Human Genetics (2018) 137:487–509
https://doi.org/10.1007/s00439-018-1902-3

ORIGINAL INVESTIGATION

A genome scan for genes underlying adult body size differences between Central African hunter-gatherers and farmers

Trevor J. Pemberton[1] · Paul Verdu[2] · Noémie S. Becker[3] · Cristen J. Willer[4] · Barry S. Hewlett[5] · Sylvie Le Bomin[2] · Alain Froment[6] · Noah A. Rosenberg[7] · Evelyne Heyer[2]

Received: 5 November 2017 / Accepted: 3 July 2018 / Published online: 14 July 2018

Abstract

The evolutionary and biological bases of the Central African "pygmy" phenotype, a characteristic of rainforest hunter-gatherers defined by reduced body size compared with neighboring farmers, remain largely unknown. Here, we perform a joint investigation in Central African hunter-gatherers and farmers of adult standing height, sitting height, leg length, and body mass index (BMI), considering 358 hunter-gatherers and 169 farmers with genotypes for 153,798 SNPs. In addition to reduced standing heights, hunter-gatherers have shorter sitting heights and leg lengths and higher sitting/standing height ratios than farmers and lower BMI for males. Standing height, sitting height, and leg length are strongly correlated with inferred levels of farmer genetic ancestry, whereas BMI is only weakly correlated, perhaps reflecting greater contributions of non-genetic factors to body weight than to height. Single- and multi-marker association tests identify one region and eight genes associated with hunter-gatherer/farmer status, and 24 genes associated with the height-related traits. Many of these genes have putative functions consistent with roles in determining their associated traits and the pygmy phenotype, and they include three associated with standing height in non-Africans (*PRKG1, DSCAM, MAGI2*). We find evidence that European height-associated SNPs or variants in linkage disequilibrium with them contribute to standing- and sitting-height determination in Central Africans, but not to the differential status of hunter-gatherers and farmers. These findings provide new insights into the biological basis of the pygmy phenotype, and they highlight the potential of cross-population studies for exploring the genetic basis of phenotypes that vary naturally across populations.

Introduction

Central African hunter-gatherers have smaller body size compared with their agricultural neighbors ("farmers" henceforth)—a phenomenon commonly termed the "pygmy" phenotype (Froment 1993; Hewlett 2014). Although no discontinuity in adult standing height exists between

Trevor J. Pemberton and Paul Verdu contributed equally to this work.

Electronic supplementary material The online version of this article (https://doi.org/10.1007/s00439-018-1902-3) contains supplementary material, which is available to authorized users.

✉ Trevor J. Pemberton
Trevor.Pemberton@umanitoba.ca

✉ Paul Verdu
verdu@mnhn.fr

✉ Noah A. Rosenberg
noahr@stanford.edu

✉ Evelyne Heyer
heyer@mnhn.fr

[1] Department of Biochemistry and Medical Genetics, University of Manitoba, Winnipeg, MB, Canada

[2] CNRS-MNHN-Université Paris Diderot, UMR 7206 Eco-Anthropologie et Ethnobiologie, Paris, France

[3] Division of Evolutionary Biology, Faculty of Biology, Ludwig-Maximilians-Universität München, Planegg-Martinsried, Germany

[4] Division of Cardiovascular Medicine, University of Michigan, Ann Arbor, MI, USA

[5] Department of Anthropology, Washington State University, Vancouver, WA, USA

[6] IRD–MNHN, UMR 208 Patrimoines locaux, Paris, France

[7] Department of Biology, Stanford University, Stanford, CA, USA

hunter-gatherers and farmers, and contemporary populations are distinguished mostly by cultural criteria rather than height (Bahuchet 1993a, b, 2012; Froment 1993; Hewlett 1996, 2014), hunter-gatherer populations do tend to have reduced stature compared to farming populations. Among hunter-gatherer males, mean standing height ranges from 142 cm for the Efe from eastern Democratic Republic of Congo (DRC) to 161 cm for the Twa from western DRC, averaging 150.6 cm across 23 Central African hunter-gatherer populations [standard deviation (SD) 6.7 cm] (Froment 1993; Hiernaux 1974). This mean is notably lower than the mean male standing height of 167.3 cm for 252 sub-Saharan farming populations that speak Niger-Congo, Nilo-Saharan, or Afro-Asiatic languages (SD 5.7 cm) (Froment 1993).

It has been hypothesized that the diminutive body size of Central African hunter-gatherers is the outcome of adaptive processes in response to hunter-gatherer lifestyles in tropical rainforests with high levels of pathogen exposure (Perry and Dominy 2009). Hypotheses for the basis of this adaptation have considered morphological adaptation to thermoregulation in a hot and humid environment (Cavalli-Sforza 1986), metabolic costs of hunting and gathering in an environment in which food is scarce (Bailey et al. 1989; Bailey 1991; Diamond 1991), and an early onset of reproduction in a context of high mortality rates (Migliano et al. 2007). Physiological studies have found distinctive glucose homeostasis, insulin secretion, and free fatty-acid profiles, and normal human growth hormone (HGH) levels between Central African hunter-gatherers and farmers (Merimee et al. 1972; Rimoin et al. 1969). Reduced growth rates compared with farmers during the first 2 years of life (Rozzi et al. 2015) as well as during adolescence (Hagino et al. 2013; Merimee et al. 1987) have been reported, possibly reflecting perturbation of the insulin-like growth factor 1 (IGF1) (Geffner et al. 1995; Hattori et al. 1996) and HGH (Bozzola et al. 2009) receptor-signaling pathways (Becker et al. 2013). Limitations on physiological data, however, and a paucity of demographic, epidemiological, and paleoanthropological data for Central Africa have contributed to the importance of genetics for assessing hypotheses about the contemporary pygmy phenotype and its evolution (Hewlett 2014).

Population-genetic studies have inferred a shared ancestral origin for West Central African hunter-gatherer populations ~ 3000 years ago (Verdu et al. 2009, 2013) and for West- and East-Central African hunter-gatherer populations ~ 25,000 years ago (Batini et al. 2011; Patin et al. 2009). Though a common origin has also been inferred for Central African hunter-gatherers and farmers, estimates for the time of their divergence range from ~ 50,000 to 70,000 years ago based upon microsatellite, targeted resequencing, and mitochondrial data (Batini et al. 2011; Patin et al. 2009; Verdu et al. 2009) to ~ 30,000–150,000 years ago based upon whole-genome or exome sequences (Hsieh et al. 2016; Lopez et al. 2018). Genetic signals of appreciable admixture with farmers have been detected in hunter-gatherers (Batini et al. 2011; Destro-Bisol et al. 2004; Patin et al. 2009, 2014, 2017; Verdu et al. 2009, 2013), correlating positively with standing height (Becker et al. 2011; Jarvis et al. 2012; Perry et al. 2014a). The genetic differences between hunter-gatherers and farmers and the correlations of standing height and admixture suggest that their contemporary adult body size differences are attributable in part to genetic factors, arguing against a view that the contemporary pygmy phenotype results solely from phenotypic plasticity in challenging nutritional and parasitic environments (Perry and Dominy 2009).

Numerous genome-wide association studies (GWAS) have sought to identify genetic factors contributing to standing height (Berndt et al. 2013; Gudbjartsson et al. 2008; Lango Allen et al. 2010; Lettre et al. 2008; Weedon et al. 2008; Wood et al. 2014), a high-heritability trait ($\geq 69\%$) (Hemani et al. 2013; Perola et al. 2007; Zaitlen et al. 2013) despite strong influences of environmental factors (Katzmarzyk and Leonard 1998; Silventoinen et al. 2000). The largest study to date identified 697 common single-nucleotide polymorphisms (SNPs) significantly associated with standing height in ~ 250,000 individuals of recent European ancestry (Wood et al. 2014). Together, these variants explained 16% of the variance and 20% of the heritability of standing height in their cohort. GWAS on admixed African–American cohorts have identified novel loci significantly associated with standing height (Carty et al. 2012; N'Diaye et al. 2011), illustrating the potential of populations with a component of recent African ancestry to contribute new information to the study of the trait. In this context, Central African hunter-gatherer and farming populations—with their marked differences in standing height that correlate with levels of genetic admixture—provide a potentially powerful system in which to understand the genetic basis of standing height. In particular, use of among-population variation in a phenotype has the potential to uncover alleles whose frequency differences among populations suggest phenotypic contributions.

Recent studies have explored the genetic basis of standing-height variation in Central Africans, comparing genomic SNPs in hunter-gatherer and neighboring farmer populations (Jarvis et al. 2012; Perry et al. 2014a). These studies uncovered signatures of differential polygenic adaptation between hunter-gatherers and farmers, primarily in genomic regions containing genes associated with immunity and metabolism. They also found differing adaptation signatures in western and eastern hunter-gatherers, suggesting a partially convergent origin of the contemporary pygmy phenotype. Perry et al. (2014a) further identified genetic admixture associations with standing height that included four regions previously implicated in standing-height variation in Europeans. However, no standing-height association was seen with the

~ 100 European standing-height-associated SNPs in these datasets. This absence of associations might reflect the combined effects of low power in modestly sized samples and possible differences in linkage disequilibrium (LD) between causal and genotyped SNPs (Shriner et al. 2009) and in effect sizes of causal alleles (Carlson et al. 2013) between Europeans and Central Africans. However, it is also compatible with the view that differences in standing height between hunter-gatherers and farmers do not result solely from general frequency differences for alleles contributing to global human standing-height variation.

Although these recent studies have focused on the differing standing height of Central African hunter-gatherers and farmers, other morphological differences also exist between these two groups. For example, hunter-gatherer body weight is proportionally smaller relative to standing height compared with farmers (Froment 2014; Lemogoum et al. 2012), a difference that cannot be wholly explained by differential nutrition levels (Dietz et al. 1989). In the context of the differing diets of hunter-gatherers and farmers and the greater seasonal variation in food availability for hunter-gatherers compared with farmers, such a difference might reflect coevolution of differences in energy usage and storage in response to different lifestyles. Craniofacial and skeletal dissimilarities have also been identified, including in skull morphology (Froment 2014; Ramirez Rozzi and Sardi 2010) and in leg and forearm length (Froment 2014; Hiernaux 1977), with hunter-gatherers generally having shorter legs and forearms relative to trunk length compared with farmers. These differences potentially reflect evolutionary changes in shared endoskeletal development pathways in response to anatomical constraints of hunting-gathering activities. Reported dissimilarities in trunk and leg length patterns accord with differential growth patterns in response to nutritional (Buschang et al. 1986; Wadsworth et al. 2002) and health (Wadsworth et al. 2002) factors that vary between hunter-gatherers and farmers (de Garine 1990; Hauspie and Pagezy 1989; Ohenjo et al. 2006), and they are consistent with the view that both shared and distinct pathways contribute to determination of upper and lower body size (Soranzo et al. 2009). Although genes contributing to trunk and leg length might be incidentally detected by genetic studies of standing-height variation in Central Africans, those distinct to each trait might fail to be identified owing to insufficient effects on standing height. Thus, analyses of variability in trunk and leg lengths in Central African hunter-gatherers and farmers might disentangle their contributions to hunter-gatherer short stature, providing a more complete picture of the genetic architecture, anatomical constraints, and evolutionary pressures underlying the contemporary pygmy phenotype.

Here, using body size and weight measurements available for > 132 individuals from seven hunter-gatherer and three farming populations from West Central Africa (Table S1) together with genome-wide SNP data, we investigate the genetic basis of standing and sitting height and their ratio as well as that of subischial leg length in Central Africans. Our full sample includes 358 hunter-gatherers and 169 farmers from 13 hunter-gatherer and seven farming populations from West- and East-Central Africa (Fig. 1; Table 1) genotyped on the Illumina Cardio-MetaboChip. This genotyping microarray contains 68,126 SNPs previously identified in well-powered GWAS conducted by consortia focused on anthropometric and cardiovascular traits in addition to 146,453 SNPs chosen to facilitate the fine-mapping of the genomic regions surrounding these 68,126 SNPs (Voight et al. 2012). This array, therefore, allows us to investigate associations with height-related variation in Central Africans of SNPs associated with standing-height variation in Europeans (Lango Allen et al. 2010; Wood et al. 2014). We first evaluate phenotypic differences between hunter-gatherers and farmers in these traits, genetic structure among hunter-gatherer and farming populations, and associations between genetic ancestry and the phenotypes. We then search for alleles and genes associated with hunter-gatherer/farmer status and with trait variation.

Results

Trait differences between hunter-gatherers and farmers

We first investigated if the height and weight traits with measurements available for our sample exhibited significant differences between hunter-gatherers and farmers. Of the 527 individuals in our full sample, we had trait measurements available for 115 hunter-gatherers and 44 farmers representing seven hunter-gatherer and three farming populations from West Central Africa (Table S1).

Among the 115 hunter-gatherers (73 males, 42 females), mean male standing height was 155.67 cm (SD 6.55) and mean female standing height was 148.95 cm (SD 5.82). For the 44 farmers (16 males, 28 females), mean male and female standing heights were 167.21 cm (SD 4.61) and 154.45 cm (SD 6.23), respectively (Fig. 2a). These values are consistent with mean standing heights reported by previous studies of 26 Central African hunter-gatherer [males = 154.85 cm (SD 3.08) and females = 146.30 cm (SD 2.87)] and 10 farming [males = 164.98 cm (SD 2.36) and females = 155.96 cm (SD 0.85)] populations (Froment 2014), highlighting the significant difference in stature that exists between hunter-gatherers and farmers ($P_{\text{males}} = 6.13 \times 10^{-8}$ and $P_{\text{females}} = 2.65 \times 10^{-4}$; Wilcoxon rank-sum test).

Consistent with the high correlations of standing height with the other height-related traits (Figure S1), subischial leg

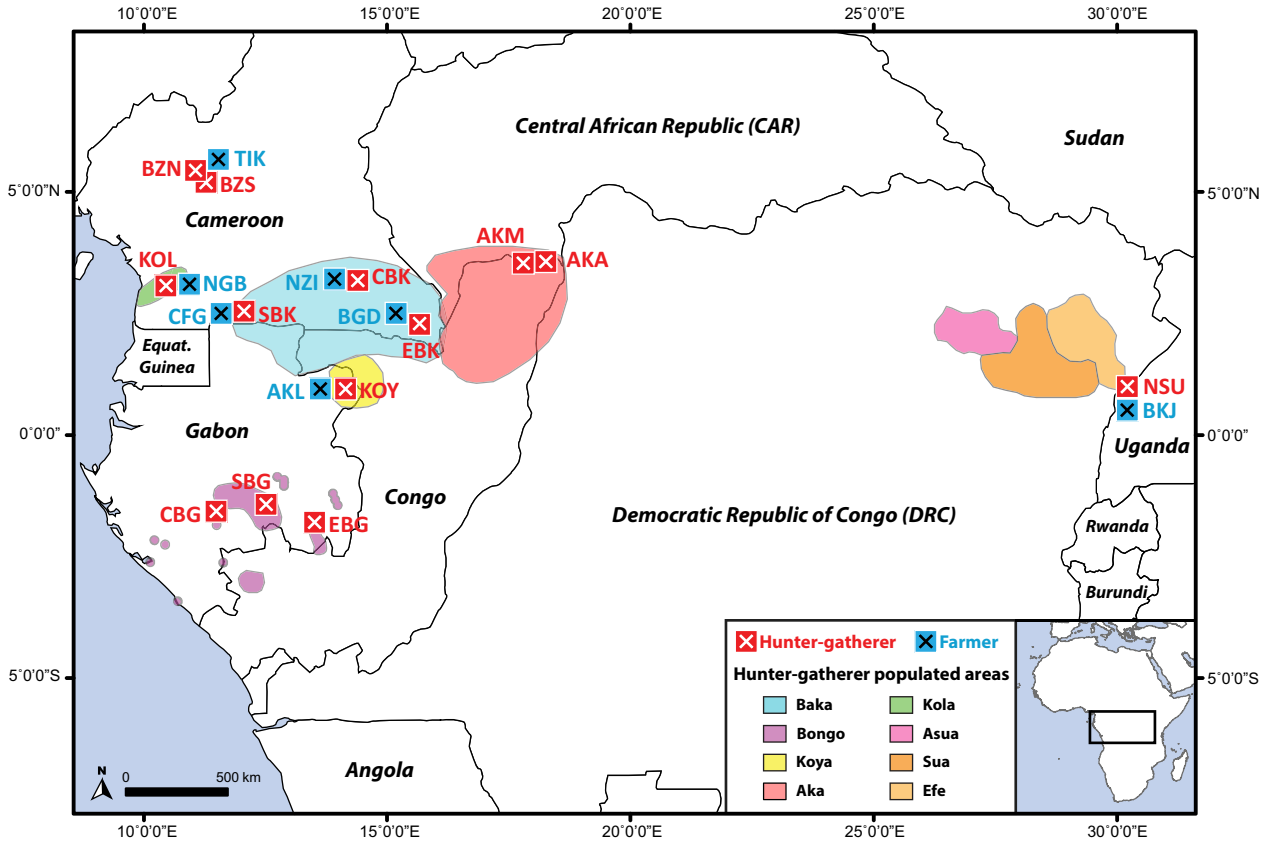

Fig. 1 Sampling locations. Hunter-gatherer/farmer status was assessed in ethnographic field work, relying on cultural criteria that do not include standing height (see "Methods"). Population codes follow Table 1. Population ranges were inferred from ethnographic field work (Bahuchet 2012)

lengths—measured as standing height minus sitting height—of male (mean 74.97 cm, SD 4.91) and female (mean 70.81 cm, SD 3.84) hunter-gatherers were significantly shorter than those of male (mean 82.65 cm, SD 5.85) and female (mean 75.44 cm, SD 3.97) farmers ($P_{\text{males}} = 0.007$ and $P_{\text{females}} = 9.38 \times 10^{-5}$; Fig. 2b). In males, sitting heights were significantly smaller for hunter-gatherers (mean 80.62 cm, SD 2.92) than for farmers (mean 86.13 cm, SD 1.28; $P = 0.002$); sitting heights of female hunter-gatherers (mean 77.74 cm, SD 2.89) and farmers (mean 78.96 cm, SD 3.51) were similar (Fig. 2c, $P = 0.073$). For the ratio of sitting to standing height (Fig. 2d), male and female hunter-gatherers exhibited a similar shift toward higher values compared with farmers. The observed difference in ratios was significant with females ($P = 0.003$) but not with males ($P = 0.203$). These findings are consistent with leg length contributing more to the pygmy phenotype than does upper body length (trunk and head) (Froment 2014).

For BMI, values were similar for male (mean 20.16, SD 1.95) and female (mean 20.06, SD 2.56) hunter-gatherers ($P = 0.587$), whereas those of male farmers (mean 22.31, SD 2.63) significantly exceeded those of female farmers (mean = 21.16, SD 3.07; $P = 0.031$; Fig. 2e). Consequently, hunter-gatherers had significantly lower BMI than farmers when considering males ($P_{\text{males}} = 2.74 \times 10^{-4}$) but not females ($P_{\text{females}} = 0.186$). Our values are generally consistent with mean BMI reported by previous studies for 26 Central African hunter-gatherer [males = 20.04 (SD 0.80) and females = 20.33 (SD 0.75)] and 10 farming [males = 21.48 (SD 0.90) and females = 21.65 (SD 1.02)] populations (Froment 2014). However, several differences are observed between mean BMI in our sample of hunter-gatherer and farming populations (Table S1) and the previous results. Our sample does not support significant male–female differences in hunter-gatherers ($P = 0.198$; Wilcoxon signed-rank test) and farmers ($P = 1$) or significant differences between male hunter-gatherers and farmers ($P_{\text{males}} = 0.258$, Wilcoxon rank-sum test) or between female hunter-gatherers and farmers ($P_{\text{females}} = 0.548$). Although the same comparisons of previously published data also do not support male–female differences in hunter-gatherers ($P = 0.458$) and farmers ($P = 0.682$), they do support significant differences between male hunter-gatherers and male farmers ($P_{\text{males}} = 1.21 \times 10^{-4}$)

Table 1 Populations and their sample sizes

Population		Region and country	Language family	Trait data	Sample size			
Code	Name				Total	Genetic dataset		
						Total	Male	Female
Hunter-gatherers								
AKA	Aka	Lobaye Prefecture, Central African Republic	Bantu C10	No	26	21	14	7
AKM	Aka (Mbati)	Lobaye Prefecture, Central African Republic	Bantu C10	No	17	8	6	2
CBK	Central/Eastern Baka	East Cameroon	Adamawa–Ubangian, Gbanzili–Sere, Ngbaka	Yes	36	29	14	15
EBK	Southeastern Baka	East Cameroon	Adamawa–Ubangian, Gbanzili–Sere, Ngbaka	Yes	11	10	10	0
SBK	Southern Baka	South Cameroon	Adamawa–Ubangian, Gbanzili–Sere, Ngbaka	No	50	41	16	25
BZN[a]	Northern Bezan	Central Cameroon	Bantoid Non-Bantu, Tikar	Yes	27	27	12	15
BZS[a]	Southern Bezan	Central Cameroon	Bantoid Non-Bantu, Tikar	Yes	48			
CBG	Central Bongo	Ngunié Province, Gabon	Bantu B30-B50	Yes	24	21	14	7
EBG	Eastern Bongo	Haut-Ogooué Province, Gabon	Bantu B60-70	Yes	30	19	13	6
SBG	Southern Bongo	Ogooué-Lolo Province, Gabon	Bantu B30	Yes	34	27	13	14
KOY	Koya	Ogooué-Ivindo Province, Gabon	Bantu B20	Yes	25	19	12	7
KOL	Kola	Littoral Region, Cameroon	Bantu A80	No	31	24	18	6
NSU	Nsua (Efe)	Bundibugyo District, Uganda	Central Sudanic, Mangbutu-Efe	No	17	11	6	5
	Total				376	257	148	109
Farmers								
AKL	Akele	Ogooué-Ivindo Province, Gabon	Bantu B20	No	12	8	7	1
CFG	Fang	South Cameroon	Bantu A70	No	36	33	19	14
NZI	Nzime	East Cameroon	Bantu A80	Yes	23	20	3	17
NGB	Ngumba	Littoral Region, Cameroon	Bantu A80	No	58	43	20	23
TIK	Tikar	Central Cameroon	Bantoid Non-Bantu, Tikar	Yes	20	20	12	8
BGD	Bangando	East Cameroon	Adamawa-Ubangian, Bangandu, Gbaya	Yes	6	6	3	3
BKJ	Konjo	Bundibugyo District, Uganda	Bantu J40	No	27	19	8	11
	Total				182	149	72	77

[a]The two Bezan populations were merged into a combined BEZ population in the 406UNRELAT dataset

and between female hunter-gatherers and female farmers ($P_{\text{females}} = 4.57 \times 10^{-4}$).

These findings reaffirm earlier observations of significant differences in standing height between Central African hunter-gatherers and farmers, and they suggest differing roles for trunk and leg length to these differences. However, BMI patterns do not fully accord with previous population-level comparisons of mean BMI in Central African hunter-gatherer and farming populations.

Genetic structure patterns

Having established that significant differences exist between hunter-gatherers and farmers for height-related measurements, we next investigated the genetic structure present among our hunter-gatherer and farmer samples. Following the removal of lower-quality SNPs and closely related individuals, our genetic dataset consisted of 406 unrelated individuals from 20 populations with genotypes at 153,798 SNPs included on the Illumina Cardio-MetaboChip (see "Methods").

In accord with earlier studies (Batini et al. 2011; Destro-Bisol et al. 2004; Jarvis et al. 2012; Patin et al. 2009; Verdu et al. 2009, 2013), multidimensional scaling (MDS) analysis of pairwise allele-sharing dissimilarities (ASD) among these individuals supported three main components to genetic differentiation among Central Africans (Fig. 3a). First, differentiation between hunter-gatherers and farmers largely determined the first MDS dimension, consistent with the expected effects of genetic drift following their divergence ~ 30,000–150,000 years ago (Batini et al. 2011; Hsieh et al. 2016; Lopez et al. 2018; Patin

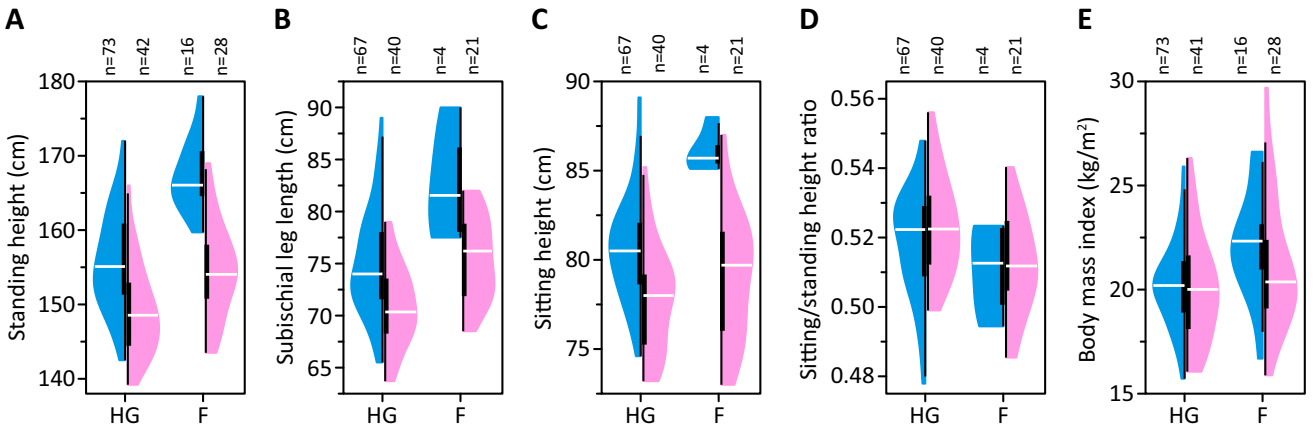

Fig. 2 Distributions of measured traits. Split violin plots (Hintze and Nelson 1998) of trait distributions for males (blue) and females (pink) are shown separately for hunter-gatherers (HG) and farmers (F). **a** Standing height. **b** Subischial leg length. **c** Sitting height. **d** Sitting/standing height ratio. **e** BMI. The numbers (*n*) of males and females with data available are provided. The mean and SD of each group appear in Table S1. Each "violin" contains a vertical black line (25–75% range) and a horizontal white line (median). The width depicts a 90°-rotated kernel density trace

et al. 2009; Verdu et al. 2009). Second, although farmers formed a tight cluster in two dimensions, reflecting the recent history of these mostly Bantu-speaking populations (de Filippo et al. 2012; Montano et al. 2011; Quintana-Murci et al. 2008) (Table 1), differentiation among hunter-gatherers was evident in the second dimension, driven partly by differences between West African Baka (CBK, EBK, SBK) and Aka (AKA, AKM) and East African Nsua (NSU). Dispersion among hunter-gatherers, measured as the variance among ASD values for all hunter-gatherer pairs, significantly exceeded that among farmers (2.07×10^{-5} and 7.76×10^{-6}, respectively; $P<10^{-16}$, one-sided F test), consistent with the expected effects of genetic drift owing to small population sizes and appreciable isolation following ancestral divergences over the last ~ 25,000 years (Batini et al. 2011; Patin et al. 2009).

Third, the degree of genetic differentiation from farmers varied among hunter-gatherer populations. This pattern accords with what might be expected with variable gene flow into hunter-gatherer populations from farming populations, where hunter-gatherers closer to the farmer cluster have greater farmer genetic ancestry than do more distant hunter-gatherers. The Bezan, Kola, and Bongo hunter-gatherers clustered nearer the farmers (mean ASD 0.2047, SD 0.0027, across 20,413 pairs of individuals) than did the Baka and Aka (mean ASD 0.2051, SD 0.0025, 16,241 pairs) and Nsua (mean ASD 0.2072, SD 0.0036, 1639 pairs); the first and third quartiles of the full distribution of ASD values across all pairs of individuals were 0.1983 and 0.2053, respectively. This pattern suggests that Bezan, Kola, and Bongo hunter-gatherers have admixed with neighboring farmers to a greater extent than have the Baka, Aka, and Nsua populations.

Genetic admixture patterns

To further explore potential signals of variable genetic admixture between hunter-gatherers and farmers suggested in our MDS analysis, we applied *STRUCTURE* (Falush et al. 2003; Pritchard et al. 2000) to four non-overlapping sets of 10,106 low-LD genome-wide SNPs. At $K=2$ (Fig. 3b), hunter-gatherers were assigned greater membership in one cluster ("red"), whereas farmers were assigned greater membership in the other cluster ("blue"). Per-individual membership proportions for the blue cluster varied appreciably among hunter-gatherer populations from 0.173 ± 0.103 in the Central/Eastern Baka (CBK) to 0.527 ± 0.096 in the Eastern Bongo (EBG). This pattern indicates that levels of farmer genetic ancestry vary across hunter-gatherer populations. In contrast, among farmers, membership proportions for the red cluster were more stable, equaling on average 0.197 ± 0.093, with generally lower and similar levels of hunter-gatherer genetic ancestry among populations. Additional population structure was not evident at higher K (Figure S2).

The general hunter-gatherer/farmer distinction and variable membership patterns accord with previous studies (Batini et al. 2011; Destro-Bisol et al. 2004; Patin et al. 2009, 2014, 2017; Perry et al. 2014a; Verdu et al. 2009, 2013), and can be interpreted as signals of asymmetric admixture between hunter-gatherers and farmers. In this view, the variable and often nontrivial membership in the blue cluster for hunter-gatherers reflects variable and sometimes considerable admixture with farmers, whereas the lower and less variable membership in the red cluster for farmers reflects comparatively low levels of hunter-gatherer introgression into the farmer gene pool.

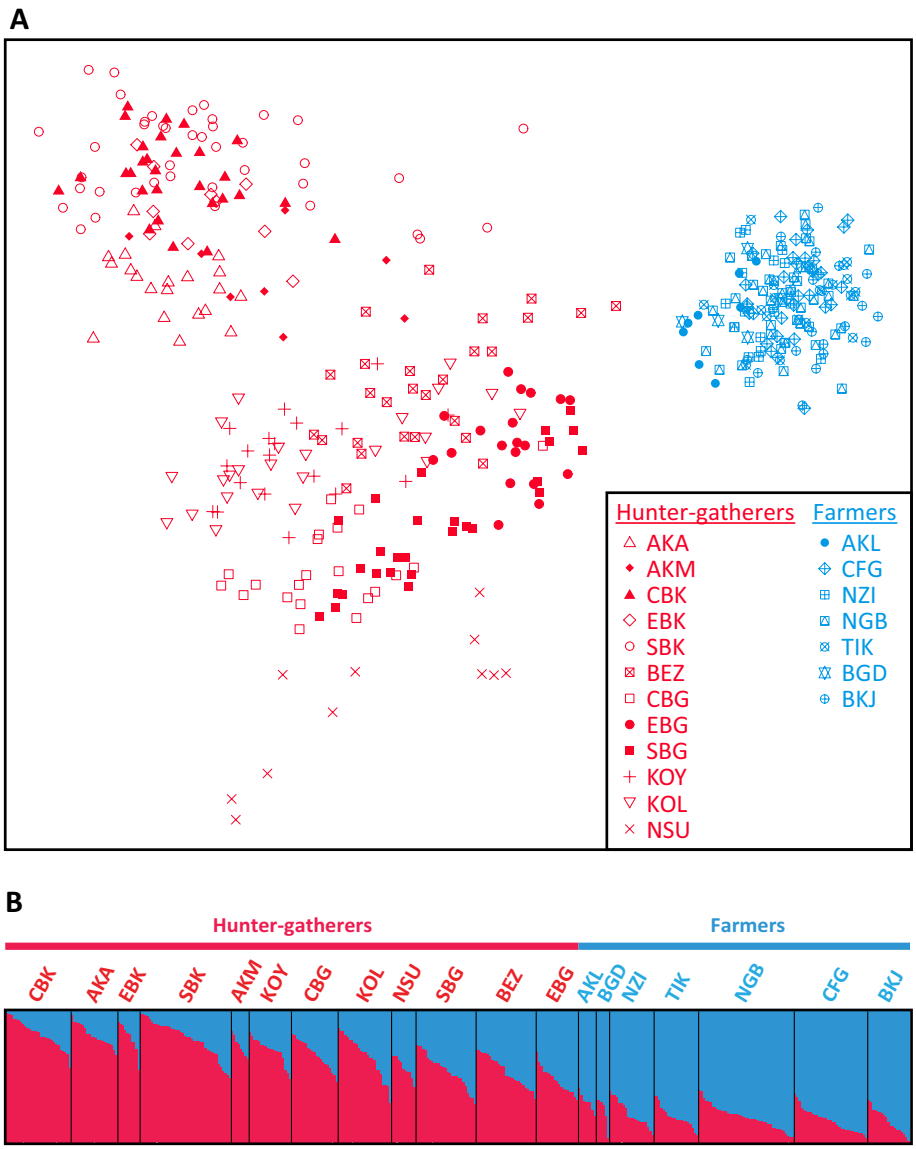

Fig. 3 Genetic structure and admixture for 406 unrelated individuals. **a** Two-dimensional MDS representation of ASD. The first dimension is plotted on the horizontal axis, and the second dimension on the vertical axis. The Spearman ρ between pairwise Euclidean distances in the MDS plot and the ASD matrix is 0.478. **b** Population structure inferred by *STRUCTURE* at $K=2$. Each individual is shown as a thin vertical line partitioned into two components representing inferred membership proportions in the two clusters. Black vertical lines separate individuals from different populations. The most salient population structure observation is the hunter-gatherer/farmer distinction

Correlation between genetic admixture and trait variation

Given the phenotypic and genetic differences observed between hunter-gatherers and farmers in our sample, we sought to establish if genetic factors might underlie observed phenotypic patterns. If genetic differences between hunter-gatherers and farmers contribute to their phenotypic differences, we might expect measurements for these traits to be correlated with per-individual levels of genetic admixture. We, therefore, separately investigated trait correlations with per-individual membership proportions in the blue "farmer" *STRUCTURE* cluster at $K=2$ (Fig. 3b; "farmer admixture" henceforth).

We observed a significant correlation between farmer admixture and standing height with the 76 males (Pearson $r=0.585$, $P=2.90\times10^{-8}$; Fig. 4a) and 57 females ($r=0.485$, $P=1.30\times10^{-4}$; Fig. 4b) in our genetic dataset with standing-height data available. The correlations remained significant when restricted to the 61 male ($r=0.311$, $P=0.015$) and 31 female hunter-gatherers ($r=0.441$, $P=0.013$). These findings accord with previous studies on the relationship between farmer admixture and standing height in Central African hunter-gatherers (Becker et al. 2011; Jarvis et al. 2012; Perry et al. 2014a) and further support an appreciable genetic component to the determination of the contemporary pygmy phenotype.

Correlations of subischial leg length and sitting height with farmer admixture were similar in males [Pearson $r=0.475$ ($P=1.41\times10^{-5}$) and $r=0.431$ ($P=8.90\times10^{-5}$), respectively; Fig. 4c, e]. Among females, the correlation of farmer admixture with sitting height was lower ($r=0.226$, $P=0.040$; Fig. 4f) than with leg length ($r=0.574$, $P=6.53\times10^{-7}$; Fig. 4d). These patterns accord with a claim

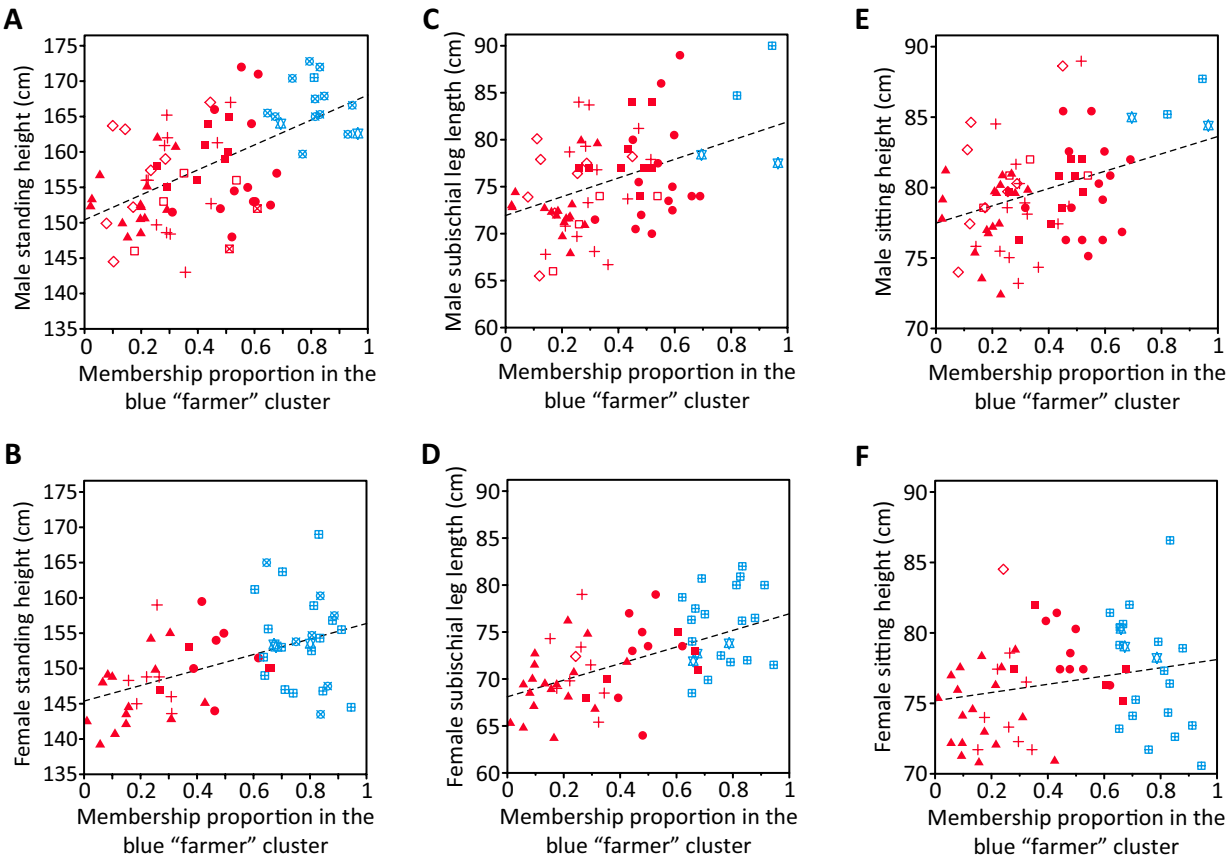

Fig. 4 Relationship between farmer admixture and height-related traits. Scatterplots are shown of individual membership proportions in the blue "farmer" *STRUCTURE* cluster at $K=2$ (Fig. 3b) and each height-related trait. **a** Standing height, 76 males ($r=0.585$, $P=3\times10^{-8}$). **b** Standing height, 57 females ($r=0.485$, $P=3\times10^{-5}$). **c** Subischial leg length, 71 males ($r=0.475$, $P=1.41\times10^{-5}$). **d** Subischial leg length, 57 females ($r=0.574$, $P=6.53\times10^{-7}$). **e** Sitting height, 71 males ($r=0.431$, $P=8.90\times10^{-5}$). **f** Sitting height, 57 females ($r=0.226$, $P=0.040$). Only individuals included in the genetic dataset with a measurement for a trait were included in a comparison. Symbols follow Fig. 3a. The trend line appears in black

that both shared and distinct pathways contribute to upper and lower body size determination (Soranzo et al. 2009), with male–female differences in correlations with non-Pygmy admixture perhaps reflecting differential contributions of genetic variants in estrogen-dependent growth pathways to trunk and leg length (Seeman 1998; Tupman 1962). Finally, BMI was positively correlated with farmer admixture in females ($r=0.262$, $P=0.015$; Figure S3B), but the correlation was not significant in males ($r=0.166$, $P=0.060$; Figure S3A). BMI correlations were weaker than those of height-related measures (Fig. 4), in accord with smaller differences in BMI between hunter-gatherers and famers (Fig. 2e) than in standing height (Fig. 2a).

The correlations with farmer admixture are consistent with a significant genetic component in the determination of height-related differences between hunter-gatherers and farmers. However, absence of strong correlations with BMI is compatible with genetic factors contributing less to observed differences in BMI between hunter-gatherers and farmers, with non-genetic factors such as diet, food security, and activity levels perhaps playing a larger role in body weight differences. Consequently, in seeking to identify genetic factors correlated with phenotypic differences between hunter-gatherers and farmers, we excluded BMI.

Association analyses of hunter-gatherer/farmer status

We have found significant genetic and phenotypic differences between hunter-gatherers and farmers, with correlations between individual trait variation and inferred farmer genetic admixture levels supporting an appreciable genetic component to the determination of the phenotypic differences. To investigate this genetic component of phenotypic differences, we first performed single- and multi-marker association tests to identify genomic regions harboring SNPs with significant allele frequency differences between

the hunter-gatherers and farmers in our genetic dataset of 406 unrelated individuals with genotypes for 153,798 SNPs.

Single-marker association tests

To account for genetic structure (Fig. 3) and background genetic relatedness among individuals due to evolutionary history within a population—which could inflate type-1 and type-2 errors (Kang et al. 2010; Marchini et al. 2004)—we used *EMMAX*, which implements a linear mixed-effect regression model that corrects per-SNP association tests for structure and relatedness via a pairwise kinship matrix (Kang et al. 2010). We evaluated how well we controlled for genetic structure using the genomic inflation factor λ, comparing association tests performed with no covariates to those performed with ethnic categorization (see "Methods") or the first two MDS dimensions (Fig. 3a) included as a covariate (Figure S4A). The MDS covariate provided the λ value closest to one (1.003), and we present the results of that analysis. We identified one SNP on the q-arm of chromosome 1 (rs1539243, $P_{\mathrm{EMMAX}} = 1.46 \times 10^{-8}$; Fig. 5c) with a significant allele frequency difference between hunter-gatherers and farmers after Bonferroni correction (family-wise error rate $\alpha = 5\%$; Fig. 5a).

Multi-marker gene-based association tests

Because individual SNPs might fail to achieve the Bonferroni significance threshold owing to insufficient power in our modest sample, we performed multi-marker gene-based association tests using *VEGAS* (Liu et al. 2010), a multivariate method that combines association signals across all SNPs located within a gene, accounting for gene size and correcting for non-independence among SNPs. Eight genes exhibited a significant association with hunter-gatherer/farmer status after Bonferroni correction ($\alpha = 5\%$; Table 2). The gene encompassing the single SNP significantly associated with hunter-gatherer/farmer status, *IKBKE* (Fig. 5c), ranked 40th across all 23,108 genes tested ($P_{\mathrm{VEGAS}} = 1.72 \times 10^{-4}$). The absence of a significant association with *IKBKE* in our gene-based test reflects the larger number of non-significant SNPs present within its transcribed region that mask the strong association signal of SNP rs1539243.

Relevance to the Central African pygmy phenotype

Many of the genes identified here as harboring SNPs with significant allele frequency differences between hunter-gatherers and farmers have been previously associated with biological processes consistent with their observed phenotypic differences. GPC6 [Mendelian Inheritance in Man (MIM) 604404] is a growth factor receptor that is important for correct growth plate formation (Bassett et al. 2006), and its abrogation causes long-bone growth retardation (Campos-Xavier et al. 2009). It has also been associated with energy and fat storage (Comuzzie et al. 2012). This association thus accords with the observation that hunter-gatherers have significantly shorter legs than farmers (Fig. 2b) and smaller ratios with sitting height (Fig. 2d), and could reflect the co-adaptation of body size with metabolic processes to support a hunter-gatherer lifestyle in the tropical rainforest (Bailey et al. 1989; Bailey 1991; Diamond 1991). Similarly, MACROD2 (MIM 611567) and GP1BB (MIM 138720) are causally implicated in disorders whose clinical features include postnatal growth deficiency (Budarf et al. 1995; Maas et al. 2007), and MACROD2 has also been associated with blood-borne IGF1 levels (Comuzzie et al. 2012). Their associations, therefore, accord with observed reductions in postnatal growth rates in hunter-gatherers relative to farmers (Hagino et al. 2013; Merimee et al. 1987; Rozzi et al. 2015) and with evidence of perturbation of IGF1-signaling in hunter-gatherers (Geffner et al. 1995; Hattori et al. 1996).

Disruption of CACNA1C (MIM 114205) function has been found to cause developmental patterning abnormalities and immunodeficiency (Splawski et al. 2005). Thus, its association might reflect the general craniofacial and skeletal dissimilarities observed between hunter-gatherers and farmers (Froment 2014; Hiernaux 1977; Ramirez Rozzi and Sardi 2010), and is consistent with prior associations between the Central African pygmy phenotype and immune-related genes (Jarvis et al. 2012; Lachance et al. 2012). Indeed, the only SNP identified in our single-marker analyses (rs1539243) is a synonymous mutation in the fourth exon of the gene encoding IKBKE (MIM 605048), an essential regulator of antiviral-signaling pathways (Sharma et al. 2003). Such associations between hunter-gatherer and farmer genetic differences and immune genes are consistent with adaptive processes related to pathogen exposure contributing to the evolution of the contemporary pygmy phenotype (Perry and Dominy 2009).

DLGAP1 (MIM 605445) has been indirectly implicated in fat and glucose metabolism (Aulchenko et al. 2007), and associated with blood-borne insulin-like growth factor binding protein 3 (IGFBP3) levels (Comuzzie et al. 2012) and age at menarche (Perry et al. 2014b). Its association is, therefore, compatible with observed perturbations in metabolic processes (Merimee et al. 1972; Rimoin et al. 1969) and IGF1-signaling (Geffner et al. 1995; Hattori et al. 1996) in hunter-gatherers, and is consistent with the pygmy phenotype being an adaptive response to the metabolic costs of hunting and gathering in the tropical rainforest (Bailey et al. 1989; Bailey 1991; Diamond 1991) and an early onset of reproduction in a context of high mortality rates (Migliano et al. 2007). Finally, PRKG1 (MIM 176894) has been implicated in proper platelet (Li et al. 2003) and cardiovascular (Pfeifer et al. 1998) function, and it has previously

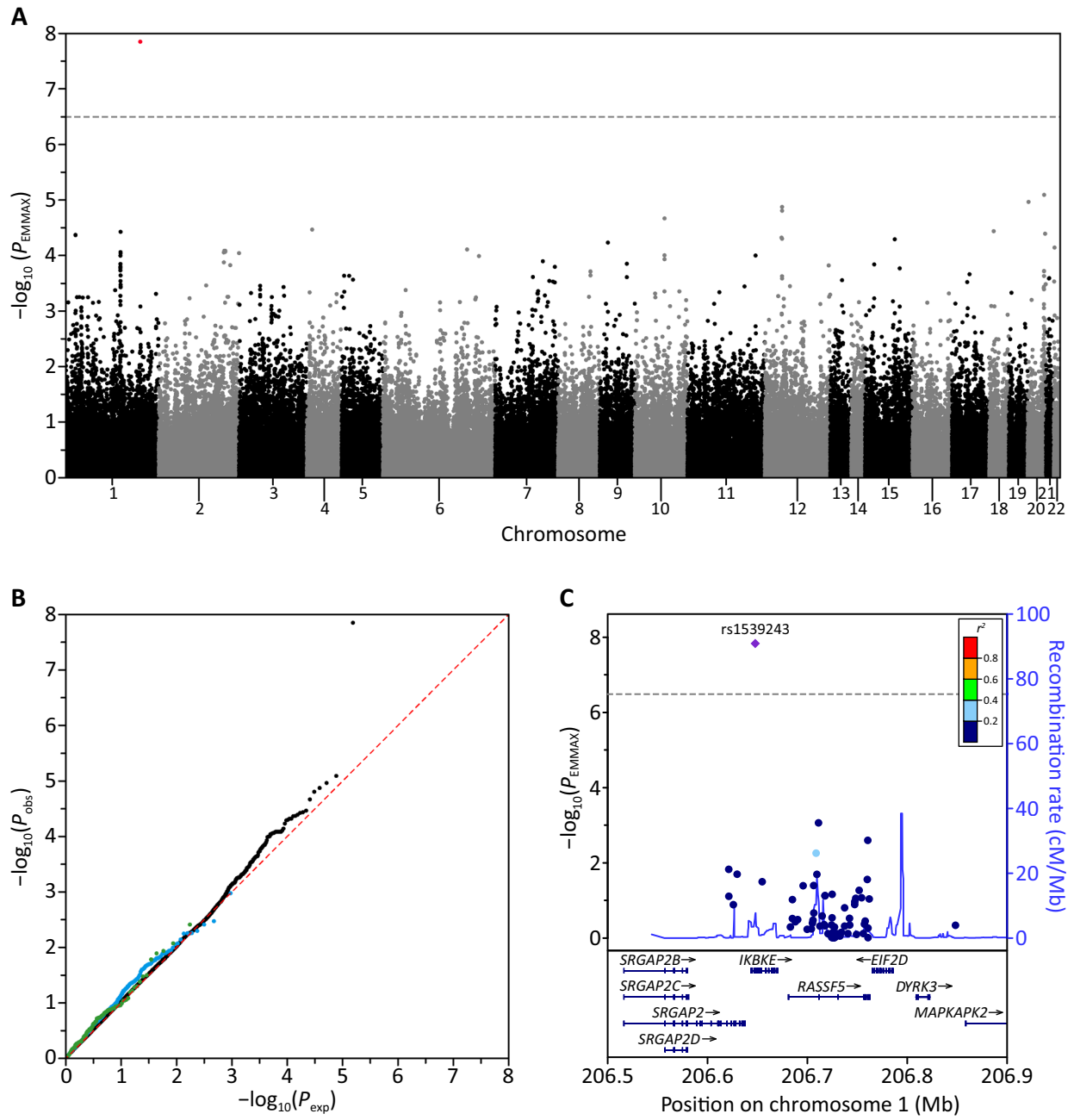

Fig. 5 *EMMAX* association signals for hunter-gatherer/farmer status. **a** Manhattan plot. **b** Quantile–quantile plot. These plots consider the 153,798 autosomal SNPs in the genetic dataset. In **a**, the gray horizontal dashed line indicates the Bonferroni $\alpha = 5\%$ significance level. In **b**, the identity line appears in red. The 173 SNPs identified by Wood et al. (2014) as significantly associated with standing-height variation in Europeans are plotted separately in green, and the 949 SNPs listed in the Cardio-MetaboChip manifest (Voight et al. 2012) as associated with standing-height variation are plotted separately in blue. **c** LocusZoom plot (Pruim et al. 2010) of the genomic region surrounding the single marginally significant SNP (rs1539243, $P_{\mathrm{EMMAX}} = 1.46 \times 10^{-8}$). Top, $-\log_{10}(P)$ values at individual SNPs, colored by their correlation (r^2) with the most significant SNP in the region (rs1539243, purple diamond), and HapMap Phase 2 recombination rates (The International HapMap Consortium 2007) (blue line). Bottom, gene locations in UCSC database release hg19 (Meyer et al. 2013)

been found to be associated with standing height (Wood et al. 2014) and gait (Adams et al. 2016) in Europeans and sitting height in Hispanics (Comuzzie et al. 2012). This association might reflect the co-adaptation of body size and cardiovascular processes in hunter-gatherers to increase endurance (Ferretti et al. 1993; Ghesquiere and Nkiama 1993) in response to hunting and gathering activities in the tropical rainforest.

Table 2 Eight genes significantly associated by *VEGAS* with hunter-gatherer/farmer status

Gene	Genomic position			Number of SNPs	P_{VEGAS}
	Chr	Begin (bp)	End (bp)		
ZNF32	10	44,124,264	44,170,147	24	$<10^{-7}$
PRKG1[a]	10	52,750,910	54,058,110	78	$<10^{-7}$
CACNA1C	12	2,162,415	2,807,115	76	1.00×10^{-7}
SMIM2	13	44,684,684	44,728,211	15	$<10^{-7}$
GPC6	13	93,879,077	95,060,273	111	1.30×10^{-6}
DLGAP1	18	3,496,029	4,455,266	165	$<10^{-7}$
MACROD2	20	13,976,145	16,033,841	183	1.20×10^{-6}
GP1BB	22	19,704,742	19,712,297	6	$<10^{-7}$

Gene names in bold were also significantly associated with height-related traits (Table 3)

Genes are ordered by chromosome and position

[a]Associated with standing-height variation in Europeans (Wood et al. 2014)

Association analyses of height-related traits

Despite our modest sample size, we identified genes associated with hunter-gatherer/farmer status that relate to skeletal growth and patterning as well as immunity and metabolism, traits that are compatible with evolutionary hypotheses for the contemporary Central African pygmy phenotype. We next sought to identify genomic regions containing SNPs significantly correlated with variation in the height-related traits that differed significantly between the hunter-gatherers and farmers in our sample (Fig. 2a–d).

Single-marker association tests

We performed per-SNP association tests with *EMMAX* to identify genomic regions harboring genes that contribute to variation in standing and sitting height and their ratio as well as subischial leg length. These analyses used 132–159 individuals with trait and SNP data available (Table S1). We again evaluated our ability to control for genetic structure using λ, comparing for each trait association tests performed with only sex as a covariate to those performed with sex and hunter-gatherer/farmer status, ethnic categorization (see "Methods"), or the first two MDS dimensions included as covariates (Figure S4B–S4E). We again chose the covariates that provided the λ closest to one: sex only [standing ($\lambda=0.964$) and sitting ($\lambda=0.928$) height], or sex and hunter-gatherer/farmer status [sitting/standing height ratio ($\lambda=0.912$) and subischial leg length ($\lambda=0.926$)]. Values of $\lambda<1$ obtained here likely reflect the limitations of our modest sample.

No SNPs were associated with standing (Figure S5) or sitting (Figure S6) height or subischial leg length (Figure S7) after Bonferroni correction, likely reflecting insufficient power with our modest sample. We did, however, identify one SNP on chromosome 3p (rs13097517, $P_{EMMAX}=6.35\times10^{-7}$; Fig. 6c) marginally associated with sitting/standing height ratio after Bonferroni correction ($\alpha=10\%$; Fig. 6a).

Multi-marker gene-based association tests

Gene-based association tests using *VEGAS* identified 24 genes associated with height-related traits after Bonferroni correction ($\alpha=5\%$; Table 3): 14 with standing height, 13 with sitting height, eight with sitting/standing height ratio, and ten with subischial leg length. Of these 24 genes, five were associated with all four traits, of which four were also significantly associated with hunter-gatherer/farmer status (Table 2). Ten were associated with more than one trait, and the remaining 14 genes were associated with only a single trait: six with standing height, six with sitting height, and two with subischial leg length.

Patterns of association among the ten genes associated with more than one trait potentially provide insights into the skeletal components and developmental pathways in which they function. For example, two of the ten are associated with sitting/standing ratio and subischial leg length (*DPYD* and *ERC2*), suggesting that their associations are driven mostly by leg length differences between hunter-gatherers and farmers (Fig. 2b). Two others are associated with standing height and sitting height (*MAGI2* and *FOXN3*), suggesting their associations are driven mostly by differences in trunk length (Fig. 2c). In this view, genes associated with multiple traits that lack one or more clearly identifiable skeletal components in common potentially encode proteins important in general skeletal development and growth.

Relevance to the Central African pygmy phenotype

Four of the five genes associated with all four height-related traits (Table 3) also show association with hunter-gatherer/farmer status (Table 2) and, as described above, have reported functions compatible with contemporary anatomical and physiological differences observed between hunter-gatherers and farmers and with evolutionary hypotheses for the pygmy phenotype: *GPC6, MACROD2, CACNA1C*, and *DLGAP1*. Among the 20 genes not overlapping those associated with hunter-gatherer/farmer status, DPYD (MIM 612779) was associated with subischial leg length and sitting/standing height ratio, and its disruption has been previously found to cause craniofacial and skeletal abnormalities (van Kuilenburg et al. 2009). Similarly, FOXN3 (MIM 602628) was associated with standing and sitting height here, and its disruption has been found to cause postnatal growth retardation and craniofacial defects in mice

Table 3 Twenty-four genes significantly associated by *VEGAS* with the height-related traits

Gene	Genomic position			Number of SNPs	P_{VEGAS}[d]			
	Chr	Begin (bp)	End (bp)		Standing height	Sitting height	Sit/stand height ratio	Subischial leg length
NRG1	8	31,497,267	32,622,558	87	$\mathbf{<10^{-7}}$	$\mathbf{2.00\times10^{-6}}$	$\mathbf{<10^{-7}}$	$\mathbf{<10^{-7}}$
CACNA1C	12	2,162,415	2,807,115	76	$\mathbf{<10^{-7}}$	$\mathbf{<10^{-7}}$	$\mathbf{<10^{-7}}$	$\mathbf{3.00\times10^{-7}}$
GPC6	13	93,879,077	95,060,273	111	$\mathbf{1.00\times10^{-6}}$	$\mathbf{<10^{-7}}$	$\mathbf{<10^{-7}}$	$\mathbf{<10^{-7}}$
DLGAP1	18	3,496,029	4,455,266	165	$\mathbf{<10^{-7}}$	$\mathbf{<10^{-7}}$	$\mathbf{<10^{-7}}$	$\mathbf{<10^{-7}}$
MACROD2	20	13,976,145	16,033,841	183	$\mathbf{<10^{-7}}$	$\mathbf{<10^{-7}}$	$\mathbf{<10^{-7}}$	$\mathbf{3.00\times10^{-7}}$
DSCAM[a]	21	41,384,342	42,219,039	72	$\mathbf{<10^{-7}}$	0.0015	$\mathbf{<10^{-7}}$	$\mathbf{<10^{-7}}$
DPYD	1	97,543,299	98,386,615	30	0.0963	3.40×10^{-5}	$\mathbf{<10^{-7}}$	$\mathbf{<10^{-7}}$
ERC2	3	55,542,335	56,502,391	36	0.0030	0.0202	$\mathbf{2.00\times10^{-7}}$	$\mathbf{8.00\times10^{-7}}$
MAGI2[b]	7	77,646,373	79,082,890	132	$\mathbf{<10^{-7}}$	$\mathbf{<10^{-7}}$	0.0086	0.0021
FOXN3	14	89,622,515	90,085,494	63	$\mathbf{<10^{-7}}$	$\mathbf{<10^{-7}}$	7.82×10^{-4}	1.83×10^{-4}
NPPA	1	11,900,375	11,907,673	618	9.30×10^{-5}	$\mathbf{<10^{-7}}$	0.0016	0.0011
DISC1	1	231,664,398	232,177,019	26	$\mathbf{<10^{-7}}$	0.0100	0.0022	1.18×10^{-4}
LMCD1	3	8,262,833	8,543,344	26	0.0594	$\mathbf{<10^{-7}}$	0.0076	0.1408
IL12A	3	159,631,189	159,924,879	16	0.0021	$\mathbf{<10^{-7}}$	0.1185	0.0542
ADIPOQ	3	186,560,462	186,576,252	32	$\mathbf{<10^{-7}}$	0.0049	0.0041	0.0011
XXYLT1	3	194,789,012	194,991,895	18	$\mathbf{<10^{-7}}$	6.70×10^{-5}	0.1610	0.0033
LNX1	4	54,326,436	54,457,753	33	$\mathbf{<10^{-7}}$	4.40×10^{-5}	0.0179	3.21×10^{-4}
EGFLAM	5	38,258,510	38,465,582	27	$\mathbf{<10^{-7}}$	8.80×10^{-5}	0.0014	5.80×10^{-6}
PACRG	6	163,148,163	163,736,524	48	0.0150	$\mathbf{<10^{-7}}$	3.70×10^{-6}	0.0012
RPS6KA2	6	166,822,853	167,275,771	51	$\mathbf{<10^{-7}}$	0.0020	1.97×10^{-4}	4.45×10^{-4}
SHANK2[c]	11	70,313,960	70,935,808	27	2.00×10^{-5}	0.0123	0.0027	$\mathbf{<10^{-7}}$
CCDC169	13	36,742,344	36,871,992	8	0.0112	0.4990	1.20×10^{-4}	$\mathbf{<10^{-7}}$
TNFRSF6B	20	62,289,162	62,330,051	24	0.0067	$\mathbf{<10^{-7}}$	3.00×10^{-5}	0.0194
COL18A1	21	46,825,096	46,933,634	12	0.1350	$\mathbf{<10^{-7}}$	0.0018	0.5910

Gene names in bold were significantly associated with hunter-gatherer/farmer status (Table 2)

Genes are ordered first by the number of traits with which they are associated and then by chromosome and position

[a]Associated with standing-height and arm-length variation in Hispanics (Comuzzie et al. 2012)

[b]Associated with standing-height variation in Europeans (Kim et al. 2010)

[c]Within 1 Mb of genomic regions identified by Perry et al. (2014a) as associated with the short stature phenotype of East Central African Batwa hunter-gatherers

[d]*VEGAS* permutation *P* values in bold are significant after Bonferroni correction ($\alpha = 5\%$)

(Samaan et al. 2010). Their associations, therefore, accord with observed reductions in postnatal growth rates in hunter-gatherers relative to farmers (Hagino et al. 2013; Merimee et al. 1987; Rozzi et al. 2015) as well as hunter-gatherer/farmer craniofacial dissimilarities (Froment 2014; Hiernaux 1977; Ramirez Rozzi and Sardi 2010), suggesting they might also have wider skeletal development functions that contribute to stature in humans.

The single SNP marginally associated with sitting/standing height ratio, rs13097517, lies in an intron of *ERC2* (MIM 617250; Fig. 6c), a regulator of neurotransmitter release (Ko et al. 2003) that was also identified in our gene-based tests (Table 3). Exploration of this SNP with the annotation program *GWAS3D* (Li et al. 2013) places it within a highly conserved genomic region with four active enhancer marks (H3K4me1, H3K27ac, P300, DHS) and sequence homology to the binding motif of transcription factor STAT5A, with a significant difference in binding affinity predicted between alleles ($P = 0.027$). The signal transducer and activator of transcription (STAT) family regulates many aspects of growth, survival, and differentiation (Darnell 1997), and includes STAT5A and its paralog STAT5B, an essential target of HGH and a critical regulator of production of IGF1 (Udy et al. 1997), a key factor in human height (Waldman and Chia 2013). This putative enhancer region lies ~ 600 kb upstream of *WNT5A*, a signaling ligand whose loss-of-function causes craniofacial and skeletal abnormalities that include shortened limbs (Oishi et al. 2003). Thus, this association might reflect the influence of rs13097517 on the binding affinity of STAT5A at this putative enhancer that modifies the expression of *WNT5A*, potentially contributing to the differential skull morphology (Froment 2014; Ramirez

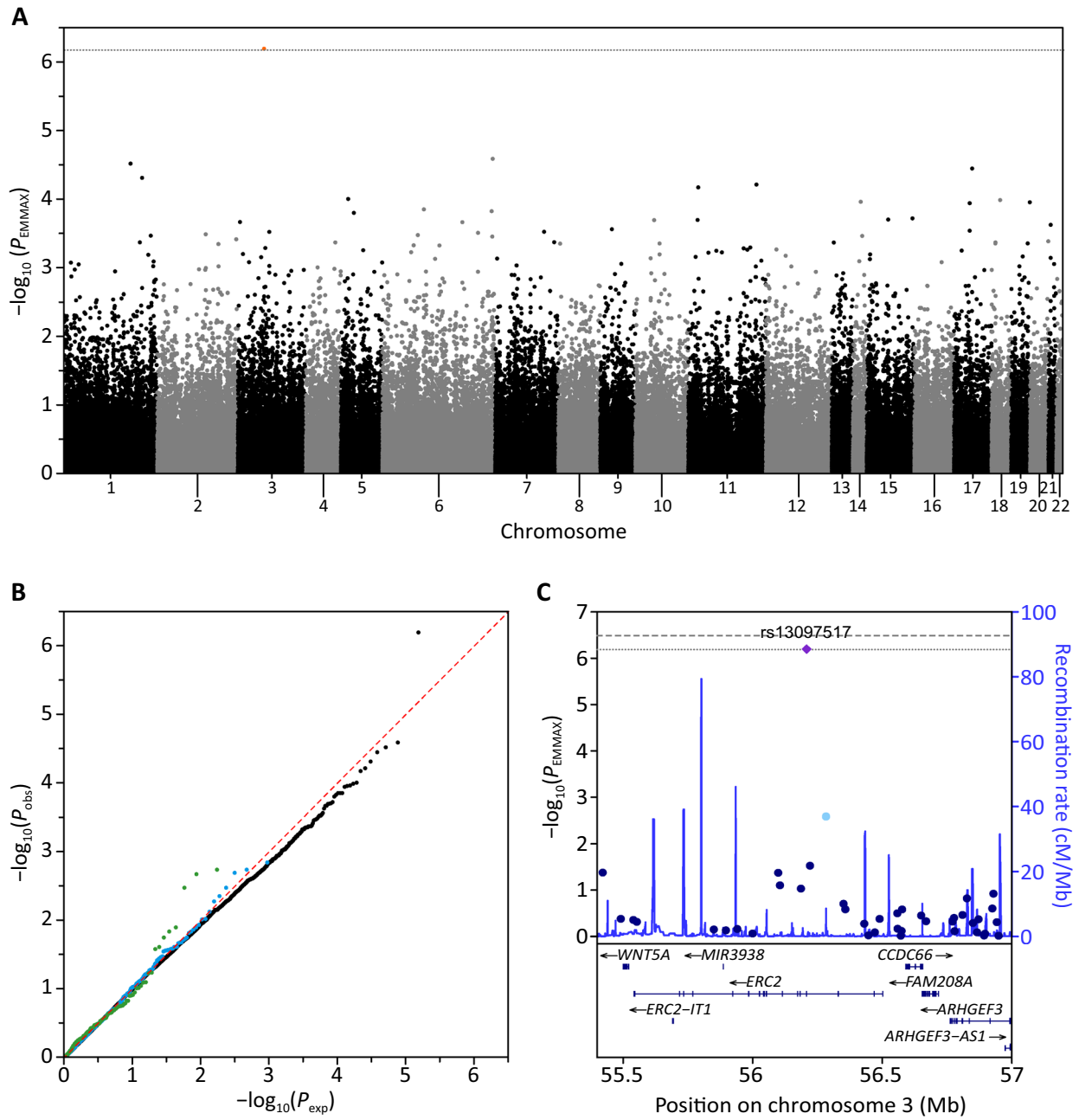

Fig. 6 *EMMAX* association signals for adult sitting/standing height ratio. **a** Manhattan plot. **b** Quantile–quantile plot. These plots consider the 153,798 autosomal SNPs in the genetic dataset, examining only the 132 individuals with measurements for both standing and sitting height (Table S1). **c** LocusZoom plot (Pruim et al. 2010) of the genomic region surrounding the single marginally significant SNP (rs13097517, $P_{\mathrm{EMMAX}} = 6.35 \times 10^{-7}$). In **a**, the gray horizontal dotted line indicates the Bonferroni $\alpha = 10\%$ significance level. The figure design follows Fig. 5

Rozzi and Sardi 2010) and leg and forearm lengths (Froment 2014; Hiernaux 1977) of hunter-gatherers compared with farmers.

The potential biological bases of associations observed between the other genes and the height-related traits remain enigmatic. Although MAGI2 (MIM 606382) has previously been associated with standing height in East Asians (Kim et al. 2010) and is associated with sitting height here (Table 3), its reported roles in renal development and function (Balbas et al. 2014) do not suggest a role in human anthropometry. Similarly, DSCAM (MIM 602523) has previously been associated with standing height and arm span in Hispanics (Comuzzie et al. 2012), and with standing height, sitting/standing height ratio, and leg length here (Table 3),

contrasting with its reported role in nervous system development (Hattori et al. 2007). Moreover, although the gene encoding SHANK2 (MIM 603290) is located within a genomic region previously associated with the pygmy phenotype in East Central African Batwa hunter-gatherers (Perry et al. 2014a) and is reported to influence serum levels of estradiol (Comuzzie et al. 2012)—a regulator of female development and metabolism (Lopez and Tena-Sempere 2016)—its reported roles in neurodevelopment (Schmeisser et al. 2012) do not give insight into its suggested role in leg length determination here (Table 3). Although associations observed with these genes might reflect the regulatory effects of mutations in enhancer elements present within their transcribed regions acting on nearby genes that have functions in human anthropometry, they might also reflect unknown functions of these genes in biological pathways and processes related to human body size.

Effect of European standing-height-associated SNPs in Central Africans

We did not observe the 949 European standing-height-associated SNPs present on the Cardio-MetaboChip (Voight et al. 2012) and polymorphic in our genetic dataset to be individually significantly associated with either hunter-gatherer/farmer status or variation in four height-related traits. However, in our gene-based tests, we did identify three genes previously associated with standing or sitting height in non-Africans (Comuzzie et al. 2012; Kim et al. 2010; Wood et al. 2014): *PRKG1* was associated with hunter-gatherer/farmer status (Table 2), *DSCAM* with standing height, sitting/standing height ratio, and subischial leg length (Table 3), and *MAGI2* with standing and sitting height (Table 3). These observations are compatible with European standing-height SNPs contributing to height determination in Central Africans, but with effect sizes too small to individually reach significance in our modest sample. Indeed, when we explore the *P* values for the 173 standing-height SNPs of Wood et al. (2014) and the 949 standing-height SNPs of Lango Allen et al. (2010), several interesting patterns emerge.

Neither the Wood et al. SNPs nor the Lango-Allen et al. SNPs departed from the uniform expectation for the *P* value distribution to a greater extent than the non-standing-height-associated SNPs included on the Cardio-MetaboChip with hunter-gatherer/farmer status (Fig. 5b) and subischial leg length (Figure S7B). However, the Wood et al. and Lango-Allen et al. SNPs do depart to a greater extent with sitting height (Figure S6B). With standing height, the Lango-Allen et al. SNPs depart to a greater extent and the Wood et al. SNPs generally do not (Figure S5B). With sitting/standing height ratio, the pattern is reversed for these two SNP sets (Fig. 6b). These patterns suggest a scenario in which European standing-height SNPs—or functional variants in strong LD with them in Central Africans—contribute to standing and sitting-height determination in Central Africans, but not to determination of leg length or categorization of hunter-gatherers and farmers.

Discussion

We have performed a joint investigation of the variability and genetic architecture of components of adult standing height in Central African hunter-gatherers and their neighboring farmers. Our findings accord with prior anthropometric observations (Froment 2014; Hiernaux 1977) that although hunter-gatherer body size is generally reduced relative to farmers, leg lengths in hunter-gatherers are significantly shorter relative to trunk length. Our results provide further support for an appreciable genetic component to the determination of body size differences between hunter-gatherers and farmers, as implied by correlations of anthropometric measures and farmer genetic admixture that replicate those reported previously for standing height only (Becker et al. 2011; Jarvis et al. 2012; Perry et al. 2014a).

Although our sample size is modest compared with traditional GWAS, our single- and multi-marker association analyses identified one genomic region and eight genes associated with hunter-gatherer/farmer status after Bonferroni correction. However, we identified only a single region marginally associated with just one of the height traits we considered (Fig. 6). Given that for each trait, some SNPs had notably lower *P* values than most SNPs tested (Figures S5–S7), the paucity of associations likely reflects lower power with smaller sample sizes (132–159) than were available for hunter-gatherer/farmer status (406). Nevertheless, gene-based multi-marker tests identified 8–14 genes enriched for SNPs exhibiting association with standing height (14), sitting height (13), sitting/standing height ratio (8), and subischial leg length (10). Quantile–quantile plots (Figure S4) and genomic inflation factors (λ) indicate that our approach appropriately corrected for structure and relatedness (Fig. 3).

Our observations of higher sitting/standing height ratios and smaller subischial leg lengths in hunter-gatherers compared with farmers (Fig. 2) potentially reflect the general nonlinear allometric scaling relationship between body size and proportions in humans (Hiernaux 1974; Huxley 1932), whereby shorter individuals tend to have shorter legs and higher sitting/standing height ratios than taller individuals. They are compatible with the study of Shea and Bailey (1996), who found body proportions in Efe hunter-gatherers to be consistent with allometric scaling relationships seen in Central African farmers. Six of the eight genes found to be significantly associated with sitting/standing height ratio were also significantly associated with standing height

(Table 3), suggesting that observed differences in body proportions between hunter-gatherers and farmers might be due in part to genetic factors that contribute both to the generally reduced body size of hunter-gatherers and to allometrically related body measurements.

We have identified novel genes and genomic regions associated with hunter-gatherer/farmer status and standing-height variation in Central Africans. Note that unlike Perry et al. (2014a), who did not uncover significant genomic regions in a similarly sized dataset of hunter-gatherers and farmers, we used association approaches rather than an admixture mapping approach that identifies genomic regions differing in inferred ancestry between hunter-gatherers and farmers. Considered together with our identification of genes and genomic regions associated with hunter-gatherer/farmer status and variation in four height-related traits, this observation suggests that cross-population GWAS are informative for exploring the genetic basis of human phenotypes that vary naturally across Central African populations.

Consistent with prior studies that investigated contributions of > 100 non-African standing-height-associated SNPs to standing-height variation among 5–230 Central African hunter-gatherers and farmers (Jarvis et al. 2012; Lachance et al. 2012; Perry et al. 2014a), none of the 949 SNPs listed in the Cardio-MetaboChip manifest (Voight et al. 2012) as associated with standing height were individually associated with height-related variation in our Central African sample. However, *P* values for these SNPs did depart from the uniform expectation to a greater extent than non-standing-height SNPs with standing and sitting height and their ratio, indicating that they or functional variants in strong LD with them might contribute to height determination in Central Africans, but with effect sizes too small to be found in our modest sample.

Insights from specific genes

Some of the genes and genomic regions identified potentially provide novel insight into the genetic basis of anatomical and physiological features of the Central African pygmy phenotype. First, a number of genes associated with the height-related traits have been reported to cause both skeletal and craniofacial abnormalities when perturbed: *MACROD2* (Maas et al. 2007), *GP1BB* (Budarf et al. 1995), *CACNA1C* (Splawski et al. 2005), *DPYD* (van Kuilenburg et al. 2009), and *FOXN3* (Samaan et al. 2010). It is, therefore, possible that craniofacial differences between hunter-gatherers and farmers (Froment 2014; Ramirez Rozzi and Sardi 2010) might in part reflect coevolution of craniofacial traits with stature.

Second, the sole SNP marginally associated with sitting/standing height ratio (rs13097517) lies in a putative STAT5A enhancer element ~ 619 kb upstream of *WNT5A*, a signaling ligand whose abrogation causes craniofacial and skeletal abnormalities including shortened limbs (Oishi et al. 2003). Although a mutational screen of the *STAT5A* gene found no evidence of association with hunter-gatherer/farmer status (Becker et al. 2013) and association of this putative enhancer with *WNT5A* expression remains unexamined, our observation is compatible with a partial contribution to hunter-gatherer short stature from perturbation of *WNT5A* expression. Intriguingly, WNT5A mediates the actions of growth factor receptor GPC6 (Yiu et al. 2011), an important determinant of long-bone growth (Bassett et al. 2006; Campos-Xavier et al. 2009) that was associated with hunter-gatherer/farmer status (Table 2) and all four height-related traits (Table 3). Perturbation of the GPC6/WNT5A pathway might thus be a general feature of hunter-gatherer short stature.

Third, two genes associated with hunter-gatherer/farmer status (Table 2) as well as all four height-related traits (Table 3) have putative roles in IGF1-signaling. *MACROD2* is implicated in postnatal growth deficiency (Maas et al. 2007) and modulation of blood-borne IGF1 levels (Comuzzie et al. 2012), and *DLGAP1* is implicated in fat and glucose metabolism (Aulchenko et al. 2007), reproductive development (Perry et al. 2014b), and modulation of blood-borne levels of IGF1-signaling component IGFBP3 (Comuzzie et al. 2012). Our identification of genes with roles in growth or metabolic pathways as well as in modulating levels of proteins with reported roles in IGF1-signaling accords with associations reported between SNPs in the *IGF1* and the growth hormone receptor genes and hunter-gatherer/farmer status (Becker et al. 2013), highlighting an intriguing possibility. The reduced body size of contemporary hunter-gatherers (Froment 1993; Hiernaux 1974) might be an evolutionary byproduct of adaptive processes acting on metabolic and developmental pathways to support a hunting-gathering lifestyle in the tropical rainforest (Bailey et al. 1989; Bailey 1991; Diamond 1991), processes that acted upon proteins that function both in IGF1-signaling and in growth and metabolism. In this view, alterations in metabolic processes (Merimee et al. 1972; Rimoin et al. 1969) and postnatal growth rates (Hagino et al. 2013; Merimee et al. 1987; Rozzi et al. 2015) in response to diet and food insecurities might have led to perturbations in IGF1-signaling (Geffner et al. 1995; Hattori et al. 1996) and an inadvertent decrease in overall stature.

Finally, two of eight genes associated with hunter-gatherer/farmer status and one gene associated with sitting height have functions in immune rather than developmental processes: *IKBKE* is an essential regulator of antiviral-signaling pathways (Sharma et al. 2003), disruption of *CACNA1C* leads to immunodeficiency (Splawski et al. 2005), and *IL12A* functions in innate and acquired immune responses (Wolf et al. 1994). Our findings, therefore, add to evidence that

adaptation in both immune- and auxologic-related processes has contributed to evolution of the contemporary pygmy phenotype (Jarvis et al. 2012; Lachance et al. 2012; Perry et al. 2014a).

Genetics of height-related traits in Central Africans and Europeans

The lack of association of the European standing-height SNPs with standing-height variation in our Central African sample is compatible with ancestry-related differences in the genetic architecture of human height in Europeans and Central Africans. Standing-height determination in Central Africans might be expected to involve both genetic factors private to Central Africans as well as those shared with Europeans. The latter would have arisen early in human evolution before migration of the ancestors of Europeans out of Africa, and would have later experienced divergent evolutionary trajectories in Central Africans and Europeans, due possibly to genetic drift or to dissimilar selective pressures leading to differences in their contemporary relationship with standing height. Indeed, Martin et al. (2017) recently argued that standing-height predictions based on European standing-height SNPs from Wood et al. (2014) are considerably shorter than the population averages for Africans. This observation highlights the biological complexity of human height as well as the imperfect transferability of European standing-height SNPs across ancestral backgrounds, potentially contributing to the absence of associations for these SNPs in our Central African sample.

Note that for some non-African SNPs, absence of significant association might also be a result of discordance in LD patterns between Europeans and Central Africans disrupting the relationship between the genotyped SNP and the trait-determining variant (Shriner et al. 2009). Indeed, for standing height, the trait-increasing allele identified here agrees with those identified by Lango Allen et al. (2010) and Wood et al. (2014) for only 47.8 and 50.3% of the European standing-height SNPs, respectively, although they are concordant for 89.6% of SNPs between the two European studies. The greater discordance between trait-increasing alleles in this Central African study and the two European studies possibly results from the same functional variants affecting standing height in Central Africans and Europeans, but with different LD patterns that weaken correlations of those variants with the variants tested.

Despite an absence of SNP-wise associations with non-African standing-height SNPs, gene-based association tests did identify three non-African standing-height genes as associated with height-related traits in Central Africans: *MAGI2* (Kim et al. 2010) and *DSCAM* (Comuzzie et al. 2012), and one associated with hunter-gatherer/farmer status, *PRKG1* (Wood et al. 2014). Thus, although non-African standing-height SNPs did not individually reach significance, our multi-marker identification of *MAGI2, DSCAM*, and *PRKG1* suggests that this outcome might result at least in part from insufficient power with our modest sample. Indeed, though not statistically significant, another non-African standing-height gene, *GPC5* (Comuzzie et al. 2012; Lango Allen et al. 2010; Wood et al. 2014), is among the top 20 genes exhibiting an association with standing height ($P_{VEGAS} = 1.60 \times 10^{-5}$), sitting height ($P_{VEGAS} = 2.30 \times 10^{-5}$), and subischial leg length ($P_{VEGAS} = 3.48 \times 10^{-5}$) in Central Africans. These associations might, however, be driven by Africa-specific height-influencing functional variants that arose separately from those identified previously in Europeans, but that nonetheless act through a common set of genes to influence human stature.

Limitations

Our study of height-related differences between Central African hunter-gatherers and farmers has used the SNPs of the Illumina Cardio-MetaboChip, a genotyping array designed to facilitate replication and fine-mapping of associations identified by large consortia studying anthropometric and cardiovascular traits in individuals of mostly recent European ancestry (Voight et al. 2012). Although this array has enabled us to investigate the role of SNPs associated with standing-height variation in Europeans (Lango Allen et al. 2010; Wood et al. 2014) in height-related variation in Central Africans, its noncomprehensive genomic distribution has limited the proportion of the genome that could be interrogated. Consequently, contributing genes and genetic variants that lie outside the regions represented by this array will have gone undetected. Moreover, limiting our analysis to SNPs included primarily for informativeness in non-Africans has potentially limited our sensitivity to detect African-specific genetic variants contributing to height-related traits in Central Africans (Conrad et al. 2006; Huang et al. 2011). Studies with denser and more uniform genomic coverage will be required to provide a more complete genome-wide perspective on the evolution and determination of the contemporary Central African pygmy phenotype.

Our results disagree in part with prior observations on BMI differences between hunter-gatherers and farmers, particularly for females (Froment 2014). This inconsistency can be explained in part by the inclusion of only rural individuals in our sample, whereas prior studies collected both rural and urban data. Body weight differs between rural and urban settings, as individuals living in urban areas are generally heavier than those in rural areas. This phenomenon affects females to a greater extent than males, and has increased over time (Abubakari et al. 2008; Fezeu et al. 2008). These urban–rural and temporal differences would be expected to influence BMI patterns reported in prior studies, particularly

for females, and to lead to apparent differences with patterns in our more recent rural-only sample. Larger studies of BMI in Central African hunter-gatherer and farming populations that account for rural and urban differences in body weight will be required to clarify the sex-specific BMI patterns in these populations.

Conclusions

Despite using a modest sample size, we have identified genomic regions and genes that are biologically meaningful in relation to both the traits themselves and to Central African hunter-gatherer and farmer phenotypic differences. Our findings highlight the need for joint analyses considering anthropometric and physiological measures in large samples of hunter-gatherers and farmers to facilitate understanding of the biological basis of the contemporary Central African pygmy phenotype. Larger-scale studies of short stature populations have potential to shed light on relative contributions of shared and distinct pathways in the development and maintenance of short stature. In particular, studying populations of different stature living in different ecologic environments will provide a greater understanding of adaptive processes underlying the variation in adult body size observed across contemporary human populations.

Methods

Samples

Our initial sample included 558 individuals from 20 Central African populations collected during fieldwork conducted between 2005 and 2009 (Fig. 1; Table 1). At each sampling site, authors (PV, BH, SLB, AF, and NB) with extensive biological and cultural anthropological fieldwork experience in the investigated communities categorized populations as hunter-gatherer ("Pygmy") or farmer ("non-Pygmy") based on cultural criteria that do not include adult standing height (Bahuchet 1993a, 2012; Fürniss and Bahuchet 1995; Hewlett 1996, 2014; Joiris 2003; Le Bomin and Mbot 2012; Turnbull 1965; Verdu 2009). Here, a community was categorized as hunter-gatherer "Pygmy" when it: (1) is recognized by outsiders (i.e., farmer "non-Pygmies") as specialized in forest activities such as hunting and gathering as well as having medical and magical knowledge of the rainforest; (2) shares complex socio-economic relationships with specific neighboring outsiders, such as exchanging forest products for iron tools; (3) is designated as "Pygmy" (or its literal local translation) or at least as "other than self" by neighboring outsiders; (4) distinguishes itself as a community with a different ethno-name from neighboring outsiders, regardless of languages spoken; (5) has differing musical practices and instruments recognized as such by neighboring outsiders.

For each individual, DNA was extracted from either whole blood buffy coats with DNeasy Blood & Tissue spin-column kits (Qiagen, Valencia, CA) or saliva collected using Oragene kits (DNA Genotek Inc., Kanata, ON), following the manufacturer protocol. For a subset of the individuals from seven hunter-gatherer [Baka (Center), Baka (East), Bezan (South), Bongo (Center), Bongo (East), Bongo (South), Koya] and three farming (Nzime, Tikar, Bangando) populations (Table S1), adult standing and sitting height were measured with a height gauge and weight with a standard weigh scale following standard anthropometric procedures (Weiner and Lourie 1981). Standing height was measured with height gauges of the same brand and model at all sites. We used electronic body weight scales with 100 g precision at all sites. Individuals in the countryside in Central Africa are lightly clothed and often do not wear shoes. Although shoes were removed during weighing, people remained clothed. To account for clothing in recorded weights, we subtracted 500 g from the measured weight as our previous experience showed that this was roughly the average weight of clothes worn by people in the communities from which our sample was collected.

Subischial leg length was calculated as standing height minus sitting height, and BMI as body weight divided by the square of standing height (kg/m^2). Accurate age data were unavailable as most of the communities do not track birth dates. Although we cannot rule out a confounding effect of osteoporotic age-related shrinking, we expect it to be minimal as elderly individuals were not considered; elderliness was determined at the time of collection via visual appearance and defined by a general lack of teeth, a hunched appearance, or reduced mobility in the absence of disease or a self-reported handicap. The ages of individuals included in our sample are estimated to be between 18 and 65 years-old.

Genotyping and quality control

Genotypes of 558 individuals (Table 1) were obtained at the University of Michigan Medical School DNA Sequencing Core (Ann Arbor, MI) using the Illumina Cardio-MetaboChip (Voight et al. 2012). Control samples (18 replicates of one individual) were used to verify genotype consistency across array batches. We focus on 196,091 of the 196,725 SNPs with independently verified genomic positions. The Illumina GenomeStudio Genotyping Module was used for genotype-calling (v.1.0; Illumina, San Diego, CA).

Following quality control procedures conducted at the genotype-calling level (Stage 1, Figure S8), the preliminary dataset contained 192,903 autosomal SNPs polymorphic in a sample of 543 hunter-gatherers and farmers. Next, we conducted population-level quality control (Stage 2, Figure

S8), creating an initial dataset of 154,106 autosomal SNPs polymorphic in the 543 individuals. Hardy–Weinberg equilibrium was evaluated separately in each population using Yates-corrected chi-squared tests (Weir 1996), and we used the same exclusion criteria as in Pemberton et al. (2010).

Hunter-gatherers and neighboring farmers interact socially and economically on a daily basis (Bahuchet and Guillaume 1982; Bahuchet 2012; Hewlett 1996; Joiris 2003). Thus, individuals from neighboring communities might have been present at the time of sampling, and some might have been wrongly categorized as hunter-gatherer or farmer. To search for categorization errors during sample collection, we performed MDS analysis (see below) in our initial dataset. We identified 11 individuals initially categorized as hunter-gatherer and five categorized as farmer who did not cluster genetically with other individuals sharing the same categorization. To be conservative, we removed these 16 putative misclassified individuals and repeated the population-level quality control procedures (stage 3, Figure S8) to create a dataset containing 154,029 autosomal SNPs polymorphic in a sample of 527 hunter-gatherers and farmers ("527RELAT" henceforth; Table 1); all individuals possessed genotypes at ≥ 94.9% of SNPs.

Pairwise relatedness among all 527 individuals in the 527RELAT dataset was evaluated using identity-by-state allele-sharing and the likelihood approach of *RELPAIR* (v.2.0.1) (Boehnke and Cox 1997; Epstein et al. 2000) following the methods of Pemberton et al. (2010) restricted to sets of 9,999 SNPs. A total of 282 pairs of individuals were inferred by *RELPAIR* to be related at a level closer than first cousins: 230 intra-population and 52 inter-population pairs. All inter-population pairs involved geographically close populations (Fig. 1); 47 with the two Bezan populations (BZN and BZS), three with two Bongo populations (CBG and EBG), and two with the nearby Fang (CFG) and Ngumba (NGB) populations.

A dataset containing no first- or second-degree relatives was created by removing one individual from each of these 282 relative pairs. To minimize exclusions, we preferentially omitted individuals present in two or more relative pairs (either intra- or inter-population). When either individual in a pair could be removed, we removed the one with more missing data. After exclusion of 121 related individuals, many with multiple relatives in the initial dataset, we repeated the population-level quality control procedures (stage 4, Figure S8), producing an unrelated dataset with 153,798 autosomal SNPs polymorphic among 406 hunter-gatherers and farmers ("406UNRELAT" henceforth; Table 1); all individuals possessed genotypes at ≥ 94.9% of SNPs. Because of small sample sizes in the 406UNRELAT dataset, we combined the two Bezan populations (11 for BZN and 17 for BZS) into one ("Bezan" henceforth; BEZ, Table 1). These samples belong to the same Bezan ethnic group of Central Cameroon, which has census size < 400, divided between two communities ~ 50 km apart and in frequent contact (PV and AF, unpublished data).

Population-genetic analyses

We performed MDS based on ASD matrices constructed for pairs of individuals in the 406UNRELAT dataset using *asd* (v1.0; https://github.com/szpiech/asd). This program considers in the calculation for a pair only those SNPs for which neither individual was missing genotypes. We applied classic metric MDS on the ASD matrix using *cmdscale* in R (v3.3.1) (R Development Core Team 2014). To study the variation in ASD represented in the first two MDS dimensions, we calculated the Spearman correlation ρ between the Euclidean distances for all pairs of individuals and their corresponding ASD values (using *cor.test* in R here and for other correlations). To evaluate dispersion among hunter-gatherers and farmers in the MDS plot, we compared the variance among their ASD values with a one-sided F test using *var.test* in R.

For the 406UNRELAT dataset, we performed *STRUCTURE* (v2.3) unsupervised model-based clustering (Falush et al. 2003; Pritchard et al. 2000), which probabilistically assigns proportions of each individual's genotypes to each of K genetic clusters, where K is set a priori, disregarding individual population categorizations. We used the admixture model with separate Dirichlet parameters in each cluster and correlated allele frequencies, and a burn-in period of 20,000 iterations followed by 10,000 iterations.

To minimize the number of linked loci owing to the gene-centric Cardio-MetaboChip design, in the *STRUCTURE* analyses we used 40,424 SNPs with minimum spacing 18.75 kb. From these 40,424 SNPs, we created four non-overlapping panels of 10,106 SNPs each with minimum marker spacing of 75 kb. For panel n ($1 \le n \le 4$), every fourth SNP along a vector of the 40,424 SNPs was selected, starting at position $n-1$. In this vector, SNPs were numbered ordered from chromosome 1–22 and by increasing genomic positions (NCBI build 37) along chromosomes. For each panel, we computed 10 *STRUCTURE* runs for K from 2 to 4, producing 40 independent replicates for each K. We identified frequent modes among the 40 replicates using the *CLUMPP* (Jakobsson and Rosenberg 2007) *Greedy* algorithm with 1000 random permutations. For each K, all pairs of runs with symmetric similarity coefficients > 0.9 were considered to belong to the same mode. For each mode, we computed individual membership proportions averaged across runs from that mode, visualizing the most frequent mode with *DISTRUCT* (Rosenberg 2004).

To investigate the relationship between farmer admixture and each measured trait in Central Africans, we calculated the Pearson correlation coefficient r between individual

membership proportions in the blue cluster at $K=2$ and trait measurements where available.

Association analyses

Single-marker tests

Per-SNP association tests were performed using *EMMAX* (Kang et al. 2010), which implements a mixed-effect regression model (Yu et al. 2006) to account for genetic structure in a sample, incorporating variance components of random polygenic effects (Fisher 1918; Ober et al. 2001) and genetic relatedness among individuals using a pairwise Balding-Nichols kinship matrix (Balding and Nichols 1995) [Eq. 7 in Kang et al. (2010)]. We constructed the matrix for all pairs of individuals in the 406UNRELAT dataset using *EMMAX*, including in the calculation only 93,821 SNPs inferred to be in linkage equilibrium ($r^2<0.5$).

Association analyses for hunter-gatherer/farmer status were performed by labeling hunter-gatherers as "cases" and farmers as "controls." We evaluated how the inclusion of ethnic categorization and the first two MDS dimensions improved control of genetic structure in our sample compared with the inclusion of no covariates using the genomic inflation factor λ (Devlin and Roeder 1999), considering in the calculation only the 93,821 SNPs inferred to be in linkage equilibrium ($r^2<0.5$). Here, ethnic categorization represents a categorical variable where hunter-gatherers and farmers who live in the same geographic area (i.e., are members of the same population pairing) are assigned the same integer (i.e., 1, 2, 3, etc.). The results presented were obtained using the covariate that provided the value of λ closest to 1.

Association analyses were performed for each measured trait, considering only individuals with data available. To control for possible population differences and sexual dimorphism in the measured traits, in our analysis of quantitative traits, we included sex as a covariate and using λ, explored how the inclusion of hunter-gatherer status (as a binary variable), ethnic categorization, or the first two MDS dimensions improved control of genetic structure in our sample compared with inclusion of no covariates. Results presented were obtained using sex and the variable with the λ value closest to 1 as covariates.

Multi-marker gene-based tests

We performed multi-marker gene-based association tests using *VEGAS* (Liu et al. 2010). Separately for each gene, *VEGAS* computes a χ^2 test statistic from *EMMAX P* values for all its SNPs (± 50 kb from ends of the transcribed region), evaluating significance by simulation from a multivariate normal distribution with mean 0 and a covariance matrix of pairwise LD among SNPs, accounting for effects of gene size in a manner found by Liu et al. (2010) to produce comparable results to permutation. Transcribed regions were defined using UCSC gene database hg19 (Meyer et al. 2013); for genes with multiple isoforms, begin and end positions were defined as the outermost positions in the union of transcribed regions of all isoforms.

Acknowledgements The authors thank the volunteers from Cameroon, Central African Republic, Gabon, and Uganda who participated in this study. They thank Peter Chines for providing independently verified Cardio-MetaboChip SNP positions, Sen Li for providing ancestral and derived allelic states for the MetaboChip SNPs, Hyun Min Kang for *EMMAX* assistance, and Frédéric Austerlitz, Erkan Buzbas, Zachary Szpiech, and Lawrence Uricchio for useful comments and discussions. This investigation was supported by the France-Stanford Center for Interdisciplinary Studies (N. A. R.); the French Assistance Publique Hôpitaux de Paris, ATM-MNHN 'Les relations Sociétés-Natures dans le long terme' 2009–2012 and the Fondation pour la Recherche Médicale and the ANR-Blanc program 'GrowingAP' (E. H.); a University of Michigan Center for Genetics in Health and Medicine postdoctoral fellowship (T. J. P.); and a Natural Sciences and Engineering Research Council of Canada Discovery Grant (RGPIN-2015-04739; T. J. P.). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Compliance with ethical standards

Ethics approval and consent to participate The study was conducted according to ethical principles of the Declaration of Helsinki. IRB approvals were obtained from the French Ministry of Higher Education and Research (2005, renewed in 2010 and 2016), Stanford (2012) and Washington State Universities (2007), and the Universities of Manitoba (2013) and Michigan (2010). Prior to sample collection, research authorizations were obtained from the Ministry of Public Health in Cameroon (2005; renewed 2010), the Ministry of Higher Education and Research in Gabon (2006 and 2007), the National Council for Science and Technology in Uganda (2007), and the Ministry of Scientific Research in the Central African Republic (2007). Informed consent was obtained from all participants.

Conflict of interest The authors declare that they have no conflicts of interest.

Availability of data and material The Illumina Cardio-MetaboChip SNP genotype datasets for the 406 unrelated individuals analyzed are available in the European Genome-Phenome archive (EGAS00001002975) and will be accessible for population-genetic studies of demography and natural selection as well as for anthropological-genetic studies of the genetic determination of anthropometrical phenotypic features in accordance with the informed consent documents used in the recruitment of individuals who participated in this study.

References

Abubakari AR, Lauder W, Agyemang C, Jones M, Kirk A, Bhopal RS (2008) Prevalence and time trends in obesity among adult West African populations: a meta-analysis. Obes Rev 9:297–311

Adams HH, Verlinden VJ, Callisaya ML, van Duijn CM, Hofman A, Thomson R, Uitterlinden AG, Vernooij MW, van der Geest JN, Srikanth V, Ikram MA (2016) Heritability and genome-wide association analyses of human gait suggest contribution of common variants. J Gerontol A Biol Sci Med Sci 71:740–746

Aulchenko YS, Pullen J, Kloosterman WP, Yazdanpanah M, Hofman A, Vaessen N, Snijders PJ, Zubakov D, Mackay I, Olavesen M, Sidhu B, Smith VE, Carey A, Berezikov E, Uitterlinden AG, Plasterk RH, Oostra BA, van Duijn CM (2007) LPIN2 is associated with type 2 diabetes, glucose metabolism, and body composition. Diabetes 56:3020–3026

Bahuchet S (1993a) L'invention des Pygmées. Cahiers d'Etudes africaines 33:153–181

Bahuchet S (1993b) La Rencontre des Agriculteurs. Les Pygmées parmi les Peuples d'Afrique Centrale, Paris

Bahuchet S (2012) Changing language, remaining Pygmy. Hum Biol 84:11–43

Bahuchet S, Guillaume H (1982) Aka-farmer relations in the Northwest Congo Basin. Cambridge University Press, Cambridge

Bailey RC (1991) The comparative growth of Efe Pygmies and African farmers from birth to age 5 years. Ann Hum Biol 18:113–120

Bailey RC, Head G, Tenike M, Owen B, Rechtman R, Zechenter E (1989) Hunting and gathering in tropical rainforest: is it possible? Am Anthropol 91:59–82

Balbas MD, Burgess MR, Murali R, Wongvipat J, Skaggs BJ, Mundel P, Weins A, Sawyers CL (2014) MAGI-2 scaffold protein is critical for kidney barrier function. Proc Natl Acad Sci USA 111:14876–14881

Balding DJ, Nichols RA (1995) A method for quantifying differentiation between populations at multi-allelic loci and its implications for investigating identity and paternity. Genetica 96:3–12

Bassett JH, Swinhoe R, Chassande O, Samarut J, Williams GR (2006) Thyroid hormone regulates heparan sulfate proteoglycan expression in the growth plate. Endocrinology 147:295–305

Batini C, Lopes J, Behar DM, Calafell F, Jorde LB, van der Veen L, Quintana-Murci L, Spedini G, Destro-Bisol G, Comas D (2011) Insights into the demographic history of African Pygmies from complete mitochondrial genomes. Mol Biol Evol 28:1099–1110

Becker NS, Verdu P, Froment A, Le Bomin S, Pagezy H, Bahuchet S, Heyer E (2011) Indirect evidence for the genetic determination of short stature in African Pygmies. Am J Phys Anthropol 145:390–401

Becker NS, Verdu P, Georges M, Duquesnoy P, Froment A, Amselem S, Le Bouc Y, Heyer E (2013) The role of *GHR* and *IGF1* genes in the genetic determination of African Pygmies' short stature. Eur J Hum Genet 21:653–658

Berndt SI, Gustafsson S, Magi R, Ganna A, Wheeler E, Feitosa MF, Justice AE, Monda KL, Croteau-Chonka DC, Day FR, Esko T, Fall T, Ferreira T, Gentilini D, Jackson AU, Luan J, Randall JC, Vedantam S, Willer CJ, Winkler TW et al (2013) Genome-wide meta-analysis identifies 11 new loci for anthropometric traits and provides insights into genetic architecture. Nat Genet 45:501–512

Boehnke M, Cox NJ (1997) Accurate inference of relationships in sib-pair linkage studies. Am J Hum Genet 61:423–429

Bozzola M, Travaglino P, Marziliano N, Meazza C, Pagani S, Grasso M, Tauber M, Diegoli M, Pilotto A, Disabella E, Tarantino P, Brega A, Arbustini E (2009) The shortness of Pygmies is associated with severe under-expression of the growth hormone receptor. Mol Genet Metab 98:310–313

Budarf ML, Konkle BA, Ludlow LB, Michaud D, Li M, Yamashiro DJ, McDonald-McGinn D, Zackai EH, Driscoll DA (1995) Identification of a patient with Bernard–Soulier syndrome and a deletion in the DiGeorge/velo-cardio-facial chromosomal region in 22q11.2. Hum Mol Genet 4:763–766

Buschang PH, Malina RM, Little BB (1986) Linear growth of Zapotec schoolchildren: growth status and yearly velocity for leg length and sitting height. Ann Hum Biol 13:225–234

Campos-Xavier AB, Martinet D, Bateman J, Belluoccio D, Rowley L, Tan TY, Baxova A, Gustavson KH, Borochowitz ZU, Innes AM, Unger S, Beckmann JS, Mittaz L, Ballhausen D, Superti-Furga A, Savarirayan R, Bonafe L (2009) Mutations in the heparan-sulfate proteoglycan glypican 6 (GPC6) impair endochondral ossification and cause recessive omodysplasia. Am J Hum Genet 84:760–770

Carlson CS, Matise TC, North KE, Haiman CA, Fesinmeyer MD, Buyske S, Schumacher FR, Peters U, Franceschini N, Ritchie MD, Duggan DJ, Spencer KL, Dumitrescu L, Eaton CB, Thomas F, Young A, Carty C, Heiss G, Le Marchand L, Crawford DC et al (2013) Generalization and dilution of association results from European GWAS in populations of non-European ancestry: the PAGE study. PLoS Biol 11:e1001661

Carty CL, Johnson NA, Hutter CM, Reiner AP, Peters U, Tang H, Kooperberg C (2012) Genome-wide association study of body height in African Americans: the Women's Health Initiative SNP Health Association Resource (SHARe). Hum Mol Genet 21:711–720

Cavalli-Sforza LL (1986) African Pygmies. Academic Press, Orlando

Comuzzie AG, Cole SA, Laston SL, Voruganti VS, Haack K, Gibbs RA, Butte NF (2012) Novel genetic loci identified for the pathophysiology of childhood obesity in the Hispanic population. PLoS One 7:e51954

Conrad DF, Jakobsson M, Coop G, Wen X, Wall JD, Rosenberg NA, Pritchard JK (2006) A worldwide survey of haplotype variation and linkage disequilibrium in the human genome. Nat Genet 38:1251–1260

Darnell JE Jr (1997) STATs and gene regulation. Science 277:1630–1635

de Garine I (1990) Adaptation biologique et bien-être psycho-culturel. Bulletins et Mémoires de la Société d 2:151–174

de Filippo C, Bostoen K, Stoneking M, Pakendorf B (2012) Bringing together linguistic and genetic evidence to test the Bantu expansion. Proc Biol Sci 279:3256–3263

Destro-Bisol G, Donati F, Coia V, Boschi I, Verginelli F, Caglia A, Tofanelli S, Spedini G, Capelli C (2004) Variation of female and male lineages in sub-Saharan populations: the importance of sociocultural factors. Mol Biol Evol 21:1673–1682

Devlin B, Roeder K (1999) Genomic control for association studies. Biometrics 55:997–1004

Diamond JM (1991) Anthropology. Why are Pygmies small? Nature 354:111–112

Dietz WH, Marino B, Peacock NR, Bailey RC (1989) Nutritional status of Efe pygmies and Lese horticulturists. Am J Phys Anthropol 78:509–518

Epstein MP, Duren WL, Boehnke M (2000) Improved inference of relationship for pairs of individuals. Am J Hum Genet 67:1219–1231

Falush D, Stephens M, Pritchard JK (2003) Inference of population structure using multilocus genotype data: linked loci and correlated allele frequencies. Genetics 164:1567–1587

Ferretti G, Atchou G, Grassi B, Marconi C, Cerretelli P (1993) Energetics of locomotion in African Pygmoids. Eur J Appl Physiol 62:7–10

Fezeu LK, Assah FK, Balkau B, Mbanya DS, Kengne AP, Awah PK, Mbanya JC (2008) Ten-year changes in central obesity and BMI in rural and urban Cameroon. Obesity (Silver Spring) 16:1144–1147

Fisher RA (1918) The correlation between relatives on the supposition of Mendelian inheritance. Trans R Soc Edinburgh 52:399–433

Froment A (1993) Adaptation biologique et variation dans l'espèce humaine: le cas des Pygmées d'Afrique. Bulletins et Mémoires de la Société d'Anthropologie de Paris 5:417–448

Froment A (2014) Human biology and health of African rainforest inhabitants. In: Hewlett BS (ed) Hunter-gatherers of the Congo Basin: cultures, histories, and biology of African Pygmies, 1st edn. Transaction Publishers, New Brunswick, pp 117–164

Fürniss S, Bahuchet S (1995) Existe-t-il des instruments de musique pygmées ? In: Dehoux V, Fürniss S, Le Bomin S, Olivier E, Voisin F (eds) Ndroje balendro, musiques, terrains et disciplines (textes offerts à Simha Arom). Peters-Selaf, Paris, pp 87–109

Geffner ME, Bersch N, Bailey RC, Golde DW (1995) Insulin-like growth factor I resistance in immortalized T cell lines from African Efe Pygmies. J Clin Endocrinol Metab 80:3732–3738

Ghesquiere JL, Nkiama E (1993) Implications of small body size to strength and endurance of Pygmy and village children in the Ituri Forest. In: 4th World conference on human ecology, Merida, Mexico

Gudbjartsson DF, Walters GB, Thorleifsson G, Stefansson H, Halldorsson BV, Zusmanovich P, Sulem P, Thorlacius S, Gylfason A, Steinberg S, Helgadottir A, Ingason A, Steinthorsdottir V, Olafsdottir EJ, Olafsdottir GH, Jonsson T, Borch-Johnsen K, Hansen T, Andersen G, Jorgensen T et al (2008) Many sequence variants affecting diversity of adult human height. Nat Genet 40:609–615

Hagino I, Hayashi K, Kawamura K, Sato H, Yamauchi T (2013) Adolescent growth spurt and growth pattern factors related to the short stature of Pygmy hunter-gatherers of Southeast Cameroon. Ann Hum Biol 40:9–14

Hattori Y, Vera JC, Rivas CI, Bersch N, Bailey RC, Geffner ME, Golde DW (1996) Decreased insulin-like growth factor I receptor expression and function in immortalized African Pygmy T cells. J Clin Endocrinol Metab 81:2257–2263

Hattori D, Demir E, Kim HW, Viragh E, Zipursky SL, Dickson BJ (2007) Dscam diversity is essential for neuronal wiring and self-recognition. Nature 449:223–227

Hauspie RC, Pagezy H (1989) Longitudinal study of growth of African babies: an analysis of seasonal variations in the average growth rate and the effects of infectious diseases on individual and average growth patterns. Acta Paediatr Scand Suppl 350:37–43

Hemani G, Yang J, Vinkhuyzen A, Powell JE, Willemsen G, Hottenga JJ, Abdellaoui A, Mangino M, Valdes AM, Medland SE, Madden PA, Heath AC, Henders AK, Nyholt DR, de Geus EJ, Magnusson PK, Ingelsson E, Montgomery GW, Spector TD, Boomsma DI et al (2013) Inference of the genetic architecture underlying BMI and height with the use of 20,240 sibling pairs. Am J Hum Genet 93:865–875

Hewlett BS (1996) Cultural diversity among African Pygmies. In: Kent S (ed) Cultural diversity among twentieth-century foragers: An African perspective, vol 1. Cambridge University Press, Cambridge, pp 215–244

Hewlett BS (2014) Introduction. Hunter-gatherers of the Congo Basin: cultures, histrories, and biology of African Pygmies. Transaction Publishers, New Brunswick, pp vxii–xxix

Hiernaux J (1974) The people of Africa. Weidenfeld and Nicolson, London

Hiernaux J (1977) Long-term biological effects of human migrations from the African savanna to the equitorial forest: a case study of human adaption to a hot and wet climate. In: Harrison GA (ed) Population structure and human variation. Cambridge University Press, London, pp 187–217

Hintze JL, Nelson RD (1998) Violin plots: a box plot-density trace synergism. Am Stat 52:181–184

Hsieh P, Veeramah KR, Lachance J, Tishkoff SA, Wall JD, Hammer MF, Gutenkunst RN (2016) Whole-genome sequence analyses of Western Central African Pygmy hunter-gatherers reveal a complex demographic history and identify candidate genes under positive natural selection. Genome Res 26:279–290

Huang L, Jakobsson M, Pemberton TJ, Ibrahim M, Nyambo T, Omar S, Pritchard JK, Tishkoff SA, Rosenberg NA (2011) Haplotype variation and genotype imputation in African populations. Genet Epidemiol 35:766–780

Huxley JS (1932) Problems of relative growth. Methuen and Company Ltd, London

Jakobsson M, Rosenberg NA (2007) CLUMPP: a cluster matching and permutation program for dealing with label switching and multimodality in analysis of population structure. Bioinformatics 23:1801–1806

Jarvis JP, Scheinfeldt LB, Soi S, Lambert C, Omberg L, Ferwerda B, Froment A, Bodo JM, Beggs W, Hoffman G, Mezey J, Tishkoff SA (2012) Patterns of ancestry, signatures of natural selection, and genetic association with stature in Western African Pygmies. PLoS Genet 8:e1002641

Joiris D (2003) The framework of Central African hunter-gatherers and neighbouring societies. Afr Stud Monogr 28:57–79

Kang HM, Sul JH, Service SK, Zaitlen NA, Kong SY, Freimer NB, Sabatti C, Eskin E (2010) Variance component model to account for sample structure in genome-wide association studies. Nat Genet 42:348–354

Katzmarzyk PT, Leonard WR (1998) Climatic influences on human body size and proportions: ecological adaptations and secular trends. Am J Phys Anthropol 106:483–503

Kim HN, Lee EJ, Jung SC, Lee JY, Chung HW, Kim HL (2010) Genetic variants that affect length/height in infancy/early childhood in Vietnamese-Korean families. J Hum Genet 55:681–690

Ko J, Na M, Kim S, Lee JR, Kim E (2003) Interaction of the ERC family of RIM-binding proteins with the liprin-alpha family of multidomain proteins. J Biol Chem 278:42377–42385

Lachance J, Vernot B, Elbers CC, Ferwerda B, Froment A, Bodo JM, Lema G, Fu W, Nyambo TB, Rebbeck TR, Zhang K, Akey JM, Tishkoff SA (2012) Evolutionary history and adaptation from high-coverage whole-genome sequences of diverse african hunter-gatherers. Cell 150:457–469

Lango Allen H, Estrada K, Lettre G, Berndt SI, Weedon MN, Rivadeneira F, Willer CJ, Jackson AU, Vedantam S, Raychaudhuri S, Ferreira T, Wood AR, Weyant RJ, Segre AV, Speliotes EK, Wheeler E, Soranzo N, Park JH, Yang J, Gudbjartsson D et al (2010) Hundreds of variants clustered in genomic loci and biological pathways affect human height. Nature 467:832–838

Le Bomin S, Mbot J (2012) The musical heritage of the Gabon Bongo. Before Farming 2012/1 article 3:1–19

Lemogoum D, Ngatchou W, Janssen C, Leeman M, Van Bortel L, Boutouyrie P, Degaute JP, Van de Borne P (2012) Effects of hunter-gatherer subsistence mode on arterial distensibility in Cameroonian pygmies. Hypertension 60:123–128

Lettre G, Jackson AU, Gieger C, Schumacher FR, Berndt SI, Sanna S, Eyheramendy S, Voight BF, Butler JL, Guiducci C, Illig T, Hackett R, Heid IM, Jacobs KB, Lyssenko V, Uda M, Boehnke M, Chanock SJ, Groop LC, Hu FB et al (2008) Identification of ten loci associated with height highlights new biological pathways in human growth. Nat Genet 40:584–591

Li Z, Xi X, Gu M, Feil R, Ye RD, Eigenthaler M, Hofmann F, Du X (2003) A stimulatory role for cGMP-dependent protein kinase in platelet activation. Cell 112:77–86

Li MJ, Wang LY, Xia Z, Sham PC, Wang J (2013) GWAS3D: detecting human regulatory variants by integrative analysis of genome-wide associations, chromosome interactions and histone modifications. Nucleic Acids Res 41:W150–W158

Liu JZ, McRae AF, Nyholt DR, Medland SE, Wray NR, Brown KM, Investigators A, Hayward NK, Montgomery GW, Visscher PM, Martin NG, Macgregor S (2010) A versatile gene-based test for genome-wide association studies. Am J Hum Genet 87:139–145

Lopez M, Tena-Sempere M (2016) Estradiol and brown fat. Best Pract Res Clin Endocrinol Metab 30:527–536

Lopez M, Kousathanas A, Quach H, Harmant C, Mouguiama-Daouda P, Hombert JM, Froment A, Perry GH, Barreiro LB, Verdu P,

Patin E, Quintana-Murci L (2018) The demographic history and mutational load of African hunter-gatherers and farmers. Nat Ecol Evol 2:721–730

Maas NM, Van de Putte T, Melotte C, Francis A, Schrander-Stumpel CT, Sanlaville D, Genevieve D, Lyonnet S, Dimitrov B, Devriendt K, Fryns JP, Vermeesch JR (2007) The *C20orf133* gene is disrupted in a patient with Kabuki syndrome. J Med Genet 44:562–569

Marchini J, Cardon LR, Phillips MS, Donnelly P (2004) The effects of human population structure on large genetic association studies. Nat Genet 36:512–517

Martin AR, Gignoux CR, Walters RK, Wojcik GL, Neale BM, Gravel S, Daly MJ, Bustamante CD, Kenny EE (2017) Human demographic history impacts genetic risk prediction across diverse populations. Am J Hum Genet 100:635–649

Merimee TJ, Rimoin DL, Cavalli-Sforza LL (1972) Metabolic studies in the African Pygmy. J Clin Investig 51:395–401

Merimee TJ, Zapf J, Hewlett B, Cavalli-Sforza LL (1987) Insulin-like growth factors in Pygmies. The role of puberty in determining final stature. N Engl J Med 316:906–911

Meyer LR, Zweig AS, Hinrichs AS, Karolchik D, Kuhn RM, Wong M, Sloan CA, Rosenbloom KR, Roe G, Rhead B, Raney BJ, Pohl A, Malladi VS, Li CH, Lee BT, Learned K, Kirkup V, Hsu F, Heitner S, Harte RA et al (2013) The UCSC Genome Browser database: extensions and updates 2013. Nucleic Acids Res 41:D64–D69

Migliano AB, Vinicius L, Lahr MM (2007) Life history trade-offs explain the evolution of human Pygmies. Proc Natl Acad Sci USA 104:20216–20219

Montano V, Ferri G, Marcari V, Batini C, Anyaele O, Destro-Bisol G, Comas D (2011) The Bantu expansion revisited: a new analysis of Y chromosome variation in Central Western Africa. Mol Ecol 20:2693–2708

N'Diaye A, Chen GK, Palmer CD, Ge B, Tayo B, Mathias RA, Ding J, Nalls MA, Adeyemo A, Adoue V, Ambrosone CB, Atwood L, Bandera EV, Becker LC, Berndt SI, Bernstein L, Blot WJ, Boerwinkle E, Britton A, Casey G et al (2011) Identification, replication, and fine-mapping of loci associated with adult height in individuals of African ancestry. PLoS Genet 7:e1002298

Ober C, Abney M, McPeek MS (2001) The genetic dissection of complex traits in a founder population. Am J Hum Genet 69:1068–1079

Ohenjo N, Willis R, Jackson D, Nettleton C, Good K, Mugarura B (2006) Health of Indigenous people in Africa. Lancet 367:1937–1946

Oishi I, Suzuki H, Onishi N, Takada R, Kani S, Ohkawara B, Koshida I, Suzuki K, Yamada G, Schwabe GC, Mundlos S, Shibuya H, Takada S, Minami Y (2003) The receptor tyrosine kinase Ror2 is involved in non-canonical Wnt5a/JNK signalling pathway. Genes Cells 8:645–654

Patin E, Laval G, Barreiro LB, Salas A, Semino O, Santachiara-Benerecetti S, Kidd KK, Kidd JR, Van der Veen L, Hombert JM, Gessain A, Froment A, Bahuchet S, Heyer E, Quintana-Murci L (2009) Inferring the demographic history of African farmers and Pygmy hunter-gatherers using a multilocus resequencing data set. PLoS Genet 5:e1000448

Patin E, Siddle KJ, Laval G, Quach H, Harmant C, Becker N, Froment A, Regnault B, Lemee L, Gravel S, Hombert JM, Van der Veen L, Dominy NJ, Perry GH, Barreiro LB, Verdu P, Heyer E, Quintana-Murci L (2014) The impact of agricultural emergence on the genetic history of African rainforest hunter-gatherers and agriculturalists. Nat Commun 5:3163

Patin E, Lopez M, Grollemund R, Verdu P, Harmant C, Quach H, Laval G, Perry GH, Barreiro LB, Froment A, Heyer E, Massougbodji A, Fortes-Lima C, Migot-Nabias F, Bellis G, Dugoujon JM, Pereira JB, Fernandes V, Pereira L, Van der Veen L et al (2017) Dispersals and genetic adaptation of Bantu-speaking populations in Africa and North America. Science 356:543–546

Pemberton TJ, Wang C, Li JZ, Rosenberg NA (2010) Inference of unexpected genetic relatedness among individuals in HapMap Phase III. Am J Hum Genet 87:457–464

Perola M, Sammalisto S, Hiekkalinna T, Martin NG, Visscher PM, Montgomery GW, Benyamin B, Harris JR, Boomsma D, Willemsen G, Hottenga JJ, Christensen K, Kyvik KO, Sorensen TI, Pedersen NL, Magnusson PK, Spector TD, Widen E, Silventoinen K, Kaprio J et al (2007) Combined genome scans for body stature in 6,602 European twins: evidence for common Caucasian loci. PLoS Genet 3:e97

Perry GH, Dominy NJ (2009) Evolution of the human pygmy phenotype. Trends Ecol Evol 24:218–225

Perry GH, Foll M, Grenier JC, Patin E, Nedelec Y, Pacis A, Barakatt M, Gravel S, Zhou X, Nsobya SL, Excoffier L, Quintana-Murci L, Dominy NJ, Barreiro LB (2014a) Adaptive, convergent origins of the pygmy phenotype in African rainforest hunter-gatherers. Proc Natl Acad Sci USA 111:E3596–E3603

Perry JR, Day F, Elks CE, Sulem P, Thompson DJ, Ferreira T, He C, Chasman DI, Esko T, Thorleifsson G, Albrecht E, Ang WQ, Corre T, Cousminer DL, Feenstra B, Franceschini N, Ganna A, Johnson AD, Kjellqvist S, Lunetta KL et al (2014b) Parent-of-origin-specific allelic associations among 106 genomic loci for age at menarche. Nature 514:92–97

Pfeifer A, Klatt P, Massberg S, Ny L, Sausbier M, Hirneiss C, Wang GX, Korth M, Aszodi A, Andersson KE, Krombach F, Mayerhofer A, Ruth P, Fassler R, Hofmann F (1998) Defective smooth muscle regulation in cGMP kinase I-deficient mice. EMBO J 17:3045–3051

Pritchard JK, Stephens M, Donnelly P (2000) Inference of population structure using multilocus genotype data. Genetics 155:945–959

Pruim RJ, Welch RP, Sanna S, Teslovich TM, Chines PS, Gliedt TP, Boehnke M, Abecasis GR, Willer CJ (2010) LocusZoom: regional visualization of genome-wide association scan results. Bioinformatics 26:2336–2337

Quintana-Murci L, Quach H, Harmant C, Luca F, Massonnet B, Patin E, Sica L, Mouguiama-Daouda P, Comas D, Tzur S, Balanovsky O, Kidd KK, Kidd JR, van der Veen L, Hombert JM, Gessain A, Verdu P, Froment A, Bahuchet S, Heyer E et al (2008) Maternal traces of deep common ancestry and asymmetric gene flow between Pygmy hunter-gatherers and Bantu-speaking farmers. Proc Natl Acad Sci USA 105:1596–1601

R Development Core Team (2014) R: a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna

Ramirez Rozzi FV, Sardi ML (2010) Diversity among African pygmies. PLoS One 5:e13620

Rimoin DL, Merimee TJ, Rabinowitz D, Cavalli-Sforza LL, McKusick VA (1969) Peripheral subresponsiveness to human growth hormone in the African Pygmies. N Engl J Med 281:1383–1388

Rosenberg NA (2004) distruct: a program for the graphical display of population structure. Mol Ecol Notes 4:137–138

Rozzi FV, Koudou Y, Froment A, Le Bouc Y, Botton J (2015) Growth pattern from birth to adulthood in African pygmies of known age. Nat Commun 6:7672

Samaan G, Yugo D, Rajagopalan S, Wall J, Donnell R, Goldowitz D, Gopalakrishnan R, Venkatachalam S (2010) Foxn3 is essential for craniofacial development in mice and a putative candidate involved in human congenital craniofacial defects. Biochem Biophys Res Commun 400:60–65

Schmeisser MJ, Ey E, Wegener S, Bockmann J, Stempel AV, Kuebler A, Janssen AL, Udvardi PT, Shiban E, Spilker C, Balschun D, Skryabin BV, Dieck S, Smalla KH, Montag D, Leblond CS, Faure P, Torquet N, Le Sourd AM, Toro R et al (2012) Autistic-like

behaviours and hyperactivity in mice lacking ProSAP1/Shank2. Nature 486:256–260
Seeman E (1998) Growth in bone mass and size—are racial and gender differences in bone mineral density more apparent than real? J Clin Endocrinol Metab 83:1414–1419
Sharma S, tenOever BR, Grandvaux N, Zhou GP, Lin R, Hiscott J (2003) Triggering the interferon antiviral response through an IKK-related pathway. Science 300:1148–1151
Shea BT, Bailey RC (1996) Allometry and adaptation of body proportions and stature in African pygmies. Am J Phys Anthropol 100:311–340
Shriner D, Adeyemo A, Gerry NP, Herbert A, Chen G, Doumatey A, Huang H, Zhou J, Christman MF, Rotimi CN (2009) Transferability and fine-mapping of genome-wide associated loci for adult height across human populations. PLoS One 4:e8398
Silventoinen K, Kaprio J, Lahelma E, Koskenvuo M (2000) Relative effect of genetic and environmental factors on body height: differences across birth cohorts among Finnish men and women. Am J Public Health 90:627–630
Soranzo N, Rivadeneira F, Chinappen-Horsley U, Malkina I, Richards JB, Hammond N, Stolk L, Nica A, Inouye M, Hofman A, Stephens J, Wheeler E, Arp P, Gwilliam R, Jhamai PM, Potter S, Chaney A, Ghori MJ, Ravindrarajah R, Ermakov S et al (2009) Meta-analysis of genome-wide scans for human adult stature identifies novel loci and associations with measures of skeletal frame size. PLoS Genet 5:e1000445
Splawski I, Timothy KW, Decher N, Kumar P, Sachse FB, Beggs AH, Sanguinetti MC, Keating MT (2005) Severe arrhythmia disorder caused by cardiac L-type calcium channel mutations. Proc Natl Acad Sci USA 102:8089–8096
The International HapMap Consortium (2007) A second generation human haplotype map of over 3.1 million SNPs. Nature 449:851–861
Tupman GS (1962) A study of bone growth in normal children and its relationship to skeletal maturation. J Bone Jt Surg Br 44-B:42–67
Turnbull C (1965) Wayward servants, the two worlds of the African Pygmies. The Natural History Press, New York
Udy GB, Towers RP, Snell RG, Wilkins RJ, Park SH, Ram PA, Waxman DJ, Davey HW (1997) Requirement of STAT5b for sexual dimorphism of body growth rates and liver gene expression. Proc Natl Acad Sci USA 94:7239–7244
van Kuilenburg AB, Meijer J, Mul AN, Hennekam RC, Hoovers JM, de Die-Smulders CE, Weber P, Mori AC, Bierau J, Fowler B, Macke K, Sass JO, Meinsma R, Hennermann JB, Miny P, Zoetekouw L, Vijzelaar R, Nicolai J, Ylstra B, Rubio-Gozalbo ME (2009) Analysis of severely affected patients with dihydropyrimidine dehydrogenase deficiency reveals large intragenic rearrangements of DPYD and a de novo interstitial deletion del(1)(p13.3p21.3). Hum Genet 125:581–590
Verdu P (2009) Anthropologie génétique des populations humaines d'Afrique Centrale: histoire du peuplement Pygmée. Ph.D. thesis, Université Pierre-et-Marie-Curie, Paris, France
Verdu P, Austerlitz F, Estoup A, Vitalis R, Georges M, Thery S, Froment A, Le Bomin S, Gessain A, Hombert JM, Van der Veen L, Quintana-Murci L, Bahuchet S, Heyer E (2009) Origins and genetic diversity of Pygmy hunter-gatherers from Western Central Africa. Curr Biol 19:312–318
Verdu P, Becker NS, Froment A, Georges M, Grugni V, Quintana-Murci L, Hombert JM, Van der Veen L, Le Bomin S, Bahuchet S, Heyer E, Austerlitz F (2013) Sociocultural behavior, sex-biased admixture, and effective population sizes in Central African Pygmies and non-Pygmies. Mol Biol Evol 30:918–937
Voight BF, Kang HM, Ding J, Palmer CD, Sidore C, Chines PS, Burtt NP, Fuchsberger C, Li Y, Erdmann J, Frayling TM, Heid IM, Jackson AU, Johnson T, Kilpelainen TO, Lindgren CM, Morris AP, Prokopenko I, Randall JC, Saxena R et al (2012) The Metabochip, a custom genotyping array for genetic studies of metabolic, cardiovascular, and anthropometric traits. PLoS Genet 8:e1002793
Wadsworth ME, Hardy RJ, Paul AA, Marshall SF, Cole TJ (2002) Leg and trunk length at 43 years in relation to childhood health, diet and family circumstances; evidence from the 1946 national birth cohort. Int J Epidemiol 31:383–390
Waldman LA, Chia DJ (2013) Towards identification of molecular mechanisms of short stature. Int J Pediatr Endocrinol 2013:19
Weedon MN, Lango H, Lindgren CM, Wallace C, Evans DM, Mangino M, Freathy RM, Perry JR, Stevens S, Hall AS, Samani NJ, Shields B, Prokopenko I, Farrall M, Dominiczak A, Johnson T, Bergmann S, Beckmann JS, Vollenweider P, Waterworth DM et al (2008) Genome-wide association analysis identifies 20 loci that influence adult height. Nat Genet 40:575–583
Weiner JS, Lourie JA (1981) Practical human biology. Academic Press, London
Weir BS (1996) Genetic data analysis II. Sinauer, Sunderland
Wolf SF, Sieburth D, Sypek J (1994) Interleukin 12: a key modulator of immune function. Stem Cells 12:154–168
Wood AR, Esko T, Yang J, Vedantam S, Pers TH, Gustafsson S, Chu AY, Estrada K, Luan J, Kutalik Z, Amin N, Buchkovich ML, Croteau-Chonka DC, Day FR, Duan Y, Fall T, Fehrmann R, Ferreira T, Jackson AU, Karjalainen J et al (2014) Defining the role of common variation in the genomic and biological architecture of adult human height. Nat Genet 46:1173–1186
Yiu GK, Kaunisto A, Chin YR, Toker A (2011) NFAT promotes carcinoma invasive migration through glypican-6. Biochem J 440:157–166
Yu J, Pressoir G, Briggs WH, Vroh Bi I, Yamasaki M, Doebley JF, McMullen MD, Gaut BS, Nielsen DM, Holland JB, Kresovich S, Buckler ES (2006) A unified mixed-model method for association mapping that accounts for multiple levels of relatedness. Nat Genet 38:203–208
Zaitlen N, Kraft P, Patterson N, Pasaniuc B, Bhatia G, Pollack S, Price AL (2013) Using extended genealogy to estimate components of heritability for 23 quantitative and dichotomous traits. PLoS Genet 9:e1003520