Individual variation of the masticatory system dominates 3D skull shape in the herbivory-adapted marsupial wombats

Vera Weisbecker[1]*, Thomas Guillerme[1]*, Cruise Speck[1], Emma Sherratt[2], Hyab Mehari Abraha[3],

Alana C. Sharp[4,5], Claire E. Terhune[6], Simon Collins[7†], Steve Johnston[7], Olga Panagiotopoulou[3]*

Affiliations:

[1] School of Biological Sciences, The University of Queensland, Australia

[2] School of Biological Sciences, The University of Adelaide, Australia

[3] Monash Biomedicine Discovery Institute, Department of Anatomy & Developmental Biology, Monash University, Australia

[4] Department of Cell & Developmental Biology, University College London, UK

[5]Institute of Ageing and Chronic Disease, University of Liverpool, UK

[6] Department of Anthropology, University of Arkansas, USA

[7] School of Agricultural and Food Sciences, The University of Queensland, Australia.

[†] Deceased

*Correspondence to:

Dr. Vera Weisbecker, email: v.weisbecker@uq.edu.au

Dr. Thomas Guillerme, email: guillert@tcd.ie

Dr. Olga Panagiotopoulou, email: olga.panagiotopoulou@monash.edu

Abstract

Background: Within-species skull shape variation of marsupial mammals is widely considered low and strongly size-dependent (allometric), possibly due to developmental constraints arising from the altricial birth of marsupials. However, species whose skulls are impacted by strong muscular stresses – particularly those produced through mastication of tough food items – may not display such intrinsic patterns very clearly because of the known plastic response of bone to muscle activity of the individual. In such cases, shape variation should not be dominated by allometry; ordination of shape in a geometric morphometric context through principal component analysis (PCA) should reveal main variation in areas under masticatory stress (incisor region/zygomatic arches/mandibular ramus); but this main variation should emerge from high individual variability and thus have low eigenvalues.

Results: We assessed the evidence for high individual variation through 3D geometric morphometric shape analysis of crania and mandibles of thre species of grazing-specialized wombats, whose diet of tough grasses puts considerable strain on their masticatory system. As expected, we found little allometry and low Principal Component 1 (PC1) eigenvalues within crania and mandibles of all three species. Also as expected, the main variation was in the muzzle, zygomatic arches, and masticatory muscle attachments of the mandibular ramus. We then implemented a new test to ask if the landmark variation reflected on PC1 was reflected in individuals with opposite PC1 scores and with opposite shapes in Procrustes space. This showed that correspondence between individual and ordinated shape variation was limited, indicating high levels of individual variability in the masticatory apparatus.

Discussion: Our results are inconsistent with hypotheses that skull shape variation within marsupial species reflects a constraint pattern. Rather, they support suggestions that individual plasticity can be an important determinant of within-species shape variation in marsupials (and possibly other mammals) with high masticatory stresses, making it difficult to understand the degree to which intrinsic constraint act on shape variation at the within-species level. We conclude that studies that link micro- and macroevolutionary patterns of shape variation might benefit from a focus on species with low-impact mastication, such as carnivorous or frugivorous species.

2

Introduction

Much of mammalian diversity is reflected in the morphology of the skull (the cranium and mandible), which caters to the reception of sensory input, accommodates a large brain, and maintains a specialized masticatory apparatus [1]. However, mammalian skull diversity does not reflect a random walk through all possible adaptive shapes, but instead evolves within complex intrinsic constraints, which can be phylogenetic, developmental, and/or genetic [2-6].

Hypotheses of how intrinsic constraints shape mammalian skull diversity have been best articulated for marsupial mammals. Their unique altricial birth, and particularly the very early onset of a long suckling phase while the skull is barely developed, is thought to constrain the capacity for the marsupial skull to evolve [7-10]. This is supported by findings of lower cranial and mandibular shape disparity compared to placentals [7, 10-12]. Because the marsupial diversity constraint is thought to be ontogenetic, it is expected to limit the ability of individuals within a species to respond to selection [13, 14]. Several studies have indeed characterized marsupial skull shape over short evolutionary time scales (within species or between close relatives) and suggested that marsupial cranial shape variation mainly changes with size as an allometric "line of least resistance" [3, 4, 12, 15-18]. This is consistent with suggestions that developmental integration may manifest in strong developmental and static allometry [19], which has been specifically posited for marsupials due to their altricial birth [4].

Suggestions of a constraint on marsupial skull shape through static allometry are at odds with several geometric morphometric studies that show only low or moderate levels of static allometry in marsupial crania [15, 20-22]. Two recent studies on kangaroos even suggested that allometry plays a lesser role in shaping cranial variation in this group, instead positing fast adaptation or individual developmental plasticity of the masticatory apparatus as the main driver [23, 24]. This is consistent with suggestions that masticatory biomechanics may impact individual cranial shape to such a degree that developmental constraints either do not contribute much to within-species shape variation, or can

4

be obscured by individual differences in mastication due to the bone's plastic response to mastication stresses [6, 25-27]. Thus, high levels of biomechanical impact on the cranium may cause an important disjunct between within-species (static) *versus* evolutionary patterns of shape variation. This would add an important caveat to the co-interpretation of within- and between-species shape variation because they may have independent sources [13].

In this study, we test the "biomechanics hypothesis" through three-dimensional (3D) geometric morphometric analysis of cranio-mandibular shape variation in three species of specialized grazing marsupial, the wombats. Wombat skulls are under high masticatory stress due to a diet of fibrous grasses [28-31], making them an ideal test case for the "biomechanics hypothesis". In particular, if masticatory stress is a driver of skull shape at the level of individuals, we should expect the main shape variation to occur in areas of high stress [the zygomatic arch, mandibular ramus, and the incisor alveolae; 28, 29-32] according to individual feeding habits. This variation should be dominated by the co-variance between cranium and mandible, be independent of size (*contra* hypotheses of an allometric "line of least resistance"). Lastly, if within- and between-species shape variation arises from different drivers, we expect that overall shape variation patterns should differ when comparing within-species and between-species variation.

For the case that individual feeding behaviour is a major determinant of within-species shape, it is also predicted that the main shape variation within our dataset should be visually similar in all species. However, this main variation should explain much of the overall variation as determined by conventional principal components analysis (PCA) [11, 33], with individuals showing different emphases on different parts of their masticatory system. However, assessing predictions of shape variation for specific shape partitions is challenging. This is because the main techniques for summarizing shape variation are ordination techniques such as PCA, which are based on the ordination of the landmark coordinate's variance/covariance, rather than the range of the actual

variation of the landmark positions. This procedure is blind to biological processes and prone to misleading conclusions [19, 34, 35], particularly where shape variation is visualized along ordination axes with low eigenvalues, as expected here. Thus, ordination-based interpretations of shape variation can lead to erroneous conclusions visually (e.g. variation is "smoothed" and shown without the context of other axes of variation [36]), procedurally (e.g. circularity when using PCA to both propose and test hypotheses), and statistically (e.g. when only a few axes are used for analyses or distinguishing groups or when the groups' main axes of variation differ).

To produce a quantitative assessment of shape variation within our sample, beyond PCA, we here use a permutation test of the prediction that landmarks of Procrustes-superimposed mandibles and crania in areas under high masticatory stress should vary more than the overall landmark distribution in hypothetical PC1 shape extremes. In addition, if our prediction of high individual variation is correct, we predict that the magnitudes of shape variation along PC1 should not consistently match those of pairs of actual landmark configurations in extreme specimens 1) along the main axis of shape variation (PC1), and 2) at extreme ends of Procrustes space. Using this procedure, we show that the morphospace of wombat shape indeed behaves as predicted, with low allometry, small eigenvalues of the first principal component, and limited correspondence between PC1 shape variation and actual specimen shapes.

Results

Disparity comparison, allometry and covariation

The northern hairy-nosed wombat had lower cranial (but not mandibular) Procrustes variance than common and southern hairy-nosed wombats (Table 1). There was no significant effect of sex on either shape or centroid size in the two species with sex data available (Additional File 1).

Table 1: Procrustes variance comparisons among wombat species using pairwise comparison tests with 1000 replicates. Upper/lower diagonals: p-values for comparisons of crania (cran.)/mandibles (mand.). respectively. Procrustes variance values for each species are on the diagonal. HN=Hairy-nosed. Bold *p*-values are significant at *p*<0.05.

	Common	Northern HN	Southern HN
Common	0.00219 (cran.)		
	0.00209 (mand.)	**0.01**	0.73
Northern HN		0.00156 (cran.)	
	0.07	0.00165 (mand.)	**0.017**
Southern HN			0.00210 (cran.)
	0.23	0.44	0.00184 (mand.)

There was either no, or very little, allometry in the cranium and mandible within all wombat species (Table 2). A small but significant effect of allometry exists when the crania or mandibles of all species are analysed together (crania: R^2=0.08; F=5.65; *p*=0.004; mandibles: R^2=0.06; F=3.30; *p*=0.007); plotting cranial and mandible shape against centroid size (Additional File 2) revealed that this is driven by the different shape of the smaller southern hairy-nosed wombats, whereas northern hairy-nosed and common wombats overlap according to size and shape.

The two-block partial least squares analyses revealed high correlation coefficients between crania and mandibles (r-PLS) around 0.90 (Table 1). 2BPLS scores were highly correlated with PC1 scores in all cases except for the cranium of southern hairy-nosed wombats, where instead PC2 and the 2BPLS scores were strongly correlated. By contrast, 2BPLS scores were only correlated with centroid size in southern hairy-nosed wombat mandibles.

7

Table 2: Summaries of the PCA analyses, Allometry analyses, two-block partial least squares (2BPLS) analyses, and correlation statistics between and PLS 1 scores with the species' PC1 scores and centroid sizes. *n*, number of specimens used in the analyses; R^2, coefficient of determination of linear models; t, t-statistic of correlation test; *p*: p-value, CW = common wombat; NHNW = northern hairy-nosed wombat; SHNW= southern hairy-nosed wombat

	n	PC1 %	R^2	Allometry F	*p*	2BPLS *n*	r-PLS	PLS1 scores~PC1 scores cor	t	*p*	PLS1 scores~Csize cor	t	*p*
Cranium													
Common	24	23.49	0.05	1.133	0.283	18	0.85	-0.83	-5.93	**0.000**	-0.046	-0.19	0.86
NHNW	23	21.56	0.11	3.079	**0.001****	11	0.93	-0.7	-2.93	**0.017**	-0.57	-2.1	0.07
SHNW	24	25.92	0.049	1.138	0.305	19	0.93	0.19	0.81	0.43	-0.27	-1.16	0.26
PC2, SHNW								0.89	8.05	**0.000**			
Mandible													
Common	21	23.79	0.59	1.188	0.263			0.99	22.508	**0.000**	0.09	0.36	0.72
NHNW	13	25.88	0.118	1.337	0.148			-0.72	-3.16	**0.012**	-0.58	-2.15	0.06
SHNW	21	18.93	0.091	1.905	**0.019***			0.89	7.67	**0.000**	-0.6	-3.08	**0.007**

Replicating the analyses reported in Table 1 and 2 without surface semilandmarks and with only fixed semilandmarks revealed very few differences. Using fixed and curve semilandmarks only did not change the significance levels of any analysis, and using fixed landmarks changed the significance levels only in cases where significance levels were already close to the p = 0.05 cut-off (Additional File 3)

PCA exploration and landmark position variation

PC1 explained little within-species shape variation in all species (18% to 25% Table 2). PCA of all wombats separated cranial and mandibular shapes of the two wombat genera on the first principal component (PC1; accounting for 45% and 30% of cranial and mandibular variation, respectively; Figure 1). PC1 of the sample of *Lasiorhinus* (accounting for 40% and 30% of cranial/ mandibular variation, respectively), described the differences between northern and southern hairy-nosed wombats with some very minor overlap (Figure 1).

Heat plots of landmark displacement between extreme PC1 shape configurations within species (Figure 2) confirm our expectation that the overall main variation relates to the muzzle, zygomatic arches, mandibular ramus at the masticatory muscle attachments (masseteric fossa, coronoid process, and angular shelf) and incisor alveolae. In addition, in all three species, longer, more ventral rostral regions coincide with less flared and more ventrally placed zygomatic arches. PC1 heatplots of all wombats and the hairy-nosed wombats only (Figure 3) – describing the differences between the two genera and hairy-nosed wombat species, respectively - show limited similarity to the within-species patterns, with the zygomatic arch or mandibular muscle attachments not showing much displacement between PC1 extremes. PC1 variation of hairy-nosed wombats differed from the within-species variation by reflecting extensive displacement in the occipital region and variation in the frontal region. In the mandible, the symphysis did not vary among hairy-nosed wombat species; the distal edge of the angular process varied most between the hairy-nosed wombats, together with a change in angle of the mandibular condyle. Relative to hairy nosed wombats, common wombats have a more

9

ventrally directed masseteric scar, anteriorly elongate nasal region, extensive ventral displacement of

the premaxillary suture, and a more dorsally placed palate. Further, the cranial and mandibular molar

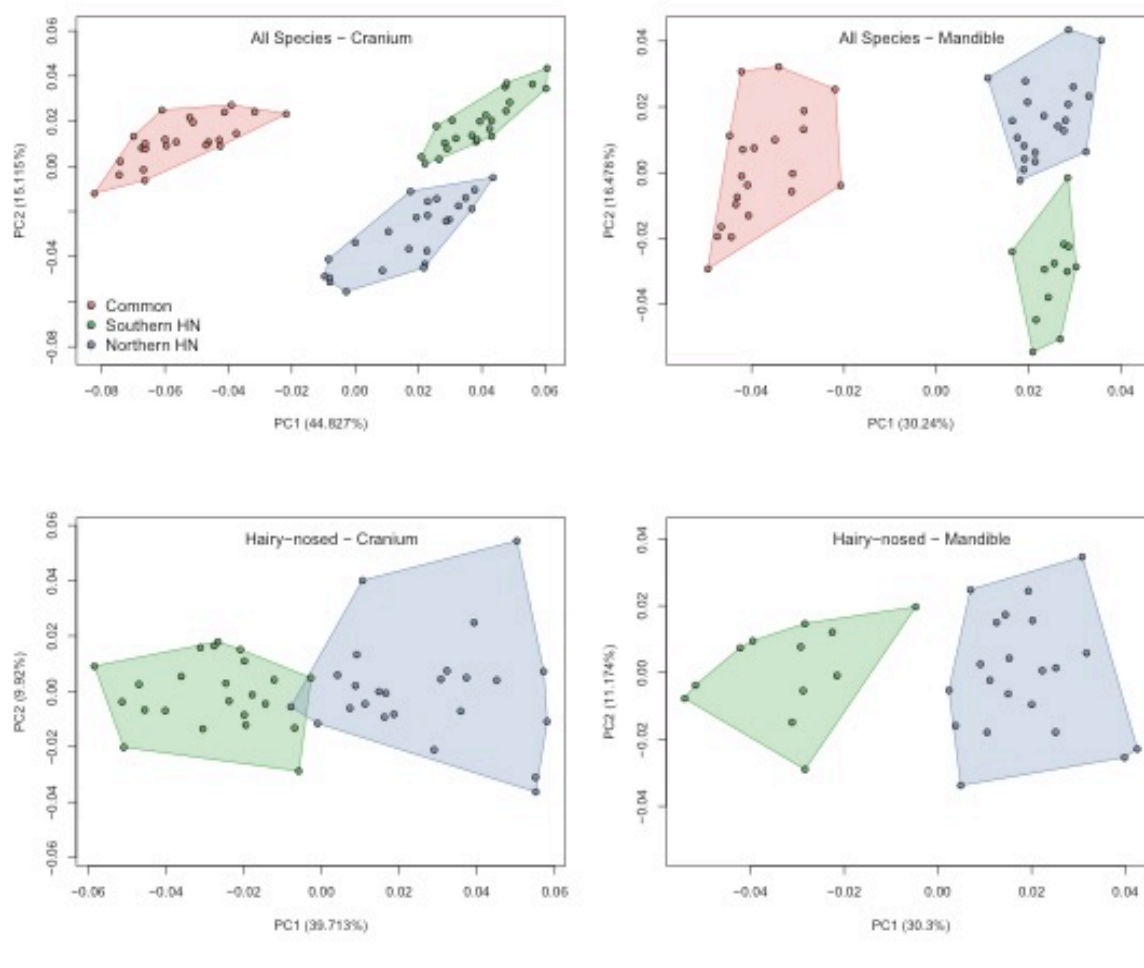rows are more dorsally situated, and the mandibular condyle is more medially directed.

Figure 1: Principal Component 1 vs. 2 plots of cranial (left) and mandibular (right) shapes, showing

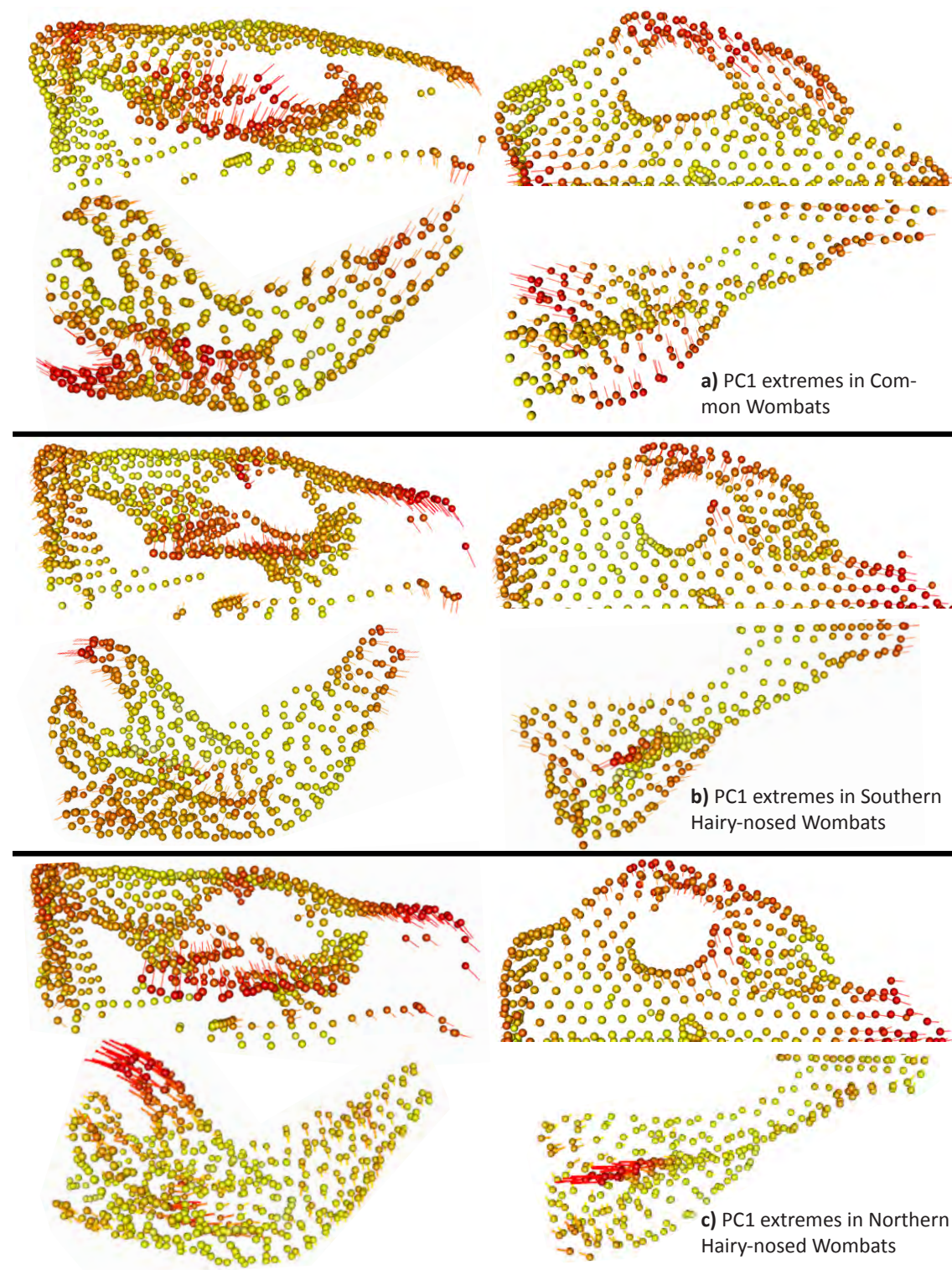distributions of specimens in the all-wombat (above) and hairy-nosed wombat (below) sample.

a) PC1 extremes in Common Wombats

b) PC1 extremes in Southern Hairy-nosed Wombats

c) PC1 extremes in Northern Hairy-nosed Wombats

Figure 2: Heat plots representing the difference in landmark position between the two most extreme specimens along PC1. Dots are the position of one landmark of one PC extreme, lines represent the displacement of the same landmark in the other. Colour heat reflects displacement magnitude (red/yellow = high/low displacement).
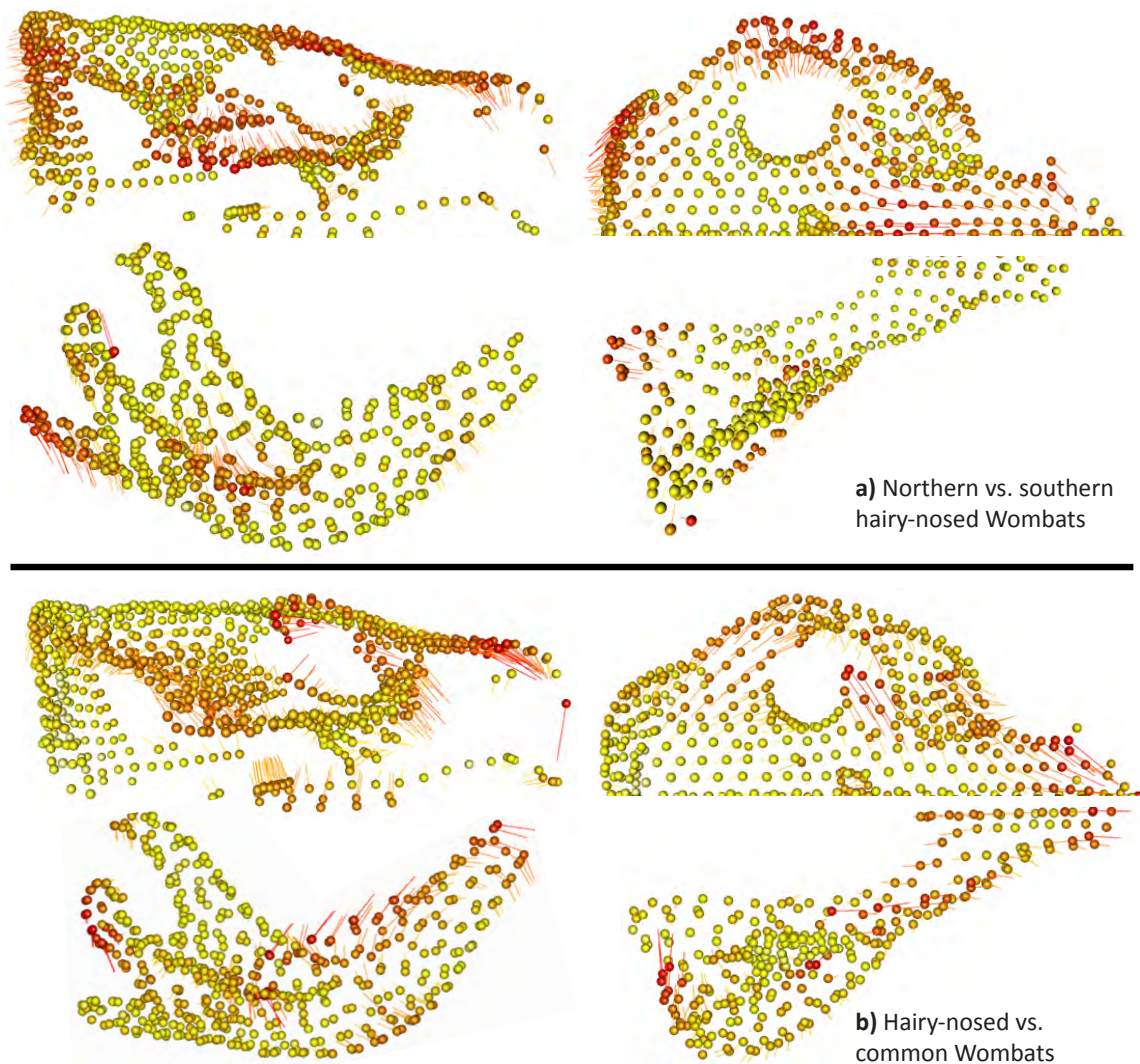
11

Figure 3: Heat plots representing the range of landmark displacement along PC1 in a) the all-wombat sample and b) in the sample of all hairy-nosed wombats, grouping the species of the genus.

Landmark position variation test

The results of our landmark variation tests are summarized in Figure 4, and the corresponding specimen configurations for each within-species comparison at 100% confidence intervals are in Additional Files 4 and 5. Few results were robust to rarefaction to minimum partition sizes (black frames in Figure 4), so we cannot exclude that the significances might be an artefact of the larger numbers of landmarks in the other partitions. The best supported hypothesis across all comparisons was that the "remainder of cranium" partitions should vary less among all landmarks. Our hypotheses

12

that the zygomatic arch, mandibular muscle attachments, and incisor area show greater landmark displacement were only partially supported, with highest support overall in comparisons of hypothetical PC1 configurations. Comparisons of specimen configurations on PC1 or Procrustes space extremes provided less support for our hypotheses, either by showing no significant displacement or by being significantly displaced but without being significantly different from the overall distribution of landmark displacements. Some comparisons between individual configurations even contradicted the visual impression from the heat plots by yielding lower displacements in the tip of rostrum/zygomatic arch partitions (yellow frames in Figure 3, see also Additional figures 4 and 5). This was particularly obvious in the whole-Procrustes space comparisons, showing that the variation according to PC1 is not a main component of shape across all of Procrustes space. Interestingly, all three hypotheses of cranial partition displacement were supported in comparisons of mean configurations of common and hairy-nosed wombats, while none was supported in mean species configurations between northern and southern hairy-nosed wombats.

	Cranium			Mandible			
	Zygomatic Arch	Tip of snout	Remainder	Masticatory muscle insertions	Symphyseal area	Remainder	
Northern HN, 100% CI	0.082	0.094	−0.174				*hypothetical PC1 projection*
95% CI	0.068	0.077	−0.143				
Southern HN, 100% CI		0.1	−0.158	0.138	0.049	−0.184	
95% CI		0.087	−0.137	0.122	0.043	−0.164	
Common, 100% CI	0.237		−0.268	0.175		−0.194	
95% CI	0.197		−0.223	0.15		−0.167	
Northern HN, 100% CI		0.117	−0.138				*PC1 extremes*
95% CI	0.131		−0.139				
Southern HN, 100% CI	0.18	0.035	−0.215	0.114	0.104	−0.217	
95% CI		0.149	−0.194	0.161		−0.137	
Common, 100% CI	0.393	−0.058	−0.333	0.197	0.061	−0.262	
95% CI	0.238		−0.259	0.137	−0.04		
Northern HN, 100% CI		0.044	−0.132		−0.042		*GPA extremes*
95% CI	−0.097	0.095					
Southern HN, 100% CI		0.141	−0.142	0.111	0.104	−0.219	
95% CI	0.172	0.086	−0.258	0.202	−0.062	−0.141	
Common, 100% CI	0.395	−0.059	−0.334	0.163		−0.148	
95% CI	0.207		−0.186		0.047	−0.138	
All wombat means (GPA)	0.13	0.073	−0.207	−0.149	0.042		
Hairy nose means (GPA)				0.058	−0.018		

Figure 4: Landmark variation test results, between the most extreme specimens (100% and 95% CI) based on the hypothetical PC1 projection, PC1 and the GPA (rows). Negative and positive values:

13

smaller and larger range of landmark variation than the whole cranium or mandible. Grey: no significant differences of landmark displacement magnitude ($p > 0.001$); Magenta: significant displacement differences undistinguishable from the full statistical distribution of landmarks (Bhattacharrya Coefficient $p > 0.001$). Green: significant differences in magnitude and statistical distribution. Black frames, significant also when rarefied. Yellow frames, results opposite to hypothesized effect.
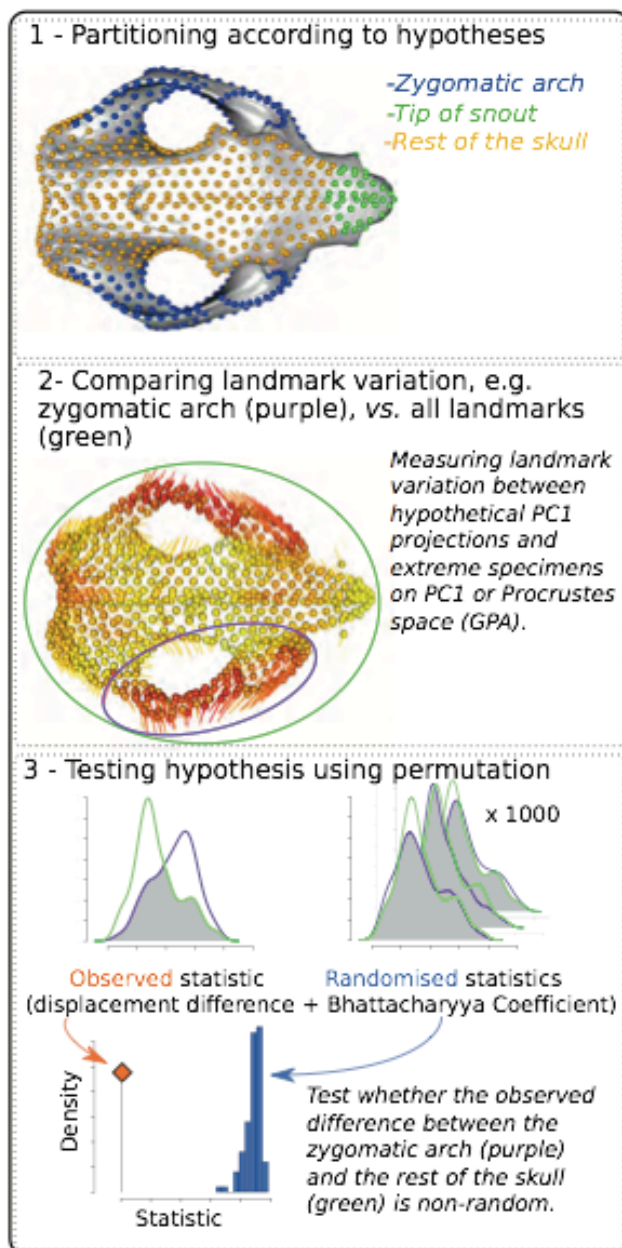

Figure 5: Outline of the landmark variation test. 1- identifying landmark partitions and statistical null. 2 – measuring distance between corresponding landmarks in pairs of specimen (visualized as

lines) 3 - Permutation test to assess whether a partition varies more than the whole of the skull based on displacement difference and Bhattacharyya Coefficient.

Discussion

Our results support our "biomechanics hypothesis" of a non-allometric, mastication-related driver of cranio-mandibular shape variation within wombat species. As expected for a skull shaped by masticatory biomechanics [29, 31, 32, 47], the heat plots of hypothetical PC1 shape extremes within species reveal strikingly uniform patterns of high landmark displacements in the zygomatic arches and rostra, also consistent with our hypothesis that these areas should vary most within species (in the mandible, displacement directions are not as uniform but still all occur in the muscle attachement sites). There is little indication that allometry plays any role in shaping within-species shape variation, with both overall shape and cranio-mandibular shape co-variation not related to centroid size (and shape co-variation instead strongly correlated with the main ordinated variation).

We also find substantial support for our prediction of high levels of individual variation within each species, as expected if individual diet or feeding habit were a main driver of skull shape. In particular, the main axes of shape variation had low eigenvalues, with PC1 explaining at most 25% of variation. Our "landmark test" of landmark displacement magnitudes confirmed that these low eigenvalues are related to inconsistencies between ordinated and individual configurations. Specifically, our hypotheses of greater displacement in the "tip of rostrum" and "zygomatic arch" partitions are best-supported in the hypothetical configurations of PC1 extremes, with less support or even contradictory evidence in actual specimens with opposite PC1 scores or on opposite ends of Procrustes space. In addition, comparisons of actual specimens between PC1 extremes and in extremes of Procrustes space showed that displacement directions (which were not measured by our landmark test) in the areas under high displacement differed substantially in the actual specimens, particularly in the mandibles. Thus, despite their overall similarity, shape variation along PC1 within species clearly arises from individual shapes in areas of high mechanic stress that only reflect parts of the variation identified by

15

PC1. This contrasts with predictions of individual shapes falling along a spectrum of change along a high-eigenvalue PC1, as expected under a scenario of constraint. Rather, it concurs with our expectation that within-species morphospaces can contain too much "noisy" variation to infer a constraint or interpret in a context of macroevolutionary patterns [6, 13, 27, 48].

While individual variability in our sample is high, shape variation within species tends to occur most frequently in areas of high masticatory stress, which also seems to cause the uniform within-species PC1 landmark displacement patterns. Remodelling of these areas according to individual mastication habit or dietary preferences thus seems a plausible source of individual variation. This is consistent with finds that the zygomatic arch grows differentially upon the onset of feeding in two marsupial species [Virginia opossums and New Guinea Quolls; 49, 50], and remodels under mechanical stress in mammals [26, 47, 51, 52]. A similar remodelling process may also cause the high shape variation in the occipital crest along our within-species PC1, which is under mechanical stress from the temporalis [31] and nuchal muscles [53]. Lastly, the genetically and spatially highly restricted population of northern hairy-nosed wombat showed only slightly lower cranial shape disparity, and similar overall landmark displacement patterns, compared to the other two wombat species; this again suggests a strong role of non-heritable feeding behaviours, rather than a role of genetic diversity [3], in shaping individual wombat skulls.

Our findings support previous concerns about allocating the contributions of individual variability and heritable adaptation in closely related mammals [6, 27]. In particular, the within-species variation occurs in parts of the skull that also adapt to a herbivorous lifestyle in macroevolutionary contexts [54-57], posing potential issues for distinguishing between individual plasticity and between-species differences across evolutionary time scales. For example, the within-species PC1 displacement patterns of cranial shapes resembles some aspects of landmark displacement between the mean of *Vombatus* and *Lasiorhinus*. However, this difference between genera has evolved over the last seven

16

million years and likely reflects the larger masseter muscle [30, 31], and different angle of incision [58] of common wombats.

The ambiguity of distinguishing within- and between-species variation in closely related species might also play a role in recent debates surrounding the existence of a possible universal Cranial Rule of Evolutionary Allometry (CREA). This posits that closely related mammals tend to have longer rostra and narrower zygomatic arches as they increase in size [59], in a pattern that is highly similar to the one we found within wombats but without allometry. CREA was suggested for kangaroo crania [18], but contested in a slightly different sample and landmarking protocol of kangaroos [23, 24]. Instead, the authors postulated that a CREA-like pattern is not allometric and instead purely due to biomechanics. CREA-like shape variation is also found in other kangaroo species, and is allometric in tammar wallabies [60] and quokkas [20], but only partially so among the rock wallabies [15]. By contrast, geometric morphometric analyses of two marsupial species with a comparatively soft diet – the insectivorous *Dromiciops gliroides* and two species of the omnivorous *Caluromys* – show little cranial variation of zygomatic arch or rostrum shape, but strong-species allometry [21, 22]. The variable association of the CREA shape variation with allometry may therefore be related to varying biomechanical stresses on individual skulls depending on a species' feeding ecology (but note that CREA has been shown within mongooses and fruit bats, which both do not feed on hard items [59]). Feeding ecology, or other biomechanical skull function, therefore seems to be an important consideration in the choice of species for intraspecific assessments of developmental constraints [see also 61].

In methodological terms, use of the landmark variation test presents a useful additional tool to the geometric morphometric toolkit for interpreting shape variation. In our specific case, the spread of variation across several low-eigenvalue ordination axes rendered visualization through scatterplots of two or three axes ambiguous (i.e. the space containing important information was more multidimensional); co-investigation of individual specimen configurations in Procrustes space

17

improved our interpretation of the morphospace. Similar procedures can also support the interpretation of other analyses (e.g. understanding the Procrustes space variation underlying patterns of allometry or other variables). Note that, like all procedures that interpret Procrustes-based landmark variation, the landmark variation test can be subject to artificial signals of directional variation due to the least-squares fitting of Procrustes superimposition [62, 63]; this can be alleviated by using high numbers and even coverage of landmarks. Note that in the case of high landmark numbers, other analyses may need to be assessed for robustness towards high numbers of landmarks [42]. Our approach also has the caveat that permutation tests have low resolution on small distributions [64], such as the occipital crest in our wombats. In such cases, conventional visual assessments of landmark displacement might be a better, if less quantitatively tractable, approach.

Conclusions

Our results point towards a mostly biomechanically caused, mastication-related drive of wombat skull shape variation, suggesting that important drivers of shape macroevolution, such as a possible constraint on marsupial skull shape, may only emerge above the species level in some mammals [6, 13, 27, 48]. This posits an important challenge in testing hypotheses of constraints, and identifying differences between heritable and epigenetic variation, within mammalian skulls [13]; this is already being acknowledged, for example in studies that account for population-level shape "noise" in tree dating [65]. More detailed ontogenetic studies of shape might allow the separation of "original" skull shape prior to the onset of feeding [as in 49, 50], might help separate epigenetic from developmental shape variation.

Methods

Data collection and landmarking

We collected 3D data from crania and mandibles of adult common wombats (*Vombatus ursinus*; 24 crania/21 mandibles), southern hairy-nosed wombats (*Lasiorhinus latifrons*; 24 crania/21 mandibles) and northern hairy-nosed wombats (*Lasiorhinus krefftii*; 23 crania/13 mandibles). Data were acquired using two computed tomography (CT) scanners (Toshiba Activion16 at the School of Veterinary Sciences, Gatton, The University of Queensland) and a Somatom Definition AS+ scanner at I-MED Radiology, Armidale). Scans were converted to 3D surface meshes in Mimics v.19 (Materialise, Belgium). Author CS repaired surface meshes of five specimens that had minor damage (Additional Table 3). All original scans and 3D data are available on MorphoSource (www.morphosource.com, Project P418). For a list of specimens, their provenance, animal ethics and permit numbers for all non-museum specimens, and repairs, refer to Additional File 6; for scan details, refer to MorphoSource. The disparate numbers of crania and mandibles derive from the fact that not all specimens had both crania and mandibles available, so that we included just crania or just mandibles to maximize sample sizes. Samples of common and southern hairy-nosed wombat were across their geographical range (respectively, east to west South Australia / from Tasmania to Queensland). We only had specimens of Northern hairy-nosed wombat samples from the single surviving population (Additional File 6). To test whether this resulted in lower shape variance, we employed Procrustes variance comparisons among the species. We also tested if sexual dimorphism of shape was a possible confounding factor in the two wombat species with sex data (southern and common wombat), by performing Procrustes ANOVAs of shape/size against sex.

Landmarks were digitized by CS in Viewbox 4 (dHAL). This involved placing 65 landmarks and 761 semilandmarks (261 curve and 500 surface) on the cranium and 35 landmarks and 542 semilandmarks (142 curve and 400 surface) on the mandible (Additional File 7, Additional File 8). Landmarks and curve semilandmarks were placed manually, and their position was used to automatically place the surface semilandmarks, based on a template of an undamaged specimen of southern hairy-nosed

19

wombat (CLVSJR5). Curve and surface semilandmarks were slid over the 3D surface of each scan (as opposed to being interpolated without surface information) according to the bending energy criterion [37, 38] in Viewbox v. 4.0.1.7. Landmark coordinates were analysed in R [39] using geomorph v. 3.1.0 [40] for standard analyses and a new package, landvR (https://github.com/TGuillerme/landvR) [41] for the permutation tests. Coordinates were scaled and superimposed using generalized Procrustes analysis (GPA) without a sliding step (as sliding was done over the 3D geometries in Viewbox), and ordinated through PCA. Separate GPAs and PCAs were done for crania and mandibles in all specimens, for hairy-nosed wombats only, and separately for each species. Data and code for all analyses are available and repeatable as R markdown files at https://github.com/TGuillerme/landmark-test.

Landmark sampling

Our landmarking protocol was designed to densely capture featureless areas of the skull, such as the cranial vault, forehead, or mandibular ramus. Although we used Viewbox to ensure that landmarks were slid over the actual 3D surface, rather than interpolated from just fixed landmarks/curve semilandmarks, the dependence of the semilandmarks of fixed landmarks raises the issue of pseudo-replication [34, 35]. However, for our landmark test, high numbers of fixed landmarks were important. To ensure our other analyses were not affected by the high number of semilandmarks [e.g. 42], we therefore re-ran all analyses with just fixed and curve semilandmarks, and just fixed landmarks.

Allometry and covariation

To test our prediction that allometry should not be a main driver of within-species cranio-mandibular shape, we first performed a Procrustes ANOVA of shape predicted by centroid size. To assess how cranio-mandibular integration related to allometry and the main ordinated variation, we also

20

conducted a two-block partial least-squares (2BPLS) analysis of cranio-mandibular covariation in specimens with crania and mandibles. This finds the axis of greatest shape covariation between the cranium and mandible and identifies how much of shape variation in each dataset is due to the covariation with the other. We then asked if cranio-mandibular co-variation was a likely driver of the main shape variation by correlating 2BPLS scores with PC1 scores within each species. We also asked if there was evidence of an allometric constraint on the co-variation, by correlating the 2BPLS scores with centroid size.

Landmark variation test

We performed a permutation-based landmark position variation test to test two hypotheses of shape variation, 1) that areas of the skull under high masticatory stresses show the greatest variation magnitudes within species and 2) that patterns of landmark displacement differ within and among species. This procedure is sketched in Figure 4, detailed in Additional File 9 and implemented in the landvR package (including tutorial vignettes) accessible at https://github.com/TGuillerme/landvR [41].

Partition selection

Based on the literature on wombat skull biomechanics [31-34, 41], we selected four partitions expected to vary most and one partition expected to not vary (as a null hypothesis). In the cranium, partitions expected to vary were 1) the zygomatic arch (262 landmarks) and 2) the "rostrum" partition (the incisor alveolae, premaxillary suture, and nasal bones immediately dorsal to this region; 66 landmarks). In the mandible, these were the mandibular "muscle attachment" partition (the masseter/pterygoid/temporalis attachments in the mandibular ramus; 142 landmarks) and the "anterior symphysis" partition (anterior to the molars at the incisor roots; 35 landmarks). Partitions not expected to vary were the rest of the cranium (501 landmarks) or mandible (400

21

landmarks) (Additional File 10). To account for different landmark numbers per partition, we also rarefied each partition to the minimum landmark number (66/cranium and 35/mandible).

Comparison between pairs of specimens and mean shapes

We sampled specimen pairs to reflect within-species variance by selecting: 1) the projection of the hypothetical extremes of configurations along PC1; 2) landmark configurations of specimens with the highest *vs.* lowest PC1 score; and 3) the most different specimens in Procrustes space (based on distance from the mean shape – see Additional text 1). To increase reliability of our results, we also tested pairs of specimens within the 95% CI of each distribution. Additionally, we tested if differences in mean shapes among species correspond to our hypotheses of shape change.

Testing differences between extreme specimens by partitions

For each specimen pair, we tested whether a partition exhibited greater displacement of landmarks compared to the rest of the cranium or mandible. For each pair, we 1) calculated the Euclidean distance between corresponding landmarks in both specimens in Procrustes space; 2) calculated the displacement difference (i.e. total landmark distance between two specimens) and the Bhattacharyya Coefficient [43-45] (i.e. overlap between the distance distributions – see Additional text 1); and 3) performed a permutation test [46] on both statistics by comparing them to the statistics of 1000 same-sized random partitions (Figure 5). As the great number of tests (96) will lead to a type I error inflation, we reduced our p-value acceptance threshold from the traditional 0.05 to 0.001 (0.1%)

Declarations

Ethics approval and consent to participate

Where specimens were not sourced from museum collections, they were sourced with ethics permissions as outlined in Additional Table 3.

Consent for publication

Not applicable

Availability of data and material

The raw CT scans for all specimens used, and their derivative 3D surface files used for landmarking, are available on MorphoSource (www.morphosource.com, Project P418). The coordinate data and code to repeat all analyses for this study is available as rmd files at https://github.com/TGuillerme/landmark-test. Our landvR package for implementation of the landmark variation test, including a step-by step vignette, is accessible on https://github.com/TGuillerme/landvR .

Competing interests

The authors declare that they have no competing interests

24

References

1.   Novacek MJ. Patterns of diversity in the mammalian skull. In: Hanken J, Hall BK, editors. The Skull. 2. Chicago: University of Chicago; 1993. p. 438-545.

2.   Marroig G, Cheverud JM, Wainwright P. Size as a line of least evolutionary resistance: Diet and adaptive morphological radiation in New World monkeys. Evolution. 2005;59:1128-42.

3.   Porto A, Sebastião H, Pavan SE, VandeBerg JL, Marroig G, Cheverud JM. Rate of evolutionary change in cranial morphology of the marsupial genus *Monodelphis* is constrained by the availability of additive genetic variation. J Evol Biol. 2015;28:973-85.

4.   Porto A, Shirai LT, de Oliveira FB, Marroig G. Size variation, growth strategies, and the evolution of modularity in the mammalian skull. Evolution. 2013;67:3305-22.

5.   Goswami A, Smaers J, Soligo C, Polly P. The macroevolutionary consequences of phenotypic integration: from development to deep time. Phil Trans R Soc B. 2014;369:20130254.

6.   Cardini A, Elton S. Does the skull carry a phylogenetic signal? Evolution and modularity in the guenons. Biol J Linn Soc. 2008;93:813-34.

7.   Goswami A, Randau M, Polly PD, Weisbecker V, Bennett CV, Hautier L, et al. Do developmental constraints and high integration limit the evolution of the marsupial oral apparatus? Int Comp Biol. 2016;56:404-15.

8.   Goswami A, Weisbecker V, Sánchez-Villagra MR. Developmental modularity and the marsupial-placental dichotomy. J Exp Zool B. 2009;312:186-95.

9.   Weisbecker V. Are monotremes primitive and marsupials inferior? In: Klieve A, Hogan L, Johnston S, Murray P, editors. Marsupials and Monotremes - Nature's Enigmatic Mammals. New York: Nova; 2015. p. 397-411.

10.   Goswami A, Polly PD, Mock OB, Sánchez-Villagra MR. Shape, variance and integration during craniogenesis: contrasting marsupial and placental mammals. J Evol Biol. 2012;25:862-72.

11.   Bennett CV, Goswami A. Statistical support for the hypothesis of developmental constraint in marsupial skull evolution. BMC Biol. 2013;11:52.

12.   Prevosti FJ, Turazzini GF, Ercoli MD, Hingst-Zaher E. Mandible shape in marsupial and placental carnivorous mammals: a morphological comparative study using geometric morphometrics. Zool J Linn Soc. 2012;164:836-55.

13.   Klingenberg CP. Studying morphological integration and modularity at multiple levels: concepts and analysis. Phil Trans R Soc B. 2014;369.

14.   Klingenberg CP. Evolution and development of shape: integrating quantitative approaches. Nat Rev Gen. 2010;11:623.

15.   Rodrigues HG, Hautier L, Evans AR. Convergent traits in mammals associated with divergent behaviors: the case of the continuous dental replacement in rock-wallabies and African mole-rats. J Mamm Evol. 2017;24:261-74.

16.   Goswami A. Phylogeny, diet, and cranial integration in australodelphian marsupials. PLOS ONE. 2007;2:e995.

17.   Shirai LT, Marroig G. Skull modularity in neotropical marsupials and monkeys: size variation and evolutionary constraint and flexibility. J Exp Zool B. 2010;314B:663-83.

18. Cardini A, Polly D, Dawson R, Milne N. Why the long face? Kangaroos and wallabies follow the same 'rule' of cranial evolutionary allometry (CREA) as placentals. Evol Biol. 2015;42:169-76.

19. Klingenberg CP, Neuenschwander BE, Flury BD. Ontogeny and individual variation: analysis of patterned covariance matrices with common principal components. Syst Biol. 1996;45:135-50.

20. Dawson R, Milne N. Cranial size and shape variation in mainland and island populations of the quokka. J Zool. 2012;288:267-74.

21. Valladares-Gómez A, Celis-Diez JL, Palma RE, Manríquez GS. Cranial morphological variation of *Dromiciops gliroides* (Microbiotheria) along its geographical distribution in south-central Chile: a three-dimensional analysis. Mamm Biol. 2017;87:107-17.

22. Magnus LZ, Machado RF, Cáceres N. Comparative ecogeographical variation in skull size and shape of two species of woolly opossums (genus *Caluromys*). Zool Anz. 2017;267:139-50.

23. Mitchell DR, Sherratt E, Ledogar JA, Wroe S. The biomechanics of foraging determines face length among kangaroos and their relatives. Proc R Soc B. 2018;285:DOI: 10.1098/rspb.2018.0845

24. Mitchell DR, Sherratt E, Sansalone G, Ledogar JA, Flavel RJ, Wroe S. Feeding Biomechanics Influences Craniofacial Morphology at the Subspecies Scale among Australian Pademelons (Macropodidae: Thylogale). J Mamm Evol. 2018:1-11.

25. Franks EM, Holton NE, Scott JE, McAbee KR, Rink JT, Pax KC, et al. Betwixt and between: intracranial perspective on zygomatic arch plasticity and function in mammals. Anat Rec. 2016;299:1646-60.

26.   Ravosa MJ, Lopez EK, Menegaz RA, Stock SR, Stack MS, Hamrick MW. Adaptive plasticity in the mammalian masticatory complex: You are what, and how, you eat. Primate craniofacial function and biology: Springer; 2008. p. 293-328.

27.   Caumul R, Polly PD, Janis C. Phylogenetic and environmental components of morphological variation: Skull, mandible, and molar shape in marmots (*Marmota*, Rodentia). Evolution. 2005;59:2460-72.

28.   Crompton AW. Masticatory motor programs in Australian herbivorous mammals: Diprotodontia. Int Comp Biol. 2011;51:271-81.

29.   Crompton AW, Lieberman DE, Owerkowicz T, Baudinette RV, Skinner J. Motor control of masticatory movements in the Southern hairy-nosed wombat (*Lasiorhinus latifrons*). In: Vinyard C, Ravosa MJ, Wall CE, editors. Primate Craniofacial Function and Biology. New York: Springer; 2008. p. 83-111.

30.   Nakajima K, Townsend G. A morphometric study of the skulls of two species of wombats (*Vombatus ursinus* and *Lasiorhinus latifrons*). Aust Mammal. 1994;17:65-72.

31.   Sharp AC, Trusler PW. Morphology of the jaw-closing musculature in the common wombat (*Vombatus ursinus*) using digital dissection and magnetic resonance imaging. PloS one. 2015;10:e0117730.

32.   Sharp AC. Comparative finite element analysis of the cranial performance of four herbivorous marsupials. J Morphol. 2015;276:1230-43.

33.   Felice RN, Randau M, Goswami A. A fly in a tube: Macroevolutionary expectations for integrated phenotypes. Evolution. 2018;72:2580-94.

34.   Bookstein FL. A method of factor analysis for shape coordinates. Am J Phys Anthrop. 2017;164:221-45.

28

35.   Bookstein FL. A newly noticed formula enforces fundamental limits on geometric morphometric analyses. Evol Biol. 2017;44:522-41.

36.   Bookstein FL. The relation between geometric morphometrics and functional morphology, as explored by Procrustes interpretation of individual shape measures pertinent to function. Anat Rec. 2015;298:314-27.

37.   Bookstein FL. Morphometric tools for landmark data: Geometry and biology. Cambridge, UK: Cambridge University Press; 1997.

38.   Gunz P, Mitteroecker P. Semilandmarks: a method for quantifying curves and surfaces. Hystrix. 2013;24:103-9.

39.   R Core Team. R: A language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing; 2018.

40.   Adams DC, Collyer ML, Kaliontzopoulou A. Geomorph: Software for Geometric Morphometric Analysis. v. 3.1.0. 2018.

41.   Guillerme T, Weisbecker V. landvR: Tools for measuring landmark position variation. 2019. Available from: https://github.com/TGuillerme/landvR/tree/v0.2.

42.   Bookstein FL. Pathologies of Between-Groups Principal Components Analysis in Geometric Morphometrics. bioRxiv. 2019:627448.

43.   Guillerme T. dispRity: A modular R package for measuring disparity. Methods Ecol Evol. 2018;0.

44.   Guillerme T, Cooper N. Effects of missing data on topological inference using a Total Evidence approach. Mol Phyl Evol. 2016;94:146-58.

45.   Bhattacharyya A. On a measure of divergence between two statistical populations defined by their probability distributions. Bull Calcutta Math Soc. 1943;35:99-109.

46.   Manly BF. Randomization, Bootstrap and Monte Carlo Methods in Biology: Chapman and Hall/CRC; 2006.

47.   Franks EM, Scott JE, McAbee KR, Scollan JP, Eastman MM, Ravosa MJ. Intracranial and hierarchical perspective on dietary plasticity in mammals. Zoology. 2017;124:30-41.

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

49.   Flores DA, Giannini N, Abdala F. Comparative postnatal ontogeny of the skull in the australidelphian metatherian *Dasyurus albopunctatus* (Marsupialia: Dasyuromorpha: Dasyuridae). J Morphol. 2006;267:426-40.

50.   Abdala F, Flores DA, Giannini NP. Postweaning ontogeny of the skull of *Didelphis albiventris*. J Mammal. 2001;82:190-200.

51.   Nogueira MR, Peracchi AL, Monteiro LR. Morphological correlates of bite force and diet in the skull and mandible of phyllostomid bats. Functional Ecol. 2009;23:715-23.

52.   Hylander WL, Johnson KR, Picq PG. Masticatory-stress hypotheses and the supraorbital region of primates. Am J Phys Anthrop. 1991;86:1-36.

53.   Arnold P, Esteve-Altava B, Fischer MS. Musculoskeletal networks reveal topological disparity in mammalian neck evolution. BMC Evol Biol. 2017;17:251.

54. Figueirido B, Serrano-Alarcón FJ, Slater GJ, Palmqvist P. Shape at the cross-roads: homoplasy and history in the evolution of the carnivoran skull towards herbivory. J Evol Biol. 2010;23:2579-94.

55. Monteiro L, R., Bonato V, Dos Reis Sérgio F. Evolutionary integration and morphological diversification in complex morphological structures: mandible shape divergence in spiny rats (Rodentia, Echimyidae). Evol & Develop. 2005;7:429-39.

56. Panagiotopoulou O, Cobb S, N. The mechanical significance of morphological variation in the macaque mandibular symphysis during mastication. Am J Phys Anthrop. 2011;146:253-61.

57. Hautier L, Lebrun R, Cox PG. Patterns of covariation in the masticatory apparatus of hystricognathous rodents: Implications for evolution and diversification. J Morphol. 2012;273:1319-37.

58. Scott G, Richardson K, Groves C. Osteological differences of the skulls of *Lasiorhinus latifrons* Owen, 1845 and *Vombatus ursinus* Shaw, 1800 (Marsupialia, Vombatidae). Aust J Zool. 1988;36:599-609.

59. Cardini A, Polly PD. Larger mammals have longer faces because of size-related constraints on skull form. Nat Comms. 2013;4:2458.

60. Hadley C, Milne N, Schmitt LH. A three-dimensional geometric morphometric analysis of variation in cranial size and shape in tammar wallaby (Macropus eugenii) populations. Aust J Zool. 2009;57:337-45.

61. Jones KE, Ruff CB, Goswami A. Morphology and biomechanics of the pinniped jaw: Mandibular evolution without mastication. Anat Rec. 2013;296:1049-63.

62. Rohlf FJ, Slice D. Extensions of the Procrustes method for the optimal superimposition of landmarks. Syst Biol. 1990;39:40-59.

63.   Zelditch ML, Swiderski DL, Sheets HD. Geometric Morphometrics for Biologists: A

Primer. New York: Academic Press; 2012.

64.   Efron B, Tibshirani RJ. An Introduction To The Bootstrap. Boca Raton: CRC Press; 1993.

65.   Álvarez-Carretero S, dos Reis M, Yang Z, Goswami A. Bayesian estimation of species

divergence times using correlated quantitative characters. 2019.