-
Common ancestral (body)				
11 Truncus arteriosus of heart	+	+	-	-
12 Divided auricle of heart	+	+	-	-
13 Limbs with more than four mesomeres	+	+	-	-
14 Pelvic girdles joined	+	+	-	-
Parallelisms or reversals				
15 Labial pit and muscular lip-fold	+	-	+	-
16 Electroreceptors mostly on snout	+	-	+	-
17 Septum ependymale dividing brain hemispheres	+	-	+	-
18 Thickened dorsal thalamus	+	-	+	-
19 Maxilla bone	-	+	-	+
20 Short dentary bone	+	-	+	-
21 Glenoid convex	+	-	+	-
22 Endolymphatic commissure and basilar papilla	-	+	+	-

The four groups are L, lungfishes (Dipnoi); T, Tetrapoda; C, coelacanth; R, ray-finned fishes (Actinopterygii). Traits were sorted by whether the common ancestor of the lungfish and tetrapod lineage is proposed (on the grounds of parsimony) to have them. States of the traits (+ meaning present and - meaning absent) were kindly compiled by P. Forey based on Rosen et al. (1981), Lagios (1982), Forey (1987), Fritsch (1987), Northcutt (1987), and Schultze (1987). For each trait two taxonomic groups are in one state and two are in the other state. Traits 1-14 favor the grouping of lungfish with tetrapods (Fig. 1B). For each of these traits, the distribution of states can be explained most simply by supposing that there was a single change from - to + (for traits 1-3, 6, 7, and 10-14) or from + to - (for traits 4, 5, 8, and 9) on the stem lineage (yz in Fig. 1B) leading to the common ancestor of extant lungfishes and tetrapods. To explain the distribution of states observed for traits 15-22, a minimum of two changes of state on different lineages is required for each trait. Not appearing in this table are traits for which three of the groups share one state while the fourth has the other state. Such trait distributions can in every case be explained by a single change on a peripheral branch in the unrooted tree. An example would be the legged state, which is unique to tetrapods, in contrast to the finned state of the limbs in the three groups of fishes. Analysis of such traits does not allow one to assign evolutionary events to the stem lineages in Fig. 1

pelvic girdles and the incorporation of more mesomeres into the limbs.

By contrast, as one would expect, there are other traits (15-22 in Table 2) whose distribution among the four extant taxa calls for a more complex explanation. In these eight cases our model requires parallelisms or reversals on two or more of the peripheral lineages in tree B, i.e., a minimum of two changes per trait, neither of which can be assigned to the stem lineage (yz).

Relationships of Fossils to Extant Lineages

This discussion has so far ignored rhipidistians, a third group of lobe-finned fishes, which are traditionally classified with coelacanths in the Crossopterygii (Romer 1966). Rhipidistians are widely be-

lieved to be directly ancestral to tetrapods (Romer 1966; Schultze 1987).

A goal for future research is to find out how rhipidistians are related to the lineages in tree B (Fig. 1). Because rhipidistians are a long-extinct group, this goal must be achieved by morphological analysis of fossils. It may be possible to link some of the rhipidistians with statistical confidence to specific lineages in tree B, for example, the yC lineage (i.e., retaining the crossopterygians as a natural monophyletic group), the zT lineage (i.e., retaining direct rhipidistian ancestry to tetrapods), or the yz lineage (i.e., implying a rhipidistian ancestry for both lungfishes and tetrapods).

Once the attachment of a rhipidistian to a specific lineage in tree B (Fig. 1) is achieved, morphological changes previously attributed to that lineage can then be ordered into two categories, those that orig-

inated before the rhipidistian junction and those that occurred afterward. Knowledge of the sequence of morphological changes involved in the colonization of land might in this way be refined.

Acknowledgments. We thank R. Murphy of the Royal Ontario Museum for the coelacanth tissue, D. Reznik of the University of California at Riverside for the lungfish tissue, S. Pääbo for the design of one primer, and W. Fitch, P. Forey, C. Patterson, D. Wake, M. Wake, and especially E. Prager for helpful advice and discussion. P. Forey kindly compiled the morphological data in Table 2. We thank D. Hillis for informing us that he has obtained evidence from nuclear gene sequences consistent with the phylogeny presented in Fig. 1B. This work received support from National Science Foundation and National Institutes of Health grants to A.C.W. and a Sloan postdoctoral fellowship to A.M.

References

- Anderson S, Bankier AT, Barrell BG, de Bruijn MHL, Coulson AR, Drouin J, Eperon IC, Nierlich DP, Roe BA, Sanger F, Schreier PH, Smith AJH, Staden R, Young IG (1981) Sequence and organization of the human mitochondrial genome. *Nature* 290:457–465
- Farris JS (1972) Estimating phylogenetic trees from distance matrices. *Am Nat* 106:645–668
- Felsenstein J (1985) Confidence limits on phylogenies: an approach using the bootstrap. *Evolution* 39:783–791
- Forey PL (1987) Relationships of lungfishes. *J Morph Suppl* 1: 75–91
- Forey PL (1988) Golden jubilee for the coelacanth *Latimeria chalumnae*. *Nature* 336:727–732
- Fritzsche B (1987) Inner ear of the coelacanth fish *Latimeria* has tetrapod affinities. *Nature* 327:153–154
- Hillis DM, Dixon MT (1989) Vertebrate phylogeny: evidence from 28S ribosomal DNA sequences. In: Fernholm B, Bremer K, Jörnvall H (eds) *The hierarchy of life*. Elsevier, Amsterdam, pp 355–367
- Hixson JE, Brown WM (1986) A comparison of the small ribosomal RNA genes from the mitochondrial DNA of the great apes and humans: sequences, structure, evolution, and phylogenetic implications. *Mol Biol Evol* 3:1–18
- Howell N, Gilbert K (1988) Mutational analysis of the mouse mitochondrial cytochrome *b* gene. *J Mol Biol* 203:607–618
- Irwin DM, Kocher TD, Wilson AC (1990) Evolution of the cytochrome *b* gene of mammals. *J Mol Evol* (in press)
- Johansen S, Guddal PH, Johansen T (1990) Organization of the mitochondrial genome of Atlantic cod, *Gadus morhua*. *Nucleic Acids Res* 18:411–419
- Kocher TD, Thomas WK, Meyer A, Edwards SV, Pääbo S, Vialablanca FX, Wilson AC (1989) Dynamics of mitochondrial DNA evolution in animals: amplification and sequencing with conserved primers. *Proc Natl Acad Sci USA* 86:6196–6200
- Lagios MD (1982) *Latimeria* and the Chondrichthyes as sister taxa: a rebuttal to recent attempts of refutation. *Copeia* 1982: 942–948
- Løvtrup S (1977) *The phylogeny of vertebrata*. Wiley, London
- Maeda N, Zhu D, Fitch WM (1984) Amino acid sequences of lower vertebrate parvalbumins and their evolution: parvalbumins of boa, turtle, and salamander. *Mol Biol Evol* 1:473–488
- Morden CW, Golden SS (1989) *psbA* genes indicate common ancestry of prochlorophytes and chloroplasts. *Nature* 337: 382–385
- Normand P, Bousquet J (1989) Phylogeny of nitrogenase sequences in *Frankia* and other nitrogen-fixing microorganisms. *J Mol Evol* 29:436–447
- Northcutt RG (1987) Lungfish neural characters and their bearing on sarcopterygian phylogeny. *J Morph Suppl* 1:277–297
- Prager EM, Wilson AC (1988) Ancient origin of lactalbumin from lysozyme: analysis of DNA and amino acid sequences. *J Mol Evol* 27:326–335
- Roe BA, Ma D-P, Wilson RK, Wong JF-H (1985) The complete nucleotide sequence of the *Xenopus laevis* mitochondrial genome. *J Biol Chem* 260:9759–9774
- Romer AS (1966) *Vertebrate paleontology*, 3rd ed. Univ of Chicago Press, Chicago
- Rosen DE, Forey PL, Gardiner BG, Patterson C (1981) Lungfishes, tetrapods, paleontology, and plesiomorphy. *Bull Am Mus Nat Hist* 167:159–276
- Saitou N, Nei M (1987) The neighbor-joining method: a new method for reconstructing phylogenetic trees. *Mol Biol Evol* 4:406–425
- Schultze H-P (1987) Dipnoans as sarcopterygians. *J Morph Suppl* 1:39–74
- Smith JLB (1939) A surviving fish of the order Actinistia. *Nature* 143:455–456
- Smith JLB (1953) The second coelacanth. *Nature* 171:99–101
- Swofford DL (1989) PAUP: phylogenetic analysis using parsimony, version 3.0b. Illinois Natural History Survey, Champaign IL
- Thomas WK, Beckenbach AT (1989) Variation in salmonid mitochondrial DNA: evolutionary constraints and mechanisms of substitution. *J Mol Evol* 29:233–245
- Turner S, Burger-Wiersma T, Giovannoni SJ, Mur LR, Pace NR (1989) The relationship of a prochlorophyte *Prochlorothrix hollandica* to green chloroplasts. *Nature* 337:380–382
- Wahlert G von (1968) *Latimeria* und die Geschichte der Wirbeltiere. Fischer, Stuttgart, West Germany
- Wiley EO (1979) Ventral gill arch muscles and the interrelationships of gnathostomes, with a new classification of the Vertebrata. *Zool J Linn Soc* 67:149–179
- Wilson AC, Cann RL, Carr SM, George M, Gyllenstein UB, Helmbachowski KM, Higuchi RG, Palumbi SR, Prager EM, Sage RD, Stoneking M (1985) Mitochondrial DNA and two perspectives on evolutionary genetics. *Biol J Linn Soc* 26:375–400
- Wilson AC, Zimmer EA, Prager EM, Kocher TD (1989) Restriction mapping in the molecular systematics of mammals: a retrospective salute. In: Fernholm B, Bremer K, Jörnvall H (eds) *The hierarchy of life*. Elsevier, Amsterdam, pp 407–419

Received May 29, 1990/Accepted June 15, 1990