THE ROYAL SOCIETY
PUBLISHING

Morphological variation in *Homo erectus* and the origins of developmental plasticity

Susan C. Antón[1], Hannah G. Taboada[1], Emily R. Middleton[1,2],
Christopher W. Rainwater[1], Andrea B. Taylor[3,4], Trudy R. Turner[5,6],
Jean E. Turnquist[7], Karen J. Weinstein[8] and Scott A. Williams[1]

[1]Department of Anthropology, New York University, New York, NY 10003, USA
[2]Department of Pathology and Anatomical Sciences, University of Missouri, M263 Medical Science Building, Columbia, MO 65212, USA
[3]Department of Orthopaedics, and [4]Department of Evolutionary Anthropology, Duke University School of Medicine, Durham, NC 27708, USA
[5]Department of Anthropology, University of Wisconsin-Milwaukee, Milwaukee, WI 53201, USA
[6]University of the Free State, Bloemfontein, South Africa
[7]Caribbean Primate Research Center and Department of Anatomy and Neurobiology (Retired), University of Puerto Rico Medical Sciences Campus, San Juan, Puerto Rico, USA
[8]Department of Anthropology and Archaeology, Dickinson College, Carlisle, PA 17013, USA

Homo erectus was the first hominin to exhibit extensive range expansion. This extraordinary departure from Africa, especially into more temperate climates of Eurasia, has been variously related to technological, energetic and foraging shifts. The temporal and regional anatomical variation in *H. erectus* suggests that a high level of developmental plasticity, a key factor in the ability of *H. sapiens* to occupy a variety of habitats, may also have been present in *H. erectus*. Developmental plasticity, the ability to modify development in response to environmental conditions, results in differences in size, shape and dimorphism across populations that relate in part to levels of resource sufficiency and extrinsic mortality. These differences predict not only regional variations but also overall smaller adult sizes and lower levels of dimorphism in instances of resource scarcity and high predator load. We consider the metric variation in 35 human and non-human primate 'populations' from known environmental contexts and 14 time- and space-restricted paleodemes of *H. erectus* and other fossil *Homo*. Human and non-human primates exhibit more similar patterns of variation than expected, with plasticity evident, but in differing patterns by sex across populations. The fossil samples show less evidence of variation than expected, although *H. erectus* varies more than Neandertals.

This article is part of the themed issue 'Major transitions in human evolution'.

1. Introduction

Homo erectus was the first hominin to exhibit extensive range expansion. Much like recent humans, this long-lived and widely dispersed species inhabited environments in equatorial Africa and more temperate Eurasia. As such, considerable work has been framed around understanding what made dispersal possible and what the broad geographic and temporal trends in variation might mean biologically for *H. erectus*. Recently, the regional variation in *H. erectus* has been described as 'human-like' [1], and by extension we have suggested that the dispersal and evolutionary longevity of the species may be related to human-like developmental (phenotypic) plasticity [2,3].

Developmental (phenotypic) plasticity is the ability to modify development in response to environmental conditions, resulting in variation in adult anatomy that is not genetically canalized [4]. Taxa with a high degree of plasticity should be able

to respond on short-term time scales to individual environmental or maternal environmental signals. Arguably this ability may also play an important role in moderating environmental influences too chronic for short-term accommodation and too short for genetic adaptation, as well as providing real advantages for occupying a broad range of environments [1]. A high degree of developmental plasticity is considered an important aspect of the human ability to occupy multiple different environmental niches.

Related to this plasticity, differences in body size, shape and dimorphism across human populations in part reflect levels of resource sufficiency and extrinsic mortality [5–8]. To be sure, body size, shape and sexual dimorphism have multifactorial causes: there is a genetic component to size and variation, and other environmental conditions such as temperature also influence the attainment of adult size. The latter is reasonably well understood, allowing consideration of other contributions to body size outcomes. Resource sufficiency includes any variable that influences the nutritional base, some of which are co-correlated with aspects of climate such as rainfall and seasonality. Extrinsic mortality can be defined generally as the external risks of mortality such as predator and parasite load, or in recent human environments, factors like homicide [9]. The theory that links shifts in body size and age at first reproduction to resource sufficiency and extrinsic mortality is relatively clear [1]. Resource sufficiency is positively correlated with extrinsic mortality and negatively correlated with adult body size; that is, decreases in resources lead to slow growth rate and small adult size, whereas increases in mortality favour early maturation usually leading to small body size. Extrinsic mortality related to predator load may differ somewhat from this expectation in instances when larger body size is advantageous for predator control or survival [10]. In these instances, early maturation but faster growth may favour the retention of large size, particularly in males. In humans, males and females are often argued to be differentially influenced especially by resource sufficiency, with human females being more strongly buffered from environmental vicissitudes and human males responding more dramatically to both resource excess and insufficiency. This difference is thought to be related to female buffering of infant brain size and to be marked in humans for this reason [11]. Such differential influence can alter dimorphism values if the female size change differs from that of males [11]. Extrapolating from living humans, this logic predicts that the skeletal record of *H. erectus* should show not only regional variations, but also overall smaller adult body sizes and lower levels of dimorphism in populations experiencing resource scarcity and high extrinsic mortality *if* the species shows human-like levels of plasticity [2].

To make the case that *H. erectus* is human-like in its plasticity and that this is meaningful for its biology, however, we have first to operationalize how we recognize human-like levels of plasticity. Importantly, we need to establish whether these patterns are unique to humans or shared with other widely dispersed non-human primates. We must also be able to reliably recognize the results of plasticity in both living populations and their skeletal remains. Previous studies have presumed that human variability is relatively high and that this degree of plasticity arose sometime during our lineage. They have further hypothesized that the point of origin for the plasticity is with *H. erectus* [1,3]. We aim to test this first by establishing whether a unique pattern is discernible

in humans that differs from other widely dispersed non-human primates and, second, by evaluating *H. erectus* relative to this pattern.

Because a major consequence of high developmental plasticity is differences in adult size *across* populations in different environmental contexts, testing whether *H. erectus* actually has human-like levels of plasticity requires comparing subunits of *H. erectus*. We thus argue that species-wide levels of variation and even region-wide levels are insufficient proxy measures, and we propose instead to examine the variation among paleodemes of *H. erectus*. Demes are local populations of polytypic species that actively interbreed with one another, that is, the smallest reproductive population of the species [12]. Paleodemes similarly refer to 'local' populations of fossil taxa that are inferred to have shared a closer gene pool than their geographically and temporally more distant relative populations. Paleodemes are thus temporally and geographically restricted fossil groupings that attempt to speak to the same local influence on past populations that demes do in the extant world [13]. The variation in form is an inherent characteristic of all biological populations and the question of variation and its significance is threaded throughout the study of *H. erectus.* However, the taxonomic level of focus and the evidence used to discuss this variation and its inferred significance for the biology of the species have shifted with time. Below we review some of the history of how the variation in *H. erectus* has been considered, what we know about the regional variation in the species and what we might need to assess to understand the population variation and the overall variability of the species.

2. A historical perspective on the meaning of variation in *H. erectus*

Most studies of fossil taxa consider variation primarily as a necessary first step in circumscribing species in order to understand overarching, species-wide themes and relationships to other taxa—the early history of *H. erectus* is no different. From its initial discovery in 1891, the variation among *H. erectus* fossils was used to infer higher level taxonomic differences. Scholars disagreed as to whether the family Hominidae could incorporate the Trinil 2 calotte—or whether, in fact, the calotte was that of an out-sized gibbon or other ape (e.g. [14–16]). When that question was resolved by the abundant fossil finds in Java and China in the 1930s, the significance in the variation among these specimens moved to the question of recognizing generic boundaries. In the early 1900s, inter-regional and sometimes intra-regional variants were recognized by generic attributions including *Sinanthropus*, *Pithecanthropus*, *Meganthropus* and even *Homo* (e.g. [17–21]). However, those most familiar with the fossils did not fully consider these designations to represent biologically significant differences (e.g. [18–21]). Weidenreich based his opinions on what he saw as a morphological bauplan shared across *Sinanthropus* and *Pithecanthropus* [19–21]. But the conclusions of other scholars were at least in part predicated on the observation that the metric variation among the Far Eastern fossils was no more than that within the only other known fossil 'men' of the time, the Neandertals [22,23]. The subsuming of these multiple genera and species into *H. erectus* by Mayr [24,25] thus formalized a biological reality already unofficially accepted by the majority of

paleoanthropologists at the time and set the stage for studies of intraspecific variation.

The remarkably complete KNM-ER 3733 and 3883 crania discovered by Richard Leakey's team in the 1970s shifted the geographic centre of the debate to East Africa and to the problem of just how much intra-specific variation *Homo erectus* could accommodate [26,27]. Leakey and Walker [26,28] argued that despite what were then significant differences in both geological age and geographic space between African and Asian fossils, the Koobi Fora specimens, nonetheless, showed the morphological neurocranial bauplan established by *H. erectus* in Asia. Indeed, they described KNM-ER 3733 as '...strikingly like that of *H. erectus* from Peking' [28, p. 573]. In contrast, other studies suggested that aspects of vault shape and specifically the apparent absence of some non-metric characters related to cranial superstructures showed that African *H. erectus* was not so clearly affiliated with Asian *H. erectus* [29]. In these instances, African *H. erectus* was often recognized as *H. ergaster* (e.g. [30–32]) even though this species designation for KNM-ER 992 originally folded in members of earlier *Homo* including KNM-ER 1805 [33]. Refinements to the chronostratigraphic framework of the regional fossil samples in the intervening decades have diminished and, in some cases, eliminated the time difference between the African and some Asian assemblages [34], thus partially removing time as an explanation for differences across geographic assemblages. Nonetheless, the original argument that KNM-ER 3733 and 3883 follow the *H. erectus* bauplan has been supported by additional fossil finds and by numerous assessments of non-metric, linear metric and three-dimensional morphometric datasets (see [35–40]).

As a result, later studies emphasize the cohesiveness of the East Asian sample and *H. erectus* as a relatively large brained and bodied taxon, despite the presence of smaller individuals across its range [36,37,40]. Average brain size in Asian *H. erectus* hovered around 1000 cc and although the Koobi Fora individuals were smaller (848 and 804 cc, respectively), they were both substantially larger than other earlier *Homo* from Koobi Fora (i.e. KNM-ER 1813 and 1470, in particular). Sealing the perception of a large *H. erectus* was the discovery by Leakey's team of the Nariokotome skeleton in 1985, with a relatively large brain (909 cc) and a large body [41]. This youth, coupled with large but less complete postcranial remains such as KNM-ER 1808, formed the basis for the supposition of large size in the species.

However, the discovery of small-sized individuals from Dmanisi, Georgia, and Koobi Fora and Olorgesailie, Kenya and possibly Gona, Ethiopia complicated the discussion of body size in the species definition and its influence on feature expression [38,42–45]. The new finds showed, in many instances, less development in cranial superstructures and sometimes more rounded and certainly absolutely thinner vaults. Despite size differences, the similarities between Georgian and African *H. erectus* as well as the overlapping of specific features with Asian specimens emphasized the association between the African and Asian fossils. The view of *H. erectus* as a widely dispersed, geographically variable and long-lived species has become broadly accepted [35,37,40]. Indeed, the view is so broadly accepted that some scholars even include *all* early African *Homo*, such as KNM-ER 1813, within *H. erectus* as well [46]. We accept *H. erectus* as a long-lived, broadly dispersed species, but we maintain that discrete characters exclude from that taxon specimens traditionally assigned to *H. habilis* and *H.*

rudolfensis (e.g. KNM-ER 1470, 1813, 1805; [47,48]). We thus agree that the global three-dimensional analyses that purportedly form the basis for subsuming all earlier *Homo* into *H. erectus* [46] are inadequate for species definition [49–51].

Homo erectus, by this definition, has larger average brain and body sizes and smaller average jaw and tooth sizes than earlier *Homo* and *Australopithecus*, and these differences have been used as evidence of *H. erectus*' acquisition of at least parts of the human 'morpho-behavioural package'. For example, larger brain and body size coupled with dental reduction is now thought to signal greater ranging and a foraging shift to a higher quality diet often inferred to have important components from animal resources [52–56]. In addition, an inferred shift in relative body size difference between the sexes, due to differential enlargement of females, has been used to argue for decreased sexual dimorphism and high energetic costs for *H. erectus* females [57]. Even though size ranges overlap across species [47] and differences in mean size may not be as great as previously thought, given evidence of smaller size in some *H. erectus* [58,59], differences in average size remain and have similar biological implications.

A parallel research trend uneasy with the idea that *H. erectus* was universally large examined the growing fossil dataset to keep regional variations visible—not to argue for specific distinctions across regions, but rather in an attempt to understand possible influences of differences in time, geography, climate and other selective factors, as well as isolation and drift [60–62]. Following Antón [40], most recent papers run multiple comparisons that consider first the entirety of *H. erectus*, in order to maximize sample size, and then subsets based on continent of origin, sometimes subdivided by time. The most common subsamples are African *H. erectus* (~1.9–0.9 Ma), which, for morphological and temporal reasons, usually includes the Dmanisi fossils and Asian *H. erectus* (1.6–perhaps 0.25 Ma), with the latter sometimes broken into earlier (Sangiran, Trinil and Zhoukoudian) and later (Ngandong/Sambungmacan) groups. When subdivided in this manner, some regional variation in size is evident both metrically and non-metrically (table 1). The Chinese fossils are more constricted across the asterionic region and have more vertical frontal squamae than the Indonesian groups [60,62]. The Asian samples are larger and thicker on average than the African and Georgian assemblages. Evidence of climatic adaptation, particularly correlating body size and minimum annual temperature, is present, but not strongly supported [59]. Most studies agree that there are some temporal trends, especially in overall size and particularly in average brain size [62–66], although a fair amount of size variation exists in all samples from all times [40,67].

By necessity, however, even these regionally based studies focus on temporally and geographically mixed samples of *H. erectus* that, by their very nature, embed multiple potential sources of variation within them, thus complicating conclusions and sometimes leading to mixed messages. This variation across the geographic range of the taxon has been used to equate the 'population' variation in *H. erectus* with that seen in extant *H. sapiens* and to argue by extension for some similarity in biology and dispersal capability in the two [2,67]. As noted above, a particular parallel has been made in aspects of the presence and degree of developmental plasticity between the two taxa as proxied by the phenotypic variation across regional samples of *H. erectus* [1–3]. But those studies that have tried to assess the variation *within* regional samples of *H. erectus* have disagreed substantially on how much

Table 1. Regional size variation in *H. erectus*.

	Africa/Georgia[a]	China[b]	Indonesia[c]
geological age			
range	1.77 – 1.5 Ma	0.78 – 0.25? Ma	1.6 – 0.25? Ma
brain size (cc)			
range	546 – 1067	855 – 1225	813 – 1251
mean	787 (with Dmanisi)	1028	1038
	863 (without Dmanisi)		933 (without Ngandong)
femur length (mm)			
range	386 – 485	378 – 413 (2)	433 – 455 (3)
mean	437.5 (with Dmanisi)	395	445
	450.5 (without Dmanisi)		
stature (cm)			
range	148 – 186	141 – 154 (2)	162 – 167 (3)
mean	164 (with Dmanisi)	148	158
	167 – 171 (without Dmanisi)		
body mass (kg)			
mean	52 (with Dmanisi)	47	52
	55 (without Dmanisi)		

[a]East African and Georgian results include both associated individuals and isolated elements. Crania include: KNM-ER 3733, 3883, 42700, KNM-WT 150000, Daka, OH 9, OH 12, Dmanisi D2280, D2282, D3444, D4500. Femora include: KNM-ER 736 and 737, OH 28 and 34, KNM-ER 1808 and adult estimates of KNM WT 15000 as well as body size estimates from Dmanisi 'large' and 'small' individuals and the Gona pelvis.

[b]Chinese results include only isolated elements. Crania include: Hexian and Zhoukoudian II, II, V, VI, X, XI, XII. Femora include: Zhoukoudian Fem I, Fem IV. Stature and body mass based on femora.

[c]Indonesian results include only isolated elements. Crania include: Ngandong 6, 7, 10, 11, 12, Sangiran 2, 4, 10, 12, 17, Trinil and Sambungmacan specimens. Femora are from Trinil but exclude Trinil I. Stature includes range as per Trinil and mean includes Ngandong Tibia B.

variation there is and what fossil variation might imply about sexual dimorphism. Spoor *et al.* [38] conclude that regional samples of *H. erectus* include nearly as much size variation as highly dimorphic apes, but Plavcan [64] argues that the variation in *H. erectus* is 'unremarkable' compared to even non-dimorphic extant primates when looking at the coefficient of variation (CV) and raw variation across the whole species. We aim to address these differing conclusions.

Regardless of which conclusion is correct, we argue that these studies are not suitable for assessing levels of population variation and by extension plasticity in *H. erectus*. First, temporally and geographically mixed regional samples subsume multiple subpopulations and are likely to mask population differences; thus we aim for paleodemes (as discussed above). Second, fossil *H. erectus* cannot be assessed in comparison with solely extant taxa but requires a frame of reference from other hominin paleodemes. And finally, developmental plasticity addresses the ability to vary in different contexts, but it is unclear whether the overall variation in a measure, that is CV, should be a good proxy for variability across populations. This can be tested (see below).

Here we attempt to operationalize a means of recognizing high developmental plasticity (a great ability to vary) that can be applied to skeletal and fossil records. We emphasize that this is an initial step in the process and that additional well-constrained extant datasets and more fossils are necessary to test the results found here.

Our aim is to provide a comparative framework in order to ask the following specific questions:

(1) What does the human pattern of population difference in mean size (body and other size measures) look like between closely related but differently resourced populations?
   (a) Do human populations in different environments show statistical differences in absolute size or in within population size variation?
   (b) Do skeletal data yield similar results?
(2) Do close populations of species of widely dispersed non-human primate (nhp) genera show the same pattern as humans?
   (a) Do nhp populations in different environments show statistical differences in absolute size or in within population size variation?
   (b) Do mixed sex CVs and skeletal data yield similar results?
(3) What does the *H. erectus* pattern of difference in size look like across paleodemes?
   (a) Do *H. erectus* paleodemes show statistical differences in absolute size and does variation within populations differ?
   (b) Do *H. erectus* paleodemes differ from variation among paleodemes of other fossil groups?

3. Material and methods

To begin to address the above inter-related issues, we compile a fossil and an extant dataset designed to consider the variation in both somatometric and osteometric (cranial and postcranial) data.

For Questions 1 and 2, we set out to assess how human and non-human primate populations vary across environments using matched datasets of osteometric and somatometric variables. We define environment broadly to include differences in climate, resource base and any aspect of extrinsic mortality (including predation, parasite loads, etc.). We purposefully limit our comparisons to intraspecific populations in closely circumscribed geographic areas in an attempt to control for similarities in ancestry and adaptation to overall climate (e.g. Bergmann's and Allen's rules). This is intended as a conservative approach minimizing differences between populations. Matched datasets are, however, challenging to assemble, and as a consequence, we consider this a preliminary effort and a call for the necessity and usefulness of detailed demic level environmental and somatic data collection from living primates. Because most of the *H. erectus* remains are craniodental, we have prioritized datasets with at least some cranial measures (see Methods section). For Question 3, we compare these to fossil hominin paleodemes across four species (table 2). All paleodemes represent adult individuals with localities/samples chosen to maximize sample size and limit temporal and geographic distribution. However, the geographic and temporal spread differs between them.

(a) Extant samples (Questions 1 and 2)

Question 1 addresses variability in living humans, and we stress that our aim was not to replicate the abundant work that has established the extensive geographic and climatic variation of recent human populations across the globe [68,69]. Instead we ask whether closely related populations in different environments yield different body sizes (as would be predicted from differing conditions) and whether the degree of variation (CV) differs across these populations. We pay close attention to sex-specific changes within and between populations and whether mixed-sex samples yield similar results, and we assess the signals from somatometric and osteometric data. To accomplish this, we compare amongst living samples and between living and skeletal samples matched for ancestry or geography.

We compare skeletal and somatometric data for six recent and one archaeological Arctic population (figure 1; table 3, electronic supplementary material, tables S1 and S2). Comparisons between these populations control for differences due to ancestry and thermal stress but vary on nutritional base and extrinsic mortality. The Point Hope assemblage represents a pre-contact, coastal group living about 200 km north of the Arctic Circle. We focus on remains from the Old Tigara cemetery at Point Hope that dates from AD 1200 to 1700. The archaeological evidence, confirmed by dental microwear, suggests a subsistence base primarily oriented around marine mammals, particularly whale hunting [71,72]. Thus these individuals were leading a completely traditional coastal lifestyle. From the archaeological assemblage at the site, this lifestyle and diet were similar to those of the coastal Inuit at the time of European contact. The diet at contact in this area consisted of 35–65% protein, 30–60% fat and very little carbohydrate [73]. We compare this Point Hope sample to a Wainwright Inuit somatometric dataset collected in the late 1960s and five other indigenous Alaskan samples collected in the earliest 1900s. The individuals sampled for the Wainwright study included all the indigenous individuals living in Wainwright Alaska in 1967–1968 [74]. These individuals were permanently domiciled in Wainwright and include only those individuals who trace both their maternal and paternal ancestry to indigenous people of the immediate area. This area of ancestry includes the coastal region north of Point Hope. The Wainwright sample had the most seasonally consistent resource base of the non-archaeological samples; at the time of data collection, there were two stores for a 40 house village. The dietary percentages were measured for Wainwright in 1971 and include 25% protein, 43.1% fat and 31.9%

carbohydrates [75]. The Wainwright population diet will thus have been more consistently available through all seasons and have included substantially greater proportions of carbohydrates than the traditional Point Hope diet. The Wainwright individuals also benefited from year-round presence of medical facilities/treatment although they were certainly more sedentary and so will not have benefited from the frame size increases in stature associated with extensive exercise [76]. The five other indigenous populations were living Alaskan natives from whom data were collected by V. Steffanson between 1906 and 1912 [70]. The populations sampled include an area surrounding Wainwright (the Barrow sample) and north of Point Hope, an area immediately inland from Wainwright and inland and north of Point Hope (Nunatagamiut), and three more eastern coastal populations from MacKenzie Valley, Victoria Island and Coronation Gulf ([70,74,77,78]; figure 1). These populations were sampled over a broader area than the Wainwright sample, and while experiencing the same general thermal stress as the Wainwright individuals and benefitting from some food stores at various outposts the individuals were not, as a rule, permanently domiciled at the outposts. They, thus, lived a more traditional lifestyle and subsistence pattern than at Wainwright, although they were all embedded into a market economy at some level, specifically seeking food resources such as flour and other staples and routinely working on whaling ships [70]. Among the samples, four are coastal samples and one (Nunagiamit) is a more inland population that also came in to the coastal areas seasonally for whale hunting and to find work on ships [70]. The inland population diet was likely to have been more seasonal than the more coastal populations and to have included more protein from terrestrial than marine mammals. At the time of Steffanson's data collection, the Barrow region was the most integrated into the market economy and likely integrated individuals from the same area as Wainwright (the outpost of which was established in 1904 [74] just before Steffanson's data collection). The MacKenzie and, especially, the Victoria Island and Coronation Gulf groups to the east were the most isolated at the time of Steffanson's visits. Steffanson and colleagues [70,79] note that Western medical treatment was available at more distant outposts just once a year. Given these parameters, we anticipate that the Wainwright sample should show relatively large size but the least dimorphism of the living samples given its more recent date, more seasonally consistent resource base and greater access to medical services. We expect that the other five populations will be more similar in size to Wainwright than to Point Hope given greater similarities in resource base. However, we expect dimorphism to be greater in these populations than in Wainwright given greater mortality risks, particularly from predation and accident.

To further consider the influence of changing environments, we use data for four populations from Boas' immigrant studies in the early 1900s ([80]; electronic supplementary material, tables S3 and S4). We matched populations for country of ancestral origin and compared between birth countries. The four populations are Central Italians born in Italy and the USA as well as Scottish individuals born in Scotland and in the USA. We chose these populations for their geographically restricted origin of the foreign-born cohorts and because they possessed relatively large US-born cohorts (>40). Boas was not concerned specifically with resource base but with the overall change in environment from foreign to USA. He used 'congested areas' of US cities, principally New York, as his locus of study and included emigres from Europe in the mid- to late 1800s and US born individuals ancestrally from those same countries. Although the immigrant areas of New York City were not particularly bountiful environments, historical data from Scotland suggest that at least the foreign-born Scots likely came from less favourable nutritional and disease conditions; the rapid industrialization of Scotland in the 1800s led by 1840 to extremely negative nutritional conditions across the country such that the 1840s were known as the 'hungry 40s', cities were particularly

Table 2. Samples by taxon (sample size).

skeletal postcranial	skeletal cranial	living somatometric
extant *Homo sapiens*		
US early 1900s	US early 1900s	US immigrants early 1900s
H-T White (*n* = 40)	H-T White (*n* = 38)	Central Italian foreign and USA (*n* = 1147)
Erie County Poorhouse (*n* = 66)	Erie County Poorhouse (*n* = 32)	Scottish foreign and USA (*n* = 176)
archaeological	archaeological	indigenous Alaskans (1968 and 1906–1912)
Point Hope, Alaska (*n* = 47)	Point Hope, Alaska (*n* = 57)	Wainwright (*n* = 80)
		Barrow (*n* = 88)
		Nunatagamiut (*n* = 119)
		MacKenzie (*n* = 87)
		Victoria Island (*n* = 49)
		Coronation Gulf (*n* = 110)
extant Old World monkeys and apes		
Chlorocebus aethiops pygerythrus	*Kenyan vervets*	*Kenyan vervets, four populations:*
Kenyan wild (*n* = 17)	wild (*n* = 22)	Naivasha (*n* = 44)
Kenyan captive (*n* = 17)	captive (*n* = 18)	Samburu (*n* = 62)
		Mosiro (*n* = 14)
		Kimana (*n* = 33)
Macaca mulatta		
Cayo Santiago (*n* = 80)	Cayo Santiago (*n* = 48)	Cayo Santiago (6–14 years; *n* = 196)
Indian wild shot (*n* = 2)	Indian wild shot (*n* = 9)	Indian somatic data (*n* = 23)
Macaca fuscata		
Troop T-1 Takagoyama, (*n* = 61)	Troop T-1 Takagoyama (*n* = 20)	*free-ranging macaques, 12 groups:*
		(listed from north to south)
		Nikko (*n* = 26)
		Shiga (*n* = 71)
		Hagachizaki (*n* = 71)
		Takahama (*n* = 30)
		Arashiyama (*n* = 68)
		Wakasa (*n* = 30)
		Ngatoro (*n* = 68)
		Awajishima (*n* = 29)
		Shodoshimo takamatsu (*n* = 40)
		Takasakiyama (*n* = 302)
		Koshima (*n* = 98)
		M. f. yakui (captive; *n* = 79)

| *Homo erectus*	*Homo neanderthalensis*	other *Homo*
cranial and postcranial		
Ngandong, Indonesia (*n* = 6)	Shanidar, Iraq (*n* = 6)	Atapuerca, Spain (*n* = 27)
Dmanisi, Georgia (*n* = 8)	Krapina, Croatia (*n* = 9)	Dinaledi, South Africa (*n* = 8)
Zhoukoudian, China (*n* = 7)	El Sidrón, Spain (*n* = 3)	
Koobi Fora, Kenya (*n* = 10)	European (*n* = 13)	
Sangiran, Indonesia (*n* = 14)	Near Eastern (*n* = 4)	
Trinil, Indonesia (*n* = 3; postcranial)		
Daka, Ethiopia (*n* = 2; postcranial)		
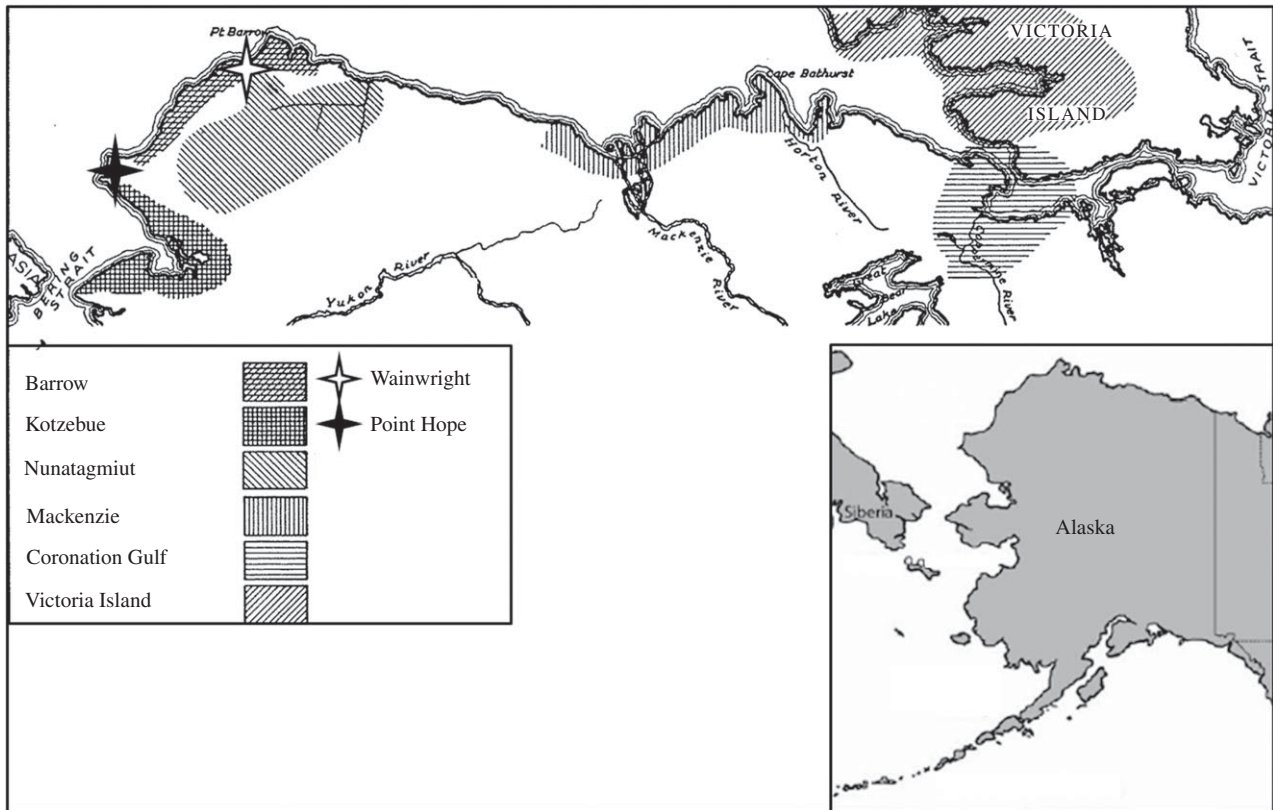
Figure 1. The geographic distribution of Arctic human populations used in this study. Adapted and redrawn from Seltzer and Stefansson [70]. Point Hope archaeological sample from AD 1200 to 1700. Wainwright population measured in 1967–1968. All other populations measured between 1906 and 1912 in the specified regions.

hard hit [81]. We expect then that the direction of change, particularly for the Scottish sample, will be that US-born cohorts of each population will be bigger than their foreign-born cohort and sexual dimorphism will be greater in US-born cohorts due to improved resources.

Finally, we compare between twentieth century skeletal samples of similar ancestry, the Hamann-Todd (H-T) collection 'Whites' and the Erie County, NY Poorhouse Collection (electronic supplementary material, tables S4 and S5). These comparisons control for area of ancestry, climate and general living conditions. While both the Hamann-Todd and the Erie County Poorhouse include individuals of lower socio-economic status [82,83] and presumably resource base, the poorhouses of the turn of the century were particularly nutritionally challenged, and likely carried a higher parasite and disease load [83]. The Erie County Poorhouse operated in Buffalo, New York from 1851 to 1926 as an almshouse, an insane asylum, and a hospital with maternity and consumptive wards with over 171 000 individuals receiving care and over 11 000 deaths occurring at the institution [84]. By the 1850s nearly all of the countries' poorhouses were failing, exposing the poor to conditions that negatively affected their occupants' health [83]. In general, these institutions were marked by poor construction, poor ventilation, poor heating, overcrowding, inadequate medical care and poor-quality food [85]. Food sources of both the institutionalized and non-institutionalized poor who may have sought care at the facility were generally considered to be low cost, high carbohydrate and highly cariogenic foods [86]. In 1856, the New York State Senate issued a statement, indicating that good health within the poorhouses was an 'impossibility' [83]. The H-T collection individuals were largely poor, working-class labourers, who were mostly first-generation US citizens from the Northern States [82]. The most frequent causes of death among these individuals were diseases of poverty and exposure (e.g. [82]). While it is not necessarily clear which group encountered worse conditions, and we expect the two groups to be largely

similar, we anticipate that the Erie sample will be somewhat smaller than the H-T and show moderate dimorphism. To compare the relationship between skeletal and somatic signals, we compare skeletal and somatometric data for the H-T medical collection 'White' sample.

Question 2 focuses on plasticity in non-human primates, and as with the human sample, we emphasize localized sub-populations of broadly dispersed species, in this case *Macaca* and *Chlorocebus*. We consider this to be a conservative approach based on the logic that broadly dispersed genera may be expected to have inherently higher levels of variability and would thus be less likely to yield a signal of remarkability compared with either humans or *H. erectus*. Indeed, both macaques and vervets are so adaptable as to often be classified as 'pest species' [87]. Future work is needed to address patterns in less widely dispersed genera.

To closely assess the intra-specific variation across living populations, we use raw somatometric data from four populations of Kenyan vervets (*Chlorocebus aethiops pygerythrus*) whose environments differ in terms of elevation, rainfall and resource availability ([88]; figure 2*a*; table 4 and electronic supplementary material, tables S6–S8). The vervets sample a species that is found in both East and South Africa, but we focus on a restricted part of its range in Kenya. The animals include groups from two highland (Naivasha and Mosiro) and two lowland (Kimana and Samburu) sites; the highland sites are colder and wetter than the lowland sites. Each pair of sites also differ in resource base with cultivated crops, which vervets are known to raid, available at Naivasha and Kimana. Thus, between sites we expect overall larger body size should ensue for Naivasha (the highest elevation) both because of climatic adaptation and because of greater nutritional sufficiency due to the greater availability of cultivated crops. The Kimana animals are expected to be somewhat larger than their lowland counterparts. We also expect greater dimorphism in both of these populations compared with their environmental counterparts due to relatively greater nutritional sufficiency.

Table 3. Sexual dimorphism values (male/female) and means by sex for Alaskan samples. Mean values that *do not* significantly differ between the living populations and the Point Hope sample (at the 0.05 level with Benjamini–Yekutieli adjustment for multiple comparisons) are indicated by '*ns*'. Significant differences within populations between the sexes are in bold italics. Stature is in cm, and all other dimensions are in mm.

	Point Hope	Wainwright	Barrow	Nunatagamiut
date	AD 1200–1700	1968–1969	1906–1912	1906–1912
resource base	precontact traditional coastal high-protein, low carbohydrate	town resident, constant higher carbohydrate, lower protein	postcontact modified traditional coastal >carbohydrate	postcontact modified traditional inland hunter
stature				
sex dimorphism	NA	1.067	1.071	1.091
male mean		*166.3*	*164.6*	*169.1*
female mean		*155.8*	*153.7*	*155.1*
bi-iliac breadth				
sex dimorphism	1.023	1.005	NA	NA
male mean	273.2	296.7		
female mean	266.9	295.2		
tibial length				
sex dimorphism	1.103	1.062	NA	NA
male mean	*359.1*	*431.9*		
female mean	*325.4*	*406.6*		
upper face ht (N-Pr)				
sex dimorphism	1.11	1.045	NA	NA
male mean	*73.3*	76.9		
female mean	*66.0*	73.6		
head length (GOL)				
sex dimorphism	1.044	1.023	1.041	1.039
male mean	*186.5*	*190.1*	*192.4*	*190.0 ns*
female mean	*178.6*	*185.8*	*184.88*	*182.7*
head breadth (Eu-EU)				
sex dimorphism	1.063	1.033	1.045	1.040
male mean	137.6	*154.6*	*151.1*	*154.5*
female mean	129.5	*149.6 ns*	*144.5 ns*	*148.5 ns*
bizygomatic breadth (Zy-zy)				
sex dimorphism	NA	1.039	1.056	1.054
male mean		*148.4*	*151.4*	*145.7*
female mean		*142.9*	*143.3*	*138.2*

To assess the relationship between skeletal and somatometric signals, we compare these data with skeletal data from Kenyan vervets housed in the National Museum of Natural History, USA, and we compare across wild and captive skeletal samples.

To assess a more temperate climatic scope, we use somatometric data for 1 captive and 12 wild populations of Japanese macaques (*M. fuscata*) distributed across the geographic range of Japan ([89]; electronic supplementary material, tables S9 and S10; figure 2*a*). These comparisons allow for a limited variation in climate and resource base while controlling for predator load. These are compared amongst themselves and with raw skeletal data from a single troop of wild macaques from Chiba Prefecture (Takagoyama Troop T-1; [90–92]). The populations vary in terms of mean annual temperature, with the more northern populations experiencing the lowest temperatures. Populations also differ in

terms of rainfall and experience different levels of seasonality (as based on the number of dry versus wet months) and live in different types of forests that are known to differ in energy availability of diets [93]. Deciduous forests have lower secondary productivity and thus are less energy dense seasonally. Environmental details and mean values are presented in the electronic supplementary material, but in brief, the Shiga group (Nagano Prefecture), lives in a deciduous forest, experiences the most seasonality in food sources, and the deepest snow and thus a reduced resource base, suggesting that we should see greater differentiation in size and the least dimorphism compared even to their climate-matched neighbours (Nikko) living under similar thermal stress. Of the Japanese macaques, the closest climate match for our skeletal sample is the Wakasa group. The Japanese macaques have no natural predators.
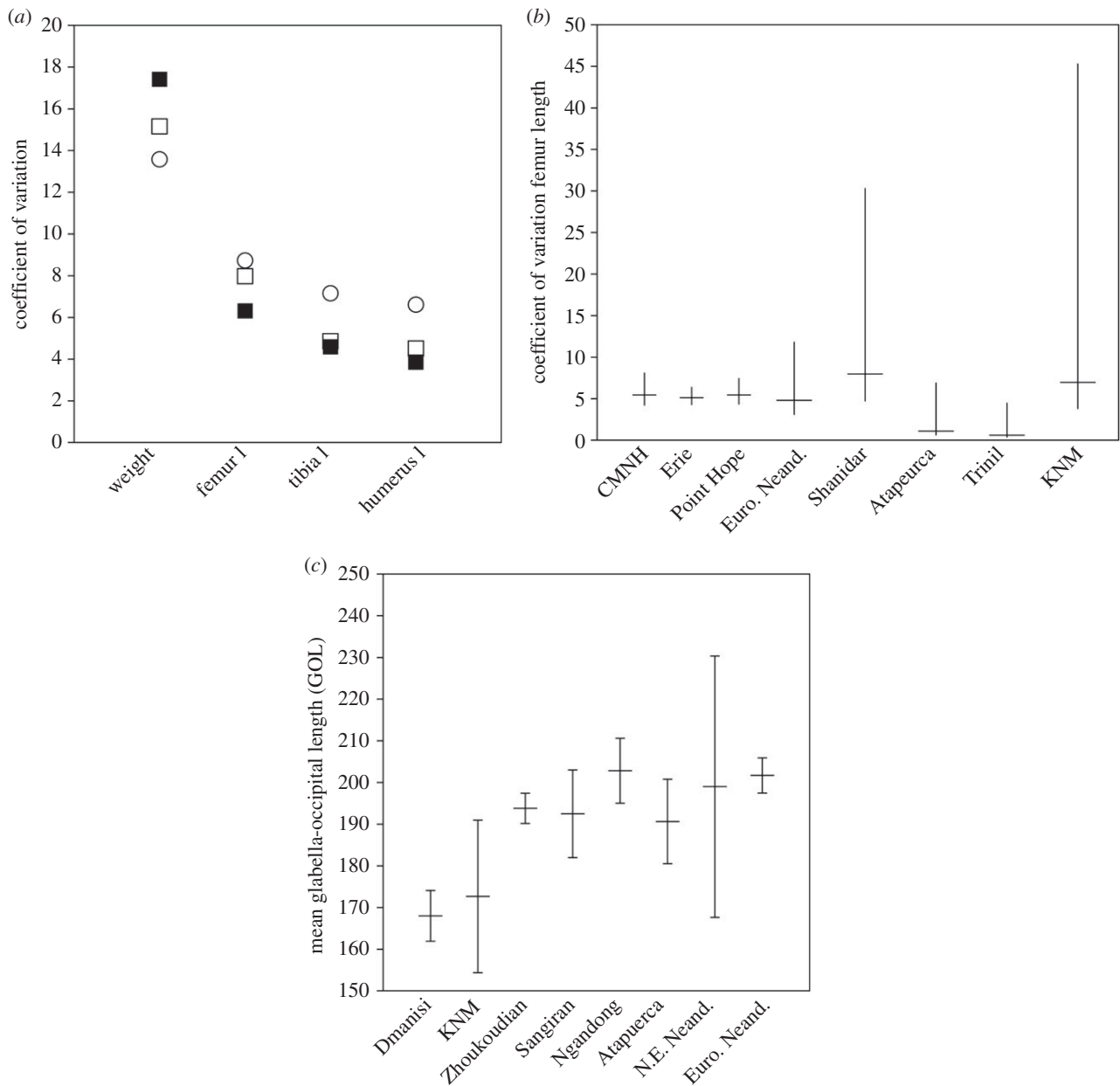
Figure 2. Comparative CV values for (*a*) postcranial somatometrics for male, Kenyan vervets from Naivasha (open circle), *M. mulatta* from Cayo Santiago (filled square) and *M. fuscata* from Wakasa (open square) and (*b*) femur length for fossil paleodemes and postcranial human samples including twentieth-century skeletal samples from the USA (CMNH, Erie) and an archaeological Arctic sample from Point Hope; 95% confidence intervals included. Note that the relative fossil CV values follow the prediction that shortest interval paleodemes (Atapeurca/Trinil) will have the lowest values. However, differences are not significant and both values seem artificially low compared to extant values. (*c*) Comparative mean values across fossil paleodemes for cranial length (GOL) with 95% confidence intervals.

We use raw data from the rhesus macaques (*M. mulatta*) from Cayo Santiago, Puerto Rico to consider patterns of variation between somatometric data collected in the early 1980s and skeletal metric data from individuals who died during that time and between the Cayo animals and Indian *M. mulatta* ([89,90,94]; figure 2*a*; electronic supplementary material, table S11). These data allow for comparison between groups of different predator load and nutritional sufficiency. The founding individuals of the Cayo Santiago colony were *M. mulatta* collected in Northern India in the 1930s whose descendants continue as a free-ranging but provisioned population on Cayo Santiago Island, Puerto Rico [95]. During the early years of the colony the animals were undernourished, but the cohort included here are animals who died in the 1970s and 1980s who had been reliably provisioned with a high protein (24–26%) monkey chow diet in addition to freely foraging on tropical plants for many generations and thus can be considered to have experienced high nutritional sufficiency

[96]. At the same time, the parasite load in the Cayo animals, while apparently asymptomatic, is reported to be as high as laboratory animals who would be symptomatic and treated for parasites [97,97]. Their predator load, however, is very low as there are no natural predators on the island. Additionally, their age at first reproduction (AFR), as would be expected from both high nutritional sufficiency and moderate extrinsic mortality, does not seem delayed (AFR = 4.27; [98]) compared to other populations (AFR = 3, 4.19, 4.5 from three populations reported in [99]). Given their region of origin, we compare a limited number of Indian-derived, wild-shot *M. mulatta* skeletons with these Cayo Santiago skeletal data. We also compare Cayo Santiago somatic data with similar data from living Indian macaques. We expect Cayo animals will be larger due to resource sufficiency and low predation, but that their size might be somewhat negatively affected by parasite load. We therefore expect a moderate amount of sexual dimorphism.

Table 4. Sexual dimorphism values (male/female) and means by sex for adult vervet samples for six of 12 variables used in the study. Bold italic values differ significantly between males and females of the same sample. Non-significant values for Mosiro may be due to small sample sizes. Weight is in kg, and all other dimensions are in mm.

	Naivasha	Mosiro	Kimana	Samburu
elevation (m)	highland	highland	lowland	lowland
	>2000	1000–1500	750–1000	600–650
rainfall (mm)	>600	400–600	200–400	350–400
vegetation	lakeside grassland	heavily grazed woodland and savannah	dry grassland and thorn scrub	dry grassland and thorn scrub
resource base	cultivated crops abundant	no cultivated crops	some cultivated crops available	no cultivated crops
weight				
sex dimorphism	1.48	1.5	1.34	1.58
male mean	*4.9*	4.2	*4.3*	*4.1*
female mean	*3.3*	2.8	*3.2*	*2.6*
body length				
sex dimorphism	1.13	1.17	1.10	1.15
male mean	*43.5*	42	*42.6*	*39.8*
female mean	*38.6*	36	*38.7*	*34.5*
chest girth				
sex dimorphism	1.15	1.16	1.12	1.18
male mean	*35.7*	33.9	*32.7*	*32.7*
female mean	*31.1*	29.1	*29.1*	*27.8*
upper arm length				
sex dimorphism	1.11	1.2	1.14	1.18
male mean	*14.5*	*15*	*13.8*	*14.6*
female mean	*13.1*	*12.5*	*12.1*	*12.4*
leg (thigh) length				
sex dimorphism	1.15	1.17	1.16	1.21
male mean	*16.9*	*16.4*	*15.7*	*16.5*
female mean	*14.7*	*14*	*13.5*	*13.6*
tail length				
sex dimorphism	1.17	1.2	1.07	1.21
male mean	*61.8*	64.8	58	*67*
female mean	*52.6*	54.1	54.2	*55.2*

(b) Fossil samples (Question 3)

To approximate the close population approach taken above, for Question 3 we apply a narrower lens to the fossil record than previous studies and attempt to construct and assess paleodemes (electronic supplementary material, tables S12–S15). Because we are interested in variation, paleodemes must contain at least two individuals in a somewhat circumscribed temporal span and geographic region. To consider the intraspecific variation across populations we require at least two, and preferably three, paleodemes per fossil taxon. These criteria severely limit our choices of fossil comparators. We consider paleodemes of *H. erectus* and, to provide a comparative framework, we also look at Neandertal and other Middle Pleistocene samples. Only Neandertals have a sufficient number of paleodemes for intraspecific comparisons and small sample sizes and preservation preclude us from considering the variation in paleodemes of earliest *Homo* at all. Each paleodeme was constructed to sample as little time as possible, although apart from Ngandong and Sima de los Huesos localities,

most are more time transgressive than the extant samples. Temporal duration is unequal across samples. The fossil samples are by necessity mixed-sex and we thus limit ourselves to sampling from just genus *Homo* among the hominins.

Homo erectus as defined here ranges in age from about 1.9 Ma to at least 250 000 years ago or younger and from Africa to Southeast Asia (figure 2b,c; electronic supplementary material, tables S12 and S14). Within this broad swath, we assess two samples that are thought to be of short or even single depositional events: Dmanisi, Georgia and Ngandong, Indonesia. The Dmanisi assemblage is dated to about 1.77 Ma and has been argued to have been deposited over no more than about 10 000 years [42]. The Ngandong assemblage, although of contested absolute age (possibly ~550 000 or 50 000 years old), is argued on all accounts to be a catastrophic depositional assemblage [100,101]. To expand the comparison, we include five potentially more time transgressive units: Zhoukoudian, China; Sangiran and Trinil, Indonesia; Daka, Ethiopia; and Koobi Fora, Kenya. We limit the

Zhoukoudian assemblage to just those hominins from layers 7 to 11 (about 750 000 years old, with a youngest age of 670 000 years; [102]), which may be comparable to the age distribution of the Neandertal cohorts (see below). The Zhoukoudian femoral specimens are restricted to layers 8–9 and should, therefore, be more time constrained. The Koobi Fora *H. erectus* specimens cover about 130 000 years of time ranging in age from about 1.5 to 1.63 Ma (for KNM-ER 3733; [103]). The Sangiran specimens are more time transgressive and include about a half million years of time from about 1.1 to 1.6 Ma [104]. The Trinil and Daka femoral assemblages (∼900 000 and 960 000 years old, respectively) are each of unknown depositional span; while they could be tightly time constrained, strictly speaking, the time between specimens is unknown and there is no close capping age in either context (see [104,105]). It should be noted that the requirement to constrain time and locality and maximize individuals per sample excludes some notable cranial fossils from the analyses. The exclusions include sites with single crania (e.g. Olorgesalie, Daka and Trinil), as well as those with too great a time span between individuals (e.g. Olduvai Gorge where OH 9 and OH 12 are separated by more than a half million years) or a time span that would be greatly expanded by including the entire assemblage (e.g. we exclude the upper Bapang Formation specimens from Sangiran and specimens from level 6 and higher at Zhoukoudian). The total *H. erectus* sample includes 50 fossil specimens.

Of these paleodemes, the Dmanisi sample provides the best evidence of extrinsic mortality and nutritional sufficiency signals. The Dmanisi site and hominin remains show evidence of carnivore activity and accumulation [106], suggesting predator load (and extrinsic mortality) was high. A few individuals also show signs of generalized stress as indicated by enamel hypoplasia [107], potentially also an indicator of issues related to high extrinsic mortality. The temperate and seasonal climatic zone of the site has also been used to infer less resource abundance or at least a more challenging resource environment compared to that of East Africa [106]. The Zhoukoudian remains are latitudinally similarly placed and should have the same seasonality and resource signal as Dmanisi. However, their direct evidence of predator load is not as great as at Dmanisi. Thus we would predict smaller size (based on resources and extrinsic mortality due to disease) in Dmanisi and moderate dimorphism.

To provide a comparative context we sample Neandertals, Sima de los Huesos and *H. naledi* (figure 2c; electronic supplementary material, tables S13 and S15). Each of these species is geographically more restricted than *H. erectus*. For *H. neanderthalensis,* we compare five samples: three from single localities (El Sidrón, Spain, Shanidar, Iraq, and Krapina, Croatia); and two from broader regions (Europe (40–80 ka) and the Near East (50–130 ka); [108–110]). While the Shanidar remains are more geographically restricted than the European sample, they span many, poorly dated metres of section that likely comprise some 30–60 000 years of time [110]. Owing to small numbers of individuals and few overlapping variables among them, the Shanidar sample is not subdivided further. The temporal range for Shanidar is thus potentially comparable to that of the European sample. Owing to fragmentation, the Krapina remains offer a more limited set of mandibular comparisons from about 130 000 years ago. Our El Sidrón sample is also limited to mandibular data and offers a more time-restricted comparison of likely about 46 000 years ago and perhaps of a single depositional event [111]. The Near Eastern Neandertals are fewer in number and cover a greater span of time (∼60 000 years) than the European individuals. We include data from the Sima de los Huesos assemblage which has been argued to be of short duration and perhaps even a catastrophic event [112]. Given their close genetic relationship to later Neandertals [113] as well as their restricted geographic and abbreviated temporal range, we believe these to be a relevant comparison. Finally, we consider a large number of *H. naledi* individuals (n = 8) whose geological

age and duration of deposition are unknown [114]. Although the published raw data are limited to sub-trochanteric femoral dimensions, these dimensions are also available for a number of other fossil paleodemes and provide a means of estimating body weight in these samples as well. As we are interested in the variation in body size, their inclusion seems warranted.

(c) Methods

Because we are interested in patterns of variation within and among taxa, rather than differentiating between taxa, we explicitly avoid using variables that show relatively little inherent variation and purposefully include those with greater variation (we would do the opposite if we were interested in circumscribing species groups). Dental variation, especially of the M1, is low relative to other variables [115], presumably due to high heritability, and we therefore exclude these variables. We include variables from both high stress (facial and joint surfaces) and low stress (neurocranial and other postcranial) environments when possible, as the former may be more variable than the latter [115]. We are further limited by the kinds of data typically collected in extant studies, and we attempt to proxy these data from skeletal samples by matching somatometric individual bone elements (e.g. thigh length to femur length; head length to glabello-occipital length (GOL)). We are similarly limited by the fragmentary nature of the fossil record and so include variables that have the highest representation in *H. erectus* samples (e.g. GOL, cranial capacity and mandibular heights and breadths). While stature and body weight estimates are seemingly easily made from skeletal remains, they require a suite of inferences about body proportions and scaling that because of compounding error estimates may obscure variation between groups. For this reason, we focus more on direct skeletal measures using directly measured stature or weight for comparisons among living groups.

The result of high developmental plasticity is variation in size *across* populations and so we compare adult means for each measure across samples as described above. However, because developmental plasticity works on a species' inherent ability to vary (i.e. its variability) there is also the question of whether a different measure *within* a single population might also be a useful proxy for assessing variability of a species. That is, should we expect species with a greater ability to vary to also show more variation within each population? If so, should each population of the species show higher CVs for any given measure than do less variable species? As variation is size dependent, we use the CV to compare the variation adjusted for relative size across measures, and we adjust CVs for small sample size as necessary [116,117]. Arguably, the *absolute* size variation *across demes* is itself more important for understanding variability within a taxon than is CV. However, it would be ideal were CV to prove to be a good proxy for variability as that would allow us to measure species variability from just one paleodeme. Such an event would greatly increase the number of taxa available for inclusion.

For the purposes of this initial evaluation we use univariate comparisons. The primary reason for this approach is the desire to contextualize the variation across *H. erectus* populations with other fossil samples and the resulting dearth of overlapping variables across all potential samples. Multivariate rank-order processes are not possible given the lack of overlapping variables across all samples. However, these may prove useful in future for more restricted samples. Additionally, while we have individual raw data measurements for our main comparative samples (vervets, *M. mulatta*, *M. fuscata* skeletal sample, Boas, H-T, Erie and Point Hope *H. sapiens* samples and the fossil samples), in order to expand potential comparisons we also use summary data from the literature for some living groups (*M. fuscata* somatometrics, Arctic *H. sapiens*). These latter studies provide only means

and standard deviations, and so Fligner-Killen (*sensu* [118,119]) or resampling methods (*sensu* [120]) are not feasible for these datasets. We therefore compare between means and CVs following Lewontin [121]. To protect our alpha values due to multiple comparisons we use the Benjamini and Yekutieli method [122] for mean values to avoid false negatives in the more conservative Bonferroni correction. Analyses were run with Bonferroni as a comparison and very few differences resulted between the two. While recognizing the inherent limits of these tests (see [118] especially for the CV ratio test), they offer an initial assessment of the question given our very small fossil sample sizes.

4. Results

(a) Question 1

Within each human group from the Arctic, males and females differ significantly on nearly every variable except bi-iliac breadth; however, they differ much less across populations (table 3 and electronic supplementary material, tables S1 and S2). None of the populations differ significantly in CV for any of the variables. The five living Alaskan populations measured between 1906 and 1912 [70] never differ from one another in stature. Sex-matched samples between populations most frequently differ in cranial breadth. The Wainwright population measured in the 1960s differs slightly more from its closest counterparts, including in head breadth and total face height from the female Barrow sample and bizygomatic breadth from the male Barrow sample; however, it also does not differ in stature from the other Arctic populations. In sex-matched comparisons across populations, male and female comparisons differ with approximately equal frequency (i.e. in most comparisons males and females differ between populations on the same number of variables, and when these comparisons are unequal about half the time males show more between-group differences and about half the time females do; see electronic supplementary material, table 2). Sexual dimorphism values for the five living Arctic groups mostly vary between 1.04 and 1.09. The Wainwright sample is always lower than or equal to the lowest values for sexual dimorphism—it is the lowest value for stature, bi-iliac breadth, tibial length, upper face height, head length and head breadth. The Coronation Gulf and Victoria Island samples, however, have lower total face height dimorphism. And while the general rule is for males and females to differ with the same frequency across populations, in the Wainright comparisons females differ with other populations twice as often as do males.

In overall size, the Point Hope Tigara sample is smaller in all mean values than all the living populations, but sexual dimorphism is usually slightly greater than the five early 1900s populations and much greater than the 1960s Wainwright sample (table 3 and electronic supplementary material, table S1). As with the living samples, CVs do not differ significantly. Postcranially, the Point Hope sample differs significantly from the Wainwright population in both absolute tibial length and bi-iliac breadth in sex-matched comparisons (the 1900s samples do not provide postcranial metrics). Cranially, the Point Hope female sample differs in greatest head length from females of all three nearest extant populations (Wainwright, Barrow and Nunagatmiut) and the males differ in breadth and length from Barrow and Wainwright and in breadth from Nunagatmiut. These absolute differences in mean values, especially in the postcranial comparisons, in part may reflect differences

between fleshed and unfleshed individuals; however, they also result in differences in sexual dimorphism values. Point Hope has equal or slightly greater dimorphism than all extant samples, and especially than the Wainwright sample, and these values should be unaffected by being taken on fleshed as opposed to unfleshed individuals (see data for H-T in electronic supplementary material for support). Greater sexual dimorphism and smaller overall size in Point Hope compared to the other samples is predicted based on greater resource sufficiency and lower extrinsic mortality due to predation and accident especially in the Wainwright sample. Point Hope females and males differ across groups with equal frequency.

As with the Alaskan groups, the foreign-born and US-born cohorts from Italy and Scotland measured by Boas show significant differences between male and female mean values for all of the available dimensions (mostly cranial, but also including stature; electronic supplementary material, table S3). As is the case in other samples, none of the samples show significant differences in CV values within or between sex-matched pairs across populations. As predicted on the basis of increased nutritional sufficiency, three of the four male and female US-born cohorts show greater mean values for stature than their sex-matched, foreign-born cohorts (US-born Scottish males show the opposite). However, except in the case of the Scottish females these differences are not significant. As a result, sexual dimorphism values differ between the Scottish US-born (1.03) and foreign-born (1.08) cohorts due to relative increase in female size in the former. Like the Arctic samples, head breadth dimorphism is the most variable across populations; and male and female comparisons differ with approximately equal frequency across groups (electronic supplementary material, table S4).

The twentieth-century skeletal samples yield similar results as the other humans, although within groups, males and females do not differ significantly from one another quite as often as in other human groups (electronic supplementary material, table S5). In both H-T and Erie samples, males and females differ from one another particularly in measures related to leg length. Beyond this, the H-T males and females differ more from one another on cranial variables than do the Erie males from females. Erie females and males differ on two of six cranial measures and H-T on five of seven measures. Sexual dimorphism values are very consistent between the two skeletal samples (and with most Arctic samples). Contrary to expectations, sex-matched comparisons across the two collections do not differ significantly for any measure and the Erie sample is absolutely larger than the H-T in all postcranial dimensions, although not significantly so.

(b) Question 2

Like humans, female and male vervets virtually never differ statistically on CV values within populations but almost always differ significantly in mean values (table 4 and electronic supplementary material, table S6). The exception to this are the Mosiro vervets whose mean values by sex sometimes do not differ significantly, likely due to small sample sizes and adjustments for multiple comparisons. Unlike humans, more significant differences accrue between females from different populations than between males; about two-thirds of differences are related to female–female differences and one-third to male–male differences (electronic supplementary material,

table S7). Additionally, sex-matched comparisons never differ on the same number of variables for males and females. As predicted based on relative nutritional sufficiency, most size measures are larger in the two populations with the greatest resource sufficiency (Naivasha and Kimana) than in their elevation-matched counterparts (Mosiro and Samburu, respectively). However, these differences do not always rise to the level of significance. The two populations with the greatest resource sufficiency (Naivasha and Kimana) have the smallest sexual dimorphism values for all variables across all populations (see table 4), contrary to expectations.

Benefitting from access to raw data for these populations, we also ran these analyses for mixed-sex samples. Mixed-sex samples tended to show *fewer* differences across populations than sex-matched comparisons, accurately revealing the differences among populations in just three of the six instances (see electronic supplementary material, table S8). While CVs were reliably elevated in the mixed-sex samples, the comparison of CVs between populations only twice identified differences between populations that had been identified through sex-matched comparisons of mean values.

The Japanese macaques come from a larger absolute area than do the Kenyan vervets examined here but from a species of more restricted total range; they show similar results as the vervets (electronic supplementary material, tables S9 and S10). Like the vervets, within-group females and males virtually never differ statistically on CV values but almost always differ significantly in mean values. The same pattern is seen in the skeletal group. As was the case with the vervets, female Japanese macaques show more significant differences across populations than do males; females differ about as often as males do, but more of these between-population comparisons show equal numbers of male and female differences than was the case in the vervets in which all comparisons showed either more male or more female differences (electronic supplementary material, tables S7 and S10). Differences between sex-matched populations are largely dictated by geographic distance, with more distant groups more strongly differentiated from one another and more significant differences occurring between female pairs. Beyond this, however, the Shiga group, which is far north (near Nagano) and experiences the most extreme seasonality, deepest snowfall and thus resource stress, also differs the most with respect to its nearby neighbours of similar thermal stress (Nikko) in terms of both the largest body size and least sexual dimorphism. Mixed-sex samples are not assessable from the published data.

The Cayo Santiago samples differ from the Indian samples and yield similar patterns of sexual dimorphism between skeletal and somatometric data (electronic supplementary material, table S11). Body weights and limb lengths for Cayo Santiago monkeys are substantially bigger than Indian somatic data. Owing to small samples sizes for Indian skeletal samples, only mixed-sex comparisons could be made. As predicted, in all cases the mixed-sex Cayo skeletal measures for humerus, femur and tibia lengths are greater than those of the Indian sample. However, these differences are not significant. The Cayo Santiago skeletal measures and measures from fleshed Cayo animals yield identical sexual dimorphism values, suggesting that skeletal and somatometric data can be directly compared. Indian somatic data are less dimorphic for tibial length (due to male size decrease) and more dimorphic for weight (due to female size decrease).

(c) Question 3

Within our paleodemes of *H. erectus*, only cranial variables differed significantly (electronic supplementary material, tables S12–S14). Five paleodemes (Dmanisi, Ngandong, Zhoukoudian, Sangiran, Koobi Fora) provided cranial data. Four *H. erectus* paleodemes (Daka, Koobi Fora, Trinil, Zhoukoudian) provided postcranial data. All but Sangiran differed in mean values of at least one cranial variable from at least one other paleodeme. The Dmanisi paleodeme is significantly smaller than both the Ngandong and Zhoukoudian paleodemes for cranial capacity and GOL (electronic supplementary material, tables S12 and S14). The Ngandong paleodeme is significantly larger than the Zhoukoudian paleodeme for biauricular breadth and than the Koobi Fora paleodeme for cranial capacity. The Sangiran paleodeme, which is also the most time transgressive of the paleodemes, does not differ from any of the other paleodemes, despite its similar or larger size ($n = 14$ for cranial variables) than other paleodemes that do yield significant differences. None of the postcranial variables differed across the *H. erectus* paleodemes. CV values do not differ across paleodemes, only approaching significance for femoral length between Trinil and the Zhoukoudian and Koobi Fora paleodemes. These results are due to seemingly unnaturally low levels of variation in Trinil (CV of 2.0 for femoral length as compared to CVs in recent humans of 5–6 for H-T White femoral lengths and 9–11 in other *H. erectus*; figure 2*b*).

Within Neandertals fewer differences are significant, but interspecifically more differences emerge across paleodemes (electronic supplementary material, tables S13 and S15). Neandertals show no significant differences (of CVs or mean values) among their five cranial and three postcranial paleodemes or with the Sima de los Huesos sample. The European Neandertal paleodeme differs from each of the *H. erectus* paleodemes with cranial remains (Dmanisi, Koobi Fora, Sangiran, Zhoukoudian, Ngandong) on cranial capacity, from the Koobi Fora and Sangiran paleodemes on mandibular corpus breadth at M1, and from Dmanisi on GOL and basion-bregma height (electronic supplementary material, table S15). The *H. erectus* Dmanisi paleodeme differs from the Neandertal Shanidar paleodeme on biorbital breadth and from Sima de los Huesos on cranial capacity. Postcranially, nine paleodemes are represented (*H. erectus*: Daka, Koobi Fora, Trinil, Zhoukoudian; *Neandertal:* European, Near Eastern, Shanidar; other *Homo:* Atapuerca, Dinaledi). Only *H. naledi* differs consistently from other paleodemes. *H. naledi* subtrochanteric dimensions differ from one of the three Neandertal postcranial paleodemes (European Neandertals) and from all three of the *H. erectus* postcranial paleodemes (Daka, Koobi Fora and Zhoukoudian) that preserve these dimensions.

5. Discussion

(a) Humans and broadly dispersed primates show similar patterns of variation between environments

Despite fairly substantial differences in sexual dimorphism between humans, on the one hand, and macaques and vervets, on the other hand (e.g. ~1.05 in humans and ~1.15 in non-human primates, tables 3 and 4), within populations, mean values between the sexes nearly always differ significantly in both the human and non-human primates we examined.

Mixed-sex CVs are reliably higher than those of single-sex samples (see also electronic supplementary material, table S8), and may, therefore, be useful indicators of relative dimorphism in fossil samples *of equal age duration*. These findings are consistent with the practice of using mixed-sex CV as a dimorphism proxy [64]. However, CV values do not differ reliably between the sexes in any of our samples.

Between populations, sex-matched comparisons show some differences (of mean values and sex dimorphism ratios) as predicted by environmental conditions of resource base and/or extrinsic mortality. While controlling for thermal stress, nutritional and lifestyle shifts in the Arctic human groups suggest that the Wainwright sample is large overall but that sexual dimorphism values are smaller than all other groups. Boas' Italian and Scottish populations show expected, but (mostly) non-significant increases in size in US-born cohorts and decreases or no change in sexual dimorphism. The non-human primates show more robust responses to environmental change with vervets and macaques showing increases in body size with increasing resource sufficiency, also often accompanied, however, by reduced sexual dimorphism. However, the CV cannot reproduce these findings and is thus not an adequate means of assessing variability of a species. Even when absolute mean values vary substantially, the CV does not vary significantly between the sexes, between groups within species, or across species for a particular variable. Although CV values do, of course, differ from variable to variable, they differ in similar ways for particular variables across species (figure 2*a*). Humans follow this pattern as much as any other primate considered here.

The pattern of which sex more frequently differs between populations, however, varies between human and non-human primates. In humans, males and females tend to differ by an equal number of variables between any pair of sample populations. In the few instances of uneven sex differences, human females are as likely as males to show more differences between populations. In contrast, in non-human primates two-thirds or more of population comparisons show unequal differences between males and females, and females are twice as likely as males to show more differences between populations. As a result, differences in sexual dimorphism between non-human primate populations are most often due to changes in female size rather than male size, whereas those in humans are as likely to be due to male as to female changes in size. These differences likely relate to competing factors acting on male and female body size in human and non-human primates. In multi-male groups with competition for mates, maintaining overall body size may be a limiting selective factor for males [123] and cerebral growth in infants should not be as strong a selective force favouring larger female body size as it is in humans [124]. Thus, female bodies may be afforded greater latitude to vary with resource stress in non-human primates—resulting in relatively smaller females size and larger dimorphism values and generally greater ability for females to vary overall. That said, in the extant human comparisons that most closely followed the environmental predictions for size and dimorphism (Wainwright versus other Alaskans and Scottish US- versus foreign-born) it was also female size differences that resulted in changes in dimorphism. Thus, while it may be harder to change dimorphism values in humans due to the similarity in patterns of change between males and females, female change is responsible for differences in dimorphism at least some of the time. We therefore conclude

that despite having very plastic life history strategies across populations [1], extant humans are not unique compared to the metric variation seen in these widely dispersed non-human primates across environments. Additionally, female size tends to differentially influence dimorphism values, particularly in non-human primates.

Our data also suggest that skeletal variables track somatometric variables and yield similar inferences about variability. The H-T White sample, Cayo Santiago and Japanese macaque data and Kenyan vervet data suggest that closely matched skeletal and somatic data are likely to yield absolutely different mean values, due to measuring skeletal versus fleshed individuals. However, these skeletal and somatometric data, nonetheless, yield similar sexual dimorphism signals. Importantly, this finding implies that comparisons of sexual dimorphism values are interpretable across somatic and skeletal samples *if* the variables are matched to measure similar anatomical regions (e.g. thigh length and femoral length). The ability to make such comparisons across skeletal and somatic samples is critical for there to be any hope of interpreting fossil skeletal signals in the context of living primates.

(b) Variation and variability in paleodemes

Paleodemes of *Homo erectus* differ more across populations of the species than do Neandertal paleodemes; indeed, Neandertal paleodemes do not differ at all within the species. Because our mixed-sex extant samples tended to show *fewer* differences across extant populations than do sex-matched comparisons, comparisons of mixed-sex fossil paleodemes should be conservative in the numbers of differences they yield. We thus anticipate that the differences across *H. erectus* paleodemes would be more numerous if we were able to assess sex-specific differences (not to mention if we were able to sample tighter time frames with more individuals represented by more complete skeletons).

Our results suggest that previous disagreements about the degree of variation in *H. erectus* may be influenced by the variables and groups considered and methods employed. Cranial length (GOL) differs the most among paleodemes of *H. erectus* (figure 2*c*), and its signal does not appear to be as influenced by the duration encompassed by the paleodeme as some other variables. On the other hand, *Pan* data show less absolute variation in this dimension (data not reported here). This result is consistent, then, with Spoor *et al.* [38] finding that *H. erectus* with the introduction of the KNM-ER 42700 calvaria showed more variation than *Pan.* However, whether this indicates increased dimorphism in *H. erectus* as they argued or increased variability/plasticity, as we might be inclined to argue, is unknowable at this time. This interpretation is likely also confounded by the fact that GOL is affected by the degree of development of both the supraorbital and occipital tori, species-specific features of *H. erectus* whose expression may differ across the sexes [60]. Despite our data for GOL variation across paleodemes, we concur with Plavcan [64] that little variation in CV is apparent between human and non-human data generally, and this seems to hold when comparing across more localized populations of extant primates as we did here, as well as across fossil paleodemes. As argued above based on our current data, we are unconvinced that this lack of difference in CVs is informative with respect to variability. But it may argue for lower dimorphism in *H. erectus,* as has been earlier argued by McHenry and others [125,126].

The most consistent postcranial differences across paleodemes are between the *H. naledi* femoral specimens, and weight as calculated by the *H. naledi* formula and other weight calculations. In all paleodemes, the *H. naledi* formula seems to produce larger weight estimates (by 10 kg or more) than other methods. The weights this formula produces for *H. naledi* itself seem inconsistent with the size of their sub-trochanteric regions, which are a bit smaller than *Australopithecus afarensis* femora (*A. afarensis* mean AP and ML sub-trochanteric dimensions are 20.2 and 28.4, respectively; *H. naledi* are 19.1 and 26.3, respectively). Despite these small subtrochanteric dimensions, the *H. naledi* weight estimates are within the size range of *H. erectus* and small-bodied modern human populations with substantially larger femoral dimensions. This may be the result of shape differences between more circular human femora (from which the weight formula were created) and the *H. naledi* sub-trochanteric regions that differ greatly in their ML versus AP dimensions. Regardless, the results highlight a conundrum regarding size in that taxon as well as the more general difficulties of estimating size in the fossil record [59].

As noted above, based on the current data, we suggest that CVs do not reflect differences in the ability of a species to vary and also rarely vary between human or non-human primates. In the fossil sample, more differences in CV are present and these differences seem most clearly explicable by the temporal duration of the paleodeme. As might be expected, paleodemes that sample shorter time spans, such as Ngandong, have lower CV values for most variables. Those sampling larger time frames, such as Sangiran and Koobi Fora, have larger CV values. Similarly, the Ngandong paleodeme also varies less on certain nonmetric cranial characters such as position and relationship of the mastoid and supramastoid crests than do other *H. erectus* samples [60]. Taken together, this lesser metric and non-metric variation may suggest that the Ngandong assemblage samples a single population (i.e. is a true deme), as was our intention here.

We started this endeavour with the most conservative approach of considering variation in genera that are broadly dispersed and may have similar underlying tendencies toward plasticity as humans (indeed both vervets and macaques are often classified as pest species). Our data suggest that humans are not unique in how much they vary across populations compared to such dispersed species, but we do not know whether either differ from more geographically localized taxa of non-human primates. Nor do we know whether widely dispersed non-primate mammals such as carnivores will show similar patterns of variation as do widely dispersed primates. For the time being, there are a few hints of unusual variability in *H. erectus* as proxied by differences across paleodemes compared with other fossil and extant taxa—but the existence of such variation and its causes remain to be confirmed when larger samples of known-locality, extant skeletal remains and additional fossils become available. Our results so far suggest that comparisons of CVs across species should yield insignificant results when sufficiently large samples are available and single measures are compared and that CV seems not to be the way to look at the inherent ability of a species to vary.

References

1. Kuzawa CW, Bragg JM. 2012 Plasticity in human life history strategy: implications for contemporary human variation and the evolution of genus *Homo*. *Curr. Anthropol.* **53**(Suppl. 6), S369–S382. (doi:10.1086/667410)

2. Antón SC, Potts R, Aiello LC. 2014 Evolution of early *Homo*: an integrated biological perspective. *Science* **345**, 1236828. (doi:10.1126/science.1236828)

3. Antón SC, Snodgrass JJ. 2012 Origins and evolution of the genus *Homo*. *Curr. Anthropol.* **53**(Suppl. 6), S479–S496. (doi:10.1086/667692)

4. West-Eberhard MJ. 2003 *Developmental plasticity and evolution*. New York: Oxford University Press.

5. Walker R *et al.* 2006 Growth rates and life histories in twenty-two small-scale societies. *Am. J. Hum. Biol.* **18**, 295–311. (doi:10.1002/ajhb.20510)

6. Chisholm JS. 1993 Death, hope, and sex: life-history theory and the development of reproductive strategies. *Curr. Anthropol.* **34**, 1–12. (doi:10.1086/204131)

7. Eveleth PB, Tanner JM. 1990 *Worldwide variation in human growth*, 2nd edn. New York, NY: Cambridge University Press.

8. Kaplin BA. 1954 Environment and human plasticity. *Am. Anthropol.* **56**, 780–800. (doi:10.1525/aa.1954.56.5.02a00050)

9. Wilson M, Daly M. 1997 Life expectancy, economic inequality, homicide, and reproductive timing in Chicago neighbourhoods. *Br. Med. J.* **314**, 1271–1274. (doi:10.1136/bmj.314.7089.1271)

10. Isbell LA. 1994 Predation on primates: ecological patterns and evolutionary consequences. *Evol. Anthropol.* **3**, 61–71. (doi:10.1002/evan.1360030207)

11. Bribiescas RG, Ellison PT, Gray PB. 2012 Male life history, reproductive effort, and the evolution of the genus *Homo*. *Curr. Anthropol.* **53**(Suppl. 6), S424–S435. (doi:10.1086/667538)

12. Simpson GG. 1961 *Principles of animal taxonomy*. New York, NY: Columbia University Press.

13. Howell FC. 1999 Paleo-demes, species, clades and extinctions in the Pleistocene hominid record. *J. Anthropol. Res.* **55**, 191–243. (doi:10.1086/jar.55.2.3631209)

14. Dubois E. 1924 On the principal characters of the cranium and the brain, the mandible and the teeth of *Pithecanthropus erectus*. *Proc. Acad. Sci.* **27**, 265–278.

15. Dubois E. 1932 The distinct organization of *Pithecanthropus* of which the femur bears evidence, now confirmed from other individuals of the described species. *Proc. Acad. Sci.* **35**, 716–722.

16. Virchow R. 1895 Sur le *Pithecanthropus*. *Z. Ethnol.* **27**, 435–440.

17. Black D. 1931 On an adolescent skull of *Sinanthropous pekinensis* in comparison with an adult skull of the same species and with other hominid skulls, recent and fossil. *Palaeontol. Sin. D* **7**, 1–114.

18. von Koengiswald GHR, Weidenreich F. 1939 The relationship between *Pithecanthropus* and *Sinanthropus*. *Nature* **144**, 926–929. (doi:10.1038/144926a0)

19. Weidenreich F. 1940 The torus occipitalis and related structures and their transformations in the course of human evolution. *Bull. Geol. Soc. China* **19**, 479–559. (doi:10.1111/j.1755-6724.1939.mp19004005.x)

20. Weidenreich F. 1941 The extremity bones of *Sinanthropus pekinensis*. *Palaeontol. Sin. D* **5**, 1–150.

21. Weidenreich F. 1943 The skull of *Sinanthropus pekinensis*; a comparative study on a primitive hominid skull. *Palaeontol. Sin. D* **10**, 1–298.

22. Boule M. 1929 Le *Sinanthropus*. *L'Anthropologie* **39**, 455.

23. Boule M, Vallois H. 1957 *Fossil men*. New York, NY: Dryden Press.

24. Mayr E. 1944 On the concepts and terminology of vertical subspecies and species. *Common Problems Genetics Paleontol. Syst. Bull.* **2**, 11–16.

25. Mayr E. 1950 Taxonomic categories of fossil hominids. *Cold Spring Harbor Symp. Quant. Biol.* **25**, 109–118. (doi:10.1101/SQB.1950.015.01.013)

26. Leakey REF. 1976 New hominid fossils from the Koobi Fora formation in Northern Kenya. *Nature* **261**, 574–576. (doi:10.1038/261574a0)

27. Leakey REF, Walker AC. 1985 Further hominids from the Plio-Pleistocene of Koobi Fora, Kenya. *Am. J. Phys. Anthropol.* **67**, 135–163. (doi:10.1002/ajpa.1330670209)

28. Leakey REF, Walker AC. 1976 *Australopithecus, Homo erectus* and the single species hypothesis. *Nature* **261**, 572–574. (doi:10.1038/261572a0)

29. Wood B. 1991 *Koobi Fora research project, volume 4, hominid cranial remains*. Oxford, UK: Clarendon Press.

30. Stringer CB. 1984 The definition of *Homo erectus* and the existence of the species in Africa and Europe. *Courier Forsch. Inst. Senckenberg* **69**, 131–143.

31. Andrews P. 1984 On the characters that define *Homo erectus*. *Courier Forsch. Inst. Senckenberg* **69**, 167–175.

32. Tattersall I. 1986 Species recognition in human paleontology. *J. Hum. Evol.* **15**, 165–175. (doi:10.1016/S0047-2484(86)80043-4)

33. Groves CP, Mazek V. 1975 An approach to the taxonomy of the Hominidae: gracile Villafranchian hominids of Africa. *Caspos Mineral. Geol.* **20**, 225–247.

34. Swisher III CC, Curtis GH, Jacob T, Getty AG, Suprijo A, Widiasmoro. 1994 Age of the earliest known hominids in Java, Indonesia. *Science* **263**, 118–1121. (doi:10.1126/science.8108729)

35. Asfaw B, Gilbert WH, Beyene Y, Hart WK, Renne PR, WoldeGabriel G, Vrba ES, White TD. 2009 Remains of *Homo erectus* from Bouri, Middle Awash, Ethiopia. *Nature* **416**, 317–320. (doi:10.1038/416317a)

36. Howells WW. 1980 *Homo erectus*: who, when and where. *Yearb. Phys. Anthropol.* **23**, 1–23. (doi:10.1002/ajpa.1330230503)

37. Rightmire GP. 1993 *The evolution of* Homo erectus, *comparative anatomical studies of an extinct human species*. New York, NY: Cambridge University Press.

38. Spoor F, Leakey MG, Gathogo PN, Brown FH, Antón SC, McDougall I, Kiarie C, Manthi FK, Leakey LN. 2007 Implications of new early *Homo* fossils from Ileret, east of Lake Turkana, Kenya. *Nature* **448**, 688–691. (doi:10.1038/nature05986)

39. Baab KL. 2008 The taxonomic implications of cranial shape variation in *Homo erectus*. *J. Hum. Evol.* **54**, 827–847. (doi:10.1016/j.jhevol.2007.11.003)

40. Antón SC. 2003 Natural history of *Homo erectus*. *Yearb. Phys. Anthropol.* **46**, 126–170. (doi:10.1002/ajpa.10399)

41. Walker A, Leakey R (eds). 1993 *The Nariokotome* Homo erectus *skeleton*. Cambridge, MA: Harvard University Press.

42. Gabunia L *et al.* 2000 Earliest Pleistocene cranial remains from Dmanisi, Republic of Georgia: taxonomy, geological setting, and age. *Science* **288**, 1019–1025. (doi:10.1126/science.288.5468.1019)

43. Lordkipanidze D *et al.* 2007 Postcranial evidence from early *Homo* from Dmanisi, Georgia. *Nature* **449**, 305–310. (doi:10.1038/nature06134)

44. Potts R, Behrensmeyer AK, Deino A, Ditchfield P, Clark J. 2004 Small mid-Pleistocene hominin associated with East African Acheulean technology. *Science* **305**, 75–78. (doi:10.1126/science.1097661)

45. Simpson SW, Quade J, Levin NE, Butler R, Dupont-Nivet G, Everett M, Semaw S. 2008 A female *Homo erectus* pelvis from Gona, Ethiopia. *Science* **322**, 1089–1092. (doi:10.1126/science.1163592)

46. Lordkipanidze D, de León MSP, Margvelashvili A, Rak Y, Rightmire GP, Vekua A, Zollikofer CP. 2013 A complete skull from Dmanisi, Georgia, and the evolutionary biology of early *Homo*. *Science* **342**, 326–331. (doi:10.1126/science.1238484)

47. Antón SC. 2012 Early *Homo*: who, when and where?. *Curr. Anthropol.* **53**(Suppl. 6), S278–S298. (doi:10.1086/667695)

48. Spoor F, Gunz P, Neubauer S, Stelzer S, Scott N, Kwekason A, Dean MC. 2015 Reconstructed *Homo habilis* type OH 7 suggests deep-rooted species diversity in early *Homo*. *Nature* **519**, 83–86. (doi:10.1038/nature14224)

49. Spoor F. 2013 Palaeoanthropology: small-brained and big-mouthed. *Nature* **502**, 452–453. (doi:10.1038/502452a)

50. Dembo M, Matzke NJ, Mooers AØ, Collard M. 2015 Bayesian analysis of a morphological supermatrix sheds light on controversial fossil hominin relationships. *Proc. R. Soc. B* **282**, 20150943. (doi:10.1098/rspb.2015.0943)

51. Schwartz JH, Tattersall I, Chi Z. 2014 Comment on 'A complete skull from Dmanisi, Georgia, and the evolutionary biology of early *Homo*'. *Science* **344**, 360. (doi:10.1126/science.1250056)

52. Aiello LC, Wheeler P. 1995 The expensive tissue hypothesis: the brain and digestive system in human and primate evolution. *Curr. Anthropol.* **36**, 199–221. (doi:10.1086/204350)

53. Antón SC, Leonard WR, Robertson M. 2002 An ecomorphological model of the initial hominid dispersal from Africa. *J. Hum. Evol.* **43**, 773–785. (doi:10.1006/jhev.2002.0602)

54. Leonard WR, Robertson ML. 1994 Evolutionary perspectives on human nutrition: the influence of brain and body size on diet and metabolism. *Am. J. Hum. Biol.* **6**, 77–88. (doi:10.1002/ajhb.1310060111)

55. Pontzer H, Rolian C, Rightmire GP, Jashashvili T, de León MSP, Lordkipanidze D, Zollikofer CP. 2010 Locomotor anatomy and biomechanics of the Dmanisi hominins. *J. Hum. Evol.* **58**, 492–504. (doi:10.1016/j.jhevol.2010.03.006)

56. Shipman P, Walker A. 1989 The costs of becoming a predator. *J. Hum. Evol.* **18**, 373–392. (doi:10.1016/0047-2484(89)90037-7)

57. Aiello LC, Key C. 2002 Energetic consequences of being a *Homo erectus* female. *Am. J. Hum. Biol.* **14**, 551–565. (doi:10.1002/ajhb.10069)

58. Graves RR, Lupo AC, McCarthy RC, Wescott DJ, Cunningham DL. 2010 Just how strapping was KNM-WT 15000? *J. Hum. Evol.* **59**, 542–554. (doi:10.1016/j.jhevol.2010.06.007)

59. Grabowski M, Hatala KG, Jungers WL, Richmond BG. 2015 Body mass estimates of hominin fossils and the evolution of human body size. *J. Hum. Evol.* **85**, 75–93. (doi:10.1016/j.jhevol.2015.05.005)

60. Antón SC. 2002 Evolutionary significance of cranial variation in Asian *Homo erectus*. *Am. J. Phys. Anthropol.* **118**, 301–323. (doi:10.1002/ajpa.10091)

61. Antón SC. 2007 Climatic influences on the evolution of early *Homo*? *Folia Primatol.* **78**, 365–388. (doi:10.1159/000105150)

62. Kidder JH, Durband AC. 2004 A re-evaluation of the metric diversity within *Homo erectus*. *J. Hum. Evol.* **46**, 297–313. (doi:10.1016/j.jhevol.2003.12.003)

63. Rightmire GP. 1981 Patterns of evolution in *Homo erectus*. *Paleobiology* **7**, 241–246.

64. Plavcan JM. 2012 Body size, size variation and sexual size dimorphism in early *Homo*. *Curr. Anthropol.* **53**(Suppl. 6), S409–S423. (doi:10.1086/667605)

65. Leigh SR. 1992 Cranial capacity evolution in *Homo erectus* and early *Homo sapiens*. *Am. J. Phys. Anthropol.* **87**, 1–13. (doi:10.1002/ajpa.1330870102)

66. Antón SC, Swisher III CC. 2001 Evolution and variation of cranial capacity in Asian *Homo erectus*.

In *A scientific life: papers in honor of Professor Dr. Teuku Jacob* (ed. E Indriati), pp. 25–39. Yogyakarta, Indonesia: Bigraf Publishing.

67. Antón SC, Spoor F, Fellmann CD, Swisher III CC. 2007 Defining *Homo erectus*: size considered. In *Handbook of Paleoanthropology, Volume 3*, Chapter 11 (eds W Henke, I Tattersall), pp. 1655–1693. Berlin, Germany: Springer.

68. Ruff CB. 2002 Variation in human body size and shape. *Annu. Rev. Anthropol.* **31**, 211–232. (doi:10.1146/annurev.anthro.31.040402.085407)

69. Leonard WR, Katzmarzyk PT. 2010 Body size and shape: climatic and nutritional influences on human body morphology. In *Human evolutionary biology* (ed. MP Muehlenbein), pp. 157–169. Cambridge, UK: Cambridge University Press.

70. Seltzer CC, Stefansson V. 1933 The anthropometry of the Western and Copper Eskimos, based on data of Vilhjalmur Stefansson. *Hum. Biol.* **5**, 313–370.

71. El-Zaatari S. 2008 Occlusal molar microwear and the diets of the Ipiutak and Tigara populations (Point Hope) with comparisons to the Aleut and Arikara. *J. Archaeol. Sci.* **35**, 2517–2522. (doi:10.1016/j.jas.2008.04.002)

72. Hilton CE, Auerbach BM, Cowgill LW, eds. 2014 *The foragers of Point Hope: the biology and archaeology of humans on the edge of the Alaskan Arctic*. Cambridge, UK: Cambridge University Press.

73. Waugh L. 1930 A study of the nutrition and teeth of the Eskimo of North Bering Sea and Arctic Alaska. *J. Dent. Res.* **10**, 387–393.

74. Jamison PL, Zegura SL. 1970 An anthropometric study of the Eskimos of Wainwright, Alaska. *Arctic Anthropol.* **7**, 125–143.

75. Draper HH. 1980 Nutrition. In *The human biology of circumpolar populations* (ed. FA Milan), pp. 257–284. Cambridge, UK: Cambridge University Press.

76. Serrat IM, Williams RM, Farnum CE. 2010 Exercise mitigates the stunting effect of cold temperature on limb elongation in mice by increasing solute delivery to the growth plate. *J. Appl. Physiol.* **109**, 1869–79. (doi:10.1152/japplphysiol.01022.2010)

77. Middleton ER. 2015 Ecogeographical influences on trunk modularity in recent humans. PhD dissertation, New York University, USA.

78. Antón SC, Carter-Menn H, DeLeon VB. 2011 Modern human origins: continuity, replacement, and masticatory robusticity in Australasia. *J. Hum. Evol.* **60**, 70–82. (doi:10.1016/j.jhevol.2010.08.004)

79. Stefánsson V. 1913 *My life with the Eskimo*. New York, NY: Macmillan Company.

80. Boas F. 1928 *Materials for the study of inheritance in Man*. Vol. VI, *Columbia University Contributions to Anthropology*. New York, NY: Columbia University Press.

81. Riggs P. 1994 The standard of living in Scotland, 1800–1850. In *Stature, living standards, and economic development* (ed. J Komlos), pp. 60–75. Chicago, IL: University of Chicago Press.

82. Cobb WM. 1935 Municipal history from anatomical records. *The Scientific Monthly* **50**, 157–162.

83. Katz MB. 1996 *In the shadow of the poorhouse: a social history of welfare in America*, 10th Anniversary edn. New York, NY: Basic Books.

84. Raines JL. 2014 The importance of documentary evidence in understanding demographic patterns at the Erie County Poorhouse (1851–1926). *Am. J. Phys. Anthropol.* **153**, 216–216.

85. Lanphear KM. 1990 Frequency and distribution of enamel hypoplasias in a historic skeletal sample. *Am. J. Phys. Anthropol.* **81**, 35–43. (doi:10.1002/ajpa.1330810106)

86. Knowles KC, Sirianni JE. 2014 Pauper diet, pauper dentition: an analysis of oral health at the Erie County Poorhouse. *Am. J. Phys. Anthropol.* **153**, 160.

87. Lee PC, Priston NE. 2005 Human attitudes to primates: perceptions of pests, conflict and consequences for primate conservation. In *Commensalism and conflict: the primate–human interface* (ed. JD Paterson). Winnipeg, Manitoba: Hignell Printing.

88. Turner TR, Anapol F, Jolly CJ. 1997 Growth, development and sexual dimorphism in vervet monkeys (*Cercopithecus aethiops*) at four sites in Kenya. *Am. J. Phys. Anthropol.* **103**, 19–35. (doi:10.1002/(SICI)1096-8644(199705)103:1<19::AID-AJPA3>3.0.CO;2-8)

89. Hamada Y, Watanabe T, Iwamoto M. 1996 Morphological variations among local populations of Japanese macaque (*Macaca fuscata*). In *Variations in the Asian macaques* (eds T Shotake, K Wada), pp. 97–115. Tokai, Japan: Tokai University Press.

90. Weinstein KJ. 2001 Comparative skeletal morphology of modern humans and macaques from high and low altitudes. PhD dissertation, University of Florida, USA.

91. Antón SC. 1994 Masticatory muscle architecture and bone morphology in primates. PhD dissertation, University of California, USA.

92. Koike S, Shimamura T. 1988 Age determination and skeletal growth of Japanese monkeys (*Macaca fuscata*) using specimens of the Takogayama T-1 Troop. *J. College Lib. Arts Saitama Univ.* **24**, 73–85.

93. Hanya G. 2010 Ecological adaptations of temperate primates: population density of Japanese macaques. In *The Japanese macaques* (eds N Nakagawa, M Nakamish, H Sugiura), pp. 79–97. New York, NY: Springer.

94. Turnquist JE, Kessler MJ. 1989 Free-ranging Cayo Santiago rhesus monkeys (*Macaca mulatta*): I. Body size, proportion, and allometry. *Am. J. Primatol.* **19**, 1–13. (doi:10.1002/ajp.1350190102)

95. Rawlins RG, Kessler MJ. 1986 *The Cayo Santiago macaques: history, behavior, and biology*. New York, NY: SUNY Press.

96. Rawlins RG, Kessler MJ. 1982 A five-year study of tetanus in the Cayo Santiago rhesus monkey colony: behavioral description and epizootiology. *Am. J. Primatol.* **3**, 23–39. (doi:10.1002/ajp.1350030103)

97. File S, Kessler MJ. 1989 Parasites of free-ranging Cayo Santiago macaques after 46 years of isolation. *Am. J. Primatol.* **18**, 231–236. (doi:10.1002/ajp.1350180306)

98. Blomquist GE. 2012 Female age of first reproduction at Cayo Santiago: heritability and shared environments. In *Bones, genetics, and behavior of rhesus macaques* Macaca mulatta *of Cayo Santiago and beyond* (ed. Q Wang). New York, NY: Springer.

99. Lee PC, Kappeler PM. 2003 Socio-ecological correlates of phenotypic plasticity in primate life history. In *Primate life histories and socioecology* (eds PM Kappeler, ME Pereria), pp. 41–65. Chicago, IL: University of Chicago Press.

100. Indriati E et al. 2011 The age of the 20 meter Solo River terrace, Java, Indonesia and the survival of *Homo erectus* in Asia. *PLoS ONE* **6**, e21562. (doi:10.1371/journal.pone.0021562)

101. Huffman OF, De Vos J, Berkhout AW, Aziz F. 2010 Provenience reassessment of the 1931–1933 Ngandong *Homo erectus* (Java), confirmation of the bone-bed origin reported by the discoverers. *PaleoAnthropology* **2010**, 1–60. (doi:10.4207/PA.2010.ART34)

102. Shen G, Gao X, Gao B, Granger DE. 2009 Age of Zhoukoudian *Homo erectus* determined with $^{26}Al/^{10}Be$ burial dating. *Nature* **458**, 198–200. (doi:10.1038/nature07741)

103. Lepre CJ, Kent DV. 2015 Chronostratigraphy of KNM-ER 3733 and other Area 104 hominins from Koobi Fora. *J. Hum. Evol.* **86**, 99–111. (doi:10.1016/j.jhevol.2015.06.010)

104. Antón SC, Swisher III CC. 2004 Early dispersals of *Homo* from Africa. *Annu. Rev. Anthropol.* **33**, 271–296. (doi:10.1146/annurev.anthro.33.070203.144024)

105. Gilbert WH, Asfaw B. 2008 Homo erectus: *Pleistocene evidence from the Middle Awash, Ethiopia* (Vol. 1). Berkeley, CA: University of California Press.

106. Tappen M, Lordkipanidze D, Bukshianidze M, Ferring R, Vekua A. 2007 Are you in or out (of Africa)? Site formation at Dmanisi and actualistic studies in Africa. In *Breathing life into fossils: taphonomic studies in honor of C.K. (Bob) Brain. Stone Age Institute Publication Series Number 2* (eds TR Pickering, K Schick, N Toth), chapter 7. Gosport, IN: Stone Age Institute Press.

107. Margvelashvili A, Zollikofer CPE, Lordkipanidze D, Tafforeau P, Ponce de León MS. In press. Comparative analysis of dentognathic pathologies in the Dmanisi mandibles. *Am. J. Phys. Anthropol.* (doi:10.1002/ajpa.22966)

108. Trinkaus E. 2014 *The Shanidar Neandertals*. New York, NY: Academic Press.

109. Heim J-L. 1980 *Les Homme Fossiles de la Ferrassie. Tome 1 Memoire 35 Archives de L'institut de Paleontologie Humaine*. Paris, France: Masson.

110. Heim J-L. 1980 *Les Homme Fossiles de la Ferrassie. Tome 2 Memoire 38 Archives de L'institut de Paleontologie Humaine*. Paris, France: Masson.

111. Wood RE et al. 2013 A new data for the Neanderthals from El Sidrón Cave (Asturias, Northern Spain). *Archaeometry* **55**, 148–158. (doi:10.1111/j.1475-4754.2012.00671.x)

112. Bocquet-Appel JP, Arsuaga JL. 1999 Age distributions of hominid samples at Atapuerca (SH) and Krapina could indicate accumulation by

catastrophe. *J. Archaeol. Sci.* **26**, 327–338. (doi:10.1006/jasc.1998.0370)

113. Meyer M *et al.* 2016 Nuclear DNA sequences from the Middle Pleistocene Sima de los Huesos hominins. *Nature* **531**, 504–507. (doi:10.1038/nature17405)

114. Berger LR *et al.* 2015 *Homo naledi*, a new species of the genus *Homo* from the Dinaledi Chamber, South Africa. *eLife* **4**, e09560. (doi:10.7554/eLife.09560)

115. Wood B, Lieberman DE. 2001 Craniodental variation in *Paranthropus boisei:* a developmental and functional perspective. *Am. J. Phys. Anthropol.* **116**, 13–25. (doi:10.1002/ajpa.1097)

116. Haldane JBS. 1955 The measurement of variation. *Evolution* **9**, 484. (doi:10.2307/2405484)

117. Sokal RR, Rohlf FJ. 2012 *Biometry*, 4th edn. New York, NY: W.H. Freeman and Co.

118. Fligner MA, Killen TJ. 1976 Distribution-free two sample tests for scale. *J. Am. Stat. Assoc.* **71**, 210–213. (doi:10.1080/01621459.1976.10481517)

119. Donnelly SM, Kramer A. 1999 Testing for multiple species in fossil samples: an evaluation and comparison of tests for equal relative variation. *Am. J. Phys. Anthropol.* **108**, 507–529. (doi:10.1002/(SICI)1096-8644(199904)108:4<507::AID-AJPA8>3.0.CO;2-0)

120. Sokal RR, Braumann CA. 1980 Significance tests for coefficients of variation and variability profiles. *Syst. Zool.* **29**, 50–66. (doi:10.2307/2412626)

121. Lewontin RC. 1966 On the measurement of relative variability. *Syst. Biol.* **15**, 141–142. (doi:10.2307/2411632)

122. Benjamini Y, Yekutieli D. 2001 The control of the false discovery rate in multiple testing under dependency. *Ann. Stat.* **29**, 1165–1188. (doi:10.1214/aos/1013699998)

123. Plavcan JM, van Schaik CP. 1997 Intrasexual competition and body weight dimorphism in anthropoid primates. *Am. J. Phys. Anthropol.* **103**, 37–68. (doi:10.1002/(SICI)1096-8644(199705)103:1<37::AID-AJPA4>3.0.CO;2-A)

124. Isler K, van Schaik CP. 2012 How our ancestors broke through the gray ceiling: comparative evidence for cooperative breeding in early *Homo. Curr. Anthropol.* **53**(S6), S453–S465. (doi:10.1086/667623)

125. McHenry HM. 1992 Body size and proportions in early hominids. *Am. J. Phys. Anthropol.* **87**, 407–431. (doi:10.1002/ajpa.1330870404)

126. McHenry HM. 1994 Behavioral ecological implications of early hominid body size. *J. Hum. Evol.* **27**, 77–87. (doi:10.1006/jhev.1994.1036)

18

rstb.royalsocietypublishing.org   *Phil. Trans. R. Soc. B* **371**: 20150236