Phylogenomic analysis of carangimorph fishes reveals flatfish asymmetry arose in a blink of the evolutionary eye

Richard C. Harrington, Brant C. Faircloth, Ron I. Eytan, W. Leo Smith, Thomas J. Near, Michael E. Alfaro and Matt Friedman

Abstract

Background: Flatfish cranial asymmetry represents one of the most remarkable morphological innovations among vertebrates, and has fueled vigorous debate on the manner and rate at which strikingly divergent phenotypes evolve. A surprising result of many recent molecular phylogenetic studies is the lack of support for flatfish monophyly, where increasingly larger DNA datasets of up to 23 loci have either yielded a weakly supported flatfish clade or indicated the group is polyphyletic. Lack of resolution for flatfish relationships has been attributed to analytical limitations for dealing with processes such as nucleotide non-stationarity and incomplete lineage sorting (ILS). We tackle this phylogenetic problem using a sequence dataset comprising more than 1,000 ultraconserved DNA element (UCE) loci covering 45 carangimorphs, the broader clade containing flatfishes and several other specialized lineages such as remoras, billfishes, and archerfishes.

Results: We present a phylogeny based on UCE loci that unequivocally supports flatfish monophyly and a single origin of asymmetry. We document similar levels of discordance among UCE loci as in previous, smaller molecular datasets. However, relationships among flatfishes and carangimorphs recovered from multilocus concatenated and species tree analyses of our data are robust to the analytical framework applied and size of data matrix used. By integrating the UCE data with a rich fossil record, we find that the most distinctive carangimorph bodyplans arose rapidly during the Paleogene (66.0–23.03 Ma). Flatfish asymmetry, for example, likely evolved over an interval of no more than 2.97 million years.

Conclusions: The longstanding uncertainty in phylogenetic hypotheses for flatfishes and their carangimorph relatives highlights the limitations of smaller molecular datasets when applied to successive, rapid divergences. Here, we recovered significant support for flatfish monophyly and relationships among carangimorphs through analysis of over 1,000 UCE loci. The resulting time-calibrated phylogeny points to phenotypic divergence early within carangimorph history that broadly matches with the predictions of adaptive models of lineage diversification.

Keywords: Adaptive radiation, Carangimorpha, Evolutionary innovation, Pleuronectiformes, UCE, Ultraconserved elements

Background

During the past decade, a series of molecular phylogenetic analyses drawing on increasingly larger samples of taxa and genetic loci have transformed our understanding of evolutionary relationships among acanthomorphs or spiny-rayed fishes [1–6], a hyperdiverse lineage that includes nearly one in three living vertebrate species. These studies support the monophyly of many clades previously recognized by morphological phylogeneticists (e.g., Tetraodontiformes, Lophiiformes), but reject the coherence of other classical groups (e.g., Scombroidei inclusive of billfishes [7], Labroidei [8]) by removing some of their core members to other, distantly related lineages [9–11]. In resolving the ‘bush’ at the top of the teleost tree of life, these molecular phylogenies have exposed striking examples of morphological, physiological,
and functional convergence among acanthomorphs [10, 11], and revealed unexpected groupings of lineages not previously regarded as closely related [11].

A well-supported radiation [1, 12, 13], variously termed Clade L [14], Carangimorpha [2, 3], or Carangimorpharia [4, 15], represents one of the most surprising features of the emerging picture of acanthomorph interrelationships. Carangimorphs include anatomically disparate lineages characterized by remarkable behavioral and anatomical novelties: eye and brain heating organs coupled with long rostra and numerous specializations for rapid swimming in Xiphioidae (marlins and swordfishes) [16]; cranial adhesion discs and commensal lifestyle in Echeneidae (remoras) [17, 18]; expanded, tactile pectoral-fin rays in Polyodontidae (threadfins) [19]; sophisticated spine-based venom delivery systems in Carangidae (scombroidin jacks) [20]; and use of water jets to capture aerial prey in Toxotidae (archerfishes) [21]. These specializations are joined by even more singular innovations in arguably the most peculiar carangimorph lineage: Pleuronectiformes (flatfishes). Flatfishes, including familiar food fish like halibut, sole, and plaice, are characterized by profound cranial asymmetry resulting from the migration of one eye to the opposite side of the skull during larval metamorphosis. This extreme developmental resculpturing of the head permits adult flatfishes to rest on the seafloor on their eyeless or ‘blind’ side, leaving both eyes of the ‘eyed’ side unobstructed by sediment.

Although carangimorph monophyly is well supported, relationships among its principal lineages are weakly supported and highly variable between studies (Fig. 1). Flatfishes provide the most concrete illustration of the uncertainty in carangimorph relationships. Ichthyologists have overwhelmingly regarded the orbital migration and associated lateralized behavior of flatfishes as unique innovations that strongly support pleuronectiform monophyly [22–30] (but see [31–33]). However, analyses that do recover a flatfish clade provide only weak statistical support for its monophyly (bootstrap support < 70 %; Fig. 1) [2, 4, 34–36]. Several published phylogenies support the monophyly of a subset of flatfishes representing Pleuronectoidei of anatomical classifications [23, 37, 38] and place Psettodes as the sister taxon of a closely related—but symmetrical—lineage that varies among analyses (Fig. 1; Centropomidae [snooks] + Xiphioidae [1]; Sphyraenidae [barracudas] [2]; Latinae [Nile perches] [3]; Nematistidae [rooesterfish] [4]; Toxotidae [38]). Although trees favoring a polyphyletic origin of Pleuronectiformes lack sufficient statistical support to reject the morphological null hypothesis of monophyly [34], some researchers have nevertheless concluded that asymmetry arose independently in Psettodes and pleuronectoids [38, 39]. These conclusions resurrect a pre-cladistic hypothesis in which generalized percomorph traits apparent in Psettodes suggest that cranial symmetry evolved within this lineage independently of that in pleuronectoids [31, 32], an inference that casts uncertainty on the phylogenetic placement and evolutionary significance of early flatfishes showing incomplete orbital migration [39].

The recent debate regarding flatfish monophyly spotlights the difficulties faced by many phylogenetic studies, particularly in the use of molecular data for radiations characterized by short internodes deep in evolutionary time. Methodological challenges, such as accounting for base compositional bias (i.e., non-stationarity [40]) and long branch attraction [41, 42], as well as natural phenomena such as horizontal gene flow and incomplete lineage sorting (ILS) can result in inference of gene trees that do not reflect a clade’s history of speciation [43–45]. Even under certain scenarios of branch length in species trees, the most frequent gene trees do not reflect the topology of the species tree (the so-called anomaly zone [46]). The use of analytical approaches to account for non-stationarity or application of the coalescent model for ILS is important for improving accuracy of gene tree and species tree estimation, but the addition of large numbers of unlinked loci may be the most direct approach for improving phylogenetic accuracy in the face of these processes [47]. Thus far, incrementally larger DNA sequence datasets of up to 23 loci have produced inconsistent increases in support for relationships among major carangimorph lineages, particularly with regard to flatfish monophyly (Fig. 1) and the identification of a symmetrical sister group.

Using recent advances in phylogenomics and high-throughput sequencing, we assembled a dataset of ultra-conserved DNA elements (UCEs) and their flanking sequences representing over 1000 loci sampled from 45 carangimorph species. Here we use this novel dataset in conjunction with the rich fossil record of Carangimorpha to: (i) establish a well-supported hypothesis of relationships among anatomically disparate carangimorph lineages; (ii) provide a statistically decisive molecular solution to the ‘pleuronectiform problem’; and (iii) estimate divergence times for carangimorphs, with an emphasis on constraining the timescale over which remarkable anatomical innovations such as flatfish asymmetry and other specialized carangimorph bodyplans likely arose.

**Methods**

We use a probe set developed for application to acanthomorph phylogenetics to generate sequence data for approximately 1200 UCE loci [48]. These loci vary in size and number of variable sites per locus, but average between 300 and 500 nucleotides in length. This approach has provided resolution across a range of evolutionary depths in phylogenetic studies of acanthomorphs.
Fig. 1 (See legend on next page.)
Ceratoscopelus warmingii

We rooted hcluster search in PartitionFinder v 1.1.1 [53, 54], applied using the Bayesian Information Criterion and model. The best-fitting partitioning scheme was obtained using the best-fitting partitioning likelihood (ML) searches for the phylogenetic tree that best fit the data using RAxML v. 8.0.19 [52], and the GTRGAMMA model. We also used RAxML to generate non-parametric bootstrap pseudoreplicates from each locus. We inferred the species tree from individual gene trees using ASTRAL v. 4.4.4 [57] under the bootstrap option. All other ASTRAL parameters were the defaults.

Gene tree-species tree analysis of UCE data

We inferred gene trees from the individual locus alignments that comprised the 75 % complete matrix described above, and we subset those gene trees based on the names of loci present in the 95 % and 100 % complete data matrices for gene tree-species tree analysis of each dataset, respectively. To infer locus-specific trees from each data matrix, we conducted 20 maximum-likelihood (ML) searches for the phylogenetic tree that best fit the data using RAxML v. 8.0.19 [52] and the GTRGAMMA model. We also used RAxML to generate ~100 non-parametric bootstrap pseudoreplicates from each locus. We inferred the species tree from individual gene trees using ASTRAL v. 4.4.4 [57] under the bootstrap option. All other ASTRAL parameters were the defaults.

Analysis of concordance among loci

Unlinked loci exhibit independent histories of sorting through populations over evolutionary timescales, and this can result in gene trees with non-matching topologies. Therefore, calculations of statistical support for clades derived from total-evidence concatenation or species tree analyses may not accurately demonstrate the genome-wide agreement on a lineage’s evolutionary history [58]. To evaluate the level of concordance among loci, and in particular the number of loci that recover a flatfish clade, we performed Bayesian Concordance Analysis (BCA) using BUCKy version 1.4.4 [59, 60]. BUCKy uses posterior tree density distributions from individual loci to estimate the proportion of loci that support branching patterns in trees (i.e., concordance factors, CF), and assembles taxa with highest CF values into clades with the least conflict in order to generate a primary concordance tree. BUCKy also applies concordance factors in a quartet-joining algorithm to construct a population tree, which is consistent with a coalescent framework species tree when gene tree discordance is due to ILS [61]. We generated gene tree distributions for each UCE locus using MrBayes version 3.2.6 [62], with MCMC...
chains run for 1 million generations using HKY model, sampling trees every 200 generations, and discarding the first 2500 as burnin. With these gene tree distributions as input, we used BUCKy to estimate the proportion of gene trees supporting possible relationships among taxa, and to generate a primary concordance and population trees. BUCKy analyses were run for 1 million generations on 4 chains, and a burnin of 100,000 generations. We performed analyses over a range of the discordance prior values (alpha values of: 0.1, 1.0, 10, and 100). Because sequencing was uneven across taxa and loci, we aimed to maximize the number of loci that informed CF estimation by using a subset of taxa, with one representative from each major lineage, including *Psettodes* and two other flatfish (*Psettodes erumei*, *Cyclosetta fimbriata*, *Citharoides macrolepis*, *Lates calcarifer*, *Centropomus mediatus*, *Toxotes jaculatrix*, *Leptobrama muelleri*, *Xiphias gladius*, *Echeneis naucratoides*, *Coryphaena hippurus*, *Trachinotus blochii*, *Alepes kleinii*, *Mene maculatus*, *Sphyraena sphyraena*, and *Polydactylus sexfilis*).

**Divergence time estimation**

We estimated divergence times of carangimorph lineages using both a relaxed molecular clock approach implemented in BEAST v 1.8 [63] and an approximate likelihood calculation in PAML [64]. Due to computational limitations in BEAST, we performed a series of four replicate analyses using different sets of 75 randomly selected loci from the 95% complete matrix (596 loci). For each set of 75 loci, we used PartitionFinder [53] to identify optimal partitioning schemes and assign loci to partitions. This produced between 5 and 8 partitions for each dataset. In BEAST, we applied an uncorrelated lognormal clock model (UCLN), GTR GAMMA site substitution model, and a birth-death speciation tree model. We constrained the topology to match that of the Bayesian and maximum likelihood analyses of the partitioned, concatenated 596 loci analyses. Calibration priors were applied to the same nodes as in BEAST analyses, using soft upper (95%) and lower (1e-300) bounds. To arrive at a partitioning scheme that accounted for variation across sites but that was simple enough to be implemented in PAML, we used APE [66] to root the Exabayes tree inferred from the 95% complete matrix, dropped the tip representing the outgroup lineage (*Ceratoscopelus warmingii*), and set the branch lengths of the resulting tree to 1.0. We then removed *Ceratoscopelus warmingii* from the concatenated, 95% complete supermatrix, and we input the supermatrix and guide tree to DendroPy [67]. We used the parsimony_score method in DendroPy to calculate the number of parsimony informative changes in each site pattern. We used kmeans clustering in R to cluster the distribution of parsimony scores by plotting the within-group sum of squares of parsimony score by a number of potential clusters (range 1–15). Visual examination of the resulting plot suggested five was the appropriate number of clusters. We then clustered the data, computed the cluster means, and assigned individual sites to one of the five clusters. We used a new program in PHYLUCE (phyluce_align_get_aligns_partitioned_by_cluster) to create a partitioned, concatenated alignment where each partition represented one of the five clusters, and we assigned an HKY85 model to each partition. We used the following values for other PAML options and model parameters: BDparas: 1, 1, 0; kappa_gamma: 6, 2; alpha_gamma: 1, 1; rgene_gamma: 2, 203.72, 1; sigma2_gamma: 2, 5, 1. We conducted 8 replicate MCMCTree runs sampling every 250 generations after 5000 generations of burnin for a total of 10000 samples per run. Tracer v 1.6 [56] was used to visually assess convergence and calculate the ESS for all model parameters. Once satisfactory ESS values were attained and convergence verified by visual inspection using Tracer, we combined log files from independent runs to calculate final posterior distributions for all parameters.

**Results**

Our sequencing generated an average of 996 loci per taxon. After alignment and alignment trimming, mean locus length was 365 nucleotides (range: 121–811), with
a mean of 97 parsimony informative sites per locus. The phylogenetic trees inferred in this study from the concatenated datasets are resolved with strong node support and exhibit near-identical topologies regardless of the matrix examined [75 % (1014 loci), 95 % (596 loci) and 100 % (97 loci) complete] and method of analysis (Bayesian and maximum likelihood) (Fig. 2, Additional file 1: Figure S1). Phylogenies inferred from coalescent gene tree species tree (GT-ST) analyses share a majority of nodes in common with results from concatenated analyses (Additional file 1: Figure S2). Disagreement is localized to the deepest divergences within Carangimorpha, which are poorly supported in GT-ST phylogenies. Best practice for the application of coalescent GT-ST methods to genomic data is currently debated [68, 69], and we therefore base our discussion on the phylogenies inferred using the 95 % complete concatenated matrix (Fig. 2). The population tree generated by BCA matched the topology of the concatenated 95 % complete matrix (Additional file 1: Figure S3). We present the phylogenies inferred using concatenation analysis of the 75 % and 100 % complete data matrices, GT-ST, and BUCKy-generated primary concordance and population trees in the extended data (Additional file 1: Figures S1-S3), and we note disagreements between phylogenies resulting from concatenated and GT-ST analyses.

There is strong support for nearly all nodes in phylogenies inferred from concatenated analyses, with Bayesian posterior probabilities (PP) of >0.99 (approximated as ‘1’ below) and bootstrap support percentages (BP) of 100 %. With the exception of the placement of Sphyraenidae, we find strong support for many of the deepest nodes within Carangimorpha, in contrast to previous studies that have inferred inconsistent and poorly supported relationships among these deeply divergent lineages [3, 4, 15, 38].

The monophyly and relationships of flatfishes
Flatfish monophyly is unambiguously supported in phylogenies resulting from all of our concatenated analyses (PP = 1, BP = 100 %), but it also receives strong support in GT-ST analyses, and flatfishes appear as a clade in both the primary concordance and population trees generated by BCA. Concordance factors calculated for the loci containing all fifteen of the subsampled taxa (557 loci contained sequence data for Psettodes and the selected representatives of pleuronectoids and each of the remaining major carangimorph lineages) show that a plurality of loci produced gene trees in which Psettodes forms a clade with the other two sampled flatfishes, to the exclusion of all other carangimorphs (Fig. 3). The genome-wide BUCKy concordance factor estimate of 0.149 [95%CI: 0.114–0.187] for the Psettodes + Pleuronectoidei clade is higher than, and non-overlapping with, concordance factors estimated for any alternative topology in which Psettodes is recovered in a clade with any non-flatfish lineage (Fig. 3). Consistent with previous morphological hypotheses [23, 26–28, 70, 71], we resolve Psettodes as the sister lineage to Pleuronectoidei, which contains all other flatfish species. The UCE-inferred phylogenies place Citharidae as the sister lineage to all remaining pleuronectoids, in agreement with previous morphological [28, 70] and molecular analyses [4]. Interrelationships among other pleuronectoid families are identical to those reported in a recent molecular phylogenetic analysis of flatfishes [15], with universally high support (PP = 1, BP > 95 %).

Carangimorph relationships
We resolve Polynemidae as the sister lineage of a monophyletic flatfish clade (Fig. 2). Although this relationship is strongly supported for some datasets (PP > 0.99 across all datasets, BP = 54, 84, and 100 % for 97, 596, and 1014 locus datasets, respectively), it has not been found in any previous molecular study of carangimorphs or on the basis of morphological features (Fig. 1). The clade containing polynemids and flatfishes is the sister lineage of a well-supported (PP = 1, BP = 100 %) and diverse clade containing species from at least nine families: Toxotidae, Leptobramidae (beach salmon), Menidae, Xiphiidae (swordfish), Istiophoridae (marlins and sailfishes), Carangidae (jacks), Rachycentridae (cobia), Coryphaenidae (dolphinfishes), and Echeneidae. Consistent with results from previous morphological [27, 72, 73] and molecular [3, 4, 15] phylogenetic analyses, we find strong support (PP = 1, BP = 100 %) for a carangiform [27] clade including jacks and echeeoids (cobia, dolphinfishes, and remoras). As presented in previous molecular phylogenies, echeeoids are nested within jacks [1–4, 34] rather than the sister lineage of carangids as hypothesized based on anatomical data [73]; we find that the carangid clades Trachinotini and Scomberoidini are more closely related to echeeoids than they are to other jacks [4, 34]. Within echeeoids, we find unambiguous support (PP = 1, BP = 100 %) for a sister-group relationship between cobia and dolphinfishes to the exclusion of remoras [1, 3, 4, 15, 17, 38, 73, 74] rather than the previously hypothesized relationship between more phenotypically similar remoras and cobia [18].

The UCE-inferred phylogeny resolves a clade of anatomically disparate fishes as the sister lineage of carangiforms (Fig. 2). Within this clade, we recover the moonfish Mene maculata, classified in the monotypic Menidae, as the sister lineage of xiphioids. This morphologically unanticipated relationship is weakly supported in some previous molecular analyses [4, 34], but receives unambiguous support in the UCE trees (PP = 1, BP = 100 %). Consistent with a clear morphological [7, 75] and molecular [3, 4, 10, 15, 34, 38] consensus, the
Fig. 2 (See legend on next page.)
monophyly of xiphoids, containing Xiphiidae and Istiophoridae, is strongly supported (PP = 1, BP = 100 %). We resolve this new clade containing moonfish and billfishes as the sister lineage of a group uniting Toxotes and Lepto-brama, each of which is the only generic representative of its family (Toxotidae, Leptobramidae). This clade is present in the results of multiple molecular analyses [4, 15, 34], although with weak support.

The only disagreement among the phylogenies from our different concatenated analyses involve the deepest nodes in Carangimorpha, and it specifically concerns the relationships between Sphyraenidae (barracudas), Centropomidae (snooks and Nile perches; comprising Centropominae and Latinae), and all remaining carangimorphs (Fig. 2, Additional file 1: Figure S1). The monophyly of both Centropominae and Latinae is consistently supported in all analyses (PP = 1, BP = 64 % in 97 loci matrix, PP = 1, BP = 100 % in 596 and 1014 matrix analyses), corroborating results from previous morphological [76, 77] and molecular analyses [4, 15, 34, 38, 78]. The phylogenies resulting from analysis of the 97 locus dataset resolve Centropomidae and Sphyraenidae as a clade, which is the sister lineage of all other carangimorphs (PP = 1, BP = 84 %), while the phylogenies inferred using the 596 and 1014 locus datasets place sphyraenids (PP = 0.61, BP = 40 %) and centropomids (PP = 0.57, BP = 44 %), respectively, as successive outgroups to all remaining carangimorphs. Neither of these alternative arrangements is well supported, indicating a degree of phylogenetic uncertainty regarding the earliest divergences within Carangimorpha even in the face of extensive genetic sampling.

The GT-ST results are congruent with nearly all major aspects of the phylogenies inferred from concatenated data (Additional file 1: Figures S1 and S2). The placement of polyenemids represents the most conspicuous difference between phylogenies inferred using concatenated and GT-ST methods. The ASTRAL-inferred GT-ST resolves threadfins as the sister lineage of the clade including toxtotids, leptobramids, menids, xiphoids, and carangiforms, rather than pleuronectiforms as in concatenated analyses, while BUCKy-inferred population tree places Polyenemidae as the sister of flatfishes. We note here, however, that the concordance factors for conflicting placement of Polyenemidae have overlapping 95 % credible intervals. Although GT-ST analyses provide strong support for the monophyly of Carangimorpha and Pleuronectiformes, deep divergences among major carangimorph lineages are not well supported.

A timescale for the diversification of Carangimorpha

Our multiple relaxed-clock molecular dating analyses produced similar and consistent estimates of divergence dates across different subsamples of UCE loci and between both analytical frameworks in the BEAST and PAML software packages, with overlapping 95 % highest
posterior density ranges for all nodes (Fig. 2, Additional file 1: Figure S4, Additional file 1: Table S3). We refer to specific branch length and divergence date estimates derived from analysis of one set of 75 UCE loci (depicted in Fig. 2); but our BEAST analyses of three other 75 UCE-loci-datasets and PAML analysis of 596 loci provide very similar node ages, branch lengths, and overlapping 95 % highest posterior densities. We infer an origin of Carangimorpha at 71.0 Ma, shortly before the end of the Late Cretaceous, although we cannot reject an earliest Paleocene (Danian) origin [95 % highest posterior density (HPD) interval: 63.44, 80.86]. Most of the strikingly anatomically divergent carangimorph lineages arose in the subsequent 15 million years, before the end of the Paleocene. This is consistent with previous molecular phylogenetic studies [3, 4] (but see Santini and Carnevale [79]) as well as the earliest fossil occurrences of specialized lineages like billfishes, jacks, flatfishes, menids, and barracudas by the end of the first stage of the Eocene [29, 80].

We estimate the pleuronectiform crown node at 61.3 Ma [95 % HPD: 54.3–69.5 Ma]. The most recent common ancestor of flatfishes and threadfins, their living symmetrical sister lineage, is 65.7 Ma [95 % HPD: 57.3–72.6 Ma]. The mean length of the flatfish stem across our posterior sample of trees is 2.97 Myr [median: 2.45 Myr; 95 % HPD: 0.47–8.35 Myr], and it provides the longest possible span over which flatfish asymmetry could have evolved. Maximum timelines for the origin of other remarkable carangimorph bodyplans are generally longer. The divergence between Menidae and Xiphioidae is estimated as 60.6 Ma [median: 95 % HPD: 56.5–66.6 Ma]. Early fossil representatives of both lineages constrain the origin of their unique bodyplans to no later than the earliest Eocene [80, 81]. Thus the mean maximum time for origin of both bodyplans is 5.89 Myr [median: 5.42 Myr; 95 % HPD: 1.91–12.48 Myr]. The difference in age between crown Echeneoidei (marking the divergence of remoras from other echeneoids) and the oldest anatomically modern remora fossils give a maximum time over which the remora adhesion disc evolved. Fossils with discs closely resembling those of modern species are known from the early Oligocene [17, 82, 83], giving a maximum evolutionary interval of 9.69 Myr [median: 9.27 Myr; 95 % HPD: 3.87–17.68 Myr]. Our ability to constrain the time over which other distinctive carangimorph bodyplans—like those of polynemids and toxotids—arose is limited by sparse fossil records [80] and sampling of modern lineages that is insufficient to constrain crown ages.

**Discussion**

**The flatfish monophyly challenge**

The recent series of studies that report conflicting interpretations of carangimorph phylogenies inferred from multi-locus sequence data and their apparent equivocal support for the monophyly (or polyphyly) of flatfishes have breathed new life into old debates about the evolutionary origins of their asymmetrical bodyplan [31–33, 84]. Our results generated from more than 1000 UCE loci provide strong molecular support for the monophyly of flatfishes and the single origin of cranial asymmetry, regardless of analytical framework applied (Fig. 2, Additional file 1: Figures S1-S3). Analysis of concordance among loci provides further insight into the longstanding difficulty of resolving relationships among carangimorph lineages using smaller molecular datasets that may not be apparent from nodal support values generated by concatenation or species tree analyses. Concordance factors calculated from our dataset estimate that a genome-wide proportion of 14.9 % of loci produce gene trees that reflect a single flatfish clade. In contrast, the estimated frequency for alternative topologies in which *Psettodes* is recovered in clades with non-flatfishes is significantly lower and non-overlapping with a monophyletic flatfish scenario (Fig. 3). Prior to this study, the largest dataset of multiple unlinked loci applied to carangimorphs [34] found weak support from concatenated analysis of 23 loci for a flatfish clade (PP = 0.65), and attributed the low nodal support to nucleotide compositional bias in protein coding genes and to a lesser extent, ILS. The proportion of loci examined by [34] that recover a flatfish clade (3 out of 23 loci, or 13 %) is within the 95 % credible interval estimated from our analyses of UCE loci (95 % CI of 11.4–18.7 %). While other phenomena (such as the non-stationarity nucleotide composition as identified by [34]) may introduce error into gene tree estimation, the short internal branches subtending successive divergences of carangimorph lineages likely resulted in substantial ILS, making their relationships difficult to recover with small datasets.

The strong support for flatfish monophyly obtained from our molecular analyses of UCE loci bolsters the morphological consensus that this remarkable innovation evolved only once. The anatomical argument for flatfish monophyly has been caricatured as reliant almost exclusively on cranial asymmetry, and thus a hypothesis formed from limited evidence. However, probable pleuronectiform synapomorphies have been identified across multiple anatomical systems, including the axial skeleton [28, 71], caudal-fin endoskeleton [28, 71], otoliths [85], patterns of epaxial muscle insertions [86], and innervations of the trunk lateral line [87]; (Fig. 4). These synapomorphies collectively represent a strong anatomical case for flatfish monophyly, independent of cranial asymmetry.

**Morphological support for relationships within carangimorpha**

Although our phylogeny agrees with many aspects of previous morphological classifications (e.g., the monophyly of xiphioids [7, 75], centropomids [77], pleuronectiforms...
Patterns of lineage diversification within carangimorphs suggest that unambiguous morphological support for some clades may prove elusive. Our molecular clock analyses indicate many specialized carangimorph groups have independent evolutionary histories that are considerably longer than those shared uniquely with their immediate sister taxa [3, 4, 38]. As a result, there was little time over which traits providing evidence for sister-group
relationships could evolve relative to the time that any synapomorphies might be overwritten by the profound morphological specializations characteristic of individual carangimorph lineages. Nevertheless, it is clear that considerable anatomical evidence for monophyly has accumulated along some of the shortest branches within the carangimorph phylogeny. This is particularly apparent for the ca. 3 Myr-long flatfish stem (discussed above).

More careful anatomical scrutiny may yield evidence for the phylogenetic relationships in the UCE-inferred trees (Fig. 2; Additional file 1: Figures S1-S3), especially because many of the sister-group pairings within Carangimorpha have never been seriously considered—and therefore investigated—from a morphological perspective. An anatocially unanticipated sister-group relationship between billfishes and the moonfish Mene is a common motif of carangimorph molecular phylogenies, and one for which we recovered strong support in the UCE-inferred trees (e.g., Fig. 2). The anatomical specializations of these groups result in strikingly different “baupläne” that may have discouraged close comparison in the past, although we note a series of derived features common to both lineages (for details of comparative materials, see Additional file 1: Table S2): considerable elongation of the second and third pelvic-fin rays (pelvic fins are lost in xiphiids), caudal hypurostegy, a consolidated hypural plate arising from the fusion of hypurals 1-4 to one another and the ural centrum (further fusion characterizes istiophorids), and posterior extensions of the gas bladder [88–90]. Similarly, the sister-group relationship between Leptobrama and toxotids was previously unpredicted on the basis of the morphology, but the two groups share a number of unusual features among carangimorphs: presence of endopterygoid teeth (uniquely), presence of ectopterygoid teeth (with Polynemidae), and presence of more anal-fin rays than dorsal-fin rays (with Polynemidae and Lactariidae [not analysed in current study]) [90, 91]. These observations are joined by sporadic reports of shared specializations in other carangimorph lineages such as similar larval colour patterns in billfishes and sphyraenids [92]; preanal canal units in polynemids, toxotids, and carangiforms, with ossified prenasals in the latter two [72]; and a series of specializations common to Mene, Lactarius and many carangiforms [91]. A systematic anatomical survey is required to determine whether the shared morphological similarities noted in this and previous studies corroborate the novel and strongly supported relationships in the carangimorph molecular phylogeny.

The origin of flatfish asymmetry: gradual but rapid
Our time-calibrated phylogeny provides the first robust constraints for the timescale over which the flatfish transformation occurred. On average, we conclude that complete orbital migration arose in no more than 2.97 Myr, although we cannot reject the possibility that it may have taken less than 470 kyr or as long as 7.96 Myr based on our posterior sample of trees. All stem pleuronectiforms identified to date show incomplete orbital migration [28, 71]. However, the identification of extinct flatfishes showing complete asymmetry but which branch outside the crown would further reduce the length of the stem over which modern pleuronectiform cranial geometry arose, thereby increasing the rate of this evolutionary transformation. Comparisons both within and outside Carangimorpha provide context for the rapid origin of the flatfish bodyplan. Mean maximum timescales for the evolution of billfishes, moonfishes, and remoras are on the order of 5–10 Myr, two to three times longer than comparable estimates for the origin of flatfish asymmetry. Similarly, the timescale of a few million years for the origin of flatfish asymmetry compares closely with the estimated age of some cichlid radiations in African rift lakes [93], upheld as examples of explosive evolutionary diversification [94], although flatfishes are arguably a more extreme morphological departure from standard acanthuriform bodyplans than even the most peculiar modern cichlids [95]. Outside of fishes, the timescale for the evolution of pleuronectiform asymmetry is substantially shorter than those estimated for the origins of the bodyplans associated with whales (ca. 20 Myr [96]) and anatomically modern humans (ca. 7 Myr). We hypothesize that the rapid evolution of flatfish asymmetry might reflect a steep peak associated with a new adaptive zone, as classically hypothesized for other rapid divergences by Simpson [97].

Diversification of carangimorpha
Our well-supported phylogeny provides new insight for exploring the origin of the anatomically diverse lineages that comprise Carangimorpha. The earliest diverging carangimorph lineage includes what are arguably the most anatomically ‘generalized’ members of the radiation: the superficially perch-like snooks and Nile perches. Interestingly, the two lineages constituting the principal carangimorph clade, excluding centropomids and sphyraenids, are differentiated along ecological lines. The first unnamed lineage is broadly associated with benthic environments, and its two constituent clades show striking adaptations to life at or near the sediment-water interface: profound asymmetry in flatfishes and free pectoral-fin rays in polynemids that serve a tactile function [19]. By contrast, the second major lineage contains pelagic groups like jacks, billfishes, and dolphinfishes. Our discovery of broad environmental divisions within carangimorphs represents the latest example of molecular analyses revealing novel clades of percomorph teleosts that share a common habitat preference or
ecology [74]. The major ecological split in carangimorph phylogeny mirrors patterns apparent at smaller spatial and temporal scales in fishes, in which initial divergences within populations often reflect partitioning between benthic and pelagic habitats and resources [98].

Our analyses suggest that anatomically modern body plans evolved by the early Eocene (ca. 50 Ma), following major diversifications in the latest Late Cretaceous and Paleocene. The rich fossil records of many individual groups [29], combined with our well-supported time-calibrated genomic perspective on phylogenetic relationships, make carangimorphs an ideal system for studying patterns of phenotypic diversification over spatial and temporal scales not reflected by young, geographically restricted clades that are often the focus of research on adaptive radiation [94, 99, 100]. The study of ancient marine percomorph radiations like the ecologically varied carangimorphs, pelagic scombriforms [74], and Antarctic notothenioidei [101] may provide insights into the generation of biological diversity that is persistent over long evolutionary timescales not provided by model systems like sticklebacks and cichlids that have diversified in spatially limited and geologically ephemeral environments [99, 102].

Conclusion
The invariably low support for monophyly of flatfishes found in previous molecular phylogenetic analyses is emblematic of a common problem in reconstructing the evolutionary history of rapidly diverging lineages throughout the Tree of Life. Although increasingly larger gene-by-gene datasets have provided valuable discoveries regarding the relationships among acanthomorph fishes and the timing of their divergences [3, 4], these datasets may not be large enough to overcome discordance due to phenomena such as incomplete lineage sorting for nodes within rapidly branching portions of the acanthomorph phylogeny. The resolution of carangimorph relationships provided by high throughput sequencing of UCE loci serves as an improved framework on which to study the evolution and diversification of fish bodyplans, and our results suggest similar phylogenomic approaches will be necessary to resolve historically difficult nodes in the acanthomorph phylogeny, as in the case of the flatfishes.

Acknowledgements
We thank A. Bentley and P. Chakrabarty for tissue samples, and J. Maclaine and O. Crimmen for access to preserved specimens. Portions of this research were conducted with high performance computational resources provided by Louisiana State University (http://www.hpc.lsu.edu) and the University of Oxford (http://www.arc.ox.ac.uk), and this study was also supported, in part, by resources and technical expertise from the Georgia Advanced Computing Resource Center, a partnership between the University of Georgia’s Office of the Vice President for Research and Office of the Vice President for Information Technology.

Funding
This research was supported by NERC grant no. NE/K026332/1 and a Philip Leverhulme Prize PLP-2012-130 (both to M.F.). NSF DEB-1742260 (to BCF) supported computational portions of this work and NSF ANT-0839007 (to TJN) supported portions of laboratory expenses.

Availability of data and material
Sequence data generated for this manuscript are archived as raw reads in the NCBI Sequence Repository (SRA) and assembled contigs in NCBI GenBank (accession numbers SAMN05784507, SAMN05786321-SAMN05785372). The data sets supporting the results of this article, including alignments of UCE loci and phylogenetic trees generated from our analyses of these data, are available in the Dryad Digital Repository (doi:10.5061/dryad.2655) [103]. Details of fossil calibrations are provided in Additional file 2.

Authors’ contributions
MF and RCH conceived the research. RCH and BCF generated sequence data. Phylogenetic analyses were performed by BCF, MA, and RCH. All authors contributed to interpretation of results and writing of the manuscript. All authors read and approved the final manuscript.

Competing interests
The authors declare that they have no competing interests.

Consent for publication
Not applicable.

Ethics approval and consent to participate
Not applicable.

Author details
1Department of Earth Sciences, University of Oxford, Oxford OX1 3AN, UK. 2Department of Ecology & Evolutionary Biology and Peabody Museum of Natural History, Yale University, New Haven, CT 06520, USA. 3Department of Biological Sciences and Museum of Natural History, Purdue University, Lafayette, IN 47907, USA. 4Department of Ecology and Evolutionary Biology, University of Kansas, Lawrence, KS 66045, USA. 5Department of Marine Biology, Texas A&M University at Galveston, Galveston, TX 77553, USA. 6Department of Geology and Biodiversity Institute and Department of Ecology and Evolutionary Biology, University of California, Berkeley, CA 94720, USA. 7Department of Paleontology and Department of Earth and Environmental Science, University of Michigan, 1109 Geddes Ave, Ann Arbor, MI 48109-1079, USA.

Received: 12 July 2016 Accepted: 30 September 2016
Published online: 21 October 2016

References
1. Smith WL, Wheeler WC. Venom evolution widespread in fishes: A phylogenetic road map for the bioprospecting of piscine venoms. J Hered. 2006;97(3):206–17.
2. Li B, Dettai A, Crusaf C, Louxou A, Desoutter-Meniger M, Lecointre G. RNF213, a new nuclear marker for acanthomorph phylogeny. Mol Phylogen Evol. 2009;50(2):345–63.
3. Near TJ, Dornburg A, Eytan RI, Keck BP, Smith WL, Kuhn KL, Moore JA, Price SA, Buebrink FT, Friedman M, et al. Phylogeny and tempo of diversification in the superradiation of spiny-rayed fishes. Proc Natl Acad Sci U S A. 2013;110(31):12738–43.
4. Betancur-R R, Broughton RE, Wiley EO, Carpenter K, Lopez JA, Li C, Holcroft NI, Arcila D, Sanciangco M, Cureton II JC, et al. The tree of life and a new classification of bony fishes. PLoS Curr. 2013; doi:10.1371/currents.tol.53b5e26640d9ccae75bb165c8c26288.

Additional files

Additional file 1: Extended descriptions of protocols for DNA isolation, library preparation, sequencing, and data pipelining, and Additional file 1: Figures S1–S4. (DOCX 6704 kb)

Additional file 2: Descriptions of fossil calibrations used in divergence dating analyses, including details of age priors and the fossil data upon which they are based. (DOCX 92 kb)
57. Mirarab S, Reaz R, Bayzid MS, Zimmermann T, Swenson MS, Warnow T. ASTRAL: genome-scale coalescent-based species tree estimation. Bioinformatics. 2014;30(17):2641–8.
58. Kubatko LS, Degnan JH. Inconsistency of phylogenetic estimates from concatenated data under coalescence. Syst Biol. 2007;56(1):17–24.
59. Ane C, Larget B, Baum DA, Smith MD, Rokas A. Bayesian estimation of concordance among gene trees. Mol Biol Evol. 2007;24(2):412–26.
60. Larget BR, Kotha SK, Dewey CN, Ane C, BUCKy: gene trees/species tree reconciliation with Bayesian concordance analysis. Bioinformatics. 2010;26(22):2910–1.
61. Degnan JH, DeGiorgio M, Bryant D, Rosenberg NA. Properties of consensus methods for inferring species trees from gene trees. Syst Biol. 2009;58(1):35–54.
62. Ronquist F, Teslenko M, van der Mark P, Ayres DL, Nieves-Aldrey J, Li C, Betancur-R R, Smith WL, Ortí G. Monophyly and interrelationships of perciform fishes long after Gondwanan rifting. Proc R Soc Lond [Biol]. 2013;280(1761):20122676.
63. Friedman M, Johanson Z, Harrington RC, Near TJ, Graham MR. On fossils, phylogenies and sequences of evolutionary change. Proc R Soc Lond [Biol]. 2014;281:20140115.
64. Saitou N, Nei M. The Neighbor-Joining method: a new method for reconstructing phylogenetic trees. Mol Biol Evol. 1987;4(4):406–25.
65. Friedman M, Johnson GD. A new species of Mene (Perciformes: Menidae) from the Paleocene of South America, with notes on the paleoenvironment and a brief review of mene fishes. J Vert Paleontol. 2005;25:770–83.