Org Divers Evol (2016) 16:467–480
DOI 10.1007/s13127-016-0275-5

ORIGINAL ARTICLE

ORGANISMS DIVERSITY & EVOLUTION

Examining the sensitivity of molecular species delimitations to the choice of mitochondrial marker

Andrew M. Ritchie[1] · **Nathan Lo**[1] · **Simon Y. W. Ho**[1]

Received: 15 September 2015 / Accepted: 4 March 2016 / Published online: 15 March 2016

Abstract Defining and understanding species diversity forms the basis of a wide range of biological and conservation work. Traditional taxonomy can be complemented and accelerated using molecular methods of species delimitation, such as the widely used Generalised Mixed Yule-Coalescent (GMYC) approach. This method uses time-calibrated phylogenetic trees in order to identify transition points between inter- and intraspecific divergence processes. Despite some important limitations, the GMYC approach appears to be robust to a wide range of dataset characteristics. It is one of the few model-based species-delimitation methods that remain practical for analysing molecular datasets with a large numbers of taxa. Most GMYC analyses have been based on datasets consisting of one or a small number of mitochondrial genes. To investigate the sensitivity of GMYC to the choice of mitochondrial marker, we compared GMYC estimates from 15 mitochondrial genes for three vertebrate datasets (cetaceans, ursids and whitefish). Despite the shared evolutionary history among mitochondrial genes, different markers exhibited substantial variation in GMYC delimitation results across all three datasets. This variability was not restricted to specific genes or taxa and extended to commonly used barcoding genes such as *COI* and *CYTB*. Using multiple concatenated markers mitigated these problems in two of the datasets, but exacerbated systematic biases present in a third. Our findings indicate the need to consider multiple markers, loci and lines of evidence when performing molecular species delimitation.

Keywords Species delineation · GMYC · Generalised Mixed Yule-Coalescent · Mitochondrial genome · Phylogenetic analysis

Introduction

Molecular data play an increasingly important role in modern taxonomic practice. Species delimitation is one area in which the availability of abundant molecular data has had a particularly transformative influence. In some cases, molecular characters are able to delimit species even when observable morphological variation is absent or misleading (Bickford et al. 2007). Moreover, the fact that molecular evolution proceeds by similar mechanisms in most organisms means that the same methods can be applied to widely different taxa. These advantages have motivated the development of numerous species-delimitation methodologies (Fujita et al. 2012; Sites and Marshall 2003). Such methods can rapidly generate and test initial species hypotheses in newly discovered groups of organisms for which detailed morphological or ecological data are not yet available. A further advantage of these methods is that they make use of explicit models of sequence evolution to quantify taxonomic uncertainty for downstream analyses. The increasing power and theoretical sophistication of these methods offers considerable prospects for complementing traditional taxonomy in producing reliable classifications for evolutionary studies under an appropriate species concept (Sites and Marshall 2003; Fujita et al. 2012; Yang and Rannala 2010, 2014).

Electronic supplementary material The online version of this article (doi:10.1007/s13127-016-0275-5) contains supplementary material, which is available to authorized users.

✉ Andrew M. Ritchie
andrew.ritchie@sydney.edu.au

[1] School of Life and Environmental Sciences, University of Sydney, Edgeworth David Building A11, Sydney, NSW 2006, Australia

One of the most widely used methods of molecular species delimitation is the Generalised Mixed Yule-Coalescent (GMYC; Pons et al. 2006; Fujisawa and Barraclough 2013). This method delimits species under a philosophy based on the unified lineage concept, which defines species as independently evolving lineages above the population level (De Queiroz 2007). This is similar to the evolutionary genetic species concept, which considers taxa as species when they are more distant than can be explained through genetic drift alone (Birky and Barraclough 2009). GMYC models the evolutionary history of a gene family as a mixture of between-species and within-species branching processes. The between-species process is assumed to follow a Yule speciation model (Yule 1924), a stochastic branching process that assumes no extinction. The within-species process is modelled under the coalescent framework (Kingman 1982). Using a time-calibrated phylogenetic tree, the method attempts to find the threshold that divides the between-species branching process from all within-species coalescence processes. The likelihoods of different threshold positions are calculated from the probabilities of the observed branching or coalescence times under the two models. Each lineage that diverges before the most likely threshold time is deemed to be a putative species.

Some evaluations have shown the GMYC approach to produce robust and plausible species hypotheses under a range of conditions (Powell et al. 2011; Talavera et al. 2013), but other studies have noted a number of limitations of the methodology. GMYC has been observed to over-split species when population sizes are large or variable and speciation is rapid (Esselstyn et al. 2012) or when the sampling scheme inadequately represents intra-species diversity (Lohse 2009; Talavera et al. 2013). The assumptions of the method, which include a pure-birth speciation process and an absence of population structure within species, have been criticised as failing to hold in the majority of cases (Hamilton et al. 2014). Furthermore, GMYC may be sensitive to different methods of phylogenetic reconstruction and divergence-time estimation in generating the input tree (Talavera et al. 2013; Tang et al. 2014). As a single-locus method, GMYC also implicitly assumes that the estimated gene tree is representative of the species tree. This is problematic when introgression or incomplete lineage sorting causes incongruence between gene trees and species trees (Funk and Omland 2003).

In spite of its various limitations, GMYC remains an attractive method for generating initial species hypotheses. Although several multi-locus methods exist that are capable of overcoming many of the problems described above (Yang and Rannala 2010; O'Meara 2010; Grummer et al. 2014; Ence and Carstens 2011), these are too computationally demanding to be applicable to large taxon sets. In addition, some of the methods require the specification of potential delimitation hypotheses on the basis of a priori knowledge that may be unavailable for novel or understudied organisms. GMYC is distinguished from other available single-locus methods (Sites and Marshall 2003; Hebert et al. 2004; e.g. Puillandre et al. 2012) by its clear modelling of evolutionary processes and ability to quantify uncertainty in the delimitation. Since its introduction, the method has been updated with a heuristic variant allowing multiple thresholds (Monaghan et al. 2009), the ability to generate confidence sets and node support values (Powell 2012), and a Bayesian framework capable of accommodating uncertainty in the underlying phylogeny (Reid and Carstens 2012). The desirable properties of GMYC ensure that continued efforts to investigate and extend its utility under different conditions are worthwhile.

One feature of GMYC studies that has received little attention is the sensitivity of the species delimitations to the choice of marker. Species are most commonly delimited using one or a small number of mitochondrial markers due to the uniparental and non-recombinant nature of animal mitochondrial inheritance (Avise et al. 1987), as well as the demonstrated ability of mitochondrial fragments such as *COI* to discriminate effectively between species in multiple branches of the tree of life (Ratnasingham and Hebert 2007). In a survey of 109 publications using GMYC published from January 2013 to December 2014, the overwhelming majority included only a single mitochondrial marker in the analysis, and for 55 % of studies, this was the only marker analysed (Fig. 1). This widespread use of protocols involving single markers calls for an improved understanding of how the choice of marker affects the results of the delimitation.

Since mitochondrial genes are completely linked, all mitochondrial genes might naively be expected to produce similar species delimitations. However, there is substantial evidence

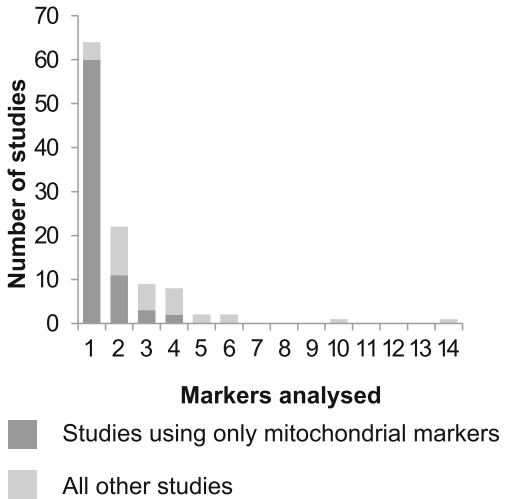

Fig. 1 Number of markers analysed in a survey of 109 studies using GMYC published between January 2013 and December 2014. *Columns* indicate the total number of studies in the sample using the relevant number of markers. The *lower dark grey portion of each column* refers to the number of studies using only mitochondrial markers, while the *upper light grey portion* refers to all other studies

that the mitogenome is subject to selection (reviewed by Galtier et al. 2009; Dowling et al. 2008). This can cause different mitochondrial genes to evolve at different rates (Saccone et al. 2000), exhibit different overall levels of among-lineage rate variation (Mueller 2006; Eo and DeWoody 2010), and produce varying estimates of evolutionary relationships and divergence times (Duchene et al. 2011; Willerslev et al. 2009; Talavera and Vila 2011). As a consequence, mitochondrial genes are known to vary in their ability to recover relationships among species (Havird and Santos 2014; Willerslev et al. 2009), and GMYC delimitation has been shown to differ between mitochondrial and nuclear genes (Tang et al. 2012). However, the properties of individual mitochondrial markers when used for species delimitation with GMYC have not been examined in detail. For well-studied taxa, the choice can reasonably be made based on existing barcode libraries or other previous work; for instance, the barcoding region of the *COI* gene is often used due to its high resolution at both inter- and intra-species levels (Hebert et al. 2003). However, the primary application for automated species delimitation is in novel or taxonomically difficult groups of organisms. The degree to which mitochondrial marker choice can affect the results of species delimitation with GMYC remains unknown.

Here, we analyse mitochondrial protein-coding and ribosomal RNA (rRNA) genes from three groups of vertebrates with varying biological and taxonomic characteristics. The results are compared among individual markers with reference both to the number of species they propose and their ability to recover currently recognised taxonomic groupings. We also assess the utility of combining multiple markers in a concatenated alignment, an approach used in some recent studies (Paz and Crawford 2012) and which might serve as a means of increasing the reliability of the delimitation. By leveraging the increasing amount of publicly available mitogenomic data and the power and speed of modern phylogenetic methods, we are able to provide important insights into the effects of marker choice on species delimitation using the GMYC method.

Materials and methods

Datasets

In order to investigate the effects of marker choice on GMYC performance in different taxonomic situations, we analysed three mitogenomic datasets (Table 1). These datasets represent a range of taxon numbers, evolutionary timescales and biological characteristics and included: (i) 357 unique sequences representing 47 extant cetacean species (whales, dolphins and porpoises); (ii) 80 unique sequences representing all eight extant species of ursids (bears); and (iii) 54 sequences from the European whitefish species complex *Coregonus lavaretus* and its close allies. We restricted our datasets to unique haplotypes because GMYC is unable to accommodate branches with zero length (Fujisawa and Barraclough 2013). Species with only a single representative were allowed to remain, as

Table 1 Numbers of unique sequences and alignment lengths of datasets used for phylogenetic reconstruction and comparison of species-delimitation results using GMYC

Gene	Cetaceans		Bears		Whitefish	
	Unique sequences	Variable/total sites	Unique sequences	Variable/total sites	Unique sequences	Variable/total sites
Concat.	357	7337/13,766	80	5008/13,883	54	4773/14,021
12S	118	392/915	28	221/937	11	139/947
16S	133	548/1487	37	388/1576	10	263/1671
ATP6	168	449/678	37	298/678	10	256/681
ATP8	102	121/189	23	101/201	–	–
COI	187	660/1542	38	530/1542	16	508/1548
COII	143	383/681	32	244/681	11	188/690
COIII	138	360/783	30	276/783	8	261/783
CYTB	186	589/1137	43	439/1137	22	415/1140
NAD1	183	483/954	33	327/948	15	385/972
NAD2	205	628/1041	35	403/1041	20	501/1047
NAD3	116	195/345	26	142/345	8	151/348
NAD4L	194	171/294	26	117/294	21	95/294
NAD4	103	784/1377	44	549/1377	10	578/1380
NAD5	222	1197/1818	48	766/1818	22	774/1836
NAD6	140	364/525	35	198/525	8	212/519

the method is robust to a moderate number of singletons (Talavera et al. 2013).

Complete mitogenome sequences were obtained from GenBank (accession numbers are available in Supplementary Table S3). The sequences of 2 ribosomal RNA genes and 13 protein-coding genes were extracted from each unique mitogenome.

Phylogenetic analyses

Sequences for each of the 15 mitochondrial genes were aligned using MUSCLE (Edgar 2004) as implemented in MEGA v6.06 (Tamura et al. 2013). Sites with >50 % gaps were removed from all alignments. The 12S and 16S rRNA genes contained some portions that could not be unambiguously aligned, and so we employed Gblocks v0.91b (Castresana 2000; Talavera and Castresana 2007) to find sequence blocks that could be aligned with confidence. We used the default specifications in Gblocks, except for allowing positions with up to 50 % gaps. Maximum-likelihood trees for individual genes and for the concatenated genes were estimated with RAxML v8.0.17 (Stamatakis 2014) using the general time-reversible substitution model with gamma-distributed rate heterogeneity among sites. For protein-coding genes, independent substitution models were assigned to the first and second versus third codon sites. For the three concatenated alignments, this was done for all protein-coding genes combined, with the RNA genes grouped together in a third partition. Input trees for GMYC were selected from the best of 200 likelihood searches. Branch support values were calculated from 100 bootstrap replicates.

Divergence-time estimates were obtained using penalised likelihood as implemented in the programme treePL (Smith and O'Meara 2012). This method penalises large rate changes between adjacent branches in the tree, based on the assumption that substitution rates change gradually through time (Sanderson 2002). Although some studies have indicated that Bayesian dating methods such as BEAST (Drummond et al. 2012) may produce more consistent chronograms for use with GMYC (Talavera et al. 2013; Weigand et al. 2013), this approach is problematic for our purposes since it requires specification of a tree prior based on either a speciation (birth-death) or population (coalescent) model, both of which are misspecified for data containing multiple individuals per species. A structured coalescent prior capable of handling datasets with discrete subpopulations has also been implemented for BEAST 2 (Bouckaert et al. 2014), but this requires users to assign sequences to demes before conducting the analysis (Vaughan et al. 2014). When using short alignments with few calibrations, the prior may have a strong influence on the results, making it difficult to isolate the effects of marker choice. For the penalised likelihood search in treePL, we first determined suitable optimization parameters in a priming step. The rate-smoothing parameter was estimated using the included random cross-validation process and the algorithm was run to convergence on each input tree. For the sake of consistency, all gene trees within each case study were analysed using an identical set of calibrations.

Cetaceans

In addition to the 357 sequences from crown cetaceans, the mitogenome of *Hippopotamus amphibius* was included in the initial dataset for the sake of calibration. The position of the root of the Cetacea-*Hippopotamus* tree was estimated using the domestic horse *Equus caballus* as an outgroup. This procedure was necessary because an estimate of the position of the root along the Cetacea-*Hippopotamus* branch was required for divergence-time estimation. To calibrate the molecular clock, we followed Steeman et al. (2009) in using a range of 53.6–56 Myr ago for the divergence of *Hippopotamus* from all cetaceans and a minimum bound of 35 Myr for the age of crown Cetacea. While GMYC does not require the input tree to be scaled in absolute times, the use of fossil calibrations allows straightforward comparison of the divergence-time estimates with those from other studies. Moreover, the inclusion of multiple fossil calibrations can potentially improve divergence-time estimation (Duchêne et al. 2014). We were unable to use potential calibrations for other nodes because those nodes were absent from one or more gene trees. *Hippopotamus* was not included in the GMYC analysis.

Bears

The grey seal *Halichoerus grypus* was included in the dataset to allow the position of the root to be estimated. The phylogenetic relationships represented by each gene tree were sufficiently different that none of the available palaeontological calibrations could be applied across all trees. Instead, we used a secondary calibration of 17.9–22.1 Myr for the crown age of bears based on molecular estimates made by Krause et al. (2008). The use of secondary calibrations is deprecated by some authors (see Hipsley and Muller 2014), but this practice is sufficient for the present study because the absolute divergence dates are not of primary interest.

European whitefish

Our third dataset focused on the whitefish *C. lavaretus*, which is suspected to be a species complex (Nelson 2006). The status of this group is deeply uncertain, with proposals ranging from a single polymorphous species (Bodaly et al. 1991; Bernatchez and Dodson 1994) to a profusion of intralacustrine species flocks (Douglas et al. 1999). Any putative species in this group are expected to have diverged recently, and no clear data on divergence times are available. For this reason, we

used an expanded dataset consisting of 96 sequences representing major salmonid taxa. Our calibrations were based on fossil and molecular estimates made by Crete-Lafreniere et al. (2012). Accordingly, we assumed that the root of the salmonid tree lay between the subfamily Thymallinae and all other salmonids. We then assigned a fixed age of 50 Myr for the MRCA of the remaining 92 sequences and a range of 16.8–22.5 Myr ago for the divergence between the genera *Brachymystax* and *Hucho*. Following divergence-time estimation using penalised likelihood, the subtree containing all sequences from *C. lavaretus*, *Coregonus oxyrinchus*, *Coregonus clupeaformis* and *Coregonus nasus*, totalling 54 sequences in the concatenated alignment, was extracted for GMYC analysis. The *ATP8* gene tree could not be analysed because there was no observed variation in this gene in *C. lavaretus*.

Species delimitation

Maximum-likelihood species delimitations were obtained for each dated tree using the single-threshold GMYC model implemented in the SPLITS package v1.0.19 (Ezard et al. 2009) for R v3.0.3 (R Core Team 2013). A multiple-threshold model is also available, but evaluations suggest that it rarely improves delimitation results (Weigand et al. 2013; Talavera et al. 2013) and we do not consider it further (but see Supplementary Figs. S2 and S3). Each GMYC analysis produces a maximum-likelihood delimitation of individuals into species and a 95 % confidence set of delimitations. Significance is determined by a likelihood-ratio test of the best-known delimitation against a null model in which all individuals belong to a single species. The true number of species in each dataset was unknown, so we evaluated the accuracy of delimitations by comparing the number of species delimited by each gene with the number of named species in the dataset. The nomenclature constitutes an independent assessment of the taxonomy, as all of the species names in our datasets have been established at least partly on the basis of morphology. We also compared the number of exact matches between the GMYC delimitation and existing species names using the comp.delimit function in SPLITS.

Results

Phylogenetics and taxonomy

In each of the three case studies, phylogenetic relationships were generally well resolved by the concatenated data. Greater uncertainty was observed in trees estimated from individual genes (Supplementary Fig. S1). Substantial incongruence was also observed between trees estimated from different mitochondrial genes.

Cetaceans

In the analysis of cetacean mitogenomes all named species were recovered as monophyletic groups (Fig. 2), although not all gene trees were able to resolve the bottlenose dolphin species *Tursiops aduncus* and *Tursiops truncatus*. However, the GMYC procedure split several named species into multiple subgroups. The fin whales (*Balaenoptera physalus*) and Indo-Pacific bottlenose dolphins (*T. aduncus*) were split into two species each. The Cuvier's and Blainville's beaked whales (*Ziphius cavirostris*, *Mesoplodon densirostris*) and the common bottlenose dolphin (*T. truncatus*) were split into three or more species each, including some that were represented by a single individual. Other relationships generally agreed with those inferred by previous mitochondrial studies (Vilstrup et al. 2011).

Bears

Relationships among bear species were well resolved by the analysis of mitogenomic data (Fig. 3) and were broadly similar to those found in other mitogenomic studies (Yu et al. 2007; Krause et al. 2008). The polar bears (*Ursus maritimus*) were nested within a paraphyletic *Ursus arctos* (brown bear) lineage group, as previously observed in several analyses (e.g. Cronin et al. 1991; Talbot and Shields 1996; Shields et al. 2000; reviewed Davison et al. 2011). As in previous studies, the polar bears formed the sister group to brown bears from the Alaskan Alexander Archipelago. Gene-tree incongruence was most noticeable in this dataset, with all branches above species level exhibiting some variation among gene trees. The species hypothesis generated from the concatenated alignment recognised five 'species' within *U. arctos*+*maritimus*: polar bears, Alexander Archipelago bears, the single Kodiak Island specimen, the single French specimen and a Northwest Eurasian clade. This analysis also resulted in almost every other bear specimen being placed in its own species, giving rise to the greatly inflated species count and poor congruence with taxonomic names (Figs. 5 and 6).

European whitefish

Relationships among higher taxa in the salmonid tree (not shown) were generally well resolved and reflected those found by Crete-Lafreniere et al. (2012). Within the *C. lavaretus* complex (Fig. 4), we found strong support for the previously observed European Northeast and Southwest clades (Østbye et al. 2005; Jacobsen et al. 2012). As in previous studies, the North Sea houting *C. oxyrinchus* was found to be non-monophyletic and present in both *C. lavaretus* lineages. All three significant GMYC analyses (*CYTB*, *NAD4* and the concatenated data) delimited these two lineages and the American lake whitefish *C. clupeaformis* as separate species.

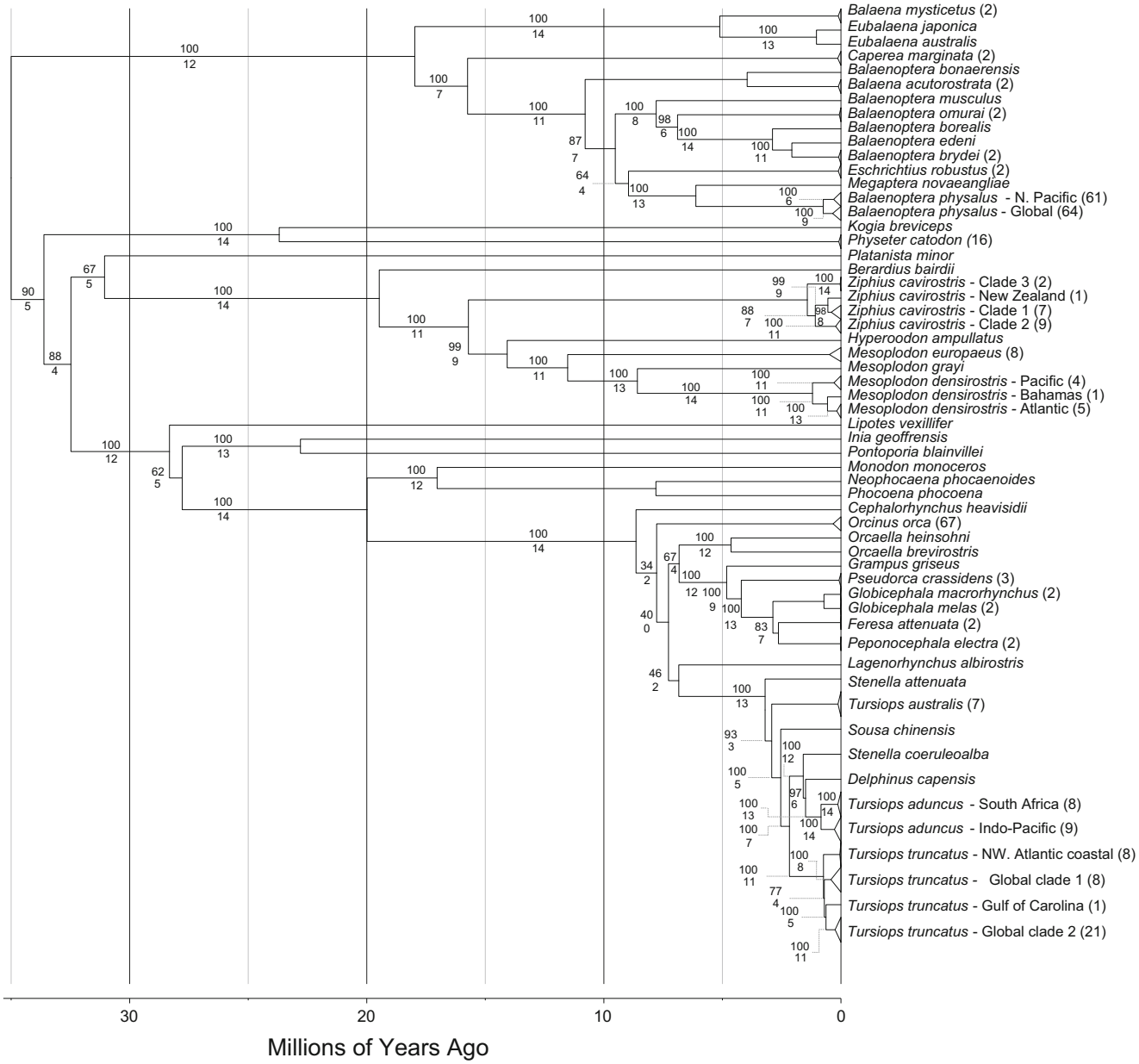

Fig. 2 Chronogram showing the relationships and divergence times for 357 cetaceans, estimated from a concatenated mitochondrial dataset comprising all 13 protein coding and 2 ribosomal RNA genes. Groups delimited as species by the GMYC analysis are shown as *triangles*. The horizontal axis shows the timescale, measured in millions of years. Bootstrap values ($n = 100$) are shown above each branch, while the numbers below each branch indicate the number of gene trees in which the branch was observed ($n = 15$). Numbers are not shown where the branch is present across all gene trees and bootstrap replicates

All other genes failed to reject a one-species null model. The position of *C. nasus* was poorly resolved.

Effects of marker choice on GMYC species delimitation

Species counts

For the bear and cetacean datasets, the estimated species counts varied substantially among individual genes (Fig. 5). The maximum-likelihood estimates for the number of cetacean species ranged from 23 (*12S*) to 67 (*COIII*) compared with the 47 named species, whereas for the bear data, the estimates ranged from 7 (*COI*, *COIII*, *NAD1*) to 22 (concatenated sequences) compared with 8 named species. The effect was less visible in whitefish because only two genes (*CYTB* and *NAD4*) produced significant delimitations, but nevertheless these two genes did identify different numbers of hypothetical species (4 and 5, respectively). The

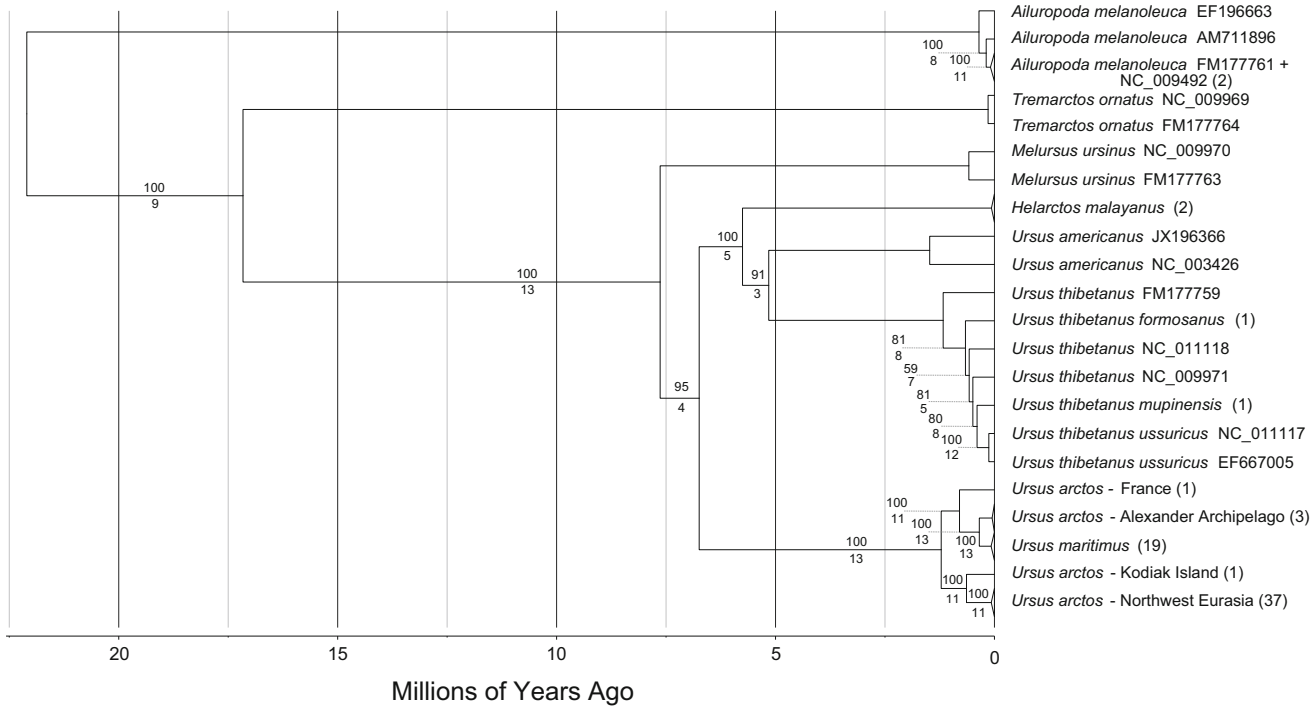

Fig. 3 Chronogram showing the relationships and divergence times for 80 bears, estimated from a concatenated mitochondrial dataset comprising all 13 protein coding and 2 ribosomal RNA genes. Groups delimited as species by the GMYC analysis are shown as *triangles*. The horizontal axis shows the timescale, measured in millions of years. Bootstrap values ($n = 100$) are shown above each branch, while the numbers below each branch indicate the number of gene trees in which the branch was observed ($n = 15$). Numbers are not shown where the branch is present across all gene trees and bootstrap replicates

tendency of individual genes to lump or split taxa was not always consistent between case studies; for example, *16S*, *COI* and *COIII* showed opposite biases in the bear and cetacean results.

Overall, the GMYC method exhibited reasonable performance in enumerating bear and cetacean species, with the number of named species falling within the 95 % confidence interval for the number of delimited species for most genes.

Fig. 4 Chronogram showing the relationships and divergence times for 54 individuals from the *Coregonus lavaretus* species complex and close allies, estimated from a concatenated mitochondrial dataset comprising all 13 protein coding and 2 ribosomal RNA genes. Groups delimited as species by the GMYC analysis are shown as *triangles*. The horizontal axis shows the timescale, measured in millions of years. Bootstrap values ($n = 100$) are included above each branch, while the numbers below each branch indicate the number of gene trees in which the branch was observed ($n = 15$). Numbers are not shown where the branch is present across all gene trees and bootstrap replicates

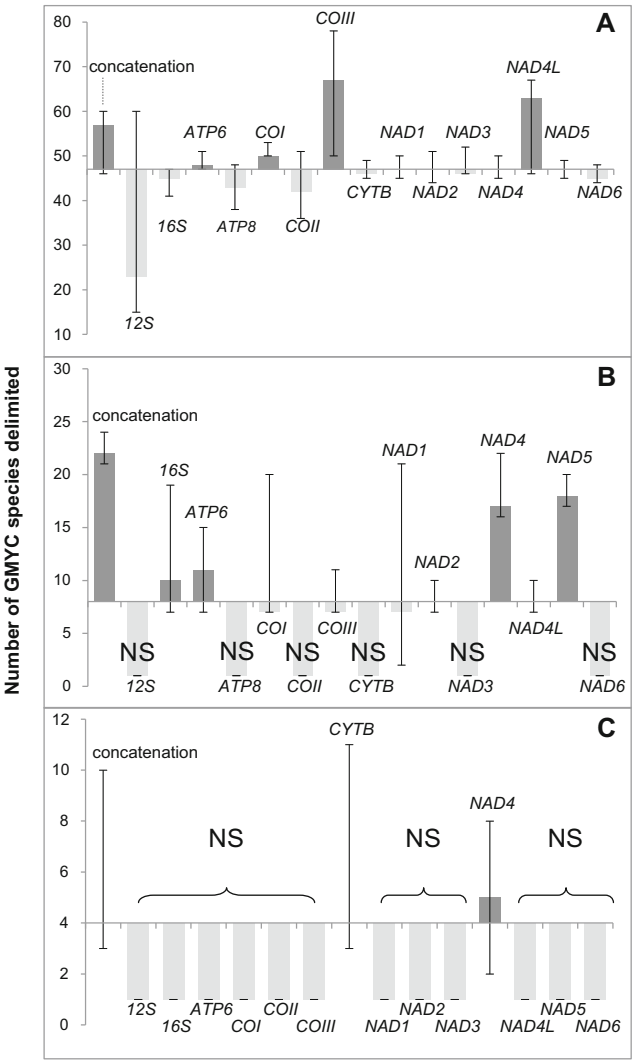

Fig. 5 Numbers of species delimited by different mitochondrial genes in GMYC analysis of three datasets: **a** cetaceans; **b** bears (Ursidae); and **c** European whitefish (*Coregonus lavaretus*) and allies. In each case, the value at which the horizontal axis crosses the vertical axis corresponds to the number of named species (47 cetaceans, 8 bears, 4 whitefish). Columns represent the number of species delimited in the most likely hypotheses identified by GMYC using individual genes, while *error bars* show the range of species counts found in the 95 % confidence set of species hypotheses generated by GMYC. Genes marked 'NS' did not provide sufficient evidence to reject the one-species null hypothesis (likelihood-ratio test, $p > 0.05$). Substantial variation is observed between delimitations given by different genes

This was also the case for whitefish, although this result is less meaningful because the taxonomy represented by current nomenclature is widely disputed (Bernatchez and Dodson 1994). Notably, both the widely used barcoding gene *COI* in cetaceans and the concatenated alignment in bears strictly overdelimited taxa with respect to the existing nomenclature. The bear and whitefish results exhibited much greater overall uncertainty than the cetacean results. Six bear genes and all but two whitefish genes failed to reject the one-species null hypothesis.

Similarity to currently proposed classifications

We found disparities between the GMYC-delimited and currently proposed classifications, with the number of exact matches varying among gene trees (Fig. 6). This measure of similarity is more sensitive than species counts to differences in tree topology. Although some genes produced species counts that agreed with the number of named species, none of the trees produced a delimitation that exactly recapitulated the currently proposed taxonomy. In cetaceans, *NAD2*, *NAD4* and *NAD5* were the most consistent, whereas *12S* was the least consistent. For the bear data, the concatenated sequence

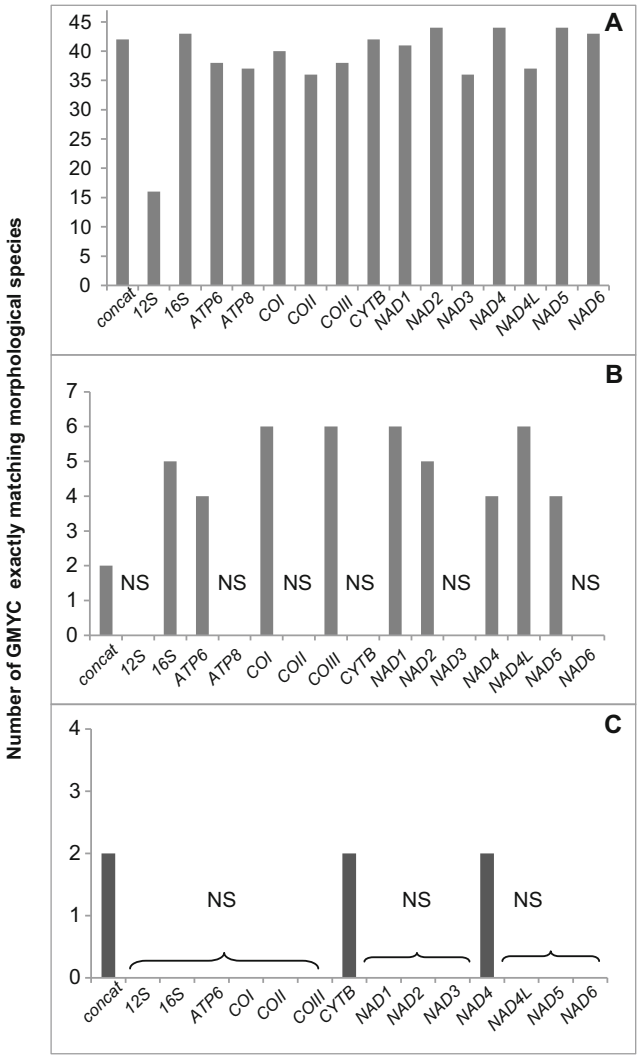

Fig. 6 Congruence of GMYC species delimitations to named taxonomy using different mitochondrial genes, as measured by the number of exact matches between GMYC entities and named species (**a** cetaceans; **b** bears; **c** European whitefish). Substantial variation in the performance of individual genes is observed. None of the GMYC estimates was able to recover all named species, even when there was a match in the number of delimited species

unexpectedly produced the worst match to the recognised taxonomy, identifying only two of the eight named species (the polar bear *U. maritimus* and the sun bear *Helarctos malayanus*).

Discussion

Species delimitation methods should, ideally, infer the same set of species regardless of the choice of genetic marker, as long as all markers share the same evolutionary history. Our analysis shows that there can be considerable variation in species delimitations using the GMYC method even among genes from a set of mitogenomes. Inferences from different mitochondrial genes varied both in the overall number of species delimited and in the number of named species recovered. These broad patterns could be seen, to some degree, in all three of our case studies of vertebrate taxa, demonstrating that the phenomenon is not due to the particular evolutionary characteristics of any one taxon. Moreover, the variation was ubiquitous across the data and was not restricted to one or a small number of genes. This suggests an innate sensitivity of the GMYC method to the choice of mitochondrial marker.

In explaining the factors responsible for the observed patterns of variation in our GMYC results, we assume that the mitogenome has not undergone recombination in the taxa analysed in our case studies, such that disagreement among genes should not be the result of differences in lineage sorting. Although this possibility cannot be excluded entirely, mitochondrial recombination has not been observed in cetaceans or bears and only in a single individual among Atlantic salmon (*Salmo salar*), a distant relative of the European whitefish (Ciborowski et al. 2007). Instead, the observed variation is probably due to factors that have produced conflicting evolutionary signals across mitochondrial genes. Our results suggest that these differences can have significant effects on downstream analysis with GMYC.

Differences in the strength and direction of selection across genes can lead to contrasting patterns of branch lengths among the gene trees. This could cause problems for the GMYC method, even supposing that branch lengths were estimated accurately. The GMYC method relies on the presence of a sharp increase in the birth rate of lineages across the speciation threshold due to a disparity between the rates of speciation and coalescence (Fujisawa and Barraclough 2013). However, an excess of slightly deleterious mutations can inflate the lengths of branches towards the tips of the tree (Ho et al. 2005, 2011; Williamson and Orive 2002). In contrast, mutational saturation can lead to underestimation of the lengths of deep branches (Phillips 2009). This might be particularly problematic in analyses of mitogenomes because of their rapid rates of evolution in animals. Moreover, inaccurate estimation of branch lengths can mislead topological inference. This can cause different mitochondrial genes to exhibit incongruent gene trees, despite their shared phylogenetic history. This cannot be ruled out for our own sequences. Using Xia's test of mutational saturation in DAMBE v5 (Xia 2013), we found evidence of saturation in the third codon positions of several bear and cetacean genes (Supplementary Table S4). When estimating tree topologies, this can be dealt with by removing or recoding saturated sites; however, in the case of species delimitation, this is likely to remove much of the within-species variation that is used by GMYC.

The sensitivity of the GMYC method to marker choice has implications for the design of future species-delimitation protocols. GMYC analyses have typically relied on single mitochondrial markers (e.g. Papadopoulou et al. 2008) or small numbers of concatenated mitochondrial genes, sometimes together with nuclear markers (e.g. Pons et al. 2006; Hendrixson et al. 2013). Our findings indicate that choosing different mitochondrial genes in such studies can produce different results, representing an additional, substantial source of uncertainty that is not normally accounted for.

Choosing the best gene for a given analysis is likely to be a difficult problem, especially in studies of novel taxa for which barcode libraries are not available. A common approach has been to exploit the properties of popular barcoding genes such as *COI* in animals. Although such markers are useful in some groups (e.g. Luo et al. 2011), they are unlikely to be truly universal (Galtier et al. 2009). The performance of *COI* in the present study was questionable, although our study differs from barcoding practice in that we analysed the entire gene rather than a fragment. For cetaceans it recovered fewer named species than did many other genes and was one of the few markers not to include the named species count in its 95 % confidence interval. Moreover, it is difficult to know how to proceed when two well-studied barcodes find similar well-resolved monophyletic groups but disagree on the delimitation; this was the case for *COI* and *CYTB* in two of the groups analysed here. If individual genes had characteristic tendencies to over- or under-split species, it would be possible to exclude problematic genes or to design systematic corrections. However, we found that most genes were not consistent in their behaviour across the three case studies, in terms of either splitting tendencies or exact matches to named species. *NAD4* and *NAD5* in particular produced delimitations that were among the closest to the named cetacean taxonomy but among the most contrasting in bears. This fact also shows that sequence length was not a sufficient predictor of performance, as *NAD5* was the longest gene analysed.

Another approach would be to include multiple mitochondrial markers in a concatenated alignment. We trialled this approach using a concatenated alignment of all mitochondrial rRNA and protein-coding genes, with mixed results. In whitefish, the concatenated alignment supported a four-species hypothesis similar to that proposed by the *CYTB* gene.

Meanwhile, the concatenated alignments for cetaceans and bears supported delimitations that substantially over-split taxa with reference to existing species names. These analyses also recovered fewer named species than did some individual mitochondrial genes.

The reasons for the over-delimitation encountered when using the concatenated alignments are uncertain. It is possible that, in some cases, the greater resolution provided by these sequences has allowed them to detect species that are currently unrecognised. In cetaceans, for example, some taxonomists have suggested that the South African *T. aduncus* clade (Natoli et al. 2004) and the Western coastal ecotype of *T. truncatus* (Moura et al. 2013) are incipient species. Additionally, some scenarios for the evolution of fin whales *B. physalus* could result in the North Pacific clade (see Fig. 2) gaining subspecies status (Archer et al. 2013). We observed the opposite situation in our analysis of the European whitefish. The two species delimited by the concatenated alignment in the *C. lavaretus* + *oxyrinchus* complex agree with two of the major refugial lineages that have been proposed for *C. lavaretus* (ØStbye et al. 2005; Jacobsen et al. 2012). However, this estimate is conservative in light of the extreme morphological and ecological variability found in this complex, which has led to a profusion of conflicting taxonomic proposals (Bernatchez and Dodson 1994). In cases like these, comparison with current nomenclature is an inadequate measure of performance.

The GMYC method involves a number of simplifying assumptions that are likely to be violated regularly in practice. If these violations were to lead to inaccurate species delimitations, adding more markers would cause the GMYC method to converge on an incorrect result with increasing levels of support. In addition, GMYC analyses are known to be sensitive to sampling practices across different taxonomic and geographical scales (Talavera et al. 2013; Bergsten et al. 2012; Dinca et al. 2015). All three groups of taxa used in the case studies are known to have several characteristics that could be problematic for GMYC analysis. For instance, the assumption of zero extinction in the Yule branching process is certainly violated by the bear and cetacean datasets, because several extinct lineages are attested from fossil evidence (see Stiller et al. 2014; Uhen 2010). High extinction rates could result in poor fit of the Yule model to the reconstructed speciation times, confusing the inference of the threshold point between among- and within-species branching processes. Evolutionary features that violate the assumptions of the coalescent model, such as population structure, rapid radiation, migration, and hybrid introgression, are also likely to have affected the results of all three case studies. For example, the *C. lavaretus* species complex consists of largely isolated populations distributed among a large number of small European lakes (Hudson et al. 2011) which might have formed relatively recently from a larger water body (Douglas et al. 1999). There is evidence of a history of extensive introgression and recent, rapid adaptive radiation between and within lake habitats (ØStbye et al. 2005; Jacobsen et al. 2012; Hudson et al. 2011). As a result, the true speciation history is unlikely to be resolvable from a single locus. Sampling considerations may also be important in some datasets. Although all three datasets were chosen to include an adequate balance of between- and within-species sampling, the proportion of diversity sampled within each individual species is likely to have varied. The ability to delimit species in the whitefish dataset in particular may be limited by unavailability of sequences for other whitefish species. However, the extent of this problem is unclear because of the possibility of multiple entities existing within the *C. lavaretus* complex. Adding linked markers in cases where sampling problems or model violations are present might not produce a better result, but they could still be helpful in allowing such systematic problems to become apparent. In the analysis of the bear data, for example, the concatenated alignment produced an extreme delimitation in which most individuals were proposed as representing distinct species. This is probably due to the strong population structure evident in the *U. arctos* samples, combined with an imbalanced sampling scheme. While this is clearly a problematic dataset, the fact that analyses using fewer markers produced apparently reasonable results suggests that single markers might fail to uncover such underlying problems.

Future development of the GMYC approach should aim to accommodate sequence data from multiple loci. The ability to incorporate unlinked markers in species delimitation with GMYC would serve to greatly mitigate the marker sensitivity problems observed in the present study, as well as overcoming the limitations of a single-locus method in analysing difficult cases such as the bear and whitefish datasets. Some authors (e.g. Weigand et al. 2013; Hendrixson et al. 2013) have previously used phylogenies estimated from concatenated mitonuclear data as input to GMYC analysis. However, this approach does not take into account the differing evolutionary histories among loci (Kubatko and Degnan 2007) and the effects on species delimitation are unclear. An alternative approach might be to analyse only synonymous substitutions or third codon sites in order to minimise the confounding effects of selection, but this method is likely to be impractical in cases where the number of informative sites is limited.

The application of Bayesian methods for species delimitation is a promising avenue for addressing the shortcomings of the GMYC method, particularly its sensitivity to marker choice. Bayesian phylogenetic methods allow estimates of species assignments to be conditioned on the uncertainty in gene-tree topologies and branch lengths (Yang and Rannala 2010, 2014; Reid and Carstens 2012; Jones et al. 2015). In addition, the simultaneous estimation of tree topologies, branch lengths and species assignments in some methods

enables the appropriate incorporation of uncertainty in the calibrations. This is a potentially important feature because rate-smoothing methods for molecular dating typically do not correctly account for calibration error (Yang 2014; Ho and Duchene 2014). A key shortcoming of Bayesian methods for species delimitation, however, is that they are not computationally feasible for large datasets. In this respect, methods based on the GMYC approach are likely to remain widely used in the near future.

Conclusions

Our results suggest a complex situation in which the species delimitations produced by a GMYC analysis are dependent both on the choice of mitochondrial marker(s) and on the specific evolutionary histories of the organisms being studied. We conclude that marker sensitivity represents an important source of uncertainty that is rarely addressed in studies using the GMYC method. This uncertainty cannot be reliably reduced by the use of any specific mitochondrial gene. Using an input tree inferred from multiple markers in concatenation might help to overcome bias caused by marker choice, though it cannot compensate for violations of the underlying assumptions of the method or for biases and imbalances in the sampling procedure. Our results also emphasise the importance of combining GMYC-based methods with other lines of evidence, such as morphology, ecology and developmental traits, in order to delimit species accurately. Further work on species-delimitation methods will help to improve our understanding of species diversity among vertebrates and other organisms.

Acknowledgements AMR was supported by an Australian Postgraduate Award. SYWH and NL were supported by the Australian Research Council.

References

Archer, F. I., Morin, P. A., Hancock-Hanser, B. L., Robertson, K. M., Leslie, M. S., Berube, M., et al. (2013). Mitogenomic phylogenetics of fin whales (*Balaenoptera physalus* spp.): genetic evidence for revision of subspecies. *PLoS ONE, 8*(5), e63396. doi:10.1371/journal.pone.0063396.

Avise, J. C., Arnold, J., Ball, R. M., Bermingham, E., Lamb, T., Neigel, J. E., et al. (1987). Intraspecific phylogeography: the mitochondrial DNA bridge between population genetics and systematics. *Annual Review of Ecology and Systematics, 18*(1), 489–522. doi:10.1146/annurev.es.18.110187.002421.

Bergsten, J., Bilton, D. T., Fujisawa, T., Elliott, M., Monaghan, M. T., Balke, M., et al. (2012). The effect of geographical scale of sampling on DNA barcoding. *Systematic Biology, 61*(5), 851–869. doi:10.1093/sysbio/sys037.

Bernatchez, L., & Dodson, J. J. (1994). Phylogenetic relationships among palearctic and nearctic whitefish (*Coregonus* sp.) populations as revealed by mitochondrial DNA variation. *Canadian Journal of Fisheries and Aquatic Sciences, 51*(S1), 240–251. doi:10.1139/f94-310.

Bickford, D., Lohman, D. J., Sodhi, N. S., Ng, P. K. L., Meier, R., Winker, K., et al. (2007). Cryptic species as a window on diversity and conservation. *Trends in Ecology & Evolution, 22*(3), 148–155. doi:10.1016/j.tree.2006.11.004.

Birky, C. W., Jr., & Barraclough, T. (2009). Asexual speciation. In I. Schön, K. Martens, & P. Dijk (Eds.), *Lost sex* (pp. 201–216). Netherlands: Springer.

Bodaly, R. A., Vuorinen, J., Ward, R. D., Luczynski, M., & Reist, J. D. (1991). Genetic comparisons of new and old world coregonid fishes*. *Journal of Fish Biology, 38*(1), 37–51. doi:10.1111/j.1095-8649.1991.tb03089.x.

Bouckaert, R., Heled, J., Kühnert, D., Vaughan, T., Wu, C.-H., Xie, D., et al. (2014). BEAST 2: a software platform for Bayesian evolutionary analysis. *PLoS Computational Biology, 10*(4), e1003537–e1003537.

Castresana, J. (2000). Selection of conserved blocks from multiple alignments for their use in phylogenetic analysis. *Molecular Biology and Evolution, 17*(4), 540–552.

Ciborowski, K. L., Consuegra, S., de Leaniz, C. G., Beaumont, M. A., Wang, J. L., & Jordan, W. C. (2007). Rare and fleeting: an example of interspecific recombination in animal mitochondrial DNA. *Biology Letters, 3*(5), 554–557. doi:10.1098/rsbl.2007.0290.

Crete-Lafreniere, A., Weir, L. K., & Bernatchez, L. (2012). Framing the salmonidae family phylogenetic portrait: a more complete picture from increased taxon sampling. *PLoS ONE, 7*(10), e46662. doi:10.1371/journal.pone.0046662.

Cronin, M. A., Amstrup, S. C., Garner, G. W., & Vyse, E. R. (1991). Interspecific and intraspecific mitochondrial DNA variation in North American bears (*Ursus*). *Canadian Journal of Zoology, 69*(12), 2985–2992. doi:10.1139/z91-421.

Davison, J., Ho, S. Y. W., Bray, S. C., Korsten, M., Tammeleht, E., Hindrikson, M., et al. (2011). Late-quaternary biogeographic scenarios for the brown bear (*Ursus arctos*), a wild mammal model species. *Quaternary Science Reviews, 30*(3–4), 418–430. doi:10.1016/j.quascirev.2010.11.023.

De Queiroz, K. (2007). Species concepts and species delimitation. *Systematic Biology, 56*(6), 879–886. doi:10.1080/10631507017010 83.

Dinca, V., Montagud, S., Talavera, G., Hernandez-Roldan, J., Munguira, M. L., Garcia-Barros, E., et al. (2015). DNA barcode reference library for Iberian butterflies enables a continental-scale preview of potential cryptic diversity. *Scientific Reports, 5*, doi:10.1038/srep12395.

Douglas, M. R., Brunner, P. C., & Bernatchez, L. (1999). Do assemblages of *Coregonus* (Teleostei: Salmoniformes) in the Central Alpine region of Europe represent species flocks? *Molecular Ecology, 8*(4), 589–603. doi:10.1046/j.1365-294x.1999.00581.x.

Dowling, D. K., Friberg, U., & Lindell, J. (2008). Evolutionary implications of non-neutral mitochondrial genetic variation. *Trends in Ecology & Evolution, 23*(10), 546–554. doi:10.1016/j.tree.2008.05.011.

Drummond, A. J., Suchard, M. A., Xie, D., & Rambaut, A. (2012). Bayesian phylogenetics with BEAUti and the BEAST 1.7. *Molecular Biology and Evolution, 29*(8), 1969–1973. doi:10.1093/molbev/mss075.

Duchene, S., Archer, F. I., Vilstrup, J., Caballero, S., & Morin, P. A. (2011). Mitogenome phylogenetics: the impact of using single regions and partitioning schemes on topology, substitution rate and divergence time estimation. *PLoS ONE, 6*(11), e27138. doi:10.1371/journal.pone.0027138.

Duchêne, S., Lanfear, R., & Ho, S. Y. W. (2014). The impact of calibration and clock-model choice on molecular estimates of divergence times. *Molecular Phylogenetics and Evolution, 78*, 277–289.

Edgar, R. C. (2004). MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Research, 32*(5), 1792–1797. doi:10.1093/nar/gkh340.

Ence, D. D., & Carstens, B. C. (2011). SpedeSTEM: a rapid and accurate method for species delimitation. *Molecular Ecology Resources, 11*(3), 473–480. doi:10.1111/j.1755-0998.2010.02947.x.

Eo, S. H., & DeWoody, J. A. (2010). Evolutionary rates of mitochondrial genomes correspond to diversification rates and to contemporary species richness in birds and reptiles. *Proceedings of the Royal Society B-Biological Sciences, 277*(1700), 3587–3592. doi:10.1098/rspb.2010.0965.

Esselstyn, J. A., Evans, B. J., Sedlock, J. L., Khan, F. A. A., & Heaney, L. R. (2012). Single-locus species delimitation: a test of the Mixed Yule-Coalescent Model, with an empirical application to Philippine round-leaf bats. *Proceedings of the Royal Society B-Biological Sciences, 279*(1743), 3678–3686. doi:10.1098/rspb.2012.0705.

Ezard, T. H. G., Fujisawa, T., & Barraclough, T. G. (2009). Splits: species limits by threshold statistics. (1.0.19 ed.).

Fujisawa, T., & Barraclough, T. G. (2013). Delimiting species using single-locus data and the Generalized Mixed Yule Coalescent approach: a revised method and evaluation on simulated data sets. *Systematic Biology, 62*(5), 707–724. doi:10.1093/sysbio/syt033.

Fujita, M. K., Leache, A. D., Burbrink, F. T., McGuire, J. A., & Moritz, C. (2012). Coalescent-based species delimitation in an integrative taxonomy. *Trends in Ecology & Evolution, 27*(9), 480–488. doi:10.1016/j.tree.2012.04.012.

Funk, D. J., & Omland, K. E. (2003). Species-level paraphyly and polyphyly: frequency, causes, and consequences, with insights from animal mitochondrial DNA. *Annual Review of Ecology, Evolution, and Systematics, 34*, 397–423. doi:10.1146/annurev.ecolsys.34.011802.132421.

Galtier, N., Nabholz, B., Glemin, S., & Hurst, G. D. D. (2009). Mitochondrial DNA as a marker of molecular diversity: a reappraisal. *Molecular Ecology, 18*(22), 4541–4550. doi:10.1111/j.1365-294X.2009.04380.x.

Grummer, J. A., Bryson, R. W., Jr., & Reeder, T. W. (2014). Species delimitation using Bayes factors: simulations and application to the *Sceloporus scalaris* species group (Squamata: Phrynosomatidae). *Systematic Biology, 63*(2), 119–133. doi:10.1093/sysbio/syt069.

Hamilton, C. A., Hendrixson, B. E., Brewer, M. S., & Bond, J. E. (2014). An evaluation of sampling effects on multiple DNA barcoding methods leads to an integrative approach for delimiting species: a case study of the North American tarantula genus *Aphonopelma* (Araneae, Mygalomorphae, Theraphosidae). *Molecular Phylogenetics and Evolution, 71*, 79–93. doi:10.1016/j.ympev.2013.11.007.

Havird, J. C., & Santos, S. R. (2014). Performance of single and concatenated sets of mitochondrial genes at inferring Metazoan relationships relative to full mitogenome data. *PLoS ONE, 9*(1), e84080. doi:10.1371/journal.pone.0084080.

Hebert, P. D. N., Cywinska, A., Ball, S. L., & DeWaard, J. R. (2003). Biological identifications through DNA barcodes. *Proceedings of the Royal Society B-Biological Sciences, 270*(1512), 313–321. doi:10.1098/rspb.2002.2218.

Hebert, P. D. N., Penton, E. H., Burns, J. M., Janzen, D. H., & Hallwachs, W. (2004). Ten species in one: DNA barcoding reveals cryptic species in the neotropical skipper butterfly *Astraptes fulgerator*. *Proceedings of the National Academy of Sciences of the United States of America, 101*(41), 14812–14817. doi:10.1073/pnas.0406166101.

Hendrixson, B. E., DeRussy, B. M., Hamilton, C. A., & Bond, J. E. (2013). An exploration of species boundaries in turret-building tarantulas of the Mojave Desert (Araneae, Mygalomorphae, Theraphosidae, Aphonopelma). *Molecular Phylogenetics and Evolution, 66*(1), 327–340. doi:10.1016/j.ympev.2012.10.004.

Hipsley, C. A., & Muller, J. (2014). Beyond fossil calibrations: realities of molecular clock practices in evolutionary biology. *Frontiers in Genetics, 5*, 138. doi:10.3389/fgene.2014.00138.

Ho, S. Y. W., & Duchene, S. (2014). Molecular-clock methods for estimating evolutionary rates and timescales. *Molecular Ecology, 23*(24), 5947–5965. doi:10.1111/mec.12953.

Ho, S. Y. W., Phillips, M. J., Cooper, A., & Drummond, A. J. (2005). Time dependency of molecular rate estimates and systematic overestimation of recent divergence times. *Molecular Biology and Evolution, 22*(7), 1561–1568. doi:10.1093/molbev/msi145.

Ho, S. Y. W., Lanfear, R., Bromham, L., Phillips, M. J., Soubrier, J., Rodrigo, A. G., et al. (2011). Time-dependent rates of molecular evolution. *Molecular Ecology, 20*(15), 3087–3101. doi:10.1111/j.1365-294X.2011.05178.x.

Hudson, A. G., Vonlanthen, P., & Seehausen, O. (2011). Rapid parallel adaptive radiations from a single hybridogenic ancestral population. *Proceedings of the Royal Society B-Biological Sciences, 278*(1702), 58–66. doi:10.1098/rspb.2010.0925.

Jacobsen, M. W., Hansen, M. M., Orlando, L., Bekkevold, D., Bernatchez, L., Willerslev, E., et al. (2012). Mitogenome sequencing reveals shallow evolutionary histories and recent divergence time between morphologically and ecologically distinct European whitefish (*Coregonus* spp.). *Molecular Ecology, 21*(11), 2727–2742. doi:10.1111/j.1365-294X.2012.05561.x.

Jones, G., Aydin, Z., & Oxelman, B. (2015). DISSECT: an assignment-free Bayesian discovery method for species delimitation under the multispecies coalescent. *Bioinformatics, 31*(7), 991–998. doi:10.1093/bioinformatics/btu770.

Kingman, J. (1982). The coalescent. *Stochastic Processes and their Applications, 13*, 235–248.

Krause, J., Unger, T., Nocon, A., Malaspinas, A.-S., Kolokotronis, S.-O., Stiller, M., et al. (2008). Mitochondrial genomes reveal an explosive radiation of extinct and extant bears near the Miocene-Pliocene boundary. *BMC Evolutionary Biology, 8*(1), 220.

Kubatko, L., & Degnan, J. (2007). Inconsistency of phylogenetic estimates from concatenated data under coalescence. *Systematic Biology, 56*, 17–24.

Lohse, K. (2009). Can mtDNA barcodes be used to delimit species? A response to Pons et al. (2006). *Systematic Biology, 58*(4), 439–441. doi:10.1093/sysbio/syp039.

Luo, A., Zhang, A., Ho, S. Y. W., Xu, W., Zhang, Y., Shi, W., et al. (2011). Potential efficacy of mitochondrial genes for animal DNA barcoding: a case study using eutherian mammals. *BMC Genomics, 12*, 84. doi:10.1186/1471-2164-12-84.

Monaghan, M. T., Wild, R., Elliot, M., Fujisawa, T., Balke, M., Inward, D. J. G., et al. (2009). Accelerated species inventory on Madagascar using coalescent-based models of species delineation. *Systematic Biology, 58*(3), 298–311. doi:10.1093/sysbio/syp027.

Moura, A. E., Nielsen, S. C. A., Vilstrup, J. T., Moreno-Mayar, J. V., Gilbert, M. T. P., Gray, H. W. I., et al. (2013). Recent diversification of a marine genus (*Tursiops* spp.) tracks habitat preference and environmental change. *Systematic Biology, 62*(6), 865–877. doi:10.1093/sysbio/syt051.

Mueller, R. L. (2006). Evolutionary rates, divergence dates, and the performance of mitochondrial genes in Bayesian phylogenetic analysis. *Systematic Biology, 55*(2), 289–300. doi:10.1080/10635150500541672.

Natoli, A., Peddemors, V. M., & Hoelzel, A. R. (2004). Population structure and speciation in the genus *Tursiops* based on microsatellite and mitochondrial DNA analyses. *Journal of Evolutionary Biology, 17*(2), 363–375. doi:10.1046/j.1420-9101.2003.00672.x.

Nelson, J. S. (2006). *Fishes of the World* (Fourthth ed.). Hoboken: Wiley.

O'Meara, B. C. (2010). New heuristic methods for joint species delimitation and species tree inference. *Systematic Biology, 59*(1), 59–73. doi:10.1093/sysbio/syp077.

ØStbye, K., Bernatchez, L., Næsje, T. F., Himberg, K. J., & Hindar, K. (2005). Evolutionary history of the European whitefish *Coregonus lavaretus* (L.) species complex as inferred from mtDNA phylogeography and gill-raker numbers. *Molecular Ecology, 14*(14), 4371–4387. doi:10.1111/j.1365-294X.2005.02737.x.

Papadopoulou, A., Bergsten, J., Fujisawa, T., Monaghan, M. T., Barraclough, T. G., & Vogler, A. P. (2008). Speciation and DNA barcodes: testing the effects of dispersal on the formation of discrete sequence clusters. *Philosophical Transactions of the Royal Society B-Biological Sciences, 363*(1506), 2987–2996. doi:10.1098/rstb.2008.0066.

Paz, A., & Crawford, A. J. (2012). Molecular-based rapid inventories of sympatric diversity: a comparison of DNA barcode clustering methods applied to geography-based vs clade-based sampling of amphibians. *Journal of Biosciences, 37*(5), 887–896. doi:10.1007/s12038-012-9255-x.

Phillips, M. J. (2009). Branch-length estimation bias misleads molecular dating for a vertebrate mitochondrial phylogeny. *Gene, 441*(1–2), 132–140. doi:10.1016/j.gene.2008.08.017.

Pons, J., Barraclough, T. G., Gomez-Zurita, J., Cardoso, A., Duran, D. P., Hazell, S., et al. (2006). Sequence-based species delimitation for the DNA taxonomy of undescribed insects. *Systematic Biology, 55*(4), 595–609. doi:10.1080/10635150600852011.

Powell, J. R. (2012). Accounting for uncertainty in species delineation during the analysis of environmental DNA sequence data. *Methods in Ecology and Evolution, 3*(1), 1–11. doi:10.1111/j.2041-210X.2011.00122.x.

Powell, J. R., Monaghan, M. T., Oepik, M., & Rillig, M. C. (2011). Evolutionary criteria outperform operational approaches in producing ecologically relevant fungal species inventories. *Molecular Ecology, 20*(3), 655–666. doi:10.1111/j.1365-294X.2010.04964.x.

Puillandre, N., Lambert, A., Brouillet, S., & Achaz, G. (2012). ABGD, automatic barcode Gap discovery for primary species delimitation. *Molecular Ecology, 21*(8), 1864–1877. doi:10.1111/j.1365-294X.2011.05239.x.

R Core Team (2013). R: A language and environment for statistical computing. http://www.R-project.org/.

Ratnasingham, S., & Hebert, P. D. N. (2007). BOLD: the barcode of life data system (www.Barcodinglife.org). *Molecular Ecology Notes, 7*(3), 355–364. doi:10.1111/j.1471-8286.2006.01678.x.

Reid, N. M., & Carstens, B. C. (2012). Phylogenetic estimation error can decrease the accuracy of species delimitation: a Bayesian implementation of the general mixed Yule-coalescent model. *BMC Evolutionary Biology, 12*, 196. doi:10.1186/1471-2148-12-196.

Saccone, C., Gissi, C., Lanave, C., Larizza, A., Pesole, G., & Reyes, A. (2000). Evolution of the mitochondrial genetic system: an overview. *Gene, 261*, 153–159.

Sanderson, M. J. (2002). Estimating absolute rates of molecular evolution and divergence times: a penalized likelihood approach. *Molecular Biology and Evolution, 19*(1), 101–109.

Shields, G. F., Adams, D., Garner, G., Labelle, M., Pietsch, J., Ramsay, M., et al. (2000). Phylogeography of mitochondrial DNA variation in brown bears and polar bears. *Molecular Phylogenetics and Evolution, 15*(2), 319–326. doi:10.1006/mpev.1999.0730.

Sites, J. W., & Marshall, J. C. (2003). Delimiting species: a Renaissance issue in systematic biology. *Trends in Ecology & Evolution, 18*(9), 462–470. doi:10.1016/s0169-5347(03)00184-8.

Smith, S. A., & O'Meara, B. C. (2012). treePL: divergence time estimation using penalized likelihood for large phylogenies. *Bioinformatics, 28*(20), 2689–2690. doi:10.1093/bioinformatics/bts492.

Stamatakis, A. (2014). RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics, 30*(9), 1312–1313. doi:10.1093/bioinformatics/btu033.

Steeman, M. E., Hebsgaard, M. B., Fordyce, R. E., Ho, S. Y. W., Rabosky, D. L., Nielsen, R., et al. (2009). Radiation of extant cetaceans driven by restructuring of the oceans. *Systematic Biology, 58*(6), 573–585. doi:10.1093/sysbio/syp060.

Stiller, M., Molak, M., Prost, S., Rabeder, G., Baryshnikov, G., Rosendahl, W., et al. (2014). Mitochondrial DNA diversity and evolution of the Pleistocene cave bear complex. *Quaternary International, 339–340*, 224–231. doi:10.1016/j.quaint.2013.09.023.

Talavera, G., & Castresana, J. (2007). Improvement of phylogenies after removing divergent and ambiguously aligned blocks from protein sequence alignments. *Systematic Biology, 56*(4), 564–577. doi:10.1080/10635150701472164.

Talavera, G., & Vila, R. (2011). What is the phylogenetic signal limit from mitogenomes? The reconciliation between mitochondrial and nuclear data in the Insecta class phylogeny. *Bmc Evolutionary Biology, 11*, doi:10.1186/1471-2148-11-315.

Talavera, G., Dinca, V., & Vila, R. (2013). Factors affecting species delimitations with the GMYC model: insights from a butterfly survey. *Methods in Ecology and Evolution, 4*(12), 1101–1110. doi:10.1111/2041-210x.12107.

Talbot, S., & Shields, G. (1996). Phylogeography of brown bears (*Ursus arctos*) of Alaska and paraphyly within the Ursidae. *Molecular Phylogenetics and Evolution, 5*, 477–494.

Tamura, K., Stecher, G., Peterson, D., Filipski, A., & Kumar, S. (2013). MEGA6: molecular evolutionary genetics analysis version 6.0. *Molecular Biology and Evolution, 30*(12), 2725–2729. doi:10.1093/molbev/mst197.

Tang, C. Q., Leasi, F., Obertegger, U., Kieneke, A., Barraclough, T. G., & Fontaneto, D. (2012). The widely used small subunit 18S rDNA molecule greatly underestimates true diversity in biodiversity surveys of the meiofauna. *Proceedings of the National Academy of Sciences, 109*(40), 16208–16212. doi:10.1073/pnas.1209160109.

Tang, C. Q., Humphreys, A. M., Fontaneto, D., & Barraclough, T. G. (2014). Effects of phylogenetic reconstruction method on the robustness of species delimitation using single-locus data. *Methods in Ecology and Evolution, 5*(10), 1086–1094. doi:10.1111/2041-210X.12246.

Uhen, M. D. (2010). The origin(s) of whales. *Annual Review of Earth and Planetary Sciences, 38*(1), 189–219. doi:10.1146/annurev-earth-040809-152453.

Vaughan, T. G., Kuehnert, D., Popinga, A., Welch, D., & Drummond, A. J. (2014). Efficient Bayesian inference under the structured coalescent. *Bioinformatics, 30*(16), 2272–2279. doi:10.1093/bioinformatics/btu201.

Vilstrup, J. T., Ho, S. Y. W., Foote, A. D., Morin, P. A., Kreb, D., Kruetzen, M., et al. (2011). Mitogenomic phylogenetic analyses of the Delphinidae with an emphasis on the Globicephalinae. *BMC Evolutionary Biology, 11*, 65. doi:10.1186/1471-2148-11-65.

Weigand, A. M., Jochum, A., Slapnik, R., Schnitzler, J., Zarza, E., & Klussmann-Kolb, A. (2013). Evolution of microgastropods (Ellobioidea, Carychiidae): integrating taxonomic, phylogenetic and evolutionary hypotheses. *BMC Evolutionary Biology, 13*, 18. doi:10.1186/1471-2148-13-18.

Willerslev, E., Gilbert, M. T. P., Binladen, J., Ho, S. Y. W., Campos, P. F., Ratan, A., et al. (2009). Analysis of complete mitochondrial genomes from extinct and extant rhinoceroses reveals lack of phylogenetic resolution. *BMC Evolutionary Biology, 9*, 95. doi:10.1186/1471-2148-9-95.

Williamson, S., & Orive, M. E. (2002). The genealogy of a sequence subject to purifying selection at multiple sites. *Molecular Biology and Evolution, 19*(8), 1376–1384.

Xia, X. H. (2013). DAMBE5: a comprehensive software package for data analysis in molecular biology and evolution. *Molecular Biology and Evolution, 30*(7), 1720–1728. doi:10.1093/molbev/mst064.

Yang, Z. (2014). *Molecular evolution: a statistical approach*. Oxford: Oxford University Press.

Yang, Z., & Rannala, B. (2010). Bayesian species delimitation using multilocus sequence data. *Proceedings of the National Academy of Sciences of the United States of America, 107*(20), 9264–9269. doi: 10.1073/pnas.0913022107.

Yang, Z., & Rannala, B. (2014). Unguided species delimitation using DNA sequence data from multiple loci. *Molecular Biology and Evolution, 31*(12), 3125–3135. doi:10.1093/molbev/msu279.

Yu, L., Li, Y.-W., Ryder, O., & Zhang, Y.-P. (2007). Analysis of complete mitochondrial genome sequences increases phylogenetic resolution of bears (Ursidae), a mammalian family that experienced rapid speciation. *BMC Evolutionary Biology, 7*(1), 198.

Yule, G. V. (1924). A mathematical theory of evolution, based on the conclusions of Dr. J. C. Willis, P.R.S. *Philosophical Transactions of the Royal Society B-Biological Sciences, 213*, 21–87.