Author for correspondence:
Magnus W. Jacobsen
e-mail: magnus.jacobsen@bios.au.dk

THE ROYAL SOCIETY
PUBLISHING

Molecular evolution

Relationship between amino acid changes in mitochondrial ATP6 and life-history variation in anguillid eels

Magnus W. Jacobsen, José Martin Pujolar and Michael M. Hansen

Department of Bioscience, Aarhus University, Ny Munkegade 114, Aarhus C 8000, Denmark

Mitochondrial genes are part of the oxidative phosphorylation pathway and important for energy production. Although evidence for positive selection at the mitochondrial level exists, few studies have investigated the link between amino acid changes and phenotype. Here we test the hypothesis that differences in two life-history related traits, migratory distance between spawning and foraging areas and larval phase duration, are associated with divergent selection within the mitochondrial ATP6 gene in anguillid eels. We compare amino acid changes among 18 species with the sequence of the putative ancestral species, believed to have shown short migratory distance and larval phase duration. We find positive correlations between both life-history related traits and (i) the number of amino acid changes and (ii) the strength of the combined physico-chemical and structural changes at positions previously identified as candidates for positive selection. This supports a link between genotype and phenotype driven by positive selection at ATP6.

1. Introduction

The mitochondrial genome (mitogenome) consists of 13 coding genes directly involved in the oxidative phosphorylation pathway and the production of energy (ATP) [1]. Although the mitogenome historically has been assumed to evolve neutrally [2], multiple studies show evidence for positive selection acting on mitochondrial genes in mammals [3,4] and fishes (e.g. [5]). So far, however, few studies have provided solid evidence for positive selection by demonstrating direct links between amino acid changes and phenotype.

Freshwater eels (*Anguilla* spp.) constitute excellent candidates for investigating positive selection at mitogenomes and its effect on phenotype. The genus encompasses several species [6], which furthermore show high effective population sizes (e.g. [7,8]). Thus, selection is likely to be a stronger evolutionary force than genetic drift in these species. Positive selection at the mitochondrial ATP6 gene has been shown between the American (*Anguilla rostrata*) and European eel (*Anguilla anguilla*) based on McDonald–Kreitman tests [8,9]. Moreover, evidence for positive diversifying selection has been reported at codon 52 in ATP6 across all species of freshwater eels, whereas such consistent patterns of selection are not present at other mitochondrial genes [8].

ATP6 is a subunit of the F_0-proton channel in ATP synthase, the fifth and final complex of the oxidative phosphorylation chain which catalyses synthesis of ATP from ADP [1]. Changes within ATP6 have been associated with differences in metabolism and selection linked to energetics [3,4,10]. In freshwater eels, selection may be linked to their migratory loops [8,11]. These involve spawning migration (from freshwater or coastal regions to the ocean) of highly variable distances in different species and consequently also a highly variable larval phase, as larvae rely on ocean currents for advection back to the continental feeding grounds [6,11]. Thus, energy production linked to the

spawning migration, or migration or metabolism during the larval stage, may underlie selection at ATP6.

The ancestor of all current eel species presumably lived in southeast Asia or Africa [6,11] around 14–18 Ma [8,12]. As most of the species found there show a short migratory loop, this is believed to be the ancestral state [11,13]. If length of spawning migration and larval phase duration underlie directional selection at ATP6, this should be reflected in more fixed non-synonymous changes and changes of stronger physico-chemical and structural importance at key positions in species with longer migratory loops. We test this by assessing the correlation of amino acid changes from the putative ancestral sequence with migratory distance or larval phase duration of species.

2. Material and methods

Twenty-one mitogenome sequences were downloaded from GenBank: 18 species of freshwater eels (GenBank accession nos NC_002707, NC_006531–NC_006547) and three species of Serrivomeridae, the family most closely related to Anguillidae [14] (NC_013436, NC_013627–NC_013628). Sequences were aligned in GENEIOUS PRO v. 5.4.6 [15] and the 13 coding genes were extracted and concatenated. The ancestral sequence of all freshwater eels was reconstructed using the Ancestral State Reconstruction software (ASR) [16] implemented in the HYPHY package on the DataMonkey server (http://www.datamonkey.org/dataupload.php) (electronic supplementary material, note S1). Synonymous and non-synonymous changes and dN/dS between the putative ancestral sequence and each species were calculated using DnaSP v. 5.1 [17].

Larval phase duration (estimated by age at metamorphosis of larvae based on Sr/Ca ratio increments in otoliths) and migratory distance to the spawning ground were obtained from the literature (electronic supplementary material, table S1). Mean migratory distances were based on estimates from Aoyama [11] or estimated in the case of the New Zealand eel (Anguilla differenbachii) (electronic supplementary material, note S2). When several estimates were available, the average among studies was used (see electronic supplementary material, note S2 and table S1). For the European and American eels, significant discrepancies exist between studies [18]. Hence, the mean estimates from Wang & Tzeng [19] were chosen, as those estimates were derived from the largest sample sizes to date and accord well with studies that have estimated larval phase duration based on ocean current transportation [20].

Correlations between genotype and either migratory distance or larval phase duration were tested using linear regression or Spearman rank analysis in PAST [21]. All data were checked for normality prior to parametric statistical analysis and a logarithmic transformation was conducted if necessary (electronic supplementary material, table S2). A possible correlation between migratory distance and larval phase duration was tested to avoid pseudoreplication using both the extant species' trait values and an analysis of independent contrasts (CAIC) correcting for phylogenetic relatedness (electronic supplementary material, note S3) [22].

Genotype was represented by different datasets: (i) total non-synonymous (amino acid changing) and synonymous changes and dN/dS considering all 13 genes, (ii) total amino acid changes at ATP6 and (iii) within five amino acid positions fixed between the American and European eels possibly under positive selection [8,9]. Differences in properties of amino acid replacements between the ancestral and species sequences were assessed using 11 physico-chemical and four structural properties [23]. The numerical differences of properties were summed to produce an overall score. This was conducted for datasets 2 and 3 and for codon position 52, which shows evidence for diversifying selection across the entire phylogeny of freshwater eels [8].

3. Results and discussion

Larval phase duration and migratory distance were highly correlated even when correcting for phylogenetic relationships (figure 1a,b), making us unable to make reliable inferences about which of these two traits selection is acting upon.

Number of amino acid changes and their overall difference in physico-chemical and structural properties within ATP6 did not correlate with either trait (figure 1f,g). However, when analysing the positions previously found to be involved in positive selection between A. anguilla and A. rostrata [8,9], significant positive correlations were observed. Currently, there is no high-resolution three-dimensional structure for ATP synthase available, which precludes prediction of the exact functional importance of the five positions. However, the positions are likely to cover important roles for protein function in the ATP6 gene in freshwater eels. Three of the positions (83, 105 and 195) are situated in transmembrane domains (electronic supplementary material, figure S1) that are highly conserved in mammals [3], and are thus likely to cover important regions. Position 195 is located in the last transmembrane domain, where mutation is associated with severe decrease in the rate of ATP production in humans [10]. Finally, positions 43 and 52 are found in a region likely to be under positive selection in mammals [3].

A significant correlation was found between the amino acid property score and duration of larval phase at codon 52. However, the association was mainly driven by A. anguilla (figure 1j, table 1). This position has been shown to be under positive diversifying selection across the 18 freshwater eels [8] and may represent an important site for protein function. However, it seems likely that positive selection involves multiple sites. This is further supported by results at the five candidate positions in ATP6 [8,9]. Within these sites, both the number and property score of derived amino acids show significant positive correlations with the examined traits (figure 1h,i). Interestingly, at these positions the derived amino acids relative to the ancestral reconstructed sequence are highly polymorphic between species with long migratory loops. This is also the case between long-looped Atlantic and Oceanian species that are sister clades (e.g. [8,12]; figure 2). This suggests that evolution at these sites is not related to the phylogenetic relationship and that the observed changes may be adaptive, occurring after species divergence. Moreover, the immediate ancestral nodes of the current long-looped Atlantic and Oceanian species also show multiple changes within the five candidate positions (figure 2). As these ancestors likely showed long migratory loops this result further supports positive selection within ATP6.

The pattern shown here could also be the outcome of relaxed purifying selection in long-looped species if they have experienced higher genetic drift. As such, evolution in the candidate positions would be slightly deleterious instead of adaptive. However, this is not supported here, as analyses of dN/dS do not show significant correlations with the length of the migratory loop (figure 1e). Moreover, the results are not related to substitution rate. In fact, species with smaller migratory loops show the highest number of substitutions from the ancestral sequence, supporting a higher substitution rate (figure 1c,d). This may reflect the fact that generation time is positively associated with food availability in eels [24], potentially leading to faster generation time and thereby
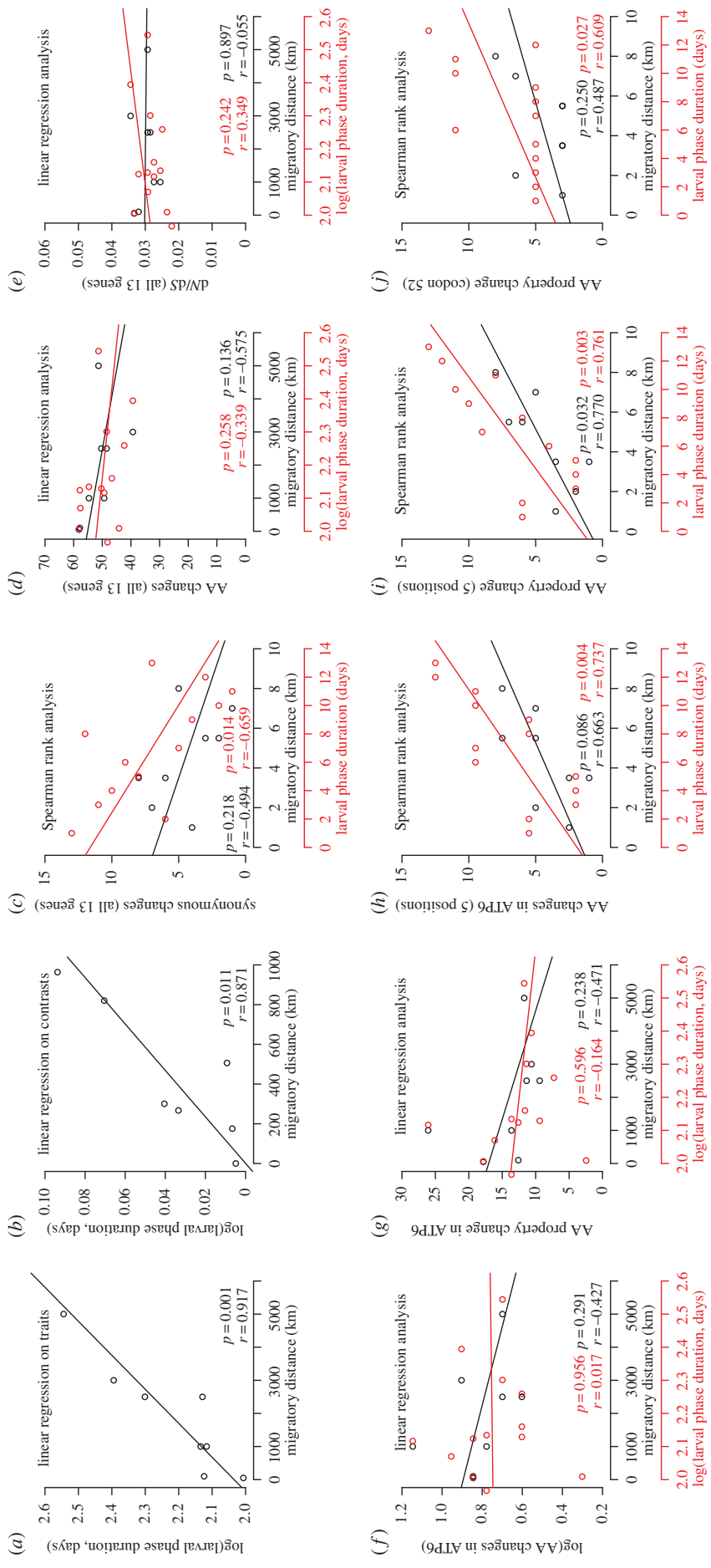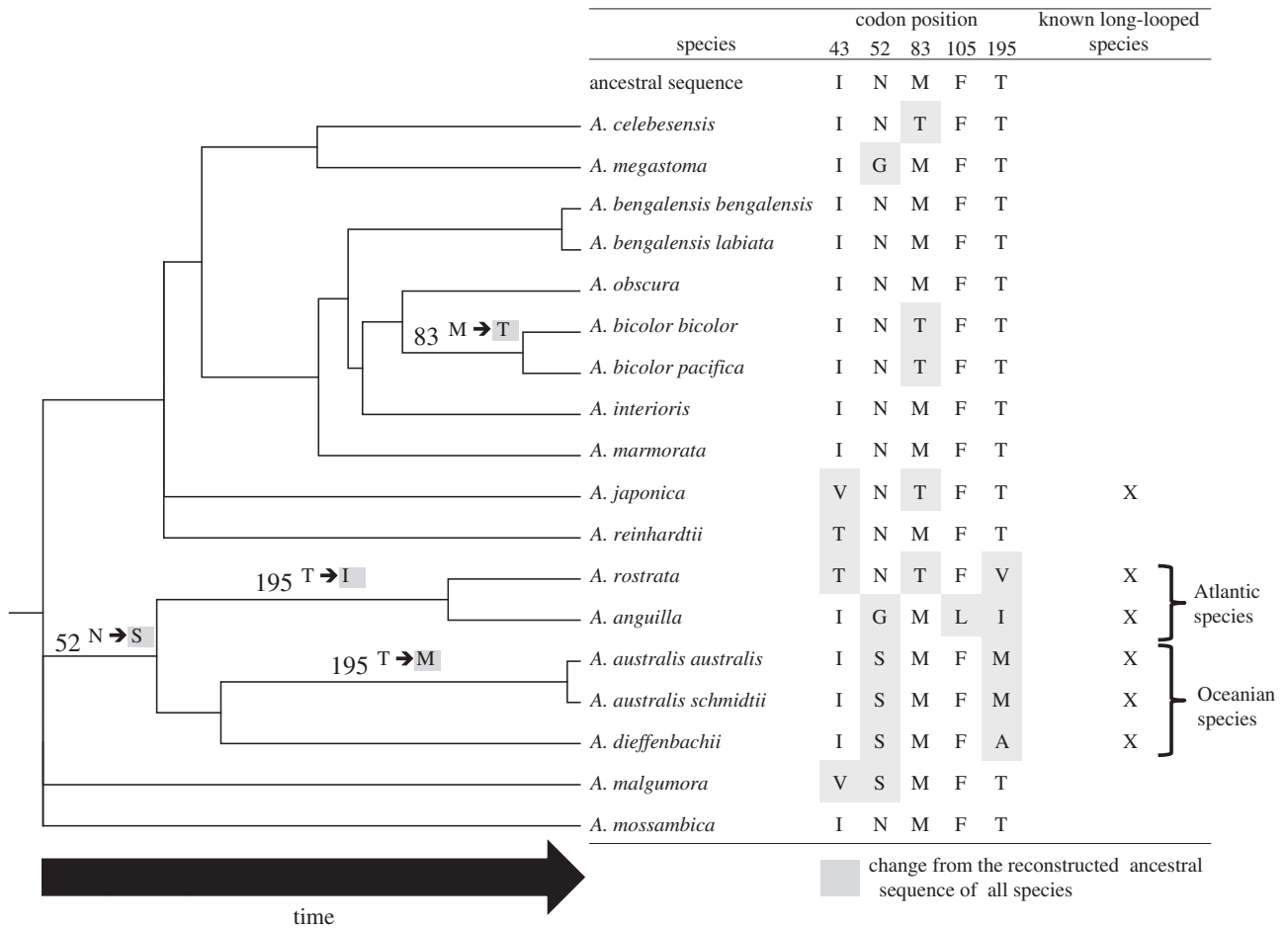
Figure 1. Correlation analyses. Each plot shows data along with the best-fitted line, correlation coefficient and *p*-values. When multiple plots are present the colours match the *x*-axis. Number of changes and property scores represent the overall difference between the current species and the putative ancestral sequence. AA denotes amino acid (non-synonymous substitutions).

Table 1. Estimates for each species of migratory distances, larval phase duration and genetic change from the putative ancestral sequence. For calculations and references, see electronic supplementary material, note S1 and table S1.

species	migratory distance mean (km)	larval phase duration mean (days)	overall number of changes (all 13 genes)		overall dN/dS (all 13 genes)	number of amino acid changes in ATP6		overall amino acid property change in ATP6		
			synonymous	non-synonymous		all sites	five sites[a]	all sites	five sites[a]	codon 52[b]
A. celebesensis	50	101.50	522.17	58.17	0.0333	7	1	17.822	3.332	0
A. megastoma	NA	NA	534.42	63.92	0.0355	7	1	16.158	3.656	3.656
A. bengalensis bengalensis/ A. nebulosa nebulosa	NA	NA	555.75	54.92	0.0308	9	0	16.118	0	0
A. bengalensis labiata/ A. nebulosa labiata	NA	117.50	554.00	57.67	0.0292	9	0	16.118	0	0
A. obscura	NA	NA	566.00	45.17	0.0235	5	0	11.109	0	0
A. bicolor bicolor	NA	92.60	631.50	48.17	0.0221	6	1	13.619	3.332	0
A. bicolor pacifica	1000	136.25	619.00	54.67	0.0255	6	1	13.619	3.332	0
A. interioris	NA	NA	539.17	46.17	0.0254	6	0	13.670	0	0
A. marmorata	1000	130.77	532.00	49.33	0.0274	14	0	26.073	0	0
A. japonica	2500	134.56	512.67	50.33	0.0293	4	2	9.397	4.555	0
A. reinhardtii	NA	144.50	511.00	46.67	0.0274	4	1	11.602	4.663	0
A. rostrata	2500	200.00	507.50	48.5	0.0285	5	3	11.355	9.228	0
A. anguilla	5000	350.20	524.00	51.33	0.0293	5	3	11.712	9.585	3.656
A. australis australis	NA	181.59	478.67	42.33	0.0249	4	2	7.275	5.148	1.816
A. australis schmidtii	NA	NA	473.17	38.83	0.0267	4	2	7.275	5.148	1.816
A. dieffenbachii	3000	248.00	354.00	39.33	0.0344	8	2	10.619	3.919	1.816
A. malgumora/ A. borneensis	100	133.00	537.83	57.83	0.0320	7	2	12.591	3.039	1.816
A. mossambica	NA	102.10	556.83	44.17	0.0235	2	0	2.446	0	0

[a]Codon positions 43, 52, 83, 105 and 195, candidates for positive selection between Atlantic eels (8,9).
[b]Candidate for diversifying selection across all species of freshwater eels (8).

species	| 43	52	83	105	195	known long-looped species
ancestral sequence	I	N	M	F	T	
A. celebesensis	I	N	T	F	T	
A. megastoma	I	G	M	F	T	
A. bengalensis bengalensis	I	N	M	F	T	
A. bengalensis labiata	I	N	M	F	T	
A. obscura	I	N	M	F	T	
A. bicolor bicolor	I	N	T	F	T	
A. bicolor pacifica	I	N	T	F	T	
A. interioris	I	N	M	F	T	
A. marmorata	I	N	M	F	T	
A. japonica	V	N	T	F	T	X
A. reinhardtii	T	N	M	F	T	
A. rostrata	T	N	T	F	V	X Atlantic species
A. anguilla	I	G	M	L	I	X
A. australis australis	I	S	M	F	M	X Oceanian species
A. australis schmidtii	I	S	M	F	M	X
A. dieffenbachii	I	S	M	F	A	X
A. malgumora	V	S	M	F	T	
A. mossambica	I	N	M	F	T	

Branch labels: 83 M → T; 195 T → I; 52 N → S; 195 T → M.

change from the reconstructed ancestral sequence of all species

time

Figure 2. Amino acid evolution in freshwater eels within the five candidate positions investigated. Ancestral changes are shown along the respective phylogenetic branches while current genotypes are shown in the table. The phylogeny is identical to that of Jacobsen *et al.* [8] using a posterior threshold of 0.90 to determine phylogenetic relationships.

substitution rate in highly productive tropical areas (e.g. [25]), such as those inhabited by most short-looped species (see electronic supplementary material, note S2 for discussion of other possible biases).

Overall, the pattern found in this study is in accordance with the hypothesis that positive selection at a few key positions drove divergence from an ancestral species with a short migratory loop [11]. According to this scenario, species with longer migratory loops show more amino acid changes and higher overall physico-chemical and structural change at ATP6 relative to the putative ancestral sequence. As changes in the migratory loop are likely crucial for understanding speciation in freshwater eels [8,11,13], selection at ATP6 may be particularly important for understanding their evolution. Selection may further lead to cytonuclear incompatibilities, as proposed for Atlantic eels [9], which could lead to reproductive isolation. Consequently, ATP6 may be a key gene underlying speciation in freshwater eels. This suggests that selection at mitochondrial genes may generally be important in groups of organisms characterized by large interspecific differences in migratory properties.

References

1. Saraste M. 1999 Oxidative phosphorylation at the fin de siecle. *Science* **283**, 1488–1493. (doi:10.1126/science.283.5407.1488)

2. Ballard JWO, Whitlock MC. 2004 The incomplete natural history of mitochondria. *Mol. Ecol.* **13**, 729–744. (doi:10.1046/j.1365-294X.2003.02063.x)

3. da Fonseca RR, Johnson WE, O'Brien SJ, Ramos MJ, Antunes A. 2008 The adaptive evolution of the mammalian mitochondrial genome. *BMC Genom.* **9**, 119. (doi:10.1186/1471-2164-9-119)

4. Fontanillas P, Depraz A, Giorgi MS, Perrin N. 2005 Nonshivering thermogenesis capacity associated to mitochondrial DNA haplotypes and gender in the greater white-toothed shrew, *Crocidura russula*. *Mol. Ecol.* **14**, 661–670. (doi:10.1111/j.1365-294X.2004.02414.x)

5. Garvin MR, Bielawski JP, Gharrett AJ. 2011 Positive Darwinian selection in the piston that powers

proton pumps in complex I of the mitochondria of Pacific salmon. *PLoS ONE* **6**, e24127. (doi:10.1371/journal.pone.0024127)

6. Tesch F. 2003 *The eel*. Oxford, UK: Blackwell Science.

7. Tseng MC, Kao HW, Hung YH, Lee TL. 2012 A study of genetic variations, population size, and population dynamics of the catadromous Japanese eel *Anguilla japonica* (Pisces) in northern Taiwan. *Hydrobiologia* **683**, 203–216. (doi:10.1007/s10750-011-0958-z)

8. Jacobsen MW, Pujolar JM, Gilbert MTP, Moreno-Mayar JV, Bernatchez L, Als TD, Lobon-Cervia J, Hansen MM. 2014 Speciation and demographic history of Atlantic eels (*Anguilla anguilla* and *A. rostrata*) revealed by mitogenome sequencing. *Heredity* **113**, 432–442. (doi:10.1038/hdy.2014.44)

9. Gagnaire PA, Normandeau E, Bernatchez L. 2012 Comparative genomics reveals adaptive protein evolution and a possible cytonuclear incompatibility between European and American eels. *Mol. Biol. Evol.* **29**, 2909–2919. (doi:10.1093/molbev/mss076)

10. Kucharczyk R, Ezkurdia N, Couplan E, Procaccio V, Ackerman SH, Blondel M, di Rago JP. 2010 Consequences of the pathogenic T9176C mutation of human mitochondrial DNA on yeast mitochondrial ATP synthase. *Biochim. Biophys. Acta* **1797**, 1105–1112. (doi:10.1016/j.bbabio.2009.12.022)

11. Aoyama J. 2009 Life history and evolution of migration in catadromous eels (genus *Anguilla*). *Aqua-BioSci. Monogr.* **2**, 1–42. See http://www.terrapub.co.jp/onlinemonographs/absm/.

12. Santini F, Kong X, Sorensen L, Carnevale G, Mehta RS, Alfaro ME. 2013 A multi-locus molecular timescale for the origin and diversification of eels (order: Anguilliformes). *Mol. Phylogenet. Evol.* **69**, 884–894. (doi:10.1016/j.ympev.2013.06.016)

13. Aoyama J, Wouthuyzen S, Miller MJ, Inagaki T, Tsukamoto K. 2003 Short-distance spawning migration of tropical freshwater eels. *Biol. Bull.* **204**, 104–108. (doi:10.2307/1543500)

14. Johnson GD, Ida H, Sakaue J, Sado T, Asahida T, Miya M. 2012 A 'living fossil' eel (Anguilliformes: Protanguillidae, fam. nov.) from an undersea cave in Palau. *Proc. R. Soc. B* **279**, 934–943. (doi:10.1098/rspb.2011.1289)

15. BioMatters. 2012 Geneious version 5.4.6. See http://www.geneious.com.

16. Pond SLK, Frost SDW. 2005 Not so different after all: a comparison of methods for detecting amino acid sites under selection. *Mol. Biol. Evol.* **22**, 1208–1222. (doi:10.1093/molbev/msi105)

17. Librado P, Rozas J. 2009 DnaSP v5: a software for comprehensive analysis of DNA polymorphism data. *Bioinformatics* **25**, 1451–1452. (doi:10.1093/bioinformatics/btp187)

18. McCleave JD. 2008 Contrasts between spawning times of *Anguilla* species estimated from larval sampling at sea and from otolith analysis of recruiting glass eels. *Mar. Biol.* **155**, 249–262. (doi:10.1007/s00227-008-1026-8)

19. Wang CH, Tzeng WN. 2000 The timing of metamorphosis and growth rates of American and European eel leptocephali: a mechanism of larval segregative migration. *Fish. Res.* **46**, 191–205. (doi:10.1016/S0165-7836(00)00146-6)

20. Kettle AJ, Haines K. 2006 How does the European eel (*Anguilla anguilla*) retain its population structure during its larval migration across the North Atlantic Ocean? *Can. J. Fish. Aquat. Sci.* **63**, 90–106. (doi:10.1139/cjz-2013-0303)

21. Hammer Ø, Harper DAT, Ryan PD. 2001 PAST: paleontological statistics software package for education and data analysis. *Palaeontol. Electron.* **4**, 9. See http://palaeo-electronica.org/2001_2001/past/issue2001_2001.htm.)

22. Purvis A, Rambaut A. 1995 Comparative analysis by independent contrasts (CAIC): an Apple Macintosh application for analysing comparative data. *Comput. Appl. Biosci.* **11**, 247–251. (doi:10.1093/bioinformatics/11.3.247)

23. Liu J, Kang X. 2012 Grading amino acid properties increased accuracies of single point mutation on protein stability prediction. *BMC Bioinform.* **13**, 44. (doi:10.1186/1471-2105-13-44)

24. Vollestad LA. 1992 Geographic variation in age and length at metamorphosis of maturing European eel: environmental effects and phenotypic plasticity. *J. Anim. Ecol.* **61**, 41–48. (doi:10.2307/5507)

25. Arai T, Chino N. 2013 Timing of maturation of a tropical eel, *Anguilla bicolor bicolor* in Malaysia. *J. Appl. Ichthyol.* **29**, 271–273. (doi:10.1111/jai.12040)