Molecular Phylogenetics and Evolution 69 (2013) 514–523

Contents lists available at ScienceDirect

Molecular Phylogenetics and Evolution

journal homepage: www.elsevier.com/locate/ympev

Genome duplication and multiple evolutionary origins of complex migratory behavior in Salmonidae

Markos A. Alexandrou [a], Brian A. Swartz [b,c], Nicholas J. Matzke [b], Todd H. Oakley [a,*]

[a] Department of Ecology, Evolution, and Marine Biology, University of California, Santa Barbara, CA 93106, USA
[b] Department of Integrative Biology and Museum of Paleontology, University of California, Berkeley, CA 94720, USA
[c] Millennium Alliance for Humanity and the Biosphere, Stanford University, Stanford, CA 94305, USA

ARTICLE INFO

Article history:
Received 23 April 2013
Revised 21 July 2013
Accepted 26 July 2013
Available online 7 August 2013

Keywords:
Evolution
Ancestral states
Phylogenetics
Time trees
Anadromy
Genome duplication
Salmon

ABSTRACT

Multiple rounds of whole genome duplication have repeatedly marked the evolution of vertebrates, and correlate strongly with morphological innovation. However, less is known about the behavioral, physiological and ecological consequences of genome duplication, and whether these events coincide with major transitions in vertebrate complexity. The complex behavior of anadromy – where adult fishes migrate up rivers from the sea to their natal site to spawn – is well known in salmonid fishes. Some hypotheses suggest that migratory behavior evolved as a consequence of an ancestral genome duplication event, which permitted salinity tolerance and osmoregulatory plasticity. Here we test whether anadromy evolved multiple times within salmonids, and whether genome duplication coincided with the evolution of anadromy. We present a method that uses ancestral character simulation data to plot the frequency of character transitions over a time calibrated phylogenetic tree to provide estimates of the absolute timing of character state transitions. Furthermore, we incorporate extinct and extant taxa to improve on previous estimates of divergence times. We present the first phylogenetic evidence indicating that anadromy evolved at least twice from freshwater salmonid ancestors. Results suggest that genome duplication did not coincide in time with changes in migratory behavior, but preceded a transition to anadromy by 55–50 million years. Our study represents the first attempt to estimate the absolute timing of a complex behavioral trait in relation to a genome duplication event.

1. Introduction

Multiple rounds of whole genome duplication have characterized the evolution of vertebrates (Jaillon et al., 2004). Specifically, in actinopterygians (ray-finned fishes), genome duplication events correlate with bursts of morphological change in extant lineages (Ohno, 1970; Taylor et al., 2001, 2003; Van de Peer et al., 2003, 2009; Braasch et al., 2009a,b) but see (Heimberg et al., 2008). However, incorporating fossil taxa into these analyses demonstrates that morphological complexity accumulated more continuously through time (Donoghue and Purnell, 2005). Whether episodic or periodic, little is known about the behavioral consequences of genome duplication in vertebrates, and if duplication events coincide with transitions in behavior and physiology (Novak et al., 2006).

The ecophysiological consequences of genome duplication remain understudied (Mable, 2004; Soltis et al., 2010; Mable et al., 2011), but evidence from plants suggests that polyploids are able to inhabit a greater diversity of environments compared to their diploid progenitors (Otto and Whitton, 2000). Generally speaking, duplication events trigger morphological and physiological changes in size and ecological tolerance (Adams and Wendel, 2005; Hu et al., 2012). For example, recent work suggests that newly formed polyploid populations are more resistant to parasites (Oswald and Nuismer, 2007). Furthermore, polyploidy affects cellular physiology by altering cell size and enhancing photosynthetic capacity (Warner and Edwards, 1993); it may even change cellular metabolic rates (Comai, 2005). Less is known about the ecophysioloical consequences of genome duplication in vertebrates. In one notable example, mating calls in gray tree frogs follow directly from changes in ploidy, which leads to reproductive isolation via female mate choice (Mable et al., 2011). The link between genome duplication and the evolution of complex traits has also been hypothesized for salmonid fishes (Ohno, 1970).

Salmonids (salmon, trout, char, whitefishes, and graylings) are one of the most well studied vertebrate groups that have undergone genomic duplication. Their migratory behaviors require major physiological plasticity, which are known popularly in Atlantic and Pacific salmon. These fishes are capable of overcoming tremendous obstacles to return from distant oceans to natal streams for spawning. This behavior has inspired centuries of research into the evolution of salmon migrations and the origin(s)

* Corresponding author. Fax: +1 805 893 4714.
E-mail address: oakley@lifesci.ucsb.edu (T.H. Oakley).

1055-7903/$ - see front matter © 2013 Elsevier Inc. All rights reserved.
http://dx.doi.org/10.1016/j.ympev.2013.07.026

of this life history strategy (McDowall, 2002). Molecular and morphological evidence places the freshwater Esocidae (pikes) as the closest relatives of salmonids, which suggests that salmonids arose in freshwater environments (Stearley, 1992; Ishiguro et al., 2003; Ramsden et al., 2003; Nelson, 2006; Broughton, 2010; Wilson and Williams, 2010). Furthermore, the Eocene salmonid †*Eosalmo driftwoodensis* is preserved in all life stages within freshwater deposits, offering strong support for non-migratory behavior in stem-salmonids as well. In addition to complex migratory behavior, all salmonids (estimated to have originated 63.2–58.1 MYA. (Crete-Lafreniere et al., 2012)) share a tetraploid genome that arose between 100 and 25 My. (Johnson et al., 1987; Sakamoto et al., 2000; Angers et al., 2002).

Beginning in the 1970s (Ohno, 1970), authors speculated that the salmonid duplication event may have provided the genetic material necessary for the evolution of migratory behavior (Allendorf and Thorgaard, 1984; Phillips and Rab, 2001; Koop et al., 2008; Norman et al., 2011, 2012). The hypothesis suggests that duplicate genes (resulting from whole genome duplication) could provide flexibility for differential optimization, as salmonids transitioned to a multi-modal life history. Recent evidence suggests that copies of the FXYD isotypes (Tipsmark, 2008; Tipsmark et al., 2008a,b,c), ATP1a1b loci (Norman et al., 2011) and other genes involved in osmoregulation (Gharbi et al., 2004, 2005; Woram et al., 2004; Danzmann et al., 2005; Norman et al., 2012) may have gained new functions following duplication events, and permitted an increased tolerance to salinity. A comprehensive analysis of salmon and pike transcriptomes has revealed an asymmetrical relaxation of selection on paralogous salmonid genes, allowing individual paralogs to evolve at faster rates (Leong et al., 2010). Many paralogs were lost, but remaining paralogs began to diverge (Leong et al., 2010). This divergence may have been a crucial factor leading to complex migratory behavior and the resulting access to marine niches with greater resources and more space. However, the absolute timing of genomic duplication in relation to the evolution of anadromy remains unknown.

Despite investigations into the consequences of genome duplication and the origin(s) of migratory behavior in salmonids, at least two important questions remain unanswered: (1) did complex migratory behavior evolve multiple times in salmonids? (2) Does the evolution of salmonid migratory behavior coincide with genome duplication? These are both phylogenetic questions. Thus, migratory behavior and ploidy can be treated as traits and studied using a time calibrated phylogenetic tree. However, one problem with this approach is that anadromy is highly variable within the Salmonidae, even at the intraspecific level. Traits with rapid evolutionary rates are difficult to reconstruct because their histories can be quickly lost (Schluter et al., 1997; Cunningham et al., 1998; Oakley and Cunningham, 2000; Oakley et al., 2005). Furthermore, there is only one temporal estimate for genomic duplication in salmonids, but confidence intervals span 75 MYA (Allendorf and Thorgaard, 1984; Crete-Lafreniere et al., 2012; Shedko et al., 2012). Despite improvement in this latter approximation, no study thus far has estimated explicitly the absolute timing of the transition from diploidy to tetraploidy.

To improve upon previous analyses and test long-standing questions about salmonid evolution, we utilized new statistical tools and copious data to evaluate the relationships among phylogeny, migratory behavior, and genome duplication. We reconstructed a phylogenetic tree of salmonids and their close relatives, and analyzed the character matrix using a mixed-model multi-partition dataset (DNA and morphology) under Bayesian and Maximum Likelihood statistical frameworks. Furthermore, we constructed a series of fossil-calibrated trees by concatenating extinct and extant taxa, thereby significantly improving upon previous divergence time estimates. We included a key freshwater stem-salmonin (†*E. driftwoodensis*), which offered unique insight into the transition to migratory behavior. We then mapped the history of migratory behavior and genome duplication onto the resulting phylogeny using Bayesian methods (Pagel et al., 2004) and stochastic character mapping (Bollback, 2006). We estimated the absolute timing of character transitions with a new method that uses ancestral character simulation data to plot the frequency of character transitions per branch on a time-calibrated phylogeny. This method improves on previous approaches because it estimates character transitions along branches instead of only at nodes. These methods allowed consideration of uncertainty in rates of trait evolution, a factor that importantly impacts ancestral state estimations (Schultz et al., 1996). Our analyses improve upon most of the limitations of previous studies and strongly support the parallel evolution of anadromy in salmonid fishes. Results indicate that complex migratory behavior evolved at least twice in salmonids (in salmon and whitefishes), but it is unclear how many times large-scale anadromy evolved, likely due to the rapid evolution of migratory behavior in crown-Salmoninae. Furthermore, we provide new estimates of ancestral tetraploidization that suggest genome duplication significantly preceded the transition to complex migratory behavior in whitefishes.

2. Materials and methods

2.1. Data collection

We constructed a data matrix from 97 species or subspecies-level salmonids and related outgroups. Our data set included 41 data partitions. We collected nucleotide sequence data from GenBank using 16 nuclear and 18 mitochondrial genes (Phillips and Oakley, 1997; Crespi and Fulton, 2004). We also assembled seven categorical data partitions, including morphological data that were compiled from three sources, excluding overlapping data (Smith, 1992; Stearley and Smith, 1993; Wilson and Li, 1999). Besides morphology, other categorical data sets included rDNA RFLPs (Phillips et al., 1992, 1994, 1995), mitochondrial RFLPs (Nielsen et al., 1998; Carrera et al., 1999), chromosome number (Phillips and Rab, 2001), Short Interspersed Nuclear Elements (SINEs) (Takasaki et al., 1997; Hamada et al., 1998), microsatellites (Olsen et al., 2000), and allozymes (Allendorf and Seeb, 2000).

To fossil-calibrate the root of the tree, we extended sampling beyond Salmonidae to include Esociformes (pikes such as *Novumbra hubbsi* and *Esox lucius*), Osmeriformes (smelts such as *Hypomesus olidus*, *Plecoglossus altivelis* and *Argentina sialis*), Gonorynchiformes (milkfishes such as *Chanos chanos*), and Cypriniformes (zebrafishes such as *Danio rerio*).

2.2. Phylogenetic analyses

We excluded data sparse taxa from further analyses. In fact, not all species-level taxa have data for all or even most data partitions: of the $97 \times 41 = 3977$ possible combinations of species taxa *x* data partition, the complete initial data matrix was 80% missing. With the exception of fossil taxa – for which only morphological data are available – we excluded from further analyses species and subspecies-level taxa containing fewer than three data partitions. Many excluded taxa were subspecies for which only chromosome numbers were available. After excluding extremely sparse taxa, the resulting matrix contained 58 taxa and was 26.6% complete as data partitions available per taxon (Supplementary Table 1). Nucleotide data collected from GenBank were aligned using the default parameters of MUSCLE (Edgar, 2004), with the exception of growth hormones and intron c; these were aligned manually (Oakley and Phillips, 1999). Several mitochondrial genomes are available for

salmonid taxa; these were separated into individual genes for alignment before concatenating with the entire dataset.

After obtaining an aligned matrix, we conducted numerous phylogenetic analyses, using Bayesian Inference (BI) and Maximum Likelihood (ML) approaches. We divided genetic data into partitions by gene and categorical data as binary or multistate. Using MrBayes 3.1 (Huelsenbeck and Ronquist, 2001), we assumed GTR + Γ + I for DNA partitions, which was the best-fit of 56 model comparisons implemented in jModeltest (Posada, 2008). For all categorical data, we assumed the Mk model of Lewis (Lewis, 2001). Only branch length parameters were linked across data partitions and we assumed default priors for all parameters. To estimate the posterior probability distribution of phylogenetic trees and rate parameters, we used Metropolis-coupled MCMC with proposal mechanisms described by Nylander et al. (Nylander et al., 2004). We conducted a run of 10,000,000 generations, sampling every 100th generation. We excluded the first 100,000 generations as "burn-in" using Tracer v1.5 (Rambaut and Drummond, 2004), after which the Markov chain was stationary. Using RAxML (Stamatakis, 2006), we did a ML search with a mixed partition model. Random starting trees were used and topological robustness was evaluated using 100 non-parametric bootstrap replicates.

We used r8s v 1.7 (Sanderson, 2003) to construct an ultrametric tree with the best scoring tree from RAxML. To this end, we used a semi-parametric Penalized Likelihood approach with a truncated Newton algorithm and bound constraints. We then used BEAST v 1.7.2 (Drummond et al., 2012) to conduct a fully parametric Bayesian analysis on the same data. With BEAST we used several different tree priors to compare the outcomes, namely the traditional Yule model (Gernhard, 2008), the standard Birth–Death model (Gernhard, 2008), and the Birth–Death model with incomplete sampling (Stadler, 2009). Furthermore, all categorical data were included in the BEAST analyses by coding it into BEAST XML input (following Pyron, 2011) using a custom R function. Extinct species taxa were included as dated, non-contemporaneous OTUs (See Table 1 for tip dates and calibrations) using the tip-dates function in BEAUTI v 1.7.2 (Drummond et al., 2012), with the morphological data encoded BEAST's XML format as in Pyron (2011) using a custom R script. The R script converts standard NEXUS-formatted morphological data (including ambiguities) to BEAST's format which requires a single alphanumeric code for each character state and each combination of ambiguous states (e.g., "?" is encoded as "Z"); the ambiguity codes are user-determined and also generated by the script. The script is available at http://oakley-web.eemb.ucsb.edu/ASMO13/. Uncertainty in the dating of fossil OTUs was accounted for with a uniform distribution covering the published age estimate of the fossil, as done in Wood et al., 2012. All BEAST runs were fully partitioned, with unlinked substitution-rates, linked trees, and clock models. Uncorrelated lognormal relaxed clocks and random starting trees were used for all analyses. Each analysis was run for 30,000,000 generations, sampled every 1000th, and checked for stationarity using Tracer v 1.5 (Rambaut and Drummond, 2004). Consensus trees with mean node values were constructed using TreeAnnotator v 1.7 (Drummond et al., 2012) after having determined the appropriate burn-in.

2.3. Trait mapping

We assigned five states to the character of migratory behavior for mapping onto our phylogenetic trees: state (0) *freshwater* – taxon members spend their entire lives in freshwater; state (1) *amphidromous* – taxon members migrate between freshwater and marine stages, but migrations are not directly coupled to spawning; state (2) *small-scale anadromy* – taxon members are born in freshwater before migrating to estuaries or other coastal environments (upriver migrations are made shortly before spawning); state (3) *large-scale anadromy* – taxon members are born in freshwater before migrating far out to sea, often following ocean currents for well over a year, rapid upriver migrations are then made shortly before spawning; and state (4) *marine* – taxon members are born and live entirely in the sea. These character states reflect our attempt to allow for specific definitions, as well as to arrive at a consensus of previous authors' opinions (see Supplementary Table 1). We are the first to use the terms *small* and *large-scale anadromy*: many previous authors do not discriminate between these categories (McDowall, 2002). By distinguishing them, we are stating our agreement with Stearley (Stearley, 1992) that the long migrations to sea by Atlantic and Pacific salmon, which often follow ocean currents, deserve their own category. However, we disagree with Stearley (Stearley, 1992) that other salmonids are not anadromous: he used the term *amphidromous* for taxa that conduct migrations a short distance to sea. Therefore, our categories of small and large-scale anadromy are consistent with most authors given that they consider many salmonids to be anadromous. Yet we also recognize that large-scale migrations of some salmon represent a different life history strategy. For character state simulations of ancestral tetraploidy (see below), we coded all diploids as 0 (outgroups) and all tetraploid as 1 (salmonids) (Allendorf and Thorgaard, 1984).

2.4. Ancestral state estimation

We used Bayesian statistics for ancestral state reconstructions, implemented in the software Bayes-Multistate (Pagel et al., 2004). A Bayesian approach aims to estimate the distribution of a particular parameter, such as the probability of a character state at a certain node. We chose this approach because it considers uncertainty in both phylogenetic tree topology and in rates of migratory evolution to reconstruct the most likely condition at the base of crown-salmonids.

Table 1
Fossil dates and calibrations.

Fossil	Formation	Reference	Age (MYA)	Calibration prior
Multiple taxa	Obere solnhofener schichten Germany	Arratia (1997)	165.2–149.85	Ln offset = 149.85 (mean 0, stdev 0.5)
†*Oldmanesox sp.*	Oldman formation, dinosaur provincial park, alberta Canada	Wilson et al. (1992)	84–70.6	Ln offset = 84 (mean 1, stdev 1)
†*Eosalmo driftwoodensis*	Driftwood creek formation, allenby formation, klondike mountain formation Canada	Wilson and Li (1999)	52–51.5	Tip date mean 51.75 uniform distribution 51.5–52
†*Salvelinus larsoni*	Lower chalk hills formation, idaho USA	Smith (1981) and Smith (1992)	9.8–7.4	Tip date mean 8.6 uniform distribution 9.8–7.4
†*Oncorhynchus lacustris*	Glenns ferry formation, idaho USA	Smith (1981)	5–2.5	Tip date mean 3.75 uniform distribution 5–2.5
†*Oncorhynchus keta*	Chalk hills formation, malheur co., oregon USA	Smith (1992) and Eiting and Smith (2007)	6.6–4.8	Ln offset = 6.6 (mean 1, stdev 1)

Fossil data used to calibrate the r8s and BEAST trees showing taxa, formations, references, ages, and parameters used for calibration.

2.5. Absolute timing of character transitions

To estimate the absolute timing of character state changes, we performed stochastic character mapping (SCM) (Nielsen, 2002; Huelsenbeck et al., 2003) on fossil-calibrated time trees. We abbreviate this approach as time tree stochastic character mapping (ttSCM). We performed ttSCM for two characters: (1) ploidy (states = tetraploid and diploid); and (2) migratory behavior (states described above under "trait mapping"). We used 150 time calibrated trees pulled from posterior distributions of MCMC chains that assumed a Yule model tree prior (details described above). For ploidy levels (ancestral tetraploidization) we used all available taxa. However, for migratory behavior we analyzed salmonids separately from their outgroups, since: (1) outgroup migratory traits are irrelevant to the question of whether anadromy evolved independently within salmonids; (2) outgroups are undersampled; and (3) outgroups are represented by additional character states that do not occur in salmonids. On each tree, we conducted 100 SCM simulations for each character. We chose the best-fit model for each character using standard Likelihood Ratio Tests to compare nested models (e.g. symmetric vs. asymmetric rate matrices) (Mooers and Schluter, 1999). We then performed SCM using make.simmap of the phytools v0.2-90 R package (Revell, 2012). Make.simmap, like simmap itself (Bollback, 2006), generates simulated character histories with the constraint that tip character states match observed tip states. Make.simmap first uses ML to determine parameter values for the model of character evolution. Next, make.simmap generates samples from a posterior distribution conditional on the ML-fitted model to reconstruct character histories on trees. This ML parameter fitting approach is different from simmap (Bollback, 2006), which samples from a joint posterior of character histories plus substitution rates. The make.simmap approach is similar in spirit to the empirical Bayes approach for ASR introduced by Pagel (Pagel et al., 2004), in which character rate models are set to ML estimates. We wrote custom R-scripts to plot histograms of the number of state changes in different time bins. In addition to estimating the timing of migratory transitions across the whole tree, we used ttSCM to demonstrate its capability to estimate character transitions on branches of interest. For this analysis, we established support for the homology of whitefish anadromy and the non-homology of whitefish and salmon anadromy. We then coded for all taxa a binary character for whitefish anadromy, with an anadromous whitefish state (1), and a state (0) for all other taxa. By doing so, we were able to estimate the absolute difference in timing between tetraploidy and a transition to anadromy (observed within the Coregoninae).

3. Results

3.1. Data collection

Taxa with the most available data included *Oncorhynchus mykiss* (data available for every partition) and *Salmo salar* (data available for 88% of partitions). The most complete data partitions included morphological data (available for 95% of taxa), chromosomal data (two characters, available for 69% of taxa), and ITS-1 (59% of taxa). The final nucleotide alignment consisted of 32,029 bps, which was concatenated with 275 categorical characters over 58 taxa.

3.2. Phylogenetic results

Bayesian and Maximum Likelihood analyses produced phylogenetic trees with strong support for most nodes (Fig. 1). The results of our supermatrix approach have implications for resolving the tree of life, because we obtained strong phylogenetic signal from a data matrix that was only 26.6% complete in terms of data partitions available. This is consistent with other studies (Sanderson et al., 2011). The placement of three freshwater salmoniform (Fig. 1) clades is particularly relevant to the multiple origins of anadromy. The nearest salmonid outgroup is a largely freshwater clade containing *Esox* and *Novumbra*. This topology is supported by other studies based on mitochondrial genome data (Ishiguro et al., 2003), and other molecular (Lopez et al., 2004) and morphological characters (Williams, 1987; Wilson and Williams, 2010). However, the position of the second freshwater clade is more surprising. The freshwater Thymallinae (graylings) emerged as sister to all other salmonids, with a posterior probability of 1.0. The Coregoninae are usually considered to be sister to other salmonids, including thymallins (Stearley and Smith, 1993). The third freshwater lineage that impacts the freshwater ancestry hypothesis is †*E. driftwoodensis*. This taxon is abundant in the Eocene Green Lake formation, which preserves all its life history stages. The stem-position of †*Eosalmo* in our tree as sister to crown-group Salmoninae is consistent with previous phylogenetic analyses (Wilson and Li, 1999). Our results, therefore, indicate multiple transitions in salmonins and coregonins from a freshwater ancestry to complex migratory behavior.

By combining molecular and morphological characters, we present the first fossil-calibrated phylogenetic framework for salmonid diversification. Including fossils in the time-calibrated tree permitted us to incorporate geological ages into the analysis, instead of simply using prior probability distributions on the dates of nodes. The tip-dating method allows more accurate placement of temporal constraints, thereby yielding more reliable estimates of divergence times (Pyron, 2011). The disadvantages of using fossil taxa to constrain nodes instead of tip dates are apparent in the case of †*Eosalmo*. In contrast to previous analyses that used †*Eosalmo* to constrain the root of the Salmonidae (Shedko et al., 2012), we placed †*Eosalmo* within our phylogeny and applied the constraint directly to the †*Eosalmo* branch. Another recent approach used †*Eosalmo* to constrain Coregoninae + Salmoninae (Crete-Lafreniere et al., 2012), which led to a more recent divergence of crown-Salmonidae. Furthermore, our approach differs from past studies in that we included dates from six fossils, three of which are OTUs that were incorporated directly into the analysis.

Trees recovered using BEAST were nearly identical to those recovered using RAxML and MrBayes (see Figs. 1 and 2), though support values for the BEAST trees were slightly higher for all nodes. One topological difference included the position of the fossil †*Salvelinus larsoni*, which was recovered as sister to the extant *Salvelinus* in BEAST, but which emerged as sister to *Parahucho perryi* in the RAxML and MrBayes analyses. This indicates that the inclusion of dating information in a total-evidence analysis might occasionally change the inference of tree topology, as suggested by Ronquist et al. (2012). This should not be surprising, since more information is being considering in an analysis with dated fossil OTUs relative to an undated analysis. The ages of fossils might constrain their possible positions in a dated tree. The phylogenetic position of †*S. larsoni* remains uncertain because of low support from all utilized methods. Divergence time estimates were similar between the r8s analysis and all BEAST trees, with the latter yielding slightly older dates (Fig. 2). Overall, the three BEAST trees (Yule, birth–death, and birth–death with incomplete sampling) produced nearly identical topologies and dates, with no obvious differences. The root of Salmonidae (using mean values from all separate BEAST analyses; Yule tips, BD tips, BD Inc. tips) was estimated at 69.6 Ma with highest posterior density (HPD) values ranging between 78.3 and 60.8 MYA across all separate speciation models. This represents a vast improvement in confidence compared to previous estimates.

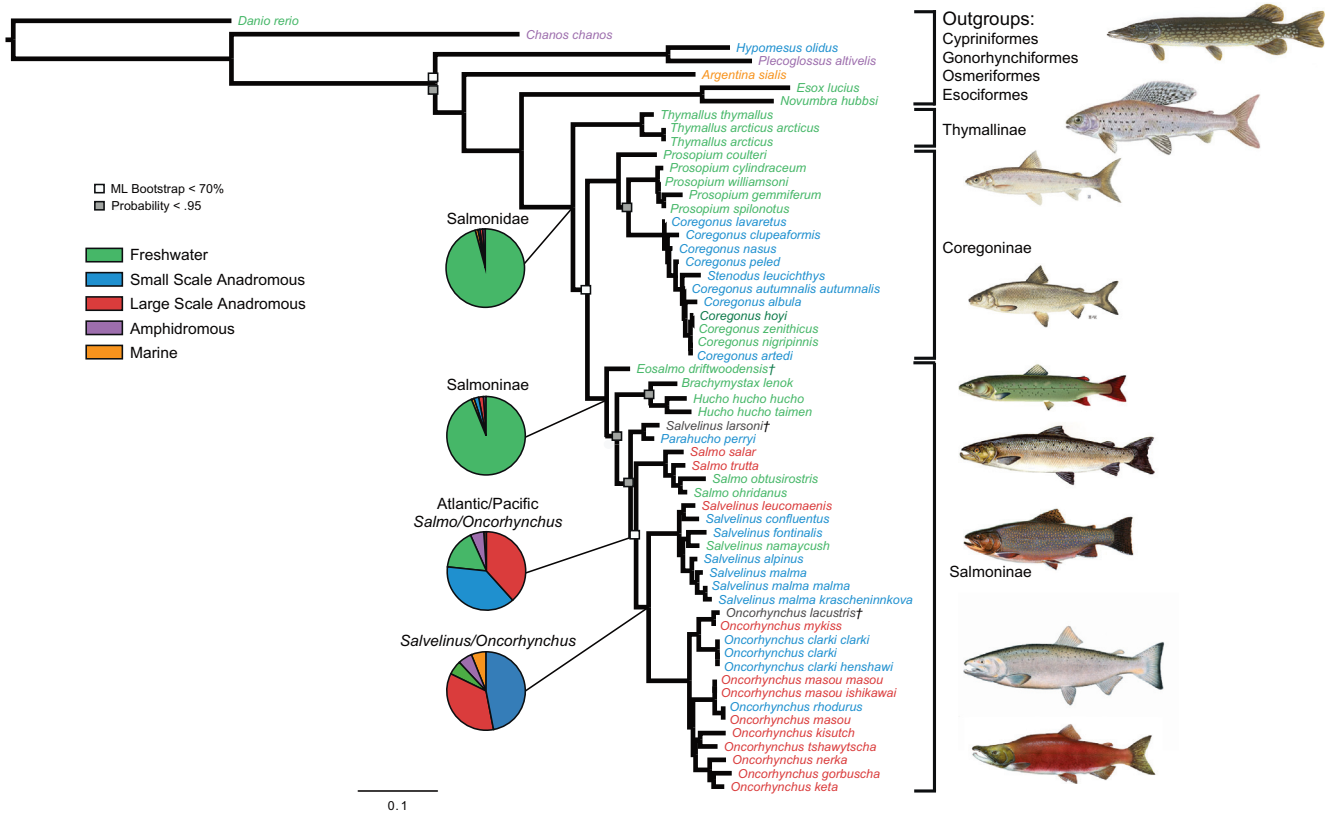

Fig. 1. Phylogeny & trait mapping. A consensus phylogeny from the RAxML analysis showing support values from bootstrap replicates, posterior probabilities, and ancestral states of migratory behavior for major nodes. Each taxon is color-coded according to habitat preference and migratory behavior. Tip labels that are not colored represent taxa that lack data on migratory behavior (both cases are extinct). Fossil taxa are denoted with a cross.

3.3. Ancestral state estimation

With a one-rate model, patterns of trait evolution were similar whether using a flat prior distribution or a prior based on the likelihood of the single rate parameter. In both cases, the ancestral states of Salmonidae and Salmoninae were reconstructed as "freshwater" with posterior probabilities greater than 0.95. Neither analysis could unambiguously reconstruct the migratory behavior of the most recent common ancestor of Atlantic and Pacific salmon (the node of common ancestry between *Salmo* and *Oncorhynchus*). See Fig. 1 for ancestral state estimates of respective nodes.

3.4. Absolute timing of character transitions

The absolute timing of salmonid tetraploidization was estimated using time tree stochastic character mapping (ttSCM). We used an equal rates (ER) model with coequal rates between diploid and tetraploid states.

Assuming an equal rates model, we plotted a histogram of transitions per branch across the tree in 10 million year time bins (Fig. 3A and B). The time bin with the highest number of transitions per branch is 80–70 MYA, with a distribution of bins ranging from 80–50 MYA. This estimate is consistent with a previous estimate of 100–25 MYA (Allendorf and Thorgaard, 1984), and with the results of a recent salmonid phylogeny (Crete-Lafreniere et al., 2012). As expected, it is also slightly older than our estimates for the root of crown-group Salmonidae (78.3–60.8 MYA), which previous approaches (those that constrained character changes to nodes) would perceive as the original timing of tetraploidization. However, because ttSCM permits character changes along branches, it provides older estimates relative to methods that constrain only nodes.

We also estimated the timing of migratory transitions over salmonid evolution. We plotted transitions per branch for three character states: freshwater, small scale anadromy, and large-scale anadromy (Fig. 4). Using a Likelihood-Ratio Test we found that a symmetric model fits migratory patterns better than ER or ARD models (p-value rejecting ARD = 0. 0.2460857, p-value rejecting ER = 1). The highest rate of transitions per branch to freshwater habitats occurred during the early history of salmonids (75–70 MYA); transitions to small-scale anadromy per lineage peaked 30–20 MYA; and transitions to large-scale anadromy peaked 10–5 MYA. Although it is possible to identify maximum time bins, the nature of the data and the Markov models leads to large numbers of character transitions through time, and therefore the time bins across the tree are very long. Early transitions to freshwater and late transitions to large-scale anadromy are consistent with an ancestral freshwater environment, and the independent origins of migratory behavior. Moreover, the intermediate timing of maximal transitions to small-scale anadromy is consistent with the possibility of common, ordered character transitions from freshwater to small-scale anadromy to large-scale anadromy.

We also estimated the timing of the transition from freshwater to small-scale anadromy on the branch leading to Coregoninae (Fig. 3C and D). The maximum time bin was 25–20 MYA, with a distribution spanning the entire length of the branch. Coupled with estimates of the transition to tetraploidy, this result further supports the hypothesis that genome duplication significantly preceded the evolution of complex migratory behavior by 55–50 MYA.

All custom scripts used in this study have been made available for download at http://oakley-web.eemb.ucsb.edu/ASMO13/. The

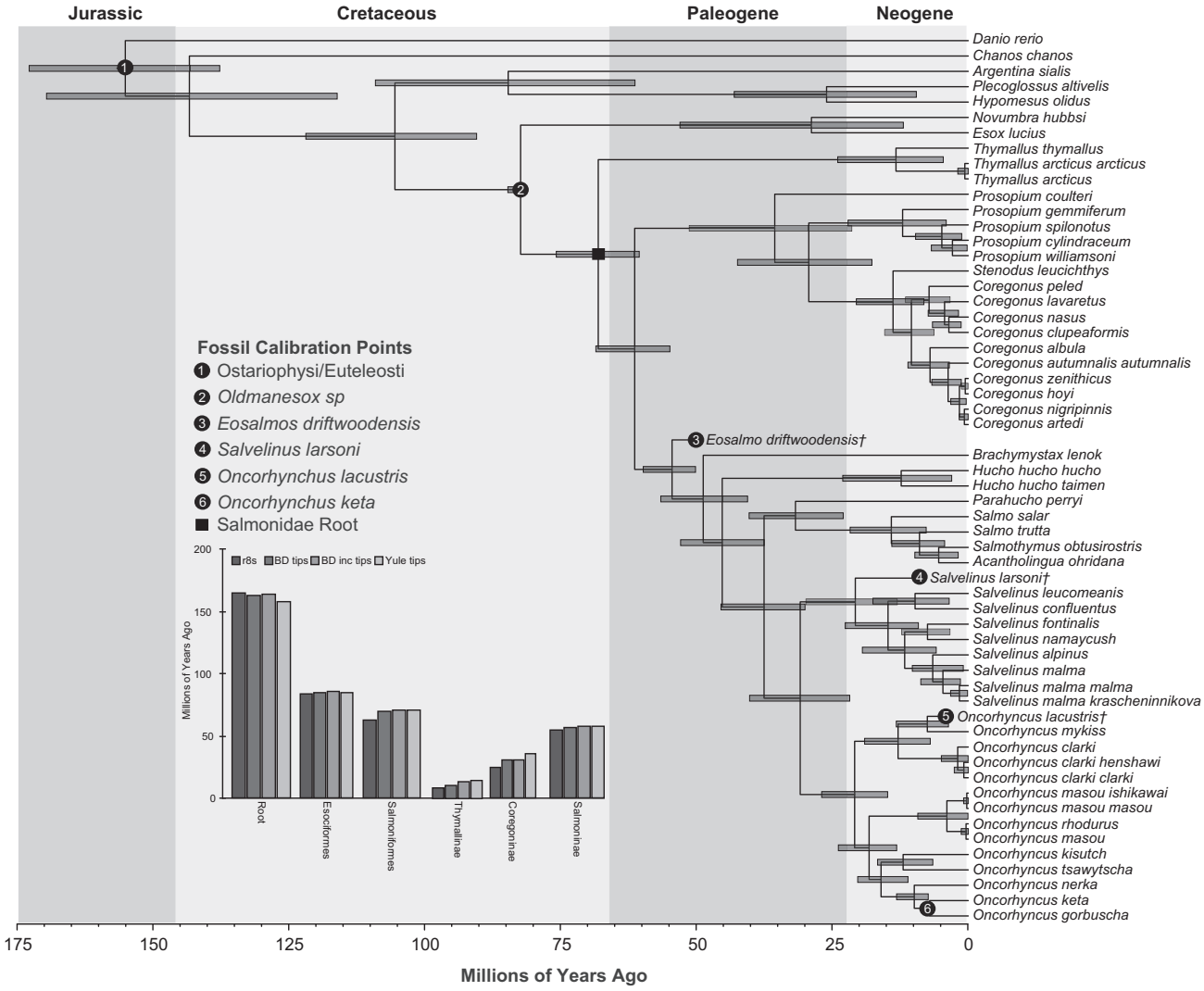

Fig. 2. Fossil-calibrated tree. A fossil-calibrated consensus phylogeny generated using the Yule model of speciation in BEAST. Fossil calibration points are noted on their respective nodes (or tips in the cases of extinct taxa included in the tree). The bars on the nodes represent confidence intervals, otherwise known as highest posterior density (HPD). The graph shows temporal variation in age for the root of the whole tree and the crown of each labeled taxonomic group, using r8s and different models in BEAST.

morphological data matrix has been made available online at Morphobank (Project ID: 608).

4. Discussion

We present the first quantitative test showing that complex migratory behavior evolved at least twice within salmonids (in Coregoninae and Salmoninae), and that genomic duplication in salmonids significantly predated transitions from freshwater to anadromous behavior. We incorporate fossil taxa by concatenating molecular and morphological characters. This approach represents an improvement on previous inferences of salmonid evolution and estimates of divergence times by placing prior calibrations directly on fossil taxa instead of constraining nodes. Furthermore, including life history information (e.g. anadromy) for key fossil taxa (e.g. †*E. driftwoodensis*), allowed us to incorporate fossil characters into a reconstruction of migratory behavior. We estimated the absolute timing of character transitions along branches (as opposed to nodes), and used this method to calculate the timing of genome duplication and the evolution of anadromy. Our results suggest that the duplication event occurred at least 55–50 MYA before a transition to anadromy within Coregonidae, and thereby preceded the origin of salmonid migrations.

4.1. Phylogeny and divergence times

A couple of recent studies used fossil constraints to calibrate the salmonid molecular clock (Allendorf and Thorgaard, 1984; Crete-Lafreniere et al., 2012; Shedko et al., 2012). The topology of the trees recovered in these studies is similar to our own, however, the divergence times differ significantly. Shedko et al. (2012) used two calibration points, one for *Oncorhynchus* (11.6–5.3 MYA) and one for Salmonidae (55.8–33.9 MYA), which led to the more recent divergence estimates. Our results also differ from Crete-Lafreniere et al. (Crete-Lafreniere et al., 2012). Despite the use of the same fossil taxa and a very similar dataset, their divergence estimates were older than the results presented here. The primary source of these differences seems to be the placement of †*Eosalmo*, and the choice of which node to constrain. We overcome this problem by including morphological data from †*Eosalmo* (and other fossils) in order to date them as tips in our tree. We also extended our sampling significantly outside of salmonids to include additional calibration points for salmoniform and euteleostean fishes.

Including fossils in a phylogeny is important because they represent distinct tips in a tree with individual branch lengths (Pyron, 2011). However, most studies use fossil information to constrain nodes among extant taxa. We used a combination of these

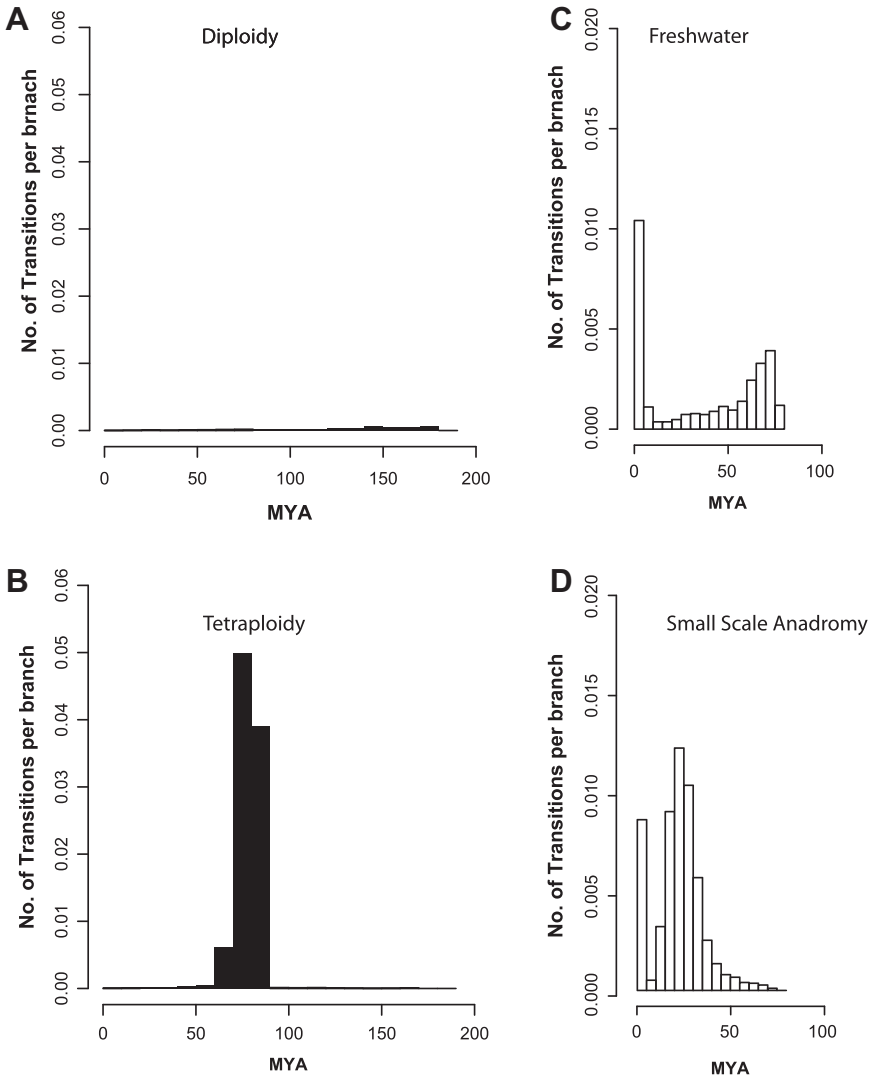

Fig. 3. Migratory character state transitions. Histograms of character simulation results showing transitions to different character states per branch. The axes have been standardized across graphs to show the number of transitions to each state on the *y*-axis and the time after the root on the *x*-axis.

calibration strategies (as in Wood et al., 2012), including prior distributions on the dates of particular nodes representing more complete fossils, and dated fossil taxa as terminal tips. The phylogenetic and ecological histories of †*Eosalmo* (as a stem-Salmoninae that inhabited lacustrine habitats) contribute strongly to the freshwater signal in crown-Salmoninae. Other fossil taxa included in the phylogeny (†*S. larsoni* and †*Oncorynchus lacustris*) are ambiguous in terms of migratory behavior. Nevertheless, using fossil taxa as tips in a phylogeny offers various advantages. Namely, the molecular clock can be calibrated automatically while co-estimating the phylogeny using extinct taxa of known age (Pyron, 2011; Heikkila et al., 2012). This approach sets our calibrated phylogeny apart from the surprisingly few other attempts that estimate divergence times in salmonids (Allendorf and Thorgaard, 1984; Crete-Lafreniere et al., 2012; Shedko et al., 2012).

4.2. Multiple origins of anadromy

The evolution of salmonid migratory behavior has interested biologists for centuries (Pennant, 1776; Fleming, 1828; Günther, 1866; Buckland, 1873; Day, 1887; Meek, 1916; Tchernavin, 1939; Hoar, 1976; Thorpe, 1982; Stearley, 1992; McDowall, 2002). Migration may present advantages such as increased access to resources and tolerance to environmental changes. Considering the benefits, it is not surprising that migratory behavior could have evolved multiple times, resulting in different variations on the same theme. We classify these variations as *large-scale anadromy* and *small-scale anadromy*, and consider them to be distinct behaviors based on the travel distance from spawning grounds to the sea. This distinction between different forms of anadromy leads to a question of whether transitions to anadromy occurred through gradual steps (i.e. from freshwater to small-scale and then large-scale anadromy), or whether the large-scale anadromy evolved directly from a freshwater ancestor. Furthermore, this remains unclear partly due to the homoplasious nature of the character, and to limitations of existing methods used to trace the evolutionary histories of rapidly evolving characters.

4.3. Genome duplication and anadromy

Cytogenetic evidence suggests that tetraploidy likely arose at the root of Salmonidae (Allendorf and Thorgaard, 1984; Phillips and Rab, 2001; Ramsden et al., 2003). Our results concur, and further show that genome duplication significantly preceded at least one origin of complex migratory behavior within salmonids. However, at least two competing hypotheses may explain the time discordance between genome duplication and migration. The first hypothesis suggests that the genome duplication event may have

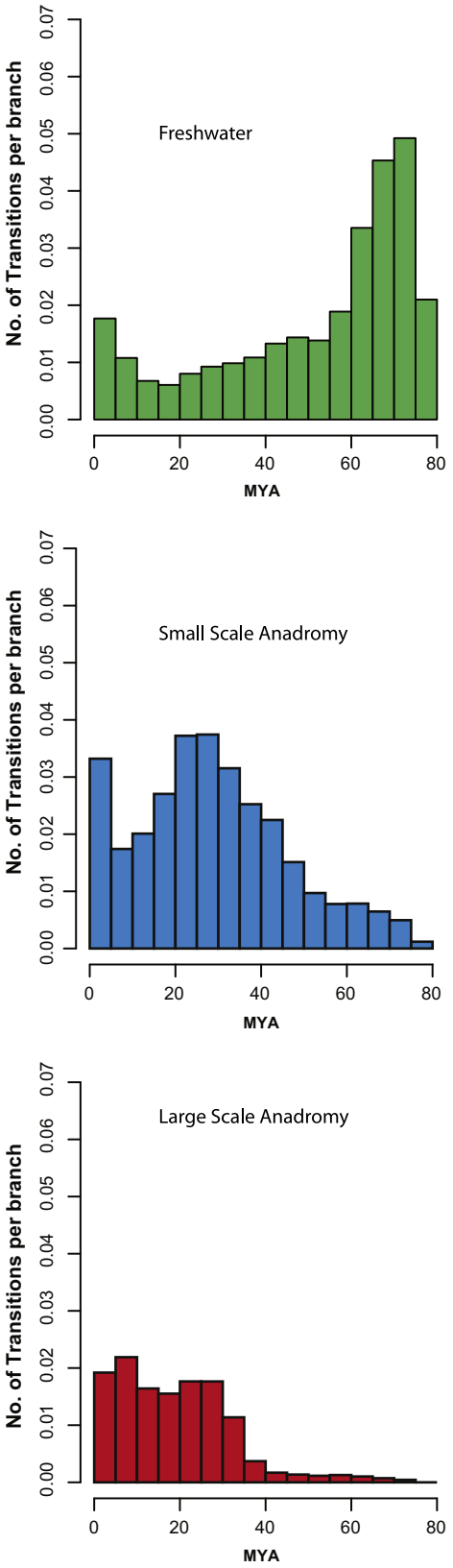

Fig. 4. Ploidy character state transitions. Histograms of character simulation results showing transitions to different character states per branch. The axes have been standardized across graphs to show the number of transitions to each state on the *y*-axis and the time after the root on the *x*-axis. (A) Transitions to diploidy. (B) Transitions to tetraploidy. (C) Transitions to freshwater. (D) Transitions to small-scale anadromy. Panels C and D reference the stem branch of the Coregoninae.

provided the genetic material necessary for the evolution of migratory behavior (Ohno, 1970; Allendorf and Thorgaard, 1984; Phillips and Rab, 2001; Koop et al., 2008; Norman et al., 2011, 2012). Duplicated genes can gain new functions over time and affect complex morphological, behavioral, and physiological traits (Lynch and Conery, 2000; Semon and Wolfe, 2007; Mable et al., 2011). In the case of salmonids, a whole genome duplication event generated a full set of new genes, providing novel gene copies with potential adaptive qualities. Some of these new gene copies could have gained new functions, optimized via selection to allow species to physiologically adapt to different environments. This optimization could have resulted in increased salinity tolerance and plasticity in osmoregulatory functions, allowing salmonids to extent out of their ancestral freshwater habitats into more productive habitats such as estuaries and oceans. Furthermore, the observed lag between genome duplication and transitions to anadromy (55–50 MYA) may have been necessary to allow time for paralogs to diverge significantly and gain new function or sub-functionalize. However, genomic evidence suggests that most duplicated genes are lost or silenced within a few million years after duplication (Lynch and Conery, 2000). Additionally, many genes that are duplicated during whole-genome duplication events become deleterious and cause epigenetic instability via non-additive gene regulation (Comai, 2005).

The second hypothesis proposes that there is no causal relationship between genome duplication and the evolution of migratory behavior. Some evidence suggests that duplicate copies of salmonid genes involved in osmoregulation may have gained new functions and permitted an increased tolerance to salinity (Gharbi et al., 2004, 2005; Woram et al., 2004; Danzmann et al., 2005; Tipsmark, 2008; Tipsmark et al., 2008a,b,c; Norman et al., 2011, 2012). Nonetheless, these duplicated genes may have arisen post genome duplication as independent events that did not result directly from tetraploidy. Additional comparative genomic data will be required to estimate the timing of adaptive gene duplications. Our results provide an estimate of the timing of tetraploidy in salmonids that can be used for such comparisons.

4.4. Time tree stochastic character mapping (ttSCM)

The ability to calculate the absolute timing of character transitions was crucial in testing whether genome duplication coincided or preceded transitions to migratory behavior. Although multiple previous authors have estimated character histories on fossil-calibrated time trees (e.g. Wiens et al., 2006; Syme and Oakley, 2012), traditional ancestral state estimation methods (e.g., maximum likelihood) usually provide probabilities of different character states at the nodes of trees, although this is mostly a matter of computational convenience and graphical display, and not intrinsic to the statistical method. In contrast, SCM simulates character histories, providing a ready-made probabilistic display of the character state histories along branches of the tree. We are unaware of previous authors using SCM on a time tree for absolute estimates of character state timing. Estimates of the absolute timing of state changes can allow comparisons of character evolution to geological events, and to the timing of other characters.

We report not the total number of state changes per time bin, but the state changes per branch during each time bin. This is because SCM simulates numerous changes on each branch that stem from the rapid rates of character evolution estimated by the Markov model. The high numbers of inferred changes often seem counter-intuitive, and may indicate that homogeneous models poorly describe the evolutionary process of many characters. Instead, models that permit various rates across a tree may better represent

character evolution (Skinner, 2010). Thus, in our analyses, when rates of evolution were high, we noted that raw histograms mainly reflected the number of lineages through time: when state changes occur on all branches, more branches yield more total changes. In this way, measuring state changes per branch reveals a more refined understanding of character history.

5. Conclusions

We present the first phylogenetic analysis to support multiple evolutionary transitions to migratory behavior in salmonids. We discovered that anadromy evolved at least twice within the group, a conclusion supported by statistically robust comparative results. Furthermore, we find support in one instance that genome duplication preceded the evolution of anadromy by 55–50 MYA. Our results suggest for future studies the benefit of combining data from Recent and fossil records to study the evolution of complex traits.

Acknowledgements

We acknowledge support from the Center for Scientific Computing at the CNSI and MRL: an NSF MRSEC (DMR-1121053) and NSF CNS-0960316. B.A.S. was supported by the Andrew W. Mellon Foundation, award no. 031424-003. N.J.M. was supported by NSF Grant DEB-0919451 "Bivalves in Time and Space", a U.C. Berkeley Wang Family Fellowship, a U.C. Berkeley Tien Fellowship, and a Google Summer of Code grant. This work was also supported by NSF Grant DEB-1046121 awarded to T.H.O.

Appendix A. Supplementary material

Supplementary data associated with this article can be found, in the online version, at http://dx.doi.org/10.1016/j.ympev.2013.07.026.

References

Adams, K.L., Wendel, J.F., 2005. Polyploidy and genome evolution in plants. Current Opinion in Plant Biology 8, 135–141.
Allendorf, F.W., Seeb, L.W., 2000. Concordance of genetic divergence among sockeye salmon populations at allozyme, nuclear DNA, and mitochondrial DNA markers. Evolution 54, 640–651.
Allendorf, F.W., Thorgaard, G.H., 1984. Tetraploidy and evolution of salmonid fishes. In: Turner, B.J. (Ed.), Evolutionary Genetics of Fishes. Plenum Press, New York, pp. 1–53.
Angers, B., Gharbi, K., Estoup, A., 2002. Evidence of gene conversion events between paralogous sequences produced by tetraploidization in Salmoninae fish. Journal of Molecular Evolution 54, 501–510.
Arratia, G., 1997. Basal teleosts and teleostean phylogeny. Palaeo Ichthyologica 7, 5–168.
Bollback, J.P., 2006. SIMMAP: stochastic character mapping of discrete traits on phylogenies. BMC Bioinformatics 7.
Braasch, I., Brunet, F., Volff, J.N., Schartl, M., 2009a. Pigmentation pathway evolution after whole-genome duplication in fish. Genome Biology and Evolution 1, 479–493.
Braasch, I., Volff, J.N., Schartl, M., 2009b. The endothelin system: evolution of vertebrate-specific ligand-receptor interactions by three rounds of genome duplication. Molecular Biology and Evolution 26, 783–799.
Broughton, R.E., 2010. Phylogeny of teleosts based on mitochondrial genome sequences. In: Nelson, J.S., Schultze, H.-P., Wilson, M.V.H. (Eds.), Origin and Phylogenetic Interrelationships of Teleosts. Verlag Dr. Friedrich Pfeil, Munchen, Germany, pp. 61–76.
Buckland, F., 1873. Familiar History of British Fishes. SPCK, London, UK.
Carrera, E., Garcia, T., Cespedes, A., Gonzalez, I., Fernandez, A., Hernandez, P.E., Martin, R., 1999. Salmon and trout analysis by PCR-RFLP for identity authentication. Journal of Food Science 64, 410–413.
Comai, L., 2005. The advantages and disadvantages of being polyploid. Nature Reviews Genetics 6, 836–846.
Crespi, B.J., Fulton, M.J., 2004. Molecular systematics of Salmonidae: combined nuclear data yields a robust phylogeny. Molecular Phylogenetics and Evolution 31, 658–679.
Crete-Lafreniere, A., Weir, L.K., Bernatchez, L., 2012. Framing the Salmonidae family phylogenetic portrait: a more comprehensive picture from increased taxon sampling. PLoS ONE 7, e46662.
Cunningham, C.W., Omland, K.E., Oakley, T.H., 1998. Reconstructing ancestral character states: a critical reappraisal. Trends in Ecology & Evolution 13, 361–366.
Danzmann, R.G., Cairney, M., Davidson, W.S., et al., 2005. A comparative analysis of the rainbow trout genome with 2 other species of fish (Arctic charr and Atlantic salmon) within the tetraploid derivative Salmonidae family (subfamily: Salmoninae). Genome 48, 1037–1051.
Day, F., 1887. British and Irish Salmonidae. Williams and Norgate, London, UK.
Donoghue, P.C.J., Purnell, M.A., 2005. Genome duplication, extinction and vertebrate evolution. Trends in Ecology & Evolution 20, 312–319.
Drummond, A.J., Suchard, M.A., Xie, D., Rambaut, A., 2012. Bayesian phylogenetics with BEAUti and the BEAST 1.7. Molecular Biology and Evolution.
Edgar, R.C., 2004. MUSCLE: a multiple sequence alignment method with reduced time and space complexity. BMC Bioinformatics 5, 1–19.
Eiting, T.P., Smith, G.R., 2007. Miocene salmon (*Oncorhynchus*) from western North America: gill raker evolution correlated with plankton productivity in the Eastern Pacific. Palaeogeography, Palaeoclimatology, Palaeoecology 249, 412–424.
Fleming, J., 1828. A history of British animals, exhibiting the descriptive characters and systematical arrangement of the genera and species of quadrupeds, birds, reptiles, fishes, mollusca, and radiata of the United Kingdom, including the indigenous, extirpated, and extinct kinds, together with periodical and occasional visitants. Edinburgh, UK.
Gernhard, T., 2008. The conditioned reconstructed process. Journal of Theoretical Biology 253, 769–778.
Gharbi, K., Semple, J.W., Ferguson, M.M., Schulte, P.M., Danzmann, R.G., 2004. Linkage arrangement of Na, K-ATPase genes in the tetraploid-derived genome of the rainbow trout (*Oncorhynchus mykiss*). Animal Genetics 35, 321–325.
Gharbi, K., Ferguson, M.M., Danzmann, R.G., 2005. Characterization of Na, K-ATPase genes in Atlantic salmon (*Salmo salar*) and comparative genomic organization with rainbow trout (*Oncorhynchus mykiss*). Molecular Genetics and Genomics 273, 474–483.
Günther, A., 1866. Catalogue of the fishes of the British Museum. London.
Hamada, M., Takasaki, N., Reist, J.D., DeCicco, A.L., Goto, A., Okada, N., 1998. Detection of the ongoing sorting of ancestrally polymorphic SINEs toward fixation or loss in populations of two species of charr during speciation. Genetics 150, 301–311.
Heikkila, M., Kaila, L., Mutanen, M., Pena, C., Wahlberg, N., 2012. Cretaceous origin and repeated tertiary diversification of the redefined butterflies. Processing Biological Science 279, 1093–1099.
Heimberg, A.M., Sempere, L.F., Moy, V.N., Donoghue, P.C.J., Peterson, K.J., 2008. MicroRNAs and the advent of vertebrate morphological complexity. Proceedings of the National academy of Sciences of the United States of America 105, 2946–2950.
Hoar, W.S., 1976. Smolt transformation: evolution, behavior, and physiology. Journal of the Fisheries Research Board of Canada 33, 1234–1252.
Hu, C., Lin, S.Y., Chi, W.T., Charng, Y.Y., 2012. Recent gene duplication and subfunctionalization produced a mitochondrial GrpE, the nucleotide exchange factor of the Hsp70 complex, specialized in thermotolerance to chronic heat stress in Arabidopsis. Plant Physiology 158, 747–758.
Huelsenbeck, J.P., Ronquist, F., 2001. MRBAYES: bayesian inference of phylogenetic trees. Bioinformatics 17, 754–755.
Huelsenbeck, J.P., Nielsen, R., Bollback, J.P., 2003. Stochastic mapping of morphological characters. Systematic Biology 52, 131–158.
Ishiguro, N.B., Miya, M., Nishida, M., 2003. Basal euteleostean relationships: a mitogenomic perspective on the phylogenetic reality of the "Protacanthopterygii". Molecular Phylogenetics and Evolution 27, 476–488.
Jaillon, O., Aury, J.M., Brunet, F., et al., 2004. Genome duplication in the teleost fish Tetraodon nigroviridis reveals the early vertebrate proto-karyotype. Nature 431, 946–957.
Johnson, K.R., Wright, J.E., May, B., 1987. Linkage relationships reflecting ancestral tetraploidy in salmonid fish. Genetics 116, 579–591.
Koop, B.F., von Schalburg, K.R., Leong, J., et al., 2008. A salmonid EST genomic study: genes, duplications, phylogeny and microarrays. BMC Genomics 9, 545.
Leong, J.S., Jantzen, S.G., von Schalburg, K.R., et al., 2010. *Salmo salar* and *Esox lucius* full-length cDNA sequences reveal changes in evolutionary pressures on a post-tetraploidization genome. BMC Genomics 11.
Lewis, P.O., 2001. A likelihood approach to estimating phylogeny from discrete morphological character data. Systematic Biology 50, 913–925.
Lopez, J.A., Chen, W.J., Orti, G., 2004. Esociform phylogeny. Copeia, 449–464.
Lynch, M., Conery, J.S., 2000. The evolutionary fate and consequences of duplicate genes. Science 290, 1151–1155.
Mable, B.K., 2004. 'Why polyploidy is rarer in animals than in plants': myths and mechanisms. Biological Journal of the Linnean Society 82, 453–466.
Mable, B.K., Alexandrou, M.A., Taylor, M.I., 2011. Genome duplication in amphibians and fish: an extended synthesis. Journal of Zoology 284, 151–182.
McDowall, R.M., 2002. The origin of Salmonid fishes: marine, freshwater... or neither? Reviews in Fish Biology and Fisheries 11, 171–179.
Meek, A., 1916. The Migrations of Fish. Edward Arnold, London, UK.
Mooers, A.O., Schluter, D., 1999. Reconstructing ancestor states with maximum likelihood: support for one- and two-rate models. Systematic Biology 48, 623–633.
Nelson, J., 2006. Fishes of the World. John Wiley & Sons Inc., New Jersey.
Nielsen, R., 2002. Mapping mutations on phylogenies. Systematic Biology 51, 729–739.

Nielsen, E.E., Hansen, M.M., Mensberg, K.L.D., 1998. Improved primer sequences for the mitochondrial ND1, ND3/4 and ND5/6 segments in salmonid fishes: application to RFLP analysis of Atlantic salmon. Journal of FishBiology 53, 216–220.
Norman, J.D., Danzmann, R.G., Glebe, B., Ferguson, M.M., 2011. The genetic basis of salinity tolerance traits in Arctic charr (*Salvelinus alpinus*). BMC Genetics 12, 81.
Norman, J.D., Robinson, M., Glebe, B., Ferguson, M.M., Danzmann, R.G., 2012. Genomic arrangement of salinity tolerance QTLs in salmonids: a comparative analysis of Atlantic salmon (*Salmo salar*) with Arctic charr (*Salvelinus alpinus*) and rainbow trout. BMC Genomics 13, 420.
Novak, A.E., Jost, M.C., Lu, Y., Taylor, A.D., Zakon, H.H., Ribera, A.B., 2006. Gene duplications and evolution of vertebrate voltage-gated sodium channels. Journal of Molecular Evolution 63, 208–221.
Nylander, J.A.A., Ronquist, F., Huelsenbeck, J.P., Nieves-Aldrey, J.L., 2004. Bayesian phylogenetic analysis of combined data. Systematic Biology 53, 47–67.
Oakley, T.H., Cunningham, C.W., 2000. Independent contrasts succeed where ancestor reconstruction fails in a known bacteriophage phylogeny. Evolution 54, 397–405.
Oakley, T.H., Phillips, R.B., 1999. Phylogeny of salmonine fishes based on growth hormone introns: Atlantic (*Salmo*) and Pacific (*Oncorhynchus*) salmon are not sister taxa. Molecular Phylogenetics and Evolution 11, 381–393.
Oakley, T.H., Gu, Z.L., Abouheif, E., Patel, N.H., Li, W.H., 2005. Comparative methods for the analysis of gene-expression evolution: an example using yeast functional genomic data. Molecular Biology and Evolution 22, 40–50.
Ohno, S., 1970. Enormous diversity in genome sizes of fish as a reflection of nature's extensive experiments with gene duplication. Transactions of the American Fisheries Society 99 (120-&).
Olsen, J.B., Bentzen, P., Banks, M.A., Shaklee, J.B., Young, S., 2000. Microsatellites reveal population identity of individual pink salmon to allow supportive breeding of a population at risk of extinction. Transactions of the American Fisheries Society 129, 232–242.
Oswald, B.P., Nuismer, S.L., 2007. Neopolyploidy and pathogen resistance. Proceedings of the Royal Society B-Biological Sciences 274, 2393–2397.
Otto, S.P., Whitton, J., 2000. Polyploid incidence and evolution. Annual Review of Genetics 34, 401–437.
Pagel, M., Meade, A., Barker, D., 2004. Bayesian estimation of ancestral character states on phylogenies. Systematic Biology 53, 673–684.
Pennant, T., 1776. British Zoology: Fishes. Wilkie and Robinson, London, UK.
Phillips, R.B., Oakley, T.H., 1997. Phylogenetic relationships among the Salmoninae based on their nuclear and mitochondrial sequences. In: Kocher, T.D., Sepien, C.A. (Eds.), Molecular Systematics of Fishes. Academic Press, San Diego, CA, pp. 145–162.
Phillips, R., Rab, P., 2001. Chromosome evolution in the Salmonidae (Pisces): an update. Biological Reviews 76, 1–25.
Phillips, R.B., Pleyte, K.A., Brown, M.R., 1992. Salmonid phylogeny inferred from ribosomal DNA restriction maps. Canadian Journal of Fisheries and Aquatic Sciences 49, 2345–2353.
Phillips, R.B., Manley, S.A., Daniels, T.J., 1994. Systematics of the salmonid genus salvelinus inferred from ribosomal DNA-sequences. Canadian Journal of Fisheries and Aquatic Sciences 51, 198–204.
Phillips, R.B., Oakley, T.H., Davis, E.L., 1995. Evidence supporting the paraphyly of Hucho (Salmonidae) based on ribosomal DNA restriction maps. Journal of Fish Biology 47, 956–961.
Posada, D., 2008. JModelTest: phylogenetic model averaging. Molecular Biology and Evolution 25, 1253–1256.
Pyron, R.A., 2011. Divergence time estimation using fossils as terminal taxa and the origins of lissamphibia. Systematic Biology 60, 466–481.
Rambaut, A., Drummond, A.J., 2004. Tracer v1.4.
Ramsden, S.D., Brinkmann, H., Hawryshyn, C.W., Taylor, J.S., 2003. Mitogenomics and the sister of Salmonidae. Trends in Ecology and Evolution 18, 607–610.
Revell, L.J., 2012. Phytools: an R package for phylogenetic comparative biology (and other things). Methods in Ecology and Evolution 3, 217–223.
Ronquist, F., Klopfstein, S., Vilhelmsen, L., Schulmeister, S., Murray, D.L., Rasnitsyn, A.P., 2012. A total-evidence approach to dating with fossils, applied to the early radiation of the Hymenoptera. Systematic Biology 61, 973–999.
Sakamoto, T., Danzmann, R.G., Gharbi, K.A., 2000. microsatellite linkage map of rainbow trout (Oncorhynchus mykiss) characterized by large sex-specific differences in recombination rates. Genetics 19, 1331–1345.
Sanderson, M.J., 2003. R8s: inferring absolute rates of molecular evolution and divergence times in the absence of a molecular clock. Bioinformatics 19, 301–302.
Sanderson, M.J., McMahon, M.M., Steel, M., 2011. Terraces in phylogenetic tree space. Science 333, 448–450.
Schluter, D., Price, T., Mooers, A.O., Ludwig, D., 1997. Likelihood of ancestor states in adaptive radiation. Evolution 51, 1699–1711.
Schultz, T.R., Cocroft, R.B., Churchill, G.A., 1996. The reconstruction of ancestral character states. Evolution 50, 504–511.
Semon, M., Wolfe, K.H., 2007. Consequences of genome duplication. Current Opinion in Genetics and Development 17, 505–512.
Shedko, S.V., Miroshnichenko, I.L., Nemkova, G.A., 2012. Phylogeny of salmonids (*Salmoniformes*: *Salmonidae*) and its molecular dating: analysis of nuclear RAG1 gene. Russian Journal of Genetics 48, 575–579.
Skinner, A., 2010. Rate heterogeneity, ancestral character state reconstruction, and the evolution of limb morphology in lerista (*Scincidae*, *Squamata*). Systematic Biology 59, 723–740.
Smith, G.R., 1981. Late cenozoic fresh-water fishes of North-America. Annual Review of Ecology and Systematics 12, 163–193.
Smith, G.R., 1992. Introgression in fishes – significance for paleontology, cladistics, and evolutionary rates. Systematic Biology 41, 41–57.
Soltis, D.E., Buggs, R.J.A., Doyle, J.J., Soltis, P.S., 2010. What we still do not know about polyploidy. Taxon 59, 1387–1403.
Stadler, T., 2009. On incomplete sampling under birth-death models and connections to the sampling-based coalescent. Journal of Theoretical Biology 261, 58–66.
Stamatakis, A., 2006. RAxML-VI-HPC: maximum likelihood-based phylogenetic analyses with thousands of taxa and mixed models. Bioinformatics 22, 2688–2690.
Stearley, R.F., 1992. Historical ecology of Salmoninae, with special reference to *Oncorhynchus*. In: Mayden, R.L. (Ed.), Systematics, Historical Ecology, and North American Freshwater Fishes. Stanford University Press, Stanford, CA, pp. 622–658.
Stearley, R.F., Smith, G.R., 1993. Phylogeny of the Pacific trout and salmon (*Oncorhynchus*) and genera of the family Salmonidae. Transactions of the American Fisheries Society 122, 1–33.
Syme, A.E., Oakley, T.H., 2012. Dispersal between shallow and abyssal seas and evolutionary loss and regain of compound eyes in cylindroleberidid ostracods: conflicting conclusions from different comparative methods. Systematic Biology 61, 314–336.
Takasaki, N., Yamaki, T., Hamada, M., Park, L., Okada, N., 1997. The salmon Smal family of short interspersed repetitive elements (SINEs): interspecific and intraspecific variation of the insertion of SINEs in the genomes of chum and pink salmon. Genetics 146, 369–380.
Taylor, J.S., Van de Peer, Y., Meyer, A., 2001. Genome duplication, divergent resolution and speciation. Trends in Genetics 17, 299–301.
Taylor, J.S., Braasch, I., Frickey, T., Meyer, A., Van de Peer, Y., 2003. Genome duplication, a trait shared by 22,000 species of ray-finned fish. Genome Research 13, 382–390.
Tchernavin, V., 1939. Ripe salmon parr: a summary of research. Proceedings of the Royal Physical Society of Edinburgh 23, 73–78.
Thorpe, J.E., 1982. Migration in salmonids with special reference to juvenile movements in freshwater. In: Brannon, E.L., Salo, E.O. (Eds.), Salmon and Trout Migratory Behavior Symposium. School of Fisheries, University of Washington, Seattle.
Tipsmark, C.K., 2008. Identification of FXYD protein genes in a teleost: tissue-specific expression and response to salinity change. American Journal of Physiology-Regulatory Integrative and Comparative Physiology 294, R1367–R1378.
Tipsmark, C.K., Baltzegar, D.A., Ozden, O., Grubb, B.J., Borski, R.J., 2008a. Salinity regulates claudin mRNA and protein expression in the teleost gill. American Journal of Physiology–Regulatory Integrative and Comparative Physiology 294, R1004–R1014.
Tipsmark, C.K., Kiilerich, P., Nilsen, T.O., Ebbesson, L.O.E., Stefansson, S.O., Madsen, S.S., 2008b. Branchial expression patterns of claudin isoforms in Atlantic salmon during seawater acclimation and smoltification. American Journal of Physiology–Regulatory Integrative and Comparative Physiology 294, R1563–R1574.
Tipsmark, C.K., Luckenbach, J.A., Madsen, S.S., Kiilerich, P., Borski, R.J., 2008c. Osmoregulation and expression of ion transport proteins and putative claudins in the gill of Southern Flounder (*Paralichthys lethostigma*). Comparative Biochemistry and Physiology a-Molecular & Integrative Physiology 150, 265–273.
Van de Peer, Y., Taylor, J.S., Meyer, A., 2003. Are all fishes ancient polyploids? Journal of Structural and Functional Genomics 3, 65–73.
Van de Peer, Y., Maere, S., Meyer, A., 2009. OPINION The evolutionary significance of ancient genome duplications. Nature Reviews Genetics 10, 725–732.
Warner, D.A., Edwards, G.E., 1993. Effects of Polyploidy on Photosynthesis. Photosynthesis Research 35, 135–147.
Wiens, J.J., Brandley, M.C., Reeder, T.W., 2006. Why does a trait evolve multiple times within a clade? Repeated evolution of snakelike body form in squamate reptiles. Evolution 60, 123–141.
Williams, R.R.G., 1987. The phylogenetic relationships of the salmoniform fishes based on the suspensorium and its muscles. University of Alberta, Edmonton.
Wilson, M.V.H., Li, G.Q., 1999. Osteology and systematic position of the Eocene salmonid *Eosalmo driftwoodensis* Wilson from western North America. Zoological Journal of the Linnean Society 125, 279–311.
Wilson, M.V.H., Williams, R.R.G., 2010. Salmoniform fishes: key fossils, supertree, and possible morphological synapomorphies. In: Nelson, J.S., Schultze, H.-P., Wilson, M.V.H. (Eds.), Origin and Phylogenetic Interrelationships of Teleosts. Verlag Dr. Friedrich Pfeil, Munchen, Germany, pp. 379–409.
Wilson, M.V.H., Brinkman, D.B., Neuman, A.G., 1992. Cretaceous esocoidei (*Teleostei*) – early radiation of the pikes in North-American fresh waters. Journal of Paleontology 66, 839–846.
Wood, H.M., Matzke, N.J., Gillespie, R.G., Griswold, C.E., 2012. Treating fossils as terminal taxa in divergence time estimation reveals ancient vicariance patterns in the palpimanoid spiders. Systematic Biology.
Woram, R.A., McGowan, C., Stout, J.A., Gharbi, K., Ferguson, M.M., Hoyheim, B., Davidson, E.A., Davidson, W.S., Rexroad, C., Danzmann, R.G., 2004. A genetic linkage map for Arctic char (*Salvelinus alpinus*): evidence for higher recombination rates and segregation distortion in hybrid versus pure strain mapping parents. Genome 47, 304–315.