



Ecogeography, genetics, and the evolution of human body form



Charles C. Roseman^{a,*}, Benjamin M. Auerbach^b

^a Department of Anthropology, University of Illinois, Urbana, IL 61802, USA

^b Department of Anthropology, The University of Tennessee, Knoxville, TN 37996, USA

ARTICLE INFO

Article history:

Received 16 January 2014

Accepted 24 July 2014

Available online 15 November 2014

Keywords:

Natural selection

Population structure

Ecogeographic variation

Bergmann's and Allen's rules

Comparative method

ABSTRACT

Genetic resemblances among groups are non-randomly distributed in humans. This population structure may influence the correlations between traits and environmental drivers of natural selection thus complicating the interpretation of the fossil record when modern human variation is used as a referential model. In this paper, we examine the effects of population structure and natural selection on postcranial traits that reflect body size and shape with application to the more general issue of how climate – using latitude as a proxy – has influenced hominin morphological variation. We compare models that include terms reflecting population structure, ascertained from globally distributed microsatellite data, and latitude on postcranial phenotypes derived from skeletal dimensions taken from a large global sample of modern humans. We find that models with a population structure term fit better than a model of natural selection along a latitudinal cline in all cases. A model including both latitude and population structure terms is a good fit to distal limb element lengths and bi-iliac breadth, indicating that multiple evolutionary forces shaped these morphologies. In contrast, a model that included only a population structure term best explained femoral head diameter and the crural index. The results demonstrate that population structure is an important part of human postcranial variation, and that clinally distributed natural selection is not sufficient to explain among-group differentiation. The distribution of human body form is strongly influenced by the contingencies of modern human origins, which calls for new ways to approach problems in the evolution of human variation, past and present.

© 2014 Elsevier Ltd. All rights reserved.

Introduction

Variation among recent and living groups of humans plays a key role in conditioning our expectations for and formulating research problems about the human fossil record (Wolpoff, 1993; Lahr and Foley, 1998; Lieberman et al., 2000; Roseman et al., 2011). Taxonomic efforts to distinguish and recognize species and systematic studies seeking to establish patterns of relatedness among them often use human variation as a benchmark for deciding how much variation is allowable in a good hypodigm, as well as which characters might be useful for reconstructing phylogeny (Kramer et al., 1995; Wood and Lieberman, 2001; Harvati and Weaver, 2006). When it comes to evolutionary processes, what constitutes evidence of the action of natural selection, gene flow, or random genetic drift in the fossil record is often arbitrated using recent human variation as a referential model (Ruff, 1991, 1994; Ackermann and Cheverud, 2004; Smith et al., 2007; Weaver et al., 2007, 2008; von Cramon-Taubadel and Lycett, 2008). Figuring prominently in

this literature is the study of ecogeographic variation: the associations between phenotypes and geography in the context of regional or global climate (Mayr, 1956; James, 1970; Graves, 1991; Holliday, 1995; Ashton et al., 2000). The distribution of body form – body size, body breadth, limb length, and the proportional relationships among them – in recent humans is foundational in the study of human adaptation and variation, and is a key part of models used to propose hypotheses about the evolution of body form in fossil hominins (Ruff, 1991, 1994).

Heretofore, most studies of the ways in which natural selection and environment affect the distributions of phenotypic states across groups assume that each group is no more genetically similar to any other group and evolve by natural selection (e.g., Graves, 1991; Ruff, 1994; Auerbach, 2007; Lovegrove and Mowoe, 2013). Genetic dependencies among the mean phenotypic values of human groups arising from a strong and complex population structure can exert substantial influences on patterns and magnitudes of morphological variation (Relethford et al., 1997; Roseman, 2004; Whitlock, 2008; Stone et al., 2011). Genetically similar groups will resemble one another in ways that are shaped by the history of population fissioning, admixture, gene flow, accumulation of neutral mutations, and fluctuations in population size.

* Corresponding author.

E-mail address: croseman@illinois.edu (C.C. Roseman).

Further complicating matters is the fact that climate and other factors that are thought to motivate natural selection on human body form are often structured geographically, perhaps in ways that are similar to patterns of genetic differences among human groups. For instance, most genetic variation in humans with recent ancestry outside of Africa is traceable back to Africa some 45–70 ka (thousands of years ago) (Henn et al., 2012). An origin in the tropics coupled with a founder effect during an expansion from Africa in a partly northward direction could profoundly affect the sampling distribution of measures of association between phenotypic variation and latitudinally distributed variables thought to cause adaptive evolution in humans in the complete absence of natural selection. These issues are similar to concerns that motivate work in phylogenetic comparative methods (e.g., Martins, 2000; O'Neill and Dobson, 2008) including recent applications to problems in species with population structure (Stone et al., 2011). In the absence of an evolutionary model of how phenotypes vary among groups of humans, we do not know how to interpret estimates of correlations or model terms (e.g., regression coefficients) that we use to summarize the relationships between climate or other natural phenomena and group means.

In what follows, we examine different scenarios for the evolution of bodily form by using computer simulations and generalized linear mixed models. The simulations provide a theoretical demonstration of how relatedness among populations can influence the distribution of phenotypes that might lead us to accept adaptive ecogeographic hypotheses more often than would be otherwise warranted. We use the mixed model approach to fit and compare models of the distribution of phenotypic means of populations that include terms reflecting both natural selection and the selectively neutral resemblances among groups arising from population structure. Comparing the fits of different models will allow us to make stronger statements about the evolutionary processes that generated human variation.

Model fitting yields estimates of both the proportion of the among-group morphological variance attributable to population history and structure (similar to a heritability estimate) and estimates of the linear relationships of the traits with latitude, used as a proxy for climate following Ruff (1994) and Holliday (1999). Comparison of the goodness of fit for the different models and the estimates of the model terms allows us to directly differentiate several models of the evolution of proposed ecogeographic patterns. Parameter estimates and the relative goodness of fit of the models guide us in making evolutionary interpretations of the evolution of modern human body form. This, in turn, provides an empirical check on the degree of confidence we have in using the results of comparative studies of modern human variation as a guide to variation in the fossil record. Morphological patterns observed among modern humans that are shown to result from population history, rather than natural selection, would have significant implications for the reconstruction and interpretation of fossil hominin body forms.

This study focuses on characteristics commonly used in research into ecogeographic patterns of human variation, including the lengths of limb elements (tibia, femur, radius, and humerus), femoral head diameter (taken to be an indicator of body mass), and a characteristic with a distribution of among-group differences that may reflect the action of natural selection motivated by climate (bi-iliac breadth) (Roberts, 1978; Ruff, 1991; Auerbach, 2012). We also examine ratios of the upper limb (brachial) and lower limb (crural) elements, both important traits that appear to be correlated with latitude to one degree or another (Trinkaus, 1981; Ruff, 1994; Holliday, 1997). All of these traits are featured in models of human thermoregulation and are thus key to problems about the relationship between climate and human evolution (Ruff, 1994; Holliday, 1999; Auerbach, 2007).

Ecogeographic variation in recent humans

Ecogeographic patterns in body shape and size have been described in a variety of endothermic organisms, including passerine birds, hares, macaques, and humans (e.g., Mayr, 1956; James, 1970; Stevenson, 1986; Graves, 1991; Holliday, 1995; Paterson, 1996; Blackburn et al., 1999; Ashton et al., 2000). The explanation of the correlation of population means for certain traits with climate (usually temperature) stems from a model based on two thermoregulatory principles developed from publications by Bergmann (1847) and by Allen (1877). Both of these principles were redefined by multiple researchers over the last century, and dubbed 'rules' by Mayr (1956). Numerous authors have argued over whether the variation is related to thermoregulation or other physiological features of organisms (Rensch, 1938; Hamilton, 1961; Ashton et al., 2000; Yom-Tov et al., 2002), but collectively agree that the pattern is present in various taxa. The most influential of the interpretations of these rules for biological anthropologists is the 'thermoregulatory imperative' and cylindrical model advocated by Ruff (1991, 1994), which followed the anthropometric work of Hiernaux (Hiernaux and Froment, 1976; Hiernaux, 1985) and (Roberts, 1978), among others (Crognier, 1981; Trinkaus, 1981). In sum, this model argues for the optimization of surface area to volume ratios for efficient regulation of body temperature.

The peopling of the globe resulted in different groups coming into contact with a variety of environments. This prehistoric set of events and processes set up what has been characterized as a set of natural experiments (Auerbach, 2007; King, 2010), suitable for studying questions about the relationships between various phenotypes (e.g., body shape, body size, and proportions) and a series of environmental variables (e.g., temperature, humidity, and seasonality). Research on humans has demonstrated empirical patterns of the morphological differences across populations that appear to support the thermoregulatory argument. For example, studies of Eurasian and African groups show associations of several phenotypes with latitude (a proxy for climate); humans from lower latitudes have relatively longer limbs (compared with torso height), longer distal limb lengths, lower body masses, and absolutely narrower bi-iliac breadths (Ruff, 1994, 2002; Holliday, 1997). Some of these associations are evident early in ontogeny (Ruff et al., 2002; Temple et al., 2011; Cowgill et al., 2012; Garofalo, 2012). Further research demonstrated potential differences in sensitivity to climatic and biomechanical factors in the limbs (Stock, 2006; Higgins and Ruff, 2011), some effects on body size or proportions arising from diachronic changes in subsistence (e.g., Malina et al., 2004; Stock and Pinhasi, 2011; Auerbach, 2011a), and that some observed variation in limb length results from general allometry (Holliday and Ruff, 2001; Sylvester et al., 2008; Auerbach and Sylvester, 2011). Even in light of the apparent effects of these additional factors, body size, shape and proportions present empirical ecogeographic patterns in Eurasia and Africa, regardless of ultimate evolutionary and environmental mechanisms.

Scientists have provided ad hoc explanations for deviations from ecogeographic expectations on the part of different phenotypes. For example, a maintenance of presumably ancestral body shape and proportions among Upper Paleolithic humans in Europe (Holliday, 1997), Yayoi in Japan (Temple et al., 2008), and Arikara in the North American Great Plains (Auerbach, 2010) all were presented by these authors (including one of the authors of this paper) as evidence for long temporal gaps between the peopling of novel environments and changes in morphology caused by selection. This argument was also made for studies of variation in the Americas, where some dimensions generally match ecogeographic expectations, such as intralimb indices, while others like body size and bi-iliac breadth have reduced among-population variation and poor

correspondence to climate (Auerbach, 2007, 2012); in these studies, Auerbach resorted to equivocating between natural selection, random genetic drift, and phenotypic plasticity as influencing these morphologies dissimilarly. None of the studies to date, though, have included explicit models of evolutionary mechanisms. In short, we are becoming more refined in our understanding of *pattern*, but we are still unable to address the evolutionary *processes* and environmental effects that cause them.

Although the exact evolutionary processes that contribute to the ecogeographic patterns in human morphological variation have not been identified with certainty, variation in the body proportions of fossil humans has been explained in terms of these patterns through an adaptationist perspective. Multiple studies have argued that Neandertal body morphology was caused by natural selection acting to optimize thermoregulation in the glacial environments of Pleistocene Europe (Trinkaus, 1981; Ruff, 1991, 2010; Holliday, 1997; Weaver, 2003; Ruff et al., 2005). Others have cited alternative factors, such as terrain (Higgins and Ruff, 2011), as influencing phenotypic states. Likewise, central African hominins (e.g., *Homo ergaster*) exhibit postcranial morphologies argued to correspond with efficient thermoregulation in tropical environments (e.g., Ruff, 1991; Potts, 1998; Weaver, 2003). The morphological differences between these hominins are readily apparent, but, as in the studies of modern human ecogeographic variation, accounts of the causes of the evolution of the characteristic form of Neandertals and *H. ergaster* are quite provisional.

We should emphasize here that this is not a study to test the empirical patterns associated with Bergmann's or Allen's rules among modern humans. Multiple studies, as cited above, have expressly assessed these, and collectively show at least some correspondence of ecogeographic expectations with the empirical patterns of human body form. We demonstrate the same with our morphological data (see morphological trait correlations with latitude in Table 1). Multiple factors could drive processes that could lead to a correspondence between the distribution of a phenotype and latitude (e.g., climate, nutrition, and disease). We are interested in using multiple sources of information to compare the plausibility of different combinations of variables as explanations for the patterns.

Models and simulations of the evolution of ecogeographic patterns

The core of our argument is based on multi-model comparison using combined morphological and molecular data and computer simulation. To demonstrate why we might want to consider models that include a population structure term from a theoretical standpoint, we use simulations of several different scenarios representing extremes of sampling strategies used in the literature. One scenario reflects a case where all groups used in the present analysis (121 groups) are evolutionarily independent, which is an implicit assumption made for evolution in the human ecogeographic

literature. Two alternative scenarios reflect the evolution of heritable phenotypes under random genetic drift when populations share common ancestry and exchange migrants. The different sampling scenarios represent both the data set under investigation in this study (121 groups; see the [Methods](#)) and a data set used in a previous study of human postcranial variation that includes a smaller number of groups (14) (Holliday, 1995).

Following the validation of the theoretical justification for being concerned about population structure, we use a Bayesian generalized linear mixed model approach to model the evolution of morphological traits in humans. We fit and compare three models, one including a term reflecting clinal selection along latitude, another reflecting random genetic drift and gene flow, and a third that combines the terms from both models. This allows us to evaluate whether natural selection or neutral evolution is responsible for among-group variance, or if a blend of the two is the best explanation for the distribution of among-group variation in humans.

The first model (latitude model) is a rendition of the clinal model, which posits that the means of populations are structured along a latitudinal gradient of optimal phenotypic states (Fig. 1A). Directional selection has long since moved populations toward the optimal state and stabilizing selection has kept them close to their latitudinally optimal state. Drift and gene flow have negligible effects on the distribution of the phenotypes in this account and all groups are regarded as having evolved independently.

Our second model (structure model) focuses on population structure in the form of an among-group covariance matrix estimated using a large panel of microsatellite genotype data. The matrix is used as a random effect. It models mean phenotypic variation among groups as a function of shared common ancestry and gene flow. This is similar to how one might use a relationship matrix from a pedigree to estimate heritability in a group of related organisms. We do not attempt to separate gene flow from shared common ancestry (a goal of Stone et al., 2011), which is not possible in this particular case. Instead, the resulting variance component summarizing the effect of the population structure is interpreted as an among-group neutral genetic variance term. The ratio of among-group genetic variance to total genetic variance has a similar interpretation to heritability within populations with all of the attendant cautions and limitations. If among-group differentiation in the phenotypes has a neutral additive genetic basis, we expect the distribution of among-group variation to be well predicted by the population structure. An illustration of a simple case involving evolution by random genetic drift along a tree is given in Fig. 1B.

The third model (latitude and structure model) that we consider is a combination of both of the first two models. It includes both the fixed effect of absolute latitude and the random effect of population structure. This model is based on one suggested for use on problems related to Bergmann's and Allen's rules by Hadfield and Nakagawa (2010). A fit of the combined model that is better than

Table 1
Correlations of morphological traits with each other and with latitude.

	BIB	FHD	HML	RML	FML	TML	Brac.	Crur.	LAT
BIB		0.65	0.37	0.25	0.33	0.22	-0.17	-0.17	0.49
FHD	0.59–0.71		0.54	0.34	0.54	0.30	-0.33	-0.47	0.48
HML	0.23–0.51	0.45–0.63		0.82	0.88	0.81	-0.13	0.02	0.12
RML	0.09–0.41	0.20–0.48	0.80–0.84		0.85	0.90	0.40	0.32	-0.16
FML	0.19–0.47	0.45–0.63	0.87–0.89	0.84–0.86		0.91	0.04	-0.00	0.07
TML	0.05–0.39	0.15–0.45	0.79–0.83	0.89–0.91	0.90–0.92		0.27	0.41	-0.20
Brac.	-0.34 to 0.00	-0.47 to -0.19	-0.31 to 0.05	0.27–0.53	-0.14 to 0.22	0.11–0.43		0.54	-0.50
Crur.	-0.34 to 0.00	-0.58 to -0.36	-0.16 to 0.20	0.17–0.47	-0.18 to 0.18	0.28–0.54	0.45–0.63		-0.63
LAT	0.38–0.60	0.27–0.59	-0.06 to 0.30	-0.33 to 0.01	-0.11 to 0.25	-0.37 to -0.03	-0.60 to -0.40	-0.70 to -0.56	

Upper right triangle: Pearson's product moments. Lower left triangle: 95% confidence intervals of Pearson's product moments. Abbreviations: Bi-iliac breadth – BIB; femoral head diameter – FHD; humerus length – HML; radius length – RML; femur length – FML; tibia length – TML; brachial index – Brac.; crural index – Crur.; and latitude – LAT.

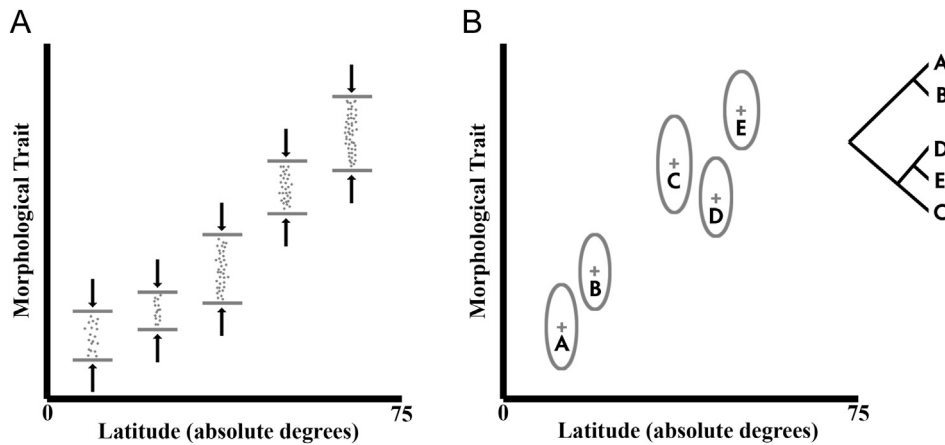


Figure 1. Plots of phenotypic states versus latitude demonstrating the theoretical underpinnings of two evolutionary models used in this study. Figure 1A (left side graph) represents a clinal model for trait variation among human groups that is motivated by natural selection as a linear function of absolute latitude. Stabilizing selection (represented by the arrows) maintains each group at a latitudinally appropriate optimum. In contrast, Figure 1B (right side graph) represents trait variation among groups geographically apportioned due solely to random genetic drift along a population history represented by a phylogenetic tree. The similarities in the geographic distribution of the groups and their relatedness led to among-group differences being latitudinally structured from random genetic drift alone.

the other two models gives us reason to think that both natural selection and neutral evolutionary forces might be influencing the among-group variation in a phenotype.

Materials and methods

Data

To meet the two general goals of this study, a large global sample of human phenotypic data consisting of dimensions associated with ecogeographic patterns of variation is necessary. We used the Goldman Data Set (1538 skeletons) (Auerbach and Ruff, 2004) and Auerbach's Americas Data Set (2749 skeletons) (Auerbach, 2007, 2010, 2011b), both of which consist mostly of postcranial measurements, and both of which were measured by one of the authors (B.M.A.). Portions of these two data sets overlap, but together they represent human populations from Eurasia, Africa, Oceania and the Americas, spanning a wide climatic range from the Equator to the Arctic. While high elevation has been shown to be associated with body form (Stinson, 1990; Ruff, 1994), only one group in the data set

used here lived at high elevation and is thus not likely to be a confounding factor. Two additional data sets were added to fill in data gaps in Africa by the gracious generosity of Trent Holliday and Chris Ruff. Holliday's data consist of modern human skeletons sampled from Europe and Africa (Holliday, 1995). A previous assessment of interobserver measurement errors between B.M.A. and Holliday showed high correspondence between the measurements taken by these two individuals (measurements in this study had less than one percent error; Auerbach, 2007). Ruff's data consist of measurements taken from two East African skeletal samples, one from Uganda and the other from Kenya.

Individuals that were of indeterminate sex and samples that were poorly provenienced were discarded, and all four data sets were reconciled. Duplicated individuals across the Goldman, Americas, and Holliday data were identified and removed. The final, aggregated skeletal data set used in this study consists of 2187 individuals (all males), aggregated into 121 groups. We used only males to maximize the global representation of the sample, and to avoid complicating effects of sexual dimorphism. These groups are shown graphically as open squares on the map in Fig. 2. A list of the

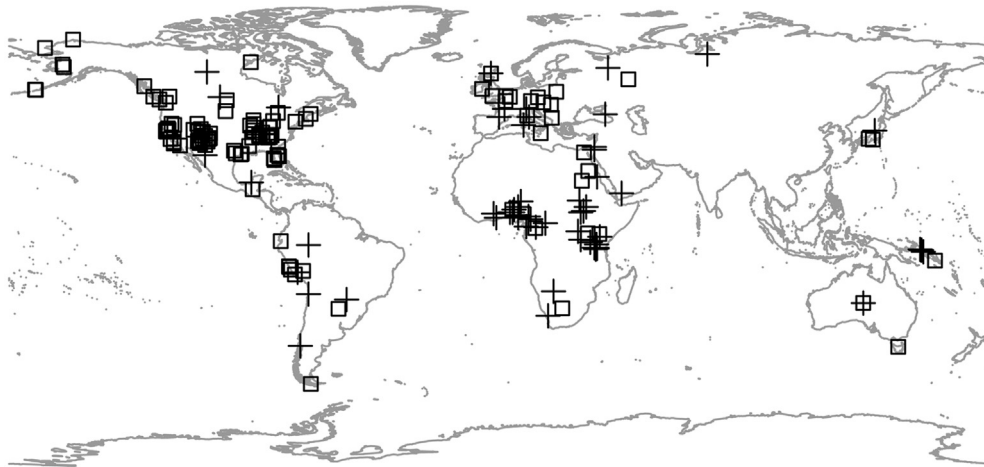


Figure 2. Map of geographic locations of sampled groups used in this study. Open squares indicate the morphological sample. In some instances, squares represent the geographic mean of more than one proximate site. Crosses indicate the sites used from the microsatellite DNA data set (sampled from Pemberton et al., 2013). SOM Table S1 contains a list of names and geographic coordinates of the samples.

groups and their sample sizes may be found in [Table S1 of the Supplementary Online Material \(SOM\)](#). We examine six dimensions from each skeleton that are commonly used in explorations of ecogeographic variation: the maximum lengths of the humerus, radius, femur, and tibia; anteroposterior diameter of the femoral head; and bi-iliac breadth. These measurements were defined by [Martin \(1928\)](#), and described in detail elsewhere ([Auerbach and Ruff, 2004; Auerbach, 2007](#)). To minimize the effects of directional bilateral asymmetry, limb bone measurements were taken from both sides and averaged ([Auerbach and Ruff, 2006](#)).

In addition to the above six raw dimensions, two indices typically used in ecogeographic studies were estimated from the individual data. Brachial and crural indices are intralimb proportions of the upper limb and lower limb, respectively. The brachial index is equal to the maximum length of the radius divided by the maximum length of the humerus. Crural indices have been defined as maximum tibial length divided by bicondylar femoral length ([Davenport, 1933](#)), but as maximum and bicondylar femoral lengths are nearly perfectly correlated with low variance ([Auerbach, 2011b](#)), and as maximum length is what we examine in the model, we chose to use the maximum femoral length rather than extrapolate bicondylar length from that dimension.

Morphological characteristics are expressed in raw millimeters or dimensionless ratios for the crural and brachial indexes. Means and standard deviations of all measurements for each group are presented in [SOM Table S1](#). All of the individual Goldman and Americas data used in this study are available upon request. The Goldman Data Set is also available for download from B.M.A.'s web site (<http://web.utk.edu/~auerbach/GOLD.htm>). Raw data shared by Holliday and by Ruff may be accessed by direct request to those individuals.

[Pemberton et al. \(2013\)](#) aggregated data from eight major data sets of human microsatellite variation at 645 loci from 5795 individuals comprising 267 populations across the world. These data may be accessed at the Rosenberg lab web site (<http://www.stanford.edu/group/rosenberglab/data/>). Of the total number of populations sampled, a subset of 59 groups, comprised of 2610 individuals, was matched to the morphological data sets (see below). The geographic locations for microsatellite group data used are shown graphically in [Fig. 2](#) (plus signs).

[Eriksson and Manica \(2011\)](#) provide persuasive evidence that many of the loci in the data set that we use may be subject to ascertainment bias leading to higher than average levels of allele similarity across Eurasia. While there are data sets that include many more loci and may be less susceptible to ascertainment, they do not have the geographic breadth of coverage of this microsatellite dataset. Our preference here is for greater geographic coverage and more sampled populations than for high precision in the estimates of the relationships, as the main source of uncertainty for our analytical approach is going to be due to the number of groups used to fit the model and the mismatches among groups (see below). Furthermore, the inherently large evolutionary variance of individual phenotypic characteristics ([Rogers and Harpending, 1983; Lynch, 1990](#)) will likely dwarf any bias or uncertainty in introduced by ascertainment bias or population mismatch.

Matching the two data sets

To set up the evolutionary models (see below) used in this analysis, we needed to match microsatellite molecular data with the phenotypic skeletal data. None of the microsatellite data come from the individuals measured for morphological traits in this study. We made matches between the genetically sampled living groups and groups represented in the skeletal data. These were

prioritized first to exact group matches when possible (for example, the San, Biaka Pygmies, French, Orkney Islanders, Japanese, and indigenous Australians). When this was not possible, geographic approximation and known linguistic relationships were used to link genetic and skeletal groups. In a few cases, there was some uncertainty as to which group in a microsatellite sample should be used for a given skeletal sample even though the geographic locations of both were known. When these instances occurred, we took the mean of the covariance among groups for all of the groups in a region. For example, in the case of Europeans without exact matches, an average of the genetic among-group covariance matrix was calculated from all of the European groups. No samples known to represent recently admixed groups were included. [Table S1](#) lists the microsatellite sample groups with their respective morphological groups; the numbers listed for the microsatellite groups are associated with populations in [Pemberton et al. \(2013\)](#).

One criticism of this method is that it is prone to error from poor matches between groups sampled for the different kinds of data. We acknowledge that uncertainty may arise because of a number of factors, especially given the temporal differences between the skeletal samples (most of which date to the last 2000 years) and living groups. In no case, however, are we matching skeletal samples with geographically distant genetic groups (e.g., we are not using microsatellite data from southern Africa for a group from North Africa). Mismatches will be local, and therefore trivial when compared with the breadth of the global data used in this study. This follows from the fact that patterns of genetic similarities among groups point to a hierarchical structure where populations are nested within clusters ([Hunley et al., 2009; Long et al., 2009](#)). So long as we assign a group to the correct low-level local or regional cluster, little harm will be done by way of mismatches.

Using regional averages of among-group variation to match multiple groups sampled for phenotypic variation, which we do in the case of North and South America and some groups in Europe, introduces a bias into the analysis. By assigning equal levels of genetic resemblance among groups in regions that are phenotypically oversampled and genomically under-sampled, we are treating those groups as though they are all evolutionarily independent with respect to one another and equally genetically similar (or different) to the remainder of the world. If there is substantial population structure within these regions, this will tend to understate the effect of population structure in the entire analysis.

Among-population relationship matrix

Using the 618 loci that were represented in all groups in our combined analysis, we estimated an among-group relationship matrix (**C**) to serve as the basis for the simulations of phenotypic evolution and as the random effect representing population structure in our mixed model. In this case, the aim is to predict the dependencies among group means, much in the same way that we might with coefficients of relatedness derived from the structure of a pedigree when we are interested in estimating heritability using an animal model ([Henderson, 1953; Thompson, 2008](#)). Our approach here is similar to the Relethford-Blangero method for estimating among-group genetic dependencies for phenotypes ([Relethford and Blangero, 1990; Relethford et al., 1997](#)). We take advantage of the fact that the sum of the lengths of the alleles at a microsatellite locus is a completely heritable quantitative trait ([Weaver et al., 2008](#)). As such, an estimate of the among-population covariance in the repeat lengths averaged across many loci gives us an estimate of the covariance in the among-group additive population values. After centering all repeat lengths on zero by subtracting the mean and standardizing all loci to unit variance by dividing through by the standard deviation of each locus, we

estimated k locus by g group matrices of group deviations from the centroid of all standardized microsatellite repeat scores (Δ). We then estimated \mathbf{C} as:

$$\mathbf{C} = \Delta \mathbf{V}^{-1} \Delta' / k$$

where \mathbf{V} is a $k \times k$ matrix in which all off-diagonal elements are set to zero (assuming no linkage among loci). The diagonal elements of \mathbf{V} are set to the average within population repeat variance. The matrix \mathbf{C} represents the expected among-group covariance of genetic population values of a quantitative trait relative to what one would expect in a single randomly mating group. Before using \mathbf{C} as a random effect, we standardized the matrix by the mean deviation of populations from the centroid (the mean of the diagonal entries).

This formulation does not allow us to distinguish between gene flow and random genetic drift acting along shared common ancestry in a population history. It does give us an expectation about how the means of groups should be genetically dependent on one another in the absence of natural selection under the assumption that the microsatellite loci are approximately neutral. A small proportion of the variance attributable to population structure could mean that natural selection is acting on the population and/or environmental (non-genetic and non-evolutionary) effects are affecting the growth of individuals in different populations.

Simulating the worldwide distribution of an ideal neutrally evolving phenotype

To assess the strength of the model fit to the available data, we simulate three different random cases to mimic the effect of population structure on the different relationship between population means and latitude. The $g \times g$ relationship matrix \mathbf{C} is estimated from the microsatellites as detailed above. A total of 100,000 random draws were each then correlated with the latitude for each of the groups and recorded to form a test distribution. We repeated this procedure once for each of three scenarios: 121 uncorrelated means, 121 correlated means drawn from the multivariate normal using our estimated \mathbf{C} matrix, and one with seven groups from Africa and another seven from Europe. The last configuration is made to approximate the sampling scheme used in Holliday's (1995) canonical study of the distribution of body form, and therefore the strength of the model fit within a latitudinally diverse but smaller data set.

Model fitting

To meet the challenges presented by multifactorial causation and non-independence, we adopt a mixed fixed and random effect model fitting approach (Henderson, 1953, 1976, 1984; Meyer, 1985; Pemberton, 2008; Thompson, 2008). Versions of it are widely used in comparative biology with applications to both interspecific (Lynch, 1991; Housworth et al., 2004) and intraspecific problems (Stone et al., 2011).

We fit and compare the three different kinds of models described in the introduction to the various phenotypes. In the first case (latitude model), we model the group means of a phenotype using absolute latitude as a fixed effect reflecting the putative effects of natural selection from climate. The second model (structure model) includes the random effect of population structure reflecting an assumption that only random genetic drift, mutation and gene flow are occurring in the population. Our third model (structure and latitude model) combines the effects of population structure and latitude into a single model as fixed and random effects, respectively.

We also use the structure model to predict absolute latitude. There may be an association between among-group relatedness and latitude, much in the same way that one can have a genotype-by-environment correlation that confounds estimates of quantitative genetic parameters within groups. This is important to know because the interpretation of the combined population structure and latitude model can be complicated if the two effects are highly confounded with one another. We might find, for instance, that the effect of latitude is substantial and yet prefer the structure-only model. This may take place if the latitude and structure terms have very similar effects on a phenotype such that combining both overfits the model.

Models were fitted using in the R statistical computing environment (R Core Team, 2013) using the MCMCglmm package (Hadfield, 2010). This method is a Bayesian approach that uses a Markov Chain Monte Carlo sampling strategy to estimate the posterior distributions for model terms and assess the goodness of fit of models. Posterior distributions for model variables were estimated by sampling every 50 iterations over a span of 100,000 iterations of a Markov Chain after a burn-in period of 100,000 iterations. We used inverse gamma distributed priors for the random effect and residual terms, each with one degree of freedom, and a mean set to half of the observed among-group variance. The random effect of population structure was made up of the matrix \mathbf{C} subject to singular value decomposition and fit using a constant covariance function following Stone et al. (2011). Multiple bouts of sampling from the Markov Chain indicated that Monte Carlo error was minimal.

The fits of the models were compared using the Deviance Information Criterion ($D.I.C.$), with lower $D.I.C.$ values reflecting better fitting models (Spiegelhalter et al., 2002). This is very similar in concept and application to the Akaike Information Criterion. We used the differences in the $D.I.C.$ ($\Delta D.I.C.$) to compare the fits of the three different models for each phenotype. According to the criteria of (Spiegelhalter et al., 2002), a value of $|\Delta D.I.C.|$ within two indicates that two models are roughly equivalent. Moreover, values in the range $2 \leq |\Delta D.I.C.| < 7$ indicate moderate support for the better fitting model and those with values of seven or greater have very strong support.

We also estimate Pearson's correlations and their 95% confidence intervals (± 2 s.e.) among group means for the morphological traits and absolute latitude to place them in the context of the rest of the literature.

All analyses were written in and using packages from the R statistical computing language (R Core Team, 2013). Packages used include MCMCglmm (Hadfield, 2010) and ape (Paradis et al., 2004). Data and simulation and analysis code, including the implementation of the mixed model using the MCMCglmm, are available from the authors.

Results

Correlations among group means and absolute latitudes along with their confidence intervals are presented in Table 1. A number of traits present robust correlations with latitude in accordance with ecogeographic predictions.

Fig. 3 shows the results of computer simulations of the evolution of a single neutral characteristic for a number of independently evolving groups equal to the number of groups used in this analysis (121), a similar number of non-independently evolving groups structured using the among-group genetic relationship matrix used in our comparative analysis here, and a small number (14) of non-independently evolving groups from Africa and Europe that mimic the geographic distribution and sampling scope of previous studies of human variation (Holliday, 1997). Each of 100,000 draws from a multivariate random normal distribution

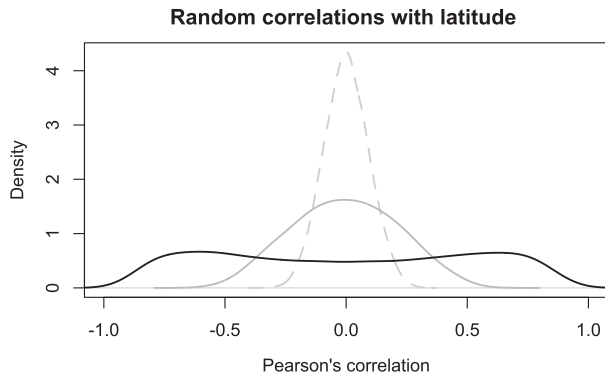


Figure 3. A density plot of the distribution of Pearson's correlations between absolute latitude and simulated neutrally evolving phenotypes. The dashed gray line represents the results of simulations involving 121 random draws from an ordinary normal distribution in each iteration. The solid gray line represents results of draws from a multivariate random normal distribution from the relationship matrix used as a random effect in the analysis of phenotypes in this study. The black line represents the results using multivariate random normals with a subset of 14 groups, seven from Africa and seven from Europe. Each distribution is made from 10,000 iterations of their respective simulations.

were correlated to absolute values of the latitudes associated with the groups sampled for genetic data. It is very clear that strong spurious relationships between population means and putative causes of natural selection – in this case those associated with latitude – will occur even when a relatively large number of groups are sampled. Very small samples (i.e., the 14 representing a latitudinal distribution) are likely to show strong correlations ($r^2 > 0.5$) of group means with latitude simply because of the presence of an underlying population structure. Note that the estimates of correlations of some traits with latitude presented in Table 1 that appear to be different from zero under the model of independent evolutionary outcomes are quite likely under the neutral model presented in the simulation.

The results of the three model fits for each morphology are presented in Table 2. For all phenotypes, the model that included population structure fits better than a model that only included the fixed effect of absolute latitude. In each case, a substantial proportion of the among-group variance in trait means is attributable to population structure. While the differences in the modal values of the posterior distributions of different traits may seem substantial, the widths of the posterior credibility intervals give us pause in declaring one trait to be more affected by population structure than another.

Combining the fixed effect of latitude with the random effect of population structure leads to moderate to high improvements (decrease of $D.I.C. > 2$) in fit over the population structure model in the case of bi-iliac breadth, radial length, tibial length, and the brachial index. In each of these cases, the 95% posterior credibility interval of the estimate of the latitude term does not include zero within its bounds, indicating a substantial effect of latitude on the traits. Radial length presents an interesting case in that the latitude-only model does not show a substantial effect of climate (latitude again as its proxy) on the distribution of group means. The effect of selection is not apparent until relatedness among groups is taken into account.

In the case of the humeral and femoral lengths, the model fits for the population structure and the structure and latitude models are almost equivalent (difference of $|D.I.C.| \leq 2$). The 95% posterior credibility interval for the latitude term overlaps with zero in both the latitude-only model and the structure and latitude model.

Femoral head diameter presents an interesting case, in that the population structure-only model appears to be the best fit among

Table 2
Model fit summaries for morphological traits.

Trait	Model	D.I.C.	Latitude (C.I.)	Structure (C.I.)
Bi-iliac breadth	Latitude	908.5	0.46 (0.29–0.60)	
	Structure Latitude and structure	849.7 847.5	0.31 (0.13–0.49)	0.68 (0.44–0.79) 0.58 (0.33–0.76)
Femoral head	Latitude	477.2	0.07 (0.05–0.09)	
	Structure Latitude and structure	450.7 456.2	0.03 (–0.01 to 0.06)	0.57 (0.33–0.71) 0.50 (0.20–0.66)
Humerus length	Latitude	930.5	0.11 (–0.25 to 0.01)	
	Structure Latitude and structure	894.1 894.1	–0.07 (–0.27 to 0.13)	0.46 (0.28–0.76) 0.54 (0.28–0.75)
Radius length	Latitude	888.1	–0.11 (–0.25 to 0.01)	
	Structure Latitude and structure	843.2 835.8	–0.20 (–0.36 to –0.04)	0.60 (0.38–0.82) 0.70 (0.41–0.82)
Femur length	Latitude	1030.0	0.10 (–0.14 to 0.32)	
	Structure Latitude and structure	995.1 993.1	–0.16 (–0.46 to 0.13)	0.58 (0.26–0.76) 0.56 (0.27–0.77)
Tibia length	Latitude	1009.9	–0.24 (–0.46 to –0.03)	
	Structure Latitude and structure	970.2 964.3	–0.38 (–0.66 to –0.13)	0.63 (0.33–0.84) 0.60 (0.34–0.82)
Brachial index	Latitude	438.5	–0.06 (–0.08 to –0.04)	
	Structure Latitude and structure	411.4 407.0	–0.05 (–0.07 to –0.02)	0.61 (0.35–0.81) 0.42 (0.18–0.75)
Crural index	Latitude	380.8	–0.07 (–0.09 to –0.05)	
	Structure Latitude and structure	338.4 351.8	–0.05 (–0.07 to –0.03)	0.71 (0.50–0.89) 0.46 (0.19–0.71)

D.I.C. values for models with the best fit for each trait are **bold**. The 'Latitude' column gives the fixed effect estimate as a slope. 'Structure' is the proportion of variance attributable to population structure. Posterior modes and 95% credibility intervals (C.I.) for each model are indicated in parentheses.

the models. When population structure is combined with the latitude term, the model has a poorer fit and the 95% posterior C.I. of the latitude term includes zero within its bounds. This stands in stark contrast to the latitude-only model, which shows a very strong and positive effect of latitude on femoral head diameter, though this is evidently spurious.

Likewise, the structure-only model for the crural index fits the data much better than either of the other two models. This is in spite of fairly strong negative estimates of the latitude term in both the latitude-only and combined latitude and structure models. This unusual result could reflect a strong confounding of population structure and latitude by a chance association between the effects of random genetic drift and latitude.

This interpretation of the crural index is all the more plausible given the results we obtained from fitting latitude as a trait using the population structure model. Doing so yields a high proportion of variance in latitude associated with population structure (mean of the posterior distribution, 0.61; C.I. 0.51–0.70). This points to dependencies between the two on account of the spatial distribution of genetic similarity among human groups.

Discussion

We draw three primary conclusions from these results. First, population structure is an important part of the global distribution of human body shape and size. Second, ecogeographic patterning is motivated by a variety of evolutionary forces. Finally, since what we take to be ecogeographic patterns are in part a result of population history and structure, referential models based on among-group variation in humans may be misleading as the population history of humans is a unique event. Taken together, our conclusions suggest that a new synthesis of comparative, experimental, and theoretical work is needed to understand hominin phenotypic evolution.

Multiple causes of among-group variation

Population structure accounts for a substantial portion of the among-group variance in all traits considered in this study. There is low support for models that do not include the term representing population structure. This indicates that the distribution of group means of traits is in large part a function of the particular population history that led to the distribution of human genetic variation that we see today. Relationships between a trait and a putative driver of natural selection or an environmental effect (e.g., nutritional status) will reflect a combination of this particular history (i.e., gene flow and random genetic drift) and whatever general effect of natural selection may be at work. The concordance between genetic distances estimated using cranial traits and those from molecular characteristics (Roseman, 2004; Harvati and Weaver, 2006; von Cramon-Taubadel, 2009) suggests that the same problem will affect evolutionary studies of morphologies from other anatomical regions in addition to aspects of postcranial form (Betti et al., 2012, 2013).

The results of the simulations of the evolution of a single neutral characteristic (Fig. 3) show that spurious relationships between the distribution of population trait means and putative selection pressures in globally distributed samples will be common if among-group variance is attributable to neutral additive differences among groups. Situations where the sampled groups are distributed along what amounts to a single geographic axis that runs parallel to a putative selective pressure will exacerbate this problem. This worst case scenario is exemplified by the circumstances represented by the solid black line in Fig. 3, where seven groups from Africa and seven from Europe were sampled (similar to the modern human sampling by Holliday, 1995, but using a different method). In this case, the latitude differences between the sampled groups lie along a major axis of human genetic variation, virtually guaranteeing substantial correlations (positive or negative) between latitude and group means. Indeed, the hypothesis of no relationship, typically the null hypothesis for tests in the human comparative literature, is a substantially less likely outcome than very strong correlations. This result will be generalizable to statistics other than Pearson's correlation coefficient.

Studies of among-group variation in humans using small numbers of sampled groups are likely to misinform researchers about the evolutionary causes of among-group variation when they use traditional methods. Even in studies that sample more groups, we will expect great variance in the estimates of terms modeling relationships between traits and hypothesized drivers of natural selection. As our sample is much larger and globally representative than is typical, it represents the present best-case scenario, and yet our estimates are beset with very high degrees of uncertainty.

This is not just a concern for studies struck from an ecogeographic approach or adaptationist mold. Our results indicate that there is substantial residual variance among groups, not all of

which may be attributable to natural selection. We know that environment can have strong effects on phenotypes related to long bone length and other skeletal traits, and the distribution of nutritional status or disease across geography may lead to phenotypic variation among groups (Wells and Stock, 2011). This poses steep challenges for the study of population structure using quantitative traits, which assume that all among-group variance is genetic and assume that all deviations from neutrality are attributable to natural selection (Brommer, 2011). While the neutral null hypothesis has been used to great effect in building a better understanding of human variation in both the cranium (Roseman and Weaver, 2007) and post-cranium (Betti et al., 2012, 2013) alike, rejecting the null hypothesis probably cannot be taken as an indicator that natural selection is uniquely responsible for the observed deviation from a neutral expectation.

We can draw two larger points from all of this. First, if among-group trait variation is substantially the result of genetic variation in groups, then human populations are not 'natural experiments' (Auerbach, 2007; King, 2010), as their entangled prehistories render them non-independent from one another. Bouts of selection in ancestral populations, along with the cumulative effects of random genetic drift and gene flow, will all shape the variation among groups. Perhaps more importantly, however, is the reinforcement of the notion that among-group human phenotypic variation is shaped by multiple environmental and evolutionary forces. This is not a new idea and it represents the core of a consensus that is coalescing in the field (e.g., Lynch, 1990; Relethford and Blangero, 1990; Relethford, 1994; Relethford et al., 1997; Harvati and Weaver, 2006; Smith et al., 2007; Wells and Stock, 2011; Auerbach, 2012; Betti et al., 2012, 2013; von Cramon-Taubadel et al., 2013). In combination with this growing body of literature, our results point toward the need to bind theory and method in an explicit way that allows us to compare complex models that reflect the diverse causes of variation.

Ecogeographic patterns and evolutionary processes

Ecogeographic patterning of human body form is not entirely attributable to clinally distributed natural selection. Population structure arising from random genetic drift, mutation, and gene flow over human population history plays a role in structuring among-group morphological differences. The picture of human body form variation that we draw from these results is multifactorial, evolutionarily dynamic, and population historically contingent.

Our results here do not overturn the previous notion that natural selection motivated by climate influences human body form (Hiernaux, 1976; Roberts, 1978; Trinkaus, 1981; Ruff, 1991). If anything, we add strength to and round out these claims by showing that they are probably not entirely the product of random genetic drift and gene flow. The distal elements of the limbs (radius length and tibia length) show signs of being subject to natural selection and population structure alike, such that shorter distal elements are associated with latitudes farther from the Equator. This accords well with adaptive accounts of human body form variation meant to explain the apparent conformation of limb proportions to Allen's rule (Holliday, 1999; Ruff, 2002). Neither the lengths of the humerus nor the femur show a strong relationship with latitude in any model.

The hypothesis that there is natural selection on bi-iliac breadth as predicted by the cylindrical model of Bergmann's rule (Ruff, 1994) is also supported in our results. Both of our estimates of the effect of latitude, however, are much smaller than the one arrived at by Ruff (1994), who reported an effect of latitude on bi-iliac breadth of 0.78 mm/degree (C.I. [\pm 2 s.e.] 0.66–0.91) in his mixed sex analysis. Reanalysis of the sexes separately yielded

estimates of 0.85 mm/degree (C.I. [± 2 s.e.] 0.69–1.01) for females and 0.74 mm/degree (C.I. [± 2 s.e.] 0.56–0.92) for males (reanalyzed using data presented in Ruff, 1994: Table 2). Our latitude-only model (males only) yielded an estimate of 0.46 mm/degree (C.I. 0.29–0.60), and we found an estimate of 0.31 mm/degree (posterior C.I. 0.13–0.49) for the combined latitude and structure model.

In the case of the latitude-only model, the difference between the two studies may be explained by the difference in sample composition, in that our sample draws more heavily on the Americas than does Ruff's (as variance in body breadth is smaller in the Americas; see Auerbach, 2012). Strictly speaking, however, this should not matter if all populations are independent from one another so long as we have roughly similar coverage across latitude, which we do. The much shallower slope of the latitude term estimated using the latitude and structure model, on the other hand, suggests that Ruff's estimate is influenced by population structure and sample selection.

In the case of the femoral head diameter, the latitude-only model shows a strong effect of latitude on the distribution of the phenotype, while the combined latitude and structure model does not (the posterior C.I. of latitude overlaps zero). This may suggest that the observed robust relationship between latitude and femoral head diameter (slope = 0.07 mm/degree latitude, $r = 0.48$; Tables 1 and 2) is an artifact of population structure. Conversely, the effect of latitude on radial length is not apparent until population structure is taken into account. In these cases, we may be more confident that latitude and population structure are not hopelessly confounded.

This multi-causal and evolutionarily dynamic view of the distribution of human body form complements recent results in the study of body mass (Foster and Collard, 2013) and other postcranial skeletal morphologies (Betti et al., 2012, 2013), which suggest that population structure and episodic bouts of selection are important for the global distribution of human body form. Betti et al. (2013) showed that there are at least regional dependencies of genetic distances estimated from pelvic morphology and geographic distances (a proxy for overall genetic distance). Foster and Collard (2013) argued that the relationship between body mass and latitude did not hold up except when extremes of latitude in the Northern Hemisphere were included in their analysis. This suggests that selection is appreciably strong only at the extremes of latitude (and, ostensibly, climatic factors) (as also argued using empirical trait distributions by Auerbach, 2007), and that other forces of evolution and environment predominate in the remainder of the comparisons.

Caution is essential, however, in interpreting the results of our analyses. The results for the crural index are a good case-in-point, where the population structure-only model is the best fit despite a substantial apparent effect of latitude in the combined model (Table 2). Because latitude is distributed in a way that can sometimes be very similar to the distribution of a neutrally evolving trait, it is possible that selection and random genetic drift might be highly confounded in some cases. This confounding of terms in the model may account for this result rather than a genuine lack of effect of latitude. As there is such high evolutionary variance for phenotypic characteristics (Rogers and Harpending, 1983; Lynch, 1990), we do not expect it to happen all of the time, but it is a concern that we all should carry forward into future investigations.

Put another way, our model assumes that the genetic differences among groups were laid out before natural selection acted on the populations. This is most certainly an oversimplification given what we know about how humans came to occupy the landscape in extremely cold environments during the Pleistocene and Holocene. The majority of the genetic variation in humans living in temperate and Arctic environments is recently derived from a tropical or near-tropical region (Henn et al., 2012). Genetic drift and directional

selection acting on ancestral populations could have occurred at the same time and our present analysis may be affected by this omission.

Implications for modeling the fossil record

A central conclusion of this study is that the evolution of human body form is contingent on population history and partly or entirely stochastic in nature. This has a number of implications for the way in which we conceptualize evolutionary problems in the hominin fossil record. First among these is the fact that the particular history that went into generating modern human variation today is unique and not referable to any arbitrary point in the past. When researchers estimate the effects of hypothesized causes of natural selection in fossil taxa phenotypes using modern humans as a referential model (Ruff, 1991; Holliday, 1995), their estimates may largely reflect this history (as also shown by Betti et al., 2012). It may not be possible to articulate generalizable principles of how human populations respond to natural selection in the absence of explicit accounts of population history.

Hominin evolution and migration have followed the same broad geographic pattern for various groups (e.g., *Homo erectus* and Neandertals) in that migrations from Africa to areas in higher latitudes are featured more than once in the fossil record (Gunz et al., 2009). This could cause a similar entanglement of geographically distributed putative drivers of selection in fossil groups as we see in Holocene *Homo sapiens*. Each hominin taxon will be subject to idiosyncratic population histories, and correlations between geographic dispersion and selection factors are likely to be a regular feature of recent hominin evolution. Moreover, recent studies of ancient DNA, for example, suggest population substructure among late Pleistocene hominins (Dalen et al., 2012; Meyer et al., 2013). Given that genetic drift shaped the morphological variation among late Pleistocene groups and humans living today (Weaver et al., 2007; Weaver, 2009), we may expect similar effects among fossil hominin groups in other phenotypes.

Referential models have also been extensively used to argue that some traits are more useful for reconstructing population and species relationships (Smith et al., 2007, 2013; von Cramon-Taubadel and Smith, 2012). Our results suggest that none of the traits examined herein are more reflective of population structure than others because the posterior credibility intervals of the estimates of the proportion of variance attributable to population structure overlap quite a bit. Rogers and Harpending (1983) pointed out that a single completely heritable characteristic has as much evolutionary information as a single dinucleotide locus, which is to say, very little. The conclusion that we draw from the combination of this theoretical insight with our empirical result is that there is no good reason to believe that we will be able to identify characteristics that are better for reconstructing phylogeny or population history. Even if some traits were more useful in a given set of populations, there is no theoretical reason to suppose that this is a general property across all groups we study. In light of this, we conclude that limiting characteristics for use in phylogenetic or population structural analyses based on criteria from comparative studies likely excludes perfectly good data.

Conclusions

Our results provide evidence for multifactorial and population historically contingent evolutionary causes of among-group variation in a range of postcranial phenotypes. We suggest here that modeling natural selection in terms of a cline that can be represented by a regression term may not be appropriate. Drawing on the entanglement of latitude and population history evident in our

results, we may profit from building models that are spatially and temporally explicit and include neutral forces of evolution and natural selection alike. This has been done in a post hoc manner in many studies of human phenotypic evolution to explain deviations from natural selection-based models of phenotypic variation in terms of population history (e.g., Auerbach, 2012), and to explain deviations from neutral models in terms of natural selection (Roseman and Weaver, 2004). The theoretical and methodological foundations of this synthesis are already available in evolutionary biology (Hansen, 1997; Hansen et al., 2008). Our study advocates for the great potential of bringing these approaches into biological anthropology, and adding the assets of functional morphological, molecular, archaeological, and paleontological approaches to bear on the study of hominin evolution.

Acknowledgments

We owe a great debt of gratitude to a number of individuals who have been instrumental in the completion of this study. Special thanks to Chris Ruff and Trent Holliday for graciously sharing their data, and for providing very useful comments on a previous version of this paper. Thanks to Graciela Cabana, Libby Cowgill, Mark Grabowski, Andrew Kramer, and Tim Weaver, who also read previous versions of this work and provided helpful insights. We are also grateful for the kind, encouraging, and thorough comments from three anonymous reviewers and Associate Editor at JHE. Elizabeth Mercer Roseman and Curtis C. Roseman provided very helpful editorial assistance. Errors and omissions are our own. Thanks the many institutions in Africa, Europe, Asia, and North America that granted access to skeletal collections used in this study. Data collection for this study was made possible, in part, by grants from the Joanna Jackson Goldman Memorial Prize and the National Science Foundation (Graduate Research Fellowship and DDIG BCS #0550673).

Appendix A. Supplementary material

Supplementary data related to this article can be found online at <http://dx.doi.org/10.1016/j.jhevol.2014.07.006>.

References

- Ackermann, R.R., Cheverud, J.M., 2004. Detecting genetic drift versus selection in human evolution. *Proc. Nat. Acad. Sci.* 101 (52), 17946–17951.
- Allen, J.A., 1877. The influence of physical conditions in the genesis of species. *Rad. Rev.* 1, 108–140.
- Ashton, K.G., Tracy, M., de Queiroz, A., 2000. Is Bergmann's rule valid for mammals? *Am. Nat.* 156 (4), 390–415.
- Auerbach, B.M., 2007. Human skeletal variation in the New World during the Holocene: effects of climate and subsistence across geography and time. Ph.D. Dissertation, Johns Hopkins University School of Medicine.
- Auerbach, B.M., 2010. Giants among us? Morphological variation and migration on the Great Plains. In: Auerbach, B.M. (Ed.), *Human Variation in the Americas: The Integration of Archaeology and Biological Anthropology*. Southern Illinois University, Carbondale, pp. 172–214.
- Auerbach, B.M., 2011a. Reaching great heights: changes in indigenous stature, body size and body shape with agricultural intensification in North America. In: Pinhasi, R., Stock, J.T. (Eds.), *Human Bioarchaeology of the Transition to Agriculture*. John Wiley & Sons, Inc., London, pp. 203–233.
- Auerbach, B.M., 2011b. Methods for estimating missing human skeletal element osteometric dimensions employed in the revised Fully technique for estimating stature. *Am. J. Phys. Anthropol.* 145 (1), 67–80.
- Auerbach, B.M., 2012. Skeletal variation among early Holocene North American humans: implications for origins and diversity in the Americas. *Am. J. Phys. Anthropol.* 149 (4), 525–536.
- Auerbach, B.M., Ruff, C.B., 2004. Human body mass estimation: a comparison of “morphometric” and “mechanical” methods. *Am. J. Phys. Anthropol.* 125 (4), 331–342.
- Auerbach, B.M., Ruff, C.B., 2006. Limb bone bilateral asymmetry: variability and commonality among modern humans. *J. Hum. Evol.* 50 (2), 203–218.
- Auerbach, B.M., Sylvester, A.D., 2011. Allometry and apparent paradoxes in human limb proportions: implications for scaling factors. *Am. J. Phys. Anthropol.* 144 (3), 382–391.
- Bergmann, C., 1847. Über die Verhältnisse der Wärmeökonomie der Thiere zu ihrer Größe. *Göttinger Stud.* 1, 595–708.
- Betti, L., von Cramon-Taubadel, N., Lycett, S.J., 2012. Human pelvis and long bones reveal differential preservation of ancient population history and migration out of Africa. *Hum. Biol.* 84 (2), 139–152.
- Betti, L., von Cramon-Taubadel, N., Manica, A., Lycett, S.J., 2013. Global geometric morphometric analyses of the human pelvis reveal substantial neutral population history effects, even across sexes. *PLoS One* 8 (2).
- Blackburn, T.M., Gaston, K.J., Loder, N., 1999. Geographic gradients in body size: a clarification of Bergmann's rule. *Divers. Distrib.* 5, 165–174.
- Brommer, J., 2011. Whither P_{ST} ? The approximation of Q_{ST} by P_{ST} in evolutionary and conservation biology. *J. Evol. Biol.* 24 (6), 1160–1168.
- Cowgill, L.W., Eleazer, C.D., Auerbach, B.M., Temple, D.H., Okazaki, K., 2012. Developmental variation in ecogeographic body proportions. *Am. J. Phys. Anthropol.* 148 (4), 557–570.
- Crognier, E., 1981. Climate and anthropometric variations in Europe and the Mediterranean area. *Annl. Hum. Biol.* 8, 99–107.
- Dalen, L., Orlando, L., Shapiro, B., Brandstrom-Durling, M., Quam, R., Gilbert, M.T.P., Fernandez-Lomana, J.C.D., Willersley, E., Arsuaga, J.L., Gotherstrom, A., 2012. Partial genetic turnover in Neandertals: continuity in the east and population replacement in the West. *Mol. Biol. Evol.* 29, 1893–1897.
- Davenport, C.B., 1933. The crural index. *Am. J. Phys. Anthropol.* 17 (3), 333–353.
- Eriksson, A., Manica, A., 2011. Detecting and removing ascertainment bias in microsatellites from the hgdp-ceph panel. *Genes Genomes Genet.* 1 (6), 479–488.
- Foster, F., Collard, M., 2013. A reassessment of Bergmann's rule in modern humans. *PLoS One* 8, e72269.
- Garofalo, E.M., 2012. Environmental and genetic effects on growth of the human skeleton: a bioarchaeological investigation. Ph.D. Dissertation, Johns Hopkins University School of Medicine.
- Graves, G.R., 1991. Bergmann's rule near the equator: latitudinal clines in body size of an Andean passerine bird. *Proc. Nat. Acad. Sci.* 88 (6), 2322–2325.
- Gunz, P., Bookstein, F.L., Mitteroecker, P., Stadlmayr, A., Seidler, H., Weber, G.W., 2009. Early modern human diversity suggests subdivided population structure and a complex Out-of-Africa scenario. *Proc. Nat. Acad. Sci.* 106, 6094–6098.
- Hadfield, J.D., 2010. MCMC methods for multi-response generalized linear mixed models: the MCMCglmm R package. *J. Stat. Softw.* 33 (2), 1–22.
- Hadfield, J.D., Nakagawa, S., 2010. General quantitative genetic methods for comparative biology: phylogenies, taxonomies and multi-trait models for continuous and categorical characters. *J. Evol. Biol.* 23 (3), 494–508.
- Hamilton, T.H., 1961. The adaptive significances of intraspecific trends of variation in wing length and body size among bird species. *Evolution* 15, 180–195.
- Hansen, T.F., 1997. Stabilizing selection and the comparative analysis of adaptation. *Evolution* 51 (5), 1341–1351.
- Hansen, T.F., Pienaar, J., Orzack, S.H., 2008. A comparative method for studying adaptation to a randomly evolving environment. *Evol.* 62 (8), 1965–1977.
- Harvati, K., Weaver, T.D., 2006. Human cranial anatomy and the differential preservation of population history and climate signatures. *Anat. Rec.* 288A (12), 1225–1233.
- Henderson, C.R., 1953. Estimation of variation and covariance components. *Biometrics* 9 (2), 226–252.
- Henderson, C.R., 1976. A simple method for computing the inverse of a numerator relationship matrix used in prediction of breeding values. *Biometrics* 32, 69–83.
- Henderson, C.R., 1984. *Applications of Linear Models in Animal Breeding*. University of Guelph Press, Guelph, Canada.
- Henn, B.M., Cavalli-Sforza, L., Feldman, M.W., 2012. The great human expansion. *Proc. Nat. Acad. Sci.* 109 (44), 17758–17764.
- Hiernaux, J., 1985. A comparison of the shoulder-hip-width sexual dimorphism in sub-Saharan Africa and Europe. In: Ghesquiere, J., Martin, R.D., Newcombe, F. (Eds.), *Human Sexual Dimorphism*. Taylor and Francis, London, pp. 191–206.
- Hiernaux, J., Froment, A., 1976. The correlations between anthropobiological and climatic variables in sub-Saharan Africa: revised estimates. *Hum. Biol.* 48 (4), 757–767.
- Hiernaux, J., 1976. Physical anthropology of the living populations of Sub-Saharan Africa. *Ann. Rev. Anthropol.* 5, 149–168.
- Higgins, R.W., Ruff, C.B., 2011. The effects of distal limb segments shortening on locomotor efficiency in sloped terrain: implications for Neandertal locomotor behavior. *Am. J. Phys. Anthropol.* 146 (3), 336–345.
- Holliday, T.W., 1995. Body size and proportions in the Late Pleistocene western Old World and the origins of modern humans. Ph.D. Dissertation, The University of New Mexico.
- Holliday, T.W., 1997. Postcranial evidence of cold adaptation in European Neandertals. *Am. J. Phys. Anthropol.* 104 (2), 245–258.
- Holliday, T.W., 1999. Brachial and crural indices of European Late Upper Paleolithic and Mesolithic humans. *J. Hum. Evol.* 36 (5), 549–566.
- Holliday, T.W., Ruff, C.B., 2001. Relative variation in human proximal and distal limb segment lengths. *Am. J. Phys. Anthropol.* 116 (1), 26–33.
- Housworth, E.A., Martins, E.P., Lynch, M., 2004. The phylogenetic mixed model. *Am. Nat.* 163 (1), 84–96.
- Hunley, K.L., Healy, M.E., Long, J.C., 2009. The global pattern of gene identity variation reveals a history of long-range migrations, bottlenecks, and local mate exchange: implications for biological race. *Am. J. Phys. Anthropol.* 139 (1), 35–46.

- James, F.C., 1970. Geographic size variation in birds and its relationship to climate. *Ecology* 51 (3), 365–370.
- King, K.A., 2010. Climate variation, biological adaptation, and postcranial metric variation in precontact North America. In: Auerbach, B.M. (Ed.), *Human Variation in the Americas: The Integration of Archaeology and Biological Anthropology*. Southern Illinois University, Carbondale, pp. 260–291.
- Kramer, A., Donnelly, S.M., Kidder, J.H., Ousley, S.D., Olah, S.M., 1995. Craniometric variation in large-bodied hominoids: testing the single-species hypothesis for *Homo habilis*. *J. Hum. Evol.* 29 (5), 443–462.
- Lahr, M.M., Foley, R.A., 1998. Towards a theory of modern human origins: geography, demography, and diversity in recent human evolution. *Yearb. Phys. Anthropol.* 41, 137–176.
- Lieberman, D.E., Pearson, O.M., Mowbray, K.M., 2000. Basicranial influence on overall cranial shape. *J. Hum. Evol.* 38 (2), 291–315.
- Long, J.C., Li, J., Healy, M.E., 2009. Human DNA sequences: more variation and less race. *Am. J. Phys. Anthropol.* 139 (1), 23–34.
- Lovegrove, B.G., Mowoe, M.O., 2013. The evolution of mammal body sizes: responses to Cenozoic climate change in North American mammals. *J. Evol. Biol.* 26 (6), 1317–1329.
- Lynch, M., 1990. The rate of morphological evolution in mammals from the standpoint of the neutral expectation. *Am. Nat.* 136 (6), 727–741.
- Lynch, M., 1991. Methods for the analysis of comparative data in evolutionary biology. *Evolution* 45 (5), 1065–1080.
- Malina, R.M., Reyes, M.E.P., Tan, S.K., Buschang, P.H., Little, B.B., Koziel, S., 2004. Secular change in height, sitting height and leg length in rural Oaxaca, southern Mexico: 1968–2000. *Annl. Hum. Biol.* 31 (6), 615–633.
- Martin, R., 1928. *Lehrbuch der Anthropologie in Systematischer Darstellung mit Besonderer Berücksichtigung der Anthropologischen Methoden für Studierende, Ärzte und Forschungsreisende*, Second edition. In: *Zweiter Band: Kraniaologie, Osteologie*. Gustav Fischer, Jena.
- Martins, E.P., 2000. Adaptation and the comparative method. *Trends Ecol. Evol.* 15, 296–299.
- Mayr, E., 1956. Geographical character gradients and climatic adaptation. *Evolution* 10 (1), 105–108.
- Meyer, K., 1985. Maximum likelihood estimation of variation components for a multivariate mixed model with equal design matrices. *Biometrics* 41 (1), 153–165.
- Meyer, M., Fu, Q., Aximu-Petri, A., Glocke, I., Nickel, B., Arsuaga, J.-L., Martínez, I., Gracia, A., Bermudez de Castro, J.M., Carbonell, E., Pääbo, S., 2013. A mitochondrial genome sequence of a hominin from Sima de los Huesos. *Nature* 505, 403–406.
- O'Neill, M.C., Dobson, S.D., 2008. The degree and pattern of phylogenetic signal in primate long-bone structure. *J. Hum. Evol.* 54, 309–322.
- Paradis, E., Claude, J., Strimmer, K., 2004. Ape: analysis of phylogenetics and evolution in R language. *Bioinformatics* 20 (2), 289–290.
- Paterson, J.D., 1996. Coming to America: acclimation in macaque body structures and Bergmann's rule. *Int. J. Primatol.* 17, 585–611.
- Pemberton, J.M., 2008. Wild pedigrees: the way forward. *Proc. R. Soc. B* 275 (1635), 613–621.
- Pemberton, T.J., DeGiorgio, M., Rosenberg, N.A., 2013. Population structure in a comprehensive genomic data set on human microsatellite variation. *Genes Genomes Genet.* 3 (5), 891–907.
- Potts, R., 1998. Environmental hypotheses of hominin evolution. *Yearb. Phys. Anthropol.* 41, 93–136.
- R Core Team, 2013. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria.** <http://www.R-project.org/>.
- Relethford, J.H., 1994. Craniometric variation among modern human populations. *Am. J. Phys. Anthropol.* 95 (1), 53–62.
- Relethford, J.H., Blangero, J., 1990. Detection of differential gene flow from patterns of quantitative variation. *Hum. Biol.* 62 (1), 5–25.
- Relethford, J.H., Crawford, M.H., Blangero, J., 1997. Genetic drift and gene flow in post-famine Ireland. *Am. J. Hum. Biol.* 69 (4), 443–465.
- Rensch, B., 1938. Some problems of geographical variation and species-formation. *Proc. Linn. Soc.* 150, 275–285.
- Roberts, D.F., 1978. *Climate and Human Variability*, Second edition. Cummings Publishing Company, Menlo Park, CA.
- Rogers, A.R., Harpending, H.C., 1983. Population structure and quantitative characters. *Genetics* 105 (4), 985–1002.
- Roseman, C.C., 2004. Detecting interregionally diversifying natural selection on modern human cranial form by using matched molecular and morphometric data. *Proc. Nat. Acad. Sci.* 101 (35), 12824–12829.
- Roseman, C.C., Weaver, T.D., 2004. Multivariate apportionment of global human craniometric diversity. *Am. J. Phys. Anthropol.* 125 (3), 257–263.
- Roseman, C.C., Weaver, T.D., 2007. Molecules versus morphology? Not for the human cranium. *Bioessays* 29 (12), 1185–1188.
- Roseman, C.C., Weaver, T.D., Stringer, C.B., 2011. Do modern humans and Neandertals have different patterns of cranial integration? *J. Hum. Evol.* 60 (6), 684–693.
- Ruff, C.B., 1991. Climate and body shape in hominid evolution. *J. Hum. Evol.* 21 (2), 81–105.
- Ruff, C.B., 1994. Morphological adaptation to climate in modern and fossil hominids. *Yearb. Phys. Anthropol.* 37, 65–107.
- Ruff, C.B., 2002. Variation in human body size and shape. *A. Rev. Anthropol.* 31 (1), 211–232.
- Ruff, C.B., 2010. Body size and body shape in early hominins: implications of the Gona Pelvis. *J. Hum. Evol.* 58 (2), 166–178.
- Ruff, C.B., Trinkaus, E., Holliday, T.W., 2002. Body proportions and size. In: Zilhão, A., Trinkaus, E. (Eds.), *Portrait of the Artist as a Child: the Gravettian Human Skeleton from the Abrigo do Lagar Velho and its Archaeological Context*, *Trabalhos de Arqueologia*, vol. 22. Instituto Português de Arqueologia, pp. 365–391.
- Ruff, C.B., Niskanen, J.-A., Jamison, P., 2005. Body mass prediction from stature and bi-iliac breadth in two high latitude populations, with application to earlier higher latitude humans. *J. Hum. Evol.* 48 (4), 381–392.
- Smith, H.F., Terhune, C.E., Lockwood, C.A., 2007. Genetic, geographic, and environmental correlates of human temporal bone variation. *Am. J. Phys. Anthropol.* 134 (3), 312–322.
- Smith, H.F., Ritzman, T., Otarola-Castillo, E., Terhune, C.E., 2013. A 3-D geometric morphometric study of intraspecific variation in the ontogeny of the temporal bone in modern *Homo sapiens*. *J. Hum. Evol.* 65 (5), 479–489.
- Spiegelhalter, D.J., Best, N.G., Carlin, B.P., van der Linde, A., 2002. Bayesian measures of model complexity and fit. *J. R. Stat. Soc. B* 64 (4), 585–639.
- Stevenson, R.D., 1986. Allen's rule in North American rabbits (*Sylvilagus*) and hares (*Lepus*) is an exception, not a rule. *J. Mammal.* 67, 312–316.
- Stinson, S., 1990. Variation in body size and shape among South American Indians. *Am. J. Hum. Biol.* 2 (1), 37–51.
- Stock, J.T., 2006. Hunter-gatherer postcranial robusticity relative to patterns of mobility, climatic adaptation, and selection for tissue economy. *Am. J. Phys. Anthropol.* 131 (2), 194–204.
- Stock, J.T., Pinhasi, R., 2011. Introduction: Changing paradigms in our understanding of the transition to agriculture: human bioarchaeology, behaviour and adaptation. In: Pinhasi, R., Stock, J.T. (Eds.), *Human Bioarchaeology of the Transition to Agriculture*. John Wiley & Sons, Inc., London, pp. 1–13.
- Stone, G.N., Nee, S., Felsenstein, J., 2011. Controlling for non-independence in comparative analysis of patterns across populations within species. *Phil. Trans. R. Soc. B* 366 (1569), 1410–1424.
- Sylvester, A.D., Kramer, P., Jungers, W.L., 2008. Modern humans are not (quite) isometric. *Am. J. Phys. Anthropol.* 137 (4), 371–383.
- Temple, D.H., Auerbach, B.M., Nakatsukasa, M., Scullin, P.W., Larsen, C.S., 2008. Variation in limb proportions between Jomon foragers and Yayoi agriculturalists from prehistoric Japan. *Am. J. Phys. Anthropol.* 137 (2), 164–174.
- Temple, D.H., Okazaki, K., Cowgill, L.W., 2011. Ontogeny of limb proportions in Late/Final Jomon period foragers. *Am. J. Phys. Anthropol.* 145, 415–425.
- Thompson, R., 2008. Estimation of quantitative genetic parameters. *Proc. R. Soc. B* 275 (1635), 679–686.
- Trinkaus, E., 1981. Neandertal limb proportions and cold adaptation. In: Stringer, C.B. (Ed.), *Aspects of Human Evolution*. Taylor and Francis, London, pp. 187–224.
- von Cramon-Taubadel, N., 2009. Congruence of individual cranial bone morphology and neutral molecular affinity patterns in modern humans. *Am. J. Phys. Anthropol.* 140, 205–215.
- von Cramon-Taubadel, N., Lycett, S.J., 2008. Brief communication: human cranial variation fits iterative founder effect model with African origin. *Am. J. Phys. Anthropol.* 136 (1), 108–113.
- von Cramon-Taubadel, N., Smith, H.F., 2012. The relative congruence of cranial and genetic estimates of hominoid taxon relationships: implications for the reconstruction of hominin phylogeny. *J. Hum. Evol.* 62 (5), 640–653.
- von Cramon-Taubadel, N., Stock, J.T., Pinhasi, R., 2013. Skull and limb morphology differentially track population history and environmental factors in the transition to agriculture in Europe. *Proc. R. Soc. B* 280, 20131337.
- Weaver, T.D., 2003. The shape of the Neandertal femur is primarily the consequence of a hyperpolar body form. *Proc. Nat. Acad. Sci.* 100 (12), 6926–6929.
- Weaver, T.D., 2009. The meaning of Neandertal skeletal morphology. *Proc. Nat. Acad. Sci.* 106, 16028–16033.
- Weaver, T.D., Roseman, C.C., Stringer, C.B., 2007. Were Neandertal and modern human cranial differences produced by natural selection or genetic drift? *J. Hum. Evol.* 53 (2), 135–145.
- Weaver, T.D., Roseman, C.C., Stringer, C.B., 2008. Close correspondence between quantitative- and molecular-genetic divergence times for Neandertals and modern humans. *Proc. Nat. Acad. Sci.* 105 (12), 4645–4649.
- Wells, J.C., Stock, J.T., 2011. Re-examining heritability: genetics, life history, and plasticity. *Trends Endocrinol. Metab.* 22 (10), 421–428.
- Whitlock, M.C., 2008. Evolutionary inference from Q_{ST} . *Mol. Ecol.* 17 (8), 1885–1896.
- Wolpoff, M.H., 1993. Reply to Dr. Foote. *Am. J. Phys. Anthropol.* 90 (3), 381–384.
- Wood, B., Lieberman, D.E., 2001. Craniodontal variation in *Paranthropus boisei*: a developmental and functional perspective. *Am. J. Phys. Anthropol.* 116 (1), 13–25.
- Yom-Tov, Y., Benjamini, Y., Kark, S., 2002. Global warming, Bergmann's rule and body mass: are they related? The chukar partridge (*Alectoris chukar*) case. *J. Zool.* 257 (4), 449–455.