Selection for increased tibia length in mice alters skull shape through parallel changes in developmental mechanisms

Colton M. Unger[1,3], Jay Devine[4], Benedikt Hallgrímsson [3,4,5], and Campbell Rolian[2,3*]

[1] Department of Biological Sciences, University of Calgary, Calgary, AB T2N 4N1, Canada.

[2] Department of Comparative Biology and Experimental Medicine, Faculty of Veterinary Medicine, University of Calgary, AB T2N4N1, Canada.

[3] McCaig Institute for Bone and Joint Health, Calgary, AB T2N4N1, Canada.

[4] Department of Cell Biology and Anatomy, University of Calgary, Calgary, AB, Canada.

[5]Alberta Children's Hospital Research Institute for Child and Maternal Health, University of Calgary, Calgary, AB, Canada.

*Corresponding author: Tel.: +1 (403) 210-3888 or Email: cprolian@ucalgary.ca

Abstract

Bones in the vertebrate cranial base and limb skeleton grow by endochondral ossification, under the control of growth plates. Mechanisms of endochondral ossification are conserved across growth plates, which increases covariation in size and shape among bones, and in turn may lead to correlated changes in skeletal traits not under direct selection. We used micro-CT and geometric morphometrics to characterize shape changes in the cranium of the Longshanks mouse, which was selectively bred for longer tibiae. We show that Longshanks skulls became longer, flatter, and narrower in a stepwise process. Moreover, we show that these morphological changes likely resulted from developmental changes in the growth plates of the Longshanks cranial base, mirroring changes observed in its tibia. Thus, indirect and non-adaptive morphological changes can occur due to developmental overlap among distant skeletal elements, with important implications for interpreting the evolutionary history of vertebrate skeletal form.

Introduction

Organismal development is a major determinant of phenotypic variation, and therefore is fundamentally related to how organisms evolve (Hendrikse, Parsons and Hallgrímsson, 2007; Hallgrímsson and Lieberman, 2008). Organisms are comprised of interrelated anatomical elements whose morphology is patterned by shared genetic pathways (i.e., pleiotropic genes) and often by the same developmental processes (Hallgrímsson and Hall, 2005; Murren, 2012). Shared genetic and/or developmental processes lead to morphological integration, that is, the tendency of sets of traits to covary more strongly internally than with traits in other sets (Cheverud, 1996). In turn, integrated individual anatomical structures contribute to the modular organization of biological systems (Wagner, Pavlicev and Cheverud, 2007; Hallgrímsson *et al.*, 2009).

If two anatomical structures are integrated due to underlying genetic phenomena, such as pleiotropy or linkage disequilibrium, then those traits are more likely to respond to selection in a concerted manner (Armbruster and Schwaegerle, 1996; Cheverud, 1996). As a result of integration, correlated responses to selection can result in phenotypic changes in some traits that are merely a consequence of covariation with other traits under selection (Gould and Lewontin, 1979; Wagner, 1984; Price and Langen, 1992; Parsons *et al.*, 2015). Understanding how developmental processes lead to correlated responses to selection is pivotal to distinguishing adaptive changes from those that are non-adaptive, or potentially even maladaptive, in analyses of phylogeny, ancestral relationships and evolutionary change within lineages (Gould and Lewontin, 1979; Riska, 1986; von Cramon-Taubadel, 2019).

The bones of the terrestrial vertebrate cranial floor (basicranium) and the postcranial skeleton represent an interesting case of integration because they are physically distant, yet both develop

by the process of endochondral ossification (EO) (De Beer, 1937; White and Wallis, 2001; Mackie *et al.*, 2008). There is therefore the potential for evolutionary changes in one structure to cause correlated phenotypic changes in the other. EO proceeds through the formation, expansion, and mineralization of a cartilaginous template, known as an anlage, that is patterned *in utero* and undergoes post-natal longitudinal expansion (Kronenberg, 2003; Mackie *et al.*, 2008; Lefebvre and Bhattaram, 2010; Roselló-Díez and Joyner, 2015). In the limbs, ossification initiates by the formation of primary and secondary ossification centers and continues into post-natal development via specialized growth plates situated at the ends of the long bones (Kronenberg, 2003; Mackie *et al.*, 2008; Lefebvre and Bhattaram, 2010). The postcranial growth plate, once formed, is comprised of three histologically distinct zones containing cartilage-producing cells (chondrocytes) in different physiological states: resting, proliferative, and hypertrophic (Roselló-Díez and Joyner, 2015).

The basicranium is comprised of three bones: the basioccipital, the basisphenoid, and the presphenoid, which make up the floor of the caudal portion of the skull in mammals and are formed by growth in the spheno-occipital (SOS) and intersphenoidal (ISS) synchondroses (Wei *et al.*, 2016). Synchondroses are structurally analogous to growth plates, however, synchondroses grow bidirectionally and have duplicated proliferative and hypertrophic zones (Wei *et al.*, 2016). Basicranial growth is thought to be a key determinant of overall skull shape. The basicranium is the first cranial skeletal element to develop and is controlled intrinsically by EO-like mechanisms, whereas the face and calvarium are influenced by, and grow in response to, hormonal regulation of surrounding tissue and brain growth, respectively (Scott, 1958; Waters and Kaye, 2002; Bastir and Rosas, 2006; Richtsmeier *et al.*, 2006). Additionally, the basicranium

supports the brain and contains critical foramina for the passage of vasculature and cranial nerves and is therefore central to proper craniofacial development (Lieberman *et al.*, 2008).

Here, we used the Longshanks mouse to study correlated evolution of cranial and post-cranial skeletal elements. The Longshanks mouse was established through artificial selection for increased tibia length relative to body mass, using an outbred CD1 stock. By generation 20, mean tibia length in two independently selected Longshanks lines had increased by 13-15% in comparison to random-bred Controls from the same genetic background with no change in average body mass (Marchini *et al.*, 2014; Castro *et al.*, 2019). Investigation of the cellular mechanisms governing limb development in Longshanks revealed structural alterations in the postnatal epiphyseal growth plate of the tibia. Specifically, the Longshanks selection regime resulted in larger tibial growth plates with larger resting and proliferative zones, without changes in cell division rate or timing of growth plate fusion compared to Controls (Marchini and Rolian, 2018). Previous analyses also suggested the tibia selection regime resulted in mice that are skeletally larger in relation to body mass, with correlated skeletal responses at the systemic level (Sparrow *et al.*, 2017), along with potentially maladaptive changes in skeletal microarchitecture (Farooq *et al.*, 2017; Cosman, Britz and Rolian, 2019).

The Longshanks experiment offers a unique opportunity to study correlated evolution in skeletal traits that were not directly under selection, in a model with known evolutionary history, under controlled laboratory settings. Given the underlying developmental relationship between the long bones and cranial base, we investigated whether selection for increased tibia length indirectly altered the shape of the Longshanks cranium. We tested the general hypothesis that selection for increased tibia length produced indirect responses in the cranium of Longshanks through changes to the shared process of endochondral ossification. Specifically, we predicted that Longshanks

crania will have a series of craniofacial morphological changes corresponding to altered synchondrosis size/architecture. To test this hypothesis, we compared the 3D shape of adult Longshanks crania from both Longshanks lines to Controls across three evenly spaced generations in the selection experiment. We also used a combination of morphometric analysis and histology to investigate cranial development in Longshanks neonates.

Materials and Methods

Animal samples

All animal procedures were approved by the Health Sciences Animal Care Committee at the University of Calgary (AC13-0077) and (AC17-0026) and performed in accordance with best practices outlined by the Canadian Council on Animal Care. For more information on the husbandry methods and selective Longshanks breeding regimen, see Marchini *et al*. (2014).

We collected 8-week old, non-breeder Longshanks mice (N=327) from generations 1, 9, and 20 across three experimental lines: Longshanks 1 (hereafter LS1), Longshanks 2 (LS2), and Control (CTL) to study changes in adult cranial shape at the beginning, middle, and end of the selection process, respectively (for group sample sizes, see Table 1). Each group was as sex and family balanced as possible to account for differences due to sexual dimorphism and/or family diversity (Karp *et al.*, 2017). To investigate the developmental basis of the Longshanks cranium, we generated postnatal day seven (P07) neonates (N=104) from F31 Longshanks mice (Table 1). The three lines have not actively undergone artificial selection since generation F22 and are maintained as experimental populations. We selected P07 as our developmental time point as this

is when Longshanks tibiae are growing fastest, and the cranial skeleton is still actively growing (Vora, Camci and Cox, 2016; Farooq *et al.*, 2017; Marchini and Rolian, 2018).

X-ray Micro-Computed Tomography (µCT)

We performed X-ray micro-computed tomography (µCT). We used a Skyscan 1173 v1.6 µCT scanner (Bruker, Kontich, Belgium) to acquire whole-body scans of the adults and separate scans of the neonate cranium and tibiae. We obtained adult samples from frozen archived carcasses at each generation, while F32 neonates were scanned the day they were euthanized. In addition, we scanned the corresponding right hindlimb of each neonate that underwent cranium scanning. Adult scans were acquired at 70-80 kV and 60-75 µA with 44.73 µm isotropic voxels and no filter, while neonates were scanned at a resolution of 17.04 µm isotropic voxels with otherwise identical parameters. Stack reconstructions were performed using NRecon v1.7.4.2 (Bruker, Kontich, Belgium).

Histology

We dissected neonate crania after scanning them and fixed them in 10% neutral buffered formalin (NBF) (Thermo Scientific) for 48 hours, with NBF replacement every 24 hours. Fixed cranium tissues were then transferred to a decalcifying solution (Cal-Ex II™, Fisher Chemical) for 72 hours with daily solution changes. After decalcification, a rectangular portion of the cranial base containing both basicranial synchondroses was dehydrated, embedded it in paraffin, and sectioned in the sagittal plane at 12 µm. Sections were deparaffinized in Slide Brite™ (Jones Scientific Products, Inc.) and subsequently stained. The slides of a specimen were stained in an alternating fashion with two stains: (1) Wiegert's Iron Haematoxylin (Sigma), 0.05% Fast-Green (FCF) (Sigma), counterstained in 0.1% Safranin-o solution (Sigma); or (2) Gill's Haematoxylin

#3 (Sigma), rinsed in 70% ethanol, and counterstained with 1% alcoholic Eosin Y (Sigma). We imaged sagittal midline sections using an Axio Scan.Z1 slide scanner (Ziess, Oberkochen, Germany) at 10x magnification and qualitatively evaluated differences in growth plate size and morphology.

Landmarking

μCT adult and neonate crania scans were subjected to a novel image registration-based pipeline to automatically detect landmarks for a geometric morphometrics (GM) shape analysis (Percival *et al.*, 2019). Automated landmarking improves data standardization and can be used to quickly process very large sample sizes while reducing intraobserver errors, such as landmark placement drift (Fruciano, 2016; Devine *et al.*, 2020). Automated landmarking involves volumetric registration using a global affine alignment of the skull volumes, followed by a dense non-linear deformation between each cranium and a reference atlas. Here, the atlas is an average volume, with a standardized landmark configuration, that best minimizes intensity differences from the rest of the sample. We used 68 3D landmarks for the adults and 50 3D landmarks for the neonates (Supplementary Figure S1 and S2; Supplementary Tables S1 and S2).

We computed the affine transformations with a multi-resolution framework, where the μCT volumes are translated, scaled, rotated, and sheared at progressively higher resolutions until their affine alignment with the atlas is maximized (Lerch, Sled and Henkelman, 2010). We computed the non-linear transformations with the multi-resolution SyN (Symmetric Normalization) algorithm (Avants *et al.*, 2011), which involves symmetrically flowing an image pair into one another along a velocity field. We then recovered, concatenated, and inverted the transformations, and finally propagated the atlas landmarks along this path to the original image

space for analysis. All image processing was performed with the open-source MINC (Medical Imaging NetCDF) toolkit (https://github.com/BIC-MNI/minc-toolkit-v2).

In addition to investigating overall neonate cranium shape, we characterized cranial base shape with two-dimensional (2D) landmarks at the sagittal midline. We used a 12 landmark set highlighting the vertices of the developing basicranial bones which provides information about the shape of the sagittal cross-section of the basicranial synchondroses (Supplemental Figure S3; Supplementary Table S3). Landmarks at the midline were placed in Amira v.5.4.2 (Visage Imaging, Berlin, Germany) by one observer (CMU) blind to the identity of the specimens. Adult tibiae lengths were quantified in Amira by calculating the distance, in mm, between two landmarks that we placed on the distal tip of the lateral malleolus and most lateral point on the proximal epiphysis, two anatomical points that were demonstrated to have high homology and repeatability (Cosman, Sparrow and Rolian, 2016). Because neonate tibia length is not fully visible in the scans due to small or absent secondary ossification centers (Moss, 1977), neonate tibia measurements were obtained from the distance, in mm, between landmarks placed on the distal and proximal ends of the ossified tibial diaphysis on the rostral edge along the sagittal midline of the tibia.

Geometric Morphometrics

Analyses were performed on the R/Rstudio computational platform (R Core Team, 2020). We investigated shape cranial differences by superimposing the adult and neonate landmark configurations into age-specific morphospaces via Generalized Procrustes Analysis (GPA). To study the influence of selection on cranial shape, we first corrected for confounding variables known to alter adult and neonatal morphology.

In the adult dataset, we controlled for the effects of sex and size. Upon regressing shape on sex, we observed that sex accounted for a small but significant amount of variation (2.2%), although there were no sex-specific differences in cranial responses to selection (data not shown). Using sex-adjusted residuals, we investigated allometry in the Longshanks cranium to parse out how much of the cranial selection response, if any, could be attributed to changes in skeletal size. While Procrustes superimposition removes scale, it does not account for differences in biological shape that are associated with size i.e. allometry (Klingenberg, 2016). Because Longshanks mice are skeletally larger in relation to body mass in the post-cranium (Sparrow *et al.*, 2017), we employed a pooled within-group analysis of covariance (ANCOVA) of cranium centroid size on body mass to determine whether the same trend exists in the Longshanks cranium. Mean centroid size after accounting for body mass was significantly different among lines at F20 ($F = 14.97$, $p < 0.001$), with Longshanks LS1 and LS2 lines having larger crania than Controls (Supplementary Figure S4) (Tukey's post-hoc test, LS1vsLS2 $p = 0.947$, LS1vsCTL $p < 0.001$, LS2vsCTL $p < 0.001$). There was no difference in mean centroid size, after controlling for covariation with body mass, among founder (F01) samples (Supplementary Figure S4) (LS1vsLS2 $p = 0.896$, LS1vsCTL $p = 0.999$, LS2vsCTL $p = 0.999$).

For the neonate dataset, we controlled for the effects of litter size but not sex, due to uncertainties in assigning sex anatomically in neonates. After regressing cranial size and tibia length on litter size, we observed a strong negative correlation (Pearson, $r = -0.72$, $p < 0.001$). LS2, which had litter sizes that were ~2 pups larger than LS1 and ~4 pups larger than controls on average, exhibited significantly smaller centroid sizes than LS1 and Controls (ANOVA, $F = 19.51$, Tukey's HSD, LS1vsLS2 $p < 0.001$, LS1vsCTL $= 0.764$, LS2vsCTL $p < 0.001$) (Supplementary Figure S5; Supplementary Table S4). Thus, we performed our neonate analyses

with Procrustes shape variables and univariate measurements, such as tibia length, adjusted for litter size effects.

Group differences in adult and neonate cranial morphology were evaluated using principal component analyses (PCA). We assessed whether group mean shapes, independent of size and/or sex, were statistically significantly different using a randomized residual (1000 permutations) Procrustes ANCOVA (Goodall, 1991; Collyer, Sekora and Adams, 2015). Post-hoc pairwise tests compared differences in least-squares means between groups (Collyer and Adams, 2018). For visualizations of cranial shape differences between lines, we used deformation heatmaps and cranial meshes with vectors of shape change that depict transformations between group means. All GM analyses were performed in R with the *geomorph*, *Morpho,* and *RRPP* packages (Schlager, 2017, 2020; Collyer and Adams, 2018; Adams, Collyer and Kaliontzopoulou, 2020).

Results

Longshanks Adults

Body mass and cranium size allometry is altered in Longshanks adults

F01 mice (founders) that had not been subjected to selection did not differ in average weight or tibia length between lines (Supplementary Figure S6; Supplementary Table S5). Moreover, random-bred F09 Controls and F20 Controls did not differ from F01 founders in terms of tibia length in ANCOVA accounting for body mass (Supplementary Figure S6; Supplementary Table S5). In contrast, LS1 and LS2 at F09 have an average of 7.3% longer tibiae compared to F09 Controls, while LS1 and LS2 at F20 have 16.4% longer tibiae on average when compared to F20 Controls (Supplementary Figure S6; Supplementary Table S5). Average body mass in our sample was stable between lines across all three generations and did not differ significantly in all but two pairwise comparisons between groups (ANOVA, $F = 3.312$) (Supplementary Figure S6; Supplementary Table S5). In contrast, at generation F20, LS1 and LS2 mice had significantly larger cranium centroid sizes than Controls (ANOVA, $F = 312.7$, Tukey's HSD, F20 LS1vsF20 CTL $p < 0.05$, F20 LS2vsF20 CTL $p < 0.05$), though the latter did not differ from F01 or F09 Controls (Tukey's HSD, F01 CTLvsF09 CTL $p = 0.929$, F09 CTLvsF20 CTL $p = 0.803$, F01 CTLvsF20 CTL $p = 1.000$).

Given that the long bones of the Longshanks post-cranial skeleton are larger than Controls at any given body mass (Sparrow *et al.*, 2017), we asked if the allometric scaling relationship between Longshanks crania and overall body mass had changed in response to 20 generations of selection. Body mass was significantly correlated with cranium centroid size in our adult dataset (Pearson, $r = 0.697$, $p < 0.001$). ANCOVA comparing mean cranium centroid size among lines

using body mass as the covariate indicates that there was a significant difference in centroid size between Controls and Longshanks at F20, however LS1 and LS2 did not differ from each other ($F = 14.97$, Tukey's HSD, LS1vsLS2 $p = 0.947$, LS1vsCTL $p < 0.001$, LS2vsCTL $p < 0.001$). Hence, Longshanks selected lines have skeletally larger crania after 20 generations of selection independent of body mass (Supplementary Figure S4).

The Longshanks cranium is longer, narrower, and flatter

Next, we asked if the fact that tibia length and cranium centroid size increases in F20 LS1 and LS2 is associated with shape differences in their cranium compared to F09 and F01 mice. We performed a principal component analysis (PCA) on the residuals of a multivariate regression of shape on sex in order to control for potential sex effects on cranial shape in our sample. Comparison of PC score means between groups demonstrates that despite overlap in skull shape, LS1 and LS2 have shifted substantially into positive PC1 space, reflecting crania that are longer and narrower with reduction in vault height (Figure 1A, Supplementary Movie M1). Post-hoc pairwise comparisons from a Procrustes ANCOVA comparing adult cranium shape by group, independent of sex effects, showed that all groups within line by generation, or within generation by line, differ in mean shape, except F01 LS2 and Controls ($F= 10.679$, F01 LS2vsF01 CTL $p = 0.203$). When comparing the Euclidean distance among group sex-adjusted PC score means, however, F20 LS1 and LS2 mice are on average over twice as far from unselected mice in morphospace (i.e., all F01 founders, F09 and F20 controls) than the latter are from each other (mean Euclidean distances 0.023 vs 0.010, Supplementary Table S6)

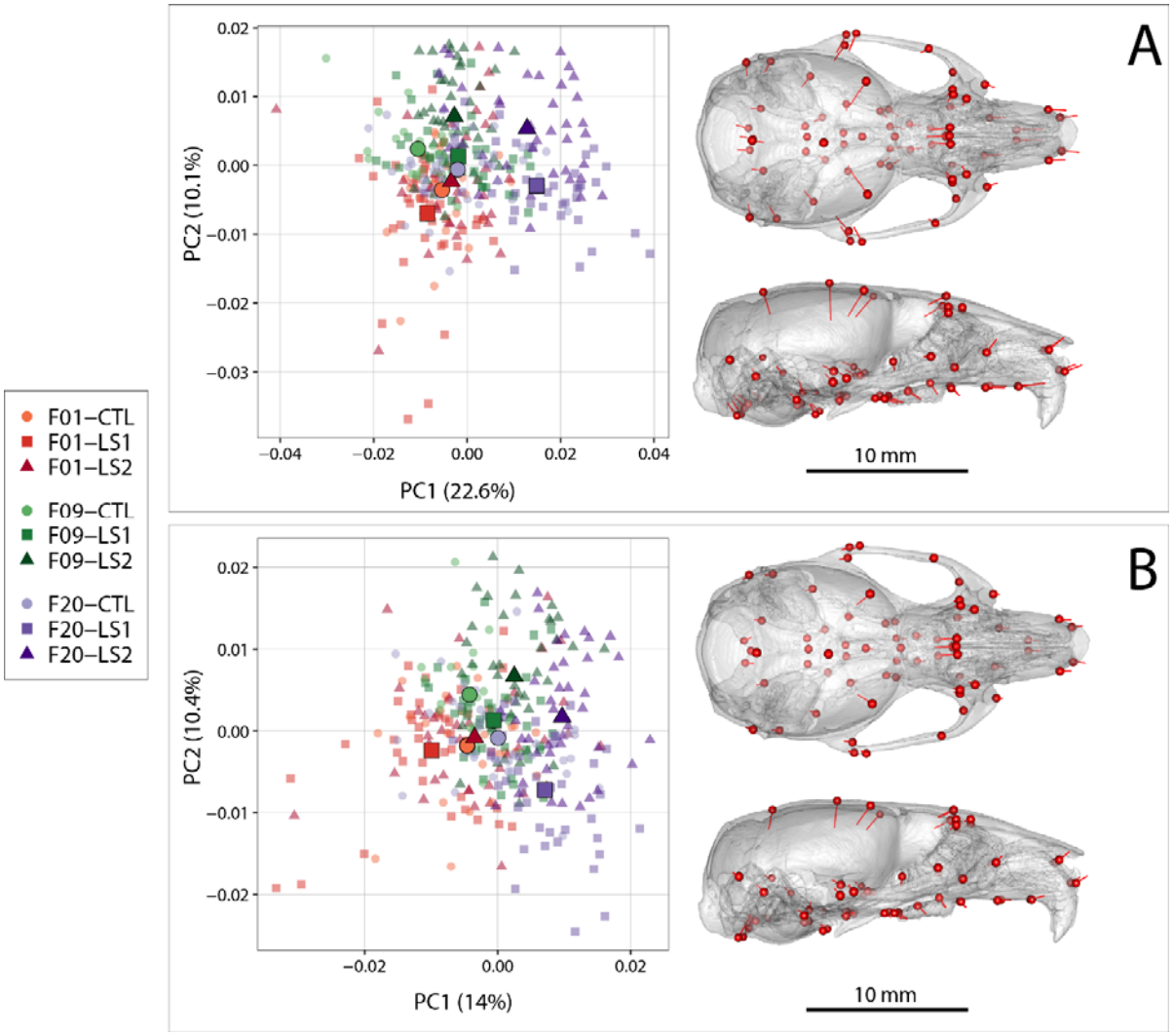


Figure 1 – Scatter plots of the first two principal components (PC) of Procrustes shape variables in adult Longshanks and Controls throughout the selection process. (A) Plot of sex-adjusted Procrustes shape variables (left), and vectors of shape change at each cranium landmark (magnified 2 times for visualization) showing shape transformations along PC1 from negative to positive scores (right). Large symbols indicate mean PC1 and PC2 scores for each respective cohort. (B) Plot of Procrustes shape variables additionally corrected for size.

275

276

277

Longshanks cranial shape differences remain after controlling for skull size and sex

Given that PC1 generally captures differences in shape primarily due to allometric effects of size (Klingenberg, 2016), and that LS1 and LS2 have larger skulls in F20, we asked if the cranium of F20 selected lines score more positively simply because they are larger, and if larger skulls are associated with different cranium shapes in F20 selected lines compared to Controls and F01 unselected lines. We compared the fitted PC1 scores of a pooled within-group regression of shape on size to log (centroid size), which shows the cranium size and shape scaling relationship, between all nine groups in our sample (Supplementary Figure S7). At any given centroid size, Longshanks F20 selected lines score more positively (longer and narrower) in predicted shape. Importantly, the slopes of the lines, which capture the scaling relationship between cranium shape and size did not differ significantly between any of the groups in our sample ($p > 0.05$). Thus, while Longshanks F20 LS1 and LS2 have larger crania at any given body mass compared to Controls, the allometric pattern within the cranium itself was not altered by selection for increased tibia length.

The difference in intercept between the fitted PC1 scores and log centroid size of LS1 and LS2 in relation to Controls (Supplementary Figure S7) suggest that while the increase in size of F20 selected crania contributes to the shape differences along PC1, it is not the only cause of shape variation. We therefore asked if differences in shape between the Longshanks and Control cohorts persist when the effect of size is removed from our sample by using multivariate regression residuals of shape on size. The PCA of shape independent of size and sex shows a marked reduction in group separation along PC1; however, F20 LS1 and LS2 still typically score more positively in PC1, corresponding to crania that are relatively longer, narrower, and have reduction in vault height (Figure 1B). Post-hoc pairwise comparisons from a Procrustes

ANCOVA comparing adult cranium shape by group independent of size and sex effects showed that all groups within line by generation, or within generation by line, differ in mean shape ($F = 8.205$, $p < 0.001$) except F01 LS2 and Controls (F01 LS2vsF01 CTL $p = 0.173$). As with sex-adjusted data, F20 Longshanks mice are substantially farther in morphospace from unselected groups than the latter are from each other (mean Euclidean distances 0.017 vs 0.010, Supplementary Table S6).

Tibia length and cranial shape are uncorrelated

After identifying underlying differences in cranial shape between F20 Longshanks and Controls irrespective of sex and size effects, we asked how this residual cranial shape variation was related to variation in tibia length, the principal selection target. To determine whether the residual cranial shape changes was a correlated response to tibia selection we performed within-group regressions on size and sex adjusted PC1 scores against tibia length (Figure 2). Differences in the slopes of within-group regressions over time, and in comparison to Controls, would reveal if LS1 and LS2 crania experienced directional selection (Price and Langen, 1992). Clusters of points for Longshanks specimens progressively increase in tibia length and adjusted cranial shape score over time (Figure 2), a trend which is not seen in Controls (Figure 2). More importantly, within cohorts, tibia length and cranium shape are not significantly correlated (Figure 2). While the population means for residual cranium shape and tibia length were changing over time in Longshanks selected lines, selected individuals with relatively longer tibiae did not necessarily also have derived cranial shapes. In other words, unlike tibia length, residual cranial shape likely did not confer a fitness advantage.

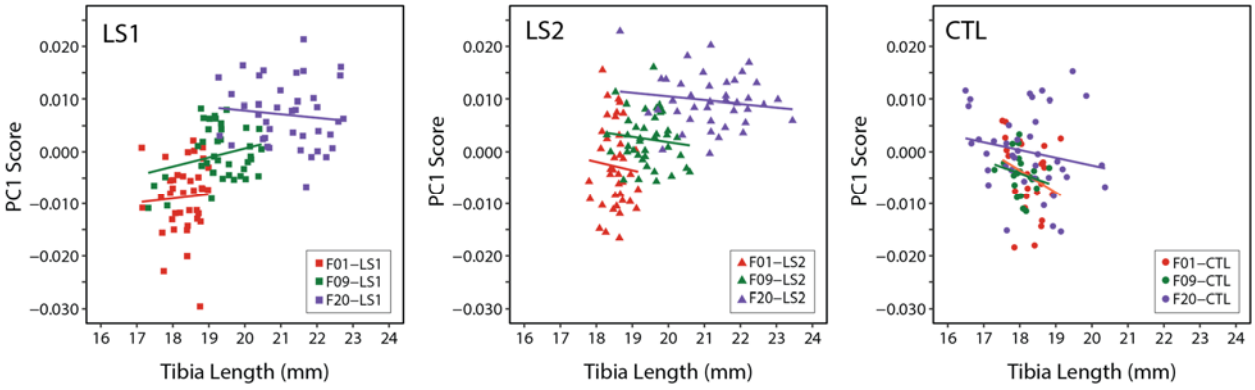


Figure 2 - Scatter plots with regression lines by cohort showing the relationships between PC1 score (cranium shape) and tibia length in adult mice throughout selection. Cranium shape data is adjusted for the effects of sex and size related allometry.

Intergenerational changes occurred in a stepwise process

F09 LS1 and LS2 score more positively along PC1 in the same direction as F20 LS1 and LS2 after removing size effects (Figures 1B, 3). In other words, F09 LS1 and LS2 appear to have intermediate shapes along PC1 between F01 groups and F20 selected lines (Figure 3). This led us to ask how intergenerational changes in cranium shape occurred throughout the selection process. We computed the mean shapes of LS1 and LS2 lines over time and compared them using deformation heat maps to track shape change between generations within a selection line. Our results show that indirect responses to selection in the Longshanks cranium occurred in a stepwise process: shape change in the first nine generations of selection contributed to the reduction in vault height, whereas the remaining 11 generations of tibia selection led to a reduction in cranial width at the zygomatic arches in parallel with snout elongation (arrows in Figure 3). In comparison, intergenerational changes in the Control lines shows virtually no change in cranial vault height between F01 and F09, and a reduction in the occipital area of the cranium from F09 to F20 (Figure 3).

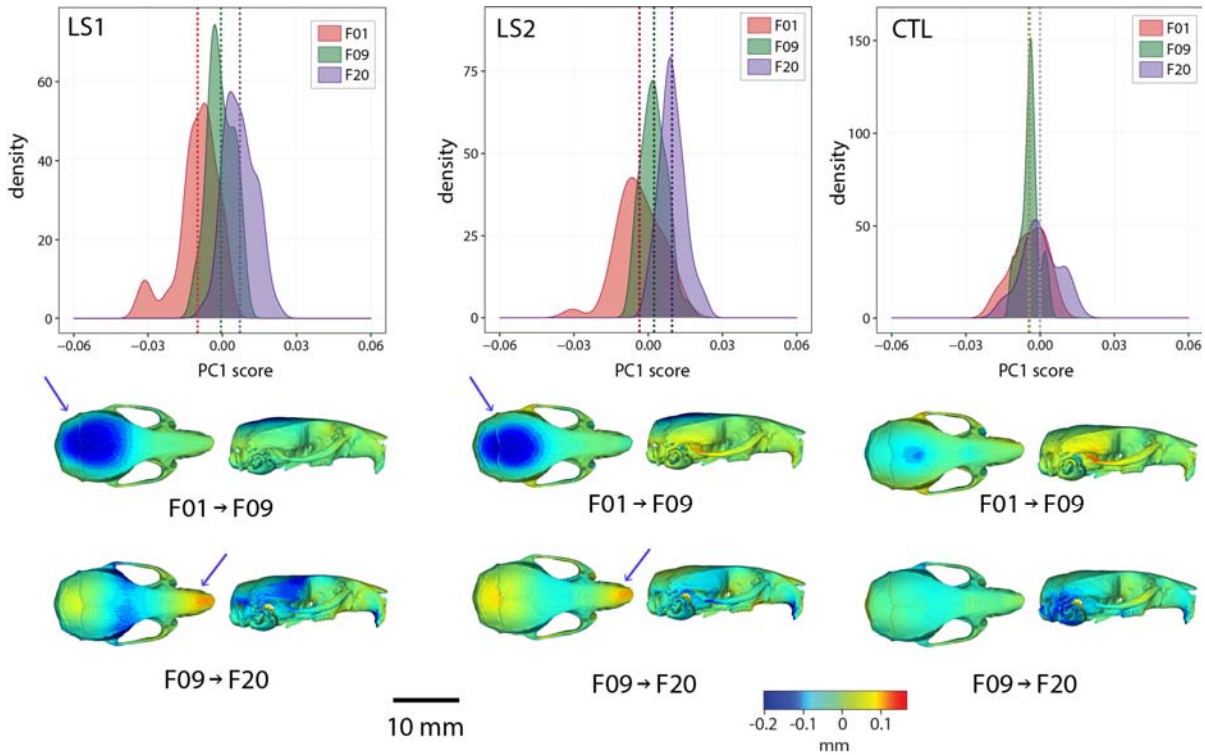


Figure 3 - Intergenerational shape changes within both Longshanks lines and Controls throughout the selection process. Top: Density plots following intergenerational shifts in mean PC1 scores within LS1 (left), LS2 (center), and CTL (right) lines for size and sex adjusted shape data. Bottom: Heatmaps showing shape transformations between mean shapes in the first 9 generations of selection (F01 to F09) and the next 11 generations (F09 to F20) after correcting for size and sex effects. Blue indicates areas of relative reduction, red indicates areas of relative expansion, and green indicates neutral areas. Longshanks independently undergo vault height reduction between generations F01 and F09, followed by snout elongation between generations F09 and F20 (blue arrows).

Longshanks P07 Neonates

Neonate crania have similar shape patterns as Longshanks adults

We investigated if we could detect the adult pattern in shape differences earlier in ontogeny. Using one-week old (P07) Longshanks neonates, we compared cranial shape at a time when the Longshanks tibia is growing most rapidly and tibia length differences are already observable (Farooq *et al.*, 2017; Marchini and Rolian, 2018) (Supplementary Figure S5, Supplementary Table S4). After regressing out litter size effects, our PCA of neonate cranial shape showed a large separation in morphospace between the selected lines and Controls (Figure 4). LS1 and LS2 cluster more closely than Controls and have skulls that are longer, narrower and have reduced vault heights (Figures 4, 5). The Procrustes ANCOVA and pairwise comparisons showed that LS1, LS2 and Control neonates significantly differ from each other in cranium shape ($F = 14.395$, LS1vsCTL $p < 0.001$, LS2vsLS1 $p < 0.001$, LS2vsCTL $p < 0.001$). In addition, we observed via deformation heatmaps that the cranial pattern seen in Longshanks adults exists by one-week post partum and becomes more marked with age (Figure 5). The LS2 selection replicate appears to have reduced magnitudes of cranial response compared to LS1 at F20 and in neonates (Figures 1B, 3, 5).

Longshanks neonate cranial bases are flatter than Controls and differ in synchondrosis shape

Given the underlying developmental relationship between the cranial base and the long bones, we asked if the neonate cranial bases differed in shape along the sagittal plane between Longshanks and Controls, where the synchondroses' primary axis of elongation exists. We performed a 2D morphometric analysis and found that groups differed in cranial base shape after

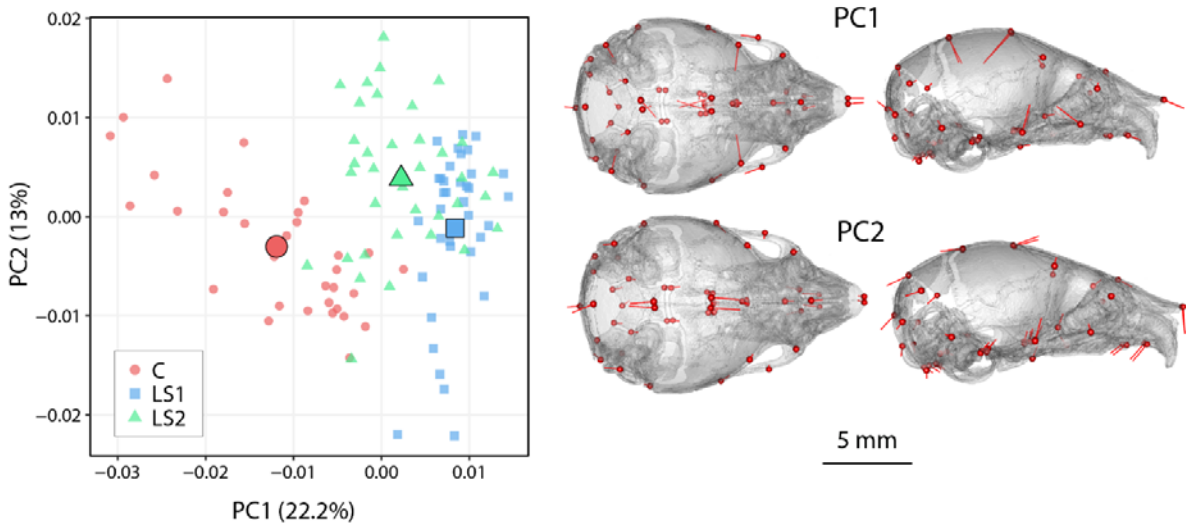


Figure 4 - Scatter plots of the first principal components (PC) in neonate Longshanks and Control cranium Procrustes shape variables at generation 32 (F32). Left: Plot of litter size adjusted Procrustes shape variables (left), large symbols indicate mean PC1 and PC2 scores for each respective cohort. Shapes of individual points indicate Longshanks lines (circle = CTL, square = LS1, and triangle = LS2). Right: Neonate cranium with vectors of shape change at each cranium landmark (magnified 4 times for visualization) showing shape transformations along PC1 (top) and along PC2 (bottom) from negative to positive scores.

removing litter size effects by Procrustes ANCOVA ($F = 20.972$, $p < 0.001$). As with the neonate cranial form, Procrustes ANCOVA and pairwise comparisons of cranial base shape showed that LS1, LS2 and Control neonate mean cranial shapes all differ from each other. Longshanks neonate cranial base shapes differed from Controls in a similar pattern, but to different extents, with LS2 assuming an intermediate position in cranial base morphospace (Figure 6A). Deformations comparing a mean Control cranial base to LS1 and LS2 means show a flattening of the cranial base in both Longshanks lines (Figure 6B). Moreover, the shape of the ISS changes in LS1 and LS2 compared to Controls, expanding dorsally to become more wedge-shaped, whereas the SOS shows no significant shape change differentiating them from the Control SOS

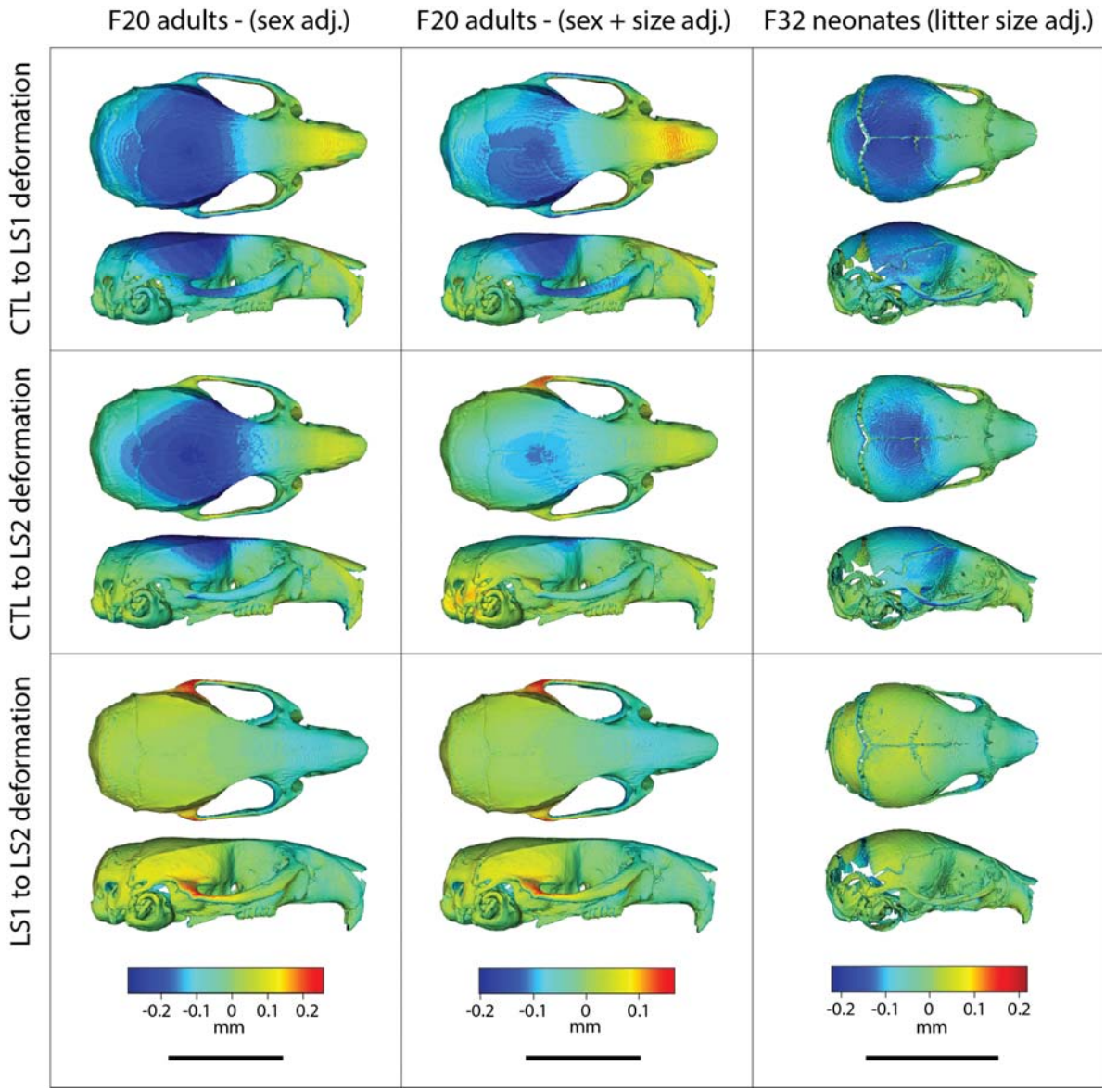


Figure 5 – Comparison of adult and neonate cranial phenotypes through shape change heatmaps. Heatmaps show the deformations required to transform between the mean shape of a given cohort to the mean shape of another. Blue indicates areas of relative reduction, red indicates areas of relative expansion, and green indicates neutral areas. Scale bar = 10 mm.

380 shape (Figure 6B). This suggests that a cellular change in the dorsal aspect of the ISS could be
381 driving cranial base flattening in the Longshanks juvenile skull, and hence potentially into
382 adulthood given the broadly similar shape changes observed at both stages. To qualitatively
383 validate the cellular changes inferred from our 2D GM analysis, we collected and imaged cranial
384 base sections near the sagittal midline. Since these developmental shape differences are still
385 subtle, we selected two representative extreme specimens that had large differences in ossified
386 tibia length, cranial base shape and cranium shape, yet comparable cranium centroid sizes so that
387 size would not confound our analysis. In agreement with our morphometric data, the SOS does
388 not differ qualitatively between these extreme specimens (Figure 6C). However, the ISS is
389 markedly larger in our Longshanks specimen compared to the Control, with larger resting and
390 proliferative zones that recapitulate the cellular differences characterized in the Longshanks
391 epiphysis (Figure 6C) (Marchini and Rolian, 2018). Crucially, the ISS is more wedge-shaped in
392 our Longshanks specimen at the cellular level, supporting the observed ISS changes at the
393 morphometric level (Figure 6B, C).
394

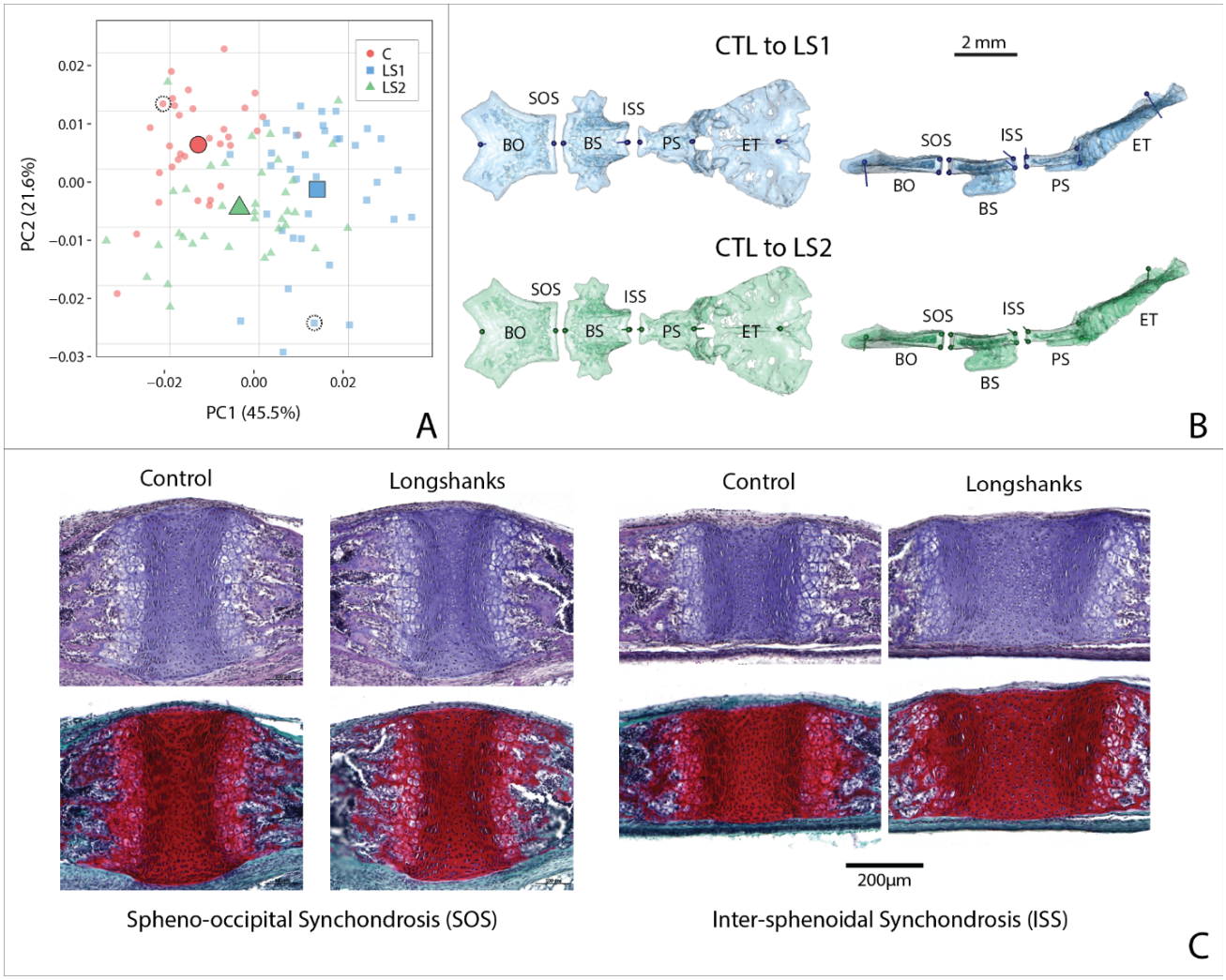


Figure 6 - Cranial base contribution to the Longshanks phenotype. (A) Scatter plot of the first two PCs of litter size-adjusted Longshanks and Control cranial base Procrustes shape variables in generation 32 (F32) neonates. Large symbols indicate mean PC1 and PC2 scores for each respective cohort. (B) Neonate cranial bases with vectors of shape change at midline cranial base landmarks (magnified 6 times for visualization) showing shape transformations to go from the mean Control cranial base to the mean LS1 (blue) and mean LS2 (green) cranial base shapes. Views in (B) are dorsal (left) and lateral (right). Abbreviations: basi-occipital bone (BO), spheno-occipital synchondrosis (SOS), basi-sphenoid bone (BS), intersphenoidal synchondrosis (ISS), presphenoid bone (PS) and ethmoid (ET). (C) Sagittal midline histological sections stained in H+E (top) and safranin-o (bottom) showing differences in synchondrosis morphology of two extreme specimens in CTL and LS1 of approximately equal centroid size (indicated by dashed circles in A).

395

Discussion

Phenotypic Changes

We investigated parallel skeletal size and shape changes in the limb and cranium in the selectively bred Longshanks mouse. Our morphometric analysis of adult cranium shape demonstrated that 20 generations of selection for longer tibiae relative to body mass are associated with the elongation of the cranium along the rostral caudal axis in Longshanks mice, independent of an overall increase in cranial size (Figure 1). In parallel, the cranium of adult Longshanks decreased in width between the zygomatic arches and reduced in vault height at the bregma and lambda (Figure 1). A similar cranium pattern was present in Longshanks neonates at P07, but with an ontogenically reduced craniofacial phenotype (Figures 4, 5). LS2 mice appear to have a more subtle phenotype than LS1 in adulthood and at P07 (Figures 1, 3, 5, 6A). This is not unexpected, as other genomic and phenotypic differences in the response to selection between LS1 and LS2 have been documented previously (Farooq *et al.,* 2017; Castro *et al.,* 2019; Cosman, Britz and Rolian, 2019).

The magnitudes of cranial shape change remain small in adult and neonate Longshanks relative to Controls, in comparison to stark morphological differences seen in skeletal mouse mutants (Munroe *et al.,* 2009; Gong, 2012; Holmes, 2012). However, in this study, we are more interested in patterns of shape change rather than magnitudes of change. Over 20 generations, the main target of selection, the tibia, increased in length by just 15%. As such, we expected that secondary cranial shape changes, while potentially significant in terms of long-term evolution, would be subtle. Moreover, selection in the tibia appears to have increased the variation in cranium measures, such as centroid size and cranial shape (Figures 3, S6). While the F20 LS1 and LS2 samples have new extreme cranium shapes not seen in earlier generations

(Supplementary Movie M2), their effects on the mean shape are dampened substantially by the fact that many F20 mice still have crania that resemble F01 random-bred mice.

The net adult cranium phenotypic change in Longshanks manifested as a sequence of two distinct evolutionary shape changes. The cranium of both selected lines consistently reduced in vault height in the first nine generations of tibia selection, and subsequently elongated and narrowed in the next eleven generations of selection (Figure 3). In comparison, the Control line shows minor cranial shape changes around the zygomatic and occipital regions, which are likely due to stochastic intergenerational variation in the Control line (i.e., drift), and/or to sampling artifacts, such as low sample size and family diversity in F09 Controls (Table 1). Similar evolutionary cranial shape changes have also been described in natural populations. Parmenter and colleagues noted that Gough Island mice, which differ significantly in body size compared to mainland relatives, have crania that are longer and narrower without differences in vault height, similar to the intergenerational changes in cranial shape between generations F09 and F20 in Longshanks (Parmenter *et al.,* 2016). The mechanisms underlying these stepwise changes are unknown, in part due to a lack of F09 developmental data. However, the transition in cranium response could be related to temporal differences in allele frequency shifts up to, and subsequent to, generation F09, and/or to the nature in which cranial bones physically interact (e.g., the degree of perturbation that must accumulate in one cranium region before adjacent structures are significantly affected).

Importantly, the cranium shape changes in adult LS1 and LS2 are uncorrelated with tibia length within each generation, despite a parallel shift in the tibia length and cranium shape population means over time (Figure 2). While we cannot say with certainty that there were no selective pressures acting on the cranium, selected individuals with relatively longer tibiae did not

necessarily have derived cranial shapes. This suggests that residual cranial shape did not confer a fitness advantage on breeders, and therefore the observed change in residual cranial shape over time in Longshanks is most likely a correlated evolutionary response to tibia selection, mediated by underlying genetic correlations due to pleiotropy and/or linkage disequilibrium (Lande, 1984).

Correlated evolutionary responses have been documented in experimental populations primarily in terms of physiological and behavioural processes, such as starvation response, developmental duration, and voluntary wheel running (Carter and Swallow, 2000; Gammie *et al.*, 2003; Davidowitz, Roff and Nijhout, 2016; Hardy *et al.*, 2018). Correlated responses in morphological phenotypes have also been characterized, but they have typically involved correlated change in the same anatomical structure, or indirect responses to body mass selection (Tobler and Nijhout, 2010). For example, selection for changes in wing spot pattern on the dorsal aspect of butterfly wings produced changes in ventral wing patterns (Beldade, Koops and Brakefield, 2002). To the best of our knowledge, our study presents the first strong case of correlated evolution of functionally unrelated and physically distant morphological structures in the mammalian skeletal system.

Developmental Correlates of Shape Change

From a developmental perspective, our investigation of Longshanks at one-week post partum revealed that Longshanks neonates have relatively flatter cranial bases, which develop by EO, compared to Controls (Figure 6). Moreover, 2D morphometrics at the midline and histology demonstrated that Longshanks neonates have a larger ISS with larger resting and proliferative

zones, especially in its dorsal aspect (Figure 6B, C), much like differences observed in the tibial proximal epiphysis of two-week old Longshanks (Marchini and Rolian, 2018). In contrast, the SOS did not show any substantive size or shape differences among groups. Taken together, these results suggest that the ISS responded to selection on the tibia independently of the SOS and may have contributed to the Longshanks cranial phenotype.

It is important to note that our ontogenetic analysis captured a single developmental stage, hence the apparent uncoupling of the synchondroses may be because the SOS has developmentally important differences between Longshanks and Controls, but at a different time in development. For example, the SOS mineralizes faster than the ISS after P10 in CD-1 mice (Wealthall and Herring, 2006). Regardless, differential timing of SOS and ISS fusion are widespread across clades, suggesting these growth centers are partly under independent control. In humans, the ISS begins fusing at birth, whereas the SOS does not fuse until adolescence (Madeline and Elster, 1995). In domestic dogs, premature fusion of the SOS is a prominent feature of brachycephalic dogs (Schmidt *et al.,* 2013). The partial developmental independence of the synchondroses is supported by knock-out studies of skeletogenic factors in rodents which have noted concerted changes in the postcranial epiphyses and synchondroses, but with differential responses in the respective synchondroses (reviewed in (Vora, 2017)). For example, in mice, Indian hedgehog ($Ihh^{-/-}$) knockout results in significantly more ectopic hypertrophic chondrocytes in the ISS than in the SOS (Young *et al.,* 2006). The intrinsic differences between the two synchondroses may ultimately be associated with their distinct embryonic origins: the pre-sphenoid and basi-sphenoid bones originate from neural crest cells that commit to endochondral ossification, whereas the basi-occiptal bone forms from mesenchymal condensations of prechordal mesoderm (McBratney-Owen *et al.,* 2008; Richtsmeier and Flaherty, 2013).

The exact ontogenetic mechanism by which the observed cranial base changes may lead to shape change across the entire cranium in Longshanks remains to be elucidated. There is extensive evidence for strong phenotypic covariation among cranial traits (i.e., morphological integration) causing indirect change throughout the cranium (Bookstein *et al.,* 2003; Goswami, 2006; Goswami *et al.,* 2012; Singh *et al.,* 2012; Bastir and Rosas, 2016; Neaux, 2016; Neaux *et al.,* 2019). For example, direct basicranium perturbations by genetic mutations in mice generated predictable shape changes throughout the cranium (Parsons *et al.,* 2015). Basicranium undergrowth by deletion of *Trsp* produces shortened faces and tall, domed calvaria. Conversely, cranial base overgrowth models, such as the *Pten* $^{-/-}$ mouse, resulted in flattened calvaria, elongated faces and reduced cranial width (Ford-Hutchinson *et al.,* 2007; Lieberman *et al.,* 2008; Parsons *et al.,* 2015). Moreover, analysis of mutant mouse models have demonstrated that cranial base length and angle are interrelated with brain and face size (Ross and Ravosa, 1993; Lieberman *et al.,* 2008). Notably, these transgenic models of basicranium perturbation also have concomitant tibia length changes (Downey *et al.*, 2009; Guntur *et al.*, 2011). While our study identified subtle evolutionary changes that occurred over time, rather than instant changes due to wholesale loss-of-function mutations, our findings are in general agreement with these cranium integration studies. Hence, correlated change among appendicular growth plates and synchondroses may be a common developmental mechanism underlying the observed overall limb and cranial shape changes, despite the physical distance and different ossification processes of the bones involved.

Potential Genetic Basis of Shape Change

Our data suggest that, at the tissue level, correlated changes in tibia length and cranial shape are associated with similar changes in growth plate morphology. At this time, the genomic basis for these correlated skeletal changes remains unknown. The vertebrate tibia and cranium are complex structures, each with a highly polygenic architecture, thus finding candidate genes that caused changes in one or the other structure, let alone correlated changes in both, is not trivial. A previous study revealed that the genomic response to tibia selection at generation F17 in Longshanks involved substantive contributions from several loci of large effect against an infinitesimal polygenic background (Castro *et al.*, 2019). By F17, LS1 and LS2 showed strong and parallel allele frequency changes in non-coding DNA in two large genomic regions spanning several topology associated domains (TADs) and containing putative cis-regulatory elements (Castro *et al.*, 2019). Several genes in or near these loci, such as *Nkx3-2* and *Chst11*, have known cranium and limb pleiotropic effects in knock-out mice and human diseases (Lettice *et al.*, 1999; Tribioli and Lufkin, 1999; Akazawa *et al.*, 2000; Klüppel *et al.*, 2005; Provot *et al.*, 2006; Chopra *et al.*, 2015). However, it is not clear whether these contributed to the overall F20 correlated cranium response, especially since allelic changes in non-coding regions are predicted to have smaller pleiotropic effects, if any, when compared to genomic changes in coding regions, which were rare at F17 (Carroll, 2008; Stern and Orgogozo, 2008; Rice and Rebeiz, 2019). Moreover, when cross-referenced against annotated mouse knockout phenotypes (Mouse Genome Informatics [MGI], http://www.informatics.jax.org), these genomic regions contain genes with singular effects on the cranium skeleton (e.g., *Tapt1 and Nfic*), hence it is also possible that the correlated response was driven in part by genetic linkage between limb- and cranium-specific genes swept along by selection. Interestingly, our query of MGI also shows that

over 15 of the genes near the altered F17 loci influence brain development and morphology, and therefore could indirectly influence the F20 net cranium shape due to brain-skull physical interactions. More work will be necessary to shed light on the underlying genetic correlates of the Longshanks cranium phenotype, especially functional genomic and molecular analyses of the cranial base at multiple developmental stages.

Limitations

The Longshanks mouse was produced under laboratory conditions, which likely do not represent the more complex multivariate selective pressures operating in nature. While the Longshanks experiment design allows us to assess skeletal change in a controlled environment without confounding variables (e.g., fluctuations in climate and/or food availability, presence of predators), it also produces an artificial context where only a single selection pressure acts on the skeleton. The simplified environment and atypically strong selection pressures in experimental evolution studies have the potential to target alleles with detrimental pleiotropic effects more frequently than in natural settings (Kawecki *et al.*, 2012). In this context, it is possible that correlated responses in the cranium to selection acting on the limb skeleton would be buffered in natural populations by counteracting selection pressures, related for example to mastication, olfaction, hearing and/or mate choice. In other words, the observed correlated responses in the cranium may have manifested in Longshanks simply because they were not under stabilizing selection for a specific cranial shape. For example, selected and random-bred mice in the Longshanks experiment were raised with a consistent diet that does not reflect the variable abundance of food types in nature. Selection for herbivory or increased bite force in rodents tends to favour wider crania with broader zygomatic arches, in which case an overlapping

selection pressure favouring individuals suited for herbivory would be in direct opposition with the correlated cranial changes we observed (Samuels, 2009; Cox *et al.*, 2012; Maestri *et al.*, 2016). Thus, while our study provides a robust example of correlated evolutionary changes in the vertebrate skeleton, it is possible that the observed limb and cranium changes are the manifestation of a path of least evolutionary resistance that only exists under controlled conditions in laboratory populations.

Conclusions

In this study, we characterized secondary skeletal responses to tibia selection that likely arose due to shared underlying developmental mechanisms between the cranium and tibia, specifically endochondral ossification. The limb and cranium are often considered separate modules in morphological analyses (Young and Hallgrímsson, 2005). Our results highlight the importance of considering evolution of the vertebrate skeleton in its entirety. Our study shows how indirect, and potentially non-adaptive, skeletal changes can occur due to developmental overlap among physically and functionally distant body parts. These findings have implications for how we reconstruct skeletal evolutionary histories of extant and extinct mammalian lineages by providing empirical evidence that skeletal traits may arise solely as side effects of selection acting elsewhere.

Acknowledgments

The authors are grateful to the Animal Resource Center staff at the University of Calgary for the continued care they provide to the Longshanks colonies. The authors are also indebted to Jason Anderson and Jessica Theodor for granting access to the SkyScan 1173 uCT scanner for data acquisition. The authors are also grateful to Dragana Ponjevic for expertise in histology and Heather Jamniczky for guidance in data analysis. This work was supported by a Natural Sciences and Engineering Research Council Canada Graduate Scholarships – Master's award to CMU, a Natural Sciences and Engineering Research Council Discovery Grant 4181932 to CR, and the Faculty of Veterinary Medicine at the University of Calgary (www.vet.ucalgary.ca).

Competing Interests

The authors declare no conflict of interest.

References

Adams, D., Collyer, M. and Kaliontzopoulou, A. (2020) 'Geomorph: Software for geometric morphometric analyses. R package version 3.2.1.' Available at: https://cran.r-project.org/package=geomorph.

Akazawa, H. *et al.* (2000) 'Targeted disruption of the homeobox transcription factor Bapx1 results in lethal skeletal dysplasia with asplenia and gastroduodenal malformation', *Genes to Cells*, 5(6), pp. 499–513. doi: 10.1046/j.1365-2443.2000.00339.x.

Armbruster, W. S. and Schwaegerle, K. E. (1996) 'Causes of covariation of phenotypic traits

among populations', *Journal of Evolutionary Biology*, 9(3), pp. 261–276. doi: 10.1046/j.1420-9101.1996.9030261.x.

Avants, B. B. *et al.* (2011) 'A reproducible evaluation of ANTs similarity metric performance in brain image registration', *NeuroImage*, 54(3), pp. 2033–2044. doi: 10.1016/j.neuroimage.2010.09.025.

Bastir, M. and Rosas, A. (2006) 'Correlated variation between the lateral basicranium and the face: A geometric morphometric study in different human groups', *Archives of Oral Biology*, 51(9), pp. 814–824. doi: 10.1016/j.archoralbio.2006.03.009.

Bastir, M. and Rosas, A. (2016) 'Cranial base topology and basic trends in the facial evolution of Homo', *Journal of Human Evolution*, 91, pp. 26–35. doi: 10.1016/j.jhevol.2015.11.001.

De Beer, G. (1937) *The development of the vertebrate skull*. Oxford: The Clarendon Press.

Beldade, P., Koops, K. and Brakefield, P. M. (2002) 'Developmental constraints versus flexibility in morphological evolution', *Nature*, 416(6883), pp. 844–847. doi: 10.1038/416844a.

Bookstein, F. L. *et al.* (2003) 'Cranial integration in Homo: Singular warps analysis of the midsagittal plane in ontogeny and evolution', *Journal of Human Evolution*, 44(2), pp. 167–187. doi: 10.1016/S0047-2484(02)00201-4.

Carroll, S. B. (2008) 'Evo-Devo and an Expanding Evolutionary Synthesis: A Genetic Theory of Morphological Evolution', *Cell*. Cell Press, pp. 25–36. doi: 10.1016/j.cell.2008.06.030.

Carter, P. A. and Swallow, J. G. (2000) 'Nesting behavior of house mice (Mus domesticus) selected for increased wheel-running activity', *Behavior Genetics*, 30(2), pp. 85–94. doi: 10.1023/A:1001967019229.

Castro, J. P. *et al.* (2019) 'An integrative genomic analysis of the Longshanks selection experiment for longer limbs in mice', *eLife*, 8. doi: 10.7554/elife.42014.

Cheverud, J. M. (1996) 'Developmental integration and the evolution of pleiotropyl', *American Zoologist*, 36(1), pp. 44–50. doi: 10.1093/icb/36.1.44.

Chopra, S. S. *et al.* (2015) 'Inherited chst11/mir3922 deletion is associated with a novel recessive syndrome presenting with skeletal malformation and malignant lymphoproliferative disease', *Molecular Genetics and Genomic Medicine*, 3(5), pp. 413–423. doi: 10.1002/mgg3.152.

Collyer, M. L. and Adams, D. C. (2018) 'RRPP: An R package for fitting linear models to high-dimensional data using residual randomization', *Methods in Ecology and Evolution*. Edited by R. Freckleton, 9(7), pp. 1772–1779. doi: 10.1111/2041-210X.13029.

Collyer, M. L., Sekora, D. J. and Adams, D. C. (2015) 'A method for analysis of phenotypic change for phenotypes described by high-dimensional data', *Heredity*, 115(4), pp. 357–365. doi: 10.1038/hdy.2014.75.

Cosman, M. N., Britz, H. M. and Rolian, C. (2019) 'Selection for longer limbs in mice increases bone stiffness and brittleness, but does not alter bending strength', *Journal of Experimental Biology*, 222(9). doi: 10.1242/jeb.203125.

Cosman, M. N., Sparrow, L. M. and Rolian, C. (2016) 'Changes in shape and cross-sectional geometry in the tibia of mice selectively bred for increases in relative bone length', *Journal of Anatomy*, 228(6), pp. 940–951. doi: 10.1111/joa.12459.

Cox, P. G. *et al.* (2012) 'Functional Evolution of the Feeding System in Rodents', *PLoS ONE*. Edited by A. Goswami, 7(4), p. e36299. doi: 10.1371/journal.pone.0036299.

von Cramon-Taubadel, N. (2019) 'Multivariate morphometrics, quantitative genetics, and neutral theory: Developing a "modern synthesis" for primate evolutionary morphology', *Evolutionary Anthropology: Issues, News, and Reviews*, 28(1), pp. 21–33. doi: 10.1002/evan.21761.

Davidowitz, G., Roff, D. and Nijhout, H. F. (2016) 'Synergism and Antagonism of Proximate Mechanisms Enable and Constrain the Response to Simultaneous Selection on Body Size and Development Time: An Empirical Test Using Experimental Evolution', *The American Naturalist*, 188(5), pp. 499–520. doi: 10.1086/688653.

Devine, J. *et al.* (2020) 'A Registration and Deep Learning Approach to Automated Landmark Detection for Geometric Morphometrics', *Evolutionary Biology*, 47(3), pp. 246–259. doi: 10.1007/s11692-020-09508-8.

Downey, C. M. *et al.* (2009) 'Osteo-chondroprogenitor-specific deletion of the selenocysteine tRNA gene, Trsp, leads to chondronecrosis and abnormal skeletal development: A putative model for Kashin-Beck disease', *PLoS Genetics*, 5(8). doi: 10.1371/journal.pgen.1000616.

Farooq, S. *et al.* (2017) 'Cortical and trabecular morphology is altered in the limb bones of mice artificially selected for faster skeletal growth', *Scientific Reports*, 7(1), pp. 10527–10527. doi: 10.1038/s41598-017-10317-x.

Ford-Hutchinson, A. F. *et al.* (2007) 'Inactivation of Pten in Osteo-Chondroprogenitor Cells Leads to Epiphyseal Growth Plate Abnormalities and Skeletal Overgrowth', *Journal of Bone and Mineral Research*, 22(8), pp. 1245–1259. doi: 10.1359/jbmr.070420.

Fruciano, C. (2016) 'Measurement error in geometric morphometrics', *Development Genes and Evolution*. Springer Verlag, pp. 139–158. doi: 10.1007/s00427-016-0537-4.

Gammie, S. C. *et al.* (2003) 'Predatory aggression, but not maternal or intermale aggression, is associated with high voluntary wheel-running behavior in mice', in *Hormones and Behavior*. Academic Press Inc., pp. 209–221. doi: 10.1016/S0018-506X(03)00140-5.

Gong, S. G. (2012) 'The Fgfr2 W290R mouse model of Crouzon syndrome', *Child's Nervous System*, 28(9), pp. 1495–1503. doi: 10.1007/s00381-012-1792-y.

Goodall, C. (1991) 'Procrustes Methods in the Statistical Analysis of Shape', *Journal of the Royal Statistical Society: Series B (Methodological)*, 53(2), pp. 285–321. doi: 10.1111/j.2517-6161.1991.tb01825.x.

Goswami, A. (2006) 'Morphological Integration in the Carnivoran Skull', *Evolution*, 60(1), pp. 169–183. doi: 10.1111/j.0014-3820.2006.tb01091.x.

Goswami, A. *et al.* (2012) 'Shape, variance and integration during craniogenesis: contrasting marsupial and placental mammals', *Journal of Evolutionary Biology*, 25(5), pp. 862–872. doi: 10.1111/j.1420-9101.2012.02477.x.

Gould, S. J. and Lewontin, R. C. (1979) 'The spandrels of San Marco and the Panglossian paradigm: a critique of the adaptationist programme', *Proceedings of the Royal Society of London. Series B. Biological Sciences*, 205(1161), pp. 581–598. doi: 10.1098/rspb.1979.0086.

Guntur, A. R. *et al.* (2011) 'Conditional ablation of Pten in osteoprogenitors stimulates FGF signaling', *Development*, 138(7), pp. 1433–1444. doi: 10.1242/dev.058016.

Hallgrímsson, B. *et al.* (2009) 'Deciphering the palimpsest: Studying the relationship between morphological integration and phenotypic covariation', *Evolutionary Biology*, 36(4), pp. 355–376. doi: 10.1007/s11692-009-9076-5.

Hallgrímsson, B. and Hall, B. K. (2005) ‘Variation and variability: Central concepts in biology’, in *Variation*. Elsevier Inc., pp. 1–7. doi: 10.1016/B978-012088777-4/50003-X.

Hallgrímsson, B. and Lieberman, D. E. (2008) ‘Mouse models and the evolutionary developmental biology of the skull’, *Integrative and Comparative Biology*, 48(3), pp. 373–384. doi: 10.1093/icb/icn076.

Hardy, C. M. *et al.* (2018) ‘Genome-Wide Analysis of Starvation-Selected Drosophila melanogaster—A Genetic Model of Obesity’, *Molecular Biology and Evolution*, 35(1), pp. 50–65. doi: 10.1093/molbev/msx254.

Hendrikse, J. L., Parsons, T. E. and Hallgrímsson, B. (2007) ‘Evolvability as the proper focus of evolutionary developmental biology’, *Evolution and Development*, 9(4), pp. 393–401. doi: 10.1111/j.1525-142X.2007.00176.x.

Holmes, G. (2012) ‘Mouse models of Apert syndrome’, *Child’s Nervous System*, 28(9), pp. 1505–1510. doi: 10.1007/s00381-012-1872-z.

Karp, N. A. *et al.* (2017) ‘Prevalence of sexual dimorphism in mammalian phenotypic traits’, *Nature Communications*, 8(1), p. 21. doi: 10.1038/ncomms15475.

Kawecki, T. J. *et al.* (2012) ‘Experimental evolution’, *Trends in Ecology and Evolution*. Elsevier, pp. 547–560. doi: 10.1016/j.tree.2012.06.001.

Klingenberg, C. P. (2016) ‘Size, shape, and form: concepts of allometry in geometric morphometrics’, *Development Genes and Evolution*. Springer Verlag, pp. 113–137. doi: 10.1007/s00427-016-0539-2.

Klüppel, M. *et al.* (2005) ‘Maintenance of chondroitin sulfation balance by chondroitin-4-

sulfotransferase 1 is required for chondrocyte development and growth factor signaling during cartilage morphogenesis', *Development*, 132(17), pp. 3989–4003. doi: 10.1242/dev.01948.

Kronenberg, H. M. (2003) 'Developmental regulation of the growth plate', *Nature*. Nature, pp. 332–336. doi: 10.1038/nature01657.

Lande, R. (1984) 'The genetic correlation between characters maintained by selection, linkage and inbreeding', *Genetical Research*, 44(3), pp. 309–320. doi: 10.1017/S0016672300026549.

Lefebvre, V. and Bhattaram, P. (2010) 'Vertebrate skeletogenesis', in *Current Topics in Developmental Biology*. Academic Press Inc., pp. 291–317. doi: 10.1016/S0070-2153(10)90008-2.

Lerch, J. P., Sled, J. G. and Henkelman, R. M. (2010) 'MRI phenotyping of genetically altered mice', *Methods in Molecular Biology*, 711, pp. 349–361. doi: 10.1007/978-1-61737-992-5_17.

Lettice, L. A. *et al.* (1999) 'The mouse bagpipe gene controls development of axial skeleton, skull, and spleen', *Proceedings of the National Academy of Sciences of the United States of America*, 96(17), pp. 9695–9700. doi: 10.1073/pnas.96.17.9695.

Lieberman, D. E. *et al.* (2008) 'Spatial packing, cranial base angulation, and craniofacial shape variation in the mammalian skull: Testing a new model using mice', *Journal of Anatomy*, 212(6), pp. 720–735. doi: 10.1111/j.1469-7580.2008.00900.x.

Mackie, E. J. *et al.* (2008) 'Endochondral ossification: How cartilage is converted into bone in the developing skeleton', *International Journal of Biochemistry and Cell Biology*. Int J Biochem Cell Biol, pp. 46–62. doi: 10.1016/j.biocel.2007.06.009.

Madeline, L. A. and Elster, A. D. (1995) 'Suture closure in the human chondrocranium: CT

assessment', *Radiology*, 196(3), pp. 747–756. doi: 10.1148/radiology.196.3.7644639.

Maestri, R. *et al.* (2016) 'Diet, bite force and skull morphology in the generalist rodent morphotype', *Journal of Evolutionary Biology*, 29(11), pp. 2191–2204. doi: 10.1111/jeb.12937.

Marchini, M. *et al.* (2014) 'Impacts of genetic correlation on the independent evolution of body mass and skeletal size in mammals', *BMC Evolutionary Biology*, 14(1), p. 258. doi: 10.1186/s12862-014-0258-0.

Marchini, M. and Rolian, C. (2018) 'Artificial selection sheds light on developmental mechanisms of limb elongation', *Evolution*, 72(4), pp. 825–837. doi: 10.1111/evo.13447.

McBratney-Owen, B. *et al.* (2008) 'Development and tissue origins of the mammalian cranial base', *Developmental Biology*, 322(1), pp. 121–132. doi: 10.1016/j.ydbio.2008.07.016.

Moss, M. L. (1977) 'A functional analysis of fusion of the tibia and fibula in the rat and mouse', *Cells Tissues Organs*, 97(3), pp. 321–332. doi: 10.1159/000144749.

Munroe, R. J. *et al.* (2009) 'Mouse H6 Homeobox 1 (Hmx1) mutations cause cranial abnormalities and reduced body mass', *BMC Developmental Biology*, 9(1), p. 27. doi: 10.1186/1471-213X-9-27.

Murren, C. J. (2012) 'The integrated phenotype', in *Integrative and Comparative Biology*. Integr Comp Biol, pp. 64–76. doi: 10.1093/icb/ics043.

Neaux, D. (2016) 'DomExp: Experimental domestication and skeleton development in captivity View project', *Article in American Journal of Physical Anthropology*. doi: 10.1002/ajpa.23163.

Neaux, D. *et al.* (2019) 'Morphological integration affects the evolution of midline cranial base, lateral basicranium, and face across primates', *American Journal of Physical Anthropology*,

170(1), pp. 37–47. doi: 10.1002/ajpa.23899.

Parmenter, M. D. *et al.* (2016) 'Genetics of skeletal evolution in unusually large mice from gough island', *Genetics*, 204(4), pp. 1559–1572. doi: 10.1534/genetics.116.193805.

Parsons, T. E. *et al.* (2015) 'Mind the Gap: Genetic Manipulation of Basicranial Growth within Synchondroses Modulates Calvarial and Facial Shape in Mice through Epigenetic Interactions', *PLOS ONE*. Edited by N. Jeffery, 10(2), p. e0118355. doi: 10.1371/journal.pone.0118355.

Percival, C. J. *et al.* (2019) 'The effect of automated landmark identification on morphometric analyses', *Journal of Anatomy*, 234(6), pp. 917–935. doi: 10.1111/joa.12973.

Price, T. and Langen, T. (1992) 'Evolution of correlated characters', *Trends in Ecology and Evolution*. Trends Ecol Evol, pp. 307–310. doi: 10.1016/0169-5347(92)90229-5.

Provot, S. *et al.* (2006) 'Nkx3.2/Bapx1 acts as a negative regulator of chondrocyte maturation', *Development*, 133(4), pp. 651–662. doi: 10.1242/dev.02258.

R Core Team (2020) 'R: A Language and environment for statistical computing (R Foundation for Statistical Computing).'

Rice, G. and Rebeiz, M. (2019) 'Evolution: How Many Phenotypes Do Regulatory Mutations Affect?', *Current Biology*. Cell Press, pp. R21–R23. doi: 10.1016/j.cub.2018.11.027.

Richtsmeier, J. T. *et al.* (2006) 'Phenotypic integration of neurocranium and brain', *Journal of Experimental Zoology Part B: Molecular and Developmental Evolution*, 306(4), pp. 360–378. doi: 10.1002/jez.b.21092.

Richtsmeier, J. T. and Flaherty, K. (2013) 'Hand in glove: Brain and skull in development and dysmorphogenesis', *Acta Neuropathologica*. Acta Neuropathol, pp. 469–489. doi:

10.1007/s00401-013-1104-y.

Riska, B. (1986) ‘Some models for development, growth, and morphometric correlation.’, *Evolution*, 40(6), pp. 1303–1311. doi: 10.1111/j.1558-5646.1986.tb05753.x.

Roselló-Díez, A. and Joyner, A. L. (2015) ‘Regulation of long bone growth in vertebrates; It is time to catch up’, *Endocrine Reviews*. Endocrine Society, pp. 646–680. doi: 10.1210/er.2015-1048.

Ross, C. F. and Ravosa, M. J. (1993) ‘Basicranial flexion, relative brain size, and facial kyphosis in nonhuman primates’, *American Journal of Physical Anthropology*, 91(3), pp. 305–324. doi: 10.1002/ajpa.1330910306.

Samuels, J. X. (2009) ‘Cranial morphology and dietary habits of rodents’, *Zoological Journal of the Linnean Society*, 156(4), pp. 864–888.

Schlager, S. (2017) ‘Morpho and Rvcg - Shape Analysis in R: R-Packages for Geometric Morphometrics, Shape Analysis and Surface Manipulations’, in *Statistical Shape and Deformation Analysis: Methods, Implementation and Applications*. Elsevier Inc., pp. 217–256. doi: 10.1016/B978-0-12-810493-4.00011-0.

Schlager, S. (2020) ‘Morpho: calculations and visualizations related to geometric morphometrics. R package version 2.8.’ Available at: https://rdrr.io/cran/Morpho/.

Schmidt, M. J. *et al.* (2013) ‘COMPARISON OF CLOSURE TIMES FOR CRANIAL BASE SYNCHONDROSES IN MESATICEPHALIC, BRACHYCEPHALIC, AND CAVALIER KING CHARLES SPANIEL DOGS’, *Veterinary Radiology & Ultrasound*, 54(5), pp. 497–503. doi: 10.1111/vru.12072.

Scott, J. H. (1958) ‘The cranial base’, *American Journal of Physical Anthropology*, 16(3), pp. 319–348. doi: 10.1002/ajpa.1330160305.

Singh, N. *et al.* (2012) ‘Morphological evolution through integration: A quantitative study of cranial integration in Homo, Pan, Gorilla and Pongo’, *Journal of Human Evolution*, 62(1), pp. 155–164. doi: 10.1016/j.jhevol.2011.11.006.

Sparrow, L. M. *et al.* (2017) ‘Gait changes in a line of mice artificially selected for longer limbs’, *PeerJ*, 2017(2). doi: 10.7717/peerj.3008.

Stern, D. L. and Orgogozo, V. (2008) ‘THE LOCI OF EVOLUTION: HOW PREDICTABLE IS GENETIC EVOLUTION?’, *Evolution*, 62(9), pp. 2155–2177. doi: 10.1111/j.1558-5646.2008.00450.x.

Tobler, A. and Nijhout, H. F. (2010) ‘Developmental constraints on the evolution of wing-body allometry in Manduca sexta’, *Evolution & Development*, 12(6), pp. 592–600. doi: 10.1111/j.1525-142X.2010.00444.x.

Tribioli, C. and Lufkin, T. (1999) ‘The murine Bapx1 homeobox gene plays a critical role in embryonic development of the axial skeleton and spleen’, *Development*, 126(24), pp. 5699–711. Available at: https://pubmed.ncbi.nlm.nih.gov/10572046/ (Accessed: 10 February 2021).

Vora, S. R. (2017) ‘Mouse models for the study of cranial base growth and anomalies’, *Orthodontics & Craniofacial Research*, 20, pp. 18–25. doi: 10.1111/ocr.12180.

Vora, S. R., Camci, E. D. and Cox, T. C. (2016) ‘Postnatal Ontogeny of the Cranial Base and Craniofacial Skeleton in Male C57BL/6J Mice: A Reference Standard for Quantitative Analysis’, *Frontiers in Physiology*, 6(JAN), p. 417. doi: 10.3389/fphys.2015.00417.

Wagner, G. P. (1984) ‘Coevolution of functionally constrained characters: Prerequisites for adaptive versatility’, *BioSystems*, 17(1), pp. 51–55. doi: 10.1016/0303-2647(84)90015-7.

Wagner, G. P., Pavlicev, M. and Cheverud, J. M. (2007) ‘The road to modularity’, *Nature Reviews Genetics*. Nat Rev Genet, pp. 921–931. doi: 10.1038/nrg2267.

Waters, M. J. and Kaye, P. L. (2002) ‘The role of growth hormone in fetal development’, *Growth Hormone and IGF Research*. Churchill Livingstone, pp. 137–146. doi: 10.1016/S1096-6374(02)00018-7.

Wealthall, R. J. and Herring, S. W. (2006) ‘Endochondral ossification of the mouse nasal septum’, *The Anatomical Record Part A: Discoveries in Molecular, Cellular, and Evolutionary Biology*, 288A(11), pp. 1163–1172. doi: 10.1002/ar.a.20385.

Wei, X. *et al.* (2016) ‘Developmental regulation of the growth plate and cranial synchondrosis’, *Journal of Dental Research*, 95(11), pp. 1221–1229. doi: 10.1177/0022034516651823.

White, A. and Wallis, G. (2001) ‘Endochondral ossification: A delicate balance between growth and mineralisation’, *Current Biology*. Cell Press. doi: 10.1016/S0960-9822(01)00359-1.

Young, B. *et al.* (2006) ‘Indian and sonic hedgehogs regulate synchondrosis growth plate and cranial base development and function’, *Developmental Biology*, 299(1), pp. 272–282. doi: 10.1016/j.ydbio.2006.07.028.

Young, N. M. and Hallgrímsson, B. (2005) ‘Serial Homology and The Evolution of Mammalian Limb Covariation Structure’, *Evolution*, 59(12), pp. 2691–2704. doi: 10.1111/j.0014-3820.2005.tb00980.x.

829 **Table 1** – Longshanks adult (F01, F09, F20) and neonate (F32) sample composition. Target
830 sample sizes of n = 40/group represent a compromise between sampling effort (e.g., scanning
831 capacity) and the ability to detect small-to-moderate effect sizes at a power of 0.8 in the case of
832 univariate analyses (Cohen's d = 0.25). For multivariate analyses (e.g., PCA), target sample sizes
833 of n = 40 with 50-68 landmarks datasets produce highly repeatable covariance matrices (average
834 RV coefficient of a sample's GPA-adjusted covariance matrix with 1000 covariance matrices
835 derived from bootstrapped data > 0.99).

Longshanks Samples (n)			
	CTL	LS1	LS2
Generation 1 (F01)	24	40	40
Generation 9 (F09)	23	40	40
Generation 20 (F20)	40	40	40
Generation 32 (F32)	32	36	36

836

837

838

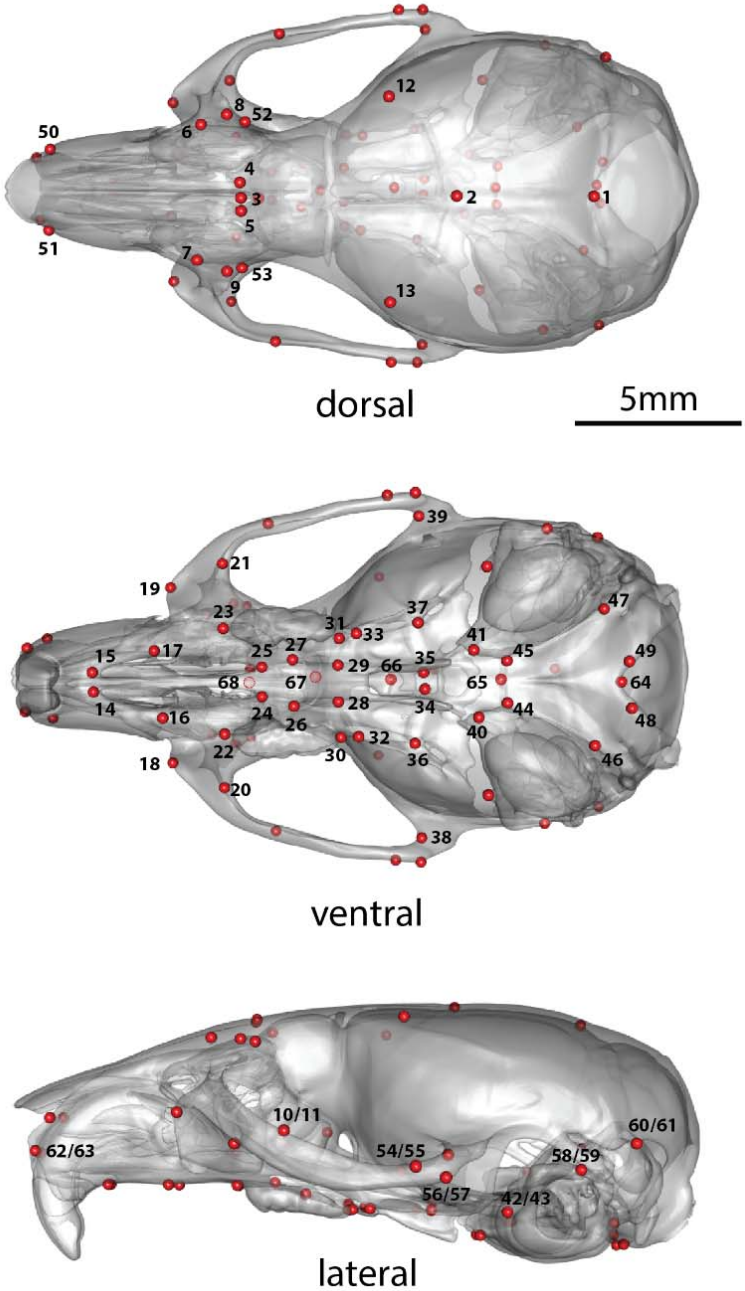
839

840

841

842

843



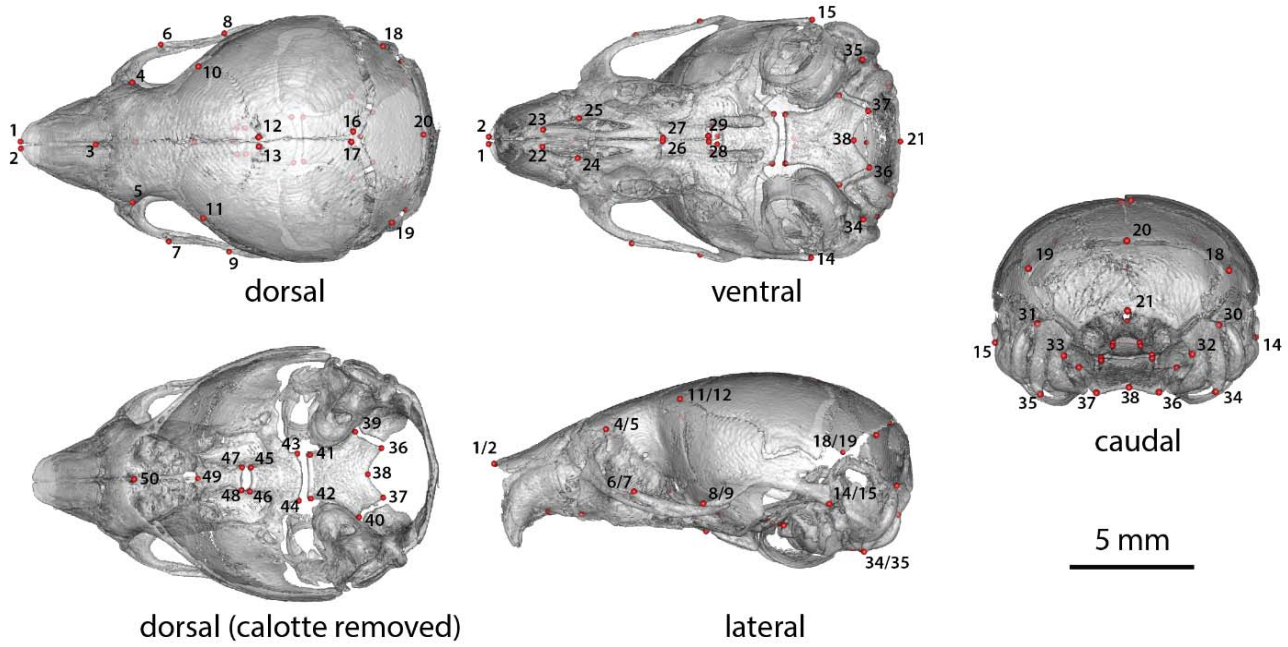
844

845

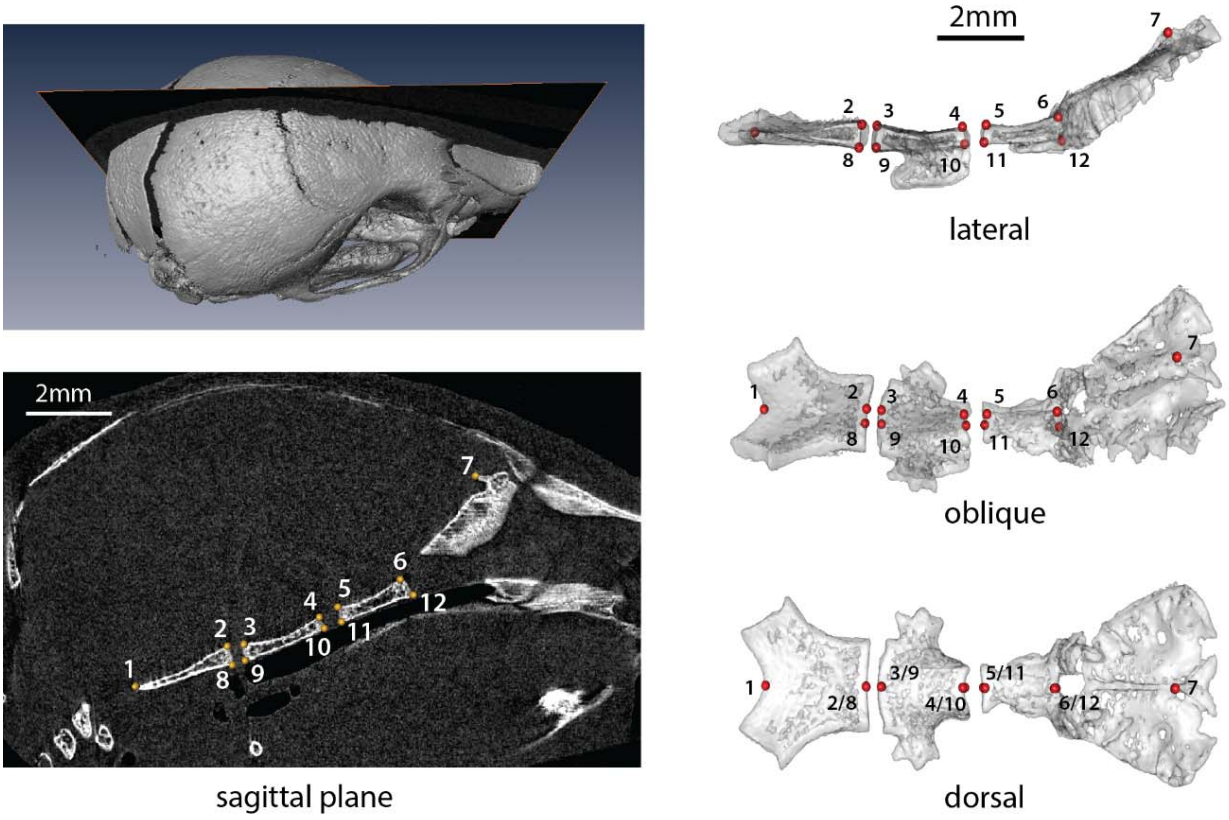
846 **Supplemental Figure 1** – Adult cranium landmarks used in this study in lateral, dorsal and

847 ventral landmark views. For landmark anatomical definitions, see Supplementary Table 1.

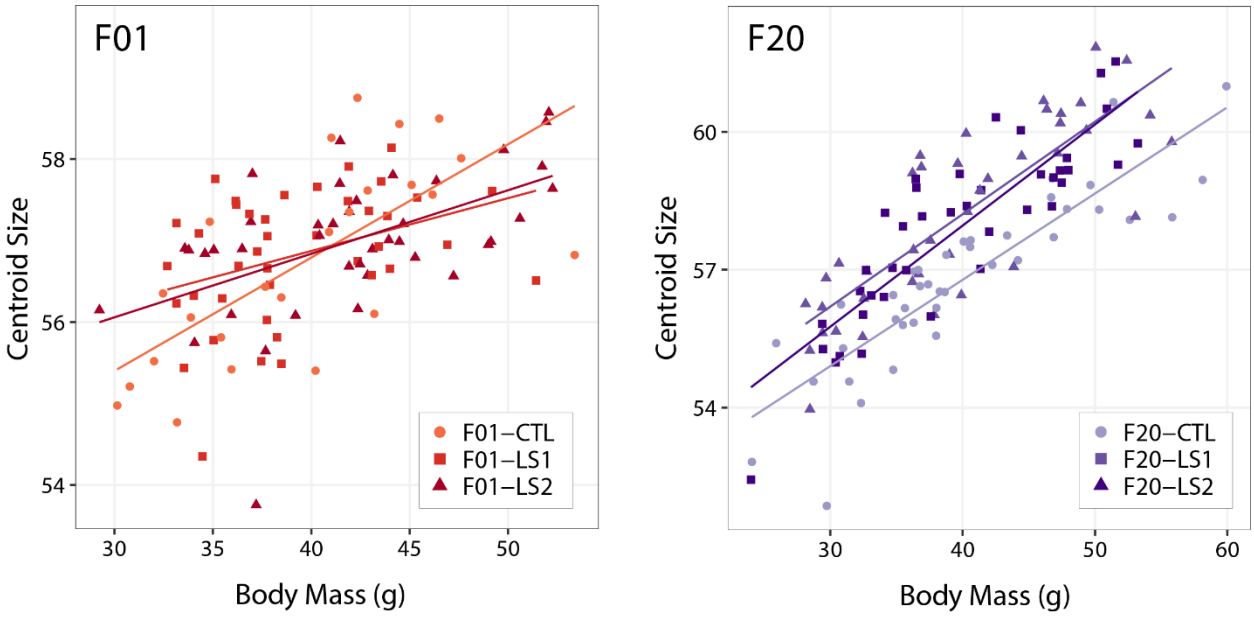
848



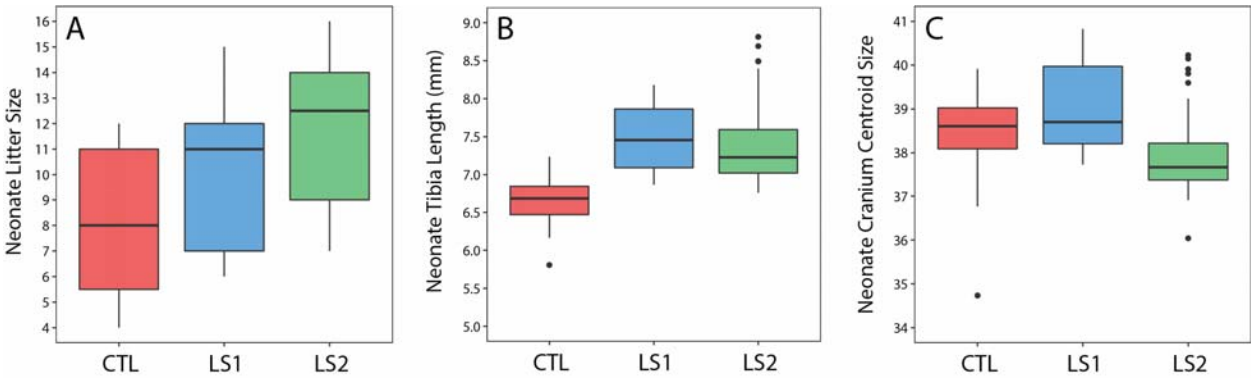
Supplemental Figure 2 – Neonate (P07) cranium landmarks used in this study in lateral, dorsal, dorsal cranial base, caudal and ventral landmark views. For landmark anatomical definitions, see Supplementary Table 2.



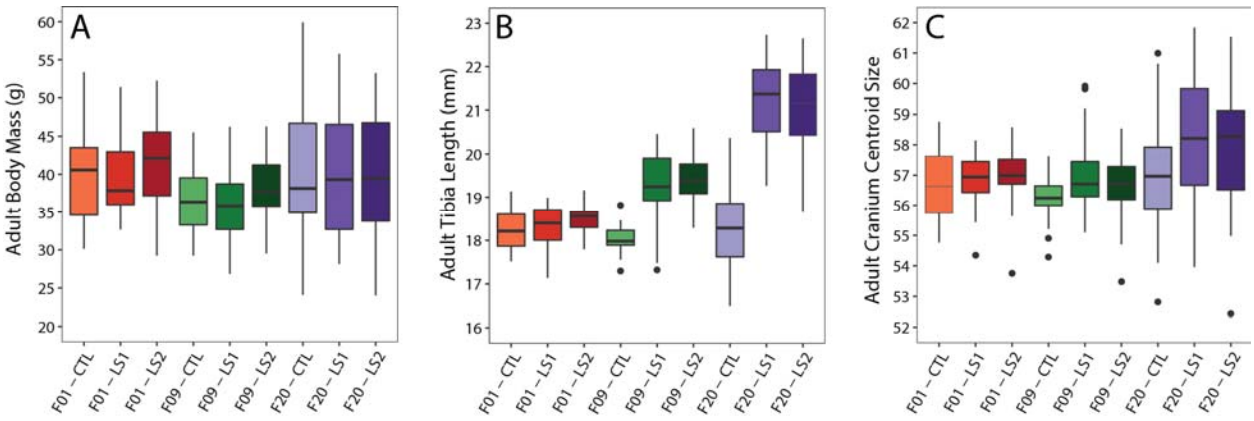
Supplemental Figure 3 – Neonate (P07) cranial base landmarks on the sagittal midline used in this study. Numbered landmarks applied to CT scan reconstruction slices at the midline (left) and landmarks numbered on a 3D cranial base mesh (right). For landmark anatomical definitions, see Supplementary Table 3.



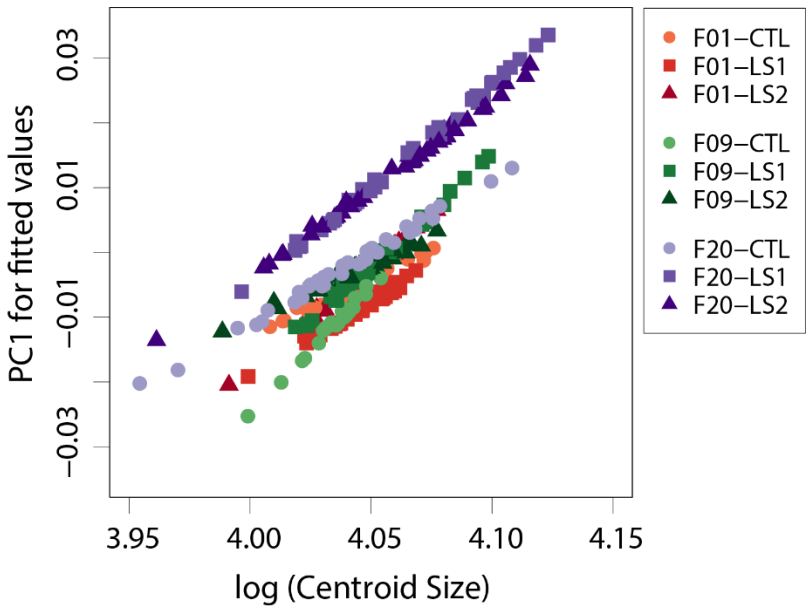
Supplemental Figure 4 – Scatter plots with regression lines by group showing the relationships between body mass and cranium size (centroid size) in adult founder mice (F01) and after 20 generations of selection (F20).



Supplemental Figure 5 - Boxplots showing differences in neonate Longshanks and Control metrics. (A) Boxplot of neonate litter sizes. (B and C) Boxplots of litter size adjusted tibia length and cranium centroid size.



Supplemental Figure 6 - Boxplots showing differences in adult Longshanks and Control metrics. (A-C) Boxplots of adult body mass, tibia length and cranial centroid size between groups.



Supplemental Figure 7 – Scatter plot of fitted PC1 scores (shape scores predicted by regression of shape on size) vs log (centroid size) showing within group patterns of cranium allometry. At any given cranium size, LS1 and LS2 are predicted to have positive shape scores (longer and narrower) shapes.

909 Table S1 – Adult cranium landmarks and their anatomical definitions.

Adult Cranium Landmarks	
Paired Landmarks (R/L)	
Lateral point on frontal suture	4/5
Lateral zygomatic-frontal suture	6/7
Posterior zygomaticofrontal junction	8/9
Posterior margin of malar process	10/11
Frontal-temporal-parietal junction	12/13
Anterior margin of incisive foramen	14/15
Medial maxilla-premaxilla junction	16/17
Anterior inferior zygomatic	18/19
Anterior temporo-zygomatic junction	20/21
Anterior superior alveoli	22/23
Posterior incisive foramen	24/25
Point along palatine-maxillary suture	26/27
Medial palatal-pterygoid junction	28/29
Posterior superior alveoli	30/31
Lateral palatal-pterygoid junction	32/33
Spheno-occipital synchondrosis	34/35
Anterior foramen ovale	36/37
Posterior temporo-zygomatic junction	38/39
Auditory-temporal-sphenoid junction	40/41
Anterior inferior auditory bulla	42/43
Occipital-auditory-sphenoid junction	44/45
Point along occipitomastoid suture	46/47
Medial occipital condyle	48/49
Anterior nasal and premaxilla	50/51
Frontal suture on orbital rim	52/53
Superior temporo-zygomatic suture	54/55
Posterior zygomatic process	56/57
Superio-posterior tympanic ring	58/59

Occipital-auditory junction	60/61
Midline superior incisor	62/63
Midline Landmarks	
Lambda	1
Bregma	2
Nasion	3
Anterior foramen magnum	64
Midline junction basioccipital and sphenoid	65
Midline junction sphenoid and presphenoid	66
Anterior junction endocranial presphenoid	67
Endocranial junction frontal and ethmoid	68

910

911 Table S2 – Neonate cranium landmarks and their anatomical definitions.

Neonate Cranium Landmarks	
Paired Landmarks (L/R)	
Medial rostral tip of nasal bones	1/2
Caudal dorsal fronto-zygomatic suture on frontal bone	4/5
Rostral dorsal zygomatic suture on zygomatic bone	6/7
Caudal dorsal zygomatic suture on zygomatic process of temporal bone	8/9
Temporal-parietal-frontal suture	10/11
Caudal medial tip of frontal bones (bregma)	12/13
Ventral caudal tip of occipital process of temporal bone	14/15
Caudal medial tip of the parietal	16/17
Caudal lateral tip of interparietal	18/19
Rostral incisor foramen	22/23
Premaxilla-maxilla suture (rostral medial maxilla)	24/25
Maxilla-palatine suture (rostral medial palatine)	26/27
Caudal medial tip of palatine	28/29
Dorsal lateral tip of exoccipital	30/31
Medial border of exoccipital at widest mediolateral span of foramen magnum	32/33
Ventral tip of paraoccipital process	34/35
Caudal medial tip of basioccipital at intra-occipital synchondrosis	36/37

Caudal lateral tip of basioccipital at intra-occipital synchondrosis	39/40
Rostral lateral tip of basioccipital at spheno-occipital synchondrosis	41/42
Caudal lateral tip of sphenoid at spheno-occipital synchondrosis	43/44
Rostral lateral tip of sphenoid at inter-sphenoidal synchondrosis	45/46
Caudal lateral tip of presphenoid at inter-sphenoidal synchondrosis	47/48
Midline Landmarks	
Caudal medial nasal bones (nasion)	3
Caudal medial border of interparietal	20
Ventral medial occipital (dorsal foramen magnum)	21
Caudal medial border of basioccipital (rostral foramen magnum)	38
Ethmoid-presphenoid suture	49
Rostral medial border of cribriform plate	50

912

913

914

915

916

917

918

919

920

921 Table S3 – Neonate cranial base landmarks and their anatomical definitions.

Sagittal Midline Neonate Cranial Base Landmarks	
Basion	1
Rostral dorsal tip of basi-occipital at spheno-occipital synchondrosis	2
Caudal dorsal tip of basi-sphenoid at spheno-occipital synchondrosis	3
Rostral dorsal tip of basi-sphenoid at intersphenoidal synchondrosis	4
Caudal dorsal tip of presphenoid at intersphenoidal synchondrosis	5
Ethmoid-presphenoid suture	6
Rostral medial border of cribriform plate	7
Rostral ventral tip of basi-occipital at spheno-occipital synchondrosis	8
Caudal ventral tip of basi-sphenoid at spheno-occipital synchondrosis	9
Rostral ventral tip of basi-sphenoid at intersphenoidal synchondrosis	10
Caudal ventral tip of presphenoid at intersphenoidal synchondrosis	11
Rostral ventral tip of the presphenoid	12

922

923

924

925

926

927

928

929

930 Table S4 – Morphometric data for neonate mice among lines and generations. Tibia and centroid size data
931 represent least squared means (SEM) and litter sizes are means (SEM). Differences in litter size were
932 determined by ANOVA, whereas tibia length and centroid size differences were assessed by ANCOVA
933 with litter size as a covariate. Superscripts denote significant differences in means ($p < 0.05$) between a
934 given group and: Controls ^{CTL}, Longshanks Line 1 ^{LS1}, Longshanks Line 2 ^{LS2}.

Neonate Morphometric Data			
Line	CTL	LS1	LS2
Ossified Tibial Diaphysis Length (mm)	6.66 (0.06) $^{LS1, LS2}$	7.50 (0.07) ^{CTL}	7.41 (0.10) ^{CTL}
Litter Size	7.94 (0.55) $^{LS1, LS2}$	10.33 (0.51) ^{CTL}	11.83 (0.51) ^{CTL}
Centroid Size	38.49 (0.17) ^{LS1}	39.05 (0.16) $^{CTL, LS2}$	37.94 (0.16) ^{LS1}

935
936
937
938
939
940
941
942
943
944

Table S5 – Morphometric data for adult mice among lines and generations. Body mass data represent means and SEM, whereas centroid size and tibia length are least squared means and SEM. Superscripts denote significant differences in means ($p < 0.05$) between a given group and: Controls ^{CTL}, Longshanks Line 1 ^{LS1}, Longshanks Line 2 ^{LS2}, from either: Generation 1 ^{F01}, Generation 9 ^{F09}, or Generation 20 ^{F20}, as determined using Tukey's HSD tests. Differences in body mass were determined by ANOVA, whereas tibia length and centroid size differences were assessed by ANCOVA with body mass as a covariate. Bold and italic superscripts indicate significant intergenerational differences and intragenerational differences, respectively.

Adult Body Mass (g)			
Generation	CTL	LS1	LS2
F01	39.61 (1.28)	39.25 (0.99)	42.13 (0.99) $^{\mathbf{F09\ CTL,\ F09\ LS2}}$
F09	36.54 (1.30) $^{\mathbf{F01\ LS2}}$	35.91 (0.99)	38.33 (0.99) $^{\mathbf{F01\ LS2}}$
F20	40.05 (0.99)	39.89 (0.99)	39.81 (0.99)
Adult Tibia Length (mm)			
Generation	CTL	LS1	LS2
F01	18.21 (0.12)	18.32 (0.09) $^{\mathbf{F09\text{-}LS1,\ F20\text{-}LS1}}$	18.30 (0.09) $^{\mathbf{F09\text{-}LS2,\ F20\text{-}LS2}}$
F09	18.23 (0.12) $^{\mathit{F09\text{-}LS1,\ F09\text{-}LS2}}$	19.51 (0.09) $^{\mathbf{F01\text{-}LS1,\ F20\text{-}LS1},\ \mathit{F09\text{-}CTL}}$	19.50 (0.09) $^{\mathbf{F01\text{-}LS2,\ F20\text{-}LS2},\ \mathit{F09\text{-}CTL}}$
F20	18.13 (0.09) $^{\mathit{F20\text{-}LS1,\ F20\text{-}LS2}}$	21.13 (0.09) $^{\mathbf{F01\text{-}LS1,\ F09\text{-}LS1},\ \mathit{F20\text{-}CTL}}$	21.15 (0.09) $^{\mathbf{F01\text{-}LS2,\ F09\text{-}LS2},\ \mathit{F20\text{-}CTL}}$
Adult Cranial Centroid Size			
Generation	CTL	LS1	LS2
F01	56.66 (0.19)	56.81 (0.15) $^{\mathbf{F20\text{-}LS1}}$	56.51 (0.15) $^{\mathbf{F20\text{-}LS2}}$
F09	56.64 (0.19)	57.49 (0.15) $^{\mathbf{F20\text{-}LS1}}$	56.82 (0.15) $^{\mathbf{F20\text{-}LS2}}$
F20	56.64 (0.15) $^{\mathit{F20\text{-}LS1,\ F20\text{-}LS2}}$	58.07 (0.15) $^{\mathbf{F01\text{-}LS1,\ F09\text{-}LS1},\ \mathit{F20\text{-}CTL}}$	57.81 (0.15) $^{\mathbf{F01\text{-}LS2,\ F09\text{-}LS2},\ \mathit{F20\text{-}CTL}}$

Supplementary Table S6: Euclidean distance between the multivariate mean PC scores of each group, based on Procrustes shape data adjusted for sex only (above diagonal), or sex and cranial centroid size (below diagonal). The only non-significant Euclidean distance, based on a post-hoc Procrustes ANCOVA, is indicated in bold.

		Mean PC scores adjusted for sex only								
Mean PC scores adjusted for centroid size and sex		F01-CTL	F01-LS1	F01-LS2	F09-CTL	F09-LS1	F09-LS2	F20-CTL	F20-LS1	F20-LS2
	F01-CTL	-	0.010	**0.007**	0.012	0.013	0.015	0.010	0.022	0.023
	F01-LS1	0.011	-	0.011	0.014	0.016	0.018	0.014	0.025	0.026
	F01-LS2	**0.007**	0.011	-	0.012	0.011	0.013	0.009	0.020	0.020
	F09-CTL	0.011	0.014	0.010	-	0.012	0.012	0.013	0.027	0.026
	F09-LS1	0.013	0.016	0.011	0.010	-	0.011	0.014	0.020	0.021
	F09-LS2	0.015	0.018	0.013	0.011	0.011	-	0.013	0.023	0.018
	F20-CTL	0.010	0.014	0.010	0.012	0.014	0.013	-	0.021	0.019
	F20-LS1	0.016	0.020	0.015	0.018	0.015	0.018	0.015	-	0.015
	F20-LS2	0.018	0.022	0.016	0.019	0.018	0.013	0.014	0.015	-

Supplementary Movie M1 - Mean cranial deformation in Longshanks. Top: deformation of an F20 Control mean cranial shape into a F20 Longshanks Line 1 (LS1) mean cranial shape, with both mean configurations scaled to the pooled mean centroid size of the total sample (n=327). Bottom: as above, but the skulls in each group have been scaled to the mean centroid size of that group. In both cases, the most significant changes relate to the flattening of the frontal and parietal regions, decreasing height of the braincase, and relative elongation of the snout in Longshanks.

Supplementary Movie M2 - Extreme cranial deformation in Longshanks. Top: deformation trajectory between one extreme Control male and one extreme LS1 male, selected from non-overlapping regions of morphospace (see Figure 1). The skull configurations are scaled to the pooled mean centroid size of the total sample (n=327). Bottom: as above, but the two skulls have been scaled to those individuals' centroid sizes, revealing substantial differences in centroid size between these two individuals as well.