JOURNAL OF morphology

Evolution of Hindlimb Bone Dimensions and Muscle Masses in House Mice Selectively Bred for High Voluntary Wheel-Running Behavior

Journal:	*Journal of Morphology*
Manuscript ID	JMOR-17-0235.R1
Wiley - Manuscript type:	Research Article
Date Submitted by the Author:	n/a
Complete List of Authors:	Castro, Alberto; University of California, Riverside, Department of Evolution, Ecology, and Organismal Biology Garland, Theodore; University of California, Riverside, Department of Evolution, Ecology, and Organismal Biology
Keywords:	bone, exercise, experimental evolution, locomotion, muscle

File = Castro_et_al_LG12_Bone_Manuscript_76.docx
formatted for *Journal of Morphology*

Evolution of Hindlimb Bone Dimensions and Muscle Masses in House Mice Selectively Bred for High Voluntary Wheel-Running Behavior

Alberto A. Castro[1] and Theodore Garland, Jr.[1*]

[1]Department of Evolution, Ecology, and Organismal Biology, University of California, Riverside, Riverside, California 92521, USA

Abbreviated Title (Running Headline): Selection on Behavior Alters Hindlimbs

* To whom proofs, correspondence, and reprint requests should be addressed.

Telephone: (951) 827-3524

Fax: (951) 827-4286 (Department)

E-mail: tgarland@ucr.edu

ABSTRACT

We have used selective breeding with house mice to study coadaptation of morphology and physiology with the evolution of high daily levels of voluntary exercise. Here, we compared hindlimb bones and muscle masses from the 11th generation of four replicate High Runner (HR) lines of house mice bred for high levels of voluntary wheel running with four non-selected control (C) lines. Mass, length, diameter, and depth of the femur, tibia-fibula, and metatarsal bones, as well as masses of gastrocnemius and quadriceps muscles, were compared by analysis of covariance with body mass or body length as the covariate. Mice from HR lines had relatively wider distal femora and deeper proximal tibiae, suggesting larger knee surface areas, and larger femoral heads. Sex differences in bone dimensions were also evident, with males having thicker and shorter hindlimb bones when compared with females. Several interactions between sex, linetype, and/or body mass were observed, and analyses split by sex revealed several cases of sex-specific responses to selection. A subset of the HR mice in two of the four HR lines expressed the mini-muscle phenotype, characterized mainly by an ~50% reduction in hindlimb muscle mass, caused by a Mendelian recessive mutation, and known to have been under positive selection in the HR lines. Mini-muscle individuals had elongated distal elements, lighter and thinner hindlimb bones, altered 3rd trochanter muscle insertion positions, and thicker tibia-fibula distal widths. Finally, several differences in levels of directional or fluctuating asymmetry in bone dimensions were observed between HR and C, mini- and normal-muscled mice, and the sexes. This study demonstrates that skeletal dimensions and muscle masses can evolve rapidly in response to directional selection on locomotor behavior.

KEY WORDS: bone; exercise; experimental evolution; locomotion; muscle

1 INTRODUCTION

Locomotion places more demands on the limbs than does any other behavior (Biewener, 1990). Limb bones transmit muscular and propulsive forces, support the axial skeleton, and exhibit phenotypic plasticity in response to loading during locomotion (Gosnell, Butcher, Maie, & Blob, 2011; Kelly, Czech, Wight, Blank, & Garland Jr., 2006; Middleton, Kelly, & Garland Jr, 2008). Therefore, limb bones may be expected to show evidence of evolutionary coadaptation with locomotor behavior and ecology.

In mammals, numerous studies have provided evidence of coadaptation between the skeleton and locomotor behavior or performance ability. For example, animals that run fast and/or for long distances are often considered "cursorial" (Gregory, 1912; Stein & Casinos, 1997). Some of the most emblematic cursorial mammalian lineages, such as Carnivora, Perissodactyla, and Artiodactyla, have evolved a high metatarsal-femur (MT/F) ratio that is postulated to increase locomotor speed and/or efficiency, and a high MT/F ratio has therefore often been used as a proxy to identify cursorial species (Carrano, 1999; Coombs Jr, 1978; Gambaryan, 1974; Garland Jr. & Janis, 1993; Hildebrand, 1974; Howell, 1944; Lovegrove & Mowoe, 2014; Smith & Savage, 1956). Other aspects of limb morphology have also been associated with cursoriality, including elongated distal limb bones, elevated foot posture, more proximally located muscle masses, more proximal muscle insertions relative to bone length (closer to hip or shoulder joint), and thinner and lighter limb bone elements (Carrano, 1999; Samuels, Meachen, & Sakai, 2013; Van Valkenburgh, 1987).

Some of the putative indicators of cursoriality seem to be associated with body size (Carrano, 1999; Middleton, Kelly, et al., 2008) and/or phylogeny (Garland Jr. & Janis, 1993), rather than only with high locomotor performance, and some workers suggest that cursorial adaptations are present only in larger-bodied animals (e.g., see Carrano, 1999). However, others

have argued that small-bodied mammals do sometimes exhibit cursorial adaptations (Steudel & Beattie, 1993). For example, elephant shrews and cursorial lagomorphs have evolved elongated distal limb bones and more gracile limb elements (Lovegrove & Mowoe, 2014; J. W. Young, Danczak, Russo, & Fellmann, 2014). Furthermore, "cursorial" lagomorphs have lower limb joint mechanical advantages, which may allow increased limb output velocity and faster cycling of limbs (J. W. Young et al., 2014). However, distal limb bone robusticity did not differ between "cursorial" and non-cursorial lagomorphs, which suggests the greater importance of bone strength versus locomotor economy at the distal end of long bones (J. W. Young et al., 2014). In any case, the rich literature associated with studies of cursoriality provides many sources for hypotheses regarding coadaptation of the skeleton with locomotor behavior.

In addition to the traits typically associated with cursoriality in mammals, increased articular surface areas around joints may be good indicators of high locomotor performance (Bramble & Lieberman, 2004; Garland Jr. & Freeman, 2005). Also, in paleo-anthropological studies, increased limb bone robusticity has been associated with increased physical activity which may co-occur in populations that have a history of heightened physical activity (Wallace et al., 2010 and references therein).

Beyond variation in limb bone and muscle sizes and proportions, asymmetry of the appendicular skeleton may impact locomotion. Fluctuating asymmetry (FA) involves small, non-directional deviations from perfect bilateral symmetry, which can be caused by environmental stress and random developmental noise (Pelabon, Hansen, Carter, & Houle, 2006; Valen, 1962). Directional asymmetry (DA) is the consistent deviation of bilateral structures such that one side is larger than the other (Carter, Osborne, & Houle, 2009). FA in limb lengths is negatively correlated with racing ability in horses (Manning & Ockenden, 1994), and in lizards hindlimb and femoral asymmetry was associated with reduced escape performance (Martín & López, 2001).

Another aspect of limb bone morphology that may influence skeletal evolution is sexual dimorphism, which sometimes results from sexual selection. In mammals, sexual size dimorphism in body mass is common, with males usually being the larger sex (Lindenfors, Gittleman, & Jones, 2007). In a study analyzing skeletal shape and size in Carnivora from the perspective of sexual selection, males generally had more robust limb elements and higher mechanical advantages, which may increase functional advantages during male competition and/or prey capture (Morris & Carrier, 2016 and references therein). In many species of mammals, including laboratory house mice as used in the present study, the pelvis is sexually dimorphic, which likely has significance for locomotion (Schutz, Donovan, & Hayes, 2009; but see Warrener, Lewton, Pontzer, & Lieberman, 2015).

One way to study the coadaptation and microevolution of the skeleton with behavior, and of genetically correlated traits in general, is via experiments in which selectively bred lines are compared with non-selected control lines (Garland Jr. & Rose, 2009; Sparrow et al., 2017). In the present study, we compared replicate lines of High Runner (HR) mice that had been selectively bred over 11 generations for voluntary wheel-running behavior with those of four non-selected Control (C) lines (Swallow, Carter, & Garland Jr., 1998). Daily wheel-running distances of the HR lines reached ~75% greater than the C lines by generation 10, mainly by increased average speed (Koteja, Garland Jr., Sax, Swallow, & Carter, 1999; Swallow et al., 1998). A subset of the mice had the mini-muscle phenotype, caused by a Mendelian recessive allele that was present at a low frequency (~7%) in the original base population. Mini-muscle mice exhibit a 50% reduction in the triceps surae and total hindlimb muscle mass, caused by a significant reduction of type IIb muscle fibers (Guderley et al., 2006; Talmadge et al., 2014). Population-genetic modeling indicates that the allele was under (unintentional) positive selection in the HR lines (Garland Jr. et

al., 2002), and so the mini-muscle phenotype is viewed as one aspect of adaptive morphological evolution in the HR lines.

The first study of skeletal materials from the HR lines was a brief communication regarding mice from generation 11, which represents the earliest available set of specimens from this selection experiment. With body mass as a covariate, HR mice had larger femoral heads and reduced directional and fluctuating asymmetry of hindlimb bone lengths (femur + tibiafibula + metartarsal), with no statistical difference in hindlimb lengths or the MT/F ratio, as compared with mice from C lines (Garland Jr. & Freeman, 2005). Additionally, males had relatively shorter hindlimb lengths and larger MT/F ratios.

Here, we extended these comparisons to body mass and length, hindlimb bone dimensions, standard mammalian measurements (ear, tail, and hindfoot lengths), and hindlimb bone and muscle masses. Additionally, we computed ratios that index the relative size of distal versus proximal limb bones, including the MT/F ratio (Garland Jr. & Freeman, 2005; Garland Jr. & Janis, 1993). We also computed various hindlimb bone morphological indices used to examine limb bone robusticity, bone density, and anatomical advantage (in-lever/out-lever lengths of hindlimb muscles) (Samuels et al., 2013; Samuels & Van Valkenburgh, 2008; Van Valkenburgh, 1987).

We formulated several hypotheses regarding limb bones and muscles of HR mice, based on basic biomechanical principles, numerous previous empirical studies of mammals (many of which focus on cursoriality: see above), paleo-anthropological studies, and previous studies of these lines of mice (see Discussion). We hypothesized that mice from the HR lines would have relatively long and gracile limb bones, more proximally located muscle insertions, and potentially reduced muscle masses (see above). Alternatively, we might expect more robust (wider or thicker) limbs in HR mice, which would serve to increase bone strength and increased bone diameters, bone depth, and bone widths at or near surface areas to reduce joint stress (Bramble &

Lieberman, 2004). Female HR mice have evolved by increased running speed, whereas males have evolved mainly by increased running speeds, but also time spent running (Garland Jr., Schutz, et al., 2011), and males are larger than females, which might lead to sex-specific evolutionary pathways in the skeleton. Therefore, another aim of our study was to examine sex differences in bone dimensions, muscle mass, and morphology, reasoning that sex-specific responses may have occurred (Garland Jr., Kelly, et al., 2011; Keeney, Meek, Middleton, Holness, & Garland Jr., 2012).

2 MATERIALS AND METHODS

Artificial Selection Model for High Voluntary Wheel-Running

Specimens were drawn from lab generation 11 of four replicate lines of a mouse colony selectively bred for high voluntary wheel-running behavior. The founding population consisted of 224 outbred, genetically variable laboratory house mice (*Mus domesticus*) of the Hsd:ICR strain (Harlan-Sprague-Dawley, Indianapolis, Indiana, USA). Four lines of mice were selected for high voluntary wheel-running (HR) and compared with four randomly bred control (C) lines (Swallow et al., 1998). Briefly, mice are weaned at 21 days of age, and then housed in same-sex groups of four per cage until age 6-8 weeks. At that point, mice are housed individually in cages with attached computer-monitored wheels (1.12 m circumference) that record revolutions in 1-min bins over six days of wheel access. For HR lines, the highest-running male and female from each family are used as breeders. The selection criterion is total revolutions on days 5 and 6. In the C lines, a male and a female are randomly chosen from each family. Each line comprises 10 breeding pairs per generation, with no sibling pairs. Mice have food and water ad lib.

Body, Bone, and Muscle Measurements

After routine wheel testing of all mice from generation 11 (i.e., each individual was given wheel access for 6 days), a random sample of males and females was chosen for study (n = 142). As sampling was random, some individuals needed to be used as breeders for the ongoing selection experiment. Mice were paired for breeding at approximately 10 weeks of age. After breeding, all individuals were housed individually without wheel access until sacrifice by carbon dioxide inhalation at a mean of 232 days of age, weighed (to nearest 0.01 g), then frozen for subsequent measurements.

After thawing, mice were again weighed and we took the following standard mammalian body measurements (Hall, 1981): body length (tip of the nostril to the end of bone in tail, to nearest 1 mm), tail length (base of tail to end bone of tail, to nearest 1 mm), hindfoot (heel to tip of nails, to nearest 0.1 mm), and ear length (notch in ear to tip, to nearest 0.1 mm). The gastrocnemius and quadriceps muscles were dissected, weighed to the nearest 0.0001 g, and averaged for subsequent statistical analyses. Mice were skinned and eviscerated, and then air dried. Dried carcasses were placed in a colony of dermestid beetles, and bones were further cleaned manually under a dissecting scope as necessary (Garland Jr. & Freeman, 2005).

Bone measurements were taken to the nearest 0.01 mm with Fowler calipers (Fowler Sylvac, ultra-cal mark III) linked to a foot pedal and a small printer. All measurements were taken by Dr. Patricia A. Freeman, blind with respect to linetype. The caliper set-up ensured that the instrument was not put down between measurements, and allowed for rapid re-measurement when needed. To reduce measurement error, three measurements taken in quick succession were averaged and recorded. Both the right and left sides were measured to allow for analysis of asymmetry (Garland Jr. & Freeman, 2005).

Eight measurements were recorded for the femur. 1) femur articular length: length from dorsal tip of head to distalmost end of the medial head. 2) length of head to third trochanter scar:

length from dorsal tip of head to distal end of trochanter muscle scar. 3) depth of femoral head: anterior-posterior diameter of the head. 4) femoral proximal width: greatest medio-lateral width of the femur at the proximal end, from the medial side of the head to the lateral side of the greater trochanter. 5) femoral width at 3rd trochanter: medial-lateral width across the femoral shaft at the third trochanter. 6) femoral least width: medial-lateral width taken on femoral shaft at its least constriction and distal to the trochanter muscle scar; measurement is similar to mid-shaft diameter. 7) femoral least depth: depth taken on femoral shaft at its least constriction and perpendicular to width. 8) femoral distal width: greatest distal width of the femur at the medial and lateral epicondyles. Six measurements were recorded for the tibia-fibula. 1) tibial length: greatest articular length of tibia, from the medial, proximal articular surface; the cup rather than the edge of the medial head to the cup, not tip of the medial malleolus of the tibia. The fibula is not part of this measurement. 2) tibial proximal width: greatest medio-lateral distance across the proximal end of the tibia. 3) tibial proximal depth: greatest antero-posterior depth, perpendicular to width. 4) Tibia-fibula least width: medial-lateral least width across tibia and fibula; measurement is similar to mid-shaft diameter. 5) tibia-fibula least depth: least depth across tibia and fibula and generally perpendicular to width. 6) tibia-fibula distal width: greatest medial-lateral width at the distal end of the tibia-fibula. In addition, we recorded the 3rd metatarsal length, measured on the dorsal surface of the metatarsal while still articulated with the proximal end of the digit (greatest length was taken because articular length was too small for calipers to grip).

We computed the ratio of metatarsal/femur length (MT/F), which shows the relative proportions of proximal and distal bone elements of the hindlimb, and relative size of the hindfoot (Garland Jr. & Freeman, 2005; Garland Jr. & Janis, 1993; Samuels et al., 2013; Samuels & Van Valkenburgh, 2008), tibia/femur ratio (T/F: also known as crural index), which indicates relative proportions of proximal and distal elements of the hind limb (Biancardi & Minetti, 2012; Samuels

et al., 2013; Samuels & Van Valkenburgh, 2008; Vanhooydonck & Van Damme, 2001), and of the length from the femoral head to the 3rd trochanter scar divided by femur length (3rd/F), which indicates changes in quadtratus femoris muscle insertion site position ("in-lever") relative to bone length ("out-lever"), which would likely affect mechanical advantage of the muscle when rotating the hip joint (Charles, Cappellari, Spence, Hutchinson, & Wells, 2016). We also computed the ratio of femoral least width divided by femoral length (FLW/F), which indicates robusticity of femur and ability to resist shearing and bending stresses and the ratio of tibia-fibula least width divided by tibial length (TFLW/T) which indicates robustness of tibia and ability to resist shearing and bending stresses (Samuels et al., 2013; Samuels & Van Valkenburgh, 2008). In addition, we computed the ratio of the tibia-fibula distal width divided by tibial length (TFDW/T) which infers relative distal hindlimb bone robustness (Morris & Carrier, 2016). Finally, because preliminary analysis revealed varying results of bone mass in males and females when either body mass or body length was used a covariate, we computed the ratio of femoral mass divided by (femoral length * [FLW^2]) and the ratio of tibia-fibula mass divided by (tibial length * [$TFLW^2$]) (e.g. see Marchini et al., 2014).

Symmetry Computations

Directional asymmetry (DA) and fluctuating asymmetry (FA) were previously reported for leg length (Garland Jr. & Freeman, 2005) computed as the sum of the lengths of the femur, tibia-fibula, and metatarsal bones, but not for the separate bone lengths and bone widths as in other studies (Auerbach & Ruff, 2006; Sarringhaus, Stock, Marchant, & McGrew, 2005). DA was computed as the right minus the left value of a trait, and FA was computed as the absolute value of the right-left difference. FA/DA was also computed since FA can be affected by DA (Palmer & Strobeck, 2003).

Statistical Analysis

As in numerous previous studies of these lines of mice, the MIXED procedure in SAS (SAS Institute, Cary, NC, USA) was used to apply nested analysis of covariance (ANCOVA) models (Garland Jr. & Freeman, 2005; Houle-Leroy, Garland Jr., Swallow, & Guderley, 2000; Houle-Leroy, Guderley, Swallow, & Garland Jr., 2003; Swallow, Koteja, Carter, & Garland Jr., 1999). Body mass was included as a covariate, except for symmetry measures and functional ratios, and results sometimes differed when body length was used instead (see Result section). We also included bone length as a covariate for bone width, mass, and depth measurements. Results for models using body length or bone length as a covariate are not shown in the tables, but are mentioned in the text when results differed from analyses using body mass as a covariate.

A cross-nested, two-way ANCOVA was used to simultaneously test the effects of linetype (High Runner vs. Control lines) and sex. Replicate line nested within linetype was a random effect, and the effect of linetype, sex, and the sex * linetype interaction were tested with 1 and 6 degrees of freedom. A main effect of the mini-muscle phenotype (Garland Jr. et al., 2002; Houle-Leroy et al., 2003; Kelly et al., 2006) was also included and tested relative to the residual variance with 1 and ~119 d.f. (or fewer in the case of missing values). In the present sample of 142 mice (not all of which had data for all traits), the number of mini-muscle individuals was 6 in HR line 3 (2 females, 4 males), 2 in HR line 6 (1 female, 1 male).

To analyze interactions, we used a cross-nested, two-way ANCOVA to simultaneously test the main effects of linetype (High Runner vs. Control lines) and sex, their interaction, and the linetype * body mass, sex * body mass, and linetype * sex * body mass interactions. Random effects included replicate line nested within linetype, sex * line(linetype), body mass * line(linetype), and body mass * sex * line(linetype). In these "full" models, the effect of linetype,

sex, sex * linetype, body mass, body mass * sex, body mass * linetype, and body mass * sex * linetype were tested with 1 and 6 degrees of freedom, whereas the effect of mini-muscle phenotype was tested with 1 and the residual d.f. (~100 for combined analyses of males and females). In addition, a main effect of the mini-muscle phenotype was included (Garland Jr. et al., 2002; Houle-Leroy et al., 2003; Kelly et al., 2006).

In these full models with all of the indicated fixed and random effects, we often obtained covariance parameter estimates of zero or near-zero for some of the interactive random effects, in which case we removed them from the model, but we always retained the line(linetype) and sex * line(linetype) random effects, given the nature of the experimental design. When the higher-order random effects were removed, then d.f. for testing the main effects and their interactions were increased, as can be seen (Online Supplemental Table 1). In the full models, we often found statistically significant ($p < 0.05$) or suggestive ($p < 0.1$) interactions involving sex and/or linetype and/or body mass, and so we then redid analyses split by sex, as our focus here is comparisons of the HR and C lines of mice. In these sex-specific models, when we did not find an interaction between linetype and body mass, we removed that interaction term and reran the analyses. For all models, outliers were removed when the standardized residual exceeded ~3. We used an α of $\leq$ 0.05 for statistical significance. For simplicity, all p values reported in tables and in the Results section, are 2-tailed.

To address the issues of inflated experiment-wise Type I error rates when making multiple comparisons, we applied the positive False Discovery Rate (pFDR Q-Value) procedure, as implemented in SAS Procedure Multtest. We applied this to the 156 P values reported in Tables 1, 2, and 3. Nominally, 68 of the 160 P values were < 0.05. The Q-Values indicated that only one of these should not be considered significant, a value of P = 0.0481, which is reported in Table 2, but not discussed.

3 RESULTS

Tables 1-3 present significance levels from results of ANOVAs and ANOVAs (using body mass as a covariate), whereas Table 4 presents Least Squares Means (group means adjusted for variation in body mass) for all of the analyses.

Body Size

Preliminary analyses indicated that a female from Control line 4 was the heaviest mouse in the data set (48.82 grams, ID = 14085), and was also a high outlier in analyses of body mass with body length as a covariate. Therefore, we concluded that this individual probably had a large amount of body fat, and we decided to exclude her measurements from all subsequent analyses that involved body mass, including when it was used as a covariate.

Although HR mice tended to be smaller than C mice, body mass and body length differences were not statistically significant (Table 1). Body mass also did not differ between linetypes when body length was included as a covariate. Mini-muscle mice had significantly reduced body mass, including when body length was a covariate, but not a reduced body length (see Table 4 for Least Squares Means). Males were significantly heavier than females, with or without body length as a covariate, but the sexes did not differ in body length (Table 1).

Standard Mammalian Measurements

Linetype differences were never significant for ear, tail or hindfoot lengths, regardless of the body-size covariate used (Table 1). Mini-muscle mice had significantly longer hindfoot lengths when body length was used as a covariate (results now shown). Males had significantly shorter tails with body mass as a covariate. Males had significantly shorter hindfeet with body

mass as a covariate, but significantly longer hindfeet with body length as a covariate. The linetype * sex interactions were not significant for any trait (Table 1).

In the models testing for interactions with body mass, the body mass * linetype * sex interaction was marginally significant for hindfoot length ($p = 0.0717$: Online Supplemental Table 1). Analyses split by sex indicated that, for females, the body mass * linetype interaction was significant ($p = 0.0075$) as was the linetype effect ($p = 0.0434$). For males, only the body mass effect was significant ($p < 0.0001$). Inspection of scatterplots showed that female HR mice had longer feet at larger body masses, as compared with female C mice.

Muscle Masses

Adjusting for body mass, quadriceps and gastrocnemius muscle mass did not differ statistically between HR and C mice, but mini-muscle individuals had significantly reduced quadriceps and gastrocnemius masses (all $p < 0.0001$). Males had heavier quadriceps and gastrocnemius when using body mass ($p = 0.0531$ and $p < 0.0001$, respectively or body length ($p < 0.0001$ and $p = 0.0318$, respectively) as a covariate (Table 1).

The interaction model for quadriceps showed a strong body mass * linetype interaction ($p = 0.0066$: Online Supplemental Table 1), with a steeper slope for HR mice. Analyses split by sex also showed this interaction ($p = 0.0618$ for females, $p = 0.0202$ for males). Inspection of scatterplots showed that HR mice tended to have lighter quadriceps at lower body mass but heavier quads at a higher mass. For gastrocnemius, the body mass * linetype interaction was also significant ($p = 0.0178$: Online Supplemental Table 1), again with HR mice having a higher slope. Analyses split by sex showed that this interaction was significant for males only ($p = 0.0164$), with their regression lines crossing at intermediate masses. For females, the effect of mass was significant ($p < 0.0001$), but linetype was not.

Bone Dimensions and Masses

Linetype differences were not significant for femur length, tibia-fibula length, 3rd metatarsal length, or leg lengths, regardless of the body size covariate used. Additionally, linetype differences were not significant for femur and tibia-fibula masses (Table 2). Mice from HR lines had increased anterior-posterior depth of the femoral head (Fig. 1: $p = 0.0366$ with body mass as covariate; $p = 0.0640$ with body length), increased femur distal width (Fig. 2: $p = 0.0176$ with body mass; $p = 0.0567$ with body length), and increased tibia proximal depth (rank $p = 0.0351$ for body mass; $p = 0.0497$ for tibia-fibula length).

Mini-muscle mice had thinner hind limb bones for many measurements in the femur and tibia-fibula (see Table 2). Femoral distal width (Fig. 2 all covariates), femoral width at 3rd trochanter (Fig. 3, all covariates), femoral least width (all covariates), and the tibia-fibula least width (all covariates) all had reduced medial-lateral width measurements. Femoral (Fig. 4) and tibia-fibula mass were significantly reduced in mini-muscle mice, regardless of covariate used (see Table 2). Mini-muscle mice also had significantly longer tibia-fibula lengths with body mass as a covariate (Fig. 5). Tibia-fibula least depth was significantly reduced when tibia-fibula length was used as a covariate. In contrast, tibia-fibula distal width was larger in mini-muscle individuals with body mass as a covariate ($p = 0.0381$).

Males had significantly shorter hindlimb bones (femur + tibia-fibula + metatarsal), as compared with females, whether body mass or length was used as a covariate (Table 2). Males had significantly greater anterior posterior depth of femoral heads ($p = 0.0002$ with body length; $p=0.0001$ with femur length), femoral width at 3rd trochanter muscle scar (Fig. 3, all covariates), and femoral least width (all covariates: Table 2). In addition, femoral proximal and distal widths were increased when using body length and femur length as covariates. Males had reduced

femoral least depths with body mass as a covariate. Femur mass was reduced in males when using body mass as a covariate (Fig. 4), but increased with femur length as a covariate. For the tibia-fibula, males had significantly increased tibial proximal width, tibial proximal depth, and tibia - fibula least width measurements for body length and tibia-fibula bone lengths as covariates (Table 2). Tibia-fibula distal width was reduced in males with body mass as the covariate, but increased with body length or tibia-fibula length as the covariate. Tibia-fibula mass was reduced in males with body mass as a covariate, but increased with body length or tibia-fibula length as a covariate (Table 2).

Several bone dimensions showed significant interactions with body mass, and full analyses are presented in Online Supplemental Table 1. Here, we discuss a few of the stronger interaction effects. For example, in the femoral head interaction models, the body mass * linetype * sex interaction was marginally significant ($p = 0.0515$) and so was the body mass * sex interaction ($P = 0.0838$, Online Supplemental Table 1). Analyses split by sex indicated that, for females (Fig. 6A), the body mass * linetype interaction was significant ($p = 0.0306$), whereas in males (Fig. 6B) only the linetype effect was significant (after removing the body mass * linetype interaction, linetype $p = 0.0298$). Inspection of Figure 6 shows that female HR mice had larger femoral heads at larger body mass when compared with C female mice, whereas male HR mice had larger femoral heads than male C mice at all body masses.

In the interaction models for femoral proximal width measurements, the body mass * sex interaction was significant ($p = 0.0242$; Online Supplemental Table 1). Analyses split by sex indicated that, for females, only the linetype effect was significant after removing the body mass * linetype interaction ($p = 0.0441$). Inspection of scatterplots showed that female HR mice had wider proximal femurs, regardless of differences in body mass, when compared with C female mice, whereas male HR and C mice did not differ, regardless of body mass.

In the interaction models for tibia proximal depth with ranked values, the body mass * linetype * sex interaction was marginally significant ($p = 0.0663$; Online Supplemental Table 1). Analyses split by sex indicated that, for females, the linetype effect was marginally significant after removing the body mass * linetype interaction ($p = 0.0673$). In males, the body mass * linetype interaction was significant ($p = 0.0324$) as was the linetype effect ($p = 0.0182$). Inspections of scatterplots (not shown) revealed that female HR mice tended to have deeper proximal tibias regardless of body mass, whereas for male the regression lines crossed at imtermediate values of body mass, with a positive slope in C mice but a negative slope for HR mice.

In the interaction models for tibia fibula least depth, the body mass * sex interaction was strongly significant ($p=0.0062$; Online Supplemental Table 1). Analyses split by sex indicated that, for females, the body mass * linetype interaction ($p = 0.0483$), but this was not so for males. Inspections of scatterplots (not shown) revealed that female HR mice generally had deeper tibias at larger masses.

Functional Ratios and Indicators of Bone Density

None of the ratios differed significantly between HR and C lines (Table 3). Mini-muscle mice had significantly increased T/F ratios, and the MT/F ratio ($p = 0.0674$) tended to be increased, suggesting increased distal limb elements relative to proximal ones. The distance from the femoral head to the 3rd trochanter muscle scar, divided by femur length (3rd/F), was significantly greater in mini-muscle individuals, indicating a change in the anatomical advantage of the quadratus femoris muscle (in-lever/out-lever; see above). Mini-muscle mice also had less robust femurs (FMW/F) and less robust tibia-fibulas (TFW/T). Mini-muscle mice tended to have reduced femoral distal widths. Finally, mini-muscle mice had increased [FM/(FL * FLW^3)] and

[TM/(TL * TFLW^2)] ratios (Table 3), suggesting increased bone density. M/F ratio and T/F ratio were significantly increased in males, suggesting increased distal limb elements relative to proximal ones Males also had more robust femurs (FMW/F), more robust tibia-fibulas (TFW/T), and increased distal tibia-fibula robustness (TFDW/T; see Table 3). Finally, males had significanttly decreased [FM/(FL * FLW^3)], suggesting that femurs were less dense than for females (Table 3).

Asymmetry

Directional asymmetry (Online Supplemental Table 2) was significantly lower in HR mice for total leg length (2-tailed p = 0.0217), (see also Garland Jr. & Freeman, 2005) and for femur length (p = 0.0311), but not for tibia-fibula or metatarsal length (Online Supplemental Table 2). The FA/DA ratio for the femur tended to be lower for HR mice (p= 0.0510). Fluctuating asymmetry was significantly lower in HR mice for tibia-fibula distal width (p = 0.0108). Tibia-fibula distal width also had increased levels of directional asymmetry in mini-muscle mice (p = 0.0236). Males had reduced directional asymmetry for 3rd metatarsal length (p = 0.0361). Males also had reduced directional asymmetry for femoral least width (p = 0.0454), with a substantially greater reduction in HR lines than in Control lines (sex * linetype interaction p = 0.0567). However, this trend was not significant when analysis was split by sex (Online Supplemental Table 1). HR males tended to have reduced directional asymmetry for tibia-fibula least width (p = 0.0871: Online Supplemental Table 1), when compared with C lines (sex * linetype interaction p = 0.0923). This trend was also apparent in FA/DA tibia-fibula least width (sex * linetype interaction p= 0.0649) and analysie split by sex for HR males (p = 0.0396: Online Supplemental Table 1).

4 DISCUSSION

We compared hindlimb bone dimensions and muscle masses of four replicate, selectively bred High Runner lines of mice with those from four non-selected Control lines at generation 11. We found several differences between the HR and C lines that would appear adaptive in the context of running long distances on a daily basis. We also found several differences between the subset of individuals that express the mini-muscle phenotype, caused by a Mendelian recessive allele (Kelly et al., 2013) and characterized by a 50% reduction in hindlimb muscle mass (Garland Jr. et al., 2002; Houle-Leroy et al., 2003), and wild-type (normal-muscled) individuals. Finally, we found differences between the sexes, including some unexpected interactions between bone dimensions and body size that differed between linetypes and/or between the sexes.

Differences between High Runner and Control Mice

In a preliminary analysis of a subset of the available bone measurements, Garland and Freeman (2005) reported increased anterior-posterior diameters of the femoral head, suggesting greater articular surface areas at the hip. In addition to confirming those results, our re-analysis also shows that HR mice have increased femoral distal widths and increased proximal tibia depths, suggesting larger knee surface areas. Functionally, larger articular surface areas may be related to increased joint mobility in mammals (Godfrey et al., 1995). We are not aware of previous studies of large or small-bodied mammals that have explored joint surface areas in relation to increased running ability (e.g., via increased stability), although studies of primates have associated joint surface areas with climbing (Godfrey, Sutherland, Boy, & Gomberg, 1991). In the genus *Homo* (as compared with *Pan* and *Australopithecus*), greater articular surface areas of the femoral head, knees, sacroiliac joint, and lumbar centra (all judged relative to body mass, as in our analyses) are suggested to be adaptions for endurance running that increase shock absorption by expanding joint

forces over larger surface areas, thus reducing joint stress from impact forces with the ground (Bramble & Lieberman, 2004). The same may be true for the hindlimbs of mice running at high speeds for many hours per day in large wheels (cf. Roach, Edke, & Griffin, 2012).

Previous studies of later generations of the selection experiment have reported increased femoral and tibiofibular mid-shaft diameters in the HR mice (Kelly et al., 2006; Wallace, Tommasini, Judex, Garland, Jr., & Demes, 2012), which may increase bone strength. We did not find this (Table 1 and 3), conceivably because differences had not evolved to a statistically detectable degree by generation 11.

Finally, we need to qualify our conclusions regarding evolutionary changes in the bones of HR mice. As explained in the Methods, all of the mice studied here were given 6 days of wheel access when young adults, followed by housing without wheels until sacrifice at 232 days of age. Thus, bones may have been affected by wheel running during that brief period, even though the mice were sexually mature (Buie, Moore, & Boyd, 2008). Moreover, at least in later generations, HR mice are more active that C mice in home cages when housed without wheels (Lynn E. Copes et al., 2015; Malisch et al., 2009). Therefore, as noted previously (Kelly et al. 2006), some of the differences we measured between HR and C mice could be caused by the intermediate phenotype of elevated activity levels, rather than by genetic differences that directly affect bone properties. On the other hand, we have also shown that week-old mice (i.e., before they locomote) from generation 45 show differences in femoral characteristics (Wallace et al., 2010). Taken as a whole, we are confident that at least some of the observed differences in skeletal properties between HR and C mice represent evolved differences, not just the result of different activity levels acting across the lifespan (see also Garland Jr. & Freeman, 2005; Kelly et al., 2006; Middleton et al., 2010; Middleton, Shubin, et al., 2008; Middleton, Kelly, et al., 2008; Schutz, Jamniczky, Hallgrímsson, & Garland Jr., 2014; Wallace et al., 2010, 2012; N. M. Young,

Hallgrímsson, & Garland Jr., 2009). Nevertheless, future studies should address the relationship between home-cage activity and bone properties by use of longitudinal sampling and also employ an immobilization model (Jämsä, Koivukangas, Ryhänen, Jalovaara, & Tuukkanen, 1999; Kodama et al., 1999).

Sex Differences

Sex hormones, growth hormones, mechanosensation, and insulin-like growth factors during puberty influence skeletal sexual dimorphism (Callewaert, Sinnesael, Gielen, Boonen, & Vanderschueren, 2010; L. E. Copes, Schutz, Dlugosz, & Garland Jr., 2017). Further, given that female mice generally run more revolutions per day and at higher average and maximum speeds in our study system (see Above), one might expect some degree of sex-specific response to selection. Indeed, several such examples have been reported, including the observation that female HR mice have evolved longer daily running distances almost entirely by increases in average running speed, whereas males also show increases in daily running duration (Garland Jr., Kelly, et al., 2011). However, only one previous study of the HR mice has examined sex differences, with Garland and Freeman (2005) reporting that males had shorter leg lengths, femurs, tibia-fibulas, and metatarsal bones when accounting for body mass as a covariate, but higher MT/F ratios. In the present study, we confirm results and also report that males have higher T/F ratios, heavier hindlimb bones, and more robust femurs and tibia-fibulas (which may increase bone strength), the latter two findings consistent with studies on skeletal sexual dimorphism in Carnivora (Morris & Carrier, 2016), rats (Kim et al., 2003), and humans (Nieves et al., 2004). Males also have shallower femurs indicating differences in the shapes of the hindlimb bones between the sexes. Males also had relatively wider distal femora and heavier hindlimb muscles when compared with females. Limb bone

morphology differs between the sexes substantially, with males having seemingly more robust hindlimb bones and larger muscles than females.

Interactions between Linetype, Sex, and Body Mass

Interaction models revealed interesting results regarding skeletal evolution, body size, and sexual dimorphism as it relates to selective breeding for high voluntary wheel running. For example, female HR mice have evolved larger femoral heads (Fig. 6A), longer hindfeet, and deeper tibia-fibulas only at larger body masses, as compared with female control mice, whereas HR males have larger femoral heads than C males at all body masses (Fig. 6B). In contrast, male HR mice have evolved altered tibia-fibula proximal depths that varied depending on body mass (Online Supplemental Table 1). Thus, allometric relations have evolved in the HR mice, and in a sex-specific way. These results imply that the genetic correlations between bone dimensions and overall body size may be more labile than is commonly assumed (see also Marchini et al., 2014).

When interactions were observed and analysis was split by sex, additional main effects were in some cases discovered (Online Supplemental Table 1). For example, female HR mice had wider proximal femurs and deeper proximal tibias (e.g., see above; near the hip and knee joint) than female control mice. Thus, sex-specific responses in the skeleton can occur even when the same selection is imposed on both sexes. In our case, we showed additional sex-specific skeletal adaptations for the selection of voluntary wheel-running, that was not previously investigated. More broadly, it seems prudent to include both sexes in skeletal evolutionary studies and comparative studies because there may be several sex-specific responses that may have important evolutionary implications. In fact, in lizard studies habitat use was a significant predictor of crus length in females but not in males (Olberding, Herrel, Higham, & Garland Jr., 2015).

Effects of the Mini-muscle Phenotype

As noted above, mini-muscle mice exhibit a 50% reduction in the triceps surae and total hindlimb muscle mass, caused by a significant reduction of type IIb muscle fibers (Guderley et al., 2006; Talmadge et al., 2014) and is evident in reduced gastrocnemius and quadriceps muscle mass. In a study of males from generation 21, mini-muscle mice were previously reported to share some traits with cursorial mammals, with thinner hindlimb bones, longer tibia-fibulas, and longer overall leg lengths (Kelly et al., 2006). In our analysis of mice from generation 11 (Table 2), mini-muscle individuals did not have significantly longer overall leg lengths, but did have longer distal limb bones relative to proximal bones (high MT/F [$P = 0.0674$] and T/F ratio [$P = 0.0069$]). A high T/F ratio and M/F ratio may promote faster running on level ground by increasing arc of hindlimb movements (Chirchir, 2015 and references therein) and are often associated with increased locomotor speed and/or efficiency (e.g. see Introduction). Mini-muscle mice have lighter hindlimb bones, as seen in many cursorial taxa, which, in principle, should reduce the muscular force required to overcome inertia through the swing phase of each stride (Carrano, 1999 and references therein), although mini-muscle individuals actually have increased costs of transport and reduced maximal sprint speeds (Dlugosz, Chappell, McGillivray, Syme, & Garland Jr., 2009). Mini-muscle mice also have thinner hindlimb bones for many measurements which may reduce rotational inertia and be reflective of reduced bone mass (J. W. Young et al., 2014); see table 2. Like cursorial lagomorphs (J. W. Young et al., 2014), mini-muscle mice have increased tibia-fibula distal widths (with body mass as a covariate), but not reduced distal limb bone robusticity (TFDW/T), suggesting the importance of maintaining bone strength at the distal limb.

Given that mini-muscle individuals have some skeletal traits similar to cursors, one might additionally expect their muscle insertion sites to be closer to the hip joints, which would serve to

increase limb output velocity at the cost of force generation (Carrano, 1999; J. W. Young et al., 2014). In contrast to this expectation, mini-muscle mice have 3rd trochanter muscle insertion sites located more distally relative to the length of the femur. The 3rd trochanter muscle scar attaches the quadratus femoris (Charles et al., 2016) and a more distal muscle insertion site may allow for greater force generation (in-lever/out-lever) (Carrano, 1999) when the hip joint is rotated during running.

Previous Studies of the Skeleton of HR Mice & Future Directions

The present study clearly shows that the skeleton can evolve rapidly when selection acts on behavior. Several previous studies that examined later generations of HR and C mice bolster the current results. Middleton and colleagues gave female HR and C mice from generation 16 access to wheels for 20 months and found that the distal width of the femur was increased as a result of selective breeding, but the fracture characteristics of the femoral neck were not affected by selective breeding or wheel running, as compared with mice housed without wheels (Middleton, Shubin, et al., 2008). In addition, the cross-sectional area of the femoral mid-shaft was increased in the HR lines with wheel access, but decreased in the controls with wheel access (genotype-by-environment interaction).

Kelly et al. (2006), studied males from generation 21, half of which were housed with wheel access from weaning for eight weeks. With body mass as a covariate, HR mice had larger femoral heads, heavier feet, and increased tibia-fibula and femoral thickness. Wheel access significantly increased hindlimb bone diameters, foot mass, and tibia-fibula mass in both HR and C lines, with no interaction between linetype and wheel access. Mice with the mini-muscle phenotype had significantly longer and thinner tibia-fibula and femoral bone measurements. However, none of the experimental factors affected the MT/F ratio. Another analysis of this

sample of mice used uCT of femoral morphology at two cortical sites and one trabecular cite (Wallace et al., 2012). HR mice had femurs with enlarged (wider) shafts, increased marrow areas, and altered mid-diaphysis shape which increased moments of inertia (resistance to bending/stress). Mini-muscle mice had reduced cortical bone area, trabecular thickness, and altered shaft shape (Wallace et al., 2012). Wheel running led to moderate periosteal enlargement but increased endocortical expansion, leading to thinner cortices and reduced metaphysis bone area. However, trabecular morphology, moments of inertia, and mid-diaphysis bone area were unaffected by exercise (Wallace et al., 2012). An additional study using this sample found that HR mice have altered semicircular canal shape (Schutz et al., 2014).

Finally, at generation 37, adult (79 days of age) female HR and C mice were housed with or without wheel access for 13-14 weeks. Both linetype and presence of the mini-muscle phenotype were significant predictors of femoral cortical cross-sectional anatomy. However, nano-indentation (micro-scale organization of materials) to measure compressive stiffness at the femoral mid-diaphysis indicated no significant effect of linetype and exercise on mean stiffness (Middleton et al., 2010).

Most studies of skeletal material from the HR selection experiment have examined aspects of the hindlimb bones (but see L. E. Copes et al., 2017; Schutz et al., 2014). A more comprehensive view of skeletal evolution in these unique lines of mice will require consideration of the forelimbs, the pectoral and pelvic girdles (e.g. see Schutz et al., 2009), the axial skeleton, and their functional associations with ligaments, tendons, and muscles. Beyond this, we will need biomechanical studies to measure kinematics and forces during wheel running, as well as studies that attempt to relate morphology to gait and stride differences (e.g. see Claghorn et al., 2017; Sparrow et al., 2017).

Acknowledgments

We thank Patricia A. Freeman for obtaining measurements and for comments on the manuscript. We also thank an anonymous reviewer, Campbell Rolian, Nicholas Schwartz, and Amanda Smolinsky for helpful comments on the manuscript. Special thanks go to Jessica Tingle for helpful discussions and revisions throughout the writing process. Supported by N.S.F. grants to T.G., most recently DEB-1655362.

FIGURE LEGENDS

Figure 1. Mean anterior-posterior depth of femoral head in relation to body mass. Larger mice had larger femoral heads, and mice from the selectively bred High Runner lines had significantly larger femoral heads for a given body size (see Table 2 for statistical analyses).

Figure 2. Mean femoral distal width in relation to body mass. Mice from the selectively bred High Runner lines had significantly broader femoral distal widths for a given body mass, sugessting increased muscle attachment area and increased articular surface area around the knee joint. Mini-muscle mice had reduced femoral distal widths (see Table 2 for statistical analyses).

Figure 3. Mean femoral width at 3rd trochanter muscle scar in relation to body mass. Males had significantly thicker femoral width at 3rd trochanter muscle scar for a given body mass, suggesting increased robustness. In addition, mini-muscle mice had reduced femoral width measurements (see Table 2 for statistical analyses).

Figure 4. Mean femoral mass in relation to body mass. Males had significantly reduced femur mass for a given body mass, and mini-muscle mice had reduced femoral masses (see Table 2 for statistical analyses).

Figure 5. Mean tibia-fibula length in relation to body mass. Males had significantly shorter tibia-fibula lengths for a given body mass, and mini-muscle mice had increased tibia-fibula lengths (see Table 2 for statistical analyses).

Figure 6. Mean anterior-posterior depth of femoral head in relation to body mass, separately by sex to illustrate the significant interaction between sex and body mass for both traits (statistical analyses are in Online Supplemental Table 1). Note that these same data are shown in Figure 1.

LITERATURE CITED

Auerbach, B. M., & Ruff, C. B. (2006). Limb bone bilateral asymmetry: variability and commonality among modern humans. *Journal of Human Evolution*, *50*, 203–218.

Biancardi, C. M., & Minetti, A. E. (2012). Biomechanical determinants of transverse and rotary gallop in cursorial mammals. *Journal of Experimental Biology*, *215*, 4144–4156.

Biewener, A. A. (1990). Biomechanics of mammalian terrestrial locomotion. *Science*, *250*, 1097.

Bramble, D. M., & Lieberman, D. E. (2004). Endurance running and the evolution of Homo. *Nature*, *432*, 345–352.

Buie, H. R., Moore, C. P., & Boyd, S. K. (2008). Postpubertal architectural developmental patterns differ between the L_3 vertebra and proximal tibia in three inbred strains of mice. *Journal of Bone and Mineral Research*, *23*, 2048–2059.

Callewaert, F., Sinnesael, M., Gielen, E., Boonen, S., & Vanderschueren, D. (2010). Skeletal sexual dimorphism: relative contribution of sex steroids, GH-IGF1, and mechanical loading. *Journal of Endocrinology*, *207*, 127–134.

Carrano, M. T. (1999). What, if anything, is a cursor? Categories versus continua for determining locomotor habit in mammals and dinosaurs. *Journal of Zoology*, *247*, 29–42.

Carter, A. J. R., Osborne, E., & Houle, D. (2009). Heritability of Directional Asymmetry in Drosophila melanogaster. *International Journal of Evolutionary Biology*, *2009*, 1–7.

Charles, J. P., Cappellari, O., Spence, A. J., Hutchinson, J. R., & Wells, D. J. (2016). Musculoskeletal geometry, muscle architecture and functional specialisations of the mouse hindlimb. *PLOS ONE*, *11*, e0147669.

Chirchir, H. (2015). A comparative study of trabecular bone mass distribution in cursorial and non-cursorial limb joints: trabecular bone mass in cursorial and non-cursorial limb joints. *The Anatomical Record*, *298*, 797–809.

Claghorn, G. C., Thompson, Z., Kay, J. C., Ordonez, G., Hampton, T. G., & Garland Jr., T. (2017). Selective breeding and short-term access to a running wheel alter stride characteristics in house mice. *Physiological and Biochemical Zoology*, *90*, 533–545.

Coombs Jr, W. P. (1978). Theoretical aspects of cursorial adaptations in dinosaurs. *The Quarterly Review of Biology*, *53*, 393–418.

Copes, L. E., Schutz, H., Dlugosz, E. M., & Garland Jr., T. (2017). Locomotor activity, hormones, and systemic robusticity: an investigation of cranial vault thickness in mouse lines bred for high endurance running. *American Journal of Physical Anthropology*, *In revision*.

Copes, Lynn E., Schutz, H., Dlugosz, E. M., Acosta, W., Chappell, M. A., & Garland Jr., T. (2015). Effects of voluntary exercise on spontaneous physical activity and food consumption in mice: Results from an artificial selection experiment. *Physiology & Behavior*, *149*, 86–94.

Dlugosz, E. M., Chappell, M. A., McGillivray, D. G., Syme, D. A., & Garland Jr., T. (2009). Locomotor trade-offs in mice selectively bred for high voluntary wheel running. *Journal of Experimental Biology*, *212*, 2612–2618.

Gambaryan, P. P. (1974). *How mammals run: anatomical adaptations*. New York: John Wiley and Sons.

Garland Jr., T., & Freeman, P. W. (2005). Selective breeding for high endurance running increases hindlimb symmetry. *Evolution*, *59*, 1851–1854.

Garland Jr., T., & Janis, C. M. (1993). Does metatarsal/femur ratio predict maximal running speed in cursorial mammals? *Journal of Zoology*, *229*, 133–151.

Garland Jr., T., 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. *Proceedings of the Royal Society B: Biological Sciences*, *278*, 574–581.

Garland Jr., T., Morgan, M. T., Swallow, J. G., Rhodes, J. S., Girard, I., Belter, J. G., & Carter, P. A. (2002). Evolution of a small-muscle polymorphism in lines of house mice selected for high activity levels. *Evolution*, *56*, 1267–1275.

Garland Jr., T., & Rose, M. R. (Eds.). (2009). *Experimental evolution: concepts, methods, and applications of selection experiments*. Berkeley: University of California Press.

Garland Jr., T., Schutz, H., Chappell, M. A., Keeney, B. K., Meek, T. H., Copes, L. E., … Eisenmann, J. C. (2011). The biological control of voluntary exercise, spontaneous physical activity and daily energy expenditure in relation to obesity: human and rodent perspectives. *Journal of Experimental Biology*, *214*, 206–229.

Godfrey, L. R., Sutherland, M., Boy, D., & Gomberg, N. (1991). Scaling of limb joint surface areas in anthropoid primates and other mammals. *Journal of Zoology*, *223*, 603–625.

Godfrey, L. R., Sutherland, M. R., Paine, R. R., Williams, F. L., Boy, D. S., & Vuillaume-Randriamanantena, M. (1995). Limb joint surface areas and their ratios in Malagasy lemurs and other mammals. *American Journal of Physical Anthropology*, *97*, 11–36.

Gosnell, W. C., Butcher, M. T., Maie, T., & Blob, R. W. (2011). Femoral loading mechanics in the Virginia opossum, *Didelphis virginiana*: torsion and mediolateral bending in mammalian locomotion. *Journal of Experimental Biology*, *214*, 3455–3466.

Gregory, W. K. (1912). Notes on the principles of quadrupedal locomotion and on the mechanism of the limbs in hoofed animals. *Annals of the New York Academy of Sciences*, *22*, 267–294.

Guderley, H., Houle-Leroy, P., Diffee, G. M., Camp, D. M., & Garland Jr., 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 and Molecular Biology*, *144*, 271–282.

Hall, E. R. (1981). *The mammals of North America* (2nd ed., Vol. 1). Wiley.

Hildebrand, M. (1974). *Analysis of vertrebrate structure*. New York: John Wiley and Sons.

Houle-Leroy, P., Garland Jr., T., Swallow, J. G., & Guderley, H. (2000). Effects of voluntary activity and genetic selection on muscle metabolic capacities in house mice *Mus domesticus*. *Journal of Applied Physiology*, *89*, 1608–1616.

Houle-Leroy, P., Guderley, H., Swallow, J. G., & Garland Jr., 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.

Howell, A. B. (1944). *Speed in animals*. Chicago: Univ. Chicago Press.

Jämsä, T., Koivukangas, A., Ryhänen, J., Jalovaara, P., & Tuukkanen, J. (1999). Femoral neck is a sensitive indicator of bone loss in immobilized hind limb of mouse. *Journal of Bone and Mineral Research*, *14*, 1708–1713.

Keeney, B. K., Meek, T. H., Middleton, K. M., Holness, L. F., & Garland Jr., T. (2012). Sex differences in cannabinoid receptor-1 (CB1) pharmacology in mice selectively bred for high voluntary wheel-running behavior. *Pharmacology Biochemistry and Behavior*, *101*, 528–537.

Kelly, S. A., Bell, T. A., Selitsky, S. R., Buus, R. J., Hua, K., Weinstock, G. M., … Pomp, D. (2013). A novel intronic single nucleotide polymorphism in the *Myosin heavy polypeptide 4* gene is responsible for the mini-muscle phenotype characterized by major reduction in hind-limb muscle mass in mice. *Genetics*, *195*, 1385–1395.

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*, *267*, 360–374.

Kim, B.-T., Mosekilde, L., Duan, Y., Zhang, X.-Z., Tornvig, L., Thomsen, J. S., & Seeman, E. (2003). The structural and hormonal basis of sex differences in peak appendicular bone strength in rats. *Journal of Bone and Mineral Research*, *18*, 150–155.

Kodama, Y., Dimai, H. P., Wergedal, J., Sheng, M., Malpe, R., Kutilek, S., … Baylink, D. J. (1999). Cortical tibial bone volume in two strains of mice: effects of sciatic neurectomy and genetic regulation of bone response to mechanical loading. *Bone*, *25*, 183–190.

Koteja, P., Garland Jr., T., Sax, J. K., Swallow, J. G., & Carter, P. A. (1999). Behavior of house mice artificially selected for high levels of voluntary wheel running. *Animal Behaviour*, *58*, 1307–1318.

Lindenfors, P., Gittleman, J. L., & Jones, K. E. (2007). Sexual size dimorphism in mammals. *Sex, Size and Gender Roles: Evolutionary Studies of Sexual Size Dimorphism*, 16–26.

Lovegrove, B. G., & Mowoe, M. O. (2014). The evolution of micro-cursoriality in mammals. *Journal of Experimental Biology*, *217*, 1316–1325.

Malisch, J. L., Breuner, C. W., Kolb, E. M., Wada, H., Hannon, R. M., Chappell, M. A., … Garland Jr., T. (2009). Behavioral Despair and Home-Cage Activity in Mice with Chronically Elevated Baseline Corticosterone Concentrations. *Behavior Genetics*, *39*, 192–201.

Manning, J. T., & Ockenden, L. (1994). Fluctuating asymmetry in racehorses. *Nature*, *370*, 185–186.

Marchini, M., Sparrow, L. M., Cosman, M. N., Dowhanik, A., Krueger, C. B., Hallgrimsson, B., & Rolian, C. (2014). Impacts of genetic correlation on the independent evolution of body mass and skeletal size in mammals. *BMC Evolutionary Biology*, *14*, 258.

Martín, J., & López, P. (2001). Hindlimb asymmetry reduces escape performance in the lizard *Psammodromus algirus*. *Physiological and Biochemical Zoology*, *74*, 619–624.

Middleton, K. M., Goldstein, B. D., Guduru, P. R., Waters, J. F., Kelly, S. A., Swartz, S. M., & Garland Jr., T. (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 Jr, T. (2008). 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 Jr., T., & Swartz, S. M. (2008). 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.

Morris, J. S., & Carrier, D. R. (2016). Sexual selection on skeletal shape in Carnivora: Sexual selection on skeletal shape in Carnivora. *Evolution*, *70*, 767–780.

Nieves, J. W., Formica, C., Ruffing, J., Zion, M., Garrett, P., Lindsay, R., & Cosman, F. (2004). Males have larger skeletal size and bone mass than females, despite comparable body size. *Journal of Bone and Mineral Research*, *20*, 529–535.

Olberding, J. P., Herrel, A., Higham, T. E., & Garland Jr., T. (2015). Limb segment contributions to the evolution of hind limb length in phrynosomatid lizards. *Biological Journal of the Linnean Society*.

Palmer, A. R., & Strobeck, C. (2003). CH 17. Fluctuating asymmetry analyses revisited. *Developmental Instability: Causes and Consequences, Oxford University Press, Oxford*, 279–319.

Pelabon, C., Hansen, T. F., Carter, A. J. R., & Houle, D. (2006). Response of fluctuating and directional asymmetry to selection on wing shape in *Drosophila melanogaster*. *Journal of Evolutionary Biology*, *19*, 764–776.

Roach, G. C., Edke, M., & Griffin, T. M. (2012). A novel mouse running wheel that senses individual limb forces: biomechanical validation and in vivo testing. *Journal of Applied Physiology*, *113*, 627–635.

Samuels, J. X., Meachen, J. A., & Sakai, S. A. (2013). Postcranial morphology and the locomotor habits of living and extinct carnivorans. *Journal of Morphology*, *274*, 121–146.

Samuels, J. X., & Van Valkenburgh, B. (2008). Skeletal indicators of locomotor adaptations in living and extinct rodents. *Journal of Morphology*, *269*, 1387–1411.

Sarringhaus, L. A., Stock, J. T., Marchant, L. F., & McGrew, W. C. (2005). Bilateral asymmetry in the limb bones of the chimpanzee *(Pan troglodytes)*. *American Journal of Physical Anthropology*, *128*, 840–845.

Schutz, H., Donovan, E. R., & Hayes, J. P. (2009). Effects of parity on pelvic size and shape dimorphism in *Mus*. *Journal of Morphology*, *270*, 834–842.

Schutz, H., Jamniczky, H. A., Hallgrímsson, B., & Garland Jr., T. (2014). Shape-shift: Semicircular canal morphology responds to selective breeding for increased locomotor activity: 3D variation in mouse semicircular canals. *Evolution*, *68*, 3184–3198.

Smith, J. M., & Savage, R. J. (1956). Some locomotory adaptations in mammals. *Zoological Journal of the Linnean Society*, *42*, 603–622.

Sparrow, L. M., Pellatt, E., Yu, S. S., Raichlen, D. A., Pontzer, H., & Rolian, C. (2017). Gait changes in a line of mice artificially selected for longer limbs. *PeerJ*, *5*, e3008.

Stein, B. R., & Casinos, A. (1997). What is a cursorial mammal? *Journal of Zoology*, *242*, 185–192.

Steudel, K., & Beattie, J. (1993). Scaling of cursoriality in mammals. *Journal of Morphology*, *217*, 55–63.

Swallow, J. G., Carter, P. A., & Garland Jr., T. (1998). Artificial selection for increased wheel-running behavior in house mice. *Behavior Genetics*, *28*, 227–237.

Swallow, J. G., Koteja, P., Carter, P. A., & Garland Jr., T. (1999). Artificial selection for increased wheel-running activity in house mice results in decreased body mass at maturity. *Journal of Experimental Biology*, *202*, 2513–2520.

Talmadge, R. J., Acosta, W., & Garland Jr., 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.

Valen, L. V. (1962). A Study of Fluctuating Asymmetry. *Evolution*, *16*, 125.

Van Valkenburgh, B. (1987). Skeletal indicators of locomotor behavior in living and extinct carnivores. *Journal of Vertebrate Paleontology*, *7*, 162–182.

Vanhooydonck, B., & Van Damme, R. (2001). Evolutionary trade-offs in locomotor capacities in lacertid lizards: are splendid sprinters clumsy climbers? *Journal of Evolutionary Biology*, *14*, 46–54.

Wallace, I. J., Middleton, K. M., Lublinsky, S., Kelly, S. A., Judex, S., Garland Jr., 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, Jr., 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.

Warrener, A. G., Lewton, K. L., Pontzer, H., & Lieberman, D. E. (2015). A wider pelvis does not increase locomotor cost in humans, with implications for the evolution of childbirth. *PloS One*, *10*, e0118903.

Young, J. W., Danczak, R., Russo, G. A., & Fellmann, C. D. (2014). Limb bone morphology, bone strength, and cursoriality in lagomorphs. *Journal of Anatomy*, *225*, 403–418.

Young, N. M., Hallgrímsson, B., & Garland Jr., T. (2009). Epigenetic Effects on Integration of Limb Lengths in a Mouse Model: Selective Breeding for High Voluntary Locomotor Activity. *Evolutionary Biology*, *36*, 88–99.

SUPPORTING INFORMATION

Online Supplemental Table 1: tests for interactions between body size, linetype, and sex, as well as analyses split by sex. Microsoft Excel file.

Online Supplemental Table 2: tests for interactions between body size, linetype, and sex, as well as analyses split by sex. Microsoft Word file.

Table 1. Analyses of body size, standard mammalian measurements, and muscle masses with use of either body mass or body length as a covariate. Significance levels (P values; bold indicates P < 0.05) are from two-way nested analysis of covariance models implemented in SAS PROC MIXED. Signs after P values indicate direction of effect: + indicates HR > C, Male > Female, or Mini-muscle > normal muscle.

Trait	N	Linetype	Sex	Sex*Linetype	Mini-Muscle	Body Size	Covariate
Degrees of Freedom		1, 6	1, 6	1, 6	1, ~19	1, ~19	**1, ~19**
Body Size (g)							
Body Mass	136	0.2326-	**<.0001+**	0.4205	**0.0297-**		
Body Mass	135	0.2871-	**<.0001+**	0.7300	**0.0148-**	**<.0001**	**BodyL**
Standard Mammalian (g)							
Hindfoot	134	0.4272-	**0.0249-**	0.3119	0.0583+	**<.0001**	**Mass**
Ear Length	135	0.1793-	0.6902-	0.3951	0.0889+	**<.0001**	**Mass**
Tail Length	136	0.3751+	**0.0335-**	0.6709	0.1838+	**<.0001**	**Mass**
Muscle Mass (g)							
Quadriceps	136	0.8939-	0.0531+	0.5929	**<.0001-**	**<.0001**	**Mass**
Gastrocnemius	133	0.5265+	**0.0041+**	0.6123	**<.0001-**	**<.0001**	**Mass**

Table 2. Analyses of bone dimensions and masses. Significance levels (P values; bold indicates P < 0.05) are from two-way nested analysis of covariance models implemented in SAS PROC MIXED. Signs after P values indicate direction of effect: + indicates HR > C, Male > Female, or Mini-muscle > normal muscle. The value marked with & is not significant after correcting for multiple comparisons with the pFDR Q-Value procedure (see Methods).

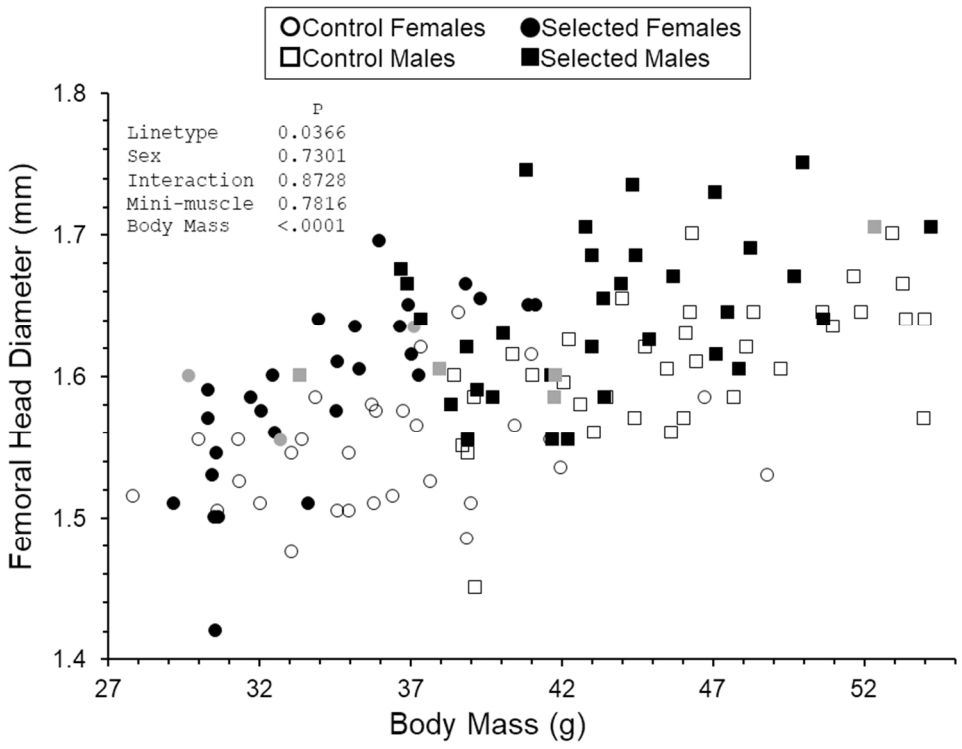
Trait	N	Linetype	Sex	Sex*Linetype	Mini-Muscle	Body Mass
Degrees of Freedom		1, 6	1, 6	1, 6	1, ~19	1, ~19
Bone Lengths						
Leg Length	130	0.9967+	**<.0001-**	0.1568	0.2087+	**<.0001**
Femur	134	0.9303-	**<.0001-**	**0.0919**	0.6813-	**<.0001**
Tibia-fibula	134	0.7818-	**<.0001-**	0.3514	**0.0274+**	**<.0001**
3rd Metatarsal	130	0.3056+	**0.0309-**	0.6962	0.1758+	**<.0001**
Femur						
A-P Depth Femoral Head	133	**0.0366+**	0.7301+	0.8728	0.7816-	**<.0001**
Femoral Distal Width	134	**0.0176+**	0.2685+	0.8537	**0.0087-**	**0.0001**
Femoral Proximal Width	132	0.0760+	0.4130+	0.5038	0.2793+	**<.0001**
Femoral Width 3rd Trochanter	133	0.4579-	**<.0001+**	0.1022	**<.0001-**	**<.0001**
Femoral Least Width	133	0.1560+	**0.0024+**	0.6184	**<.0001-**	**0.0093**
Femoral Least Depth	134	0.4576+	**0.0209-**	0.8432	0.3656-	**<.0001**
Femoral Head to 3rd Trochanter	134	0.5368+	**0.0046-**	0.2522	0.1068+	**0.0080**
Tibia-Fibula						
Tibial Proximal Depth[1]	134	**0.0351+**	0.2562+	0.6720	0.9993-	**0.0001**
Tibial Proximal Width[1]	134	0.1480+	0.2856+	0.5342	0.9973-	**<.0001**
Tibia-fibula Least Width	132	0.1946+	0.8511+	0.8987	**<.0001-**	**<.0001**
Tibia-fibula Least Depth	133	0.9432+	&0.0481+	0.3615	0.1344-	**0.0020**
Tibia-fibula Distal Width	133	0.6212+	**0.0105-**	0.3128	**0.0381+**	**<.0001**
Bone Masses						
Femur	134	0.6365+	**0.0010-**	0.7405	**0.0038-**	**<.0001**
Tibia-fibula	134	0.4927+	**0.0090-**	0.4863	**0.0055-**	**<.0001**

[1] variable was rank-transformed for statistical analyses.

Table 3. Analyses of functional ratios and indicators of bone density. Significance levels (P values; bold indicates P < 0.05 or P < 0.1 for interaction terms) are from two-way nested analysis of variance models implemented in SAS PROC MIXED. Body mass was not used as a covariate in these analyses. Signs after P values indicate direction of effect: + indicates HR > C, Male > Female, or Mini-muscle > normal muscle.

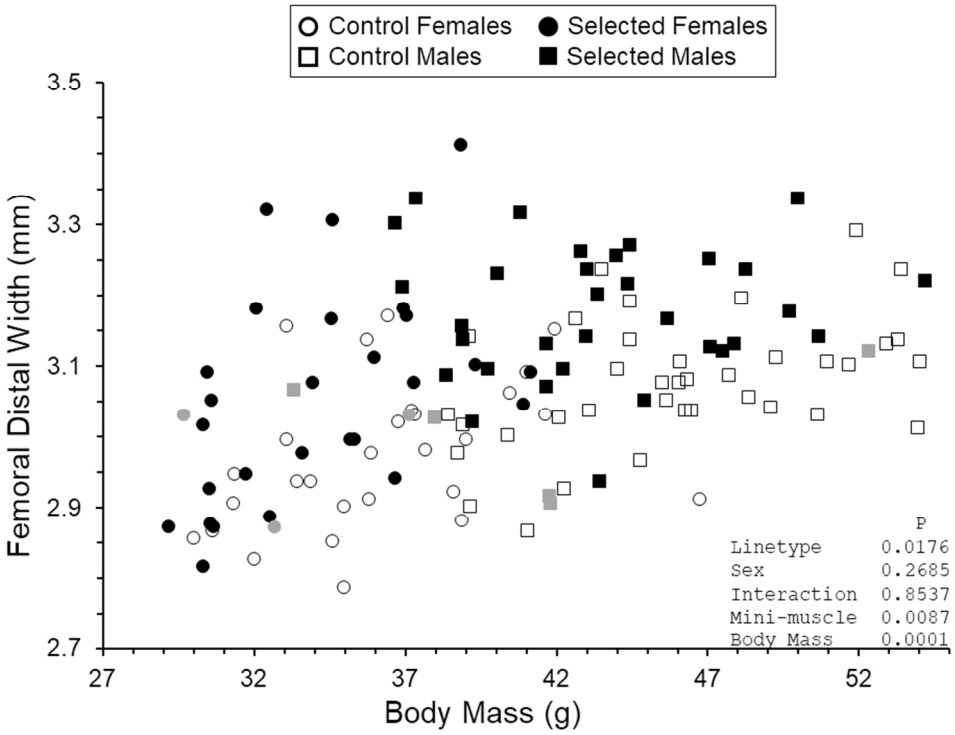
Trait	N	Linetype	Sex	Sex*Linetype	Mini-Muscle
Degrees of Freedom		**1, 6**	1, 6	1, 6	1, ~19
MT/F	135	0.6636+	**0.0001+**	0.5765	0.0674+
T/F	139	0.7975-	**0.0023+**	0.3722	**0.0069+**
3rd/F	139	0.4257+	0.5833+	0.6812	**0.0239+**
FMW/F	139	0.3371+	**<.0001+**	0.5329	**<.0001-**
TFW/T	138	0.3150+	**0.0007+**	0.9425	**<.0001-**
FDW/F	138	0.1435+	**0.0002+**	0.8563	0.0837-
TFDW/T	138	0.5418+	**0.0004+**	0.7574	0.9530+
FM/(FL * FLW^2)	139	0.1885-	**<.0001-**	0.6714	**0.0061+**
TM/(TL * TFLW^3)	138	0.1804-	0.7141-	0.4622	**<.0001+**

		Female		Male		Female		Male		Muscle		Muscle	
	Units	Mean	SE	Mean	SE	Mean	SE	Mean	SE	Mean	SE	Mean	SE
1	g	33.94	1.49	44.51	1.45	32.44	1.41	41.80	1.35	39.99	0.79	36.35	1.75
2	g	33.88	1.10	44.33	1.08	32.84	1.06	42.79	1.01	39.93	0.55	36.99	1.25
3	mm	19.66	0.14	19.41	0.13	19.84	0.14	19.43	0.11	19.42	0.06	19.74	0.17
4	mm	16.62	0.12	16.51	0.12	16.43	0.12	16.45	0.09	16.37	0.04	16.64	0.16
5	mm	105.72	1.21	102.32	1.15	106.14	1.17	103.40	0.94	103.31	0.38	105.48	1.58
6	g	0.1458	0.0047	0.1568	0.0044	0.1467	0.0046	0.1548	0.0037	0.1994	0.0018	0.1027	0.0057
7	g	0.1292	0.0048	0.1451	0.0045	0.1337	0.0047	0.1474	0.0041	0.1823	0.0025	0.0954	0.0052
8	mm	43.61	0.33	41.44	0.32	43.42	0.33	41.63	0.30	42.34	0.19	42.70	0.33
9	mm	16.65	0.19	15.55	0.18	16.52	0.19	15.65	0.17	16.12	0.11	16.06	0.18
10	mm	19.29	0.14	18.38	0.13	19.19	0.14	18.40	0.12	18.66	0.07	18.97	0.15
11	mm	7.66	0.05	7.54	0.04	7.69	0.05	7.60	0.04	7.58	0.02	7.66	0.06
12	mm	1.57	0.02	1.57	0.02	1.62	0.02	1.62	0.02	1.60	0.01	1.59	0.02
13	mm	2.95	0.04	2.98	0.03	3.05	0.04	3.09	0.03	3.07	0.02	2.96	0.04
14	mm	3.62	0.03	3.61	0.02	3.68	0.03	3.65	0.02	3.62	0.01	3.66	0.03
15	mm	2.16	0.04	2.55	0.04	2.18	0.04	2.48	0.03	2.46	0.01	2.23	0.05
16	mm	1.69	0.04	1.84	0.03	1.75	0.03	1.88	0.03	1.89	0.02	1.69	0.04
17	mm	1.50	0.03	1.42	0.03	1.52	0.03	1.45	0.02	1.49	0.01	1.46	0.03
18	mm	6.86	0.10	6.47	0.09	6.85	0.09	6.60	0.08	6.61	0.04	6.78	0.11
19	rank	53.12	11.34	62.36	10.51	74.39	11.19	90.50	9.37	70.10	4.51	70.09	12.72
20	rank	54.18	13.19	60.59	12.43	72.75	13.11	88.04	11.58	68.91	6.90	68.87	13.49
21	mm	0.98	0.02	0.97	0.02	1.00	0.02	1.00	0.01	1.04	0.01	0.94	0.02
22	mm	1.20	0.02	1.25	0.02	1.21	0.02	1.24	0.02	1.24	0.01	1.21	0.02
23	mm	2.68	0.03	2.62	0.02	2.69	0.03	2.64	0.02	2.63	0.01	2.68	0.03
24	g	0.0642	0.0016	0.0567	0.0015	0.0652	0.0016	0.0571	0.0013	0.0637	0.0007	0.0579	0.0020
25	g	0.0511	0.0015	0.0479	0.0014	0.0527	0.0015	0.0487	0.0013	0.0521	0.0008	0.0481	0.0016
26	ratio	0.4639	0.0055	0.4838	0.0054	0.4682	0.0053	0.4856	0.0052	0.4711	0.0033	0.4797	0.0055
27	ratio	1.1599	0.0087	1.1824	0.0086	1.1608	0.0084	1.1761	0.0081	1.1575	0.0049	1.1821	0.0098
28	ratio	0.4149	0.0053	0.4154	0.0052	0.4179	0.0051	0.4214	0.0049	0.4110	0.0026	0.4238	0.0059
29	ratio	0.1017	0.0025	0.1185	0.0024	0.1054	0.0024	0.1203	0.0023	0.1173	0.0013	0.1056	0.0028
30	ratio	0.0506	0.0010	0.0532	0.0010	0.0518	0.0010	0.0545	0.0009	0.0558	0.0006	0.0493	0.0011
31	ratio	0.1381	0.0015	0.1425	0.0015	0.1395	0.0014	0.1436	0.0014	0.1409	0.0009	0.1410	0.0016
32	ratio	0.001288	0.000037	0.001122	0.000037	0.001235	0.000036	0.001053	0.000034	0.001118	0.000021	0.001231	0.000043
33	ratio	0.002744	0.000058	0.002759	0.000057	0.002690	0.000055	0.002647	0.000052	0.002569	0.000027	0.002850	0.000069



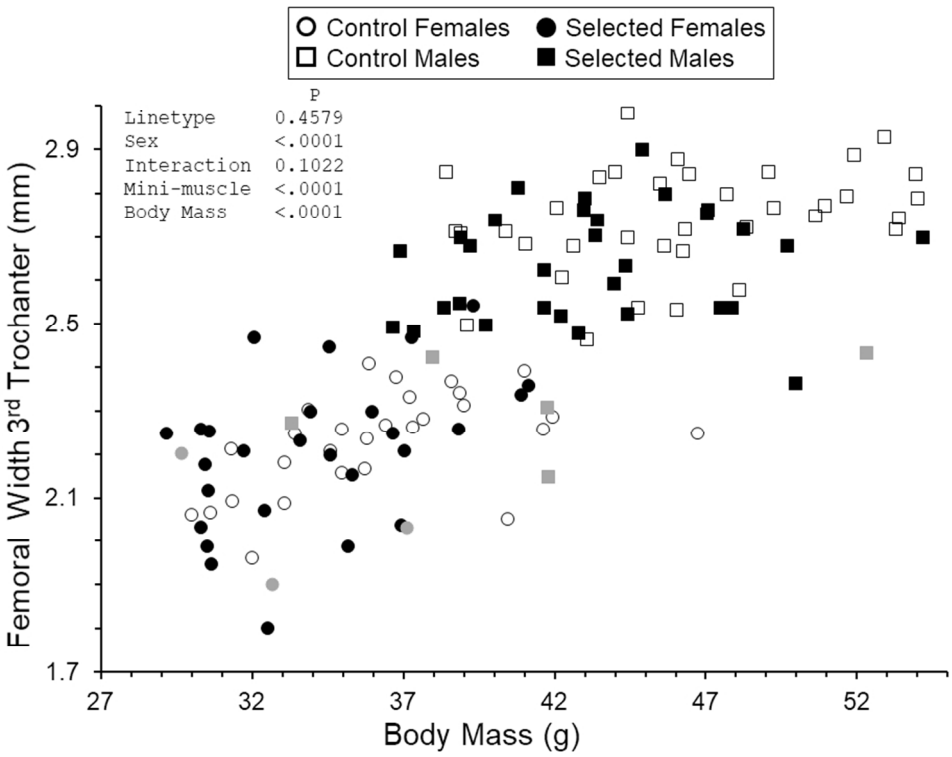
Slide1.TIF

254x190mm (96 x 96 DPI)



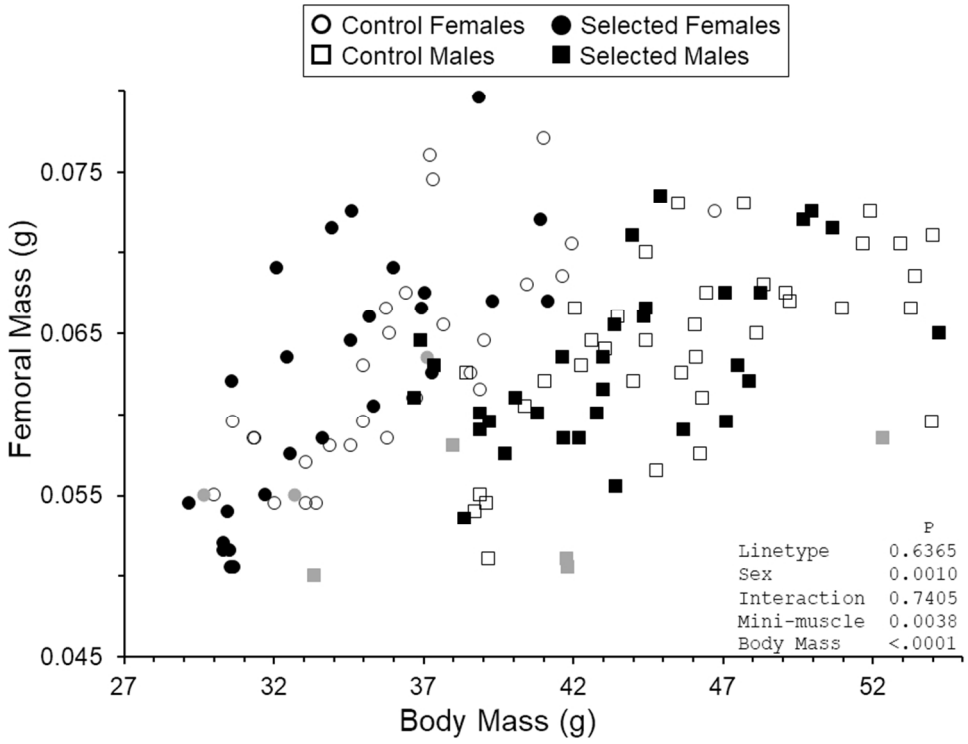
Slide2.TIF

254x190mm (96 x 96 DPI)



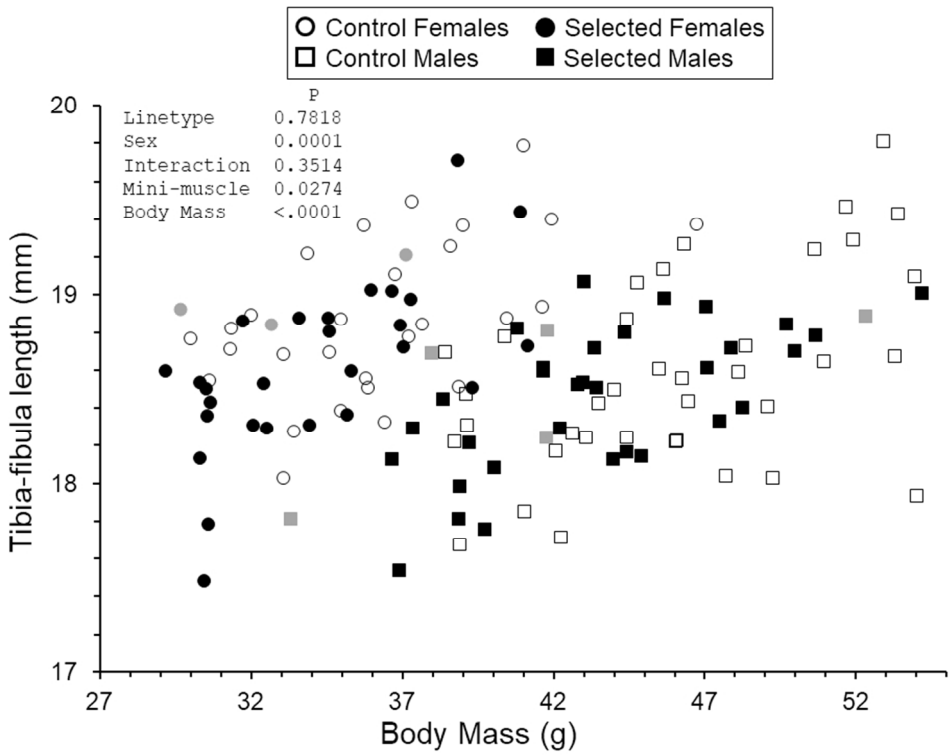
Slide3.TIF

254x190mm (96 x 96 DPI)



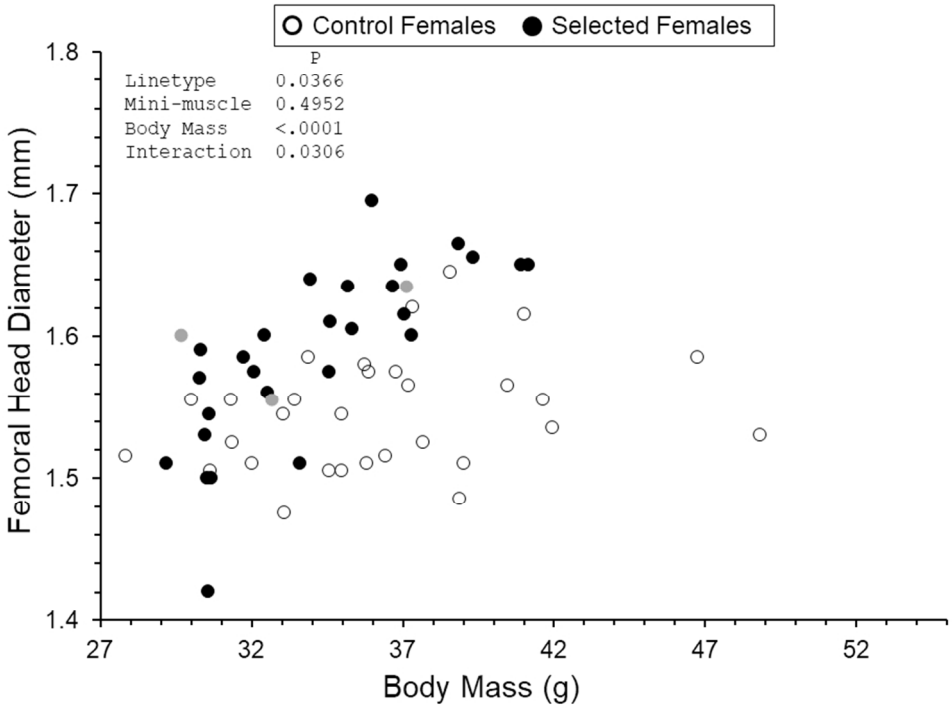
Slide4.TIF

254x190mm (96 x 96 DPI)



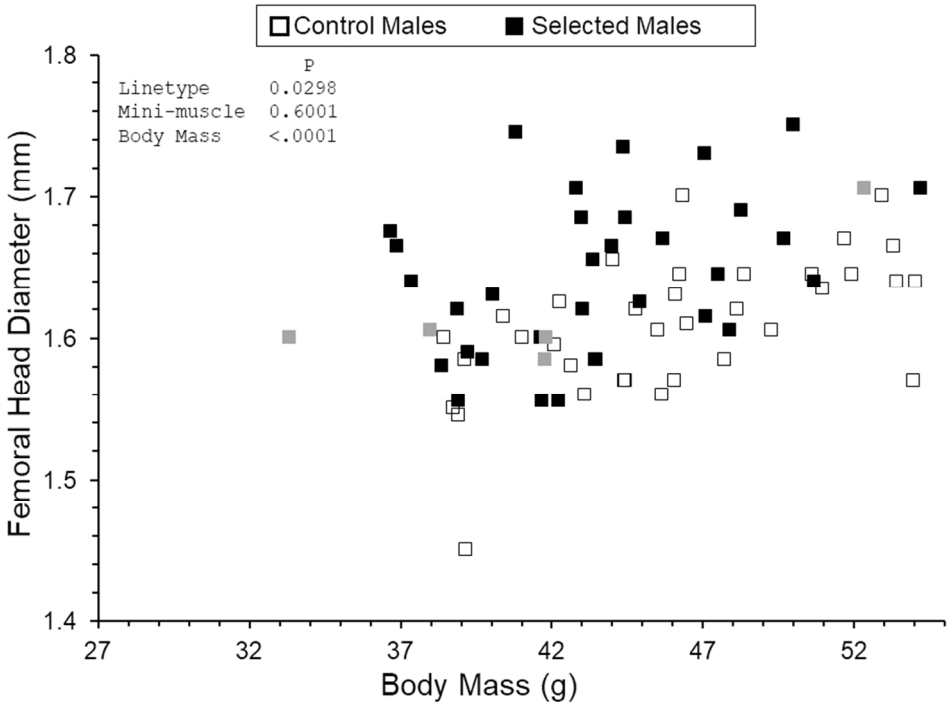
Slide5.TIF

254x190mm (96 x 96 DPI)



Slide6A.TIF

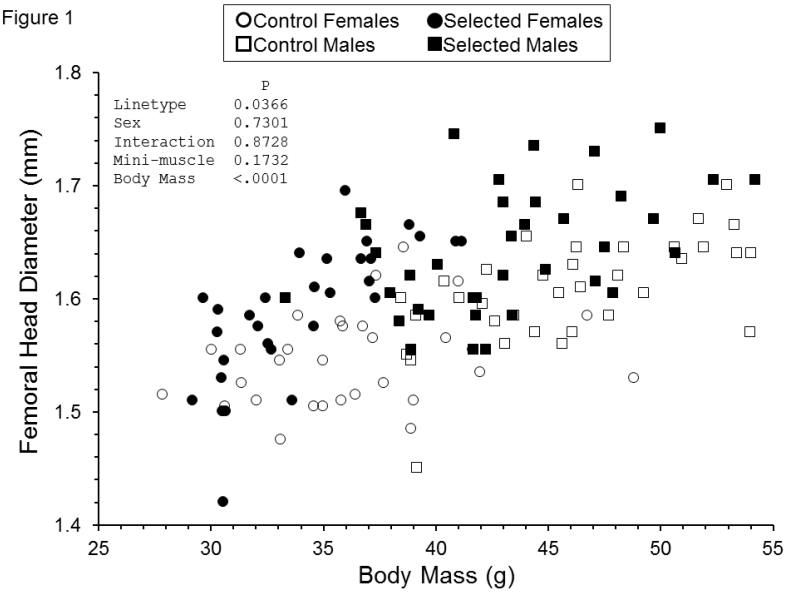
254x190mm (96 x 96 DPI)

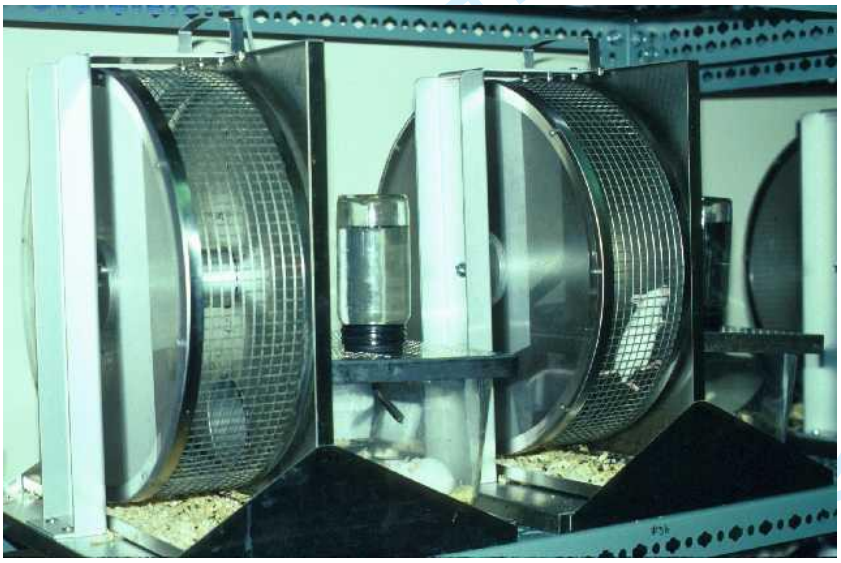


Slide6B.TIF

254x190mm (96 x 96 DPI)

Figure 1


Control Females
Selected Females
Control Males
Selected Males
p
Linetype 0.0366
Sex 0.7301
Interaction 0.8728
Mini-muscle 0.1732
Body Mass <.0001
Femoral Head Diameter (mm)
1.8
1.7
1.6
1.5
1.4
25
30
35
40
45
50
55
Body Mass (g)



We used selective breeding of house mice to study coadaptation of morphology with the evolution of high voluntary exercise. We found that skeletal dimensions and muscle masses can evolve rapidly in response to directional selection on locomotor behavior.