Molecular Phylogenetics and Evolution xxx (2012) xxx–xxx

Contents lists available at SciVerse ScienceDirect

Molecular Phylogenetics and Evolution

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

Species tree of a recent radiation: The subfamily Delphininae (Cetacea, Mammalia)

Ana R. Amaral [a,b,*], Jennifer A. Jackson [c], Luciana M. Möller [d], Luciano B. Beheregaray [b,d], M. Manuela Coelho [a]

[a] Universidade de Lisboa, Faculdade de Ciências, Centro de Biologia Ambiental, Campo Grande, 1749-016 Lisbon, Portugal
[b] Department of Biological Sciences, Macquarie University, Sydney, NSW 2109, Australia
[c] British Antarctic Survey, High Cross, Madingley Road, Cambridge CB3 0Et, England, United Kingdom
[d] School of Biological Sciences, Flinders University, Adelaide, SA 5001, Australia

ARTICLE INFO

Article history:
Received 30 June 2011
Revised 4 April 2012
Accepted 4 April 2012
Available online xxxx

Keywords:
Delphinidae
Incomplete lineage sorting
Hybridization
Rapid radiation
Speciation
Species tree

ABSTRACT

Lineages undergoing rapid radiations provide exceptional opportunities for studying speciation and adaptation, but also represent a challenge for molecular systematics because retention of ancestral polymorphisms and the occurrence of hybridization can obscure relationships among lineages. Dolphins in the subfamily Delphininae are one such case. Non-monophyly, rapid speciation events, and discordance between morphological and molecular characters have made the inference of phylogenetic relationships within this subfamily very difficult. Here we approach this problem by applying multiple methods intended to estimate species trees using a multi-gene dataset for the Delphininae (*Sousa*, *Sotalia*, *Stenella*, *Tursiops*, *Delphinus* and *Lagenodelphis*). Incongruent gene trees obtained indicate that incomplete lineage sorting and possibly hybridization are confounding the inference of species history in this group. Nonetheless, using coalescent-based methods, we have been able to extract an underlying species-tree signal from divergent histories of independent genes. This is the first time a molecular study provides support for such relationships. This study further illustrates how methods of species-tree inference can be very sensitive both to the characteristics of the dataset and the evolutionary processes affecting the evolution of the group under study.

1. Introduction

Species complexes undergoing rapid radiation provide an exceptional opportunity to investigate the processes of speciation and adaptation. They also represent a major challenge in molecular systematics because relationships among lineages can be hidden by incomplete lineage sorting and/or introgressive hybridization (Maddison and Knowles, 2006; Wiens et al., 2006). During a rapid radiation, the coalescent pattern of individual gene phylogenies may not match the true pattern of speciation due to incomplete lineage sorting (Hudson, 1992). As a result, many gene trees will be discordant between each other and from the actual species tree (reviewed in Degnan and Rosenberg (2009) and Knowles (2009)). Furthermore, in cases of rapid radiations, the intrinsic barriers that prevent gene flow between species may have had insufficient time to fully develop, leading to hybridization among recently evolved lineages (Seehausen, 2004). The extent to which gene flow persists throughout the process of speciation remains unclear, although it has been documented in a number of recent studies (Niemiller et al., 2008; Quesada et al., 2007; Savolainen et al., 2006). The existence of hybridization events in the evolutionary history of a group means that such taxa will not follow the expected procedure of divergence from a common ancestor through a bifurcating tree (Hennig, 1966), resulting in discordant gene trees.

Incomplete lineage sorting and hybridization are only two of the evolutionary processes that can lead to discordance between gene trees and species trees. Horizontal gene transfer and gene duplication can also lead to such incongruence (Maddison, 1997). However, incomplete lineage sorting and hybridization are the processes that have been more thoroughly studied in a phylogenetic context, leading to the development of several methods that incorporate the stochastic sorting of lineages in the estimation of species trees from gene trees (Kubatko et al., 2009; Liu, 2008; Liu and Pearl, 2007; Maddison, 1997; Maddison and Knowles, 2006). Analyses of empirical and simulated data suggest that these methods can accurately estimate species trees even when high levels of discordance between gene trees exist (Brumfield et al., 2008; and Farrell, 2008; Liu et al., 2008), and are therefore useful for investigating relationships in species complexes associated with rapid diversification (Belfiore et al., 2008; Dolman and Hugall, 2008).

In this study we use four different methods for estimating species trees, applying these to a subfamily of dolphins that has likely radiated recently and presents highly confounding phylogenetic

* Corresponding author at: Faculdade de Ciências da Universidade de Lisboa, Departamento de Biologia Animal, Campo Grande, 1749-016 Lisboa, Portugal. Fax: +351 217500028.
E-mail address: aramaral@fc.ul.pt (A.R. Amaral).

1055-7903/$ - see front matter © 2012 Elsevier Inc. All rights reserved.
http://dx.doi.org/10.1016/j.ympev.2012.04.004

relationships. Firstly, we use a concatenation approach combining multigene data to estimate a species tree. The reasoning behind this approach is that by combining the data, a dominant signal will emerge, resulting in a more strongly supported phylogenetic estimate, which is assumed to be the species tree (de Queiroz and Gatesy, 2007; Gadagkar et al., 2005; Huelsenbeck et al., 1996; Rokas and Carroll, 2005). This approach has recently been criticized (Degnan and Rosenberg, 2006; Kubatko and Degnan, 2007), mainly because if substantial variation in single-gene histories exists, this variance is not incorporated, and phylogenetic signals from the most variable loci will tend to dominate, misleading inference of the true species evolutionary history. Therefore, we also use three different methods that incorporate the coalescent in the estimation of species trees: a summary statistic method, the 'minimize deep coalescence' method (Maddison and Knowles, 2006), and two probabilistic methods that combine Bayesian models in a coalescent framework as implemented in the programs BEST (Liu, 2008) and ∗BEAST (Heled and Drummond, 2010). These three methods differ in the way the information from the coalescent is incorporated, in how coalescent times are summarized, and in the incorporation of uncertainty in the estimated species tree (Knowles, 2009; Liu et al., 2009). The comparison between these different methods allows us to estimate a species tree from a recent and likely rapid radiation and to explore the evolutionary processes that have shaped the evolution of this dolphin group.

The subjects of this study, dolphins of the subfamily Delphininae (family Delphinidae) have likely arisen through a rapid radiation (McGowen et al., 2009; Kingston et al., 2009; Steeman et al., 2009; Slater et al., 2010). Incongruence between mitochondrial and nuclear phylogenies and incomplete lineage sorting (Kingston et al., 2009), uncertainty in the placement of taxa (Xiong et al., 2009), and a rise in net diversification rate within the Delphinidae (Steeman et al., 2009; Slater et al., 2010) all support a rapid radiation of these taxa. Nevertheless, the speciation and extinction dynamics within the Delphininae have not yet been explicitly tested (Slater et al., 2010). The Delphininae therefore exemplify the challenges of inferring species boundaries and phylogenetic relationships described above.

The rapid radiation resulting in extant Delphinidae is estimated to have occurred during the mid- to late Miocene (11–15 mya) (Barnes et al., 1985; McGowen et al., 2009; Steeman et al., 2009). The potential drivers of this radiation are thought to be related to social structure, growth and reproductive characteristics (Gygax, 2002), trophic diversification (Lipps and Mitchell, 1976), and/or climatic changes during the glacial periods of the Pleistocene (Steeman et al., 2009). Delphinidae is the largest cetacean family, composed of at least 37 species (Caballero et al., 2008). Although several phenetic morphological (e.g., (Flower, 1883; Mead, 1975; Muizon, 1988; True, 1889) and cladistic molecular (Caballero et al., 2008; LeDuc et al., 1999; May-Collado and Agnarsson, 2006; Kingston et al., 2009; McGowen et al., 2009; Vilstrup et al., 2011) studies have been conducted, the evolutionary relationships within this family remain unclear, particularly within the subfamily Delphininae (*Sotalia*, *Sousa, Stenella, Tursiops, Delphinus*, and *Lagenodelphis*) (Fig. 1, Supplementary material, Appendix B). The monophyly of the genera *Stenella* and *Tursiops* has been questioned for more than a century (True, 1889) due to a complex of cranial characters not shared by all species of the genus *Stenella*, some of which may actually be more closely related to *Tursiops* or *Delphinus* than to their congeners. Recent phylogenetic studies based on newly sequenced mitochondrial genomes and Amplified Fragment Length Polymorphism (AFLP) markers have supported the polyphyly of *Stenella* and *Tursiops* (Kingston et al., 2009; Xiong et al., 2009). However, two other studies using nuclear gene sequences and a supermatrix approach to infer the phylogeny of whales and dolphins have recovered *Tursiops* as monophyletic (McGowen et al., 2009; Steeman et al., 2009). The number of species within the genera *Tursiops* and *Delphinus* has also been a point of contention. Within *Tursiops*, most recent studies recognize two species, the common bottlenose dolphin (*T. truncatus*) and the Indo-Pacific bottlenose dolphin (*T. aduncus*). However, recent molecular evidence based on mitochondrial DNA suggested that *aduncus*-type dolphins occurring off South Africa (Natoli et al., 2004), as well as coastal bottlenose dolphins from southern Australia, may actually be different species (Möller et al., 2008; Natoli et al., 2004), with the latter being more closely related to *Lagenodelphis hosei* than to the *T. truncatus* or *aduncus* types (see Fig. 1 and Supplementary material, Appendix B). Within *Delphinus*, there are two currently recognized species, the short-beaked common dolphin (*D. delphis*) and the long-beaked common dolphin (*D. capensis*). In addition, an extremely long-beaked form from the Arabian Sea is considered a subspecies (*D. capensis tropicalis*) (Jefferson and Van Waerebeek, 2002), and common dolphins in the Black Sea are also recognized as a subspecies (*D. delphis ponticus*) (Perrin, 2009). However, recent morphological and molecular evidence suggests that at least in some geographical areas this classification (which is based primarily on morphological characters such as beak length and coloration) may not be valid (Amaral et al., 2007a; Murphy et al., 2006).

The use of phylogenetic methods that do not capture the complex nature of DNA evolution in cetaceans has been cited as a reason for the poorly clarified evolutionary relationships within the Delphininae (Xiong et al., 2009; May-Collado and Agnarsson, 2006). Moreover, the datasets used so far have proved to be insufficient to resolve the phylogenetic tree of the subfamily. Nearly all molecular studies conducted so far have focused on the mitochondrial genome (using the cytochrome *b* gene or full genome, e.g., Xiong et al., 2009; May-Collado and Agnarsson, 2006; LeDuc et al., 1999). Four other studies have included DNA sequences from multiple nuclear loci. Two were aimed at elucidating phylogenetic relationships within the order Cetacea and used supermatrix approaches (McGowen et al., 2009; Steeman et al., 2009), the third applied phylogenetic methods to single-locus and concatenated datasets (Caballero et al., 2008), and the fourth used AFLP markers (Kingston et al., 2009).

Here we utilize a multi-gene dataset, including one mitochondrial gene (cytochrome *b*) and 13 nuclear loci (3 introns and 10 anonymous) to estimate a species tree for the subfamily Delphininae. For the first time, coalescent-based methods that account for gene tree heterogeneity were used. The following questions were addressed. (1) Are the coalescent-based methods for species tree estimation able to consistently resolve relationships within the Delphininae? (2) Do these relationships differ from previously published mtDNA and nuclear DNA (nuDNA) phylogenies, particularly with respect to the paraphyly of the genera *Tursiops* and *Stenella*? (3) Are there differences between the species tree topologies obtained with the different methods used? In a broader context, our study contributes to recent analytical debates concerning gene trees and species trees and helps to clarify the evolutionary history of a rapid radiation of a globally distributed and charismatic group of organisms.

2. Material and methods

2.1. Sample acquisition, DNA extraction, amplification, and sequencing

A total of 22 individual samples comprising 12 species belonging to the subfamily Delphininae were used (Table 1). Additionally, *Globicephala melas* (Globicephalinae) and *Phocoena phocoena* (Phocoenidae) were used as outgroups in the phylogenetic analyses (Caballero et al., 2008). Samples were obtained as skin or muscle

Please cite this article in press as: Amaral, A.R., et al. Species tree of a recent radiation: The subfamily Delphininae (Cetacea, Mammalia). Mol. Phylogenet. Evol. (2012), http://dx.doi.org/10.1016/j.ympev.2012.04.004

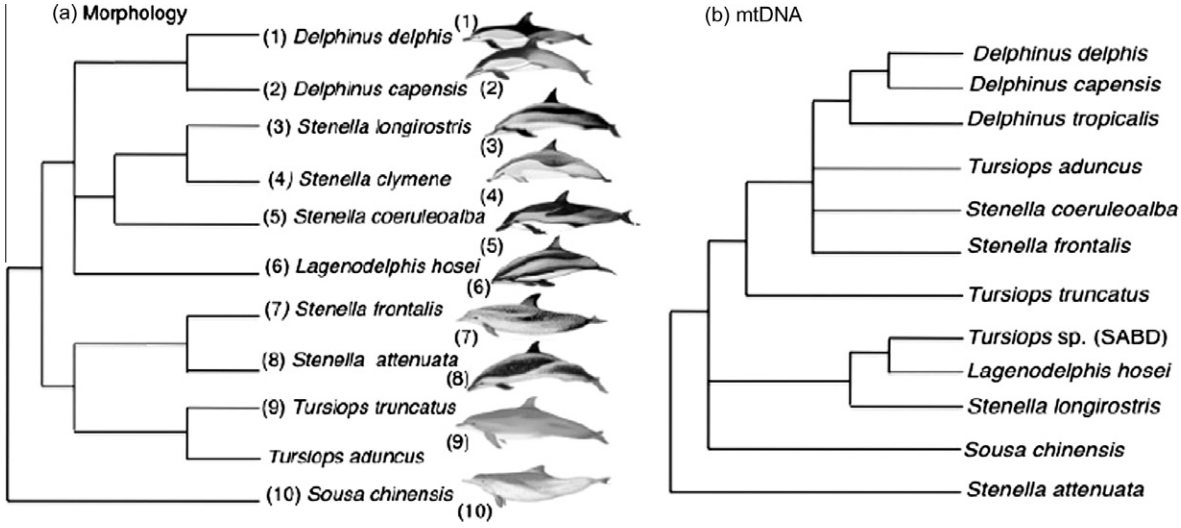

Fig. 1. Relationships between the nominal species in the subfamily Delphininae based on (a) recent morphological analyses (Perrin et al., 1987; Perrin, 2009) and (b) mitochondrial cytochrome *b* gene (adapted from LeDuc et al. (1999) and Möller et al. (2008).

tissue from dead stranded animals or from free-ranging animals using biopsy darts. Some samples were received from the Southwest Fisheries Science Center, Marine Mammal and Turtle Research Sample Collection (SWFSC-NOAA, La Jolla, CA) as extracted DNA. Recognized experts made all species identifications.

DNA was extracted following standard phenol–chloroform extraction protocols (Sambrook et al., 1989). The polymerase chain reaction (PCR) was used to amplify one fragment of the mitochondrial genome (the cytochrome *b* gene), three nuclear introns (BTN, CHRNA1, and PLP) and ten anonymous nuclear loci (Table 2). Anonymous markers are non-coding regions of the genome, randomly collected and presumably dispersed across the chromosomes, thereby representing wide and potentially unbiased variation across the genome. These loci were developed from clone sequences selected from a genomic library created for the common dolphin *D. delphis* (Amaral et al., 2010). The PCR reactions were performed in 25-μL reactions containing 10–100 ng DNA, 0.2 mM each dNTP, 0.3 μM each primer, 1 U Taq Polymerase, and 1× Taq buffer. PCR products were separated on 1.0% agarose gels, stained with ethidium bromide, and visualized with ultraviolet light. PCR products were cleaned with exonuclease I and shrimp alkaline phosphatase to remove free nucleotides and primers and sequenced in both directions (BigDye Terminator CycleSequencing: Applied Biosystems) on an ABI 3730xl automated sequencer (Applied Biosystems).

Sequences were manually edited and aligned using Sequencher v. 4.2 (Gene Code Corporation). Some alignments for the nuclear loci required gaps: in Del_02 a gap of 90 bp was required; in Del_05 a gap of 8 bp was required; in Del_08 several simple gaps were required due to length variations of AT repeats, as well as an additional gap of 36 bp; and in Del_11 a gap of 19 bp was required. In BTN a simple gap of 2 bp was required. Alignments were confirmed using CLUSTALX v. 2.0.10 with the default parameter settings. (data deposited in the Dryad Repository: http://dx.doi.org/10.5061/dryad.6dr0475t).

The direct sequencing of the PCR products for the anonymous loci Del_02, Del_05, Del_10 and Del_12 frequently generated continuous overlap of signals between sequences amplified with the

Table 1
Species included in this study. classification follows Rice (1988).

Species	Number of samples sequenced	Geographical location	Institution[a]
Family Delphinidae			
Subfamily Delphininae			
Sotalia fluviatilis (Tucuxi)	1	Brazil	SWFSC
Sousa chinensis (Indo-Pacific humpback dolphin)	2	Hong Kong	SWFSC
Lagenodelphis hosei (Fraser's dolphin)	1	Hawaii	SWFSC
Stenella longirostris (Spinner dolphin)	2	East Pacific	SWFSC
S. attenuata (Pantropical spotted dolphin)	2	Mexico	SWFSC
S. coeruleoalba (Striped dolphin)	2	Portugal	ICN
S. frontalis (Atlantic spotted dolphin)	1	Portugal	ICN
Delphinus delphis (Short-beaked common dolphin)	2	Portugal	ICN
D. capensis (Long-beaked common dolphin)	2	East Pacific	SWFSC
D. c. tropicalis (Arabian common dolphin)	1	Arabian Sea	SWFSC
Tursiops truncatus (Common bottlenose dolphin)	2	Portugal	ICN
T. aduncus (Indo-Pacific bottlenose dolphin)	2	Australia	MQ
Tursiops sp. (Southern Australian bottlenose dolphin)	2	Australia	MQ
Subfamily Globicephalinae			
Globicephala melas (Long-finned pilot whale)	1	Portugal	ICN
Family Phocoenidae			
Phocoena phocoena (Harbor porpoise)	1	Portugal	ICN

[a] Institutional abbreviations: ICN – Instituto de Conservação da Natureza (Portugal); SWFSC – Southwest Fisheries Research Center (USA); MQ – Macquarie University.

Table 2
List of loci and primers used in this study and variable sites obtained.

Locus	Primers	Total (variable) sites	References
Cyt*b*	L14724 P2	1120 (290)	LeDuc et al. (1999)
BTN	But-b1s BTNr4	754 (4)	Lyons et al. (1999)
PLP	PLP-F PLP-R	750 (20)	Lyons et al. (1999)
CHRNA1	CHRNA1F CHRNA1R	357 (14)	Roca et al. (2001)
Del_02	Del_02F Del_02R	923 (25)	Amaral et al. (2010)
Del_04	Del_04F Del_04R	636 (26)	Amaral et al. (2010)
Del_05	Del_05F Del_05R	750 (16)	Amaral et al. (2010)
Del_08	Del_08F Del_08R	806 (41)	Amaral et al. (2010)
Del_10	Del_10F Del_10R	402 (15)	Amaral et al. (2010)
Del_11	Del_11F Del_11R	572 (26)	Amaral et al. (2010)
Del_12	Del_12F Del_12R	729 (40)	Amaral et al. (2010)
Del_14	Del_14F Del_14R	318 (2)	Amaral et al. (2010)
Del_15	Del_15F Del_15R	780 (32)	Amaral et al. (2010)
Del_17	Del_17F Del_17R	739 (21)	Amaral et al. (2010)

forward and reverse primers. This indicated the existence of a length polymorphism in the amplified region. We cloned the PCR product (performed at Macrogen, Inc.) and sequenced between 8–10 cloned fragments to determine the two allelic sequences. We then used the program Indelligent (http://ctap.inhs.uiuc.edu/dmitriev/indel.asp) to check if the two allelic sequences identified matched the expected ones. Although an error in PCR can create artificial heterozygosity, we performed PCRs for cloning and direct sequencing independently and still found consistency between the overlapping signals on the direct and cloned sequences. The results thus appeared to be unaffected by PCR artefacts. The program PHASE v. 2.1 (Stephens and Donnelly, 2003; Stephens et al., 2001) was used to infer alleles from heterozygous individuals, setting the phase-certainty threshold to 90%.

2.2. Dataset construction

Both combined and separate analyses of the major partitions were conducted. Three different datasets were constructed: "mtDNA" including the cytochrome *b* gene; "nuDNA" including the three introns and the 10 anonymous loci; and "mtDNA + nuDNA" including the cytochrome *b* gene, the three introns and the 10 anonymous loci.

In order to test whether heterozygote sites, allele size polymorphisms, and insertions/deletions would have any influence on the estimated phylogenetic trees, we performed analyses considering: (i) heterozygous sites coded using the IUPAC ambiguity code and gaps treated as missing "new states" or coded with a binary code (0 or 1) indicating their presence or absence; and (ii) including information from heterozygotes and allele size polymorphisms by including all allelic sequences from each individual in the input matrix. Gene tree topologies did not vary substantially with the inclusion of gap partitions and all allelic sequences. Numbers of equally parsimonious trees obtained, tree lengths, bootstrap nodal support, and posterior probability values were altered only slightly. Therefore, subsequent analyses were performed with heterozygous sites coded as IUPAC ambiguity codes and gaps treated as missing data. Using all allele sequences would be computationally very demanding, particularly for estimating species trees using the program BEST.

2.3. Phylogenetic analyses

Modeltest v. 3.7 (Posada and Crandall, 1998) was used to infer the best-fitting nucleotide substitution model for each locus. Models of evolution were chosen for subsequent analyses according to a second-order Akaike Information Criterion (AIC_c), with branch lengths included as additional parameters and a correction for small sample sizes employed (Hurvich and Tsai, 1995; Posada and Buckley, 2004). Nucleotide substitution models for each locus are given in Table 3.

MrBayes v. 3.1. (Huelsenbeck and Ronquist, 2001) was used to estimate Bayesian phylogentic trees for each locus and for the concatenated dataset. Four simultaneous Metropolis-Coupled MCMC chains (one cold and three heated) were run for 2 million generations, with trees sampled at intervals of 100 generations. Random trees were used to begin each Markov chain, and a molecular clock was not enforced. Convergence was assessed by the standard deviation of split frequencies (values below 0.01) and by the achievement of stationarity of the log-likelihood values of the cold chain. The first 2000 trees were discarded as "burn-in" after examining the variation in log-likelihood scores over time. The cytochrome *b* dataset was partitioned by codon positions in the Bayesian analysis, assuming that there might be differences in the molecular evolution of the different positions (Shapiro et al., 2005). Phylogenetic trees for each locus were also obtained under the maximum parsimony (MP) criterion. Details of this analysis and results are included in Supplementary material (Appendix A).

We tested for incongruence among loci by performing 'crossed' Shimodaira-Hasegawa tests on maximum likelihood trees generated in PAUP* (Shimodaira and Hasegawa, 1999), whereby the highest likelihood topologies obtained with individual datasets were compared against each other (using the AIC_c-preferred nucleotide substitution model for each dataset), also including comparisons of individual datasets against the ML phylogenies obtained by all nuclear loci concatenated, and all loci concatenated (e.g., Delsuc et al., 2002). With this analysis we tested whether the mitochondrial gene tree was more discordant to the nuclear gene trees than nuclear gene trees were to each other.

2.4. Estimation of species trees

In addition to the concatenated analysis performed using Bayesian Inference methods and Maximum Parsimony (Supplementary

Table 3
Modeltest minimum AIC models (taxa and character-corrected) for each locus. *AIC was incalculable for this locus.

Locus	Nucleotide substitution model	AICc	Weight
Del_02	K81uf	2630	0.16
Del_04	JC	1912	0.29
Del_05	TIM	2133	0.11
Del_08	JC	2084	1.00
Del_10	GTR + I	1198	0.22
Del_11	K81uf	1723	0.17
Del_12	JC	1891	1.00
Del_14	N/A*		
Del_15	HKY	2358	0.47
Del_17	HKY	2136	0.09
BTN	JC	680	1.00
CHRNA1F	GTR	1087	0.15
PLP	K81 + I	2302	0.04
CYTB	TRN + G (0.418)	7572	0.51

material, Appendix A), three different coalescent-based methods were used to estimate the species tree from the 13 nuclear gene trees obtained: Minimize Deep Coalescence (MDC), Bayesian Estimation of Species Trees (BEST) and *BEAST. The MDC approach seeks the species tree that minimizes the number of incomplete lineage sorting (deep coalescence) events that must be inferred to explain observed gene trees (Maddison, 1997). This approach was implemented in Mesquite v. 2.72 (Maddion and Maddison, 2009) using the individual gene trees estimated using MP and Bayesian Inference as described above. Following the methods proposed by Linnen and Farrell (2008), tree searches were first performed using the following options: contained polytomies automatically resolved, branch lengths of contained trees included, and tree rearrangements made by subtree pruning and regrafting (SPR). A second search was performed without automatically resolving polytomies and without including branch lengths to evaluate the sensitivity of the species tree to this search strategy.

Bayesian Estimation of Species Trees (BEST) uses a Bayesian hierarchical model to estimate a distribution of species trees from vectors of estimated gene trees across multiple loci, under a multispecies coalescent model (Liu and Pearl, 2007). BEST uses MrBayes to generate a posterior distribution of gene trees across loci using a prior based on an approximate species tree; it then estimates a species tree from the joint posterior distribution of gene trees using a uniform prior method. The analysis was implemented in BEST v. 2 using the partitioned combined dataset described above. In order to minimize over-parameterization of the individual datasets in the shared Bayesian framework, one transition:transversion ratio and set of base frequencies was co-estimated for all datasets combined. Sensitivity of each dataset to this approach was determined by comparing the maximum likelihood phylogeny from each dataset (as estimated in PAUP* using the AICc-preferred nucleotide substitution models summarized in Table 3) with the phylogeny produced under the parameters of the combined nucleotide substitution model using Shimodaira–Hasegawa testing (Shimodaira and Hasegawa, 1999), and this nucleotide substitution model was not found to be a significantly worse fit for any dataset.

We explored a number of prior choices in BEST. To investigate the impact of population size prior choice on species tree inference, two analyses were performed with alternative inverse gamma priors, chosen so as to bound the sequence-based estimates of θ calculated in DNAsp (Rozas et al., 2003). These were $\alpha = 2$ and $\beta = 0.001$ and 0.002, corresponding to $\theta = 0.001$ and 0.002, respectively. Each analysis was run for 20 million generations and summarized over 10–12 independent runs, with 2 million generations discarded as burn-in. In order to examine the impact of using different branch length priors, we further performed two analyses using a coalescent (uniform clock) branch length prior and inverse gamma priors of $\theta = 0.001$ and $\theta = 0.002$. We also tested the phylogenetic impact of widening the range of the mutation rate prior for the coalescent prior analyses, since inspection of the posterior mutation rates of these analyses revealed some median values close to the prior boundaries. The standard prior (which allowed relative rates to be uniform over 0.5–1.5 of the mean value) was therefore modified to a wider range of 0.1–2. A BEST analysis using the inverse gamma prior of $\theta = 0.0015$ (and the coalescent tree prior and wider mutation rate prior) was also carried out using ten independent runs over 20 million generations, since this theta value is closest to the true estimate derived for the population. Posterior summary distributions were inspected for convergence and mixing using the program TRACER v1.5 (Rambaut and Drummond, 2007).

The third method used was implemented in the software package *BEAST (Heled and Drummond, 2010). Although *BEAST also estimates species trees under the multispecies coalescent, there are several modeling differences when compared to the method implemented in BEST. *BEAST coestimates the species tree and all gene trees simultaneously in one Bayesian MCMC analysis, instead of the two steps required by BEST. Moreover, an outgroup is not required, population size does not have to be assumed constant over the branches of the trees, and two different species tree priors are available: Yule process and Birth–Death process (Heled and Drummond, 2010). We ran 100 million MCMC generations sampling every 10,000 generations, choosing the Yule process as species tree prior and the Piecewise constant and linear model for population size estimates. The HKY model of nucleotide substitution was chosen for all loci with the exception of Del_10 and BTN, for which the GTR model was chosen. A relaxed molecular clock with an uncorrelated lognormal distribution was chosen. The program TRACER v.1.5 was run to ensure mixing and convergence of the posterior distribution and parameters by examining effective sample size (ESS) values. TreeAnnotator v.1.6.1. (Rambaut and Drummond, 2010) was subsequently used to summarize the obtained trees in a single, consensus tree that represents the posterior distribution.

3. Results

3.1. nuDNA – separate analyses of nuclear loci

A total of 8516 bp of nuclear DNA was obtained for the 15 species analyzed across 13 nuclear loci (data deposited in the Dryad Repository: doi:10.5061/dryad.6dr0475t). Amplification of some loci was unsuccessful for some species (Del_04, Del_05 and Del_08 for *Lagenodelphis hosei*; Del_05 and Del_17 for *Phocoena phocoena* and Del_17 for *Globicephala melas*). The crossed SH tests identified a conflict of signal among the data partitions, i.e., individual gene trees (see Supplementary material, Table S1). Fragment lengths obtained varied from 357 bp (CHRNA1) to 923 bp (Del_02). Levels of polymorphism obtained were low (Table 2). MP and Bayesian trees obtained for each locus had low resolution and presented highly discordant genealogies (Supplementary Figs. S1 and S2, Bayesian trees only).

3.2. mtDNA

A 1120-bp fragment of the cytochrome *b* gene (hereafter Cyt*b*) was sequenced for the 15 species analyzed in this study. The Bayesian tree resulted in nearly the same topology as the MP tree (Fig. 2a, Bayesian tree only). This tree is similar to the one presented in previous studies (Möller et al., 2008 (Fig. 1a); LeDuc et al., 1999). The genus *Delphinus* was the only genus that was rendered monophyletic. The genus *Stenella* was paraphyletic, with *S. coeruleoalba* and *S. frontalis* more closely related to *Delphinus* spp. than with its congeners. The genus *Tursiops* was also paraphyletic, with the southern Australian bottlenose dolphin (hereafter *Tursiops* sp.) clustering with *L. hosei*.

3.3. nuDNA (introns + anonymous loci)

The concatenation of all nuDNA loci resulted in a total of 8516 bp. Bayesian and MP trees resulted in very similar topologies, with all but three branches in the Bayesian tree having posterior probabilities of 100% (Fig. 2b). The nuDNA phylogeny differed from the one obtained with mtDNA in the order of branching relationships, but the genera *Stenella* and *Tursiops* were still not monophyletic. Southern Australian *Tursiops* sp. clustered with *T. aduncus* and not with *L. hosei* as it did in the mtDNA tree. *L. hosei* was more closely related with *S. coeruleoalba* and *Delphinus* spp. in this nuDNA phylogeny. The differences in taxon position between mtDNA

Please cite this article in press as: Amaral, A.R., et al. Species tree of a recent radiation: The subfamily Delphininae (Cetacea, Mammalia). Mol. Phylogenet. Evol. (2012), http://dx.doi.org/10.1016/j.ympev.2012.04.004

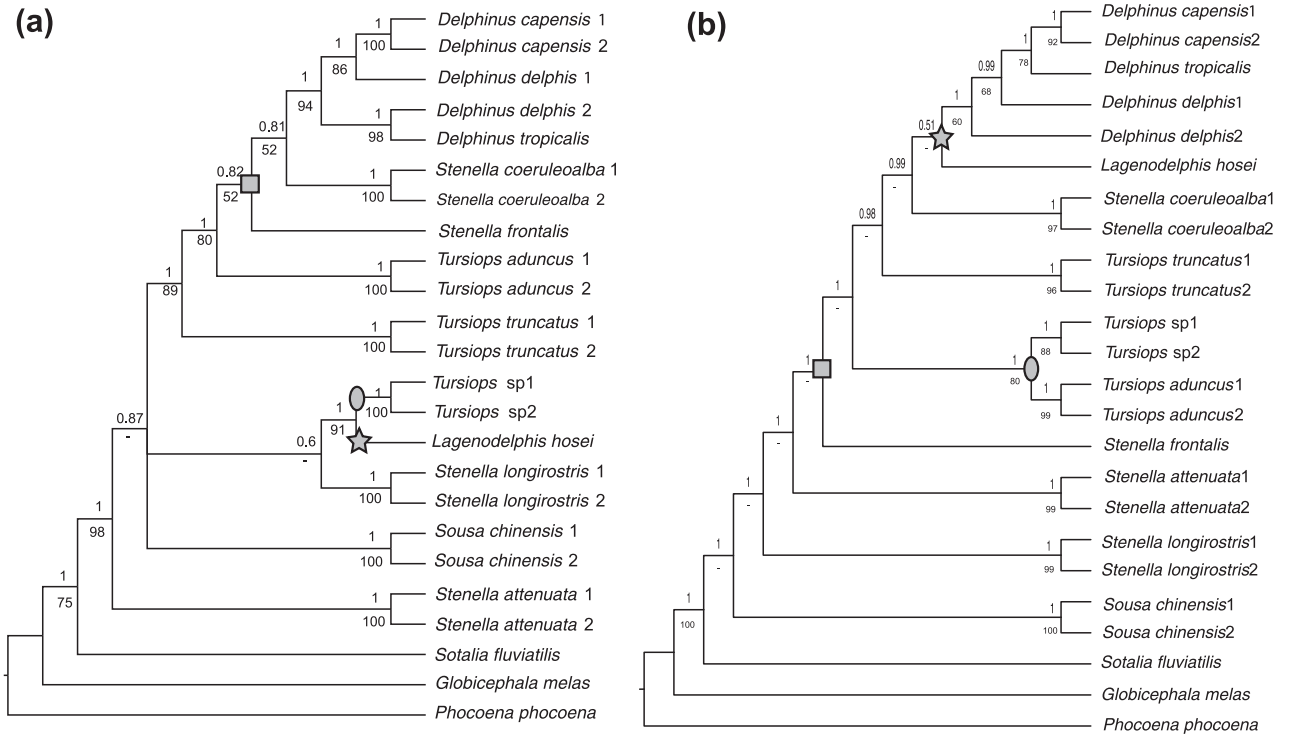

Fig. 2. Majority rule consensus trees generated in MrBayes for the cytochrome *b* gene (a) and for the 13 concatenated nuclear loci (b). Posterior probability values are above nodes and bootstrap support values obtained in the maximum parsimony analysis are below nodes, with (–) indicating lack of support of the particular branch in this analysis. Symbols indicate clades that differ between both trees (see Section 3.3 for further details).

and nuDNA trees are represented in the trees (Figs. 2a and b, respectively). The sister taxon of *Delphinus* was not resolved.

3.4. Species trees

For the MDC analyses we decided to discard locus Del_14 since a very low number of variable sites available (two) provided insufficient information to estimate phylogeny. The MDC method for species tree estimation, as implemented in Mesquite, yielded three species trees with score 94 when individual gene trees estimated with the MP method were used (Fig. 3a, Table 4), and four species trees with a score of 71 when Bayesian individual gene trees were used (Fig. 3b, Table 4). Topologies obtained were similar, differing only in the branching order within Clade A (Figs. 3a-b), which includes the genus *Delphinus*, *S. coeruleoalba*, *S. longirostris*, and *L. hosei*. In both trees the genus *Tursiops* is rendered a monophyletic sister group to *S. attenuata* and *S. frontalis* (Clade B). *Sousa chinensis* and *S. fluviatilis* occupy a basal position in both trees. Not resolving polytomies automatically resulted in one tree with score 549 when MP gene trees were used, and one tree with score 634 when Bayesian gene trees were used. These trees differed from those obtained when polytomies were automatically resolved (results not shown) and the MDC scores were much higher (Table 4). The species trees presented here were estimated using branch lengths since this is usually recommended in order for the fit to reflect the actual history. Excluding branch length information in the estimation of the species tree had no effect on tree topology.

No individual locus rejected the nucleotide substitution model estimated from the combined dataset, so this model (Ti:Tv = 4.4915, rates = equal, base frequencies from combined dataset) was applied to all loci in the BEST framework. The BEST analyses averaged over 10–12 runs did not achieve full convergence (effective sample sizes, ESS, higher than 100) for all parameters, suggesting that incomplete mixing was achieved for most analyses. However, nearly all parameters in the coalescent tree prior analysis achieved convergence (with some mutation rate priors not converging), while neither mutation rates nor tree priors converged adequately in the exponential tree prior analysis (Supplementary material, Appendix C).

Choice of different population size, branch length and range of mutation rate priors had some influence on the topology of the species trees estimated by BEST (results not shown). These differences relate mostly to the position of the genus *Tursiops* and of *Lagenodelphis hosei*. The analyses that resulted in the highest levels of convergence of the likelihood parameter (as given by ESS values, Supplementary material, Appendix C) were the ones where the coalescent branch length prior and a wider ranging mutation rate prior were used. Here, using different theta priors resulted in identical species tree topologies. Although support for most branches is quite low, these trees are very similar to the MDC species trees in that Clades A and B are also recovered. The only difference is that the spotted dolphins (*Stenella attenuata* and *S. frontalis*) cluster together (Clade C) but do not have a sister taxon relationship with *Tursiops* (Fig. 4a).

The *BEAST analysis achieved convergence, with the posterior distribution and all parameters having ESS values higher than 200. The tree obtained has exactly the same topology as the tree obtained in BEST, with Clades A and C being supported by higher posterior probabilities and Clade B by lower posterior probabilities (Fig. 4b). The low levels of support reflect the uncertainty in resolving relationships within these clades, which is likely due to the low number of variable characters found.

4. Discussion

This is the first study to use a species tree approach that accounts for gene tree heterogeneity to infer phylogenetic relationships within Delphininae using DNA sequences from several nuclear loci. Although individual gene trees were unresolved and highly incongruent, using coalescent-based species tree methods

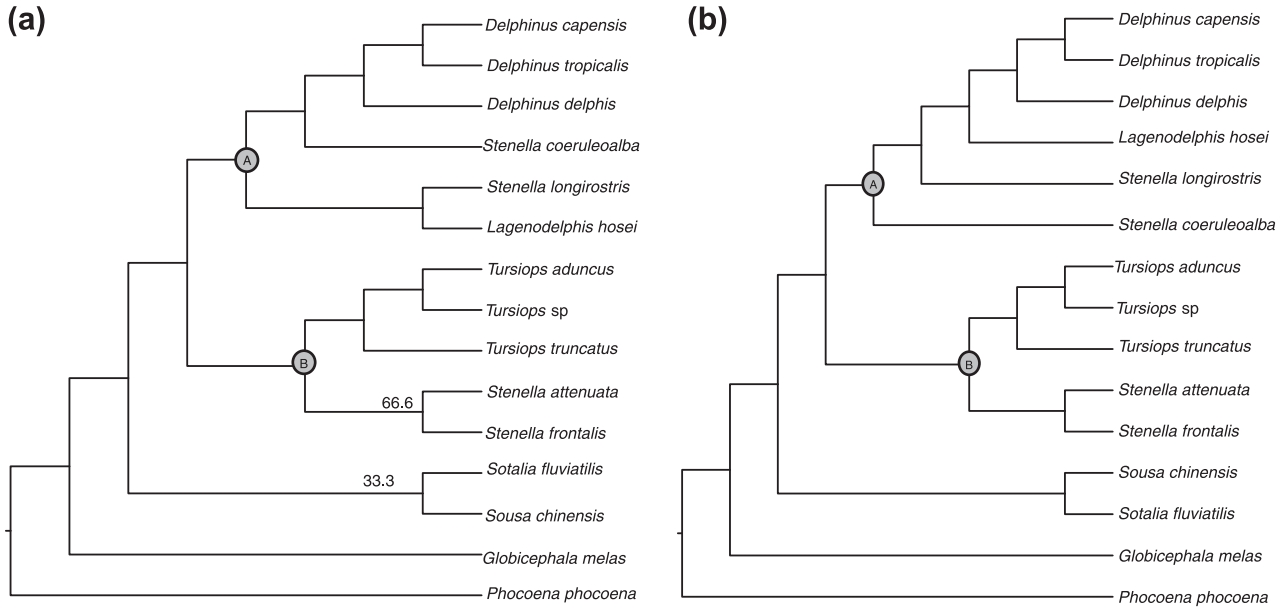

Fig. 3. Species tree estimated with minimize deep coalescence method for (a) individual trees obtained with maximum parsimony and (b) individual trees obtained with Bayesian Inference method. In 'a', number above node indicates the percentage of Minimizing Deep Coalescence trees that contained that clade. Clade labels A and B indicate clades that are discussed further in Section 3.4.

we suggest a new hypothesis for relationships within Delphininae that has not been found in previous molecular studies (LeDuc et al., 1999; Caballero et al., 2008; Möller et al., 2008; Xiong et al., 2009; Kingston et al., 2009; Steeman et al., 2009; McGowen et al., 2009). These relationships are similar to those supported by morphological data (Fig. 1) (Perrin et al., 1987; Perrin, 2009). The topology of the species tree obtained differed between the concatenation and the coalescent-based approaches. Among these latter methods, differences found were not related with topology of the tree, but with the success of parameter convergence and prior choice.

4.1. Phylogenetic relationships

We provide an example of how phylogenetic studies based on single gene trees may prove misleading, particularly in recent species radiations. The phylogeny obtained with mtDNA differed from the phylogeny obtained by concatenating all nuclear loci, and both differed from the species trees estimated by methods that account for gene tree heterogeneity. Although MDC and the Bayesian species trees obtained in *BEAST and BEST differed in the placement of *Tursiops*, the overall topology is similar to the morphology-based relationships (Fig. 1). Uncertainty in resolving relationships within Clade A, which may be due to the low levels of variability found in the 13 nuclear loci used, together with the short branch lengths obtained, reflects the early history of these species and further suggests that they have rapidly radiated.

Most of the controversy surrounding the taxonomy of members of the subfamily Delphininae has arisen from the disagreement between the taxonomy originally established by morphological characters (e.g., Flower, 1883) and the phylogenetic relationships subsequently supported by molecular studies (e.g., LeDuc et al., 1999). Such studies, however, were based on mtDNA, single-locus phylogenies, or AFLPs (Kingston et al., 2009) and were likely recovering an incomplete species history.

4.1.1. Genus Tursiops

The two species presently included in the genus *Tursiops* share several morphological similarities, with a short beak distinctly marked off from the prenarial adipose elevation and less numerous and larger teeth distinguishing them from other Delphininae genera (Flower, 1883; Perrin et al., 2007; True, 1889; Wang et al., 2000). However, all molecular studies conducted using mtDNA and AFLPs have recovered the genus as polyphyletic (e.g., Kingston et al., 2009; LeDuc et al., 1999). The species trees obtained in this study with the coalescent-based methods have all recovered the genus as monophyletic, although with low support. These results support the findings obtained in two recent studies that used a supermatrix approach (McGowen et al., 2009; Steeman et al., 2009). The recently proposed new species of bottlenose dolphin from southern Australia (*Tursiops* sp.) (Möller et al., 2008) grouped with *T. aduncus* in all analyses that included the nuDNA dataset. This relationship is strongly discordant with the mtDNA tree but supports the future revision of its taxonomic status. This discordance between nuDNA and mtDNA should be further explored to clarify, for instance, whether the new species may have arisen through hybridization.

4.1.2. Genus Stenella

The spotted dolphins, *S. attenuata* and *S. frontalis*, are morphologically very similar both in coloration and in skull characters (Perrin et al., 1987), but most molecular phylogenetic studies conducted to date have found them to be non-monophyletic (LeDuc et al., 1999; this study, Fig. 1). However, the coalescent-based species trees obtained in this study have recovered a sister relationship between these two taxa, supporting a recent phylogenetic study based on AFLP markers (Kingston et al., 2009). More importantly, the MDC species trees grouped these two species with the genus *Tursiops*, a relationship that has never before been recovered. Interestingly, the two studies referred to above that used a supermatrix approach did not recover this relationship, as did the phylogenetic tree resulting from the concatenation approach presented in Fig. 2b.

The genus *Stenella* is rendered polyphyletic in all analyses, supporting previous molecular phylogenetic studies (LeDuc et al., 1999; Caballero et al., 2008; Xiong et al., 2009; Kingston et al., 2009; McGowen et al., 2009; Steeman et al., 2009) and further suggesting that this group needs considerable taxonomic revision. High support for this paraphyly is given in the *BEAST tree

Table 4
Deep coalescence scores obtained for each locus in the analysis based on individual gene trees obtained with maximum parsimony (MP) and Bayesian Inference (BI): (a) using the auto-resolve polytomies option and (b) non auto-resolving polytomies.

	Nuclear loci											Total score
	Del02	Del 04	Del05	Del 08	Del10	Del11	Del12	Del15	Del17	BTN	CHRNA1	PLP	
(a) Species trees													
MP													
1	8	1	0	20	8	4	6	12	7	9	16	3	97
2	8	1	0	21	8	4	8	13	6	9	13	3	97
3	7	1	0	21	8	4	9	14	6	8	13	3	97
BI													
1	2	1	5	21	0	3	6	11	7	10	3	2	71
2	2	1	6	20	0	4	6	10	7	10	3	2	71
3	2	2	5	20	0	3	6	12	6	10	3	2	71
4	2	2	6	19	0	4	6	11	6	10	3	2	71
(b) nonAR													
MP	58	53	58	49	63	54	62	36	47	45	60	49	634
BI	40	46	55	43	48	50	50	40	43	41	57	36	549

(Fig. 4b). Species within this genus are both morphologically and genetically very dissimilar (LeDuc et al., 1999; Perrin, 1997), suggesting that this may be an artificial assemblage. As can be seen in all trees, *S. coeruleoalba* is more closely related to the genus *Delphinus* than with its congeners, and *S. frontalis* and *S. attenuata* form a different, divergent group from *S. longirostris*.

4.1.3. Genera Delphinus, Lagenodelphis and Sousa

The genus *Delphinus* is rendered monophyletic in all analyses. However, in the mtDNA and nuDNA phylogenies, the position of the *tropicalis* form varied. In the species tree framework, all individuals from one species are forced to be monophyletic by the structure of the data input, so these relationships could not be clarified. The sister taxon affinities of this genus could also not be elucidated. The skulls of *L. hosei* and *S. coeruleoalba* show a strong resemblance to that of *Delphinus* spp. with regard to the presence of deep palatal grooves, a derived characteristic that no other delphinid species possesses (Dolar, 2009; personal communication, W.F. Perrin). The skull of *S. coeruleoalba* shares additional similarities with that of species of *Delphinus* (Amaral et al., 2009). In fact, most phylogenetic trees obtained in this study, including mtDNA and nuDNA trees, place *S. coeruleoalba* as the sister taxon to *Delphinus*.

The position of *Sousa chinensis* varied between the phylogenetic analyses but its inclusion in the subfamily Delphininae is supported, as suggested by other molecular phylogenies (LeDuc et al., 1999; Caballero et al., 2008).

4.2. Comparison of methods

The different species tree methods used in this study resulted in somewhat different topologies. The tree obtained with the concatenation approach (Fig. 2b) differed from the trees obtained with the coalescent-based methods that take into account gene tree heterogeneity, despite having strongly-supported branches. It has been suggested that the statistical advantage conferred by increasing sample size (number of sites) may result in a presumed improvement in phylogenetic accuracy and branch support (Gadagkar et al., 2005). This possibly explains the fact that the two studies using a supermatrix approach (McGowen et al., 2009; Steeman et al., 2009) recovered the genus *Tursiops* as monophyletic since they used a considerably larger number of sites than the present study. However, the failure of such approaches to explicitly model relationships between gene trees and species trees will likely result in an incorrect phylogeny estimate (Degnan and Rosenberg, 2006; Kubatko and Degnan, 2007; Degnan and Rosenberg, 2009). This difference between a method based on summary statistics and parameter-rich Bayesian probabilistic models, as implemented in BEST and *BEAST, is likely explained by the characteristics of the dataset used, particularly the low number of variable sites obtained. This lack of variability caused problems with parameter convergence in the BEST analysis, which in turn was reflected in the low support obtained for most branches. Similar results were obtained in a recent study of the genus *Oriza*, which was also characterized by low levels of sequence divergence (Cranston et al., 2009). Although the analysis in *BEAST achieved convergence much better than the analysis in BEST, the resulting tree still had low support for most branches. In contrast, MDC gene trees are obtained using maximum parsimony (Maddison and Knowles, 2006), which performs optimally under conditions of relatively low sequence divergence, since it cannot account for unobserved substitutions (e.g., Steel and Penny, 2000).

Additionally, the multispecies coalescent model implemented in BEST and *BEAST assumes that incomplete lineage sorting (deep coalescence) is the only evolutionary process causing the

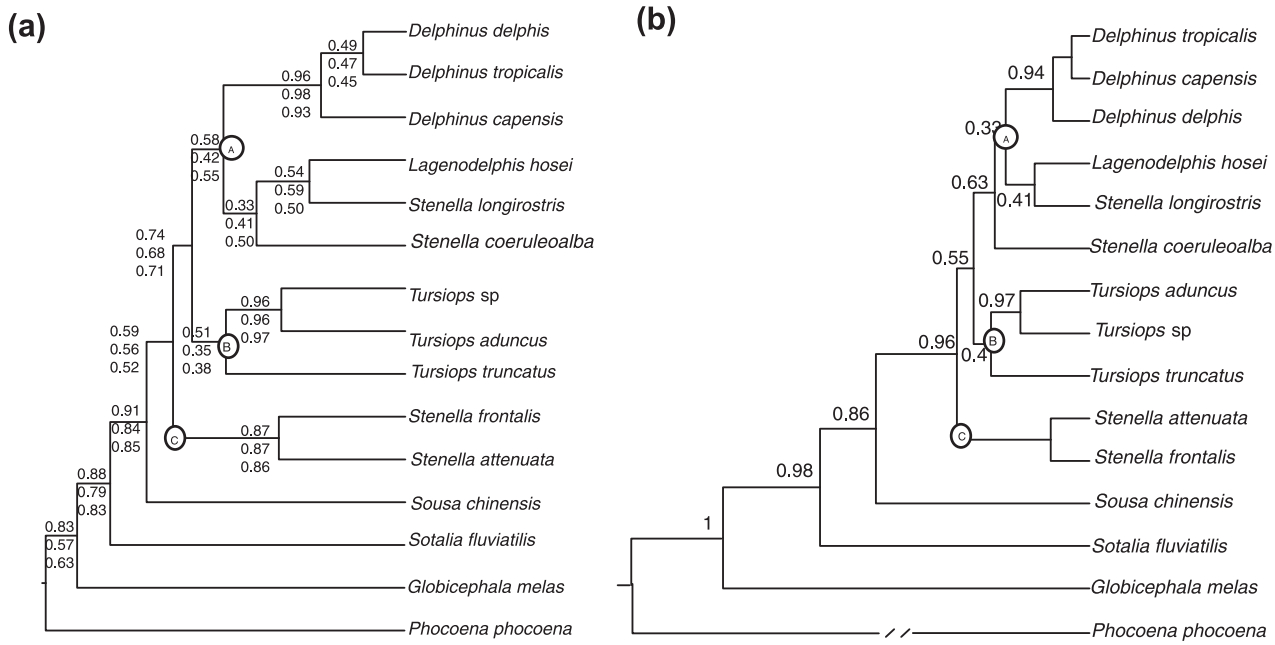

Fig. 4. Species trees estimated with (a) BEST using a coalescent branch length prior, a wider mutation rate prior, and population size priors of $\theta = 0.001$, $\theta = 0.0015$, and $\theta = 0.002$; and (b) with the *BEAST method. In (a) posterior probability values are above and below nodes and correspond to the trees obtained with the different theta values, from top to bottom, respectively. In (b) posterior probability values are above nodes. Node heights represent mean heights. Clade labels A and B indicate clades that are discussed further in the text (Section 3.4).

incongruence between gene trees. However, it is possible that hybridization is also playing a role in the evolutionary history of Delphininae, which may bias these methods by altering gene tree branch lengths, which in turn will restrict the corresponding speciation times and mislead the species tree estimation (Liu and Pearl, 2007). Although hybridization can also bias the MDC reconstruction, it has been suggested that this method is more robust to the presence of gene flow as long as it is not the major force driving the evolutionary history of the species (Maddison and Knowles, 2006; Liu et al., 2009). However, this method still has some caveats. Firstly, there were differences in the estimated trees when polytomies were not automatically resolved (although MDC scores were substantially worse), which suggests that the method still has difficulties in handling uncertainties in individual gene trees. Secondly, this method is unable to provide a measure of support for the relationships.

We found that BEST was particularly sensitive to prior choice when estimating the species tree for this dataset. Different branch length, range of mutation rates and theta priors highly influenced the resulting trees (Fig. 4a). Although we have managed to find the "combination" of priors that lead to a more robust species tree estimation, where theta values no longer influenced tree topology, this process was time-consuming and does not guarantee that the best species tree is estimated since not all parameters converged to satisfaction. Although some studies have been quite robust to prior choice (Liu and Pearl, 2007; Brumfield et al., 2008), others have also found BEST to be sensitive to the choice of theta for estimating species tree in their dataset (Linnen and Farrell, 2008). We therefore suggest that a thorough exploration of the priors choices be done when using BEST, as these can be highly dependent on the dataset used.

Finally, sampling may have also influenced the differences obtained with the methods used. It has been shown that the species tree methods used in this study can in fact be sensitive to sampling schemes (e.g., Linnen and Farrell, 2008). Several species within Delphinidae present cosmopolitan distributions and higher intraspecific than interspecific genetic variability (e.g., Amaral et al., 2007b; Forcada, 2009). It is possible that having sampled individuals from different geographical locations or even more individuals per species could have resulted in different tree topologies. The sensitivity of the methods to different sampling schemes could not be evaluated in this study due to the difficulty of accessing samples, but the fact that each method resulted in a different topology indicates that such sensitivity may exist.

4.3. Incomplete lineage sorting, hybridization, or both?

Incongruence between mtDNA and nuDNA phylogenies has been described in other animal groups (Goncalves et al., 2007; McCracken and Sorenson, 2005; Peters et al., 2007; Shaw, 2002) and may be due to incomplete lineage sorting, hybridization, or both.

Gene tree heterogeneity is common in cases of rapid speciation such as the one that has likely given rise to the Delphininae. Incomplete lineage sorting will make the genealogical histories of individual gene loci appear misleading or uninformative about the relationships among species due to retention and stochastic sorting of ancestral polymorphism (Pamilo and Nei, 1988). However, a genetic polymorphism shared among lineages can also result from a gene copy introduced to the population via gene flow if the lineages exchange members, which can be particularly common if they occur in sympatry. It is often very difficult to distinguish between these two processes, and methods that estimate species trees taking into account both the presence of incomplete lineage sorting and hybridization are still in their infancy (Kubatko, 2009). Our analyses of multiple, independent loci suggest that a rapid series of divergences, characterized by short internodes, occurred during the early stages of diversification of the Delphininae (Supplementary Figs. S1 and S2), which suggests that incomplete lineage sorting is affecting the inference of phylogenetic relationships in this group. This was also clear in other molecular studies (Amaral et al., 2007b), where the failure to recover monophyletic groups was attributed to this process. However, using phylogenetic methods that account for this process has not yielded a fully

Please cite this article in press as: Amaral, A.R., et al. Species tree of a recent radiation: The subfamily Delphininae (Cetacea, Mammalia). Mol. Phylogenet. Evol. (2012), http://dx.doi.org/10.1016/j.ympev.2012.04.004

resolved species phylogeny. One of the reasons may be that other factors, such as hybridization, are affecting species history, thereby further confounding the inference of phylogenetic relationships. Hybridization in cetaceans has been reported to occur both in captivity and in the wild (Bérubé, 2002) and could in fact be more common than previously thought. The low variability found in the cetacean nuclear genome, which has been attributed to its slow pace of evolution (e.g., Jackson et al., 2009), is also likely to affect inference of phylogenetic relationships in this group. In order to fully resolve the evolutionary history of the Delphininae, a higher number of informative nuclear markers, coupled with species tree approaches will certainly

5. Conclusion

This study illustrates the complexity of inferring phylogenetic relationships in a group where incomplete lineage sorting, possibly coupled with hybridization events, and a low polymorphism in the nuclear genome make attempts to reconstruct species history challenging. Nevertheless, through the application of different species tree methods we were able to extract an underlying species tree signal from divergent histories of independent genes, which resulted in a new hypothesis of relationships. This new species tree, which is concordant with the relationships established based on morphology, opens up a new avenue of investigation for molecular geneticists, where more genes will have to be used in order to fully clarify the true species phylogeny of this group. Our study shows that the use of multiple loci is likely to result in a more realistic depiction of lineage history than the use of one or a few loci, particularly if analyzed in a coalescent context. Furthermore, our results emphasize the need for coalescent methods that can be applied at the interface of phylogenetic and population processes and that account for both recent rapid speciation events and gene flow between lineages.

Acknowledgments

The authors would like to thank Kelly Robertson, the Southwest Fisheries Science Center Marine Mammal and Turtle Molecular Research Sample Collection (SWFSC), Thomas Jefferson and Marine Sequeira (ICNB) for providing tissue samples. We also thank W.F. Perrin, E. Jockusch, and two anonymous reviewers for comments on the manuscript. Funding was provided by Macquarie University and by Fundação para a Ciência e Tecnologia (Portugal) trough a grant to A.R. Amaral. Part of this work was carried out by using the resources of the Computational Biology Service Unit from Cornell University, which is partially funded by Microsoft Corporation.

Appendix A. Supplementary material

Supplementary data associated with this article can be found, in the online version, at http://dx.doi.org/10.1016/j.ympev.2012.04.004.

References

Amaral, A.R., Sequeira, M., Cedeira-Martínez, J., Coelho, M.M., 2007a. New insights on population genetic structure of *Delphinus delphis* from the northeast Atlantic and phylogenetic relationships within the genus inferred from two mitochondrial markers. Mar. Biol. 151, 1967–1976.

Amaral, A.R., Sequeira, M., Coelho, M.M., 2007b. A first approach to the usefulness of COI barcodes in the identification of closely related delphinid cetacean species. Mar. Freshwater Res. 58, 505–510.

Amaral, A.R., Coelho, M.M., Marugan-Lobon, J., Rohlf, F.J., 2009. Cranial shape differentiation in three closely related delphinid cetacean species: Insights into evolutionary history. Zoology 112, 38–47.

Amaral, A.R., Silva, M.C., Möller, L.M., Beheregaray, L.B., Coelho, M.M., 2010. Anonymous nuclear markers for cetacean species. Conserv. Genet. 11, 1143–1146.

Barnes, L.G., Domning, D.P., Ray, C.E., 1985. Status of studies on fossil marine mammals. Mar. Mammal Sci. 1, 15–53.

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

Bérubé, M., 2002. Hybridism. In: Perrin, W.F., Wursig, B., Theweissen, J.G.M. (Eds.), Encyclopedia of marine mammals. Academic Press Inc., San Diego, Calif., pp. 596–600.

Brumfield, R.T., Liu, L., Lum, D.E., Edwards, S.V., 2008. Comparison of species tree methods for reconstructing the phylogeny of bearded manakins (Aves: Pipridae, *Manacus*) from multilocus sequence data. Syst. Biol. 57, 719–731.

Caballero, S., Jackson, J., Mignucci-Giannoni, A.A., Barrios-Garrido, H., Beltran-Pedreros, S., Montiel-Villalobos, M.G., Robertson, K.M., Baker, C.S., 2008. Molecular systematics of South American dolphins *Sotalia*: sister taxa determination and phylogenetic relationships, with insights into a multilocus phylogeny of the Delphinidae. Mol. Phylogenet. Evol. 46, 252–268.

Cranston, K.A., Hurwitz, B., Ware, D., Stein, L., Wing, R.A., 2009. Species trees from highly incongruent gene trees in rice. Syst. Biol. 58, 489–500.

de Queiroz, A., Gatesy, J., 2007. The supermatrix approach to systematics. Trends Ecol. Evol. 22, 34–41.

Degnan, J.H., Rosenberg, N.A., 2006. Discordance of species trees with their most likely gene trees. PLoS Genet. 2, 762–768.

Degnan, J.H., Rosenberg, N.A., 2009. Gene tree discordance, phylogenetic inference and the multispecies coalescent. Trends Ecol. Evol. 24, 332–340.

Delsuc, F., Scally, M., Madsen, O., Stanhope, M.J., de Jong, W.W., Catzeflis, F.M., Springer, M.S., Douzery, E.J.P., 2002. Molecular phylogeny of living xenarthrans and the impact of character and taxon sampling on the placental tree rooting. Mol. Biol. Evol. 19, 1656–1671.

Dolar, M.L., 2009. Fraser's dolphin *Lagenodelphis hosei*. In: Perrin, W.F., Würsig, B., Thewissen, J.G.M. (Eds.), Encyclopedia of Marine Mammals. Academic Press Inc., San Diego, pp. 469–472.

Dolman, G., Hugall, A.F., 2008. Combined mitochondrial and nuclear data enhance resolution of a rapid radiation of Australian rainbow skinks (Scincidae: *Carlia*). Mol. Phylogenet. Evol. 49, 782–794.

Flower, W.H., 1883. On the characters and divisions of the family Delphinidae. Proc. Zool. Soc. Lond., 466–513.

Forcada, J., 2009. Distribution. In: Perrin, W.F., Würsig, B., Thewissen, J.G.M. (Eds.), Encyclopedia of Marine Mammals. Academic Press Inc., San Diego, pp. 316–321.

Gadagkar, S.R., Rosenberg, M.S., Kumar, S., 2005. Inferring species phylogenies from multiple genes: concatenated sequence tree versus consensus gene tree. J. Exp. Zool. Part. B 304B, 64–74.

Goncalves, H., Martinez-Solano, I., Ferrand, N., Garcia-Paris, M., 2007. Conflicting phylogenetic signal of nuclear vs mitochondrial DNA markers in midwife toads (Anura, Discoglossidae, *Alytes*): deep coalescence or ancestral hybridization? Mol. Phylogenet. Evol. 44, 494–500.

Gygax, L., 2002. Evolution of group size in the superfamily Delphinoidea (Delphinidae, Phocoenidae and Monodontidae): a quantitative comparative analysis. Mammal Rev. 32, 295–314.

Heled, J., Drummond, A.J., 2010. Bayesian inference of species trees from multilocus data. Mol. Biol. Evol. 27, 570–580.

Hennig, W., 1966. Phylogenetic Systematics. Illinois Press, Urbana-Champaign, IL.

Hudson, R.R., 1992. Gene trees, species trees and the segregation of ancestral alleles. Genetics 131, 509–512.

Huelsenbeck, J.P., Ronquist, F., 2001. MRBAYES: Bayesian inference of phylogenetic trees. Bioinformatics 17, 754–755.

Huelsenbeck, J.P., Bull, J.J., Cunningham, C.W., 1996. Combining data in phylogenetic analysis. Trends Ecol. Evol. 11, 152–158.

Hurvich, C.M., Tsai, C.L., 1995. Model selection for extended quasi-likelihood models in small samples. Biometrics 51, 1077–1084.

Jackson, J.A., Baker, C.S., Vant, M., Steel, D.J., Medrano-Gonzalez, L., Palumbi, S.R., 2009. Big and slow: phylogenetic estimates of molecular evolution in baleen whales (suborder Mysticeti). Mol. Biol. Evol. 26, 2427–2440.

Jefferson, T.A., Van Waerebeek, K., 2002. The taxonomic status of the nominal dolphin species *Delphinus tropicalis* van Bree, 1971. Mar. Mammal Sci. 18, 787–818.

Kingston, S.E., Adams, L.D., Rosel, P.E., 2009. Testing mitochondrial sequences and anonymous nuclear markers for phylogeny reconstruction in a rapidly radiating group: molecular systematics of the Delphininae (Cetacea: Odontoceti: Delphinidae). BMC Evol. Biol. 9, 245.

Knowles, L.L., 2009. Estimating species trees: methods of phylogenetic analysis when there is incongruence across genes. Syst. Biol. 58, 463–467.

Kubatko, L.S., 2009. Identifying hybridization events in the presence of coalescence via model selection. Syst. Biol. 58, 478–488.

Kubatko, L.S., Degnan, J.H., 2007. Inconsistency of phylogenetic estimates from concatenated data under coalescence. Syst. Biol. 56, 17–24.

Kubatko, L.S., Carstens, B.C., Knowles, L.L., 2009. STEM: species tree estimation using maximum likelihood for gene trees under coalescence. Bioinformatics 25, 971–973.

LeDuc, R.G., Perrin, W.F., Dizon, A.E., 1999. Phylogenetic relationships among the delphinid cetaceans based on full cytochrome b sequences. Mar. Mammal Sci. 15, 619–648.

Linnen, C., Farrell, B., 2008. Comparison of methods for species tree inference in the aawfly genus *Neodiprion* (Hymenoptera: Diprionidae). Syst. Biol. 57, 876–890.

Lipps, J.H., Mitchell, E., 1976. Trophic model for the adaptive radiations and extinctions of pelagic marine mammals. Paleobiology 2, 147–155.

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.
Liu, L., Pearl, D.K., Brumfield, R.T., Edwards, S.V., 2008. Estimating species trees using multiple-allele DNA sequence data. Evolution 62, 2080–2091.
Liu, L., Yu, L.L., Kubatko, L., Pearl, D.K., Edwards, S.V., 2009. Coalescent methods for estimating phylogenetic trees. Mol. Phylogenet. Evol. 53, 320–328.
Lyons, L.A., Kehler, J.S., O'Brien, S.J., 1999. Development of comparative anchor tagged sequences (CATS) for canine genome mapping. J. Hered. 90, 15–26.
Maddion, W.P., Maddison, D.R., 2009. Mesquite: A Modular System for Evolutionary Analysis. Version 2.72. <http://mesquiteproject.org>.
Maddison, W.P., 1997. Gene trees in species trees. Syst. Biol. 46, 523–536.
Maddison, W.P., Knowles, L.L., 2006. Inferring phylogeny despite incomplete lineage sorting. Syst. Biol. 55, 21–30.
May-Collado, L., Agnarsson, I., 2006. Cytochrome b and Bayesian inference of whale phylogeny. Mol. Phylogenet. Evol. 38, 344–354.
McCracken, K.G., Sorenson, M.D., 2005. Is homoplasy or lineage sorting the source of incongruent mtDNA and nuclear gene trees in the stiff-tailed ducks (*Nomonyx-Oxyura*)? Syst. Biol. 54, 35–55.
McGowen, M.R., Spaulding, M., Gatesy, J., 2009. Divergence date estimation and a comprehensive molecular tree of extant cetaceans. Mol. Phylogenet. Evol. 53, 891–906.
Mead, J.G., 1975. Anatomy of the external nasal passages and facial complex in the Delphinidae (Mammalia: Cetacea). Smithsonian Contrib. Zool. 207, 1–72.
Möller, L.M., Kerstin, B., Charlton-Robb, K., Beheregaray, L., 2008. Multi-gene evidence for a new bottlenose dolphin species in southern Australia. Mol. Phylogenet. Evol. 49, 674–681.
Muizon, C.D., 1988. Les relations phylogénétiques des Delphinida (Cetacea, Mammalia). Annales de Paléontologie 74, 159–227.
Murphy, S., Herman, J.S., Pierce, G.J., Rogan, E., Kitchener, A.C., 2006. Taxonomic status and geographical cranial variation of common dolphins (*Delphinus*) in the eastern north Atlantic. Mar. Mammal Sci. 22, 573–599.
Natoli, A., Peddemors, V.M., Rus-Hoelzel, A., 2004. Population structure and speciation in the genus *Tursiops* based on microsatellite and mitochondrial DNA analyses. J. Evolution. Biol. 17, 363–375.
Niemiller, M.L., Fitzpatrick, B.M., Miller, B.T., 2008. Recent divergence with gene flow in Tennessee cave salamanders (Plethodontidae: *Gyrinophilus*) inferred from gene genealogies. Mol. Ecol. 17, 2258–2275.
Pamilo, P., Nei, M., 1988. Relationships between gene trees and species trees. Mol. Biol. Evol. 5, 568–583.
Perrin, W.F., 1997. Development and homologies of head stripes in the delphinoid cetaceans. Mar. Mammal Sci. 13, 1–43.
Perrin, W.F., 2009. Common dolphins, *Delphinus delphis* and *D. capensis*. In: Perrin, W.F., Würsig, B., Thewissen, J.G.M. (Eds.), Encyclopedia of Marine Mammals. Academic Press Inc., San Diego, pp. 255–259.
Perrin, W.F., Mitchell, E.D., Mead, J.G., Caldwell, D.K., Caldwell, M.C., Vanbree, P.J.H., Dawbin, W.H., 1987. Revision of the spotted dolphins, *Stenella* spp. Mar. Mammal Sci. 3, 99–170.
Perrin, W.F., Robertson, K.M., Van Bree, P.J.H., Mead, J.G., 2007. Cranial description and genetic identity of the holotype specimen of *Tursiops aduncus* (Ehrenberg, 1832). Mar. Mammal Sci. 23, 343–357.
Peters, J.L., Zhuravlev, Y., Fefelov, I., Logie, A., Omland, K.E., 2007. Nuclear loci and coalescent methods support ancient hybridization as cause of mitochondrial paraphyly between gadwall and falcated duck (*Anas* spp.). Evolution 61, 1992–2006.
Posada, D., Buckley, T., 2004. Model selection and model averaging in phylogenetics: advantages of akaike information criterion and Bayesian approaches over likelihood ratio tests. Syst. Biol. 53, 793–808.
Posada, D., Crandall, K.A., 1998. MODELTEST: testing the model of DNA substitution. Bioinformatics 14, 817–818.
Quesada, H., Posada, D., Caballero, A., Moran, P., Rolan-Alvarez, E., 2007. Phylogenetic evidence for multiple sympatric ecological diversification in a marine snail. Evolution 61, 1600–1612.
Rambaut, A., Drummond, A.J., 2007. Tracer v1.4, Available from: <http://beast.bio.ed.ac.uk/Tracer>.
Rambaut, A., Drummond, A.J., 2010. TreeAnnotator v.1.6.1. Available from: <http://beast.bio.ed.ac.uk/TreeAnnotator>.
Roca, A.L., Georgiadis, N., Pecon-Slattery, J., O'Brien, S.J., 2001. Genetic evidence for two species of elephant in Africa. Science 293, 1473–1477.
Rokas, A., Carroll, S.B., 2005. More genes or more taxa? The relative contribution of gene number and taxon number to phylogenetic accuracy. Mol. Biol. Evol. 22, 1337–1344.
Rozas, J., Sanchez-DelBarrio, J.C., Messeguer, X., Rozas, R., 2003. DnaSP, DNA polymorphism analyses by the coalescent and other methods. Bioinformatics 19, 2496–2497.
Sambrook, J., Fritsch, E.F., Maniatis, T., 1989. Molecular Cloning: A Laboratory Manual, vols. 1-3. Cold Spring Harbor Laboratory Press, NY.
Savolainen, V., Anstett, M.C., Lexer, C., Hutton, I., Clarkson, J.J., Norup, M.V., Powell, M.P., Springate, D., Salamin, N., Baker, W.J., 2006. Sympatric speciation in palms on an oceanic island. Nature 441, 210–213.
Seehausen, O., 2004. Hybridization and adaptive radiation. Trends Ecol. Evol. 19, 198–207.
Shapiro, B., Rambaut, A., Drummond, A.J., 2005. Choosing appropriate substitution models of the phylogenetic analysis of protein-coding sequences. Mol. Biol. Evol. 23, 7–9.
Shaw, K.L., 2002. Conflict between nuclear and mitochondrial DNA phylogenies of a recent species radiation: what mtDNA reveals and conceals about modes of speciation in Hawaiian crickets. Proc. Natl. Acad. Sci. USA 99, 16122–16127.
Shimodaira, H., Hasegawa, M., 1999. Multiple comparisons of log-likelihoods with applications to phylogenetic inference. Mol. Biol. Evol. 16, 1114–1116.
Slater, G.J., Price, S.A., Santini, F., Alfaro, M.E., 2010. Diversity versus disparity and the radiation of modern cetaceans. Proc. R. Soc. B-Biol. Sci. 277, 3097–3104.
Steel, M., Penny, D., 2000. Parsimony, likelihood and the role of models in molecular phylogenetics. Mol. Biol. Evol. 17, 839–850.
Steeman, M.E., Hebsgaard, M.B., Fordyce, R.E., Ho, S.Y.W., Rabosky, D.L., Nielsen, R., Rahbek, C., Glenner, H., Sorensen, M.V., Willerslev, E., 2009. Radiation of extant cetaceans driven by restructuring of the oceans. Syst. Biol. 58, 573–585.
Stephens, M., Donnelly, P., 2003. A comparison of Bayesian methods for haplotype reconstruction from population genotype data. Am. J. Hum. Genet. 73, 1162–1169.
Stephens, M., Smith, N.J., Donnelly, P., 2001. A new statistical method for haplotype reconstruction from population data. Am. J. Hum. Genet. 68, 978–989.
True, F.W., 1889. Contributions to the natural history of the cetaceans, a review in the family Delphinidae. Bull. U.S. Nat. Mus. 36, 191.
Vilstrup, J.T., Ho, S.Y.W., Foote, A.D., Morin, P.A., Kreb, D., Krutzen, M., Parra, G.J., Robertson, K.M., de Stephanis, R., Verborgh, P., Willerslev, E., Orlando, L., Gilbert, M.T.P., 2011. Mitogenomic phylogenetic analyses of the Delphinidae with an emphasis on the Globicephalinae. BMC Evol. Biol 11, 65.
Wang, J.Y., Chou, L.S., White, B.N., 2000. Osteological differences between two sympatric forms of bottlenose dolphins (genus *Tursiops*) in Chinese waters. J. Zool. 252, 147–162.
Wiens, J.J., Engstrom, T.N., Chippindale, P.T., 2006. Rapid diversification, incomplete isolation, and the "speciation clock" in North American salamanders (Genus *Plethodon*): Testing the hybrid swarm hypothesis of rapid radiation. Evolution 60, 2585–2603.
Xiong, Y., Brandley, M.C., Xu, S.X., Zhou, K.Y., Yang, G., 2009. Seven new dolphin mitochondrial genomes and a time-calibrated phylogeny of whales. BMC Evol. Biol. 9, 13.

Please cite this article in press as: Amaral, A.R., et al. Species tree of a recent radiation: The subfamily Delphininae (Cetacea, Mammalia). Mol. Phylogenet. Evol. (2012), http://dx.doi.org/10.1016/j.ympev.2012.04.004