S1. Phylogenetic comparative methods

We used genetic and morphological data from previous publications to compute ancestral state reconstructions and estimate the relative body size for this study [55]. The phylogenetic structure of 151 species in the mouse-related clade of rodents was used to reconstruct the ancestral states, then trimmed to only include the Dipodoidea clade, as depicted in Figure 1. Phylogenetic reconstructions for the 151 species tree are available as Supplementary Figure S5.

Metatarsal fusion in extant taxa was determined by examining preserved skeletal preparations. To estimate fusion in extinct ancestors, we used the `ancThresh` function in the `phytools` package for `R`. This function uses a threshold model, which assumes that the observed discrete states are ordered and determined by underlying, unobserved continuous trait variation [57]. Unfused, partially fused, and fused metatarsals were considered discrete states with each extant species fitting into a given state with 100% probability. The probability with which an ancestor would exhibit each of the discrete character states is indicated with a pie chart (Figure 1 A).

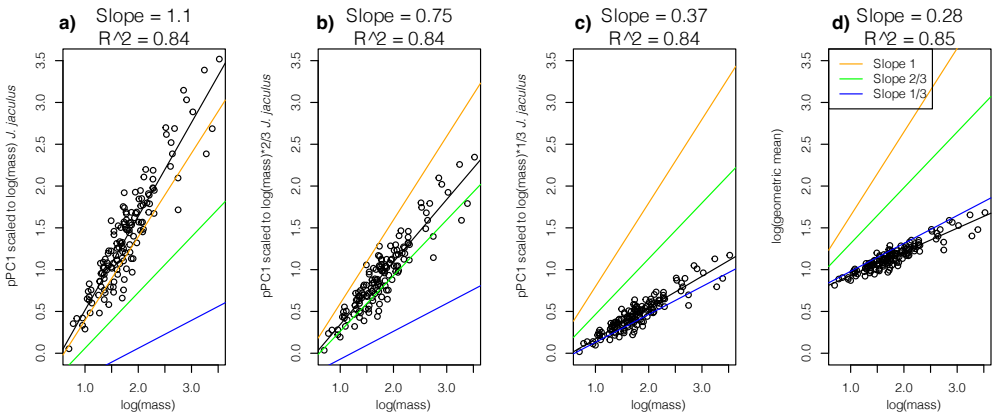

Figure S1. PC1 and geometric mean scale with body mass. a) pPC1 is scaled so that the value for *J. jaculus* matches with the log(mass) value for that species and scales well with mass^1. b) pPC2 is scaled to log(mass)$*2/3$ and scales well with mass$^{2/3}$. c) pPC2 is scaled to log(mass)$*1/3$ and scales well with mass$^{1/3}$. d) Geometric mean scales well with mass$^{1/3}$.

For large datasets, the empirically measured body size of each individual specimen may not be available or may vary according to the conditions in which the animal lived (i.e. wild-caught or lab-reared). In such cases, body size can be estimated using a principal components analysis [17]. For each species, bones are measured throughout the body, and when analyzed together, the first principal component frequently aligns best with changes in body size. Because the proportions of the forelimbs and the hind limbs vary greatly within Dipodoidea [17], including many non-dipodoid rodent outgroups provides a more accurate body size estimate. We used the `phyl.pca` function in the `phytools` package for `R` to perform a phylogenetically informed principal components analysis on the log-transformed long bone measurements (femur, tibia, metatarsus, humerus, ulna, metacarpus) in the dataset. We then compared pPC1 scores to previously published body mass ranges. We found that pPC1 scales with body mass^1 (Figure S1 A), so it can be used as a proxy for overall body size.

We used the skeletal measurements and pPC1 to compare the patterns of metatarsal length and body size evolution in Dipodoidea (Figure 1). Because these data are continuous measurements, we used the `contMap` function in the `phytools` package for `R` to statistically reconstruct estimates of these values for each common ancestor of the extant taxa [58]. We used the geometric

mean of the measurements as a linear representation of body size because it scales well with mass$^{1/3}$ (Figure S1 D).

S2. Model meshing

STL metatarsal models were reconstructed in MeshLab to achieve continuous, closed surfaces. Bone remeshing was performed in Meshmixer (Autodesk, San Rafael, CA) to reduce mesh density and smooth irregular surfaces. Defects such as self-intersecting faces and non-manifold edges were identified in Rhinoceros (Version 6, McNeel Europe, Barcelona, Spain) and remodeled in Meshmixer and MeshLab. Surface meshes fully closed and without defects were scaled to the desired size and transformed into polysurfaces in Rhinoceros.

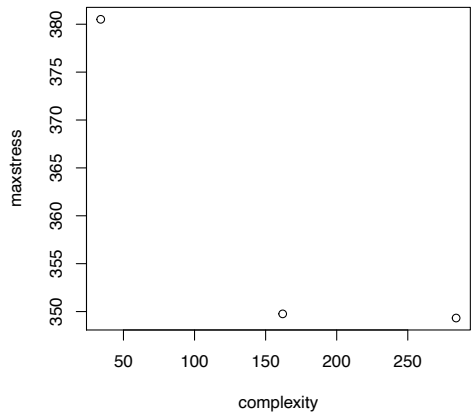

Figure S2. A convergence test was used to determine sufficient element size for the 3D meshes.

We used Hypermesh (Hyperworks 2020, Altair, Troy, MI) to scale the external and internal surface models according to the conditions in Table 1. Surface meshes were converted into solid meshes with quadratic tetrahedral elements. We performed a convergence test (Figure S2) to determine appropriate element sizes: the maximum element size was set to 1 mm, the minimum to 0.05 mm. The maximum deviation was set to 0.1 mm and the maximum angle was set to 15° for all scaled meshes Care was taken to preserve overall bone geometry. For unscaled meshes, the mesh was generated at the size of the *J. jaculus*, then scaled to the appropriate size. Thus, the number of elements is preserved, but the element sizes reflect the size of each species. Unfused metatarsals have more elements than fused metatarsals because of the greater complexity of the geometry.

S3. Second Moment of Area

We used 3D Slicer [30] to segment all of our STL models into 1000 slices along the lengths of each metatarsus. We then used a plugin for Slicer, called Segment Geometry [59], to measure the second moment of area at 1000 locations along the long axis of the metatarsus. Second moments were calculated about the neutral bending axis in our simulations, in addition to both the major and minor axes, which were determined independently for each sample. We directly compared *Jaculus jaculus* and *Euchoreutes naso* because these models have nearly the same outer diameters

Species	Scale	# 3D elements	Shortest elem (mm)	Longest elem (mm)	Shortest elem (%)	Longest elem (%)	Volume (mm^3)
NI	S	131136	1.04E-02	4.56E-01	3.09E-02	1.35	105.41
NI	U	149629	2.33E-03	1.77E-01	6.90E-03	0.52	7.71
ST	S	139166	8.74E-03	7.74E-01	2.59E-02	2.29	130.85
ST	U	162133	2.68E-03	1.56E-01	7.93E-03	0.46	1.29
EN	S	71085	1.32E-03	4.39E-01	3.91E-03	1.30	60.79
EN	U	81381	1.04E-03	3.57E-01	3.08E-03	1.06	32.07
JJ	R	71340	2.30E-03	4.04E-01	6.82E-03	1.20	65.16
AM	S	110643	5.10E-03	4.66E-01	1.51E-02	1.38	99.93
AM	U	90988	1.52E-02	4.41E-01	4.49E-02	1.31	87.56

Table S1. Features of the tetrahedral STL models. The minimum and maximum 1-dimensional length of any tetrahedral element is reported both in millimeters and as a percentage of total length of the model. Models with multiple unfused metatarsals have more elements. Elements located in the diaphysis are the most simple and largest, whereas the smaller elements are located in the epiphyses. In the Scale column, S stands for scaled to *Jaculus jaculus*, U represents unscaled, and R represents Reference.

when scaled to the same length. The major and minor axes are determined at each slice of the bone, so are not constrained to the sagittal and horizontal planes.

S4. Material model

The material properties for bone were used for all regions of bone, and did not incorporate information regarding the radiodensity of bone. Regions of radiodensity in tibiae have been shown to indicate bipedalism versus quadrupedalism in marsupials [60], but the present study is focusing on the effects of shape changes associated with fusion.

We used material properties from previous femur mechanical characterization experiments [37,61] into an elastic model with transversely isotropic and Hookean behavior without criteria for failure. Where there were differences in material properties between tension and compression, we selected the lower value to provide a conservative estimate for the behavior. We used the longitudinal Young's Modulus (16.4) for tension and the transverse Young's Modulus (11.7) for compression, and all other values were equivalent in compression and tension. The shear modulus (G23) was 4.0, the shear modulus (G13) was 3.5, the Poisson's ratio (v21=v31) was 0.235, and the Poisson's ratio (v32) was 0.376.

We used material properties from previous femur mechanical characterization experiments [37,61] into an elastic model without criteria for failure. Where there were differences in material properties between tension and compression, we selected the lower value to provide a conservative estimate for the behavior. We used the longitudinal Young's Modulus value for tension and the transverse Young's Modulus value for compression, and all other values were equivalent in compression and tension.

LS-DYNA enables specific and complex models of material properties, including highly non-linear behaviors. An isotropic asymmetric plastic material with different behavior for tension and compression was used for the cortical bone (MAT 124). For the plate, a rigid material with properties of a standard steel (Young's modulus 200 GPa) was chosen (MAT 20).

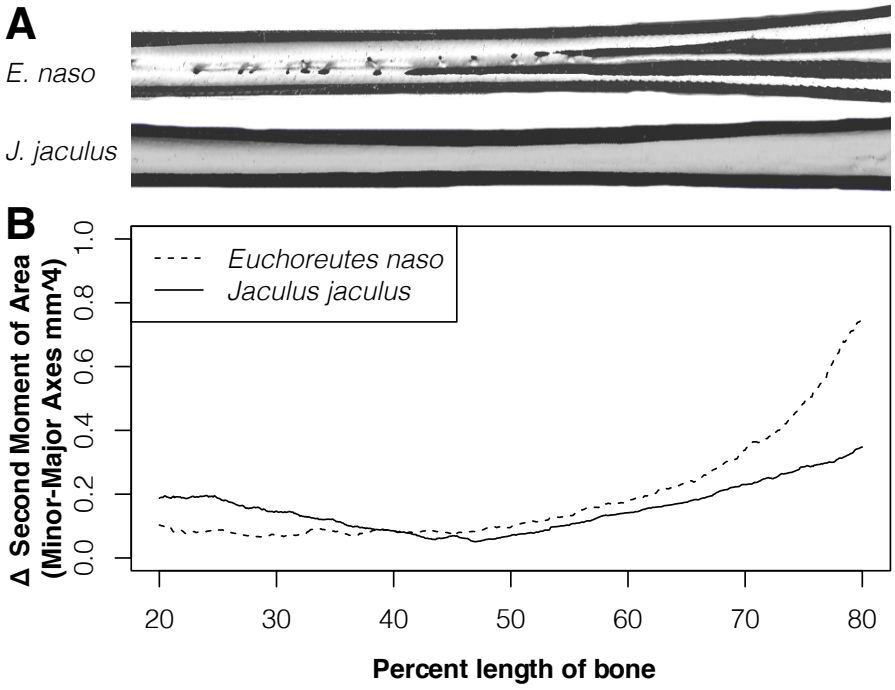

Figure S3. A) Longitudinal cross-sections show the morphology of the medullary cavities of a partially fused (*Euchoreutes naso*) species and a fully fused (*Jaculus jaculus*). B) The difference in the second moment of area between the major and minor axes is computed with respect to position on the bone. Greater differences correspond to less radial symmetry in the cross-sectional shape at that portion of the bone. For *Euchoreutes naso*, the difference in second moment of area between major and minor axes exceeds that of *Jaculus jaculus* in distal areas where vertical columns are present and the metatarsals begin to separate into distinct medullary cavities.

Species	Museum Number	Fusion	Locomotion	MT length (mm)	Cortical Thickness (mm)
Napeozapus insignis	MCZ 64235	Unfused	Fac. Biped	13.16	1.92
Salpingotus thomasi	AMNH 244431	Unfused	Biped	6.77	1.84
Euchoreutes naso	MSB 227347	Partial	Biped	27.20	2.04
Jaculus jaculus	**CFS JJ0039**	**Fused**	**Biped**	**33.74**	**2.13**
Allactaga sibirica	AMNH M-55980	Fused	Biped	33.89	2.64

Table S2. List of specimens used in this study and dimensions. *Jaculus jaculus* was used as the reference species.

Species	Body Size (geometric mean)	Unscaled Plate Displacement
N. insignis	12.93	2.25
S. thomasi	6.5	1.13
E. naso	15.85	2.75
J. jaculus	**17.26**	**3.00**
A. sibirica	21.44	3.73

Table S3. Scaling conditions for each species, with respect to *Jaculus jaculus* geometric mean values derived from [55]. Displacements are reported in millimeters.

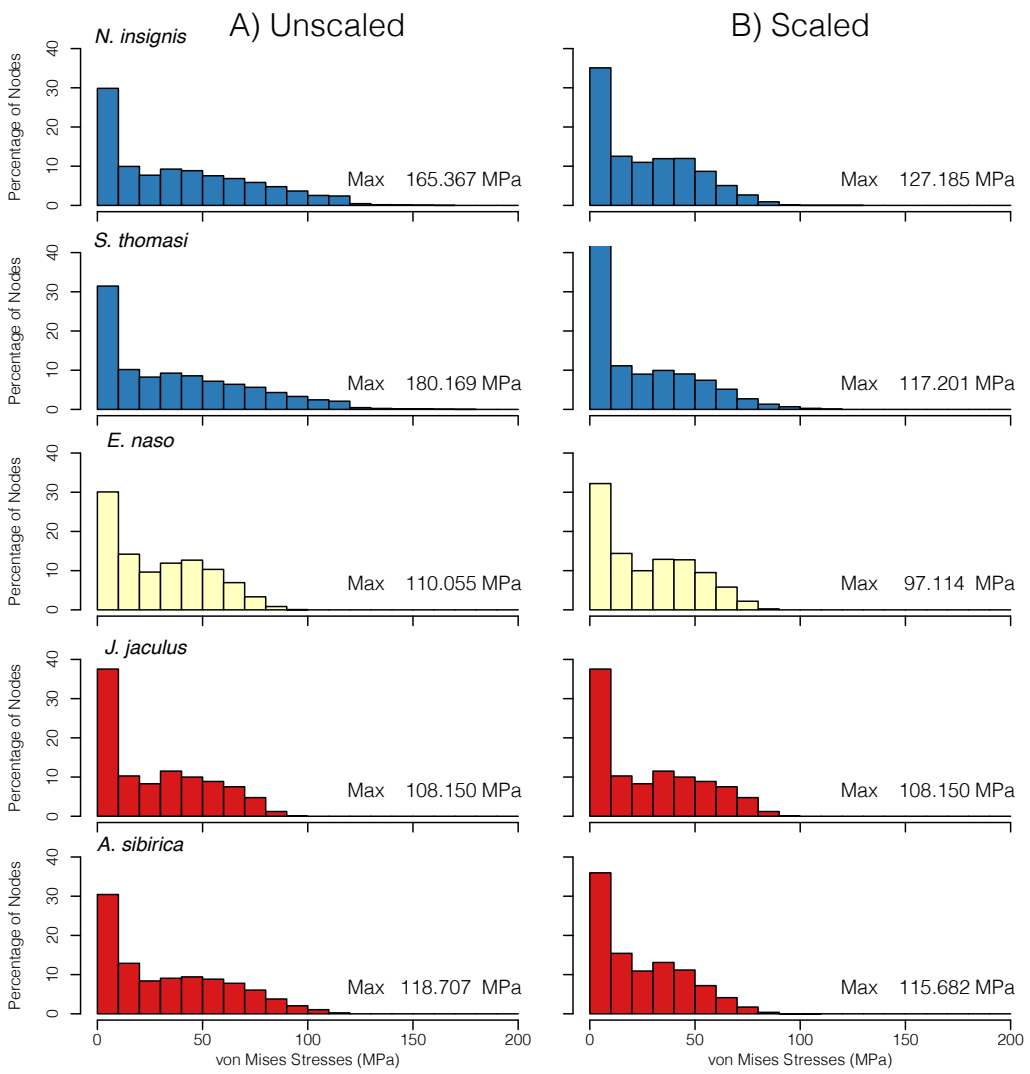

Figure S4. Histograms of peak von Mises stress for every node indicate stress concentration for each model. Colors indicate metatarsal fusion. Reference Table 1 for details regarding each model.

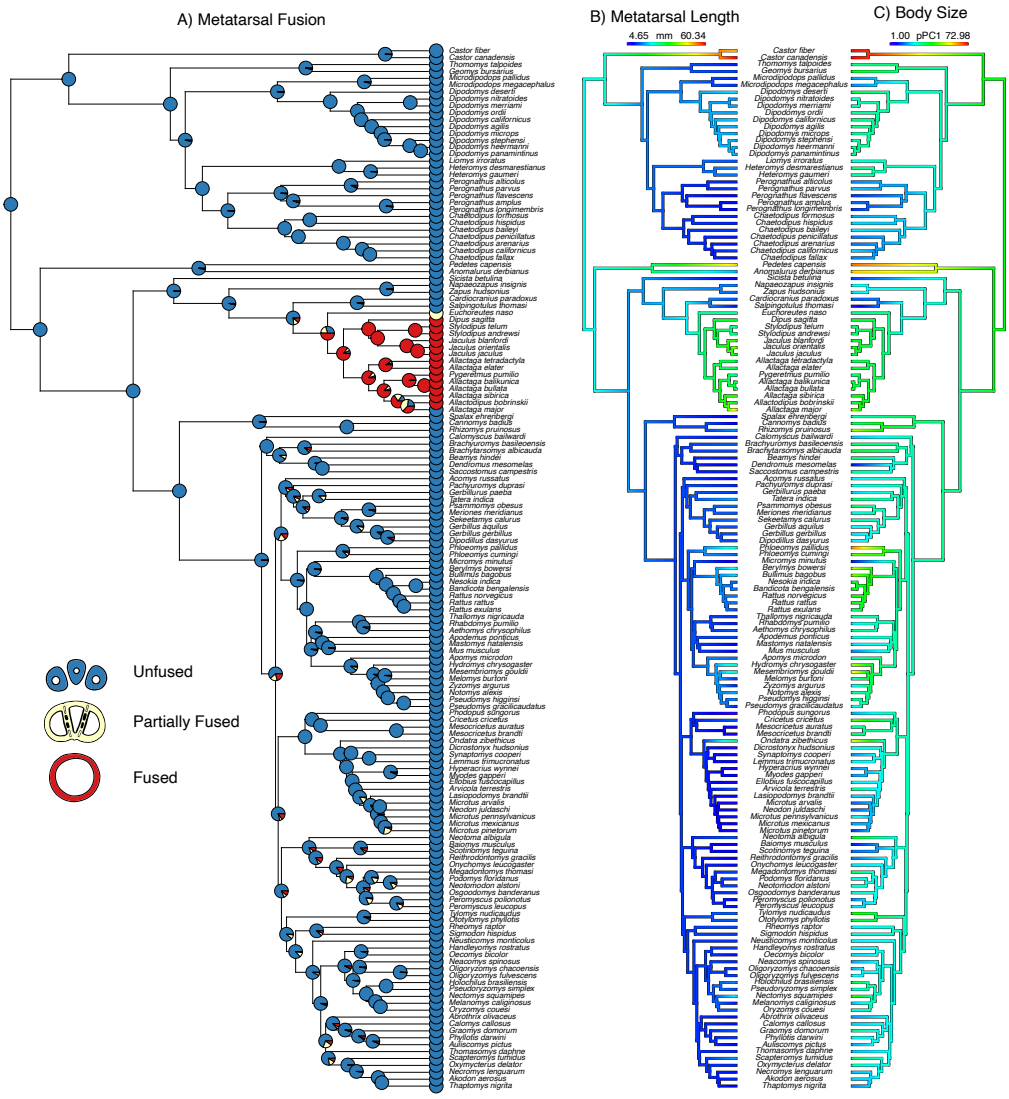

Figure S5. Ancestral state reconstructions for all 151 mouse-related rodent species used to generate the trimmed Dipodoidea phylogenites in Figure 1. A) Metatarsal fusion was reconstructed using a threshold model. The pie charts depict the probability that an ancestor would have unfused, partially fused, or fully fused metatarsals. B) Metatarsal length was reconstructed as a continuous character. C) Body size was reconstructed as a continuous character.