The functional diversity of marsupial limbs is influenced by both ecology and developmental constraint

Spencer K. PEVSNER[1,3], David M. GROSSNICKLE[2], Zhe-Xi LUO[3]

[1]School of Earth Sciences, University of Bristol, Bristol, UK

[2]Department of Biology, University of Washington, Seattle, USA

[3]Department of Organismal Biology & Anatomy, University of Chicago, Chicago, USA

ABSTRACT

Extant marsupials are less ecologically diverse than placentals, and this is reflected by placentals exhibiting a greater diversity of locomotor modes, including powered flight and fully aquatic swimming. One proposed explanation for this discrepancy is that the development of more disparate marsupial forelimbs is prevented by the neonate's crawl to the pouch, which requires precocious forelimb development for climbing adaptations. To test predictions of this Developmental Constraint Hypothesis, we pursue a comparative morphometric study on osteological traits of mammalian limbs, with an emphasis on functional differentiation of marsupial limbs among locomotor modes. We apply multivariate analyses to a large dataset of limb metrics and a diverse sample of mammals, with the placental sample limited to taxa whose locomotor modes are exhibited in marsupials. Overall, we do not find consistent evidence in support of the Developmental Constraint Hypothesis. Diprotodontia serves as an exception, with comparisons of their forelimbs to hind limbs supporting the Developmental Constraint Hypothesis. Our results suggest that developmental constraints on marsupial forelimbs may have limited marsupial diversity to some degree. Despite this, the marsupial locomotor groups show unexpectedly high levels of morphological differentiation relative to placentals of the same locomotor modes, indicating that ecological functions may overcome developmental constraints on a macroevolutionary scale.

Keywords: developmental constraint – ecomorphology – functional morphology – locomotion – Mammalia – phylogenetic comparative methods

INTRODUCTION

Extant mammals exhibit considerable functional diversity; the breadth of ecomorphotypes displayed by mammals is spectacular, far greater than other major clades of extant tetrapod vertebrates (Nowak, 1999; MacDonald, 2002). This diversity is exemplified by their numerous locomotor behaviors, including running, burrowing, climbing, gliding, powered flight, swimming, semi-aquatic and fully aquatic life (Hildebrand & Goslow, 2001). Although this functional diversity is most visible in placental mammals, marsupials have also achieved considerable diversity, displaying semi-fossorial, semi-aquatic, terrestrial, scansorial, arboreal, saltatorial, and gliding adaptations (Nowak, 2005). For example, *Chironectes* is a semi-aquatic opossum with webbed feet (pedes) well-adapted for swimming, *Notoryctes* has especially stout limbs that are well adapted to digging and its subterranean life, and *Petaurus* is a glider with a patagium (skin membrane) that acts as an airfoil. However, several derived locomotor modes of placental mammals have never evolved in marsupials, including fully aquatic swimming, as exemplified in placentals by whales and manatees, and powered flight, as seen in bats (Nowak, 1999; Nowak, 2005).

The conventional hypothesis proposed to explain this discrepancy in locomotor diversity between placentals and marsupials is that marsupials are constrained developmentally by the necessity for neonates to make a rigorous crawl from the birth canal to the teat. We refer to this as the Developmental Constraint Hypothesis (DCH). The neonate crawl requires forelimbs to develop much earlier in ontogeny in marsupials than in placentals (Lillegraven, 1975; Gemmell et al., 2002; Sears, 2004; Keyte & Smith, 2010; Cooper & Steppan, 2010; Kelly & Sears, 2011a; Kelly & Sears, 2011b). Additionally, this crawl requires the forelimbs to have a critical climbing ability – the neonate must be able to crawl to the maternal pouch to begin suckling, and their precocially developed forelimbs must have some traits that allow climbing, such as digits capable of grasping. According to the DCH, this requirement of well-developed precocial forelimbs with climbing ability precludes the development of disparate limb structures like flippers or true wings, thus preventing marsupials from entering a wider range of ecological niches with fully aquatic locomotion, or powered flight (Lillegraven, 1975; Sears, 2004; Cooper & Steppan, 2010; Kelly & Sears, 2011a; Kelly & Sears, 2011b).

The DCH is a rigorous hypothesis that offers clear predictions, as outlined concisely by Sears (2004). One prediction is that the morphological disparity of placental limbs and girdles will greatly exceed the disparity of marsupials (Fig. 1). A second prediction is that the magnitude of constraints on forelimb morphology will vary within marsupial phylogeny (Fig. 1), because styles of crawls by the newborns are different among marsupial groups (Gemmell et al., 2002). Sears (2004) predicted that the effects of the DCH should be most pronounced in Diprotodontia, as the newborns of this group show the most rigorous crawl from the birth canal to the teat. As such, they should experience the strongest developmental selection for climbing ability, which should diminish the potential for forelimb evolutionary experimentation and disparity (Gemmell et al., 2002; Sears, 2004; Martín-Serra & Benson, 2020). By contrast, the newborns of Peramelemorphia have a much shorter crawl that requires much less climbing than the newborns of other marsupial orders, so they are expected to experience less developmental constraint on their forelimb morphology (Gemmell et al., 2002; Garland et al., 2017).

Many previous studies have supported the DCH (Lillegraven, 1975; Gemmell et al., 2002; Sears, 2004; Keyte & Smith, 2010; Cooper & Steppan, 2010; Kelly & Sears, 2011a; Kelly & Sears, 2011b). For instance, Sears (2004) determined that marsupials have a lower level of morphological diversity than placentals in the girdle structures, and inferred developmental constraint for marsupials. Additionally, Sears (2004) found that the marsupial species that perform more rigorous neonatal crawls show similar patterns of ontogenetic shape change in their scapulae, while such shared patterns were not observed in marsupial taxa that did not make such extensive crawls. Further, Cooper & Steppan (2010) and Kelly & Sears (2011a) found evidence for the DCH by examining limb elements via linear morphometrics. Cooper & Steppan's (2010) approach is similar to that of Sears (2004) by examining levels of disparity; but Cooper & Steppan examined a greater number of limb elements than shoulder and pelvic structure. They found that marsupial forelimbs occupy a far more limited region of morphospace compared to placental forelimbs, and additionally found that this discrepancy is less pronounced in hind limbs. Kelly & Sears (2011a) compared the morphospace occupation of limb elements in different taxa representing different locomotor modes across placental and marsupial mammals, finding that placentals extend further from the origin of the morphospace than marsupials for each locomotor mode examined. Kelly & Sears (2011b) continued to build upon the results of Sears (2004), examining covariance of limb traits (following Young & Hallgrimsson, 2005, and Schmidt & Fischer, 2009), and they found further evidence of forelimb differences among marsupial groups that vary with how demanding their crawls are. This suggests the presence of locomotor-based pressures on forelimb functional morphology in marsupials. Beyond studies on limbs and girdles, evidence for developmental constraints on marsupial disparity has also been observed in jaws (Fabre et al., 2021; Conith et al., 2021) as well as in ossification heterochrony patterns in postcranial material (Weisbecker et al., 2008) and integration of limb elements (Bennett & Goswami, 2011; Goswami et al., 2014; Kelly et al., 2019), suggesting that marsupial reproductive behavior may broadly influence many aspects of marsupial morphology.

An alternative hypothesis to explain the discrepancy in the morphological diversity between marsupials and placental mammals is that on a macroevolutionary scale, selective pressures associated with ecological function have counter-balanced or superseded the influence of developmental constraints that may or may not exist (Kirsch, 1977a; Kirsch, 1977b; Sánchez-Villagra 2013; Martín-Serra & Benson, 2020). Notably, ecological drivers of morphological diversification might exert a strong influence on adult phenotype, even if there was an ontogenetic developmental constraint on limbs in neonates of marsupials. This hypothesis was first proposed by Kirsch (1977a) in response to Lillegraven (1975), and it opposed the DCH by invoking the ecological specialization of marsupials such as *Notoryctes*, the marsupial mole, and *Chironectes*, the water opossum. This hypothesis has also been supported by a recent study by Martín-Serra & Benson (2020) on the morphological disparity of marsupial forelimbs and hind limbs. Martín-Serra & Benson (2020) tested predictions of the DCH, such as the expectation of greater morphological disparity in marsupial hind limbs than in forelimbs, and their results do not support the DCH. They instead concluded that there must be other evolutionary drivers that are overpowering the developmental constraints on the neonates of marsupials – however, their study did not include placentals for comparison to marsupials. Sánchez-Villagra (2013) further raised the possibility that the paleobiogeography of marsupials had been a contingency in their early evolution. This may have constrained their diversification, beyond potential developmental biases resulting from the neonatal crawls of these mammals. Under this alternative hypothesis to the DCH, we predict that marsupial forelimb morphology would be no less correlated with locomotor mode than hind limbs, because the forelimb is no more constrained by development than the hind limb, or that morphologies of both limbs are strongly correlated with locomotor modes more directly relevant to ecologies.

In contrast to many of the previous studies on the DCH, we test the DCH from an ecomorphology-first perspective by directly incorporating the ecological variable of locomotor mode into phylogenetic comparative analyses. Many skeletal traits of the mammalian limb are strongly correlated with locomotor modes (Argot, 2001; Argot, 2002; Polly, 2007; Samuels & Van Valkenburgh, 2008; Salton & Sargis, 2009; Chen & Wilson, 2015; Meng et al., 2017; Nations et al., 2019; Grossnickle et al., 2020; Weaver & Grossnickle, 2020), due in part to shared biomechanical functional adaptation by taxa with the same locomotor mode. For instance, the length of the ulnar olecranon process, which approximates the inlever length of the triceps brachii muscle, correlates strongly with locomotor modes because of its strong influence on the speed and strength of various forearm movements used during locomotion (Hildebrand & Goslow, 2001). Thus, simple linear measurements that reflect biomechanical traits can help to capture a key aspect of functional diversity (Hildebrand & Goslow, 2001; Cooper & Steppan, 2010; Chen & Wilson, 2015). Here, we apply limb correlates of locomotor modes (i.e., linear measurements of skeletal limb traits) to perform “function-informed morphometrics” (Feilich & López-Fernández, 2019). With this morphometric approach, we examine the ecomorphological diversification of mammalian forelimbs and hind limbs. This approach can be useful in examining the influence of ecological diversification on the morphological evolution of mammals (Chen & Wilson, 2015; Grossnickle et al., 2020; Weaver & Grossnickle, 2020).

We use functional traits of limb (including the manus and pes) and girdle bones (Table S1) to test two predictions of the DCH on the expected influence of forelimb developmental constraints on marsupial ecological diversity (Fig. 1). First, the DCH predicts that the marsupial forelimbs will demonstrate less ecomorphological diversification than marsupial hind limbs (Sears, 2004), and this is supported if marsupial hind limbs outperform marsupial forelimbs as correlates of locomotor modes. Second, the DCH predicts that marsupial forelimbs will demonstrate less ecomorphological diversification than placental forelimbs. This will be supported if placental forelimbs outperform marsupial forelimbs as correlates of locomotor modes. These predictions are similar to those tested by previous studies (Sears, 2004; Cooper & Steppan, 2010; Kelly & Sears, 2011b; Martín-Serra & Benson, 2020), but our predictions specifically refer to ecomorphological diversification, in addition to the morphological disparity that had been the focus of previous studies. Although we also measured morphological disparity for testing hypotheses, we emphasized ecomorphological differentiation of limb morphologies by incorporating a key ecological variable - locomotor modes - into phylogenetic comparative analyses. Our analyses incorporate measurements of forelimbs (including the scapula) and hind limbs (including the pelvis) across a taxonomically diverse sample of marsupials and placentals. Our results provide novel insight on the DCH and highlight the importance of incorporating ecology and phylogeny into tests of the DCH.

MATERIALS AND METHODS

Our dataset consists of 142 mammal species, with 31 skeletal measurements (Fig. S1) collected for each species (one specimen per species). Specimens were measured in the mammalogy collections of the Field Museum of Natural History, American Museum of Natural History, Yale Peabody Museum of Natural History, and the Burke Museum of Natural History and Culture. We measured 63 marsupial species, representing 40 genera. This dataset encompasses representatives from all extant marsupial orders: Paucituberculata, Didelphimorphia, Diprotodontia, Microbiotheria, Notoryctemorphia, Peramelemorphia, and Dasyuromorphia (Fig. 2). In addition, data for 79 placental species (70 genera) was obtained from Weaver & Grossnickle (2020), which built upon the dataset of Chen & Wilson (2015). Our full dataset consists of mostly small-to-medium sized mammals (< 40 kg; Jones et al., 2009); and this dataset is built using measurements that make it comparable to other datasets that examined only smaller mammals and were informative in parsing locomotor and other ecomorphological diversities of small mammals (Chen & Wilson, 2016; Nations et al. 2019; Grossnickle et al., 2020; Weaver & Grossnickle, 2020). In order to highlight the functional similarities or differences between marsupials and placentals datasets within each locomotor mode, we excluded placental taxa that display locomotor modes not also found in extant marsupial taxa, such as graviportal and fully aquatic taxa (see justification below). This has limited the body size range of the taxa included in our dataset.

The 31 measured skeletal traits, which are illustrated in Figure S1, include many that were measured by previous studies (Samuels & Van Valkenburgh, 2008; Chen & Wilson, 2015; Meng et al., 2017; Grossnickle et al., 2020; Weaver & Grossnickle, 2020). For each specimen examined, photographs were taken of the skull, girdles, and limb elements. Skull lengths were collected because these are common proxies for body size. The limb and girdle skeletal elements examined included the scapula, humerus, radius, ulna, metacarpals, manual phalanges (hereafter collectively referred to as ‘forelimb’ measurements), pelvis, femur, tibia, fibula, calcaneum, metatarsals, and pedal phalanges (hereafter collectively referred to as ‘hind limb’ measurements). Linear measurements were collected from photographs with scale bars using *ImageJ* (Schindelin et al., 2012). To size-correct the data, we converted the measurements into log-shape ratios (Mosimann, 1970), which involves dividing each measurement by the geometric mean of all measurements (forelimb or hind limb) and then $\log_{10}$-transforming the resulting ratio (Claude, 2013). Separate geometric means were calculated for forelimb measurements and hind limb measurements to ensure that forelimb effects were not affecting hind limb data, and vice versa. Both geometric mean measurements did include skull lengths to more fully account for overall body size. Log-shape ratios were used in all downstream analyses.

Our study includes hand and feet traits that are directly linked to substrate preference in mammalian ecology (Argot, 2001; Argot, 2002; Bloch & Boyer, 2003; Luo et al., 2003; Kirk et al., 2008; Chen & Wilson, 2015; Meng et al., 2017; Nations et al., 2019). With the exception of Cooper & Steppan (2010), manus and pes skeletal traits are not included in some previous disparity studies of marsupial girdles and limbs (Sears, 2004; Kelly & Sears, 2011b; Martín-Serra & Benson, 2020). One reason for their exclusion in other analyses may be that these skeletal traits can be challenging to measure because in many museum specimens these parts of the skeleton are still bound in mammal skins. By including measurements of manual and pedal skeletal elements, our dataset is more comprehensive in coverage of these functional traits that are directly related to substrate preferences in the ecological niches of these mammals (Kirk et al., 2008; Nations et al., 2019).

We assigned all species in our sample to one of four locomotor modes: terrestrial, arboreal (including scansorial), glider, and saltatorial. Although some species in our sample are commonly classified as semi-aquatic, semi-fossorial, or fossorial, we grouped all of these taxa into the terrestrial locomotor mode due to small sample sizes for marsupials of each locomotor category. For example, *Chironectes minimus* is the only semi-aquatic marsupial species in our sample. Terrestrial taxa are defined as living primarily on the ground; arboreal (and scansorial) as frequently climbing or living in the trees; gliding as having the ability to glide; and saltatorial as primarily locomoting by leaping or bounding. Our classification scheme is subjective and could influence results. Thus, we tested the sensitivity of our analyses to our locomotor mode classification scheme by repeating some analyses using alternative schemes: classifying arboreal and scansorial taxa separately, condensing locomotor modes to tree-dwelling and ground-dwelling only, and expanding categorization to include semi-fossorial/fossorial (see Supporting Information). Sources for locomotor mode classifications include Nowak (1999), Chen & Wilson (2015; and sources within), the Animal Diversity Web (https://animaldiversity.org), and the International Union for Conservation of Nature (IUCN) Red List (https://www.iucnredlist.org).

To adequately compare the ecomorphological diversity in marsupials and placentals, we only included placentals with locomotor modes that are also found in marsupials (i.e., terrestrial, arboreal, gliding, and saltatorial). This ensures that the ecomorphological breadth of the two samples being compared are similar, because we are comparing the ecomorphological differentiation of taxa of four specific locomotor modes. Further, Sears (2004) also analyzed marsupials and placentals of these four locomotor modes, noting that they have evolved independently in multiple mammalian orders and provide adequate sampling for statistical analyses.

We employed three types of multivariate analyses to examine ecomorphological diversification of mammalian limbs. First, we used a phylogenetic flexible discriminant analysis (pFDA; Motani & Schmitz, 2011; Schmitz & Motani, 2011; Angielczyk & Schmitz, 2014) to examine how well limb traits differentiate among locomotor modes. Discriminant analyses can be used effectively on large datasets of skeletal measurements for modern taxa with known locomotor modes (Samuels & Van Valkenburgh, 2008; Samuels et al., 2013; Chen & Wilson, 2015; Kilbourne, 2017), and pFDAs in particular have been used in previous studies to examine differentiation in limb morphologies among mammals of various locomotor modes (Maga & Beck, 2017; Grossnickle et al., 2020). Here, we apply discriminant analyses separately on marsupials and placentals, and separately on forelimbs and hind limbs, and use them to examine how many species are correctly classified into their known locomotor modes using posterior probabilities. We interpret larger percentages of correctly classified species to indicate greater ecomorphological differentiation among locomotor modes. In conducting the pFDAs, we calculated optimal Pagel’s lambda values (Motani & Schmitz, 2011; Schmitz & Motani, 2011; Angielczyk & Schmitz, 2014), which adjusts the strength of the phylogenetic bias that is being removed from the analysis. In most analyses, the optimal Pagel’s lambda value was zero (in which case the analyses are FDAs instead of pFDAs), so we repeated the analyses with a

Pagel’s lambda value of 0.1 to test the influence of phylogeny on the results (Angielczyk & Schmitz, 2014; Grossnickle et al., 2020).

Second, we used phylogenetic multivariate analyses of variance (pMANOVAs) to test for morphological differences among taxa of different locomotor modes (Fig. 4), using the *RRPP* package (Collyer & Adams, 2018) for R (R Core Team 2021). Post-hoc pairwise comparisons were used to examine differences in morphologies between individual locomotor modes. To test the influence of phylogeny on the pMANOVA results, we also performed MANOVAs.

Third, we calculated morphological disparity, including pairwise comparison (with *p*-values calculated from simulations) using the *geomorph* R package (Adams & Otárola-Castillo, 2013). We also conducted a variation on the disparity test in which we removed two measurements (fibular proximal width and metatarsal length) that we were concerned were biasing our results (see Supporting Information for more details). All three types of analyses were repeated using each subset of measurements (forelimb, hind limb) as well as various mammalian subclades, including placentals, marsupials, and marsupial subclades. The specific marsupial subclades examined are Didelphimorphia, Diprotodontia, ‘Ameridelphia’ [Paucituberculata + Didelphimorphia], and Australidelphia [Microbiotheria + Notoryctemorphia + Peramelemorphia + Dasyuromorphia + Diprotodontia]. Didelphimorphia and Diprotodontia were selected because they each include a large enough sample to be independently examined – in particular, Diprotodontia is very ecomorphologically diverse, containing all of the marsupial gliders and saltators. While the other marsupial orders do not contain a sample size large enough to be examined on their own, their effects can be inferred by examining the larger supergroups of ‘Ameridelphia’ and Australidelphia and comparing their results to those of

Didelphimorphia and Diprotodontia. In this respect, ‘Ameridelphia’, while paraphyletic, is useful for determining the influence of the DCH on Paucituberculata.

The phylogeny used in the pFDAs and pMANOVAs is the species-level mammal tree of Upham et al. (2019) (Fig. 2). We generated a maximum clade credibility tree (Drummond et al., 2012) using 1000 randomly selected trees from the posterior distribution of the ‘completed trees’ analysis, and then pruned the tree to the species included in our analyses (placentals, marsupials, or marsupial subgroups).

Data availability

The data from this study are available in the Supporting Information in Table S1.

RESULTS

Based on the skeletal limb traits and multivariate analyses, marsupials and placentals show equally strong morphological differentiation for the examined locomotor modes when examining all traits. This is most apparent in the pFDA (Fig. 3, Table 1), MANOVA (Table 2), and disparity results (Fig. 4; Tables S2, S4, and S5), although it is not supported by the pMANOVA results (Tables 2, S6 and S7). The pFDA results are similar for Marsupialia and Placentalia – the analyses correctly classify the locomotor mode of a majority of species. Discriminatory power decreases slightly when examining forelimb or hind limb traits alone (Table 1), but the analyses are still able to correctly classify most species. For forelimb and hind limb results, marsupials show no difference between misclassification percentages while placentals show greater misclassification using their hind limb traits than their forelimb traits, but the difference in placentals – 1.3% when phylogeny is not taken into account and 5% when it is – is a nominal number of new taxa being misclassified by the analysis, so this difference is negligible. The similarity of results between marsupials and placentals is also apparent in our morphological disparity results (Fig. 4). Across all traits and forelimb traits alone, marsupials and placentals show near identical levels of disparity that are not significantly different (Table S2). For hind limb traits, the difference between their respective disparity levels runs counter to expectations of the DCH; marsupials show greater disparity levels than placentals, with that difference being statistically significant (Fig. 4; Table S2).

Table 1. The percentages of species correctly classified into locomotor modes by the phylogenetic flexible discriminant analyses (pFDAs). 'All' analyses include all forelimb and hind limb traits. The optimal lambda (λ) value is zero for each analysis, and thus we repeated the analyses to examine the influence of phylogeny by incorporating a lambda value of 0.1.

	FDA (λ = 0)			**pFDA (λ = 0.1)**		
	All	**Forelimb**	**Hind limb**	**All**	**Forelimb**	**Hind limb**
Marsupialia (*n* = 63)	98.41%	93.65%	93.65%	98.41%	92.06%	92.06%
'Ameridelphia' (*n* = 26)	100%	96.15%	100%	100%	96.15%	100%
Didelphimorphia (*n* = 23)	100%	95.65%	100%	100%	95.65%	100%
Australidelphia (*n* = 37)	100%	97.3%	100%	100%	94.59%	100%
Diprotodontia (*n* = 23)	100%	100%	100%	100%	100%	100%
Placentalia (*n* = 79)	93.67%	86.1%	84.8%	93.67%	87.3%	82.3%

Table 2. Statistical results for MANOVAs and pMANOVAs, comparing multivariate morphologies of taxa of different locomotor modes. For marsupials, australidelphians, diprotodontians, and placentals, the analyses test for morphological differences between four locomotor modes (terrestrial, arboreal, saltatorial, and gliding), and for ameridelphians and didelphimorphians the analyses test for differences between terrestrialists and arborealists (due to the lack of gliders and saltators).

	MANOVA				**pMANOVA**			
	Forelimb		**Hind limb**		**Forelimb**		**Hind limb**	
	***p*-value**	***F*-stat**	***p*-value**	***F*-stat**	***p*-value**	***F*-stat**	***p*-value**	***F*-stat**
Marsupialia (*n* = 63)	<0.001	5.488	<0.001	7.599	0.167	1.367	0.208	1.423
'Ameridelphia' (*n* = 26)	0.059	2.526	0.042	3.393	0.210	1.387	0.157	1.508
Didelphimorphia (*n* = 23)	0.075	2.633	0.043	4.914	0.208	1.425	0.167	1.448
Australidelphia (*n* = 37)	<0.001	4.069	<0.001	5.125	0.461	0.993	0.517	0.884
Diprotodontia (*n* = 23)	<0.001	5.748	0.004	6.842	0.313	1.169	0.192	1.483
Placentalia (*n* = 79)	<0.001	5.927	<0.001	3.503	0.003	2.882	0.001	3.094

Although most of the various marsupial subclades each show similar patterns with regard to forelimb and hind limb functional diversity (Tables 1 and 2), diprotodontian disparity results do not show the same trend. For diprotodontans, all traits and hind limb traits alone show high values of disparity while forelimb traits alone show low disparity levels; this supports predictions of the DCH (Fig. 4). While ‘Ameridelphia’ and Didelphimorphia alone show pMANOVA results with a lower *p*-value for hind limb traits than for forelimb traits, and Australidelphia pMANOVA results show a lower *p*-value for forelimb traits than for hind limb traits, all three of these differences are within a small margin of 0.05 (Table 2). Additionally, the pFDA results for Australidelphia alone show that hind limb characters are better in predicting locomotor mode than forelimb characters, correctly predicting the locomotor mode of all taxa, compared to misclassifying 5.41% of taxa using forelimb characters (Table 1). Disparity values for Didelphimorphia were also consistently low, for all three trait categories (Fig. 4). When Diprotodontia alone was examined, the pMANOVA tests showed a lower *p*-value for both hind limb traits than for forelimb traits, by a margin of ~0.12 (Table 2), relatively wider than the margin seen for ‘Ameridelphia’, Didelphimorphia, and Australidelphia. These results are consistent with those of the pFDAs, with forelimb and hind limb traits both correctly predicting the locomotor modes of all taxa (Table 1).

The results for ‘Ameridelphia’ and Australidelphia show mostly similar patterns to those of their major subclades, Didelphimorphia and Diprotodontia, respectively. The morphological disparity results for ‘Ameridelphia’ track closely with those of Didelphimorphia, maintaining low levels across all three trait categories (Fig. 4). By contrast, pMANOVA results for Australidelphia show a lower value for forelimb traits than for hind limb traits, but both traits are far higher than those of ‘Ameridelphia’, pushing them further from statistical significance (Table 2). The pFDA results for Australidelphia are very similar to those of ‘Ameridelphia,’ with all traits and hind limb traits alone correctly predicting the locomotor mode of all taxa, while forelimb traits misclassify a small number of taxa (Table 1). While the percentage of taxa misclassified using forelimb traits increased when phylogeny was incorporated, the number of taxa involved in this increase is small enough to be considered negligible (Table 1). Overall disparity patterns for australidelphians and diprotodontians are similar, with hind limb disparity levels being much greater than those of forelimbs in both groups. However, diprotodontians have relatively low forelimb disparity (less than that of all marsupials) whereas australidelphians have relatively high forelimb disparity (greater than that of all marsupials) (Fig. 4).

One aspect of our pMANOVA results that should be noted is that the magnitude of our *p*-values is likely being affected by the smaller sample size of our marsupial dataset (and to a greater extent the marsupial subclades) compared to our placental dataset. While this does not influence the relative forelimb versus hind limb patterns in our results for this analysis, this effect could help to explain the large rift in magnitude between placentals and marsupials and their subclades (Table 2). Further, the incorporation of phylogeny has an especially large influence on the marsupial results, as indicated by the MANOVA results that show considerably greater statistical significance than the pMANOVA results (Table 2). This is likely due in large part to the limited number of independent evolutions of locomotor modes in marsupials, especially for saltators (represented only by macropods) and gliders (represented by two closely related clades) (Fig. 2).

In analyses in which we used varying locomotor classification schemes, the results of individual clades are consistent with our main results. This includes our results for analyses comparing just tree-dwelling and ground-dwelling taxa (i.e., two locomotor groups), analyses that separate scansorial taxa from arboreal taxa as a separate locomotor mode, and analyses that further separate semi-fossorial/fossorial from the terrestrial taxa (Table 3). For the pMANOVA Marsupialia tree-dwelling versus ground-dwelling taxa comparison, we find for forelimbs a *p*-value of 0.059 (*F*-stat 5.827) and for hind limbs a *p*-value of 0.058 (*F*-stat 5.671). For the pMANOVA Marsupialia comparison that includes semi-fossorial/fossorial, we find for forelimbs a *p*-value of 0.089 (*F*-stat 4.716) and for hind limbs a *p*-value of 0.075 (*F*-stat 4.622). The similarity in results among these various analyses suggests that our choice of locomotor categorization scheme does not significantly influence our conclusions.

Table 3. Statistical significance (*p*-value) results for pMANOVAs and the percentages of correctly classified species of phylogenetic flexible discriminant analyses (pFDAs) with arboreal and scansorial separated into two locomotor mode categories. 'All' analyses include all forelimb and hind limb traits.

Sample	**pMANOVA**		**FDA (λ = 0)**			**pFDA (λ = 0.1)**		
	Forelimb	**Hind limb**	**All**	**Forelimb**	**Hind limb**	**All**	**Forelimb**	**Hind limb**
Marsupialia	0.205	0.262	98.41%	92.06%	93.65%	98.41%	92.06%	90.48%
'Ameridelphia'	0.178	0.281	100%	92.31%	100%	94.62%	92.31%	100%
Didelphimorphia	0.172	0.370	100%	95.66%	100%	100%	95.66%	100%
Australidelphia	0.444	0.621	100%	100%	97.30%	70.3%	97.30%	97.30%
Diprotodontia	0.312	0.180	100%	100%	100%	100%	100%	100%
Placentalia	0.004	0.001	92.41%	84.81%	81.01%	91.14%	87.34%	77.21%

DISCUSSION

Lack of consistent evidence for developmental constraints on marsupial forelimbs

A prediction of the Developmental Constraint Hypothesis (DCH) is that marsupial hind limbs should show a greater degree of ecomorphological differentiation than forelimbs due to the lack of developmental constraints, and therefore should outperform forelimbs as predictors of locomotor mode. Additionally, the DCH predicts that this discrepancy between forelimbs and hind limbs should be less pronounced in placentals (Fig. 1). However, across our three sets of analyses, we consistently fail to find strong support for such predictions of the DCH. This is most apparent in marsupial versus placental comparisons (Tables 1 and 2, Fig. 4). Our disparity results suggest that the forelimbs of marsupials and placentals are not as disparate as we would expect under the DCH. In contrast to the initial prediction based on DCH, marsupial forelimb traits show greater predictive ability via pFDAs than placental forelimb traits with regards to locomotor mode (Table 1). It is noteworthy that our analyses show similar levels of disparity for all traits between marsupials and placentals in the small body size range (<40 kg), in contrast to the prediction of the DCH that marsupials should show far lower disparity levels than placentals. In particular, for forelimb traits, where the DCH predicts that placentals have relatively high disparity levels compared to marsupials, our results show that disparity levels between placentals and marsupials are very similar (Fig. 4, Table S2).

Our morphological disparity results are directly comparable to those of Sears (2004), Cooper & Steppan (2010), and Martín-Serra & Benson (2020). Sears (2004) found that the disparity of marsupial scapulae was far lower than that of placental scapulae, and Cooper & Steppan (2010) found that the disparities of marsupial forelimbs and hind limbs were far lower than those of placental forelimbs and hind limbs, and that marsupial forelimbs showed lower disparity than their hind limbs (Table 4). While our disparity results are congruent with those of

Sears (2004) and Cooper & Steppan (2010) in regards to marsupial forelimbs showing lower disparity than marsupial hind limbs, our placental disparity results are not significantly different from our marsupial placental disparity results, and this is true when examining all traits, forelimb traits, and hind limb traits (Fig. 4). One of the primary findings of Martín-Serra & Benson (2020) was that forelimbs and hind limbs in marsupials showed similar levels of disparity, contrary to the predictions of the DCH (Table 4). In contrast to their findings, our disparity results suggest the opposite, that forelimbs consistently show lower disparity levels than hind limbs within marsupials. Additionally, while Martín-Serra & Benson (2020) did not consider direct placental-marsupial comparisons in their disparity tests (Table 4); their analyses of marsupial subclades show that diprotodontians have overall higher levels of disparity than Didelphimorphia and Dasyuromorphia, and that this difference is more pronounced in the forelimb elements than in the hind limb elements. While our disparity results also demonstrate Diprotodontia to have higher disparity levels than Didelphimorphia, we find that the relative disparity levels between the two groups are much more similar in forelimbs than in hind limbs.

Table 4. Results of relevant previous studies on the DCH broken down by methodology and metrics used.

Study	**Groups examined**	**Morphological marker**	**Elements examined**	**Type of comparison**	**Disparity result**	**Overall interpretation of DCH**
Sears, 2004	Marsupials & placentals (no subclades)	Linear measurements	Girdles	PCA & disparity	Marsupial disparity << placental disparity	Favored DCH

Cooper & Steppan, 2010	Marsupials & placentals (no subclades)	Linear measurements	Stylopod, zeugopod and autopod	PCA & disparity	Marsupial forelimbs more restricted than those of placentals	Favored DCH
Kelly & Sears, 2011a	Marsupials & placentals (no subclades)	Linear measurements	Girdles, stylopod, and zeugopod	Patterns of correlation, PCA, & disparity	Marsupial disparity << placental disparity	Favored DCH
Martín-Serra & Benson, 2020	Marsupial subclades	3D geometric morphometrics	Girdles, stylopod, and zeugopod	PCA, disparity, evolutionary rates, & integration	Marsupial forelimb disparity = marsupial hind limb disparity	Did not favor DCH
This study	Marsupials & placentals (including subclades)	Linear measurements	Girdles, stylopod, zeugopod, and autopod	MANOVA & pMANOVA, FDA & pFDA, disparity	Marsupial overall disparity = placental overall disparity; marsupial forelimb disparity < marsupial hind limb disparity	Favor DCH only for Diprotodontia

Further, we found far less consistent support of the DCH than Kelly & Sears (2011a). While Kelly & Sears (2011a) did not examine comparisons between marsupial subclades (Table 4), their results comparing marsupial forelimbs and hind limbs to those of placentals are directly comparable to our results. In comparisons of marsupial forelimbs to those of placentals, our disparity results are consistent with those of Kelly & Sears (2011a), which suggested that placentals include lineages with more derived morphologies compared to marsupials, and indicated greater morphological specialization in placental forelimbs over marsupial forelimbs (Table 4). This trend is also present in our forelimb disparity results (Fig. 4).

However, the results of Kelly & Sears (2011a) also suggested that placental forelimbs are morphologically more derived than marsupial hind limbs, indicating greater specialization in placental forelimbs than marsupial hind limbs - this is not replicated in our disparity results, which show marsupial hind limbs as having far greater disparity than placental forelimbs (Fig. 4). Further, the aforementioned patterns observed in Kelly & Sears (2011a) do not appear in our pFDA results that show the predictive power of marsupial forelimbs and hindlimbs outweighing that of placentals (Table 1).

In contrast to some previous studies that tested the DCH by examining limb and girdle morphologies (Sears, 2004; Kelly & Sears, 2011a; Kelly & Sears, 2011b; Martín-Serra & Benson, 2020), we have included autopod metrics in our analysis. By incorporating the autopod into analyses, our dataset includes a number of measurements that are functionally informative of locomotor mode, including phalangeal, metacarpal, and metatarsal length and width (Weisbecker & Warton, 2006; Kirk et al., 2008; Chen & Wilson, 2015; Nations et al., 2019). These additional autopod measurements may offer an explanation as to why our results differ from those of some previous studies (Sears, 2004; Kelly & Sears, 2011a; Kelly & Sears, 2011b). To test this, we repeated analyses in which we removed the autopod traits. We found that analyses without autopod data yielded more similar results to those of previous analyses that support the DCH (Tables S3–S4). In particular, the difference in forelimb disparity between marsupials and placentals is far more prominent when the autopod is removed, with disparity in placentals being greater than that of marsupials. Our pFDA results show greater support for the DCH across all marsupial subclades, because the percentage of correctly classified taxa based on forelimb traits alone decreased while the percentage based on hind limb traits remained the same relative to results that included the autopod (Tables S3–S4). This greater support for the DCH is more in line with the results from previous work (Sears, 2004; Kelly & Sears 2011a; Kelly & Sears, 2011b). The difference in results that comes from the inclusion of the autopod may have a number of explanations. For example, it is possible that developmental constraints have a more significant impact on the girdle, stylopod, and zeugopod than on the autopod. Another possibility is that increased digit functional diversity in marsupials is compensating for constraints placed on the rest of the forelimb. Both explanations warrant further research, and could be critical in understanding the impact developmental constraints have on the marsupial autopod compared to the rest of the forelimb.

One possible explanation for the greater-than-expected morphological disparity in marsupial forelimbs is that forelimbs are often involved in additional functions besides locomotion, such as food manipulation or social behavior (Fischer, 1998); selective pressures associated with these other functions could help to generate increased morphological disparity in marsupials. If this is the case, then the association between morphology and locomotor mode might be weakened because of the additional influences on morphology, decreasing the power of marsupial forelimbs to predict locomotor function. However, our discriminant analyses contradict this expectation because they show that the predictive power of marsupial forelimbs is greater than that of placental forelimbs (Table 3). Thus, forelimb morphologies of marsupials are heavily influenced by locomotor mode even if they are also involved in additional non-locomotory functions.

Support for the DCH is strongest among marsupial subclade comparisons

Despite the lack of consistent support for the DCH, some analyses do support the DCH. For instance, the pMANOVA results indicate that placental forelimb morphologies (but not marsupial forelimbs) show significant differentiation among locomotor modes (Table 2). Further, the predictions of the DCH are slightly better supported by marsupial-only analyses; pMANOVAs, pFDAs, and morphological disparity values all indicate that marsupial hind limbs show greater ecomorphological differentiation than marsupial forelimbs, but this is supported to different extents by each analysis (Tables 1 and 2, Figs. 3 and 4). For instance, the pMANOVA results for the marsupial subclades, with the exception of Diprotodontia alone, are roughly equivalent for forelimb and hind limb traits (Table 2); however, the pMANOVA results for Diprotodontia show a greater relative discrepancy between the *p*-values for forelimb traits and hind limb traits, indicating a lower *p*-value for the latter, in line with the DCH predictions for marsupials (Fig. 1). The pMANOVA results for placentals are statistically significant for both forelimb and hind limb traits, suggesting that the relative levels of ecomorphological diversification are about the same for each set of limbs (Table 2). This is in line with predictions of the DCH because the lack of any developmental constraint in placentals means that we might expect similar ecomorphological differentiation in forelimbs and hind limbs (Fig. 1). Morphological disparity results are consistent with this pattern; marsupials as a whole have low disparity in their forelimbs and high disparity in the hind limbs, while placentals have similar levels of disparity in both limbs (Fig. 4). Thus, the DCH is supported by some of our within-marsupial and within-placental results.

The members of Diprotodontia make the most rigorous crawls of all marsupial taxa (Sears, 2004) due to the longer relative distance the newborn is required to navigate between the urogenital opening and the teat, so the effects of developmental constraints are expected to be strongest in this clade. This prediction is supported by the forelimb versus hind limb disparity results (Fig. 4), which show that diprotodontian disparity is considerably greater in hind limbs than forelimbs, as well as by the relative pMANOVA p-values for diprotodontian forelimbs and hind limbs, with hind limbs showing a lower value than hind limbs separated by a relatively wide margin (Table 2). This is consistent with results of Sears (2004) who found that ontogenetic shape change in diprotodontian scapulae correlated with extensive crawls to the teat, indicating stronger constraints on diprotodontians. However, our disparity results are incongruent with those from Martín-Serra & Benson (2020), which showed similar levels of disparity between forelimbs and hind limbs in diprotodontians, while our analysis shows higher disparity for the hind limb than forelimb (Fig. 4). In contrast to our disparity and pMANOVA results, our pFDA results do not support the expectation for greater forelimb constraints on Diprotodontia because results are similar for forelimbs and hind limbs (Table 1). The similarities in values for forelimbs and hind limbs suggest that in diprotodontians each set of limbs is equally good at predicting locomotor mode. Because this differs from our disparity and pMANOVA results, we can therefore say that the evidence favors the DCH in Diprotodontia for most of our analyses.

It is worth noting that all known saltatorial and gliding marsupials are within Diprotodontia, which heightens the morphological variety among the limbs of this order. This may have contributed to the greater difference in disparity between forelimbs and hind limbs observed relative to other marsupial orders (Fig. 4). Thus, much of the morphological limb diversity in marsupials is contained within this single order. In combination with the particular morphological adaptations that go with specialized locomotor modes such as saltatorial and gliding, this may help to explain why Diprotodontia displays the highest relative support for the DCH among marsupials. While it could be argued that such increased limb morphology specialization could lead to more forelimb disparity and thus less support for the DCH, a great deal of the locomotory specializations involved in these locomotor modes occurs in the hind limbs, and not the forelimbs, a difference that can greatly accentuate the difference in limb disparity between the forelimbs and the hind limbs. The lack of such locomotory specializations in other marsupial orders may have contributed to the lower levels of difference in disparity between forelimbs and hind limbs therein.

Although Diprotodontia shows support for the DCH, the more phylogenetically inclusive Australidelphia [Microbiotheria + Notoryctemorphia + Peramelemorphia + Dasyuromorphia + Diprotodontia] demonstrates less support for the DCH. In our pMANOVA results, Australidelphia shows the *p*-value for forelimb traits is lower than that of hind limb traits (Table 2). Similarly, our pFDA results show a fairly small difference in predictive ability between forelimb and hind limb traits when phylogeny is incorporated, suggesting that the predictive ability of hind limbs is roughly equivalent to that of forelimbs (Table 1). However, our disparity results show some support for the DCH in Australidelphia, where forelimb traits show low disparity, while hind limb traits show high disparity. Here, we only see support for the predictions of the DCH in one of our three analyses. Taken together, we can note that the differences between the Australidelphia results and the Diprotodontia results can inform us as to the effects of the other clades within Australidelphia besides Diprotodontia – namely, Dasyuromorphia, Peramelemorphia, Notoryctemorphia, and Microbiotheria (Fig. 2). From the relatively higher support for the DCH in Diprotodontia compared to Australidelphia, we can infer that these other australidelphian clades may be less strongly influenced by developmental constraints, which is consistent with the predictions of the DCH as Diprotodontians make the most rigorous crawls of the clades involved. The sample sizes of these groups are small (e.g., Notoryctemorphia and Microbiotheria each contain a single extant genus in our dataset), making it challenging to examine these groups in greater detail. The weaker support for the DCH in Australidelphia, and in particular Notoryctemorphia, lends further credence to the suggestion of Kirsch (1977a) that the highly derived hands of *Notoryctes* are not conducive to a rigorous crawl from the urogenital opening to teat, and that marsupial forelimb morphology was not restricted by these crawls. Additionally, Kirsch noted that bandicoots, the members of Peramelemorphia, do not exert the manus in reaching the teat. The weak support for the DCH observed from our study is consistent with Kirsch's expectations.

In contrast to Diprotodontia (and Australidelphia), Didelphimorphia forelimbs and hind limbs have very similar morphological disparity levels (Fig. 4), suggesting that their forelimb diversity is not constrained relative to hind limbs. This could be interpreted as being consistent with the DCH, because neonatal crawls to the pouch in marsupials of Didelphimorphia are not as arduous as in other marsupial clades under some lower levels of developmental constraints on didelphimorphians (Sears, 2004). Additionally, when we directly compare our 'Ameridelphia' results to our Didelphimorphia results, we can infer how developmental constraints might affect Paucituberculata, which is another 'Ameridelphia' group. We find that the difference between the results is small. This is expected given that Paucituberculata contains only three living genera, each of which is represented in our dataset, so our small sample size should correspond to a small difference. However, our pMANOVA perform worse when Paucituberculata is included, suggesting that developmental constraints may have a stronger influence on Paucituberculata than on Didelphimorphia.

Multiple factors influence the functional diversity of marsupial limbs

We summarize our conclusions in Figure 5, which shows how the predictions of the DCH perform when comparing placentals to marsupials, and to the major marsupial subclades. The majority of comparisons display either mixed or failed support for the predictions of the DCH, except for Diprotodontia. Additionally, our results suggest a variation in the strength of the DCH between marsupial subclades. According to Sears (2004), the observed trends of marsupial-placental disparity are not something we would expect if the DCH were the only major constraint on forelimb functional morphology in marsupials, because marsupial forelimb morphology should be more constrained than placental forelimb morphology, while hind limb morphology should be roughly the same in both groups. However, we observe forelimb disparity to be similar in both groups, while marsupial hind limb disparity is much greater than placental hind limb disparity. This indicates to us that the DCH is not the only factor limiting the range of forelimb morphologies, and DCH varies in strength among marsupial clades.

We do not dispute the established ontogenetic presence of a developmental constraint in the forelimbs in marsupials through the accelerated development of forelimbs, which necessarily are used for the climb to the maternal teat (Keyte and Smith, 2010). Instead, we suggest that in addition to the DCH, alternate hypotheses should be given greater consideration for their potential to explain the discrepancies between the ecological niches occupied by placentals and marsupials. For example, it is possible that the historical geographic distribution of the Marsupialia versus the Placentalia had a greater influence over the ecological niches that each group had an opportunity to exploit (Sánchez-Villagra, 2013). Placentals have historically inhabited a greater proportion of the Earth’s landmass, and so were likely exposed to a greater range of ecological niches than the marsupials. Alternatively, it is possible that ecological pressures may have overcome constraints like the DCH to a far greater degree than has previously been realized, as proposed by Martín-Serra & Benson (2020). If this was the case, then the ontogenetic developmental constraints that have been observed in marsupials may be the result, not the cause, of their limited ecology. Influenced by the traditional assumption of the primacy of the DCH, tests of these alternate hypotheses have yet to be fully explored (e.g., Sánchez-Villagra, 2013). Although we appreciate the quantification of morphological diversity and comparisons among groups, our and other recent studies demonstrate the lack of universal patterns within major clades and the difficulty of extrapolating from extant patterns of disparity in adults to developmental mechanisms that could have influenced evolutionary trends across deep time. Our investigation reveals that tests of alternate hypotheses are indeed important and necessary in order to elucidate the true nature of the relationship between developmental constraints and ecological diversity in mammalian functional morphology.

CONCLUSIONS

Our results indicate that while the Developmental Constraint Hypothesis may have played a role in preventing marsupials from evolving such specialized locomotor modes as power-flying and fully aquatic swimming, its effect on constraining the functional diversity among other locomotor modes may have been far more limited. While the predictions of the DCH are somewhat supported by within-marsupial and within-placental comparisons of forelimbs and hind limbs (Tables 1-2, Figs. 3-4), they are not as well supported by comparisons between placentals and marsupials, a result that seems to be particularly driven by the inclusion of autopod traits in our analyses (Tables S3-S4). Additionally, the strength of the within-group comparison of the DCH varies with marsupial subclades; this demonstrates phylogenetic differences in the strength of the DCH. Overall, our analyses fail to consistently support the DCH, with results for Diprotodontia being an exception. We believe that further testing of alternative hypotheses regarding the limitations of occupied ecomorphotypes in marsupials would be illuminating. Additionally, this study only covers limb morphology; however, there are other aspects of biology that still warrant investigation in order to more fully test the DCH. While there is indeed an ontogenetic element of accelerated forelimb development in marsupials, we believe our results show that it is unlikely to be the sole reason why members of Marsupialia have not evolved developed wings, flippers, or other extremely specialized forelimbs. Developmental constraints in marsupials may have been counter-balanced, or even overpowered by ecological pressures in less specialized locomotor modes.

ACKNOWLEDGEMENTS

For access to museum specimen collections, we would like to thank Lawrence Heaney, Bruce Patterson, and Adam Ferguson of Division of Mammals of the Field Museum of Natural History; Eric Sargis and Kristof Zyskowski from the Yale Peabody Museum of Natural History; Marisa Surovy, Sara Ketelsen, Nancy B. Simmons, and Eleanor Hoeger from the American Museum of Natural History; Helen F. James and Darrin P. Lunde from the Smithsonian National Museum of Natural History; and Sharlene E. Santana and Jeff Bradley from the Burke Museum of Natural History and Culture. We deeply appreciate the assistance and suggestions for improvement for this work from Kenneth Angielczyk, Michael Coates, Christine Janis, Michael Granatosky, Kathryn Stanchak, Chris Law, Rachel Roston, Lucas Weaver, William Brightly, William Richardson, and Rhys Charles. Finally, we would like to thank Charles Pevsner for his assistance in the creation of some of the figures in this work. This research has been supported by the Metcalf Undergraduate Fellowship from the University of Chicago, awarded for an internship in China by the University of Chicago Center in Beijing, and by the Ecology and Evolution Undergraduate Research Fellowship from the Biological Science Collegiate Division, with oversight by Catherine Pfister.

REFERENCES

Adams DC, Otárola-Castillo E. 2013. geomorph: an r package for the collection and analysis of geometric morphometric shape data. *Methods in Ecology and Evolution* **4**: 393–399.

Angielczyk KD, Schmitz L. 2014. Nocturnality in synapsids predates the origin of mammals by over 100 million years. *Proceedings of the Royal Society B: Biological Sciences* **281**: 20141642.

Argot C. 2001. Functional-adaptive anatomy of the forelimb in the Didelphidae, and the paleobiology of the Paleocene marsupials *Mayulestes ferox* and *Pucadelphys andinus*. *Journal of Morphology* **247**: 51–79.

Argot C. 2002. Functional-adaptive analysis of the hind limb anatomy of extant marsupials and the paleobiology of the Paleocene marsupials *Mayulestes ferox* and *Pucadelphys andinus*. *Journal of Morphology* **253**: 76–108.

Bennett VC, Goswami A. 2011. Does developmental strategy drive limb integration in marsupials and monotremes? *Mammalian Biology* **76**: 79–83.

Chen M, Wilson GP. 2015. A multivariate approach to infer locomotor modes in Mesozoic mammals. *Paleobiology* **41**: 280-312.

Collyer ML, Adams DC. 2018. RRPP: An R package for fitting linear models to high-dimensional data using residual randomization. *Methods in Ecology and Evolution* **9**: 1772–1779.

Conith AJ, Meagher MA, Dumont ER. 2021. The influence of divergent reproductive strategies in shaping modularity and morphological evolution in mammalian jaws. *Journal of Evolutionary Biology* **00**: 1-16.

Cooper WJ, Steppan SJ. 2010. Developmental constrain on the evolution of marsupial forelimb morphology. *Australian Journal of Zoology* **58**: 1-15.

Drummond AJ, Suchard MA, Xie D, Rambaut A. 2012. Bayesian Phylogenetics with BEAUti and the BEAST 1.7. *Molecular Biology and Evolution* **29**: 1969–1973.

Fabre A-C, Dowling C, Miguez RP, Fernandez V, Noirault E, Goswami A. 2021. Functional constraints during development limit jaw shape evolution in marsupials. *Proceedings of the Royal Society B: Biological Sciences* **288**: 20210319.

Feilich KL, López-Fernández H. 2019. When does form reflect function? Acknowledging and supporting ecomorphological assumptions. *Integrative and Comparative Biology* **59**: 358-370.

Fischer MS. 1998. Die Lokomotion von *Procavia capensis* (Mammalia: Hyracoidea): Zur Evolution des Bewegungssystems bei Säugetieren. *Abhandlungen des Naturwissenschaftlichen Verein Hamburg (NF)* **33**: 1-188.

Garland K, Marcy A, Sherratt E, Weisbecker V. 2017. Out on a limb: bandicoot limb co-variation suggests complex impacts of development and adaptation on marsupial forelimb evolution. *Evolution & Development* **19**: 69-84.

Gemmell RT, Veitch C, Nelson J. 2002. Birth in marsupials. *Comparative Biochemistry and Physiology Part B: Biochemistry and Molecular Biology* **131**: 621-630.

Goswami A, Smaers JB, Soligo C, Polly PD. 2014. The macroevolutionary consequences of phenotypic integration: from development to deep time. *Philosophical Transactions of the Royal Society B: Biological Sciences* **369**: 20130254.

Grossnickle DM. 2020. Feeding ecology has a stronger evolutionary influence on functional morphology than on body mass in mammals. *Evolution* **74**: 610-628.

Grossnickle, DM, Chen M., Wauer JGA, Pevsner SK, Weaver LN, Meng Q-J, Liu D, Zhang Y-G, Luo Z-X. 2020. Incomplete convergence of gliding-mammal skeletons. *Evolution* **74**: 2662-2680.

Hildebrand, M, Goslow Jr GE. 2001. in Analysis of Vertebrate Structure (5th Edition), John Wiley, New York, pp. 1-635.

Jones KE, et al. 2009. PanTHERIA: a species-level database of life history, ecology, and geography of extant and recently extinct mammals. *Ecology* **90**: 2648–2648.

Kelly, EM, Sears KE. 2011a. Limb specialization in living marsupial and eutherian mammals: constraints on mammalian limb evolution. *Journal of Mammalogy* **92**: 1038–1049.

Kelly EM, Sears KE. 2011b. Reduced phenotypic covariation in marsupial limbs and the implications for mammalian evolution. *Biological Journal of the Linnean Society* **102**: 22-36.

Kelly EM, Marcot JD, Selwood L, Sears KE. 2019. The development of integration in marsupial and placental limbs. *Integrative Organismal Biology* **1**: 1-11.

Keyte AL, Smith KK. 2010. Developmental origins of precocial forelimbs in marsupial neonates. *Development* **137**: 4283-4294.

Kilbourne BM. 2017. Selective regimes and functional anatomy in the mustelid forelimb: Diversification toward specializations for climbing, digging, and swimming. *Ecology and Evolution* **7**: 8852-8863.

Kirk EC, Lemelin P, Hamrick MW, Boyer DM, Bloch JI. 2008. Intrinsic hand proportions of euarchontans and other mammals: implications for the locomotor behavior of plesiadapiforms. *Journal of Human Evolution* **55**: 278–299.

Kirsch JAW. 1977a. Biological aspects of the marsupial-placental dichotomy: A reply to Lillegraven. *Evolution* **31**: 898-200.

Kirsch JAW. 1977b. The six-percent solution: Second thoughts on the adaptedness of the Marsupialia: Features of their physiology and diversity suggest that marsupials represent an alternative but not inferior kind of mammal, valuable in understanding the course of mammalian evolution. *American Scientist* **65**: 276-288.

Lillegraven JA. 1975. Biological considerations of the marsupial-placental dichotomy. *Evolution* **29**: 707-722.

Luo Z-X, Ji Q, Wible JR, Yuan C-X. 2003. An Early Cretaceous tribosphenic mammal and metatherian evolution. *Science* **302**: 1934-1940.

MacDonald DW. 2002. in The Encyclopedia of Mammals (2nd Revised Edition), Oxford University Press, Oxford, UK.

Maga AM, Beck RMD. 2017. Skeleton of an unusual, cat-sized marsupial relative (Metatheria: Marsupialiformes) from the middle Eocene (Lutetian: 44-43 million years ago) of Turkey. *PLoS ONE* **12**: e0181712.

Martín-Serra A, Benson RBJ. 2020. Developmental constraints do not influence long-term phenotypic evolution of marsupial forelimbs as revealed by interspecific disparity and integration patterns. *The American Naturalist* **195**: 547-560.

Meng Q-J, Grossnickle DM, Liu D, Zhang Y-G, Neander AI, Ji Q, Luo Z-X. 2017. New gliding mammaliaforms from the Jurassic. *Nature* **548**: 291-296.

Motani R, Schmitz L. 2011. Phylogenetic Versus Functional Signals in the Evolution of Form–Function Relationships in Terrestrial Vision. *Evolution* **65**: 2245–2257.

Nations JA, Heaney LR, Demos TC, Achmadi AS, Rowe KC, Esselstyn JA. 2019. A simple skeletal measurement effectively predicts climbing behavior in a diverse clade of small mammals. *Biological Journal of the Linnean Society* **128**: 323-336.

Nowak RM. 1999. Walker's Mammals of the World (Vol. 1-2, 6th Edition), Johns Hopkins University Press, Baltimore.

Nowak RM. 2005. Walker's Marsupials of the World, Johns Hopkins University Press, Baltimore, pp. 1-266.

Polly, P. D. 2007. Limbs in mammalian evolution. Chapter 15, pp. 245-268, in Fins into Limbs: Evolution, Development, and Transformation, Brian K Hall (ed.). University of Chicago Press: Chicago. Pp. 1-772.

R Core Team 2020. R: A language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing.

Salton JA, Sargis EJ. 2009. Evolutionary morphology of the Tenrecoidea (Mammalia) hind limb skeleton. *Journal of Morphology* **270**: 367–387.

Samuels JX, Van Valkenburgh B. 2008. Skeletal indicators of locomotor adaptations in living and extinct rodents. *Journal of Morphology* **269**: 1387-1411.

Samuels JX, Meachen JA, Sakai SA. 2013. Postcranial morphology and the locomotor habits of living and extinct carnivorans. *Journal of Morphology* **274**: 121-146.

Sánchez-Villagra MR. 2013. Why are there fewer marsupials than placentals? On the relevance of geography and physiology to evolutionary patterns of mammalian diversity and disparity. *Journal of Mammalian Evolution* **20**: 279-290.

Schindelin J, Arganda-Carreras I, Frise E, Kaynig V, Longair M, Pietzsch T, Preibisch S, Rueden C, Saalfeld S, Schmid B, Tinevez J-Y, White DJ, Hartenstein V, Eliceiri K, Tomancak P, Cardona, A. 2012. Fiji: an open-source platform for biological-image analysis. *Nature Methods* **9**: 676-682.

Schmidt M, Fischer MS. 2009. Morphological Integration in Mammalian Limb Proportions: Dissociation Between Function and Development. *Evolution* **63**: 749–766.

Schmitz L, Motani R. 2011. Nocturnality in Dinosaurs Inferred from Scleral Ring and Orbit Morphology. *Science* **332**: 705–708.

Sears KE. 2004. Constraints on the morphological evolution of marsupial shoulder girdles. *Evolution* **58**: 2353-2370.

Upham NS, Esselstyn JA, Jetz W. 2019. Inferring the mammal tree: Species-level sets of phylogenies for questions in ecology, evolution, and conservation. *PLoS Biology* **17**: e3000494.

Weaver LN, Grossnickle DM. 2020. Functional diversity of small-mammal postcrania is linked to both substrate preference and body size. *Current Zoology* **66**: 539-553.

Weisbecker V, Goswami A, Wroe S, Sanchez-Villagra MR. 2008. Ossification heterochrony in the therian postcranial skeleton and the marsupial–placental dichotomy. *Evolution* **62**: 2027–2041.

Weisbecker V, Warton DI. 2006. Evidence at hand: diversity, functional implications and locomotor prediction in intrinsic hand proportions of diprotodontian marsupials. *Journal of Morphology* **267**: 1469–1485.

Young NM, Hallgrímsson B. 2005. Serial Homology and the Evolution of Mammalian Limb Covariation Structure. *Evolution* **59**: 2691–2704.

FIGURE CAPTIONS

Figure 1. Functional diversity patterns in skeletal limb morphologies as predicted by the Developmental Constraint Hypothesis (DCH). The DCH predicts that morphological disparity and functional diversity are constrained by precocial forelimb development in marsupials due to the necessity for the neonatal crawl. This can be tested by morphometric comparison between forelimbs and hind limbs, between placentals and marsupials, and among marsupial groups with varying intensity of crawls. The DCH predicts that marsupial forelimbs have lower functional diversity levels than that of placental forelimbs. Ameridelphian marsupials have a less rigorous neonatal crawl than those of Australidelphia and thus are predicted to have relatively greater forelimb disparity than australidelphians. Silhouette images are from www.phylopic.org. Marsupialia, 'Ameridelphia', and Australidelphia images by Sarah Werning (CC BY 3.0; http://creativecommons.org/licenses/by/3.0/). All other images are not copyrighted.

Figure 2. Phylogeny of the taxa used in the analyses of this study, pruned from a maximum clade credibility tree based on a sample of trees from Upham et al. (2019). Silhouettes are from www.phylopic.org. The Paucituberculata, Microbiotheria, Dasyuromorphia, Notoryctemorphia and Didelphimorphia images are by Sarah Werning (CC BY 3.0; http://creativecommons.org/licenses/by/3.0/). All other silhouette images are not copyrighted.

Figure 3. The flexible discriminant analysis (FDA) morphospace plots of the first two discriminant axes for marsupials and placentals, which are classified into four locomotor modes. These analyses demonstrate that the terrestrial, arboreal, gliding and saltatorial locomotor modes can be sufficiently

distinguished by our limb skeletal measurements for the small-to-medium sized marsupials and placentals (<40 kg) in our sample. The different colors represent different locomotor modes, using the same color legend as in Figure 2. Percentages represent the amount of between-group variance explained by each axis. *Petaurus*, *Macropus*, and *Sarcophilus* silhouettes are by Sarah Werning, Glaucomys silhouette is by Chloé Schmidt (CC BY 3.0; http://creativecommons.org/licenses/by/3.0/). All other silhouettes are not copyrighted.

Figure 4. Comparison of morphological disparity (variance) in forelimb, hind limb and all limbs between placentals and marsupials, and among different marsupial groups. The forelimb-hind limb discrepancy in trait variance in 'Ameridelphia' and in Didelphimorpha is closest to the pattern of placentals (i.e., similar disparity levels for forelimbs and hind limbs). The forelimb disparity (variance) is most comparable between diprotodontian marsupials and placentals. These two observations on disparity pattern are inconsistent with DCH expectations. But forelimb-hind limb discrepancy in Australidelphia and Diprotodontia clades are consistent with DCH expectation.

Figure 5. Summary of test results on the predictions of the Developmental Constraint Hypothesis (DCH), at different phylogenetic levels. The DCH is somewhat supported by forelimb-hind limb comparison within placentals, within marsupials, and within subclades of marsupials, but we find little support for the DCH in comparisons of placentals either to marsupials as a whole, or to subclades of marsupials. The size of the red and blue circles represents expected functional diversity, as outlined in Figure 1. Check marks indicate greater support for the DCH, 'X's indicate a failure to support the DCH, and question marks indicate mixed support for the DCH. The arrows indicate the groups being compared that correspond to each mark, with red arrows being between-group comparisons, while purple arrows

are within-group comparisons. Marsupialia, ‘Ameridelphia’, and Australidelphia images are by Sarah Werning (CC BY 3.0; http://creativecommons.org/licenses/by/3.0/). All other images not copyrighted.