Physical activity alters limb bone structure but not entheseal morphology

Ian J. Wallace [a], Julia M. Winchester [b], Anne Su [c], Doug M. Boyer [d], Nicolai Konow [e, f, *]

[a] Department of Human Evolutionary Biology, Harvard University, Cambridge, MA 02138, USA
[b] Interdepartmental Doctoral Program in Anthropological Sciences, Stony Brook University, Stony Brook, NY 11794, USA
[c] School of Health Sciences, Cleveland State University, Cleveland, OH 44115, USA
[d] Department of Evolutionary Anthropology, Duke University, Durham, NC 27708, USA
[e] Department of Ecology and Evolutionary Biology, Brown University, Providence, RI 02912, USA
[f] Department of Biological Sciences, University of Massachusetts Lowell, Lowell, MA 01854, USA

A B S T R A C T

Studies of ancient human skeletal remains frequently proceed from the assumption that individuals with robust limb bones and/or rugose, hypertrophic entheses can be inferred to have been highly physically active during life. Here, we experimentally test this assumption by measuring the effects of exercise on limb bone structure and entheseal morphology in turkeys. Growing females were either treated with a treadmill-running regimen for 10 weeks or served as controls. After the experiment, femoral cortical and trabecular bone structure were quantified with μCT in the mid-diaphysis and distal epiphysis, respectively, and entheseal morphology was quantified in the lateral epicondyle. The results indicate that elevated levels of physical activity affect limb bone structure but not entheseal morphology. Specifically, animals subjected to exercise displayed enhanced diaphyseal and trabecular bone architecture relative to controls, but no significant difference was detected between experimental groups in entheseal surface topography. These findings suggest that diaphyseal and trabecular structure are more reliable proxies than entheseal morphology for inferring ancient human physical activity levels from skeletal remains.

© 2017 Elsevier Ltd. All rights reserved.

1. Introduction

Physical activity levels are a critical aspect of human biological and cultural variation (Leonard, 2008; Pontzer et al., 2012). Physical activity is essential for individual health and survival (Lieberman, 2013), it can be a driving force behind phenotypic evolution (Wallace et al., 2010; Raichlen and Polk, 2013), and it is a nexus that links features of economy, technology, and social relations (Kelly, 2013). For these reasons, paleoanthropological investigations of the lifeways of our ancient ancestors often aim to glean information about their levels of physical activity, typically by analyzing their skeletal remains (e.g., Villotte et al., 2010; Lieverse et al., 2013; Shaw and Stock, 2013; Chirchir et al., 2015; Ruff et al., 2015). Two skeletal features are assumed to be especially informative: (1) the quantity and distribution of bone within the limb elements, and (2) the morphology of muscle and tendon attachments sites, or entheses (Jurmain et al., 2012; Larsen, 2015). Individuals with

robust limb bones and/or rugose, hypertrophic entheses are inferred to have been highly active during life, whereas individuals with gracile limb bones and/or smooth, modest-sized entheses are assumed to have been more sedentary.

In this study, we experimentally test the model within which paleoanthropologists deduce physical activity levels from human skeletal remains by measuring the effects of treadmill-running exercise on limb bone structure and entheseal morphology in turkeys. If the paleoanthropological model is valid, then animals treated with exercise should exhibit enhanced bone structure and/or altered entheseal form. Turkeys and other bipedal birds are good locomotor models for human locomotion because the species exhibit similar patterns of limb movement (Gatesy and Biewener, 1991), basic body center-of-mass dynamics and pendular exchange (Biewener and Daley, 2007), and energetic costs associated with locomotion (Rubenson et al., 2006). Bipedal birds are also a well-accepted system for investigating links between activity-induced bone loading, limb muscle stresses, and bone functional adaptation (e.g., Roberts et al., 1997; Judex and Zernicke, 2000; Gabaldon et al., 2004; Pontzer et al., 2006).

* Corresponding author.
E-mail address: nicolai_konow@uml.edu (N. Konow).

2. Methods

All experimental procedures were reviewed and approved by the Institutional Animal Care and Use Committee of Brown University. The experiment was designed and carried out for an unrelated study. Thus, no animals were harmed for the specific purpose of gathering the data reported here.

Ten 1-year-old female Eastern wild turkeys (*Meleagris gallopavo*) were obtained from a licensed breeder and transported to the animal care facility at Brown University. Animals were group-housed in a well-ventilated, climate-controlled room on a 12:12 light:dark cycle. The room was fitted with a pen (6 × 10 × 3 m; width × length × height) with rubber flooring. Hay was provided for bedding, and food and water were available ad libitum. When left undisturbed in the pen, all animals were generally quiescent and active only for brief periods for socialization, eating, and drinking, similar to the behavior of turkeys observed in prior studies (Adams et al., 1997; Fritton et al., 2000).

Animal subjects were designated as either runners or non-runner controls (n = 5/group). Runners were exercised on a declined motor-driven treadmill at a speed of 2.5 m s^{-1} for 30 min a day, four days a week, for 10 weeks, on average taking 141,912 (standard deviation = 7,280) strides during the exercise treatment. Each stride generated peak vertical substrate reaction forces between approximately 150% and 250% of the animal's body weight (Roberts and Scales, 2002). This exercise routine constitutes a substantial increase in high-magnitude limb loading events relative to the normally sedentary activity patterns of laboratory turkeys (Fritton et al., 2000), and it is consistent with prior exercise studies that have revealed statistically significant musculoskeletal tissue changes in galliform birds (Buchanan and Marsh, 2001, 2002), as well as in mammals (Wallace et al., 2015). All subjects were euthanized by 18−19 months of age and their femora were extracted. The use of animals between one and two years old is ideal for examining the skeletal effects of physical activity, as the adaptability of turkey bones to mechanical loading has been shown to be high during this ontogenetic period (Rubin et al., 1992).

Femora were scanned at a 45-μm^3 voxel size using a Nikon XT H 225 ST μCT system. Cortical and trabecular bone structural properties were quantified from μCT images using the BoneJ plugin (Doube et al., 2010) for ImageJ software. Images were thresholded to extract the bone phase using the "optimise threshold" option in BoneJ. Cortical bone parameters were measured from mid-diaphyseal transverse cross-sectional μCT images and included cortical bone area, maximum second moments of area, and minimum second moments of area. In standard beam analysis, cortical area approximates a bone cross section's internal resistance to axial compression and tension, and maximum and minimum second moments of area describe resistance to bending around principal axes. Trabecular bone structural properties were quantified in the lateral condyles of distal femoral epiphyses. This location was chosen for analysis due to its close proximity to the entheseal site examined (see below). Volumes of interest (VOI) were defined in μCT image stacks as the largest centered sphere to fit completely within each region of trabecular bone. Bone volume fraction is the percentage of bone voxels present in the VOI relative to the total number of voxels in the VOI. Trabecular thickness is the average strut thickness. Trabecular connectivity density characterizes the redundancy of trabecular connections, normalized to the volume of the VOI.

Entheseal morphology was assessed at the lateral epicondyles of distal femora (Fig. 1). The lateral epicondyle is the attachment site for the tendon of the lateral gastrocnemius muscle, a key extensor of the ankle. This site was selected for analysis because during decline running in turkeys, the lateral gastrocnemius is activated at high levels (generating forces >70% of the animal's body weight) to function as a brake, a mechanical function that is facilitated by muscle fibers undergoing lengthening contractions during stance in order to absorb mechanical energy (Gabaldon et al., 2004; Roberts and Azizi, 2010). Thus, if muscle activity associated with decline running influences turkey entheseal morphology, it is reasonable to expect that the lateral epicondyles would be especially affected.

The shape of entheseal surfaces was quantified using morphological topographic analysis (Evans, 2013). Topographic metrics analyzed included Dirichlet normal energy (which describes surface curvature; Bunn et al., 2011), relief index (which describes surface relief; M'Kirera and Ungar, 2003; Boyer, 2008), and orientation patch count rotated (which describes surface complexity; Evans et al., 2007) (Fig. 1). Quantifying these aspects of shape together provides a holistic assessment of surface morphology, as has been well illustrated in studies of tooth shape (e.g., Bunn et al., 2011; Winchester et al., 2014).

For topographic analyses, Amira software was used to segment 3D external surfaces of distal femora from μCT images, and then to crop the surfaces to isolate the lateral epicondyles (Fig. 1). In order to ensure comparability of lateral epicondyle surfaces between animals, a standard cropping protocol was developed. Lateral epicondyle morphology can be generally described as laterally convex with limited degrees of gentle curvature leading to sharper flexure around the anterior, posterior, and inferior borders. Anterior, posterior, and inferior margins of epicondyle surfaces were placed at the lines of transition between blunter and sharper degrees of curvature. Junctions between posterior and superior margins were placed along posterior margins at heights equal to the furthest observed extents of the gastrocnemius entheses, and superior margins were placed as straight lines from posterior to anterior margins. All polygons within the defined margins were cropped to produce isolated epicondyle surfaces. Surfaces were then simplified to 20,000 polygon faces and smoothed with a single iteration prismatic shapes (aggressive) method using Geomagic software.

After epicondyle surfaces were cropped, simplified, and smoothed, topographic variables were quantified using the MorphoTester application (Winchester, 2016). Dirichlet normal energy was measured as a sum of per-polygon curvature or bending values across the epicondyle surface. Bending per polygon is quantified as the change in normal vector map relative to change in polygon vertices, multiplied by polygon area to account for polygon scaling (for more details see Bunn et al., 2011; Winchester, 2016). MorphoTester parameters for Dirichlet normal energy calculation included condition number checking and outlier removal of energy times polygon area values above the 99.9th percentile. Implicit fair smoothing was not used. Relief index was calculated as the natural log of the square root of surface area divided by the square root of projection area, resulting in a dimensionless quantity. Orientation patch count rotated was measured as a count of patches across the epicondyle surface with differing facing or aspect on the epicondyle surface plane, averaged across eight rotations to account for rotation of that plane around a central perpendicular axis (approximately mediolateral here). This topographic variable was measured with a minimum patch count of six, meaning that surface patches comprising fewer than six polygons are not counted toward the surface patch total.

Shapiro-Wilk tests were used to determine if data followed a normal distribution, and Levene's tests were used to assess the equality of group variances. Depending on the results of these tests, differences between the two experimental groups were analyzed with independent samples t-tests, Mann−Whitney U-tests, or Welch's t-tests. Significance level for tests was $p < 0.05$ and tests were two-tailed.