CHAPTER 18

Molecular evidence on the origin of and the phylogenetic relationships among the major groups of vertebrates

Rafael Zardoya and Axel Meyer

Vertebrates are a good model to study macroevolutionary patterns and processes because they possess a comparatively well known fossil record (Carroll 1997). Thanks to the detailed investigations of several generations of morphologists and paleontologists over the last two centuries, it has been possible to reconstruct the phylogeny of vertebrates with some degree of confidence (Fig. 18.1). A robust phylogenetic framework of vertebrates is fundamental for comparative studies in this group.

The first major landmark in vertebrate evolution was the origin of jaws from mandibular branchial arches, and dates back to the Cambrian, 540–505 million years ago (mya). Accordingly, vertebrates have been traditionally classified into Agnatha (hagfishes and lampreys) and Gnathostomata (jawed vertebrates) (Fig. 18.1). Among the latter, the major distinction, based on the nature of the skeleton, is between Chondrichthyes (cartilaginous fishes) and Osteichthyes (bony fishes) (Fig. 18.1). Bony fishes are divided into Actinopterygii (ray-finned fishes) and Sarcopterygii (lobe-finned fishes and tetrapods) (Fig. 18.1). The origin of the major lineages within bony fishes as well as the transition to life on land date back to the Devonian, 408–360 mya. The major evolutionary novelty within tetrapods was the origin of the amniote egg, a type of egg with a semipermeable shell, a large amount of yolk, and several embryonic membranes involved in respiration, feeding, and waste disposal. Accordingly, extant tetrapods are classified into Lissamphibia (caecilians, salamanders, and frogs) and Amniota (mammals and reptiles, that is, turtles, squamates, crocodiles, and birds) (Fig. 18.1). The origin of amniotes dates back to the Pennsylvanian, 325–280 mya.

Although the general framework of vertebrate relationships is well supported by morphological and paleontological evidence, and new fossil discoveries continue to refine it, some key branching events in the vertebrate tree remain controversial (see **polytomies**, that is, collapsed nodes in Fig. 18.1). Appearance of vertebrate taxa in the fossil record suggests that the origin of the major lineages of vertebrates occurred as rapid radiation events within a relatively short time frame several hundred million years ago (Carroll 1997). In such cases, it is difficult to recover the exact branching pattern because there was little time during the radiation to accumulate shared derived characters that define lineages, and later much of the phylogenetic information was obliterated during the independent evolution of each lineage. In the particular case of vertebrates, the relative phylogenetic positions of jawless vertebrates, lobe-finned fishes, amphibians, turtles, and monotremes are still under debate.

Recently, the advent of molecular techniques (in particular, the **polymerase chain reaction** or PCR, and automated DNA sequencing) has made it possible to tackle these phylogenetic questions from a different perspective. Two molecular markers, mitochondrial DNA and nuclear rRNA genes, are
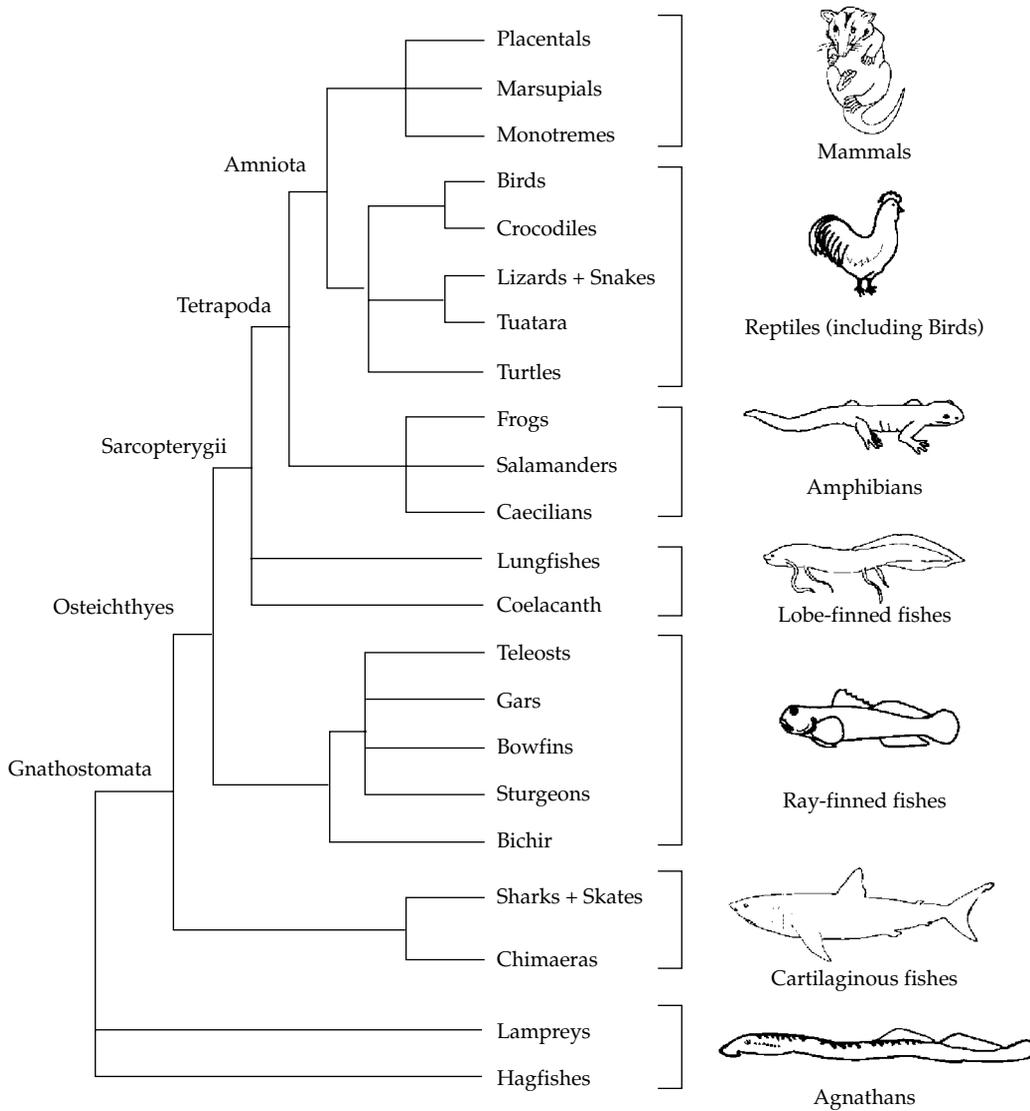
Figure 18.1 Phylogenetic relationships among the main lineages of vertebrates based on morphological and paleontological evidence (e.g. Carroll 1988; Cloutier and Ahlberg 1997).

the most widely used in phylogenetic studies among distantly related taxa. Phylogenetic analyses of mitochondrial genomes and nuclear rRNA data have largely corroborated the traditional morphology-based phylogeny of vertebrates (Fig. 18.1), and they are contributing to the resolution of some long-standing controversies in vertebrate large-scale systematics. In this chapter, we review the molecular data that have been collected with the explicit goal of resolving competing hypotheses that have been postulated to explain the origin of the main lineages of vertebrates.

The living sister group of jawed vertebrates

The origin of jaws, a key innovation that allowed gnathostomes to grasp and bite prey, was one of the

major events in the history of vertebrates. Among extant vertebrates only the hagfishes (Mixiniformes; 43 species) and the lampreys (Petromyzontiformes; 41 species) remain jawless. The phylogenetic relationships of hagfishes, lampreys, and jawed vertebrates have given rise to one of the longest controversies in vertebrate systematics. There are three problems affecting the recovery of the exact relationships between these taxa (Mallat and Sullivan 1998). First, these three lineages of vertebrates appeared in a time window of less than 40 million years, back in the Cambrian, so it is difficult to find shared derived characters between them. Second, there is a rather poor fossil record to trace this event, and third, both lampreys and hagfishes not only have retained many primitive characters, but also, and because of their apparently fast evolutionary rates, they have accumulated numerous unique characters.

Traditional classifications of vertebrates united hagfishes and lampreys as cyclostomes (Fig. 8.2a). Such grouping was supported by the presence in both taxa of horny teeth, a respiratory velum, and a complex "tongue" apparatus (Delarbre *et al*. 2000). However, more recent morphological analyses found many seemingly shared derived characters between lampreys and jawed vertebrates (Janvier 1981) to the exclusion of hagfishes (Fig. 18.2b). The basal position of hagfishes with respect to lampreys

and jawed vertebrates was also supported in a recent analysis of agnathan fossils from China (Shu *et al*. 1999).

Several molecular studies have been conducted to resolve the question of the living sister group of gnathostomes. Phylogenetic analyses of the nuclear 18S and 28S rRNA genes recovered a monophyletic cyclostome (hagfish + lamprey) **clade** with high statistical support (Mallat and Sullivan 1998; Mallat *et al*. 2001; Zardoya and Meyer 2001b) (Fig. 18.2a). Several other nuclear loci also support the cyclostome hypothesis (Kuraku *et al*. 1999) (Fig. 18.2a). In contrast, phylogenetic analyses of mitochondrial protein-coding genes seemed to support lampreys as the closest sister group of jawed vertebrates (Rasmussen *et al*. 1998) (Fig. 18.2b). An important problem in the reconstruction of early vertebrate phylogeny using molecular data is that the hagfish branch is extremely long (Zardoya and Meyer 2001b). This circumstance could spuriously pull the highly divergent hagfish sequence toward the outgroup (the next more divergent sequence in the tree), and may explain the phylogeny recovered by Rassmussen *et al*. (1998). In fact, more recent phylogenetic analyses with additional taxa demonstrated that mitochondrial evidence can support both competing hypotheses on the relative position of jawless vertebrates depending on the choice of method of phylogenetic inference (Delarbre *et al*. 2000).
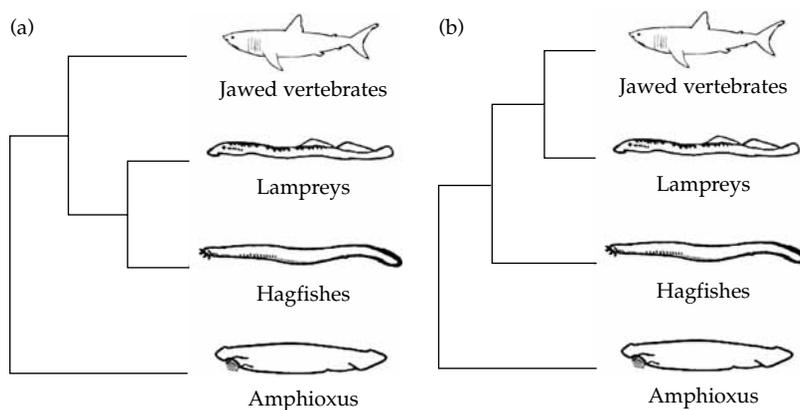

Figure 18.2 Competing hypotheses on the phylogenetic relationships of extant hagfishes, lampreys, and jawed vertebrates. (a) The Cyclostome hypothesis: hagfishes and lampreys are sister group taxa and equally distant to jawed vertebrates. (b) The Vertebrate hypothesis: lampreys are the living sister group of jawed vertebrates.

The origin of tetrapods

The origin of land vertebrates occurred as a rapid adaptive radiation back in the Devonian (408–360 mya). The transition from life in water to life on land was a complex evolutionary event that involved successive morphological, physiological, and behavioral changes. It is well established that early tetrapods evolved from lobe-finned fishes (Sarcopterygii), and recent fossil discoveries have shown that panderichthyids are their closest sister group (Cloutier and Ahlberg 1997; Ahlberg and Johanson 1998). The closest relatives of panderichthyids and tetrapods are osteolepiforms (Cloutier and Ahlberg 1997; Ahlberg and Johanson 1998). The other two major groups within sarcopterygians are Dipnomorpha and Actinistia. Dipnomorphs include the extinct porolepiforms, and the air-breathing extant dipnoi (lungfishes). Actinistia or coelacanths were a highly successful group of lobe-finned fishes during the Devonian which now are represented by only two surviving species (*Latimeria chalumnae* and *Latimeria menadoensis*). Although most recent morphological and paleontological evidence support lungfishes as the closest living sister group of tetrapods (Cloutier and Ahlberg 1997; Ahlberg and Johanson 1998),

there is still no general agreement regarding which group of sarcopterygians, the Actinistia or the Dipnomorpha, is the one most closely related to the tetrapod lineage. The controversy will continue until new relevant fossils of intermediate forms connecting the three groups are discovered, and agreement among paleontologists about the **homology** of some characters (e.g. the choanae) is achieved (Cloutier and Ahlberg 1997).

Molecular data from lungfishes, the coelacanth, and tetrapods, the only living sarcopterygians, have been collected to bear on the phylogenetic question. There are three competing hypotheses to explain phylogenetic relationships among the living lineages of sarcopterygians: (1) lungfishes as closest living relatives of tetrapods (Fig. 18.3a), (2) the coelacanth as living sister group of tetrapods (Fig. 18.3b), and (3) lungfish and coelacanth equally closely related to tetrapods (Fig. 18.3c).

The first molecular data set that supported lungfishes as the closest living relatives of tetrapods (Fig. 18.3a) was based on two fragments of the mitochondrial 12S rRNA and cytochrome *b* genes (Meyer and Wilson 1990). Further support for this hypothesis was obtained from the phylogenetic analysis of complete 12S and 16S rRNA mitochondrial genes (Hedges *et al*. 1993). However, a reanalysis of this
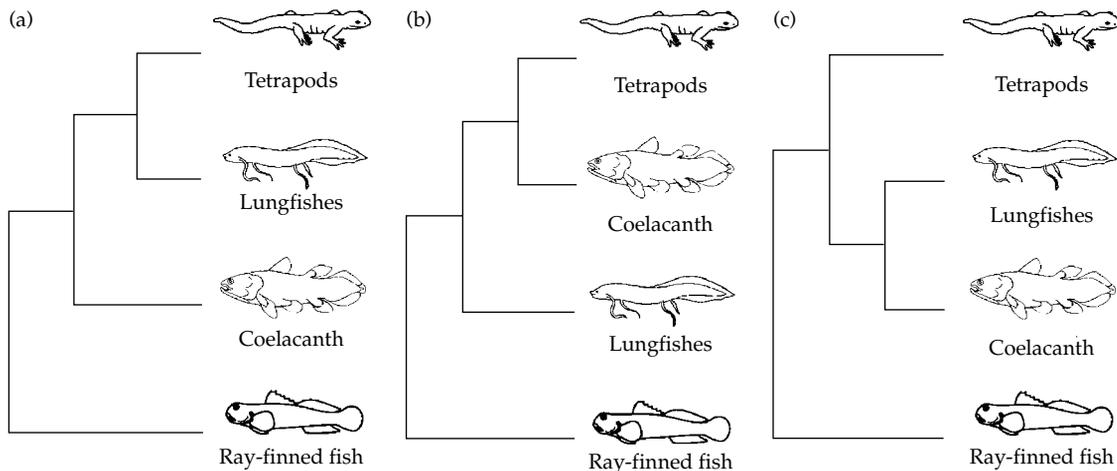

Figure 18.3 Alternative hypotheses on the phylogenetic relationships of living lobe-finned fishes and tetrapods. (a) Lungfishes as the closest living sister group of tetrapods. (b) Coelacanths as the closest living relatives of tetrapods. (c) Lungfishes and coelacanths equally closely related to tetrapods.

data set with more taxa resulted in an unresolved lungfish + coelacanth + tetrapod trichotomy (Zardoya *et al*. 1998). Phylogenetic analyses of a data set that combined all mitochondrial protein-coding genes recovered lungfishes as sister group of tetrapods (Zardoya *et al*. 1998) (Fig. 18.3a). However, this data set could not statistically reject a lungfish + coelacanth clade (Fig. 18.3c) but could do so for the traditional hypothesis that has been favored by most textbooks for the past 50 years (Fig. 18.3b). Phylogenetic analyses of a data set that combined all mitochondrial tRNA genes supported a close relationship between lungfishes and the coelacanth (Zardoya *et al*. 1998) (Fig. 18.3c). When the mitochondrial protein-coding gene data set was combined with the rest of the mitochondrially encoded (rRNA and tRNA) genes, it also supported lungfishes as the closest living sister group of tetrapods (Zardoya *et al*. 1998). Phylogenetic analyses of nuclear 28S rRNA gene sequences favored a lungfish + coelacanth clade (Zardoya and Meyer 1996). The phylogenetic analyses of the combined mitochondrial and 28S rRNA nuclear data sets were not entirely conclusive. Depending on the method of phylogenetic inference used, both a lungfish + coelacanth or a lungfish + tetrapod clade were supported (Zardoya *et al*. 1998). The hypothesis of the coelacanth as closest living sister group of tetrapods (Fig. 18.3b) received the least support in all phylogenetic analyses of molecular data. Recent phylogenetic analyses of a nuclear gene, the myelin DM20 again supported lungfishes as the sister group of tetrapods (Tohyama *et al*. 2000) (Fig. 18.3a). Moreover, the lungfish + tetrapod clade is supported by a single **deletion** in the amino acid sequence of a nuclear-encoded gene RAG2 that is shared by lungfishes and tetrapods (Venkatesh *et al*. 2001). Overall, most molecular evidence supports lungfishes as the closest living sister group of tetrapods. However, further work on nuclear genes is required to discard definitively a lungfish + coelacanth relationship.

Phylogenetic relationships of modern amphibians

There is little controversy that modern amphibians (Lissamphibia) are a monophyletic group that likely appeared in the Permian (280–248 mya) (Duellman and Trueb 1994; but see Carroll 1988). However, it is still debated whether the extinct temnospondyls (e.g. Trueb and Cloutier 1991) or the extinct lepospondyls (Laurin and Reisz 1997; Laurin 1998) are their stem group. Moreover, there is no general agreement regarding the phylogenetic relationships among the three living orders of amphibians, that is, Gymnophiona (caecilians), Caudata (salamanders), and Anura (frogs). Most morphological and paleontological studies support that salamanders are the living sister group of frogs (and form the clade Batrachia) and that caecilians are more distantly related to both (Trueb and Cloutier 1991; Duellman and Trueb 1994) (Fig. 18.4a). However, other studies seem to suggest that salamanders can be the sister group of caecilians to the exclusion of frogs (Carroll 1988) (Fig. 18.4b). The latter hypothesis finds support in Laurin's (1998) analysis, although the author suggests that this may be a spurious result. Because all three lineages of extant amphibians acquired their distinctive body plans early on their evolutionary history, there are few reliable shared derived characters among them. Moreover, a rather poor Permian–Triassic fossil record complicates the search for putative Lissamphibian relatives (Carroll 1988).

Early phylogenetic analyses of this question using nuclear and mitochondrial rRNA data suggested that caecilians are the closest living relatives of salamanders to the exclusion of frogs (Feller and Hedges 1998) (Fig. 18.4b). However, phylogenetic analyses of the complete mitochondrial genomes of a salamander (*Mertensiella luschani*), a caecilian (*Typhlonectes natans*), and a frog (*Xenopus laevis*) supported with high statistical support the Batrachia hypothesis, that is, a sister group relationship of salamanders and frogs (Zardoya and Meyer 2001a) (Fig. 18.4a), in agreement with most morphological evidence rather than with earlier molecular studies. The current overall support of the Batrachia hypothesis both from morphology and molecules provides a phylogenetic framework that will be helpful in many comparative studies of living amphibians. Unfortunately, molecular data cannot provide information on the question of the closest relative of Lissamphibia which requires the
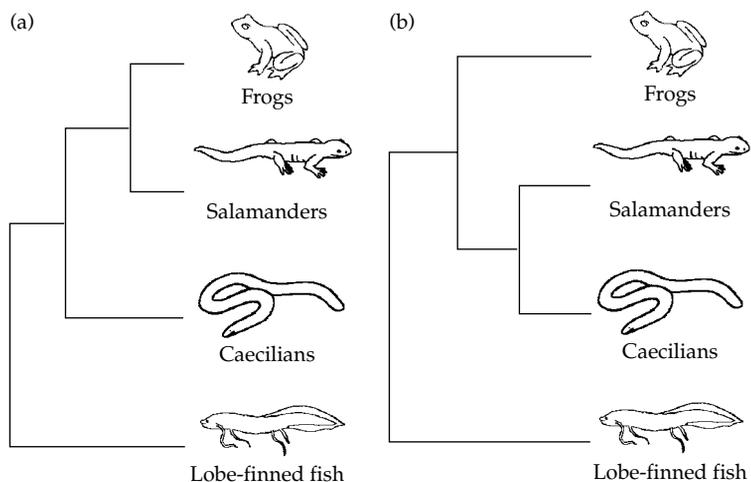
Figure 18.4 Phylogenetic relationships of the three living groups of amphibians. (a) The Batrachia hypothesis: frogs as the closest living sister group of salamanders. (b) Caecilians as the closest extant relatives of salamanders.

discovery of key Permian–Triassic fossils, and their rigorous phylogenetic analysis.

The phylogenetic position of turtles and amniote relationships

Living amniotes are classified, based on the presence and type of temporal fenestration of the skull, in anapsids (those that show a completely roofed skull), diapsids (those that have two fenestrae in the temporal region of the skull), and synapsids (those that present a single lower temporal hole in their skulls). Turtles are considered to be the only living representatives of anapsids; the tuatara, lizards, snakes, crocodiles, and birds are diapsids; and mammals are synapsids. The classical view of amniote phylogeny supported by morphological and fossil data considers synapsids as the most basal lineage, and places diapsids in a derived position relative to anapsids (Laurin and Reisz 1995; Lee 1997) (Fig. 18.5a).

However, turtles exhibit such a unique morphology that they only share a few characters with any other group of amniotes. As a result, it is difficult to determine the exact phylogenetic position of turtles within the amniota. Recent analyses reveal that support for the anapsid affinities of turtles is rather weak (deBraga and Rieppel 1997). Alternatively, turtles might be the closest living relatives of

the Lepidosauria (tuatara, lizards, and snakes) (Fig. 18.5b) (deBraga and Rieppel 1997), or the sister group of Archosauria (crocodiles and birds) (Fig. 18.5c) (Hennig 1983).

If turtles are anapsid reptiles, their closest relatives may be procolophonids (Laurin and Reisz 1995) or pareiasaurs (Gregory 1946; Lee 1997). If turtles are placed as advanced diapsid reptiles, they may be closely related to extinct Sauropterygia (marine plesiosaur and pliosaur reptiles), and Lepidosauria would be their closest living relatives (deBraga and Rieppel 1997) (Fig. 18.5b). However, there is also morphological evidence that places plesiosaurs and pliosaurs basal to the Archosauria (Merck 1997). Hence, it is possible that turtles could be closely related to the Archosauria (Hennig 1983) (Fig. 18.5c) rather than to the Lepidosauria. In both cases, if turtles are diapsids, then the anapsid condition of the turtle skull was developed secondarily.

Although early analyses of mitochondrial complete 12S and 16S rRNA gene data sets supported the traditional anapsid position of turtles, that is, outside diapsids (e.g. Hedges 1994) (Fig. 18.5a), recent reanalyses of the same genes with additional taxa (including representatives of the two major lineages of turtles, Pleurodira and Cryptodira) recover a turtle + Archosauria clade with moderately high statistical support (Zardoya and Meyer 1998) (Fig. 18.5c). However, alternative hypotheses, that is, turtles as
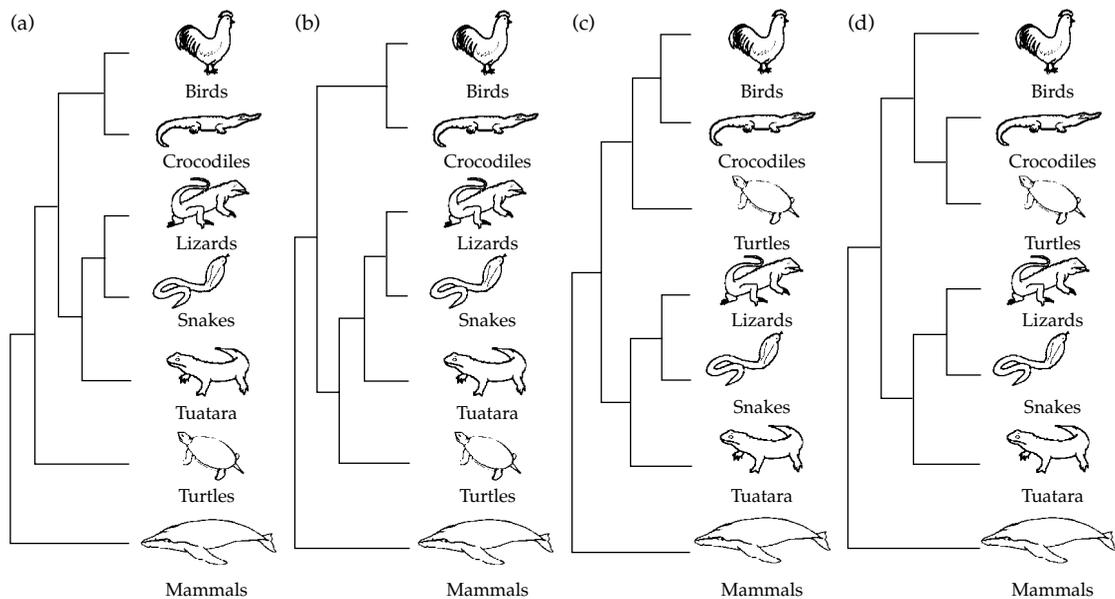
Figure 18.5 The phylogenetic position of turtles within living amniotes. (a) Turtles as the only living representatives of anapsid reptiles, and basal to diapsid reptiles, that is, Lepidosauria (tuatara, snakes, and lizards) and Archosauria (crocodiles and birds). (b) Turtles have diapsid affinities, and are the sister group of Lepidosauria. (c) Turtles have diapsid affinities, and are the sister group of Archosauria. (d) Turtles have archosaurian affinities, and are the sister group of crocodiles.

anapsids (Fig. 18.5a) or turtles as sister group of lepidosaurs (Fig. 18.5b), could not be statistically rejected based on this data set (Zardoya and Meyer 1998). Recent phylogenetic analyses of relatively large mitochondrial and nuclear sequence data sets further support the diapsid affinities of turtles, and only differ on their relative position with respect to Lepidosauria and Archosauria. Molecular evidence based on complete mitochondrial protein-coding genes confirmed the archosaurian affinities of turtles, and statistically rejected alternative hypotheses (Kumazawa and Nishida 1999; Janke *et al*. 2001) (Fig. 18.5c). Phylogenetic analyses of a data set including complete mitochondrial protein-coding, rRNA, and tRNA genes also strongly supported the phylogenetic position of turtles as sister group of archosaurs (Zardoya and Meyer 2001b) (Fig. 18.5c). In agreement with mitochondrial evidence, nuclear pancreatic polypeptide data support archosaurs as the living sister group of turtles (Platz and Conlon 1997).

Phylogenetic analyses of eleven nuclear proteins, as well as the nuclear 18S and 28S rRNA genes, supported crocodiles as the closest living relatives of turtles (Hedges and Poling 1999) even to the exclusion of birds (Fig. 18.5d). Furthermore, a phylogenetic analysis that combined mitochondrial and nuclear data also recovered a crocodile + turtle grouping (Cao *et al*. 2000). However, morphological and paleontological evidences clearly support the monophyly of archosaurs. Interestingly, both crocodiles and turtles show significantly long branches which might introduce biases in the phylogenetic analyses. Hence, a sister group relationship of crocodiles and turtles needs to be treated as tentative, and further molecular work lies ahead.

The origin of placental mammals

Mammals are tetrapods well suited for life on land. In particular, the acquisition of a placenta, a membrane to nourish the fetus, was a major breakthrough in mammalian evolution that may partly explain their success and rapid radiation after the mass extinction of dinosaurs in the uppermost Cretaceous. The classical view of mammalian phylogeny supported by morphological and fossil data
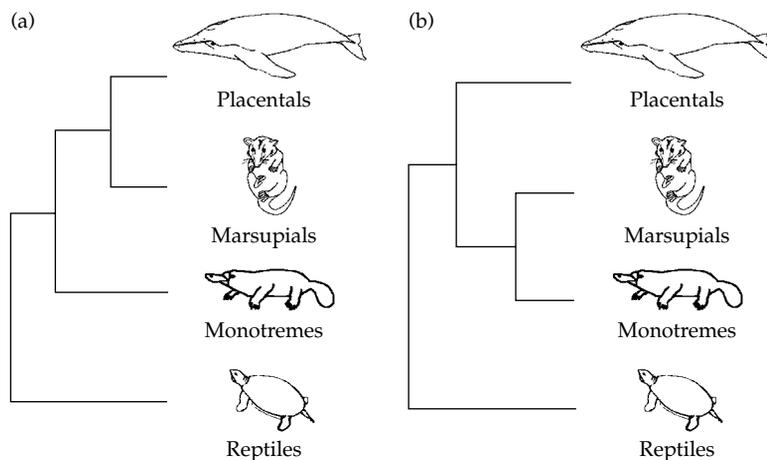
Figure 18.6 Alternative hypotheses explaining the phylogenetic relationships of monotremes, marsupials, and placental mammals. (a) The Theria hypothesis: marsupials are the closest living sister group of placentals. (b) The Marsupionta hypothesis: the marsupials are closely related to monotremes, and both groups are equally distant to placentals.

places monotremes (the platypus and echidnas) as the most basal lineage of mammals, with the marsupials as closest relatives of eutherians (placental mammals) (Carroll 1988) (Fig. 18.6a). In this regard, there are many morphological features that have been interpreted as shared derived characters between marsupials and placentals. However, some workers advocate a close relationship of monotremes and marsupials (the Marsupionta hypothesis) based on similar tooth-replacement patterns (Gregory 1947; Kühne 1973), to the exclusion of placentals (Fig. 18.6b). The poor fossil record for monotremes (Carroll 1988) complicates the analysis of their phylogenetic relationships to marsupials and placentals, and leaves open the question on the closest relative of Eutheria.

The complete nucleotide sequences of the mitochondrial genomes of the platypus and the opossum were determined to clarify the debate based on morphological data (Janke *et al*. 1996). Phylogenetic analyses of a data set that combined the inferred amino acid sequences of the mitochondrial protein-coding genes favored with high statistical support that monotremes and marsupials are sister groups (Janke *et al*. 1996) (Fig. 18.6b). Further phylogenetic analyses with additional tetrapod taxa (including the wallaroo, *Macropus robustus*) seemed to confirm

the mitochondrial support for the Marsupionta hypothesis (Janke *et al*. 2001). However, some workers have noted that the support of the mitochondrial protein data for the Marsupionta hypothesis varies considerably depending on the outgroup and phylogenetic methods used (Wadell *et al*. 1999). Phylogenetic analyses of DNA-hybridization data on several amniotes also supported the monotreme + marsupial clade (Kirsch and Mayer 1998). However, it has been suggested that monotremes show relatively high GC contents, and that this bias might be shared by marsupials, but not by placentals (Kirsch and Mayer 1998). If confirmed, a base-compositional bias rather than a true phylogenetic signal could be responsible for the monotreme + marsupial grouping in the DNA-hybridization analyses. Recently, a large nuclear gene, the mannose 6-phosphate/insulin-like growth factor II receptor, was sequenced from representatives of all three mammalian groups to clarify the controversy (Killian *et al*. 2001). Phylogenetic analyses of this nuclear gene sequence favored with statistical support that marsupials are the sister group of eutherians to the exclusion of monotremes (Fig. 18.6a). Hence, new molecular data seems to corroborate morphological evidence. Future molecular studies (including, for example, the complete

mitochondrial genome of an echidna, as well as more nuclear gene sequence data), and the confrontation of molecular and morphological phylogenies will certainly improve our understanding of the origin of placental mammals.

Summary

Vertebrates offer the opportunity to study long-term evolutionary patterns and processes because their phylogeny is comparatively well known. Most major events that have occurred throughout the evolution of vertebrates are well documented in the fossil record, and phylogenetic analyses of such paleontological data have made it possible to reconstruct rather resolved trees that explain vertebrate phylogenetic relationships. However, some nodes in the vertebrate tree, at the origin of the main lineages, remain controversial. The origin of the main lineages of vertebrates was accompanied by key morphological innovations and rapid radiation events. Gaps in the fossil record associated to some of these events, difficulties in the interpretation of what has been preserved, and the existence of few morphological shared derived characters between the putative living sister groups and the radiated groups hamper the inference of the exact phylogenetic relationships between the taxa involved in the origin of the major groups of vertebrates.

Recent advances in molecular techniques as well as more powerful phylogenetic algorithms and faster computers have made it possible to infer phylogenetic relationships using sequence data. Two molecular markers, mitochondrial DNA and nuclear rRNA genes, have been widely applied to phylogenetic inference of vertebrate relationships. Besides the corroboration of the traditional morphology-based phylogeny, new molecular data have been particularly helpful in discerning among alternative hypotheses to explain the origin of the major lineages of vertebrates. Examples are the recent molecular evidence that supports a sister group relationship of hagfishes and lampreys, that groups lungfishes with tetrapods to the exclusion of coelacanths, that favors the Batrachia hypothesis (salamanders as sister group of frogs), that places turtles as diapsid reptiles, and that suggests marsupials as the closest relatives of placental mammals to the exclusion of monotremes.

In most cases, molecular data corroborate morphological evidence, but in some cases molecular and morphological signals conflict. Ultimately, comparisons of conflicting signals should enable evolutionary biologists to detect anomalies that result in misinterpreting one of the two types of data. Understanding the sources of signal conflict will definitively improve phylogenetic inference and may contribute to settling open debates on vertebrate systematics.