**Appendix 2 = Methodology, results, and discussion of sub-analyses**

Analysis of Molecular Data

 The analysis of molecular data specified the same parameters as the initial test of model selection, outlined in the main text. Results from this analysis are figured in Supplementary Figure 1.

Analysis of Morphological Data

In molecular phylogenetics, sources of systematic error are those which may cause an analysis to fail to accurately model evolutionary processes (Phillips et al., 2004). Systematic errors in molecular phylogenetics have been thoroughly, though not exhaustively, investigated (Simmons and Freudenstein, 2003; Gruber et al., 2007; Li et al., 2008; Liu et al., 2010). In morphological phylogenetics, possible systematic errors are not well-understood, and their potential to mislead phylogenetic topologies has not been well-explored (Dávalos et al., 2014). Encouragingly, recent experimental work by Zou and Zhang (20166) suggests morphological characters can provide nearly as much phylogenetic utility as molecular characters. if appropriately generated. Morphological characters are thought to be more susceptible to convergence than are molecular characters (Givnish and Sytsma, 1997). However, this may be caused by the difficulty in devising characters that accurately model phenotypic evolution, and the limited amount of morphological characters typically devisable, whereas the significantly higher amount of available molecular characters may mask the random convergence in a dataset. Additionally, the use of non-independent characters remains one of the most pernicious aspects of phylogenetic analyses (Zou and Zhang, 2016). To combat the non-independence of characters in morphological phylogenies, Zou and Zhang (2016) suggest using fewer characters, and preferentially selecting characters with multiple states (but see Brazeau, 2008 for a view on common misapplications of multi-state characters). The approach recommended by Zou and Zhang (2016) was followed in the present study.

Among paleontologists, parsimony has been the preferred method of phylogenetic inference of morphological datasets, likely owing to its greater precision (O’Reilly et al., 2016) and computational simplicity. However, Bayesian inference has been shown to be a similarly more accurate estimator of phylogeny than when applied to parsimony, even with morphological datasets (Wright and Hillis, 2014; O’Reilly et al., 2016, 2018a, 2018b; Puttick et al., 2018). In the present study, greater attention is paid to the Bayesian inference analyses. The present study allows for a comparison between the two methods, as applied to a dataset involving molecular and morphological data. The Mk model is the default model for Bayesian inference of morphological data, and was employed in the Bayesian analysis of morphological data. The stepping-stone method was used to explore the fit of the rates (equal, gamma, invgamma) to the morphological data. The stepping-stone analysis was run for 1,100,000 generations and used 50 steps. The gamma rates model had the highest marginal likelihood score, and was thus selected for the morphological data analyzed by Bayesian inference. Besides model selection, the other parameters of the Bayesian inference analysis of morphological data followed those previously outlined for the analyses of molecular data. PAUP software v4.0b10 (Swofford, 2002) was used to run the parsimony analysis using the heuristic search criterion. For this analysis, *Canis* and *Hesperocyon* were defined as outgroup taxa. Analysis of the morphology dataset employed the TBR algorithm, 100 random addition sequences, and 10,000 max trees. Bremer Decay Indices (Bremer, 1988), calculated manually in PAUP, were used to assess the strength of support for internal nodes. The 50% majority-rule consensus trees for each analysis are reported. Using Mesquite, characters were optimized onto a tree with the same topology as the 50% majority-rule consensus tree (excepting the sole polytomy at the base of Mustelidae + Procyonidae) from the morphological analysis, using both DELTRAN and ACCTRAN algorithms. Unambiguous synapomorphies were identified as those characters that appeared at the same node using both modes of optimization.

Total Evidence Analysis

In addition to the total evidence analysis using Parsimony, outlined in the main text, a total evidence analysis was performed using Bayesian inference on the same dataset. Bayesian inference was performed in MrBayes v3.2.6 (Huelsenbeck, 2001). The molecular models were the same ones selected for the molecular analysis. Rates of character evolution were specified to vary across sites and were assigned a gamma distribution. To allow each partition to vary according to its own set of parameters, all parameters were unlinked. The rate prior was set to variable for all partitions (prset ratepr=variable), to allow for different rates of evolution across partitions. *Canis lupus* was specified as the outgroup for both analyses.

Results

Molecular Only

The molecular only topology (Supplementary Figure 1) recovers a monophyletic Pinnipedia, with otariids and *Odobenus* sharing a closer relationship with one another to the exclusion of Phocidae. Musteloids are identified as the sister group of pinnipeds. Mephitids are the first family to diverge from the rest of the musteloid stock, followed by ailurids. Procyonids and mustelids are recovered as sister clades, with *Taxidea* diverging at the base of mustelids. All relationships are very strongly supported (pp=1.00). The topology of the molecular-only partition resembles that of other phylogenetic analyses of molecular datasets (Flynn and Nebdal, 1998; Flynn et al., 2005; Arnason, 2006; Fulton and Strobeck, 2007; Sato et al., 2010). This is unsurprising, though still noteworthy, as the inclusion of additional otariid and phocid taxa did nothing to disrupt the inter-familial relationships identified or their level of support from posterior probability values. The strong posterior probability values suggest the relationships identified in the molecular-only analysis can also be confidently used to construct a backbone for the later parsimony analyses.

Morphology Only, Bayesian inference---. A crown group Pinnipedia, including odobenids, otariids, phocids, and desmatophocids is recovered, though poorly supported (pp=0.54), in the morphological analysis (Supplementary Figure 2). Within this crown group, an unconventional coupling of otariids and phocids is recovered. This aberrant otariid-phocid pairing is reasonably well-supported (pp=0.72). Desmatophocids are recovered as sister to this group (pp=0.72). Monachines are not recovered as monophyletic, but along the sequence leading to a well-supported monophyletic Phocinae (pp=1.00). *Puijila* is recovered as the earliest diverging member of the pinniped lineage lineage producing Pinnipedia, after its divergence from the common ancestor of Musteloidea and Pinnipedia (pp=0.82). A clade resembling ‘Oligobuninae’, though also including *Pseudobassaris*, is recovered as the sister group to pinnipeds, though with poor support (pp=0.13).

Without a molecular backbone constraint, most previously published morphological phylogenetic analyses (excepting Kohno, 1996) recover Ursidae as the sister to pinnipeds (Berta and Wyss, 1994; Tedford et al., 1994; Furbish, 2015). In contrast, the morphological analysis in the present study recovered musteloids as sister to pinnipeds, which is consistent with analyses of molecular data (Flynn et al., 2005; Arnason, 2006; Fulton and Strobeck, 2007; Sato et al., 2010). However, the relationships within Musteloidea, found in the morphological analysis, are not in total agreement with the molecular analysis, as mephitids are recovered within the mustelid radiation and ailurids are excluded from the group entirely. Unexpectedly, the morphological analysis also recovered otariids and phocids as sister taxa. To our knowledge, otariids and phocids have not previously been recovered as sister groups, to the exclusion of odobenids, in any published phylogenetic analysis, though an early divergence of odobenids was suggested by Lento et al. (1995). In the present analysis, this unconventional relationship is presumably caused by a lack of transitional phocids and otariids. *Prototaria*, the earliest-diverging odobenid in the analysis, is a conspicuous link between the hypothetical last common ancestor of pinnipeds and later-diverging odobenids (Kohno, 1994, Kohno et al., 1995). On the other hand, *Devinophoca* and *Thalassoleon*, the earliest-diverging phocid and otariid, respectively, in the analysis, share many features with one another that may have arisen in parallel as the two lineages became more aquatically specialized. Many of these shared features are absent in *Prototaria*, and *Proneotherium*, another early-diverging obodenid, but appear in later-diverging odobenids like Imagotaria. *Pinnarctidion* is recovered as sister to the pinniped crown group, rathern than with the Desmatophocidae. *Desmatophoca* and *Allodesmus* display several proposed pinniped synapomorphies (Berta and Wyss, 1994) and appear well-removed from pinnipedimorphs. On the other hand, *Pinnarctidion* retains many features otherwise thought to characterize stem pinnipeds. These include: retention of an intrabullar pseudoseptum; presence of pseudosylvian sulcus; presence of fossa muscularis; presence of protocone shelf on P3 ; retention of fovea for ligamentum teres femoris; unreduced trochanteric fossa; presence of nasolabialis fossa; small antorbital process; large basioccipital embayment for the inferior petrosal sinus.

Desmatophocidae has been consistently recovered as sister to Phocidae within the proposed ‘Phocoidea’ (Wyss and Flynn, 1993; Berta and Wyss, 1994; Kohno, 1996b; Deméré and Berta, 2002; Boessenecker and Churchill, 2018; Tonomori et al., 2018), with some of these analyses recovering Odobenidae as the successive sister taxon to this group (Wyss and Flynn, 1993; Berta and Wyss, 1994; Deméré and Berta, 2002). The present analyses are ambiguous with regards to the phylogenetic position of Desmatophocidae. Although several characters have been recognized as synapormorphies for ‘Phocoidea’, many of these are cranial characters for which ontogenetic information is not yet available. The taxonomic position of Desmatophocidae and confirmation of the validity of ‘Phocoidea’ must await further phylogenetic revision of the Otarioidea and Desmatophocidae within the context of Pinnipedia.

The most notable transgression in the morphological analysis is the outdated placement of mephitids (Mephitis + Promephitis) as sister to the lutrine group. While morphological phylogenies consistently identify skunks as mustelids (excepting the phenetic inferences of Radinsky, 1973), molecular evidence overwhelmingly favors a branching event of Mephitidae towards the base of Musteloidea (Dragoo and Honeycutt, 1997; Flynn et al., 2005; Sato et al., 2010). In spite of their mustelid-like morphology, extant mephitids display a mosaic of ancestral/primitive (retention of postprotocrista of M1, short palate) and uniquely-derived (reduction of osseus tentorium, presence of mastoid sinus, absence of entepicondylar foramen of the humerus, reduction of baculum) features that reveal their independent derivation from the musteloid stock.

Morphology Only, Parsimony---.

The parsimony analysis of the Morphology Only dataset, without constraints, recovered 48 most parsimonious trees with 1036 steps. In the strict consensus tree (Supplementary Figure 3 *Puijila* and *Potamotherium* are reconstructed as the most basal members of Pinnipedimorpha group. A procyonid-mustelid clade is recovered as the sister group to the pinnipedimorphs, to the exclusion of ailurids. As in the Bayesian analysis of morphological data, Mephitidae is recovered as sister to Lutrinae. While the tree, in general, is moderately well-resolved (consistency fork index [CFI]=0.797), the base of pinnipeds is not well-resolved, recovering a polytomy of phocids, otarioids, odobenids and desmatophocids.

Character optimization indicates that *Kolponomos*’ allocation to the pinnipedimorph branch, while well-supported, occurs despite a high number of reversals. *Kolponomos* appears to possess many character states otherwise restricted to Odobenidae, though it is unclear if these are phylogenetically-driven or the result of convergence upon a benthic feeding style. *Kolponomos* also displays numerous autapomorphies, and shares several character states exclusively with Ailurus, as well as several other character states that do not easily conform to the discrete character states of this analysis. Its peculiar morphology makes it difficult to align Kolponomos with any other arctoids, though its arctoid affinities are well-established (Stirton, 1960; Tedford et al., 1994; Kohno, 1996a). In a parsimony analysis, Boessenecker and Churchill (2013) identified five unequivocal synapomorphies of the Odobenidae, three of which are present in *Kolponomos* (large, thickmargined, dorsoventrally-elliptical narial opening; dorsoventrally thick and laterally broad pterygoid strut; triple-rooted M1 ), and two of which are not presently codeable for *Kolponomos* as they would require computer tomography (CT) or destructive preparation (possession of a large epitympanic recess; bony tentorium appressed to petrosal). Additionally*, Kolponomos* may possess the equivocal clade synapormorphy suggested in their analysis: the presence of well developed cuspules on the P1-2 lingual cingulum, though the appearance and configuration of these cuspules in *Kolponomos* deviates from the typical odobenid form. Kohno (2006) identified another unambiguous synapomorphy of Odobenidae – antorbital process split by maxillary-frontal suture – that is absent in *Kolponomos. Kolponomos* also does not possess any of the proposed synapomorphies for crown pinnipeds. *Kolponomos* appears to have converged upon an Odobenus-like morphology, but may still have some bearing on early pinniped evolution. At the present moment, it is probably most appropriate to place it in Arctoidea Incertae sedis, until its postcrania or internal cranial architecture become available for study

Total Evidence Analysis, Bayesian---.

The results of the total evidence analysis using Bayesian inference (TEA) are displayed in Supplementary Figure 4. The TEA recovers a topology in agreement with molecular analyses on the relationships between extant families. Otariidae and Odobenidae are recovered as sister groups, rounded out by Phocidae in a monophyletic Pinnipedia. Musteloids are identified as the sister group to pinnipeds. Many of the other relationships are not strongly supported by posterior probability values. Within the odobenines, a clade of early-diverging pinnipedimorphs is recovered. Within the clade are the desmatophocids and enaliarctines, along with *Puijila, Potamotherium, Semantor*, and *Kolponomos. Enaliarctos* is recovered as paraphyletic, as the different species are distributed in a clade also containing *Kolponomos.* Besides a moderately well-supported coupling of *Potamotherium* and *Puijila* (pp=0.75), all of these relationships are very poorly supported (pp=0.51). Monophyly of the Desmatophocidae + *Pinnarctidion* is moderately well-supported in the total evidence analysis (pp=0.66), and a closer association between *Allodesmus* and *Desmatophoca* is strongly supported.

Discussion

Total Evidence Analyses

Total evidence analyses (TEAs), while commonly applied to many other vertebrate groups (Larson and Dimmick, 1993; O’Leary, 1999; Schulte et al., 2003; Weksler, 2006; Maddin et al., 2012), have only recently begun to permeate the study of carnivorans (Finarelli, 2008; Churchill et al., 2014; Furbish, 2015; Wang et al., 2017). Molecular characters are often found to be more suitable than morphological characters in phylogenetic inference (Page and Holmes, 1998; Wake et al., 2011; Davalos et al., 2012; 2014; Springer et al., 2013; Zou and Zhang, 2016). Molecular characters are easily defined and coded, with four strictly-delineated, mutuallyexclusive character states. Differentiation of morphological data into discrete characters is made difficult by the greater integration of morphological structures (Olson and Miller, 1958; Cheverud, 1995). An important assumption of cladistics is the supposed independence of characters, but such rules are necessarily transgressed by morphological phylogenies (Riedl, 1978; Leamy et al., 1999; Cardini and Elton, 2008). However, molecular data alone offers an incomplete picture on the details underlying the evolutionary history of a clade. In TEAs, the high proportion of molecular data constrains the topology by overwhelming the morphological data, so that morphological characters do not exert a strong influence on the phylogenetic position of those taxa with molecular data available. Rather, the morphological characters allow the lacunae of the lineages to be filled in by fossil extinct taxa. If a topology or phylogenetic placement is accurate, fossils may offer direct evidence of ancestral states, break up long branches (Cobbett et al., 2007), and allow for accurate description of homoplasy. The TEA in the present study using Bayesian inference produces a very unconventional topology that is highly unlikely, in light of fossil evidence (Deméré et al., 2003; Churchill et al., 2014). The combined morphological and molecular dataset is unable to recover a similar topology to either the morphological or molecular datasets, which in isolation, produce reasonable, though highly discrepant topologies. This is most conspicuous on the pinniped branch, where Early and Middle Miocene taxa are observed towards the terminal nodes, and extant, highly derived taxa are recovered towards the stem of each group. An otariid-odobenid grouping is very well-supported by molecular data, but the TEA is not able to accommodate the fossil extinct taxa into this clade without producing an illogical topology. Prototaria and Proneotherium are well-established early-diverging odobenids, connecting Odobenus to the base in the analyses of the morphological partitions. However, the only fossil otariid included in the analysis, Thalassoleon, is already highly derived, as is Pithanotaria, another early-diverging otariid that is not as completely known, and was not included in the present study. Eotaria appears to represent a more conspicuous transitional form that would theoretically connect the otariids to the base of the tree, but this taxon is known only from isolated mandibles (Boessenecker and Churchill, 2015; Velez-Juarbe, 2017). Due to this fossil otariid gap, there is less morphological disparity between the aquatically-specialized crown otariids and Odobenus than between Prototaria and Thalassoleon. Consequentially, Odobenus, and later-diverging crown otariids, are forced towards the base of their respective branches. The same pattern is observed within the phocid clade, in which the middle Miocene fossil taxon Devinophoca appears as the latest-diverging phocid, and highly derived extant taxa are pushed toward the base. When Parsimony is applied to the same dataset, the tree topology more closely resembles the topologies produced by the analyses of the other datasets, recovering less-derived fossil extinct taxa with transitional morphologies towards the stem of their identified lineage clade (e.g., Devinophoca at the base of Phocidae), as expected.

Evolution of early-diverging pinnipeds

If *Enaliarctos* and *Pteronarctos*, both flippered taxa, are stem pinnipeds, this suggests the last common ancestor of crown pinnipeds was flippered. Even in this scenario, parallelisms still appear rampant within the three crown pinniped families. Most striking is the trend towards homodonty and reduction in the length of the tooth row (Boessenecker, 2011). The earliest phocid and odobenid known from cranio-dental material display heterodonty, though their carnassials are becoming more homodont and falling in file with the remaining postcanines (Kohno, 1994; Koretsky and Halec, 2002). *Enaliarctos*, though it retains multicuspidate and disparately shaped postcanines, is trending toward a more homodont dentition. A transitional series, with regards to homodonty, can already be established within *Enaliarctos*, from the more heterodont *E. mealsi and E. barnesi,* to the nearly homodont *E. mitchelli* (Berta, 1991). Likewise*, Puijila* and *Potamotherium* display reduced molars, indicating a reduction of the tooth row. While reduction and/or loss of the M2 / M2 also characterises neomustelids, the M1 of mustelids is rarely as reduced, and is commonly larger than their premolars. The M1 of *Puijila* and *Potamotherium* is buccolingually narrow, and similar in size and shape to the M1 observed in USNM 314295 (*Enaliarctos barnesi*). Evidently, selective pressures favouring homodonty were already affecting these early pinniped ancestors, as they became increasingly specialized for an aquatic existence. While the dentition of *Enaliarctos* suggests a raptorial feeding style (Adam and Berta, 2002; Churchill and Clementz, 2016) in which the jaws and teeth are the only means of capturing prey (Hocking et al., 2017), its skeletal morphology suggests it was likely capable of manipulating its prey with its forelimbs (Berta and Ray, 1990; Hocking et al., 2017). It appears likely *Enaliarctos* still practiced, at least facultatively, though possibly habitually, a semi-aquatic feeding style (Hocking et al., 2017), in which the prey may be captured under water and consumed or processed at the surface or on land, as seen in many modern otter taxa. If *Enaliarctos* is emblematic of the hypothetical last common ancestor of crown pinnipeds, then crown pinnipeds ostensibly developed many of their extreme aquatic specializations in parallel. Hocking et al. (2017) suggests aquatic mammals are limited to a small number of feeding styles, partly dependent on their degree of aquatic adaptation. Within the framework of Hocking et al., (2017) *Enaliarctos*, and other pinnipedimorphs may have employed both semi-aquatic and raptorial feeding styles. A transition to obligate raptorial feeding in such organisms would not be unsurprising. Such an event could have transpired as oceans cooled during the Miocene, when the earliest known fossil extinct taxa confidently ascribed to Otariidae, Phocidae, and Odobenidae appear. Specializations associated with raptorial feeding in the crown pinniped families appear to have arisen independently in otariids and phocids, as suggested by the present dataset, which identifies independent transitions towards homodonty in otariids and phocids (Table 3). Convergent evolution of aquatic feeding behaviours and related morphologies within closely-related lineages has previously been documented in the fossil record of Ziphiidae (Cetacea) (Bianucci et al., 2016).