Interrelationships of Lamniform Sharks: Testing Phylogenetic Hypotheses with Sequence Data

GAVIN J. P. NAYLOR
Department of Biology
Osborn Memorial Laboratory
Yale University
New Haven, Connecticut 06520

ANDREW P. MARTIN
Biological Sciences
University of Nevada Las Vegas
Las Vegas, Nevada 89154

ERIK G. MATTISON and WESLEY M. BROWN
Department of Biology
University of Michigan
Ann Arbor, Michigan 48109

I. Introduction

Phylogenetic reconstruction involves estimating relationships from patterns of character-state covariation seen among taxa. The endeavor would be straightforward were each evolutionary lineage to acquire its own set of unique traits at birth and then pass them on immutably to all descendents. If this were the case, phylogenetic reconstruction would require no more than a search for the evolutionary tree which accounted for the distribution of traits as a perfectly nested set. Unfortunately, evolution is not so simple. Character-state changes occur with markedly different probabilities across both characters and taxa, traits frequently revert to previous conditions, lineages occasionally coalesce, and identical character states arise in multiple lineages by parallel or convergent evolution. In many cases these vagaries conspire to confound or bias inferences about evolutionary history. It is important, when estimating phylogeny, to explore the strengths and limitations of data in light of these potentially confounding influences.

In the past decade, much has been made of the power of molecular sequences for inferring evolutionary history (Avise, 1994). Arguments often promote a notion that molecular data are intrinsically better templates than are morphological or behavioral data for recording the tell-tale imprint of evolutionary history. The authors believe that DNA sequence data offer advantages for phylogenetic reconstruction, but do not subscribe to the view that they are intrinsically “better” than other forms of data. The advantages seen for DNA sequence data are: (1) A large number of potentially informative, heritable, and discrete characters can be obtained. This can be useful when the group under investigation is conservative, is characterized by a scarcity of good morphological characters, or has been subjected to repeated bouts of evolutionary parallelism of phenotypic characters. (2) Protein-encoding DNA sequence data can be broken down into different constraint categories based on a knowledge of the genetic code. First, second, and third codon positions can be recognized as can two- and four-fold degenerate sites. These categories can be treated as distinct classes of data and analyzed separately. Morphological and behavioral traits cannot be broken down in this way. (3) Because distinct classes of data can be recognized, differences in the evolutionary dynamics among classes can be investigated. Observations in a given class can be pooled across a number of sites and sub-
nucleotide bases G, A, T, and C are often found in unequal proportions. Such bias in base composition distinguishes different lineages, all lineages should be similarly affected. However, if bias differs across lineages—a condition known as "deviation from stationarity"—there is a tendency for phylogenetic methods to group taxa by the similarity of base composition, regardless of their historical relatedness (Lockhart et al., 1992, 1994). Finally, the various processes and constraints that act on DNA sequences can interact to increase the variance in the evolutionary rate seen for a given class of site, making the task of fitting a model of evolutionary change to the data particularly difficult. For example, redundancy of the genetic code renders third codon positions more free to vary than other codon positions. However, sites free to vary also accumulate compositional biases. The bias, in turn, restricts the amount of evolutionary change that can be recorded. Third position sites can thus range from appearing to be highly variable, fast-evolving sites that record a large number of evolutionary changes to highly constrained, slowly evolving sites that seldom record an event, depending on the evenness of base composition.

Phylogenetic inferences derived from molecular data should be critically evaluated in light of these shortcomings so that character-state covariances due to site saturation, rate variation, and base compositional effects are not mistaken for evolutionary signal.

**A. Assumptions**

In phylogenetic inference, a model of evolutionary transformation between character states is applied to a distribution of character states for a group of taxa to yield a tree that best explains the data (Sober, 1988). Different tree-building algorithms invoke different models. Some models are very specific and restrictive (e.g., distance cluster analysis of corrected genetic distances among taxa). Others are more general and have fewer restrictions (e.g., cladistic parsimony). There is generally a trade-off: the more restrictive the model the more explanatory power is reaped; however, restrictions increase so does the likelihood that assumptions of the model will be violated (Huelsenbeck and Hillis, 1993). Because of the heterogeneous nature of evolutionary change both among characters and among lineages, it is perhaps best to rely on models that can be (1) applied across different types of characters, and (2) modified to include restrictive assumptions when suggested by the data. Cladistic parsimony has been championed as a method that requires few restrictive assumptions (Farris, 1983) and, therefore, as a method that should be widely applicable across a broad range of evolutionary dynamics, such as those seen at the different positions of a codon. The authors subscribe to this view, but feel it important to outline the assumptions that are implicit in parsimony analyses. This is done to underscore that the conclusions are inferences...
contingent on data fitting the implied model. Parsimony requires that homoplasy be randomly distributed among taxa. When this is the case, the true historical "signal" (if it is the most influential source of character-state covariance among taxa) should overshadow any "noise" due to homoplasy. In keeping with this view, it is assumed (1) that incorrect inferences are due to stochastic error associated with a small sample of characters, and (2) that erroneous inferences should disappear as more data are collected. This argument is appealing. However, highly structured, nonhistorical sources of character-state covariation among taxa can often dilute or eradicate the signal due to shared history. In some cases these can even swamp any phylogenetic signal with a positively misleading signal. For example, as previously alluded to, nucleotide base compositional frequencies can vary across taxa in such a way that distantly related organisms have more similar base compositions than do close relatives. In such situations, parsimony will be predisposed to incorrectly group the distantly related taxa together because of their base compositional similarity (Loomis and Smith, 1990; Penny et al., 1990; Sidow and Wilson, 1990, 1991; Lockhart et al., 1992; Hasegawa and Hashimoto, 1993; Steel et al., 1993). However, such nonrandom distributions of homoplasy do not necessarily preclude the effective use of parsimony. A careful inspection of data can help identify classes of characters that have the potential to be misleading. Problems so identified can sometimes be ameliorated by judicious use of differential weighting schemes (Hillis et al., 1994; Huelsenbeck et al., 1994). For example, in a data set of six taxa where the two most divergent forms share a base composition profile comprising 45% G, 45% C, 5% T, and 5% A, while the remainder share a profile of 5% G, 5% C, 45% T, and 45% A, parsimony will be predisposed to link the two most divergent forms together as sister taxa. However, if the nucleotides are recoded as either purines or pyrimidines, all six taxa are rendered an unbiased 50:50 purine:pyrimidine base composition. In essence, the transformation of data results in an amplification of the phylogenetic signal to noise ratio by bringing the data more into line with parsimony's requirement for the random distribution of homoplasy.

B. Fossils and Phylogenies

Investigation of the sources of character-state covariation in molecular data sets is best accomplished for groups in which the phylogeny is known (e.g., bacteriophage (Hillis et al., 1992), mice (Sage et al., 1993), corn (Kellog and Birchler, 1993)) or can be corroborated by independent means (e.g., higher-order vertebrate classes). In most cases these "model groups" may not lead to predictive results that can be widely applied to different groups of organisms because they either focus on unusual genomes in contrived conditions (e.g., bacteriophage) or address issues that arise when highly divergent lineages are compared (e.g., vertebrate classes). Successes and pitfalls encountered in the analysis of higher taxa may have little relevance for analyses at lower taxonomic ranks because higher taxa generally differ in so many ways that it is impossible to attribute patterns of character-state covariation to any specific subset of biological influences. Ideally, character-state covariation in molecular data sets is best explored by describing patterns of molecular evolution for a group whose phylogeny can be corroborated by independent means and then expanding the taxonomic scope to include taxa that are biologically similar.

The success of these studies often depends on information about the history of the group derived from the fossil record. Although first appearances of fossil lineages that lead to extant forms do not provide information about phylogeny, positive correlation between the age of lineages estimated from fossils and the age of lineages determined from phylogenetic analysis of DNA sequences can provide a gauge for the accuracy of phylogenetic inference. Although a lack of significant correlation between age and clade ranks may be indicative of a poor fossil record (Norell and Novacek, 1992) or of a grossly inaccurate reconstruction of phylogeny, a significant positive correlation between age and clade ranks most likely reflects correspondence between phylogenetic pattern and evolutionary history recorded in the paleontological record (Norell and Novacek, 1992). Times of first appearance in the fossil record have been documented for diversifying lineages in a number of groups (e.g., Bryozoa (Jackson and Cheetham, 1994); cartilages (Lundberg, 1992); sharks (Maisey, 1984; Cappetta, 1987)). When the fossil record for such groups is dense and continuous, the first appearance times of different lineages can be used to calibrate rates of molecular evolution and to investigate rate heterogeneity within and among taxonomic groups (Martin et al., 1992). This can be important for testing alternative hypotheses of molecular evolution (Gillespie, 1986b; Kimura, 1983) and for establishing the phylogenetic utility of specific genes at various levels of taxonomic differentiation (Graybeal, 1994; Friedlander et al., 1994).

C. Sharks and the Order Lamniformes

The fossil record of sharks is dense, relatively continuous, and consists almost entirely of teeth (Maisey, 1984; Cappetta, 1987). Many of these teeth are distinc-
tive enough to allow identification of the fossil lineages that gave rise to extant forms, making sharks a model group for the type of fossil-calibrated molecular systematics study described earlier. In general, sharks are a morphologically conservative group. Phylogenetic hypotheses based on morphology have been hampered by a scarcity of shared derived character states (Fechhelm and McEachran, 1984; Compagno, 1988). Molecular sequences may provide a much needed source of shared derived character information with which to infer phylogenetic hypotheses for different shark groups (e.g., Martin, 1993).

One group of sharks that is particularly well represented in the fossil record is the order Lamniformes, which originated 124–140 million years ago (Maisey, 1984; Cappetta, 1987). Paleontological work indicates that the order was at its most diverse in the middle and late Cretaceous, but subsequently suffered repeated bouts of extinction. Extant lamniform sharks thus constitute a relictual assemblage of highly divergent lineages. The differentiation among contemporary species is reflected in their classification. There are 16 recognized species classified in 10 genera and seven families. Five of the genera and four of the families are monotypic. The order comprises the relatively well-known, endothermic superpredators (Lamnidae), i.e., the great white, the two makos, the porbeagle, and the salmon shark; the three species of thresher sharks (Alopiidae) with their extremely long caudal fins, which they use like whips to stun and kill schooling fishes (Compagno, 1984); the whale-like, filter-feeding, basking shark (Cetorhinidae), which can attain lengths of up to 30 feet; the sluggish, benthic sand tiger sharks (Odontaspididae); the deep-water crocodile and goblin sharks (Pseudocarchariidae and Mitsukurinidae, respectively); and the recently discovered megamouth shark (Megachasma).

Attempts to estimate the evolutionary relationships among these extant taxa based on morphological characters have yielded conflicting hypotheses (Maisey, 1985, Fig. 1A; Compagno, 1990; Fig. 1B). In order to evaluate these alternate hypotheses and to investigate the correspondence between phylogenetic inference and paleontological information, the sequences of the mitochondrial protein-encoding cytochrome b and NADH 2 genes have been determined for all but 2 of the 16 extant lamniform species and the data have been subjected to phylogenetic analysis. Particular attention has been paid to issues that might confound covariance of character states due to shared ancestry. Results based solely on these sequence data suggest a new phylogenetic hypothesis for the order. When inferred branch lengths based on sequence data are contrasted with first appearance information from the fossil record, the new hypothesis shows a better fit to the fossil record than do the hypotheses of either Compagno (1990) or Maisey (1985).

II. Materials and Methods

Fresh tissue samples were obtained for all but 2 of the 16 lamniform species. Tissues from Odontaspis noronhai or from Carcharias tricuspidatus were not obtained. Where possible, multiple specimens were sequenced for each species. A list of specimens sequenced is presented in the Appendix with corresponding locality data. Sequences for the mitochondrial NADH 2 and cytochrome b genes were obtained. Both were amplified using polymerase chain reaction (PCR), using a different protocol for each.

The NADH 2 gene was amplified in two steps. The double-stranded DNA product was made by subjecting a total DNA preparation to 30 cycles of PCR amplification in a 100-μl reaction using Perkin-Elmer Taq polymerase and two universal NADH 2 primers (Kocher et al., 1995). The product was chloroform extracted, precipitated with ammonium acetate and ethanol, washed in a 70% ethanol/10mM TRIS, 1μM EDTA (TE) solution, air dried, resuspended in TE, and run out on a low melting point gel. The band, visualized under low intensity ultraviolet light, was excised, purified (Gene-Clean; United States Biochemical), and stored in 80 μl 0.1 TE. A single-stranded DNA product was then made by using the double-stranded product as the template in a second 20-cycle PCR reaction, to which only one of the two original amplification primers was added. The resulting single-stranded product was cleaned and concentrated using fourflushes of 0.1 TE solution through ultrafree microconcentration tubes (Millipore Corporation), then sequenced using the Sequenase protocol (USB) employing dideoxy-NTP termination reactions (Sanger et al., 1977, 1980) in conjunction with a series of sequencing primers spaced at approximately 150-bp intervals along the fragment.

The cytochrome b sequence was amplified in a 25-μl reaction using 12.5 pmol of the primers GluDG and Cb1211H (Falumbi et al., 1991), Perkin-Elmer buffer, 200 μM of each nucleotide, and 1 μl of Perkin-Elmer Taq polymerase. An initial 30 amplification cycles were carried out at 94°C for 30 sec, 52°C for 15 sec, and 72°C for 60 sec. One microliter of the amplified product was used to seed a second, 150-μl reaction containing 75 pmol of the GluDG primer and 7.5 pmol of biotin-labeled Cb1211H. After another 35 cycles of amplification, the DNA was precipitated with 7.5 M ammonium acetate and 50% ethanol, pelleted by centrifugation at
high speed for 10 min, washed once with ethanol, air dried, and resuspended in 40 μl of water. For each sample, 20 μl of Dynal streptavidin beads was washed with 50 μl of binding and washing (BW) buffer (4 M NaCl, 10 mM Tris, 1 mM EDTA, 0.1% Nonidet P-40). The beads were resuspended in 40 μl of BW buffer, combined with the DNA, and the solution was incubated for 1 hr with slow rotation at 45°C to allow the biotin-labeled DNA to bind to the streptavidin beads. The beads were then washed once with 50 μl of BW buffer, twice with 50 μl of sterile, distilled water, and resuspended in 12 μl of water. The sample was boiled for 15 sec and quickly put on the magnet to remove the beads from solution. The solution containing the non-biotin-labeled DNA was collected, diluted with 28 μl of water, labeled, and stored at 4°C. Following heat denaturation, the beads were incubated at room temperature for 10 min in 0.1 N NaOH, washed twice with 50 μl of sterile, distilled water, and resuspended in 40 μl of water. Both strands were sequenced using a battery of primers and the Sequenase protocol (USB).

In most cases, both genes were sequenced for each individual. However, there were instances where NADH 2 was sequenced from one individual whereas...
cytochrome b was sequenced from a different individual of the same species (see Appendix). In these cases, the genes from the different individuals were combined to represent that species. Such use could pose a problem if within-species polymorphism was so great as to render a species paraphyletic with respect to other taxa. However, in all cases in which multiple individuals were sequenced there was very little within-species sequence variation (<0.5%). In fact, in most cases, sequences from replicate individuals of the same species were identical.

A data set comprising NADH 2 and cytochrome b sequences for the 14 lamniform taxa, seven carcharhiniform outgroup taxa (GenBank Access numbers U91417-U91447, L08031-L08034 and L08036-L08042) and corresponding GenBank sequences for carp (Cyprinus carpio), loach (Crossostoma lacustre) and trout (Oncorhynchus mykiss) was assembled. Sequences were aligned with Clustal W (Thompson et al., 1994), checked visually with the codon coloring feature of Aligner (Eernisse, 1992), and subjected to a number of different analyses. The authors relied predominantly on cladistic parsimony for phylogenetic reconstruction and shied away from methods with more restrictive assumptions. A number of analyses were carried out to evaluate the overall signal and the signal at different codon positions of each gene. These data subsets have been assessed for internal concordance and combinability. The reliability of inferred clades has been evaluated whether two data sets were combinable, the authors evaluated whether the partition between the two data subsets was statistically different from random expectation. Parsimony analyses were carried out for each data subset separately. The length of the most parsimonious tree (MPT) for the first data subset was added to that of the second. This combined tree length was then contrasted with a distribution of comparable tree lengths generated by randomly partitioning the original data (Farris et al., 1995); y values reflecting the placement of the original data partition in the distribution were computed and are shown in bold type, whereas the number of parsimony informative sites are shown in regular text. (A) Contrasts among codon positions between the two genes. Bootstrap consensus trees are presented for each of the six gene/codon subsets to provide a sense of the phylogenetic signal intrinsic to each data subset. The retention index for the MPT for each data subset is also shown. Loxy, Isurus oxyrinchus; I.pau, Isurus paucus; C.car, Carcharodon carcharias; L.nas, Lamna nasus; L.dit, Lamna ditropis; C.max, Cetorhinus maximus; M.pe1, Megachasma pelagios; P.kam, Pseudecharcharias kamoharae; O.fer, Odontaspis ferox; A.sup, Alloptis superciliosus; A.vul, A. vulpinus; A.pe1, A. pelagicus; C.lau, Carcharias taurus; M.ows, Mitsukurina owstoni; G.cuv, Galeocerdo cuvier; N.br, Negaprion brevirostris; P.gla, Prionace glauca; C.por, Carcharhinus porosus; C.plu, C. plumbeus; S.tib, Sphyra tiburo; S.lew, S. lewini; carp, Cyprinus carpio; loach, Crossostoma lacustre; trout, Oncorhynchus mykiss. (B) Contrasts among codon positions within genes.

III. Results and Discussion

A. Combinability

Although mitochondrial genes are inherited as a single unit and are generally not subject to recombination, it is possible that differential selection among genes over time could render their respective hierarchical signals incongruent or noncombinable. In order to evaluate this aspect, combinability tests were conducted. (Farris et al., 1995) among the three codon positions of NADH 2 and cytochrome b. Sequence data were sorted by gene and codon position (six data subsets). Tests were then conducted to assess whether the partitions between pairs of subsets were statistically different from random expectation. Parsimony analyses were carried out for each data subset separately. The length of the most parsimonious tree (MPT) for the first data subset was added to that of the second. This combined tree length was then contrasted with a distribution of comparable tree lengths generated by randomly partitioning the original data. Results (Fig. 2) suggest that the signals from all pairs of subsets are highly combinable. The lowest p value for combinability was 0.31 between the third position sites of the two genes.

B. Nucleotide Composition

Base composition was assessed for variable sites at first, second, and third codon positions for each gene separately (Table I). Noticeable differences were seen, both among codon positions and between genes. Base compositional stationarity was evaluated for all shark
13. Lamniform Sharks

A CYTB pos 1 mpt ri = 0.58

<table>
<thead>
<tr>
<th></th>
<th>CYT B pos 1</th>
<th>CYT B pos 2</th>
<th>CYT B pos 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.82</td>
<td>pars inf sites: CYT B = 116 ND2 = 151 TOTAL = 267</td>
<td>pars inf sites: CYT B = 116 ND2 = 66 TOTAL = 182</td>
<td>pars inf sites: CYT B = 39 ND2 = 151 TOTAL = 190</td>
</tr>
<tr>
<td>0.94</td>
<td>pars inf sites: CYT B = 39 ND2 = 66 TOTAL = 105</td>
<td>pars inf sites: CYT B = 39 ND2 = 66 TOTAL = 105</td>
<td>pars inf sites: CYT B = 39 ND2 = 151 TOTAL = 190</td>
</tr>
<tr>
<td>0.53</td>
<td>pars inf sites: CYT B = 39 ND2 = 66 TOTAL = 105</td>
<td>pars inf sites: CYT B = 39 ND2 = 66 TOTAL = 105</td>
<td>pars inf sites: CYT B = 39 ND2 = 151 TOTAL = 190</td>
</tr>
</tbody>
</table>

mpt ri = 0.60 mpt ri = 0.72 mpt ri = 0.46

NADH 2 pos 1

NADH 2 pos 2

NADH 2 pos 3

pars inf sites:
CYT B = 116
ND2 = 151
TOTAL = 267
pars inf sites:
CYT B = 116
ND2 = 66
TOTAL = 182
pars inf sites:
CYT B = 39
ND2 = 151
TOTAL = 190
pars inf sites:
CYT B = 39
ND2 = 66
TOTAL = 105
pars inf sites:
CYT B = 39
ND2 = 66
TOTAL = 105
pars inf sites:
CYT B = 39
ND2 = 151
TOTAL = 190

pars inf sites:
CYT B = 39
ND2 = 66
TOTAL = 105
pars inf sites:
CYT B = 39
ND2 = 66
TOTAL = 105
pars inf sites:
CYT B = 39
ND2 = 151
TOTAL = 190
pars inf sites:
CYT B = 39
ND2 = 66
TOTAL = 105
pars inf sites:
CYT B = 39
ND2 = 66
TOTAL = 105
pars inf sites:
CYT B = 39
ND2 = 151
TOTAL = 190

pars inf sites:
CYT B = 39
ND2 = 66
TOTAL = 105
pars inf sites:
CYT B = 39
ND2 = 66
TOTAL = 105
pars inf sites:
CYT B = 39
ND2 = 151
TOTAL = 190
pars inf sites:
CYT B = 39
ND2 = 66
TOTAL = 105
pars inf sites:
CYT B = 39
ND2 = 66
TOTAL = 105
pars inf sites:
CYT B = 39
ND2 = 151
TOTAL = 190
pars inf sites:
CYT B = 39
ND2 = 66
TOTAL = 105
pars inf sites:
CYT B = 39
ND2 = 66
TOTAL = 105
pars inf sites:
CYT B = 39
ND2 = 151
TOTAL = 190
pars inf sites:
CYT B = 39
ND2 = 66
TOTAL = 105
pars inf sites:
CYT B = 39
ND2 = 66
TOTAL = 105
pars inf sites:
CYT B = 39
ND2 = 151
TOTAL = 190
pars inf sites:
CYT B = 39
ND2 = 66
TOTAL = 105
pars inf sites:
CYT B = 39
ND2 = 66
TOTAL = 105
pars inf sites:
CYT B = 39
ND2 = 151
TOTAL = 190
pars inf sites:
CYT B = 39
ND2 = 66
TOTAL = 105
pars inf sites:
CYT B = 39
ND2 = 66
TOTAL = 105
pars inf sites:
CYT B = 39
ND2 = 151
TOTAL = 190
pars inf sites:
CYT B = 39
ND2 = 66
TOTAL = 105
pars inf sites:
CYT B = 39
ND2 = 66
TOTAL = 105
pars inf sites:
CYT B = 39
ND2 = 151
TOTAL = 190
pars inf sites:
CYT B = 39
ND2 = 66
TOTAL = 105
pars inf sites:
taxa (note that carp, loach, and trout, having clearly different base compositions, were excluded from this analysis). Base compositions at first and second codon positions did not differ significantly among the taxa surveyed. Third position sites of both genes were highly significantly different among taxa ($p < 0.00000001$). However, when these third position sites were recoded as either purines or pyrimidines, deviations from stationarity were ameliorated (see Table I). Note that the $\chi^2$ tests presented are intended only as coarse descriptors of base compositional stationarity; structured patterns of base compositional difference are not always reflected by such tests. For example, the lamnid taxa have a lower percentage of adenine and a higher percentage of guanine at first position sites (of both genes) than do any of the other sharks surveyed (Table I), yet $p$ values suggest that deviations from stationarity are not significant.

**C. Saturation**

Transitions and transversions for each codon position were plotted separately for each gene (Fig. 3) against the average number of nucleotide differences between pairs of taxa (mean raw distance). Although there is an element of nonindependence in such plots (part of the information is the same for both the ordinate and the abscissa), differences in trend are detectable. In both genes, third positions are more variable than first positions, which are, in turn, more variable than second positions. This is expected based on the relative redundancy of the genetic code. The NADH 2 gene registers more changes than does the cytochrome b gene, reflecting perhaps a diminished level of constraint in NADH 2 relative to cytochrome b. There is clear evidence for signal saturation in transitions at third positions of both genes. Slight saturation is seen for the first position transitions of NADH 2; however, this is restricted to the distant comparisons between sharks and the teleost outgroups. Second positions show a considerable amount of scatter, particularly in cytochrome b, possibly reflecting a relatively constrained character-state space at this position (Irwin et al., 1991). It is possible that a distance which accommodated multiple substitutions might be a more sensitive indicator of saturation. This was tested by plotting various such distance measures against the average number of nucleotide differences used herein. All distances examined exhibited a linear relationship (Fig. 4) with the “uncorrected” distance used, suggesting that the presented estimates of saturation would not be greatly affected.

Transition:transversion ratios derived from pairwise comparisons among taxa vary according to taxonomic depth of the comparison and to codon position. For pairwise contrasts within the Lamniformes, the mean observed transition:transversion ratio was approximately 5:1 for first position sites and 3:1 for second and third position sites. Somewhat unexpectedly, given the apparent differences in substitution rate
TABLE I  Base Composition Profiles for Cytochrome B and NADH 2

<table>
<thead>
<tr>
<th>Taxon</th>
<th>First position sites</th>
<th>Second position sites</th>
<th>Third position sites</th>
<th>Third position R/Y</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>A</td>
<td>C</td>
<td>G</td>
<td>T</td>
</tr>
<tr>
<td>Carp</td>
<td>0.25</td>
<td>0.32</td>
<td>0.27</td>
<td>0.15</td>
</tr>
<tr>
<td>Loach</td>
<td>0.21</td>
<td>0.29</td>
<td>0.31</td>
<td>0.19</td>
</tr>
<tr>
<td>Trout</td>
<td>0.19</td>
<td>0.31</td>
<td>0.33</td>
<td>0.17</td>
</tr>
<tr>
<td>C.cuv</td>
<td>0.35</td>
<td>0.26</td>
<td>0.15</td>
<td>0.24</td>
</tr>
<tr>
<td>Nbre</td>
<td>0.31</td>
<td>0.29</td>
<td>0.17</td>
<td>0.22</td>
</tr>
<tr>
<td>Pgl</td>
<td>0.30</td>
<td>0.29</td>
<td>0.18</td>
<td>0.22</td>
</tr>
<tr>
<td>C.por</td>
<td>0.29</td>
<td>0.24</td>
<td>0.19</td>
<td>0.27</td>
</tr>
<tr>
<td>C.plu</td>
<td>0.30</td>
<td>0.29</td>
<td>0.18</td>
<td>0.22</td>
</tr>
<tr>
<td>S.lew</td>
<td>0.29</td>
<td>0.32</td>
<td>0.18</td>
<td>0.21</td>
</tr>
<tr>
<td>S.tib</td>
<td>0.29</td>
<td>0.27</td>
<td>0.18</td>
<td>0.26</td>
</tr>
<tr>
<td>Loxy</td>
<td>0.27</td>
<td>0.36</td>
<td>0.21</td>
<td>0.17</td>
</tr>
<tr>
<td>Ipau</td>
<td>0.28</td>
<td>0.36</td>
<td>0.21</td>
<td>0.15</td>
</tr>
<tr>
<td>C.car</td>
<td>0.32</td>
<td>0.31</td>
<td>0.19</td>
<td>0.17</td>
</tr>
<tr>
<td>Lnas</td>
<td>0.32</td>
<td>0.32</td>
<td>0.18</td>
<td>0.09</td>
</tr>
<tr>
<td>Ldit</td>
<td>0.29</td>
<td>0.33</td>
<td>0.20</td>
<td>0.18</td>
</tr>
<tr>
<td>C.max</td>
<td>0.33</td>
<td>0.33</td>
<td>0.15</td>
<td>0.20</td>
</tr>
<tr>
<td>M.pe1</td>
<td>0.36</td>
<td>0.26</td>
<td>0.13</td>
<td>0.25</td>
</tr>
<tr>
<td>A.pe1</td>
<td>0.33</td>
<td>0.33</td>
<td>0.17</td>
<td>0.17</td>
</tr>
<tr>
<td>A.vul</td>
<td>0.35</td>
<td>0.30</td>
<td>0.14</td>
<td>0.20</td>
</tr>
<tr>
<td>A.sup</td>
<td>0.30</td>
<td>0.34</td>
<td>0.18</td>
<td>0.18</td>
</tr>
<tr>
<td>M.ows</td>
<td>0.35</td>
<td>0.30</td>
<td>0.13</td>
<td>0.21</td>
</tr>
<tr>
<td>Oler</td>
<td>0.36</td>
<td>0.29</td>
<td>0.15</td>
<td>0.21</td>
</tr>
<tr>
<td>Pkam</td>
<td>0.34</td>
<td>0.27</td>
<td>0.16</td>
<td>0.23</td>
</tr>
<tr>
<td>C.tau</td>
<td>0.31</td>
<td>0.33</td>
<td>0.17</td>
<td>0.19</td>
</tr>
<tr>
<td>Mean</td>
<td>0.30</td>
<td>0.31</td>
<td>0.19</td>
<td>0.20</td>
</tr>
</tbody>
</table>

χ² = 39.6

p = 0.08

χ² = 10.8

p = 1.0

χ² = 163.8

p < 0.00000001

χ² = 34.7

p > 0.5

χ² = 8.4

p = 1.0

χ² = 204

p < 0.00000001

χ² = 18.2

p > 0.5

*Frequencies are presented for each codon position for both genes separately. χ² values, together with corresponding p values, are presented to summarize the extent of deviation from base compositional stationarity. Note that deviations from stationarity are highly significantly different in third codon positions of both genes. However, a recoding of third position character states as either "purine" (R) or "pyrimidine" (Y) effectively eliminates the deviation from stationarity at this position. Taxon labels as in Fig. 2.*
FIGURE 3  Relationship between genetic distance and the number of transitional and transversional differences at different codon positions for each gene. Plots of the number of observed transitional (open squares) and transversional (filled circles) differences between all pairwise contrasts in the data set were plotted against the mean absolute distance between pairs. Note that transitions appear saturated at third positions of both genes.
between the two genes, the transition:transversion ratios at respective codon positions were similar for both genes.

D. Phylogenetic Analysis

The most parsimonious tree resulting from the analysis in which all sites were weighted equally is shown in Fig. 5A. The number of steps required to break down each inferred clad (Bremer, 1988) is depicted on the most parsimonious tree, whereas the corresponding bootstrap consensus tree is presented to show the core hierarchical signal reflected by the data. Equivalent transversion parsimony analyses are shown in Fig. 5B.

In order to expose any latent phylogenetic bias that might be caused by structured deviations from base compositional stationarity, data were subjected to two types of analysis designed to minimize such influences. In the first analysis, third position sites were recoded according to their status as purine or pyrimidine to ameliorate deviations from stationarity (see earlier discussion) and then the resultant data matrix was subjected to parsimony analysis (Fig. 6A). In the second analysis, data were subjected to LogDet transformation and subsequent neighbor-joining (Fig. 6B). This approach has been shown to be resilient to the influence of base compositional differences among taxa when sites evolve independently (Lake, 1994; Steel, 1994).

Both approaches consistently show the Lamniformes to be monophyletic, the Lamnidae to constitute a monophyletic group within the order, the basking shark (Cetorhinus maximus) to be the sister taxon to the Lamnidae, each of the genera Lamna and Isurus to constitute monophyletic groups within the Lamnidae, and the great white shark C. carcharias to be the sister taxon to the genus Isurus. Both analyses also suggest that Alopias pelagicus and A. vulpinus are sister taxa and that Pseudocarcharias kamoharai, Odontaspis ferox, Megachasma pelagios, and the three thresher sharks (Alopiidae) constitute a monophyletic group. When trees resulting from these two analyses are contrasted with the bootstrap consensus tree based on equal weighting of all sites (Fig. 5A), there is considerable congruence. The bootstrap consensus is somewhat less resolved (as expected) and does not show sister group relationships between C. carcharias and the genus Isurus or between the thrasher sharks A. pelagicus and A. vulpinus. From these results, it would seem that the bootstrap consensus based on equal weighting of all sites is not adversely skewed by deviations from stationarity and can be viewed (for these data, at least) as a conserva-
FIGURE 5  Phylogenetic inferences based on parsimony analysis of the sequence data. Trees in the left column correspond to the most parsimonious trees (MPTs), or a strict consensus when multiple MPTs resulted from analysis. The number of steps required to break down each inferred clade (Bremer, 1988) is depicted directly on the trees. Trees in the right column correspond to the consensus of 100 bootstrap replicates. Numbers on the consensus correspond to the percentage occurrence of the clade among replicates. The bootstrap consensus trees are presented to provide a sense of the core hierarchical signal in each data set. (A) Analyses conducted on data comprising both transitions and transversions. (B) Analyses using transversional differences only.

tive estimate of the historical signal in the presented data set.

Somewhat surprisingly, there is little evidence in any of the analyses to support a monophyletic Alopi-
idae. Nevertheless, there is extensive morphological support for such a grouping. In light of this, the au-
thors propose the phylogenetic hypothesis presented in Fig. 7. This hypothesis is based on the bootstrap con-

E. Testing Hypotheses with DNA Sequence Data

Although the data do not provide robust support for any one fully resolved cladogram, they can be used to evaluate competing hypotheses concerning lamniform relationships. Each of the hypotheses was fitted to the data set separately, using the constraints option of PAUP. A heuristic search was carried out with the hypothesis under test implemented as a constraint. The number of steps required by each hypothesis was contrasted with the number of steps required by the most parsimonious tree for the data set. The fewer “extra” steps required, the better the fit of the hypothesis to the data and, therefore, the more credible the hypothesis. The following 10 hypotheses were evaluated.

Hypothesis 1: Monophyly of Alopidiidae  
Hypothesis 2: Monophyly of the filter-feeding trait (Maisey, 1985)  
Hypothesis 3: Monophyly of Odontaspidae  
Hypothesis 4: Sister group relationship between Lamnidae and Cetorhinidae (Compagno, 1990)  
Hypothesis 5: Sister group relationship between a monophyletic Alopidiidae and a clade comprising the Lamnidae and the Cetorhinidae (Compagno, 1990)  
Hypothesis 6: New hypothesis proposed herein (Fig. 7)  
Hypothesis 7: Compagno’s (1990) hypothesis of relationships for all members within the Lamniformes (Fig. 1A)
Hypothesis 8: LogDet tree (Fig. 6B)
Hypothesis 9: MPT derived from equally weighted transitions and transversions (Fig. 5A)
Hypothesis 10: MPT derived from tranversion parsimony analysis (Fig. 5B)

Figure 9 presents the fit of these hypotheses to both the raw data set and to the data set reflecting transversional differences only. Of all of the hypotheses tested, the sister group relationship between the Lamnidae and the Cetorhinidae (H4) requires the fewest extra steps. Indeed, this relationship requires no extra steps when applied to the raw data and only one extra step when applied to the transversion data. The monophyly of the thresher sharks (Alopiidae) was the next most tenable hypothesis, requiring 16 extra steps when fitted to the raw data and 2 extra steps when applied to the transversion data. Given that the cost of invoking monophyly for the Alopiidae is low relative to other hypotheses, and given that there is strong morphological evidence to support monophyly, it is likely that the molecular-based inference suggesting a nonmonophyletic Alopiidae is erroneous.

Maisey's hypothesis (H2) for the sister group relationship between the two filter feeders, Cetorhinus and Megachasma, is not supported; neither is the monophyly of the family Odontaspidae (H3). Compagno's proposed sister group relationship between the Alopiidae and a clade containing the Lamnidae and Cetorhinidae (H5) is also not supported. It should be noted that as hypotheses of relationship become more restrictive or "explain more relationships," so they are prone to cost extra steps. Thus, for example, it would be inappropriate to assert that Compagno's (1990) hypothesis (H7) for the relationships among members of the entire order is less well supported than Maisey's hypothesis for the monophyly of the filter-feeding trait (H2) because Compagno's hypothesis forwards relationships for all 14 taxa while Maisey's forwards a relationship for just one pair. A more appropriate comparison to evaluate Compagno's hypothesis might be made by contrasting the extra steps required to fit Compagno's tree with those required to fit the LogDet hypothesis (H8) to the data. Both of these hypotheses are of the same "explanatory magnitude." When this is carried out, the cost of fitting Compagno's tree is 56 extra steps for the raw data and 18 extra steps for the transversion data. In contrast, the LogDet hypothesis requires 14 extra steps for the raw data and 21 extra steps for the transversion data. Thus, Compagno's hy-
HYPOTHESIS

<table>
<thead>
<tr>
<th>Hypothesis</th>
<th>Tis + Tvs</th>
<th>Tvs only</th>
</tr>
</thead>
<tbody>
<tr>
<td>H1: monophyly of Alopidae</td>
<td>4957 (16)</td>
<td>1421 (2)</td>
</tr>
<tr>
<td>H2: monophyly of filter-feeding</td>
<td>4982 (41)</td>
<td>1435 (16)</td>
</tr>
<tr>
<td>H3: monophyly of Odontaspididae</td>
<td>4975 (34)</td>
<td>1433 (14)</td>
</tr>
<tr>
<td>H4: Sister relationship between Lamnidae and Cetorhinidae</td>
<td>4941 (0)</td>
<td>1420 (1)</td>
</tr>
<tr>
<td>H5: ((Alopidae (Lamnidae + Cetorhinidae))</td>
<td>4978 (37)</td>
<td>1433 (14)</td>
</tr>
<tr>
<td>H6: New hypothesis</td>
<td>4988 (19)</td>
<td>1421 (2)</td>
</tr>
<tr>
<td>H7: Compagno 1990 hypothesis</td>
<td>4997 (56)</td>
<td>1437 (18)</td>
</tr>
<tr>
<td>H8: Log Det tree</td>
<td>4955 (14)</td>
<td>1440 (21)</td>
</tr>
<tr>
<td>H9: (eq. wt. parsimony) 1 MPT</td>
<td>4977 (56)</td>
<td>1435 (16)</td>
</tr>
<tr>
<td>H10: (transition parsimony) 2 MPTs</td>
<td>4997 / 5001 (60)</td>
<td>(60)</td>
</tr>
</tbody>
</table>

FIGURE 9 The number of steps required when different hypotheses are fitted to data. There are 10 hypotheses (H1 to H10) and two versions of the data set (one comprising transitions and transversions, the other transversions only). The number of steps required for each hypothesis is shown in its corresponding cell. The number of "extra" steps beyond the number required by the MPT is depicted in italics. Hypotheses H1 to H5 propose relationships for subsets of the taxa rather than for all 14 taxa as is the case for hypotheses H6 to H10. Each hypothesis was fitted to data as a constraint. A heuristic search was implemented to determine the shortest tree length consistent with the constraint.

Hypothesis has a better fit to the transversion data, but a worse fit to the raw data. This highlights the fact that hypotheses can vary in their fit according to the type of analysis employed. Indeed, when the two most parsimonious trees resulting from transversion parsimony analysis are fitted to the raw data set, they require more extra steps than do any of the other hypotheses proposed (56 and 60 extra steps, respectively).

F. Correspondence to the Fossil Record

Because the fossil record of lamniform teeth is dense and continuous, it is reasonable to expect a correspondence between branch lengths inferred from sequence data and first appearances of corresponding lineages in the fossil record. A meaningful correspondence requires (1) that the phylogenetic inference based on molecular data be correct, (2) that the fossil record be sufficiently well sampled to ensure that segments of lineages are not "missed," and (3) that fossil teeth be correctly assigned to their corresponding lineages. There is no way of evaluating conditions 1 and 2 a priori. However, it is known that the task of correctly assigning teeth to lineages (condition 3) has not been straightforward. Extant lamniforms represent the "pruned down" remnants of a greater lamniform diversity which has, for the most part, become extinct. Most of the tooth lineages in the fossil record represent
evolutionary "cul-de-sacs" that have no extant descendants. Many of these "cul-de-sac" lineages show evolutionary parallelism with lineages that have led to extant forms (Capetta, 1987). This has made the task of identifying and tracking anagenetic change particularly difficult and requires careful evaluation of within- and among-species variation at successive increments along lineages (Espinosa-Arrubarrena, 1987). As more of these detailed stratigraphic studies are completed, a more accurate picture of lamniform evolution should emerge from the fossil record.

The authors contrasted inferred amount of molecular change (i.e., branch length) and first appearance estimates derived from the fossil record for the new hypothesis, for Compagno's (1990) hypothesis, and for Maisey's (1985) hypothesis. Sequence data were fitted to each of the hypotheses separately using the "enforced molecular clock" option of the maximum likelihood platform of PAUP* 4.0. Inferred branch lengths were contrasted with corresponding first appearance estimates derived from the fossil record (Fig. 10). Considerable correspondence is seen for the new hypothesis ($r^2 = 0.63$) and for Compagno's (1990) hypothesis ($r^2 = 0.62$). Note that while the data appear to fit a "molecular clock," considerable leverage is exerted by taxa at the extremes of the distribution. Less correspondence is seen for Maisey's hypothesis ($r^2 = 0.5$). This may be due to the fact that Maisey's hypothesis is substantially less resolved.

In the plots corresponding to the new hypothesis and to Compagno's hypothesis, the basking shark lineage Cetorhinus (g) falls conspicuously above the regression line. Indeed, $r^2$ increases to 0.88 for the new hypothesis and 0.68 for Compagno's (1990) hypothesis if data for Cetorhinus are excluded. The aberrant placement of Cetorhinus could reflect an accelerated rate of molecular evolution, an incorrect first appearance estimate in the fossil record, or an incorrect placement of the taxon in both phylogenies. Because the teeth of Cetorhinus are so different from those of any other extant lamniform, we suspect that transitional forms (Capetta, 1987). Indeed, Compagno (1990) has suggested that Cetorhinus may have evolved its distinctive feeding apparatus from odontaspidid-like features by jaw-size exaggeration, acquisition of papillose gill rakers, and modification of jaw protrusion mechanisms for suction feeding. Such a scenario is loosely consistent with the shared derived placement of Megachasma in a clade with Odontaspis, Pseudocarcharias, and Alopias.

The presented sequence data do not support a monophyletic Odontaspididae. Although O. noronhai has not yet been sampled, it is likely that Carcharias taurus and O. ferox are nonmonophyletic. A number of authors have already suggested this (Compagno, 1990; Capetta, 1987). Indeed, Compagno (1990) presents the Odontaspididae as two basally adjacent paraphyletic lineages (Fig. 1b). We recommend a reexamination of morphological characteristics for the group. Based on molecular data and the past failure to identify compelling shared-derived features for Odontaspis and Carcharias, we anticipate the splitting of the Odontaspididae into two families to better reflect their phylogenetic distinctness.

Our new hypothesis asserts that the three species of thresher shark (Alopiidae), Pseudocarcharias, Megachasma, and Odontaspis, constitute an ancient monophyletic assemblage. This grouping is found in 74% of bootstrap replicates, implying considerable reliability. However, supporting morphological evidence for this relationship has yet to be discovered. We note that this inference is at odds with prevailing views (Maisey, 1985).

### G. Implications of New Hypothesis

Perhaps the most striking implication of our presented hypothesis (Fig. 7) is that filter feeding has evolved twice within extant Lamniformes and that the similarities in jaw articulation seen between Megachasma and Cetorhinus (Maisey, 1985) are the consequence of convergence. This hypothesis (independent origins of filter feeding) is more in keeping with Compagno's (1990) view of lamniform phylogeny than it is with Maisey's (1985). However, our tree differs from Compagno's, Megachasma clusters with Odontaspis, Pseudocarcharias, and the Alopiidae in our presented tree (Fig. 7) whereas Megachasma is the sister group to a clade containing Lamnidae, Cetorhinus, and the Alopiidae in Compagno's tree (Fig. 1b).

Morphological data show support for some aspects of the presented hypothesis (Fig. 7). Cetorhinus and the Lamnidae are united by a number of derived traits (i.e., lunate tail, peduncle depressed with lateral keels, enlarged gill slits, presence of ectethmoid processes on the chondrocranium to limit jaw protrusion, and suborbital shelves with prominent lateral wings behind orbital notches (Compagno, 1990)). The implied sister group relationship between Megachasma and Pseudocarcharias has less morphological support. However, possible synapomorphies include similar intestinal valve counts, dorsal fin shapes, and a diel vertical migration habit. Compagno (1990) has suggested that Megachasma may have evolved its distinctive feeding apparatus from odontaspidid-like features by jaw-size exaggeration, acquisition of papillose gill rakers, and modification of jaw protrusion mechanisms for suction feeding. Such a scenario is loosely consistent with the shared derived placement of Megachasma in a clade with Odontaspis, Pseudocarcharias, and Alopias.

The presented sequence data do not support a monophyletic Odontaspididae. Although O. noronhai has not yet been sampled, it is likely that Carcharias taurus and O. ferox are nonmonophyletic. A number of authors have already suggested this (Compagno, 1990; Capetta, 1987). Indeed, Compagno (1990) presents the Odontaspididae as two basally adjacent paraphyletic lineages (Fig. 1b). We recommend a reexamination of morphological characteristics for the group. Based on molecular data and the past failure to identify compelling shared-derived features for Odontaspis and Carcharias, we anticipate the splitting of the Odontaspididae into two families to better reflect their phylogenetic distinctness.

Our new hypothesis asserts that the three species of thresher shark (Alopiidae), Pseudocarcharias, Megachasma, and Odontaspis, constitute an ancient monophyletic assemblage. This grouping is found in 74% of bootstrap replicates, implying considerable reliability. However, supporting morphological evidence for this relationship has yet to be discovered. We note that this inference is at odds with prevailing views (Maisey,
13. Lamniform Sharks

<table>
<thead>
<tr>
<th>Lineage</th>
<th>Age MYBP</th>
<th>Stage</th>
<th>Fossil Representative</th>
<th>Source</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mitsukurina</td>
<td>124-112</td>
<td>Aptian</td>
<td><em>Anomodontodon principalis</em></td>
<td>Cappetta, 1975</td>
</tr>
<tr>
<td>Carcharias</td>
<td>112-97</td>
<td>Albian</td>
<td><em>C. striata</em></td>
<td>Dalinkevich, 1935</td>
</tr>
<tr>
<td>Odontaspis</td>
<td>83-74</td>
<td>Campanian</td>
<td><em>O. australis</em></td>
<td>Cappetta &amp; Case, 1975</td>
</tr>
<tr>
<td>Alopaia</td>
<td>55</td>
<td>Ypresian</td>
<td><em>A. crocollari</em></td>
<td>Ward, 1978</td>
</tr>
<tr>
<td>Mesoselachus</td>
<td>42</td>
<td>Bartonian</td>
<td><em>unpub. material</em></td>
<td>Ward Collection (pers. comm.)</td>
</tr>
<tr>
<td>Lamnidae</td>
<td>42-38</td>
<td>Bartonian</td>
<td><em>unpub. material</em></td>
<td>Ward Collection (pers. comm.)</td>
</tr>
<tr>
<td>Cetorhinus</td>
<td>35-29</td>
<td>Rupelian</td>
<td><em>C. parvus</em></td>
<td>Leriche, 1908</td>
</tr>
<tr>
<td>Carcharodon</td>
<td>65</td>
<td>Danian</td>
<td><em>Cetorhinus</em></td>
<td>Applegate &amp; Espinosa, 1996</td>
</tr>
<tr>
<td>Heterus</td>
<td>55</td>
<td>Ypresian</td>
<td><em>Heterus</em></td>
<td>Leriche, 1905</td>
</tr>
<tr>
<td>Lamnidae</td>
<td>35-29</td>
<td>Pliocene</td>
<td><em>Lamnidae</em></td>
<td>Herman, 1974</td>
</tr>
<tr>
<td>Lamnidae</td>
<td>29-25</td>
<td>Artian</td>
<td><em>Lamnidae</em></td>
<td>Espinosa, 1987</td>
</tr>
</tbody>
</table>

**FIGURE 10** A comparison of inferred branch length and first appearance times in the fossil record. (A) First appearance estimates in millions of years before present (MYBP) for lamniform lineages seen in the fossil record. The geological stage and the fossil taxon representing each lineage are also shown. Note that the first appearance time for the *Carcharodon* lineage is controversial. Whereas Purdy (1996) and Applegate and Espinosa-Arzarbarrena (1996) suggest an early Paleocene origination for the lineage, other workers suggest a later origination for the lineage (D. Ward, personal communication). (B) Phylograms inferred for hypotheses when fitted to sequenced data under an enforced molecular clock. Relative branch lengths are depicted directly on the phylograms. Taxon labels are as in Fig. 2. Letters in lowercase, bold type correspond to similarly labeled lineages in A. (C) Relationship between first appearance times (from A) and inferred branch length (from B). Letters in lowercase, bold type correspond to those in A and B. See text for details.
1985; Compagno, 1990), which suggest a relatively recent shared ancestry between the Alopiidae and the Lamnidae with respect to Odontaspis, Pseudoceorhayias, and Megachasma. Shared derived features linking the Alopiidae and the Lamnidae include an erect first dorsal fin, the partial extension of stiffening cartilagenous elements into the fin web (semiplesodic fins), and intestinal valve counts increasing to a range of 33–55 (Compagno, 1990). In addition, both the lamnids (Carey et al., 1985) and the alopids exhibit adaptations for endothermy [A. supercilious has a conspicuous vascular heat exchanger behind the eye (G. J. P. Naylor, personal observation)]. If the presented DNA-based hypothesis is correct and the two families do not constitute a monophyletic group (together with Cetorhinus), then these features are likely convergent.

The presented phylogram (Fig. 10B) is made up of long terminal branches connected by short internodal segments. This pattern suggests a history of early diversification followed by directional selection along divergent trajectories leading to highly autopomorphic extant taxa; a pattern reflected by both morphological and molecular characters. Such patterns of diversification are known to be problematic for phylogenetic analysis. Thus, although we stand by the proposed hypothesis (Fig. 7) as the best estimate of phylogeny for the data at hand, we acknowledge that no single fully resolved tree is decisively supported by the data and that the best-fitting phylogenetic hypothesis for the group may change as more data are collected.

Acknowledgements

It has taken 10 years to acquire the tissue samples required for this project. We are grateful to all those who have either donated tissue directly or assisted Gavin Naylor collecting samples in the field: John Stevens (Australia); M. Miya and K. Yano (Japan); Janine Caira, Jack Casey, José Castro, Don De Maria, Ken Goldman, Nancy Kohler, John Morissey, Lisa Natsnson, Heidi Robek, Wes Pratt, and Greg Skomal (North America); Gergy Cliff and the shark net inspection crews of the Natal Sharks Board (South Africa); Shou-Jeng Joung, Che-Tsung Chen and “Mr. Chen” (at the Nan Fan Ao fish Janding Taiwan); and David Sims (United Kingdom). Most of the field work for this project was supported by NSF Grant BSR-87-08121 to G. Vermeij and G.J.P.N. The laboratory work was supported by a Sloan Postdoctoral Fellowship (to G.J.P.N.), NSF grant DEB-92-20640 to W. Brown, and the University of Nevada–Las Vegas (to A.P.M.). G.J.P.N. acknowledges access to laboratory equipment and chemicals generously made available by Tom Dowling to facilitate completion of the project while at Arizona State University. Thanks to Tom Kocher and a particularly attentive anonymous reviewer for critical comments which led to the improvement of the original manuscript.

References


13. Lamniform Sharks


### APPENDIX I: Collection Data for Specimens Sequenced

<table>
<thead>
<tr>
<th>Taxon</th>
<th>ID*</th>
<th>Locality</th>
<th>.Tissue sampled by</th>
<th>Gene</th>
<th>Tissue NADH Cyto-chrome b</th>
</tr>
</thead>
<tbody>
<tr>
<td>Alopias pelagicus</td>
<td>954</td>
<td>Sao, fish landing, Taiwan</td>
<td>G.J.P.N.</td>
<td>x</td>
<td>--</td>
</tr>
<tr>
<td>Alopias pelagicus</td>
<td>925</td>
<td>Sao, fish landing, Taiwan</td>
<td>G.J.P.N.</td>
<td>x</td>
<td>--</td>
</tr>
<tr>
<td>Alopias pelagicus</td>
<td>Alstai81</td>
<td>Taiwan</td>
<td>C. Chen</td>
<td>x</td>
<td>x</td>
</tr>
<tr>
<td>Alopias pelagicus</td>
<td>Alpe</td>
<td>Taiwan</td>
<td>C. Chen</td>
<td>--</td>
<td>x</td>
</tr>
<tr>
<td>Alopias superciliosus</td>
<td>871</td>
<td>37° 13N 74° 31W, off Virginia coast</td>
<td>G.J.P.N.</td>
<td>x</td>
<td>--</td>
</tr>
<tr>
<td>Alopias superciliosus</td>
<td>874</td>
<td>35° 31N 74° 46W, off North Carolina coast</td>
<td>G.J.P.N.</td>
<td>x</td>
<td>--</td>
</tr>
<tr>
<td>Alopias superciliosus</td>
<td>1001</td>
<td>Sao, fish landing, Taiwan</td>
<td>G.J.P.N.</td>
<td>x</td>
<td>--</td>
</tr>
<tr>
<td>Alopias superciliosus</td>
<td>Alstai81</td>
<td>Taiwan</td>
<td>C. Chen</td>
<td>--</td>
<td>x</td>
</tr>
<tr>
<td>Alopias superciliosus</td>
<td>622</td>
<td>28° 10N 79° 55W, off Florida coast</td>
<td>G.J.P.N.</td>
<td>--</td>
<td>x</td>
</tr>
<tr>
<td>Alopias vulpinus</td>
<td>429</td>
<td>Bay shore, Long Island</td>
<td>G.J.P.N.</td>
<td>x</td>
<td>x</td>
</tr>
<tr>
<td>Alopias vulpinus</td>
<td>Alstai81</td>
<td>Japan</td>
<td>M. Miya</td>
<td>x</td>
<td>x</td>
</tr>
<tr>
<td>Carcharodon carcharias</td>
<td>864</td>
<td>South shore, Long Island</td>
<td>G.J.P.N.</td>
<td>x</td>
<td>--</td>
</tr>
<tr>
<td>Carcharodon carcharias</td>
<td>1060</td>
<td>Pacifica, California</td>
<td>Cal Academy</td>
<td>x</td>
<td>--</td>
</tr>
<tr>
<td>Carcharodon carcharias</td>
<td>1064</td>
<td>Farallon Islands, California</td>
<td>K. Goldman</td>
<td>x</td>
<td>x</td>
</tr>
<tr>
<td>Carcharodon carcharias</td>
<td>677</td>
<td>Off Natal Coast, South Africa</td>
<td>G. Cliff</td>
<td>--</td>
<td>x</td>
</tr>
<tr>
<td>Ceterhinus maximus</td>
<td>1058</td>
<td>Barbican, Plymouth, UK</td>
<td>D. Sims</td>
<td>x</td>
<td>x</td>
</tr>
<tr>
<td>Ceterhinus maximus</td>
<td>CemaAlSI</td>
<td>Australia</td>
<td>J. Stevens</td>
<td>x</td>
<td>x</td>
</tr>
<tr>
<td>Carcharhinus plumbeus</td>
<td>556</td>
<td>Off Makapu point, Oahu, Hawaii</td>
<td>G.J.P.N.</td>
<td>--</td>
<td>x</td>
</tr>
<tr>
<td>Carcharhinus plumbeus</td>
<td>912</td>
<td>35° 25N 74° 53W, off Cape Hatteras</td>
<td>G.J.P.N.</td>
<td>x</td>
<td>--</td>
</tr>
<tr>
<td>Carcharodon pereosus</td>
<td>477</td>
<td>Fish landing, Port of Spain, Trinidad</td>
<td>G.J.P.N.</td>
<td>x</td>
<td>x</td>
</tr>
<tr>
<td>Carcharhinus pereosus</td>
<td>916</td>
<td>36° 26N 75° 41W, off North Carolina coast</td>
<td>G.J.P.N.</td>
<td>x</td>
<td>x</td>
</tr>
<tr>
<td>Carcharias taurus</td>
<td>1064</td>
<td>Steinhardt aquarium</td>
<td>Cal Academy</td>
<td>--</td>
<td>x</td>
</tr>
<tr>
<td>Carcharias taurus</td>
<td>627</td>
<td>31° 29N 80° 52W, off Georgia coast</td>
<td>G.J.P.N.</td>
<td>x</td>
<td>--</td>
</tr>
<tr>
<td>Carcharias taurus</td>
<td>SA1</td>
<td>Off Natal Coast, South Africa</td>
<td>G. Cliff</td>
<td>--</td>
<td>x</td>
</tr>
<tr>
<td>Galeocerdo cuvier</td>
<td>888</td>
<td>30° 02N 80° 38W, off Floridz coast</td>
<td>G.J.P.N.</td>
<td>x</td>
<td>--</td>
</tr>
<tr>
<td>Galeocerdo cuvier</td>
<td>553</td>
<td>Off Makapu point, Oahu, Hawaii</td>
<td>G.J.P.N.</td>
<td>--</td>
<td>x</td>
</tr>
<tr>
<td>Isurus oxyrinchus</td>
<td>237</td>
<td>Ocean City, Maryland</td>
<td>G.J.P.N.</td>
<td>x</td>
<td>x</td>
</tr>
<tr>
<td>Isurus oxyrinchus</td>
<td>873</td>
<td>35° 31N 74° 46W, off Virginia coast</td>
<td>G.J.P.N.</td>
<td>--</td>
<td>x</td>
</tr>
<tr>
<td>Isurus oxyrinchus</td>
<td>412</td>
<td>Point Pleasant, New Jersey</td>
<td>G.J.P.N.</td>
<td>--</td>
<td>x</td>
</tr>
<tr>
<td>Isurus paucus</td>
<td>614</td>
<td>Florida, Keys</td>
<td>D. de Maria</td>
<td>x</td>
<td>x</td>
</tr>
<tr>
<td>Isurus paucus</td>
<td>JIC#10</td>
<td>Off South Florida coast</td>
<td>J. L. Castro</td>
<td>--</td>
<td>x</td>
</tr>
<tr>
<td>Isurus paucus</td>
<td>JIC#8</td>
<td>Off South Florida coast</td>
<td>J. L. Castro</td>
<td>--</td>
<td>x</td>
</tr>
<tr>
<td>Lamna ditropis</td>
<td>1062</td>
<td>Off North California coast</td>
<td>Cal Academy</td>
<td>--</td>
<td>x</td>
</tr>
<tr>
<td>Lamna ditropis</td>
<td>Lad#:MM12</td>
<td>Japan</td>
<td>M. Miya</td>
<td>x</td>
<td>x</td>
</tr>
<tr>
<td>Lamna nasus</td>
<td>633</td>
<td>Gulf of Maine</td>
<td>G.J.P.N.</td>
<td>x</td>
<td>x</td>
</tr>
<tr>
<td>Mitsukurina owstoni</td>
<td>1057</td>
<td>Off coast of Tasmania</td>
<td>J. Stevens</td>
<td>x</td>
<td>x</td>
</tr>
<tr>
<td>Mitsukurina owstoni</td>
<td>Miow JI</td>
<td>Japan</td>
<td>M. Miya</td>
<td>x</td>
<td>x</td>
</tr>
<tr>
<td>Megachasma pelagios</td>
<td>Mepe</td>
<td>Japan</td>
<td>K. Yano, M. Miya</td>
<td>x</td>
<td>x</td>
</tr>
<tr>
<td>Negaprion brevirostris</td>
<td>617</td>
<td>25° 18N 81° 56W, SW coast of Florida</td>
<td>G.J.P.N.</td>
<td>--</td>
<td>x</td>
</tr>
<tr>
<td>Negaprion brevirostris</td>
<td>558</td>
<td>Off Cosgrove, Florida Keys</td>
<td>D. de Maria</td>
<td>x</td>
<td>--</td>
</tr>
<tr>
<td>Odontaspis ferox</td>
<td>JIC#11</td>
<td>South Car Caye, Bahamas</td>
<td>J. L. Castro</td>
<td>--</td>
<td>x</td>
</tr>
<tr>
<td>Prionace glauca</td>
<td>540</td>
<td>220 miles S S.W off Oahu, Hawaii</td>
<td>G.J.P.N.</td>
<td>--</td>
<td>x</td>
</tr>
<tr>
<td>Prionace glauca</td>
<td>920</td>
<td>38° 48N 72° 58W, off New Jersey coast</td>
<td>G.J.P.N.</td>
<td>--</td>
<td>x</td>
</tr>
<tr>
<td>Pseudeocharacias kamoharai</td>
<td>1033</td>
<td>Sao, fish landing, Taiwan</td>
<td>S. J. Joung</td>
<td>x</td>
<td>--</td>
</tr>
<tr>
<td>Pseudeocharacias kamoharai</td>
<td>1034</td>
<td>Sao, fish landing, Taiwan</td>
<td>S. J. Joung</td>
<td>x</td>
<td>--</td>
</tr>
<tr>
<td>Sphyraena lewini</td>
<td>Sple</td>
<td>Hawaii</td>
<td>A.P.M.</td>
<td>--</td>
<td>x</td>
</tr>
<tr>
<td>Sphyraena lewini</td>
<td>886</td>
<td>35° 10N 80° 11W, off Florida coast</td>
<td>G.J.P.N.</td>
<td>x</td>
<td>--</td>
</tr>
<tr>
<td>Sphyraena tiburo</td>
<td>500</td>
<td>Marzanilla Bay, Trinidad</td>
<td>G.J.P.N.</td>
<td>--</td>
<td>x</td>
</tr>
<tr>
<td>Sphyraena tiburo</td>
<td>501</td>
<td>Marzanilla Bay, Trinidad</td>
<td>G.J.P.N.</td>
<td>--</td>
<td>x</td>
</tr>
</tbody>
</table>

*Assigned field number in G.J.P.N. or A.P.M. data base.*