UC Riverside

UC Riverside Electronic Theses and Dissertations

Title

Skeletal Evolution and Muscle Physiology in House Mice Selectively Bred for High Levels of Voluntary Wheel-Running Behavior

Permalink

https://escholarship.org/uc/item/7pd0d5mq

Author

Castro, Alberto A

Publication Date

2021

Supplemental Material

https://escholarship.org/uc/item/7pd0d5mq#supplemental

Peer reviewed|Thesis/dissertation

eScholarship.org Powered by the California Digital Library
University of California

UNIVERSITY OF CALIFORNIA
RIVERSIDE

Skeletal Evolution and Muscle Physiology in House Mice Selectively Bred for High Levels of Voluntary Wheel-Running Behavior

A Dissertation submitted in partial satisfaction
of the requirements for the degree of

Doctor of Philosophy

in

Evolution, Ecology, and Organismal Biology

by

Alberto Arturo Castro

December 2021

Dissertation Committee:
Dr. Theodore Garland Jr., Chairperson
Dr. Natalie C. Holt
Dr. Timothy E. Higham

Copyright by
Alberto Arturo Castro
2021

The Dissertation of Alberto Arturo Castro is approved:

_____

_____

_____

Committee Chairperson

University of California, Riverside

Acknowledgements

I would like to extend my thanks to the special people that have made this dissertation possible! For starters, thanks to Dr. Ted Garland for helping me with writing, statistics, and teaching me to think like a scientist. I will always remember Murphy's Law ("*Anything that can go wrong will go wrong*") and to always have a backup plan! Thanks also to Dr. Natalie Holt for inviting me to her lab with open arms. Thanks to Dr. Tim Higham, Dr. Tomasz Owerkowicz, Dr. Mark Springer, Dr. David Reznick, Dr. Rich Cardullo, Dr. Wendy Saltzman, Dr. Kim Hammond, Dr. Mike Fugate, Dr. Sara Becker, Dr. Lynn Copes, Dr. Patricia Freeman, Dr. Jerry Claghorn, Dr. Sudha Bansodhe and Dr. Alex Karaktosis for your scientific input, pedagogical expertise, and valuable advice. Thank you to my undergraduate mentees Hannah Rabitoy, Aayushi Trivedi, Saad Ahmed and Holland McClendon for your hard work and enthusiasm for science! Thank you to all the friends that I have made at UCR both within the biology department and the campus community. Thanks also to my fellow lab mates and friends in EEOB for making graduate school one of the most fun experiences! Thank you to my friends from San Bernardino and beyond for helping me laugh when I felt the blues! Thanks to Robert, Felix, and Gilbert Caballero for being reliable friends! Special thanks to my parents for your unwavering support and believing in my potential! Finally, thanks to my fraternal twin, Patricia Castro for always looking out for me, Daniel Aguilar, for reminding me to appreciate my accomplishments, and Claudia Bruno for helping me grow as a person while pursuing my doctoral degree. Thanks to everyone, it really does take a village!

The copyright material is held jointly with co-authors for each chapter and specific contributions can be found at the beginning of each chapter. Chapters 1, 2, and 3 are fully published articles and full citations are below.

Castro, A. A., and T. Garland Jr. 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.

Castro, A. A., H. Rabitoy, G. C. Claghorn, and T. Garland. 2021. Rapid and longer-term effects of selective breeding for voluntary exercise behavior on skeletal morphology in house mice. Journal of Anatomy 238:720–742.

Castro, A. A., F. A. Karakostis, L. E. Copes, H. E. McClendon, A. P. Trivedi, N. E. Schwartz, and T. Garland Jr. 2021. Effects of selective breeding for voluntary exercise, chronic exercise, and their interaction on muscle attachment site morphology in house mice. Journal of Anatomy joa.13547.

All manuscripts © by John Wiley and Sons.

ABSTRACT OF THE DISSERTATION

Skeletal Evolution and Muscle Physiology in House Mice Selectively Bred for High Levels of Voluntary Wheel-Running Behavior

by

Alberto Arturo Castro

Doctor of Philosophy, Graduate Program in Evolution, Ecology, and Organismal Biology
University of California, Riverside, December 2021
Dr. Theodore Garland, Jr., Chairperson

Experimental evolutionary studies are excellent approaches to study the coadaptation of the musculoskeletal with locomotor behavior in real time. My dissertation examines mice from 4 replicate High Runner (HR) lines bred for voluntary wheel running as young adults and their 4 non-selected Control (C) lines.

Chapter 1 involved the analysis of skeletal data from generation 11 and found that HR mice have evolved larger hip and knee surface areas, which would lower stress (force per unit area), acting on the hindlimb during running. This chapter demonstrates that skeletal dimensions and muscle masses can evolve rapidly in response to directional selection on locomotor behavior.

Chapter 2 examined the rapid and longer-term effects of selective breeding on the skeleton. HR mice reached an apparent selection limit between generations 16-28, running ~3-fold more than C mice. Analysis of bone data from generations 11, 16, 21, 37, 57, and 68 revealed unique results. I 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.

Chapter 3 highlighted aspects of phenotypic plasticity achieved through exercise. Muscle attachment site morphology reflects muscle mass and function and therefore, paleontologists have routinely used characteristics of muscle entheses to infer the past loading history of individual specimens. I used mice from generation 57 that were housed with or without wheels for throughout ontogeny to quantify the genetic differences in muscle attachment site morphology between HR and C mice, as well as plastic changes resulting from chronic exposure to exercise. My results demonstrate that muscle attachment site morphology can be (but is not always) affected by chronic exercise.

Chapter 4 investigated a negative correlation between average running speed and time spent running on wheels that exists among the HR lines. I hypothesized that this trade-off may be related to evolved changes in muscle physiology and used *in-situ* preparations to quantify muscle contractile properties, including speed and endurance. I found that muscle-level speed and endurance do trade-off in these mice, but not in a way that maps to the observed organismal-level speed-endurance trade-off.

Table of Contents

List of Figures

List of Tables

Introduction

Locomotion (active movement through the environment) is vital for animal survival and reproductive success (Dickinson *et al.*, 2000; Cloyed *et al.*, 2021). Animals locomote to flee from predators, forage for food and resources, and when searching for mates. During locomotion, limb bones transmit muscular and propulsive forces, and support the axial skeleton, and respond to loading (Biewener, 1990; Kelly *et al.*, 2006; Middleton *et al.*, 2008b). Therefore, the appendicular skeleton may be expected to show evidence of evolutionary coadaptation with locomotor behavior and ecology.

Perhaps the most emblematic example of coadaptation of locomotor behavior with skeletal morphology involves "cursorial" mammals, or those that run fast and/or long distances (Gregory, 1912; Stein & Casinos, 1997). Some of the most 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, especially for extinct animals (Figure 0.1) (Howell, 1946; Maynard Smith & Savage, 1956; Gambaryan, 1974; Hildebrand, 1974; Coombs, 1978; Garland & Janis, 1993; Carrano, 1999; Lovegrove & Mowoe, 2014). Cursorial mammals have also evolved elongated distal limb bones, elevated foot posture, more proximally located muscles, and lighter and gracile bone elements: these traits are presumed to improve running ability and/or locomotor efficiency.

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). In studies of humans, muscle attachment sites have been used to infer physical activity of ancient populations, with larger attachment sites assumed to indicate larger muscles (Robb, 1998; Drapeau, 2008). 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).

Trade-offs have long been held as cornerstones in evolutionary biology and in many sub-fields of organismal biology (see Schmidt-Nielson, 1984; Garland & Carter, 1994; Ackerly *et al.*, 2000; Martin *et al.*, 2015; Agrawal, 2020). Trade-offs (and constrains) can affect phenotypic evolution in various ways (e.g., see Gustafsson *et al.*, 1995; Roff & Fairbairn, 2007; Garland *et al.*, 2022) and they have been suggested to be key drivers in speciation events (Miles *et al.*, 2018). Multiple types of trade-offs have been recognized (Cohen *et al.*, 2020; Mauro & Ghalambor, 2020; Garland *et al.*, 2022): perhaps the most common type involves allocation constraints. For example, if the energy available to an organism is limited, then spending more on one function (e.g., disease resistance) means less is available for other functions (e.g., reproduction). Another common type of trade-off occurs when features that enhance performance of one task decrease performance of another (Garland et al., 2022), which are termed functional conflicts.

In the locomotor system, the most well-known trade-off that occurs at the organismal performance level is the negative relationship between speed and endurance. For example, among 12 species of closely related lacertid lizards, the residuals for speed and endurance capabilities showed a negative relationship (accounting for variation in body size) (Vanhooydonck *et al.*, 2001). However, that trade-off is not apparent among species of phyrosomatid lizards (Albuquerque *et al.*, 2015; see also Toro *et al.*, 2004; Goodman *et al.*, 2007). Many studies have also tested for trade-offs at the level of variation among individuals. For example, statistically significant trade-offs were detected between speed-related and endurance-related events in a study of 1,369 elite human athletes participating in heptathlon and decathlon events (Careau & Wilson, 2017). In the Hsd:ICR strain of laboratory house mice, residual sprint speed and swimming endurance are uncorrelated phenotypically, but negatively correlated genetically (Dohm *et al.*, 1996). When present, the organismal-level trade-off between speed and endurance is thought to be underpinned by a muscle-level trade off in speed and endurance caused by the stereotyped combination of myosin isoforms and oxidative capacities in different muscle fibers (e.g., see discussion and references in Garland, 1988).

Mammalian muscle fiber types vary along a continuum of contractile and metabolic properties (for a review see Schiaffino & Reggiani, 2011). At one end of the spectrum, Type I fibers contract slowly, use oxidative metabolism, have low power outputs, and are fatigue resistant. At the other end of the spectrum, Type IIb fibers contract rapidly, use glycolysis, have high power outputs, and fatigue rapidly (Komi,

1984; Rome *et al.*, 1988; Esbjörnsson *et al.*, 1993). Muscle fiber type variation has clear links with locomotor diversity. For instance, the predominance of Type I fibers in the forelimb muscles of slow-moving sloths (Spainhower *et al.*, 2018) contrasts with the predominance of Type IIb fibers in the hindlimb muscles of fast-sprinting cheetahs (Williams *et al.*, 1997). The spectrum of locomotor performance variation among lizard species also seems to relate to variation in muscle fiber types (Bonine *et al.*, 2005; Vanhooydonck *et al.*, 2014; Albuquerque *et al.*, 2015; Scales & Butler, 2016)

Phenotypic plasticity is defined as the ability of one genotype to produce more than one phenotype when exposed to different environments. In general, such plasticity occurs in response to environmental effects, which can occur at any point after the formation of the zygote (or even before, affecting gametes) and can encompass numerous abiotic and biotic factors that can influence phenotypic expression (for a review see Garland, & Kelly, 2006). Physical conditioning (e.g., exercise through running, weight-lifting) is a form of phenotypic plasticity that receives much biomedical attention (Fluck, 2006). For example, mice and rats that given chronic wheel access or forced treadmill exercise typically undergo cardiac hypertrophy (e.g., see Swallow *et al.*, 2005).

Bone is a dynamic and metabolically active organ composed of calcium phosphate minerals and type I collagen. Bone remodeling, the actions of osteoclasts and osteoblasts during bone resorption and formation, is essential for mineral and mechanical homeostasis of the skeleton (Doherty *et al.*, 2015). In mice and rats, limb loading experienced during exercise induces bone formation, retards bone loss, and enhances bone structure and ultimately strength (e.g., Plochocki *et al.*, 2008). Some studies of rats

and mice (both with voluntary wheel-running and tower climbing required to obtain food) have shown increases in tibia-fibula and femoral thickness and mass (Newhall *et al.*, 1991; Mori *et al.*, 2003). However, other studies have reported no significant differences in bone morphology and mechanical properties between sedentary and exercised mice (Peacock *et al.*, 2018).

Bone performance during loading is governed by both genetic factors and loading history (Figure 0.2) (Middleton *et al.*, 2008a), but also by age, sex, and epigenetic factors (discussed in Wallace *et al.*, 2012). Thus, controversy arises when paleontologists attempt to infer the past loading history (i.e., activity levels) of individuals from their bone morphology (e.g., Lieberman *et al.*, 2004). In fact, recent studies investigating the skeletal response to loading using genetically distinct mouse strains suggest that bone shape and responsiveness to loading is largely determined by genetic background rather than exercise per se (Wallace *et al.*, 2015; Peacock *et al.*, 2018). These results suggest that differences in bone morphology among fossil remains reflect genetic variation, in addition to the loading history of individuals during their lifetimes (Cowgill, 2009; Wallace *et al.*, 2010; Shaw & Stock, 2013; Peacock *et al.*, 2018).

Broad comparative studies of living species (and/or paleontological comparisons) can reveal correlations between behavior and performance capacities on one hand with morphology on the other. However, such studies have various limitations. For example, unless all species can be raised under common conditions they cannot account for environmental effects (e.g., phenotypic plasticity) or possible genotype-by-environment

interactions (Garland, 2003; Garland *et al.*, 2005). Moreover, broad comparative studies are inherently correlational, not experimental.

Experimental evolutionary approaches can be used to study evolution in action (in real time) by determining the sequence of phenotypic and behavioral changes that occur during adaptation in response to a defined selective regime (Garland, 2001; Garland & Rose, 2009). These approaches are well-suited to study coadaptation of the skeleton with locomotor behavior and body size (Kemp *et al.*, 2005; Marchini *et al.*, 2014), the role of exercise in shaping bone morphology (including genotype*environment interactions), and the biomechanical basis for trade-offs in lower-level traits, such as skeletal muscles.

My research uses two genetically differentiated groups of mice, each with four independent replicate lines. Mice from the 4 High Runner (HR) lines are bred for voluntary wheel running during 6 days of wheel access as young adults and are compared with 4 non-selected Control (C) lines (Swallow *et al.*, 1998). The HR mice have increased their daily wheel-running distance ~3-fold, have increased endurance and maximal aerobic capacity (VO_2 max) during forced treadmill exercise, larger hearts, and larger brains (Garland, 2003; Wallace & Garland, 2016). With the HR mouse model, I can study coadaptation of the musculoskeletal system with locomotor behavior in real time with animals raised under common conditions and over many generations of selective breeding. We can think of the HR mice as "cursorial" from a behavioral perspective, because they run at high speeds and for many hours daily (Garland *et al.*, 2011). Therefore, I am experimentally testing biomechanical theories established by

functional morphologists and paleontologists decades ago (Gregory, 1912; Maynard Smith & Savage, 1956; Stein & Casinos, 1997).

In my first dissertation chapter, I analyzed skeletal data from generation 11 and found that HR mice have evolved larger hip and knee surface areas, which would lower stress (force per unit area), acting on the hindlimb during running. This confirmed my general hypothesis, but not my specific one, and pointed the way to key additional studies.

In my second chapter, I analyzed bone data from HR mice across several generations both before and after selection limits were reached. HR mice reached an apparent selection limit between generations 16-28, running ~3-fold more than C mice (Careau *et al.*, 2013), and I hypothesized that bone dimensions would stop evolving after this point. To test this, I analyzed appendicular bone data from generations 11, 16, 21, 37, 57, and 68. I 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.

In my third chapter, I examined both the genetically evolved differences between the HR lines and their non-selected controls and phenotypic plasticity in response to chronic exercise. Given that muscle attachment site morphology is assumed to reflect muscle mass and function, morphologists have routinely used characteristics of muscle entheses to infer the past loading history (i.e., activity levels) of individual specimens. For example, in studies of humans, muscle attachment site hypertrophy is assumed to indicate increased habitual use of muscles, although little experimental evidence warrants

such practices (e.g., see Zumwalt, 2006). I used a sample of mice from generation 57 that was housed with or without wheels for 12 weeks starting at weaning. I quantified the genetic differences in muscle attachment site morphology between HR and C mice, as well as plastic changes resulting from chronic exposure to exercise, which will be relevant for such controversies. My results demonstrate that muscle attachment site morphology can be (but is not always) affected by chronic exercise.

In my fourth chapter, I investigated a negative correlation between average running speed and time spent running on wheels that exists among the HR lines but not among the C lines (Garland *et al.*, 2011) (Figure 0.3). I hypothesized that the evolution of this trade-off is related to evolved changes in lower-level traits, specifically in skeletal muscles. I used *in situ* preparations to measure muscle fatigue, twitch characteristics, and the force-velocity curve. I did find a trade-off between muscle speed and stamina, but it does not parallel the one observed among the HR lines of mice for wheel-running behavior.

References

Ackerly, D.D., Dudley, S.A., Sultan, S.E., Schmitt, J., Coleman, J.S., Linder, C.R., *et al.* 2000. The evolution of plant ecophysiological traits: Recent advances and future directions. *BioScience* **50**: 979.

Agrawal, A.A. 2020. A scale-dependent framework for trade-offs, syndromes, and specialization in organismal biology. *Ecology* **101**.

Albuquerque, R.L., Bonine, K.E. & Garland, T. 2015. Speed and endurance do not trade off in Phrynosomatid lizards. *Physiol. Biochem. Zool.* **88**: 634–647.

Biewener, A.A. 1990. Biomechanics of mammalian terrestrial locomotion. *Sci. New Ser.* **250**: 1097–1103.

Bonine, K.E., Gleeson, T.T. & Garland Jr., T. 2005. Muscle fiber-type variation in lizards (Squamata) and phylogenetic reconstruction of hypothesized ancestral states. *J. Exp. Biol.* **208**: 4529–4547.

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

Careau, V. & Wilson, R.S. 2017. Performance trade-offs and ageing in the 'world's greatest athletes.' *Proc. R. Soc. B Biol. Sci.* **284**: 20171048.

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. *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. *J. Zool.* **247**: 29–42.

Cloyed, C.S., Grady, J.M., Savage, V.M., Uyeda, J.C. & Dell, A.I. 2021. The allometry of locomotion. *Ecology* **102**.

Cohen, A.A., Coste, C.F.D., Li, X., Bourg, S. & Pavard, S. 2020. Are trade-offs really the key drivers of ageing and life span? *Funct. Ecol.* **34**: 153–166.

Coombs, W.P. 1978. Theoretical aspects of cursorial adaptations in dinosaurs. *Q. Rev. Biol.* **53**: 393–418.

Cowgill, L.W. 2009. The ontogeny of Holocene and Late Pleistocene human postcranial strength. *Am. J. Phys. Anthropol.* NA-NA.

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.

Doherty, A.H., Ghalambor, C.K. & Donahue, S.W. 2015. Evolutionary physiology of bone: Bone metabolism in changing environments. *Physiology* **30**: 17–29.

Dohm, M.R., Hayes, J.P. & Garland Jr., T. 1996. Quantitative genetics of sprint running speed and swimming endurance in laboratory house mice *Mus domesticus*. *Evolution* **42**: 355–350.

Drapeau, M.S.M. 2008. Enthesis bilateral asymmetry in humans and African apes. *HOMO* **59**: 93–109.

Esbjörnsson, M., Sylvén, C., Holm, I. & Jansson, E. 1993. Fast twitch fibres may predict anaerobic performance in both females and males. *Int. J. Sports Med.* **14**: 257–263.

Fluck, M. 2006. Functional, structural and molecular plasticity of mammalian skeletal muscle in response to exercise stimuli. *J. Exp. Biol.* **209**: 2239–2248.

Gambaryan, P.P. 1974. *How mammals run: Anatomical adaptations*. John WIley and Sons, New York.

Garland Jr., T. 1988. Genetic basis of activity metabolism. I. Inheritance of speed, stamina, and antipedator displays in the Garter Snake *Thampnophis Sirtalis*. *Evolution* **42**.

Garland Jr., T. 2001. Phylogenetic comparison and artificial selection. In: *Hypoxia* (R. C. Roach, P. D. Wagner, & P. H. Hackett, eds), pp. 107–132. Springer US, Boston, MA.

Garland Jr., T. 2003. Selection experiments: An under-utilized tool in biomechanics and organismal biology. In: *Vertebrate Biomechanics and Evolution*, p. 35. Bios Scientific Publisher, Oxford.

Garland Jr., T., Bennett, A.F. & Rezende, E.L. 2005. Phylogenetic approaches in comparative physiology. *J. Exp. Biol.* **208**: 3015–3035.

Garland Jr., T. & Carter, P.A. 1994. Evolutionary Physiology. *Annu. Rev. Physiol.* **56**: 579–621.

Garland Jr., T., Downs, C. & Ives, A.R. 2022. Perspective: Trade-offs (and constraints) in organismal biology. *Physiol. Biochem. Zool.* **95**: 717897.

Garland Jr., T. & Janis, C.M. 1993. Does metatarsal/femur ratio predict maximal running speed in cursorial mammals? *J. Zool.* **229**: 133–151.

Garland Jr., T. & Kelly, S.A. 2006. Phenotypic plasticity and experimental evolution. *J. Exp. Biol.* **209**: 2344–2361.

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. *Proc. R. Soc. B Biol. Sci.* **278**: 574–581.

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

Goodman, B.A., Krockenberger, A.K. & Schwarzkopf, L. 2007. Master of them all: Performance specialization does not result in trade-offs in tropical lizards. *Evol. Ecol. Res.* **9**: 527–546.

Gregory, W.K. 1912. Notes on the principles of quadrupedal locomotion and on the mechanism of the limbs in hoofed animals. *Ann N Acad Sci* **22**: 267–294.

Gustafsson, L., Qvarnström, A. & Sheldon, B.C. 1995. Trade-offs between life-history traits and a secondary sexual character in male collared flycatchers. *Nature* **375**: 311–313.

Hildebrand, M. 1974. *Analysis of vertebrate structure*. John WIley and Sons, New York.

Howell, A.B. 1946. Speed in animals. *J. Gen. Psychol.* **34**: 107–110.

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. *J. Morphol.* **267**: 360–374.

Kemp, T.J., Bachus, K.N., Nairn, J.A. & Carrier, D.R. 2005. Functional trade-offs in the limb bones of dogs selected for running versus fighting. *J. Exp. Biol.* **208**: 3475–3482.

Komi, P.V. 1984. Physiological and biomechanical correlates of muscle function: effects of muscle structure and stretch-shortening cycle on force and speed. *Exerc. Sport Sci. Rev.* **12**: 81–121.

Lieberman, D.E., Polk, J.D. & Demes, B. 2004. Predicting long bone loading from cross-sectional geometry. *Am. J. Phys. Anthropol.* **123**: 156–171.

Lovegrove, B.G. & Mowoe, M.O. 2014. The evolution of micro-cursoriality in mammals. *J. Exp. Biol.* **217**: 1316–1325.

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 Evol. Biol.* **14**: 15.

Martin, L.B., Ghalambor, C.K. & Woods, H.A. 2015. *Integrative organismal biology*. Wiley Blackwell, Hoboken, New Jersey.

Mauro, A.A. & Ghalambor, C.K. 2020. Trade-offs, pleiotropy, and shared molecular pathways: A unified view of constraints on adaptation. *Integr. Comp. Biol.* **60**: 332–347.

Maynard Smith, J. & Savage, R.J.G. 1956. Some locomotory adaptations in mammals. *J. Linn. Soc. Lond. Zool.* **42**: 603–622.

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. *Integr. Comp. Biol.* **48**: 394–410.

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.

Miles, M.C., Goller, F. & Fuxjager, M.J. 2018. Physiological constraint on acrobatic courtship behavior underlies rapid sympatric speciation in bearded manakins. *eLife* **7**: e40630.

Mori, T., Okimoto, N., Sakai, A., Okazaki, Y., Nakura, N., Notomi, T., *et al.* 2003. Climbing exercise increases bone mass and trabecular bone turnover through transient regulation of marrow osteogenic and osteoclastogenic potentials in mice. *J. Bone Miner. Res.* **18**: 2002–2009.

Newhall, K.M., Rodnick, K.J., van der Meulen, M.C., Carter, D.R. & Marcus, R. 1991. Effects of voluntary exercise on bone mineral content in rats. *J. Bone Miner. Res.* **6**: 289–296.

Peacock, S.J., Coats, B.R., Kirkland, J.K., Tanner, C.A., Garland Jr., T. & Middleton, K.M. 2018. Predicting the bending properties of long bones: Insights from an experimental mouse model. *Am. J. Phys. Anthropol.* **165**: 457–470.

Plochocki, J.H., Rivera, J.P., Zhang, C. & Ebba, S.A. 2008. Bone modeling response to voluntary exercise in the hindlimb of mice. *J. Morphol.* **269**: 313–318.

Robb, J.E. 1998. The interpretation of skeletal muscle sites: a statistical approach. *Int. J. Osteoarchaeol.* **8**: 363–377.

Roff, D.A. & Fairbairn, D.J. 2007. The evolution of trade-offs: Where are we? *J. Evol. Biol.* **20**: 433–447.

Rome, L.C., Funke, R.P., Alexander, R.M., Lutz, G., Aldridge, H., Scott, F., *et al.* 1988. Why animals have different muscle fibre types. *Nature* **335**: 824–827.

Scales, J.A. & Butler, M.A. 2016. Adaptive evolution in locomotor performance: How selective pressures and functional relationships produce diversity. *Evolution* **70**: 48–61.

Schiaffino, S. & Reggiani, C. 2011. Fiber types in mammalian skeletal muscles. *Physiol. Rev.* **91**: 1447–1531.

Schmidt-Nielson, K. 1984. *Scaling: why is animal size so important?* Cambridge University Press, Cambridge.

Shaw, C.N. & Stock, J.T. 2013. Extreme mobility in the Late Pleistocene? Comparing limb biomechanics among fossil Homo, varsity athletes and Holocene foragers. *J. Hum. Evol.* **64**: 242–249.

Spainhower, K.B., Cliffe, R.N., Metz, A.K., Barkett, E.M., Kiraly, P.M., Thomas, D.R., *et al.* 2018. Cheap labor: Myosin fiber type expression and enzyme activity in the forelimb musculature of sloths (*Pilosa: Xenarthra*). *J. Appl. Physiol.* **125**: 799–811.

Stein, B.R. & Casinos, A. 1997. What is a cursorial mammal? *J. Zool.* **242**: 185–192.

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

Swallow, J.G., Rhodes, J.S. & Garland Jr, T. 2005. Phenotypic and evolutionary plasticity of organ masses in response to voluntary exercise in house mice1. *Integr. Comp. Biol.* **45**: 426–437.

Toro, E., Herrel, A. & Irschick, D. 2004. The evolution of jumping performance in Caribbean *Anolis* lizards: Solutions to biomechanical trade-offs. *Am. Nat.* **163**: 844–856.

Vanhooydonck, B., James, R.S., Tallis, J., Aerts, P., Tadic, Z., Tolley, K.A., *et al.* 2014. Is the whole more than the sum of its parts? Evolutionary trade-offs between burst and sustained locomotion in lacertid lizards. *Proc. R. Soc. B Biol. Sci.* **281**: 20132677.

Vanhooydonck, B., Van Damme, R. & Aerts, P. 2001. Speed and stamina trade-off in lacertid lizards. *Evolution* **55**: 1040.

Wallace, I.J. & Garland Jr., T. 2016. Mobility as an emergent property of biological organization: Insights from experimental evolution: Mobility and biological organization. *Evol. Anthropol. Issues News Rev.* **25**: 98–104.

Wallace, I.J., Judex, S. & Demes, B. 2015. Effects of load-bearing exercise on skeletal structure and mechanics differ between outbred populations of mice. *Bone* **72**: 1–8.

Wallace, I.J., Middleton, K.M., Lublinsky, S., Kelly, S.A., Judex, S., Garland Jr., T., *et al.* 2010. Functional significance of genetic variation underlying limb bone diaphyseal structure. *Am. J. Phys. Anthropol.* **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. *Am. J. Phys. Anthropol.* **148**: 24–35.

Williams, T.M., Dobson, G.P., Mathieu-Costello, O., Morsbach, D., Worley, M.B. & Phillips, J.A. 1997. Skeletal muscle histology and biochemistry of an elite sprinter, the African cheetah. *J. Comp. Physiol. [B]* **167**: 527–535.

Zumwalt, A. 2006. The effect of endurance exercise on the morphology of muscle attachment sites. *J. Exp. Biol.* **209**: 444–454.

Figure 0.1. MT/F Ratio. Skeleton of a badger and a hyena which are both Carnivora but have distinct locomotor behaviors. Morphologists have routinely used the MT/F ratio as an index of cursoriality which quantifies the degree to which distal elements of the hindlimb are elongated relative to the more proximal ones. Cursorial species have higher MT/F ratios when compared with non-cursorial species which indicates increased speed, stamina, agility, or efficiency.

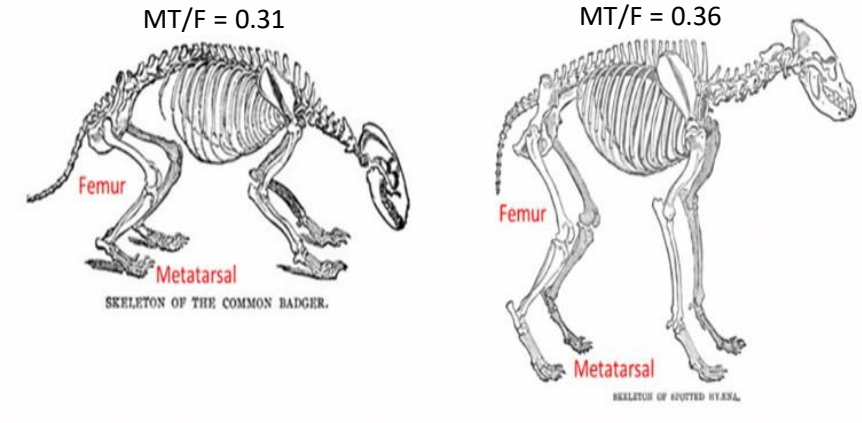


Figure 0.2. Bone Performance. Bone performance during loading is governed by bone structure, shape, and form. Bone morphology is determined by both genetics and past loading history (e.g., exercise). Selection can act on the motivation to do exercise (gray line) and the physiological response to exercise (dotted line).

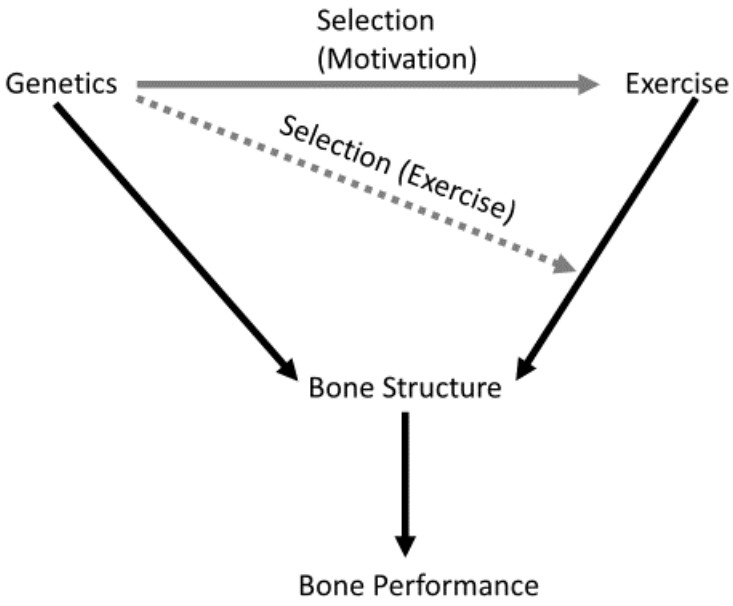
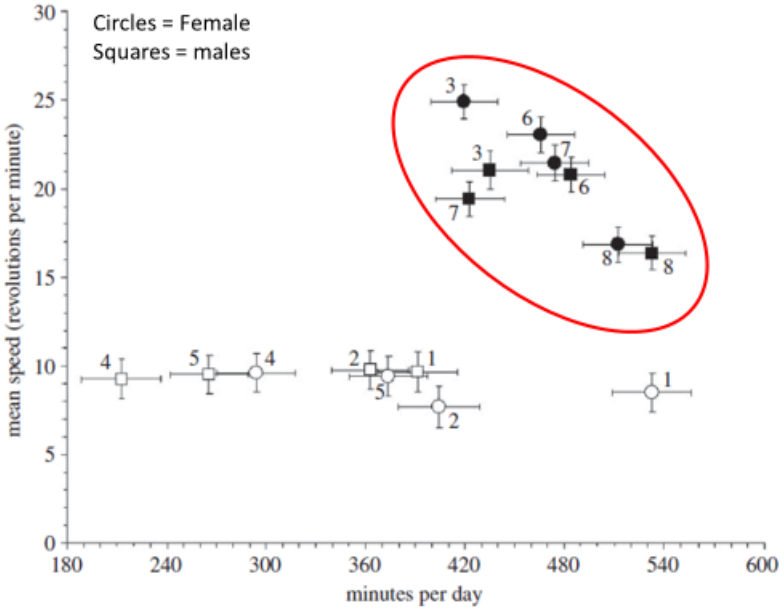


Figure 0.3. Trade-offs in HR mice. Evolution of a trade-off between average running speed and time spent running on wheels among HR mice.



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

Acknowledgements

We thank Patricia A. Freeman for obtaining measurements and for comments on the manuscript (Dr. Freeman declined to be a coauthor on this manuscript). We also thank an anonymous reviewer, Campbell Rolian, Nicole 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.

Author Contributions

TG and PAF collected data. AAC and TG analyzed the data. AAC and TG wrote the manuscript. All authors edited the manuscript.

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.

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 (Kelly *et al.*, 2006; Middleton *et al.*, 2008a; Gosnell *et al.*, 2011). 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 (Howell, 1944; Smith & Savage, 1956; Gambaryan, 1974; Hildebrand, 1974; Coombs Jr, 1978; Garland & Janis, 1993; Carrano, 1999; Lovegrove & Mowoe, 2014). 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 (Van Valkenburgh, 1987; Carrano, 1999; Samuels *et al.*, 2013).

Some of the putative indicators of cursoriality seem to be associated with body size (Carrano, 1999; Middleton *et al.*, 2008a) and/or phylogeny (Garland & 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; Young *et al.*, 2014). Furthermore, "cursorial" lagomorphs have lower limb joint mechanical advantages, which may allow increased limb output velocity and faster cycling of limbs (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 (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 & 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 (Valen, 1962; Pelabon *et al.*, 2006). Directional asymmetry (DA) is the consistent deviation of bilateral structures such that one side is larger than the other (Carter *et al.*, 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 *et al.*, 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 *et al.*, 2009; but see Warrener *et al.*, 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 & 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 *et al.*, 1998). Daily wheel-running distances of the HR lines reached ~75% greater than the C lines by generation 10, mainly by increased average speed (Swallow *et al.*, 1998; Koteja *et al.*, 1999). 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 *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 & 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 & Janis, 1993; Garland & Freeman, 2005). 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) (Van Valkenburgh, 1987; Samuels & Van Valkenburgh, 2008; Samuels *et al.*, 2013).

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 *et al.*, 2011b), 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 *et al.*, 2011a; Keeney *et al.*, 2012).

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. All experimental procedures were approved by the University of Wisconsin-Madison institutional animal care and use committee.

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 & 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 & 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 like 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 & Janis, 1993; Garland & Freeman, 2005; Samuels & Van Valkenburgh, 2008; Samuels *et al.,* 2013), tibia/femur ratio (T/F: also known as crural index), which indicates relative proportions of proximal and distal elements of the hind limb (Vanhooydonck & Van Damme, 2001; Samuels & Van Valkenburgh, 2008; Biancardi & Minetti, 2012; Samuels *et al.*, 2013), 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 *et al.*, 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 & Van Valkenburgh, 2008; Samuels *et al.*, 2013). 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 & 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 (Sarringhaus *et al.*, 2005; Auerbach & Ruff, 2006). 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 (Swallow *et al.*, 1999; Houle-Leroy *et al.*, 2000, 2003; Garland & Freeman, 2005). 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 *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 *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. 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 ≤ 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 1.1, 1.2, and 1.3 (S1.1 represents least square means and standard error for Tables 1.1,1. 2, and 1.3. 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 1.2, but not discussed.

Results

Tables 1.1-1.3 present significance levels from results of ANOVAs and ANOVAs (using body mass as a covariate), whereas S1.1 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 1.1). 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 S1.1 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 1.1).

Standard Mammalian Measurements

Linetype differences were never significant for ear, tail or hindfoot lengths, regardless of the body-size covariate used (Table 1.1). 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 1.1).

In the models testing for interactions with body mass, the body mass * linetype * sex interaction was marginally significant for hindfoot length ($p = 0.0717$: S1.2). 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 1.1).

The interaction model for quadriceps showed a strong body mass * linetype interaction ($p = 0.0066$: S1.2), 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 ($p = 0.0178$: S1.2), 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 1.2). Mice from HR lines had increased anterior-posterior depth of the femoral head (Fig. 1.1: $p = 0.0366$ with body mass as covariate; $p = 0.0640$ with body length), increased femur distal width (Fig. 1.2: $p = 0.0176$ 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 1.2). Femoral distal width (Fig. 1.2 all covariates), femoral width at 3rd trochanter (Fig. 1.3, all covariates), femoral least width (all covariates), and the tibia-fibula least width (all covariates) all had reduced medial-lateral width measurements. Femoral (Fig. 1.4) and tibia-fibula mass were significantly reduced in mini-muscle mice, regardless of covariate used (see Table 1.2). Mini-muscle mice also had significantly longer tibia-fibula lengths with body mass as a covariate (Fig. 1.5). 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 1.2). 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. 1.3, all covariates), and femoral least width (all covariates: Table 1.2). 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. 1.4) 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 1.2). 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 1.2).

Several bone dimensions showed significant interactions with body mass, and full analyses are presented in S1.2. 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 ($P = 0.0838$, S1.2). 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 ($p = 0.0242$; S1.2). 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 ($p = 0.0663$; S1.2). 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 ($p=0.0062$; S1.2). 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 1.3). 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 (3^{rd}/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 1.3). Finally, males had significanttly decreased [FM/(FL * FLW^3)], suggesting that femurs were less dense than for females (Table 1.3).

Asymmetry

Directional asymmetry (Appendix 1.1) was significantly lower in HR mice for total leg length (2-tailed $p = 0.0217$), (see also Garland & Freeman, 2005) and for femur length ($p = 0.0311$), but not for tibia-fibula or metatarsal length (Appendix 1.1). 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 was not significant when analysis was split by sex (S1.2). HR males tended to have reduced directional asymmetry for tibia-fibula least width ($p = 0.0871$: S1.2), when compared with C lines (sex * linetype interaction $p = 0.0923$). This was also observed in FA/DA tibia-fibula least width (sex * linetype interaction $p= 0.0649$) and analyses split by sex for HR males ($p = 0.0396$: S1.2).

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 can be interpreted as 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 *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 *et al.,* 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 *et al.*, 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 *et al.*, 2012), which may increase bone strength. We did not find this (Table 1.1 and 1.3), 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 *et al.*, 2008). Moreover, at least in later generations, HR mice are more active that C mice in home cages when housed without wheels (Malisch *et al.*, 2009; Copes *et al.*, 2015). 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 & Freeman, 2005; Kelly *et al.*, 2006; Middleton *et al.*, 2008b; a, 2010; Young *et al.*, 2009; Wallace *et al.*, 2010, 2012; Schutz *et al.*, 2014). 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 *et al.*, 1999; Kodama *et al.*, 1999).

Sex Differences

Sex hormones, growth hormones, mechanosensation, and insulin-like growth factors during puberty influence skeletal sexual dimorphism (Callewaert *et al.*, 2010; Copes *et al.*, 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 *et al.*, 2011a). 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. 1.6), 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.1.6). In contrast, male HR mice have evolved altered tibia-fibula proximal depths that varied depending on body mass (S1.2). 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 (S1.2). 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 *et al.*, 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 1.2), 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 *et al.*, 2009). Mini-muscle mice also have thinner hindlimb bones for many measurements which may reduce rotational inertia and be reflective of reduced bone mass (Young *et al.*, 2014); see Table 1.2. Like cursorial lagomorphs (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; 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 quadtratus 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 *et al.*, 2008b). 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 μCT 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 (Schutz *et al.*, 2014; but see Copes *et al.*, 2017). 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).

References

Auerbach, B.M. & Ruff, C.B. 2006. Limb bone bilateral asymmetry: variability and commonality among modern humans. *J. Hum. Evol.* **50**: 203–218.

Biancardi, C.M. & Minetti, A.E. 2012. Biomechanical determinants of transverse and rotary gallop in cursorial mammals. *J. Exp. Biol.* **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. *J. Bone Miner. Res.* **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. *J. Endocrinol.* **207**: 127–134.

Carrano, M.T. 1999. What, if anything, is a cursor? Categories versus continua for determining locomotor habit in mammals and dinosaurs. *J. Zool.* **247**: 29–42.

Carter, A.J.R., Osborne, E. & Houle, D. 2009. Heritability of Directional Asymmetry in Drosophila melanogaster. *Int. J. Evol. Biol.* **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. *Anat. Rec.* **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. *Physiol. Biochem. Zool.* **90**: 533–545.

Coombs Jr, W.P. 1978. Theoretical aspects of cursorial adaptations in dinosaurs. *Q. Rev. Biol.* **53**: 393–418.

Copes, L.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. *Physiol. Behav.* **149**: 86–94.

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. *Am. J. Phys. Anthropol.* **In revision**.

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. *J. Exp. Biol.* **212**: 2612–2618.

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

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? *J. Zool.* **229**: 133–151.

Garland Jr., T., Kelly, S.A., Malisch, J.L., Kolb, E.M., Hannon, R.M., Keeney, B.K., *et al.* 2011a. How to run far: multiple solutions and sex-specific responses to selective breeding for high voluntary activity levels. *Proc. R. Soc. B Biol. Sci.* **278**: 574–581.

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.

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

Garland Jr., T., Schutz, H., Chappell, M.A., Keeney, B.K., Meek, T.H., Copes, L.E., *et al.* 2011b. The biological control of voluntary exercise, spontaneous physical activity and daily energy expenditure in relation to obesity: human and rodent perspectives. *J. Exp. Biol.* **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. *J. Zool.* **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. *Am. J. Phys. Anthropol.* **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. *J. Exp. Biol.* **214**: 3455–3466.

Gregory, W.K. 1912. Notes on the principles of quadrupedal locomotion and on the mechanism of the limbs in hoofed animals. *Ann. N. Y. Acad. Sci.* **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. *Comp. Biochem. Physiol. B Biochem. Mol. Biol.* **144**: 271–282.

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

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

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*. *J. Appl. Physiol.* **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. *Am. J. Physiol. - Regul. Integr. Comp. Physiol.* **284**: R433–R443.

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

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. *J. Bone Miner. Res.* **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. *Pharmacol. Biochem. Behav.* **101**: 528–537.

Kelly, S.A., Bell, T.A., Selitsky, S.R., Buus, R.J., Hua, K., Weinstock, G.M., *et al.* 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. *J. Morphol.* **267**: 360–374.

Kim, B.-T., Mosekilde, L., Duan, Y., Zhang, X.-Z., Tornvig, L., Thomsen, J.S., *et al.* 2003. The structural and hormonal basis of sex differences in peak appendicular bone strength in rats. *J. Bone Miner. Res.* **18**: 150–155.

Kodama, Y., Dimai, H.P., Wergedal, J., Sheng, M., Malpe, R., Kutilek, S., *et al.* 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. *Anim. Behav.* **58**: 1307–1318.

Lindenfors, P., Gittleman, J.L. & Jones, K.E. 2007. Sexual size dimorphism in mammals. *Sex Size Gend. Roles Evol. Stud. Sex. Size Dimorphism* 16–26.

Lovegrove, B.G. & Mowoe, M.O. 2014. The evolution of micro-cursoriality in mammals. *J. Exp. Biol.* **217**: 1316–1325.

Malisch, J.L., Breuner, C.W., Kolb, E.M., Wada, H., Hannon, R.M., Chappell, M.A., *et al.* 2009. Behavioral Despair and Home-Cage Activity in Mice with Chronically Elevated Baseline Corticosterone Concentrations. *Behav. Genet.* **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., *et al.* 2014. Impacts of genetic correlation on the independent evolution of body mass and skeletal size in mammals. *BMC Evol. Biol.* **14**: 258.

Martín, J. & López, P. 2001. Hindlimb asymmetry reduces escape performance in the lizard *Psammodromus algirus*. *Physiol. Biochem. Zool.* **74**: 619–624.

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. *J. Anat.* **216**: 121–131.

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. *Integr. Comp. Biol.* **48**: 394–410.

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.

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., *et al.* 2004. Males have larger skeletal size and bone mass than females, despite comparable body size. *J. Bone Miner. Res.* **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. *Biol. J. Linn. Soc.*

Palmer, A.R. & Strobeck, C. 2003. CH 17. Fluctuating asymmetry analyses revisited. *Dev. Instab. Causes Consequences Oxf. Univ. Press Oxf.* 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*. *J. Evol. Biol.* **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. *J. Appl. Physiol.* **113**: 627–635.

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

Samuels, J.X. & Van Valkenburgh, B. 2008. Skeletal indicators of locomotor adaptations in living and extinct rodents. *J. Morphol.* **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)*. *Am. J. Phys. Anthropol.* **128**: 840–845.

Schutz, H., Donovan, E.R. & Hayes, J.P. 2009. Effects of parity on pelvic size and shape dimorphism in *Mus*. *J. Morphol.* **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. *Zool. J. Linn. Soc.* **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? *J. Zool.* **242**: 185–192.

Steudel, K. & Beattie, J. 1993. Scaling of cursoriality in mammals. *J. Morphol.* **217**: 55–63.

Swallow, J.G., Carter, P.A. & Garland Jr., T. 1998. Artificial selection for increased wheel-running behavior in house mice. *Behav. Genet.* **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. *J. Exp. Biol.* **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. *Mech. Dev.* **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. *J. Vertebr. Paleontol.* **7**: 162–182.

Vanhooydonck, B. & Van Damme, R. 2001. Evolutionary trade-offs in locomotor capacities in lacertid lizards: are splendid sprinters clumsy climbers? *J. Evol. Biol.* **14**: 46–54.

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. *Am. J. Phys. Anthropol.* **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. *Am. J. Phys. Anthropol.* **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. *J. Anat.* **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. *Evol. Biol.* **36**: 88–99.

Table 1.1 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 (mm)							
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 1.2. 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 1.3. 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 * $TFLW^2$)	138	0.1804-	0.7141-	0.4622	**<.0001+**

Figure 1.1. Femoral Head. 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 1.2 for statistical analyses).

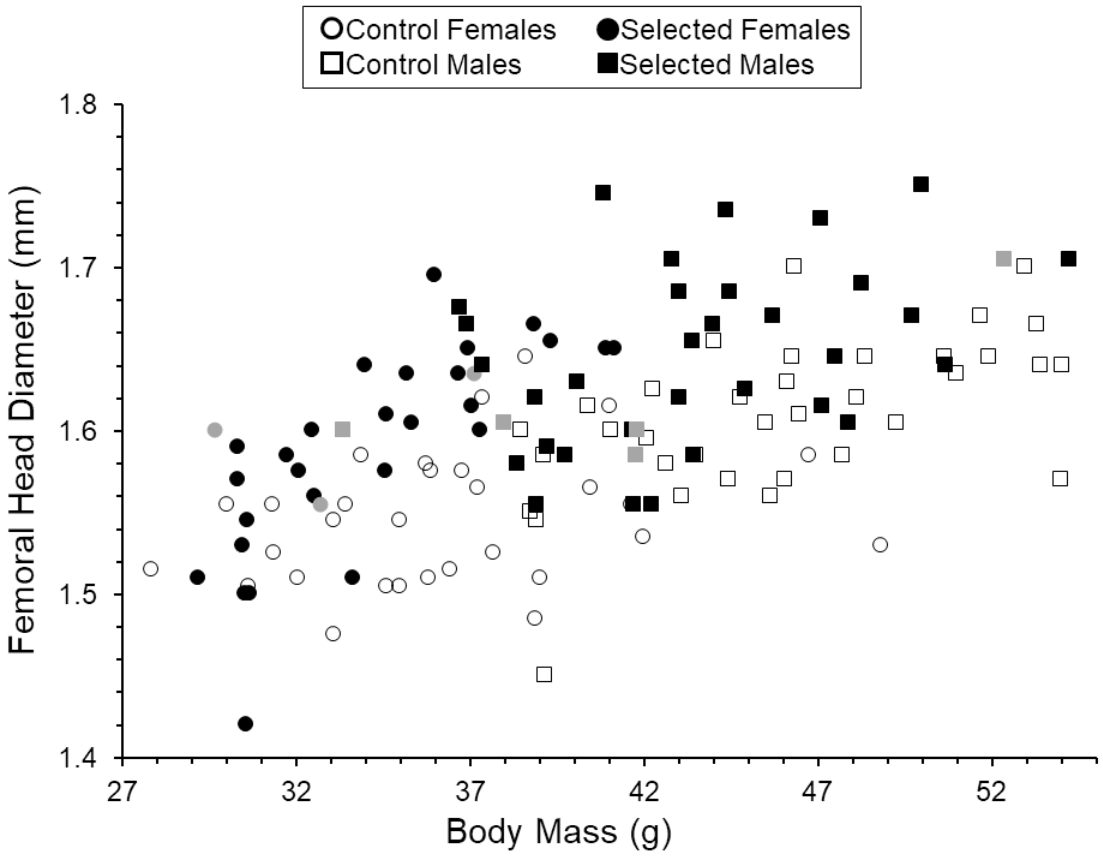


Figure 1.2. Femoral Distal Width. Mean femoral distal width in relation to body mass. Mice from the selectively bred High Runner lines had significantly broader femoral distal widths for a given body mass, sugessting increased muscle attachment area and increased articular surface area around the knee joint. Mini-muscle mice had reduced femoral distal widths (see Table 1.2 for statistical analyses).

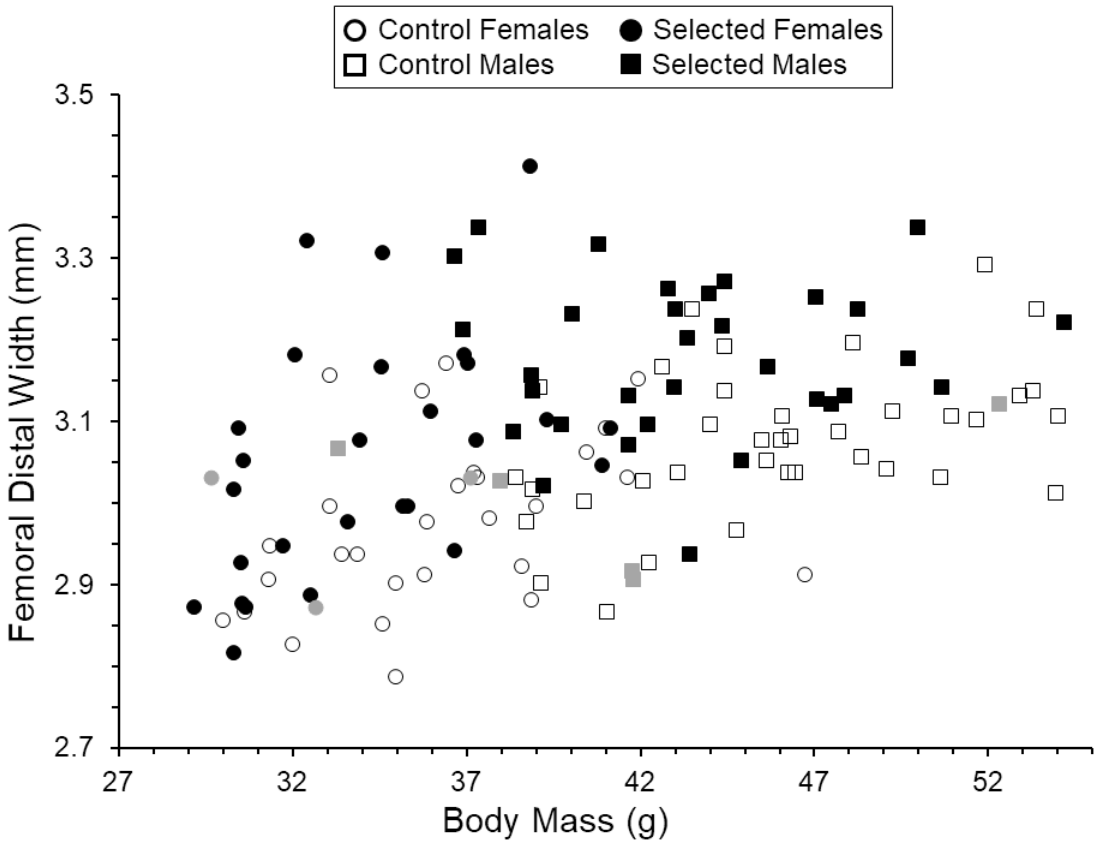


Figure 1.3. Femoral Width 3rd Trochanter. Mean femoral width at 3rd trochanter muscle scar in relation to body mass. Males had significantly thicker femoral width at 3rd trochanter muscle scar for a given body mass, suggesting increased robustness. In addition, mini-muscle mice had reduced femoral width measurements (see Table 1.2 for statistical analyses).

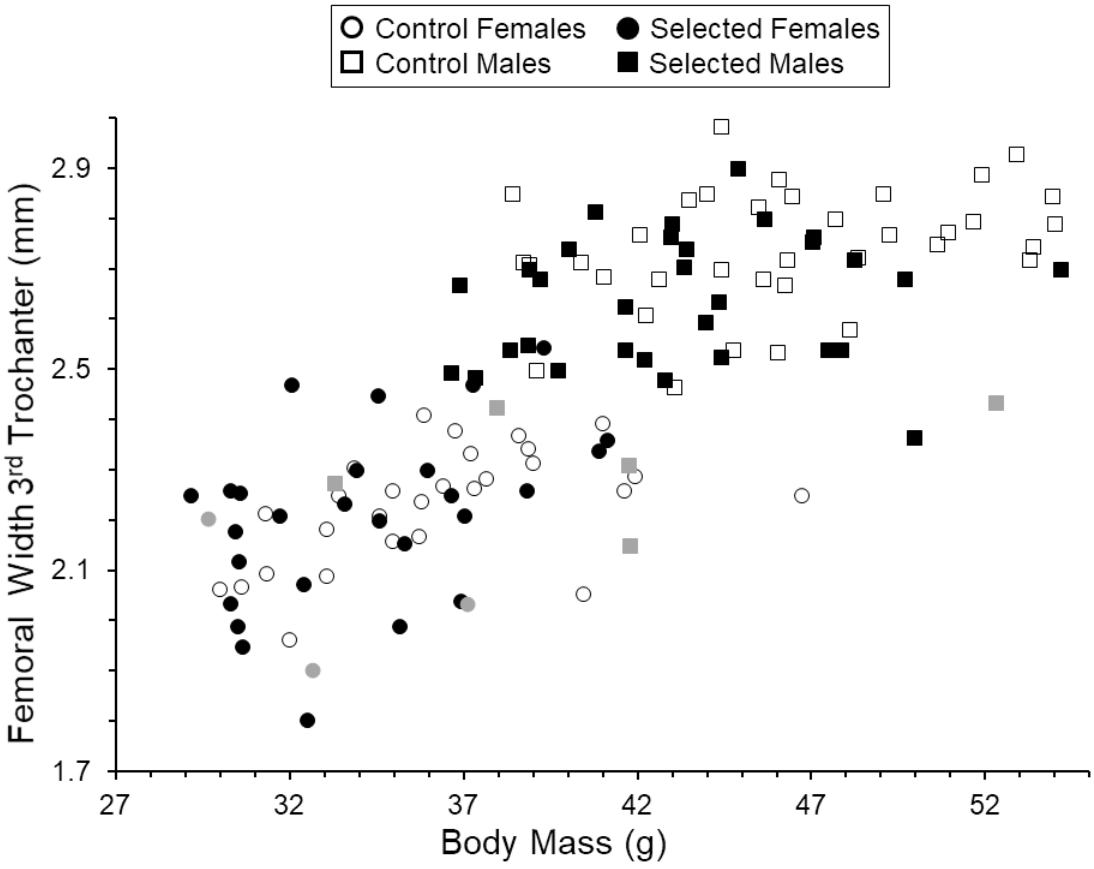


Figure 1.4. Femoral Mass. Mean femoral mass in relation to body mass. Males had significantly reduced femur mass for a given body mass, and mini-muscle mice had reduced femoral masses (see Table 1.2 for statistical analyses).

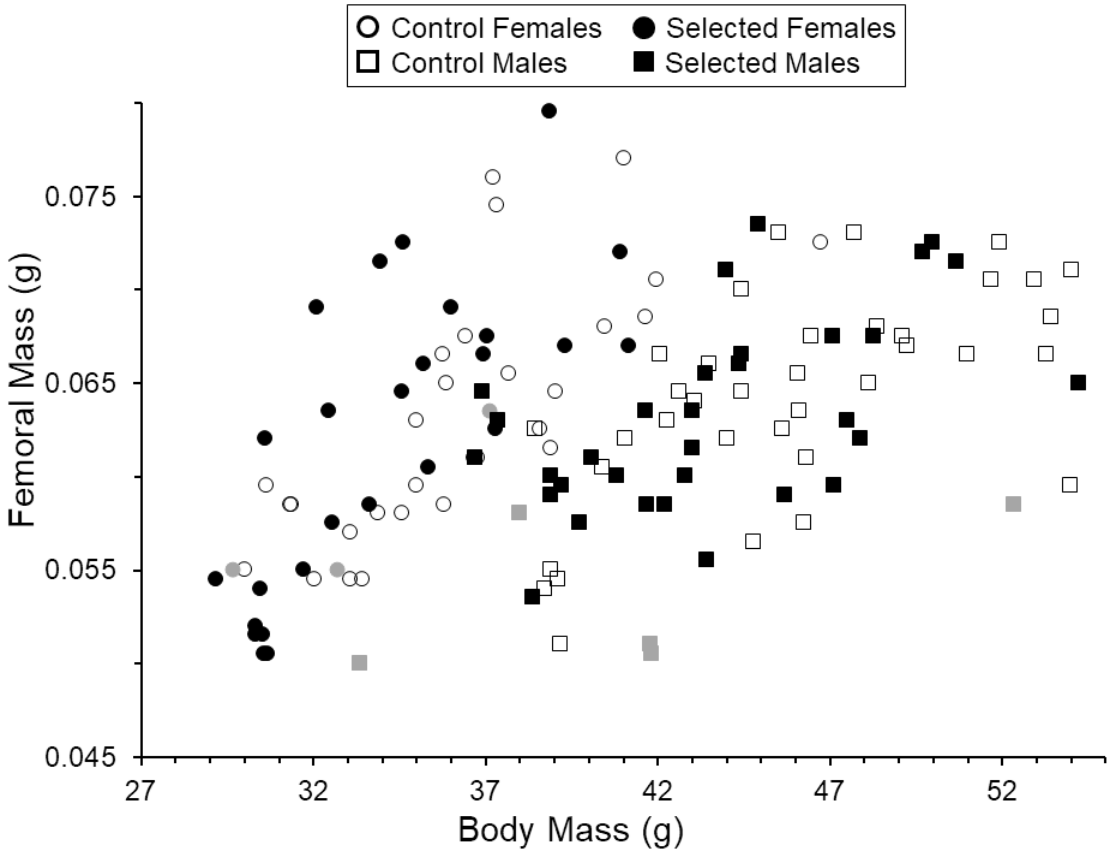


Figure 1.5. Tibia-fibula Length. Mean tibia-fibula length in relation to body mass. Males had significantly shorter tibia-fibula lengths for a given body mass, and mini-muscle mice had increased tibia-fibula lengths (see Table 1.2 for statistical analyses).

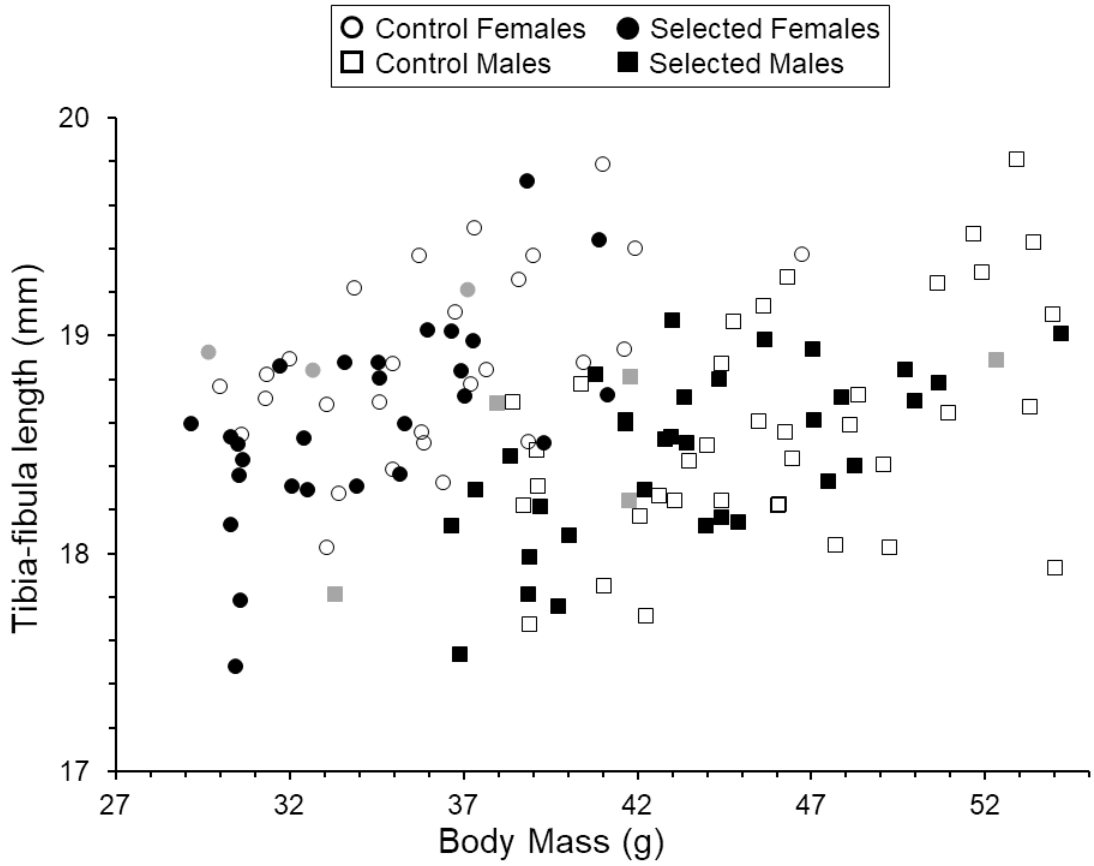
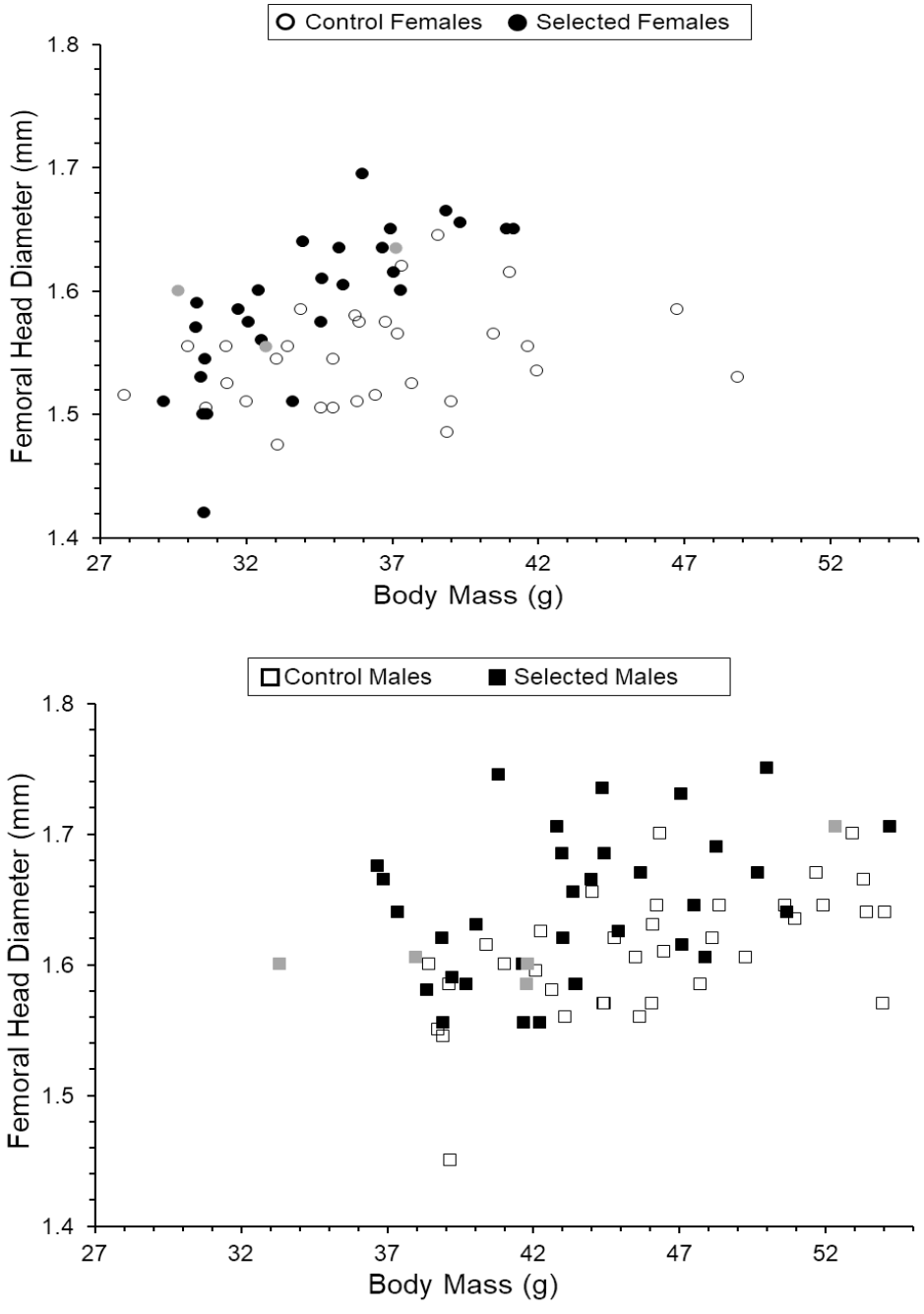


Figure 1.6. Femoral Head: Sex-specific. Mean anterior-posterior depth of femoral head in relation to body mass, separately by sex to illustrate the significant interaction between sex and body mass for both traits (statistical analyses are in S1.2). Note that these same data are shown in Figure 1.1.



Appendices

Appendix 1.1. Asymmetry Analyses. 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. 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
Degrees of Freedom		**1, 6**	1, 6	1, 6	1, ~19
DA Leg	134	**0.0217-**	0.4412+	0.3885	0.2340-
DA Femur L	137	**0.0311-**	0.9000-	0.1286	0.4762-
DA Tibia-fibula L	137	0.8981+	0.8958+	0.6055	0.4298-
DA 3rd Metatarsal	135	0.1398+	**0.0361-**	0.2592	0.7564+
DA Femoral Head	137	0.6190-	0.1875+	0.6215	0.8323-
DA Femoral Least Width	139	0.8551-	**0.0454-**	**0.0567**	0.8748-
DA Femoral Distal Width	135	0.1530-	0.7582+	0.5918	0.9447-
DA Tibia-fibula Least Width	136	0.4843-	0.6799-	**0.0923**	0.4814+
DA Tibia-fibula Distal Width	136	0.6963-	0.5590+	0.6426	**0.0236+**
FA Leg	134	0.0865-	0.2770-	0.7845	0.2586+
FA Femur L	138	0.0993-	0.2472-	0.9080	0.2131+
FA Tibia-fibula L	139	0.2538-	0.8835-	0.8816	0.8535+
FA 3rd Metatarsal	135	0.7827-	0.0806+	0.1785	0.4161-
FA Femoral Head	138	0.8409+	0.7462+	0.7363	0.1165+
FA Femoral Least Width	139	0.9712+	0.2062+	0.3773	0.4716+
FA Femoral Distal Width	135	0.3993-	0.6279+	0.9473	0.7901-
FA Tibia-fibula Least Width	136	0.3849+	0.2557+	0.1546	0.1934-
FA Tibia-fibula Distal Width	136	**0.0108-**	0.8846-	0.8235	0.0666+
FA/DA Leg	133	0.1230-	0.1783+	0.3295	0.2375-
FA/DA Femur L	135	0.0510-	0.4426+	0.2369	0.3549-
FA/DA Tibia-fibula L	133	0.5766-	0.2693-	0.5146	0.7018-
FA/DA 3rd Metatarsal	127	0.6255+	0.1990-	0.4861	0.4614+
FA/DA Femoral Head	128	0.6587-	0.1250+	0.5736	0.8764-
FA/DA Femoral Least Width	132	0.7714-	0.0743-	**0.0997**	0.9553+

FA/DA Femoral Distal Width	135	0.6163-	0.2418-	0.9592	0.2141-
FA/DA Tibia-fibula Least Width	133	0.2623-	0.9409-	**0.0649**	0.3467+
FA/DA Tibia-fibula Distal Width	126	0.9550+	0.5792+	0.7144	0.3473+

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, and Theodore Garland, Jr.*

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

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.

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 17-27, 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.

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 (Van Valkenburgh, 1987; Garland & Janis, 1993). These types of associations are a cornerstone of ecomorphology (Van Der Klaauw, 1948; Samuels & Van Valkenburgh, 2008; Samuels *et al.*, 2013; e.g., Jones, 2016).

Perhaps the most emblematic example of coadaptation (Huey & Bennett, 1987; Angilletta Jr. *et al.*, 2006; Foster *et al.*, 2018) 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 (Howell, 1944; Maynard Smith & Savage, 1956; Gambaryan, 1974; Hildebrand, 1974; Coombs Jr, 1978; Garland & Janis, 1993; Stein & Casinos, 1997; Carrano, 1999; Lovegrove & Mowoe, 2014). 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 (Samuels & Van Valkenburgh, 2008; Lovegrove & Mowoe, 2014; Young *et al.*, 2014; Vianey-Liaud *et al.*, 2015), 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 (Middleton *et al.*, 2008a; Marchini *et al.*, 2014). 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 (Garland *et al.*, 2011; Careau *et al.*, 2013), 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 (Garland and Freeman 2005; Castro and Garland 2018). 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 increases bone strength and reduces the risk of bone fractures (Wallace and 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 (Jenkins & Camazine, 1977;

Álvarez *et al.*, 2013). 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 Results)? 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 Jr *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 (Kelly *et al.*, 2006; Castro & Garland, 2018). 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.

Materials and Methods

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, Indianapolis, Indiana, USA). Mice were randomly bred for two generations and then separated into 8 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 (Swallow *et al.*, 1998; Hiramatsu *et al.*, 2017). 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^{o} C, with ad lib access to food and water at all times. Photoperiod is 12L:12D with, the light phase beginning at 0700 hours and the dark phase at 1900 hours.

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 (lab designation), which remains polymorphic for the mini-muscle phenotype, which involves numerous differences in muscles, organs, and the skeleton (Garland Jr *et al.*, 2002; Syme *et al.*, 2005; Kelly *et al.*, 2006, 2017), 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.; Quincy, MA) 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.

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 S2.1-2.3. The authors elect not to share data.

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 (Appendix 2.1), many of which are the same as the measurements in previous studies (Kelly *et al.*, 2006; Castro & Garland, 2018), which allows for multi-generational comparisons. For three measurements (see Appendix 2.1), 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 & Van Valkenburgh, 2008; Samuels *et al.*, 2013) (Table 2.1). Morphometric indices were used to examine limb bone robusticity and anatomical advantage (in-lever/out-lever lengths) of various muscles on the appendicular skeleton.

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 (Garland, & Freeman, 2005; Kelly *et al.*, 2006; Middleton *et al.*, 2008b, 2010; Castro & Garland, 2018). 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 (Kelly *et al.*, 2006; Middleton *et al.*, 2008b; Castro & Garland, 2018). 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 Statistical

Analysis). The multi-generational comparisons graphs report P-values, age at sacrifice, and least square means with associated standard error bars, including 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 (Kelly *et al.*, 2006; Middleton *et al.*, 2008b; Copes *et al.*, 2018), but we only included data for mice that did not have wheel-access, as in the present study.

Statistical Analysis

We used the MIXED procedure in SAS (SAS Institute, Cary, NC, USA) 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 (Swallow *et al.*, 1999; Houle-Leroy *et al.*, 2000, 2003). Furthermore, the main effect of the mini-muscle phenotype (Garland Jr *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 2-tailed.

Considering all of the main analyses reported here, done separately by sex (S2.1,2 .2, 2.3), 288 P values were produced, 61 of which had nominal P values ≤ 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 values ≤ 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 Results), we discuss all P values that were nominally $P \leq 0.05$. Thus, all p values reported in the text are the nominal ones, not adjusted for multiple comparisons.

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 Jr *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 (Kelly *et al.*, 2006; Castro & Garland, 2018; 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 generation 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 Introduction).

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 3-5% (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 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-values < 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) has 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, 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.

Results

Significance levels from ANCOVAs of skeletal dimensions (using body length as a covariate) are presented in S2.1 and S2.2, whereas significance levels from ANOVAs

of morphological indices are presented in S2.3. In addition, S2.1-2.3 present Least Square Means for all statistical analysis.

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; S2.1) 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 (S2.1).

Skeletal Dimensions

Linetype differences were not statistically significant for individual bone lengths of the forelimb, hindlimb, scapula or pelvis, for either sex (S2.1 and S2.2). HR male mice had thinner distal ilia (P = 0.0481) (Figure 2.1) and lighter pelvises (P = 0.0500) (Figure 2.2) when compared with C male mice. However, HR females did not differ in the breadth of the distal ilium (P = 0.2263) (Figure 2.1), pelvis mass (P = 0.8532) (Figure 2.2), or other pelvis dimensions when compared with C female mice (S2.1). HR mice also had thinner femoral greater trochanters (P =0 .0012 for males; P = 0.0141 for females) when compared with C mice (Figure 2.3). For females, HR mice tended to have thicker tibia-fibula mid-shaft diameters (P = 0.0544) and heavier tibia-fibulas (P =

0.0699) (S2.1). In the forelimb, HR males had lighter humeri (P = 0.0219) (Figure 2.4) when compared with C males, but females did not significantly differ (P = 0.9361) (Figure 2.4) (S2.2). For females, HR mice tended to have thicker humeri (P = 0.0732) (S2.2).

Mini-muscle male mice had longer lower ilia (P = 0.0058) (Figure 2.5), but shorter pubic bones (P=0.0395), when compared with normal muscled male mice (S2.1). Mini-muscle female mice, however, did not differ in the length of the lower ilia (P = 0.5093) (Figure 2.5) 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 2.1) and lighter pelvises (P = 0.0074) (Figure 2.1) (S2.1). 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) (S2.1). 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 (S2.2).

Morphometric Indices

Results for the functional indices are presented in S2.3 (descriptions in Table 1). HR males had higher HRI indices (P = 0.0420) when compared with C male mice, suggesting more robust humeri (S2.3). 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 S2.3). For females, mini-muscle mice had higher 3^{rd}/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 (S2.3).

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 2.2 shows that mini-muscle female mice have lighter pelvises at larger body length values only.

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 2.6). 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 had larger knees when compared with HR mice) (Figure 2.6). Finally, HR female mice had thicker distal femora throughout the generations sampled when compared with C female mice, but especially so at generation 16.

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 2.7). 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 2.7). 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 vs. 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 2.7).

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 (S2.4). However, the Type I error rate for mini-muscle was unaffected under the Mini50 conditions, but was inflated under the MiniFix scenario (S2.4). 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) (S2.4). 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.

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.

High Runner vs 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 (S2.1). 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) (Appendix 2.2), 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 (S2.1) and humeri (S2.2) 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).

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 (S2.1-2.3). For example, mini-muscle males had lighter pelves, 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 (S2.1). 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 (S2.2), 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 (Polly, 2007; Álvarez *et al.*, 2013; Charles *et al.*, 2016b). 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 (0.04-62kg). 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) (Maynard Smith & Savage, 1956; Polly, 2007;

Álvarez *et al.*, 2013; Lewton, 2015) since muscle fiber length 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 (Kelly *et al.*, 2006; Claghorn *et al.*, 2017; 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).

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 2.6). 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 vs 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 greater than 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 2.7) 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 2.7).

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 (S2.4).

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.

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 generation 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 2.2). 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 (Garland, & Freeman, 2005; Kelly *et al.*, 2006; Castro & Garland, 2018), 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.

Conclusions

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; Wallace *et al.*, 2012; Schutz *et al.*, 2014). However, after selection limits occurred, some skeletal adaptations were lost.

The apparent loss of skeletal coadaptations might be explained in various 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; Miyatake, 2002; Hannon *et al.*, 2011; Hiramatsu, 2017).

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 2.7). 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 2.6), as well as additional simulations (S2.4).

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.

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. *Physiol. Biochem. Zool.* **79**: 282–294.

Biancardi, C.M. & Minetti, A.E. 2012. Biomechanical determinants of transverse and rotary gallop in cursorial mammals. *J. Exp. Biol.* **215**: 4144–4156.

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

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. *J. Exp. Zool. Part Ecol. Genet. Physiol.* **311A**: 189–198.

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

Bult, A. & Lynch, C.B. 1996. Multiple selection responses in house mice bidirectionally selected for thermoregulatory nest-building behavior: Crosses of replicate lines. *Behav. Genet.* **26**: 439–446.

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. *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. *J. Zool.* **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. *J. Morphol.* **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. *J. Anat.*, doi: 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. *Physiol. Biochem. Zool.* **90**: 533–545.

Coombs Jr, W.P. 1978. Theoretical aspects of cursorial adaptations in dinosaurs. *Q. Rev. Biol.* **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. *Am. J. Phys. Anthropol.* **166**: 442–458.

Cowgill, L.W. 2009. The ontogeny of holocene and late pleistocene human postcranial strength. *Am. J. Phys. Anthropol.* 16–37.

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.

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. *J. Exp. Biol.* **212**: 2612–2618.

Foster, K.L., Garland, 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. *Biol. J. Linn. Soc.*, doi: 10.1093/biolinnean/bly146.

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

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? *J. Zool.* **229**: 133–151.

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. *Proc. R. Soc. B Biol. Sci.* **278**: 574–581.

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.

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

Gregory, W.K. 1912. Notes on the principles of quadrupedal locomotion and on the mechanism of the limbs in hoofed animals. *Ann. N. Y. Acad. Sci.* **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. *Comp. Biochem. Physiol. B Biochem. Mol. Biol.* **144**: 271–282.

Hannon, R.M., Meek, T.H., Acosta, W., Maciel, R.C., Schutz, H. & Garland, T. 2011. Sex-specific heterosis in line crosses of mice selectively bred for high locomotor activity. *Behav. Genet.* **41**: 615–624.

Hildebrand, M. 1974. *Analysis of vertebrate structure*. John Wiley and Sons, New York.

Hiramatsu, L. 2017. Physiological and genetic causes of a selection limit for voluntary wheel-running in mice. University of California, Riverside, Riverside, 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. *Physiol. Behav.* **179**: 235–245.

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. *J. Linn. Soc. Lond. Zool.* **41**: 259–271.

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*. *J. Appl. Physiol.* **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. *Am. J. Physiol. - Regul. Integr. Comp. Physiol.* **284**: R433–R443.

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

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. *J. Zool.* **181**: 351–370.

Jones, K.E. 2016. New insights on equid locomotor evolution from the lumbar region of fossil horses. *Proc. R. Soc. B Biol. Sci.* **283**: 20152947.

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. *J. Morphol.* **267**: 360–374.

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. *J. Exp. Biol.* **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. *Anat. Rec.* **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. *Am. J. Phys. Anthropol.* **168**: 543–551.

Lovegrove, B.G. & Mowoe, M.O. 2014. The evolution of micro-cursoriality in mammals. *J. Exp. Biol.* **217**: 1316–1325.

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 Evol. Biol.* **14**: 258.

Maynard Smith, J. & Savage, R.J.G. 1956. Some locomotory adaptations in mammals. *J. Linn. Soc. Lond. Zool.* **42**: 603–622.

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. *J. Morphol.* **183**: 87–115.

Meachen-Samuels, J. & Van Valkenburgh, B. 2009. Forelimb indicators of prey-size preference in the Felidae. *J. Morphol.* **270**: 729–744.

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. *J. Anat.* **216**: 121–131.

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. *Integr. Comp. Biol.* **48**: 394–410.

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.

Miyatake, T. 2002. Circadian rhythm and time of mating in *Bactrocera cucurbitae* (Diptera: Tephritidae) selected for age at reproduction. *Heredity* **88**: 302–306.

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

Oxnard, C.E. 1967. The functional morphology of the primate shoulder as revealed by comparative anatomical, osteometric and discriminant function techniques. *Am. J. Phys. Anthropol.* **26**: 219–240.

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

Rose, M.R. 2005. The effects of evolution are local: Evidence from experimental evolution in *Drosophila*. *Integr. Comp. Biol.* **45**: 486–491.

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

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

Schneider, C.A., Rasband, W.S. & Eliceiri, K.W. 2012. NIH Image to ImageJ: 25 years of image analysis. *Nat. Methods* **9**: 671–675.

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

Schutz, H., Jamniczky, H.A., Hallgrímsson, B. & Garland, 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.

Schwartz, N.L., Patel, B.A., Garland, 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. *J. Anat.* **233**: 193–203.

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.

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. *Physiol. Behav.* **198**: 27–41.

Stein, B.R. & Casinos, A. 1997. What is a cursorial mammal? *J. Zool.* **242**: 185–192.

Swallow, J.G., Carter, P.A. & Garland Jr., T. 1998. Artificial selection for increased wheel-running behavior in house mice. *Behav. Genet.* **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. *J. Exp. Biol.* **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. *J. Appl. Physiol.* **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. *Mech. Dev.* **134**: 16–30.

Van Der Klaauw, C.J. 1948. Ecological Morphology. In: *Ecological Studies and Reviews*. Bibliotheca Biotheoretica.

Van Valkenburgh, B. 1987. Skeletal indicators of locomotor behavior in living and extinct carnivores. *J. Vertebr. Paleontol.* **7**: 162–182.

Vanhooydonck, B. & Van Damme, R. 2001. Evolutionary trade-offs in locomotor capacities in lacertid lizards: are splendid sprinters clumsy climbers? *J. Evol. Biol.* **14**: 46–54.

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: *Evolution of the Rodents* (P. G. Cox & L. Hautier, eds), pp. 539–588. Cambridge University Press, Cambridge.

Wallace, I.J. & Garland Jr., T. 2016. Mobility as an emergent property of biological organization: Insights from experimental evolution: Mobility and biological organization. *Evol. Anthropol. Issues News Rev.* **25**: 98–104.

Wallace, I.J., Judex, S. & Demes, B. 2015. Effects of load-bearing exercise on skeletal structure and mechanics differ between outbred populations of mice. *Bone* **72**: 1–8.

Wallace, I.J., Middleton, K.M., Lublinsky, S., Kelly, S.A., Judex, S., Garland Jr., T., *et al.* 2010. Functional significance of genetic variation underlying limb bone diaphyseal structure. *Am. J. Phys. Anthropol.* **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. *Am. J. Phys. Anthropol.* **148**: 24–35.

Young, J.W., Danczak, R., Russo, G.A. & Fellmann, C.D. 2014. Limb bone morphology, bone strength, and cursoriality in lagomorphs. *J. Anat.* **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. *Evol. Biol.* **36**: 88–99.

Table 2.1. 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.

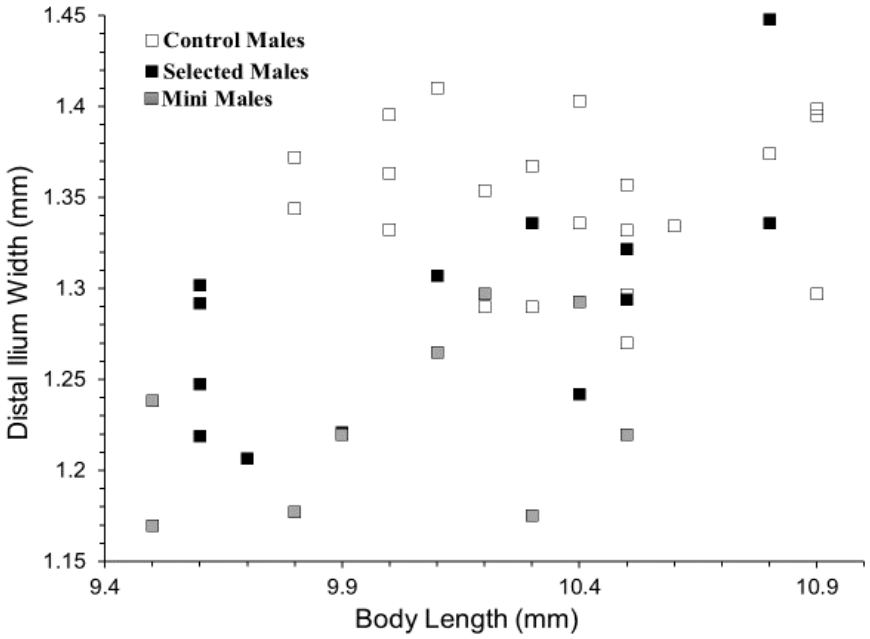
Morphometric Indices	Definition and Functional Significance
Metatarsal Femur Ratio (MT/F)	Relative proportions of the proximal and distal hindlimb (or relative size of the hindfoot) Classically used as an indicator of "cursoriality" (3rd Metatarsal Length/Femur Length) (Garland & Janis, 1993; Samuels *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) (Vanhooydonck & Van Damme, 2001; Biancardi & Minetti, 2012; Samuels *et al.*, 2013)
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 & Van Valkenburgh, 2008; Samuels *et al.*, 2013)
Tibia Robusticity Index (TRI)	Robusticity of the tibia and ability to resist shearing and bending stresses (Tibia-fibula Mid-Shaft Diameter/Tibia Length) (Samuels & Van Valkenburgh, 2008; Samuels *et al.*, 2013)
Humerus Robusticity Index (HRI)	Robusticity of the humerus and ability to resist shearing and bending stresses (Humerus Mid-Shaft Diameter/Humerus Length) (Samuels & Van Valkenburgh, 2008; Samuels *et al.*, 2013)
Ulna Robusticity Index (URI)	Robusticity of the ulna and ability to resist shearing and bending stresses (Ulna Mid-Shaft Diameter/Ulna Length) (Samuels & Van Valkenburgh, 2008; Samuels *et al.*, 2013)
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)	Robusticty of the distal hindlimb and relative size of the ankle joint (Tibia-fibula Distal Width/Tibia Length) (Morris & Carrier, 2016; Castro & Garland, 2018)
Ischium Anatomical Advantage (IA)	Anatomical advantage of the bicpes femoris, semimembranosus, and semitendinosus when extending the hip joint (Ischium Length/Femur Length) (Young *et al.*, 2014; Charles *et al.*, 2016a; Morris & Carrier, 2016)
Gluteal Anatomical Advantage (GA)	Anatomical advantage of the glutues maximus when rotating the hip joint (Greater Trochanter Height/Femur Length) (Samuels & Van Valkenburgh, 2008; Samuels *et al.*, 2013; Charles *et al.*, 2016a)
Third Trochanter Anatomical Advantage (3rd/F)	Anatomical advantage of the quadtratus femoris when rotating the hip joint (Femoral Head to 3rd Trochanter Muscle Scar /Femur Length) (Charles *et al.*, 2016a; Castro & Garland, 2018)
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) (Samuels & Van Valkenburgh, 2008; Samuels *et al.*, 2013; Charles *et al.*, 2016a)

Table 2.2. The number of statistically significiant (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	Genera-tion	Males	Females	# studied	# signif-icant	% signif-icant
Garland and Freeman 2005; Castro and Garland 2018	11	X	X	58	3	5
Middleton et al. 2008b	16		X	9	2	22
Kelly et al. 2006; Young et al. 2009	21	X		26	8	30
Middleton et al. 2010	37		X	6	0	0
Copes et al. 2018; Lewton et al. 2019; Present Study	57		X	3	0	0
Present Study	68	X	X	110	7	6

Figure 2.1. Distal Ilium Width. Figure 1A and 1B Mean distal Ilium width in relation to body length and split by sex. For male mice 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 for either linetype or mini-muscle. Mini-muscle mice are in grey.



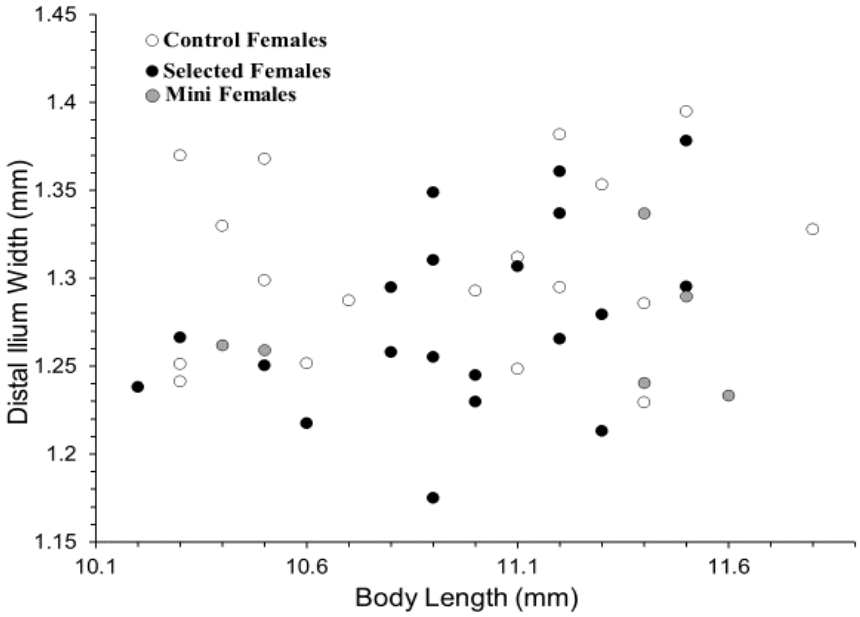
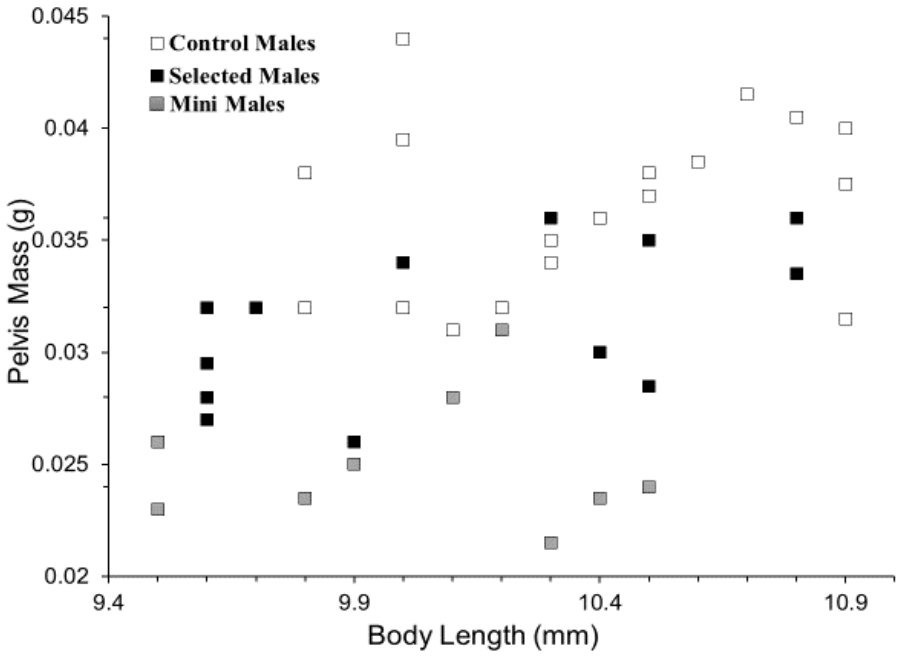


Figure 2.2. Pelvis Mass. Figure 2A and 2B Mean pelvis mass in relation to body length and split by sex. For male mice only, the effect of body length was positive and statistically significant. HR males had significantly lighter pelvises for a given body length, and mini-muscle males had lighter pelvses when compared to normal-muscled males. For females, there was no signifcant effect for linetype, but mini-muscle mice have lighter pelvises at a larger body lengths only. Mini-muscle mice are in grey.



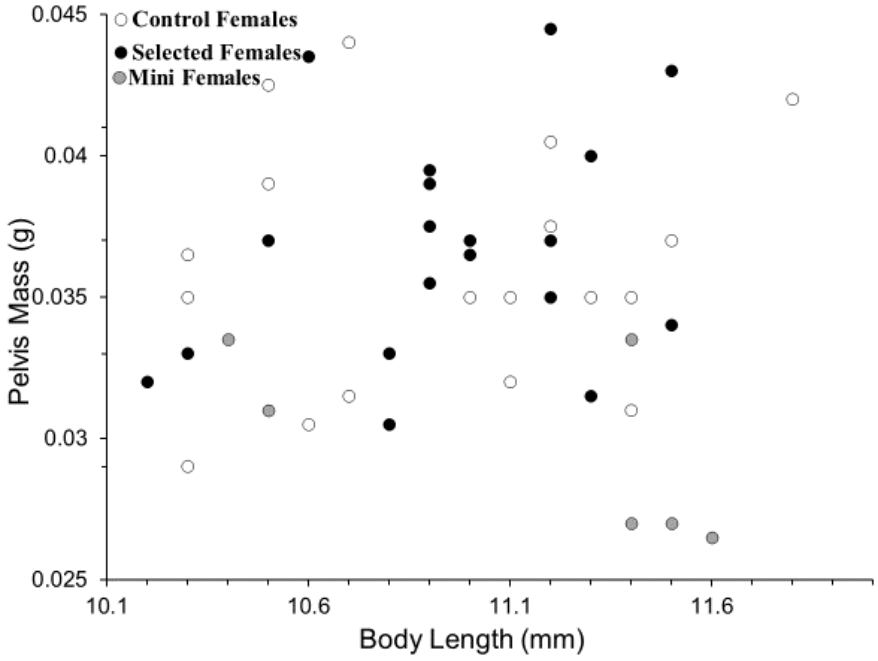


Figure 2.3. Greater Trochanter Breadth. Figure 3A and 3B Mean femoral greater trochanter in relation to body length and split by sex. HR male and female mice had significantly thinner femoral greater trochanters for a given body length, and mini-muscle females had thinner femoral greater trochanters when compared to normal-muscled females. Mini-muscle mice are in grey.

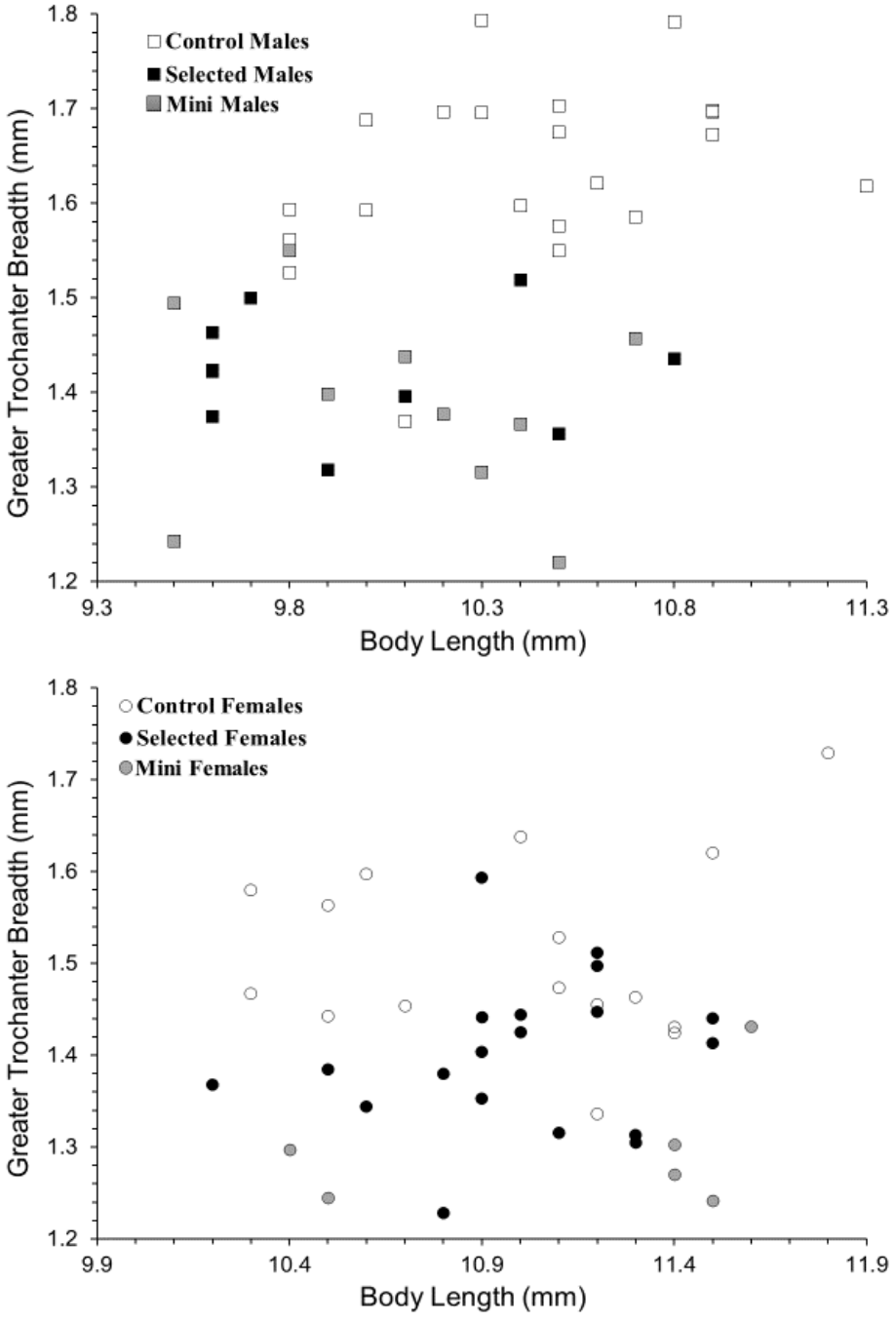


Figure 2.4. Humerus Mass. Figure 4A and 4B Mean humerus mass in relation to body length and split by sex. For male mice only, the effect of body length was positive and statistically significant. HR males had significantly lighter humeri for a given body length, and mini-muscle males had lighter humeri when compared to normal-muscled males. For females, there was no signifcant effect for either linetype or mini-muscle. Mini-muscle mice are in grey.

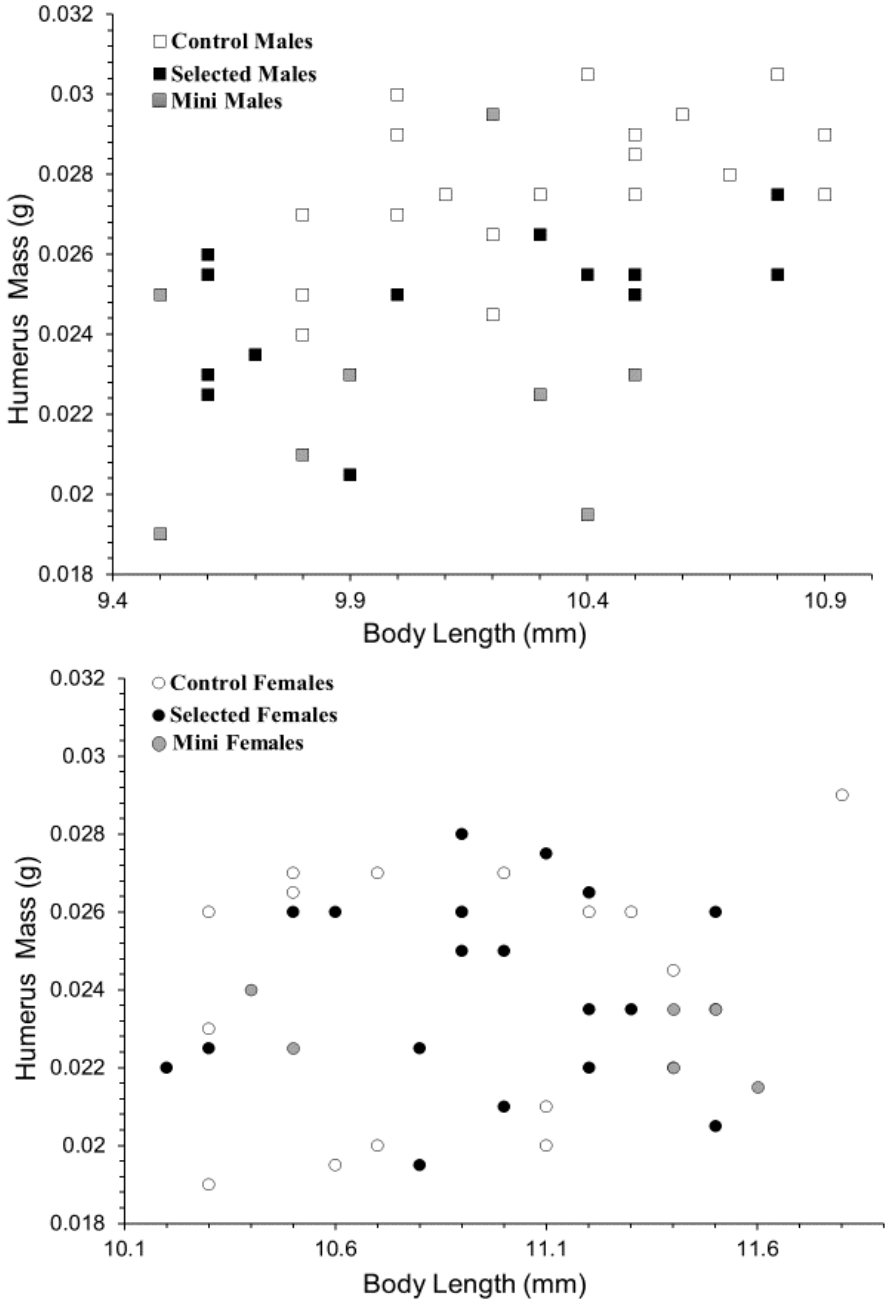


Figure 2.5. Lower Ilium Length. Figure 5A and 5B Mean lower ilium length in relation to body length and split by sex. For both sexes, the effect of body length was positive and statistically significant. Mini-muscle males had significantly longer lower ilia for a given body length. For females, there was no signifcant effect for either linetype or mini-muscle. Mini-muscle mice are in grey.

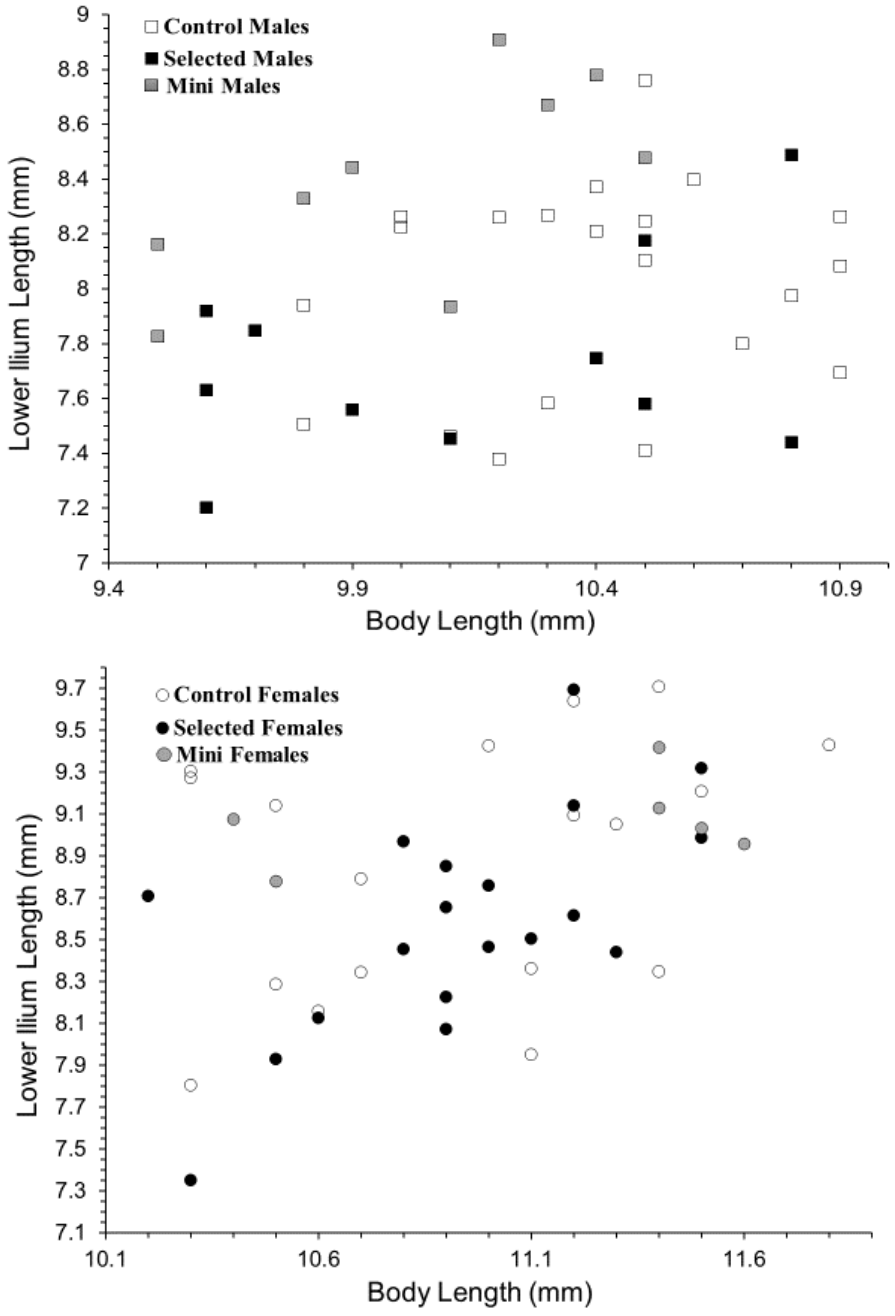
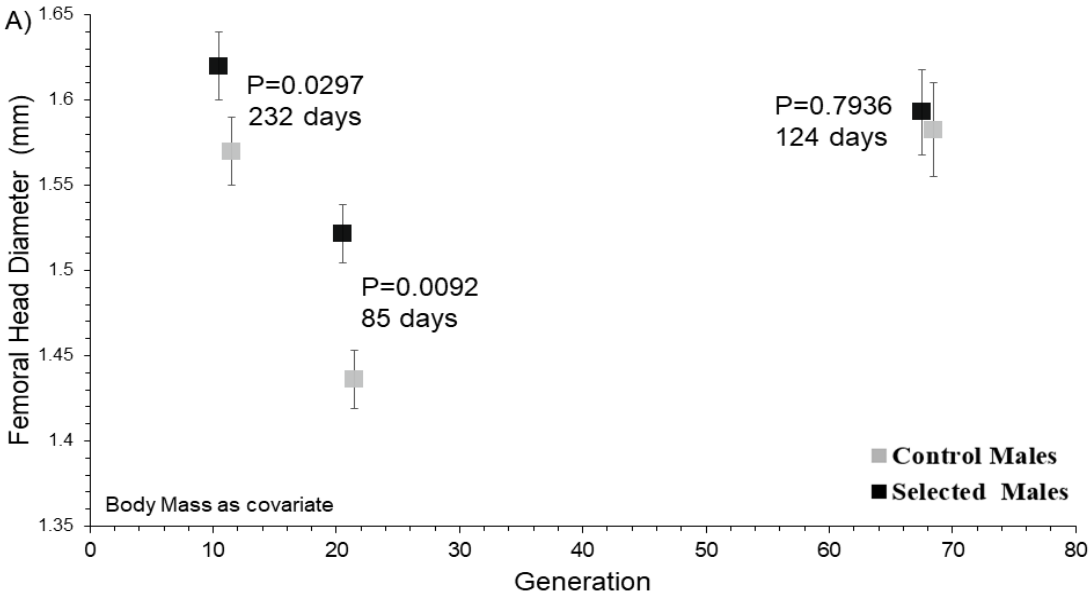
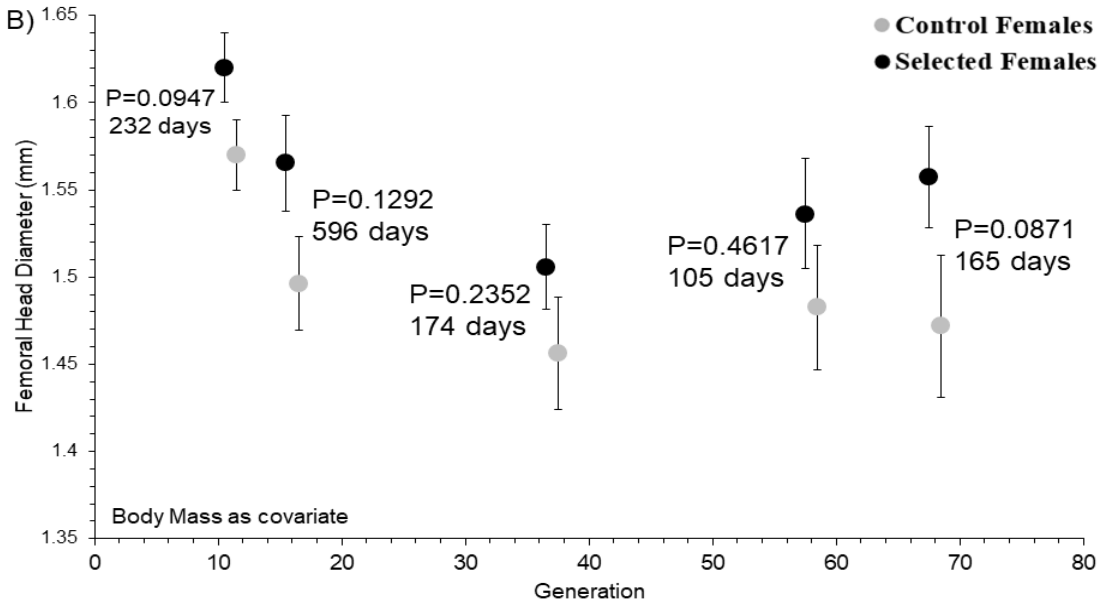
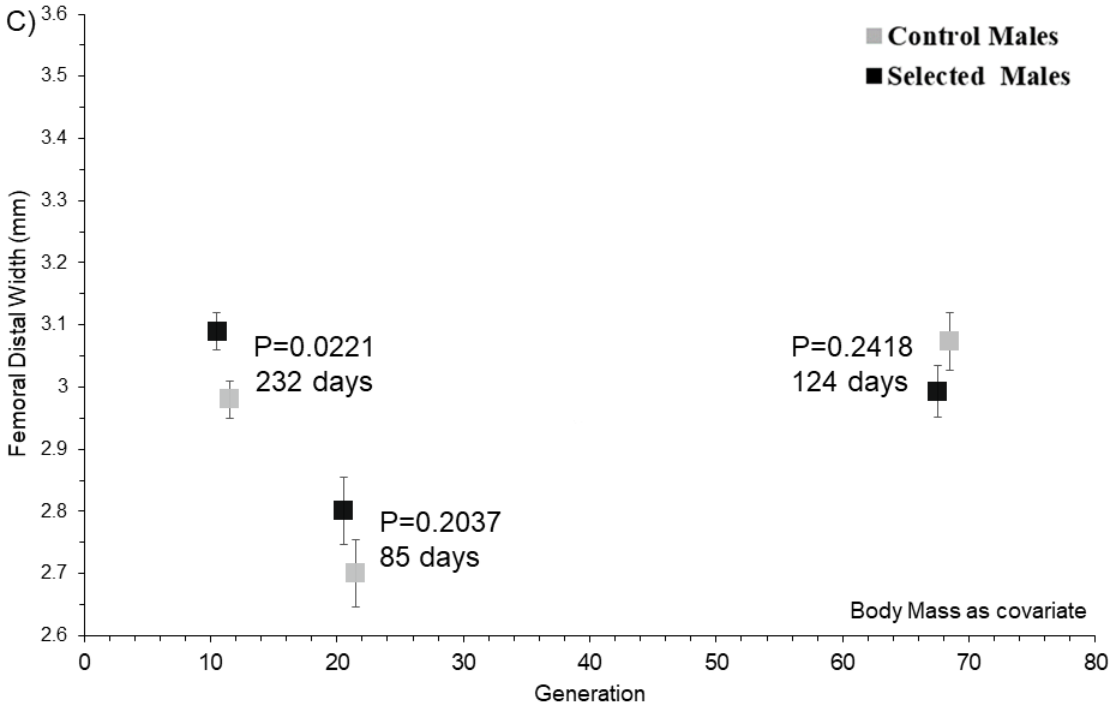


Figure 2.6. Femoral Dimensions by Generation. Figure 2.6 Mean femoral dimensions across generations. Three femoral dimensions are plotted across generations separately for males and females. Values are least square means and associated standard error bars, with body mass as a covariate and age presented in the graphs.
A) and B): HR male mice had thicker femoral heads when compared with C mice in earlier generations, but these differences were not appraent in the later generation. HR female mice have thicker femoral heads throughout the generations sampled although these differences were not always significant.
C) and D): HR male mice had thicker distal femora when compared with C mice in earlier generations, but these differences were not appraent in the later generation. HR female mice have thicker distal femora throughout the generations sampled, but especially so at generation 16.


A)
Femoral Head Diameter (mm)
1.65
1.6
1.55
1.5
1.45
1.4
1.35
P=0.0297
232 days
P=0.0092
85 days
P=0.7936
124 days
Body Mass as covariate
Control Males
Selected Males
0
10
20
30
40
50
60
70
80
Generation


B)
Control Females
Selected Females
P=0.0947
232 days
P=0.1292
596 days
P=0.2352
174 days
P=0.4617
105 days
P=0.0871
165 days
Body Mass as covariate
Femoral Head Diameter (mm)
Generation
1.65
1.6
1.55
1.5
1.45
1.4
1.35
0
10
20
30
40
50
60
70
80


C)
Control Males
Selected Males
Femoral Distal Width (mm)
3.6
3.5
3.4
3.3
3.2
3.1
3
2.9
2.8
2.7
2.6
P=0.0221
232 days
P=0.2037
85 days
P=0.2418
124 days
Body Mass as covariate
0
10
20
30
40
50
60
70
80
Generation

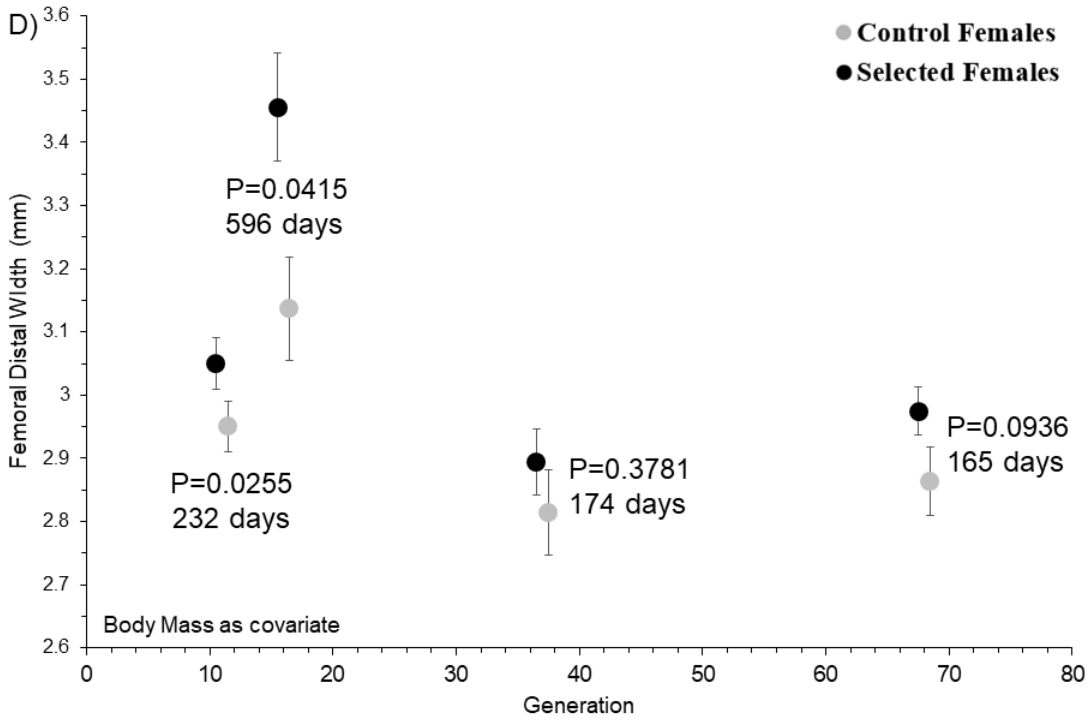
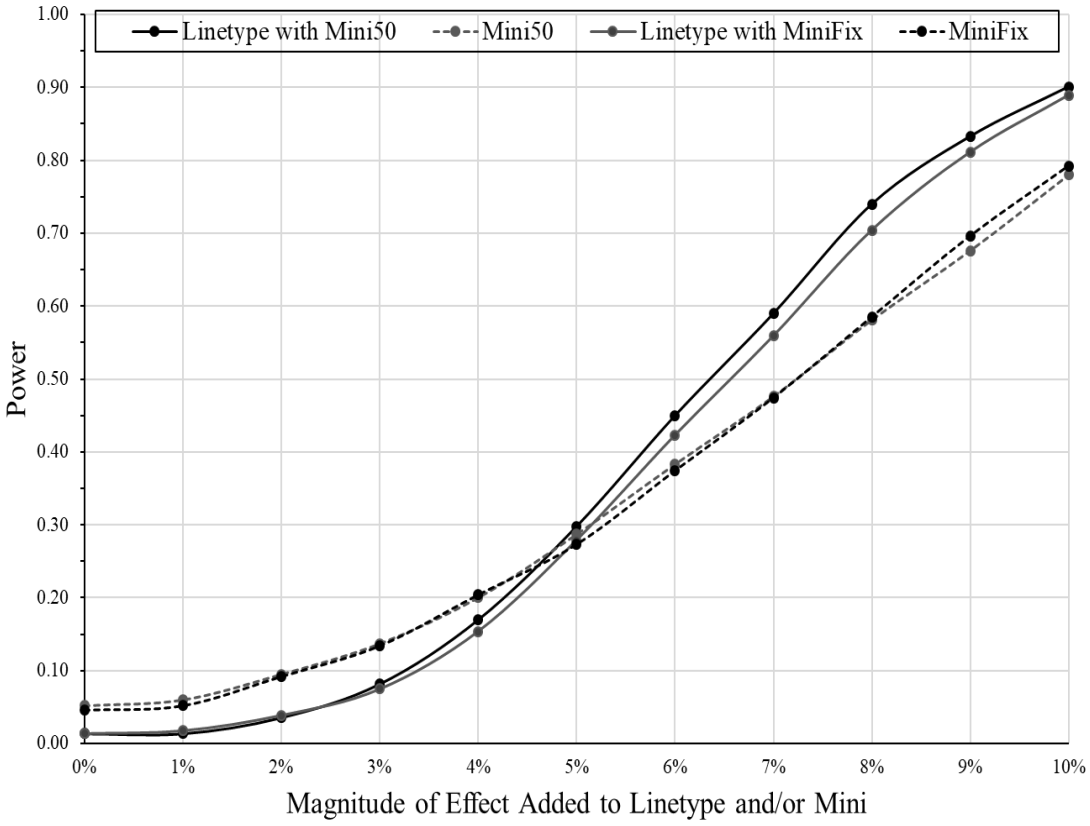

D)
Control Females
Selected Females
Femoral Distal Width (mm)
3.6
3.5
3.4
3.3
3.2
3.1
3
2.9
2.8
2.7
2.6
P=0.0415
596 days
P=0.0255
232 days
P=0.3781
174 days
P=0.0936
165 days
Body Mass as covariate
0
10
20
30
40
50
60
70
80
Generation

Figure 2.7. Power Curve. Figure 7 Variable One Power Curve. 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 vs. 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 rate for the mini-muscle effects were very close to the expected 5%, but for the HR vs C effect it was greatly deflated (~1.4%).



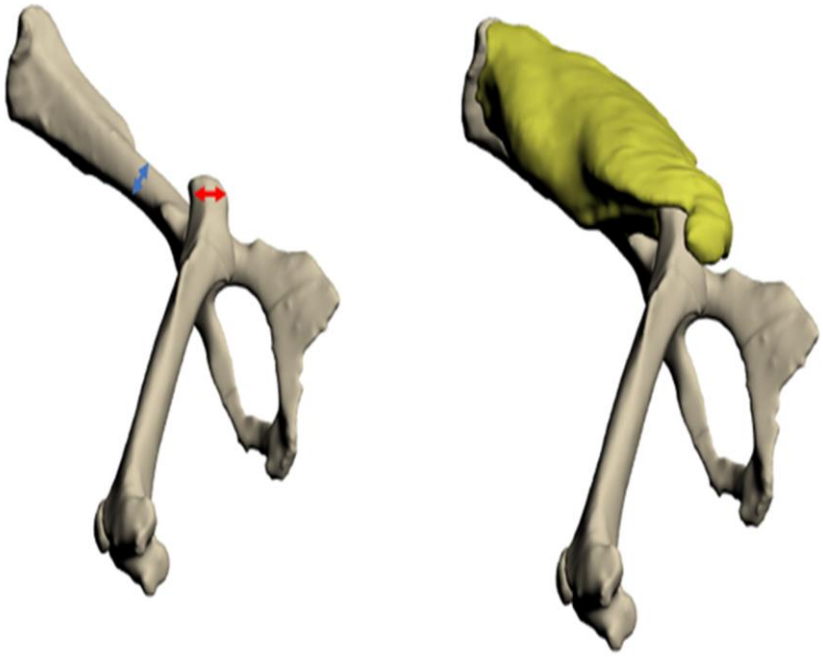
Appendices

Appendix 2.1. Skeletal Measurements. Skeletal morphometrics taken for mouse specimens.
* Denotes skeletal traits measured with calipers

Skeletal Measurement	Definition and Functional Significance
Femur Length	Dorsal tip of the femoral head to the distal most end of the medial condyle
Femoral Head Diameter	Medial-lateral width of the femoral head
Femoral Width at 3rd Trochanter	Breadth of the 3rd trochanter at the widest point
Femoral Distal Width	Breadth of the distal femur across the medial and lateral epicondyles
Femoral Greater Trochanter Breadth	Medial-lateral width of the greater trochanter on the femur
Femoral Mid-shaft Diameter	Medial-lateral width at the mid-point of the femur shaft
Femoral Head to 3rd Trochanter Muscle Scar	Dorsal tip of the femoral head to the distalmost end of the 3rd trochanter muscle scar *
Femoral Greater Trochanter Length	Height of the greater trochanter on the femur *
Femoral Mass	Mass of the femur
Tibia Length	Superior articular surface of the lateral condyle to the tip of the medial malleolus
Tibial Proximal Width	Greatest medio-lateral distance across the proximal end of the tibia, includes spike on the fibula)
Tibia-fibula Distal Width	Greatest medial-lateral width at the distal end of the tibia
Tibia-fibula Mid-Shaft Diameter	Medial-lateral width at the mid-point of the tibia-fibula shaft
Tibia-fibula Mass	Mass of the tibia-fibula
3rd Metatarsal Length	Greatest length of the 3rd metatarsal (not including phalanges and tarsals)
Calcaneum Length	Tip of the calcaneum heel to the articulation point *
Pelvis Mass	Mass of the pelvis
Pelvis Length	Ramus of ischium to the most proximal end of iliac crest
Lower Ilium Length	Lower end of the proximal ilium to the most lateral tip of the ilio-pectineal eminence
Least Distal Width of Ilium	Smallest breadth at the distal end of the ilium
Greatest Proximal Width	Greatest medial-lateral width at the proximal end of the

of Ilium	ilium
Ischium Length	Most Lateral point of the ischium to the acetabulum
Greatest Distal Width of Ischium	Breadth of the ischium from the most lateral point of the ischium tuberosity to the obturator foramen
Pubis Length	Most lateral point of the descending ramus of the pubis to the obturator foramen
Least Width of Pubis	Smallest breadth of the pubis
Scapula Length	Midpoint of scapular spine to the acromion process
Scapula Width	Greatest breadth of the scapula blade across the superior and inferior angles
Humerus Length	Humerus Head to the trochlea
Humerus Head Width	Medial-lateral diameter of the humerus head
Humerus Distal Width	Breadth of the distal humerus across the medial and lateral epicondyles
Humerus Mid-Shaft Diameter	Medial-lateral width at the mid-point of the humerus shaft
Humerus Mass	Mass of the humerus
Ulna Length	Growth plate on the olecranon to the styloid process
Olecranon Length	Growth plate on the olecranon to the distal end of the olecranon process
Radius Length	Radial head to the styloid process
Styloid Width	Greatest medial-lateral width at the distal end of the radius and ulna
Ulna Mid-Shaft Diameter	Medial-lateral width at the mid-point of the ulna shaft
Metacarpal Length	Greatest length of the 3rd metacarpal (not including phalanges and carpals)

Appendix 2.2. 3D Model of Mouse Gluteal Muscles. As modeled by Charles et al. 2016, illustrating the gluteal muscles that act to rotate and extend the hip joint during locomotion. The gluteal muscles originate on the ilium and insert on the femoral greater trochanter. The blue arrow and red arrow indicate our measurements for the least distal width of ilium and femoral greater trochanter breadth. The 3D muscle and bone models are in the lateral view.



Effects of selective breeding for voluntary exercise, chronic exercise, and their interaction on muscle attachment site morphology in house mice

Alberto A. Castro[a], Fotios Alexandros Karakostis[b], Lynn E. Copes[c], Holland E. McClendon[a], Aayushi P. Trivedi[a], Nicole E. Schwartz[a], Theodore Garland, Jr.[a,*]

[a] *Department of Evolution, Ecology, and Organismal Biology, University of California, Riverside, Riverside, CA 92521, USA*

[b] *Paleoanthropology, Department of Geosciences, Senckenberg Centre for Human Evolution and Palaeoenvironment, University of Tübingen, Tübingen 72070, Germany*

[c] *Department of Medical Sciences, Frank H. Netter MD School of Medicine, Quinnipiac University, Hamden, CT 06518, USA*

Acknowledgments

Supported by NSF grant DEB-1655362 to TG.
We thank Campbell Rolian and an anonymous referee for comments on the manuscript.

Author Contributions

LEC and TG provided the specimens. LEC provided the CT scans. AAC, FAK, and NLS developed the measurement protocol. AAC, FAK, HEM, and APT collected the data. AAC, TG, and FAK analyzed the data and drafted the manuscript. All authors edited the final version of the manuscript.

ABSTRACT

Skeletal muscles attach to bone at their origins and insertions, and the interface where tendon meets bone is termed the attachment site or enthesis. Mechanical stresses at the muscle/tendon-bone interface are proportional to the surface area of the bony attachment sites, such that a larger attachment site will distribute loads over a wider area. Muscles that are frequently active and/or are of larger size should cause attachment sites to hypertrophy (training effect); however, experimental studies of animals subjected to exercise have provided mixed results. To enhance our ability to detect training effects (a type of phenotypic plasticity), we studied a mouse model in which 4 replicate lines of High Runner (HR) mice have been selectively bred for 57 generations. Selection is based on the average number of wheel revolutions on days 5 & 6 of a 6-day period of wheel access as young adults (6-8 weeks old). Four additional lines are bred without regard to running and serve as non-selected controls (C). On average, mice from HR lines voluntarily run ~3 times more than C mice on a daily basis. For this study, we housed 50 females (half HR, half C) with wheels (Active group) and 50 (half HR, half C) without wheels (Sedentary group) for 12 weeks starting at weaning (~3 weeks old). We tested for evolved differences in muscle attachment site surface area between HR and C mice, plastic changes resulting from chronic exercise, and their interaction. We used a precise, highly repeatable method for quantifying the three-dimensional (3D) surface area of four muscle attachment sites: the humerus deltoid tuberosity (the insertion point for the spinodeltoideus, superficial pectoralis, and acromiodeltoideus), the femoral third trochanter (the insertion point for the quadratus femoris), the femoral lesser trochanter

(the insertion point for the iliacus muscle), and the femoral greater trochanter (insertion point for the middle gluteal muscles). In univariate analyses, with body mass as a covariate, mice in the Active group had significantly larger humerus deltoid tuberosities than Sedentary mice, with no significant difference between HR and C mice and no interaction between exercise treatment and linetype. These differences between Active and Sedentary mice were also apparent in the multivariate analyses. Surface areas of the femoral third trochanter, femoral lesser trochanter, and femoral greater trochanter were unaffected by either chronic wheel access or selective breeding. Our results, which used robust measurement protocols and relatively large sample sizes, demonstrate that muscle attachment site morphology can be (but is not always) affected by chronic exercise experienced during ontogeny. However, contrary to previous results for other aspects of long bone morphology, we did not find evidence for evolutionary coadaptation of muscle attachments with voluntary exercise behavior in the HR mice.

Introduction

Bone is a dynamic and metabolically active organ composed of calcium phosphate minerals and type I collagen. Bone modeling and remodeling, the actions of osteoclasts and osteoblasts during bone resorption and formation, is essential for the mineral and mechanical homeostasis of the skeleton (Frost, 2003; Doherty *et al.*, 2015; Katsimbri, 2017). Mechanical forces acting on the skeleton cause strain and microdamage to bone tissue, which is responded by osteocytes (mechanosensory cells that sense fluid flow associated with strain) that translate mechanical strain to biochemical signals, and initiates bone remodeling (Bonewald, 2007; Yu *et al.*, 2017). On one hand, increases in mechanical load cause changes in shape and material properties of bones that lead to increased stiffness and strength (Frost, 2003; Ruff *et al.*, 2006; Hart *et al.*, 2017). On the other, structural and mechanical changes occur on the skeleton during paralysis, unloading, and/or disuse (lower levels of strain and stress), that leads to declines in bone mass and mechanical integrity (Morey-Holton & Globus, 1998; Kodama *et al.*, 1999; Ruff *et al.*, 2006; Maupin *et al.*, 2019).

Physical conditioning (e.g., exercise through running, weightlifting) is important for the maintenance of adequate bone mass and strength. In mammals, exercise induces bone formation and retards bone loss, enhancing bone structure and ultimately strength (Rubin & Lanyon, 1984; Eliakim *et al.*, 1997; Lieberman, 2003; Plochocki *et al.*, 2008). For example, in studies of rats and mice, both voluntary wheel running and tower climbing required to obtain food can increase the thickness and mass of the tibia-fibula and femur (Newhall *et al.*, 1991; Notomi *et al.*, 2001; Mori *et al.*, 2003).

The extent of training effects depend on genetic factors (Middleton *et al.*, 2008a; Peacock *et al.*, 2018), as well as age, sex, and epigenetic factors (discussed in Wallace *et al.*, 2012). For example, one study used two outbred strains of mice (ICR vs. CD1) to examine the effects of exercise (30 minutes of treadmill-running, 5 days a week) on skeletal structure and mechanics (Wallace *et al.*, 2015). ICR mice that ran had significantly improved diaphyseal bone quantity, enhanced trabecular morphology, and increased femoral mechanical strength (as compared with sedentary controls). However, CD1 mice that ran (same regime) had reduced femoral structural strength (diaphyseal resistance to fracture). As another example, Peacock et al. (2018) studied the effects of genetics and exercise on bone properties in three inbred mouse strains: high bone density (C3H/He), low bone density (C57BL/6), and a high-runner strain homozygous for the *Myh4*$^{\text{Minimsc}}$ allele (see Methods). Although several interstrain differences were observed, femoral bone cross-sectional geometry and bending mechanics were not significantly different between exercised (wheel access for a 7-week period) and sedentary mice.

The origins and insertions of muscles adhere to the skeleton directly by aponeurosis or via tendons, both of which are termed muscle attachments (or entheses) (Benjamin *et al.*, 1986, 2002). Mechanical stresses at the muscle/tendon-bone interface are proportional to surface area of the bony attachment sites, such that larger entheses will distribute loads over a wider surface area (Biewener, 1992). Recent studies have investigated the micromechanics and microstructure of muscle entheses, finding variation in material properties (hard-to-soft interface) and collagen/fiber orientation that leads to concentrated compliance zones at the micrometer level when the attachments are loaded

(Deymier *et al.*, 2017; Rossetti *et al.*, 2017). Therefore, as with other aspects of long bone morphology, one would expect that intense and sustained physical activity might cause growth and/or remodeling of attachment sites (i.e., training effects due to increases in muscle activation and/or muscle size). This expectation has served as the basis for studies attempting to reconstruct the physical activity levels of an organism from fossilized bones based on their muscle attachment site morphology (Hawkey & Merbs, 1995; Schlecht, 2012; Foster *et al.*, 2014; Becker, 2020; Karakostis *et al.*, 2021). However, experimental studies of the effects of physical activity on muscle attachment site morphology have generated mixed results.

For example, mature female sheep wearing weighted backpacks were given treadmill exercise for one hour, 5 days/week, over 90 days. Muscle attachment site morphology did not differ between exercised and sedentary sheep, and muscle mass did not correlate with muscle attachment size or complexity within the sedentary group (not analyzed in the active group) (Zumwalt, 2006). In another study, muscle enthesis size, periosteal growth rate, and muscle architecture of the upper forelimb were compared among sedentary mice and those housed with either wheels or a climbing tower (exercise) for 11 weeks, beginning either at 25 or 46 days of age (Rabey *et al.*, 2015). Both types of exercise increased the periosteal growth rate of the deltoid tuberosity and significantly altered fiber lengths and physiological cross-sectional areas of the shoulder muscles; however, muscle attachment surface area, length, and diameter of the deltoid crest were not significantly altered (Rabey *et al.*, 2015). Finally, muscle entheseal topography, diaphyseal bone dimensions, and trabecular architecture of the femur were

compared in growing female turkeys that were either trained on a declined treadmill (inclined running group not included) or remained sedentary over a 10-week period (Wallace *et al.*, 2017). Trabecular thickness of the distal femoral metaphysis (knee) and second moments of inertia of the femoral mid-shaft were increased in exercised turkeys, but muscle enthesis topography was unaltered (but see below).

Although the foregoing studies suggest that muscle attachment site morphology may be generally less plastic than other aspects of bone morphology (e.g., length, thickness, density), more recent studies indicate that it can respond to physical activity, loading, or direct muscle stimulation. For example, with the same individual turkeys as in Wallace et al. (2017), and an additional group of incline runners, Karakostis et al. (2019b) found distinctive multivariate patterns involving three different entheses that distinguish controls from running groups (both inclined and declined running groups). Furthermore, a study of adult rats found that in vivo electrical muscle stimulation (over 28 days) caused changes (relative to non-stimulated controls) that differed among muscle entheses and reflected repetitive muscle recruitment (Karakostis *et al.*, 2019a).

Here, we compare mice from four replicate lines that have been 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; Careau *et al.*, 2013). All four HR lines evolved rapidly and reached selection limits after ~17-27 generations, depending on replicate line and sex (Garland *et al.*, 2011; Careau *et al.*, 2013), at which point HR mice run approximately three-fold more wheel revolutions per day than C mice. These high levels of physical activity should enhance statistical power

to detect training effects (phenotypic plasticity) in muscle attachment site morphology. In addition, we tested for differences between the HR and C lines to study possible coadaptation of the skeleton with physical activity behavior.

Although previous studies of HR mice have found size and shape differences between the long bones of HR and C mice (e.g., see Garland & Freeman, 2005; Wallace *et al.*, 2010; Schwartz *et al.*, 2018), which is part of their overall "mobility" phenotype (Wallace & Garland, 2016), no studies of muscle enthesis morphology are available for this animal model. Thus, our objectives were to quantify any evolved differences in muscle attachment site morphology between HR and C mice, plastic changes resulting from chronic exposure to exercise (active vs. sedentary experimental treatments), and potential genotype-by-environment interactions, as have been observed for some other skeletal traits (Middleton *et al.*, 2008a). We used a precise, highly repeatable method to measure the three-dimensional (3D) surface area of muscle attachment sites that was first introduced in a study of human hand entheses (Karakostis & Lorenzo, 2016). Since then, this method, named "Validated Entheses-based Reconstruction of Activity" (V.E.R.A.) (Karakostis & Harvati, 2021), has been applied successfully in other experimental studies using small mammals and birds (Karakostis *et al.*, 2019a; b). Our large sample size, which is substantially greater than in other experimental studies of muscle entheses, allows us to reliably apply robust linear statistical models and probability testing.

We had several hypotheses regarding the effects of exercise and genetics on muscle attachment morphology. 1: HR mice will have evolved larger (increased surface area) muscle attachment sites, which would reduce stress acting on the muscle insertion

sites, a lower-lever trait that may be beneficial for endurance running. 2: HR and C Mice housed with wheels throughout ontogeny will have enlarged muscle attachments due to the dynamic loads experienced when running (cf., Roach *et al.*, 2012). 3: HR mice may have altered phenotypic plasticity, which would be detected as a statistical interaction between the main effects of linetype and activity housing condition (i.e., a genotype-by-environment interaction). Hypothesis 3 implies comparing the alternatives of "more pain, more gain" versus the "principle of initial value." In the former, one expects a greater amount of training effort (e.g., greater daily wheel-running distance) to be associated with a stronger training response. Given that HR mice run more than C mice, one might generally expect to find larger training effects in the former, regardless of the trait in question (e.g., see Garland & Kelly, 2006). On the other hand, the principle of initial value expects an inverse relationship between the initial value of a trait and the magnitude of response to training (Koch *et al.*, 2005; Middleton *et al.*, 2008a). Thus, if mice from the HR lines were to have innately larger muscle entheses (i.e., without training), then we would expect any increase caused by training to be blunted as compared with the effect observed for mice from the non-selected Control lines.

Methods

High Runner mouse model

Mice from the 4 High Runner (HR) lines are bred for high voluntary wheel running and are compared with 4 non-selected Control (C) lines (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, Indianapolis, Indiana, USA). Mice were randomly bred for two generations and then separated into 8 closed lines, which consist of 10 breeding pairs per line per generation. 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 6-8 weeks of age. 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 (Swallow *et al.*, 1998; Careau *et al.*, 2013; Hiramatsu, 2017). In the HR lines, the highest-running (average revolutions during days 5 and 6 of a 6-day trial) male and female from each family are chosen as breeders. Running on days 5 and 6 is used as the selection criterion 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. Mice are kept at room temperatures of approximately 22°C, with ad lib access to food and water. Photoperiod is 12L:12D.

Mouse specimens, physical activity, and wheel access

We studied 100 female mice (evenly sampled from the 8 lines except for line 6: see Section 2.3) from generation 57 of the selection experiment (Copes *et al.*, 2015,

2018). We chose female mice to remove the confounding effects of skeletal sexual dimorphism (Nieves *et al.*, 2004; Castro & Garland, 2018) and because female mice generally run more revolutions per day and at higher average and maximum speeds in our study system. Briefly, mice were housed individually beginning at weaning (21 days of age), with food and water ad lib. At approximately 24 days of age, half of the mice were given access to wheels (attached to their cages as described above) for 12 weeks (see Figure 1). Furthermore, physical activity measures within the home cage were recorded daily using passive infrared motion detection sensors over a 23.5 hour period (Copes *et al.*, 2015, 2018). Across the 12 weeks of either wheel access (Active group) or being housed without wheels (Sedentary), 3 individual mice died of natural causes and so were not included in our study. Therefore, our sample consisted of 25 Sedentary C mice, 23 Sedentary HR mice, 24 Active C mice, and 25 Active HR mice (Figure 1). None of the mice in this study were used as breeders for the selection experiment. All experiments were approved by the University of California, Riverside Institutional Animal Care and Use Committee.

The wheel running and home-cage activity data sampled over 12 weeks are presented and analyzed in Copes et al. (2018). As expected, HR mice ran significantly further than C mice each week, due to both longer duration of running and higher average running speeds. Mice with wheel access had lower home-cage activity than Sedentary mice. Sedentary HR mice had significantly higher levels of home-cage activity than Sedentary C mice throughout the course of the experiment, due to moving during more

intervals each day. Among Active mice, linetype did not have a significant effect on home-cage activity measures (Copes *et al.*, 2015, 2018; Lewton *et al.*, 2019).

Dissection and bone preparation

At 15 weeks of age, 97 mice were euthanized, weighed, and dissected for tissues (Figure 3.1). The mass of the triceps surae muscle was used to identify individual mice with the mini-muscle phenotype (Garland *et al.*, 2002). In our selection experiment, the "mini-muscle" phenotype occurred in a subset of the mice, characterized by 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 is caused by a novel intronic single nucleotide polymorphism in the *Myosin heavy polypeptide 4* gene (Kelly *et al.*, 2013) that behaves as a Mendelian recessive allele. This allele was present in the starting population of 224 mice at a frequency of ~7%, and population-genetic modeling indicates that the mini-muscle phenotype was (unintentionally) under positive selection in the HR lines (Garland *et al.*, 2002). The mini-muscle phenotype eventually became fixed in one HR line, but remains polymorphic in another. In our sample of 97 mice (not all of which had data for all traits), the number of mini-muscle individuals was all 11 in HR line #3 and 5 of 11 in HR line #6. After dissection of the triceps surae muscles, mice were disembowled and the carcasses were soaked in a 1% solution of enzymatic detergent (Tergazyme) to dissolve flesh from bone (Copes *et al.*, 2018; Selvey *et al.*, 2018) (Figure 1).

Selection of muscle entheses

For this study, we analyzed muscle entheses on the humerus and femur because they are the largest of the mouse long bones, with a considerable amount of attached muscle mass that is activated during exercise (Benjamin *et al.*, 2002; Bab *et al.*, 2007; Charles *et al.*, 2016). Previous studies of mice and rats have routinely shown that long bones respond to mechanical loading, often achieved through exercise (Newhall *et al.*, 1991; Mori *et al.*, 2003; Yang *et al.*, 2007; Plochocki *et al.*, 2008), including in the HR and C mice (e.g., see Kelly *et al.*, 2006; Middleton *et al.*, 2008b; Young *et al.*, 2009; Wallace *et al.*, 2012). Furthermore, the few experimental studies that have investigated the effects of exercise on muscle attachment site morphology have included those found on long bones (Zumwalt, 2006; Rabey *et al.*, 2015; Wallace *et al.*, 2017; Karakostis *et al.*, 2019a; b). However, muscles that attach on limb bones can have different roles during exercise (e.g., the quadratus femoris muscle functions to stabilize the hip and to counter the medial rotation generated by the gluteal muscles during extension) and may not be directly involved with load bearing per se.

When choosing which muscle attachments to analyze (especially considering the use of dry bone specimens), certain criteria were deemed necessary for measurements to be taken, including: 1) the enthesis must be clearly defined and homologous across specimens (e.g., see Bab *et al.*, 2007; Karakostis *et al.*, 2018); 2) the enthesis must not be damaged across multiple specimens (e.g., the femoral distal condyles frequently broke off); and 3) the muscles that attach on the enthesis must be clearly defined in terms of function and morphology in mice (see below). The muscle attachment sites that met

these criteria for the femur include the femoral lesser trochanter, which serves as the insertion point for the iliacus muscle (origin is on the iliac crest of the pelvis) and functions to flex the hip joint (activated during the swing phase), the femoral third trochanter which serves the insertion point for the quadratus femoris muscle (origin is on the pubis bone of the pelvis) and functions to stabilize and rotate the hip joint laterally, and the femoral greater trochanter, which serves as the insertion point for the middle gluteal muscles (origin is on the lateral aspect of iliac crest) and functions to extend the hip joint (activated during the stance phase and providing propulsion) (Charles *et al.*, 2016). On the humerus, the humeral deltoid tuberosity is a clearly defined, prominent ridge that serves as the insertion point for the spinodeltoideus, superficial pectoralis, and acromiodeltoideus muscles, with the spinedeltoideus inserting along most of the lateral surface and the superficial pectoralis and acromiodeltoideus inserting along the medial surface (Bab *et al.*, 2007; Rabey *et al.*, 2015). Although the deltoid muscles (spinedeltoideus and acromiodeltoideus) function primarily as shoulder extensors (activated during the swing phase), the pectoralis muscle (superficial pectoralis) functions during retraction, bringing the mouse forward during locomotion on the supporting limb. All the muscles considered are either involved in weight bearing, stabilization, rotation, retraction and/or protraction during mouse locomotor behavior (Clarke & Still, 1999).

μCT scanning and image segmentation

As described by Copes et al. (2018), the right femur and humerus were µCT scanned at 12-µm resolution using a small animal preclinical microtomography scanner Viva-CT40, Scanco Medical AG, (Basserdorf, Switzerland) housed at the University of

Calgary, Calgary, Alberta. For each specimen, the raw data were reconstructed as 16-bit TIFF image sequential stacks using ImageJ software (Schneider *et al.*, 2012). Image stacks were imported into Thermo Scientific AMIRA 5.6 Software, Thermo Fisher Scientific (Waltham, Massachusetts, U.S.A) for visualization and segmentation. The *Isosurface* module was used to create surface renderings of the humerus and femur to examine the external morphology of the muscle entheses. Next, an orthographic perspective (objects are displayed proportional to their true size) was used and each bone was virtually re-oriented along its long axis using *Align Principal Axes*. Afterwards, the *OrthoSlice* module was used to segment out the total bone area for each bone and 3D surface models were created and exported as .stl files using the *Label Field* module (cf., Schwartz *et al.*, 2018).

3D reconstruction of muscle entheses

The 3D surface models of the right humerus and femur were separately imported and rendered into MeshLab (CNR-INC, Rome, Italy). Following the V.E.R.A. method (Karakostis & Harvati, 2021), four entheseal surfaces were delineated using image filtering techniques based on surface elevation, coloration, and surface complexity (Karakostis & Lorenzo, 2016; Karakostis *et al.*, 2018, 2019a; b, 2021; Karakostis & Harvati, 2021). In this study, we provide a pipeline figure of the V.E.R.A method based on the humerus deltoid tuberosity (Figure 3.2). Furthermore, we illustrate the delineated models of all the muscle attachments described above (see Selection of muscle entheses) (Figure 3.3). In this study, we could not include the "*equalize vertex colors*" filter because our scans lacked color information and primarily focused on the presence of

distinctive surface elevation (i.e., projecting, or depressed bone surface) and irregularities (Karakostis & Lorenzo, 2016; Karakostis *et al.*, 2021) (Figure 3.2A). Previous inter-method tests have confirmed the precise applicability of this methodology on 3D scans lacking color information (Karakostis *et al.*, 2018). For each of the muscle attachments, we applied the "*Discrete Curvatures*" filter, a surface curvature filter in MeshLab that color-maps the 3D surface of the bone depending on its elevation and irregularity (Figure 3.2B). After this, the "*Z-Painting*" tool was used to model the general boundary of the muscle attachment (Figure 3.2C) (including a flatter region circling it) (Figure 3.2.D) followed by using the "*Invert Selection*" filter to crop it from the rest of the whole bone 3D surface model. The "*Curvature Principal Directions*" filter was applied on the individual muscle attachment models (Figure 3.2E) to separate out the flatter area (shown in dark blue) surrounding the entheseal surface (Figure 3.2F). Before quantifying the delineated muscle attachment 3D surface models, we transformed the scale on MeshLab based on linear measurements from the whole bone 3D surface models. Finally, the "*Compute Geometric Measurements*" filter was used to measure the external surface area of the muscle attachments in mm^2 (Karakostis & Lorenzo, 2016; Karakostis *et al.*, 2018, 2019a; b, 2021; Karakostis & Harvati, 2021) (Figure 3.3).

Repeatability of measurements

All measurements were blind with respect to activity, linetype, and mini-muscle status. Prior to taking measurements from the segmented 3D models, we tested repeatability across a sub-sample of 10 specimens to account for intra- and inter-observer measurement error, akin to other studies using the V.E.R.A. approach, which found that

the maximum mean precision error was 0.62% (see Karakostis & Lorenzo, 2016). The repeatability test used to check inter-observer error involved only two observers (FAK and AAC). Furthermore, each specimen was measured twice, and both observers applied the V.E.R.A method on these specimens over two days. More specifically, we used a repeated-measures ANOVA for the inter-observer measurements and non-parametric paired tests (Wilcoxon) across all possible repetitions. In all scenarios, the p values were >0.05, indicating repeatability. In our study, AAC delineated and quantified all the muscle attachments using the V.E.R.A. approach. In addition, we did not measure any individual muscle enthesis that were damaged (e.g., while being prepped using Tergazyme) and/or had a significant amount of osteoporosis (e.g., see Appendix 3.1), which resulted in variable sample sizes for both univariate and multivariate analyses. Final samples sizes are indicated in Table 3.1.

Statistical analysis

We used the MIXED Procedure in SAS (SAS Institute, Cary, NC, USA) to apply nested analysis of covariance models with replicate line as a random effect nested within linetype, yielding 1 and 6 d.f. for testing the effect of linetype (Swallow *et al.*, 1999; Houle-Leroy *et al.*, 2000, 2003). Likewise, the main effects of activity (wheel access) and the interaction between activity and linetype were tested with 1 and 6 d.f. (see Copes *et al.*, 2018; Lewton *et al.*, 2019). The main effect of the mini-muscle phenotype (see Methods) was included and tested relative to the residual variance with 1 and ~62-80 d.f. (depending on the muscle enthesis). All analyses of muscle entheses included body mass (recorded at dissections) as a covariate.

In addition, we used measures of physical activity as covariates. We went back to the primary paper for this data set (Copes et al. 2018) and used the weekly average activity variables, as reported in their supplemental materials (e.g., wk8run6, wk8htot6). We then summed these variables across all 12 weeks to obtain a measure of the total "volume" of wheel-running distance and of spontaneous physical activity in the home cages. We used these as covariates (along with body mass) for each of the four attachment measures. We analyzed the sedentary and active mice separately, the former using only cage activity.

For multivariate analyses, the four entheseal surface area measurements were subjected to principal components analysis (PCA) that included body mass (5 principal components total). PCA is an exploratory technique used to reveal multivariate patterns of variation in a sample, without using *a priori* group classification (Field, 2013). Scores on the five PCs were subjected to the same analyses described above for the individual measures. Sample sizes were reduced because computation of PC scores requires all measurements for a given mouse (i.e., listwise deletion of missing data).

In all analyses (univariate or multivariate), outliers were removed when the standardized residual exceeded ~3.0 and we used an α of ≤ 0.05 for statistical significance. For univariate analyses, two low outliers were removed for the humerus deltoid tuberosity (MouseID = 60653 and 60625) and one low outlier for the femoral third trochanter (MouseID = 60714). For multivariate analyses, one low outlier was removed from PC1 (MouseID = 60235). All p values reported are two-tailed (Table 3.1).

Results

Significance levels from ANCOVAs of muscle entheses (using body mass as a covariate) and from ANOVAs using PC scores are presented in Table 3.1. Table 3.2 presents least square means and Table 3.3 shows the results of a PCA of body mass and the four attachment areas.

Body size

Body mass was not significantly different when comparing HR vs C mice, active vs sedentary mice, or mini-muscle mice vs normal-muscled individuals (Table 3.1). However, HR mice tended to be lighter ($p=0.0938$) than C mice and active mice tended to weigh less than sedentary mice ($p=0.0922$).

Muscle attachments

Figures 2 and 3A show a representative deltoid tuberosity image. With body mass as a covariate, mice from the active group had significantly larger humerus deltoid tuberosities (increased surface areas) than sedentary mice ($p=0.0346$; Figure 3.4), with no significant linetype effect or interaction between activity and linetype (see Table 3.1 for full statistical results and Table 3.2 for least squares means for the experimental groups). Mice with the mini-muscle phenotype had significantly smaller femoral third trochanters ($p=0.0119$) when compared with normal-muscled individuals (Figure 3.5), with no significant linetype or activity effect. Morphology of the femoral lesser trochanter and femoral greater trochanter were not affected by chronic wheel access and did not differ between HR and C mice (Figure 3.5).

For the mice with wheel access, neither measure of physical activity was ever a statistically significant predictor of attachment surface area. Similarly, for the mice without wheels, home-cage activity was ever a significant predictor of attachment surface area.

Principal components analysis

Table 3.3 shows the results of a PCA of body mass and the four attachment areas. PC 1 accounted for 40.6% of the total variance and mainly reflected humerus deltoid tuberosity surface area and body mass (loading strongly in the same direction). PC 2 accounted for 28.5% of the variance that contrasted the femoral lesser trochanter and femoral greater trochanter surface areas with the femoral third trochanter surface area. PC 3 (12.1% of variance) was mostly related to femoral greater trochanter surface area, PC 4 (9.9%) reflected a contrast between the deltoid tuberosity area and the other entheses, and PC 5 (8.9%) was mostly body mass and humerus deltoid tuberosity (Table 3.3).

Figure 3.6 shows scores for PC 1 differed significantly between mini- and normal-muscled mice ($p=0.0315$), whereas scores for PC 4 (9.90% of total variance) showed an effect of activity ($p=0.0211$, Table 3.3). Figure 3.6 shows scores for PC 1 against PC 4, separately for C and HR mice, illustrating the separation by mini-muscle status and activity group.

Discussion

We studied the sizes of four muscle entheses from a unique model system that includes four replicate High Runner (HR) lines of house mice that have been selectively bred for wheel-running behavior and four non-selected Control lines. For the present study, half of the experimental subjects were housed with wheels (Active group) and half without wheels (Sedentary group) for 12 weeks starting at weaning. Thus, we studied phenotypic plasticity, evolved differences, and their interaction. With body mass as a covariate, mice in the Active group had significantly larger humerus deltoid tuberosities when compared with Sedentary mice (i.e., a training effect or phenotypic plasticity). We did not find any overall differences between the HR and C lines, but the subset HR individuals with the mini-muscle phenotype had significantly reduced femoral third trochanters. Consistent with most previous studies of skeletal traits in the HR and C mice, we did not find evidence for differential training responses (i.e., any linetype and activity group interactions) (e.g., see Kelly *et al.*, 2006; Middleton *et al.*, 2008b; Wallace *et al.*, 2012). We discuss our results primarily in the context of previous studies of the causal relationship between physical activity and muscle attachment morphology.

Effects of chronic exercise: phenotypic plasticity

As reviewed in the Introduction, some of the previous experimental studies using sheep, mice, and turkeys have not found statistically significant effects of chronic exercise on muscle attachment morphology (Zumwalt, 2006; Rabey *et al.*, 2015; Wallace *et al.*, 2017). Results of those studies would not support use of muscle enthesis morphology as an indicator of physical activity levels, e.g., in fossils. However, more

recent studies that used a combination of precise 3D quantification (i.e., the V.E.R.A. protocol) and multivariate principal components analysis (rather than univariate analyses of single attachments) have found effects of physical activity on muscle attachment surface areas (see Introduction and Karakostis *et al.*, 2019a; b). In the present study, we used this highly repeatable 3D method (see Karakostis & Lorenzo, 2016; Karakostis & Harvati, 2021), and included both univariate and multivariate analyses to facilitate comparison of our results with previous studies.

The humerus deltoid tuberosity is a clearly defined, prominent ridge that serves as an insertion point for three muscles of the forelimb (spinodeltoideus, superficial pectoralis, and acromiodeltoideus muscles) that are involved in shoulder extension (spinodeltoideus and acromiodeltoideus) and retraction (superficial pectoralis) during locomotion (Rabey *et al.*, 2015 and Methods). In univariate analyses, with body mass as a covariate, mice in the Active group had significantly larger humerus deltoid tuberosities (increased surface area) when compared with Sedentary mice (Figure 3.4). In multivariate analyses, PC 4 scores, which represent variation in the deltoid tuberosity relative to the other entheses (Table 3.3), were also significantly different between Active and Sedentary mice (Table 3.1). A larger surface area of the deltoid tuberosity suggests increased surface area available to dissipate stress (force per unit area) that results from elevated muscular activity and/or increased muscle size (muscle dimensions were not quantified in this study). Thus, we provide clear evidence that physical activity levels experienced across post-weaning ontogeny can cause muscle attachment hypertrophy, although not necessarily in all attachments. In our experimental model, daily access to

large, rat-sized wheels resulted in an enlarged muscle attachment site on the forelimb, presumably induced by muscular forces required to turn the wheels during initial acceleration, when running, when decelerating, and/or when climbing/hanging in wheels.

Although we found experimental evidence of muscle attachment hypertrophy caused by voluntary exercise, only one of four (humerus deltoid tuberosity) showed such an effect. This finding is consistent with previous studies reporting that certain attachment sites were influenced much more by physical activity than others (e.g., Karakostis *et al.*, 2017, 2019a; b). These differential responses might, in part, be explained by differences in muscle function and loading during voluntary wheel running. For example, three muscles (spinodeltoideus, acromiodeltoideus, superficial pectoralis) insert on the humerus deltoid tuberosity, whereas the femoral greater, lesser, and third trochanters only have one muscle insertion. Moreover, all these muscles differ in patterns of activation and force production during locomotion (see Methods). Further, the forelimb and hindlimb in general must experience somewhat different loading patterns. Finally, the humerus and femur may inherently differ in their responsiveness to a given pattern and/or intensity of loading.

In any case, our results emphasize the point that, in studies attempting to interpret the physical activity levels of fossil specimens from muscle attachments (e.g., Hawkey & Merbs, 1995; Schlecht, 2012; Foster *et al.*, 2014; Becker, 2020; Karakostis *et al.*, 2021) special caution is necessary in the selection of entheses analyzed. For instance, previous applications of the V.E.R.A. protocols to fossil hominin hand skeletons (Karakostis *et al.*, 2018; Karakostis & Harvati, 2021) focused exclusively on entheses that have shown

consistent differences across documented humans with distinct lifelong occupational activities (Karakostis *et al.*, 2017). Similar recommendations have been made for any skeletal feature used for reconstructing activity in the past (e.g., see Lieberman *et al.*, 2004; Wallace *et al.*, 2012; Copes *et al.*, 2018; Peacock *et al.*, 2018).

Exercise throughout ontogeny vs. exercise as adults

In our study, mice were granted wheel access shortly after weaning (~24 days of age) and throughout ontogeny (12-week treatment), which includes the critical period before sexual maturity, during which bones grow and are more likely to adapt to mechanical loads caused by exercise as compared with mature mice (e.g., see Gardinier *et al.*, 2018). Similar to our study, Rabey et al. (2015) gave mice exercise (housed in caged with activity wheels or with 1 m tall wire-mesh tower) when they were either 25 or 46 days old (two cohorts). Likewise, Wallace et al. (2017) and Karakostis et al. (2019b) used 1-year old female Eastern wild turkeys (maturity occurs at ~15 months of age). However, others focused on adults: Karakostis et al. (2019a) used eight-week-old rats (maturity occurs at ~6 weeks of age) and Zumwalt (2006) used adult female sheep that were all at least four years of age (or older). Studies in rats and mice have shown that senescence leads to reductions in bone mass and mineral density, declines in maximal muscular force production, and changes in muscle properties (Ferguson *et al.*, 2003; Horner *et al.*, 2011; Holt *et al.*, 2016; Sobolev *et al.*, 2017; Hill *et al.*, 2020), all of which could affect muscle entheses. However, experiments on the effects of exercise on muscle attachment morphology that include comparisons of young, mature, and aged animals are lacking.

Effects of selective breeding for high voluntary wheel running

Selection experiments and experimental evolutionary approaches (Garland & Rose, 2009) are well-suited for study of microevolution and coadaptation of the skeleton with locomotor behavior, locomotor performance, and body size (Middleton *et al.*, 2008a; Marchini *et al.*, 2014). Here, we found no evidence of evolutionary coadaptation of muscle attachment sites with voluntary exercise behavior (HR vs. C) in mice. However, previous studies have shown several examples of skeletal changes in the HR lines of mice (Kelly *et al.*, 2006; Schutz *et al.*, 2014; Castro *et al.*, 2021). For example, by generation 11, male and female HR mice had evolved larger knee and hip surface areas (adjusting for body size), which, all else being equal, would reduce stress (i.e., force per unit area) acting on limb joints for mice running long distances on wheels (Garland & Freeman, 2005; Castro & Garland, 2018). Another study of the same specimens from generation 11 found that femurs from HR mice have larger total nutrient canal area (due to increased average cross-section size but not the number of canals) than those from C mice (Schwartz *et al.*, 2018). The fact that selection limits were not reached until ~17-27 generations suggests that further skeletal evolution was probable. Indeed, by generation 21, HR males (females were not studied) had larger femoral heads (adjusting for body size and as reported for generation 11), but also had evolved thicker femurs and tibia-fibulas (measured but not significantly different 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; Castro & Garland, 2018). However, at generation 68 few differences were found between HR and C mice, and some of the differences in bone

dimensions identified in earlier generations were no longer statistically significant (Castro *et al.*, 2021). We are not aware of any other selection experiments targeting locomotion or physical activity that have tested for correlated responses in muscle attachment sites.

As compared with other aspects of skeletal anatomy, relatively few interspecific comparisons of mammals have involved muscle attachment sites, and those that did not directly tested for relations with aspects of locomotor performance or ecology (e.g., skeletal correlates with daily movement distances or maximal running sprint speeds, Garland & Janis, 1993; Harris & Steudel, 1997; Kelly *et al.*, 2006). For example, muscle attachment sites are enlarged in the forelimbs of species of rodents and carnivorans that regularly dig and swim (Samuels & Van Valkenburgh, 2008; Samuels *et al.*, 2013) and in the forelimbs of carnivorans that specialize on larger prey (Meachen-Samuels & Van Valkenburgh, 2009). Another study compared the hindlimb muscle attachment morphology of the extinct the Santacrucian (Early Miocene) sloths with that of extant Xenarthran species and found that the former has thicker and more robust muscle attachment sites, suggesting increased muscle size and force generation capabilities for climbing behaviors, regardless of their comparatively larger body size (Toledo *et al.*, 2015). The evolutionary relationships between muscle attachment sizes and locomotor behavior or performance require further study.

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 drastic reduction of type IIb muscle fibers

(Garland *et al.*, 2002; Guderley *et al.*, 2006; Talmadge *et al.*, 2014). We found that mini-muscle mice had significantly reduced surface areas of the femoral third trochanter when compared with normal muscled mice (Table 3.1, also reflected in scores on PC 1). Likewise, at generations 11 and 68 for both sexes, mini-muscle mice had significantly thinner femoral third trochanters (Castro & Garland, 2018; Castro *et al.*, 2021). These results suggest that the quadratus femoris muscle, which functions to stabilize and rotate the hip, has reduced surface area for attachment at the insertion site, likely due to the greatly reduced hindlimb musculature.

Experimental models and future directions

Future investigations of correlations between physical activity and muscle attachment size would benefit from a deeper understanding of how muscle entheses respond to the magnitude of applied loads (e.g., see Rossetti *et al.*, 2017), as well as their frequency (for studies in long bones see Hart *et al.*, 2017; Yang *et al.*, 2017; Berman *et al.*, 2019; DeLong *et al.*, 2020). Moreover, prolonged paralysis or disuse of muscles have detrimental effects on bone mass and strength (unloading) (Morey-Holton & Globus, 1998; Kodama *et al.*, 1999; DeLong *et al.*, 2020), that may also characterize the bony response of attachments. For example, one study used 40 mature male mice to investigate the effects of unloading (achieved via paralysis of shoulder muscles for 21 days) on the mechanical properties of the humerus-supraspinatus muscle attachment, finding an increased risk of fracture and significant bone loss at the millimeter scale, as well as changes at the micro and nanometer scales (Deymier *et al.*, 2019). In addition, a preliminary study comparing muscle attachment size and ground reaction forces in CD1

wild-type, and myostatin-deficient (muscular hypertrophy) mice found that while myostatin-deficient mice had expanded muscle attachments (humerus deltoid tuberosity and femoral third trochanter), both groups experienced similar vertical ground reaction forces (Schmitt *et al.*, 2010). Thus, the relationship between muscle enthesis size and limb loading experienced during locomotion is unclear, highlighting the need for future kinematic and biomechanical studies (e.g., see Claghorn *et al.*, 2017; Sparrow *et al.*, 2017; Abraham *et al.*, 2021).

References

Abraham, A.C., Fang, F., Golman, M., Oikonomou, P. & Thomopoulos, S. 2021. The role of loading in murine models of rotator cuff disease. *J. Orthop. Res.* jor.25113.

Bab, I., Hajb-Yonissi, C., Gabet, Y. & Muller, R. 2007. *Micro-tomographic atlas of the mouse skeleton*. Springer, New York.

Becker, S.K. 2020. Osteoarthritis, entheses, and long bone cross-sectional geometry in the Andes: Usage, history, and future directions. *Int. J. Paleopathol.* **29**: 45–53.

Benjamin, M., Evans, E.J. & Copp, L. 1986. The histology of tendon attachments to bone in man. *J. Anat.* 12.

Benjamin, M., Kumai, T., Milz, S., Boszczyk, B.M., Boszczyk, A.A. & Ralphs, J.R. 2002. The skeletal attachment of tendons—tendon 'entheses.' *Comp. Biochem. Physiol. A. Mol. Integr. Physiol.* **133**: 931–945.

Berman, A.G., Hinton, M.J. & Wallace, J.M. 2019. Treadmill running and targeted tibial loading differentially improve bone mass in mice. *Bone Rep.* **10**: 100195.

Biewener, A.A. 1992. Overview of structural mechanics. In: *Biomechanics (Structures and Systems): A Practical Approach*, pp. 1–20. Oxford University Press, New York.

Bonewald, L.F. 2007. Osteocytes as dynamic multifunctional cells. *Ann. N. Y. Acad. Sci.* **1116**: 281–290.

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. *Evolution* **67**: 3102–3119.

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. *J. Morphol.* **279**: 766–779.

Castro, A.A., Rabitoy, H., Claghorn, G.C. & Garland, T. 2021. Rapid and longer-term effects of selective breeding for voluntary exercise behavior on skeletal morphology in house mice. *J. Anat.* **238**: 720–742.

Charles, J.P., Cappellari, O., Spence, A.J., Wells, D.J. & Hutchinson, J.R. 2016. Muscle moment arms and sensitivity analysis of a mouse hindlimb musculoskeletal model. *J. Anat.* **229**: 514–535.

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. *Physiol. Biochem. Zool.* **90**: 533–545.

Clarke, K. & Still, J. 1999. Gait analysis in the mouse. *Physiol. Behav.* **66**: 723–729.

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

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. *Am. J. Phys. Anthropol.* **166**: 442–458.

DeLong, A., Friedman, M.A., Tucker, S.M., Krause, A.R., Kunselman, A., Donahue, H.J., *et al.* 2020. Protective effects of controlled mechanical loading of bone in C57BL6/J mice subject to disuse. *JBMR Plus* **4**.

Deymier, A.C., An, Y., Boyle, J.J., Schwartz, A.G., Birman, V., Genin, G.M., *et al.* 2017. Micro-mechanical properties of the tendon-to-bone attachment. *Acta Biomater.* **56**: 25–35.

Deymier, A.C., Schwartz, A.G., Cai, Z., Daulton, T.L., Pasteris, J.D., Genin, G.M., *et al.* 2019. The multiscale structural and mechanical effects of mouse supraspinatus muscle unloading on the mature enthesis. *Acta Biomater.* **83**: 302–313.

Doherty, A.H., Ghalambor, C.K. & Donahue, S.W. 2015. Evolutionary physiology of bone: Bone metabolism in changing environments. *Physiology* **30**: 17–29.

Eliakim, A., Raisz, L.G., Brasel, J.A. & Cooper, D.M. 1997. Evidence for increased bone formation following a brief endurance-type training intervention in adolescent males. *J. Bone Miner. Res.* **12**: 1708–1713.

Ferguson, V.L., Ayers, R.A., Bateman, T.A. & Simske, S.J. 2003. Bone development and age-related bone loss in male C57BL/6J mice. *Bone* **33**: 387–398.

Field, A. 2013. *Discovering Statistics Using SPSS*, 4th ed. SAGE, London, UK.

Foster, A., Buckley, H. & Tayles, N. 2014. Using enthesis robusticity to infer activity in the past: A review. *J. Archaeol. Method Theory* **21**: 511–533.

Frost, H.M. 2003. Bone's mechanostat: A 2003 update. *Anat. Rec.* **275A**: 1081–1101.

Gardinier, J.D., Rostami, N., Juliano, L. & Zhang, C. 2018. Bone adaptation in response to treadmill exercise in young and adult mice. *Bone Rep.* **8**: 29–37.

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? *J. Zool.* **229**: 133–151.

Garland Jr., T. & Kelly, S.A. 2006. Phenotypic plasticity and experimental evolution. *J. Exp. Biol.* **209**: 2344–2361.

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. *Proc. R. Soc. B Biol. Sci.* **278**: 574–581.

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.

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

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. *Comp. Biochem. Physiol. B Biochem. Mol. Biol.* **144**: 271–282.

Harris, M.A. & Steudel, K. 1997. Ecological correlates of hind-limb length in the Carnivora. *J. Zool.* **241**: 381–408.

Hart, N.H., Nimphius, S., Rantalainen, T., Ireland, A., Siafarikas, A. & Newton, R.U. 2017. Mechanical basis of bone strength: Influence of bone material, bone structure and muscle action. *J Musculoskelet Neuronal Interact* **17**: 114–139.

Hawkey, D.E. & Merbs, C.F. 1995. Activity-induced musculoskeletal stress markers (MSM) and subsistence strategy changes among ancient Hudson Bay Eskimos. *Int. J. Osteoarchaeol.* **5**: 324–338.

Hill, C., James, R.S., Cox, Val.M., Seebacher, F. & Tallis, J. 2020. Age-related changes in isolated mouse skeletal muscle function are dependent on sex, muscle, and contractility mode. *Am. J. Physiol.-Regul. Integr. Comp. Physiol.* **319**: R296–R314.

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

Holt, N.C., Danos, N., Roberts, T.J. & Azizi, E. 2016. Stuck in gear: Age-related loss of variable gearing in skeletal muscle. *J. Exp. Biol.* **219**: 998–1003.

Horner, A.M., Russ, D.W. & Biknevicius, A.R. 2011. Effects of early-stage aging on locomotor dynamics and hindlimb muscle force production in the rat. *J. Exp. Biol.* **214**: 3588–3595.

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*. *J. Appl. Physiol.* **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. *Am. J. Physiol.-Regul. Integr. Comp. Physiol.* **284**: R433–R443.

Karakostis, F.A. & Harvati, K. 2021. New horizons in reconstructing past human behavior: Introducing the “Tübingen University Validated Entheses-based Reconstruction of Activity” method. *Evol. Anthropol. Issues News Rev.* evan.21892.

Karakostis, F.A., Hotz, G., Scherf, H., Wahl, J. & Harvati, K. 2018. A repeatable geometric morphometric approach to the analysis of hand entheseal three-dimensional form. *Am. J. Phys. Anthropol.* **166**: 246–260.

Karakostis, F.A., Hotz, G., Scherf, H., Wahl, J. & Harvati, K. 2017. Occupational manual activity is reflected on the patterns among hand entheses. *Am. J. Phys. Anthropol.* **164**: 30–40.

Karakostis, F.A., Jeffery, N. & Harvati, K. 2019a. Experimental proof that multivariate patterns among muscle attachments (entheses) can reflect repetitive muscle use. *Sci. Rep.* **9**: 16577.

Karakostis, F.A. & Lorenzo, C. 2016. Morphometric patterns among the 3D surface areas of human hand entheses. *Am. J. Phys. Anthropol.* **160**: 694–707.

Karakostis, F.A., Reyes-Centeno, H., Franken, M., Hotz, G., Rademaker, K. & Harvati, K. 2021. Biocultural evidence of precise manual activities in an Early Holocene individual of the high-altitude Peruvian Andes. *Am. J. Phys. Anthropol.* **174**: 35–48.

Karakostis, F.A., Wallace, I.J., Konow, N. & Harvati, K. 2019b. Experimental evidence that physical activity affects the multivariate associations among muscle attachments (entheses). *J. Exp. Biol.* **222**: jeb213058.

Katsimbri, P. 2017. The biology of normal bone remodelling. *Eur. J. Cancer Care (Engl.)* **26**: e12740.

Kelly, S.A., Bell, T.A., Selitsky, S.R., Buus, R.J., Hua, K., Weinstock, G.M., *et al.* 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. *J. Morphol.* **267**: 360–374.

Koch, L.G., Green, C.L., Lee, A.D., Hornyak, J.E., Cicila, G.T. & Britton, S.L. 2005. Test of the principle of initial value in rat genetic models of exercise capacity. *Am. J. Physiol.-Regul. Integr. Comp. Physiol.* **288**: R466–R472.

Kodama, Y., Dimai, H.P., Wergedal, J., Sheng, M., Malpe, R., Kutilek, S., *et al.* 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.

Lewton, K.L., Ritzman, T., Copes, L.E., Garland, T. & Capellini, T.D. 2019. Exercise-induced loading increases ilium cortical area in a selectively bred mouse model. *Am. J. Phys. Anthropol.* **168**: 543–551.

Lieberman, D.E. 2003. Optimization of bone growth and remodeling in response to loading in tapered mammalian limbs. *J. Exp. Biol.* **206**: 3125–3138.

Lieberman, D.E., Polk, J.D. & Demes, B. 2004. Predicting long bone loading from cross-sectional geometry. *Am. J. Phys. Anthropol.* **123**: 156–171.

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. **14**: 15.

Maupin, K.A., Childress, P., Brinker, A., Khan, F., Abeysekera, I., Aguilar, I.N., *et al.* 2019. Skeletal adaptations in young male mice after 4 weeks aboard the International Space Station. *Npj Microgravity* **5**: 21.

Meachen-Samuels, J. & Van Valkenburgh, B. 2009. Forelimb indicators of prey-size preference in the Felidae. *J. Morphol.* **270**: 729–744.

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. *Integr. Comp. Biol.* **48**: 394–410.

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.

Morey-Holton, E.R. & Globus, R.K. 1998. Hindlimb unloading of growing rats: A model for predicting skeletal changes during space flight. *Bone* **22**: 83S-88S.

Mori, T., Okimoto, N., Sakai, A., Okazaki, Y., Nakura, N., Notomi, T., *et al.* 2003. Climbing exercise increases bone mass and trabecular bone turnover through transient regulation of marrow osteogenic and osteoclastogenic potentials in mice. *J. Bone Miner. Res.* **18**: 2002–2009.

Newhall, K.M., Rodnick, K.J., van der Meulen, M.C., Carter, D.R. & Marcus, R. 1991. Effects of voluntary exercise on bone mineral content in rats. *J. Bone Miner. Res.* **6**: 289–296.

Nieves, J.W., Formica, C., Ruffing, J., Zion, M., Garrett, P., Lindsay, R., *et al.* 2004. Males have larger skeletal size and bone mass than females, despite comparable body size. *J. Bone Miner. Res.* **20**: 529–535.

Notomi, T., Okimoto, N., Okazaki, Y., Tanaka, Y., Nakamura, T. & Suzuki, M. 2001. Effects of tower climbing exercise on bone mass, strength, and turnover in growing rats. *J. Bone Miner. Res.* **16**: 166–174.

Peacock, S.J., Coats, B.R., Kirkland, J.K., Tanner, C.A., Garland, T. & Middleton, K.M. 2018. Predicting the bending properties of long bones: Insights from an experimental mouse model. *Am. J. Phys. Anthropol.* **165**: 457–470.

Plochocki, J.H., Rivera, J.P., Zhang, C. & Ebba, S.A. 2008. Bone modeling response to voluntary exercise in the hindlimb of mice. *J. Morphol.* **269**: 313–318.

Rabey, K.N., Green, D.J., Taylor, A.B., Begun, D.R., Richmond, B.G. & McFarlin, S.C. 2015. Locomotor activity influences muscle architecture and bone growth but not muscle attachment site morphology. *J. Hum. Evol.* **78**: 91–102.

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. *J. Appl. Physiol.* **113**: 627–635.

Rossetti, L., Kuntz, L.A., Kunold, E., Schock, J., Müller, K.W., Grabmayr, H., *et al.* 2017. The microstructure and micromechanics of the tendon–bone insertion. *Nat. Mater.* **16**: 664–670.

Rubin, C.T. & Lanyon, L.E. 1984. Regulation of bone formation by applied dynamic loads. *J. Bone Jt. Surg.* **66**: 6.

Ruff, C., Holt, B. & Trinkaus, E. 2006. Who's afraid of the big bad Wolff?: "Wolff's law" and bone functional adaptation. *Am. J. Phys. Anthropol.* **129**: 484–498.

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

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

Schlecht, S.H. 2012. Understanding entheses: Bridging the gap between clinical and anthropological perspectives. *Anat. Rec. Adv. Integr. Anat. Evol. Biol.* **295**: 1239–1251.

Schmitt, D., Zumwalt, A.C. & Hamrick, M.W. 2010. The relationship between bone mechanical properties and ground reaction forces in normal and hypermuscular mice. *J. Exp. Zool. Part Ecol. Genet. Physiol.* **313A**: 339–351.

Schneider, C.A., Rasband, W.S. & Eliceiri, K.W. 2012. NIH Image to ImageJ: 25 years of image analysis. *Nat. Methods* **9**: 671–675.

Schutz, H., Jamniczky, H.A., Hallgrímsson, B. & Garland, 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.

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. *J. Anat.* **233**: 193–203.

Selvey, H., Doll, A. & Stephenson, J. 2018. Exploring skeletal preparation techniques: Recuration of Botswana Mammals from a 1969 expedition using Tergazyme. *Biodivers. Inf. Sci. Stand.* **2**: e26185.

Sobolev, A.P., Mannina, L., Costanzo, M., Cisterna, B., Malatesta, M. & Zancanaro, C. 2017. Age-related changes in skeletal muscle composition: A pilot nuclear magnetic resonance spectroscopy study in mice. *Exp. Gerontol.* **92**: 23–27.

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.

Swallow, J.G., Carter, P.A. & Garland Jr., T. 1998. Artificial selection for increased wheel-running behavior in house mice. *Behav. Genet.* **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. *J. Exp. Biol.* **202**: 2513–2520.

Talmadge, R.J., Acosta, W. & Garland, T. 2014. Myosin heavy chain isoform expression in adult and juvenile mini-muscle mice bred for high-voluntary wheel running. *Mech. Dev.* **134**: 16–30.

Toledo, N., Bargo, M.S. & Vizcaíno, S.F. 2015. Muscular reconstruction and functional morphology of the hind limb of santacrucian (Early Miocene) sloths (Xenarthra, Folivora) of Patagonia: Santacrucian Sloths Hind Limb Functional Morphology. *Anat. Rec.* **298**: 842–864.

Wallace, I.J. & Garland Jr., T. 2016. Mobility as an emergent property of biological organization: Insights from experimental evolution: Mobility and biological organization. *Evol. Anthropol. Issues News Rev.* **25**: 98–104.

Wallace, I.J., Judex, S. & Demes, B. 2015. Effects of load-bearing exercise on skeletal structure and mechanics differ between outbred populations of mice. *Bone* **72**: 1–8.

Wallace, I.J., Middleton, K.M., Lublinsky, S., Kelly, S.A., Judex, S., Garland Jr., T., *et al.* 2010. Functional significance of genetic variation underlying limb bone diaphyseal structure. *Am. J. Phys. Anthropol.* **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. *Am. J. Phys. Anthropol.* **148**: 24–35.

Wallace, I.J., Winchester, J.M., Su, A., Boyer, D.M. & Konow, N. 2017. Physical activity alters limb bone structure but not entheseal morphology. *J. Hum. Evol.* **107**: 14–18.

Yang, H., Embry, R.E. & Main, R.P. 2017. Effects of loading duration and short rest insertion on cancellous and cortical Bone adaptation in the mouse tibia. *PLOS ONE* **12**: e0169519.

Yang, L., Zhang, P., Liu, S., Samala, P.R., Su, M. & Yokota, H. 2007. Measurement of strain distributions in mouse femora with 3D-digital speckle pattern interferometry. *Opt. Lasers Eng.* **45**: 843–851.

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. *Evol. Biol.* **36**: 88.

Yu, K., Sellman, D.P., Bahraini, A., Hagan, M.L., Elsherbini, A., Vanpelt, K.T., *et al.* 2017. Mechanical loading disrupts osteocyte plasma membranes which initiates mechanosensation events in bone. *J. Orthop. Res.*, doi: 10.1002/jor.23665.

Zumwalt, A. 2006. The effect of endurance exercise on the morphology of muscle attachment sites. *J. Exp. Biol.* **209**: 444–454.

Table 3.1 Significance levels (p values; bold indicates p<0.05, unadjusted for multiple comparisons) from two-way nested analysis of covariance models implemented in SAS PROC MIXED. Signs following p values indicate direction of effect: + indicates HR lines > C or active > sedentary mice or mini > than non-mini.

Trait	**N**	**Linetype**	**Activity**	**Activity*Linetype**	**Mini-Muscle**	**Body Mass**
Degrees of Freedom		1, 6	1, 6	1, 6	1, ~62-80	1, ~62-68
Body Mass	97	*0.0938-*	*0.0922-*	0.5230	0.4307-	
Humerus Deltoid Tuberosity (mm^2)	80	0.3705-	**0.0346+**	0.9894	0.3045-	**0.0088**
Femoral Third Trochanter (mm^2)	80	0.3478-	0.3818-	0.4567	**0.0073-**	**0.0358**
Femoral Lesser Trochanter (mm^2)	86	0.9791-	0.8989-	0.4195	0.5155+	0.2144
Femoral Greater Trochanter (mm^2)	91	0.7878+	0.4467+	0.6389	0.7576+	**0.0302**
PC 1	64	0.4275	0.3102	0.1213	**0.0315-**	
PC 2	65	0.3283	0.6678	0.9300	0.0927	
PC 3	65	0.2731	0.2406	0.3385	0.5004	

PC 4	65	0.5523	**0.0212-**	0.6398	0.8259
PC 5	65	0.5475	0.2148	0.5231	0.1797

Table 3.2 Least-squares means and standard errors from SAS PROC MIXED, corresponding to statistical tests presented in Table 3.1.

Trait	**Control Lines**				**High-Runner Lines**				**Mini-Muscle**			
	Sedentary		Active		Sedentary		Active		Normal		Mini	
	Mean	*SE*	Mean	*SE*	Mean	*SE*	Mean	*SE*	Mean	*SE*	Mean	*SE*
Body Mass (g)	28.24	1.28	26.93	1.28	24.79	1.13	24.14	1.13	26.53	0.78	25.52	1.28
Humerus Deltoid Tuberosity (mm^2)	6.71	0.28	7.15	0.29	6.39	0.24	6.83	0.24	6.93	0.16	6.61	0.29
Femoral Third Trochanter (mm^2)	2.94	0.19	2.71	0.18	2.63	0.15	2.61	0.15	3.04	0.08	2.41	0.20
Femoral Lesser Trochanter (mm^2)	1.12	0.11	1.16	0.09	1.17	0.09	1.11	0.09	1.09	0.06	1.18	0.12
Femoral Greater Trochanter (mm^2)	3.53	0.20	3.58	0.19	3.52	0.16	3.71	0.16	3.54	0.09	3.62	0.20

Table 3.3 Principal components analysis of body mass and the four muscle attachment areas along with component correlations with principal components (factor loadings). Scores on all five PCs were analyzed statistically (see Table 3.1).

PC	*Eigenvalue*	*% of Variance*	*Cumulative %*
1	2.03	40.6	40.6
2	1.42	28.5	69.1
3	0.61	12.1	81.2
4	0.49	9.9	91.1
5	0.45	8.9	100.0

Trait	*PC 1*	*PC 2*	*PC 3*	*PC 4*	*PC 5*
Body Mass	0.775	-0.267	0.368	-0.104	-0.427
Humerus Deltoid Tuberosity (mm^2)	0.832	0.002	-0.014	-0.378	0.406
Femoral Third Trochanter (mm^2)	0.514	-0.686	-0.106	0.476	0.167
Femoral Lesser Trochanter (mm^2)	0.311	0.789	0.391	0.328	0.144
Femoral Greater Trochanter (mm^2)	0.613	0.509	-0.555	0.080	-0.223

Figure 3.1. Experimental Design. Timeline and experimental design for mouse specimens used in this study.

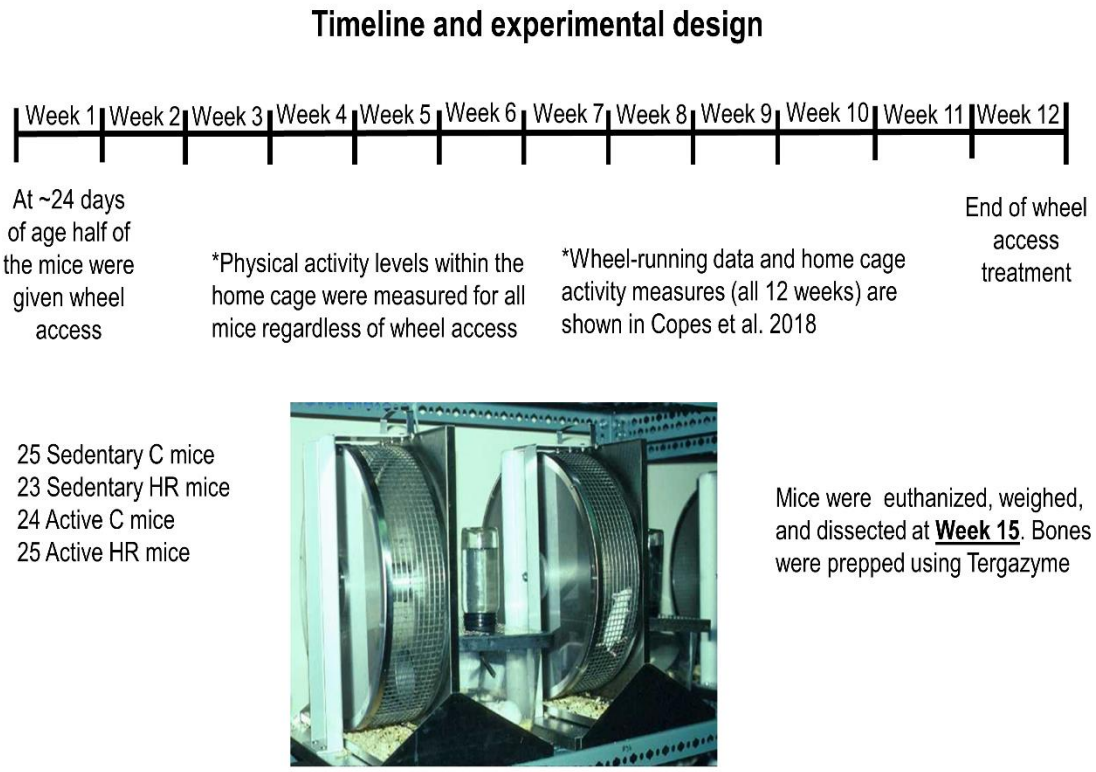


Figure 3.2. V.E.R.A. Illustration of the applied V.E.R.A. 3D measuring protocols (Karakostis & Lorenzo, 2016; Karakostis & Harvati, 2021). A: 3D surface models of the right humerus bone (shown in left) and the right humerus deltoid tuberosity (shown in right) in the caudal-lateral view using MeshLab. Color information was not available for these scans. B: Application of the "Discrete Curvatures" filter, a surface curvature filter that color-mapped the 3D surface of the right humerus deltoid tuberosity depending on its elevation and irregularity. 3D surface models of the right humerus deltoid tuberosity are examined over 360° and are depicted in the caudal-lateral (shown in right) and ventral-lateral (shown in left) views. C: Selection of the right humerus deltoid tuberosity's border and interior surface based on the application of the "Discrete Curvatures" filter (elevation and irregularities are shown in red and blue pigmentation for these specimens). 3D surface models of the right humerus deltoid tuberosity are examined over 360° and are depicted in the caudal-lateral (shown in right) and ventral-medial (shown in left) views. D: Additional selection of a very thin flat region (red arrow) around the right humerus deltoid tuberosity (muscle attachment: see E). The 3D surface model of the right humerus deltoid tuberosity was examined over 360° and is depicted in the caudal lateral view. Subsequently, the area selection was inverted to the rest of the bone surface (using the "Invert Selection" option), which was removed. E: Color-mapping of the muscle attachment surface based on the application of the "Curvature Principal Directions" filter (followed by selecting "Principal Component Analysis"), a filter that highlights the surrounding flat region (shown in dark blue) of the right humerus deltoid tuberosity based on the principal direction of curvature. The 3D surface model of the right humerus deltoid tuberosity was examined over 360° and is depicted in the caudal lateral view. F: Removal of the flatter area surrounding the right humerus deltoid tuberosity (see Figure 1E). The "Compute Geometric Measurements" filter was used to measure the external surface area of the muscle entheses in mm^2 (shown in the left). The 3D surface model of the right humerus deltoid tuberosity was the superimposed on the bone model (A) for final verification (shown in the right).

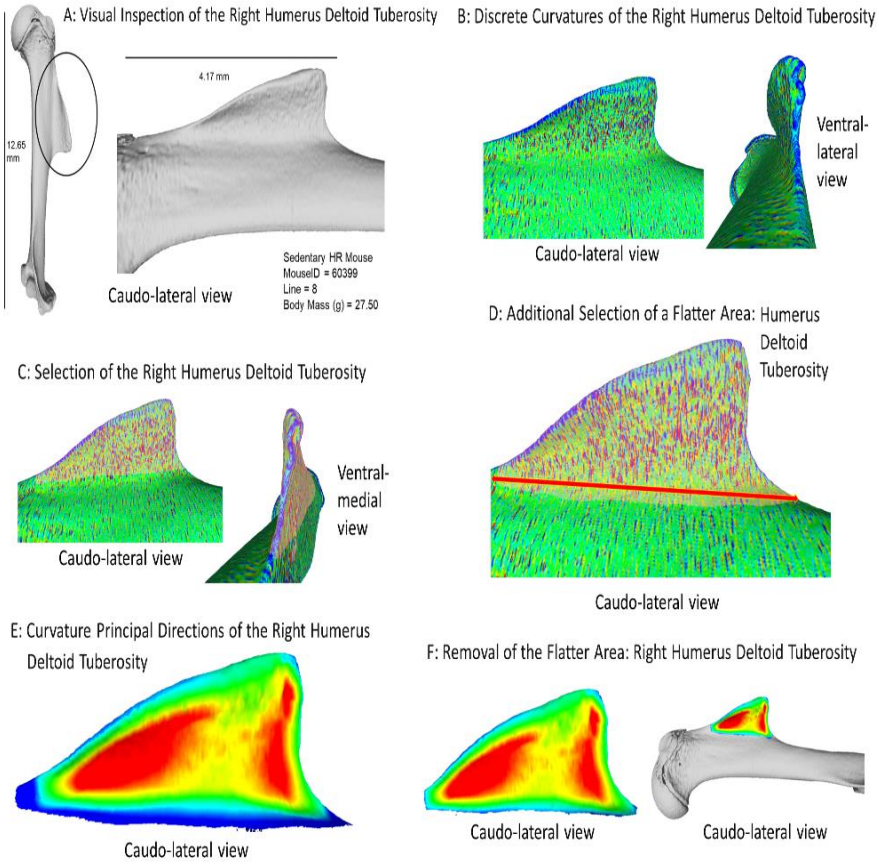

A: Visual Inspection of the Right Humerus Deltoid Tuberosity
12.65 mm
4.17 mm
Caudo-lateral view
Sedentary HR Mouse
MouseID = 60399
Line = 8
Body Mass (g) = 27.50
B: Discrete Curvatures of the Right Humerus Deltoid Tuberosity
Ventral-lateral view
Caudo-lateral view
C: Selection of the Right Humerus Deltoid Tuberosity
Ventral-medial view
Caudo-lateral view
D: Additional Selection of a Flatter Area: Humerus Deltoid Tuberosity
Caudo-lateral view
E: Curvature Principal Directions of the Right Humerus Deltoid Tuberosity
Caudo-lateral view
F: Removal of the Flatter Area: Right Humerus Deltoid Tuberosity
Caudo-lateral view
Caudo-lateral view

Figure 3.3. Surface Models. A: 3D surface models of the humerus deltoid tuberosity on the humerus (shown in right) and the delineated muscle attachment (shown in left) using MeshLab. B: 3D surface models of the femoral lesser trochanter on the femur (shown in right) and the delineated muscle attachment (shown in left) using MeshLab. C: 3D surface models of the femoral third trochanter on the femur (shown in right) and the delineated muscle attachment (shown in left) using MeshLab. D: 3D surface models of the femoral greater trochanter on the femur (shown in right) and the delineated muscle attachment (shown in left) using MeshLab.

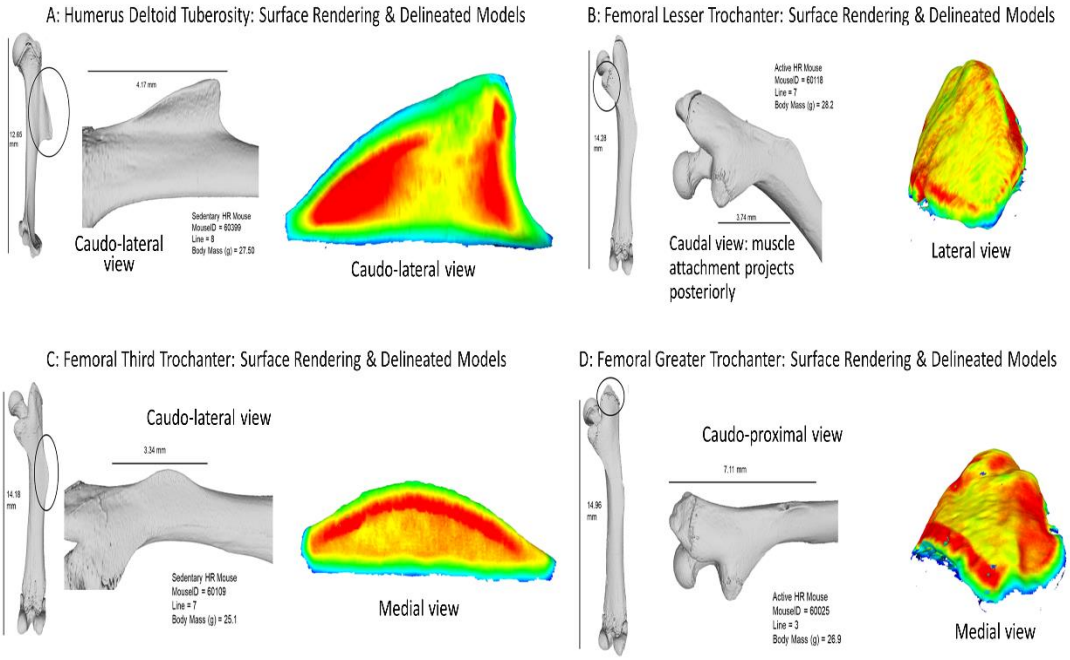


Figure 3.4. Humerus Deltoid Tuberosity. Humerus deltoid tuberosity surface area in relation to body mass. Larger mice had larger deltoid tuberosities, and Active mice (housed with long-term wheel access) had significantly larger muscle attachments for a given body size when compared with Sedentary mice (see Table 3.1 for statistical analyses).

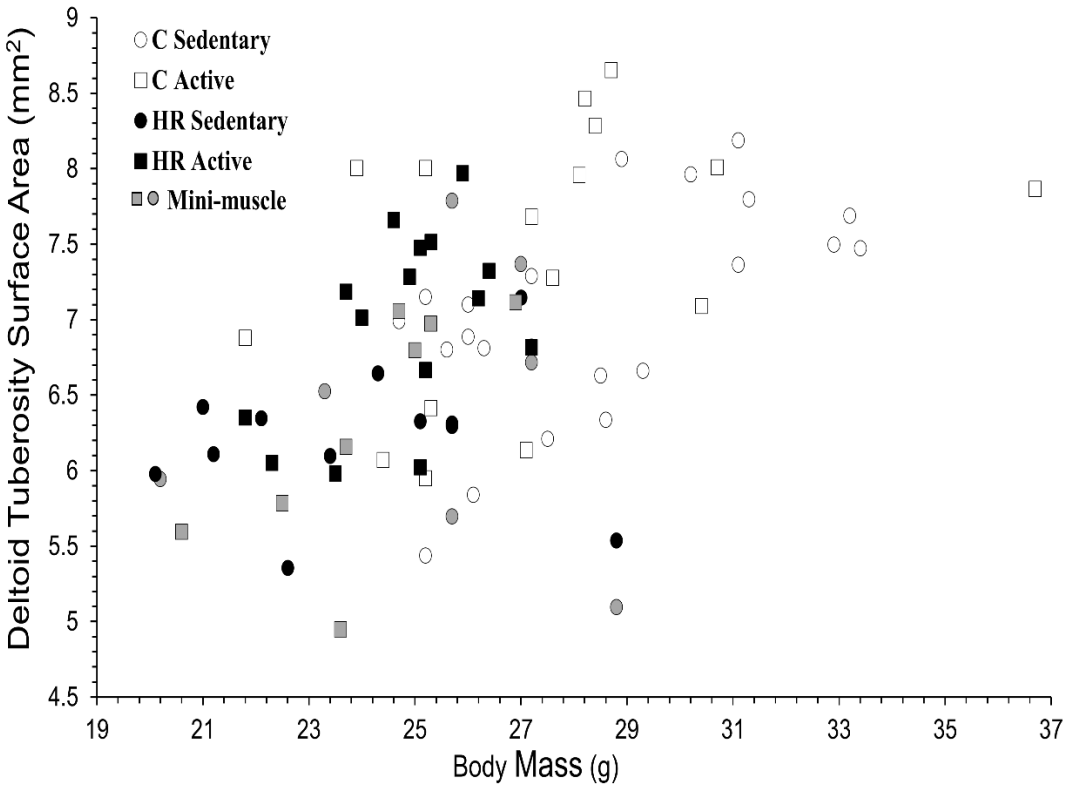
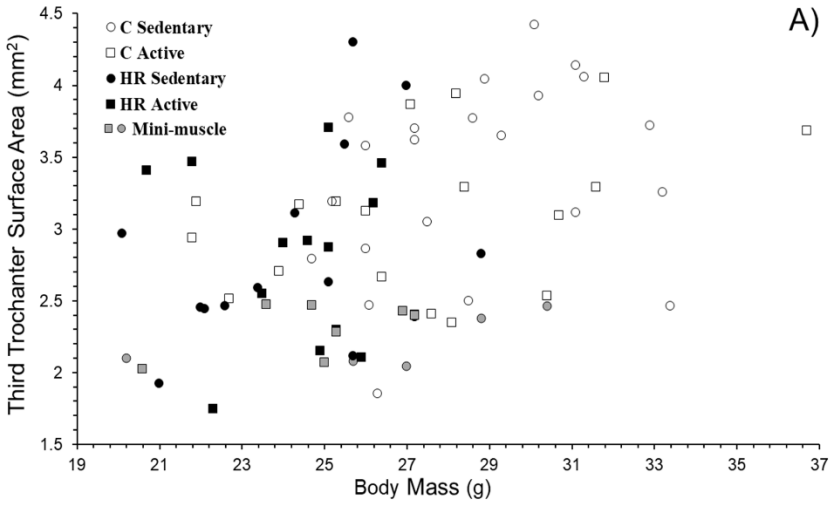
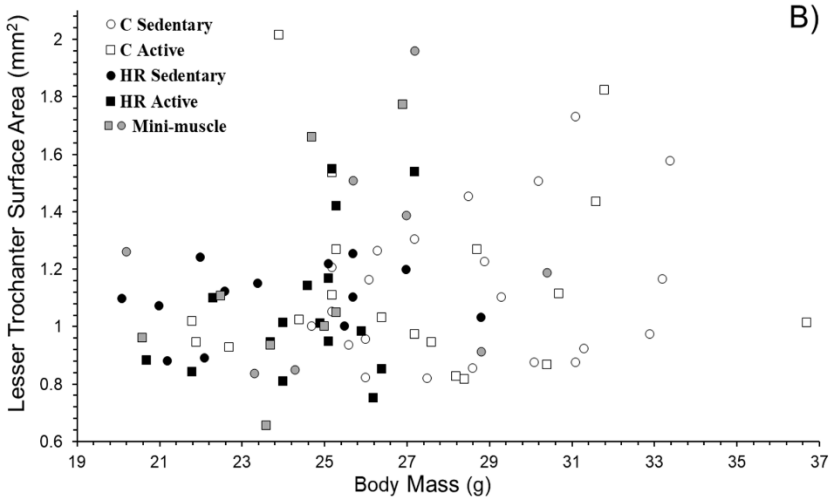


Figure 3.5. Femoral Attachments. A: Femoral third trochanter surface area in relation to body mass. Large mice had larger entheses and mini-muscle mice had significantly reduced entheses when compared with normal-muscled mice, with no significant effect of either linetype or activity. B: Femoral lesser trochanter in relation to body mass. There were no significant effects of linetype, activity or body mass. C: Femoral greater trochanter surface area in relation to body mass. Large mice had larger entheses, with no significant effect of either linetype or activity.


A)
C Sedentary
C Active
HR Sedentary
HR Active
Mini-muscle
Third Trochanter Surface Area (mm²)
Body Mass (g)


B)
C Sedentary
C Active
HR Sedentary
HR Active
Mini-muscle
Lesser Trochanter Surface Area (mm²)
Body Mass (g)

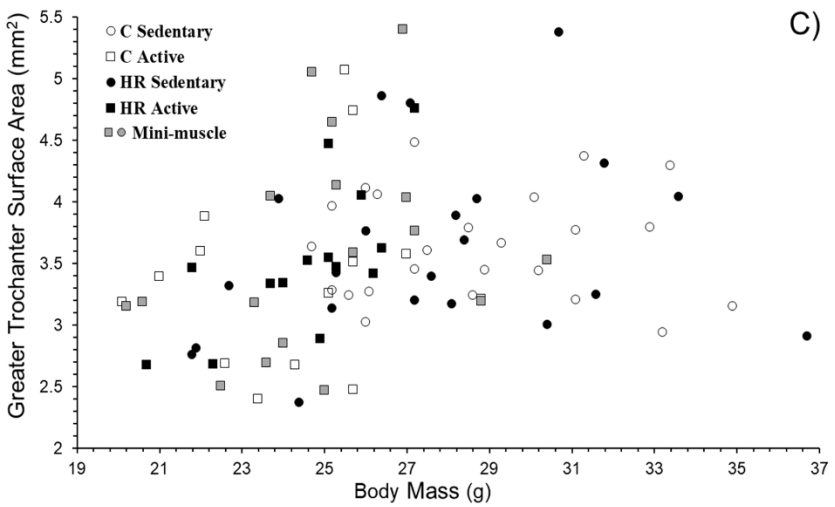
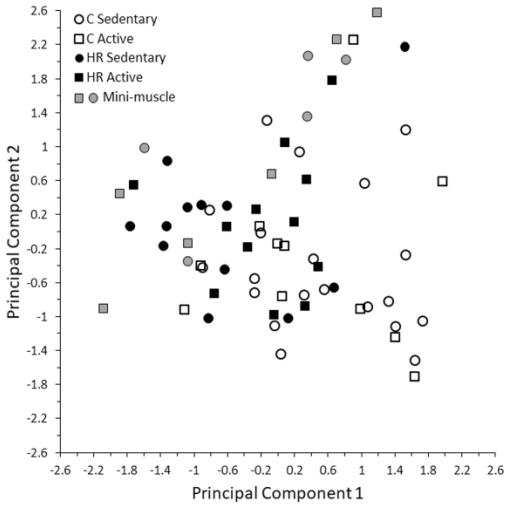
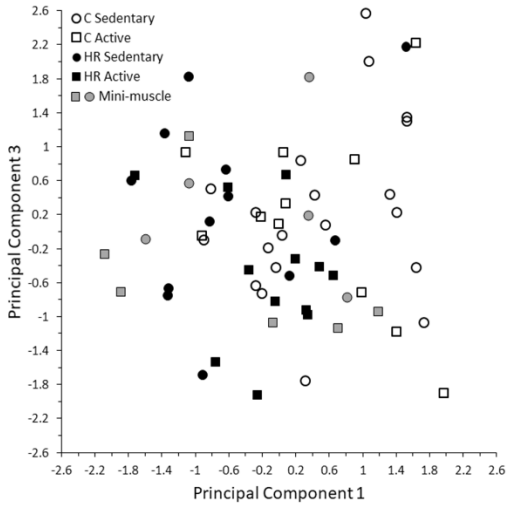
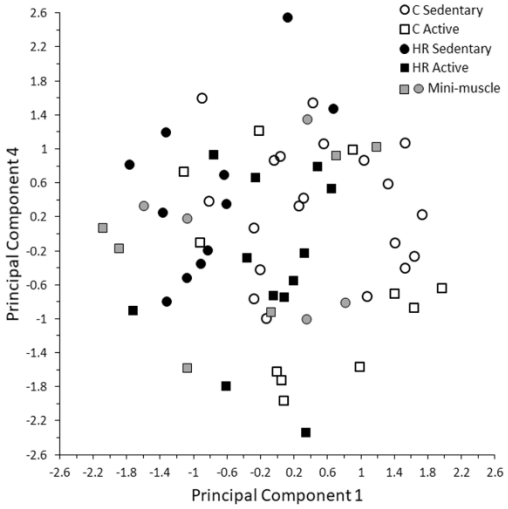
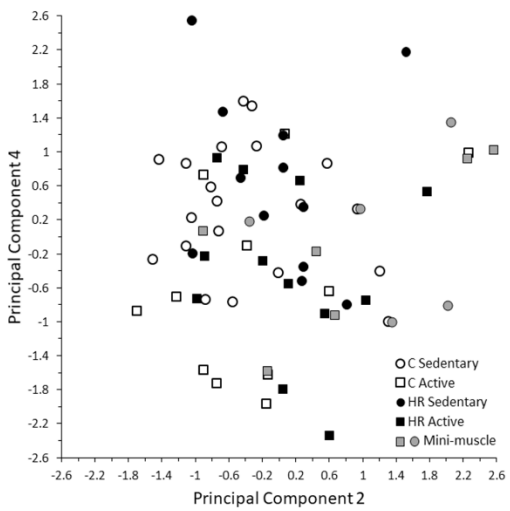

C)
C Sedentary
C Active
HR Sedentary
HR Active
Mini-muscle
Greater Trochanter Surface Area (mm²)
Body Mass (g)

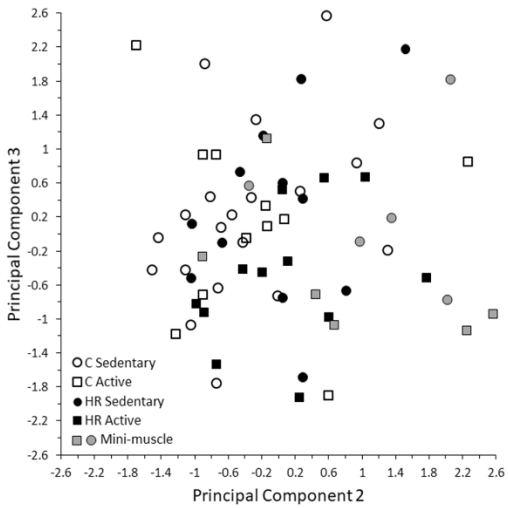
Figure 3.6. PCA. Pairwise plots of scores on principal components (see Table 3.3). Mice from the Active group (housed with long-term wheel access: squares) had significantly lower scores for PC 4, and mice with the mini-muscle phenotype (all in the HR lines: gray symbols) had lower scores for PC 1 (Table 3.1).


C Sedentary
C Active
HR Sedentary
HR Active
Mini-muscle
Principal Component 2
Principal Component 1


C Sedentary
C Active
HR Sedentary
HR Active
Mini-muscle
Principal Component 3
Principal Component 1


C Sedentary
C Active
HR Sedentary
HR Active
Mini-muscle
Principal Component 4
Principal Component 1

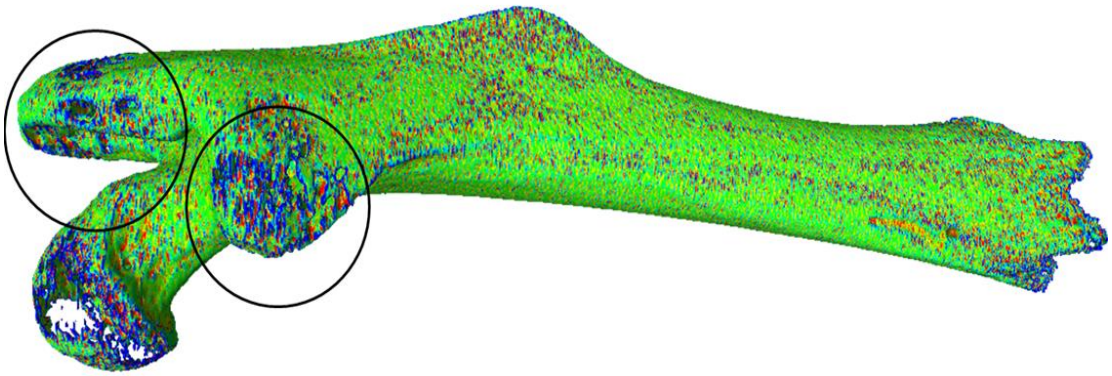

C Sedentary
C Active
HR Sedentary
HR Active
Mini-muscle
Principal Component 4
Principal Component 2


C Sedentary
C Active
HR Sedentary
HR Active
Mini-muscle
Principal Component 3
Principal Component 2

Appendices

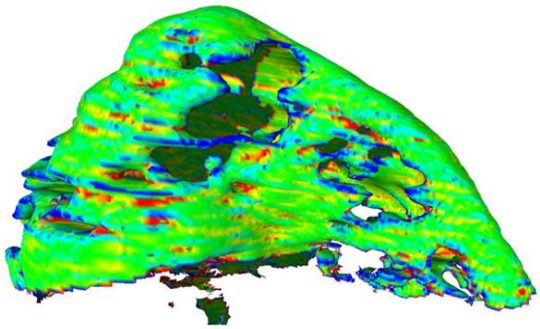
Appendix 3.1. Osteoporosis on a Mouse Femur

Osteoporosis in the femur of MouseID = 60515

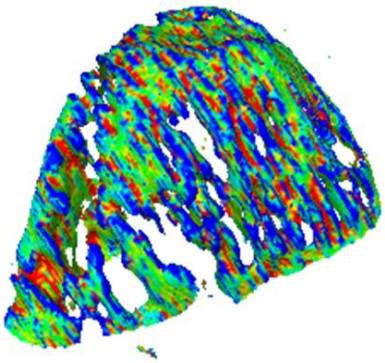


Osteoporosis on muscle attachments of the femur

Femoral Greater Trochanter



Femoral Lesser Trochanter



Trade-offs in muscle physiology in selectively bred High Runner mice

Alberto A. Castro, Theodore Garland, Jr., Saad Ahmed, Natalie C. Holt*

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

Acknowledgements

We thank members of the Garland lab for helping to obtain the animals used here and Allyn Nguyen for helpful discussions. Supported by U.S. National Science Foundation grant IOS-1121273 to T.G. and N.H.

Author Contributions

A.A.C., T.G., and N.C.H designed the experiments. AAC, SA, and N.C.H. conducted the experiments and collected data. A.A.C. analyzed the data and drafted the manuscript. A.A.C., T.G., and N.C.H revised and edited the manuscript. All authors approved the final version of the manuscript.

Abstract

A trade-off between locomotor speed and endurance seems to occur in many taxa and is thought to be underpinned by a muscle-level trade off related to myosin isoforms and oxidative capacities in different muscle fiber types. Among four replicate High Runner (HR) lines of mice selectively bred for high levels of voluntary wheel-running behavior, a negative correlation between average running speed and time spent running has evolved, and we hypothesize that this trade-off is related to changes in muscle physiology. We studied HR lines at generation 90, at which time one of these lines (L3) is fixed for the mini-muscle phenotype (caused by a Mendelian recessive allele), another is polymorphic (L6), and the others (L7, L8) lack mini-muscle individuals. Mini-muscle individuals have greatly diminished hindlimb muscle mass (~50%), few Type IIb fibers, and generally run faster but for fewer minutes daily on wheels. We used *in-situ* preparations to quantify the contractile properties, including speed and endurance, of the triceps surae complex. Maximal shortening velocity varied significantly, being lowest in mini-muscle mice (L3 Mini=25.2, L6 Mini=25.5 mm s^{-1}), highest in normal-muscled lines L6 and L8 (40.4- and 50.3-mm s^{-1}), and intermediate in L7 (37.2 mm s^{-1}). Endurance, measured as the slope of the decline in force over a series of 90 tetanic contractions, also varied significantly, being shallowest in the mini-muscle mice (L3 Mini=-0.00348, L6 Mini=-0.00238), steepest in lines L6 and L8 (-0.01676 and -0.01853), and intermediate in L7 (-0.01145). Therefore, muscle-level speed and endurance do trade-off in these mice, but not in a way that maps to the observed organismal-level speed-endurance trade-off.

Introduction

Trade-offs, limits to adaptation, and multiple solutions, have long been held as cornerstones in evolutionary biology and in many sub-fields of organismal biology, (Garland & Carter, 1994; Ackerly *et al.*, 2000; Martin *et al.*, 2015; Agrawal, 2020). Multiple types of trade-offs have been recognized (Cohen *et al.*, 2020; Mauro & Ghalambor, 2020; Garland *et al.*, 2022): perhaps the most common type involves allocation constraints. For example, if the energy available to an organism is limited, then spending more on one function (e.g., disease resistance) means less is available for other functions (e.g., reproduction). Another common type of trade-off occurs when features that enhance performance of one task decrease performance of another (Garland et al., 2022). Such functional conflicts are apparent in the biomechanics of bone and muscle function [e.g., relative lengths of in-levers and out-levers (Santana, 2016), force-velocity trade-offs in muscle (Herrel *et al.*, 2009; Schaeffer & Lindstedt, 2013)].

In the locomotor system, the most commonly studied trade-off at the organismal performance level is the negative relationship between speed and endurance. For example, among 12 species of closely related lacertid lizards, the speed and endurance capabilities are negatively related, after accounting for variation in body size (Vanhooydonck *et al.*, 2001). However, that trade-off is not apparent among species of phrynosomatid lizards (Albuquerque *et al.*, 2015; see also Toro *et al.*, 2004; Goodman *et al.*, 2007). Many studies have also tested for trade-offs at the level of variation among individuals. For example, statistically significant trade-offs were detected between speed-related and endurance-related events in a study of 1,369 elite human athletes

participating in heptathlon and decathlon events (Careau & Wilson, 2017). When present, the organismal-level trade-off between speed and endurance is thought to be underpinned by a muscle-level trade off in speed and endurance caused by the stereotyped combination of myosin isoforms and oxidative capacities in different muscle fibers (e.g., see Garland, 1988).

Mammalian muscle fiber types vary along a continuum of contractile and metabolic properties (for a review see Schiaffino & Reggiani, 2011). At one end of the spectrum, Type I fibers contract slowly, use oxidative metabolism, have low power outputs, and are fatigue resistant. At the other end of the spectrum, Type IIb fibers contract rapidly, use glycolysis, have high power outputs, and fatigue rapidly (Komi, 1984; Gleeson & Harrison, 1988; Rome *et al.*, 1988; Esbjörnsson *et al.*, 1993). Muscle fiber type variation has clear links with locomotor diversity. For instance, the predominance of Type I fibers in the forelimb muscles of slow-moving sloths (Spainhower *et al.*, 2018) contrasts with the predominance of Type IIb fibers in the hindlimb muscles of fast-sprinting cheetahs (Williams *et al.*, 1997). The spectrum of locomotor performance variation among lizard species also seems to relate to variation in muscle fiber types (Bonine *et al.*, 2005; Vanhooydonck *et al.*, 2014; Albuquerque *et al.*, 2015; Scales & Butler, 2016)

The purpose of the present study was to test whether a muscle-level trade-off underlies the negative relationship between the duration of daily running and the average running speed that has evolved among four replicate lines of high runner (HR) mice that have been selectively bred for 90 generations based on the average number of wheel

revolutions on days 5 & 6 of a 6-day period of wheel access (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), at which point HR mice run approximately three-fold more wheel revolutions per day than those from four replicate Control (C) lines. At generation 43, Garland et al. (2011) reported a significant negative correlation between average running speed and time spent running on wheels among the HR lines but not among the C lines. In the base population, these two traits were positively correlated both phenotypically and genetically (Swallow *et al.*, 1998).

Here, we quantified the speed and endurance properties of an important locomotor muscle group in mice, the triceps surae complex (which includes the lateral and medial heads of the gastrocnemius, the soleus, and the plantaris). More specifically, we used *in situ* preparations to measure isometric contractile properties (twitch and tetanic properties), the force-velocity relationship (including maximal shortening velocity), and endurance properties (fatigue in response to hundreds of isometric contractions). We hypothesized that contractile speed and stamina would trade-off among the HR lines in a way that parallels the documented whole organismal variation in running speed versus duration (Garland et al. 2011).

Materials and Methods

HR Mice Model

Mice from the 4 High Runner (HR) lines are bred for voluntary wheel running during 6 days of wheel access as young adults and are compared with 4 non-selected Control (C) lines (Swallow *et al.*, 1998). Briefly, the founding population was 224 laboratory house mice (*Mus domesticus*) of the outbred, genetically variable Hsd:ICR strain (Harlan-Sprague-Dawley, Indianapolis, Indiana, USA). Mice were randomly bred for two generations and then separated into 8 closed lines, which consist of 10 breeding pairs. 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 6-8 weeks of age. 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 (Swallow *et al.*, 1998; Careau *et al.*, 2013; Hiramatsu, 2017). 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. Sibling mating is not allowed. Mice are kept at room temperatures of approximately 22°C, with ad lib access to food and water. Photoperiod is 12L:12D with, the light phase beginning at 0700 hours and the dark phase at 1900 hours.

Mouse Specimens

To examine whether trade-offs in muscle function underlie the trade-off between average running speed and duration that has evolved among the HR lines (Garland *et al.*,

2011), we studied all four of the HR lines (lab designated as L3, L6, L7, and L8). Female mice (N = 31) from generation 90 of the selection experiment were housed 4/cage beginning at weaning (~21 days of age). We chose HR females because they generally run greater daily distances and at higher average speeds as compared with HR males (Garland *et al.*, 2011), thus making it more likely that muscle-based trade-offs might be relevant.

The "mini-muscle" phenotype presently occurs in a subset of the HR mice, characterized by to a 50% reduction in triceps surae (Garland *et al.*, 2002) and total hindlimb muscle mass (Houle-Leroy *et al.*, 2003), which is caused primarily by a dramatic reduction in type IIb muscle fibers (Guderley *et al.*, 2006; Talmadge *et al.*, 2014). Population-genetic modeling indicates that the mini-muscle phenotype was (unintentionally) under positive selection in the HR lines (Garland *et al.*, 2002). One of the HR lines (L3) has become is fixed for the mini-muscle phenotype (caused by a Mendelian recessive allele), another is polymorphic (L6), and the other two (L7, L8) lack mini-muscle individuals. In our sample of 31 mice (not all of which had data for all traits), the number of mini-muscle individuals was all of 6 in L3 and 6 of 13 in L6. Therefore, we had five total groups: L3 Mini, L6 Mini, L6, L7, and L8.

All mice were housed at room temperature with food and water ad lithium. All experiments were approved by the University of California- Riverside Institutional Animal Care and Use Committee.

Surgery Procedure

The twitch, force-velocity, and endurance properties of the left triceps surae (calf) muscle group, a muscle group key to locomotor performance (Clarke & Still, 1999), were determined *in situ*. Mice were anaesthetized (SomnoSuite Low-flow Anesthesia System, Kent Scientific, Torrington, CT, USA) and maintained 1.5-5% isoflurane anesthesia. The depth of anesthesia was continually monitored, and the dosage adjusted to maintain a sufficient depth. Body temperature was monitored using a thermometer inserted into the rectum and maintained throughout surgery via an integrated system that continuously adjusted the temperature of the heat pad placed under the animal (RightTemp System, Kent Scientific, Torrington, CT, USA). The sciatic nerve was surgically exposed, and a bipolar nerve cuff for electrical stimulation was placed around it. Mineral oil was applied at the attachment site to maintain moisture and the incision was closed. The proximal end of femur was exposed and clamped into a custom-made stereotaxic frame. Next, the Achilles tendon was exposed distally, Kelvar thread tied tightly around it, and the calcaneus cut. The end of the free tendon was attached to the lever arm of a servomotor (305C-LR Dual-Mode Lever System, Aurora Scientific, Aurora, ON, CA), allowing for measurements of muscle force, length, and velocity in the triceps surae complex (Ranatunga, 1984; Claflin & Faulkner, 1989; Zhan *et al.*, 1999; Syme *et al.*, 2005; Holt *et al.*, 2016; Javidi *et al.*, 2020).

Muscle Stimulation and Isometric Contractile Properties

All recordings and data processing were performed using data acquisition software (IgorPro 7, WaveMetric, Lake Oswego, OR, USA). The muscles were

stimulated (High-Power, Biphase Stimulator, Aurora Scientific, Aurora, ON, CA) by applying supramaximal square wave pulses (amplitude 1-2 mA, pulse duration 0.1 ms) to the sciatic nerve (Holt & Azizi, 2014). Optimum muscle length was determined by stimulating the muscle with a single square wave pulse to elicit twitch contractions at a range of lengths. The length that yielded peak twitch force was determined and defined as optimum length (L_0). All subsequent contractions were performed at this length.

Next, an isometric tetanic contraction (pulse frequency 80 Hz, stimulation duration 300ms) was performed to determine peak tetanic force (F_0) (Table 4.1). This isometric contraction was repeated at regular intervals to monitor muscle performance (e.g., decline in force due to fatigue and/or tissue decay) (Holt & Azizi, 2014; Holt *et al.*, 2016). If force had dropped below 90% of its initial value by the first control isometric tetani, the experiment was terminated.

Muscle Force-Velocity Properties

To determine the relationship between force and muscle velocity, isotonic tetanic contractions were performed at a range of forces (0.1-0.9 F_0). Peak shortening velocity was determined at each of these force levels (Appendix 4.1) and force-velocity curves were constructed (Marsh & Bennett, 1985, 1986; Bennett *et al.*, 1989; Askew & Marsh, 1997; Zhan *et al.*, 1999; Syme *et al.*, 2005; Holt *et al.*, 2016; Alcazar *et al.*, 2019; Javidi *et al.*, 2020).

Muscle Endurance Properties

The force-generating capacity of the triceps surae muscle over multiple repeated tetanic contractions was used to assess endurance capabilities (cf. Renaud & Kong, 1991;

Zhan *et al.*, 1999; James *et al.*, 2004; Syme *et al.*, 2005). The *in situ* muscle preparation eliminated the effects of the central nervous system while maintaining blood supply and, therefore, provided an assessment of the muscular basis of endurance. The fatigue test was run using a standard procedure of repeated isometric tetanic contractions (Allen *et al.*, 2008) and the same stimulation parameters as previous isometric tetanic contractions. One contraction was performed every 5 seconds until force dropped below 50% of its initial value, or for a maximum of 50 minutes or 500 contractions). However, due to data logging limitations these contractions had to be performed in 10-minute bouts. At the end of each bout the data were saved, and a new bout started.

Dissection and Muscle Dimensions

Once the endurance contraction protocol was completed, an overdose of isoflurane anesthesia was administered. The length of the Achilles tendon, triceps surae muscle complex, and muscle tendon unit were measured to the nearest 0.1 mm with digital calipers while the mouse was still in the stereotaxic frame (L_0). Mice were then removed from the frame, decapitated, and weighed. The triceps surae complex was dissected out and weighed to the nearest 0.0001 g.

Isometric Contractile Properties

Muscle anatomical cross-sectional area (Anatomical CSA) (not accounting for pennation angle or fiber length) was determined from muscle mass and muscle length assuming a density of 1.06 kg/L (Mendez & Keys, 1960) (Table 4.1). Subsequently, we calculated the peak tetanic stress (Stress = F_0/CSA) of the triceps surae muscle group (Askew & Marsh, 1997; Zhan *et al.*, 1999; Syme *et al.*, 2005; Holt *et al.*, 2016) (Table

4.1). Peak tetanic force (F_0) was also normalized to body mass (F_0 Mass) to assess the capacity of the muscle complex to support body weight during locomotion (Table 4.1).

Time series force data from twitch contractions were used to calculate the time from onset of muscle activation (at one second) to peak twitch force (TP_{tw}) and time from peak twitch force to 50% relaxation (TR_{50}) (Marsh & Bennett, 1985, 1986; Bennett *et al.*, 1989; Askew & Marsh, 1997; Syme *et al.*, 2005; Nguyen *et al.*, 2020) (Table 4.1). All isometric contractile force measurements were corrected for passive force.

F-V Curve Fitting

The force-velocity data were normalized by dividing active force by peak tetanic force (F/F_0) and velocity was normalized by dividing shortening velocity mm s^{-1} (absolute value) for the different isotonic contractions by muscle length (V_{norm}) (Table 4.1). After plotting the force-velocity points for individual mice, we fitted the force-velocity curve using different equations. We chose not to rely on a single force-velocity curve fit as none of the commonly used fits have a mechanistic basis, and, likely due to the triceps surae being a composite of multiple muscles, the force-velocity curves characterized here were relatively linear compared to previously observed curves (Marsh & Bennett, 1986; for a review see Alcazar *et al.*, 2019). We fitted using the Hill rectangular-hyperbola equation: $(P + a) (v + b) = b (P0 + a)$ (Hill, 1938), the Marsh-Bennett hyperbolic linear equation: $V = B(l – F/F_0) / (A + F/F_0) + C(1 – F/F_0)$ (Marsh & Bennett, 1986; Askew & Marsh, 1997), and using second-order polynomials: $f(x)=Ax^2+Bx+C$. Maximal shortening velocity mm s^{-1} (V_{max}) (Table 4.1) values were

determined for 3 fits for all mice and were visually rendered to check for outlier points that might lead to poor estimates.

Endurance Properties

The fatigue test included a series of repeated isometric tetanic contractions, one contraction every 5 seconds per 10-minute interval. From these graphs, we took active force measurements for each single contraction. Next, each endurance trial was concatenated to create endurance profiles for each mouse by creating a wave based on all the active force measurements (~200-500 contractions) (Appendix 4.2). Endurance ($Endur_{0\text{-}90}$) was quantified as the linear fit (slope) of the decline in force over the first 90 tetanic contractions (one contraction every 5 seconds) (Table 4.1). After the first 90 contractions, we quantified the average force that was sustained (Sustained F) over a series of tetanic contractions without a decrease in force (Table 4.1). The sustained isometric force (Sustained F) was also normalized to peak isometric force (F_0) to quantify the decline in active force.

Statistical Analysis

Isometric Contractile Properties

To compare the five groups (4 HR lines, with HR line 6 divided into those with and without the mini-muscle phenotype), we used the MIXED Procedure in SAS (SAS Institute, Cary, NC, USA) to apply analysis of covariance models with age as the covariate. Analyses of muscle dimensions also included body mass (except for variables that were normalized) as a covariate. We calculated an *a priori* contrast comparing L3 mini and L6 mini with L6, L7, and L8. For post-hoc comparisons within the mini- and within the normal-muscled groups, we examined Differences of Least Squares Means from SAS Procedure MIXED, with adjustment for multiple comparisons. Specifically, we employed Scheffe's procedure because this is the most conservative multiple-range comparison for unequal sample sizes. In all analyses, outliers were removed when the standardized residual exceeded ~3.0 and we used an α of ≤ 0.05 for statistical significance.

Force-Velocity Repeated Measures

Multiple force-velocity points were obtained for each individual mouse, so we used repeated-measures models in SAS Procedure MIXED to test for effects of group on both absolute shortening velocity (V_{max}) and normalized velocity (V_{norm}) (Table 4.2). Covariates were age, relative force (F/F_0), and z-transformed relative force squared (orthogonal polynomial). Individual was treated as a random effect nested within line. Furthermore, we included the interaction between force (F/F_0) and group (F/F_o*group) to test for differences in slopes. Initially, we also included the interaction between

(Zfnorm2) and group (Zfnorm2*group) to test for differences in curvature, but this interaction was not significant, so it was removed from the final model we present.

Least-square means generated from the repeated-measures analyses were estimated at $F/F_0 = 0$ to estimate maximal shortening velocity (mm s^{-1}) values from the 2nd degree polynomials for both V_{norm} and V_{max} (Table 4.2). We used a formal outlier test (Cook & Sanford, 1999) to make decisions about removing outliers (individual data points).

Correlations of Muscle Traits

To examine covariation of muscle performance metrics among the five groups, we examined bivariate scatterplots and calculated Pearson pairwise correlation coefficients for V_{norm}, $Endur_{0\text{-}90}$, Stress, TP_{tw}, TR_{50}, and SustainedF.

Results

Significance levels from ANCOVAs of body mass, muscle dimensions, and isometric contractile (tetanic and twitch) properties of the triceps surae complex in HR mice (using body mass and age as a covariate when appropriate) are shown in Figures 4.1 and 4.2 (full analyses are in S4.1). Table 4.2 and Figure 4.3 illustrate the results of force-velocity analysis (including representative traces from all groups). Figure 4.4 depicts the significance values from the endurance metrics (including representative traces) and Table 4.3 shows the pairwise correlation for the primary muscle contractile characteristics.

Body Size and Muscle dimensions

Average body mass varied significantly among groups (P=0.0011) (Figure 4.1A). With body mass as a covariate, muscle length (Figure 4.1B), tendon length, and the muscle-tendon unit (MTU) length were not significantly different among groups. As expected, relative triceps surae muscle mass varied among groups (P<<0.0001) (Figure 4.1C), with the mini-muscle mice (L3 Mini and L6 Mini) having significantly lighter muscles (LS Means of 0.052 g and 0.046 g, respectively) when compared with normal-muscled mice (L6=0.105 g, L7=0.095 g, L8=0.114 g) (*a priori* contrast P<<0.0001). *Post hoc* comparisons indicated no statistically significant differences between the two mini-muscle groups or among the three normal-muscle groups. The pattern for anatomical cross-sectional area was similar to that of muscle mass (Figure 4.1D).

Isometric contractile properties

Stress was not significantly different among groups (Figure 4.2A). F_0 Mass (peak tetanic force normalized to body mass; Table 4.1) was significantly different among groups ($P<0.0001$) (Figure 4.2B), with the main difference being that mini-muscle mice (L3 Mini and L6 Mini) had significantly lower values (both 0.046 N/g) when compared with the other groups (0.082 N/g for L6, 0.087 N/g for L7, and 0.100 N/g for L8) (*a priori* $P<<0.0001$).

TP_{tw} (s), time from muscle activation to peak twitch force (Table 4.1), ranged from an average of 0.021 (s) for L3 Mini to 0.025 (s) for L6 Mini but was not significantly different among groups (Figure 4.2C). TR_{50}, time from peak twitch force to half relaxation (Table 4.1) also did not differ among groups (Figure 4.2D).

Force-Velocity Repeated Measures

Figure 3 depicts force-velocity traces from individual mice per group, along with the three different curve-fits (see Methods). The second-order polynomials provided the most reliable fit for the force-velocity points and estimation of maximal shortening velocity (mm s^{-1}) (Figure 4.3A-4.3E). The Hill equation forced a curve when none existed, and the Marsh Bennet equation often generated convex shapes (Figure 4.3A-4.3E). Therefore, the force-velocity points were analyzed using second-order polynomials.

For absolute velocity (V_{max}: Figure 4.3F), the effect of group was highly significant (both $P<<0.0001$), as was the effect of normalized force (F/F_0) (both $P<<0.0001$), the z-transformation of normalized force (ZFnorm2) ($P<<0.0001$), and the

interaction between F/F_o*line ($P<<0.0001$) (Table 4.2). The interaction between F/F_o*line indicates differences in slope of the F-V curve among the groups. The *a priori* contrast between mini- and normal-muscled groups was highly significant ($P<< 0.0001$). In addition, the post hoc comparisons indicated that L8 was significantly higher than L6 ($P=0.0005$) and L7 ($P<<0.0001$). Results were similar for normalized velocity (V_{norm}) (Table 4.2).

Endurance

Figure 4 illustrates the fatigue test and endurance waves for representative individual mice from each of the five groups. The slight recovery in active force at Contraction # 100, 200, 300, and 400 occur due to the need to save data and restart the protocol every 100 contractions which gave the muscle a slightly longer recovery time between contractions (see Methods). At a gross level, L3 Mini and L6 Mini individuals had endurance waves with minimal drops in active force as compared with the other three groups (e.g., Figure 4.4A and 4.4B versus Figure 4.4C, 4.4D, 4.4E). Given these differences, we took the slope of the first 90 tetanic contractions as our measure of fatigue, and also measured sustained force during later contractions when force had stabilized at its apparent lowest value.

$Endur_{0\text{-}90}$ was significantly different among groups ($P<<0.0001$), being shallowest in the mini-muscle mice (L3 Mini=-0.00348, L6 Mini=-0.00238), steepest in normal lines L6 and L8 (-0.01676 and -0.01853, respectively), and intermediate in L7 (-0.01145) (Figure 4.4F). The *a priori* contrast between mini- and normal-muscled groups was highly significant ($P<<0.0001$). Sustained F/ F_0 (sustained isometric force normalized to

peak tetanic force) also differed among groups (P<0.0001), with mini-muscle groups having higher values (0.98 N for L3 Mini and 0.92 N for L6 Mini) when compared with L6 (0.44 N), L7 (0.47 N), and L8 (0.36 N) groups (Figure 4.4G). The *a priori* contrast between mini- and normal-muscled groups was also highly significant (P<<0.0001).

Pairwise Pearson's Correlations

Table 4.3 provides correlations for the five group Least Squares Means for force-velocity (V_{norm}), endurance ($Endur_{0\text{-}90}$, Sust. F), and isometric contractile properties (Stress, TP_{tw}, TR_{50}). Of the 15 correlations, only the correlation between V_{norm} and $Endur_{0\text{-}90}$ (r = -0.993) was statistically significant (P=0.0010) (Figure 4.5). Mini-muscle mice (L3 Mini and L6 Mini) have the highest endurance ($Endur_{0\text{-}90}$) but slowest muscles (V_{norm}), L6 and L8 have the lowest endurance but fastest muscles, and L7 is intermediate.

Discussion

Objectives

The purpose of the present study was to test whether a muscle-level trade-off underlies the negative relationship between the duration of daily running and the average running speed that has evolved among four replicate lines of high runner (HR) mice (Garland *et al.*, 2011). We did this by using *in situ* muscle preparations of the triceps surae complex, to measure isometric contractile properties, the force-velocity relationship, and endurance properties. Overall, we investigated trade-offs in muscle physiology and provided insight on the extent to which physiological trade-offs constrain evolutionary responses

Muscle Morphometrics

As in previous studies mini-muscle mice had about ~50% reduced triceps surae muscle mass when compared with normal-muscled individuals (e.g., see Garland *et al.*, 2002; Houle-Leroy *et al.*, 2003; Syme *et al.*, 2005), with no differences in muscle length (Figure 4.1). The same was true for the anatomical cross-sectional area (Anatomical CSA cm^2), which was calculated based on muscle length and mass, without accounting for pennation angle of muscle fibers (Mendez & Keys, 1960).

Isometric Contractile Comparisons

The calculation of Stress (N/cm^2) (Table 1, Figure 4.2A) was based on the whole triceps surae complex (not individual muscles) and Anatomical CSA. This will lead to higher estimates of Stress as compared with physiological cross-sectional area because we are using whole muscle length to calculate Anatomical CSA, ignoring the pennation

angle in the muscles of this group, which means that physiological cross-sectional area would be larger. A previous study of isolated muscles in HR mice found that the medial gastrocnemius muscle had Stress values ranging from 16.70 N/cm^2 to 17.8 N/cm^2 (Zhan *et al.*, 1999). Furthermore, studies of isolated calf muscles (soleus and extensor digitorium longus) in CD-1 mice reported variable Stress values that depended on age and fatigue (James *et al.*, 2004; Hill *et al.*, 2020), but were on average lower than Stress values for the triceps surae complex. Stress values in our experiment ranged from an average of 25.9 N/cm^2 for L6 to 36.6 N/cm^2 for L8 and as these relatively high values were expected, they demonstrate that that the muscle preparations were healthy.

TP_{tw} (s), time from muscle activation to peak twitch force, ranged from an average of 0.021 (s) for L3 Mini to 0.025 (s) for L6 Mini, but was not significantly different among groups (Figure 4.2C). Previous studies in HR mice (including mini-muscle individuals) reported faster values for the isolated soleus and medial gastrocnemius muscles (Syme *et al.*, 2005). Likewise, half relaxation times (TR_{50}) in a study of isolated muscles in ICR outbred mice were reported to be 0.023 (s) for the soleus (Askew & Marsh, 1997), which is on average, slower than the relaxation times of the whole triceps surae muscle for HR mice (ranging from 0.012 s for L7 to 0.013 s for L8). Furthermore, Syme et al. (2005), found that the medial gastrocnemius muscle in mini-muscle mice had slower twitch periods when compared with normal-muscled mice, but no significant differences for the soleus muscle. We likely did not find such an effect due to the triceps surae complex being a composite of four muscles (lateral and medial gastrocnemius,

plantaris, and soleus), with each muscle having different fiber types (e.g., Zhan *et al.*, 1999; Schaeffer & Lindstedt, 2013) and because the Achilles tendon was set in series.

Force-Velocity Properties

The first study examining force-velocity properties in the medial gastrocnemius of HR mice was at generation 10 (before discovery of the mini-muscle) and found no significant difference in maximum shortening velocity V_{max} or maximum power output between HR and C mice (Zhan *et al.*, 1999). Subsequently, Syme et al. (2005), reported significant increases in the force-velocity curvature, but reduced muscle mass and maximum isotonic power of the medial gastrocnemius muscle in mini-muscle mice. However, such differences were not apparent for the soleus muscle, despite the increase in muscle mass for mini-muscle individuals (Syme *et al.*, 2005). In the present study, force-velocity relationships differed among groups (differences in slope but not curvature of the F-V relationship) (Table 4.2). Additionally, V_{max} varied significantly, being lowest in mini-muscle mice (L3 Mini=25.2, L6 Mini=25.5 mm s^{-1}), highest in normal-muscled lines L6 and L8 (40.4- and 50.3-mm s^{-1}), and intermediate in L7 (37.2 mm s^{-1}) (Figure 4.3). Our estimate of V_{max} of the whole triceps surae muscle is slightly slower than the V_{max} estimates in previous studies of the soleus muscle in mice, which generally report values ~60-63 mm s^{-1} (Asmussen & Maréchal, 1989; Maréchal & Beckers-Bleukx, 1993; Askew & Marsh, 1997). The variance in V_{max} among isolated muscles in house mice are attributed to differences in fiber type and myosin isoform composition (e.g., see Maréchal & Beckers-Bleukx, 1993; Zhan *et al.*, 1999; Hill *et al.*, 2020).

Endurance Properties

The soleus and extensor digitorum longus muscles in mice generally fatigue within the first 100 tetanic contractions (Cabelka *et al.*, 2019) or within 100-500 seconds (e.g., see Pagala *et al.*, 1998; Zhao *et al.*, 2005), with the soleus generally being more fatigue resistant. Such differences in muscle fatigue are, at least in part, attributed to muscle fiber type composition, with Type I fiber abundance being positively correlated with fatigue resistance (e.g., see references in Garland, 1988; Schiaffino & Reggiani, 2011). The first study examining endurance properties in muscles from HR mice was at generation 10, and while voluntary exercise on wheels for 2 months improved muscle fatigue-resistance, no significant differences were found between HR and C mice (mini-muscle individuals were not present in the sample) (Zhan *et al.*, 1999). Subsequently, Syme et al. (2005) reported that the soleus in mini-muscle individuals fatigued at about half the rate of normal-muscled individuals. Likewise, in our study, endurance, measured as the slope of the decline in force over a series of 90 tetanic contractions, varied significantly for the triceps surae muscle, being shallowest in the mini-muscle mice (L3 Mini=-0.00348, L6 Mini=-0.00238), steepest in lines L6 and L8 (-0.01676 and -0.01853), and intermediate in L7 (-0.01145) (Figure 4.4F). Sustained F/F_0 (sustained isometric force normalized to peak tetanic force) was higher in mini-muscle mice (Figure 4.4G), likely due to the prevalence of fatigue-resistant muscle fibers (Talmadge *et al.*, 2014).

Trade-offs and Experimental Studies

Despite the biomechanical rationale for, and evolutionary importance of, speed-endurance trade-offs, experimental evidence for either is inconsistent (both at the

organismal level and in muscles). On one hand, trade-offs at the muscle level can sometimes be related to organismal-level performance trade-offs. For example, muscle specializations in the "the roll-snap" behavior (the rapid snapping of their wings together above their back) of bearded manakins can be explained by female preference for muscle speed that influenced assortative mating (Miles *et al.*, 2018).

On the other hand, trade-offs at the level of subordinate traits, such as muscles, can be at odds with speed and endurance metrics at the organismal level. For example, at the organismal-level, one study reported an absence of a trade-off between burst swimming performance and endurance capacity in African clawed frogs (Wilson *et al.*, 2002) and another found marginal evidence for a trade-off between burst (speed and acceleration) and sustained locomotion in lacertid lizards (Vanhooydonck *et al.*, 2014). At the muscle-level, studies of these same specimens (frogs and lizards respectively) have revealed significant trade-offs between muscular power output and fatigue resistance (Wilson *et al.*, 2002; Vanhooydonck *et al.*, 2014). As another example, trade-offs between maximum power output and fatigue resistance were apparent in individual mouse EDL muscles (Wilson & James, 2004).

Experimental Evolution and Trade-offs in HR mice

Selection experiments and experimental evolution can be used to study evolution in real time by determining the sequence of phenotypic and behavioral changes that occur during adaptation to a defined selective regime (Garland, 2003; Garland & Rose, 2009; Marchini *et al.*, 2014; Biesiadecki *et al.*, 2020). However, few studies, have used these approaches to elucidate mechanisms that underlie trade-offs (including discrepancies

between trade-offs at the organismal level and those found among lower-level traits). For example, functional trade-offs between running and fighting appear to have emerged as greyhounds and pit bulls were being developed by artificial selection (Pasi & Carrier, 2003; Kemp *et al.*, 2005).

At generation 43, Garland et al. (2011) reported a significant negative correlation between average running speed and time spent running on wheels among the four replicate HR lines but not among the C lines. Moreover, at the level of individual variation, the speed-duration correlation was lower (less positive) in the HR lines as compared with the C lines. Line 3 mini-muscle mice (mini-muscle status was unknown for L6) ran for fewer minutes per day on wheels but at faster average speeds. Mice from L8 ran for longer durations on wheels (minutes per day) but for the slowest speeds (revolutions per minute), and L7 was intermediate. Our muscle performance data show the opposite: mini-muscle mice (L3 Mini and L6 Mini) have the highest endurance ($Endur_{0\text{-}90}$) but slowest muscles (V_{norm}), L6 and L8 have the lowest endurance but fastest muscles, and L7 is intermediate.

Muscle and organismal level trade-offs might not replicate one another for several reasons. For starters, female HR mice run more intermittently when compared with female C mice (Girard *et al.*, 2001), and such differences in running behavior could "mask" muscle-level trade-offs. Furthermore, maximal running speeds on wheels (cf. Roach *et al.*, 2012) are well below maximal sprint speeds (e.g., see Dohm *et al.*, 1996; Girard *et al.*, 2001; Dlugosz *et al.*, 2009). For example, Line 3 individuals also have reduced maximal sprint speeds when measured on a racetrack (Dlugosz *et al.*, 2009).

Although metrics of maximal sprint speed are relatively closely related to aspects of muscle properties (e.g., see Komi, 1984) other traits are also important (discussed in Garland, 1988), and measures of endurance encompass many additional lower-level traits besides muscle physiology, including oxygen transport and delivery, thermoregulatory abilities, and additional cellular biochemical processes (discussed in Jones & Lindstedt, 1993; Schiaffino & Reggiani, 2011; Vanhooydonck *et al.*, 2014). Moreover, individuals may at least in part, compensate for the functional constraints imposed upon muscles by activating several agonistic muscles (discussed in Wilson & James, 2004) or by changing gaits (e.g., duty factor or stride length) (Claghorn *et al.*, 2017).

References

Ackerly, D.D., Dudley, S.A., Sultan, S.E., Schmitt, J., Coleman, J.S., Linder, C.R., *et al.* 2000. The evolution of plant ecophysiological traits: Recent advances and future directions. *BioScience* **50**: 979.

Agrawal, A.A. 2020. A scale-dependent framework for trade-offs, syndromes, and specialization in organismal biology. *Ecology* **101**.

Albuquerque, R.L., Bonine, K.E. & Garland, T. 2015. Speed and endurance do not trade off in Phrynosomatid lizards. *Physiol. Biochem. Zool.* **88**: 634–647.

Alcazar, J., Csapo, R., Ara, I. & Alegre, L.M. 2019. On the shape of the force-velocity relationship in skeletal muscles: The linear, the hyperbolic, and the double-hyperbolic. *Front. Physiol.* **10**.

Allen, D.G., Lamb, G.D. & Westerblad, H. 2008. Skeletal muscle fatigue: Cellular mechanisms. *Physiol. Rev.* **88**: 287–332.

Askew, G.N. & Marsh, R.L. 1997. The effects of length trajectory on the mechanical power output of mouse skeletal muscles. *J. Exp. Biol.* **200**: 3119–3131.

Asmussen, G. & Maréchal, G. 1989. Maximal shortening velocities, isomyosins and fibre types in soleus muscle of mice, rats and guinea-pigs. *J. Physiol.* **416**: 245–254.

Bennett, A.F., Garland Jr., T. & Else, P.L. 1989. Individual correlation of morphology, muscle mechanics, and locomotion in a salamander. *Am. J. Physiol.-Regul. Integr. Comp. Physiol.* **256**: R1200–R1208.

Biesiadecki, B.J., Brotto, M.A., Brotto, L.S., Koch, L.G., Britton, S.L., Nosek, T.M., *et al.* 2020. Rats genetically selected for low and high aerobic capacity exhibit altered soleus muscle myofilament functions. *Am. J. Physiol.-Cell Physiol.* **318**: C422–C429.

Bonine, K.E., Gleeson, T.T. & Garland Jr., T. 2005. Muscle fiber-type variation in lizards (Squamata) and phylogenetic reconstruction of hypothesized ancestral states. *J. Exp. Biol.* **208**: 4529–4547.

Cabelka, C.A., Baumann, C.W., Collins, B.C., Nash, N., Le, G., Lindsay, A., *et al.* 2019. Effects of ovarian hormones and estrogen receptor α on physical activity and skeletal muscle fatigue in female mice. *Exp. Gerontol.* **115**: 155–164.

Careau, V. & Wilson, R.S. 2017. Performance trade-offs and ageing in the 'world's greatest athletes.' *Proc. R. Soc. B Biol. Sci.* **284**: 20171048.

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. *Evolution* **67**: 3102–3119.

Claflin, D.R. & Faulkner, J.A. 1989. The force-velocity relationship at high shortening velocities in the soleus muscle of the rat. *J. Physiol.* **411**: 627–637.

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. *Physiol. Biochem. Zool.* **90**: 533–545.

Clarke, K. & Still, J. 1999. Gait analysis in the mouse. *Physiol. Behav.* **66**: 723–729.

Cohen, A.A., Coste, C.F.D., Li, X., Bourg, S. & Pavard, S. 2020. Are trade-offs really the key drivers of ageing and life span? *Funct. Ecol.* **34**: 153–166.

Cook, R.D. & Sanford, W. 1999. *Applied regression including computing and graphics*. John Wiley & Sons, Inc., New York, NY.

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. *J. Exp. Biol.* **212**: 2612–2618.

Dohm, M.R., Hayes, J.P. & Garland Jr., T. 1996. Quantitative genetics of sprint running speed and swimming endurance in laboratory house mice *Mus domesticus*. *Evolution* **42**: 355–350.

Esbjörnsson, M., Sylvén, C., Holm, I. & Jansson, E. 1993. Fast twitch fibres may predict anaerobic performance in both females and males. *Int. J. Sports Med.* **14**: 257–263.

Garland Jr., T. 1988. Genetic basis of activity metabolism. I. Inheritance of speed, stamina, and antipedator displays in the Garter Snake *Thampnophis Sirtalis*. *Evolution* **42**.

Garland Jr., T. 2003. Selection experiments: An under-utilized tool in biomechanics and organismal biology. In: *Vertebrate Biomechanics and Evolution*, p. 35. Bios Scientific Publisher, Oxford.

Garland Jr., T. & Carter, P.A. 1994. Evolutionary Physiology. *Annu. Rev. Physiol.* **56**: 579–621.

Garland Jr., T., Downs, C. & Ives, A.R. 2022. Perspective: Trade-offs (and constraints) in organismal biology. *Physiol. Biochem. Zool.* **95**: 717897.

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. *Proc. R. Soc. B Biol. Sci.* **278**: 574–581.

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.

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

Girard, I., McAleer, M.W., Rhodes, J.S. & Garland Jr., T. 2001. Selection for high voluntary wheel-running increases speed and intermittency in house mice *(Mus domesticus)*. *J. Exp. Biol.* **204**: 4311–4320.

Gleeson, T.T. & Harrison, J.M. 1988. Muscle composition and its relation to sprint running in the lizard *Dipsosaurus dorsalis*. *Am. J. Physiol.-Regul. Integr. Comp. Physiol.* **255**: R470–R477.

Goodman, B.A., Krockenberger, A.K. & Schwarzkopf, L. 2007. Master of them all: Performance specialization does not result in trade-offs in tropical lizards. *Evol. Ecol. Res.* **9**: 527–546.

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. *Comp. Biochem. Physiol. B Biochem. Mol. Biol.* **144**: 271–282.

Herrel, A., Podos, J., Vanhooydonck, B. & Hendry, A.P. 2009. Force-velocity trade-off in Darwin's finch jaw function: A biomechanical basis for ecological speciation? *Funct. Ecol.* **23**: 119–125.

Hill, A.V. 1938. The heat of shortening and the dynamic constants of muscle. *Proc. R. Soc. Lond. Ser. B - Biol. Sci.* **126**: 136–195.

Hill, C., James, R.S., Cox, Val.M., Seebacher, F. & Tallis, J. 2020. Age-related changes in isolated mouse skeletal muscle function are dependent on sex, muscle, and contractility mode. *Am. J. Physiol.-Regul. Integr. Comp. Physiol.* **319**: R296–R314.

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

Holt, N.C. & Azizi, E. 2014. What drives activation-dependent shifts in the force–length curve? *Biol. Lett.* **10**: 20140651.

Holt, N.C., Danos, N., Roberts, T.J. & Azizi, E. 2016. Stuck in gear: Age-related loss of variable gearing in skeletal muscle. *J. Exp. Biol.* **219**: 998–1003.

Houle-Leroy, P., Guderley, H., Swallow, J.G. & Garland Jr., T. 2003. Artificial selection for high activity favors mighty mini-muscles in house mice. *Am. J. Physiol.-Regul. Integr. Comp. Physiol.* **284**: R433–R443.

James, Rob.S., Wilson, R.S. & Askew, G.N. 2004. Effects of caffeine on mouse skeletal muscle power output during recovery from fatigue. *J. Appl. Physiol.* **96**: 545–552.

Javidi, M., McGowan, C.P. & Lin, D.C. 2020. Estimation of the force-velocity properties of individual muscles from measurement of the combined plantarflexor properties. *J. Exp. Biol.* **223**: jeb.219980.

Jones, J.H. & Lindstedt, S.L. 1993. Limits to maximal performance. *Annu. Rev. Physiol.* **55**: 547–569.

Kemp, T.J., Bachus, K.N., Nairn, J.A. & Carrier, D.R. 2005. Functional trade-offs in the limb bones of dogs selected for running versus fighting. *J. Exp. Biol.* **208**: 3475–3482.

Komi, P.V. 1984. Physiological and biomechanical correlates of muscle function: effects of muscle structure and stretch-shortening cycle on force and speed. *Exerc. Sport Sci. Rev.* **12**: 81–121.

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 Evol. Biol.* **14**: 15.

Maréchal, G. & Beckers-Bleukx, G. 1993. Force-velocity relation and isomyosins in soleus muscles from two strains of mice (C57 and NMRI). *Eur. J. Physiol.* **424**: 478–487.

Marsh, R.L. & Bennett, A.F. 1986. Thermal dependence of contractile properties of skeletal muscle from the lizard *Sceloporus occidentalis* with comments on methods for fitting and comparing force-velocity curves. *J. Exp. Biol.* **126**: 63–77.

Marsh, R.L. & Bennett, A.F. 1985. Thermal dependence of isotonic contractile properties of skeletal muscle and sprint performance of the lizard *Dipsosaurus dorsalis*. *J. Comp. Physiol. B* **155**: 541–551.

Martin, L.B., Ghalambor, C.K. & Woods, H.A. 2015. *Integrative organismal biology*. Wiley Blackwell, Hoboken, New Jersey.

Mauro, A.A. & Ghalambor, C.K. 2020. Trade-offs, pleiotropy, and shared molecular pathways: A unified view of constraints on adaptation. *Integr. Comp. Biol.* **60**: 332–347.

Mendez, J. & Keys, A. 1960. Density and composition of mammalian muscle. *Metabolism* **9**: 184–188.

Miles, M.C., Goller, F. & Fuxjager, M.J. 2018. Physiological constraint on acrobatic courtship behavior underlies rapid sympatric speciation in bearded manakins. *eLife* **7**: e40630.

Nguyen, A., Balaban, J.P., Azizi, E., Talmadge, R.J. & Lappin, A.K. 2020. Fatigue resistant jaw muscles facilitate long-lasting courtship behaviour in the southern alligator lizard (*Elgaria multicarinata*). *Proc. R. Soc. B Biol. Sci.* **287**: 20201578.

Pagala, M.K., Ravindran, K., Namba, T. & Grob, D. 1998. Skeletal muscle fatigue and physical endurance of young and old mice. *Muscle Nerve* **21**: 1729–1739.

Pasi, B.M. & Carrier, D.R. 2003. Functional trade-offs in the limb muscles of dogs selected for running vs. fighting: Functional trade-offs of running vs. fighting. *J. Evol. Biol.* **16**: 324–332.

Ranatunga, K.W. 1984. The force-velocity relation of rat fast- and slow-twitch muscles examined at different temperatures. *J. Physiol.* **351**: 517–529.

Renaud, J.M. & Kong, M. 1991. The effects of isotonic contractions on the rate of fatigue development and the resting membrane potential in the sartorius muscle of the frog, *Rana pipiens*. *Can. J. Physiol. Pharmacol.* **69**: 1754–1759.

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. *J. Appl. Physiol.* **113**: 627–635.

Rome, L.C., Funke, R.P., Alexander, R.M., Lutz, G., Aldridge, H., Scott, F., *et al.* 1988. Why animals have different muscle fibre types. *Nature* **335**: 824–827.

Santana, S.E. 2016. Quantifying the effect of gape and morphology on bite force: Biomechanical modelling and *in vivo* measurements in bats. *Funct. Ecol.* **30**: 557–565.

Scales, J.A. & Butler, M.A. 2016. Adaptive evolution in locomotor performance: How selective pressures and functional relationships produce diversity. *Evolution* **70**: 48–61.

Schaeffer, P.J. & Lindstedt, S.L. 2013. How animals move: Comparative lessons on animal locomotion. In: *Comprehensive Physiology* (R. Terjung, ed), p. c110059. John Wiley & Sons, Inc., Hoboken, NJ, USA.

Schiaffino, S. & Reggiani, C. 2011. Fiber types in mammalian skeletal muscles. *Physiol. Rev.* **91**: 1447–1531.

Spainhower, K.B., Cliffe, R.N., Metz, A.K., Barkett, E.M., Kiraly, P.M., Thomas, D.R., *et al.* 2018. Cheap labor: Myosin fiber type expression and enzyme activity in the forelimb musculature of sloths (*Pilosa: Xenarthra*). *J. Appl. Physiol.* **125**: 799–811.

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

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. *J. Appl. Physiol.* **99**: 1308–1316.

Talmadge, R.J., Acosta, W. & Garland, T. 2014. Myosin heavy chain isoform expression in adult and juvenile mini-muscle mice bred for high-voluntary wheel running. *Mech. Dev.* **134**: 16–30.

Toro, E., Herrel, A. & Irschick, D. 2004. The evolution of jumping performance in Caribbean *Anolis* lizards: Solutions to biomechanical trade-offs. *Am. Nat.* **163**: 844–856.

Vanhooydonck, B., James, R.S., Tallis, J., Aerts, P., Tadic, Z., Tolley, K.A., *et al.* 2014. Is the whole more than the sum of its parts? Evolutionary trade-offs between burst and sustained locomotion in lacertid lizards. *Proc. R. Soc. B Biol. Sci.* **281**: 20132677.

Vanhooydonck, B., Van Damme, R. & Aerts, P. 2001. Speed and stamina trade-off in lacertid lizards. *Evolution* **55**: 1040.

Williams, T.M., Dobson, G.P., Mathieu-Costello, O., Morsbach, D., Worley, M.B. & Phillips, J.A. 1997. Skeletal muscle histology and biochemistry of an elite sprinter, the African cheetah. *J. Comp. Physiol. [B]* **167**: 527–535.

Wilson, R.S. & James, R.S. 2004. Constraints on muscular performance: trade–offs between power output and fatigue resistance. *Proc. R. Soc. Lond. B Biol. Sci.* **271**.

Wilson, R.S., James, R.S. & Damme, R.V. 2002. Trade-offs between speed and endurance. *J. Exp. Biol.* **205**: 8.

Zhan, W.-Z., Swallow, J.G., Garland, T., Proctor, D.N., Carter, P.A. & Sieck, G.C. 1999. Effects of genetic selection and voluntary activity on the medial gastrocnemius muscle in house mice. *J. Appl. Physiol.* **87**: 2326–2333.

Zhao, X., Yoshida, M., Brotto, L., Takeshima, H., Weisleder, N., Hirata, Y., *et al.* 2005. Enhanced resistance to fatigue and altered calcium handling properties of sarcalumenin knockout mice. *Physiol. Genomics* **23**: 72–78.

Table 4.1. Definitions of muscle dimensions and contractile properties of the triceps surae complex in HR mice.

Abbreviations	Definition and Functional Significance
F_0 (N)	Peak isometric and tetanic force of the triceps surae complex
Anatomical CSA (cm^2)	Anatomical cross-sectional area of the triceps surae complex [((Triceps surae mass (g)'/1000)/1060)/ (Triceps surae length (mm)'/1000) *10000)] (Mendez & Keys, 1960)
Stress (N/cm^2)	Peak tetanic stress of the triceps surae complex ['F_0'/'Anatomical CSA'] (Askew & Marsh, 1997; Zhan *et al.*, 1999; Syme *et al.*, 2005; Holt *et al.*, 2016)
F_0 Mass (N/g)	Peak tetanic force of the triceps surae complex normalized to body mass ['F_0'/'Body Mass']
TP_{tw} (ms)	Time from muscle activation to peak twitch force (Marsh & Bennett, 1985, 1986; Bennett *et al.*, 1989; Askew & Marsh, 1997; Syme *et al.*, 2005)
TR_{50} (ms)	Time from peak twitch force to 50% relaxation (Marsh & Bennett, 1985, 1986; Bennett *et al.*, 1989; Askew & Marsh, 1997; Syme *et al.*, 2005)
F/F_0 (N)	Active force of isotonic contractions divided by F_0 (0.1-0.9 F_0) (Marsh & Bennett, 1985, 1986; Bennett *et al.*, 1989; Askew & Marsh, 1997; Syme *et al.*, 2005; Holt *et al.*, 2016; Alcazar *et al.*, 2019)
V_{norm}	Shortening velocity of isotonic contractions (0.1-0.9 F_0) divided by muscle length (Marsh & Bennett, 1985, 1986; Bennett *et al.*, 1989; Askew & Marsh, 1997; Syme *et al.*, 2005; Holt *et al.*, 2016; Alcazar *et al.*, 2019)
V_{max} ($mm\ s^{-1}$)	Maximal shortening velocity (Marsh & Bennett, 1985, 1986; Askew & Marsh, 1997; Zhan *et al.*, 1999; Syme *et al.*, 2005; Holt *et al.*, 2016; Alcazar *et al.*, 2019)
$Endur_{0\text{-}90}$	Linear fit (slope) of the first 90 isometric contractions of the triceps surae complex
Sustained F (N)	Sustained isometric force of the muscle after first 90 contractions of the fatigue protocol and after tetanic force stops to decline
Sustained F/ F_0 (N)	Sustained isometric force normalized to peak tetanic force ['Sustained F'/'F_0']

Table 4.2. Repeated-measures analyses for force-velocity measurements based on second-order polynomial fits for absolute velocity (V_{max}) and normalized velocity (V_{norm}).

V_{max} N=192	**P-Value**	**F-Value**	**d.f.**	**V_{norm} N=184**	**P-Value**	**F-Value**	**d.f.**
Group	**<0.0001**	102.44	4,180		**<0.0001**	56.48	4,172
F/F_0	**<0.0001**	995.63	1,180		**<0.0001**	722.36	1,172
ZFnorm2	**<0.0001**	39.67	1,180		**<0.0001**	24.18	1,172
Age	**<0.0001**	85.85	1,180		**<0.0001**	101.71	1,172
F/F_0 * Group	**<0.0001**	19.52	4,180		**<0.0001**	9.01	4,172
	LSM	**SE**			**LSM**	**SE**	
L3 Mini	25.1848	0.5173			1.8309	0.05417	
L6 Mini	25.5339	0.8492			1.9046	0.07414	
L6	40.4354	1.6776			3.1393	0.1577	
L7	37.1715	1.6129			2.7271	0.09728	
L8	50.2800	1.3021			3.4822	0.12800	
Sol. Fixed Effect				**Sol. Fixed Effect**			
	F/F_0	**Zfnorm2**	**Age**		**F/F_0**	**Zfnorm2**	**Age**
	-45.56	1.99	0.14		-3.07	0.14	0.015

Table 4.3. Pairwise Pearson correlations for Least Squares Means of force-velocity, endurance properties, and isometric contractile properties (N = 5). The only statistically significant correlation is between estimated maximal shortening velocity (V_{norm}) and endurance (see Figure 4.5).

		V_{norm}	$Endur_{0-90}$	Stress	TP_{tw}	TR_{50}	Sust. F
V_{norm}	Correlation		-.993	-.266	.143	-.148	-.671
	Sig. (2-tailed)		.001	.666	.818	.812	.215
$Endur_{0-90}$	Correlation			.340	-.040	.111	.610
	Sig. (2-tailed)			.576	.949	.859	.274
Stress	Correlation				.341	.466	.287
	Sig. (2-tailed)				.574	.429	.640
TP_{tw}	Correlation					-.261	-.505
	Sig. (2-tailed)					.672	.385
TR_{50}	Correlation						.784
	Sig. (2-tailed)						.117

Figure 4.1. Muscle Dimensions and Body Mass. A: Least square means and standard errors of body mass for each line (L3 Mini, L6 Mini, L6, L7 and L8). Age was positively associated with body size and both L6 Mini and L7 mice were significantly lighter when compared with the other lines. B: Least square means and standard errors of triceps surae muscle length for each line. Neither age or body mass was associated with muscle length, and the lines did not differ significantly. C: Least square means and standard errors of triceps surae muscle mass for each line. Triceps surae muscle mass was positively associated with body mass, and mini-muscle mice (L3 Mini and L6 Mini) had significantly lighter muscles when compared with the other lines (L6, L7 and L8), with L7 having intermediate values. D: Least square means and standard errors of Anatomical CSA for each line. Mini-muscle mice (L3 Mini and L6 Mini) had significantly lower Anatomical CSA values when compared with the other lines.

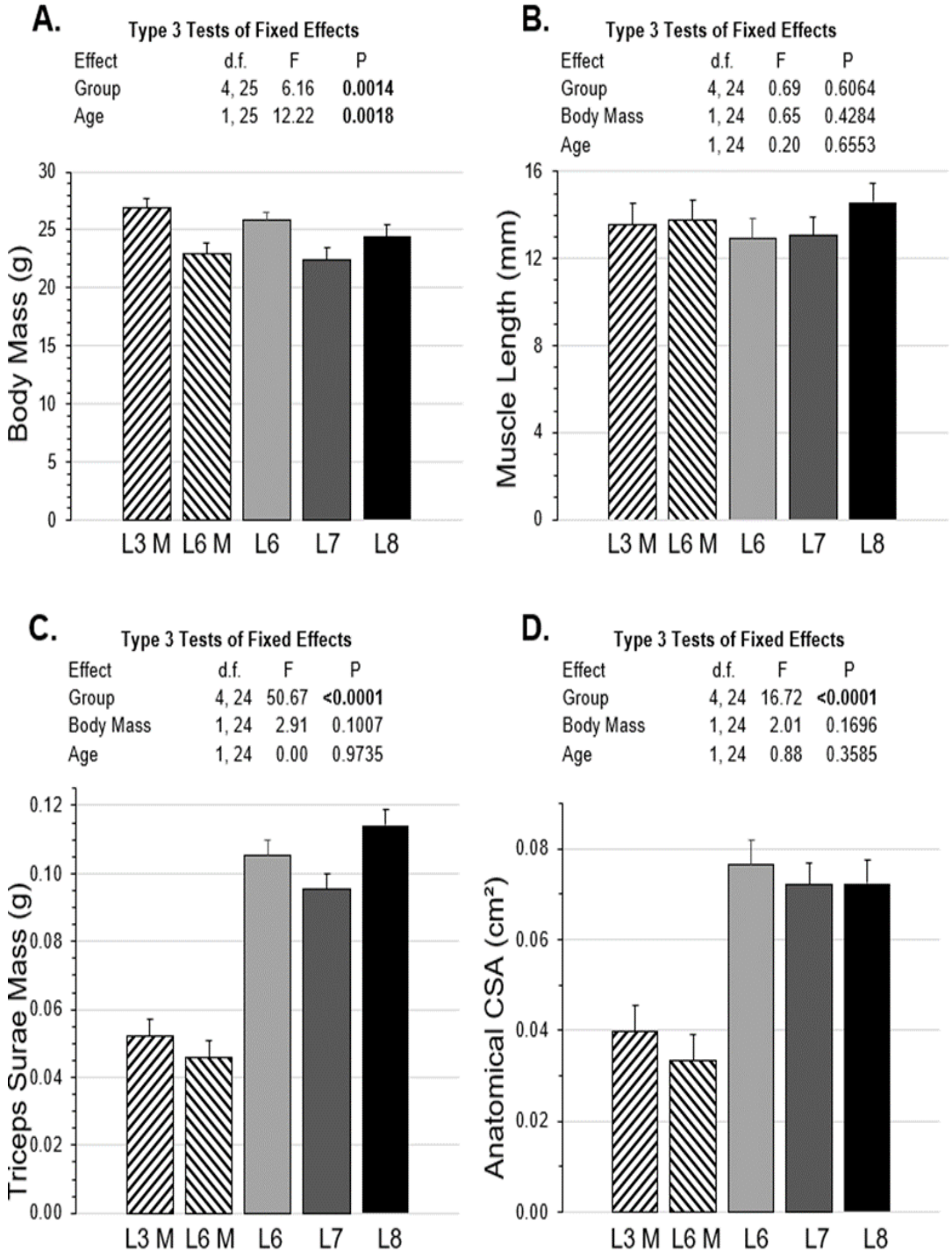

A.
Type 3 Tests of Fixed Effects
Effect d.f. F P
Group 4, 25 6.16 0.0014
Age 1, 25 12.22 0.0018
Body Mass (g)
0
5
10
15
20
25
30
L3 M
L6 M
L6
L7
L8
B.
Type 3 Tests of Fixed Effects
Effect d.f. F P
Group 4, 24 0.69 0.6064
Body Mass 1, 24 0.65 0.4284
Age 1, 24 0.20 0.6553
Muscle Length (mm)
0
4
8
12
16
L3 M
L6 M
L6
L7
L8
C.
Type 3 Tests of Fixed Effects
Effect d.f. F P
Group 4, 24 50.67 <0.0001
Body Mass 1, 24 2.91 0.1007
Age 1, 24 0.00 0.9735
Triceps Surae Mass (g)
0.00
0.02
0.04
0.06
0.08
0.10
0.12
L3 M
L6 M
L6
L7
L8
D.
Type 3 Tests of Fixed Effects
Effect d.f. F P
Group 4, 24 16.72 <0.0001
Body Mass 1, 24 2.01 0.1696
Age 1, 24 0.88 0.3585
Anatomical CSA (cm²)
0.00
0.02
0.04
0.06
0.08
L3 M
L6 M
L6
L7
L8

Figure 4.2. Isometric Contractile Properties. A: Least square means and standard errors of Stress for each line (L3 Mini, L6 Mini, L6, L7 and L8). Neither age or body mass was associated with Stress and the lines did not differ significantly. B: Least square means and standard errors of F_0 Mass for each line. Mini-muscle mice (L3 Mini and L6 Mini) had significantly lower F_0 Mass values when compared with the other lines (L6, L7, and L8). C: Least square means and standard errors of TP_{tw} for each line. Neither age or body mass was associated with $TP_{tw,}$ and the lines did not differ significantly. D: Least square means and standard errors of TR_{50} for each line. Neither age or body mass was associated with $TR_{50,}$ and the lines did not differ significantly.

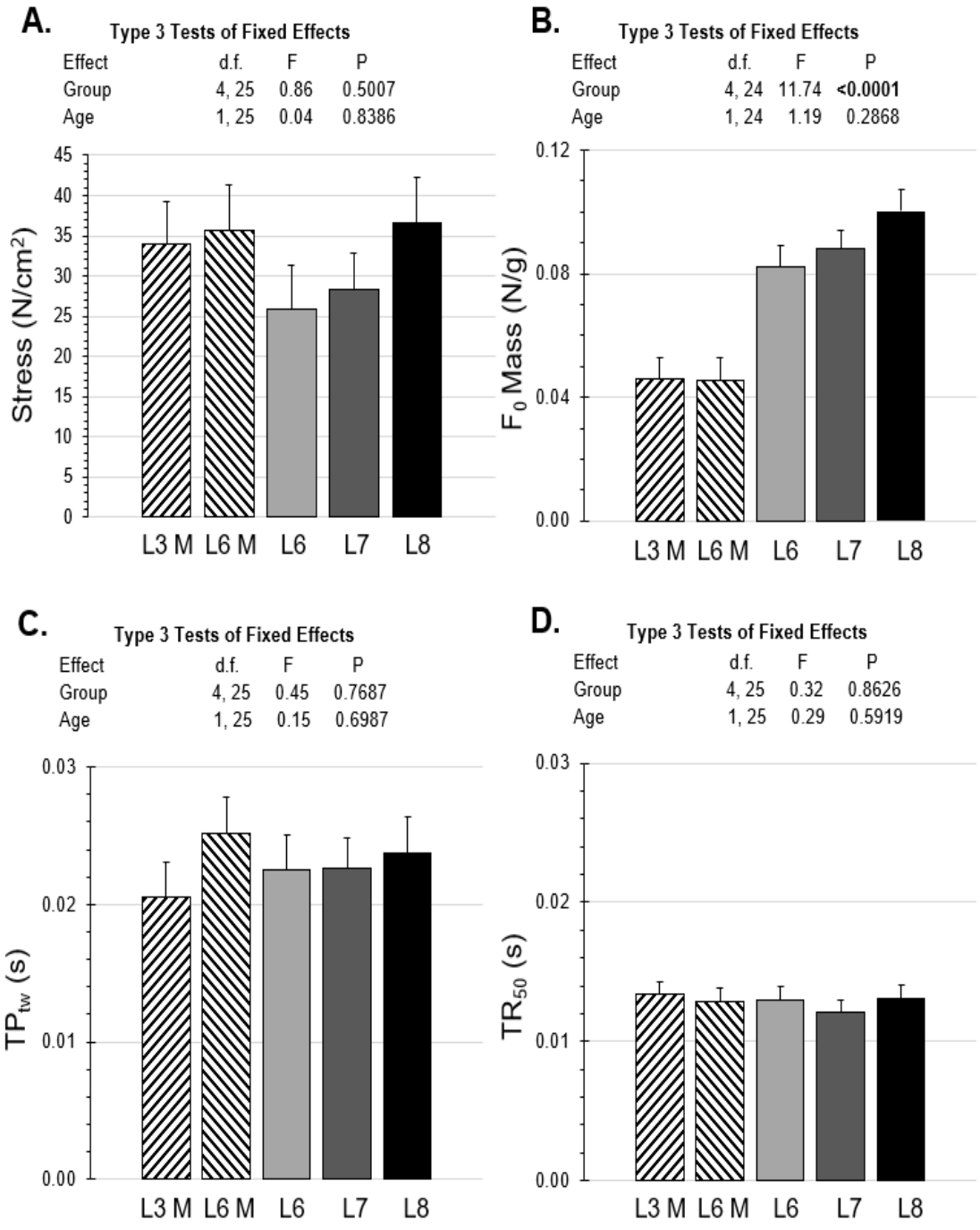

A.
Type 3 Tests of Fixed Effects
Effect d.f. F P
Group 4, 25 0.86 0.5007
Age 1, 25 0.04 0.8386
Stress (N/cm²)
0
5
10
15
20
25
30
35
40
45
L3 M
L6 M
L6
L7
L8
B.
Type 3 Tests of Fixed Effects
Effect d.f. F P
Group 4, 24 11.74 <0.0001
Age 1, 24 1.19 0.2868
F0 Mass (N/g)
0.00
0.04
0.08
0.12
L3 M
L6 M
L6
L7
L8
C.
Type 3 Tests of Fixed Effects
Effect d.f. F P
Group 4, 25 0.45 0.7687
Age 1, 25 0.15 0.6987
TPtw (s)
0.00
0.01
0.02
0.03
L3 M
L6 M
L6
L7
L8
D.
Type 3 Tests of Fixed Effects
Effect d.f. F P
Group 4, 25 0.32 0.8626
Age 1, 25 0.29 0.5919
TR50 (s)
0.00
0.01
0.02
0.03
L3 M
L6 M
L6
L7
L8

Figure 4.3. Force-Velocity A: Representative force-velocity trace for L3 Mini, with (F/F_0) on the x-axis and absolute shortening velocity on the y-axis. The force-velocity points were curve-fitted using the Hill equation, Marsh-Bennet equation, and second-order polynomials. Maximal shortening velocity mm s^{-1} estimates using the three fits are visually rendered. B: Representative force-velocity trace for L6 Mini. C: Representative force-velocity trace for L6. D: Representative force-velocity trace for L7. E: Representative force-velocity trace for L8. F: Least square means and standard errors of V_{max} for each line. V_{max} was positively associated with age, and mini-muscle mice (L3 Mini and L6 Mini) had significantly lower V_{max} values when compared with the other lines (L6, L7 and L8), with L7 having intermediate V_{max} values.

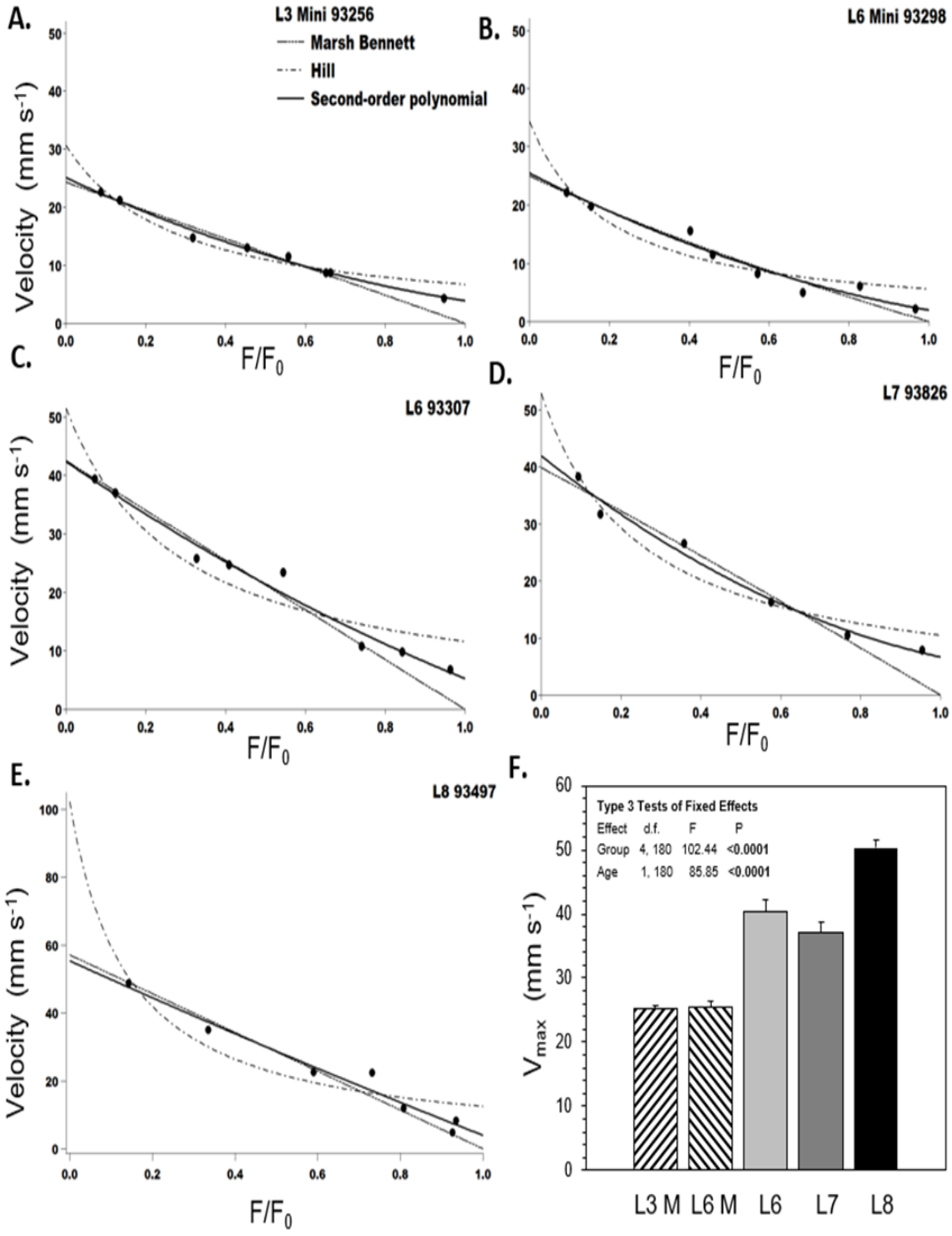

A.
L3 Mini 93256
Marsh Bennett
Hill
Second-order polynomial
Velocity (mm s-1)
F/F0
B.
L6 Mini 93298
C.
L6 93307
D.
L7 93826
E.
L8 93497
F.
Type 3 Tests of Fixed Effects
Effect d.f. F P
Group 4, 180 102.44 <0.0001
Age 1, 180 85.85 <0.0001
Vmax (mm s-1)
L3 M L6 M L6 L7 L8

Figure 4.4. Endurance A: Representative endurance trace wave profile for L3 Mini, with Contraction # on the x-axis and isometric force on the y-axis. The linear fit ($Endur_{0-90}$) of the decline in force over the first 90 tetanic contractions and the average sustained force (Sustained F) are visually rendered. B: Representative endurance trace wave profile for L6 Mini. C: Representative endurance trace wave profile for L6. D: Representative endurance trace wave profile for L7. E: Representative endurance trace wave profile for L8. L8 mice all fatigued within the first 200 contractions. F: Least square means and standard errors of $Endur_{0-90}$ for each line. $Endur_{0-90}$ was positively associated with age, and mini-muscle mice (L3 Mini and L6 Mini) had significantly lower $Endur_{0-90}$ values when compared with the other lines (L6, L7 and L8), with L7 having intermediate $Endur_{0-90}$ values. G: Least square means and standard errors of Sustained F/ F_0 for each line. Mini-muscle mice (L3 Mini and L6 Mini) had significantly lower Sustained F/ F_0 values when compared with the other lines.

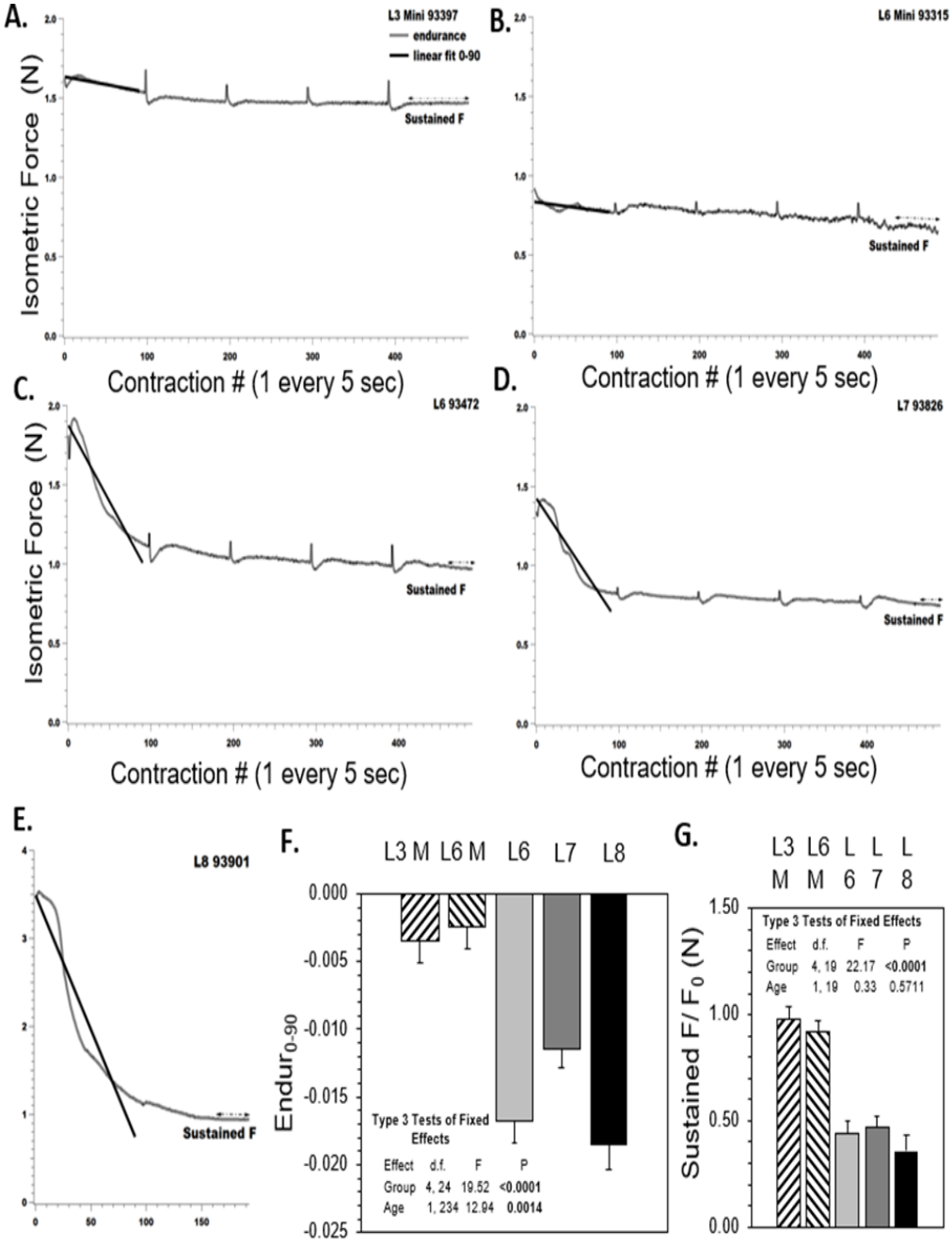
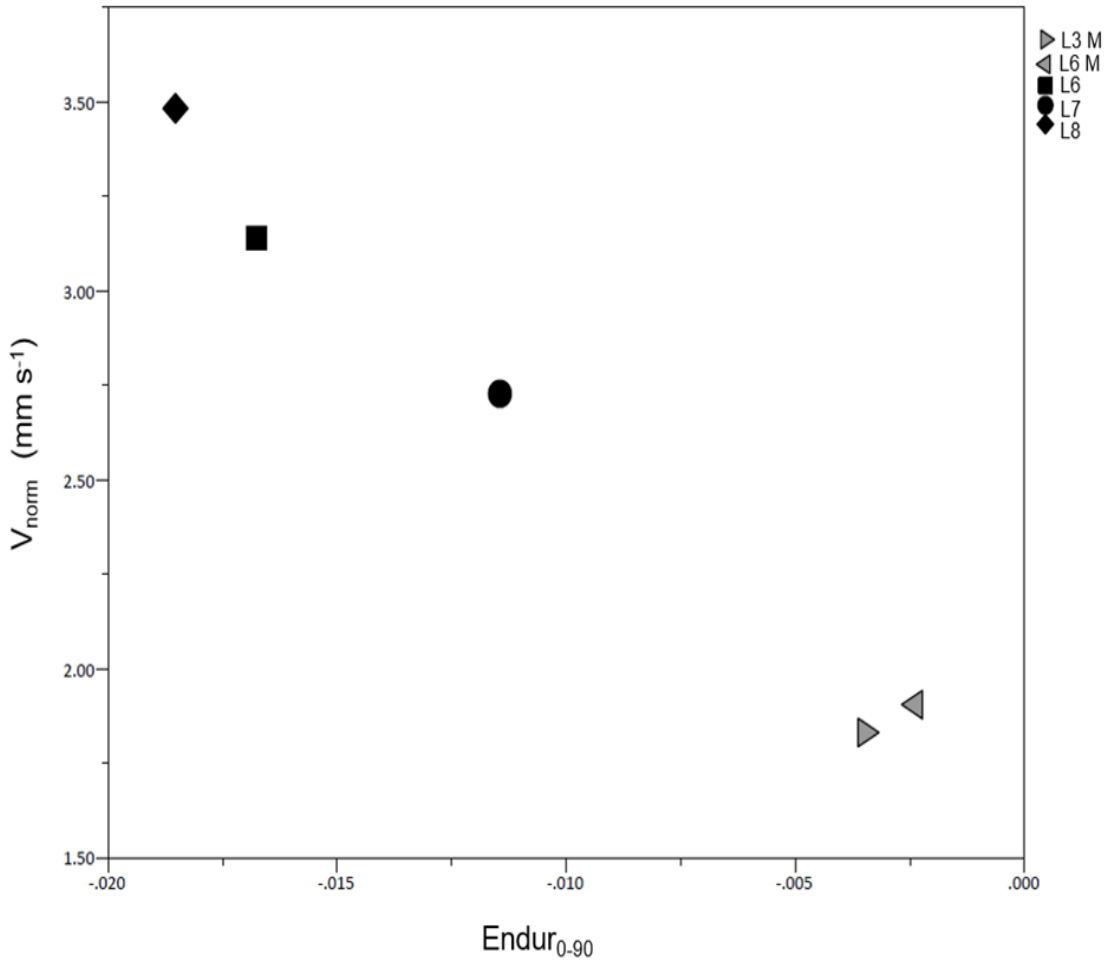
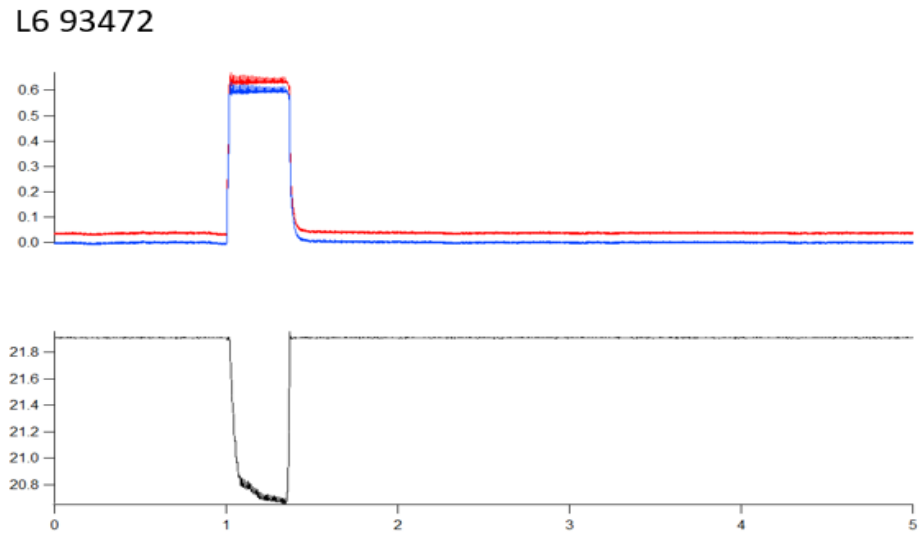

A.
L3 Mini 93397
endurance
linear fit 0-90
Sustained F
Isometric Force (N)
Contraction # (1 every 5 sec)
B.
L6 Mini 93315
Sustained F
Contraction # (1 every 5 sec)
C.
L6 93472
Sustained F
Isometric Force (N)
Contraction # (1 every 5 sec)
D.
L7 93826
Sustained F
Contraction # (1 every 5 sec)
E.
L8 93901
Sustained F
F.
L3 M L6 M L6 L7 L8
Endur0-90
Type 3 Tests of Fixed Effects
Effect d.f. F P
Group 4, 24 19.52 <0.0001
Age 1, 234 12.94 0.0014
G.
L3 M L6 M L 6 L 7 L 8
Sustained F/ F0 (N)
Type 3 Tests of Fixed Effects
Effect d.f. F P
Group 4, 19 22.17 <0.0001
Age 1, 19 0.33 0.5711

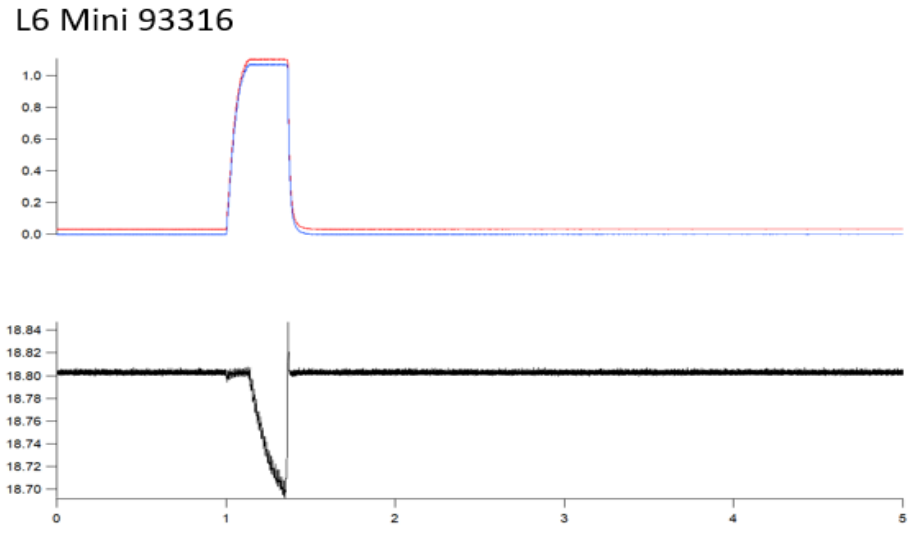
Figure 4.5. Correlation Matrix Correlation matrix between V_{norm} (normalized maximal shortening velocity) and $Endur_{0\text{-}90}$ (linear slope of the first 90 contractions). V_{norm} and $Endur_{0\text{-}90}$ have a negative relationship. Mini-muscle mice (L3 Mini and L6 Mini) have the highest endurance ($Endur_{0\text{-}90}$) but slowest muscles (V_{norm}). L6 and L8 groups have the lowest endurance but fastest muscles. L7 group is intermediate.


L3 M
L6 M
L6
L7
L8
3.50
3.00
2.50
2.00
1.50
-.020
-.015
-.010
-.005
.000
V_{norm} (mm s^{-1})
$Endur_{0-90}$

Appendices

Appendix 4.1. Force and Length Traces
Sample traces of individual force-velocity points in IgorPro 7 used to generate force velocity graphs.

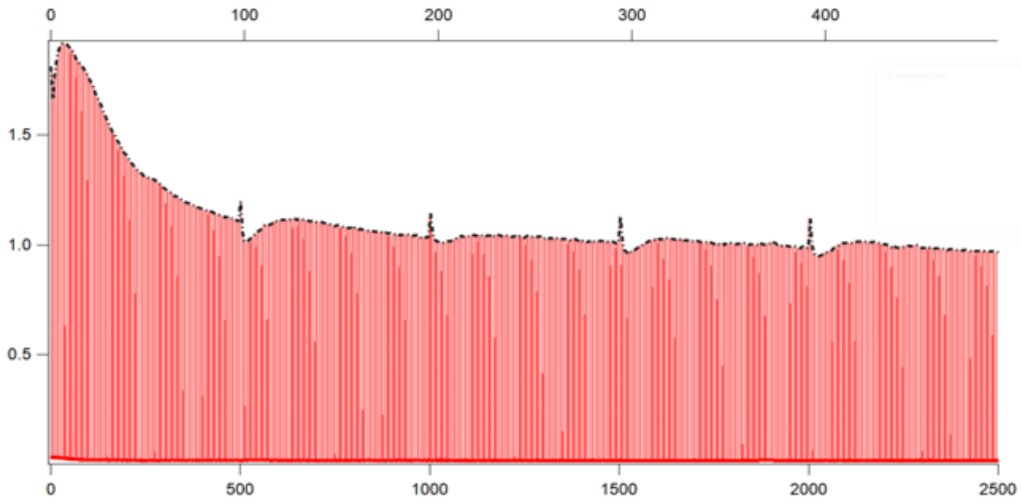




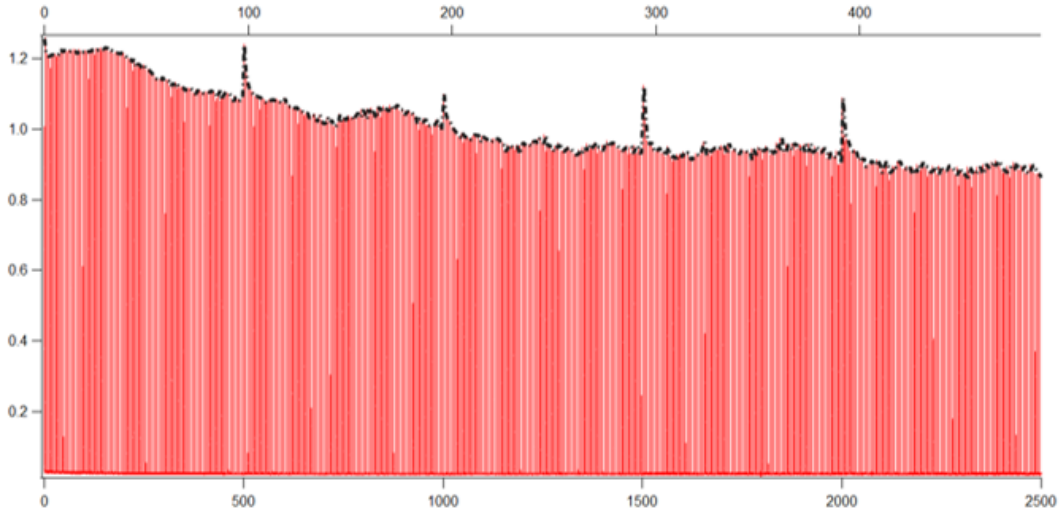
Appendix 4.2. Endurance Traces

Sample traces of repeated isometric contractions in IgorPro 7 used to generate endurance wave profiles.

L6 93472



L6 Mini 93316



Conclusion

Observing a cheetah chasing down prey at high speeds or a bear climbing up a tree to forage, biologists (and people in general) are fascinated with the locomotor abilities of animals. What are the traits that allow a cheetah to cycle their limbs at such high frequencies? Or the characteristics of the bear's paws that allow it to grip on a complex substrate without falling? From a biomechanics perspective, understanding the coadaptation of the musculoskeletal system with locomotor behavior is of key importance, and many comparative studies have revealed correlations between morphological traits and locomotor behavior. For example, cursorial mammals have long and gracile legs that allows an increase in stride length and presumably speed (Carrano 1999). However, such studies are not experimental and unless all species are raised under common conditions they cannot account for possible environmental effects (i.e., phenotypic plasticity, such as training caused by exercise), including potential genotype-by-environmental interactions.

Artificial selection for increased wheel-running behavior in house mice led to multiple correlated responses relating to locomotor performance, morphology, and physiology that are part of their mobility phenotype (Wallace and Garland, 2016). The High Runner (HR) lines of mice have increased treadmill endurance (Meek et al. 2009), increased maximal exercise-induced aerobic capacity (VO_2 max) (Rezende et al. 2006; Garland et al. 2017; Hiramatsu 2017), and achieve faster running speeds on wheels (Garland et al. 2011). Furthermore, HR mice have narrower stance widths, indicating more upright postures (Claghorn et al. 2017), and female HR mice run more intermittently (males were not studied) (Girard et al. 2001). HR mice also have smaller

and leaner bodies (Swallow et al. 1999), increased ventricle mass, which suggests increased stroke volume and presumably cardiac output (Kolb et al. 2010), and larger brains and midbrains, suggesting altered neurobiological function (Kolb et al. 2013).

The "mini-muscle" phenotype presently occurs in a subset of the HR mice, characterized by to a 50% reduction in triceps surae (Garland et al. 2002) and total hindlimb muscle mass (Houle-Leroy et al. 2003), which is caused primarily by a dramatic reduction in type IIb muscle fibers (Guderley et al. 2006; Talmadge et al. 2014). Population-genetic modeling indicates that the mini-muscle phenotype was (unintentionally) under positive selection in the HR lines (Garland et al. 2002). Mini-muscle mice also have larger heart ventricles, increased aerobic capacity (VO_2 max), increased myoglobin concentrations, increased muscle capillarity, higher muscle mitochondrial densities, and increased muscular endurance (Houle-Leroy et al. 2003; Syme et al. 2005; Rezende et al. 2006; Wong et al. 2009; Hiramatsu et al. 2017; Kelly et al. 2017)

I studied mice from the HR model to simultaneously study both the musculoskeletal changes that occurred due to selective breeding and the plasticity of bone and muscle that occurs in response to chronic exercise, including genotype * environment interactions. In my first dissertation chapter, I analyzed skeletal data from generation 11 and, in agreement with to previous studies (Garland and Freeman 2005; Kelly et al. 2006; Middleton et al. 2008), found that HR mice have evolved larger hip and knee surface areas, which would lower stress (force per unit area) acting on the hindlimb during running. The larger articular surface areas of the hip and knee joints in the HR mice are

like the changes documented in the fossil record when comparing the genus Homo with the older *Pan* and *Australopithecus* (Bramble and Lieberman 2004). Thus, my results have relevance for understanding human behavioral and skeletal evolution, given that the evolution of endurance-running capacity has been claimed as a major feature of human biology

Many of the longest-running selection experiments involve laboratory populations of *Drosophila*. Some of these studies have investigated the interrelations of life history traits over a wide variety of selection regimes (Rose 2005; Burke et al. 2016), as well as many generations of laboratory adaptation (Bieri and Kawecki 2003; Kawecki and Mery 2003; Simões et al. 2008, 2019). 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). Little is known about the potential continued adaptation of the skeleton after selection limits have been reached, or if there is a difference in the behavioral response relative to the skeletal phenotype. For my second dissertation chapter, I studied mice from generations 11, 16, 21, 37, 57, and 68 to test whether the skeleton continued to evolve after selection limits were reached. Contrary to my expectation, I 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. I hypothesized 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.

The extent to which various aspects of bone morphology respond to loading are currently under debate in both the paleontological and biomedical community (e.g., Ruff et al. 2006). Some studies have demonstrated that exercise, such as weightlifting or running, makes bones stronger and more robust (Notomi et al. 2001; Mori et al. 2003; Gardinier et al. 2018), while others find little evidence for such effects (Wallace et al. 2015; Peacock et al. 2018). These differing results cast doubt on the practice of attempting to infer the past loading history of individuals from skeletal materials. Results of my muscle attachment study are relevant to these controversies. For my third chapter, I used a sample of mice from generation 57 that were housed with or without wheels for 12 weeks starting at weaning. I used a precise, highly repeatable method for quantifying the three-dimensional (3D) surface area of four muscle attachment sites. In univariate analyses, with body mass as a covariate, mice in the Active group had significantly larger humerus deltoid tuberosities than Sedentary mice, with no significant difference between HR and C mice and no interaction between exercise demonstrating that muscle attachment site morphology can be (but is not always) affected by chronic exercise experienced during ontogeny.

In the locomotor system, the most studied trade-off at the organismal performance level is the negative relationship between speed and endurance that is thought to be underpinned by variation in muscle fiber type composition (see Garland et al. 2022). Trade-offs in organismal performance and in lower-level traits such as muscles may have evolutionary significance. For example, performance trade-offs in the "roll-snap" behavior of male bearded manakins can be explained by female preference for muscle

speed that likely influenced divergence in courtship display (Miles et al. 2018). For my fourth dissertation chapter, I investigated the muscular underpinnings of a negative correlation between average running speed and time spent running on wheels that exists among the HR lines but not among the C lines (Garland et al. 2011). I used *in situ* preparations to measure muscle endurance, twitch characteristics, and force-velocity curves. Maximal shortening velocity varied significantly, being lowest in mini-muscle mice, highest in normal-muscled lines L6 and L8, and intermediate in L7. Endurance, measured as the slope of the decline in force over a series of 90 tetanic contractions, also varied significantly, being shallowest in the mini-muscle mice, steepest in lines L6 and L8, and intermediate in L7. Therefore, muscle-level speed and endurance do trade-off in these mice, but not in a way that maps to the observed organismal-level speed-endurance trade-off.

Overall, my dissertation 1: elucidated mechanisms of skeletal evolution, both during short and long-term selection for voluntary exercise 2: unraveled the effects of chronic exercise on bone morphology, specifically in muscle attachments and 3: explored trade-offs between speed and endurance at the level of muscle physiology. Results have implications for evolutionary physiology, functional morphology, and biomechanics.

References

Bieri J. and T.J. Kawecki. 2003. Genetic architecture of differences between populations of cowpea weevil *Callosobruchus maculatus* evolved in the same environment. Evolution 57:274–287.

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

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

Carrano M.T. 1999. What, if anything, is a cursor? Categories versus continua for determining locomotor habit in mammals and dinosaurs. J Zool 247:29–42.

Claghorn G.C., Z. Thompson, J.C. Kay, G. Ordonez, T.G. Hampton, and T. Garland Jr. 2017. Selective breeding and short-term access to a running wheel alter stride characteristics in house mice. Physiol Biochem Zool 90:533–545.

Gardinier J.D., N. Rostami, L. Juliano, and C. Zhang. 2018. Bone adaptation in response to treadmill exercise in young and adult mice. Bone Rep 8:29–37.

Garland Jr. T., M.D. Cadney, and R.A. Waterland. 2017. Early-life effects on adult physical activity: Concepts, relevance, and experimental approaches. Physiol Biochem Zool 90:1–14.

Garland Jr. T., C. Downs, and A.R. Ives. 2022. Perspective: Trade-offs (and constraints) in organismal biology. Physiol Biochem Zool 95:717897.

Garland Jr. T. and P.W. Freeman. 2005. Selective breeding for high endurance running increases hindlimb symmetry. Evolution 59:1851–1854.

Garland Jr. T., S.A. Kelly, J.L. Malisch, E.M. Kolb, R.M. Hannon, B.K. Keeney, S.L. Van Cleave, et al. 2011. How to run far: Multiple solutions and sex-specific responses to selective breeding for high voluntary activity levels. Proc R Soc B Biol Sci 278:574–581.

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

Girard I., M.W. McAleer, J.S. Rhodes, and T. Garland Jr. 2001. Selection for high voluntary wheel-running increases speed and intermittency in house mice *(Mus domesticus)*. J Exp Biol 204:4311–4320.

Guderley H., P. Houle-Leroy, G.M. Diffee, D.M. Camp, and T. Garland Jr. 2006. Morphometry, ultrastructure, myosin isoforms, and metabolic capacities of the "mini muscles" favoured by selection for high activity in house mice. Comp Biochem Physiol B Biochem Mol Biol 144:271–282.

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

Hiramatsu L., J.C. Kay, Z. Thompson, J.M. Singleton, G.C. Claghorn, R.L. Albuquerque, B. Ho, et al. 2017. Maternal exposure to Western diet affects adult body composition and voluntary wheel running in a genotype-specific manner in mice. Physiol Behav 179:235–245.

Houle-Leroy P., H. Guderley, J.G. Swallow, and T. Garland Jr. 2003. Artificial selection for high activity favors mighty mini-muscles in house mice. Am J Physiol-Regul Integr Comp Physiol 284: R433–R443.

Kawecki T.J. and F. Mery. 2003. Evolutionary conservatism of geographic variation in host preference in *Callosobruchus maculatus*. Ecol Entomol 28:449–456.

Kelly S.A., P.P. Czech, J.T. Wight, K.M. Blank, and T. Garland Jr. 2006. Experimental evolution and phenotypic plasticity of hindlimb bones in high-activity house mice. J Morphol 267:360–374.

Kelly S.A., F.R. Gomes, E.M. Kolb, J.L. Malisch, and T. Garland Jr. 2017. Effects of activity, genetic selection and their interaction on muscle metabolic capacities and organ masses in mice. J Exp Biol 220:1038–1047.

Kolb E.M., S.A. Kelly, and T. Garland Jr. 2013. Mice from lines selectively bred for high voluntary wheel running exhibit lower blood pressure during withdrawal from wheel access. Physiol Behav 112–113:49–55.

Kolb E.M., S.A. Kelly, K.M. Middleton, L.S. Sermsakdi, M.A. Chappell, and T. Garland Jr. 2010. Erythropoietin elevates but not voluntary wheel running in mice. J Exp Biol 213:510–519.

Meek T.H., B.P. Lonquich, R.M. Hannon, and T. Garland Jr. 2009. Endurance capacity of mice selectively bred for high voluntary wheel running. J Exp Biol 212:2908–2917.

Middleton K.M., S.A. Kelly, and T. Garland Jr. 2008. Selective breeding as a tool to probe skeletal response to high voluntary locomotor activity in mice. Integr Comp Biol 48:394–410.

Miles M.C., F. Goller, and M.J. Fuxjager. 2018. Physiological constraint on acrobatic courtship behavior underlies rapid sympatric speciation in bearded manakins. eLife 7:e40630.

Mori T., N. Okimoto, A. Sakai, Y. Okazaki, N. Nakura, T. Notomi, and T. Nakamura. 2003. Climbing exercise increases bone mass and trabecular bone turnover through transient regulation of marrow osteogenic and osteoclastogenic potentials in mice. J Bone Miner Res 18:2002–2009.

Notomi T., N. Okimoto, Y. Okazaki, Y. Tanaka, T. Nakamura, and M. Suzuki. 2001. Effects of tower climbing exercise on bone mass, strength, and turnover in growing rats. J Bone Miner Res 16:166–174.

Peacock S.J., B.R. Coats, J.K. Kirkland, C.A. Tanner, T. Garland, and K.M. Middleton. 2018. Predicting the bending properties of long bones: Insights from an experimental mouse model. Am J Phys Anthropol 165:457–470.

Rezende E.L., T. Garland Jr., M.A. Chappell, J.L. Malisch, and F.R. Gomes. 2006. Maximum aerobic performance in lines of *Mus* selected for high wheel-running activity: effects of selection, oxygen availability and the mini-muscle phenotype. J Exp Biol 209:115–127.

Rose M.R. 2005. The effects of evolution are local: Evidence from experimental evolution in *Drosophila*. Integr Comp Biol 45:486–491.

Ruff C., B. Holt, and E. Trinkaus. 2006. Who's afraid of the big bad Wolff?: "Wolff's law" and bone functional adaptation. Am J Phys Anthropol 129:484–498.

Simões P., I. Fragata, J. Santos, M.A. Santos, M. Santos, M.R. Rose, and M. Matos. 2019. How phenotypic convergence arises in experimental evolution. Evolution 73:1839–1849.

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

Swallow J.G., P. Koteja, P.A. Carter, and T. Garland Jr. 1999. Artificial selection for increased wheel-running activity in house mice results in decreased body mass at maturity. J Exp Biol 202:2513–2520.

Syme D.A., K. Evashuk, B. Grintuch, E.L. Rezende, and T. Garland Jr.,. 2005. Contractile abilities of normal and “mini” triceps surae muscles from mice (*Mus domesticus*) selectively bred for high voluntary wheel running. J Appl Physiol 99:1308–1316.

Talmadge R.J., W. Acosta, and T. Garland. 2014. Myosin heavy chain isoform expression in adult and juvenile mini-muscle mice bred for high-voluntary wheel running. Mech Dev 134:16–30.

Wallace I.J. and T. Garland Jr.,. 2016. Mobility as an emergent property of biological organization: Insights from experimental evolution: Mobility and biological organization. Evol Anthropol Issues News Rev 25:98–104.

Wallace I.J., S. Judex, and B. Demes. 2015. Effects of load-bearing exercise on skeletal structure and mechanics differ between outbred populations of mice. Bone 72:1–8.

Wong L.E., T. Garland, S.L. Rowan, and R.T. Hepple. 2009. Anatomic capillarization is elevated in the medial gastrocnemius muscle of mighty mini mice. J Appl Physiol 106:1660–1667.