

UC Riverside

UC Riverside Previously Published Works

Title

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

Permalink

<https://escholarship.org/uc/item/4bm4b3gt>

Journal

American Journal of Biological Anthropology, 165(3)

ISSN

0002-9483

Authors

Peacock, Sarah J
Coats, Brittney R
Kirkland, J Kyle
et al.

Publication Date

2018-03-01

DOI

10.1002/ajpa.23363

Peer reviewed



Published in final edited form as:

Am J Phys Anthropol. 2018 March ; 165(3): 457–470. doi:10.1002/ajpa.23363.

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 Jr⁵, and Kevin M Middleton¹

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

²Department of Basic Sciences, Saint Louis College of Pharmacy

³Department of Kinesiology, Michigan State University

⁴Blivet Solutions, Inc

⁵Department of Biology, University of California, Riverside

Abstract

Objective—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 seven-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 between 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

Mouse model; Biomechanics; Loading; Skeletal plasticity

Introduction

Bone response to loading has been a frequent research focus due to both its clinical and biological relevance (Burr et al., 1985; Lanyon, 1992; Duncan and Turner, 1995; Lieberman, 2003; Sugiyama et al., 2012; Meakin et al., 2015). Because intense and/or chronic skeletal loading can induce an anabolic response (Rubin, 1984; Rubin and Lanyon, 1985; Moss, 1997; Hamrick et al., 2006), 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 et al., 2006; Middleton et al., 2008a; b; Simons et al., 2011), including humans and other primates (Ruff et al., 1984; Demes et al., 1991, 2000; Trinkaus and Ruff, 1999; Polk et al., 2000; Stock and Pfeiffer, 2001; Holt, 2003; Maggiano et al., 2008; Ruff, 2009; Shaw and Stock, 2013). 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 et al., 2004a; Schmitt et al., 2010; Wallace et al., 2012, 2014, 2017; Maki, 2013; Niinimäki et al., 2017). Contrasting studies suggest that genetic background (Gilbert and Gill, 1990; Wescott, 2005, 2006; Wescott and Srikanta, 2008; Cowgill, 2014; Tallman and Winburn, 2015), age (Bouxsein et al., 1994), adaptations to climate (Pearson, 2000; Weaver, 2003), and body mass (Agostini and Ross, 2011) may have more profound influences on cross-sectional geometry.

The present 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 et al., 2002; Carlson et al., 2008; Hallgrímsson and Lieberman, 2008; Middleton et al., 2008a; Gross et al., 2010) 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^{Minimsc}* (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.

Methods

Experimental design

We compared bone phenotypes in three strains of mice: (1) high-runner (HR) mice homozygous for the *Myh4^{Minimsc}* 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 (Houle-Leroy et al., 2003; Guderley et al., 2006, 2008; Bilodeau et al., 2009; Talmadge et al., 2014). Male MM mice were chosen at weaning (~21 d.) from eight dams (two per dam) that had undergone 12 generations of inbreeding with continued selection for high levels of voluntary wheel running (see S11 for additional information about inbreeding protocol and Swallow et al. (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 et al., 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.

Voluntary wheel-running behavior

Half of each group of mice was randomly assigned to either a wheel access group or a non-wheel 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 et al., 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 ($n = 7$ non-wheel control; $n = 4$ wheel access). One MM histological section was damaged and could not be analyzed ($n = 8$ non-wheel 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 were computed using a magnet attached to the wheel 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 software. Revolutions were summed for each daily 24-hour 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 (Fig. 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.

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 (Fig. 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 (Fig. 2B). The images were then manually segmented to remove medullary contents (Fig. 2C), and a threshold was applied to yield a black and white image of the cortex (Fig. 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).

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/minute at the mid-femoral diaphysis until fracture, as determined by the point where the load rate dropped by 40% (Fig. 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 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 Jr 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, 2015), 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 (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 (Fig. 3).

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 et al., 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 SI 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 (Fig. 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 (Fig. 4). We summarized the differences between posterior distributions for comparisons of interest using the median and 95% highest 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 as well as 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 approximately 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 ($\hat{R} \approx 1$; 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 (Hoffman and Gelman, 2014; Carpenter et al., 2015). 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 Supplemental Information.

Results

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 – -1.5 g; Fig. 4), C3H/He = -4.0 g (95% HDI: -6.9 – -1.2 g), C57BL/6 = -4.5 g (95% HDI: -6.7 – -2.2 g). This pattern of reduced body mass with access to running wheels has been previously observed in HR mice (e.g., Swallow et al., 1998b, 2001; Koteja et al., 1999; 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 Supplementary Information Table 1.

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 SI 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 (Fig. 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 (Fig. 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 (Fig. 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 (HDI that do not include 0) between all combinations of strains (Fig. 6a, b). In general, we find that C3H/He mice have mediolaterally 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 (Fig. 6).

Mechanical testing

Three-point bending analyses revealed moderate differences in estimated biomechanical traits but a similar overall pattern as the morphometric analyses above (Fig. 7; SI Table 1). Within mouse strains, only two of twenty-one comparisons deviated from the null difference of 0: ultimate strain in C57BL/6 and ultimate toughness in C3H/H3 (Fig. 7a). These results contrast with pairwise comparisons among the strains for sedentary and wheel access treatments (Fig. 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 (Fig. 7).

Total distance

In the wheel access treatment only, analyses of gross and histomorphometric traits as well as 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 SI). 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 (SI Fig. 1). Although some median slope estimates deviated substantially from 0 (SI Figs. 2, 3), they were associated with very wide 95% HDIs (SI Fig. 1).

Discussion

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 (Sheng et al., 1999; Akhter et al., 2000;

Kodama et al., 2000), 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 (Fig. 5). For example, C3H/He mice showed the highest values for cross-sectional area, which is consistent with their observed comparative cross-sectional morphology (Figs. 2E, 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 (Fig. 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 et al., 2015) mouse strains as well as suggested by studies on human populations (Weiss, 2003; Cowgill, 2010; Meiring et al., 2013), 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 as well as 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 (Wescott, 2006; Wallace et al., 2010, 2015).

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 (Swallow et al., 1998a; Novak et al., 2012), 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) non-inbred 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 et al., 1998; Kodama et al., 2000; Kelly et al., 2006; Plochocki et al., 2008).

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 (Fig. 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 (Fig. 7, Supplementary 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 (Keller, 1994; Van Der Meulen et al., 2001; Jepsen et al., 2003, 2015; Donnelly et al., 2010).

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 (Fig. 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 et al., 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 (Van Der Meulen et al., 2001; Jepsen et al., 2015). 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 (Wescott, 2006; Cowgill et al., 2010; Wallace et al. 2010; Garland et al., 2011; Osborne et al., 2012; Maki, 2013; Cowgill, 2014).

Plasticity vs. 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 (Ruff and Hayes, 1983, 1984; Bridges, 1991; Dobson and Trinkaus, 2002; Lieberman et al., 2004b; Weiss, 2005; Shackelford et al., 2013). 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 (Hylander, 1979; Susman, 1979; Carter and Orr, 1992; Ruff et al., 2006; Marchi, 2008; Shaw and Stock, 2009; Meloro et al., 2013; Evteev et al., 2014). 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 (Lieberman, 1997; Weiss, 2003; Cowgill, 2010; Wallace et al., 2010). 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 and colleagues (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 (West-Eberhard, 2003; Garland Jr and Kelly, 2006; Kelly et al., 2012) 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.

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.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

Grant sponsorship: This project was supported by an EARDA award from National Institutes of Health National Institute of Arthritis and Musculoskeletal Diseases (KMM; 5G11HD052368-03), the University of Missouri Department of Pathology and Anatomical Sciences (SJP, KMM), the University of Missouri Bond Life Sciences Fellowship program (SJP), the University of Missouri Gus T. Ridgel Fellowship (SJP), the California State University, San Bernardino Department of Biology (BRC, KMM), Associated Students Incorporated of California State University, San Bernardino (BRC), the University of Missouri School of Medicine (JKK), and the National Science Foundation (TG, Jr.; IOS-1121273).

The authors wish to thank Cindy Chrisler for assistance with animal husbandry. We 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 manuscript.

References

- Agostini GM, Ross AH. The Effect of Weight on the Femur: A Cross-Sectional Analysis*. *J Forensic Sci.* 2011; 56:339–343. [PubMed: 21210806]
- Akhter MP, Cullen DM, Pedersen EA, Kimmel DB, Recker RR. Bone Response to In Vivo Mechanical Loading in Two Breeds of Mice. *Calcif Tissue Int.* 1998; 63:442–449. [PubMed: 9799831]
- Akhter MP, Iwaniec UT, Covey MA, Cullen DM, Kimmel DB, Recker RR. Genetic Variations in Bone Density, Histomorphometry, and Strength in Mice. *Calcif Tissue Int.* 2000; 67:337–344. [PubMed: 11000349]
- Beamer WG, Donahue LR, Rosen CG, Baylink DJ. Genetic Variability in Adult Bone Density Among Inbred Strains of Mice. *Bone.* 1996; 18:397–403. [PubMed: 8739896]
- Beamer WG, Donahue LR, Rosen CJ. Genetics and bone: Using the mouse to understand man. *J Musculoskelet Neuronal Interact.* 2002; 2:225–231. [PubMed: 15758440]
- Bilodeau GM, Guderley H, Joannis DR, Garland T. Reduction of type IIb myosin and IIB fibers in tibialis anterior muscle of mini-muscle mice from high-activity lines. *J Exp Zool A Ecol Genet Physiol.* 2009; 311:189–98. [PubMed: 19177556]
- Bouxsein ML, Myburgh K, van der Meulen MCH, Linderberger E, Marcus R. Age-Related Differences in Cross-Sectional Geometry of the Forearm Bones in Healthy Women. *Calcif Tissue Int.* 1994; 54:113–118. [PubMed: 8012866]
- Bramanti B, Thomas, Haak, Antanaitis-Jacobs, Haidle I, Lueth CJ, Hiller T, Matsumura J, Forster S, Burger P. Genetic discontinuity between local hunter-gatherers and Europe's first farmers. *Science.* 2009; 326:137–140. [PubMed: 19729620]
- Bridges, PS. Skeletal Evidence of Changes in Subsistence Activities Between the Archaic and Mississippian Time Periods in Northwestern Alabama. In: Powell, ML, Bridges, PS., Mires, AMW., editors. *What Mean These Bones? Studies in Southeastern Bioarchaeology.* The University of Alabama Press; 1991. p. 89-101.
- Brooks S, Gelman A. General Methods for Monitoring Convergence of Iterative Simulations. *J Comput Graph Stat.* 1998; 7:434–455.
- Burr DB, Martin RB, Schaffler MB, Radin EL. Bone remodeling in response to in vivo fatigue microdamage. *J Biomech.* 1985; 18:189–200. [PubMed: 3997903]
- Carlson KJ, Lublinsky S, Judex S. Do different locomotor modes during growth modulate trabecular architecture in the murine hind limb? *Integr Comp Biol.* 2008; 48:385–393. [PubMed: 21669800]
- Carpenter B, Gelman A, Hoffman M, Lee D, Goodrich B, Betancourt M, Brubaker M, Guo J, Li P, Riddell A. Stan: a probabilistic programming language. *J Stat Softw.* 2015
- Carter DR, Orr TE. Skeletal development and bone functional adaptation. *J Bone Miner Res.* 1992; 7(Suppl 2):S389–95. [PubMed: 1485546]

- Cowgill LW. The ontogeny of Holocene and Late Pleistocene human postcranial strength. *Am J Phys Anthropol.* 2010; 141:16–37. [PubMed: 19530138]
- Cowgill, LW. Femoral Diaphyseal Shape and Mobility: An Ontogenetic Perspective. In: Carlson, KJ., Marchi, D., editors. *Reconstructing Mobility*. Boston, MA: Springer US; 2014. p. 193–208.
- Cowgill LW, Warrener A, Pontzer H, Ocobock C. Waddling and toddling: the biomechanical effects of an immature gait. *Am J Phys Anthropol.* 2010; 143:52–61. [PubMed: 20310059]
- Demes B, Jungers WL, Selpien K. Body size, locomotion, and long bone cross-sectional geometry in indriid primates. *Am J Phys Anthropol.* 1991; 86:537–47. [PubMed: 1776660]
- Demes B, Jungers WL, Walker C. Cortical bone distribution in the femoral neck of strepsirhine primates. *J Hum Evol.* 2000; 39:367–79. [PubMed: 11006046]
- Demes B, Stern JT, Hausman MR, Larson SG, Mcleod KJ, Rubin CT. Patterns of Strain in the Macaque Ulna During Functional Activity. *Am J Phys Anthr.* 1998; 106:87–100.
- Dobson SD, Trinkaus E. Cross-sectional geometry and morphology of the mandibular symphysis in Middle and Late Pleistocene Homo. *J Hum Evol.* 2002; 43:67–87. [PubMed: 12098211]
- Donnelly E, Chen DX, Boskey AL, Baker SP, Van Der Meulen MCH. Contribution of mineral to bone structural behavior and tissue mechanical properties. *Calcif Tissue Int.* 2010; 87:450–460. [PubMed: 20730582]
- Doube M, Kłosowski MM, Arganda-Carreras I, Cordelières FP, Dougherty RP, Jackson JS, Schmid B, Hutchinson JR, Shefelbine SJ. BoneJ: Free and extensible bone image analysis in ImageJ. *Bone.* 2010; 47:1076–9. [PubMed: 20817052]
- Duncan RL, Turner CH. Mechanotransduction and the functional response of bone to mechanical strain. *Calcif Tissue Int.* 1995; 57:344–358. [PubMed: 8564797]
- Evtcev A, Cardini AL, Morozova I, O’Higgins P. Extreme climate, rather than population history, explains mid-facial morphology of Northern Asians. *Am J Phys Anthropol.* 2014; 153:449–62. [PubMed: 24374801]
- Frost HM. From Wolff’s law to the Utah paradigm: Insights about bone physiology and its clinical applications. *Anat Rec.* 2001; 262:398–419. [PubMed: 11275971]
- Garland T Jr, Freeman PW. Selective Breeding for High Endurance Running Increases Hindlimb Symmetry. *Evolution (N Y).* 2005; 59:1851–1854.
- Garland T Jr, Kelly S. Phenotypic plasticity and experimental evolution. *J Exp Biol.* 2006; 209:2344–2361. [PubMed: 16731811]
- Gelman A, Hill J, Yajima M. Why We (Usually) Don’t Have to Worry About Multiple Comparisons. *J Res Educ Eff.* 2012; 5:189–211.
- Gelman A, Rubin D. A single series from the Gibbs sampler provides a false sense of security. *Bayesian Stat.* 1992; 4:625–631.
- Gelman A, Tuerlinckx F. Type S error rates for classical and Bayesian single and multiple comparison procedures. *Comput Stat.* 2000; 15:373–390.
- Gere, J., Timoshenko, S. *Mechanics of materials*. Brooks, editor. Pacific Grove, CA: Cole; 2001.
- Gilbert, R., Gill, GW. Metric Technique For Identifying American Indian Femora. In: Gill, GW., Rhine, S., editors. *Skeletal Attribution of Race*. Albuquerque, NM: Maxwell Museum of Anthropology; 1990. p. 97–99.
- Green, EL. *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; 1981.
- Gross TS, Poliachik SL, Prasad J, Bain SD. The effect of muscle dysfunction on bone mass and morphology. *J Musculoskelet Neuronal Interact.* 2010; 10:25–34. [PubMed: 20190377]
- Guderley H, Houle-Leroy P, Diffie GM, Camp DM, Garland T. Morphometry, ultrastructure, myosin isoforms, and metabolic capacities of the “mini muscles” favoured by selection for high activity in house mice. *Comp Biochem Physiol B Biochem Mol Biol.* 2006; 144:271–82. [PubMed: 16707270]

- Guderley H, Joanisse DR, Mokas S, Bilodeau GM, Garland T. Altered fibre types in gastrocnemius muscle of high wheel-running selected mice with mini-muscle phenotypes. *Comp Biochem Physiol B Biochem Mol Biol*. 2008; 149:490–500. [PubMed: 18226573]
- Hallgrímsson B, Lieberman DE. Mouse models and the evolutionary developmental biology of the skull. *Integr Comp Biol*. 2008; 48:373–384. [PubMed: 21669799]
- Hamrick MW, Skedros J, Pennington C, McNeil P. Increased osteogenic response to exercise in metaphyseal versus diaphyseal cortical bone. *J Musculoskelet Neuronal Interact*. 2006; 6:258–263. [PubMed: 17142947]
- Hervella M, Izagirre N, Alonso S, Fregel R, Alonso A, Cabrera VM, De La Rúa CN, Hofreiter M. Ancient DNA from Hunter-Gatherer and Farmer Groups from Northern Spain Supports a Random Dispersion Model for the Neolithic Expansion into Europe. *PLoS One*. 2012; 7:e34417. [PubMed: 22563371]
- Hoffman M, Gelman A. The no-U-turn sampler: Adaptively setting path lengths in Hamiltonian Monte Carlo. *J Mach Learn Res*. 2014; 15:1593–1623.
- Holt BM. Mobility in Upper Paleolithic and Mesolithic Europe: evidence from the lower limb. *Am J Phys Anthropol*. 2003; 122:200–15. [PubMed: 14533179]
- Houle-Leroy P, Guderley H, Swallow JG, Garland T. Artificial selection for high activity favors mighty mini-muscles in house mice. *Am J Physiol Regul Integr Comp Physiol*. 2003; 284:R433–43. [PubMed: 12529284]
- Hylander WL. The functional significance of primate mandibular form. *J Morphol*. 1979; 160:223–40. [PubMed: 458862]
- Jepsen KJ, Akkus O, Majeska RJ, Nadeau JH. Hierarchical relationship between bone traits and mechanical properties in inbred mice. *Mamm Genome*. 2003; 14:97–104. [PubMed: 12584605]
- Jepsen KJ, Silva MJ, Vashishth D, Guo XE, van der Meulen MC. Establishing Biomechanical Mechanisms in Mouse Models: Practical Guidelines for Systematically Evaluating Phenotypic Changes in the Diaphyses of Long Bones. *J Bone Miner Res*. 2015; 30:951–966. [PubMed: 25917136]
- Keller TS. Predicting the compressive mechanical behavior of bone. *J Biomech*. 1994; 27:1159–1168. [PubMed: 7929465]
- Kelly SA, Czech PP, Wight JT, Blank KM, Garland T Jr. Experimental Evolution and Phenotypic Plasticity of Hindlimb Bones in High-Activity House Mice. *J Morphol*. 2006; 374:360–374.
- Kelly SA, Panhuis TM, Stoeckl AM. Phenotypic plasticity: Molecular mechanisms and adaptive significance. *Compr Physiol*. 2012; 2:1417–1439. [PubMed: 23798305]
- Kodama Y, Umemura Y, Nagasawa S, Beamer WG, Donahue LR, Rosen CR, Baylink DJ, Farley JR. Exercise and mechanical loading increase periosteal bone formation and whole bone strength in C57BL/6J mice but not in C3H/HeJ mice. *Calcif Tissue Int*. 2000; 66:298–306. [PubMed: 10742449]
- Konigsberg LW, Frankenberg S. Bayes in biological anthropology. *Am J Phys Anthropol*. 2013; 152(Suppl):153–184. [PubMed: 24190509]
- Koteja P, Swallow JG, Carter P, Garland T Jr. Energy cost of wheel running in house mice: implications for coadaptation of locomotion and energy budgets. *Physiol Biochem Zool*. 1999; 72:238–249. [PubMed: 10068627]
- Kruschke, J. *Doing Bayesian data analysis: a tutorial with R, JAGS, and Stan*. Boston, MA: Academic Press; 2015.
- Lanyon LE. The success and failure of the adaptive response to functional load-bearing in averting bone fracture. *Bone*. 1992; 13:S17–S21.
- Lewontin R. The Apportionment of Human Diversity. *Evol Biol*. 1972; 6:381–398.
- Lieberman DE. Making Behavioral and Phylogenetic Inferences from Hominid Fossils: Considering the Developmental Influence of Mechanical Forces. *Annu Rev Anthropol*. 1997; 26:185–210.
- Lieberman DE. Optimization of bone growth and remodeling in response to loading in tapered mammalian limbs. *J Exp Biol*. 2003; 206:3125–3138. [PubMed: 12909694]
- Lieberman DE, Polk JD, Demes B. Predicting long bone loading from cross-sectional geometry. *Am J Phys Anthropol*. 2004a; 123:156–71. [PubMed: 14730649]

- Lieberman DE, Polk JD, Demes B. Predicting Long Bone Loading from Cross-Sectional Geometry. *Am J Phys Anthropol.* 2004b; 123:156–171. [PubMed: 14730649]
- Maggiano IS, Schultz M, Kierdorf H, Sosa TS, Maggiano CM, Tiesler Blos V. Cross-sectional analysis of long bones, occupational activities and long-distance trade of the Classic Maya from Xcambó--archaeological and osteological evidence. *Am J Phys Anthropol.* 2008; 136:470–7. [PubMed: 18383157]
- Maki, JM. Unpublished doctoral dissertation. Washington University; St. Louis: 2013. The Biomechanics of Spear Throwing: An Analysis of the Effects of Anatomical Variation on Throwing Performance, with Implications for the Fossil Record.
- Marchi D. Relationships between lower limb cross-sectional geometry and mobility: the case of a Neolithic sample from Italy. *Am J Phys Anthropol.* 2008; 137:188–200. [PubMed: 18470890]
- McElreath, R. Statistical Rethinking: A Bayesian Course with Examples in R and Stan. Boca Raton: CRC Press; 2015.
- Meakin LB, Udeh C, Galea GL, Lanyon LE, Price JS. Exercise does not enhance aged bone's impaired response to artificial loading in C57Bl/6 mice. *Bone.* 2015; 81:47–52. [PubMed: 26142929]
- Meijer J, Robbers Y. Wheel running in the wild. *Proc R Soc B.* 2014; 281:20140210.
- Meiring RM, Avidon I, Norris SA, McVeigh JA. 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.* 2013; 57:522–530. [PubMed: 24012701]
- Meloro C, Elton S, Louys J, Bishop LC, Ditchfield P. Cats in the forest: predicting habitat adaptations from humerus morphometry in extant and fossil Felidae (Carnivora). *Paleobiology.* 2013; 39:323–344.
- Van Der Meulen MCH, Jepsen KJ, Mikic B. Understanding Bone Strength: Size Isn't Everything. *Bone.* 2001; 29:101–104. [PubMed: 11502469]
- Middleton KM, Goldstein BD, Guduru PR, Waters JF, Kelly SA, Swartz SM, TG. Variation in within-bone stiffness measured by nanoindentation in mice bred for high levels of voluntary wheel running. *J Anat.* 2010; 216:121–131. [PubMed: 20402827]
- Middleton KM, Kelly SA, Garland T Jr. Selective breeding as a tool to probe skeletal response to high voluntary locomotor activity in mice. *Integr Comp Biol.* 2008a; 48:394–410. [PubMed: 21669801]
- Middleton KM, Shubin CE, Moore DC, Carter PA, Garland T Jr, Swartz SM. The relative importance of genetics and phenotypic plasticity in dictating bone morphology and mechanics in aged mice: Evidence from an artificial selection experiment. *Zoology.* 2008b; 111:135–147. [PubMed: 18221861]
- Morgan T, Garland T Jr, Carter PA. Ontogenies in Mice Selected for High Voluntary Wheel-Running Activity. I. Mean Ontogenies. *Evolution (N Y).* 2003; 57:646–657.
- Moss ML. The functional matrix hypothesis revisited. 1. The role of mechanotransduction. *Am J Orthod Dentofac Orthop.* 1997; 112:8–11.
- Muggeo VMR. Estimating regression models with unknown break-points. *Stat Med.* 2003; 22:3055–71. [PubMed: 12973787]
- Nakagawa S, Schielzeth H. Repeatability for Gaussian and non-Gaussian data: a practical guide for biologists. *Biol Rev.* 2010; 85:935–956. [PubMed: 20569253]
- Niinimäki S, Narra N, Harkonen L, Abe S, Nikander R, Hyttinen J, Knusel C, Sievanen H. The relationship between loading history and proximal femoral diaphysis cross-sectional geometry. *Am J Hum Biol.* 2017 Feb.:1–14.
- Novak CM, Burghardt PR, Levine JA. The use of a running wheel to measure activity in rodents: relationship to energy balance, general activity, and reward. *Neurosci Biobehav Rev.* 2012; 36:1001–14. [PubMed: 22230703]
- Osborne DL, Weaver CM, McCabe LD, McCabe GP, Novotny R, Van Loan MD, Going S, Matkovic V, Boushey CJ, Savaiano DA. 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.* 2012; 51:888–95. [PubMed: 22944607]
- Pearson OM. 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 1. *Curr Anthropol.* 2000; 41:569–607. [PubMed: 11624671]

- Plochocki JH, Rivera JP, Zhang C, Ebba SA. Bone Modeling Response to Voluntary Exercise in the Hindlimb of Mice. *J Morphol.* 2008; 269:313–318. [PubMed: 17957711]
- Polk JD, Demes B, Jungers WL, Biknevicius AR, Heinrich RE, Runestad JA. A comparison of primate, carnivoran and rodent limb bone cross-sectional properties: are primates really unique? *J Hum Evol.* 2000; 39:297–325. [PubMed: 10964531]
- R Core Team. R: A language and environment for statistical computing. 2015. Available from: <http://www.r-project.org/>
- Rosenberg NA, Pritchard JK, Weber JL, Cann HM, Kidd KK, Zhivotovsky LA, Feldman MW. Genetic Structure of Human Populations. *Science.* 2002; 298:2381–2385. [PubMed: 12493913]
- Rubin CT. Skeletal strain and the functional significance of bone architecture. *Calcif Tissue Int.* 1984; 36:S11–S18. [PubMed: 6430509]
- Rubin CT, Lanyon LE. Regulation of bone mass by mechanical strain magnitude. *Calcif Tissue Int.* 1985; 37:411–417. [PubMed: 3930039]
- Ruff C. Relative limb strength and locomotion in *Homo habilis*. *Am J Phys Anthropol.* 2009; 138:90–100. [PubMed: 18711733]
- Ruff C, Holt B, Trinkaus E. Who's afraid of the big bad Wolff?: "Wolff's law" and bone functional adaptation. *Am J Phys Anthropol.* 2006; 129:484–98. [PubMed: 16425178]
- Ruff CB, Hayes WC. Cross-sectional geometry of Pecos Pueblo femora and tibiae--a biomechanical investigation: II. Sex, age, side differences. *Am J Phys Anthropol.* 1983; 60:383–400. [PubMed: 6846511]
- Ruff CB, Hayes WC. Bone-mineral content in the lower limb. Relationship to cross-sectional geometry. *J Bone Joint Surg Am.* 1984; 66:1024–1031. [PubMed: 6480631]
- Ruff CB, Larsen CS, Hayes WC. Structural changes in the femur with the transition to agriculture on the Georgia coast. *Am J Phys Anthropol.* 1984; 64:125–36. [PubMed: 6465303]
- Schmitt D, Zumwalt AC, Hamrick MW. The relationship between bone mechanical properties and ground reaction forces in normal and hypermuscular mice. *J Exp Zool Part A Ecol Genet Physiol.* 2010; 313:339–351.
- Schrieffer JL, Robling AG, Warden SJ, Fournier AJ, Mason JJ, Turner CH. A comparison of mechanical properties derived from multiple skeletal sites in mice. *J Biomech.* 2005; 38:467–475. [PubMed: 15652544]
- Shackelford L, Marshall F, Peters J. Identifying donkey domestication through changes in cross-sectional geometry of long bones. *J Archaeol Sci.* 2013; 40:4170–4179.
- Shaw CN, Stock JT. Intensity, repetitiveness, and directionality of habitual adolescent mobility patterns influence the tibial diaphysis morphology of athletes. *Am J Phys Anthropol.* 2009; 140:149–59. [PubMed: 19358289]
- Shaw CN, Stock JT. Extreme mobility in the Late Pleistocene? Comparing limb biomechanics among fossil *Homo*, varsity athletes and Holocene foragers. *J Hum Evol.* 2013; 64:242–249. [PubMed: 23453436]
- Sheng MHC, Baylink DJ, Beamer WG, Donahue LR, Rosen CJ, Lau KHW, Wergedal JE. 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.* 1999; 25:421–429. [PubMed: 10511108]
- Simons ELR, Hieronymus TL, O'Connor PM. Cross sectional geometry of the forelimb skeleton and flight mode in peleciform birds. *J Morphol.* 2011; 272:958–971. [PubMed: 21567447]
- Skoglund P, Malmstrom H, Omrak A, Raghavan M, Valdiosera C, Gunther T, Hall P, Tambets K, Parik J, Sjogren K-G, Apel J, Willerslev E, Stora J, Gotherstrom A, Jakobsson M. Genomic Diversity and Admixture Differs for Stone-Age Scandinavian Foragers and Farmers. *Science.* 2014; 344:747–750. [PubMed: 24762536]
- Skoglund P, Malmstrom H, Raghavan M, Stora J, Hall P, Willerslev E, Gilbert MTP, Gotherstrom A, Jakobsson M. Origins and Genetic Legacy of Neolithic Farmers and Hunter-Gatherers in Europe. *Science.* 2012; 336:466–469. [PubMed: 22539720]
- Stock J, Pfeiffer S. Linking structural variability in long bone diaphyses to habitual behaviors: Foragers from the southern African Later Stone Age and the Andaman Islands. *Am J Phys Anthropol.* 2001; 115:337–348. [PubMed: 11471132]

- Sugiyama T, Meakin LB, Browne WJ, Galea GL, Price JS, Lanyon LE. 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. *J Bone Miner Res.* 2012; 27:1784–93. [PubMed: 22431329]
- Susman RL. Comparative and functional morphology of hominoid fingers. *Am J Phys Anthropol.* 1979; 50:215–36. [PubMed: 443358]
- Swallow JG, Carter PA, Garland T. Artificial Selection for Increased Wheel-Running Behavior in House Mice. *Behav Genet.* 1998a; 28:227–237. [PubMed: 9670598]
- Swallow JG, Garland T Jr, Carter P, Zhan W, Sieck G. Effects of voluntary activity and genetic selection on aerobic capacity in house mice (*Mus domesticus*). *J Appl Physiol.* 1998b; 84:69–76. [PubMed: 9451619]
- Swallow JG, Koteja P, Carter P, Garland T Jr. Food consumption and body composition in mice selected for high wheel-running activity. *J Comp Physiol.* 2001; 171:651–659. [PubMed: 11765974]
- Syme DA, Evashuk K, Grintuch B, Rezende EL, Garland T. Contractile abilities of normal and “mini” triceps surae muscles from mice (*Mus domesticus*) selectively bred for high voluntary wheel running. *J Appl Physiol.* 2005; 99:1308–16. [PubMed: 15947032]
- Tallman SD, Winburn AP. Forensic Applicability of Femur Subtrochanteric Shape to Ancestry Assessment in Thai and White American Males. *J Forensic Sci.* 2015; 60:1283–9. [PubMed: 25845441]
- Talmadge RJ, Acosta W, Garland T. Myosin heavy chain isoform expression in adult and juvenile mini-muscle mice bred for high-voluntary wheel running. *Mech Dev.* 2014; 134:16–30. [PubMed: 25217814]
- Trinkaus E, Ruff CB. Diaphyseal Cross-sectional Geometry of Near Eastern Middle Palaeolithic Humans: The Tibia. *J Archaeol Sci.* 1999; 26:1289–1300.
- Turner CH. Yield Behavior of Bovine Cancellous Bone. *J Biomech Eng.* 1989; 111:256–260. [PubMed: 2779192]
- Turner, CH., Burr, DB. *Bone Mechanics Handbook.* 7. 2001. Experimental Techniques for Bone Mechanics; p. 1-35.
- Wallace, IJ., Demes, B., Judex, S. Ontogenetic and Genetic Influences on Bone's Responsiveness to Mechanical Signals. In: Percival, CJ., Richtmeier, JT., editors. *Building Bones: Bone Formation and Development in Anthropology.* Cambridge: Cambridge University Press; 2017. p. 205-232.
- Wallace IJ, Demes B, Mongle C, Pearson OM, Polk JD, Lieberman DE. Exercise-induced bone formation is poorly linked to local strain magnitude in the sheep tibia. *PLoS One.* 2014; 9:e99108. [PubMed: 24897411]
- Wallace IJ, Judex S, Demes B. Effects of load-bearing exercise on skeletal structure and mechanics differ between outbred populations of mice. *Bone.* 2015; 72:1–8. [PubMed: 25460574]
- Wallace IJ, Middleton KM, Lublinsky S, Kelly SA, Judex S, Garland T, Demes B. Functional significance of genetic variation underlying limb bone diaphyseal structure. *Am J Phys Anthropol.* 2010; 143:21–30. [PubMed: 20310061]
- Wallace IJ, Tommasini SM, Judex S, Garland T, Demes B. Genetic variations and physical activity as determinants of limb bone morphology: An experimental approach using a mouse model. *Am J Phys Anthropol.* 2012; 148:24–35. [PubMed: 22331623]
- Weaver TD. The shape of the Neandertal femur is primarily the consequence of a hyperpolar body form. *PNAS.* 2003; 100:6926–6929. [PubMed: 12761384]
- Weiss E. Effects of rowing on humeral strength. *Am J Phys Anthropol.* 2003; 121:293–302. [PubMed: 12884311]
- Weiss E. Humeral cross-sectional morphology from 18th century Quebec prisoners of war: Limits to activity reconstruction. *Am J Phys Anthropol.* 2005; 126:311–7. [PubMed: 15386299]
- Wescott D, Srikanta D. Testing assumptions of the Gilbert and Gill method for assessing ancestry using the femur subtrochanteric shape. *HOMO J Comp Hum Biol.* 2008; 59:347–63.
- Wescott DJ. Population variation in femur subtrochanteric shape. *J Forensic Sci.* 2005; 50:286–293. [PubMed: 15813538]

- Wescott DJ. Effect of mobility on femur midshaft external shape and robusticity. *Am J Phys Anthropol.* 2006; 130:201–13. [PubMed: 16365861]
- West-Eberhard, M. *Developmental Plasticity and Evolution*. New York: Oxford University Press; 2003. Plasticity; p. 34-55.

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

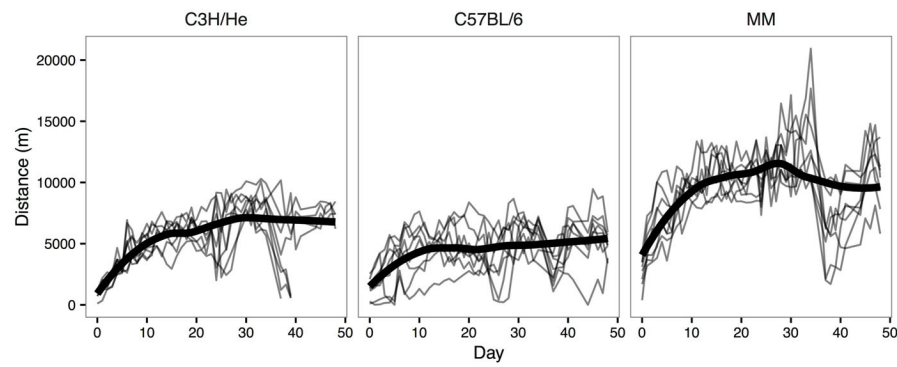


Figure 1. Mean daily locomotor distance

Individual (thin lines) and mean per strain (thick lines) running distances for three mouse strains during the seven-week experimental period. Mean total distance in MM mice across the experimental period was 419 km, 285 km in C3H/He, and 199 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.

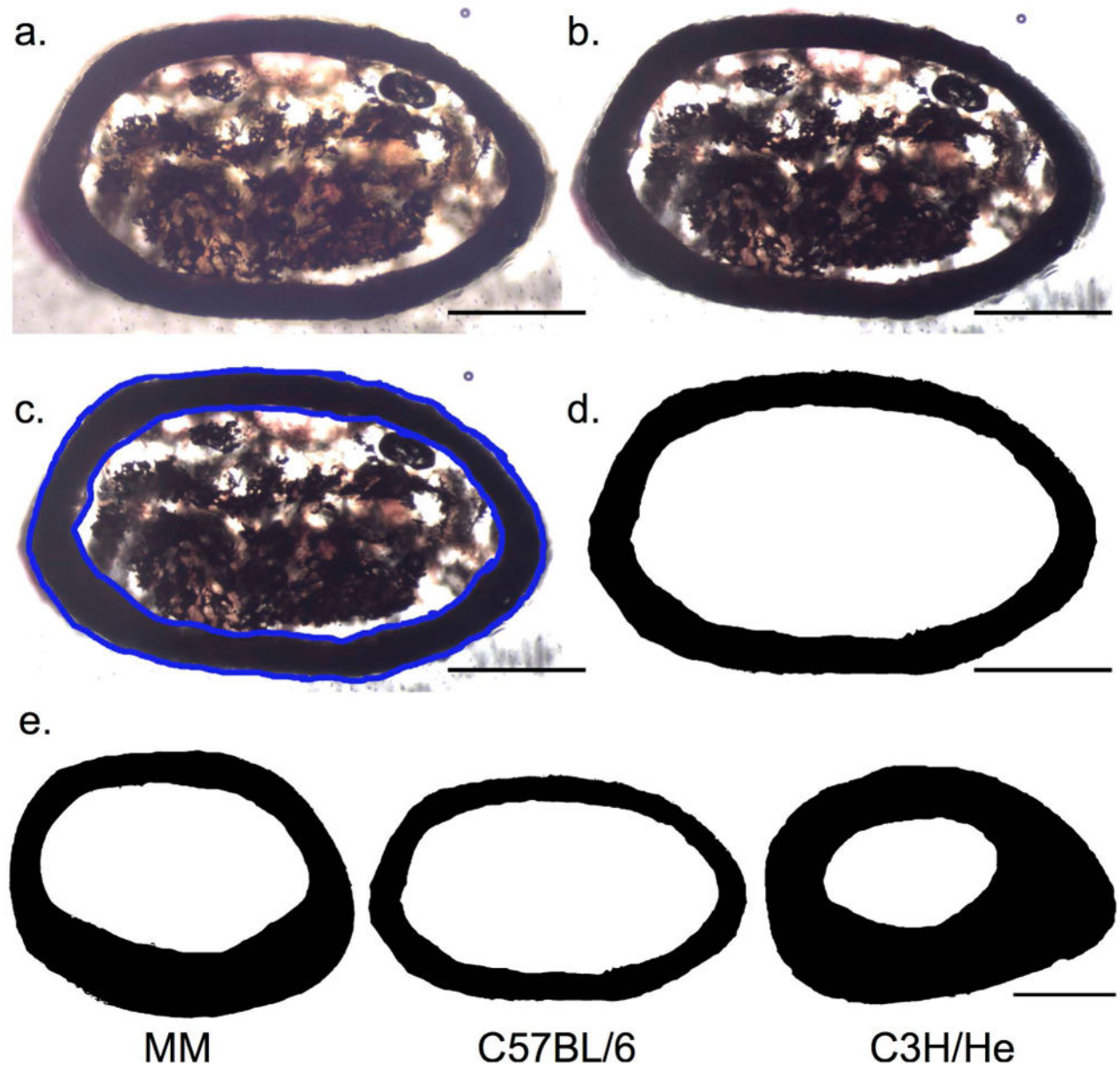


Figure 2. Comparative femoral cross-sections

(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.

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

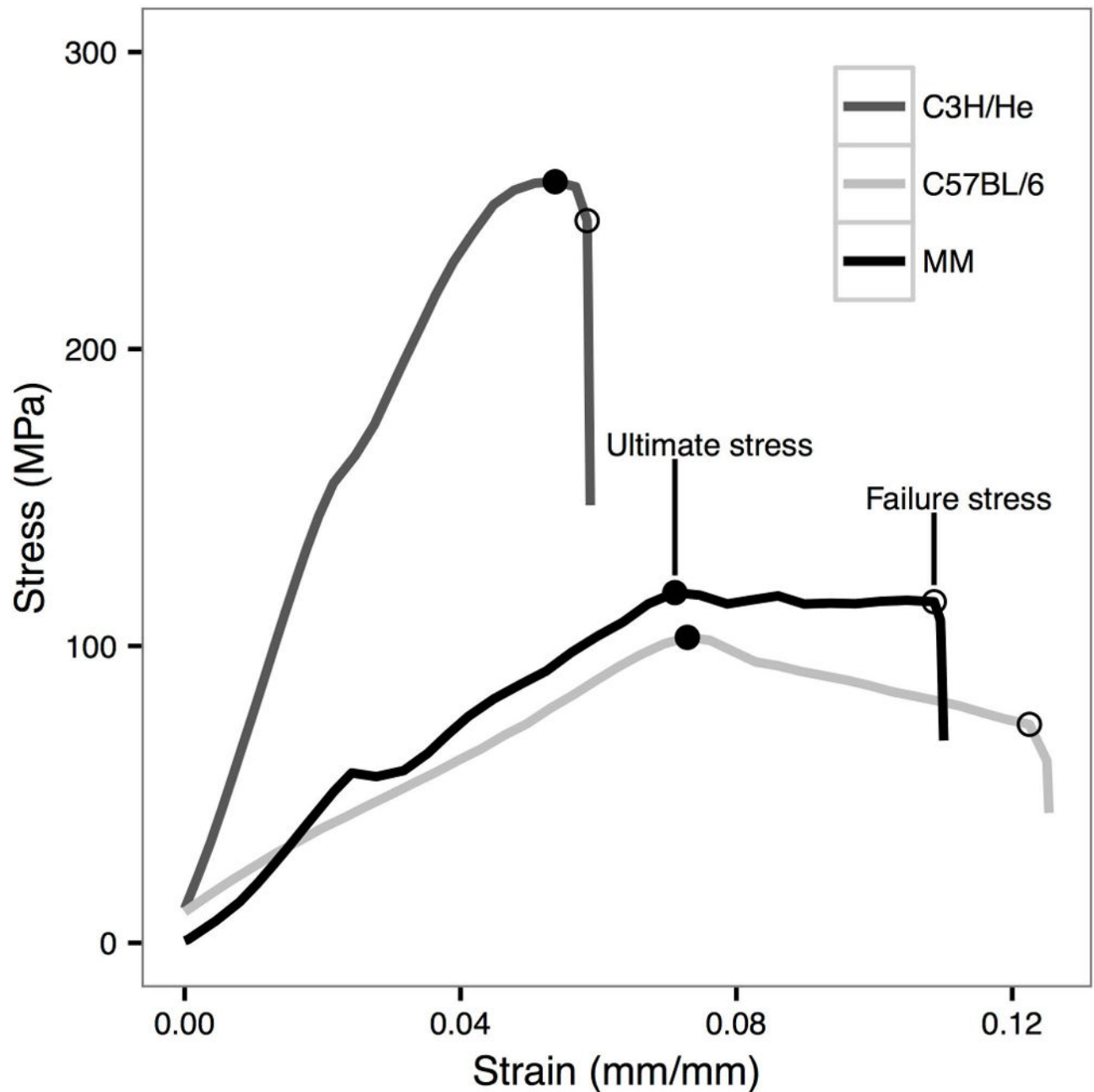


Figure 3. Comparative stress-strain curves

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.

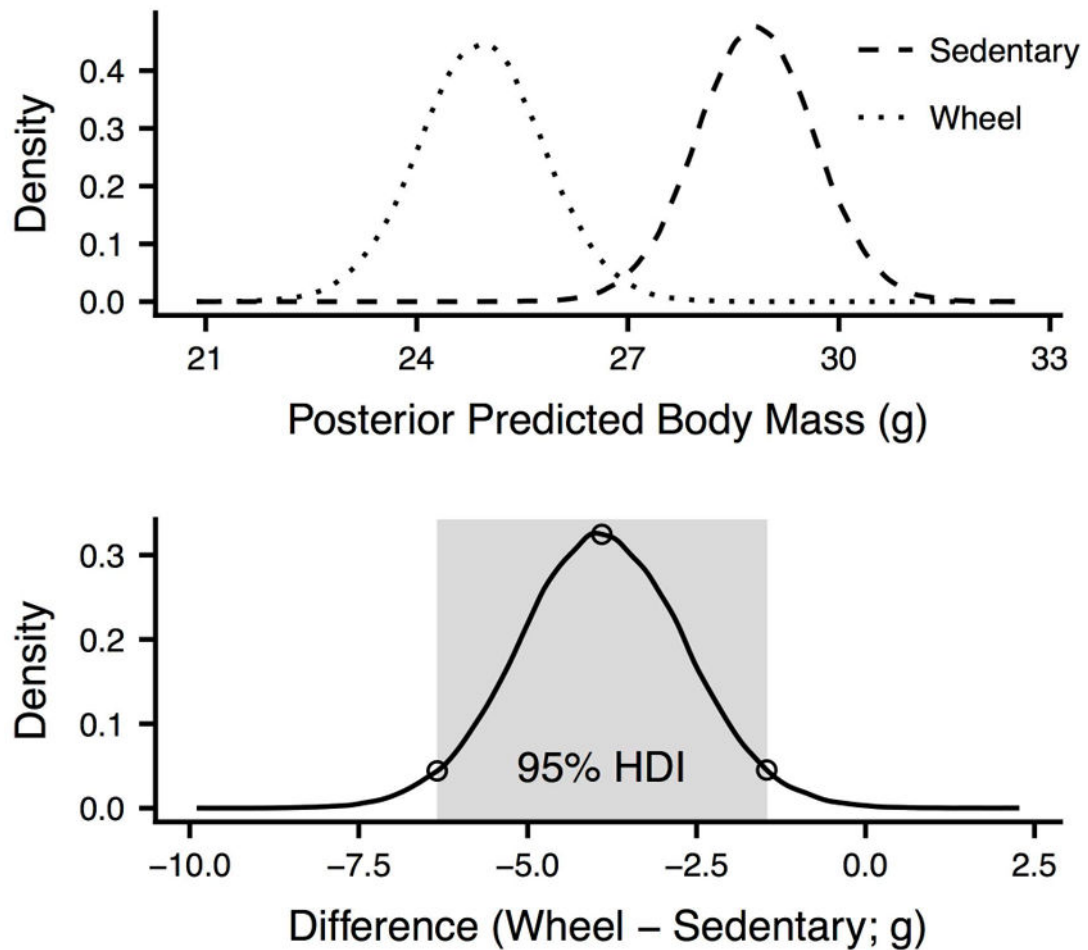


Figure 4. Differences of posterior distributions

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 (Figs. 5, 6, 7).

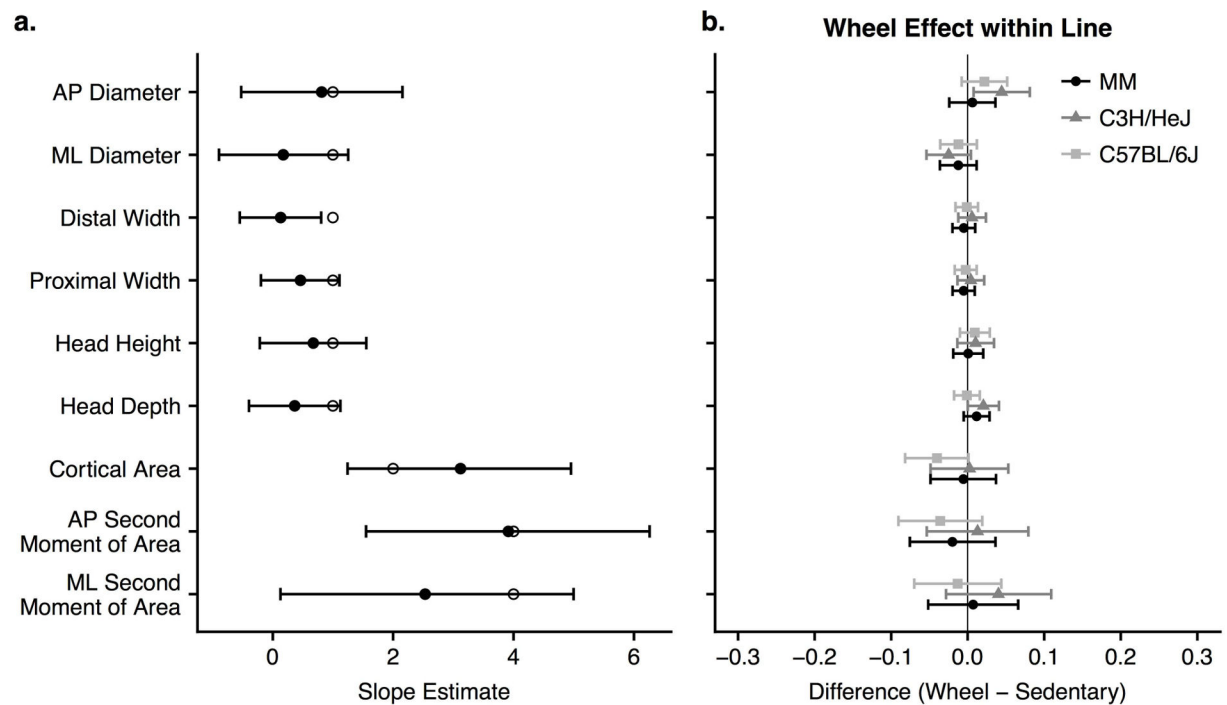


Figure 5. Summarized linear models of gross and histomorphometry

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 a 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.

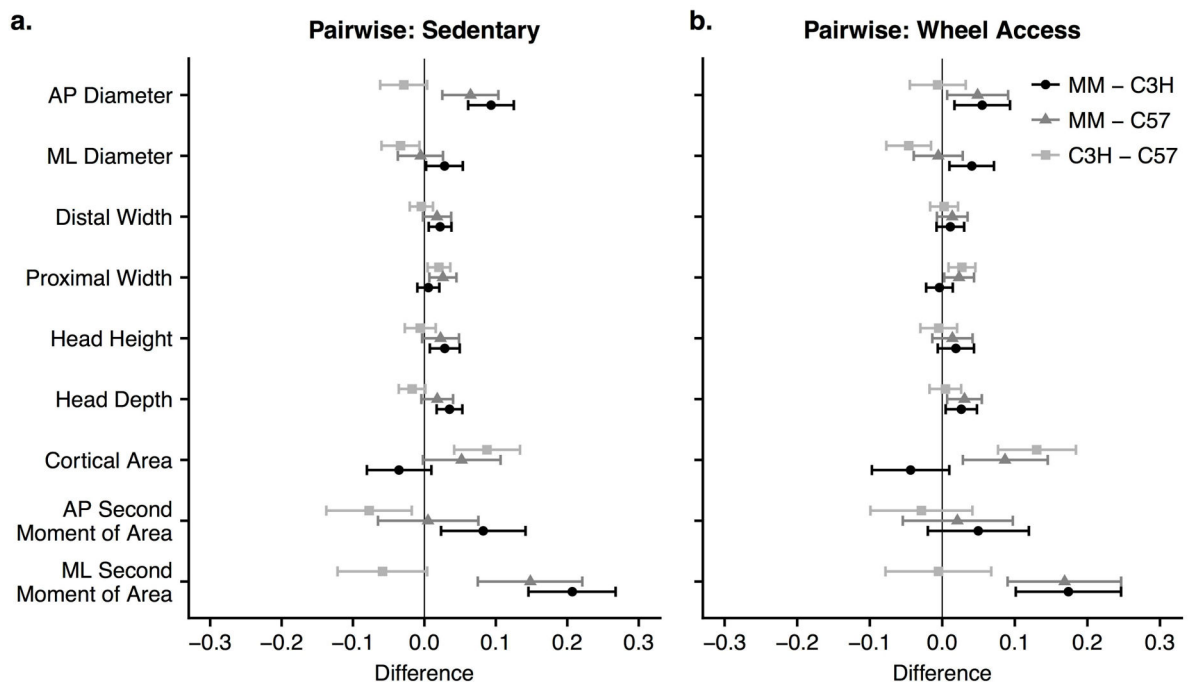


Figure 6. Summarized linear models of gross and histomorphometry

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.

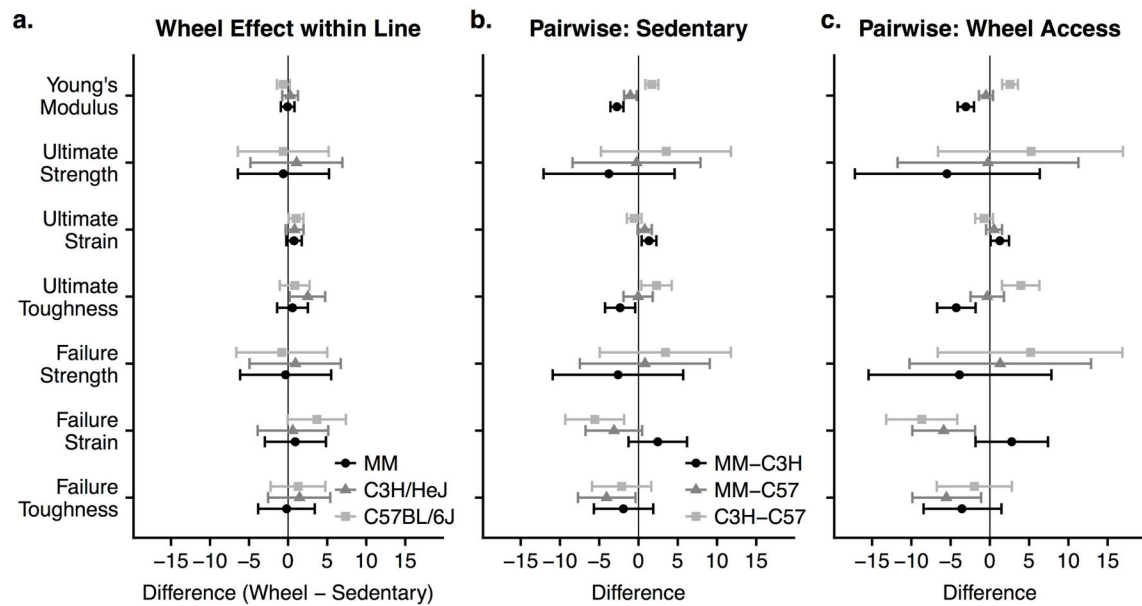


Figure 7. Summarized linear models of biomechanical measures

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