Received: 15 June 2020 | Revised: 24 August 2020 | Accepted: 28 September 2020

DOI: 10.1111/joa.13341

ORIGINAL ARTICLE

ANATOMICAL SOCIETY Journal of Anatomy WILEY

Rapid and longer-term effects of selective breeding for voluntary exercise behavior on skeletal morphology in house mice

Alberto A. Castro | Hannah Rabitoy | Gerald C. Claghorn | Theodore Garland Jr.

Department of Evolution, Ecology, and Organismal Biology, University of California, Riverside, CA, USA

Correspondence
Theodore Garland, Jr., Department of Evolution, Ecology, and Organismal Biology, University of California, Riverside, CA, USA
Email: tgarland@ucr.edu

Funding information
National Science Foundation, Grant/Award Number: DEB-1655362

Abstract

Selection experiments can elucidate the varying course of adaptive changes across generations. We examined the appendicular skeleton of house mice from four replicate High Runner (HR) lines bred for physical activity on wheels and four non-selected Control (C) lines. HR mice reached apparent selection limits between generations 17 and 27, running ~3-fold more than C. Studies at generations 11, 16, and 21 found that HR mice had evolved thicker hindlimb bones, heavier feet, and larger articular surface areas of the knee and hip joint. Based on biomechanical theory, any or all of these evolved differences may be beneficial for endurance running. Here, we studied mice from generation 68, plus a limited sample from generation 58, to test whether the skeleton continued to evolve after selection limits were reached. Contrary to our expectations, we found few differences between HR and C mice for these later generations, and some of the differences in bone dimensions identified in earlier generations were no longer statistically significant. We hypothesize that the loss of apparently coadapted lower-level traits reflects (1) deterioration related to a gradual increase in inbreeding and/or (2) additional adaptive changes that replace the functional benefits of some skeletal changes.

KEYWORDS

adaptation, artificial selection, behavior, bone, coadaptation, experimental evolution, locomotion

1 | INTRODUCTION

Locomotion, active movement through the environment (Dickinson et al., 2000), is vital for animal survival and reproductive success. Animals locomote to flee from predators, forage for food and resources, and when searching for mates. Locomotion places more demands on the skeleton than any other behavior (Biewener, 1990). For instance, limb bones transmit muscular and propulsive forces, support the axial skeleton, and respond to loading during locomotion. Given that locomotion can play a vital role in survival and reproduction, skeletal traits are often correlated with aspects of locomotor behavior, performance, and ecology (Garland & Janis, 1993; Van Valkenburgh, 1987). These types of associations are a cornerstone of ecomorphology (Samuels et al., 2013; Samuels & Van Valkenburgh, 2008; Van Der Klaauw, 1948; e.g., Jones, 2016).

Perhaps the most emblematic example of coadaptation (Angilletta et al., 2006; Foster et al., 2018; Huey & Bennett, 1987) of locomotor behavior with skeletal morphology involves "cursorial" mammals, or those that run fast and/or for long distances (Gregory, 1912; Stein & Casinos, 1997). Within multiple phylogenetic lineages, cursorial mammals have convergently evolved relatively long and tapered limbs, a high metatarsal–femur ratio (MT/F), more proximally located muscles, hinge-like joints that limit motion to the parasagittal plane, fused distal limb bones, and the loss of lateral digits: these traits are presumed to improve running ability and/or locomotor efficiency (Carrano, 1999; Coombs Jr., 1978; Gambaryan, 1974;

Garland & Janis, 1993; Hildebrand, 1974; Howell, 1944; Lovegrove & Mowoe, 2014; Maynard Smith & Savage, 1956; Stein & Casinos, 1997). Another example of skeletal coadaptation occurs in the genus *Homo* (as compared with *Pan* and *Australopithecus*), where larger articular surface areas occur across various hindlimb joints and are thought to improve capabilities for endurance running (Bramble & Lieberman, 2004). Studies of both humans and mice have also shown that increased limb bone robusticity co-occurs in populations with elevated levels of terrestrial mobility, which is partly a result of genetic differences among populations (i.e., present in juveniles before onset of locomotor activities; Cowgill, 2009; Wallace et al., 2010, 2015).

Species of wild small mammals that frequently run at maximal sprint speeds or partake in cost-effective long-distance locomotion (e.g., cursorial elephant shrews, lagomorphs, rodents) have evolved longer and more gracile bones (Lovegrove & Mowoe, 2014; Samuels & Van Valkenburgh, 2008; Vianey-Liaud et al., 2015; Young et al., 2014), as well as having reductions in lower limb joint mechanical advantages, which allows for increased limb output velocity and faster cycling of limbs (Samuels & Van Valkenburgh, 2008; Young et al., 2014).

Selection experiments and experimental evolutionary approaches (Garland & Rose, 2009) are well-suited to study the coadaptation and microevolution of the skeleton with locomotor behavior, locomotor performance, and body size (Marchini et al., 2014; Middleton et al., 2008a). Although, long-term selection studies have traditionally used *Drosophila* as research models (Rose, 2005; Simões et al., 2008, 2019, and references therein; Burke et al., 2016), few, if any, have investigated the correlated changes in skeletal phenotypes as a result of long-term selection for locomotor behavior in vertebrate models. Here, we compare mice from four, replicate lines selectively bred for high levels of voluntary activity (wheel-running behavior: High Runner or HR lines) with those from four non-selected Control (C) lines (Swallow et al., 1998). The HR lines evolved rapidly and reached selection limits after ~17–27 generations, depending on replicate line and sex (Careau et al., 2013; Garland et al., 2011), at which point HR mice run approximately three-fold more wheel revolutions per day than C mice.

Remarkably, the external bone dimensions of the HR mice evolved rapidly as a correlated response to selection. For example, by generation 11, male and female HR mice evolved larger knee and hip surface areas accounting for body size, which, all else being equal, would reduce stress (i.e., force per unit area) acting on limb joints for mice running large distances on wheels (Castro & Garland, 2018; Garland & Freeman, 2005). In addition, various allometric relationships with body size have evolved in sex-specific ways (e.g., female HR mice have larger femoral heads, longer hindfeet, and deeper tibias only at larger body masses). These results may be surprising, given that only 11 generations of selection had been imposed, and that the selection was on behavior, not directly on skeletal dimensions. The fact that selection limits were not reached until ~17–27 generations suggests that further skeletal evolution is probable. Indeed, by generation 21, HR males (females were not studied) had larger femoral heads (as reported for generation 11), but also had evolved thicker femurs and tibia-fibulas (measured but not significant at generation 11) along with heavier feet and longer metatarsals and metacarpals (not weighed at generation 11; Kelly et al., 2006; Young et al., 2009). Broader shafts for femurs and tibia-fibulas increase bone strength and reduce the risk of bone fractures (Wallace & Garland, 2016), which makes intuitive sense as an adaptation in HR mice that run at relatively high speeds for long durations, thus, frequently loading their hindlimbs. Heavier feet could confer better gripping ability on the wire mesh wheels (Kelly et al., 2006), although this has not yet been measured.

The purpose of the present study was to analyze key traits in the appendicular skeleton of HR mice sampled from generation 68, which was well past the point that selection limits for wheel-running behavior were reached. This data set allows us to address four general questions. First, has the skeleton continued to evolve after wheel-running behavior plateaued? Second, have skeletal phenotypes in the HR mice changed from what was reported in previous generations? To address the first and second question, we include analysis of limb bone dimensions, graphs of least square means, and standard errors for bone phenotypes using published data from previous generations, as well as the current data set for generation 68. For these comparisons, we were primarily interested in skeletal traits around the knee and hip joints, which were reported to be significantly different between HR and C mice (see previous paragraph).

Third, what additional aspects of the skeleton may have evolved in response to continued selection? We include analysis of the pelvis and scapula to have a more comprehensive view of skeletal evolution in these unique lines of mice. The scapula and the pelvis (ilium, ischium, and pubis) are flat bones that connect limb bones to the vertebral column, act to transmit body weight onto limb bones, and serve as attachment sites for many muscles important during locomotion (Polly, 2007). Biomechanical analysis of hip joint structure and pelvic dimensions has revealed a structure-to-function relationship with aspects of locomotor behavior in mammals (Álvarez et al., 2013; Jenkins & Camazine, 1977). Furthermore, the shape and size of the scapula along with its muscular attachments directly reflect locomotor behavior among species of mammals (Maynard Smith & Savage, 1956; Oxnard, 1967; Polly, 2007). Therefore, coadaptation of the pelvic and pectoral girdles with locomotor behavior is probable.

Fourth, can genes of major effect and their increase in frequency among the High-Runner mouse model explain losses of apparently coadapted lower-level traits in skeletal dimensions (see Section 3)? A major result of our selection experiment was the presence of the "mini-muscle" phenotype, which occurred in a subset of the HR mice leading to a 50% reduction in triceps surae and total hindlimb muscle mass, primarily caused by a significant reduction in type IIb muscle fibers (Guderley et al., 2006; Talmadge et al., 2014). The phenotype was caused by a Mendelian recessive allele that was present in the original population (~7%), and so the mini-muscle phenotype was (unintentionally) under positive selection (Garland et al., 2002). The mini-muscle phenotype has drastic effects on skeletal phenotypes

and, generally, the hindlimb bones of the mini-muscle mice are more gracile when compared with normal-muscled mice (Castro & Garland, 2018; Kelly et al., 2006). The increasing frequency of the mini-muscle phenotype in two of the HR lines, eventually going to fixation in one HR line (Houle-Leroy et al., 2003), may cause a reduction in statistical power to detect differences between HR and C lines. Therefore, we conducted computer simulations to explore the statistical consequences of increasing the frequency of the mini-muscle phenotype.

2 | MATERIALS AND METHODS

2.1 | High runner mouse model

We used specimens from generation 68 of an on-going selection experiment that breeds for high voluntary wheel-running behavior in house mice (Swallow et al., 1998). The founding population was 224 laboratory house mice (*Mus domesticus*) of the outbred, genetically variable Hsd:ICR strain (Harlan-Sprague-Dawley). Mice were randomly bred for two generations and then separated into eight closed lines, which consist of at least 10 breeding pairs. Four of these lines have been selectively bred for high voluntary wheel-running (HR) and compared with four non-selected control (C) lines. During the routine selection protocol, mice are weaned at 21 days of age and housed in groups of 4 individuals of the same sex until age 6–8 weeks. Mice are then housed individually in cages attached to computer-monitored wheels (1.12-m-circumference, 35.7-cm-diameter, and 10-cm-wide wire-mesh running surface) with a recording sensor that counts wheel revolutions in 1-min intervals over 6 days of wheel access (Hiramatsu et al., 2017; Swallow et al., 1998). In the HR lines, the highest-running male and female from each family are chosen as breeders. The selection criterion is total wheel revolutions on days 5 and 6 to avoid potential effects of neophobia. In the C lines, a male and a female are randomly chosen from each family. Sibling mating is not allowed in any line. Mice are kept at room temperatures of approximately 22°C, with ad lib access to food and water at all times. Photoperiod is 12L:12D with, the light phase beginning at 07:00 h and the dark phase at 19:00 h.

2.2 | DigiGait testing and wheel access

Here, we used 50 male and 50 female mice from a previous study investigating gait differences between the HR and C lines (Claghorn et al., 2017). Our initial sample included 6 males and 6 females from each line, except for HR line 6 (laboratory designation), which remains polymorphic for the mini-muscle phenotype, which involves numerous differences in muscles, organs, and the skeleton (Garland et al., 2002; Kelly et al., 2006, 2017; Syme et al., 2005), for which we used 8 males (3 mini) and 8 females (1 mini). Mice were raised as in the routine selection experiment (see above), except that their toes were not clipped for identification and they were housed individually beginning at weaning. When mice were ~6 weeks of age, the DigiGait Imaging System (Mouse Specifics, Inc.) was used to record stride characteristics (see Claghorn et al., 2017). The University of California–Riverside Institutional Animal Care and Use Committee approved all experimental conditions and protocols.

2.3 | Dissection and bone preparation

Following the gait analyses of Claghorn et al. (2017), 86 (43 males, 43 females) of the 100 mice were used as breeders to produce the next generation. Of the 43 females that were paired, 4 did not give birth. Bone dimensions may have been affected by pregnancy and/or parturition (e.g., see Schutz et al., 2009), but we did not attempt to account for this in statistical analyses due to the greatly unbalanced sample size (39 gave birth, 4 did not). Males were killed by carbon dioxide inhalation at ~4 months of age and females at ~5.5 months of age, which was approximately 21 days after weaning of their pups. Mice were weighed to the nearest 0.01 g and we measured body length (tip of nostril to anus) for each individual mouse; carcasses were subsequently frozen. Later, mice were defrosted, skinned, and eviscerated, and their carcasses were soaked in a 1% solution of enzymatic detergent (Tergazyme) to dissolve flesh from bone (Copes et al., 2018). Bones were then air-dried and manually cleaned under a microscope to remove excess tissue not dissolved by the Tergazyme. During this process, eight carcasses were damaged and therefore not included in the present analyses. In addition, various individual bones were damaged and could not be measured. Final samples sizes are indicated in the Tables 2–4. The authors elect not to share data.

2.4 | Bone imaging and caliper measurements

Bones were photographed individually on a black background with two fluorescent lamps, using a Nikon D60 camera with a 50-mm lens placed ~15 cm above the bones. When photographing, measurement error can occur because of parallax and variance in specimen orientation. For each bone, we chose a highly repeatable element orientation that was positioned at the center of the focal plane and kept at a fixed lens distance (see Schutz et al., 2009) and a scale bar was included.

We used ImageJ (Schneider et al., 2012) to take skeletal measurements from the digital images (Online Supplemental S1 and S2), many of which are the same as the measurements in previous studies (Castro & Garland, 2018; Kelly et al., 2006), which allows for multi-generational comparisons. For three measurements (see Online Supplemental S1), we used hand-held digital calipers (FineSource Electronic Digital Caliper) to facilitate rotation for identification of muscle insertions, and caliper measurements were taken to the nearest 0.01 mm. All measurements were blind with respect

to both linetype and sex. Both the right and left sides were measured to increase accuracy by analysis of mean values. All measurements were checked in two ways: first, we divided the right measure by the left (R/L); second, we computed the right-left difference and divided by the mean value of the measurement ((R-L)/(Mean of R and L)). If the R/L ratio exceeded 1.05 or 0.95, or if the second ratio exceeded −0.05 or 0.05, then photographs were re-measured in Image J or the bone was re-photographed if the bone orientation appeared inappropriate. In addition, we weighed (twice) the air-dried pelvis, humerus, tibia, and femur to the nearest 0.001 g. We computed several morphometric indices that reflect locomotor function (Van Valkenburgh, 1987), many of which are used routinely in ecomorphological studies of mammals (Samuels et al., 2013; Samuels & Van Valkenburgh, 2008; Table 1). Morphometric indices were used to examine limb bone robusticity and anatomical advantage (in-lever/out-lever lengths) of various muscles on the appendicular skeleton.

2.5 | Multi-generational comparisons

A few of the bone measurements taken from generation 68 are directly comparable to those taken in previous studies of these mice (Castro & Garland, 2018; Garland & Freeman, 2005; Kelly et al., 2006; Middleton et al., 2008b, 2010). We therefore compared key measurements for the femur because many of our previous studies have found significant differences between HR and C mice for the femoral head diameter and the femoral distal width (Castro & Garland, 2018; Kelly et al., 2006; Middleton et al., 2008b). In some instances, we used bone data that were not previously reported but were measured in previous generations or measured by us for the present study. Furthermore, we re-analyzed data sets in the exact same fashion across all generations (see Section 2.6). The multi-generational comparisons graphs report *p*-values, age at sacrifice, and least square means with associated standard error bars, including

TABLE 1 Morphometric indices used to interpret function of skeletal traits when comparing the linetypes (HR vs. C) and the mini-muscle phenotype (Normal vs. Mini) in male and female mice

Morphometric indices	Definition and functional significance
Metatarsal–femur ratio (MT/F)	Relative proportions of the proximal and distal hindlimb (or relative size of the hindfoot) Classically used as an indicator of "cursoriality" (3rd Metatarsal length/femur length; Garland & Janis, 1993; Samuels et al., 2013)
Crural index (CI)	Relative proportions of the proximal and distal hindlimb Moment arm of the distal limb, with higher values indicating faster running speeds (Tibia length/femur length; Biancardi & Minetti, 2012; Samuels et al., 2013; Vanhooydonck & Van Damme, 2001)
Brachial index (BI)	Relative proportions of the proximal and distal forelimb (Radius length/humerus length) Lower values indicate increased arboreality while higher values indicate "cursoriality" (Meachen-Samuels & Van Valkenburgh, 2009; Samuels et al., 2013)
Femoral robusticity index (FRI)	Robusticity of the femur and ability to resist shearing and bending stresses (Femoral mid-shaft diameter/femur length; Samuels et al., 2013; Samuels & Van Valkenburgh, 2008)
Tibia robusticity index (TRI)	Robusticity of the tibia and ability to resist shearing and bending stresses (Tibia-fibula mid-shaft diameter/tibia length; Samuels et al., 2013; Samuels & Van Valkenburgh, 2008)
Humerus robusticity index (HRI)	Robusticity of the humerus and ability to resist shearing and bending stresses (Humerus mid-shaft diameter/humerus length; Samuels et al., 2013; Samuels & Van Valkenburgh, 2008)
Ulna robusticity index (URI)	Robusticity of the ulna and ability to resist shearing and bending stresses (Ulna mid-shaft diameter/ulna length; Samuels et al., 2013; Samuels & Van Valkenburgh, 2008)
Scapula breath ratio (SBR)	Relative scapula proportions indicate if the scapula is broader than it is longer Lower values indicate increased "cursoriality" and arboreality (Scapula width/scapula length; Kimes et al., 1981; Polly, 2007)
Distal hindlimb robusticity index (DRI)	Robusticity of the distal hindlimb and relative size of the ankle joint (Tibia-fibula distal width/tibia length; Castro & Garland, 2018; Morris & Carrier, 2016)
Ischium anatomical advantage (IA)	Anatomical advantage of the biceps femoris, semimembranosus, and semitendinosus when extending the hip joint (Ischium length/femur length; Charles et al., 2016a; Morris & Carrier, 2016; Young et al., 2014)
Gluteal anatomical advantage (GA)	Anatomical advantage of the gluteus maximus when rotating the hip joint (Greater trochanter height/femur length; Charles et al., 2016a; Samuels et al., 2013; Samuels & Van Valkenburgh, 2008)
Third trochanter anatomical advantage (3rd/F)	Anatomical advantage of the quadratus femoris when rotating the hip joint (Femoral head to 3rd trochanter muscle scar/femur length; Castro & Garland, 2018; Charles et al., 2016a)
Calcaneum anatomical advantage (CA)	Anatomical advantage of the gastrocnemius when plantarflexing the ankle joint (Calcaneum length/3rd metatarsal length; Charles et al., 2016a; Morris & Carrier, 2016)
Olecranon mechanical advantage (OA)	Anatomical advantage of the triceps brachii when extending the elbow joint (Olecranon length/ulna length; Charles et al., 2016a; Samuels et al., 2013; Samuels & Van Valkenburgh, 2008)

body mass as a covariate. We split the graphs by sex because of unequal sampling from previous studies, skeletal sexual dimorphism, and because our current analysis for generation 68 was also split by sex (see below). Many of the previous skeletal studies included sets of mice that had experienced long-term access to wheels and often showed phenotypic plasticity of bones caused by chronic exercise (Copes et al., 2018; Kelly et al., 2006; Middleton et al., 2008b), but we only included data for mice that did not have wheel access, as in the present study.

2.6 | Statistical analysis

We used the MIXED procedure in SAS (SAS Institute) to apply nested analysis of covariance models with replicate line as a random effect nested within linetype, yielding 1 and 6 d.f. for the effect of linetype for males, but 1 and 5 d.f. for females due to a lack of female mini-muscle individuals in line 6 (Houle-Leroy et al., 2000, 2003; Swallow et al., 1999). Furthermore, the main effect of the mini-muscle phenotype (Garland et al., 2002; Houle-Leroy et al., 2003; Kelly et al., 2006) was included and tested relative to the residual variance with 1 and ~35 d.f. (or fewer depending on sex and skeletal trait). In the present sample of 92 mice (not all of which had data for all traits), the number of mini-muscle individuals was all 12 in HR line 3 (6 females, 6 males) and 4 of 12 in HR line 6 (all males). We split analyses by sex because male and female mice were dissected at different ages and, hence, are not directly comparable. Exploratory analyses revealed that body length was a better predictor of bone dimensions compared to body mass. Therefore, all analyses of skeletal dimensions included body length (recorded at dissections) as a covariate, except for the morphometric indices, all of which are ratios.

To explore possible allometric differences in skeletal dimensions (see Castro & Garland, 2018), we tested for the linetype * body length interaction, mini * body length interaction, and a third model simultaneously including both of the interactions. Initial models included the line (linetype) term and the body length * line (linetype) term as random effects, but the covariance parameter estimates were often zero or near-zero in these full models with all the indicated fixed and random effects. Therefore, final models did not include the body length * line (linetype) random effect, but the line (linetype) random effect was always included given the experimental design (Castro & Garland, 2018). When one or more interaction terms were significant, we used AICc (Akaike information criterion, corrected for small sample size) to compare models (including those with the main effects only), with smaller AICc values indicating a better fit. In addition, we graphed each of the skeletal measurements with body length to verify interactions when present.

In all analyses, outliers were removed when the standardized residual exceeded ~3.0 and we used an α of ≤0.05 for statistical significance. All p values reported are two tailed.

Considering all of the main analyses reported here, done separately by sex (Tables 2–4), 288 p values were produced, 61 of which had nominal $p \leq 0.05$. To address the likelihood of inflated experiment-wise Type I error rates when making so many comparisons on related data, we applied the adaptive False Discovery Rate procedure, as implemented in SAS Procedure Multtest. This indicated that only the lowest 18 would have a corrected $p \leq 0.05$, with the cutoff being $p < 0.004$. However, given that our simulations to explore statistical power (see next section) indicated generally deflated Type I error rates for $\alpha = 0.05$ (see Section 3), we discuss all p values that were nominally $p < 0.05$. Thus, all p values reported in the text are the nominal ones, not adjusted for multiple comparisons.

2.7 | Simulations to explore statistical power

Line-specific changes in the frequency of a gene of major effect on muscle mass, such as the gene causing the "mini-muscle" phenotype, may reduce statistical power to detect general differences between the HR and C lines. Therefore, we conducted simulations to explore the potentially confounding effect that the mini-muscle phenotype has on our ability to detect linetype differences.

As noted above, one unique feature of the High Runner mouse selection experiment was the discovery of the mini-muscle phenotype, characterized primarily by an approximately 50% reduction in the mass of the triceps surae muscle (Garland et al., 2002) and of the rest of the thigh muscles (Houle-Leroy et al., 2003). Previous studies have shown that mini-muscle individuals differ from normal-muscled individuals in most bone measurements (Castro & Garland, 2018; Kelly et al., 2006; Schwartz et al., 2018). Moreover, the frequency of mini-muscle individuals has changed across the generations sampled for the present study. At generation 11 (Castro & Garland, 2018), there were 6 individuals in HR line #3 and 2 individuals in HR line #6 that were mini-muscle mice. However, the mini-muscle phenotype eventually went to 100% (fixation of the recessive allele) in HR line #3 by approximately generations 36–38 (Syme et al., 2005) and has since remained polymorphic in HR line #6. Once fixed in HR line #3, a confounding occurs between the variable denoting presence/absence of the mini-muscle phenotype and line membership, with replicate line used as a random effect nested within linetype (see Section 1).

We created 1,000 random data sets that contained 80 mice (10 for each line). The dependent variable (variable One) had a mean of 10 and a standard deviation of 1. We investigated the power to detect linetype effects by multiplying the dependent variable by values ranging from 1.01 (1% increase in the HR lines) to 1.1 (10% increase). We chose these values because the differences in bone dimensions between HR and C mice range from 3 to 5% (Garland & Freeman, 2005; Kelly et al., 2006; the present study; see Castro & Garland, 2018). We used the same procedure to increase values for individuals with the mini-muscle phenotype, which again approximates the magnitude of differences that have been observed. When altering values for hypothetical individuals with the mini-muscle phenotype, we did so under two relevant scenarios. First, we modeled a frequency of 50% mini-muscle individuals in

TABLE 2 *p*-values, least squares means, and associated standard errors from SAS Procedure Mixed, corresponding to body length, body mass, and skeletal dimensions of the hindlimb. *p*-values in bold are nominally significant at $p < 0.05$, unadjusted for multiple comparisons

Bone trait	♂ *N*	Linetype	Mini	Body mass	HR	MM	Least square mean (LSM)	SE of LSM	♀ *N*	Linetype	Mini	Body mass	HR	MM	Least square mean (LSM)	SE of LSM
Body mass	49	0.0810	0.0992		**0**		35.7740	1.6481	46	0.2053	0.1727		**0**		35.7930	1.9253
					1		31.2922	1.4663					**1**		32.9995	1.4395
						0	34.8383	1.0694						**0**	32.3957	0.9596
						1	32.2279	1.6220						**1**	36.3967	2.6759
Body length	50	0.1547	0.7269		**0**		10.3970	0.1591	46	0.6303	0.3808		**0**		10.9742	0.1395
					1		10.0897	0.1233					**1**		11.0442	0.1006
						0	10.2069	0.0944						**0**	10.9200	0.0683
						1	10.2798	0.1878						**1**	11.0983	0.1892
Body mass W/length	49	0.0974	**0.0453**	**0.0033**	**0**		35.2288	1.4467	46	0.1555	0.1903	**0.0025**	**0**		35.6511	1.7551
					1		31.5686	1.2640					**1**		32.7271	1.3146
						0	34.8614	0.9233						**0**	32.4359	0.8746
						1	31.9361	1.4484						**1**	35.9424	2.4428
Hindlimb length	42	0.9102	0.4523	**0.0023**	**0**		40.3736	0.7683	39	0.4762	0.5554	**<.0001**	**0**		40.0006	0.7679
					1		40.4975	0.7086					**1**		40.5878	0.5698
						0	40.1725	0.5001						**0**	40.6380	0.3775
						1	40.6986	0.7249						**1**	39.9505	1.0702
Pelvis length	45	0.5871	0.6151	**<.0001**	**0**		34.8614	0.9233	45	0.9598	0.5755	**0.0007**	**0**		20.9797	0.5625
					1		31.9361	1.4484					**1**		21.0093	0.4153
						0	19.7775	0.2519						**0**	21.2320	0.2759
						1	19.9168	0.3321						**1**	20.7571	0.7810
Lower ilium length	45	0.4106	**0.0058**	**0.0475**	**0**		8.2632	0.1620	44	0.4644	0.5093	**0.0040**	**0**		9.0912	0.3379
					1		8.0860	0.1333					**1**		8.8244	0.2525
						0	7.8842	0.0987						**0**	8.7894	0.1685
						1	8.4649	0.1835						**1**	9.1262	0.4694
Ischium length	45	0.1298	0.7113	0.1798	**0**		4.4153	0.0764	45	0.2781	0.8304	0.3749	**0**		4.2164	0.1195
					1		4.2521	0.0608					**1**		4.0716	0.0890
						0	4.3152	0.0455						**0**	4.1248	0.0595
						1	4.3522	0.0898						**1**	4.1632	0.1656

(Continues)

TABLE 2 (Continued)

Bone trait	♂ *N*	Linetype	Mini	Body mass	HR	MM	Least square mean (LSM)	SE of LSM	♀ *N*	Linetype	Mini	Body mass	HR	MM	Least square mean (LSM)	SE of LSM
Pubis length	44	0.9249	**0.0395**	**0.0003**	**0**		5.8668	0.0926	43	0.4828	0.5998	**0.0126**	**0**		6.7904	0.1455
					1		5.8780	0.0762					**1**		6.6810	0.1063
						0	5.9911	0.0565						**0**	6.7921	0.0723
						1	5.7537	0.1037						**1**	6.6793	0.1987
Greatest prox. width of ilium	45	0.9735	**0.0136**	0.2248	**0**		3.1415	0.0575	45	0.5312	**0.0374**	0.1492	**0**		3.2792	0.1283
					1		3.1392	0.0436					**1**		3.3652	0.0959
						0	3.2452	0.0331						**0**	3.5295	0.0640
						1	3.0355	0.0706						**1**	3.1150	0.1783
Least distal width of ilium	44	**0.0481**	**0.0163**	**0.0024**	**0**		1.3172	0.0158	44	0.2263	0.7994	0.1594	**0**		1.3083	0.0268
					1		1.2703	0.0117					**1**		1.2715	0.0197
						0	1.3220	0.0090						**0**	1.2949	0.0133
						1	1.2655	0.0194						**1**	1.2848	0.0368
Greatest distal width of ischium	45	0.8172	0.0574	**0.0204**	**0**		3.1685	0.0772	44	0.7970	0.2594	**0.0345**	**0**		3.0449	0.1115
					1		3.1457	0.0622					**1**		3.0748	0.0835
						0	3.2535	0.0463						**0**	3.1555	0.0550
						1	3.0607	0.0897						**1**	2.9641	0.1558
Least width of pubis	45	0.7880	**0.0036**	0.2325	**0**		1.2221	0.0516	45	0.4499	**0.0303**	0.1302	**0**		0.7251	0.0571
					1		1.2036	0.0445					**1**		0.7717	0.0425
						0	1.3000	0.0325						**0**	0.8443	0.0284
						1	1.1257	0.0547						**1**	0.6525	0.0791
Pelvis mass	44	**0.0500**	**0.0074**	**0.0305**	**0**		0.0337	0.0016	45	0.8532	0.0699	0.3639	**0**		0.0327	0.0026
					1		0.0288	0.0014					**1**		0.0332	0.0020
						0	0.0338	0.0010						**0**	0.0366	0.0013
						1	0.0287	0.0017						**1**	0.0293	0.0036
Femur length	46	0.6493	0.7223	**0.0281**	**0**		15.4516	0.2764	43	0.9418	0.9431	**0.0003**	**0**		16.0143	0.4580
					1		15.2730	0.2583					**1**		15.9791	0.3449
						0	15.3257	0.1858						**0**	15.9719	0.2290
						1	15.3989	0.2446						**1**	16.0215	0.6398
Femoral head to 3rd trochanter	47	0.7709	0.9078	0.0950	**0**		6.6605	0.1628	44	0.9751	0.2746	0.4418	**0**		6.7857	0.2520
					1		6.7255	0.1455					**1**		6.7774	0.1888
						0	6.7018	0.1059						**0**	6.5721	0.1259
						1	6.6842	0.1581						**1**	6.9911	0.3505

(Continues)

TABLE 2 (Continued)

Bone trait	♂ *N*	Linetype	Mini	Body mass	HR	MM	Least square mean (LSM)	SE of LSM	♀ *N*	Linetype	Mini	Body mass	HR	MM	Least square mean (LSM)	SE of LSM
Greater trochanter height	48	0.7033	0.4140	0.3215	**0**		2.7773	0.0970	44	0.3583	0.5145	0.2871	**0**		2.6711	0.0809
					1		2.8238	0.0722					**1**		2.5899	0.0583
						0	2.8550	0.0561						**0**	2.5921	0.0401
						1	2.7461	0.1158						**1**	2.6688	0.1091
Femoral head diameter	47	0.8438	0.2878	0.3842	**0**		1.5831	0.0280	43	0.1844	0.9097	0.2566	**0**		1.4994	0.0387
					1		1.5904	0.0237					**1**		1.5589	0.0287
						0	1.6032	0.0176						**0**	1.5324	0.0193
						1	1.5703	0.0294						**1**	1.5259	0.0534
Femoral width at 3rd trochanter	47	0.5711	**<.0001**	0.7633	**0**		2.1990	0.0690	43	0.5279	**0.0092**	0.2705	**0**		1.8753	0.1003
					1		2.1449	0.0612					**1**		1.9433	0.0753
						0	2.3980	0.0447						**0**	2.1171	0.0501
						1	1.9459	0.0679						**1**	1.7015	0.1397
Femoral distal width	47	0.1597	0.8404	**0.0285**	**0**		3.0686	0.0419	42	0.2633	0.4655	**0.0490**	**0**		2.9155	0.0497
					1		2.9837	0.0349					**1**		2.9781	0.0360
						0	3.0310	0.0260						**0**	2.9203	0.0248
						1	3.0213	0.0452						**1**	2.9734	0.0671
Fem. greater trochanter breadth	47	**0.0012**	0.2325	0.4692	**0**		1.6200	0.0293	41	**0.0141**	**0.0155**	0.3152	**0**		1.4593	0.0307
					1		1.4146	0.0214					**1**		1.3474	0.0210
						0	1.5423	0.0169						**0**	1.4572	0.0152
						1	1.4922	0.0353						**1**	1.3495	0.0398
Femoral mid-shaft diameter	47	0.5922	**0.0001**	0.5360	**0**		1.7829	0.0560	43	0.1945	0.1112	0.3552	**0**		1.6172	0.0651
					1		1.7409	0.0508					**1**		1.7146	0.0487
						0	1.8644	0.0369						**0**	1.7455	0.0325
						1	1.6593	0.0526						**1**	1.5863	0.0904
Femoral mass	44	0.2025	**0.0019**	0.0636	**0**		0.0449	0.0022	44	0.3848	0.1347	0.6738	**0**		0.0413	0.0030
					1		0.0408	0.0019					**1**		0.0442	0.0022
						0	0.0468	0.0014						**0**	0.0461	0.0015
						1	0.0389	0.0023						**1**	0.0393	0.0042
Tibia length	47	0.3272	0.7298	0.9959	**0**		18.1197	0.2956	44	0.7935	0.5395	**0.0042**	**0**		18.1889	0.3621
					1		17.7176	0.2518					**1**		18.0891	0.2710
						0	17.8631	0.1865						**0**	17.9710	0.1807
						1	17.9742	0.3096						**1**	18.3069	0.5035

(Continues)

TABLE 2 (Continued)

Bone trait	♂	*N*	Linetype	Mini	Body mass	HR	MM	Least square mean (LSM)	SE of LSM	♀	*N*	Linetype	Mini	Body mass	HR	MM	Least square mean (LSM)	SE of LSM
Tibia proximal width		48	0.3342	**0.0330**	0.4348	**0**		4.0543	0.0796		43	0.2423	0.1054	0.8692	**0**		3.4745	0.0872
						1		3.9452	0.0705						**1**		3.5897	0.0637
							0	4.0832	0.0516							**0**	3.6381	0.0434
							1	3.9164	0.0781							**1**	3.4261	0.1188
Tib fib distal width		48	0.3204	0.6006	0.9028	**0**		2.1849	0.0362		44	0.2442	0.3065	0.9061	**0**		2.1903	0.0471
						1		2.2318	0.0268						**1**		2.2521	0.0346
							0	2.1953	0.0209							**0**	2.1852	0.0234
							1	2.2214	0.0434							**1**	2.2572	0.0646
Tib fib mid-shaft diameter		48	0.7262	**<.0001**	0.2964	**0**		2.8160	0.0368		42	0.0544	**<.0001**	0.4332	**0**		2.6212	0.0380
						1		2.8322	0.0274						**1**		2.7156	0.0268
							0	2.9852	0.0213							**0**	2.7963	0.0189
							1	2.6631	0.0440							**1**	2.5405	0.0504
Tibia-fibula mass		43	0.2100	**0.0030**	0.6752	**0**		0.0360	0.0015		43	0.0699	**0.0049**	0.5180	**0**		0.0299	0.0018
						1		0.0334	0.0012						**1**		0.0340	0.0014
							0	0.0374	0.0009							**0**	0.0361	0.0009
							1	0.0320	0.0016							**1**	0.0278	0.0025
3rd metatarsal length		45	0.5276	0.8966	0.2002	**0**		7.1252	0.1359		41	0.3760	0.0739	0.0563	**0**		6.8133	0.1278
						1		7.0143	0.1036						**1**		6.6883	0.0933
							0	7.0581	0.0811							**0**	6.5784	0.0644
							1	7.0813	0.1569							**1**	6.9232	0.1732
Calcaneus length		47	0.7506	0.2327	0.8631	**0**		3.9604	0.0875		38	0.1664	0.5077	0.0654	**0**		3.8896	0.0678
						1		3.9253	0.0640						**1**		3.7780	0.0474
							0	4.0183	0.0509							**0**	3.8658	0.0346
							1	3.8675	0.1064							**1**	3.8018	0.0885

TABLE 3 *p*-values, least squares means, and associated standard errors from SAS Procedure Mixed, corresponding to skeletal dimesnsions of the forelimb. *p*-values in bold are nominally significant at $p < 0.05$, unadjusted for multiple comparisons

Bone trait	♂	*N*	Linetype	Mini	Body mass	HR	MM	Least square mean (LSM)	SE of LSM	♀	*N*	Linetype	Mini	Body mass	HR	MM	Least square mean (LSM)	SE of LSM
Scapula length		45	0.2106	0.2015	0.2194	**0**		12.0334	0.1597		42	0.1718	0.3569	**0.0068**	**0**		11.8888	0.0954
						1		11.7558	0.1320						**1**		11.7365	0.0692
							0	12.0137	0.0981							**0**	11.7480	0.0478
							1	11.7755	0.1744							**1**	11.8773	0.1289
Scapula breath		45	0.2096	**0.0037**	**0.0242**	**0**		8.1990	0.1748		43	0.3160	0.7354	**0.0081**	**0**		8.2366	0.2345
						1		7.8806	0.1547						**1**		7.9755	0.1761
							0	8.2996	0.1126							**0**	8.0461	0.1172
							1	7.7800	0.1734							**1**	8.1660	0.3267
Humerus length		45	0.2494	0.5001	**0.0330**	**0**		13.0367	0.2350		41	0.4924	0.3471	**0.0005**	**0**		13.1006	0.4286
						1		12.6469	0.2092						**1**		12.7836	0.3233
							0	12.9175	0.1520							**0**	12.6335	0.2141
							1	12.7661	0.2320							**1**	13.2507	0.5999
Humerus head diameter		46	0.5587	0.1786	0.9835	**0**		1.8432	0.0443		41	0.4654	0.7700	0.1515	**0**		1.7692	0.0478
						1		1.8761	0.0346						**1**		1.8069	0.0352
							0	1.8986	0.0261							**0**	1.7984	0.0239
							1	1.8208	0.0517							**1**	1.7777	0.0655
Humerus distal width		46	0.4081	0.8569	0.6508	**0**		2.3581	0.0758		42	0.5601	0.7558	0.3204	**0**		2.3696	0.0717
						1		2.4367	0.0550						**1**		2.4138	0.0501
							0	2.4071	0.0425							**0**	2.3759	0.0355
							1	2.3877	0.0938							**1**	2.4075	0.0946
Humerus mid-shaft diameter		46	0.1771	**0.0252**	0.1246	**0**		0.9775	0.0261		41	0.0732	0.1077	**0.0053**	**0**		0.9338	0.0393
						1		1.0253	0.0203						**1**		1.0226	0.0292
							0	1.0406	0.0153							**0**	1.0266	0.0197
							1	0.9622	0.0305							**1**	0.9298	0.0543
Humerus mass		42	**0.0219**	**0.0087**	**0.0177**	**0**		0.0261	0.0007		44	0.9361	0.5766	0.4255	**0**		0.0236	0.0017
						1		0.0235	0.0005						**1**		0.0234	0.0013
							0	0.0261	0.0004							**0**	0.0242	0.0009
							1	0.0235	0.0009							**1**	0.0228	0.0024

(Continues)

TABLE 3 (Continued)

Bone trait	♂	*N*	Linetype	Mini	Body mass	HR	MM	Least square mean (LSM)	SE of LSM	♀	*N*	Linetype	Mini	Body mass	HR	MM	Least square mean (LSM)	SE of LSM
Ulna length		42	0.2242	0.8735	0.1545	**0**		15.0660	0.1560		42	0.4774	**0.0469**	**0.0189**	**0**		14.8414	0.2063
						1		14.8075	0.1184						**1**		14.6830	0.1520
							0	14.9201	0.0945							**0**	14.4474	0.1032
							1	14.9533	0.1805							**1**	15.0771	0.2833
Olecranon length		43	0.8420	0.8417	**0.0061**	**0**		2.0651	0.0448		42	0.4553	0.4240	0.2383	**0**		1.9153	0.0518
						1		2.0537	0.0343						**1**		1.9571	0.0377
							0	2.0652	0.0270							**0**	1.9669	0.0259
							1	2.0536	0.0511							**1**	1.9055	0.0705
Radius length		40	0.0789	0.0553	0.3861	**0**		12.5513	0.1450		28	0.9558	0.3824	0.0874	**0**		12.1595	0.2014
						1		12.2022	0.1104						**1**		12.1716	0.1454
							0	12.1740	0.0813							**0**	12.0349	0.1035
							1	12.5796	0.1835							**1**	12.2962	0.2683
Styloid width		38	0.3377	0.1149	0.4202	**0**		1.8766	0.0298		28	0.8443	0.8190	**0.0364**	**0**		1.8045	0.0288
						1		1.8407	0.0226						**1**		1.7982	0.0178
							0	1.8230	0.0171							**0**	1.7972	0.0153
							1	1.8942	0.0384							**1**	1.8056	0.0332
Ulna mid-shaft diameter		43	0.8143	0.5738	**0.0284**	**0**		1.0423	0.0287		42	0.6245	0.8831	0.5948	**0**		0.9969	0.0531
						1		1.0339	0.0206						**1**		0.9692	0.0395
							0	1.0268	0.0169							**0**	0.9772	0.0265
							1	1.0494	0.0339							**1**	0.9889	0.0734

TABLE 4 *p*-values, least squares means, and associated standard errors from SAS Procedure Mixed, corresponding to morphological indices. *p*-values in bold are nominally significant at $p < 0.05$, unadjusted for multiple comparisons

Morph. index	♂	*N*	Linetype	Mini	HR	MM	Least square mean (LSM)	SE of LSM	♀	*N*	Linetype	Mini	HR	MM	Least square mean (LSM)	SE of LSM
MT/F		44	0.8576	0.0975	**0**		0.4509	0.0138		41	0.7130	0.3150	**0**		0.4253	0.0105
					1		0.4543	0.0122					**1**		0.4212	0.0077
						0	0.4637	0.0090						**0**	0.4154	0.0053
						1	0.4416	0.0134						**1**	0.4311	0.0142
CI		47	0.4179	0.9278	**0**		1.1702	0.0132		42	0.6748	0.4211	**0**		1.1401	0.0137
					1		1.1565	0.0101					**1**		1.1340	0.0100
						0	1.1625	0.0079						**0**	1.1289	0.0069
						1	1.1641	0.0155						**1**	1.1451	0.0186
BI		38	0.6182	0.7310	**0**		0.9484	0.0193		28	0.4689	0.7002	**0**		0.9317	0.0326
					1		0.9588	0.0146					**1**		0.9577	0.0243
						0	0.9486	0.0098						**0**	0.9542	0.0166
						1	0.9587	0.0269						**1**	0.9352	0.0448
FRI		47	0.9545	**0.0003**	**0**		0.1156	0.0034		43	0.1887	0.1384	**0**		0.1004	0.0049
					1		0.1153	0.0029					**1**		0.1078	0.0036
						0	0.1221	0.0022						**0**	0.1096	0.0024
						1	0.1088	0.0034						**1**	0.0985	0.0068
TRI		48	0.8440	0.3118	**0**		0.0694	0.0021		41	0.9120	0.4268	**0**		0.0660	0.0027
					1		0.0700	0.0019					**1**		0.0663	0.0020
						0	0.0706	0.0014						**0**	0.0678	0.0014
						1	0.0687	0.0020						**1**	0.0645	0.0038
HRI		45	**0.0420**	**0.0314**	**0**		0.0747	0.0022		41	0.0560	0.0709	**0**		0.0717	0.0037
					1		0.0814	0.0017					**1**		0.0809	0.0028
						0	0.0811	0.0013						**0**	0.0814	0.0019
						1	0.0750	0.0025						**1**	0.0711	0.0051
URI		42	0.9939	0.8031	**0**		0.0695	0.0021		42	0.7752	0.5816	**0**		0.0669	0.0032
					1		0.0694	0.0015					**1**		0.0659	0.0024
						0	0.0691	0.0013						**0**	0.0677	0.0016
						1	0.0698	0.0025						**1**	0.0651	0.0044

(Continues)

TABLE 4 (Continued)

Morph. index	♂ *N*	Linetype	Mini	HR	MM	Least square mean (LSM)	SE of LSM	♀ *N*	Linetype	Mini	HR	MM	Least square mean (LSM)	SE of LSM
SBR	44	0.2522	0.0748	**0**		0.6869	0.0128	43	0.5975	0.9586	**0**		0.6909	0.0170
				1		0.6665	0.0109				**1**		0.6813	0.0127
					0	0.6895	0.0081					**0**	0.6854	0.0085
					1	0.6639	0.0136					**1**	0.6867	0.0235
DRI	48	0.1053	0.7247	**0**		0.1207	0.0026	44	0.3908	0.9059	**0**		0.1200	0.0046
				1		0.1269	0.0022				**1**		0.1243	0.0034
					0	0.1233	0.0016					**0**	0.1218	0.0023
					1	0.1243	0.0028					**1**	0.1226	0.0063
IA	45	0.3541	0.8773	**0**		0.2846	0.0053	43	0.2686	0.9222	**0**		0.2633	0.0067
				1		0.2782	0.0042				**1**		0.2550	0.0049
					0	0.2819	0.0032					**0**	0.2587	0.0033
					1	0.2809	0.0062					**1**	0.2596	0.0092
GA	45	0.6742	0.4091	**0**		0.1810	0.0056	42	0.3938	0.6223	**0**		0.1683	0.0057
				1		0.1839	0.0041				**1**		0.1629	0.0040
					0	0.1858	0.0033					**0**	0.1636	0.0028
					1	0.1791	0.0068					**1**	0.1676	0.0076
3rd/F	47	0.7589	0.8049	**0**		0.4321	0.0090	42	0.8873	**0.0415**	**0**		0.4230	0.0076
				1		0.4357	0.0075				**1**		0.4241	0.0055
					0	0.4352	0.0056					**0**	0.4119	0.0038
					1	0.4326	0.0098					**1**	0.4352	0.0103
CA	45	0.3094	0.0700	**0**		0.5456	0.0132	36	0.2874	0.0663	**0**		0.5802	0.0124
				1		0.5627	0.0093				**1**		0.5647	0.0087
					0	0.5714	0.0077					**0**	0.5890	0.0065
					1	0.5369	0.0159					**1**	0.5558	0.0159
OA	42	0.7788	0.5250	**0**		0.1370	0.0031	42	0.2217	**0.0409**	**0**		0.1291	0.0031
				1		0.1382	0.0023				**1**		0.1334	0.0023
					0	0.1390	0.0019					**0**	0.1361	0.0016
					1	0.1362	0.0037					**1**	0.1264	0.0043

both lines 3 and 6, which corresponds approximately to the situation around generation 11 (Castro & Garland, 2018). Second, we modeled mini-muscle being fixed in line 3 and still polymorphic at a frequency of 30% in line 6, which approximates frequencies for mice from generation 68 (present study). Simulated data were analyzed in SAS Proc Mixed, using the same syntax as for the real bone data. We quantified statistical power for $\alpha = 0.05$, defined as the probability of rejecting the null hypothesis when it is false (1—Type II error rate), by recording the number of $p < 0.05$ for the linetype variable and for the mini-muscle variable out of 1,000 data sets.

The foregoing simulations did not involve adding among-line variance to the set of four HR lines or to the set of four C lines. This is a reasonable approximation of the situation during the early generations of the selection experiment, but eventually random genetic drift and multiple adaptive responses (in the HR lines) have led to significant among-line variance (Garland et al., 2011; e.g., see Careau et al., 2013), which should decrease the power for detecting linetype effects. Therefore, we also analyzed simulated data for which we added among-line variance for both HR and C mice. Specifically, we added or subtracted the following values, which sum to zero within both HR and C lines: Line 1 −0.1, Line 2 +0.9, Line 3 −0.095, Line 4 −0.05, Line 5 +0.06, Line 6 +0.08, Line 7 −0.04, and Line 8 +0.055, ranging −0.1 to +0.09. We analyzed additional data sets in which we multiplied those values by 2, 3, or 4, thus, further increasing the amount of among-line variance. Simultaneously, we investigated the power to detect linetype and/or mini-muscle effects by multiplying the dependent variable by the values 1.05 (5%, increase) and 1.09 (9% increase) if individuals were HR mice and/or mini-muscle individuals. These data sets were analyzed the same way as described above.

3 | RESULTS

Significance levels from ANCOVAs of skeletal dimensions (using body length as a covariate) are presented in Tables 2 and 3, whereas significance levels from ANOVAs of morphological indices are presented in Table 4. In addition, Tables 1–3 present least square means for all statistical analysis.

3.1 | Body size

Neither body mass nor body length differed significantly between HR and C mice and was not different when comparing mini-muscle mice with normal-muscled individuals for either sex. However, HR male mice tended to be lighter ($p = 0.0810$; Table 2) than C male mice and mini-muscle male mice tended to weigh less ($p = 0.0992$) than normal muscled male mice. Similarly, with body length as a covariate, HR male mice tended to be lighter ($p = 0.0974$; Table 2) than C male mice and mini-muscle male mice weighed less ($p = 0.0453$) than normal muscled male mice. For female mice, we detected no significant linetype or mini-muscle differences in body mass and length-adjusted body mass (Table 2).

3.2 | Skeletal dimensions

Linetype differences were not statistically significant for individual bone lengths of the forelimb, hindlimb, scapula, or pelvis, for either sex (Table 2). HR male mice had thinner distal ilia ($p = 0.0481$; Figure 1a) and lighter pelvises ($p = 0.0500$; Figure 2a) when compared with C male mice. However, HR females did not differ in the breadth of the distal ilium ($p = 0.2263$; Figure 1b), pelvis mass ($p = 0.8532$; Figure 2b), or other pelvis dimensions when compared with C female mice (Table 2). HR mice also had thinner femoral greater trochanters ($p = 0.0012$ for males; $p = 0.0141$ for females) when compared with C mice (Figure 3a,b). For females, HR mice tended to have thicker tibia-fibula mid-shaft diameters ($p = 0.0544$) and heavier tibia-fibulas ($p = 0.0699$; Table 2). In the forelimb, HR males had lighter humeri ($p = 0.0219$; Figure 4a) when compared with C males, but females did not significantly differ ($p = 0.9361$; Figure 4b; Table 3). For females, HR mice tended to have thicker humeri ($p = 0.0732$; Table 3).

Mini-muscle male mice had longer lower ilia ($p = 0.0058$; Figure 5a), but shorter pubic bones ($p = 0.0395$), when compared with normal muscled male mice (Table 1). Mini-muscle female mice, however, did not differ in the length of the lower ilia ($p = 0.5093$; Figure 5b) or pubic bones ($p = 0.5998$) when compared with normal-muscled female mice. Moreover, mini-muscle male mice had narrower distal ilia ($p = 0.0163$; Figure 1a) and lighter pelvises ($p = 0.0074$; Figure 2a; Table 2). For both sexes of mini-muscle mice, the pubis was thinner ($p = 0.0036$ for males; $p = 0.0303$ for females) when compared with normal-muscled mice. Mini-muscle male mice also had thinner femoral mid-shaft diameters ($p = 0.0001$) and proximal tibias ($p = 0.0330$), and lighter femora ($p = 0.0019$; Table 2). Similarly, mini-muscle mice had thinner femoral third trochanters for both males and females ($p < 0.0001$ and $p = 0.0038$, respectively), as well as thinner tibia-fibula mid-shaft diameters ($p < 0.0001$ for both males and females), and lighter tibia-fibulas ($p = 0.0030$ and $p = 0.0049$, respectively). In the forelimb, mini-muscle male mice had thinner scapulae ($p = 0.0037$) and lighter humeri ($p = 0.0087$) when compared with normal-muscled male mice (Table 3). In contrast, mini-muscle females had longer ulnas ($p = 0.0469$) when compared with normal-muscled females. For the rest of the forelimb skeletal traits, mini-muscle females did not differ significantly from normal-muscled females (Table 3).

3.3 | Morphometric indices

Results for the functional indices are presented in Table 4 (descriptions in Table 1). HR males had higher HRI indices ($p = 0.0420$) when compared with C male mice, suggesting more robust humeri (Table 4). In addition, mini-muscle males had reduced HRI indices ($p = 0.0314$) when compared with normal-muscled males, suggesting less robust humeri. Likewise, the FRI indices were reduced in mini-muscle male mice ($p = 0.0003$), indicating less robust femurs (Table 4). For females, mini-muscle mice had higher

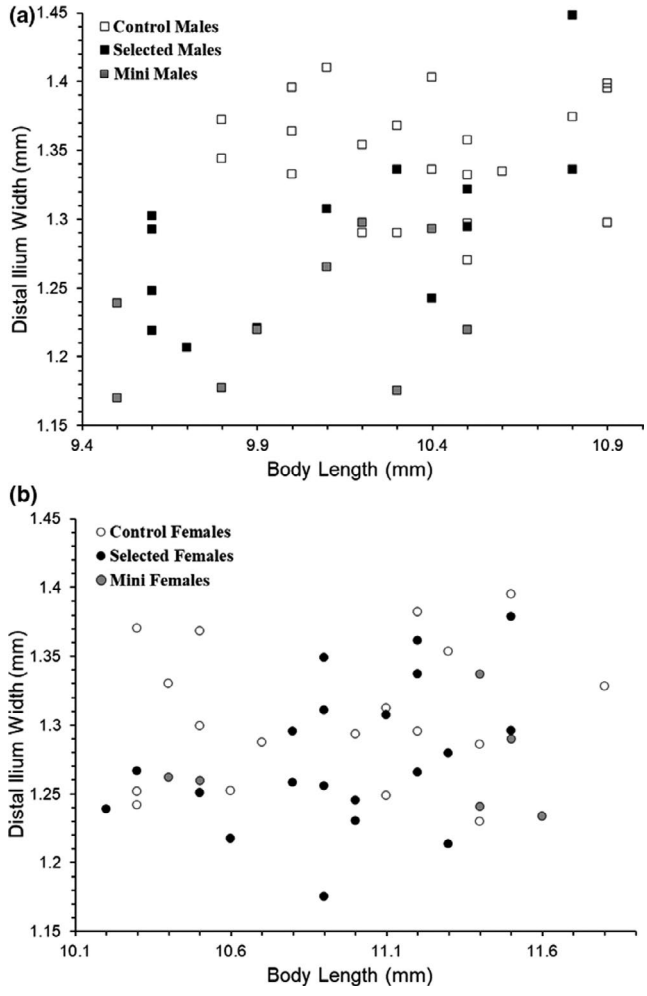

FIGURE 1 (a and b) Mean distal Ilium width in relation to body length. For males only, the effect of body length was positive and statistically significant. HR males had significantly narrower distal ilia for a given body length, and mini-muscle males had thinner distal ilia when compared to normal-muscled males. For females, there was no signifcant effect of either linetype or mini-muscle

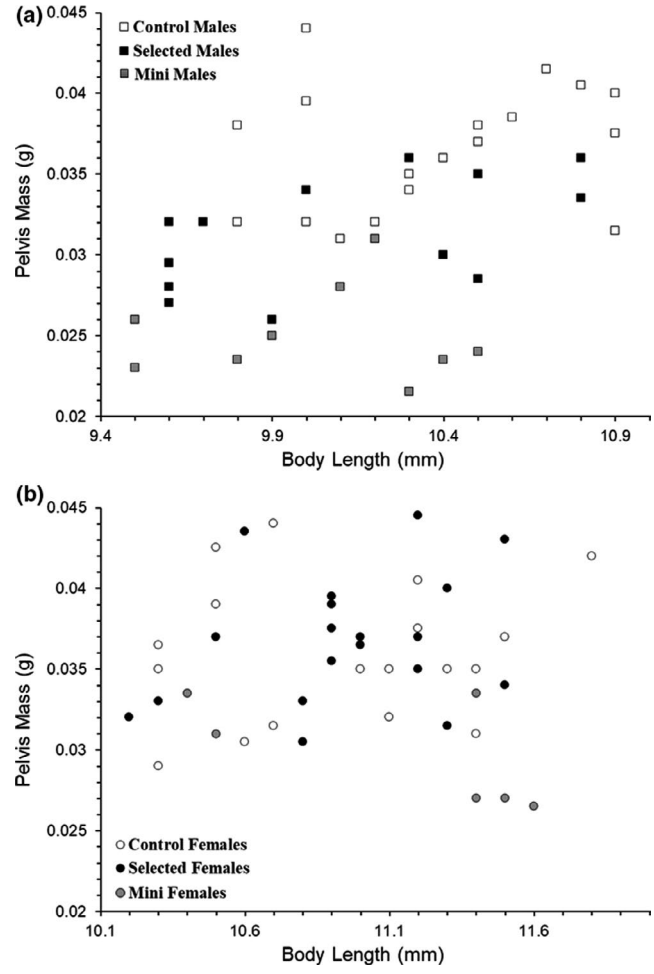

FIGURE 2 (a and b) Mean pelvis mass in relation to body length. For males only, the effect of body length was positive and statistically significant. HR males had significantly lighter pelvises for a given body length, and mini-muscle males had lighter pelvises when compared to normal-muscled males. For females, there was no signifcant effect of linetype, but mini-muscle mice had lighter pelvises only at larger body lengths

3rd/F indices ($p = 0.0415$), indicating increased anatomical advantage (in-lever/out-lever lengths) of the quadratus femoris muscle. In contrast, mini-muscle female mice had reduced OA indices ($p = 0.0409$) when compared with normal-muscled female mice, suggesting reduced anatomical advantage of the triceps brachii (Table 4).

3.4 | Interactions with body length

Overall, we found little statistical support for models that included interactions with body length and full analyses of interactions between skeletal dimensions and body length are presented in Supporting Information (results not shown). In the interaction models for pelvis mass, the mini * body length interaction was significant ($p = 0.0412$) for female mice only. Inspection of Figure 2a, shows that mini-muscle female mice have lighter pelvises at larger body length values only.

3.5 | Multi-generational comparisons

We graphed the least square means for skeletal dimensions after adjusting for body size, and with associated standard error bars over generation time for both HR and C mice. Variation in bone dimensions in all of the populations from each generation can be partly attributed to differences in age and body size with older mice weighing more. However, we describe general differences between HR and C mice over generations for which we had femoral data. HR male mice had significantly thicker femoral heads when compared with C mice in earlier generations 11 and 22, but these differences were not apparent in generation 68 (Figure 6). However, HR female mice had thicker femoral heads throughout most of the generations sampled when compared with C female mice, athough the results were not always statistically significant. HR male mice had significantly broader distal femora when compared with C mice in earlier generations, but these differences were not apparent in generation 68 (e.g., C mice

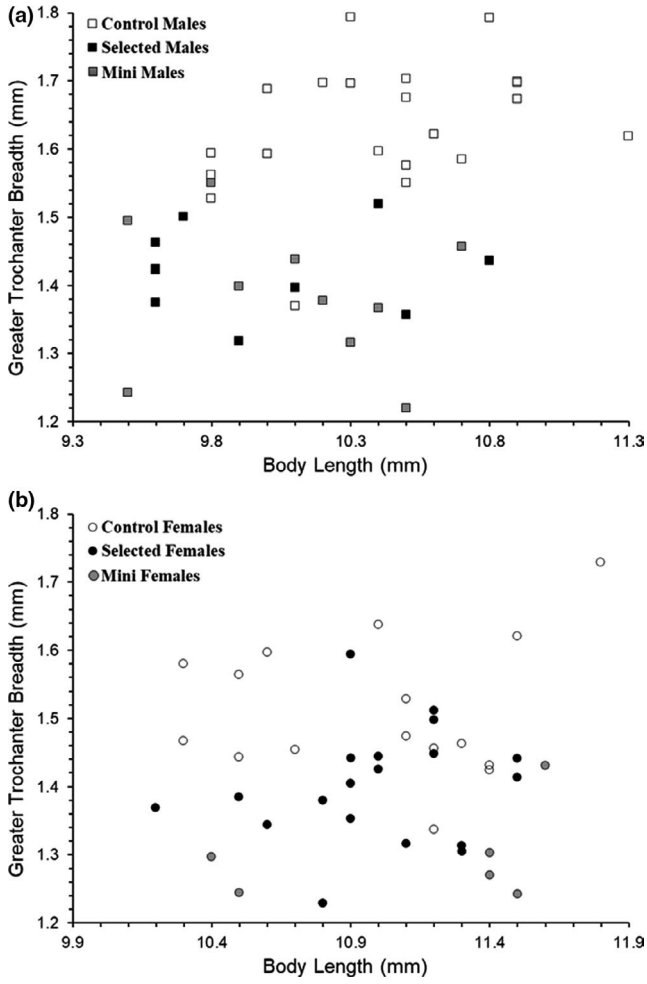

FIGURE 3 (a and b) Mean femoral greater trochanter in relation to body length. HR male, and female mice had significantly thinner femoral greater trochanters for a given body length, and mini-muscle females had thinner femoral greater trochanters as compared with normal-muscled females

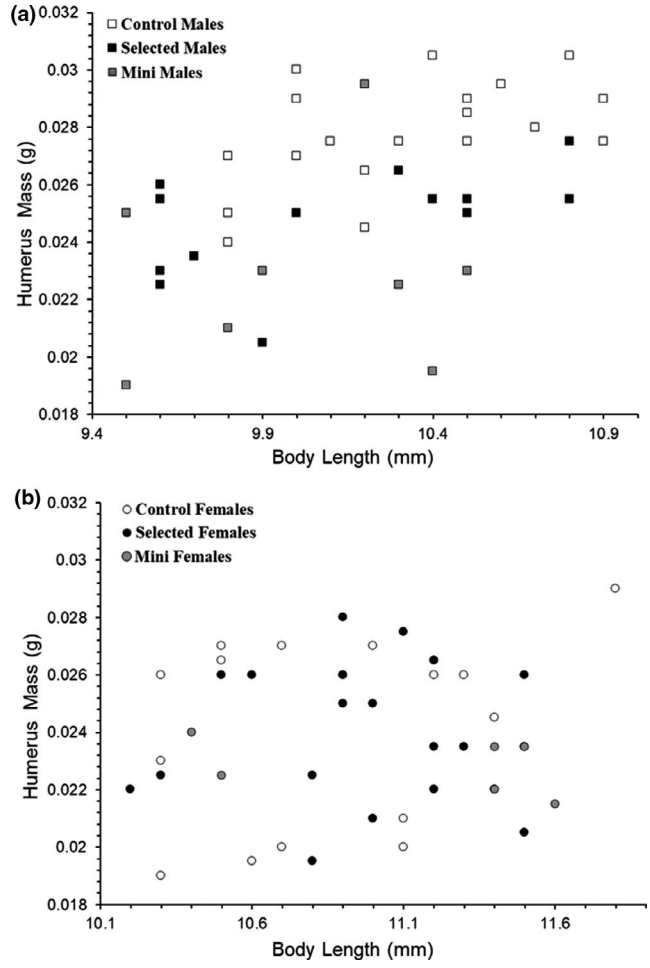

FIGURE 4 (a and b) Mean humerus mass in relation to body length. For males only, the effect of body length was positive and statistically significant. HR males had significantly lighter humeri for a given body length, and mini-muscle males had lighter humeri when compared to normal-muscled males. For females, there was no signifcant effect of either linetype or mini-muscle

had larger knees when compared with HR mice; Figure 6). Finally, HR female mice had thicker distal femora throughout the generations sampled when compared with C female mice, but especially so at generation 16.

3.6 | Simulations to explore statistical power

For simulations under the null hypothesis, the Type I error rate for the mini-muscle effect was very close to the expected 5% (Figure 7). In contrast, the Type I error rate for the linetype effect was only 1.4%.

As expected, power increased with the magnitude of the simulated difference between HR and C lines or between mini-muscle and normal individuals (Figure 7). Overall, the difference in mini-muscle frequency had little effect on the power to detect either a linetype effect or a mini-muscle effect. However, when the magnitude of the HR versus C difference was 6% or more, the power to detect a linetype effect was slightly higher (never >0.04) when mini-muscle frequency was 50% in both HR lines (Mini50 in Figure 7).

When we added among-line variance under the null hypothesis, the Type I error rate for the linetype effect decreased when the magnitude of among-line variance was increased (Online Supplemental S3). However, the Type I error rate for mini-muscle was unaffected under the Mini50 conditions, but was inflated under the MiniFix scenario (Online Supplemental S3). As expected, the power to detect linetype effects decreased as the magnitude of among-line variance was increased, and in similar ways under both scenarios for the mini-muscle phenotype (MiniFix and Mini50; Online Supplemental S3). Similarity, the power to detect mini-muscle effects decreased when the magnitude of among-line variance was increased but was similar between the two frequencies of mini-muscle phenotype that we considered.

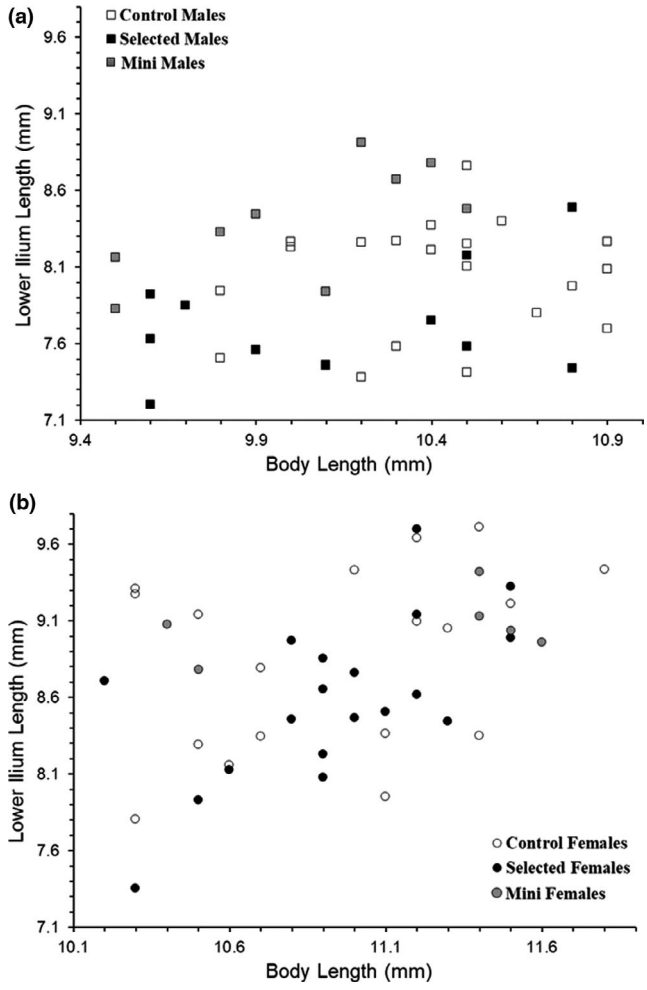

FIGURE 5 (a and b) Mean lower ilium length in relation to body length. For both sexes, the effect of body length was positive and statistically significant. Mini-muscle males had significantly longer lower ilia for a given body length. For females, there was no signifcant effect of either linetype or mini-muscle

4 | DISCUSSION

We compared skeletal dimensions of four replicate lines of house mice that have been selectively bred for high levels of voluntary wheel-running behavior with those of four non-selected Control lines at generation 68. With body length as a covariate, mice from the High Runner lines showed relatively few differences from the C lines. This general result was surprising because previous studies at generations 11 (Castro & Garland, 2018), 16 (Middleton et al., 2008b), and 21 (Kelly et al., 2006) showed relatively more skeletal differences between the HR and C lines for linear dimensions, many of which appeared to be adaptive with respect to endurance-running ability. In the Conclusions, we offer several possible explanations for this pattern.

4.1 | High runner versus control lines

A previous study of female mice from generation 57 found that the cortical bone area of the distal ilium (mm^2) of the pelvis was not significantly different between HR and C mice, controlling for body size (Lewton et al., 2019). The present results for females are consistent with that study; however, we found that HR male mice (not studied at generation 57) had significantly thinner distal ilia when compared with C male mice (Table 2). In addition, we found that the femoral greater trochanter was significantly thinner in HR mice for both sexes. What is the functional significance of these differences? Based on the 3D reconstructions of mouse hindlimbs in Charles et al. (2016a; Online Supplemental S4), the gluteal muscles originate on the ilium and insert on the femoral greater trochanter, acting to rotate, and abduct the hip joint during locomotion. Therefore, having both thinner femoral greater trochanters and distal ilia may indicate reduced muscular forces required to rotate and abduct the hip joint during sustained locomotion (Carrano, 1999). Further study with electromyography and sonomicrometry, combined with kinematics, might be used to test these ideas, but were beyond the scope of the present study.

Controlling for body length, HR male mice had lighter pelvises (Table 2) and humeri (Table 3) than those of C male mice. These differences may reduce the muscular forces required during sustained locomotion (i.e., reduce the kinetic energy required to overcome inertia through the swing phase of each stride; Carrano, 1999).

4.2 | Mini-muscle phenotype

In the present study, mini-muscle mice tended to be smaller in body mass, and in general had thinner and lighter bones (Tables 2–4). For example, mini-muscle males had lighter pelvises, thinner distal ilia and pubic bones, and mini-muscle mice of both sexes had thinner proximal ilia when compared to normal-muscled mice, resulting in a more gracile pelvis overall (Table 2). In addition, mini-muscle males had thinner and lighter femora and mini-muscle mice of both sexes had thinner and lighter tibia-fibulas when compared to normal-muscled mice, resulting in a more gracile hindlimb overall. Functionally, a more gracile pelvis, femur, and tibia-fibula should reduce muscular forces required to overcome inertia through the swing phase of each stride (Carrano, 1999). Moreover, mini-muscle male mice had narrower scapulae and humeri (Table 3), which may allow forelimb muscles to produce larger movements of the humerus during locomotion, as is found in some cursorial mammals (Hopwood, 1947; Maynard Smith & Savage, 1956).

In mammals (including mice), the gluteal muscles originate on the ilium and function to extend and externally rotate the hip joint during locomotion (Álvarez et al., 2013; Charles et al., 2016b; Polly, 2007). Based on phylogenetic analysis, Álvarez et al (2013) found that cursorial mammals have an elongated ilium, wide ramus of ischium, and a reduced pectineal tuberosity, although that data set only encompassed small- to medium-sized mammals (0.04–62 kg). Several authors have suggested that an elongated ilium increases the moment arm and, therefore, mechanical advantage of the hip and knee extensors (including the gluteal muscles; Álvarez et al., 2013; Lewton, 2015; Maynard Smith & Savage, 1956; Polly, 2007) since muscle fiber length

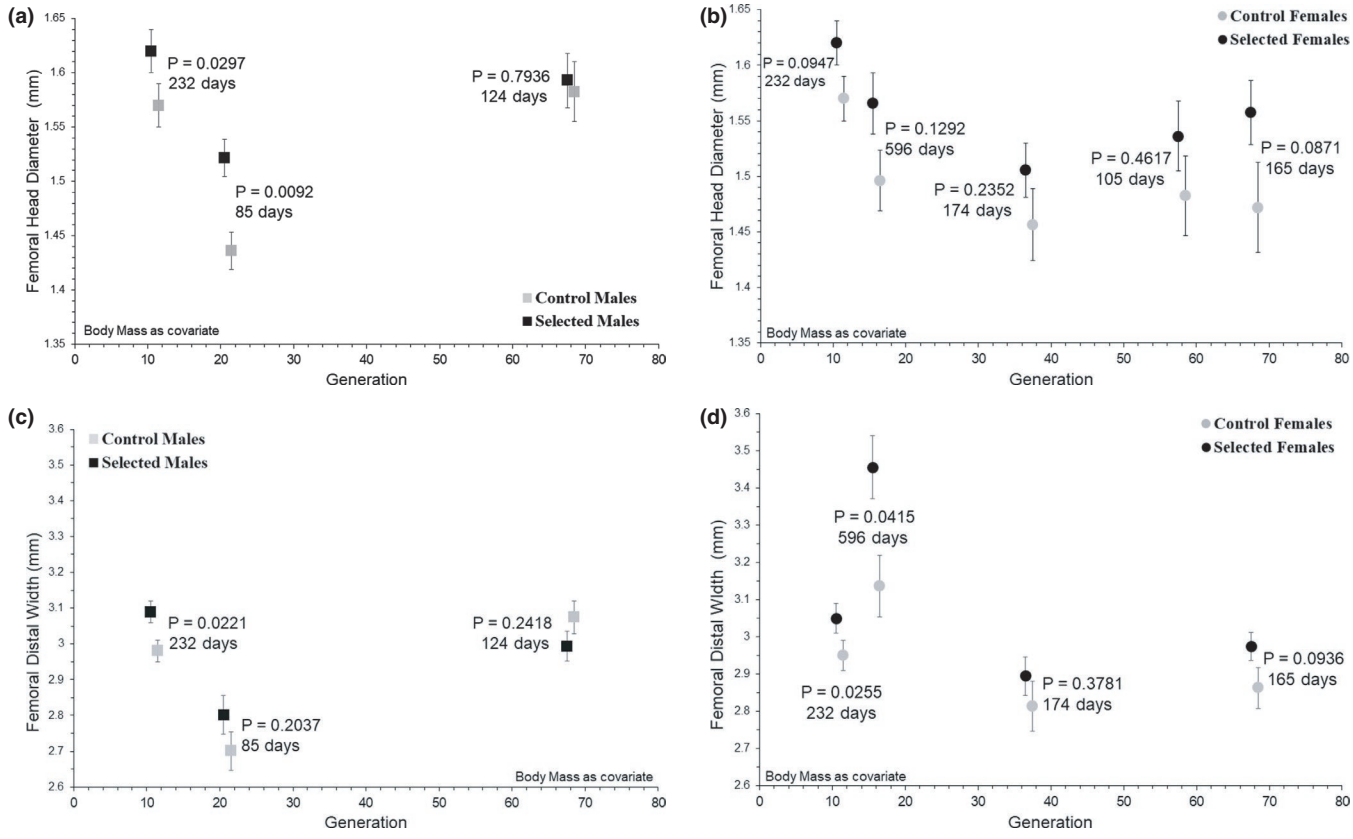

FIGURE 6 Mean femoral dimensions across generations. Femoral dimensions are plotted across generations separately for males and females. Values are least square means and associated standard error bars, with body mass as a covariate. Age at sampling varied among generations. (a and b: HR male mice had thicker femoral heads when compared with C mice in earlier generations, but these differences were not apparent at generation 69. HR female mice had thicker femoral heads throughout the generations sampled, although these differences were not always statistically significant. (c and d): HR male mice had thicker distal femora when compared with C mice at generation 11, but this difference was not statistically significant at generations 21 or 69. HR female mice had thicker distal femora throughout the generations sampled, but especially so at generation 16

positively correlates with the moment arm of muscles (McClearn, 1985). We found that mini-muscle male mice had longer lower ilium lengths when compared to normal-muscled males. An elongated lower ilium likely increases the moment arm of the knee extensors and gluteal muscles, which would be beneficial during initial propulsion when running.

Overall, the effects of the mini-muscle phenotype on the appendicular skeleton could reduce mechanical costs when running on wheels and are akin to cursorial adaptations in the skeleton of mammals (as suggested by Kelly et al., 2006). Indeed, mini-muscle individuals run at higher average and maximal speeds on wheels (Claghorn et al., 2017; Kelly et al., 2006; Singleton & Garland, 2019), but, surprisingly, they have higher costs of transport when running on wheels, in combination with reduced maximal sprint speeds (which are substantially higher than wheel-running speeds) when chased along a racetrack (Dlugosz et al., 2009).

4.3 | Simulations to explore statistical power

A concern regarding the results from the present study at generation 68 as compared with earlier ones was that the increased frequency of the mini-muscle phenotype would cause a reduction in statistical power to detect linetype effects. Such a reduction could explain why some differences between HR and C lines, detected at generation 11 and/or 21, were no longer statistically significant (e.g., see Figure 6). Therefore, we simulated data with a frequency of 50% mini-muscle individuals in both HR lines #3 and #6, reflective of earlier generations (Garland & Freeman, 2005), and compare it with data simulated to have HR line #3 fixed for the mini-muscle phenotype, with a frequency 30% for HR line #6, which approximates frequencies from mice at generation 68.

Results from our first set of simulations showed that when the magnitude of the HR versus C difference exceeds 6%, the power to detect linetype differences was slightly higher when the mini-muscle frequency was 50% in both HR lines that have it, although the increase in power was never >4%. Moreover, the average difference in least squares means for linear bone dimensions was only 5%. Hence, the statistical power to detect general differences between the HR and C lines may have been higher in earlier generations, but the effect is so small (Figure 7) that it probably cannot account for apparent loss of some differences in later generations. Furthermore, the power to detect a mini-muscle effect was not affected by the frequency of the mini-muscle phenotype (Figure 7).

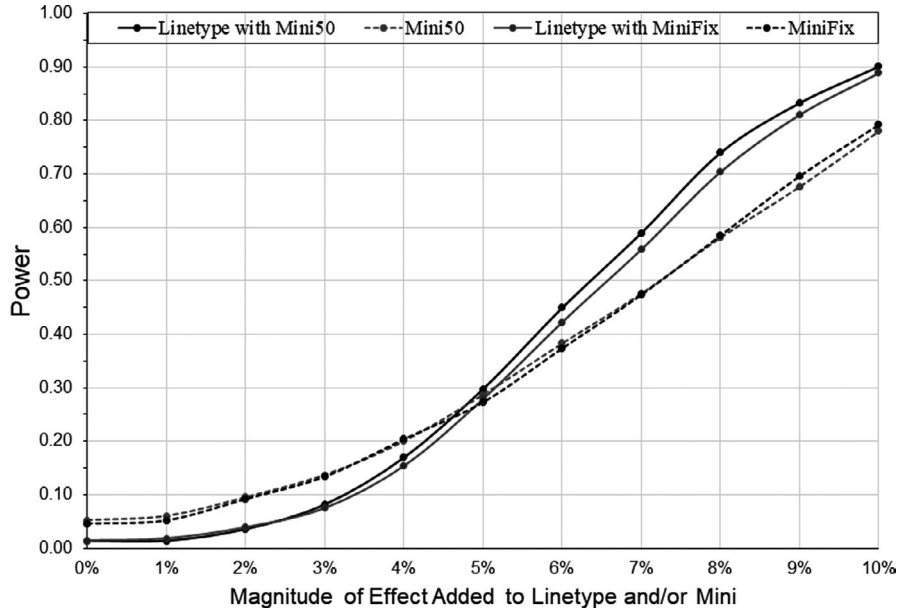

FIGURE 7 Power increased with the magnitude of the simulated difference between HR and C lines or between mini-muscle and normal individuals. When the magnitude of the HR versus C difference exceeded 6%, the power to detect a linetype effect was slightly higher when in the Mini50 models when compared with the MiniFix models, although the increase in power was never greater than ~0.04. The power to detect a mini-muscle effect also increased with the magnitude of the effect, but the power showed little difference between Mini50 and MiniFix. Under the null hypothesis, the Type I error rates for the mini-muscle effects were very close to the expected 5%, but for the HR versus C effect it was greatly deflated (~1.4%)

TABLE 5 The number of statistically significant (nominal $p < 0.05$, not adjusted for multiple comparisons) linear bone dimensions in comparisons between mice from the HR and C lines. The skeletal traits analyzed include linear dimensions of the forelimb and hindlimb measured as a part of many previous studies. "#studied" is the combined tally for both sexes. For example, in the present study, 55 traits were measured for both males and females. Here, "significant" refers to analyses with no correction for multiple comparisons in any study

References	Generation	Males	Females	#studied	#significant	%significant
Garland and Freeman (2005) and Castro and Garland (2018)	11	X	X	58	3	5
Middleton et al. (2008b)	16		X	9	2	22
Kelly et al. (2006) and Young et al. (2009)	21	X		26	8	30
Middleton et al. (2010)	37		X	6	0	0
Copes et al. (2018) and Lewton et al. (2019); Present study	57		X	3	0	0
Present study	68	X	X	110	7	6

In a second set of simulations, we added among-line variance to model the likely effects of random genetic drift and, in the selected lines, possible multiple solutions (Garland et al., 2011). Increased among-line variance should reduce the power to detect linetype effects. As expected, increasing the magnitude of among-line variance decreased the power to detect both linetype and mini-muscle effects (Online Supplemental Table S3).

A surprising result for the simulations under the null hypothesis was that the Type I error rate for detecting the linetype effect (but not for detecting the mini-muscle effect) was only ~1% for $\alpha = 0.05$, suggesting that our analysis of the linetype effects may be generally underpowered. As expected, simulations with increased among-line variance caused the Type 1 error rate for linetype effects to decrease even further, emphasizing a potential reduction in power across generations.

4.4 | Changing effects of long-term selection on bone dimensions

Aside from the High Runner mouse experiment, no other long-term selection experiment using vertebrates has investigated how morphological traits can coadapt with locomotor behavior over tens of generations (>60). This experiment provides a unique opportunity to investigate how such coadaptation may differ before and after selection limits have been attained. Limits for wheel-running behavior

occurred at approximately ~17–27 generations (Careau et al., 2013). The first published studies of skeletal traits were from generation 11, well before the selection limit, and they reported reduced levels of asymmetry in the HR lines, and that the HR mice had larger knee and hip joints (Garland & Freeman, 2005). Subsequent studies at generations 16 and 21 confirmed the differences in knee and hip joints, and also found that HR mice had evolved thicker mid-shafts of the hindlimb and heavier feet (Kelly et al., 2006; Middleton et al., 2008b; Wallace & Garland, 2016).

Considering the additional studies published since generation 21, the overall pattern suggests coadaptation of limb bone dimensions with running behavior that became more apparent across generations prior to the selection limits, and then diminished after the limits were reached. Specifically, during earlier generations, relatively more linear bone dimensions were found to be significantly different between HR and C lines (5%, 22%, and 30% for generations 11, 16, and 21, respectively) as compared with the 6% found significant at generation 68 (Table 5). For example, as shown in Figure 6, HR male mice had evolved larger femoral heads and distal femoral by at least generation 11, and maintained this difference at generation 21 (Castro & Garland, 2018; Garland & Freeman, 2005; Kelly et al., 2006), but these differences were not evident at generation 68. In contrast, for female mice, the differences in limb diameters of the knee and hip joint remained relatively constant throughout the generations sampled.

4.5 | Summary and future directions

We investigated coadaptation of the appendicular skeleton with locomotor behavior in the unique High Runner mouse model across many generations of selective breeding, including both before and after selection limits for the behavior had been attained. This is the first study in a vertebrate that considers the extent to which skeletal traits (and morphological traits in general) coadapt with locomotor behavior over many generations of artificial selection. In previous studies, we found that the skeleton evolves rapidly as a response to directional selection for high levels of voluntary wheel running (as early as generation 11). Further evidence of skeletal coadaptation was found at generations 16 (Middleton et al., 2008b) and 21 (Kelly et al., 2006; Schutz et al., 2014; Wallace et al., 2012). However, after selection limits occurred, some skeletal adaptations were lost.

The apparent loss of skeletal coadaptations might be explained in at least two ways. First, deterioration may have occurred due to a gradual increase in inbreeding, which was compensated by other adaptive changes that replaced the functional benefits of some skeletal changes (e.g., adaptive changes in muscles, Bilodeau et al., 2009 or gaits, Claghorn et al., 2017). This possibility could be addressed by compiling and comparing the cross-generational trajectories of multiple other traits related to endurance capacity (e.g., heart mass, skeletal muscle mass). Another approach would be to cross the replicate selected lines and test for heterosis (hybrid vigor) in various traits (e.g., Bult & Lynch, 1996; Hannon et al., 2011; Hiramatsu, 2017; Miyatake, 2002).

Second, increases in the frequency of a gene of major effect on muscle mass (causing the "mini-muscle" phenotype), which also has numerous effects on skeletal dimensions, reduced statistical power to detect differences between HR and C lines. At the same time, increased levels of among-line variation within the HR and/or C lines could have reduced the statistical power to detect linetype effects. The former possibility seems not to be the case based on the simulations we present (Figure 7). The latter possibility deserves more study, based on the magnitude of the difference in linetype least squares means relative to the standard errors (e.g., Figure 6b), as well as additional simulations (Online Supplemental S3).

ACKNOWLEDGEMENTS

We thank S. Bansode, D. Buenaventura, M. Needham, and D. Thai for assistance with skeletal preparations and preliminary measurements. D. Hillis provided R code for simulations. Supported by NSF grant DEB-1655362 to TG. The authors declare no conflicts of interest.

AUTHOR CONTRIBUTIONS

AAC, HR, and GCC collected data. AAC, HR, and TG analyzed the data. AAC, HR, and TG wrote the manuscript. All authors edited the manuscript.

ORCID

Theodore Garland https://orcid.org/0000-0002-7916-3552

REFERENCES

Álvarez, A., Ercoli, M.D. & Prevosti, F.J. (2013) Locomotion in some small to medium-sized mammals: A geometric morphometric analysis of the penultimate lumbar vertebra, pelvis and hindlimbs. *Zoology*, 116, 356–371. https://doi.org/10.1016/j.zool.2013.08.007

Angilletta Jr., M.J., Bennett, A.F., Guderley, H., Navas, C.A., Seebacher, F. & Wilson, R.S. (2006) Coadaptation: A unifying principle in evolutionary thermal biology. *Physiological and Biochemical Zoology*, 79, 282–294. https://doi.org/10.1086/499990

Biancardi, C.M. & Minetti, A.E. (2012) Biomechanical determinants of transverse and rotary gallop in cursorial mammals. *Journal of Experimental Biology*, 215, 4144–4156. https://doi.org/10.1242/jeb.073031

Biewener, A.A. (1990) Biomechanics of mammalian terrestrial locomotion. *Science*, 250, 1097. https://doi.org/10.1126/science.2251499

Bilodeau, G.M., Guderley, H., Joanisse, D.R. & Garland Jr., T. (2009) Reduction of type IIb myosin and IIB fibers in tibialis anterior muscle of mini-muscle mice from high-activity lines. *Journal of Experimental Zoology Part A: Ecological Genetics and Physiology*, 311A, 189–198. https://doi.org/10.1002/jez.518

Bramble, D.M. & Lieberman, D.E. (2004) Endurance running and the evolution of Homo. *Nature*, 432, 345–352. https://doi.org/10.1038/nature03052

Bult, A. & Lynch, C.B. (1996) Multiple selection responses in house mice bidirectionally selected for thermoregulatory nest-building behavior: Crosses of replicate lines. *Behavior Genetics*, 26, 439–446. https://doi.org/10.1007/BF02359488

Burke, M.K., Barter, T.T., Cabral, L.G., Kezos, J.N., Phillips, M.A., Rutledge, G.A. et al. (2016) Rapid divergence and convergence of life-history in experimentally evolved *Drosophila melanogaster*: Rapid divergence and convergence. *Evolution*, 70, 2085–2098.

Careau, V., Wolak, M.E., Carter, P.A. & Garland Jr., T. (2013) Limits to behavioral evolution: The quantitative genetics of a complex trait under directional selection: quantitative genetics of a selection limit. *Evolution*, 67, 3102–3119.

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.

Castro, A.A. & Garland Jr., T. (2018) Evolution of hindlimb bone dimensions and muscle masses in house mice selectively bred for high voluntary wheel-running behavior. *Journal of Morphology*, 279, 766–779.

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

Charles, J.P., Cappellari, O., Spence, A.J., Wells, D.J. & Hutchinson, J.R. (2016b) Muscle moment arms and sensitivity analysis of a mouse hindlimb musculoskeletal model. *Journal of Anatomy*, 229(4), 514–535. https://doi.org/10.1111/joa.12461

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., Dlugsoz, E.M., Judex, S. & Garland Jr., T. (2018) Locomotor activity, growth hormones, and systemic robusticity: An investigation of cranial vault thickness in mouse lines bred for high endurance running. *American Journal of Physical Anthropology*, 166, 442–458.

Cowgill, L.W. (2009) The ontogeny of Holocene and Late Pleistocene human postcranial strength. *American Journal of Physical Anthropology*. https://doi.org/10.1002/ajpa.21107

Dickinson, M.H., Farley, C.T., Full, R.J., Koelh, M.A.R., Kram, R. & Lehman, S. (2000) How animals move: An integrative view. *Science*, 288, 100–106. https://doi.org/10.1126/science.288.5463.100

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.

Foster, K.L., Garland Jr., T., Schmitz, L. & Higham, T.E. (2018) Skink ecomorphology: Forelimb and hind limb lengths, but not static stability, correlate with habitat use and demonstrate multiple solutions. *Biological Journal of the Linnean Society*, https://doi.org/10.1093/biolinnean/bly146

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

Garland Jr., T. & Freeman, P.W. (2005) Selective breeding for high endurance running increases hindlimb symmetry. *Evolution*, 59, 1851–1854. https://doi.org/10.1111/j.0014-3820.2005.tb01832.x

Garland Jr., T. & Janis, C.M. (1993) Does metatarsal/femur ratio predict maximal running speed in cursorial mammals? *Journal of Zoology*, 229, 133–151. https://doi.org/10.1111/j.1469-7998.1993.tb02626.x

Garland Jr., T., Kelly, S.A., Malisch, J.L., Kolb, E.M., Hannon, R.M., Keeney, B.K. et al. (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. https://doi.org/10.1098/rspb.2010.1584

Garland Jr., T., Morgan, M.T., Swallow, J.G., Rhodes, J.S., Girard, I., Belter, J.G. et al. (2002) Evolution of a small-muscle polymorphism in lines of house mice selected for high activity levels. *Evolution*, 56, 1267–1275. https://doi.org/10.1111/j.0014-3820.2002.tb01437.x

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

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.

Hannon, R.M., Meek, T.H., Acosta, W., Maciel, R.C., Schutz, H. & Garland, Jr., T. (2011) Sex-specific heterosis in line crosses of mice selectively bred for high locomotor activity. *Behavior Genetics*, 41, 615–624. https://doi.org/10.1007/s10519-010-9432-3

Hildebrand, M. (1974) *Analysis of vertebrate structure*. John Wiley and Sons.

Hiramatsu, L. (2017) *Physiological and genetic causes of a selection limit for voluntary wheel-running in mice*. University of California.

Hiramatsu, L., Kay, J.C., Thompson, Z., Singleton, J.M., Claghorn, G.C., Albuquerque, R.L. et al. (2017) Maternal exposure to Western diet affects adult body composition and voluntary wheel running in a genotype-specific manner in mice. *Physiology & Behavior*, 179, 235–245. https://doi.org/10.1016/j.physbeh.2017.06.008

Hopwood, A.T. (1947) Contributions to the study of some African Mammals. III. Adaptations in the bones of the fore-limb of the lion, leopard, and cheetah. *Journal of the Linnean Society of London, Zoology*, 41, 259–271. https://doi.org/10.1111/j.1096-3642.1940.tb02076.x

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. https://doi.org/10.1152/ajpregu.00179.2002

Howell, A.B. (1944) *Speed in animals*. University of Chicago Press.

Huey, R. & Bennett, A.F. (1987) Phylogenetic studies of coadaptation: Preferred temperature versus optimal performance temperatures of lizards. *Evolution*, 41, 18.

Jenkins, F.A. & Camazine, S.M. (1977) Hip structure and locomotion in ambulatory and cursorial carnivores. *Journal of Zoology*, 181, 351–370. https://doi.org/10.1111/j.1469-7998.1977.tb03249.x

Jones, K.E. (2016) New insights on equid locomotor evolution from the lumbar region of fossil horses. *Proceedings of the Royal Society B: Biological Sciences*, 283, 20152947. https://doi.org/10.1098/rspb.2015.2947

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. https://doi.org/10.1002/jmor.10407

Kelly, S.A., Gomes, F.R., Kolb, E.M., Malisch, J.L. & Garland Jr., T. (2017) Effects of activity, genetic selection and their interaction on muscle metabolic capacities and organ masses in mice. *Journal of Experimental Biology*, 220, 1038–1047.

Kimes, K.R., Siegel, M.I. & Sadler, D.L. (1981) Alteration of scapular morphology through experimental behavioral modification in the laboratory mouse *Mus musculus*. *Cells Tissues Organs*, 109, 161–165.

Lewton, K.L. (2015) Pelvic form and locomotor adaptation in strepsirrhine primates: Primate pelvic shape and locomotion. *Anatomical Record*, 298, 230–248.

Lewton, K.L., Ritzman, T., Copes, L.E., Garland Jr., T. & Capellini, T.D. (2019) Exercise-induced loading increases ilium cortical area in a selectively bred mouse model. *American Journal of Physical Anthropology*, 168, 543–551. https://doi.org/10.1002/ajpa.23770

Lovegrove, B.G. & Mowoe, M.O. (2014) The evolution of micro-cursoriality in mammals. *Journal of Experimental Biology*, 217, 1316–1325. https://doi.org/10.1242/jeb.095737

Marchini, M., Sparrow, L.M., Cosman, M.N., Dowhanik, A., Krueger, C.B., Hallgrimsson, B. et al. (2014) Impacts of genetic correlation on the independent evolution of body mass and skeletal size in mammals. *BMC Evolutionary Biology*, 14, 258. https://doi.org/10.1186/s12862-014-0258-0

Maynard Smith, J. & Savage, R.J.G. (1956) Some locomotory adaptations in mammals. *Journal of the Linnean Society of London, Zoology*, 42, 603–622. https://doi.org/10.1111/j.1096-3642.1956.tb02220.x

McClearn, D. (1985) Anatomy of raccoon *Procyon lotor* and coati *Nasua narica and N. nasua* forearm and leg muscles: Relations between fiber length, moment-arm length, and joint-angle excursion. *Journal of Morphology*, 183, 87–115.

Meachen-Samuels, J. & Van Valkenburgh, B. (2009) Forelimb indicators of prey-size preference in the Felidae. *Journal of Morphology*, 270, 729–744. https://doi.org/10.1002/jmor.10712

Middleton, K.M., Goldstein, B.D., Guduru, P.R., Waters, J.F., Kelly, S.A., Swartz, S.M. et al. (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. https://doi.org/10.1111/j.1469-7580.2009.01175.x

Middleton, K.M., Kelly, S.A. & Garland Jr., T. (2008a) Selective breeding as a tool to probe skeletal response to high voluntary locomotor activity in mice. *Integrative and Comparative Biology*, 48, 394–410. https://doi.org/10.1093/icb/icn057

Middleton, K.M., Shubin, C.E., Moore, D.C., Carter, P.A., Garland Jr., T. & Swartz, S.M. (2008b) The relative importance of genetics and phenotypic plasticity in dictating bone morphology and mechanics in aged mice: Evidence from an artificial selection experiment. *Zoology*, 111, 135–147. https://doi.org/10.1016/j.zool.2007.06.003

Miyatake, T. (2002) Circadian rhythm and time of mating in *Bactrocera cucurbitae* (Diptera: Tephritidae) selected for age at reproduction. *Heredity*, 88, 302–306. https://doi.org/10.1038/sj.hdy.6800044

Morris, J.S. & Carrier, D.R. (2016) Sexual selection on skeletal shape in Carnivora. *Evolution*, 70, 767–780. https://doi.org/10.1111/evo.12904

Oxnard, C.E. (1967) The functional morphology of the primate shoulder as revealed by comparative anatomical, osteometric and discriminant function techniques. *American Journal of Physical Anthropology*, 26, 219–240.

Polly, P.D. (2007) Limbs in mammalian evolution. In: Hall, B.K. (ed.) *Fins into limbs: Evolution, development, and transformation*. University of Chicago Press, pp. 245–268.

Rose, M.R. (2005) The effects of evolution are local: Evidence from experimental evolution in *Drosophila*. *Integrative and Comparative Biology*, 45, 486–491. https://doi.org/10.1093/icb/45.3.486

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. https://doi.org/10.1002/jmor.20077

Samuels, J.X. & Van Valkenburgh, B. (2008) Skeletal indicators of locomotor adaptations in living and extinct rodents. *Journal of Morphology*, 269, 1387–1411. https://doi.org/10.1002/jmor.10662

Schneider, C.A., Rasband, W.S. & Eliceiri, K.W. (2012) NIH Image to ImageJ: 25 years of image analysis. *Nature Methods*, 9, 671–675. https://doi.org/10.1038/nmeth.2089

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. https://doi.org/10.1111/evo.12501

Schwartz, N.L., Patel, B.A., Garland, Jr., T. & Horner, A.M. (2018) Effects of selective breeding for high voluntary wheel-running behavior on femoral nutrient canal size and abundance in house mice. *Journal of Anatomy*, 233, 193–203. https://doi.org/10.1111/joa.12830

Simões, P., Fragata, I., Santos, J., Santos, M.A., Santos, M., Rose, M.R. et al. (2019) How phenotypic convergence arises in experimental evolution. *Evolution*, 73, 1839–1849. https://doi.org/10.1111/evo.13806

Simões, P., Santos, J., Fragata, I., Mueller, L.D., Rose, M.R. & Matos, M. (2008) How is repeatable is adaptive evolution? The role of geographical origin and founder effects in laboratory adaptation. *Evolution*, 62, 1817–1829.

Singleton, J.M. & Garland Jr., T. (2019) Influence of corticosterone on growth, home-cage activity, wheel running, and aerobic capacity in house mice selectively bred for high voluntary wheel-running behavior. *Physiology & Behavior*, 198, 27–41.

Stein, B.R. & Casinos, A. (1997) What is a cursorial mammal? *Journal of Zoology*, 242, 185–192. https://doi.org/10.1111/j.1469-7998.1997.tb02939.x

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.

Syme, D.A., Evashuk, K., Grintuch, B., Rezende, E.L. & Garland Jr., T. (2005) Contractile abilities of normal and "mini" triceps surae muscles from mice *Mus domesticus* selectively bred for high voluntary wheel running. *Journal of Applied Physiology*, 99, 1308–1316.

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. https://doi.org/10.1016/j.mod.2014.08.004

Van Der Klaauw, C.J. (1948) Ecological studies and reviews. IV. Ecological morphology. *Bibliotheca Biotheoretica*, 4, 25–110.

Van Valkenburgh, B. (1987) Skeletal indicators of locomotor behavior in living and extinct carnivores. *Journal of Vertebrate Paleontology*, 7, 162–182. https://doi.org/10.1080/02724634.1987.10011651

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. https://doi.org/10.1046/j.1420-9101.2001.00260.x

Vianey-Liaud, M., Hautier, L. & Marivaux, L. (2015) Morphological disparity of the postcranial skeleton in rodents and its implications for palaeobiological inferences: the case of the extinct Theridomyidae (Rodentia, Mammalia). In: Cox, P.G. & Hautier, L. (eds.) *Evolution of the rodents*. Cambridge University Press, pp. 539–588.

Wallace, I.J. & Garland Jr., T. (2016) Mobility as an emergent property of biological organization: Insights from experimental evolution: Mobility and biological organization. *Evolutionary Anthropology: Issues, News, and Reviews*, 25, 98–104. https://doi.org/10.1002/evan.21481

Wallace, I.J., Judex, S. & Demes, B. (2015) Effects of load-bearing exercise on skeletal structure and mechanics differ between outbred populations of mice. *Bone*, 72, 1–8. https://doi.org/10.1016/j.bone.2014.11.013

Wallace, I.J., Middleton, K.M., Lublinsky, S., Kelly, S.A., Judex, S., Garland Jr., T. et al. (2010) Functional significance of genetic variation underlying limb bone diaphyseal structure. *American Journal of Physical Anthropology*, 143, 21–30. https://doi.org/10.1002/ajpa.21286

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. https://doi.org/10.1002/ajpa.22028

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. https://doi.org/10.1007/s11692-009-9053-z

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section.

How to cite this article: Castro AA, Rabitoy H, Claghorn GC, Garland T. Rapid and longer-term effects of selective breeding for voluntary exercise behavior on skeletal morphology in house mice. *J. Anat.* 2021;238:720–742. https://doi.org/10.1111/joa.13341