© 2015. Published by The Company of Biologists Ltd | Journal of Experimental Biology (2015) 218, 3002-3009 doi:10.1242/jeb.118729

RESEARCH ARTICLE

Focal enhancement of the skeleton to exercise correlates with responsivity of bone marrow mesenchymal stem cells rather than peak external forces

Ian J. Wallace[1,*,‡,§], Gabriel M. Pagnotti[2,‡], Jasper Rubin-Sigler[3], Matthew Naeher[1], Lynn E. Copes[4], Stefan Judex[2], Clinton T. Rubin[2] and Brigitte Demes[3]

ABSTRACT

Force magnitudes have been suggested to drive the structural response of bone to exercise. As importantly, the degree to which any given bone can adapt to functional challenges may be enabled, or constrained, by regional variation in the capacity of marrow progenitors to differentiate into bone-forming cells. Here, we investigate the relationship between bone adaptation and mesenchymal stem cell (MSC) responsivity in growing mice subject to exercise. First, using a force plate, we show that peak external forces generated by forelimbs during quadrupedal locomotion are significantly higher than hindlimb forces. Second, by subjecting mice to treadmill running and then measuring bone structure with μCT, we show that skeletal effects of exercise are site-specific but not defined by load magnitudes. Specifically, in the forelimb, where external forces generated by running were highest, exercise failed to augment diaphyseal structure in either the humerus or radius, nor did it affect humeral trabecular structure. In contrast, in the ulna, femur and tibia, exercise led to significant enhancements of diaphyseal bone areas and moments of area. Trabecular structure was also enhanced by running in the femur and tibia. Finally, using flow cytometry, we show that marrow-derived MSCs in the femur are more responsive to exercise-induced loads than humeral cells, such that running significantly lowered MSC populations only in the femur. Together, these data suggest that the ability of the progenitor population to differentiate toward osteoblastogenesis may correlate better with bone structural adaptation than peak external forces caused by exercise.

KEY WORDS: Physical activity, Ground reaction forces, Bone adaptation, Cortical bone, Trabecular bone, Osteoprogenitor cells, Mechanical loading

INTRODUCTION

Numerous aspects of organismal structure and function can be positively affected by load-bearing exercise, not least of which is the skeleton. Studies of humans and other animals frequently demonstrate the potential for activities such as running and jumping to promote bone formation, delay bone loss, and, ultimately, enhance skeletal structure and strength (e.g. Woo et al., 1981; Biewener and Bertram, 1994; Lieberman et al., 2003; McKay et al., 2005; Weeks et al., 2008). Yet, as potent an anabolic stimulus as exercise may be, the parameters of the regimen that stimulate bone's adaptive response remain difficult to identify, nor is it clear if all elements of the skeleton are equally responsive to loading.

It is often presumed that skeletal benefits of exercise are proportional to load magnitudes, and thus the greatest structural enhancement is expected where the greatest functional challenges occur (Tommerup et al., 1993; Nikander et al., 2006; Wilks et al., 2009). A well-known example of focal adaptation to physical demand is the humeral hypertrophy of the playing arms of tennis players and baseball pitchers relative to their non-playing arms (Jones et al., 1977; Haapasalo et al., 2000; Warden et al., 2014). Implicitly, these athletes demonstrate the ability of bone to perceive and respond to biophysical signals in a site-specific manner, perhaps to maintain an optimal mechanical environment to cells regulating the response (Rubin, 1984).

Notwithstanding load magnitudes, it is also possible that the site-specificity of skeletal adaptation to exercise is defined to some extent by regional variation in the mechanoresponsiveness of bone progenitor cells. The mesenchymal stem cell (MSC) is a pluripotent progenitor that has the ability to differentiate into cells that form bone, fat, tendon, muscle and ligament. Accumulating evidence indicates that fate selection of marrow-based MSCs is influenced by mechanical signals, biasing differentiation of MSCs toward bone and away from fat (Luu et al., 2009a; Ozcivici et al., 2010; Chan et al., 2012; Pagnotti et al., 2012). It has further been shown that bone marrow in some skeletal elements may contain MSCs with greater osteogenic potential than those located in the marrow of other elements (Risbud et al., 2006; Volk et al., 2012), which would suggest that some bones may be more responsive to exercise-induced mechanical signals than others.

In this study, we used a mouse model to examine the site-specificity of the adaptive response to treadmill-running exercise in five weight-bearing limb bones (humerus, radius, ulna, femur and tibia) and a non-weight-bearing cranial bone (interparietal). If the presumption that higher loads cause more structural enhancement is true, then anabolic adaptation would be expected to be greatest in those skeletal regions subject to the highest loads. Alternatively, if the adaptive response to exercise failed to correlate with load magnitude, perhaps it would better associate with a biological rather than a mechanical parameter, specifically, the osteoprogenitor population.

RESULTS

Limb forces generated during locomotion

To assess the relationship between the magnitude of loads produced by exercise and the site-specific osteogenic response, we first

[1]Department of Anthropology, Stony Brook University, Stony Brook, NY 11794, USA. [2]Department of Biomedical Engineering, Stony Brook University, Stony Brook, NY 11794, USA. [3]Department of Anatomical Sciences, Stony Brook University, Stony Brook, NY 11794, USA. [4]Department of Medical Sciences, Quinnipiac University, Hamden, CT 06518, USA.
*Present address: Department of Human Evolutionary Biology, Harvard University, Cambridge, MA 02138, USA.
‡These authors contributed equally to this work

§Author for correspondence (iwallace@fas.harvard.edu)

Received 21 December 2014; Accepted 20 July 2015

Journal of Experimental Biology

Journal of Experimental Biology (2015) 218, 3002-3009 doi:10.1242/jeb.118729

List of symbols and abbreviations

A_{cortex}	cortical area
A_{marrow}	marrow area
A_{total}	total area
I_{max}	maximum second moment of area
I_{min}	minimum second moment of area
MSC	mesenchymal stem cell
N_{trab}	trabecular number
S_{trab}	trabecular separation
T_{cortex}	mean cortical thickness
T_{dip}	mean diploë thickness
T_{trab}	trabecular thickness
T_{total}	mean total thickness
V_{bone}	bone volume
V_{total}	total volume

determined whether the forelimbs or hindlimbs experience greater external forces during mouse quadrupedal locomotion. To this end, the vertical component of peak ground reaction forces generated by steady-state symmetrical gaits was measured in a sample of outbred female mice (N=10; 6–7 weeks old) using a force plate. Over a range of speeds chosen by the animals, involving both walking and running gaits, peak forces generated by the forelimbs were significantly higher than those experienced by the hindlimbs (P<0.0001; Fig. 1). Therefore, if the skeleton's adaptive response to exercise correlates with load magnitude, then elevated locomotor activity would be expected to augment bone structure more in the forelimbs than the hindlimbs.

Effects of running on skeletal structure

The skeletal effects of exercise were established through an experiment in which a second cohort of female mice from the same outbred stock were either treated with a treadmill-running regimen for 4 weeks (N=20) or served as sedentary controls (N=20), beginning shortly after weaning (4 weeks of age). Animals assigned to the treadmill group ran 5 days week^{-1} for 30 min day^{-1}.

At the beginning of the experiment, body mass was not significantly different between the experimental groups (P=0.15; Fig. 2A). In both groups, body mass increased steadily over the course of the experiment, with the most rapid increase occurring during the first week. By the end of the experiment, body mass remained similar between controls and runners (P=0.49). Limb muscle mass was also not significantly different between groups at the end of the experiment (triceps brachii: P=0.32, quadriceps: P=0.29, triceps surae: P=0.49; Fig. 2B). Home-cage activity quantified at the end of week 3 of the exercise program was comparable between groups (Fig. 2C), such that the total number of ambulatory movements over 24 h did not significantly differ between controls and runners (P=0.50; Fig. 2D). Thus, any differences in skeletal structure between groups can be confidently attributed to exercise treatment, rather than group differences in body size, muscle mass or cage activity.

The effects of treadmill running on skeletal structure (quantified with μCT) were found to vary greatly across anatomical regions (Table 1; Fig. 3). In the cranium, where loads produced by running were presumably lowest, the thickness of interparietal cortical bone was not significantly altered by exercise treatment (P=0.19; Fig. 3A,B), nor was interparietal total thickness (P=0.20). Similarly, in the forelimb, where external loads generated by running were highest, exercise failed to significantly augment mid-diaphyseal cortical bone structure in either the humerus or radius (Fig. 3C). By contrast, in the ulna, running led to 6.9±3.0% greater

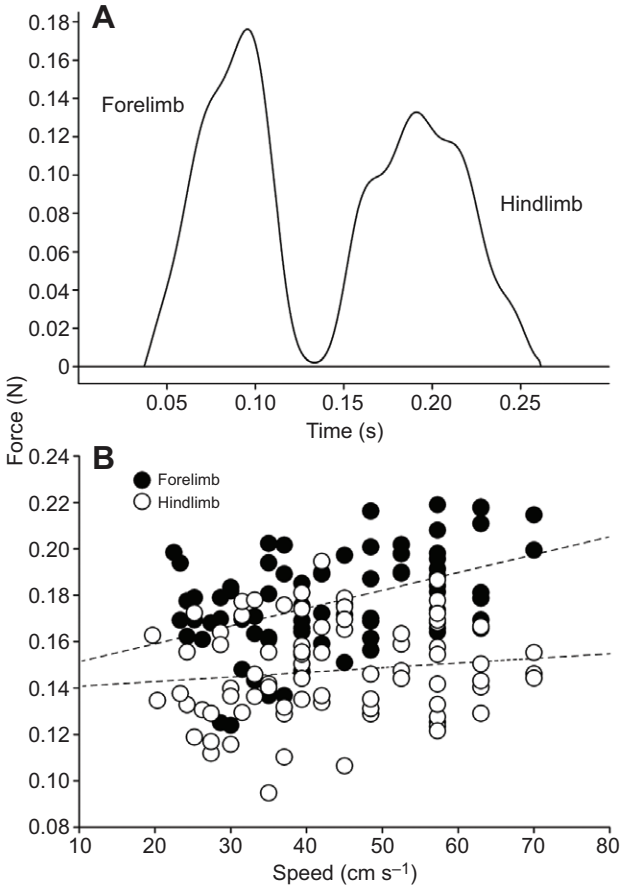

Fig. 1. Mouse forelimbs generated higher peak vertical ground reaction forces than hindlimbs during quadrupedal locomotion. (A) Representative vertical ground reaction force trace from a 22.3 g mouse running at 39.4 cm s^{-1}. (B) Bivariate plot of peak vertical ground reaction forces sustained by forelimbs and hindlimbs versus subject speed. Plotted lines were determined by least-squares regression. P<0.0001 for ANCOVA comparison between forelimb and hindlimb forces with speed as a covariate.

mid-diaphyseal cortical area (P=0.024) and 7.9±2.8% greater total (subperiosteal) area (P<0.01), as well as a 15.7±7.2% increase in maximum second moment of area (P=0.036) and a 15.4±6.6% increase in minimum second moment of area (P=0.026). In the hindlimb, where running-induced external load magnitudes were lower than in the forelimb, exercise resulted in several significant cortical bone structural enhancements in both the femur and tibia. Compared with controls, runners had femora with 11.4±3.1% larger mid-diaphyseal cortical area (P<0.001), a 9.7±2.6% thicker cortex (P<0.001) and 5.9±2.7% larger total area (P=0.032). In runners, femoral mid-diaphyseal maximum and minimum second moments of area were increased by 17.2±6.0% (P<0.01) and 16.2±5.6% (P<0.01), respectively. In the tibia, relative to controls, runners had 11.0±3.3% larger mid-diaphyseal cortical area (P<0.01), a 3.6±1.5% thicker cortex (P=0.025) and 9.1±3.1% larger total area (P<0.01). Tibial mid-diaphyseal maximum and minimum second moments of area were enhanced in runners by 20.2±6.4% (P<0.01) and 21.0±6.6% (P<0.01), respectively.

Responsivity of trabecular bone to exercise was also observed to vary among limb elements (Table 2; Fig. 3D). In the forelimb, exercise failed to significantly affect trabecular morphology in the distal humeral epiphysis. By contrast, in the hindlimb, trabecular bone volume fraction in the distal femoral metaphysis was 24.1±11.4% greater in runners compared with controls (P=0.041) and

Journal of Experimental Biology (2015) 218, 3002-3009 doi:10.1242/jeb.118729

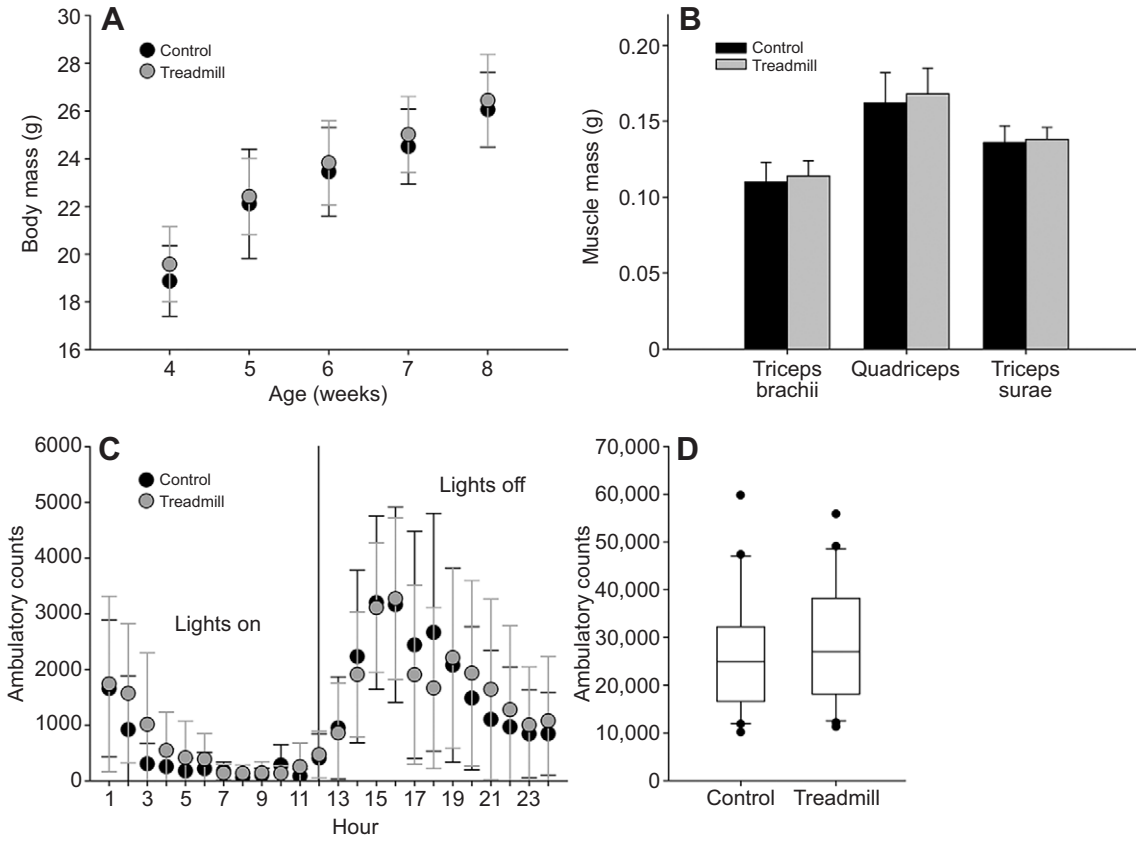

Fig. 2. Treadmill exercise did not alter mouse body mass, muscle mass or home-cage activity level. (A) Change in mean body mass (±s.d.) among controls and treadmill runners during the experiment. (B) Mean limb muscle mass (+s.d.) among controls and runners at the end of the experiment. (C) Mean hourly ambulatory counts (±s.d.) among controls and runners throughout a 24 h period on a non-running day. (D) Total number of ambulatory counts among controls and runners during the 24 h. Boxes, median±first and third quartile; whiskers, median±range of non-outliers; circles, outliers. $P>0.05$ for all comparisons between controls and treadmill runners.

trabecular number was 12.2±5.8% higher (P=0.043). Moreover, in the proximal tibial metaphysis, trabeculae were 9.3±4.1% thicker as a result of running (P=0.028).

Marrow-derived mesenchymal stem cells

The responsivity to exercise of marrow-derived MSCs in the humerus and femur was determined by treating a third cohort of female mice from the same outbred stock (N=6) to 5 days of treadmill running (30 min day^{-1}) beginning at 6 weeks of age and comparing MSC populations at the end of the experimental period (quantified with flow cytometry) with those obtained from a sample of age-matched controls (N=17). In the humerus, MSCs in the bone marrow population, identified as positive for SCA-1, c-Kit, CD90.2, CD105 and CD44 surface antigens (Fig. 4A,B), did not differ between runners and controls (P=0.39; Fig. 4C), whereas in the femur, running significantly lowered the number of MSCs (P=0.044), implying that they had been signaled away from their progenitor state. The number of MSCs relative to marrow cavity volume was not significantly different between the two bones in controls (P=0.61; Fig. 4D), indicating that baseline density of MSCs in the humerus and femur was similar.

DISCUSSION

This study demonstrates the adaptive capacity of the mouse skeleton to be site-specific, with some regions being more responsive than others to exercise-induced loads. Treadmill running stimulated significant cortical bone structural enhancement in the femur, tibia and ulna, whereas cortical structure failed to respond in the humerus and radius, despite external load magnitudes being significantly higher in the forelimb as opposed to the hindlimb. Trabecular structure was also significantly enhanced by exercise in the femur and tibia, but exercise failed to initiate an anabolic response in the humerus. Cranial bone morphology was not influenced by running, emphasizing the absence of a systemic stimulus for bone growth. These results are inconsistent with any hypothesis predicting that skeletal adaptation is proportional to external load magnitudes and oppose the idea that exercise releases a systemic humor that stimulates a generic response.

If skeletal adaptation to exercise was not correlated with load magnitude, what biological parameter might have modulated variation in bone responsivity? That the density of MSCs per unit volume of marrow space was comparable in the humerus and femur of controls undermines any conclusion that bases the differential structural response of those two bones on a difference in baseline availability of osteoprogenitors. There was, however, a significant reduction in the number of MSCs in the femora of exercised animals compared with controls, but no significant difference was observed in the humerus. First, these data imply that cortical and trabecular enhancement of the femur might be related to the recruitment of MSCs toward osteoblastogenesis (Luu et al., 2009a). Although the cortical response primarily affected the periosteal surface rather than the endosteal surface, this does not preclude the possibility that marrow-derived MSCs contributed to the observed cortical growth, as there is evidence that MSCs migrate outside the marrow cavity to

Journal of Experimental Biology (2015) 218, 3002-3009 doi:10.1242/jeb.118729

Table 1. Mouse cortical bone properties in controls and treadmill runners

Region	Measure	Control	Treadmill	*P*
Interparietal	T_{total}	0.260±0.026	0.270±0.019	0.20
	T_{dip}	0.104±0.015	0.108±0.012	0.36
	T_{cortex}	0.078±0.007	0.081±0.006	0.19
Humerus	A_{total}	0.86±0.05	0.87±0.05	0.43
	A_{marrow}	0.31±0.05	0.31±0.05	0.76
	A_{cortex}	0.55±0.04	0.56±0.04	0.21
	T_{cortex}	0.212±0.020	0.218±0.023	0.37
	I_{max}	0.063±0.008	0.065±0.008	0.30
	I_{min}	0.038±0.005	0.040±0.005	0.22
Radius	A_{total}	0.33±0.02	0.33±0.03	0.34
	A_{marrow}	0.08±0.01	0.08±0.02	0.54
	A_{cortex}	0.25±0.02	0.25±0.02	0.44
	T_{cortex}	0.166±0.014	0.167±0.013	0.82
	I_{max}	0.008±0.001	0.009±0.001	0.31
	I_{min}	0.007±0.001	0.007±0.001	0.39
Ulna	A_{total}	0.33±0.03	0.36±0.03	**0.008**
	A_{marrow}	0.06±0.02	0.06±0.01	0.12
	A_{cortex}	0.28±0.03	0.30±0.03	**0.024**
	T_{cortex}	0.193±0.017	0.197±0.014	0.50
	I_{max}	0.017±0.004	0.020±0.003	**0.036**
	I_{min}	0.004±0.001	0.005±0.001	**0.026**
Femur	A_{total}	1.72±0.13	1.82±0.15	**0.032**
	A_{marrow}	0.90±0.10	0.91±0.10	0.78
	A_{cortex}	0.82±0.08	0.91±0.08	**0.0007**
	T_{cortex}	0.203±0.019	0.223±0.015	**0.0007**
	I_{max}	0.210±0.036	0.246±0.043	**0.007**
	I_{min}	0.125±0.020	0.146±0.024	**0.007**
Tibia	A_{total}	0.84±0.08	0.92±0.08	**0.006**
	A_{marrow}	0.26±0.04	0.28±0.04	0.27
	A_{cortex}	0.58±0.06	0.64±0.05	**0.002**
	T_{cortex}	0.21±0.01	0.22±0.01	**0.025**
	I_{max}	0.060±0.013	0.073±0.012	**0.003**
	I_{min}	0.043±0.009	0.052±0.009	**0.003**

Values are means±s.d. T_{total}, mean total thickness (mm); T_{dip}, mean diploë thickness (mm); T_{cortex}, mean cortical thickness (mm); A_{total}, total area (mm^2); A_{marrow}, marrow area (mm^2); A_{cortex}, cortical area (mm^2); I_{max}, maximum second moment of area (mm^4); I_{min}, minimum second moment of area (mm^4). Bold *P*-values represent statistically significant ($P<0.05$) differences between controls and treadmill runners.

Table 2. Mouse trabecular bone properties in controls and treadmill runners

Bone	Trait	Control	Treadmill	*P*
Humerus	V_{bone}/V_{total}	49.4±5.3	50.5±3.9	0.44
	N_{trab}	10.6±1.1	10.5±1.2	0.81
	T_{trab}	65.9±4.6	66.2±4.2	0.87
	S_{trab}	103±15	102±14	0.82
Femur	V_{bone}/V_{total}	15.1±5.7	18.7±5.2	**0.041**
	N_{trab}	5.5±1.1	6.2±0.9	**0.043**
	T_{trab}	45.2±5.1	48.2±6.1	0.10
	S_{trab}	195±38	172±27	**0.035**
Tibia	V_{bone}/V_{total}	12.6±6.4	15.9±5.9	0.09
	N_{trab}	5.5±1.2	5.9±1.1	0.25
	T_{trab}	40.5±5.6	44.2±4.8	**0.028**
	S_{trab}	198±40	184±32	0.22

Values are means±s.d. V_{bone}/V_{total}, bone volume fraction (%); N_{trab}, trabecular number ($1\ mm^{-1}$); T_{trab}, trabecular thickness (μm); S_{trab}, trabecular separation (μm). Bold *P*-values represent statistically significant ($P<0.05$) differences between controls and treadmill runners.

be recruited at the periosteum and other sites (Luu et al., 2009b). Second, these data indicate that MSCs in the humerus were less responsive to mechanical signals, either because of reduced cellular mechanosensitivity per se (Volk et al., 2012), or perhaps due to aspects of the marrow niche (e.g. adjacency to hematopoietic stem cells, higher concentration of adipocytes, marrow viscosity) that confounded the ability of MSCs to commit to an osteoblast lineage (Adler et al., 2014; Metzger et al., 2015). These interpretations must be considered with caution, because changes in MSC numbers alone are not a direct indication of differentiation activity toward osteoblast production, and we did not directly show here that osteoblast numbers increased with exercise treatment. Nevertheless, the observed associations between changes in MSCs and bone structure are compelling.

Teleologically, it is not clear why mouse forelimb bones would be less responsive to exercise than hindlimb bones, although it may relate to maintaining a lightly built architecture for the appendage that is also routinely used for manipulation. Yet, from a broad evolutionary perspective, it is easier to imagine how it could be advantageous in some instances for certain bones to be more or less affected by mechanical signals. In migratory birds, for example, if routinely large wing forces were to lead to equally large gains in bone mass, then this would probably deleteriously increase energetic costs of flight (Dumont, 2010). Similarly, if woodpecker crania were to become progressively thickened with every bout of tree drumming, then this could constrict tiny foramina and damage vital nerves and blood vessels. At the other extreme, bones that rarely sustain large loads, such as elements of the inner ear, would be likely to erode away over time (Schutz et al., 2014). Thus, evolution of bone adaptability can be expected to have involved some degree of fine-tuning of responsivity across different skeletal locations in order to achieve site-specific compromises between maintaining equilibrium mechanical conditions while still ensuring that bones preserve optimal structure for their respective functional roles (Lieberman and Crompton, 1998; Currey, 2002; Lieberman et al., 2003). Bone progenitors are perhaps just one of several cells that orchestrate such regional variation in skeletal responsivity.

It is important to emphasize that to characterize the mechanical environment of the forelimbs and hindlimbs during exercise, we relied on a force plate to quantify only peak external limb forces. We did not attempt, however, to calculate internal forces (stresses) or measure *in vivo* strains experienced by the individual bones during running. Lacking bone stress and strain data, we cannot exclude the possibility that differences in mechanical signals experienced locally by forelimb and hindlimb bones were not as great as suggested by external limb forces. Indeed, this seems likely in light of evidence that peak bone stresses/strains are generally similar between homologous limb elements (e.g. humerus, femur) among quadrupeds (Biewener, 1983) and across different limb elements among a variety of animals (Rubin and Lanyon, 1984; Biewener, 1990). Such broad uniformity in local mechanical signals is possible due to numerous suites of evolutionary adaptations related to anatomy (e.g. skeletal allometry, bone curvature) and locomotor kinematics and kinetics (e.g. speed, joint posture) that modulate variation in external forces (Biewener, 1990). However, even with internal forces and tissue deformation magnitudes being broadly similar, the difference in osteogenic response between proximal forelimb and hindlimb elements of the mice in this study is clear. Therefore, it is plausible that our conclusion that the responsiveness of bone to exercise-induced loading was not proportional to external force magnitude could also apply to local stresses and strains; but a rigorous test of this hypothesis would require a different approach than that adopted here.

Beyond possible implications of this study for the physiological response of bone to exercise, this study is also relevant for paleontologists, including paleoanthropologists, who attempt to

Journal of Experimental Biology (2015) 218, 3002-3009 doi:10.1242/jeb.118729

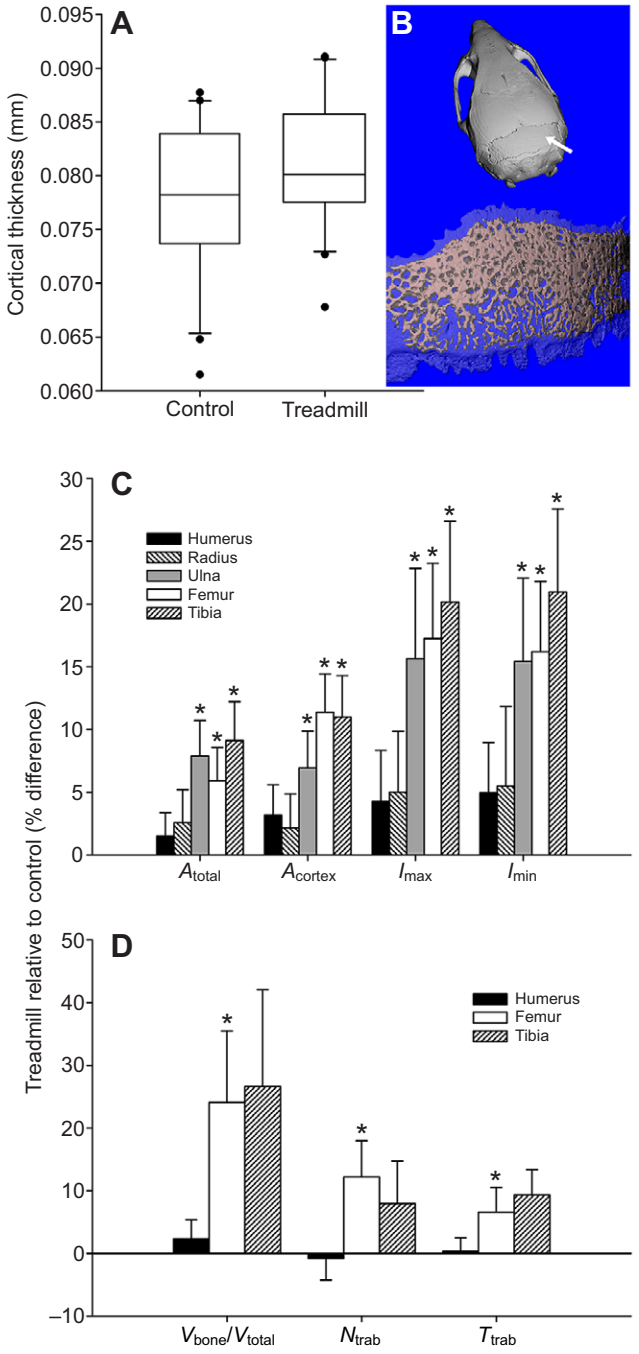

Fig. 3. Skeletal adaptation to treadmill exercise was site-specific. (A) Interparietal cortical bone thickness among controls and treadmill runners. Boxes are median±first and third quartile; whiskers, median±range of non-outliers; circles, outliers. (B) μCT images of a mouse cranium showing the location of the interparietal (top, arrow) and the separation of cortical bone from diploë (bottom). Relative differences in (C) limb bone cortical properties and (D) trabecular properties between controls and runners. A_{total}, total area; A_{cortex}, cortical area; I_{max} and I_{min}, maximum and minimum second moments of area; V_{bone}, bone volume; V_{total}, total volume; N_{trab}, trabecular number; T_{trab}, trabecular thickness. Differences between group means are plotted as percentage difference+s.d. of the sampling distribution of the relative difference. *P<0.05 for comparisons between controls and treadmill runners. P>0.05 for all other differences.

deduce the physical activity patterns of ancient species based on the morphology of fossils (Ruff, 2005). Frequently, researchers compare the diaphyseal structure of hominin upper limbs with that of lower limbs in order to infer the emphasis placed on activities involving upper limb loading, such as arboreal climbing, relative to lower limb loading, particularly bipedal locomotion (e.g. Ruff, 2008, 2009). This approach assumes that upper and lower limb bones respond in the same way to loading, such that their diaphyseal structure similarly reflects loading history. The results of this study do not support this presumed homogeneity in bone responsiveness across limbs. Moreover, paleoanthropologists have interpreted thick cranial bones in ancient hominins as evidence of high physical activity levels, based on the assumption that physical activity systemically enhances skeletal structure (e.g. Lieberman, 1996, 1998). This assumption is also not supported by our findings. Overall, the results of this study suggest that effects of physical activity on the skeleton can be unpredictable and, likewise, that caution is necessary when using fossils to glean information about the behavioral patterns of ancient species.

In summary, this study provides yet another example of the importance of load-bearing exercise in influencing skeletal structure and strength, in this case, in a site-specific manner. Ultimately, this study supplies evidence (albeit preliminary) that the MSC, the cell that must be recruited to form bone, may be differentially responsive to mechanical signals, either because of variation in cellular mechanosensitivity or because the niche in which it resides is influencing the ability of the progenitor to commit to a particular tissue lineage. Additional research of this subject is warranted because it could yield important insights pertinent to the design of exercise programs and technologies aimed at improving human skeletal structure and strength with biophysical signals. It would also aid our understanding of bone adaptation to mechanical loads, thus providing a more solid basis for interpreting bone structure from a functional point of view.

MATERIALS AND METHODS

Animals

All experiments and procedures were reviewed and approved by the Stony Brook University Institutional Animal Care and Use Committee. At weaning age, female Hsd:ICR mice (N=73) were acquired from Harlan Sprague Dawley (Indianapolis, IN, USA) and housed individually in standard cages and maintained on a 12 h:12 h light:dark cycle with free access to water and chow. At 4 weeks of age, animals were divided into three groups: cohort 1 was used to measure limb forces (N=10), cohort 2 was used to assess the skeletal effects of treadmill exercise (N=40), and cohort 3 was used to examine the responsivity of marrow-derived MSCs to exercise (N=23). Young animals were used because skeletal adaptability to exercise is typically greatest during the growing years (Rubin et al., 1992; Lieberman et al., 2003; Main et al., 2014). At the conclusion of each experiment, animals were killed by cervical dislocation preceded by inhalation of either CO_2 (cohorts 1 and 2) or isoflurane (cohort 3).

Limb force recordings

The design, calibration and use of the force plate has been described in detail elsewhere (Riskin et al., 2009; Wallace et al., 2015) and is only briefly summarized here. The plate consists of a stiff fiberfoam surface (28 cm×28 cm) supported at two opposite ends by aluminium beams with spring blade elements fitted with strain gauges. Signals from gauges were amplified, digitized and saved to a computer. The plate was modified for collection of single-limb contacts by affixing a thin wooden strip to the plate surface and covering the rest of the plate with a runway. Animals in cohort 1 traveled freely in a clear tunnel (44 cm long×7 cm wide) placed over the runway. The relationship between vertical force magnitude and voltage output was determined by applying a range of known weights to the wooden strip. Regressions of force to voltage were linear (R^2>0.999). Ground reaction forces recorded at 1000 Hz and video recorded in lateral view at 500 Hz were synchronized with a ProCapture motion analysis system (Xcitex, Inc., Woburn, MA, USA). Animal speed was determined from

Journal of Experimental Biology (2015) 218, 3002-3009 doi:10.1242/jeb.118729

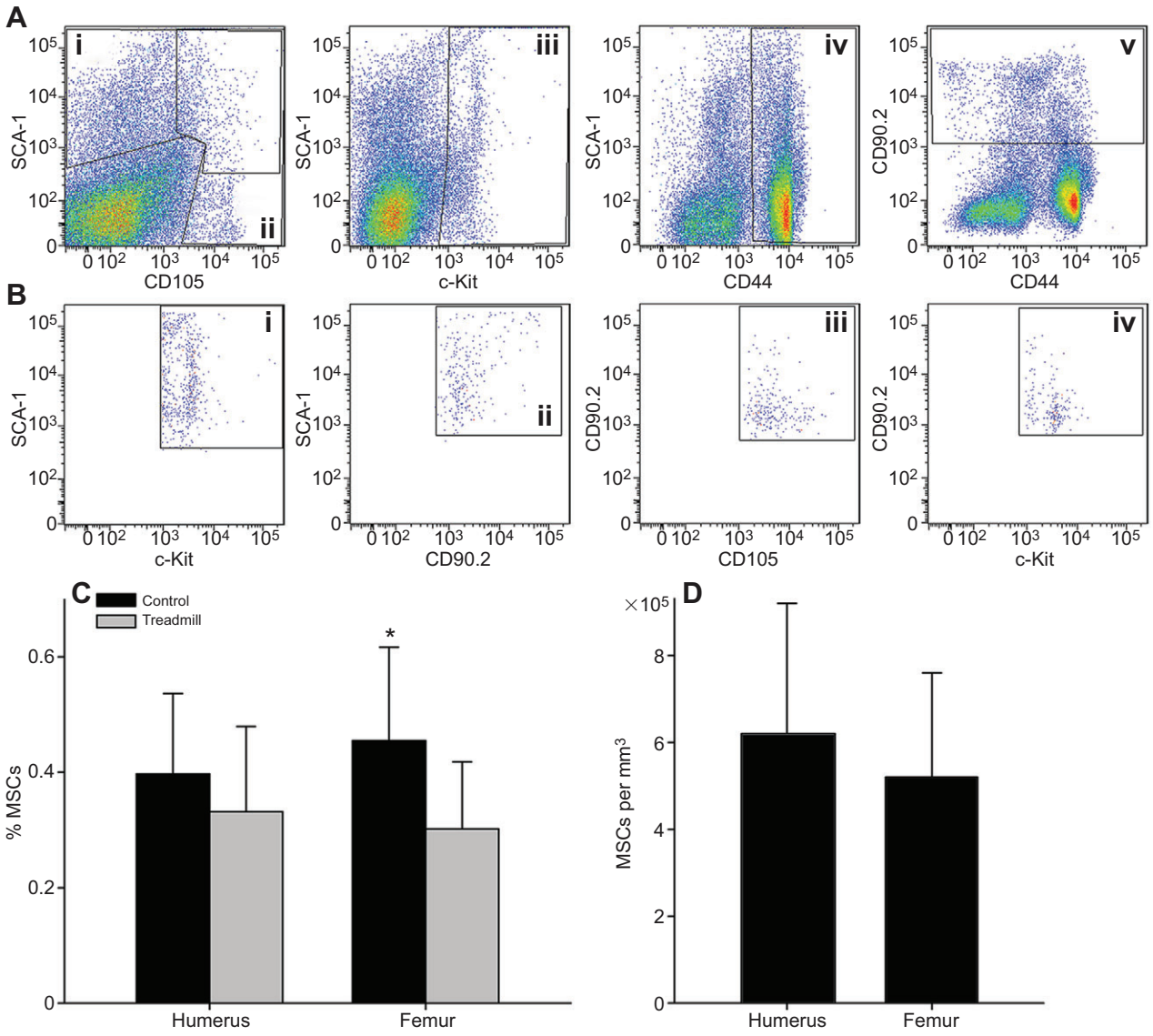

Fig. 4. Responsiveness of bone marrow mesenchymal stem cells to treadmill exercise. (A) Flow cytometric analysis illustrating the marrow cell population of a representative femur, with cells that positively stained with fluorescent antibody surface markers within the gates shown: (i) 10.3% SCA-1$^+$; (ii) 3.65% CD105$^+$; (iii) 4.63% c-Kit$^+$; (iv) 60.9% CD44$^+$; (v) 6.09% CD90.2$^+$. (B) Subgating of surface markers selecting for discrete MSC populations: (i) 99.4% SCA-1$^+$, c-Kit$^+$; (ii) 99.3% SCA-1$^+$, c-Kit$^+$, CD90.2$^+$; (iii) 100% SCA-1$^+$, c-Kit$^+$, CD90.2$^+$, CD105$^+$; (iv) 100% SCA-1$^+$, c-Kit$^+$, CD90.2$^+$, CD105$^+$, CD44$^+$. (C) Percentage of total gated cells identified as MSCs in the humerus and femur of controls and treadmill runners. *P<0.05 for between-group comparison. (D) The number of MSCs per unit of marrow cavity volume in the humerus and femur of controls (P>0.05). Data are means+s.d.

video as the time required for a fixed anatomical landmark (i.e. the nose) to pass between markers on either side of the runway. Force data were filtered with a 30 Hz low-pass Butterworth filter and extracted with ProAnalyst software (Xcitex, Inc.). Data from between 5 and 10 forelimb and hindlimb contacts were collected per animal (N=78 total for each limb).

Treadmill exercise

Runners in cohorts 2 and 3 (N=20 and 6, respectively) were exercised 5 days week^{-1} on a six-lane treadmill (Columbus Instruments, Columbus, OH, USA) at a rate of 12 m min^{-1} for 30 min day^{-1}. Exercise treatment of cohort 2 began at 4 weeks of age and lasted for 4 weeks, and treatment of cohort 3 began at 6 weeks of age and lasted for 1 week. Controls in cohorts 2 and 3 (N=20 and 17, respectively) were handled similar to exercising mice but did not run.

Home-cage activity

Home-cage activity of animals in cohort 2 was quantified with a 16-chamber Opto-M3 system (Columbus Instruments), in which each cage was placed into an apparatus that cast a grid of infrared beams (1.27 cm^2) above the cage floor. No additional stimulus was supplied by the system and animals retained free access to water and chow. Activity was quantified as the total number of times that a new beam was broken. Recurrent interruptions of beams caused by non-ambulatory movements (e.g. scratching, grooming) were disregarded.

Muscle mass

In cohort 2, right and left triceps brachii, quadriceps femoris and triceps surae muscles were removed with a scalpel and weighed on an electronic balance (Acculab VI-3mg; Sartorius AG, Göttingen, Germany). Right and left muscle mass values were averaged. Triceps brachii was cut at its proximal origins at the humeral neck and greater tubercle and the caudal scapular border, and at its distal insertion at the ulnar olecranon process. Quadriceps was cut proximally from its origins at the cranial aspect of the femur, the femoral neck, and the greater and third trochanters, and distally at its insertion at the patellar ligament. Triceps surae was cut at its proximal

origins at the medial and lateral femoral epicondyles and the fibular head, and at its distal insertion at the calcaneal tendon.

Microcomputed tomography

In cohort 2, right limb bones were scanned in distilled water at a 10 μm^3 voxel size (70 kVp, 114 μA, 150 ms integration time) using a μCT 40 scanner (Scanco Medical AG, Brüttisellen, Switzerland). Cortical bone structure was assessed in a 300-μm-long region of the tibial diaphysis and in 600-μm-long regions of the diaphyses of other limb bones. The humeral volume of interest (VOI) began 100 μm distal to the deltoid tuberosity, and the VOI in the other four limb elements was centered at half bone length. A smaller VOI for the tibia relative to other bones was deemed appropriate given the large degree of longitudinal structural variation along the shaft of that element (Bab et al., 2007). Trabecular bone was assessed in a 300-μm-long region of the distal humeral epiphysis starting 100 μm proximal to the cortical–trabecular bone interface, in a 1100-μm-long region of the distal femoral metaphysis starting 850 μm proximal to the growth plate, and in a 600-μm-long region of the proximal tibial metaphysis starting 500 μm distal to the growth plate. Although most studies of trabecular bone in rodents focus on bone tissue in the metaphysis rather than the epiphysis (Bouxsein et al., 2010), the humeral distal epiphysis was analyzed in this study because of the low quantity of trabecular bone in mouse humeral metaphyses (Bab et al., 2007). Volumes were segmented using a constrained 3D-Gaussian filter to reduce noise [support=1, sigma=0.1 (diaphysis) and 0.5 (epiphysis and metaphysis)] and thresholded to extract the bone phase. Threshold values chosen for cortical and trabecular bone (593.1 and 428.1 mg HA/cm^3, respectively) were determined empirically to achieve maximal concordance between raw and thresholded images (Judex et al., 2004). Trabecular bone was isolated from the cortical shell using an automated algorithm (Lublinsky et al., 2007). Interparietals were scanned using the same scanner settings as those used for limb bones. Volumes were filtered (3D Gauss; support=1, sigma=0.1) and thresholded at 634.4 mg HA/cm^3. Cortical layers were separated from the diploë using an automated algorithm (Copes, 2012). All bone properties were computed using the internal imaging code supplied by the scanner manufacturer, and were defined as recommended (Bouxsein et al., 2010). Structural data from the femur have been reported previously (Wallace et al., 2015).

In cohort 3, right humeri and femora were scanned dry at a 20 μm^3 voxel size (70 kVp, 114 μA, 150 ms integration time). μCT image processing followed the same protocol as in cohort 2. Marrow cavity volume was calculated using the BoneJ plugin for ImageJ (Doube et al., 2010). The humeral marrow cavity VOI was defined proximally by the growth plate and distally by the coronoid-olecranon notch. The femoral VOI was defined proximally by the most projecting point of the lesser trochanter and distally by the growth plate.

Flow cytometry

Cellular constituents of bone marrow were quantified in left humeri and femora of animals in cohort 3 immediately after death. Using a 25AWG needle, bone marrow was flushed from each endocortical cavity with supplemented medium, then mixed, and filtered through a 40 μm mesh (BD Biosciences, San Jose, CA, USA). Supplemented medium was prepared with Dulbecco's modified essential medium (Life Technologies, Inc., Gaithersburg, MD, USA), 2% fetal bovine serum (Invitrogen, Carlsbad, CA, USA), 1% penicillin/streptomycin and 10 mmol l^{-1} HEPES buffer (Life Technologies, Inc.). Bone marrow erythrocytes were lysed with 1× PharmLyse (BD Biosciences). Samples were diluted in Dulbecco's phosphate-buffered saline at a concentration of 1:100. A hemocytometer was used to quantify the number of cells in each sample. This enabled back calculation of the total number of marrow cells within each bone. A single-cell suspension containing 2×10^6 cells from each sample was then prepared to identify MSC populations through direct immunofluorescence staining with flow cytometry. Specifically, fluorescent antibody probes for SCA-1, c-Kit, CD90.2, CD105 and CD44 (BD Biosciences) were designated as the unique MSC surface identifiers (Pagnotti et al., 2012), and populations positive for all five markers were sub-gated accordingly using FlowJo cytometric analytical software (TreeStar, Inc., Ashland, OR, USA). The percentage of MSCs within each 2×10^6-cell sample was used to calculate the total number of MSCs within each bone.

Statistics

Statistical analyses were performed in SPSS (v20; IBM Corp., Armonk, NY, USA). Analysis of covariance was used to compare peak vertical ground reaction forces experienced by the forelimbs and hindlimbs of animals in cohort 1, with subject speed as the covariate (Wallace and Demes, 2008). Independent-samples *t*-tests were used to assess differences in body mass, muscle mass and bone traits between controls and exercised animals in cohort 2. Mann–Whitney *U*-tests were used to compare home-cage activity between controls and runners in cohort 2 (because of non-normal distribution of data), as well as MSC populations between controls and runners in cohort 3 (because of uneven sample sizes). A paired-sample *t*-test was used to compare MSC density in the humeri and femora of controls in cohort 3. Significance level for tests was $P<0.05$ and tests were two-tailed. Relative differences between group means were calculated as percentage difference±s.d. of the sampling distribution of the relative difference (Wallace et al., 2015).

Acknowledgements

We thank D. Riskin and S. Swartz for allowing us to use their force plate, A. Tuthill and T. Zimmerman for facilitating the experiments and S. Lublinsky for providing technical assistance with μCT analyses. We also thank A. Biewener and two anonymous reviewers for very helpful comments on earlier versions of this article.

Competing interests

The authors declare no competing or financial interests.

Author contributions

I.J.W., G.M.P., C.T.R. and B.D. conceived of the study. I.J.W., G.M.P., S.J., C.T.R. and B.D. designed the experiments. I.J.W., G.M.P. and B.D. performed the experiments. I.J.W., G.M.P., J.R.-S., M.N., L.E.C. and B.D. collected and analyzed data. I.J.W., G.M.P., C.T.R. and B.D. drafted the manuscript. All authors contributed to the interpretation of the findings and helped to edit the manuscript.

Funding

Supported primarily by a grant from the L.S.B. Leakey Foundation, as well as grants from the National Science Foundation (BCS 0935321), the National Aeronautics and Space Administration (NNX12AL25G and NNX08BA35G), the National Institutes of Health (AR 043498) and Sigma Xi. Deposited in PMC for release after 12 months.

References

Adler, B. J., Kaushansky, K. and Rubin, C. T. (2014). Obesity-driven disruption of haematopoiesis and the bone marrow niche. *Nat. Rev. Endocrinol.* **10**, 737-748.

Bab, I. A., Hajbi-Yonissi, C., Gabet, Y. and Müller, R. (2007). *Micro-tomographic Atlas of the Mouse Skeleton.* New York: Springer.

Biewener, A. A. (1983). Locomotory stresses in the limb bones of two small mammals: the ground squirrel and chipmunk. *J. Exp. Biol.* **103**, 131-154.

Biewener, A. A. (1990). Biomechanics of mammalian terrestrial locomotion. *Science* **250**, 1097-1103.

Biewener, A. A. and Bertram, J. E. A. (1994). Structural response of growing bone to exercise and disuse. *J. Appl. Physiol.* **72**, 946-955.

Bouxsein, M. L., Boyd, S. K., Christiansen, B. A., Guldberg, R. E., Jepsen, K. J. and Müller, R. (2010). Guidelines for assessment of bone microstructure in rodents using micro-computed tomography. *J. Bone Miner. Res.* **25**, 1468-1486.

Chan, M. E., Adler, B. J., Green, D. E. and Rubin, C. T. (2012). Bone structure and B-cell populations, crippled by obesity, are partially rescued by brief daily exposure to low-magnitude mechanical signals. *FASEB J.* **26**, 4855-4863.

Copes, L. (2012). *Comparative and Experimental Investigations of Cranial Robusticity in Mid-Pleistocene Hominins.* Ph.D. dissertation, Arizona State University.

Currey, J. D. (2002). *Bones: Structure and Mechanics.* Princeton: Princeton University Press.

Doube, M., Kłosowski, M. M., Arganda-Carreras, I., Cordelières, F. P., Dougherty, R. P., Jackson, J., Schmid, B., Hutchinson, J. R. and Shefelbine, S. J. (2010). BoneJ: free and extensible bone image analysis in ImageJ. *Bone* **47**, 1076-1079.

Dumont, E. R. (2010). Bone density and the lightweight skeletons of birds. *Proc. R. Soc. B Biol. Sci.* **277**, 2193-2198.

Haapasalo, H., Kontulainen, S., Sievänen, H., Kannus, P., Järvinen, M. and Vuori, I. (2000). Exercise-induced bone gain is due to enlargement in bone size without a change in volumetric bone density: a peripheral quantitative computed tomography study of the upper arms of male tennis players. *Bone* **27**, 351-357.

Jones, H. H., Priest, J. D., Hayes, W. C., Tichenor, C. C. and Nagel, D. A. (1977). Humeral hypertrophy in response to exercise. *J. Bone Jt. Surg. Am.* **59**, 204-208.

Journal of Experimental Biology

Judex, S., Garman, R. A., Squire, M. E., Donahue, L. R. and Rubin, C. T. (2004). Genetically based influences on the site-specific regulation of trabecular and cortical bone morphology. *J. Bone Miner. Res.* **19**, 600-606.
Lieberman, D. E. (1996). How and why humans grow thin skulls: experimental evidence for systemic cortical robusticity. *Am. J. Phys. Anthropol.* **101**, 217-236.
Lieberman, D. E. (1998). Neandertal and early modern human mobility patterns. In *Neandertals and Modern Humans in Western Asia* (ed. T. Akazawa, K. Aoki and O. Bar-Yosef), pp. 263-275. New York: Plenum Press.
Lieberman, D. E. and Crompton, A. W. (1998). Responses of bone to stress: constraints on symmorphosis. In *Principles of Animal Design: The Optimization and Symmorphosis Debate* (ed. E. R. Weibel, C. R. Taylor and L. Bolis), pp. 78-86. Cambridge: Cambridge University Press.
Lieberman, D. E., Pearson, O. M., Polk, J. D., Demes, B. and Crompton, A. W. (2003). Optimization of bone growth and remodeling in response to loading in tapered mammalian limbs. *J. Exp. Biol.* **206**, 3125-3138.
Lublinsky, S., Ozcivici, E. and Judex, S. (2007). An automated algorithm to detect the trabecular-cortical bone interface in micro-computed tomographic images. *Calcif. Tissue Int.* **81**, 285-293.
Luu, Y. K., Capilla, E., Rosen, C. J., Gilsanz, V., Pessin, J. E., Judex, S. and Rubin, C. T. (2009a). Mechanical stimulation of mesenchymal stem cell proliferation and differentiation promotes osteogenesis while preventing dietary-induced obesity. *J. Bone Miner. Res.* **24**, 50-61.
Luu, Y. K., Pessin, J. E., Judex, S., Rubin, J. and Rubin, C. T. (2009b). Mechanical signals as a non-invasive means to influence mesenchymal stem cell fate, promoting bone and suppressing the fat phenotype. *BoneKey* **6**, 132-149.
Main, R. P., Lynch, M. E. and van der Meulen, M. C. H. (2014). Load-induced changes in bone stiffness and cancellous and cortical bone mass following tibial compression diminish with age in female mice. *J. Exp. Biol.* **217**, 1775-1783.
McKay, H. A., MacLean, L., Petit, M., MacKelvie-O'Brien, K., Janssen, P., Beck, T. and Khan, K. M. (2005). "Bounce at the Bell": a novel program of short bouts of exercise improves proximal femur bone mass in early pubertal children. *Br. J. Sports Med.* **39**, 521-526.
Metzger, T. A., Kreipke, T. C., Vaughan, T. J., McNamara, L. M. and Niebur, G. L. (2015). The in situ mechanics of trabecular bone marrow: the potential for mechanobiological response. *J. Biomech. Eng.* **137**, 011006.
Nikander, R., Sievänen, H., Uusi-Rasi, K., Heinonen, A. and Kannus, P. (2006). Loading modalities and bone structures at nonweight-bearing upper extremity and weight-bearing lower extremity: a pQCT study of adult female athletes. *Bone* **39**, 886-894.
Ozcivici, E., Luu, Y. K., Rubin, C. T. and Judex, S. (2010). Low-level vibrations retain bone marrow's osteogenic potential and augment recovery of trabecular bone during reambulation. *PLoS ONE* **5**, e11178.
Pagnotti, G. M., Adler, B. J., Green, D. E., Chan, M. E., Frechette, D. M., Shroyer, K. R., Beamer, W. G., Rubin, J. and Rubin, C. T. (2012). Low magnitude mechanical signals mitigate osteopenia without compromising longevity in an aged murine model of spontaneous granulosa cell ovarian cancer. *Bone* **51**, 570-577.
Risbud, M. V., Shaprio, I. M., Guttapalli, A., Di Martino, A., Danielson, K. G., Beiner, J. M., Hillibrand, A., Albert, T. J., Anderson, D. G. and Vaccaro, A. R. (2006). Osteogenic potential of adult human stem cells of the lumbar vertebral body and the iliac crest. *Spine* **31**, 83-89.
Riskin, D. K., Bahlman, J. W., Hubel, T. Y., Ratcliffe, J. M., Kunz, T. H. and Swartz, S. M. (2009). Bats go head-under-heels: the biomechanics of landing on a ceiling. *J. Exp. Biol.* **212**, 945-953.
Rubin, C. T. (1984). Skeletal strain and the functional significance of bone architecture. *Calcif. Tissue Int.* **36**, S11-S18.
Rubin, C. T. and Lanyon, L. E. (1984). Dynamic strain similarity in vertebrates: an alternative to allometric limb bone scaling. *J. Theor. Biol.* **107**, 321-327.
Rubin, C. T., Bain, S. D. and McLeod, K. J. (1992). Suppression of the osteogenic response in the aging skeleton. *Calcif. Tissue Int.* **50**, 306-313.
Ruff, C. B. (2005). Mechanical determinants of bone form: insights from skeletal remains. *J. Musculoskelet. Neuronal Interact.* **5**, 202-212.
Ruff, C. (2008). Femoral/humeral strength in early African *Homo erectus*. *J. Hum. Evol.* **54**, 383-390.
Ruff, C. (2009). Relative limb strength and locomotion in *Homo habilis*. *Am. J. Phys. Anthropol.* **138**, 90-100.
Schutz, H., Jamniczky, H. A., Hallgrímsson, B. and Garland, T., Jr (2014). Shape-shift: semicircular canal morphology responds to selective breeding for increased locomotor activity. *Evolution* **68**, 3184-3198.
Tommerup, L. J., Raab, D. M., Crenshaw, T. D. and Smith, E. L. (1993). Does weight-bearing exercise affect non-weight-bearing bone? *J. Bone Miner. Res.* **8**, 1053-1058.
Volk, S. W., Wang, Y. and Hankenson, K. D. (2012). Effects of donor characteristics and ex vivo expansion on canine mesenchymal stem cell properties: implications for MSC-based therapies. *Cell Transplant.* **21**, 2189-2200.
Wallace, I. J. and Demes, B. (2008). Symmetrical gaits of *Cebus apella*: implications for the functional significance of diagonal sequence gait in primates. *J. Hum. Evol.* **54**, 783-794.
Wallace, I. J., Judex, S. and Demes, B. (2015). Effects of load-bearing exercise on skeletal structure and mechanics differ between outbred populations of mice. *Bone* **72**, 1-8.
Warden, S. J., Mantila Roosa, S. M., Kersh, M. E., Hurd, A. L., Fleisig, G. S., Pandy, M. G. and Fuchs, R. K. (2014). Physical activity when young provides lifelong benefits to cortical bone size and strength in men. *Proc. Natl. Acad. Sci. USA* **111**, 5337-5342.
Weeks, B. K., Young, C. M. and Beck, B. R. (2008). Eight months of regular in-school jumping improves indices of bone strength in adolescent boys and girls: the POWER PE study. *J. Bone Miner. Res.* **23**, 1002-1011.
Wilks, D. C., Winwood, K., Gilliver, S. F., Kwiet, A., Chatfield, M., Michaelis, I., Sun, L. W., Ferretti, J. L., Sargeant, A. J., Felsenberg, D. and Rittweger, J. (2009). Bone mass and geometry of the tibia and the radius of master sprinters, middle and long distance runners, race-walkers and sedentary control participants: a pQCT study. *Bone* **45**, 91-97.
Woo, S. L. Y., Kuei, S. C., Amiel, D., Gomez, M. A., Hayes, W. C., White, F. C. and Akeson, W. H. (1981). The effect of prolonged physical training on the properties of long bone: a study of Wolff's Law. *J. Bone Jt. Surg. Am.* **63**, 780-787.