Molecular Phylogenetics and Evolution 60 (2011) 305–316

Contents lists available at ScienceDirect

Molecular Phylogenetics and Evolution

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

Multilocus analyses of an Antarctic fish species flock (Teleostei, Notothenioidei, Trematominae): Phylogenetic approach and test of the early-radiation event

Karel Janko [a,*], Craig Marshall [b], Zuzana Musilová [a,1], Jeroen Van Houdt [c], Arnaud Couloux [d], Corinne Cruaud [d], Guillaume Lecointre [e]

[a] Laboratory of Fish Genetics, Institute of Animal Physiology and Genetics of the Academy of Sciences of Czech Republic, Rumburská 89, 277 21 Liběchov, Czech Republic
[b] Department of Biochemistry and Genetics Otago, University of Otago, P.O. Box 56, Dunedin, New Zealand
[c] Katholieke Universiteit Leuven, Laboratory of Animal Diversity and Systematics, Ch. Deberiotstraat 32, B-3000 Leuven, Belgium
[d] Genoscope, Centre National de Sequençage. 2, rue Gaston Crémieux, CP5706, 91057 Evry Cedex, France
[e] UMR 7138 CNRS-UPMC-MNHN-IRD, Département Systématique et Evolution, Muséum National d'Histoire Naturelle, CP39, 57 rue Cuvier, 75005 Paris cedex 05, France

ARTICLE INFO

Article history:
Received 20 July 2010
Revised 22 January 2011
Accepted 7 March 2011
Available online 12 March 2011

Keywords:
Species tree versus gene tree
Multilocus phylogeny
Diversification rate
Evolutionary radiation
Antarctic fish

ABSTRACT

Clades that have undergone episodes of rapid cladogenesis are challenging from a phylogenetic point of view. They are generally characterised by short or missing internal branches in phylogenetic trees and by conflicting topologies among individual gene trees. This may be the case of the subfamily Trematominae, a group of marine teleosts of coastal Antarctic waters, which is considered to have passed through a period of rapid diversification. Despite much phylogenetic attention, the relationships among Trematominae species remain unclear. In contrast to previous studies that were mostly based on concatenated datasets of mitochondrial and/or single nuclear loci, we applied various single-locus and multilocus phylogenetic approaches to sequences from 11 loci (eight nuclear) and we also used several methods to assess the hypothesis of a radiation event in Trematominae evolution. Diversification rate analyses support the hypothesis of a period of rapid diversification during Trematominae history and only a few nodes in the hypothetical species tree were consistently resolved with various phylogenetic methods. We detected significant discrepancies among trees from individual genes of these species, most probably resulting from incomplete lineage sorting, suggesting that concatenation of loci is not the most appropriate way to investigate Trematominae species interrelationships. These data also provide information about the possible effects of historic climate changes on the diversification rate of this group of fish.

1. Introduction

Evolutionary radiations, lineages that have undergone a phase of rapid cladogeneses, are often characterised in phylogenetic reconstruction by "soft polytomies" (Maddison, 1989): putative speciation events that occur very rapidly in the past genealogical tree and lead either to short or absent internal branches in the reconstructed phylogenetic trees (i.e. unresolved phylogenetic relationships) or to conflicting topologies among gene trees when using multiple loci. Such topological conflicts among gene trees are generally not supported statistically because the time elapsed since these cladogeneses is longer than the time between any two cladogenetic events and the time during which derived states can be erased by multiple substitutions is greater than the time during which these derived states accumulated. For clusters of recent cladogenetic events the lack of congruence among gene trees can also be due to the failure of individual gene genealogies to coalesce along internal branches leading to incomplete lineage sorting. Phylogenetic studies often attribute a finding of conflicting genealogies and/or a lack of phylogenetic resolution coupled with short internal branches to rapid diversification in the taxon of interest and a high rate of incomplete lineage sorting (e.g. Lopez-Fernandez et al., 2005; Belfiore et al., 2008). However, exact tests for periods of increased rates of diversification are usually absent from such studies.

The Trematominae, a tribe of Notothenioid teleosts from Antarctica widely used as a model of cold adaptation, fall into this category as it has a problematic phylogeny. The species flock is strikingly diverse in morphology and has evolved a variety of ecological forms including species of benthic, pelagic or cryopelagic habits. Ruber and Zardoya (2005) applied a γ statistic (Pybus and Harvey, 2000) to published phylogenetic data and tested for evidences of periods with increased diversification rates in several groups of marine fish. They showed that Trematominae

* Corresponding author.
E-mail address: janko@iapg.cas.cz (K. Janko).
URL: http://www.genoscope.fr (C. Cruaud).
[1] Present address: Department of Zoology, Faculty of Science, Charles University in Prague, Vinicna 7, 128 44 Praha, Czech Republic.

1055-7903/$ - see front matter © 2011 Published by Elsevier Inc.
doi:10.1016/j.ympev.2011.03.008

phylogenies deviate significantly from rate constancy and found negative γ values that support a scenario of early radiation in Trematominae history.

The Trematominae contains 15 described species in three genera; *Trematomus*, *Pagothenia* and *Cryothenia*. The latter two genera are embedded within the genus *Trematomus* (Sanchez et al., 2007; Kuhn and Near, 2009). Several studies have focused on the phylogeny of Trematominae using morphological (Balushkin, 2000; Klingenberg and Ekau, 1996; Ekau, 1991) and molecular approaches (Ritchie et al., 1996; Near, 2004; Sanchez et al., 2007; Near and Cheng, 2008; Kuhn and Near, 2009). Whereas initial molecular studies focused on one or a few mitochondrial genes that led to many unresolved relationships, the latter three studies increased the resolution power by concatenating the data from several mtDNA loci and one nuclear locus resulting in an almost fully resolved phylogenetic tree (especially Kuhn and Near (2009), which used data derived from 13 of 15 known species). Both morphological and molecular work suggested that pelagic species evolved from a benthic ancestor on at least two occasions. This suggests a high evolutionary versatility in the Trematominae. From morphological data, Balushkin (2000) postulated the existence of a cryopelagic clade (including two species foraging under or within the platelet ice; *T. newnesi* and *P. borchgrevinki*), which was recently corroborated by Kuhn and Near (2009). In contrast, Ritchie et al. (1996) and Sanchez et al. (2007) inferred that *T. scotti* and *T. newnesi* are basal species and concluded that the Trematominae crown-group diversified after their split from the common ancestor. Using standard mtDNA molecular clocks, Ritchie et al. (1996) suggested that average percentage sequence divergence among the crown-group species corresponds to 3.4 Mya, and proposed that the main Trematominae radiation coincided with deglaciation of coastal waters during the Pliocene warming. However, Near (2004) used fossil-based calibration and penalised likelihood to create an ultrametric tree and showed that the Trematominae as a whole are much older (mean estimate = 9.4 Mya), and that the most recent common ancestor of the crown-group (which should include *T. newnesi*; see Kuhn and Near, 2009) is as old as 7.4 Mya.

Despite the apparent resolution of Trematominae relationships in Kuhn and Near (2009), the signal of rapid diversification early in the history of Trematominae suggests that reservations about concatenating the data from unlinked loci apply to this group and one must be aware of possible conflicts among gene trees. Although increasing the number of loci and characters that are put into a single concatenated data set can improve the phylogenetic resolution, it may also lead to false estimates of species trees (Edwards et al., 2007), especially when the lengths of the internal branches are short and ancestral population sizes large. This is because concatenation assumes the same phylogenetic history for all sites in the dataset (Belfiore et al., 2008). Such an assumption can easily be violated especially in the case of incomplete lineage sorting (Takahata, 1989; Felsenstein, 2004; Knowles and Chan, 2008).

The primary aim of this study is to investigate potential phylogenetic conflicts among unlinked loci as well as to reconstruct a Trematominae phylogeny explicitly taking into account the possible effects of stochastic lineage sorting. We achieve this by inspecting conflicts among individual gene trees as well as by applying Bayesian estimation of species trees (BEST; Liu, 2008) that allows for stochastic differences of topology of individual gene trees resulting from lack of gene lineage coalescence between speciation events. BEST is particularly suitable for data where incongruence in the phylogenetic signal of individual genes is expected. It is claimed to perform better and produce fewer artefactual topologies than traditional concatenation (Edwards et al., 2007).

The second aim of the study is to investigate the character of the putative signal of changes in diversification rate in all loci. This is important since the γ statistic applied previously by Ruber and Zardoya, although correctly depicting the general signal of a decreasing diversification rate, does not assess whether this reduction in rate fits to a gradual or an abrupt decrease of speciation tempo. Furthermore, similar γ values may be generated by a simple decrease speciation rate or by a combination of decreasing speciation rate and increasing extinction rates (Rabosky and Lovette, 2008; Quental and Marshall, 2009). We used several methods including non-parametric approaches that allow testing of complex scenarios.

Previous works proposed particular extrinsic factors that could explain the origin of Antarctic *Notothenioidei* in general (rev. in Eastman and Clarke, 1998) and Trematominae in particular (see Ritchie et al. (1996) for hypothesised link with Pliocene warming). Pleistocene climatic oscillations might have also strongly affected the rate of diversification either acting as a diversity pump (Avise, 2000), or inducing mass extinctions, particularly of benthic species (Thatje et al., 2005; Janko et al., 2007). Therefore, our final aim is to compare observed patterns of diversification with the timing of several known extrinsic factors or events to better understand Trematominae radiation.

2. Material and methods

2.1. Taxon sampling

The Trematominae contains two genera with 15 described species; 11 species of the genus *Trematomus* (one missing in our study – *T. tokarevi*), two species of the genus *Pagothenia* (one missing – *P. brachysoma*) and two of *Cryothenia* (*one* missing – *C. peninsulae*). The 12 species used in this study are listed in Table 1. We used several outgroup species; namely *Chionodraco hamatus*, *Notothenia rossii*, *Patagonotothen tessellata* and *Lepidonotothen squamifrons*.

Table 1
Specimens of Trematominae used for sequencing in this study.

Species	Specimen number	Geographical location
T. scottii	1371	Weddell Sea
	867	Weddell Sea
T. bernacchii	TNB1-5	Terra Nova Bay
	CH 629	Cape Hallett
P. borchgrevinki	Psp. B 5	Terre Adelie
	CA 113b	Cape Armitage
T. nicolai	CR 296	Cape Roberts
	1369	Weddell Sea
	CA436	Cape Armitage
T. newnesi	Tn399TA6	Terre Adelie
	Tn401TA2	Terre Adelie
	CA450	Cape Armitage
T. hansoni	1216	South Georgia
	CH1026	Cape Hallett
T. lepidorhinus	1368	Weddell Sea
	565	Terra Nova Bay
T. eulepidotus	759	Terra Nova Bay
	1370	Weddell Sea
T. loennbergii	139	Terre Adelie
	426	Terre Adelie
T. vicarius	1214	South Georgia
	256	South Georgia
T. pennellii	CH610	Cape Hallett
	CR321	Cape Roberts
C. amphitreta	USNM 385901	McMurdo

2.2. Molecular methods

Genomic DNA was extracted from a fin clip using Dneasy Tissue Kits (Qiagen), following the manufacturer's protocol. Primers used in this study are shown in Table 2. The PCR reaction volume (25 μL) contained PPP master mix (Top-Bio) including $MgCl_2$ (2.5 μM), both forward and reverse primers (0.4 μM) and 1 μL of DNA sample. The reaction scheme for all fragments consisted of an initial denaturation step of 94 °C (2 min), followed by 35–40 cycles of denaturation at 94 °C for 1 min, annealing for 1 min at 46–55 °C (depending on the fragment and species) and extension at 72 °C for 1 min. The terminal extension was at 72 °C for 8 min. The reactions were performed on a thermo-cycler PTC-200 (MJ Research). PCR products were purified using the QIAquick PCR Purification Kit (Qiagen) and Macherey–Nagel Kit and directly used as a template for a sequencing reaction using the ABI PRISM Big-Dye™ Terminator v3.1 and v1.1 Ready Reaction Cycle Sequencing Kits (Applied Biosystems). Sequences were cleaned by DyeEX and than were read on a 3130 Genetic Analyser (Applied Biosystems) automatic sequencer.

We sequenced two or three individuals per species in our analysis, except for a single specimen of *C. amphitreta*. Sequences were obtained for eight nuclear markers (introns in Gnrh, Irbp and S7 genes, internal transcribed spacer 1 and 2 (ITS1, ITS2), three anonymous nuclear loci T1, T5, T16), and four mitochondrial genes (cytochrome b, NADH2, 12s and 16 ribosomal genes) from all the available species of Trematominae (see Table 2 for primers). We successfully sequenced all species in all loci except for *C. amphitreta*, which did not amplify either the Irbp intron or ITS2 locus. For the final analyses, we have chosen a single specimen per species, which were homozygous in all markers or heterozygous in no more than a single position in each locus to avoid problems with phase distinction. This is also because analyses of diversification rates require a single tip per species in the tree.

Sequences determined in this work were deposited with Genbank and have the accession numbers.

2.3. Phylogenetic reconstructions of individual gene trees and tests of topology

Sequences were aligned using the program ClustalW with the default settings as implemented in the BioEdit software package (Hall, 1999) and were then manually checked and edited. The amino acid translation of the coding sequences was examined for stop codons. All nuclear loci were treated separately but since mtDNA is effectively one locus, we concatenated all mtDNA markers into a single haplotype for each individual. We used Akaike information criterion (Posada and Crandall, 1998) to select the models that best fit the analysed sequence partitions using Modeltest version 3.7 for the Maximum likelihood (ML) analysis and MrModeltest (Nylander et al., 2004) for analyses with MrBayes. We evaluated the most likely model separately for each of the three codon positions in coding genes. ML analyses were conducted in PAUP* 4b10 (Swofford, 2003 with bootstrap analysis of 500 pseudoreplications. Bayesian analyses were performed in MrBayes, version 3.1 (Ronquist and Huelsenbeck, 2003) and model parameter values were "unlinked" among partitions with the same assumed molecular evolutionary model (Ronquist and Huelsenbeck, 2003). The analyses were performed with 10^7 generations, sampling every 100 trees and using two parallel runs with four chains. To assess the convergence of model parameters and tree topology, we examined the relationship between the likelihood score and the number of generations and also inspected whether the standard deviation of split frequencies between the two runs was less than 0.05. Generations before reaching convergence were discarded as the burn-in.

To test predefined phylogenetic hypotheses (Fig. 1) we used constrained ML trees provided by PAUP and tested whether the ML trees for these hypotheses are part of the confidence set of trees applying the approximately unbiased (AU) test (Shimodaira, 2002). Namely, we tested two hypotheses of basal species in the trematomin phylogeny, i.e. (1) *T. scotti* as the sister group of all trematomins (as proposed by Near (2004)), and (2) *T. scotti* plus *T. newnesi* as the most external lineages of the remaining trematomins

Table 2
Primers used for amplification of analysed loci.

Gene		Primer		Source
		F	R	
GNRH	Gonadotrophin releasing hormone = intron in protein coding nuclear gene	AATGCACCACATGCTAACAAGGC	CGCACCATCACTCTGCTGTTCGC	Hassan et al. (2002)
IRBP	Interphotoreceptor retinoid-binding protein gene 2 (module 1) including part of exon 1, intron A and part of exon 2	TGTGGAGCACTTTGTTCCAGAC	CTCCATCAAAGAANTAAGAGCAGAA	Dettai and Lecointre (2008)
ITS1	Internal transcribed spacer 1 = ribosomal nuclear gene	AAAAAGCTTCCGTAGGTGAACCTGCG	AGCTTGCTGCGTTCTTCATCGA	Presa et al. (2002)
		CTTGACTATCTAGAGGAAGT		
ITS2	Internal transcribed spacer 2 = ribosomal nuclear gene	CTACGCCTGTCTGAGTGTC	ATATGCTTAAATTCAGCGGG	Presa et al. (2002)
S7_i	First intron in ribosomal nuclear protein coding gene	TGGCCTCTTCCTTGGCCGTC	AAC TCG TCT GGC TTT TCG CC	Chow and Hazama (1998)
T1	Anonymous locus	AGAGAGCTTTGCCGCCTTGCTGGA	CAGAGGCCTGCTGTGGTTACGAGC	This study
T5	Anonymous locus	TGGATGAAGCGACCGAGCTTCTG	GCAAAGAACAGAAACGTTCTCTC	This study
T16	Anonymous locus	TCATCGGTACAGATGCAGCGGTGC	TGTGTGAATCAGGTAGTCAGTCCA	This study
CYTB	Cytochrome b = protein coding mitochondrial gene	TGAGGkGGrTTTTCGGTAGATA	GATrTAnGGrTCCTCAaCGGG	Janko et al. (2007)
			GGCATnCCrCCAATTCAGGT	
NADH2	NADH dehydrogenase subunit 2 = protein coding mitochondrial gene	CTACCTGAAGAGATCAAAAC	CGCGTTTAGCTGTTAACTAA	Kocher et al. (1995)
12S	12S ribosomal subunit	AA ACT GGG ATT AGA TAC CCC ACT AT	TGA CTG CAG AGG GTG ACG GGC GGT GTG T	Kocher et al. (1989)
16S	16S ribosomal subunit	AGC ATC TCC CTT ACA CTG AGA AGT	GTT AAG GAG AGG ACT TGA ACC TCT	Near et al. (2003)

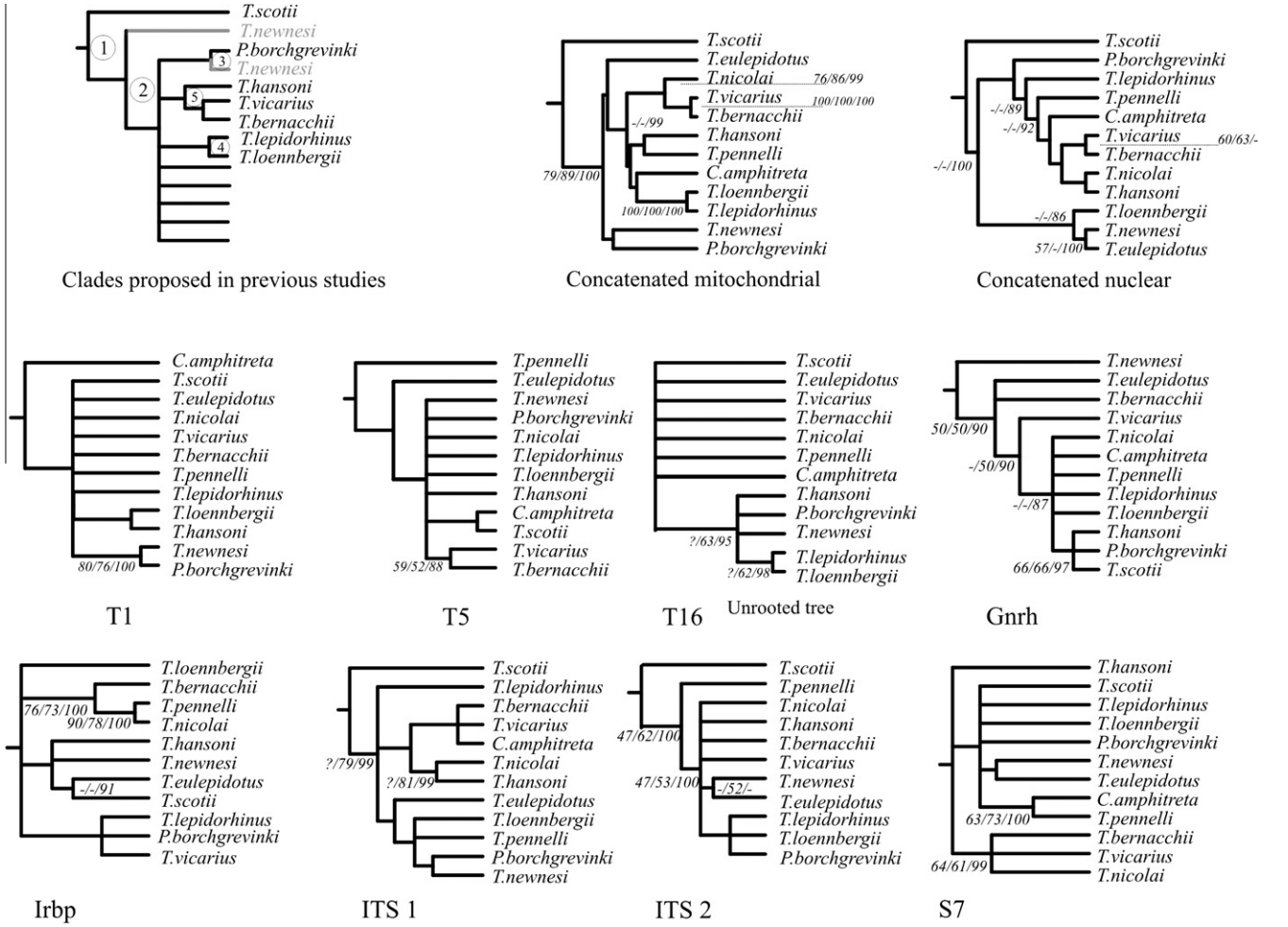

Fig. 1. Phylograms for Trematominae fish. The upper left phylogram denotes the clades, which were proposed in previous work and are tested by the AU test in this study (numerals 1–5 refer to clade description in the text). Note that *T. newnesi* is indicated twice, since its position was proposed either as basal, or as a member of the cryopelagic clade. The remaining phylograms represent ML trees with branch lengths corrected by PL penalised likelihood from analyses of individual loci as well as concatenated datasets. Outgroups were removed from Maximum Likelihood trees before entering the PL analysis. The numbers in nodes represent the node support in the following order: bootstrap support from Maximum parsimony, Maximum Likelihood and Bayesian posterior probabilities, '–' indicates that the given clade was not observed in that analysis. '?' marks supports positions where the bootstrap analysis for Maximum parsimony had not completed and was unavailable. Only support values higher than 50% for bootstrap or 0.80 for posterior probabilities are shown.

(proposed by Ritchie et al. (1996) and corroborated by Sanchez et al. (2007)). (3) We also tested for signal in support of the (*P. borchgrevinki*, *T. newnesi*) cryopelagic clade (Balushkin, 2000). Further, we tested two clades that were proposed in previous studies, i.e. (*T. lepidorhinus*, *T. loennbergii*) (4) and ((*T. bernacchii*, *T. vicarius*), *T. hansoni*) (5). Finally (6), using the same rationale, we tested the alternative topology of (*T. nicolai* (*T. bernacchii*, *T. vicarius*)), which was found in some genes during this study.

2.4. Phylogenetic reconstructions based on multilocus dataset

2.4.1. Concatenation

We performed phylogenetic analyses for two concatenated data sets including all nuclear, and all markers together. The data were divided into respective partitions including one from each of the three codon positions of the coding sequences of cyt b and ND2 genes, as well as single partitions from individual nuclear genes. The concatenated data sets were analysed with 100 million generations with sampling every 1000 trees. Each analysis was replicated to ensure that convergence was reached. We used the AWTY software on-line (Nylander et al., 2008) with the default settings to monitor the stability of posterior probabilities for individual clades during the analysis. Based on AWTY results, the burn-in for Bayesian trees was set to 25% in both nuclear and complete concatenates.

2.4.2. Filtered Supernetwork construction

In order to evaluate how much of the phylogenetic noise (resulting from multiple substitutions, convergences, or incomplete lineage sorting, among other processes) is present in the data, all single-locus trees resulting from Maximum Likelihood analysis were entered into SplitsTree 4.5 (Huson and Bryant, 2006) except for the T16 locus for which we did not have an outgroup sequence. We applied the homoplasy filter (Holland et al., 2008), which treats each split in a gene tree as a binary character, and examines the level of homoplasy by evaluating the number of character state changes required to explain the observed character states in individual gene trees – referred to as a distortion score. It requires two user-defined parameters, MinNumberTrees (MNT) and MaxDistortionScore (MDS). Splits that have a distortion score greater than MDS in more than MNT are filtered out from the analysis.

2.4.3. Analysis using Bayesian estimation of species tree

BEST software (Liu, 2008) generates the most probable gene trees for each individual gene and also gives the set of possible spe-

cies trees for the analysis. The data set of all loci was analysed in BEST by two parallel runs with four chains for 100 million generations and sampled every 1000 trees, treating all mtDNA sequences as single locus. We used independent gamma distributions as the prior of θ, setting the effective population sizes of uniparentally inherited and haploid mtDNA loci as one fourth that of autosomal markers (following Liu and Pearl, 2007). The trees then obtained were summarised in MrBayes software using the 'sumt' command. We also analysed the stability of posterior probabilities for individual clades during the two independent runs of BEST using the AWTY software on-line (Wilgenbusch et al. (2004) with the default settings). Based on AWTY results, the burn-in for BEST and Bayesian trees was set for 55%.

2.5. Estimates of diversification rates

We evaluated putative changes in diversification rates in all loci separately (treating mtDNA as single locus). An ultrametric tree for subsequent analyses of diversification rate was obtained by the Penalised Likelihood method (PL; Sanderson, 2002) as implemented in r8s software (Sanderson, 2003), which involves a roughness penalty for the autocorrelation between the substitution rates of ancestral and descendant branches of the tree. The ML trees were entered into PL, but before doing so, we extracted all outgroup taxa to avoid misleading artefacts during divergence-times estimation (Paradis, 2006). We subsequently evaluated the history of diversification in each locus with a number of methods that test for the departures from a constant diversification rate. They all are implemented either in Ape (Paradis, 2006) or Laser (Rabosky, 2006a): packages for the R programming environment (see below). The reason for using multiple methods is that each of them relies on different model assumptions, which may be violated by real data, and we consider an outcome as conclusive if it is corroborated by the different methods applied here. In each case, the time was scaled according to the first branching in *Trematomus* gene trees, setting the value for the first node to −1 with respect to the present time, which is referred to as 0 (Rabosky, 2006a). The advantage of this approach is that it does not rely on absolute dating of nodes, which is difficult in Notothenioids in the absence of a good fossil record (see Near (2004) for discussion of this issue).

We used the goodness-of-fit (GOF) approach by Paradis (1998) using two complementary tests to compare the distribution of the branching times with the theoretical expectation of constant diversification, i.e. Cramér-von Mises test (Stephens, 1974) and Anderson–Darling test (Stephens, 1982), which gives more weight to the tail of the distribution. It has been suggested that rejection of the null model with the latter test suggests an abrupt change in a diversification rate of the lineage, while rejection by both tests implies a gradual change of diversification rate (Paradis, 1998).

The logarithm of the number of reconstructed lineages was plotted against the branch length distance from the root on the ultrametric trees (Fig. 4, Lineage Through Time – LTT; Nee et al., 1992). If the diversification rate is constant, a straight line is expected. If the diversification rate is decelerating, the nodes would appear closer to the root (their plot would be concave), while an accelerating speciation rate would result in a convex curve. LTT also proved useful to detect the signatures of mass extinctions, which are manifested by antisigmoidal shape (Crisp and Cook, 2009). However, this method also has serious limitations. For example in LTT, a concave pattern may result from incomplete taxon-sampling (note that our study omits three of 15 known species), whereas a convex pattern can result from background extinctions, which are more likely to eliminate older lineages i.e. the 'pull towards the present' (Nee et al., 1994). To mitigate these potential problems, we applied several additional methods that can directly evaluate the effect of incomplete taxon-sampling or even varying extinction-rate over time. We first describe the rationale of such methods and subsequently explain how we dealt with these confounding issues.

The γ statistic of Pybus and Harvey (2000) was calculated from the relative positions of intervals between cladogenetic events for each gene tree. Under the pure birth model and complete sampling of the reconstructed phylogeny, this test follows a normal distribution. Negative γ values indicate the prevalence of nodes closer to the root and suggest a decreasing tempo of diversification in time, whereas positive γ values suggest either an increasing tempo or a non-zero extinction rate.

The survivorship model of Paradis (1997) was used to assess how well temporal shifts in diversification rate fit the data for each locus in three models of lineage evolution. Model A assumes a constant net rate of diversification through time ($\delta = \lambda - \mu$; where λ is speciation rate and μ extinction rate). Model B assumes a gradual change of δ following a Weibull distribution with a shape parameter β. Values of $\beta > 1$ indicate that δ decreased through time, while $\beta < 1$ suggest an increasing diversification rate. The last model, C, assumes that diversification rate changed abruptly from δ_1 to δ_2 at a time T_c that was varied in 50 intervals between −1 and 0 and the highest likelihood determined. Model A is a special case of both models B and C and therefore LRT was used to assess the superiority of the latter models to model A. Akaike information criterion (AIC; Akaike, 1973) was used to select the better of models B and C, which are not nested.

Rabosky's (2006b) method is particularly suitable to test for deviations from the null model under a non-zero background extinction rate and has been shown to perform better than other statistics especially when the diversification rate changed abruptly at some time. We examined two rate-constant (pure-birth and birth–death (bd)), and three rate-variable (DDX, DDL, Yule2rate) models. The pure-birth model assumes no extinction rate, while the bd model assumes a constant extinction/speciation ratio. DDX fits exponential, and DDL logistic variants of the density-dependent speciation rate model to the branching times. The Yule2rate model assumes that the clade diversified under some constant rate until a given period when it switched abruptly to another diversification rate. AIC is used here to select the best-fit model. ΔAIC_{RC} is used as a test statistic evaluating how much more likely the rate-variable models are relative to rate-constant ones ($\Delta AIC_{RC} = AIC_{RC} - AIC_{RV}$; where RC and RV indices refer to rate-constant and rate-variable models respectively) and its significance is assessed by simulation (see below).

Rabosky and Lovette (2008) developed a method to disentangle the impact of changing extinction rate on observed signal of diversification rate decrease. It fits three diversification models to the data: SPVAR (exponentially decreasing speciation rate, constant extinction); EXVAR (exponentially increasing extinction rate, constant speciation); and BOTHVAR (both speciation and extinction varied through time). Again, we could not use the LRT, since the former two models are not nested, and AIC was applied to select the best fitting model.

2.6. Evaluation of the effect of incomplete taxon-sampling

In our collection, three of 15 described Trematominae species were unavailable and this may affect the estimated parameters by underestimating the number of recent nodes in the phylogeny (Pybus and Harvey, 2000). To account for such effect on the γ distribution, we applied a Monte Carlo simulation and generated 1000 phylogenies of 15 extant species with only 12 sampled tips (11 sampled tips in Irbp and ITS2 loci due to lack of the data from *Cryothenia*) under the pure birth model using Phylogen (Rambaut, 2002). Genie (Pybus and Rambaut, 2002) was used to estimate the corresponding γ values and the upper and lower 2.5 percen-

tiles. Simulated trees were imported into LASER to obtain the null distribution of Rabosky's (2006b) ΔAIC_{RC} values in the absence of missing taxa. To evaluate the potential effect of missing species on the survivorship analysis, we followed the suggestion by Paradis (1997). Briefly, the survivorship analysis in Ape requires a vector of branching times observed in the tree. Some of the branching times may be entered as "censored", which indicates that species descendent from those events are known to have survived since the censoring time, but there is no precise information as to when they diversified from their sister lineages. The estimate of branching time is therefore entered into the analysis as distributed according to Eq. (2) of Paradis (1997). For that purpose, we assumed that both missing species have existed at least 1% of the total time of the existence of Trematominae group.

3. Results and discussion

Sequences from this work were deposited in Genbank with accession numbers JN008753 to JN008863.

The variability in all markers is given in detail in Table 3. In contrast to mtDNA loci, we found relatively low variability in nuclear markers with a small proportion of parsimony-informative sites.

3.1. Extracting phylogenetic information

Generally, the phylogenetic signal in individual markers was flat with low support values and we observed much incongruence in reconstructed relationships among loci (Fig. 1). In particular, we found that some nuclear markers suggest different trees in each of the methods used (MP, ML, Bayesian methods) with low levels of support (bootstrapping or Bayesian posterior probabilities). Although several clades emerged more or less frequently with high support, none of them could be accepted for all of the individual gene trees (see below).

The Filtered Split Network analysis summarises the (in)congruencies among individual gene trees. It clearly demonstrates that although several clades were either present or consistent with more than one gene tree (and allowed the weighting of those splits according to their frequencies), none of them is consistent with every partial gene tree which suggests significant among-locus phylogenetic conflict. This is demonstrated by our finding that low filtering power identified many splits but resulted in unreadable phylogenetic patterns, whereas increasing the strength of the homoplasy filter eliminated incongruent phylogenetic signals introduced by individual gene trees, but at the cost of a reduced number of resolved nodes (Fig. 2). Only a few splits seem to persist with higher strengths of filtering. These were also recovered in phylogenetic analyses based on all genes simultaneously and define the basal position of *T. scotti* and monophyly respectively of the (*P. borchgrevinki*, *T. newnesi*), (*T. loennbergii*, *T. lepidorhinus*) and ((*T. bernacchii*, *T. vicarius*) *T. nicolai*) clades. Whereas previous publications (see Kuhn and Near (2009) for review) generally agree with the first three topologies/clades; the last clade is unexpected since previous papers suggest monophyly for ((*T. bernacchii*, *T. vicarius*) *T. hansoni*) clade.

AU tests are in agreement with the conclusion that even some of the most well-supported clades at some loci are contradicted by phylogenetic relationships at other loci. Of the four clades above, all except the basal position of *T. scotti* are in significant conflict with at least one other locus (Table 4). We found that the previously hypothesised ((*T. bernacchii*, *T. vicarius*) *T. hansoni*) clade was also inconsistent with three loci (Gnrh, Irbp and mitochondrial DNA).

Two phylogenetic approaches to dealing with a dataset comprising all sequenced nucleotide positions led to quite different outcomes. Concatenation of all genes greatly improved the phylogenetic resolution as in Kuhn and Near (2009) and highlighted the presence of several significantly supported clades (Fig. 3). However, as noted above, we did not find the ((*T. bernacchii*, *T. vicarius*) *T. hansoni*) clade of Kuhn and Near (2009). In contrast, the topology of the species tree estimated by BEST was unresolved and polytomies were present in the consensus species tree. BEST identified only two clades with high support, i.e. (*T. bernacchii*, *T. vicarius*) and (*T. nicolai*, *T. pennellii*) (see Fig. 3) whereas the stability of observed posterior probabilities for the other clades in the species tree during BEST analysis was much lower than during Bayesian analysis based on a concatenated data set. On the other hand, although concatenation seemed to resolve most nodes with high support, it showed much worse lnL scores in fitting the data to the defined species tree than did the BEST analysis (lnL around −15,650 in BEST vs. −16,700 in the concatenation data set).

In the case of short intervals between speciation events preventing complete lineage sorting of loci, BEST should detect the polytomy, reflecting contradictory genealogies at individual loci, whereas concatenation could lead to an artefactual topology, however well-supported (Belfiore et al., 2008). This is because the concatenation has one strong assumption, which is relaxed in the BEST: all loci are assumed to have the same evolutionary history and therefore the resulting clades in species tree inferred by concatenation may be affected by an overwhelming phylogenetic

Table 3
Summary of the sequence data for all loci.

Locus	Length of fragment	Constant sites*	Variable sites*	Parsimony informative sites*	CI	RI	AIC
gnrh	381	374	7	3	0.92	0.89	TrN + I
irbp	757	723	34	11	0.9	0.7	TrN + I
its1	515	475	40	10	0.63	0.6	TrN + I
its2	408	388	20	4	0.69	0.62	K81uf + I
s7	619	587	32	3	1	1	F81 + I
T1	621	606	15	2	1	1	F81
T5	346	329	17	2	0.97	0.93	K81uf
T16	271	261	10	2	1	1	TrNef
12s + 16s	918	847	71	15	0.91	0.66	K80 + G
NADH2	1050	768	282	131	0.73	0.49	
1st position	–	–	–	–	–	–	GTR + I
2nd position	–	–	–	–	–	–	GTR + G
3rd position	–	–	–	–	–	–	GTR + G
cytb	525	388	137	81	0.65	0.44	
1st position	–	–	–	–	–	–	K80 + I + G
2nd position	–	–	–	–	–	–	HKY + G
3rd position	–	–	–	–	–	–	HKY + I

Consistency (CI) and retention (RI) indices are indicated as well as the best models selected under Akaike information criterion. Asterisks mark information on data without outgroups. Note also that no outgroup was available for locus T16.

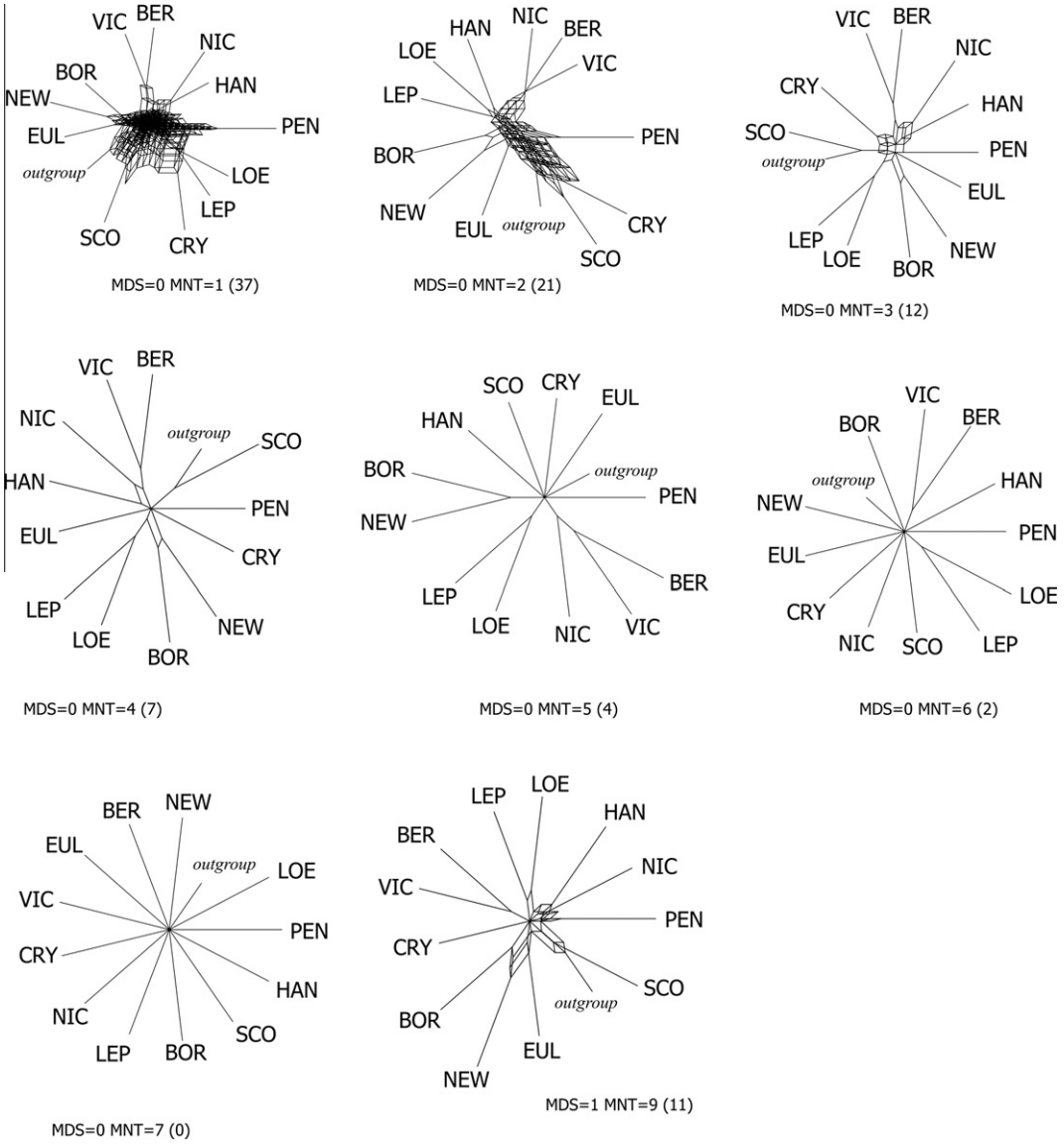

Fig. 2. Homoplasy filter analyses based on individual gene trees from seven nuclear gene trees (T16 excluded for the absence of outgroup) and one mitochondrial tree are shown. For each filter setting, we describe its minimum distortion score (MDS), minimum number of trees consistent with presented splits (MNT) as well as total number of splits in parentheses. Abbreviations: *BER = Trematomus bernacchii*, *EUL = T. eulepidotus*, *HAN = T. hansoni*, *LEP = T. lepidorhinus*, *LOE = T. loennbergii*, *NEW = T. newnesi*, *NIC = T. nicolai*, *BOR = Pagothenia borchgrevinki*, *PEN = T. pennellii*, *SCO = T. scottii*, *VIC = T. vicarius*, *CRY = C. amphitreta*.

Table 4
Results of maximum likelihood analysis of alternative topologies using the AU test.

Topological contraints	gnrh	irbp	its1	its2	S7	T1	T5	T16	mtDNA
No constraint (ML tree; see Fig. 1)	662.46*	1427.09*	1422.5*	773.50*	1450.6*	1154.62*	713.53*	446.5*	10675.39*
(*T. borchgrevinki*, *T. newnesi*)	687.12	1432.91	ML*	773.73	1450.77	ML*	ML*	ML*	ML*
Basal position of *T. scotti*	674.58*	1429.88*	ML*	ML*	1451.04*	1158.38*	715.43*	ML*	ML*
Basal position of *T. scotti* and *T. newnesi*	675.19	1432.81*	1431.18*	777.70	1451.2*	1173.71	713.53*	451.92*	10676.44*
(*T. loennbergii*, *T. lepidorhinus*)	ML*	1431.34*	1427.8*	ML*	ML*	1154.62	ML*	ML*	ML*
((*T. bernacchii*, *T. vicarius*) *T. nicolai*)	667.24	1453.37	1428.76*	773.50*	ML*	ML*	713.53*	ML*	ML*
((*T. bernacchii*, *T. vicarius*) *T. hansoni*)	681.28	1447.52	1428.89*	773.50*	1456.16*	1154.62*	713.53*	446.5*	10707.21

For each tree in respective locus, either ML or constrained, we indicate the corresponding – lnL values. Asterisks indicate that the tree was included in the 0.95 confidence set for the AU test. 'ML' marks the situations where the constrained tree was consistent with the ML tree.

signal from one or a few loci. By definition, concatenation does not take into account the possibility that different genes may have followed different phylogenetic trajectories.

It should be considered that BEST may also report a failure to resolve branches partly because it manages discrepancies emerging from loci with a very low number of informative sites. In such situations the relative weight of a single homoplastic substitution is high, whereas loci of greater variation might provide better resolution. However, the Trematominae phylogenetic tree has been investigated from mitochondrial DNA sequences for some time

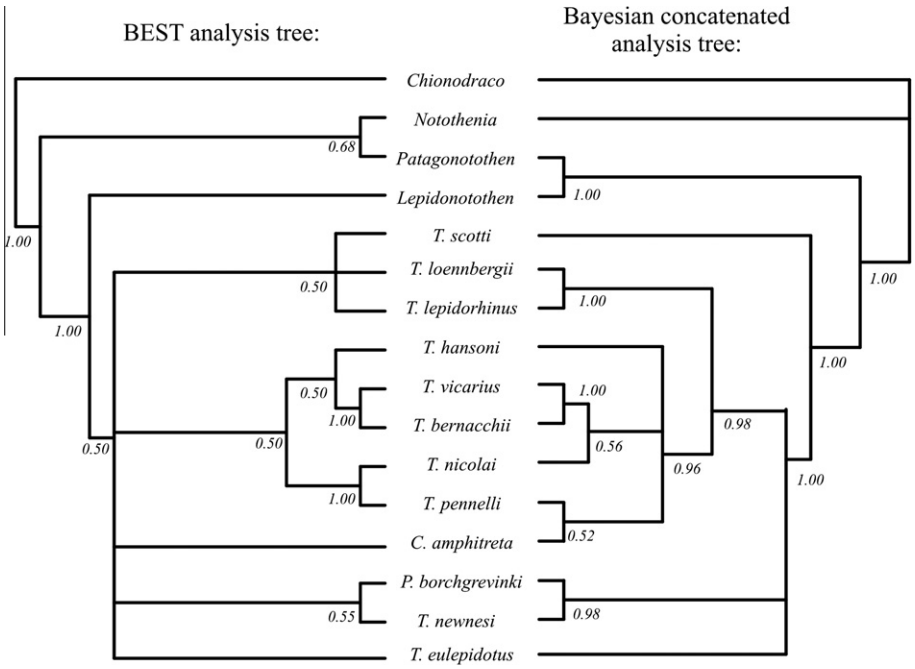

Fig. 3. Bayesian estimated species tree (BEST, left) and Bayesian concatenation tree from all loci (right) from analyses run under the same initial conditions and substitution models.

(Ritchie et al., 1996) and there is a need to find non-mitochondrial gene markers. Highly variable nuclear loci are rare, and this work is a first attempt to fill this gap.

Nonetheless, in concordance with Kuhn and Near (2009) and some other previous studies, the inspection of individual gene trees as well as their comparison by Filtered Supernetworks point to the basal position of *T. scotti* as well as the existence of (*T. lepidorhinus*, *T. loennbergii*) and (*T. newnesi*, *P. borchgrevinki*) clades. We observed contradictory signals for the sister species to the (*T. vicarius*, *T. bernacchii*) clade. Our data raise question about the existence of a ((*T. vicarius*, *T. bernacchii*) *T. hansoni*) clade since the split defining the monophyly of ((*T. vicarius*, *T. bernacchii*) *T. nicolai*) was consistent with most loci and persisted until a high level of filtering was introduced. However, BEST clearly suggested that *T. nicolai* is a sister species of *T. pennellii*. It seems therefore, that identification of most clades largely depends on the locus selected and the phylogenetic method applied. The Trematominae species tree resists resolution despite an increased number of loci. Significant conflict among individual gene trees and resulting low resolution of species tree reconstructed by BEST may either suggest a hard polytomy (in the sense of Maddison, 1989) or a rapid series of sequential cladogenetic events.

3.2. Detection of the signal of a radiation event and the potential effect on phylogenetic analyses

All the analyses of diversification rate reported here suggest that the Trematominae passed through a period of rapid cladogeneses and that the diversification tempo decreased towards the present (consistent with finding of Ruber and Zardoya (2005)). We observed two shapes of LTT. At Irbp, ITS1 and S7, the LTT were concave suggesting initially rapid diversification and a subsequent decrease in speciation tempo (Fig. 4). At other loci (Gnrh, ITS2, T1, T5 and mtDNA) LTT were sigmoid following a straight line for about half of the life of the Trematominae at which time there is a change indicative of sudden increase in the number of lineages. We shall see below that such patterns are largely consistent with the predictions of several applied models. It is notable that these patterns do not fit the pull-towards-present shape typical of high background extinction (Nee et al., 1994). There is no evidence of a mass extinction event since we did not observe the antisigmoidal shape characteristic of such events (Crisp and Cook, 2009).

The results of other statistical analyses are listed in Table 5. A GOF analysis resulted in rejection of the null model of constant diversification rate in all loci under the Cramér–von Mises test and in all nuclear loci under the Anderson–Darling test. Similarly, all of the observed γ values were negative, suggesting decreasing diversification rate towards the present and most of the values were significantly different from zero even after correction for missing species (Table 5).

When taking into account missing species, the survivorship analysis rejected the constant diversification rate (model A) in favour of the model with gradual decrease (model B) at all but two loci (Irbp and mitochondria, where the AIC scores of model B were still lower than of model A). Survivorship analyses also rejected

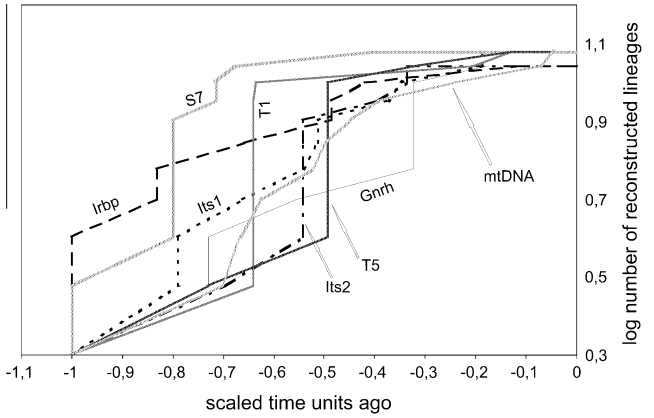

Fig. 4. Semi-logarithmic plot of lineages through time with time scaled relative to oldest branching event for each locus. Note that for Irbp and ITS2 we had data from only 11 species due to the lack of amplification in *C. amphitreta*.

Table 5
Five tests of the null hypothesis of a constant rate of lineage accumulation in Trematominae. Asterisks indicate significant results taking into account missing taxa (see Section 2).

	Survivorship analysis			γ-Statistics	GOF	Rabosky (2006b)	Rabosky and Lovette (2008)					
	Model A LnL (AIC)	Model B LnL (AIC)	Model C LnL (AIC)			Best-fit model (ΔAIC_{RC})	Birth–Death	SPVAR	EXVAR	BOTHVAR	$\Delta\lambda$	$\Delta\mu$
Gnrh	−2.28 (6.55)	0.756* (2.48)	2.902* (−1.80)	−1.44	0.01/0.05	Yule2rate a = 0.32 (2.7)	−51.5 (81.2)	−9.9 (0)	−10.05 (0.2)	−9.95 (1.0)	−0.1	10^{-6}
	δ = 0.22	Beta = 1.94	a = 0.33; δ_1 = 0.6; δ_2 = 5.2									
IRBP	−6.19 (14.38)	−1.87 (7.74)	−2.46* (8.92)	−2.13	0.01/0.05	DDX (15.6*)	−54.7 (85.3)	−10.67 (0)	−13.21 (5)	−10.7 (2.1)	−0.47	10^{-6}
	δ = 0.14	Beta = 2.59	a = 0.74; δ_1 = 0.8; δ_2 = 5.2									
ITS1	−4.148 (10.29)	0.196* (3.609)	a = 0.51; 0.582* (2.837)	−1.99	0.01/0.05	DDL (5.2*)	−55.3 (92.1)	−8.44 (0)	−13.92 (11.7)	−8.44 (1.9)	−0.78	10^{-6}
	δ = 0.18	Beta = 2.35	δ_1 = 0.9; δ_2 = 5.7									
ITS2	−3.98 (9.96)	2.74* (−1.69)	5.63* (−7.19)	−2.57*	0.01/0.01	DDL (11.9*)	−53.1 (74.8)	−10.1 (0)	−11.6 (3)	−10.1 (2)	−0.43	10^{-6}
	δ = 0.18	Beta = 3.04	a = 0.54; δ_1 = 0.6; δ_2 = 11.4									
S7	−8.175 (18.35)	5.4* (−6.79)	6.165* (−8.33)	−4.12*	0.01/0.01	DDL (44.2*)	−55.9 (89.2)	−7.7 (0)	−14.5 (13.9)	−7.71 (2.3)	−0.78	10^{-6}
	δ = 0.13	Beta = 6.04	a = 0.8; δ_1 = 0.5; δ_2 = 17.4									
T1	−5.145 (12.29)	0.59* (2.81)	11.714* (−19.42)	−3.01*	0.01/0.01	Yule2rate a = 0.64 (22.6*)	−51.4 (79.8)	−10.3 (0)	−10.8 (0.9)	−10.28 (2)	−0.23	10^{-6}
	δ = 0.17	Beta = 2.91	a = 0.64; δ_1 = 0.5; δ_2 = 22.3									
T5	−3.642 (9.28)	1.819* (0.36)	9.668* (−15.33)	−2.63*	0.01/0.01	Yule2rate a = 0.49 (12.6*)	−54.9 (90)	−9.23 (0)	−12.37 (6.3)	−9.22 (1.9)	−0.67	10^{-6}
	δ = 0.19	Beta = 2.61	a = 0.5; δ_1 = 0.4; δ_2 = 12.7									
mtDNA	−2.778 (7.55)	−1.19 (6.37)	0.466* (3.067)	−1.24	0.01/–	DDL (1.4)	−52.6 (75.7)	−10.71 (0)	−11.1 (0.8)	−10.7 (2)	−0.19	10^{-6}
	δ = 0.21	Beta = 1.66	a = 0.5; δ_1 = 1.2; δ_2 = 5.8									

For the survivorship analysis, the net speciation rate (δ) of model A, β values in the model B and the three parameters of the model C, i.e. the two diversification rates δ_1, δ_2 and a, the scaled time of the rate shift, are listed. Asterisks indicate the best models according to LRT. Observed p-values in goodness-of-fit statistic are given for both Cramér–von Mises/Anderson–Darling tests, respectively. In the analysis with the method of Rabosky (2006b), we indicate the best selected model as well as the relevant ΔAIC_{RC} value followed by an asterisk, where significant. For the Yule2rate model, we list the value of parameter a, as above. Finally we show the results of fitting constant-rate (birth–death) and variable-rate (SPVAR, EXVAR, BOTHVAR) models to *Trematomus* phylogenies with the Maximum log-likelihoods as well as ΔAIC scores (parentheses), which show the difference between the AIC score of the considered model and the best model. $\Delta\lambda$ and $\Delta\mu$ indicate net change in speciation and extinction rates between the first branching event and the present day under SPVAR and EXVAR models, respectively.

model A in favour of model C assuming abrupt change at all loci without exception. Model C obtained lower AIC scores than model B at all but the Irbp locus and the best estimates of the timing of the diversification rate change varied between −0.33 (Gnrh) and −0.8 (S7) scaled time units, but mostly were close to −0.5 in other loci (see Table 5). Similarly, positive values of ΔAIC_{RC} (Rabosky, 2006b) for all loci indicated the best fit of the rate-variable models for Trematominae. It was significant after taking into account the null distribution with unsampled species at all loci except Gnrh and mtDNA (Fig. 5). Rabosky's (2006b) method selected the model of abrupt decrease in diversification rate in three loci (Gnrh, T1 and T5), which is generally consistent with results of the survivorship analysis.

Rabosky and Lovette's (2008) method chose SPVAR as the best-fitting model at all loci (Table 5), suggesting that an increasing extinction rate did not contribute to the observed explosive early-radiation pattern. This is corroborated by the absence of pull-towards-present pattern in LTT and by fact that the best-fit parameterisation of the EXVAR model suggested almost no variation of the extinction rate through time. We have to stress that the estimates of extinction rate must be taken with a pinch of salt since several processes may affect the results. Variation of diversification rate among lineages tends to inflate the estimated rate of extinction (Rabosky, 2009a), whereas phylogenetic clustering of extinct lineages leads to a signal apparently of low extinction Rabosky (2009b). However, the extent to which both mechanisms may bias our results is questionable since they are expected to have more effect in large phylogenies. The former mechanism is also unlikely to explain near-zero extinction rate estimates from smaller species-level phylogenies (Rabosky, 2009a) such as in this study. Another possible bias may come from a mass extinction event that could leave a pattern consistent with an early burst of speciation (Crisp and Cook, 2009) that is usually characterised by low estimates of background extinction. However, as discussed above, our observed LTT deviated from the pattern expected for a mass extinction.

Statistical evidence for periods of rapid accumulation of cladogenetic events in the history of the Trematominae suggests that incomplete lineage sorting might explain the contrasts observed among individual gene trees. Such incomplete lineage sorting probably has influenced also the inference of the divergence rate itself. This is because the distribution of nodes on individual genetrees does not precisely reflect true speciation times, but

Fig. 5. Histogram distribution of simulated ΔAIC_{RC} values of 1000 pure-birth phylogenies of 15 extant species with just 12 sampled. Arrows point at values observed at respective loci.

rather the splits of genealogical lineages leading to contemporary species/clades. Stochastic variability of coalescence may also explain the differences in timings of diversification rate change among loci. To our knowledge, however, such an effect has not been taken into account in previous analytical studies.

It must be kept in mind however, that several other processes, such as multiple substitutions saturating the phylogenetic signal, and hybridisation might explain the observed polytomies and incongruence. Saturation of the phylogenetic signal was rejected for mtDNA variability (which was the most variable marker in the present analysis) in a broader taxonomic sampling context (Chen et al., 1998; Sanchez et al., 2007) and relatively high values of consistency and retention indices (Table 3) negate any likely role for homoplasy. On the other hand, we observed low sequence variability and a low proportion of parsimony informative sites among variable sites. These findings suggest that most of the variability is restricted to terminal branches of individual gene trees especially in nuclear markers. In this context, some conflicting clades among loci could be supported by shared nucleotides gained by chance independently within each of them (convergent homoplasies having a strong impact on the topologies because of the low numbers of parsimony-informative sites). This is consistent with what could be expected from a "hard polytomy", i.e. according to Maddison (1989) "an event of multiple speciation, in which several species arose independently from a wide ranging, more or less uniform parent species, and thus any similarities shared by only some of these daughter species are to be accounted for by convergence". However it is still difficult to distinguish the result of a hard polytomy from the simple fact that the available nuclear genes are too conserved to answer the question at hand. In other words nuclear genes are too conserved here to allow the claim for a "hard polytomy" (no nodes at all in the genealogical tree), which is an extreme hypothesis. It is more prudent (i.e. conservative) to provisionally keep the null hypothesis (or general model) of a very rapid series of cladogenetic events.

As to introgressive hybridisation, we found no evidence for hybrids either in the phylogeographic study of four trematomin species using one mitochondrial and one nuclear marker (Janko et al., 2007), or in the microsatellite analysis of six Trematominae species (Van de Putte et al., 2009). The data presented here are also inconsistent with introgressive hybridisation, since we observed no reticulations in networks when the homoplasy filter was set according to the HF2 filter setting of Holland et al. (2008) (Fig. 3). This level of filtering keeps only splits with no homoplasy consistent with 75% or more of the gene trees and was suggested as sufficiently powerful to disentangle stochastic lineage sorting and hybridisation. In fact, no reticulation was observed even under the more relaxed HF3 setting (keeping only splits with no homoplasy consistent with 50% or more of the genetrees). However, these are only indicative findings and we plan a study specifically aimed to test for the presence of introgressive hybridisation among Trematominae using highly variable markers such as microsatellites on large population samples for all available species.

In any case, given that the observed data indicate a soft polytomy, there is no clearcut evidence as to whether such a soft polytomy is indicative of a rapid sequence of separate speciation events or whether it implies a hard polytomy (in the sense of Maddison, 1989) in Trematominae. As the second hypothesis cannot be positively documented because of the lack of variability in the nuclear genes, we provisionally keep the first one as the null model. Our study reinforces the importance of locus-by-locus phylogenetic inference and the usage of coalescent-based methods in order to increase the chance of avoiding falsely positive phylogenetic signals.

3.3. Insights into the evolution of the Trematominae species flock

Trematominae represent a geographically isolated monophyletic group with a sympatric distribution of most species, variable morphology and a diverse array of ecological adaptations. We do not intend to speculate whether it fits to any definition of adaptive radiation, given the arguments raised by Olson and Arroyo-Santos (2009) against such classifications, but this group meets the definition of a rapid diversification *sensu* Ribbink (1984) characterised by a slowing of speciation tempo as niches fill (Coyne and Orr, 2004). Although the causes of rapid speciations such as those seen in the Trematominae may be entirely intrinsic, it is of interest to discuss possible extrinsic stimuli since the first molecular study of this group (Ritchie et al., 1996) explicitly mentions the potential role of climate changes. Linking the radiation to known climatic events by absolute dating is difficult given the problematic application of molecular dating to Antarctic fish and the scarcity of a notothenioid fossil record (see Near (2004) for discussion of this issue). However, diversification-rate analyses provide useful qualitative tool to evaluate two candidate external stimuli: Pleistocene glaciation cycles (Thatje et al., 2005), and Pliocene warming (Ritchie et al., 1996).

Populations of Antarctic shelf organisms have been dramatically affected during the Pleistocene glaciations when the shelf habitat was repeatedly reduced by advancing glaciers (Thatje et al., 2005). We have documented significant bottlenecks in ancestral populations of contemporary benthic fish species, at least some of which survived in fragmented refugia (Janko et al., 2007). Pleistocene might have also affected the local biodiversity of shelf species, either through mass extinctions due to habitat deterioration or through an increased speciation rate due to allopatric fragmentation of ancestral taxa in isolated refugia. Both processes are expected to leave traces in the Trematominae phylogeny. There is a consensus that the origin of this tribe is pre-Pleistocene (see above) and therefore we should have observed a signal of increasing extinction rate if there were mass extinctions in the Pleistocene. In contrast, if diversification increased during the Pleistocene, we should expect an increase in the tempo of diversification with time. We found neither of these signals in our data which support instead a decrease in the tempo of diversification with time, thereby ruling out an increased speciation rate during the Pleistocene. Furthermore, our data do not show any evidence of varying extinction rates or a mass extinction event from which it may be concluded, counter-intuitively, that the Pleistocene did not notably affect the distributions of nodes in Trematominae phylogeny. It seems that although the fast and significant changes of Antarctic climate during ice ages greatly affected the demographic histories of local fish species (Janko et al., 2007), the Pleistocene did not affect their diversity to a large extent. Such an interpretation must be considered with caution due to the reservations noted above about estimating extinction rates without fossils but it is the most thorough analysis for Antarctic fish so far.

Ritchie et al. (1996) proposed that the Pliocene warming induced the radiation of the Trematominae based on standard mtDNA molecular dating. Although problems with application of molecular clocks to Antarctic fish (Near, 2004) impede a robust test of this hypothesis, it is interesting to note that in our data several loci supported a scenario of abrupt change in diversification rate about halfway between the time of the MRCA of contemporary Trematominae and the present (see Table 5). Our results therefore support the hypothesis that the radiation of the Trematominae was linked to some relatively well-defined period of time with little speciation afterwards and are qualitatively consistent with Ritchie et al.'s (1996) proposal.

3.4. Implications for comparative studies of evolutionary radiations

McCune (1997) defined two indices (the speciation rate (SR) and time-for-speciation (TFS)), which may be used to compare the evolutionary rates of various organismal groups (see e.g. McCune, 1997; Near and Benard, 2004). In case of a symmetrical tree topology like that of the Trematominae, these indices have respective formulas: $SR_{ln} = \ln n/t$ and $TFS_{ln} = t(\ln 2) * (\ln n)^{-1}$, where t is the age of the most recent common ancestor of the clade and n is the number of extant species. One general outcome of our analysis is that such comparisons of the diversification rates among species flocks of very different ages are quite misleading because both SR and TFS estimators assume constant diversification rate.

We applied these estimators to the Trematominae using the fossil-based dating of Near (2004) and found that $SR_{ln} = 0.276$ and $TFS_{ln} = 2.508$. This is slow in comparison to other documented adaptive radiations (McCune, 1997), and the slowest among fish radiations including allopatrically-speciated groups (Near and Benard, 2004). However, we have seen above that diversification rate was decreasing with time in Trematominae. Under model B of Paradis (1997) for example, one may estimate the initial values of diversification rate under the Weibull distribution of the rate of change of this parameter (diversification rate $\delta = \alpha\beta(\alpha t)^{(\beta-1)}$, where t is time and α and β are the two parameters of the Paradis (1997) model). Estimated instantaneous rate at the beginning of the radiation depended on examined locus and varied between 0.26 and 1.75 speciations per Mya. It appears that the divergence rates were greatly underestimated with constant-rate estimators ($SR_{ln} = 0.276$) as compared to an instantaneous rate at the beginning of the radiation. This suggests that comparisons of species flocks of different ages (e.g. McCune, 1997; Near and Benard, 2004) may result in somewhat different estimates for ancient versus recent radiations, even if their dynamics were identical. Essentially different groups were 'caught' in different phases of their evolution and their comparisons based on indices assuming constant diversification rates may be strongly biased.

Funding

This work was supported by the Grant Agency of Academy of Sciences [B60045602 and KJB600450903 to K.J.]; the Synthesys project conducted at the Muséum National d'Histoire Naturelle (Paris, France), part of the European sixth framework program [FR-TAF-325 for K.J.]; the Centre for Biodiversity Research [LC06073-MSMT to Z.M.]; the Institute of Animal Physiology and Genetics receives continuous support from IRP IAPG [No. AV0Z50450515]. This work was also supported by the "service de systématique moléculaire" of the Muséum National d'Histoire Naturelle (UMS 2700, département "Systématique & Evolution") and the ANTFLOCKS project funded by the French national research agency (ANR-07-BLAN-213). It was also supported by the "Consortium National de Recherche en Génomique" through the agreement no. 2005/67 between the Genoscope and the Muséum National d'Histoire Naturelle on the project "Macrophylogeny of life".

Author contribution

KJ, GL, and CM designed the study and were involved in sample collection. JVH identified several new nuclear loci and designed suitable primers for these regions. KJ, CC and AC carried out most of the molecular work. ZM contributed by sequencing of 3 nuclear loci and by standard phylogenetic analyses of individual nuclear markers. KJ performed the final phylogenetic analyses of mtDNA and multilocus datasets as well as calculating diversification rates. The main text was writen by KJ and subsequently ameliorated by CM, GL, ZM and JVH.

Acknowledgments

We thank to Agnes Dettaï and Barbara R Holland for inspiring comments and help with the network reconstructions. We are grateful to Chris C. Cheng for providing us with *Cryothenia* material. We also thank for technical support the French polar institute "Institut Paul Emile Victor" and the IPEV program 281 (ICOTA) and to Alfred Wegener Institute, Bremerhaven, Germany, for supporting the Polarstern Cruise ANT XXIII-8. We thank Antarctica New Zealand and the Latitudinal Gradients Project for supporting the work of K066. We acknowledge with gratitude the support of The National Science Foundation Grant OPP 0132032 to the ICEFISH2004 cruise (Chief Scientist Bill Detrich, Northwestern University).

References

Akaike, H., 1973 Information theory and an extension of the maximum-likelihood principle. In: Petrovm, B.N., Csáki, F. (Eds.), Second Int. Symp. Inform. Theory, Tsahkadzor, Armenia, USSR. Akadémiai Kiadó, Budapest, pp. 267–281.

Avise, J.C., 2000. Phylogeography: The History and Formation of Species. Harvard University Press, Cambridge.

Balushkin, A.V., 2000. Morphology, classification, and evolution of notothenioid fishes of the Southern Ocean (Notothenioidei, Perciformes). J. Ichthyol. 40, 74–109.

Belfiore, N.M., Liu, L., Moritz, C., 2008. Multilocus phylogenetics of a rapid radiation in the Genus Thomomys (Rodentia: Geomyidae). Syst. Biol. 57 (2), 294–310.

Chen, W.-J., Bonillo, C., Lecointre, G., 1998. Phylogeny of the Channichthyidae (Notothenioidei, Teleostei) based on two mitochondrial genes. In: di Prisco, G., Pisano, E., Clarke, A. (Eds.), Fishes of Antarctica: A Biological Overview. Springer-Verlag Italia, Milan, pp. 287–298.

Chow, S., Hazama, K., 1998. Universal primer for S7 ribosomal protein gene intron in fish. Mol. Ecol. 7, 1255–1256.

Coyne, J.A., Orr, H.A., 2004. Speciation. Sinauer Associates, Sunderland, MA.

Crisp, M.D., Cook, L.G., 2009. Explosive radiation or cryptic mass extinction? Interpreting signatures in molecular phylogenies. Evolution 63, 2257–2265.

Dettai, A., Lecointre, G., 2008. New insights into the organization and evolution of vertebrate IRBP genes and utility of IRBP gene sequences for the phylogenetic study of the Acanthomorpha (*Actinopterygii*: Teleostei). Mol Phylogenet Evol 48, 258–269.

Eastman, J.T., Clarke, A., 1998. Radiations of Antarctic and non-Antarctic fish. In: di Prisco, G., Pisano, E., Clarke, A. (Eds.), Fishes of Antarctica. A biological Overview. Springer-Verlag, Milano, pp. 3–26.

Edwards, S.V., Liu, L., Pearl, D.K., 2007. High resolution species trees without concatenation. Proc. Natl. Acad. Sci. USA 104, 5936–5941.

Ekau, W., 1991. Morphological adaptations and mode of life in high Antarctic fish. In: di Prisco, G., Pisano, E., Clarke, A. (Eds.), Fishes of Antarctica. A biological Overview. Springer-Verlag, Milano, pp. 23–39.

Felsenstein, J., 2004. Coalescents and species trees. In: Inferring Phylogenies. Sinauer Associates, Sunderland, Massachusetts, pp. 488–495.

Hall, T.A., 1999. BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. Nucl. Acids. Symp. Ser. 41, 95–98.

Hassan, M., Lemaire, C., Fauvelot, C., Bonhomme, F., 2002. Seventeen new exon-primed intron-crossing polymerase chain reaction amplifiable introns in fish. Mol. Ecol. Notes 2, 334–340.

Holland, B.R., Benthin, S., Lockhart, P.J., Moulton, V., Huber, K.T., 2008. Using supernetworks to distinguish hybridization from lineage-sorting. BMC Evol. Biol. 8, 202.

Huson, D.H., Bryant, D., 2006. Application of phylogenetic networks in evolutionary studies. Mol. Biol. Evol. 23, 254–267 (software available from www.splitstree.org).

Janko, K., Lecointre, G., DeVries, A., Couloux, A., Cruaud, C., Marshall, C., 2007. Did glacial advances during the Pleistocene influence differently the demographic histories of benthic and pelagic Antarctic shelf fishes? Inferences from intraspecific mitochondrial and nuclear DNA sequence diversity. BMC Evol. Biol. 7, 220.

Klingenberg, C.P., Ekau, W., 1996. A combined morphometric and phylogenetic analysis of an ecomorphological trend: pelagization in Antarctic fishes (Perciformes: Nototheniidae). Biol. J. Linn. Soc. 59, 143–177.

Knowles, L.L., Chan, Y.-H., 2008. Resolving species phylogenies of recent evolutionary radiations. Ann. Missouri Bot. Gard. 95, 224–231.

Kocher, T.D., Thomas, W.K., Meyer, A., Edwards, S.V., Pääbo, S., Villablanca, F.X., Wilson, A.C., 1989. Dynamics of mitochondrial DNA evolution in animals: amplification and sequencing with conserved primers. Proc. Natl. Acad. Sci. USA 86, 6196–6200.

Kocher, T.D., Conroy, J.A., McKaye, K.R., Stauffer, J.R., Lockwood, S.F., 1995. Evolution of NADH dehydrogenase subunit 2 in east African cichlid fish. Mol. Phylogenet. Evol. 4, 420–432.

Kuhn, K., Near, T.J., 2009. Phylogeny of *Trematomus* (Notothenioidei: Nototheniidae) inferred from mitochondrial and nuclear gene sequences. Antarctic Sci. 2009, 1–6.

Liu, L., 2008. BEST: Bayesian estimation of species trees under the coalescent model. Bioinformatics 24, 2542–2543.

Liu, L., Pearl, D.K., 2007. Species trees from gene trees: reconstructing Bayesian posterior distributions of a species phylogeny using estimated gene tree distributions. Syst. Biol. 56, 504–514.

Lopez-Fernandez, H., Honeycutt, R.L., Winemiller, K.O., 2005. Molecular phylogeny and evidence for an adaptive radiation of geophagine cichlids from South America (Perciformes: Labroidei). Mol. Phylogenet. Evol. 34, 227–244.

Maddison, W., 1989. Reconstructing character evolution on polytomous cladograms. Cladistics 5, 365–377.

McCune, A.R., 1997. How fast is speciation? Molecular, geological, and phylogenetic evidence from adaptive radiations of fishes. In: Givnish, T.J., Sytsma, K.J. (Eds.), Molecular Evolution and Adaptive Radiation. Cambridge Univ. Press, Cambridge, UK, pp. 585–610.

Near, J.T., 2004. Estimating divergence times of notothenioid fishes using a fossil calibrated molecular clock. Antarctic Sci. 16, 37–44.

Near, J.T., Benard, M.F., 2004. Rapid allopatric speciation in longperch darters (Percidae: Percina). Evolution 58, 2798–2808.

Near, J.T., Cheng, C.-H.C., 2008. Phylogenetics of notothenioid fishes (Teleostei: Acanthomorpha): inferences from mitochondrial and nuclear gene sequences. Mol. Phylogenet. Evol. 47, 832–840.

Near, T.J., Pesavento, J.J., Cheng, C.-H.C., 2003. Mitochondrial DNA, morphology and the phylogenetic relationships of Antarctic icefishes (Notothenioidei: Channichthyidae). Mol. Phylogenet. Evol. 28, 87–98.

Nee, S., Harvey, P.H., Mooers, A.Ø., 1992. Tempo and mode of evolution revealed from molecular phylogenies. Proc. Natl. Acad. Sci. USA 89, 8322–8326.

Nee, S., Holmes, E.C., May, R.M., Harvey, P.H., 1994. Extinction rates can be estimated from molecular phylogenies. Philos.Trans. Roy. Soc. Lond. B 344, 77–82.

Nylander, J., Ronquist, F., Huelsenbeck, J., Nieves-Aldrey, J., 2004. Bayesian phylogenetic analysis of combined data. Syst Biol 53, 47–67.

Nylander, J.A., Wilgenbusch, J.C., Warren, D.L., Swofford, D.L., 2008. AWTY (are we there yet?): a system for graphical exploration of MCMC convergence in Bayesian phylogenetics. Bioinformatics 15, 581–583.

Olson, M.E., Arroyo-Santos, A., 2009. Thinking in continua: beyond the "adaptive radiation metaphor". BioEssays 31, 1337–1346.

Paradis, E., 1997. Assessing temporal variations in diversification rates from phylogenies: estimation and hypothesis testing. Proc. Roy. Soc. B 264, 1141–1147.

Paradis, E., 1998. Testing for constant diversification rates using molecular phylogenies: a general approach based on statistical tests for goodness of fit. Mol. Biol. Evol. 15, 476–479.

Paradis, E., 2006. Analysis of phylogenetics with R. Springer Science + Business Media, LLC.

Posada, D., Crandall, K.A., 1998. MODELTEST: testing the model of DNA substitution. Bioinformatics 14, 817–818.

Presa, P., Pardo, B.G., Martinez, P., Bernatchez, L., 2002. Phylogeographic congruence between mtDNA and rDNA ITS markers in brown trout. Mol. Biol. Evol. 19, 2161–2175.

Pybus, O.G., Harvey, P.H., 2000. Testing macroevolutionary models using incomplete molecular phylogenies. Proc. Roy. Soc. B 267, 2267–2272.

Pybus, O.G., Rambaut, A.P., 2002. GENIE (Genealogy Interval Explorer), Version 3.0. Department of Zoology, University of Oxford, Oxford, UK.

Quental, T.B., Marshall, C.R., 2009. Extinction during evolutionary radiations: reconciling the fossil record with molecular phylogenies. Evolution 63, 3158–3167.

Rabosky, D.L., 2006a. LASER: a maximum likelihood toolkit for detecting temporal shifts in diversification rates from molecular phylogenies. Evol. Bioinf. Online 2, 257–260.

Rabosky, D.L., 2006b. Likelihood methods for inferring temporal shifts in diversification rates. Evolution 60, 1152–1164.

Rabosky, D.L., 2009a. Extinction rates should not be estimated from molecular phylogenies. Evolution 64, 1816–1824.

Rabosky, D.L., 2009b. Heritability of extinction rates links diversification patterns in molecular phylogenies and fossils. Syst. Biol. 58, 629–640.

Rabosky, D.L., Lovette, I.J., 2008. Explosive evolutionary radiations: decreasing speciation or increasing extinction over time? Evolution 62, 1866–1875.

Rambaut, A.P., 2002. PhyloGen, Version 1.0. Department of Zoology, University of Oxford, Oxford.

Ribbink, A.J., 1984. Is the species flock concept tenable? In: Echelle, A.A., Kornfield, I. (Eds.), Evolution of Fish Species Flocks. University of Maine at Orono Press, Orono, pp. 21–25.

Ritchie, P.A., Bargelloni, L., Meyer, A., Taylor, J.A., Macdonald, J.A., Lambert, D.M., 1996. Mitochondrial phylogeny of trematomid fishes (Nototheniidae, Perciformes) and the evolution of Antarctic fish. Mol. Phylogenet. Evol. 5, 383–390.

Ronquist, F., Huelsenbeck, J.P., 2003. MrBayes 3: Bayesian phylogenetic inference under mixed models. Bioinformatics 19, 1572–1574.

Ruber, L., Zardoya, R., 2005. Rapid cladogenesis in marine fishes revisited. Evolution 59, 1119–1127.

Sanchez, S., Dettaï, A., Bonillo, C., Ozouf-Costaz, C., Detrich, B., Lecointre, G., 2007. Molecular and morphological phylogenies of the Nototheniidae, with on taxonomic focus on the Trematominae. Polar Biol. 30, 155–166.

Sanderson, M.J., 2002. Estimating absolute rates of molecular evolution and divergence times: a penalized likelihood approach. Mol. Biol. Evol. 19, 101–109.

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.

Shimodaira, H., 2002. An approximately unbiased test of phylogenetic tree selection. Syst. Biol. 51, 492–508.

Stephens, M.A., 1974. EDF statistics for goodness of fit and some comparisons. J. Am. Stat. Assoc. 69, 730–737.

Stephens, M.A., 1982. Anderson–Darling test for goodness of fit. In: Kotz, S., Johnson, N.L. (Eds.), Encyclopedia of Statistical Science, vol. 1. John Wiley and Sons, New York, pp. 81–85.

Swofford, D.L., 2003. PAUP*. Phylogenetic Analysis Using Parsimony (*and Other Methods). Version 4. Sinauer Associates, Sunderland, Massachusetts.

Takahata, N., 1989. Gene genealogy in three related populations: consistency probability between gene and population trees. Genetics 122, 957–966.

Thatje, S., Hillenbrand, C.-D., Larter, R., 2005. On the origin of Antarctic marine benthic community structure. Trends Ecol. Evol. 20, 534–540.

Van de Putte, A.P., Van Houdt, J.K.J., Maes, G.E., Janko, K., Koubbi, P., Rock, J., Volckaert, F.A.M., 2009. Species identification in the trematomid family using nuclear genetic markers. Polar Biol. 32, 1731–1741.

Wilgenbusch, J.C., Warren, D.L., Swofford, D.L., 2004. AWTY: a system for graphical exploration of MCMC convergence in Bayesian phylogenetic inference. Available from: <http://ceb. csit. fsu. edu/awty>.