

Predicting the bending properties of long bones: Insights from an experimental mouse model

Sarah J. Peacock¹  | Brittney R. Coats² | J. Kyle Kirkland³ |
Courtney A. Tanner⁴ | Theodore Garland, Jr.⁵  | Kevin M. Middleton¹ 

¹Department of Pathology and Anatomical Sciences, University of Missouri, Columbia, Missouri

²ITW David Speer Academy, Chicago, Illinois

³Department of Kinesiology, Michigan State University, East Lansing, Michigan

⁴Blivet Solutions, Inc., Chicago, Illinois

⁵Department of Evolution, Ecology, and Organismal Biology, University of California, Riverside, California

Correspondence

Sarah J. Peacock, Department of Pathology & Anatomical Sciences, University of Missouri, Medical Sciences M263, 1 Hospital Dr., Columbia, MO 65212, USA. Email: sjpd58@mail.missouri.edu

Funding information

National Institutes of Health National Institute of Arthritis and Musculoskeletal Diseases, Grant/Award Number: 5G11HD052368-03; University of Missouri Department of Pathology and Anatomical Sciences; University of Missouri Bond Life Sciences Fellowship Program; University of Missouri Gus T. Ridgel Fellowship; California State University, San Bernardino Department of Biology; Associated Students Incorporated of California State University, San Bernardino; University of Missouri School of Medicine; National Science Foundation, Grant/Award Number: IOS-1121273

Abstract

Objectives: Analyses of bone cross-sectional geometry are frequently used by anthropologists and paleontologists to infer the loading histories of past populations. To address some underlying assumptions, we investigated the relative roles of genetics and exercise on bone cross-sectional geometry and bending mechanics in three mouse strains: high bone density (C3H/He), low bone density (C57BL/6), and a high-runner strain homozygous for the *Myh4*^{Minimsc} allele (MM).

Methods and Materials: Weanlings of each strain were divided into exercise (wheel) or control (sedentary) treatment groups for a 7-week experimental period. Morphometrics of the femoral mid-diaphysis and mechanical testing were used to assess both theoretical and ex vivo bending mechanics.

Results: Across all measured morphological and bending traits, we found relatively small effects of exercise treatment compared to larger and more frequent interstrain differences. In the exercised group, total distance run over the experimental period was not a predictor of any morphological or bending traits. Cross-sectional geometry did not accurately predict bone response to loading.

Discussion: Results from this experimental model do not support hypothesized associations among extreme exercise, cross-sectional geometry, and bending mechanics. Our results suggest that analysis of cross-sectional geometry alone is insufficient to predict loading response, and questions the common assumption that cross-sectional geometry differences are indicative of differential loading history.

KEYWORDS

biomechanics, loading, mouse model, skeletal plasticity

1 | INTRODUCTION

Bone response to loading has been a frequent research focus due to both its clinical and biological relevance (Burr, Martin, Schaffler, & Radin, 1985; Duncan and Turner, 1995; Lanyon, 1992; Lieberman, 2003; Meakin, Udeh, Galea, Lanyon, & Price, 2015; Sugiyama et al., 2012). Because intense and/or chronic skeletal loading can induce an anabolic response (Hamrick, Skedros, Pennington, & McNeil, 2006; Moss, 1997; Rubin, 1984; Rubin and Lanyon, 1985), analyses of bone morphology and cross-sectional geometry have been used as proxies

for inferring a bone's functional response to loading in a wide range of vertebrates (Kelly, Czech, Wight, Blank, & Garland, 2006; Middleton, Kelly, & Garland, 2008a, Middleton et al., 2008b; Simons, Hieronymus, & O'connor, 2011), including humans and other primates (Demes, Jungers, & Selpien, 1991; Demes, Jungers, & Walker, 2000; Holt, 2003; Maggiano et al., 2008; Polk et al., 2000; Ruff, Larsen, & Hayes, 1984; Ruff, 2009; Shaw and Stock, 2013; Stock and Pfeiffer, 2001; Trinkaus and Ruff, 1999). However, the extent to which cross-sectional phenotypes can be used to reconstruct past behaviors has been questioned by other studies (Demes et al., 1998; Lieberman, Polk, & Demes, 2004;

Maki, 2013; Niinimäki et al., 2017; Schmitt, Zumwalt, & Hamrick, 2010; Wallace, Tommasini, Judex, Garland, & Demes, 2012; Wallace et al., 2014; Wallace, Demes, & Judex, 2017). Contrasting studies suggest that genetic background (Cowgill, 2014; Gilbert and Gill, 1990; Tallman and Winburn, 2015; Wescott, 2005, 2006; Wescott and Srikanta, 2008), age (Bouxsein, Myburgh, van der Meulen, Linderberger, & Marcus, 1994), adaptations to climate (Pearson, 2000; Weaver, 2003), and body mass (Agostini and Ross, 2011) may have more profound influences on cross-sectional geometry.

This study was designed to investigate the potential limitations of using cross-sectional morphology to reconstruct past behaviors, by exploring the extent to which a natural behavior, long-term voluntary exercise (Meijer and Robbers, 2014), leads to measurably different skeletal phenotypes and how those bone cross-sectional phenotypes behave during *ex vivo* loading. Due to the difficulties in controlling for other environmental factors that may influence bone phenotype and the logistical and ethical constraints of using human or other primates in an experimental setting, we chose to use a mouse model. Mice are generally accepted to be an appropriate model system for human skeletal biology (Beamer, Donahue, & Rosen, 2002; Carlson, Lublinsky, & Judex, 2008; Gross, Poliachik, Prasad, & Bain, 2010; Hallgrímsson and Lieberman, 2008; Middleton et al., 2008a) due to their well-understood genetics and similar plastic response to loading as primates. Additionally, shorter experimental timescales and greater capacity to control factors like diet and genetic background allow mice to be a more feasible model system than primates for experiments of this kind.

We studied the effect of exercise on femora of homozygous *Mhy4^{Minimisc}* (MM) mice, originally from one of four closed lines of mice that had been artificially selected for high levels of voluntary running, and compared results with those for two inbred strains, which are common models for high (C3H/He) and low (C57BL/6) bone mass. We tested three hypotheses: (1) mice with different genetic backgrounds will exhibit variation in cross-sectional femoral morphology that can be modulated by differences between strains and by loading caused by wheel exercise, (2) femora of mice allowed wheel access will show a corresponding anabolic loading response, and (3) inferences of bending mechanics based on cross-sectional geometry will accurately predict results from mechanical testing.

2 | METHODS

2.1 | Experimental design

We compared bone phenotypes in three strains of mice: (1) high-runner (HR) mice homozygous for the *Mhy4^{Minimisc}* allele (hereafter “MM”), (2) a high bone density phenotype, C3H/HeNHsd, and (3) a low bone density phenotype, C57BL/6NHsd. MM mice display an approximately 50% reduction in the mass of the triceps surae muscle complex resulting from the near complete absence of type IIb muscle fibers (Guderley, Houle-Leroy, Diffée, Camp, & Garland, 2006; Guderley, Joannisse, Mokas, Bilodeau, & Garland, 2008; Bilodeau, Guderley, Joannisse, & Garland, 2009; Houle-Leroy, Guderley, Swallow, & Garland, 2003; Talmadge, Acosta, & Garland, 2014). Male MM mice

were chosen at weaning (~21 days) from eight dams (two per dam) that had undergone 12 generations of inbreeding with continued selection for high levels of voluntary wheel running (see Supporting Information 1 for additional information about inbreeding protocol and Swallow, Carter, and Garland (1998a) for details about the selection experiment prior to founding of the MM inbred line at generation 51). After 12 generations of brother-sister mating, we estimate that at least 93% of alleles are homozygous, following the equations adapted from Green (1981) for residual heterozygosity at the *i*th generation under brother-sister mating:

$$\text{Heterozygosity} = 1 - F_i/2^{(i-1)}$$

where F_i is the *i*th observation in the Fibonacci sequence ($F_i = F_{i-1} + F_{i-2}$ given that $F_1 = 1$ and $F_2 = 2$). 93% homozygosity is likely an underestimate because after 51 generations of selection, the source HR line was likely already homozygous at many loci.

The remaining two inbred strains C3H/HeNHsd (hereafter “C3H/He”) and C57BL/6NHsd (hereafter “C57BL/6”) mice were purchased from a commercial vendor (Harlan Laboratories) at weaning. These strains were chosen because they are standard models for the study of skeletal loading responses. They represent the ends of a continuum of bone mineralization among inbred strains (Beamer, Donahue, Rosen, & Baylink, 1996) and provide context for interpretation of the bending properties of MM bones.

Rodent chow (Teklad Rodent Diet 8604; Harlan) and water were supplied *ad libitum*. After a seven-week experimental period starting at 21 days, mice were euthanized via CO₂ inhalation. All procedures were approved by the California State University, San Bernardino Institutional Animal Care and Use Committee.

2.2 | Voluntary wheel-running behavior

Half of each group of mice was randomly assigned to either a wheel access group or a nonwheel control (sedentary) group ($n = 8$ of each strain in each group). Mice without wheels were housed in groups of four in standard cages according to strain, while mice given wheels were individually housed in cages that included a 0.73 m circumference running wheel (Nalgene Cages, Bend, OR; see Morgan, Garland, & Carter, 2003 for additional details). During the course of the experiment, five C3H/He mice died for unknown reasons, leaving a final group size of 11 ($n = 7$ nonwheel control; $n = 4$ wheel access). One MM histological section was damaged and could not be analyzed ($n = 8$ nonwheel control; $n = 7$ wheel access). All C57BL/6 mice were analyzed ($n = 16$).

Although single housing induces some stress and may alter baseline activity, it is necessary in order to measure an individual mouse's activity. Single housing of sedentary animals was not possible due to space constraints. Mice were free to use the exercise wheel as much or as little as they chose at all times. The number of wheel revolutions was computed using a magnet attached to the wheel which triggered a reed switch connected to a digital I/O interface (USB-120A; Eagle Appliances Ltd., Cape Town, South Africa). Time stamps for each revolution were recorded to the nearest millisecond, using custom-written

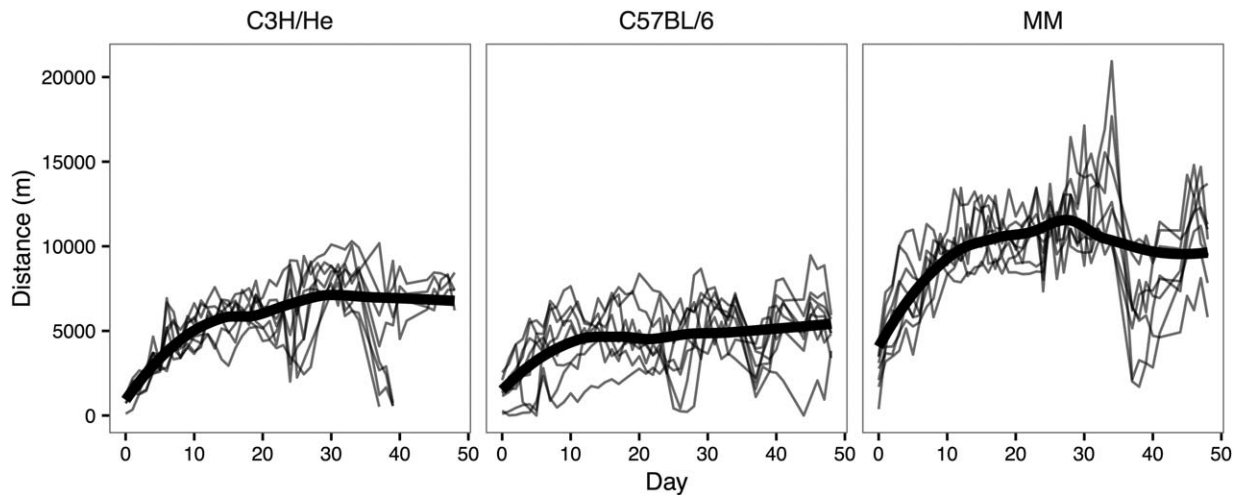


FIGURE 1 Mean daily locomotor distance. Individual (thin lines) and mean per strain (thick lines) running distances for three mouse strains during the 7-week experimental period. Mean total distance in MM mice across the experimental period was 419 km, 285 km in C3H/He, and 199 km in C57BL/6. A transient drop in running from days 37 to 40, most obvious in MM mice, coincided with two injections of calcium binding fluorochrome markers for a related experiment not described here

software. Revolutions were summed for each daily 24-h period starting at 1200 (noon) and ending at 1159 using a custom R package (<http://github.com/kmiddleton/binning>) and converted to meters for analysis (Figure 1). Cages were changed weekly, which resulted in slightly shorter running durations on those days; however, this gap occurred during the light phase when mice ran minimally.

2.3 | Gross morphometrics and histomorphometry

Following euthanasia, both hind limbs of each mouse were dissected. In preparation for whole-bone morphometrics, left femora were manually defleshed, and traits were measured with digital calipers to 0.01 mm: femur length (superior articular surface of the femoral head to the farthest extent of the distal condyles), anteroposterior and mediolateral diameters at the mid-femoral diaphysis, proximal width (greater trochanter to the medial femoral head), distal width (across the distal condyles), and femoral head proximodistal height and anteroposterior depth. Left femora were subsequently frozen in saline for use in mechanical testing (see below).

Right femora were dried and defleshed using dermestid beetles. Once cleaned of residual muscle tissues, the bones were individually embedded in low viscosity epoxy (Epo-Thin; Buehler, Ltd., Lake Bluff, IL) under a vacuum and allowed to cure overnight. Approximately 1-mm-thick sections were cut from the mid-diaphysis of each femur below the third trochanter using a low-speed saw (IsoMet; Buehler, Ltd., Lake Bluff, IL).

The cross-sections were mounted on slides and imaged using light microscopy (Figure 2a). Each image was captured digitally and then prepared for analysis of cortical cross-sectional geometry (GNU Image Manipulation Program; www.gimp.org). Image color levels were adjusted to reduce background color and to improve bone contour visibility (Figure 2b). The images were then manually segmented to remove medullary contents (Figure 2c), and a threshold was applied to yield a black and white image of the cortex (Figure 2d). Repeatability of cross-sectional area measurements using this method was 0.99, estimated as

the ratio of within group to total variance from a linear mixed model (Nakagawa and Schielzeth, 2010). Thus, despite slight variation in choice of threshold value and manual segmentation, this method yields highly repeatable results. Cross-sectional images were analyzed for area, second moments of area, and section modulus following Lieberman et al. (2004), using BoneJ (Doubé et al., 2010).

2.4 | Mechanical testing

Mechanical testing was performed on freshly thawed bones that had been previously frozen in 0.9% saline (see above). The authors acknowledge that the material properties of bone have been shown to change with hydration level. However, all bones were treated similarly and can be assumed to be equally moist. As our interest is not in the specific values of the mechanical properties tested, but rather in the differences between strains, this is sufficient for our needs. To test bending behavior, left femora were placed into a custom-built 3-point bending apparatus (following Turner and Burr, 2001) with a span of 6 mm and a support radius of 0.75 mm to avoid shear (Turner and Burr, 2001). Following the methods of Kodama et al. (2000), femora were loaded via a materials testing machine (Instron 5942; Instron, Inc., Norwood, MA) anteroposteriorly at a rate of 10 mm/min at the mid-femoral diaphysis until fracture, as determined by the point where the load rate dropped by 40% (Figure 3).

Raw load–displacement data were converted into flexure stress (σ) and flexure strain (ϵ) following Turner (1989) and Turner and Burr (2001), using the equations:

$$\sigma = \frac{P \cdot s \cdot r}{4 \cdot I}$$

$$\epsilon = \frac{6 \cdot D \cdot d}{s^2}$$

where P is the load, s is the span length (6 mm), r is the radius of the bone in the direction of loading, I is the second moment of area about

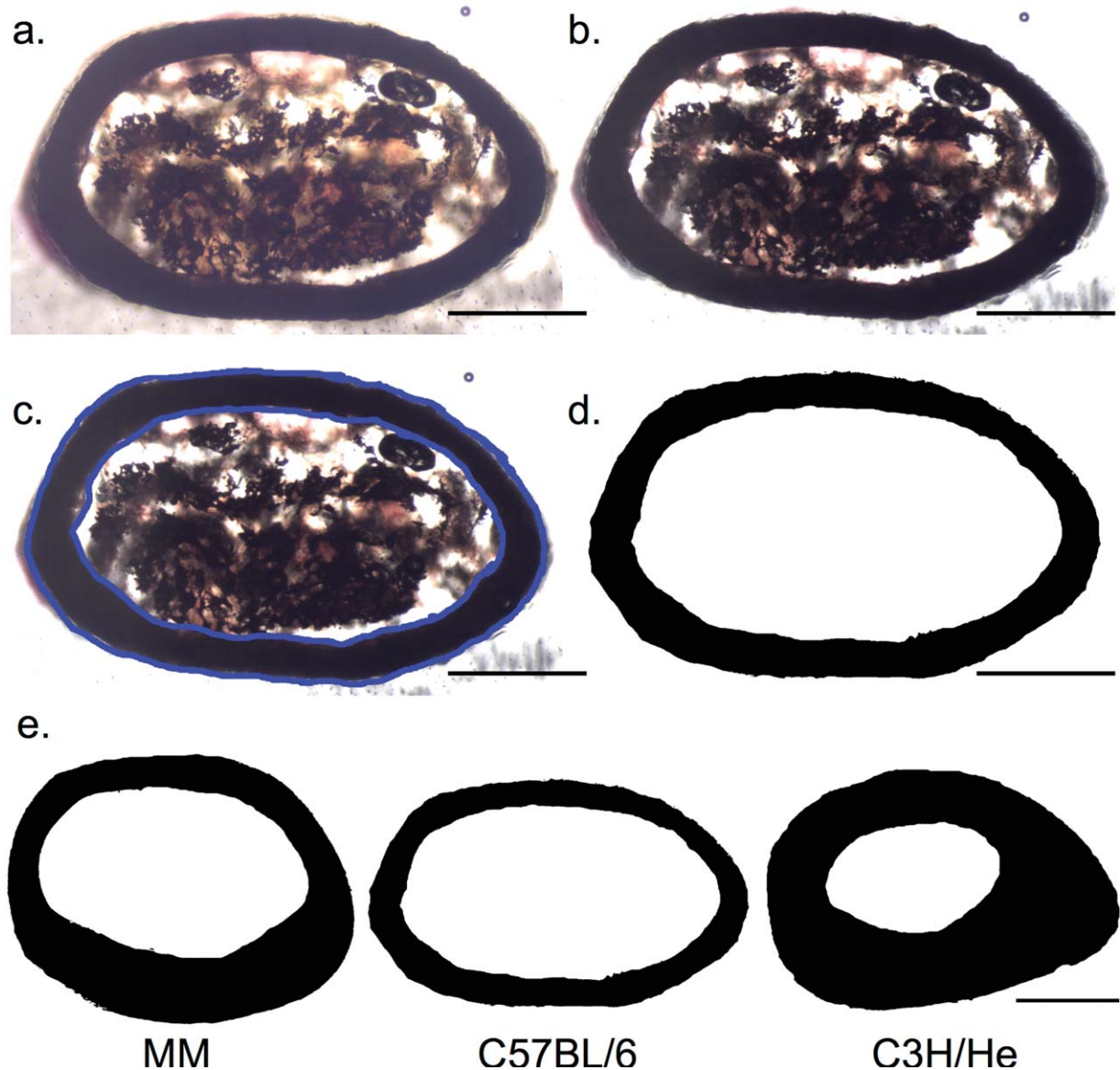


FIGURE 2 Comparative femoral cross-sections. (a) An example of the raw femoral cross-section image data prior to thresholding. (b) The same cross-section image after color adjustment to reduce excess background color. (c) An outline has been added to the cross-section to indicate the areas of bone that will be kept as a silhouette; everything else is removed before analysis. (d) The final silhouette after background removal and conversion to black and white. (e) Comparative cross-sectional morphology of MM, C57BL/6, and C3H/He femora from mice of approximately the same body size (about 27 g). The C3H/He femur has significantly greater cortical thickness than the other two strains, which is not unexpected given that C3H/He mice are the high bone mass phenotype mouse model. In comparison, the C57BL/6 femur appears much more gracile, while the MM femur is somewhat intermediate to C3H/He and C57BL/6 morphology. Scale bar equals 0.5 mm

the axis of loading, D is the vertical displacement under load, and d is the diameter in the direction of loading. I was determined via histomorphometric analysis of the contralateral limb using BoneJ (see above). Although it would have been ideal to use the same leg for both tests, the logistical constraints of undecalcified bone histology precluded this. Previous studies in HR mice have shown significantly lower limb skeleton asymmetry than in randomly bred controls (Garland and Freeman, 2005), so we assume that minimal bias was introduced by using cross sectional data from the contralateral limb.

The yield point of a stress–strain curve is the point which separates the elastic and plastic loading regions (Gere and Timoshenko, 2001). This point was identified algorithmically with piecewise regression in R (R Core Team, 2017), using the package “segmented” (Muggeo, 2003). We used the slope of the elastic portion of the stress–strain curve (the portion of the curve to the left of the yield point) as an estimate of Young’s modulus, a measure of the intrinsic material stiffness. From the stress–strain curve, we estimated ultimate strength (maximum value of the stress–strain curve), ultimate strain

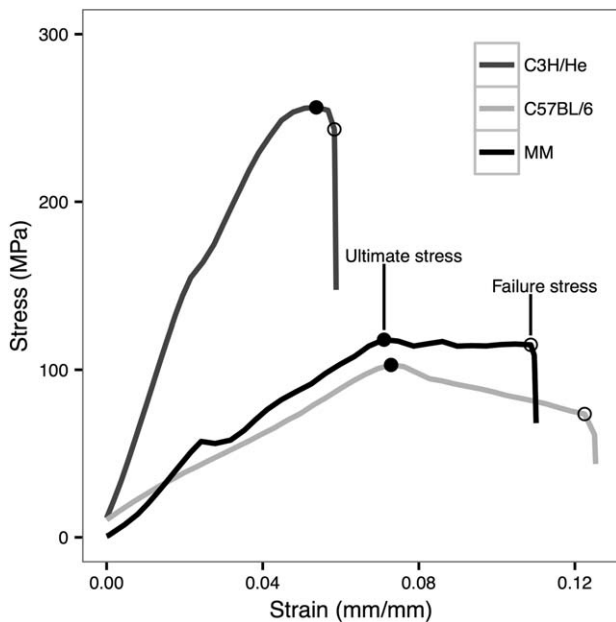


FIGURE 3 Comparative stress–strain curves. Three representative stress–strain curves are shown for C3H/He, C57BL/6, and MM strains. The slope of each curve in the elastic region before ultimate stress (filled circle) is Young’s modulus, an estimate of the stiffness of the bone. Failure stress (open circle) was marked as the point where rate of loading dropped 40%. Young’s modulus is similar in MM and C57BL/6 femora, but significantly higher in C3H/He

(strain at ultimate strength), failure strength (stress at failure), failure strain (strain at failure), and modulus of toughness (area under the stress–strain curve) both to ultimate strength and to failure (Figure 3).

2.5 | Statistical analysis

We used a Bayesian inference framework for all statistical analyses. Bayesian analysis has gained traction in recent years with the development of tools to facilitate model definition and has been suggested as applicable for questions in physical anthropology (Konigsberg and Frankenberg, 2013). Benefits of a Bayesian approach for our analysis include (1) allowing comparisons of posterior distributions of differences between strains or exercise vs. sedentary treatment without reliance on specific post-hoc tests or intermediate ANOVA or ANCOVA models (Kruschke, 2015; McElreath, 2015), (2) obviating the need for post-hoc multiple comparisons adjustments (Gelman and Tuerlinckx, 2000; Gelman, Hill, & Yajima, 2012), and (3) facilitating the interpretation of parameter estimates relative to a value of interest (e.g., difference = 0), using a 95% highest density interval (HDI), which includes the most credible values for the parameter estimate. In this case, a 95% HDI, which excludes 0, constitutes credible evidence that parameter value is different from 0 (Kruschke, 2015).

We used three types of linear models: (1) Prediction of single traits (total distance run, body mass, femoral length) by strain, by wheel access, or by the full factorial model (one- or two-way ANOVA-like

models), (2) prediction of traits by strain and wheel access with femoral length as a continuous covariate (an ANCOVA-like model), and (3) prediction of traits by total cumulative distance for animals in the wheel treatment group only (results presented in Supporting Information, 2). The authors recognize that body mass is a more commonly-used covariate than femoral length within anthropological research; however, as body mass is more highly influenced by activity levels, and considering that variation in activity is an important component of our experimental design, we decided that using femur length was more appropriate in the context of this study. In the final set of models, we were interested in the slope parameter predicting traits from total cumulative distance. These models were run for each mouse strain separately, to avoid confounding strain with total distance (Figure 1). We neither calculated nor report *p* values, but in their stead include summaries of posterior distributions of models: priors conditioned on the observed data. Reporting parameter estimates rather than significance levels for specific null hypothesis tests allows us to avoid multiple comparison issues (Gelman et al., 2012; Kruschke, 2015).

In a Bayesian framework, each model generates posterior distributions of parameter estimates from observations (data) conditioned on priors using Markov chain Monte-Carlo (MCMC), where the priors are the initial probabilities for each possible value of each parameter. Because each MCMC sample contains simultaneous estimates for all parameters in a model, differences between estimates (e.g., effect of wheel access) can be calculated directly from posterior distributions (Figure 4). We summarized the differences between posterior distributions for comparisons of interest using the median and 95% highest

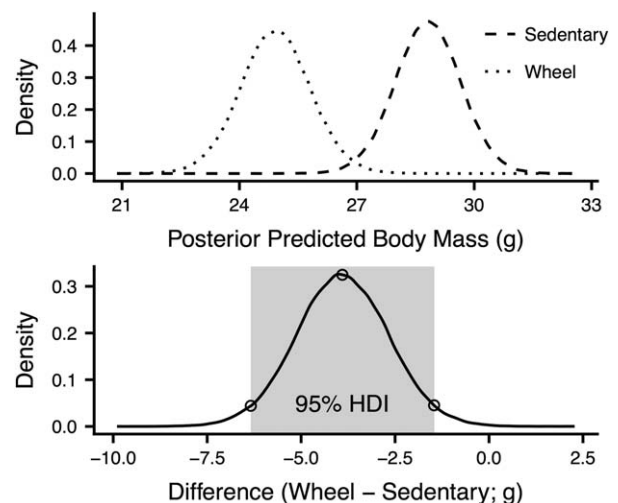


FIGURE 4 Differences of posterior distributions. The top panel shows the posterior distributions for estimated body mass in sedentary (dashed line) and wheel access (dotted line) groups of MM mice. Because each Bayesian MCMC iteration produces estimates for all estimated parameters simultaneously, each pairs of values (here, body mass in sedentary and wheel access MM mice) is jointly estimated. The difference of those pairs represents the most credible estimate of their difference (lower panel). The 95% highest density interval (HDI) is summarized by the lower bound, the median, and the upper bound (Figures 5–7)

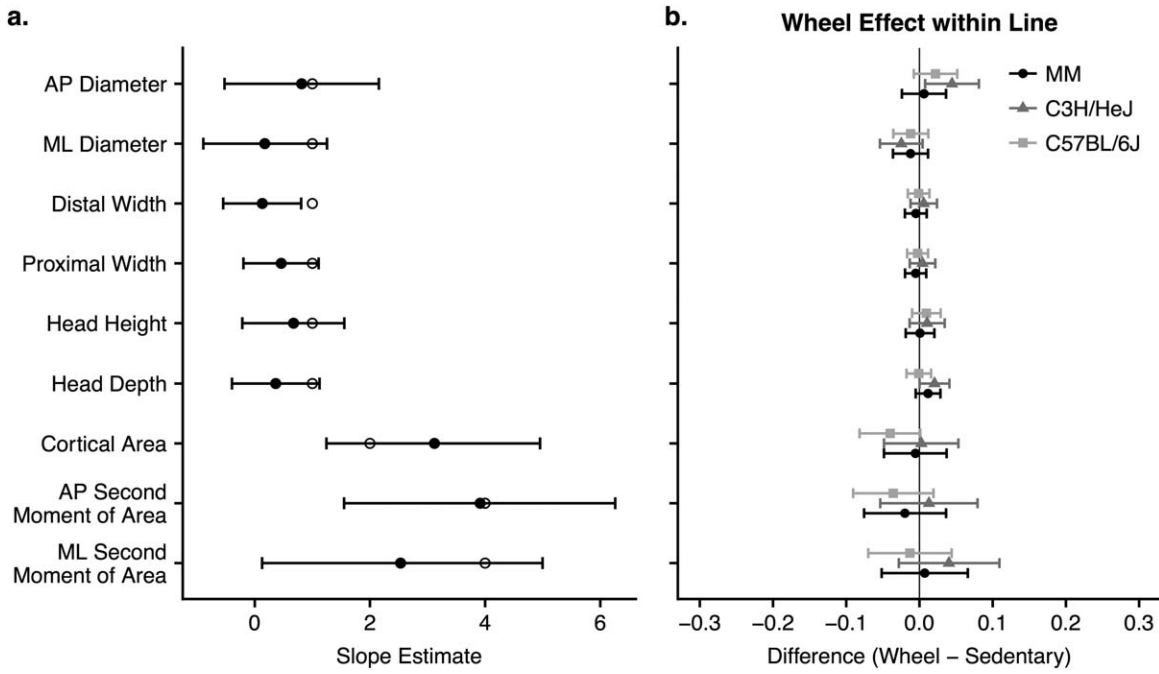


FIGURE 5 Summarized linear models of gross and histomorphometry. Results of Bayesian linear models with femoral length as a continuous covariate are summarized by the median (solid circle) and 95% HDI (bars). (a) Slope estimates for the covariate femoral length are compared to the null hypothesized isometric slope (open circle). In only one trait (distal femoral width) did the 95% HDI not include the null slope value. (b) Summaries of posterior distributions of differences of wheel: no wheel treatment groups for each mouse strain show that very few traits show an HDI that deviates from 0, indicating very small differences between treatment groups. The range of the x axis is equal between Figures 5b and 6 to facilitate comparisons

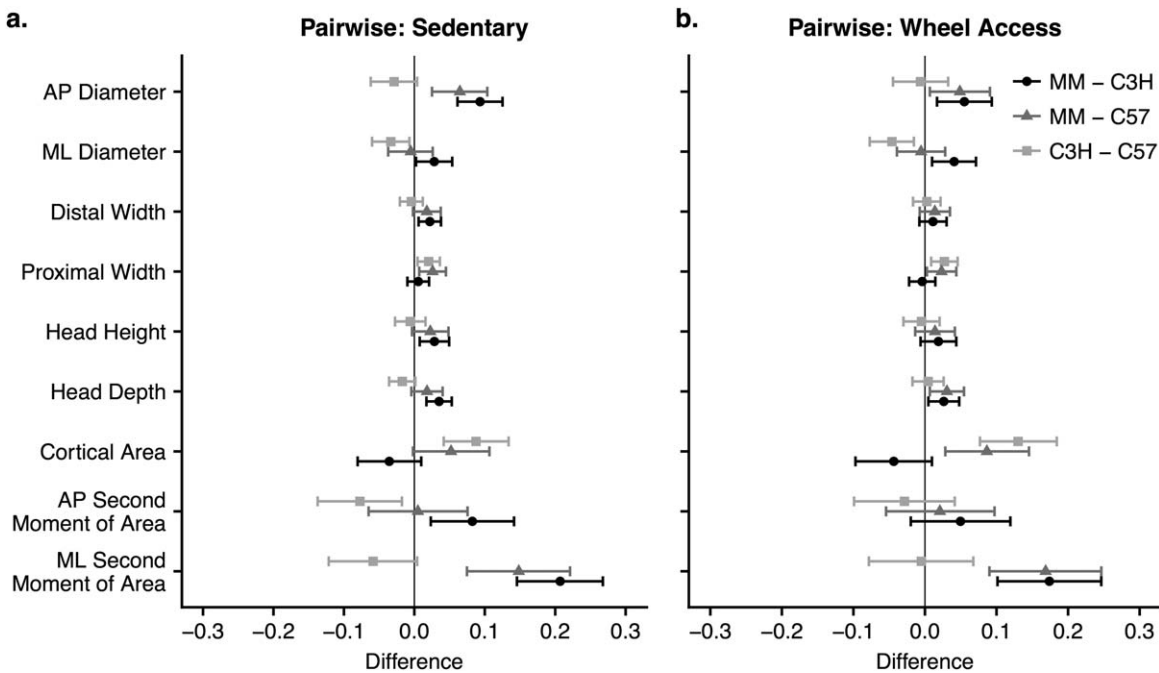


FIGURE 6 Summarized linear models of gross and histomorphometry. In pairwise comparisons of differences of HDIs for (a) sedentary and (b) wheel access groups, each line summarizes the difference between one pair of strains in one treatment group (e.g., MM: C3H/He). HDIs that do not cross 0 are interpreted as having strong credibility that the difference deviates from 0. For example, AP diameter is larger in MM than in either C3H/He or C57BL/6 in both sedentary and wheel access treatments, but the latter two are not different from each other. The range of the x axis is equal between Figures 5b and 6 to facilitate comparisons

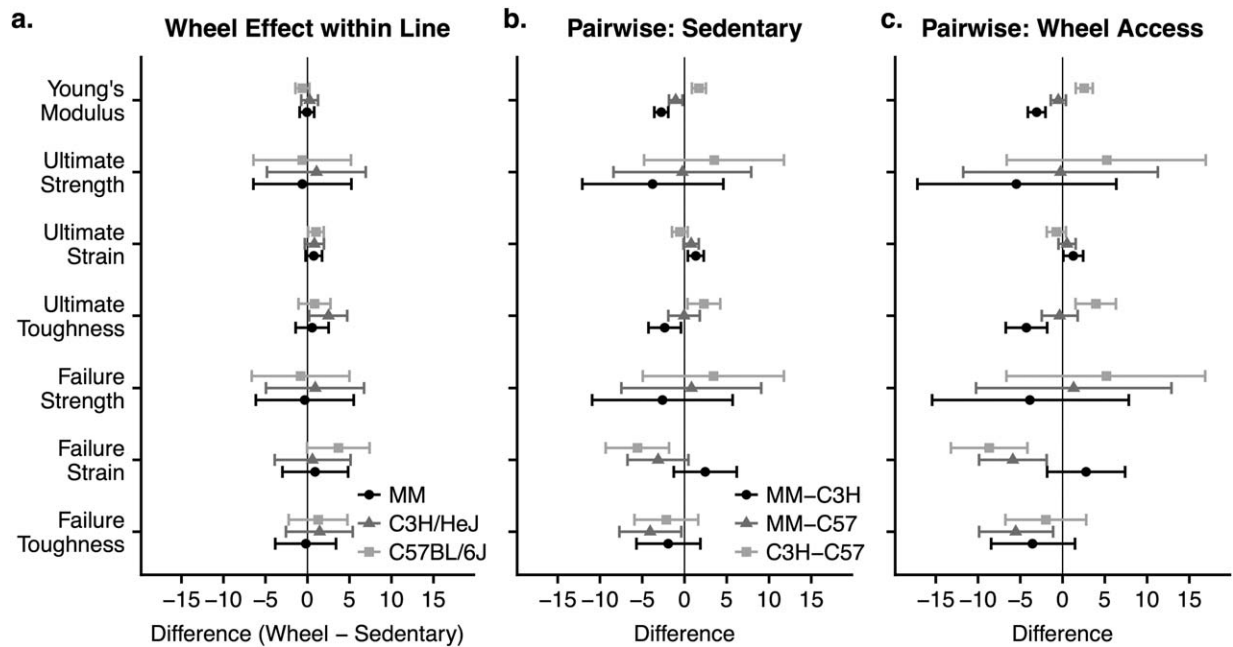


FIGURE 7 Summarized linear models of biomechanical measures. (a) Biomechanical measures of resistance to bending reveal few wheel effects, but, as for gross and histomorphometry, (b,c) many differences among strains in both sedentary and wheel access treatments. Estimates for strain have been multiplied by 75 to avoid overplotting medians with their 95% HDIs

density interval (HDI). When the parameter of interest (e.g., difference = 0) falls outside the HDI, we interpret this situation as plausible evidence that the parameter estimate does not include the value in question. For example, Figure 4 shows the posterior densities for body mass estimates in wheel access and sedentary MM mice and the difference between the posterior estimates. The 95% HDI for the difference does not encompass 0, which we interpret as credible evidence for lower mass in the wheel access treatment group. The most credible difference is ~ 4 g.

With the exception of body mass, which was estimated on the original scale, all morphological variables were \log_{10} -transformed and centered before analyses. Centering (subtracting the grand mean from all values) improves model fitting while preserving estimation of slopes for covariates on the original scale. Cumulative distance run was standardized (Z) prior to analysis, because centering did not adequately account for the very large variance differences in cumulative distances between strains.

For all parameters, we specified broad but weakly regularizing priors that temper the effects of extreme values: normal distributions centered on 0 for most parameters, normal distributions centered on null hypothesized isometric slopes for continuously varying covariates, and Cauchy distributions with a shape parameter of 2 for standard deviations. Models were initially run with four replicate chains for 2,000 MCMC iterations with 50% for warmup to assess suitability of prior specifications and check for adequate sampling, judged from trace plots, and ratio of between chain to within chain variance (*Check* : $\hat{R} \approx 1$; Gelman and Rubin, 1992; Brooks and Gelman, 1998). To fully explore posterior distributions, final models were run as a single chain for 200,000 iterations, also with 50% warmup, yielding 100,000

samples from the posterior distribution for all parameters. HDI were calculated from these distributions.

All statistical analyses were performed in R version 3.3.2 (R Core Team, 2016). Bayesian inference was carried out using the *rstan* package for R (version 2.14.1; <https://cran.r-project.org/web/packages/rstan/index.html>), an interface to Stan, a Hamiltonian Monte Carlo sampler (Carpenter et al., 2015; Hoffman and Gelman, 2014). We used the *rethinking* R package (version 1.59; <https://github.com/rmcelreath/rethinking>), which includes convenience functions for building, sampling, and summarizing models (McElreath, 2015). All code for models and analysis is included as Supporting Information.

3 | RESULTS

3.1 | Body mass

Across all strains, mice with access to a running wheel were about 4 g lighter than the sedentary controls: MM = -3.9 g (95% HDI of Wheel vs Sedentary: -6.3 to -1.5 g; Figure 4), C3H/He = -4.0 g (95% HDI: -6.9 to -1.2 g), C57BL/6 = -4.5 g (95% HDI: -6.7 to -2.2 g). This pattern of reduced body mass with access to running wheels has been previously observed in HR mice (e.g., Koteja, Swallow, Carter, & Garland, 1999; Swallow, Garland, Carter, Zhan, & Sieck, 1998b; Swallow, Koteja, Carter, & Garland, 2001; and subsequent); however, here we find a similar pattern and magnitude in the two fully inbred mouse strains as well. Descriptive statistics (means and standard errors of the mean) for all traits analyzed are included in Supporting Information, Table 1.

3.2 | Gross morphometrics and histomorphometry

Bayesian linear models for the gross morphometric (with femoral length as a covariate) and histomorphometric analyses are summarized in Figures 5 and 6 and Supporting Information, Table 1. Overall, we find that 95% HDIs of covariate estimates (i.e., slopes of trait vs femoral length) encompass the null hypothesized isometric slope value, with the exception of distal femoral width, which scales with negative allometry (Figure 5a). Across all measures of limb bone robusticity (diameters, widths, areas, and second moments of area), we find relatively few effects of activity *within* the three strains of mice (Figure 5b). Indeed, the only HDIs that did not include the null hypothesized difference of 0 were AP femoral diameter and femoral head depth in C3H/He mice. In stark contrast, we find widespread differences in morphometric traits *between* strains (Figure 6). Because the linear model included a strain by activity interaction term, differences were calculated pairwise between the strains separately for each activity group. Pairwise comparisons among both sedentary and exercise treatment groups show differences (HDIs that do not include 0) between all combinations of strains (Figure 6a,b). In general, we find that C3H/He mice have medio-laterally narrower femora than C57BL/6 or MM but are larger in cross-sectional area than C57BL/6 (but not MM). Lower cross-sectional areas in C3H/He mice lead to lower overall second moments of area than in the other two strains (Figure 6).

3.3 | Mechanical testing

Three-point bending analyses revealed moderate differences in estimated biomechanical traits but a similar overall pattern as the morphometric analyses above (Figure 7 and Supporting Information, Table 1). Within mouse strains, only 2 of 21 comparisons deviated from the null difference of 0: ultimate strain in C57BL/6 and ultimate toughness in C3H/H3 (Figure 7a). These results contrast with pairwise comparisons among the strains for sedentary and wheel access treatments (Figure 7b,c), wherein we find far-more HDIs that exclude 0 (16 of 42). These results demonstrate that estimated Young's modulus is credibly higher in C3H/He mice than in either MM or C57BL/6, and that MM is also credibly higher than C57BL/6. Furthermore, C3H/He strain mice show consistent patterns indicative of their relatively stiff bones, including reduced ultimate strain and higher modulus of toughness relative to MM, and lower failure strain than C57BL/6 (Figure 7).

3.4 | Total distance

In the wheel access treatment only, analyses of gross and histomorphometric traits and biomechanical loading traits were repeated separately for each strain with standardized total distance run as a continuous covariate. It was necessary to analyze each strain separately because strain is confounded with total distance (see "Voluntary running behavior" in Supporting Information). In all morphometric and bending traits, HDIs of the linear regression slope included the null value of zero, indicating that there was minimal association between total cumulative distance and the measured traits (Supporting Information, Figure 1). Although some median slope estimates deviated substantially from 0

(Supporting Information, Figures 2 and 3), they were associated with very wide 95% HDIs (Supporting Information, Figure 1).

4 | DISCUSSION

4.1 | Cross-sectional morphology and the exercise response

The goals of this study were to examine the connections between locomotor behavior and observed cross-sectional geometry after several weeks of "training" opportunity and the relationship between inferred and directly measured bone loading performance. Our overall aim was to better define the limits of inference about bone mechanics and locomotor behavior from observed bone morphology.

Previous studies have shown significant differences in cross-sectional properties and loading responses between C3H/He and C57BL/6 mice (Akhter et al., 2000; Kodama et al., 2000; Sheng et al., 1999), and our results agree with these and are here extended to include MM mice. Our first hypothesis, that mice with different genetic backgrounds will exhibit differences in cross-sectional morphology and histomorphometrics, was widely supported (Figure 5). For example, C3H/He mice showed the highest values for cross-sectional area, which is consistent with their observed comparative cross-sectional morphology (Figures 2e and 6). C3H/He mice also exhibited the lowest values for both anteroposterior and mediolateral second moments of area, which is also predicted given their smaller femoral diameters than C57BL/6 or MM strains (Figure 6). MM mice had higher mediolateral second moment of area values than both C3H/He and C57BL/6 mice but were similar to C57BL/6 mice in their anteroposterior second moment of area values.

In general, we find much more prevalent differences in whole-bone measurement and cross-sectional properties comparing among strains than we do within strains with exercise treatment (see below). This result is relevant to our understanding of how to best interpret observed variation in bone phenotype among archaeological populations. If bone shape and response to loading are driven more by genetic background than by exercise, as has been demonstrated by research using both inbred (Middleton et al., 2008b; Wallace et al., 2010, 2012) and outbred (Wallace, Judex, & Demes, 2015) mouse strains and suggested by studies on human populations (Cowgill, 2010; Meiring, Avidon, Norris, & McVeigh, 2013; Weiss, 2003), it becomes very difficult to extrapolate population-level changes in activity based solely on observed bone morphology.

Although we recognize that there is greater genetic variation within human populations than among them (Lewontin, 1972; Rosenberg et al., 2002), we do expect archaeological samples representative of different human populations to have some degree of among-group genetic differences, which may involve the skeletal system. In this context, we view genetically distinct mouse strains as informative for considering the expected results for archeological human populations with different genetic backgrounds. Studies have found genetic differences between archaeological populations with different subsistence strategies (Bramanti et al., 2009; Hervella et al., 2012; Skoglund et al., 2012,

2014). As these populations differ in their genetic backgrounds and in their habitual behaviors, any observed changes in their skeletal morphology cannot be definitively ascribed to one cause over the other. Therefore, we suggest caution in interpreting cross-sectional differences in fossil assemblages as resulting from different loading environments when other underlying population-level differences and selection histories remain unknown (Wallace et al., 2010, 2015; Wescott, 2006).

The second hypothesis, that mice allowed wheel access will show a corresponding bone loading response that can be quantified histomorphometrically, was not widely supported. A potential explanation for this result may be that wheel running, although a common locomotor behavior in mice (Novak, Burghardt, & Levine, 2012; Swallow et al., 1998a) does not provide an adequate loading environment (external or internal) to produce the expected morphological responses (but see Kelly et al., 2006). If the muscle forces generated by voluntary running on a wheel are not enough to induce a plastic response, then no difference would be seen between the bones of mice with wheel access and those without. In the context of our study, it is possible that the increased activity associated with the mice given wheel access was not sufficient to reach threshold for bone remodeling (Frost, 2001).

Previous studies using this mouse model system have produced mixed results, with some studies showing widespread effects of activity (Kelly et al., 2006) while others have not (Middleton et al., 2008b, 2010). These three studies differed from the one described here in use of (1) a larger diameter wheel (1.1 m circumference), which may alter gait patterns, (2) noninbred animals of both sexes and from different generations in the artificial selection protocol, and (3) the use of a different covariate for statistical analysis (here we used femur length rather than body mass, which allowed us to avoid any analytical complications arising from the impact of exercise on body mass). The inbreeding process in the MM mice used here may have reduced the magnitude of phenotypic plasticity in response to habitual exercise; however, we doubt this as other studies have found an effect of exercise on bone phenotype using inbred strains (Akhter, Cullen, Pedersen, Kimmel, & Recker, 1998; Kelly et al., 2006; Kodama et al., 2000; Plochocki, Rivera, Zhang, & Ebba, 2008).

4.2 | Relationship between cross-sectional geometry and mechanical loading behavior

Our third hypothesis, that predictions of bending mechanics based on cross-sectional geometry would reflect the results of mechanical testing, was unsupported. The bones were loaded anteroposteriorly (i.e., about the mediolateral axis), and mediolateral second moment of area is the variable used to predict loading response in that direction. Histomorphometric analysis indicated that, based on their mediolateral second moments of area, the femora of MM mice should be the most resistant to bending and C3H/He the least, with C57BL/6 mice intermediate (Figure 6). However, the results of mechanical testing indicate that C3H/He femora are the most resistant to bending, while MM and C57BL/6 femora are not different from one another (difference ≈ 0).

We found that C3H/He mice have significantly stiffer bones that are able to absorb more energy to maximum strength than both MM and C57BL/6 mice (Figure 7 and Supporting Information, Table 1). This result is supported by the observation that C3H/He mice have relatively high bone mineralization (Beamer et al., 1996; Sheng et al., 1999). Our findings support other studies that indicate that mineralization is an important component of bone strength (Donnelly, Chen, Boskey, Baker, & van Der Meulen, 2010; Keller, 1994; Jepsen, Akkus, Majeska, & Nadeau, 2003; Jepsen, Silva, Vashishth, Guo, & van der Meulen, 2015; van Der Meulen, Jepsen, & Mikic, 2001).

Similarly, although the values of mediolateral second moments of area suggested that MM femora should be more resistant to bending than those of C57BL/6 mice, mechanical testing resulted in no difference in loading response between the two strains. Thus, while the cross-sectional geometry of MM mice differs grossly from that of C57BL/6 mice (Figure 2e), the bones of MM mice do not actually behave very differently when loaded. These results suggest that the changes in bone morphology associated with the MM genotype do not actually provide them with a biomechanical (i.e., locomotor fitness in terms of resistance to bending) advantage related to increased voluntary locomotion. Furthermore, lack of differences in bending resistance between C57BL/6 and MM mice suggest that differences in hind limb muscle mass and/or mass-specific contractile strength (Syme, Evashuk, Grintuch, Rezende, & Garland, 2005) may not necessarily be correlated with differences in bone biomechanical behavior in this system. The observed similarities between MM and C57BL/6 bones in response to mechanical testing despite differences in femoral cross-sectional shape suggest that MM and C57BL/6 mice may have similar mineralization profiles.

Overall, our results support the idea that estimations of bone loading response that are made based solely on cross-sectional geometry, without consideration of bone mineralization, will be incomplete (Jepsen et al., 2015; van Der Meulen et al., 2001). While the expected phenotypic variation within inbred mouse strains is more analogous to a group of identical siblings than to heterogeneous human populations, the importance of bone mineralization on loading behavior is likely to be true on a population level as well as an individual or family group level. The implication of these results is that differences in bone morphology among populations, whether driven by genetic background or plastic responses to muscle forces, may not indicate differences in bone functional properties or behavioral patterns (Cowgill, 2014; Cowgill, Warrener, Pontzer, & Ocobock, 2010; Garland et al., 2011; Maki, 2013; Osborne et al., 2012; Wallace et al., 2010; Wescott, 2006).

4.3 | Plasticity versus genetic adaptation

Our results have important implications for the interpretation of loading mechanics in archaeological and fossilized remains. Observation of bone cross-sectional properties is a technique frequently used to infer past behavior and mechanical loading history (Bridges, 1991; Dobson and Trinkaus, 2002; Lieberman et al., 2004; Ruff and Hayes, 1983, 1984; Shackelford, Marshall, & Peters, 2013; Weiss, 2005). However, as demonstrated here and elsewhere (Akhter et al., 2000; Schrieffer

et al., 2005), the mechanical properties of bone are dependent on both their size and their mineral content. As such, histomorphometric analyses without accompanying assessment or knowledge of the mineral composition of the bone in question will only convey a partial sense of its expected response to loading. Because mineralization data are not available from fossilized remains, predictions of mechanical properties made from fossil bones should be made with caution.

Skeletal morphology, including how morphology changes in response to environmental influences, is often hypothesized to be adapted to a particular function or behavior (Carter and Orr, 1992; Evteev, Cardini, Morozova, & O'Higgins, 2014; Hylander, 1979; Marchi, 2008; Meloro, Elton, Louys, Bishop, & Ditchfield, 2013; Ruff, Holt, & Trinkaus, 2006; Shaw and Stock, 2009; Susman, 1979). However, because evolutionary adaptations and individual-level plastic responses can result in similar phenotypes, it can be difficult to determine how to best interpret observed morphological patterns. Reconstructions of behavior in fossil remains present a particular challenge because, at the time of analysis, the fossilized bones have been decoupled from the selective history and/or loading environments that formed them. Individual bone phenotypes of fossils are often interpreted as being purely indicative of plastic responses to loading regimes that correspond to habitual activity patterns. However, because little is known about the genetic backgrounds of most of these populations, it is difficult to definitively determine whether the observed morphology is derived from individual-level plastic responses or population-level genetic history. Our results have shown it is quite possible for genetically based differences within a species to dwarf the degree of environmental differences. While we are only able to characterize three genotypes here and thus cannot make more general statements about the relative importance of genetic variation and environmental variation on a population wide scale, we do consider it likely that differences observed between populations reflect some degree of genetically based differences rather than differences due to behavior alone.

In other words, an osteological correlate of increased activity could either indicate that the population themselves led an active lifestyle, or that their ancestors did (Cowgill, 2010; Lieberman, 1997; Wallace et al., 2010; Weiss, 2003). Similarly, because evolutionary adaptations are mediated by the genome, a particular phenotype may be apparent even in the absence of a specific environmental stimulus. For example, Wallace et al. (2010, 2012) showed that at least some of the limb bone morphology of mice artificially selected for increased voluntary activity (the high-runner lines, from which the MM strain used here were derived) reflects evolutionary adaptation based in genetics, rather than individual plastic responses to loading. An additional complication is that an individual's capacity for plasticity is under genetic control (Garland and Kelly, 2006; Kelly, Panhuis, & Stoehr, 2012; West-Eberhard, 2003) and may also be sex-dependent. Thus, while the specific morphology itself may be the result of phenotypic plasticity, the range of possible plastic responses may be genetically constrained. These factors are important to consider when interpreting behavior from the fossil record, because they demonstrate that alterations in behavior do not necessarily correspond directly to alterations in bone morphology.

5 | CONCLUSION

Our results highlight two main concerns associated with inferring past behaviors from the cross-sectional morphology of fossilized bones. First, we found that observed differences in femoral morphology were more reflective of differences in genetic background than differences in activity. Second, we found that our predictions of bone behavior under loading, based solely on femoral cross-sectional geometry and without consideration of genetically based differences in bone material properties, were incorrect. Together these results suggest that analysis of skeletal morphology alone is not sufficient to predict loading behavior. Our intention here is not to suggest that skeletal geometry has no impact on how a bone behaves when loaded; we recognize the importance of bone shape and the role of plastic adaptation in adjusting skeletal morphology in response to strain. Rather, our goal is to draw attention to the limitations associated with reconstructing past behaviors from fossilized remains and to help underscore the importance of bone mineral content for understanding bone strength.

Tests of the underlying genetic or environmental determinants of bone morphology, such as those discussed previously, are generally not possible for fossils. Instead, comparisons between the morphology of fossil populations and populations with known behaviors, such as the studies by Stock and Pfeiffer (2001) and Shaw and Stock (2013), may provide a viable alternative. However, such analyses should be made with the understanding that they are incomplete. Further studies of bone loading response in controlled experimental settings using animals with known genetic backgrounds will be critical in furthering our understanding of phenotypic variation in fossil populations.

ACKNOWLEDGEMENTS

The authors wish to thank Cindy Chrisler for assistance with animal husbandry. They also acknowledge Monica Ahsan, Elizabeth King, and Karthik Panchanathan for fruitful discussion of Bayesian analysis. Kristina Aldridge, Scott Maddux, Ian Wallace, and an anonymous reviewer provided helpful comments on earlier drafts of this article.

ORCID

Sarah J. Peacock  <http://orcid.org/0000-0002-1064-9482>

Theodore Garland, Jr.  <http://orcid.org/0000-0002-7916-3552>

Kevin M. Middleton  <http://orcid.org/0000-0003-4704-1064>

REFERENCES

- Agostini, G. M., & Ross, A. H. (2011). The effect of weight on the femur: A cross-sectional analysis*. *Journal of Forensic Sciences*, *56*, 339–343.
- Akhter, M. P., Cullen, D. M., Pedersen, E. A., Kimmel, D. B., & Recker, R. R. (1998). Bone response to in vivo mechanical loading in two breeds of mice. *Calcified Tissue International*, *63*, 442–449.
- Akhter, M. P., Iwaniec, U. T., Covey, M. A., Cullen, D. M., Kimmel, D. B., & Recker, R. R. (2000). Genetic variations in bone density, histomorphometry, and strength in mice. *Calcified Tissue International*, *67*, 337–344.
- Beamer, W. G., Donahue, L. R., Rosen, C. G., & Baylink, D. J. (1996). Genetic variability in adult bone density among inbred strains of mice. *Bone*, *18*, 397–403.

- Beamer, W. G., Donahue, L. R., & Rosen, C. J. (2002). Genetics and bone: Using the mouse to understand man. *Journal of Musculoskeletal and Neuronal Interactions*, 2, 225–231.
- Bilodeau, G. M., Guderley, H., Joanisse, D. R., & Garland, T. (2009). Reduction of type IIb myosin and IIB fibers in tibialis anterior muscle of mini-muscle mice from high-activity lines. *Journal of Experimental Zoology. Part A, Ecological Genetics and Physiology*, 311, 189–198.
- Bouxsein, M. L., Myburgh, K., van der Meulen, M. C. H., Linderberger, E., & Marcus, R. (1994). Age-related differences in cross-sectional geometry of the forearm bones in healthy women. *Calcified Tissue International*, 54, 113–118.
- Bramanti, B. T., Haak, Antanaitis-Jacobs, Haidle, I., Lueth, C. J., Hiller, T., ... Burger, P. (2009). Genetic discontinuity between local hunter-gatherers and Europe's first farmers. *Science*, 326, 137–140.
- Bridges, P. S. (1991). Skeletal evidence of changes in subsistence activities between the archaic and Mississippian time periods in northwestern Alabama. In M. L. Powell, P. S. Bridges, A. M. W. Mires (Eds.), *What mean these bones? Studies in southeastern bioarchaeology* (pp. 89–101). The University of Alabama Press.
- Brooks, S., & Gelman, A. (1998). General methods for monitoring convergence of iterative simulations. *Journal of Computational and Graphical Statistics*, 7, 434–455.
- Burr, D. B., Martin, R. B., Schaffler, M. B., & Radin, E. L. (1985). Bone remodeling in response to in vivo fatigue microdamage. *Journal of Biomechanics*, 18, 189–200.
- Carlson, K. J., Lublinsky, S., & Judex, S. (2008). Do different locomotor modes during growth modulate trabecular architecture in the murine hind limb? *Integrative and Comparative Biology*, 48, 385–393.
- Carpenter, B., Gelman, A., Hoffman, M., Lee, D., Goodrich, B., Betancourt, M., Brubaker, M., Guo, J., Li, P., & Riddell, A. (2015). Stan: A probabilistic programming language. *Journal of Statistical Software*.
- Carter, D. R., & Orr, T. E. (1992). Skeletal development and bone functional adaptation. *Journal of Bone and Mineral Research*, 7 Suppl 2, S389–S395.
- Cowgill, L. W. (2010). The ontogeny of Holocene and Late Pleistocene human postcranial strength. *American Journal of Physical Anthropology*, 141, 16–37.
- Cowgill, L. W. (2014). Femoral diaphyseal shape and mobility: An ontogenetic perspective. In: K. J. Carlson, D. Marchi (Eds.), *Reconstructing mobility* (pp. 193–208). Boston, MA: Springer US.
- Cowgill, L. W., Warrener, A., Pontzer, H., & Ocobock, C. (2010). Waddling and toddling: The biomechanical effects of an immature gait. *American Journal of Physical Anthropology*, 143, 52–61.
- Demes, B., Jungers, W. L., & Selpien, K. (1991). Body size, locomotion, and long bone cross-sectional geometry in indriid primates. *American Journal of Physical Anthropology*, 86, 537–547.
- Demes, B., Jungers, W. L., & Walker, C. (2000). Cortical bone distribution in the femoral neck of strepsirhine primates. *Journal of Human Evolution*, 39, 367–379.
- Demes, B., Stern, J. T., Hausman, M. R., Larson, S. G., Mcleod, K. J., & Rubin, C. T. (1998). Patterns of strain in the macaque ulna during functional activity. *American Journal of Physical Anthropology*, 106, 87–100.
- Dobson, S. D., & Trinkaus, E. (2002). Cross-sectional geometry and morphology of the mandibular symphysis in Middle and Late Pleistocene Homo. *Journal of Human Evolution*, 43, 67–87.
- Donnelly, E., Chen, D. X., Boskey, A. L., Baker, S. P., & van Der Meulen, M. C. H. (2010). Contribution of mineral to bone structural behavior and tissue mechanical properties. *Calcified Tissue International*, 87, 450–460.
- Doube, M., Kłosowski, M. M., Arganda-Carreras, I., Cordelières, F. P., Dougherty, R. P., Jackson, J. S., ... Shefelbine, S. J. (2010). BoneJ: Free and extensible bone image analysis in ImageJ. *Bone*, 47, 1076–1079.
- Duncan, R. L., & Turner, C. H. (1995). Mechanotransduction and the functional response of bone to mechanical strain. *Calcified Tissue International*, 57, 344–358.
- Evteev, A., Cardini, A. L., Morozova, I., & O'Higgins, P. (2014). Extreme climate, rather than population history, explains mid-facial morphology of Northern Asians. *American Journal of Physical Anthropology*, 153, 449–462.
- Frost, H. M. (2001). From Wolff's law to the Utah paradigm: Insights about bone physiology and its clinical applications. *Anatomical Record*, 262, 398–419.
- Garland, T., Jr., & Freeman, P. W. (2005). Selective breeding for high endurance running increases hindlimb symmetry. *Evolution (New York)*, 59, 1851–1854.
- Garland, T., Jr., & Kelly, S. (2006). Phenotypic plasticity and experimental evolution. *Journal of Experimental Biology*, 209, 2344–2361.
- Garland, T., Jr, Kelly, S. A., Malisch, J. L., Kolb, E. M., Hannon, R. M., Keeney, B. K., ... Middleton, K. M. 2011. How to run far: multiple solutions and sex-specific responses to selective breeding for high voluntary activity levels. *Proc Biol Sci*, 278:574–581.
- Gelman, A., Hill, J., & Yajima, M. (2012). Why we (usually) don't have to worry about multiple comparisons. *Journal of Research on Educational Effectiveness*, 5, 189–211.
- Gelman, A., & Rubin, D. (1992). A single series from the Gibbs sampler provides a false sense of security. *Bayesian Statistics*, 4, 625–631.
- Gelman, A., & Tuerlinckx, F. (2000). Type S error rates for classical and Bayesian single and multiple comparison procedures. *Computational Statistics*, 15, 373–390.
- Gere, J., & Timoshenko, S. (2001). *Mechanics of materials*. In Brooks (Ed.). Pacific Grove, CA: Cole.
- Gilbert, R., & Gill, G. W. (1990). Metric technique for identifying American Indian femora. In G. W. Gill, S. Rhine (Eds.), *Skeletal attribution of race* (pp. 97–99). Albuquerque, NM: Maxwell Museum of Anthropology.
- Green, E. L. (1981). *Genetics and probability in animal breeding experiments: A primer and reference book on probability, segregation, assortment, linkage and mating systems for biomedical scientists who breed and use genetically defined laboratory animals for research*. New York: Oxford University Press.
- Gross, T. S., Poliachik, S. L., Prasad, J., & Bain, S. D. (2010). The effect of muscle dysfunction on bone mass and morphology. *Journal of Musculoskeletal and Neuronal Interactions*, 10, 25–34.
- Guderley, H., Houle-Leroy, P., Diffie, G. M., Camp, D. M., & Garland, T. (2006). Morphometry, ultrastructure, myosin isoforms, and metabolic capacities of the “mini muscles” favoured by selection for high activity in house mice. *Comparative Biochemistry and Physiology Part B: Biochemistry & Molecular Biology*, 144, 271–282.
- Guderley, H., Joanisse, D. R., Moka, S., Bilodeau, G. M., & Garland, T. (2008). Altered fibre types in gastrocnemius muscle of high wheel-running selected mice with mini-muscle phenotypes. *Comparative Biochemistry and Physiology Part B: Biochemistry & Molecular Biology*, 149, 490–500.
- Hallgrímsson, B., & Lieberman, D. E. (2008). Mouse models and the evolutionary developmental biology of the skull. *Integrative and Comparative Biology*, 48, 373–384.
- Hamrick, M. W., Skedros, J., Pennington, C., & McNeil, P. (2006). Increased osteogenic response to exercise in metaphyseal versus

- diaphyseal cortical bone. *Journal of Musculoskeletal and Neuronal Interactions*, 6, 258–263.
- Hervella, M., Izagirre, N., Alonso, S., Fregel, R., Alonso, A., Cabrera, V. M., ... Hofreiter, M. (2012). Ancient DNA from hunter-gatherer and farmer groups from Northern Spain supports a random dispersion model for the neolithic expansion into Europe. *PLoS One*, 7, e34417.
- Hoffman, M., & Gelman, A. (2014). The no-U-turn sampler: Adaptively setting path lengths in Hamiltonian Monte Carlo. *Journal of Machine Learning Research*, 15, 1593–1623.
- Holt, B. M. (2003). Mobility in Upper Paleolithic and Mesolithic Europe: Evidence from the lower limb. *American Journal of Physical Anthropology*, 122, 200–215.
- Houle-Leroy, P., Guderley, H., Swallow, J. G., & Garland, T. (2003). Artificial selection for high activity favors mighty mini-muscles in house mice. *American Journal of Physiology-Regulatory, Integrative and Comparative Physiology*, 284, R433–R443.
- Hylander, W. L. (1979). The functional significance of primate mandibular form. *Journal of Morphology*, 160, 223–240.
- Jepsen, K. J., Akkus, O., Majeska, R. J., & Nadeau, J. H. (2003). Hierarchical relationship between bone traits and mechanical properties in inbred mice. *Mammalian Genome*, 14, 97–104.
- Jepsen, K. J., Silva, M. J., Vashishth, D., Guo, X. E., & van der Meulen, M. C. (2015). Establishing biomechanical mechanisms in mouse models: Practical guidelines for systematically evaluating phenotypic changes in the diaphyses of long bones. *Journal of Bone and Mineral Research*, 30, 951–966.
- Keller, T. S. (1994). Predicting the compressive mechanical behavior of bone. *Journal of Biomechanics*, 27, 1159–1168.
- Kelly, S. A., Czech, P. P., Wight, J. T., Blank, K. M., & Garland Jr, T. (2006). Experimental evolution and phenotypic plasticity of hindlimb bones in high-activity house mice. *Journal of Morphology*, 374, 360–374.
- Kelly, S. A., Panhuis, T. M., & Stoehr, A. M. (2012). Phenotypic plasticity: Molecular mechanisms and adaptive significance. *Comprehensive Physiology*, 2, 1417–1439.
- Kodama, Y., Umemura, Y., Nagasawa, S., Beamer, W. G., Donahue, L. R., Rosen, C. R., ... Farley, J. R. (2000). Exercise and mechanical loading increase periosteal bone formation and whole bone strength in C57BL/6J mice but not in C3H/HeJ mice. *Calcified Tissue International*, 66, 298–306.
- Konigsberg, L. W., & Frankenberg, S. (2013). Bayes in biological anthropology. *American Journal of Physical Anthropology*, 152 Suppl, 153–184.
- Koteja, P., Swallow, J. G., Carter, P., & Garland Jr, T. (1999). Energy cost of wheel running in house mice: Implications for coadaptation of locomotion and energy budgets. *Physiological and Biochemical Zoology*, 72, 238–249.
- Kruschke, J. (2015). *Doing Bayesian data analysis: A tutorial with R, JAGS, and Stan*. Boston, MA: Academic Press.
- Lanyon, L. E. (1992). The success and failure of the adaptive response to functional load-bearing in averting bone fracture. *Bone*, 13, S17–S21.
- Lewontin, R. (1972). The apportionment of human diversity. *Evolutionary Biology*, 6, 381–398.
- Lieberman, D. E. (1997). Making behavioral and phylogenetic inferences from hominid fossils: Considering the developmental influence of mechanical forces. *Annual Review of Anthropology*, 26, 185–210.
- Lieberman, D. E. (2003). Optimization of bone growth and remodeling in response to loading in tapered mammalian limbs. *Journal of Experimental Biology*, 206, 3125–3138.
- Lieberman, D. E., Polk, J. D., & Demes, B. (2004a). Predicting long bone loading from cross-sectional geometry. *American Journal of Physical Anthropology*, 123, 156–171.
- Maggiano, I. S., Schultz, M., Kierdorf, H., Sosa, T. S., Maggiano, C. M., & Tiesler Blos, V. (2008). Cross-sectional analysis of long bones, occupational activities and long-distance trade of the Classic Maya from Xcambó-archaeological and osteological evidence. *American Journal of Physical Anthropology*, 136, 470–477.
- Maki, J. M. (2013). The biomechanics of spear throwing: An analysis of the effects of anatomical variation on throwing performance, with implications for the fossil record (unpublished doctoral dissertation). St. Louis: Washington University.
- Marchi, D. (2008). Relationships between lower limb cross-sectional geometry and mobility: The case of a Neolithic sample from Italy. *American Journal of Physical Anthropology*, 137, 188–200.
- McElreath, R. (2015). *Statistical rethinking: A Bayesian course with examples in R and Stan*. Boca Raton: CRC Press.
- Meakin, L. B., Udeh, C., Galea, G. L., Lanyon, L. E., & Price, J. S. (2015). Exercise does not enhance aged bone's impaired response to artificial loading in C57Bl/6 mice. *Bone*, 81, 47–52.
- Meijer, J., & Robbers, Y. (2014). Wheel running in the wild. *Proceedings of the Royal Society of London B: Biological Sciences*, 281, 20140210.
- Meiring, R. M., Avidon, I., Norris, S. A., & McVeigh, J. A. (2013). A two-year history of high bone loading physical activity attenuates ethnic differences in bone strength and geometry in pre-/early pubertal children from a low-middle income country. *Bone*, 57, 522–530.
- Meloro, C., Elton, S., Louys, J., Bishop, L. C., & Ditchfield, P. (2013). Cats in the forest: Predicting habitat adaptations from humerus morphology in extant and fossil Felidae (Carnivora). *Paleobiology*, 39, 323–344.
- van Der Meulen, M. C. H., Jepsen, K. J., & Mikic, B. (2001). Understanding bone strength: Size isn't everything. *Bone*, 29, 101–104.
- Middleton, K. M., Goldstein, B. D., Guduru, P. R., Waters, J. F., Kelly, S. A., & Swartz, S. M., Jr T. G. (2010). Variation in within-bone stiffness measured by nanoindentation in mice bred for high levels of voluntary wheel running. *Journal of Anatomy*, 216, 121–131.
- Middleton, K. M., Kelly, S. A., & Garland, T. Jr (2008a). Selective breeding as a tool to probe skeletal response to high voluntary locomotor activity in mice. *Integrative and Comparative Biology*, 48, 394–410.
- Middleton, K. M., Shubin, C. E., Moore, D. C., Carter, P. A., Garland, T., Jr., & Swartz, S. M. (2008b). The relative importance of genetics and phenotypic plasticity in dictating bone morphology and mechanics in aged mice: Evidence from an artificial selection experiment. *Zoology*, 111, 135–147.
- Morgan, T., Garland, T., Jr., & Carter, P. A. (2003). Ontogenies in mice selected for high voluntary wheel-running activity. I. Mean ontogenies. *Evolution (New York)*, 57, 646–657.
- Moss, M. L. (1997). The functional matrix hypothesis revisited. 1. The role of mechanotransduction. *American Journal of Orthodontics and Dentofacial Orthopedics*, 112, 8–11.
- Muggeo, V. M. R. (2003). Estimating regression models with unknown break-points. *Statistics in Medicine*, 22, 3055–3071.
- Nakagawa, S., & Schielzeth, H. (2010). Repeatability for Gaussian and non-Gaussian data: A practical guide for biologists. *Biological Reviews*, 85, 935–956.
- Niinimäki, S., Narra, N., Harkonen, L., Abe, S., Nikander, R., Hyttinen, J., Knusel, C., & Sievanen, H. (2017). The relationship between loading history and proximal femoral diaphysis cross-sectional geometry. *American Journal of Human Biology*, 1–14.
- Novak, C. M., Burghardt, P. R., & Levine, J. A. (2012). The use of a running wheel to measure activity in rodents: Relationship to energy balance, general activity, and reward. *Neuroscience & Biobehavioral Reviews*, 36, 1001–1014.

- Osborne, D. L., Weaver, C. M., McCabe, L. D., McCabe, G. P., Novotny, R., Van Loan, M. D., ... Savaiano, D. A. (2012). Body size and pubertal development explain ethnic differences in structural geometry at the femur in Asian, Hispanic, and white early adolescent girls living in the U.S. *Bone*, *51*, 888–895.
- Pearson, O. M. (2000). Activity, climate, and postcranial robusticity implications for modern human origins and activity, climate, and postcranial robusticity implications for modern human origins and scenarios of adaptive change. *Current Anthropology*, *41*, 569–607.
- Plochocki, J. H., Rivera, J. P., Zhang, C., & Ebba, S. A. (2008). Bone modeling response to voluntary exercise in the hindlimb of mice. *Journal of Morphology*, *269*, 313–318.
- Polk, J. D., Demes, B., Jungers, W. L., Biknevicius, A. R., Heinrich, R. E., & Runestad, J. A. (2000). A comparison of primate, carnivoran and rodent limb bone cross-sectional properties: Are primates really unique?. *Journal of Human Evolution*, *39*, 297–325.
- R Core Team. (2017). R: A Language and Environment for Statistical Computing. ISBN 3-900051-07-0. Available from <http://www.r-project.org/>, Vienna, Austria.
- Rosenberg, N. A., Pritchard, J. K., Weber, J. L., Cann, H. M., Kidd, K. K., Zhivotovskiy, L. A., & Feldman, M. W. (2002). Genetic structure of human populations. *Science*, *298*, 2381–2385.
- Rubin, C. T. (1984). Skeletal strain and the functional significance of bone architecture. *Calcified Tissue International*, *36*, S11–S18.
- Rubin, C. T., & Lanyon, L. E. (1985). Regulation of bone mass by mechanical strain magnitude. *Calcified Tissue International*, *37*, 411–417.
- Ruff, C. (2009). Relative limb strength and locomotion in *Homo habilis*. *American Journal of Physical Anthropology*, *138*, 90–100.
- Ruff, C., Holt, B., & Trinkaus, E. (2006). Who's afraid of the big bad Wolff?: "Wolff's law" and bone functional adaptation. *American Journal of Physical Anthropology*, *129*, 484–498.
- Ruff, C. B., & Hayes, W. C. (1983). Cross-sectional geometry of Pecos Pueblo femora and tibiae—a biomechanical investigation: II. Sex, age, side differences. *American Journal of Physical Anthropology*, *60*, 383–400.
- Ruff, C. B., & Hayes, W. C. (1984). Bone-mineral content in the lower limb. Relationship to cross-sectional geometry. *Journal of Bone and Joint Surgery. American Volume*, *66*, 1024–1031.
- Ruff, C. B., Larsen, C. S., & Hayes, W. C. (1984). Structural changes in the femur with the transition to agriculture on the Georgia coast. *American Journal of Physical Anthropology*, *64*, 125–136.
- Schmitt, D., Zumwalt, A. C., & Hamrick, M. W. (2010). The relationship between bone mechanical properties and ground reaction forces in normal and hypermuscular mice. *Journal of Experimental Zoology. Part A, Ecological Genetics and Physiology*, *313*, 339–351.
- Schriefer, J. L., Robling, A. G., Warden, S. J., Fournier, A. J., Mason, J. J., & Turner, C. H. (2005). A comparison of mechanical properties derived from multiple skeletal sites in mice. *Journal of Biomechanics*, *38*, 467–475.
- Shackelford, L., Marshall, F., & Peters, J. (2013). Identifying donkey domestication through changes in cross-sectional geometry of long bones. *Journal of Archaeological Science*, *40*, 4170–4179.
- Shaw, C. N., & Stock, J. T. (2009). Intensity, repetitiveness, and directionality of habitual adolescent mobility patterns influence the tibial diaphysis morphology of athletes. *American Journal of Physical Anthropology*, *140*, 149–159.
- Shaw, C. N., & Stock, J. T. (2013). Extreme mobility in the Late Pleistocene? Comparing limb biomechanics among fossil *Homo*, varsity athletes and Holocene foragers. *Journal of Human Evolution*, *64*, 242–249.
- Sheng, M. H. C., Baylink, D. J., Beamer, W. G., Donahue, L. R., Rosen, C. J., Lau, K. H. W., & Wergedal, J. E. (1999). Histomorphometric studies show that bone formation and bone mineral apposition rates are greater in C3H/HeJ (high-density) than C57BL/6J (low-density) mice during growth. *Bone*, *25*, 421–429.
- Simons, E. L. R., Hieronymus, T. L., & O'Connor, P. M. (2011). Cross sectional geometry of the forelimb skeleton and flight mode in pelecaniiform birds. *Journal of Morphology*, *272*, 958–971.
- Skoglund, P., Malmstrom, H., Omrak, A., Raghavan, M., Valdiosera, C., Gunther, T., ... Jakobsson, M. (2014). Genomic diversity and admixture differs for stone-age Scandinavian foragers and farmers. *Science*, *344*, 747–750.
- Skoglund, P., Malmstrom, H., Raghavan, M., Stora, J., Hall, P., Willerslev, E., ... Jakobsson, M. (2012). Origins and genetic legacy of neolithic farmers and hunter-gatherers in Europe. *Science*, *336*, 466–469.
- Stock, J., & Pfeiffer, S. (2001). Linking structural variability in long bone diaphyses to habitual behaviors: Foragers from the southern African Later Stone Age and the Andaman Islands. *American Journal of Physical Anthropology*, *115*, 337–348.
- Sugiyama, T., Meakin, L. B., Browne, W. J., Galea, G. L., Price, J. S., & Lanyon, L. E. (2012). Bones' adaptive response to mechanical loading is essentially linear between the low strains associated with disuse and the high strains associated with the lamellar/woven bone transition. *Journal of Bone and Mineral Research*, *27*, 1784–1793.
- Susman, R. L. (1979). Comparative and functional morphology of hominoid fingers. *American Journal of Physical Anthropology*, *50*, 215–236.
- Swallow, J. G., Carter, P. A., & Garland, T. (1998a). Artificial selection for increased wheel-running behavior in house mice. *Behavioral Genetics*, *28*, 227–237.
- Swallow, J. G., Garland, T., Jr Carter, P., Zhan, W., & Sieck, G. (1998b). Effects of voluntary activity and genetic selection on aerobic capacity in house mice (*Mus domesticus*). *Journal of Applied Physiology*, *84*, 69–76.
- Swallow, J. G., Koteja, P., Carter, P., & Garland Jr, T. (2001). Food consumption and body composition in mice selected for high wheel-running activity. *Journal of Comparative Physiology*, *171*, 651–659.
- Syme, D. A., Evashuk, K., Grintuch, B., Rezende, E. L., & Garland, T. (2005). Contractile abilities of normal and "mini" triceps surae muscles from mice (*Mus domesticus*) selectively bred for high voluntary wheel running. *Journal of Applied Physiology*, *99*, 1308–1316.
- Tallman, S. D., & Winburn, A. P. (2015). Forensic applicability of femur subtrochanteric shape to ancestry assessment in Thai and White American males. *Journal of Forensic Sciences*, *60*, 1283–1289.
- Talmadge, R. J., Acosta, W., & Garland, T. (2014). Myosin heavy chain isoform expression in adult and juvenile mini-muscle mice bred for high-voluntary wheel running. *Mechanisms of Development*, *134*, 16–30.
- Trinkaus, E., & Ruff, C. B. (1999). Diaphyseal cross-sectional geometry of near eastern middle palaeolithic humans: The tibia. *Journal of Archaeological Science*, *26*, 1289–1300.
- Turner, C. H. (1989). Yield behavior of bovine cancellous bone. *Journal of Biomechanics Engineering*, *111*, 256–260.
- Turner, C. H., & Burr, D. B. (2001). Experimental techniques for bone mechanics. In *Bone mechanics handbook*. 7th ed. p 1–35.
- Wallace, I. J., Demes, B., & Judex, S. (2017). Ontogenetic and genetic influences on bone's responsiveness to mechanical signals. In C. J. Percival, J. T. Richtmeier (Eds.), *Building bones: Bone formation and development in anthropology* (pp. 205–232). Cambridge: Cambridge University Press.
- Wallace, I. J., Demes, B., Mongle, C., Pearson, O. M., Polk, J. D., & Lieberman, D. E. (2014). Exercise-induced bone formation is poorly linked to local strain magnitude in the sheep tibia. *PLoS One*, *9*, e99108.

- Wallace, I. J., Judex, S., & Demes, B. (2015). Effects of load-bearing exercise on skeletal structure and mechanics differ between outbred populations of mice. *Bone*, 72, 1–8.
- Wallace, I. J., Middleton, K. M., Lublinsky, S., Kelly, S. A., Judex, S., Garland, T., & Demes, B. (2010). Functional significance of genetic variation underlying limb bone diaphyseal structure. *American Journal of Physical Anthropology*, 143, 21–30.
- Wallace, I. J., Tommasini, S. M., Judex, S., Garland, T., & Demes, B. (2012). Genetic variations and physical activity as determinants of limb bone morphology: An experimental approach using a mouse model. *American Journal of Physical Anthropology*, 148, 24–35.
- Weaver, T. D. (2003). The shape of the Neandertal femur is primarily the consequence of a hyperpolar body form. *Proceedings of the National Academy of Sciences*, 100, 6926–6929.
- Weiss, E. (2003). Effects of rowing on humeral strength. *American Journal of Physical Anthropology*, 121, 293–302.
- Weiss, E. (2005). Humeral cross-sectional morphology from 18th century Quebec prisoners of war: Limits to activity reconstruction. *American Journal of Physical Anthropology*, 126, 311–317.
- Wescott, D., & Srikanta, D. (2008). Testing assumptions of the Gilbert and Gill method for assessing ancestry using the femur subtrochanteric shape. *HOMO - Journal of Comparative Human Biology*, 59, 347–363.
- Wescott, D. J. (2005). Population variation in femur subtrochanteric shape. *Journal of Forensic Sciences*, 50, 286–293.
- Wescott, D. J. (2006). Effect of mobility on femur midshaft external shape and robusticity. *American Journal of Physical Anthropology*, 130, 201–213.
- West-Eberhard, M. (2003). Plasticity. In *Developmental plasticity and evolution* (pp. 34–55). New York: Oxford University Press.

SUPPORTING INFORMATION

Additional Supporting Information may be found online in the supporting information tab for this article.

How to cite this article: Peacock SJ, Coats BR, Kirkland JK, Tanner CA, Garland T Jr, Middleton KM. Predicting the bending properties of long bones: Insights from an experimental mouse model. *Am J Phys Anthropol*. 2017;00:1–14. <https://doi.org/10.1002/ajpa.23363>