

ORIGINAL ARTICLE

Pinniped phylogenetic relationships inferred using AFLP markers

KK Dasmahapatra¹, JI Hoffman² and W Amos²

¹Galton Laboratory, Department of Biology, University College London, London, UK and ²Department of Zoology, University of Cambridge, Cambridge, UK

Amplified fragment length polymorphisms (AFLPs) are widely used for phylogenetic reconstruction in plants but their use in animal taxa has been limited, and in mammals rare. In addition, their use has been largely limited to shallow relationships amongst species or subspecies. Here, we genotype 23 pinniped species for 310 AFLP markers and find a strong phylogenetic signal, with individuals coclustering within species, and overall a good agreement between our phylogeny and those constructed using mitochondrial DNA and nuclear sequences even at nodes ~15 million years old. Although supporting the existing ideas about pinniped relationships, our data shed light on relationships

within the hitherto relatively unresolved Phocine species group, and provide further supporting evidence for raising two subspecies of *Zalophus californianus*, *Z. c. californianus* and *Z. c. wollebaeki*, to species level. Plotting AFLP divergence time estimates against those based on both mtDNA and nuclear sequences we find strong linear relationships, suggesting that the different markers are evolving in a clocklike fashion. These data further emphasize the utility of AFLP markers as general tools for phylogenetic reconstruction.

Heredity advance online publication, 11 March 2009; doi:10.1038/hdy.2009.25

Keywords: AFLP; pinniped; phylogeny; molecular clock; amplified fragment length polymorphism

Introduction

Biologists require robust phylogenies to put findings in an evolutionary context, and recent decades have seen extensive use of usually sequence-based data for phylogenetic inference. Although multiple nuclear sequences are used for constructing phylogenies in multicellular organisms (Steppan *et al.*, 2004; Mallarino *et al.*, 2005; Nazari *et al.*, 2007), mitochondrial (mtDNA) genes such as *cytochrome b*, *cytochrome oxidase I* and the mitochondrial control region are still probably the most widely used genetic markers for this phylogenetic reconstruction (Folmer *et al.*, 1994; Simmons and Weller, 2001). More recently, a short section of the mitochondrial *cytochrome oxidase I* gene has been put forward as a universal candidate locus for species identification purposes or DNA barcoding (Hebert *et al.*, 2003; Stoeckle and Hebert, 2008).

The widespread use of mtDNA genes can be attributed to a number of factors: ease of amplification, high evolutionary rate, low recombination rate and smaller population size compared to nuclear loci. However, there are factors that may complicate interpretation of the resulting trees. First, as mitochondria are maternally inherited, mtDNA only reflects female history. In most cases this does not impact on phylogenetic reconstruction, but many species, especially among mammals,

show strong female phylopatriy (Greenwood, 1980), creating strong substructure within a species evident at cytoplasmic but not nuclear genes (Ruppell *et al.*, 2003; Nies *et al.*, 2005). Second, interspecific hybridization may cause branch-swapping between related species and affect mitochondrial and nuclear genomes differentially, with mtDNA apparently crossing species boundaries much more readily (Chan and Levin, 2005; Linnen and Farrell, 2007). Third, there is mounting evidence of direct selection on mtDNA genes and of indirect selection arising from disequilibrium with other maternally inherited genes and cytoplasmic elements, both of which could confuse the resulting phylogenetic picture (Ballard and Whitlock, 2004; Hurst and Jiggins, 2005).

These disadvantages of relying solely on mtDNA genes mean that studies using multigene phylogenies integrating information from both nuclear and mitochondrial sequences are becoming more prevalent in plants and animals (Beltrán *et al.*, 2007; Alfaro *et al.*, 2008; Hugall *et al.*, 2008). Relatively fast evolving sequences are required to determine species-level relationships, and nuclear coding regions are preferable as frequent indels in non-coding regions create length polymorphism requiring post-PCR cloning before sequencing. However, it has not been possible to find nuclear coding regions that both meet these criteria and that can be amplified reliably across a wide range of taxa. As a result, so far there are no universal 'nuclear barcode' genes (Dasmahapatra and Mallet, 2006).

Amplified fragment length polymorphisms (AFLPs) are predominantly nuclear in origin and have rapid rates of evolution (Vos *et al.*, 1995; Mueller and Wolfenbarger, 1999). Additionally, they are readily amplifiable in both prokaryotic and eukaryotic taxa with little need for optimization.

Correspondence: Dr KK Dasmahapatra, Department of Biology, University College London, Wolfson House, 4 Stephenson Way, London NW1 2HE, UK.

E-mail: k.dasmahapatra@ucl.ac.uk

Received 21 November 2008; revised 9 February 2009; accepted 10 February 2009

Among other applications, AFLPs are currently widely used to create robust phylogenetic inferences especially within taxa, in which interspecific relationships have been difficult to resolve or when appropriate nuclear genes are scarce (Barluenga *et al.*, 2006; Mendelson and Simons, 2006; Pellmyr *et al.*, 2007). This widespread application is despite the restrictions on the use of AFLP markers in phylogenetic analyses resulting from their unknown mutation characteristics and anonymous nature (Felsenstein, 2004; Koopman, 2005). However, there is bias in application with most usage largely limited to sub-generic and subspecies groupings with a heavy bias towards plant taxa (Bensch and Akesson, 2005).

An ISI Web of Science search in October 2007 using 'AFLP AND phylog*' as search terms in the 'Topic' field retrieved 525 records of which 421 used AFLPs for phylogenetic purposes. Although these search terms do not recover all AFLP-based phylogenetic studies, the recovered dataset is large and unbiased. By examining the titles, abstracts and where necessary the paper itself, each record was broadly classified into taxonomic groupings. This analysis revealed that 61% of the studies were on plants and only 14% on animals, with the remainder on fungi and bacteria. Within plants and animals, only 5% were conducted above the generic level. In particular, there were just eight family-level studies on vertebrates. This included one mammalian study and four studies on rapidly radiating cichlid fish. Our analysis clearly indicates that AFLP markers are rarely used to construct higher level phylogenies.

Here, we test the utility of AFLP markers to recover both family- and species-level relationships within a mammalian order. As our test case, we use a group with a relatively well-represented molecular phylogeny: the Pinnipedia. Pinnipeds, or seals, are a diverse group of aquatic mammals (Riedman, 1990) comprising 34 extant species spread across three distinct families: the Odobenidae, the Otariidae and the Phocidae (Rice, 1998). It is estimated that this group of mammals originated 23–26 million years ago (MYA), but much of the current species diversity is of relatively recent origin, having arisen during and after the Pleistocene (Arnason *et al.*, 2006; Higdon *et al.*, 2007). Pinnipeds have been intensively studied due to their ecological importance as top predators in marine ecosystems and conservation concerns facing many species. Considerable effort has gone into elucidating the phylogeny of this group of mammals using both mtDNA (Wynen *et al.*, 2001; Arnason *et al.*, 2006) and nuclear sequences (Fulton and Strobeck, 2006).

Higdon *et al.*'s (2007) phylogenetic analysis of the 34 extant pinniped species based on all available GenBank pinniped sequences is, to date, the most comprehensive reported analysis of pinniped species relationships. The Mustelidae have been shown to be the sister taxa to a monophyletic pinniped clade (Fulton and Strobeck, 2006; Sato *et al.*, 2006), and within the pinnipeds the Odobenidae are sister to the Otariidae. Although basal relationships are strongly supported, many of the more terminal branches have weak support. Within the Otariids there is no evidence to support monophyly of either sea lions or fur seals and relationships among many species are difficult to resolve. On the other main pinniped branch, the Phocidae, relatively rapid speciation, especially within the Phocinae, has led to uncertainty about the exact relationships among several species. The analyses

in Higdon *et al.* (2007) are dominated by mitochondrial genes, and from a nuclear perspective the species coverage is low with an average of only 5.5 species' represented at each nuclear gene. A further limitation of this study is that where sequences were available, each species was represented by a single sequence. This does not allow the evaluation of intraspecific genetic variation and may not be sufficient sampling for species involved in rapid radiations such as in the genus *Phoca* (Arnason *et al.*, 2006), or many Otariid species which are known to hybridize with one another (Rice, 1998; Lancaster *et al.*, 2007). Therefore, a well-supported nuclear phylogeny of pinnipeds remains elusive.

We use AFLP markers to recover both family- and species-level relationships in a mammalian order and compare the resulting phylogeny with published versions based on both mtDNA (Wynen *et al.*, 2001; Arnason *et al.*, 2006) and nuclear (Fulton and Strobeck, 2006) sequences to test the utility of AFLPs for higher-level phylogenetics. At the same time we assess the relative rates of change of the different marker types.

Methods

A total of 23 pinniped species comprising 8 otariids, 14 phocids and the walrus (Table 1) were used in the analysis. Where possible, five representatives of each species were used. The Ursidae and Mustelidae are the putative outgroups to the pinnipeds, therefore, two polar bear (*Ursus maritimus*) and two badger (*Meles meles*) samples were added as representatives of these taxa. DNA was extracted by Proteinase K digestion using an adapted Chelex 100 protocol (Walsh *et al.*, 1991) followed by standard phenol–chloroform purification (Sambrook *et al.*, 1989).

The AFLP protocol was similar to that used in Vos *et al.* (1995) and Dasmahapatra *et al.* (2008). Briefly, 100–400 ng of genomic DNA was first digested using *TaqI* (5 U in a 10 µl volume, 65 °C for 2 h) and then with *EcoRI* (5 U in a 20 µl volume, 37 °C for 2 h). *EcoRI* and *TaqI* adapters (Ajmone-Marsan *et al.*, 1997; Madden *et al.*, 2004) were ligated on to the digested DNA using T4 DNA ligase (1 U in a 50 µl volume, 37 °C for 3 h) and then diluted 10 times with low TE (10 mM Tris-HCl, 0.1 mM EDTA, pH 8). For the preamplification, 5 µl of ligation mix was added to a 50 µl PCR reactions containing Tris-HCl (10 mM, pH 8.3), MgCl₂ (1.5 mM), KCl (50 mM), dNTPs (0.2 mM), *Taq* polymerase (1 U) and 50 ng each of the preamplification primers *EcoRI*-A and *TaqI*-C carrying one selective nucleotide. Following 30 preamplification cycles (30 s at 94 °C, 60 s at 50 °C and 60 s at 72 °C) the products were diluted 10 times with low TE. For the selective amplification, 2.5 µl of diluted preamplification product was added to a 12.5 µl reaction containing Tris-HCl (10 mM, pH 8.3), MgCl₂ (1.5 mM), KCl (50 mM), dATPs, dTTP and dGTP (0.2 mM each), dCTP (0.04 mM), α³²P-dCTP (0.1 µCi) *Taq* polymerase (0.2 U), *EcoRI* selective primer (*EcoRI*-AXX, 5 ng) and *TaqI* selective primer (*TaqI*-CXX, 30 ng). Samples were subjected to 13 selective amplification cycles (30 s at 94 °C, 60 s at 65 °C, reducing by 0.7 °C each cycle, and 60 s at 72 °C), followed by further 23 cycles (30 s at 94 °C, 60 s at 56 °C and 60 s at 72 °C). The samples were genotyped using eight AFLP primer combinations (*TaqI*-CAC with *EcoRI*-ACA, *TaqI*-CCA with *EcoRI*-ACA, *TaqI*-CGA with *EcoRI*-ACA,

Table 1 Taxonomy of the 24 pinniped taxa used in the phylogenetic analysis

Taxonomic rank	Species	Common name	No. of samples
<i>Pinnipedia</i>		Seals	
Family: Odobenidae		Walrus	
	<i>Odobenus rosmarus</i>	Walrus	5
Family: Otariidae		Sea lions and fur seals	
Subfamily: Callorhinae		—	
	<i>Callorhinus ursinus</i>	Northern fur seal	5
Subfamily: Otariinae		—	
	<i>Arctocephalus gazella</i>	Antarctic fur seal	5
	<i>Eumetopias jubatus</i>	Steller sea lion	5
	<i>Zalophus californianus californianus</i>	California sea lion	5
	<i>Zalophus californianus wollebaeki</i>	Galapagos sea lion	5
	<i>Arctocephalus galapagoensis</i>	Galapagos fur seal	5
	<i>Arctocephalus forsteri</i>	New Zealand fur seal	3
	<i>Arctocephalus australis</i>	South American fur seal	5
Family: Phocidae		True seals	
Subfamily: Monachinae		Southern true seals	
Tribe: Miroungini		Elephant seals	
	<i>Mirounga leonina</i>	Southern elephant seal	5
Tribe: Monachini		Monk seals	
	<i>Monachus monachus</i>	Mediterranean monk seal	2
Tribe: Lobodontini		Antarctic seals	
	<i>Hydrurga leptonyx</i>	Leopard seal	5
	<i>Leptonychotes weddelli</i>	Weddell seal	5
	<i>Lobodon carcinophagus</i>	Crabeater seal	5
	<i>Ommatophoca rossi</i>	Ross seal	5
Subfamily: Phocinae		Northern true seals	
Tribe: Erignathini		—	
	<i>Erignathus barbatus</i>	Bearded seal	4
Tribe: Cystophorini		—	
	<i>Cystophora cristata</i>	Hooded seal	5
Tribe: Phocini		—	
	<i>Phoca largha</i>	Spotted seal	3
	<i>Phoca hispida</i>	Ringed seal	3
	<i>Phoca vitulina</i>	Harbour seal	5
	<i>Phoca caspica</i>	Caspian seal	4
	<i>Pagophilus groenlandicus</i>	Harp seal	5
	<i>Halichoerus grypus</i>	Grey seal	5

Samples were supplied by the following individuals and institutions: K Robertson, National Marine Fisheries Service, USA; Terry Burke, University of Sheffield, UK; K Acevedo-Whitehouse, Marine Mammal Centre, USA; British Antarctic Survey, UK; Sea Mammal Research Unit, UK; Paul Thompson, University of Aberdeen, UK; J Wolf, University of Bielefeld, Germany; J Bickham, Texas A&M University.

TaqI-CTG with EcoRI-ACA, TaqI-CAG with EcoRI-ACA, TaqI-CAC with EcoRI-AAC, TaqI-CAC with EcoRI-ATG, TaqI-CAC with EcoRI-AGC). All adapter and primer sequences are described in Madden *et al.* (2004). AFLP products were resolved by electrophoresis through 6% acrylamide gels and visualized by autoradiography.

AFLP profiles were assessed and scored by eye. Samples with odd AFLP profiles, where most bands are not observed in other profiles, were excluded. Only loci that amplified consistently, that is, sharp bands with minimal size variation, allowing them to be scored across the entire dataset were used. Three hundred and ten AFLP loci, sized 100–350 bp, were polymorphic either within or between species and could be scored unambiguously. It was assumed that AFLP bands that were the same size across species represented homologous markers. To assess our genotyping error rate, the original DNA extractions of 20 samples representing five species were reamplified at all eight primer combinations and rerun. The 310 loci scored above were identified in the AFLP profiles and the loci independently scored blind by two people. Genotyping error rate was estimated by comparing the original genotypes with the error-checking genotypes and calculating the number of mismatches (Bonin *et al.*, 2004).

The final character matrix (provided in the Supplementary information) used as input for the phylogenetic analysis consisted of 310 binary characters representing the presence and absence genotypes of 109 individuals from 23 pinniped species and two outgroup species at 310 AFLP bands. Distance methods were used to create phylogenetic trees, and the results compared with those obtained with parsimony and Bayesian methods. As the various pinniped species may have diverged up to 25 MYA, Nei–Li (Nei and Li, 1979) and Jaccard (Jaccard, 1901) distances were used in the distance analysis. These distance measures are best suited for AFLPs as they take account of the sharing of presence alleles but are not affected by the sharing of absent or zero alleles, which are more likely to be homoplasious. PhylTools (Buntjer *et al.*, 2002) was used to generate Nei–Li and Jaccard distance matrices for each of the 1000 bootstrapped datasets, and a consensus neighbour-joining (NJ) tree generated using the NEIGHBOR and CONSENSE modules in PHYLIP (Felsenstein, 1989). For the bootstraps, starting trees for each replicate were obtained by randomizing the input order of the species. Branch lengths were obtained using the FITCH module.

Parsimony methods are generally thought to perform poorly with AFLP data, most obviously because bands may be lost independently in more than one lineage,

giving rise to multiple null alleles that are not identical by descent (Pellmyr *et al.*, 2007). None the less, we chose to include one parsimony analysis for comparison,

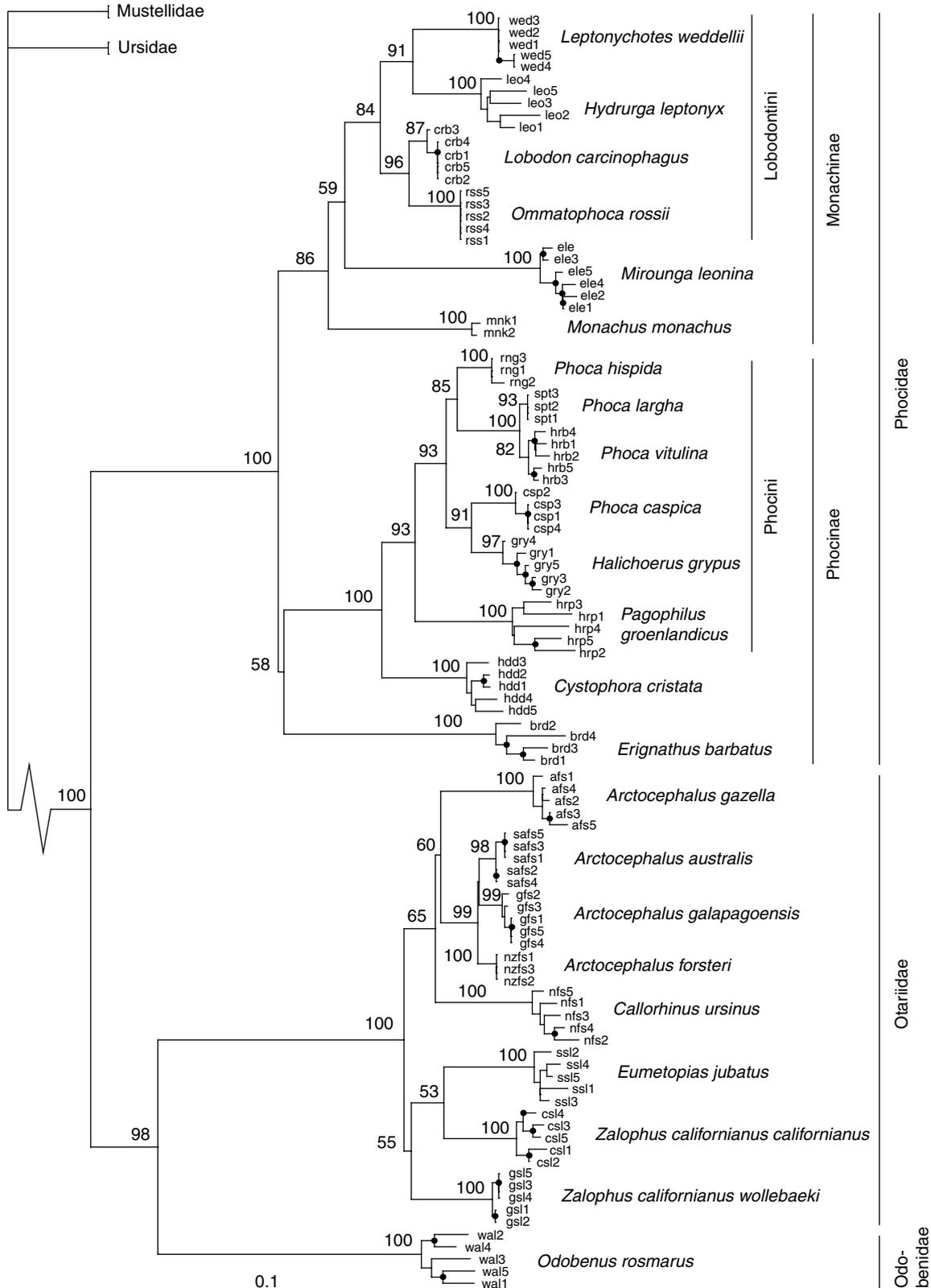


Figure 1 Consensus neighbour-joining phylogeny of pinnipeds based on Nei-Li distances calculated from 310 amplified fragment length polymorphism (AFLP) markers. Numbers represent per cent bootstraps support for branching events obtained from 1000 replications and are shown only when >50%. Intraspecific branches with >50% bootstrap support are labelled with •. Scale bar indicates Nei-Li distance.

choosing Dollo parsimony (Le Quesne, 1974), which assumes that mutations from band to null are more likely than mutations that create new bands. As species clusters are strongly supported in the NJ analyses (Figure 1), to reduce computational time a single representative from each species was randomly chosen for inclusion in the parsimony analysis. We conducted 1000 bootstrap pseudoreplicates using the DOLLOP module in PHYLIP, followed by CONSENSE to create the consensus tree.

Although Bayesian methods of phylogenetic inference from AFLP markers using models that accurately represent the evolution of AFLP markers have been developed (Luo *et al.*, 2007), the computational costs with the current algorithm is high and so the model is not practically applicable (Luo *et al.*, 2007; Koopman *et al.*, 2008). Instead, we opted to analyse our data with MrBayes 3.1.2 (Huelsenbeck and Ronquist, 2001) using the restriction sites model. As it is designed for restriction site data, this model is an oversimplification of the more complex evolutionary processes underlying AFLP polymorphisms. The option 'noabsencesites' was used to correct for coding bias in AFLP datasets (Ronquist *et al.*, 2005; Koopman *et al.*, 2008). Analysis with Mr Bayes was carried out on the same individuals as were used in the parsimony analysis. The Bayesian analysis consisted of two runs of 10 000 000 generations with a sample frequency of 1000. Defaults were used for all other parameters. A burnin of 5000 samples (5000 000 generations) was used based on the criteria of stable log likelihoods and convergence values (standard deviations of split frequencies). Posterior probabilities were calculated based on the remaining 5000 trees per run.

As mtDNA and nuclear sequences exist for a number of pinniped species, we were able to examine how rates of AFLP band evolution relate to rates evolution at mtDNA loci as well as at specific nuclear loci. Mitochondrial *cytochrome b* (*cytb*) sequences representing all species used in the AFLP analysis were downloaded from GenBank. For 10 of these species, sequences from 5 nuclear genes (*IRBP* exon 1, *FES*, *GHR*, *CHRNA1* and *RHO1*) were also available (Fulton and Strobeck, 2006). The final 360 bp *cytb* alignment contained 157 sequences representing 23 pinniped species. Each nuclear alignment (*IRBP* exon 1—1183 bp; *FES*—362 bp; *GHR*—618 bp; *CHRNA1*—360 bp; *RHO1*—266 bp) comprised one sequence from each of the 10 species. Raw percentage sequence divergences at *cytb* and the nuclear genes were calculated among all possible species pairs using SITES (Hey and Wakeley, 1997). We compared average pairwise AFLP Jaccard distances among pinniped species with *cytb* sequence divergences and nuclear sequence divergence (averaged over the five loci). Jaccard instead of Nei–Li distances were used to allow comparison with another study (Buntjer *et al.*, 2002), for which only Jaccard distances were reported (see below). In practice, Jaccard and Nei–Li distance measures are highly correlated, with an r^2 of 0.99 in our dataset. Buntjer *et al.* (2002) reported pairwise Jaccard distances among nine species within the subfamily Bovinae, for which 870 bp of *cytb* sequence were also available on GenBank. This allowed us to make a comparison of AFLP band and mtDNA evolution rates between two different mammalian orders.

The significance of the correlations between pairwise distance matrices was tested with Mantel tests implemented using the software *zt* (Bonnet and Van de Peer, 2002).

Results

Our estimated genotyping error rate was 0.5%, typically the level found in several other AFLP studies (Jones *et al.*, 1997; Arens *et al.*, 1998; Bonin *et al.*, 2004). Only 6 of the 310 AFLP bands scored were size homologous between pinnipeds and the two outgroup species, implying a relatively rapid rate of evolution for AFLP markers. However, these six bands are polymorphic within the pinnipeds, and their size homology is likely the result of size homoplasy rather than being genuinely homologous loci. Such tenuous overlap between the two groups meant that we were unable to infer which of the two outgroup taxa is closest to the pinnipeds. None the less, inclusion of the outgroups allowed confirmation that the pinnipeds form a monophyletic group, excluding the Ursidae and the Mustelidae, and that the Odobenidae are sister to the Otariidae and not the Phocidae (Figure 1).

NJ trees using Nei–Li and Jaccard distances were identical in topology with only minor differences in branch support (not shown), and we show the former (Figure 1). Both parsimony and Bayesian trees (Figure 2) are less well resolved compared to the NJ analysis. Bootstrap supports for most branches in the parsimony tree are weak. Although relationships within the Phocinae are concordant between the parsimony and NJ trees, there are a number of discrepancies within the Monachinae and Otariidae. However, except for the basal positioning of *Callorhinus ursinus* within the Otariidae, none of these discrepancies are strongly supported in the parsimony analysis. In the Bayesian tree, although posterior probabilities tend to be high at bifurcating nodes, several polytomies are present within both the otariids and phocids. Relationships within the Phocinae are again concordant between Bayesian and NJ trees. Of the three analyses, the tree from the NJ analysis is best resolved, therefore we focus on the NJ tree.

In the NJ tree, all conspecific individuals cocluster with strong support, indicating good resolution at the level of species (Figure 1). In general, more basal relationships, for example at the base of both the Phocinae and the Monachinae, receive relatively weaker bootstrap support (BS = 58 and 86%, respectively). The otariids form a strongly supported monophyletic group, but within-otariid species relationships remain poorly resolved in many instances. The subfamily divisions between fur seals (*Arctocephalus/Callorhinus*) and the sea lions are weakly supported in the NJ analysis (BS = 65 and 55%, respectively), however, as discussed later, this is probably an artefact of incomplete species sampling. The only strongly supported otariid clade contains *A. galapagoensis*, *A. australis* and *A. forsteri*. The two subspecies of *Zalophus californianus*, the California (*Z. c. californianus*) and Galapagos sea lions (*Z. c. wolfebaeki*) were found to be strongly divergent and are not sister taxa. Instead, there is weak support for a sister relationship between the *E. jubatus* and *Z. c. californianus* (BS = 53%).

Monophyly of the phocids is strongly supported (Figure 1). Within the Phocidae, monophyletic groupings

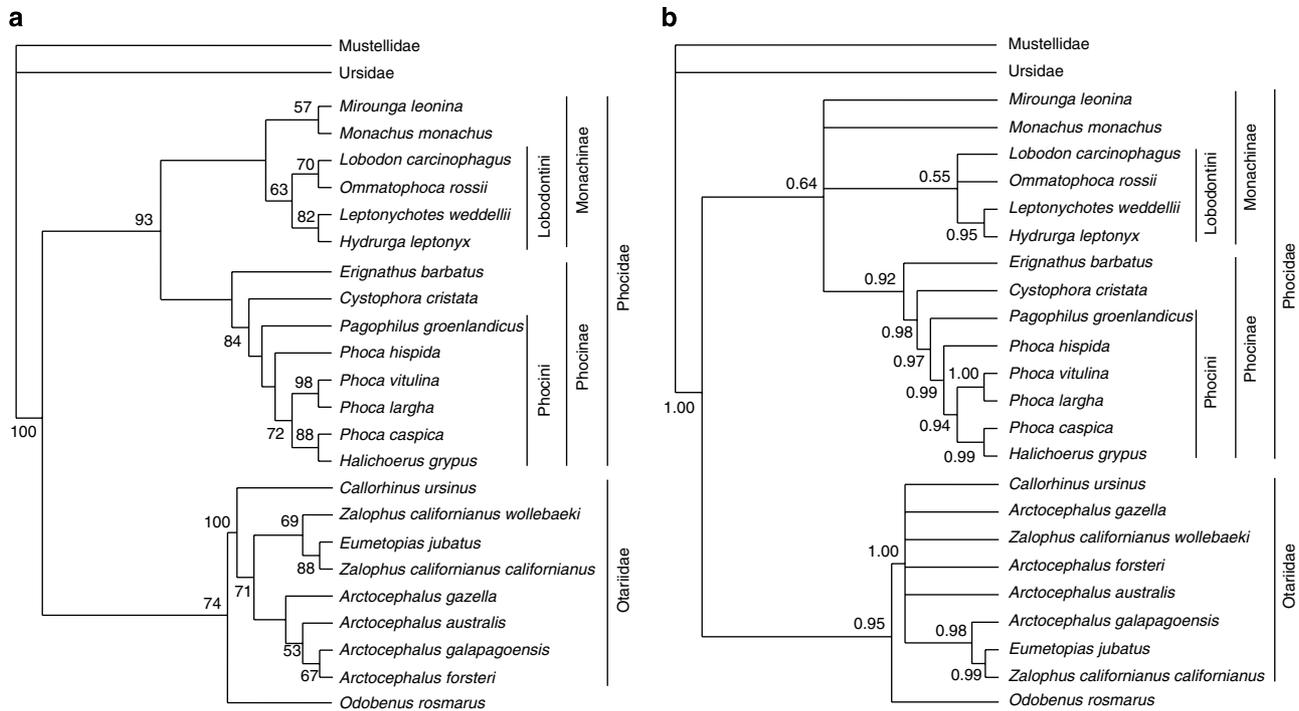


Figure 2 (a) Consensus parsimony tree with per cent bootstrap support. (b) Bayesian tree with posterior probabilities.

of both the Phocinae and Monachinae are recovered (BS = 86 and 58%, respectively), although support for the latter is weak. In the Monachinae, remaining relationships are consistent with traditional divisions: the monk seals, represented by *M. monachus*, are the basal taxa, followed by *M. leonina* and then the Lobodontini. Within the Lobodontini there is strong support for sister relationships between *H. leptonyx* and *L. weddellii* (BS = 91%), and between the *L. carcinophagus* and *O. rossii* (BS = 96%). In the Phocinae, *E. barbatus* was the deepest branching lineage (BS = 53%), followed by *C. cristata* (BS = 100%) and then the Phocini (BS = 93%). Interspecific relationships within the Phocini were well supported, with *P. groenlandicus* being the basal taxon (BS = 93%). A clade containing *P. hispida*, *P. largha* and *P. vitulina* was recovered (BS = 85%) within which the *P. hispida* was basal (BS = 100%). There was also strong support for sister relationships between *P. caspica* and *H. grypus* (BS = 91%).

There are strong correlations between pairwise interspecific AFLP Jaccard distances and interspecific *cytb* sequence divergences within both the pinnipeds (Figure 3a: $r = 0.87$, Mantel test, $P = 0.00001$) and the Bovinae (Figure 3a: $r = 0.89$, Mantel test, $P = 0.0003$). Both relationships are linear, and over the evolutionary distance investigated neither AFLP distances nor mtDNA sequence divergences appear to be significantly saturating. The two slopes (pinnipeds: 0.164 ± 0.006 ; Bovinae: 0.157 ± 0.014) do not differ significantly from each other (ANOVA: $F_{(1,264)} = 0.15$, $P = 0.7$) indicating that rate of AFLP-to-*cytb* evolution is similar in both groups of species. A strong correlation (Figure 3b: $r = 0.89$, $P = 0.00005$) is also present between pairwise interspecific AFLP Jaccard distances and interspecific average nuclear divergences.

Discussion

Here we use AFLP genotypes from a wide range of pinnipeds species to explore the effectiveness of these markers to recover useful phylogenetic information at the scale of a mammalian order. We obtain a generally well-supported tree, with all individuals within each species coclustering. Compared with the equivalent trees obtained using mitochondrial and nuclear sequence data we find close agreement as well as enhanced species resolution within the Phocinae. We also report striking correlations among inferred rates of evolutionary change based on AFLP, mtDNA and nuclear sequences.

Size homology of AFLP markers

Probably the main consideration that appears to limit the use of AFLPs for phylogenetic inference is concern about the assumption of size homology of AFLP markers. This concern is particularly relevant for studies carrying out deeper than species-level analyses (Bussell *et al.*, 2005; Althoff *et al.*, 2007). Simulated AFLP profiles of human and mice genomes show that only 66% of same-sized bands are homologous within a species (Althoff *et al.*, 2007). In the same study, comparison of simulated AFLP profiles of three *Drosophila* species show that only 59 and 19% of bands are homologous between species estimated to have diverged 2 and 6 MYA, respectively. These results suggest that AFLP markers should carry little or no phylogenetic signal when comparisons are carried out between taxa that have diverged more than a few million years ago. In contrast to this expectation we find sufficient phylogenetic signal even as far back as ~15 MYA (Higdon *et al.*, 2007) enabling verification of the Odobenidae/Otariidae sister relationship as well as basal relationships within the Phocidae. AFLP markers

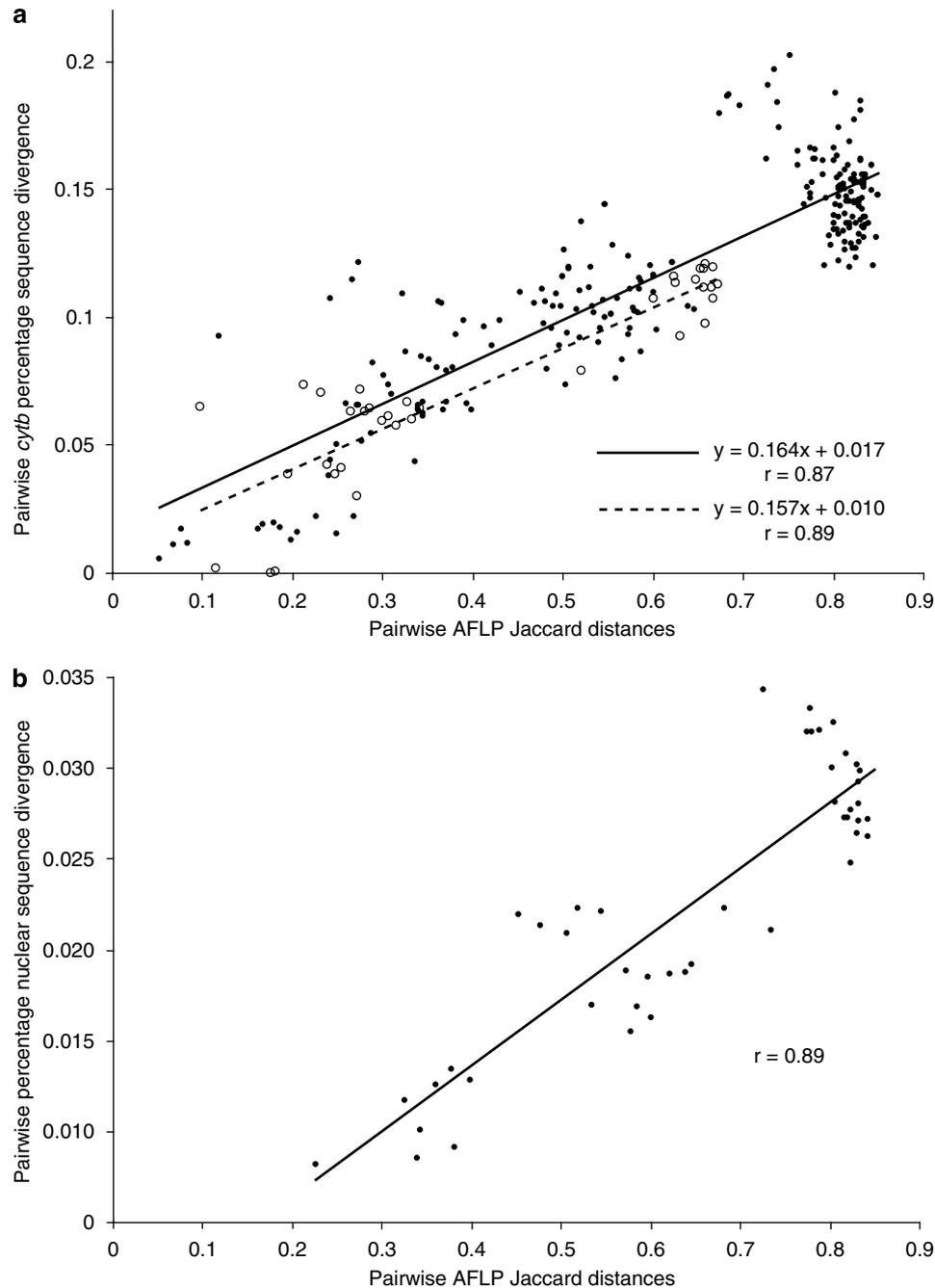


Figure 3 (a) Pairwise interspecific amplified fragment length polymorphism (AFLP) distances plotted against raw percentage pairwise interspecific *cytochrome b* (*cytb*) sequence divergences for 23 pinniped species (closed circles and solid line) and 9 Bovinae species (open circles and dashed line) with regression lines. (b) Pairwise interspecific AFLP distances plotted against raw percentage pairwise interspecific nuclear sequence divergence (averaged over five loci) for 10 pinniped species (Fulton and Strobeck, 2006).

fail to recover pinniped outgroup relationships, but this is not surprising given that fewer than 20% of bands are shared among the deepest pinniped lineages (Figure 3a), corresponding to 20–26 MYA (Higdon *et al.*, 2007). Extrapolating the linear relationship in Figure 3a, one would expect very few shared homologous bands between the outgroup taxon and pinnipeds as this split is dated to 30–41 MYA (Higdon *et al.*, 2007).

Homology of AFLP bands should decrease with the phylogenetic distance over which comparisons are carried out (Althoff *et al.*, 2007). As a result, phylogenetic signal-to-noise ratio should also decrease with increasing

phylogenetic distance. However, in both the Pinnipedia and Bovinae we find no evidence for an increase in the scatter of points about the regression line indicative of greater noise at larger distances (Figure 3a). We have only scored bands >100 bp, and although this will reduce the frequency of homoplasious bands compared to levels reported in Althoff *et al.* (2007) where bands >50 bp were used, it will not have eliminated the problem. A plausible reason for the discrepancy may lie in the fundamental difference in methodology. Althoff *et al.* (2007) used simulated AFLP profiles with no marker selection, whereas the results presented here are based

on real AFLP profiles with the genotyping involving a degree of marker selection. In reality, size homoplasious bands representing different loci may show different amplification intensities as well as some conformation or sequence dependent differences in electrophoretic mobility. Although such variation will not present in studies using simulated AFLP profiles, it may to a certain extent be captured in studies using real AFLP genotypes. For example, here we excluded loci showing widely variable levels of amplification across the species and loci that appeared to show minor size variation less than 1 bp. Marker selection by scorers is inherent in actual AFLP genotyping (Bonin *et al.*, 2004) and may substantially reduce the frequency of size homoplasy from those expected based on simulated results. Although this is clearly an important area for further investigation, there is no reason to believe that depth of phylogenetic resolution achieved here is specific to the pinniped study system. In any case, as the effect of a greater proportion of homoplasious bands at greater distances is to introduce noise, use of larger numbers of AFLP bands could help provide sufficient signal at these deeper phylogenetic levels.

Comparing AFLP, mtDNA and nuclear sequence tree topologies

Both the parsimony and Bayesian approaches used here inadequately deal with the AFLP data and result in poorly resolved trees. Loss of an AFLP band to create a null allele can occur independently in different lineages. This presents a problem for parsimony analyses, where null alleles in different species are assumed to have a common origin, and the problem worsens the longer the period over which phylogenetic inferences are being drawn. The Bayesian analysis implemented here is imperfect as it uses an oversimplified model of AFLP marker evolution. Distance methods using Nei–Li or Jaccard distances may make less use of available information but do not assume that null alleles are homologous and generally make fewer assumptions about AFLP marker evolution. As such, distance-based methods are likely to be more reliable than parsimony and current Bayesian methods when applied to AFLP datasets such as ours. Consequently, despite the fact that all three methods give broadly similar results, the remaining discussion is based predominantly on results of the distance analysis (Figure 1).

Most interspecific relationships within otariids are not strongly supported in the AFLP analyses. Direct comparison of otariid relationships found in the AFLP and Fulton and Strobeck's (2006) phylogenies based on nuclear sequences is difficult as sampling of species is patchy and not complementary. Previous mtDNA and nuclear sequence-based analyses have questioned splitting the Otariidae into fur seals and sea lions (Arnason *et al.*, 2006; Higdon *et al.*, 2007). In mtDNA phylogenies (Wynen *et al.*, 2001; Arnason *et al.*, 2006), *C. ursinus* is basal to the whole group. We found that although *C. ursinus* is basal to all other Otariidae using parsimony, the NJ analysis weakly support placing it at the base of all fur seals included in our analyses. The AFLP NJ analysis appears to support monophyletic fur seal and sea lion clades. However, the position of *C. ursinus* within the AFLP phylogeny is only weakly supported.

Excluding *C. ursinus*, the four fur seal species used in this study (*A. gazella*, *A. galapagoensis*, *A. australis* and *A. fosteri*) also form a monophyletic group in Wynen *et al.* (2001), as do the two sea lion species (*Z. californianus* and *E. jubatus*). Therefore, it is possible that the monophyly of both fur seals and sea lions that we observe is a consequence of our limited sampling of extant otariid species. Like Wynen *et al.* (2001), we also place *A. gazella* basally with respect to the *A. galapagoensis*/*A. australis*/*A. fosteri* group, but relationships within the *A. galapagoensis*/*A. australis*/*A. fosteri* group are more difficult to resolve. A close relationship among these three *Arctocephalus* species is also consistent with morphological analysis (Brunner, 2003).

Several authorities have called for the elevation of the three subspecies of *Zalophus californianus*, *Z. c. californianus*, *Z. c. wollebaeki* and *Z. c. japonicus* (now believed to be extinct) to species level (Sivertson, 1953; Rice, 1998), and recent nuclear and mitochondrial evidence supports such a reclassification (Wolf *et al.*, 2007). However, following Scheffer (1958), they are still widely considered to be populations at the subspecies level. Our AFLP data lend further clear support for classifying the Galapagos sea lion (*Z. c. wollebaeki*) as a species separate from the California sea lion (*Z. c. californianus*). The average AFLP–Nei distance between *Z. c. californianus* and *Z. c. wollebaeki* is similar to that between the *Z. c. californianus* and *E. jubatus*, the sister species based on mtDNA (Wynen *et al.*, 2001) indicating that the two subspecies of *Z. californianus* are genetically distinct from each other. These added data justify inclusion of *Z. c. wollebaeki* in the vulnerable category on the IUCN Red List (<http://www.iucnredlist.org/>). However, its phylogenetic distinctiveness from *Z. c. californianus* may mean that *Z. c. wollebaeki* warrants greater conservation consideration.

Species relationships within the Monachinae are concordant among trees obtained using AFLP, mtDNA and nuclear sequence datasets (Arnason *et al.*, 2006; Fulton and Strobeck, 2006; Higdon *et al.*, 2007). All three datasets support the basal divergence between *Monachus* and the rest of the Monachinae, and between *Mirounga* and the Lobodontini, as well as a sister relationship between *L. weddelli* and *H. leptonyx*. Relationships within the Lobodontini have traditionally been contentious. Morphological analyses fail to find adequate resolution (Bininda-Emonds and Russell, 1996), and although some molecular studies support an *O. rossii*, *L. carcinophagus* and *L. weddelli*, *H. leptonyx* relationship (Arnason *et al.*, 2006; Fulton and Strobeck, 2006), others support *L. carcinophagus* as sister to the remaining taxa (Fyler *et al.*, 2005; Higdon *et al.*, 2007). In all cases, support for these different scenarios has been weak. Our data provide rather strong support for a different relationship: (*O. rossii*, *L. carcinophagus*), (*L. weddelli*, *H. leptonyx*). Clearly, this is an area where further research is desirable.

Within the Phocinae, most recent studies have recovered *E. barbatus* as the basal taxon followed by *C. cristata*, and then a branch separating *P. groenlandicus* and *H. fasciata* from the remaining taxa (Arnason *et al.*, 2006; Fulton and Strobeck, 2006; Higdon *et al.*, 2007). However, more terminal relationships among the remaining six Phocini species were poorly resolved due to weak phylogenetic signal in this part of the tree. Using AFLP genotypes, we also recover *E. barbatus* and *C. cristata* as the deepest branches within the Phocinae, but in contrast

to previous studies, we also have strong support for species relationships within the Phocini. Within this group, *P. groenlandicus* is the basal species, a conclusion also supported by most other recent studies (Arnason *et al.*, 2006; Higdon *et al.*, 2007). Our data offer stronger support to the sister relationships between *P. vitulina* and *P. largha*, and between *P. caspica* and *H. grypus* compared with previous studies (Arnason *et al.*, 2006; Higdon *et al.*, 2007). Additionally, we also find good support for a sister relationship between *P. hispida* and the (*P. vitulina* + *P. largha*) clade. Our data strongly support placing the grey seal (*Halichoerus grypus*) amidst species belonging to the genus *Phoca*, and reinforces the suggestion of Arnason *et al.* (1995, 2006) for abolishing the name *Halichoerus* and inclusion of the grey seal within the genus *Phoca*.

Evolutionary rates

We have shown that AFLP-based distances between species have a strong linear relationship with both mitochondrial *cytb* and average nuclear sequence divergences. It is unlikely that selective forces could influence rates of evolution at both mtDNA locus and multiple nuclear markers in the same fashion. Consequently, the strong correlation among the three datasets seems likely to reflect genuine similarity in the estimated divergence times, and hence the *cytb*, the five nuclear loci (*IRBP* exon 1, *FES*, *GHR*, *CHRNA1* and *RHO1*) and the AFLP markers are evolving in a reasonably clocklike manner. Different selection pressures and mutation rate variation across lineages have the potential to cause deviations from clocklike behaviour at specific loci. Although such factors may affect individual AFLP loci, they are unlikely to affect all of the several hundred markers constituting a typical AFLP profile in a consistent manner. Therefore, the clocklike behaviour of AFLP markers is not unexpected, although to the best of our knowledge, this is the first time it has been shown. What is remarkable is that this clocklike behaviour occurs over an evolutionary timescale of ~26 million years (Higdon *et al.*, 2007). The strong correlation among the three marker types also offers support for the assumption that mtDNA in pinnipeds has not suffered major distortions of evolutionary rate due to natural selection, giving further confidence to the estimated divergence times reported in (Arnason *et al.*, 2006).

We have further shown that the rate of AFLP-to-*cytb* evolution is almost identical in two different mammalian orders. As AFLP studies do not, in general, report pairwise distance matrices, it was not possible to carry out a wide-ranging analysis based on a more extensive range of taxa, but this will be an interesting area for future research. This consistency between two mammalian orders also indicates that AFLP markers may provide a useful phylogenetic tool for investigating deeper species relationships.

In conclusion, although AFLP markers are widely used in plant taxonomy especially in rapidly evolving taxa, their use in animal taxa has been limited. Despite concerns over AFLP band homoplasy, we have shown that AFLP markers are useful tools for elucidating relationships within animal taxa over a timescale spanning ~26 million years. Our analysis of pinniped phylogenetic relationships using AFLP markers yields results that are largely concordant with those based on mtDNA and nuclear sequences, and at the same time

provides some additional resolution within rapidly radiating groups.

Acknowledgements

We thank the anonymous referees for their helpful comments. This research was funded by a Biotechnology and Biological Sciences Research Council PhD studentship.

References

- Ajmone-Marsan P, Valentini A, Cassandro M, Vecchiotti-Antaldi G, Bertoni G, Kuiper M (1997). AFLPTM markers for DNA fingerprinting in cattle. *Anim Genet* **28**: 418–426.
- Alfaro ME, Karns DR, Voris HK, Brock CD, Stuart BL (2008). Phylogeny, evolutionary history, and biogeography of Oriental-Australian rear-fanged water snakes (Colubroidea: Homalopsidae) inferred from mitochondrial and nuclear DNA sequences. *Mol Phylogenet Evol* **46**: 576–593.
- Althoff DM, Gitzendanner MA, Segraves KA (2007). The utility of amplified fragment length polymorphisms in phylogenetics: a comparison of homology within and between genomes. *Syst Biol* **56**: 477–484.
- Arens P, Coops H, Jansen J, Vosman B (1998). Molecular genetic analysis of black poplar (*Populus nigra* L.) along Dutch rivers. *Mol Ecol* **7**: 11–18.
- Arnason U, Bodin K, Gullberg A, Ledje C, Mouchaty S (1995). A molecular view of pinniped relationships with particular emphasis on the true seals. *J Mol Evol* **40**: 78–85.
- Arnason U, Gullberg A, Janke A, Kullberg M, Lehman N, Petrov EA *et al.* (2006). Pinniped phylogeny and a new hypothesis for their origin and dispersal. *Mol Phylogenet Evol* **41**: 345–354.
- Ballard JWO, Whitlock MC (2004). The incomplete natural history of mitochondria. *Mol Ecol* **13**: 729–744.
- Barluenga M, Stölting KN, Salzburger W, Muschick M, Meyer A (2006). Sympatric speciation in Nicaraguan crater lake cichlid fish. *Nature* **439**: 719–723.
- Beltrán M, Jiggins CD, Brower AVZ, Bermingham E, Mallet J (2007). Do pollen feeding and pupal-mating have a single origin in *Heliconius*? Inferences from multilocus sequence data. *Biol J Linn Soc* **92**: 221–239.
- Bensch S, Åkesson M (2005). Ten years of AFLP in ecology and evolution: why so few animals? *Mol Ecol* **14**: 2899–2914.
- Bininda-Emonds ORP, Russell AP (1996). A morphological perspective on the phylogenetic relationships of the extant phocod seals (Mammalia: Carnivora: Phocidae). *Bohne Zoologische Monographien* **41**: 1–256.
- Bonin A, Bellemain E, Eidesen PB, Pomparon F, Brochmann C, Taberlet P (2004). How to track and assess genotyping errors in population genetics studies. *Mol Ecol* **13**: 3261–3273.
- Bonnet E, Van de Peer Y (2002). Zt: a software for simple and partial Mantel tests. *J Stat Softw* **7**: 1–12.
- Brunner S (2003). Fur seals and sea lions (Otariidae): identification of species and taxonomic review. *System Biodivers* **1**: 339–439.
- Buntjer JB, Otsen M, Nijman IJ, Kuiper MTR, Lenstra JA (2002). Phylogeny of bovine species based on AFLP fingerprinting. *Heredity* **88**: 46–51.
- Bussell JD, Waycott M, Chappill JA (2005). Arbitrarily amplified DNA markers as characters for phylogenetic inference. *Perspect Plant Ecol Evol Syst* **7**: 3–26.
- Chan KMA, Levin SA (2005). Leaky prezygotic isolation and porous genomes: rapid introgression of maternally inherited DNA. *Evolution* **59**: 720–729.
- Dasmahapatra KK, Lacy RC, Amos W (2008). Estimating levels of inbreeding using AFLP markers. *Heredity* **100**: 286–295.
- Dasmahapatra KK, Mallet J (2006). DNA barcodes: recent successes and future prospects. *Heredity* **97**: 254–255.
- Felsenstein J (1989). PHYLIP—phylogeny inference package (Version 3.2). *Cladistics* **5**: 164–166.

- Felsenstein J (2004). *Inferring Phylogenies*. Sinauer Associates Inc.: Sunderland, Massachusetts.
- Folmer O, Black M, Hoeh W, Lutz R, Vrijenhoek R (1994). DNA primers for amplification of mitochondrial cytochrome c oxidase subunit I from diverse metazoan invertebrates. *Mol Marine Biol Biotechnol* **3**: 294–299.
- Fulton TL, Strobeck C (2006). Molecular phylogeny of the Arctoidea (Carnivora): effect of missing data on supertree and supermatrix analyses of multiple gene data sets. *Mol Phylogenet Evol* **41**: 165–181.
- Fyler CA, Reeder TW, Berta A, Antonelis G, Aguilar A, Androukaki E (2005). Historical biogeography and phylogeny of monachine seals (Pinnipedia: Phocidae) based on mitochondrial and nuclear DNA data. *J Biogeogr* **32**: 1267–1279.
- Greenwood PJ (1980). Mating systems, philopatry and dispersal in birds and mammals. *Anim Behav* **28**: 1140–1162.
- Hebert PDN, Cywinska A, Ball SL, deWaard JR (2003). Biological identifications through DNA barcodes. *Proc Biol Sci* **270**: 313–321.
- Hey J, Wakeley J (1997). A coalescent estimator of the population recombination rate. *Genetics* **145**: 833–846.
- Higdon JW, Bininda-Emonds ORP, Beck RMD, Ferguson SH (2007). Phylogeny and divergence of the pinnipeds (Carnivora: Mammalia) assessed using a multigene dataset. *BMC Evol Biol* **7**: 216.
- Huelsenbeck JP, Ronquist F (2001). MRBAYES: Bayesian inference of phylogenetic trees. *Bioinformatics* **17**: 754–755.
- Hugall AF, Foster R, Hutchinson M, Lee MSY (2008). Phylogeny of Australasian agamid lizards based on nuclear and mitochondrial genes: implications for morphological evolution and biogeography. *Biol J Linn Soc* **93**: 343–358.
- Hurst GDD, Jiggins FM (2005). Problems with mitochondrial DNA as a marker in population, phylogeographic, and phylogenetic studies: the effects of inherited symbionts. *Proc Biol Sci* **272**: 1525–1534.
- Jaccard P (1901). Étude comparative de la distribution florale dans une portion des Alpes et des Jura. *Bulletin del la Société Vaudoise des Sciences Naturelles* **37**: 547–579.
- Jones CJ, Edwards KJ, Castaglione S, Winfield MO, Sala F, vandeWiel C et al. (1997). Reproducibility testing of RAPD, AFLP and SSR markers in plants by a network of European laboratories. *Mol Breed* **3**: 381–390.
- Koopman WJM (2005). Phylogenetic signal in AFLP data sets. *Syst Biol* **54**: 197–217.
- Koopman WJM, Wissemann V, De Cock K, Van Huylbroeck J, De Riek J, Sabatino GJH et al. (2008). AFLP markers as a tool to reconstruct complex relationships: a case study in *Rosa* (Rosaceae). *Am J Bot* **95**: 353–366.
- Lancaster ML, Gemmel NJ, Negro S, Goldsworthy S, Sunnucks P (2007). Ménage à trois on Macquarie Island: hybridization among three species of fur seal (*Arctocephalus* spp.) following historical population extinction. *Mol Ecol* **15**: 3681–3692.
- Le Quesne WJ (1974). The uniquely evolved character concept and its cladistic application. *Syst Zool* **23**: 513–517.
- Linnen CR, Farrell BD (2007). Mitonuclear discordance is caused by rampant mitochondrial introgression in Neodiprion (Hymenoptera: Diprionidae) sawflies. *Evolution* **61**: 1417–1438.
- Luo RY, Hipp AL, Larget B (2007). A Bayesian model of AFLP marker evolution and phylogenetic inference. *Stat Appl Genet Mol Biol* **6**: 1–30.
- Madden JR, Lowe TJ, Fuller HV, Coe RL, Dasmahapatra KK, Amos W et al. (2004). Neighbouring male spotted bowerbirds are not related, but do maraud each other. *Anim Behav* **68**: 751–758.
- Mallarino R, Bermingham E, Willmott KR, Whinnett A, Jiggins CD (2005). Molecular systematics of the butterfly genus *Ithomia* (Lepidoptera: Ithomiinae): a composite phylogenetic hypothesis based on seven genes. *Mol Phylogenet Evol* **34**: 625–644.
- Mendelson TC, Simons JN (2006). AFLPs resolve cytonuclear discordance and increase resolution among barcheek darters (Percidae : Etheostoma : Catonotus). *Mol Phylogenet Evol* **41**: 445–453.
- Mueller UG, Wolfenbarger LL (1999). AFLP genotyping and fingerprinting. *Trends Ecol Evol* **14**: 389–394.
- Nazari V, Zakharov EV, Sperling FAH (2007). Phylogeny, historical biogeography, and taxonomic ranking of Parnassiinae (Lepidoptera, Papilionidae) based on morphology and seven genes. *Mol Phylogenet Evol* **42**: 131–156.
- Nei M, Li WH (1979). Mathematical model for studying genetic variation in terms of restriction endonucleases. *Proc Natl Acad Sci USA* **76**: 5269–5273.
- Nies G, Zachos FE, Hartl GB (2005). The impact of female philopatry on population differentiation in the European roe deer (*Capreolus capreolus*) as revealed by mitochondrial, DNA and allozymes. *Mamm Biol* **70**: 130–134.
- Pellmyr O, Segraves KA, Althoff DM, Balcazar-Lara M (2007). The phylogeny of yuccas. *Mol Phylogenet Evol* **43**: 493–501.
- Rice DW (1998). *Marine Mammals of the World: Systematics and Distribution*. Society for Marine Mammalogy: Lawrence, KS.
- Riedman M (1990). *The Pinnipeds: Seals, Sea Lions and Walruses*. Society for Marine Mammalogy: Lawrence, KS.
- Ronquist F, Huelsenbeck JP, van der Mark P (2005). *MrBayes 3.1 Manual draft 5/26/2005*. School of Computational Science, Florida State University: Tallahassee, FL, USA.
- Ruppell O, Stratz M, Baier B, Heinze J (2003). Mitochondrial markers in the ant *Leptothorax rugatulus* reveal the population genetic consequences of female philopatry at different hierarchical levels. *Mol Ecol* **12**: 795–801.
- Sambrook J, Fritsh EF, Maniatis T (1989). *Molecular Cloning: A Laboratory Manual*, 2nd edn. Cold Spring Harbor Laboratory Press: New York.
- Sato JJ, Wolsan M, Suzuki H, Hosoda T, Yamaguchi Y, Hiyama K et al. (2006). Evidence from nuclear DNA sequences sheds light on the phylogenetic relationships of Pinnipedia: single origin with affinity to Musteloidea. *Zool Sci* **23**: 125–146.
- Scheffer VB (1958). *Seals, Sea Lions and Walruses: A Review of the Pinnipedia*. Stanford University Press: Stanford.
- Simmons RB, Weller SJ (2001). Utility and evolution of cytochrome b in insects. *Mol Phylogenet Evol* **20**: 196–210.
- Sivertson E (1953). A new species of sea lion, *Zalophus wollebaeki*, from the Galapagos Islands. *K Norske Vidensk Selsk Forh* **26**: 1–3.
- Steppan SJ, Adkins RM, Anderson J (2004). Phylogeny and divergence-date estimates of rapid radiations in muroid rodents based on multiple nuclear genes. *Syst Biol* **53**: 533–553.
- Stoeckle MY, Hebert PDN (2008). Bar code of life: DNA tags help classify life. *Sci Am* **10**: 82–88.
- Vos P, Hogers R, Bleeker M, Reijans M, van de Lee T, Hornes M et al. (1995). AFLP: a new technique for DNA fingerprinting. *Nucleic Acids Res* **23**: 4407–4414.
- Walsh P, Metzger DA, Higuchi R (1991). Chelex 100 as a medium for simple extraction of DNA for PCR-based typing from forensic material. *Biotechniques* **10**: 506–513.
- Wolf JBW, Tautz D, Trillmich F (2007). Galápagos and Californian sea lions are separate species: genetic analysis of the genus *Zalophus* and its implications for conservation management. *Front Zool* **4**: 20.
- Wynen LP, Goldsworthy SD, Insley SJ, Adams M, Bickham JW, Francis J et al. (2001). Phylogenetic relationships within the eared seals (Otariidae: Carnivora): implications for the historical biogeography of the family. *Mol Phylogenet Evol* **21**: 270–284.

Supplementary Information accompanies the paper on Heredity website (<http://www.nature.com/hdy>)