Multiple markers and multiple individuals refine true seal phylogeny and bring molecules and morphology back in line

Tara Lynn Fulton*, † and Curtis Strobeck
Department of Biology, University of Alberta, Edmonton, Alberta, Canada T6G 2E9

Despite decades of study, some aspects of Phocidae (Pinnipedia, Carnivora) phylogeny still remain unresolved. Using the largest novel dataset to date, including all extant phocids and comprising 15 nuclear and 13 mitochondrial genes, we illustrate the utility of including multiple individuals per species in resolving rapid radiations, and provide new insight into phocid phylogeny. In line with longstanding morphological views, Pusa is recovered as monophyletic for the first time with genetic data. The data are also used to explore the relationship between genetic distance and taxonomic rank. Intraspecific sampling also highlights the discrepancy between molecular and morphological rates of evolution within Phocidae.

Keywords: rapid radiation; taxonomy; species tree

1. INTRODUCTION

The phylogeny of the true seals (Phocidae, Pinnipedia, Carnivora) remains to be completely resolved, despite decades of study. Higher-level classifications (figure 1) are widely accepted based on both morphological and molecular data (King 1966; Burns & Fay 1970; Davis et al. 2004; Ærnason et al. 2006; Higdon et al. 2007) but some species relationships remain contentious.

Only two phocid tribes, Lobodontini and Phocini, contain multiple genera. Both clades have proved difficult to resolve, probably resultant from rapid radiations. Similar situations are common across Carnivora, including the origin of Pinnipedia (Flynn et al. 2005; Fulton & Strobeck 2006; Sato et al. 2006; Ærnason et al. 2007). Even when large nuclear datasets are applied, some species relationships remain difficult to disentangle, such as for bears (Pagès et al. 2008), fur seals and sea lions (Yonezawa et al. 2009) and mustelids (Koepfli et al. 2008). The northern true seals (Phocaeanae) are comprised of pagoophilic (ice-loving) species, while the ‘southern’ true seals (Monachinae) include pagophilic (Lobodontini), temperate (Miroguni) and warm water (Monachini) tribes. Thus, unraveling these close relationships (Maddison & Knowles 2006; Carstens & Knowles 2007). Here, we employ multiple nuclear loci plus complete mitochondrial (mt) genomes for two individuals per species to generate the largest dataset yet applied to estimate the Phocidae species tree.

Using multiple individuals also provides the opportunity to address phylogenetic issues from a molecular perspective. Within Phocini, Halichoerus (grey seal) is consistently recovered within Pusa (ringed, Baikal and Caspian seals), leading to the argument that it should be reassigned to Phoca (harbour and spotted seals, plus Pusa as a subgenus) on genetic grounds (Arnason et al. 1995, 2006). Conversely, it has been argued to be the only species warranting its own genus within all Phocini based on dentition and snout morphology (Burns & Fay 1970). Assessment of inter- and intraspecific variation levels provide insight into this taxonomic issue, as well as highlight intriguing differences between the rates of morphological and molecular evolution across Phocidae.

2. MATERIAL AND METHODS

Forty-seven species were included in this study including seven non-pinniped carnivores. Two individuals from each phocid species were sequenced for each gene, except Phoca sibirica and Monachus monachus (one individual each), and Mirounga leonina (no individuals) for the nuclear dataset and M. monachus and Mirounga angustirostris (one individual each) for the mt dataset, due to sample availability. The same two individuals were not always represented, though the same individuals were used when possible. Fifteen nuclear genes from 14 unlinked regions were selected (see the

* Author for correspondence (tara.fulton@ualberta.net).
† Present address: 320 Mueller Laboratory, The Pennsylvania State University, University Park, PA 16802, USA.

electronic supplementary material, table S1). Extraction, amplification and sequencing conditions are listed in the electronic supplementary material. Alignment details, indel information and model selection methodology are listed in the electronic supplementary material. Base composition homogeneity across taxa was assessed in PAUP* v.4.0b10 (Swofford 2003) for each data-set as a whole and partitioned by codon position. The nuclear data showed no bias among taxa ($p = 0.9999$), while mtDNA data did ($p = 0$), but this bias was not observed after the mt 3rd codon position bases were removed ($p = 0.9999$). The mt 3rd codon position bases were thus excluded. Locus congruence was tested using the ILD test (Farris et al. 1995) as the partition homogeneity test (100 replicates) in PAUP*. No locus was significantly incongruent at $p = 0.01$, nor were the mt and nuclear partitions from each other ($p = 0.84$).

Maximum likelihood (ML) tree searching and bootstrapping (MLBP) was performed using RAxML v.7.0.0 (Stamatakis 2006a,b). Bayesian inference (BI) was performed in MrBayes v.3.1.2 (Huelsenbeck & Ronquist 2001; Ronquist & Huelsenbeck 2003). Details are included in the electronic supplementary material.

The effect of intraspecific sampling was examined by randomly excluding different sets of individuals (including all possible combinations in Phocina) so each species was only represented by a single taxon. Maximum parsimony bootstrap (MPBP) searches were performed in PAUP*.

Inter- and intraspecific variation was measured by the distance (logdet for mtDNA, GTR for nuclear) between each taxon pair in PAUP*.

3. RESULTS

(a) Sequencing and phylogenetic results

GenBank accession numbers GU167671–GU167877 and GU174591–GU174608 were assigned to new sequences (see the electronic supplementary material, table S3). Amplification could not be achieved in six cases (see the electronic supplementary material). Only relationships within Phocina and Lobodontini were not clearly resolved (figure 1).
(c) **Inter- and intraspecific variation**

Pair-wise nuclear distances between taxa were approximately 15–20 times lower in all comparisons than for mtDNA, when LogDet distance was used for both datasets. Average genetic distance between relevant taxa provided comparisons within species (between individuals), within genus (between species, with individuals averaged/species), between genera of the same subtribe (between species, with individuals averaged/species), between subtribes (genus comparisons, with species averaged/genus), and between tribes (genera averaged within subtribe). Genetic distances were not corrected for physical distance within the species range, population size or structure or any other factors that may affect intraspecific genetic distance.

4. **DISCUSSION**

As with other molecular studies (Davis et al. 2004; Fyler et al. 2005; Arnason et al. 2006; Fulton & Strobeck 2006; Higdon et al. 2007), we confirm the generally accepted subfamily classification structure of Phocidae (figure 1). But unlike other molecular studies, we provide the first molecular evidence for the Phocina that is congruent with morphology, supporting *Pusa* as monophyletic. There are two Phocini subtribes (Chapskii 1955): Histriophocina (*Histriophoca, Pagophilus*) and Phocina (*Phoca, Pusa, Halichoerus*). *Pusa* was historically considered a subgenus of *Phoca*, but was elevated by Scheffer (1958) to generic status. Both *Pusa* (King 1966; de Muizon 1982) and *Phoca* (*Pusa*; Chapskii 1955; McLaren 1960; Burns & Fay 1970) have been applied since, and all the species are morphologically similar. Previous molecular work placed the grey seal (*Halichoerus*) within the *Phoca–Pusa* group, often rendering *Pusa* paraphyletic (figure 2). The use of multiple individuals per species proved key for unraveling the rapid Phocini divergences. Here, *Pusa* is supported as monophyletic (figure 1), in agreement with morphology. Placement of *Halichoerus* within *Pusa* can be recovered when particular combinations including only one individual per species are used, indicating that it is probably an artefact of something like compositional or rate heterogeneity (Ho & Jermiin 2004). Because of the power of intraspecific sampling in identifying this variation in phylogenetic reconstruction based on sampling, it would probably benefit all studies of recent rapid radiations to include multiple individuals per species (see also Syring et al. 2007; Brumfield et al. 2008; Willyard et al. 2009), to highlight areas of uncertainty and increase resolution.
Figure 3. Genetic versus ‘taxonomic’ distance. For (a) nuclear genes with GTR distance and (b) mitochondrial genes with logdet distance, a series of pairwise distances were calculated from between individuals (a,a1,a2,j) up to between tribes (i,n). Distances were averaged for nested taxonomic comparisons. The approximate breaks in genetic distance between different taxonomic levels are indicated by dashed lines and outliers are noted. The range of genetic distance within and between species for nuclear loci was overlapping and a natural break was not observed (a, Phocinae spp.; a1, Pusa spp.; a2, other Phocinae spp.; b, Phoca, c, Pusa; d, Pusa versus Halichoerus; e, Phoca versus Halichoerus; i, Pusa versus Phoca; g, Histiophocina; h, Phocini; i, Phocinae; j, Monachinae spp.; k, Mirounga; l, Monachus, m, Lobodontini; triangles, within species; diamonds, within genus; circles, within tribe; squares, within subfamily).

Intraspecific sampling, particularly in combination with new gene tree/species tree analytical advances, can be a powerful tool in resolving recent rapid radiations (Carstens & Knowles 2007). These methods are primarily designed to estimate the species tree by reconciling conflict between genes when deep coalescences occur (Degnan & Rosenberg 2009). However, the data here were selected to represent a range of evolutionary rates to address phylogeny at all levels from species to subfamily. Inclusion of slowly evolving markers in this case does not lend well to analyses that first examine genes separately, as each nuclear gene alone provides very little resolution (see the electronic supplementary material). In a Bayesian framework (i.e. Liu 2008), analysing the many nuclear genes here with individually low information content leads to incredibly wide posterior distributions and difficulty in achieving convergence of all parameters. In a likelihood framework (i.e. Kubatko et al. 2009), choosing the single ML tree to represent each gene’s history ignores thousands of nearly equally probable topologies and is not in the intended spirit of the method, which is to resolve incongruence, not to extract weak signals from individual genes. Instead, a concatenation approach, as presented here, seems synergistic, illustrating what appears to be an emergent signal (Gatesy & Baker 2005), highlighting the importance of sampling for this clade and providing an important step towards resolving Phocina relationships. However, application of new species tree estimation methods to a more focused dataset for Phocina including many more individuals per species will probably now be another key to improved resolution for Phocina.

Intraspecific sampling also provided insight into taxonomic distinctions and a comparison of morphological versus molecular differentiation. All Phocina species are close genetically, particularly at nuclear loci, where their interspecific variation level is more similar to the intraspecific variation level of other phocids (figure 3a). *Halichoerus* is no more distant from any of the *Pusa* genera than *Pusa* species are to one another (figure 3; see also Davis et al. 2004). Although *Phoca* is more genetically distinct, the distance is still much smaller than between any other genera (except *Hydrurga* and *Leptonychotes*). From a purely genetic perspective, it is appropriate to roll all three genera, *Pusa, Halichoerus* and *Phoca*, into one (*Phoca, Linnaeus 1758*) as suggested previously (Arnason et al. 1995, 2006). But *Halichoerus* is morphologically distinct (Burns & Fay 1970; Nowak 1999), thus it seems premature to revoke its generic status, though its exact phylogenetic placement is not conclusive. The low genetic and high morphological divergence in *Halichoerus* is contrasted by the case of the monk seals (*Monachus*), which exhibit high morphological similarity, but extreme genetic divergence (figure 3). Monk seals are unquestionably monophyletic, but appear almost morphologically stagnant, resembling their extinct relatives (Hendey 1972; Repenning & Ray 1977; Wyss 1988). Whether the sister group of *Halichoerus* is *Pusa* (generally preferred genetically) or *Phoca + Pusa* (more preferred morphologically) depends on the analysis performed, but it is clear that including multiple individuals and multiple loci is critical to solidifying the phylogeny of this tribe, bringing molecules and morphology in line and opening an avenue of investigation into shifts in the rate of morphological evolution across Phocidae.

In contrast, DNA evidence for the Antarctic Lobodontini supports a movement away from traditional morphological groups. Morphological analyses using primarily cranial and dental characters (Hendey 1972; de Muizon 1982) placed the four Antarctic species into...
two clades: leopard + crabeater and Weddell + Ross. Molecules strongly disagree with morphology, placing the leopard and Weddell seals together (figure 1), but vary in placing the crabeater (Fyler et al. 2005; Higdon et al. 2007) or Ross (Davis et al. 2004; Arnason et al. 2006; Fulton & Strobeck 2006) seal as most basal. AFLP analysis (Dasmahapatra et al. 2009) supported the same two sister groupings as recovered here. Leopard and Weddell seals are very close genetically (figure 3), and share some superficial morphological similarity, such as their spotted coats. The Weddell seal has also been called the ‘sea leopard’ (Scheffer 1958) and the ‘false sea leopard’ (Hince 2000). Leopard seals are highly predatory but also rely heavily on krill (Rogers 2002). The crabeater seal relies almost solely on krill (Bengtson 2002) and both have specialized dentition for this food source, thus, a re-evaluation of morphology with a decreased focus on dental characters would be useful. The placement of the Ross and crabeater seals is less clear. Analysing nuclear loci alone or in combination with mtDNA provided a novel hypothesis for sequence analyses: that these species are sister (figure 1). A morphology + DNA supertree also recovered this topology (Bininda-Emonds et al. 1999), thus, this grouping may be more congruent with morphology than thought. Whatever the final resolution, the Lobodontini lineages diverged rapidly from one another, presumably near the time of their entry into the Antarctic (Fulton & Strobeck in press). Increasing nuclear markers and morphological review will be required to solidify their relationships and investigate the processes of how species invade such a divergent environment from their ancestral habitat, bringing us closer to understanding the links between environment, specialization and speciation.

Thanks to B. Shapiro, W. Murphy, I. Delisle, D. Coltman, S. Moore, G. Wilson and two reviewers for helpful comments and to the many people/organizations who provided samples. This work was funded by an NSERC research grant to CS and fellowship to TLF.

REFERENCES


Fulton, T. L. & Strobeck, C. In press. Multiple fossil calibrations, nuclear loci and mitochondrial genomes provide new insight into biogeography and divergence timing for true seals (Phocidae, Pinnipedia). J. Biogeogr.


King, J. E. 1983 *Seals of the world*, 2nd edn. London and Oxford, UK: British Museum (Natural History) and Oxford University Press.


