Check for updates

Received: 28 January 2022 | Revised: 25 October 2022 | Accepted: 27 October 2022

DOI: 10.1002/ajpa.24660

BRIEF COMMUNICATION

AMERICAN JOURNAL OF BIOLOGICAL ANTHROPOLOGY
The Official Journal of the American Association of Biological Anthropologists

WILEY

Examination of runs of homozygosity in relation to height in an endogamous Namibian population

Natalie Swinford[1] | Brenna Gallagher[1] | Jacob Sheehama[2] | Meng Lin[3] | Sean Prall[4] | Brooke Scelza[5] | Brenna M. Henn[1,6]

[1]Department of Anthropology, University of California, Davis, California, USA

[2]School of Medicine, University of Namibia, Windhoek, Namibia

[3]Anschutz Medical Campus, University of Colorado, Aurora, Colorado, USA

[4]Department of Antrhopology, University of Missouri, Columbia, Missouri, USA

[5]Department of Anthropology, University of California, Los Angeles, California, USA

[6]UC Davis Genome Center, University of California, Davis, California, USA

Correspondence
Brenna M. Henn, Department of Anthropology, University of California, Davis, CA, USA.
Email: bmhenn@ucdavis.edu

Funding information
National Institutes of Health, Grant/Award Number: R35GM133531; National Science Foundation, Grant/Award Number: BCS-1534682

Abstract

Objectives: Height is a complex, highly heritable polygenic trait subject to both genetic composition and environmental influences. Recent studies suggest that a large proportion of height heritability is determined by the cumulative effect of many low allele frequency variants across the genome. Previous research has also identified an inverse relationship between height and runs of homozygosity (ROH); however, this has yet to be examined within African populations. We aim to identify this association within the Himba, an endogamous Namibian population who are recently bottlenecked, resulting in elevated haplotype sharing and increased homozygosity.

Materials and Methods: Here, we calculate the fraction of the genome composed of long runs of homozygosity (F_{ROH}) in a sample of 245 adults and use mixed effects models to assess its effect on height.

Results: We find that Himba adults exhibit increased homozygosity. However, in contrast to previous studies in other populations, we do not find a significant effect of F_{ROH} on height within the Himba. We further estimated heritability of height, noting both an enrichment of distant relatives and greater developmental homogeneity across households; we find that $h_g^2 = 0.59$ ($SE \pm 0.146$), comparable to estimates reported in Europeans.

Discussion: Our results may be due to other environmental variables we were not able to include, measurement error, or low statistical power, but may also imply that phenotypic expression resulting from increased homozygosity may vary from population to population.

KEYWORDS
height, population genetics, runs of homozygosity

1 | INTRODUCTION

Height is a complex, polygenic trait determined by genetic and environmental factors. On average, height is estimated to be roughly 80% heritable (Silventoinen et al., 2003; Visscher et al., 2008; Yang et al., 2010). Genome-wide association studies (GWAS) use a minimum effect rate threshold to identify critical variants for complex traits such as height and diseases. However, considering only highly significant single nucleotide polymorphisms (SNPs) that pass this significance threshold explains only a small proportion of height

This is an open access article under the terms of the Creative Commons Attribution-NonCommercial License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited and is not used for commercial purposes.
© 2022 The Authors. *American Journal of Biological Anthropology* published by Wiley Periodicals LLC.

26927691, 0, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/ajpa.24660, Wiley Online Library on [24/11/2022]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

heritability. Early GWA studies identified ~50 variants significantly associated with height, but together these only explained ~5% of the variation (Yang et al., 2010). Later studies increased the number of SNPs found to reach genome-wide significance in association with height to ~700 variants, yet still these SNPs together only capture ~16% of the variance (Wood et al., 2014). By using all common SNPs, not just those reaching genome-wide significance, the percentage of variance explained increases to 60% (Wood et al., 2014; Yang et al., 2010). More recently, studies have sought to understand the role of rare and low-frequency SNPs in the genetic architecture of this phenotype and have suggested that rare variants can capture the missing heritability from previous estimates (Marouli et al., 2017; Wainschtein et al., 2021). Marouli et al. (2017) identified 83 rare and low minor allele frequency (MAF) coding SNPs affecting height, and Wainschtein et al. (2021) report that 47% of height heritability is explained by variants with low MAF values ranging from 0.0001 to 0.1. Roughly half of the genetic variance of height can be explained by the combined effect of multiple alleles of relatively small effect size across the genome. Because height is largely determined by the cumulative effect of individually insignificant variants scattered across the genome, runs of homozygosity (ROH) resulting from identically inherited haplotypes would capture the expression of recessive alleles associated with height. Furthermore, a negative correlation between ROH and height has previously been identified within several, mostly European, populations, where offspring of first cousins exhibited a height reduction of 1.2–3.0 cm (Joshi et al., 2015; McQuillan et al., 2012). This suggests a strong influence of multiple recessive alleles across the genome. However, research is comparatively lacking in assessing this relationship in global, and particularly underrepresented, populations.

Population bottlenecks and endogamy can lead to increased levels of identity by descent (IBD) sharing, which in turn can lead to increased ROH as they increase the likelihood that individuals may inherit identical segments from parents who share a common ancestor. The more recent this common ancestor is shared, the longer the ROH segments in the offspring are expected to be, and research has shown that longer segments of ROH carry a significantly higher proportion of deleterious variants compared with shorter runs (Mooney et al., 2018; Szpiech et al., 2013). The fraction of the genome composed of ROH is mathematically denoted as F_{ROH} and can be used to estimate the inbreeding coefficient, F, which is a technical term used in population genetics referring to the expected proportion of genomic autozygosity (Ceballos et al., 2018). Elevated levels of F_{ROH} have previously been associated with several quantitative traits, including cardiometabolic disease phenotypes, height, and lung function (Ceballos et al., 2020; Joshi et al., 2015; McQuillan et al., 2012; Nabulsi et al., 2003; Yengo et al., 2019).

Here we investigate the relationship between height and F_{ROH} among a population of Himba living in northwestern Namibia. Specifically, our study was conducted in the community of Omuhonga. The Himba are a Bantu-speaking, seminomadic, agro-pastoralist population whose previous and current demographic includes various factors that are known to increase F_{ROH}. They have experienced a recent bottleneck estimated to have begun approximately 60 generations ago and reaching a minimum effective population size 12 generations ago (Swinford et al., 2022). They are endogamous and have social norms promoting consanguineous marriage (Harpending & Pennington, 1990; Scelza et al., 2021). However, "love matches," divorce, polygyny, and concurrent partnerships (including extra-marital partnerships) are also common (Scelza et al., 2020; Scelza, 2011). This practice of concurrency results in a high rate of extrapair paternity (48%) and a large number of second degree relationships within the population (Scelza et al., 2020; Williams et al., 2020). It has also been observed that the Himba exhibit elevated background levels of haplotype sharing, evidenced by the consistent overestimation of relationship type by common IBD-based methods of relationship inference (Williams et al., 2020). Therefore, Himba population history provides a unique opportunity to examine the effects of F_{ROH} on height within an underrepresented African population.

When investigating the effect of F_{ROH} on height, it is important to consider possible confounding variables. Fosterage occurs when a child is raised by an adult other than the parent, typically another family member, and is a common practice within Himba culture, with 39% of women and 42% of men reporting having been fostered at some point (Prall & Scelza, 2017). Fosterage occurs for a variety of reasons, with variation in the duration of and age at which fosterage begins. A woman may choose to foster out one or more children before moving to the household of a new husband (i.e., after remarriage) or if she has many children. Similarly, a woman may choose to foster out the eldest of her small children if she has multiple babies due to short interbirth intervals. Fosterage can also occur by request. In these cases, an adult who would like to bring in a fostered child to help with household work asks the biological parents to foster out a child to them. In addition, fosterage occurs in cases of parental death (Prall & Scelza, 2017; Scelza & Silk, 2014). Previous research has shown that fosterage can negatively impact development, notably height (Prall & Scelza, 2017). Therefore, to account for the potential masking effects of fosterage on the association of F_{ROH} and height, we include fosterage status as a variable in our models. We expect to find elevated levels of F_{ROH} within the Himba and to consequently observe a negative correlation between F_{ROH} and height in Himba adults.

2 | RESULTS

In order to characterize the relationship between F_{ROH} and height, we utilized genome-wide SNP array data from 245 Himba adults and built a mixed effects model, where height (i.e., outcome) was predicted by F_{ROH} using the covariates of sex, year of birth (YOB), fosterage status, and a genetic relatedness matrix (GRM). In our dataset, height was normally distributed for males ($n = 99$, mean = 175.2, SD = 5.8) and females ($n = 146$, mean = 164.3, SD = 5.2) (Figure 1a). Himba adult height is comparable to many European-descent populations and is normally distributed. F_{ROH} was calculated using a minimum 1500 kilobase (kb) threshold—a level comparable to a pedigree-based estimate of the F coefficient (Ceballos et al., 2018). Thus, a F_{ROH} 1500 value of

26927691, 0, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/ajpa.24660, Wiley Online Library on [24/11/2022]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

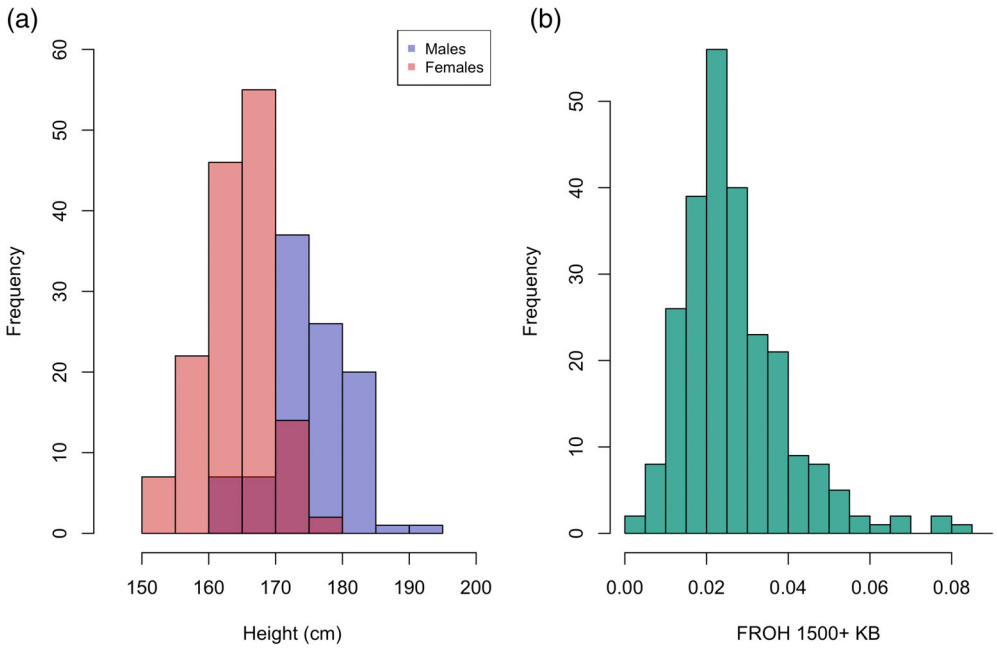

FIGURE 1 (a) Height distribution for individuals ($n = 245$) with females ($n = 146$, mean = 164.2, SD = 5.24) represented in pink, males ($n = 99$, mean = 175.2, SD = 5.79) represented in purple, and the maroon color indicating overlap between the male and female height distributions. Individuals' heights are represented as the average of their height measurements when individuals had multiple height measurements in the dataset. (b) the distribution of F_{ROH} values for individuals (mean = 0.026).

0.0625, or 6.25%, represents the expected level of homozygosity for the offspring of first cousins in a population that does not contain elevated background haplotype sharing. This measure of F_{ROH} 1500 has also been implemented in previous studies to examine the relationship between homozygosity and height (Joshi et al., 2015; McQuillan et al., 2012) and between homozygosity and various other phenotypes in the UK Biobank (Yengo et al., 2019). Here, we find elevated levels of F_{ROH} 1500 within the Himba with an overall average F_{ROH} value of 0.0266, or 2.66% (Figure 1b), which is higher than the expectation for offspring of second cousins but less than the expectation for offspring of first cousins (Bittles, 2002). This finding is congruent with previous reports of elevated IBD sharing among Himba (Williams et al., 2020).

Fosterage is defined as children leaving their natal homes to live with another adult from the community, typically a family member (Prall & Scelza, 2017). Research using a similar Himba dataset found that fosterage status negatively affected growth among children, continuing into adulthood for women who had been fostered (Prall & Scelza, 2017) and thus we included this variable in our models. In this sample of Himba adults, fosterage status includes any individual who was fostered at any point in their childhood (males $n = 36$, females $n = 31$). Although children varied in age and length of fosterage and the reason for fosterage, these data were missing for most individuals in our dataset, and thus these specific details of child fosterage could not be incorporated into our models. Therefore, we only included a binary variable describing if an individual had ever been fostered.

Relatedness among individual samples is often a concern in genetic architecture of traits because spurious results can occur when household features affecting the trait are then correlated with the genetic features of a family. However, we do not expect this to be an issue in this dataset due to the complex structure of households. The pedigrees in this community are extremely reticulated (Swinford et al., 2022). This is likely due to the practice of concurrency which leads to a very high rate extra-pair-paternity (Scelza et al., 2020). In addition, the practice of fosterage is very common. Therefore, rather than households clustering genetically, households can consist of both children and adults who are unrelated to each other or related to different degrees. With this type of mixing among households, we do not expect the presence of related individuals in the dataset to cause confounding, but we have also included a GRM in our mixed effects models to help correct for any possible confounding.

TABLE 1 Summary statistics for baseline mixed effects model

Variable	*p* value	Effect size	Standard error
Year of birth	0.79	−0.0036	0.014
Sex: Male	0***	11.06	0.779
Fosterage: Yes	0.95	−0.063	1.016
Sex: Male, fosterage: Yes	0.85	−0.267	1.459

***highly significant *p*-value ($p < 0.001$).

We built a baseline model which used fosterage status, sex, YOB (meant to capture fluctuating environmental effects such as food availability during periods of drought, as well as the shrinking effects of age), and a GRM to isolate any effect of F_{ROH} by comparing our baseline and F_{ROH}-incorporated models. Our baseline model did not reveal a significant association between fosterage status and height ($p = 0.95$, Table 1) nor between YOB and height ($p = 0.79$, Table 1) within Himba adults. Backwards elimination of either variable did not improve significance of the other variable. We found similar results in our F_{ROH}-incorporated model, which used log-transformed F_{ROH} (for normalization), fosterage status, sex, YOB, and a GRM to predict height. This model did not reveal any effect of F_{ROH} 1500 on height ($p = 0.98$, Figure 2, Table 2). Similarly, neither fosterage status ($p = 0.95$) nor YOB ($p = 0.80$) had a significant effect on adult height. The only covariate exhibiting significance was the expected fixed effect of sex on height ($p = 2 \times 10^{-16}$), but there was no interaction between fosterage status and sex on height ($p = 0.462$). In addition, backwards elimination of the YOB and fosterage variables did

26927691, 0, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/ajpa.24660, Wiley Online Library on [24/11/2022]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

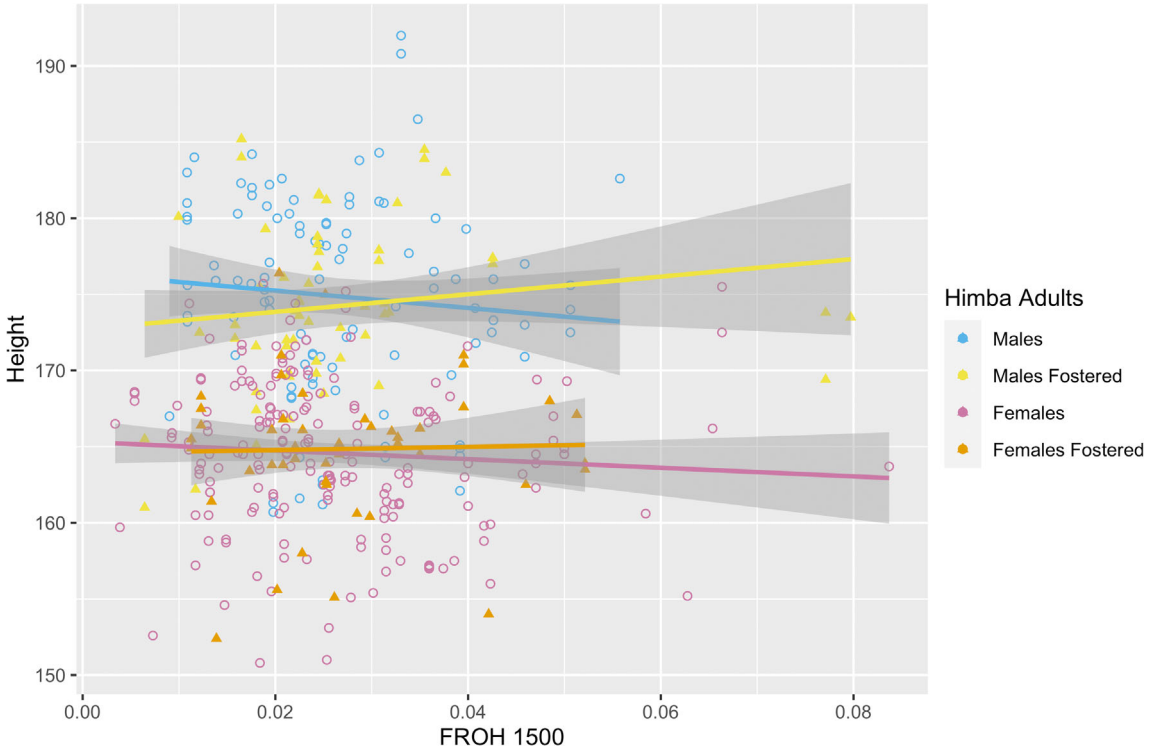

FIGURE 2 Scatter plot showing the relationship between F_{ROH} and height for unfostered females (open pink circles), fostered females (solid orange triangles), unfostered males (open blue circles), and fostered males (solid yellow triangles). The model prediction is shown for fostered and unfostered males (yellow and blue lines, respectively) as well as fostered and unfostered females (orange and pink lines, respectively).

TABLE 2 Summary statistics for mixed effects model with F_{ROH}

Variable	*p* value	Effect size	Standard error
Log-transformed F_{ROH} 1500*	0.98	0.020	0.724
Year of birth	0.80	0.004	0.014
Sex: Male	0***	11.06	0.784
Fosterage: Yes	0.95	−0.064	1.017
Sex: Male, fosterage: Yes	0.86	−0.264	1.463

Note: The effect size and standard error of F_{ROH} is not detransformed here because height and F_{ROH} are not linear in the model.
***highly significant *p*-value ($p < 0.001$).

not reveal any new significance in the model covariates. These patterns were consistent among all models using different ROH thresholds to calculate F_{ROH}. In other words, we did not find an effect of F_{ROH} on height for any size threshold used to call ROH (Figure 2).

We also explored using Garlic (Szpiech et al., 2017) to call ROH because Garlic performs its own length classifications to determine short, medium, and long segments. However, the length thresholds chosen by Garlic were almost identical to the length thresholds we chose and therefore would not have changed the conclusions of the model. We also do not find an effect of F_{ROH} on height at a single time point. Rather than including multiple measurements per individual and varying the intercept by individual ID, we tested for an association between F_{ROH} and height for all individuals measured in a single year. However, because not every individual was measured every year, the sample size is reduced. The year with the most measurements was 2016 ($n = 180$). We ran a linear model on this data with a principal component correction for relatedness, but still found no significant association.

One potential concern could be that height in the Himba is attenuated by environmental or developmental variability related to drought, fosterage or other stochastic events. We estimated the narrow sense heritability of height in our dataset using a linear mixed model GRM approach consisting of ~370 K SNPs (*Methods*). Because overall relatedness within the Himba population is high, it would be impossible to remove all relatives from our dataset. Thus, our estimate is a combination of both a family based heritability as well as unrelated individuals, analogous to a K_{IBD} approach (Zaitlen et al., 2013). We estimated heritability to be $h^2_g = 0.59$ (SE ± 0.146). This estimate is comparable to previous studies which use common variants among unrelated individuals in Europeans and Ugandans, $h^2_g = 0.50 - 0.60$ (Heckerman et al., 2016; Wood et al., 2014; Yang et al., 2010), but slightly lower than $h^2_{IBD} = 0.69$ in the Icelandic population. Our heritability results suggest that environmental attributes are not unusually large for the Himba, though greater sample sizes are needed to increase precision in the estimate.

3 | DISCUSSION

We identified increased levels of F_{ROH} within our cohort of Himba adults, where the average F_{ROH} estimate is 0.0266. Our results suggest an intermediate level of homozygosity: the mean F_{ROH} value is higher than that expected for progeny of second cousins (0.0125), but

26927691, 0, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/ajpa.24660, Wiley Online Library on [24/11/2022]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

less than that expected for progeny of first cousins (0.0625) (Bittles, 2002). ROH are ubiquitous; however, the global distribution of ROH is not uniform among populations as a general trend is observed between decreased genomic diversity and increased distance from Africa (Ceballos et al., 2018). ROH do not always follow this distinct stepwise pattern and are dependent upon demographic factors including consanguinity and reduced effective population size (Ceballos et al., 2018; Pemberton et al., 2012), both of which have likely affected the Himba. Previous work has reported a societal preference for first-cousin arranged marriages within the Himba (Scelza et al., 2021), but unmarried relationships are common and accepted, and a large proportion of births occur through either nonmarital or extra-marital partnerships (Scelza et al., 2020). The Himba have also experienced a relatively recent bottleneck that reached a minimum population size approximately 12 generations ago (Swinford et al., 2022). Population bottlenecks reduce the genetic diversity in a population. Thus, even if first-cousin marriages do not result in offspring, children resulting from extra-pair partnerships would still be subject to this overall background level of haplotype sharing. Unrelated partners may still share more IBD than would be expected due to elevated levels of IBD sharing across the entire population, resulting from population bottleneck and endogamy over time.

Previous research continues to demonstrate the saliency of examining variants with low allele frequency when investigating human height, as studies have shown that common SNPs only explain a fraction (~50%) of the heritability (Wood et al., 2014; Yang et al., 2010) and an understanding of the genetic architecture contributing to height is lacking in global populations. For example, Rotwein (2020) concludes that four out of five significant GWAS hits associated with human height were most prevalent in European populations compared with others due to ascertainment bias. Therefore, SNPs thought to contribute to or tag causal SNPs for height may not be informative when considering diverse populations. Heritability of height has also been strongly associated with a collection of low allele frequency variants scattered across the genome, especially in regions of low linkage disequilibrium (Wainschtein et al., 2021). Therefore, the likelihood of inheriting low-MAF deleterious genotypes together is higher in individuals with longer ROH and thus more likely to result in reduced height (Ceballos et al., 2018; Mooney et al., 2018). However, contrary to previous findings that have associated elevated levels of ROH with decreased height (Joshi et al., 2015; McQuillan et al., 2012), we do not find this association within the Himba, even when controlling for fosterage status, sex, and different age cohorts (YOB). This result may be due to other environmental variables we were not able to include, measurement error, or low statistical power.

Previous work has reported an effect of fosterage on height among Himba children, especially girls, and this effect was replicated among an adult sample (Prall & Scelza, 2017). Here, the effect on adults was not replicated. There are several reasons why the two sets of results might differ. First, the two datasets were mostly nonoverlapping, as individuals included here must have both anthrometrics and DNA to be included. Second, here we use repeated measures for adults when available, which should better account for measurement error. Third, sample sizes for adults, particularly women for whom the effect was strongest in the initial analysis, are small in both datasets, as the 2017 analysis was focused mainly on differences among children. Finally, because fosterage mainly affects children, its effects on adult stature may be muted by myriad other factors that occur during adolescence, including age at first reproduction (which can occur before growth is complete), and age at marriage which would signal another change in household and consequent resource access and workload.

We include YOB as a covariate in our models to not only account for the overall effect of age on height, but also to capture any differences in environment for different age cohorts. Previous studies have found that drought can lead to a permanent loss of height in young children (Hoddinott & Kinsey, 2001). Because the Himba have also experienced periods of drought (Bollig, 1997), we aimed to control for this effect; however, we did not observe any evidence of such effects in our data. More detailed data on whether drought occurred during critical periods in a person's growth trajectory might reveal differences we do not see here.

The LCT gene, associated with lactase persistence, is known to be a source of beneficial calorie intake (Gerbault et al., 2011) and may have differential effects on growth patterns. If the lactase persistence genotype of the LCT gene continues to express throughout childhood into adulthood, the ability to digest milk into adulthood may lead to nutritional advantages over individuals who are not able to digest milk, potentially affecting adult height. Future studies might examine the prevalence of lactase persistence within the Himba, as recent research has identified the presence of LCT gene alleles within other African pastoralist groups (Hollfelder et al., 2020). Therefore, unknown environmental and socioecological factors that extend beyond the realm of our covariates may be outweighing a relationship between height and F_{ROH}, thus masking a relationship between these two variables.

Another reason we were unable to replicate previous results could be due to sampling error during data collection. Variation in measurements from the raw height data may have been dependent on inconsistent, oftentimes vertical hairstyles sported by Himba women. This, as well as inconsistency in measurement technique over the course of an 8-year period, may be a potential source of error and should be taken into consideration. Although the variance in height measurements for individuals with multiple height measurements is mostly very low (median = 0.72), higher variances, likely due to the aforementioned reasons, increases the average individual variance to 3.12. Another likely reason that we were unable to replicate previous results is that this dataset is underpowered. Previous studies (Joshi et al., 2015; McQuillan et al., 2012) have used thousands of individuals from variable environments (i.e., different countries, lifestyles, diets). Although our sample comes from a much more homogenous environment, our sample size of 245 individuals is underpowered to detect significant effects of F_{ROH} on height. In our power simulations assuming effect sizes described in Europeans, a sample size of 250 resulted in 15%–20% power of detection, while a sample sizes of 1000 and 2000 resulted in 50%–53% and 80%–84% power, respectively. Based on these results, we would need either a much larger sample size or F_{ROH} to have a much larger effect

26927691, 0, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/ajpa.24660, Wiley Online Library on [24/11/2022]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

than previously reported in order to detect a significant effect of F_{ROH} on height. Because our power simulations used the effect size of F_{ROH} and proportion of variance explained by F_{ROH} reported in urban European cohorts with thousands of individuals from more heterogenous environments, and our study uses individuals from a more homogenous environment, our power calculations are likely conservative.

Although an effect of F_{ROH} on height may be too small to detect in our dataset, elevated F_{ROH} in this population could have detectable effects on other phenotypes depending on the effect size of individual variants or location of ROH segments. F_{ROH} highlights the expression of recessive deleterious alleles scattered across the genome. If phenotypes other than height are affected by variants with larger effect sizes, then the effect of F_{ROH} on these other traits could be detected. Furthermore, if a deleterious recessive variant that disrupts the function of an important gene or causes loss of function of that gene is found in ROH, the phenotypic, medical, or fitness effect could be very large and thus more easily detected. Therefore, variance in the location of ROH segments and individual variant effect sizes are important to consider when investigating the effect of recessive variants on phenotypes, and future studies should consider investigating a possible effect of F_{ROH} on other phenotypes.

Our work sheds light on the complex association between homozygous genotypes and height, demonstrating the need for further examination of the effect level of homozygosity on height and epigenetic influence within the Himba. Unlike previously reported findings with larger samples (Joshi et al., 2015; McQuillan et al., 2012), we do not find a significant effect of F_{ROH} on height in this smaller sample. This may be due to being underpowered but may also imply that phenotypic expression resulting from increased homozygosity may vary from population to population.

4 | MATERIALS AND METHODS

4.1 | Ethics statement

Ethical approval for this study was granted by the University of California, Los Angeles (IRB-10-000238), the State University of New York, Stony Brook (IRB-636415-12), and was approved by the Namibian Ministry of Home Affairs and the University of Namibia Office of Academic Affairs and Research. Chief Basekama Ngombe provided permission to work in the community and local approval of the study. Community leaders were actively involved in discussions regarding what genetic data could be used for, who would have access to it, and whether there was a for-profit element involved (there was not). Individual informed consent, and for minors parental assent, was obtained orally from all participants given low rates of literacy in the community. Care was taken to protect participants' privacy, for example via a double-blind procedure for DNA collection (Scelza et al., 2020). These data were collected as part of the longitudinal Kunene Rural Health and Demography Project, which has been working in the community since 2010.

4.2 | Study population

The Himba reside in the Kunene Region of Northwestern Namibia and Southern Angola, but our study was conducted specifically in the community of Omuhonga. The Himba are a seminomadic, agropastoralist, nonmarket-integrated population, who are closely related to the Herero. Himba social norms promote a preference for first-cousin consanguinity in arranged marriages, particularly the first, but some of these marriages may never be consummated, and "love matches" also exist. Divorce and polygyny are common, and extramarital concurrent partnerships are frequent and socially accepted for both men and women. Thus, women have children with multiple men. These children, unless fostered out, reside with their mother, and when a woman is married, her husband becomes the social father of all her children, even if he is not the biological father.

4.3 | Genetic data

DNA was collected with informed consent from individuals via saliva using the prepIT-L2P kit and protocol. Individuals were then genotyped on one of two arrays, H3Africa or MEGAex. These datasets previously underwent quality control measures and were filtered using PLINK for missingness greater than 5%, a MAF less than or equal to 1%, and a Hardy–Weinberg equilibrium exact test with a p-value below 0.0001 as described in Scelza et al (2020) and documented in dbGaP phs001995.v1.p1. We thinned the H3Africa array to match the SNP density of the MEGAex array and selected autosomes only (H3Africa: 755,660 SNPs, MEGAex: 755,423 SNPs). Each dataset was then filtered to keep adults (18+ years old) who had associated height data.

To ensure that there were no batch effects across platforms or due to thinning the H3Africa SNP density, we performed a test (also in Swinford et al., 2022) where we thinned the array 10 separate times, identified ROH using a 1500 kb minimum threshold, and calculated F_{ROH} for each individual (as described below in *Methods: Measuring Homozygosity*). We then took the absolute value of the difference between the F_{ROH} values for each test and the original F_{ROH} value (used in analyses) and calculated a root-mean-squared-error value of 0.00058. Thus, there is no significant difference between tests and therefore no batch effects due to different SNPs.

4.4 | Measuring homozygosity

We used a custom R script to calculate the length of the genome that was tested on each array and then divided the total ROH identified in the genome by the length of the genome to calculate F_{ROH} for individuals (https://github.com/hennlab/HimbaHeight). ROH was identified using PLINK at a 1500 kb length threshold, allowing for a scanning window of 50 SNPs, a maximum of two missing SNPs and one heterozygote. We also called ROH using a 500 kb threshold and a 5000 kb threshold to test our model using F_{ROH} calculated using different

26927691, 0, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/ajpa.24660, Wiley Online Library on [24/11/2022]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

length thresholds. After F_{ROH} calculations were complete in the separate H3Africa and MEGAex datasets, they were merged to create a single dataset.

We explored using Garlic to estimate ROH. We used the error (0.001), winsize (30), auto-winsize, and auto-winsize-step (5) parameters to run Garlic. The selected thresholds chosen by Garlic for each dataset were ~530 kb and ~1900 kb, which are very similar to the thresholds we used.

4.5 | Height and fosterage data

Height data was collected over a period of 8 years, spanning from 2010 to 2018, resulting in multiple height measurements for most individuals. We removed height measurements that were taken before individuals turned 18 years old and accounted for varying numbers of height measurements among individuals in our models. We removed outliers that fell above or below three standard deviations from the mean. This resulted in the removal of one male whose height was greater than three standard deviations above the mean. Next, we merged in fosterage status data and filtered out 15 individuals who did not have associated fosterage data. After these quality filtering steps, our final dataset consisted of 245 individuals.

4.6 | Heritability

We used GCTA to create a GRM and calculate narrow sense heritability. We used the merged dataset that includes all individuals across both platforms. The H3Africa and MEGAex arrays were merged using PLINK and SNPs that had a missingness rate greater than 5% across platforms were removed, as in Scelza et al. (2020), resulting in a final SNP count of 368,749 SNPs. We created a GRM for all 245 individuals and calculated heritability using the "reml" flag in GCTA. Because the GCTA requires a file with a single phenotype value per individual for this function, we used the mean height value for individuals who have multiple height measurements.

4.7 | Statistical modeling

The baseline model was created to isolate any effect of F_{ROH} on height by assessing the effect of all the variables except for F_{ROH} on height and was created using raw height data instead of averaged heights for individuals. Using raw height scores allowed us to account for multiple measurements per individual, whereas estimated height values resulting from averaging varying height measurements could potentially mask a signal between F_{ROH} and height. This baseline mixed effects model was run using the coxme package in R and predicts height using YOB, sex, fosterage status, and the GRM (see *Methods: Heritability*) (https://github.com/hennlab/HimbaHeight). We allowed for a relationship between sex and fosterage and varied the intercept by individual ID to account for multiple height measurements per adult. To assess whether F_{ROH} effects height, we created a second model by adding the F_{ROH} 1500 variable into our baseline model.

To include our GCTA-generated GRM into the model, we read the binary file into R and converted the values to a matrix. We then converted this matrix to a BDS matrix for use in the mixed effects model function, lmekin(). Random chance, especially in small datasets, can sometimes produce a small amount of collinearity in the matrix resulting in GRM being nonpositive definite. The model cannot handle these values; thus, we bended the matrix, by eigendecomposing it and correcting one of slightly negative eigenvalues to a slightly positive value (10e-4) and made sure the bended matrix is highly similar to the original GRM via Mantel correlation (~1), prior to including it in the model.

4.8 | Power calculations

We calculated our power for detection of a significant effect of F_{ROH} on height using basic simulations in R (https://github.com/hennlab/HimbaHeight). Our code takes in a user-assigned sample size and assigns individuals a F_{ROH} value drawn from the observed distribution of our sample. Height values are simulated using the effect size of F_{ROH} and non-F_{ROH} effects from gaussian noise, and with the proportion of variance attributed to F_{ROH} that are reported in previous work (0.4%, McQuillan et al., 2012). Specifically, the portion of height affected by F_{ROH} can be calculated as F_{ROH} multiplied by the expected effect size (−0.12 SD of height per 1% increase of F_{ROH}, McQuillan et al., 2012), and the remaining portion of height is drawn from a normal distribution. We then ran 10 replicates of a linear model where simulated height is the response variable and simulated F_{ROH} is the predictor at samples sizes of $n = 250$, 1000, and 2000. We report the p-values in ranges for different sample sizes.

AUTHOR CONTRIBUTIONS

Natalie Swinford: Formal analysis (equal); investigation (equal); methodology (equal); supervision (equal); visualization (equal); writing – original draft (supporting); writing – review and editing (lead). **Brenna Gallagher:** Formal analysis (equal); investigation (equal); methodology (supporting); visualization (equal); writing – original draft (lead). **Jacob Sheehama:** Project administration (equal). **Meng Lin:** Formal analysis (supporting). **Sean Prall:** Data curation (equal); investigation (supporting); methodology (supporting). **Brooke Scelza:** Data curation (equal); funding acquisition (equal); investigation (supporting). **Brenna Henn:** Conceptualization (equal); funding acquisition (equal); investigation (supporting); resources (equal); supervision (equal).

ACKNOWLEDGMENTS

This research was funded by the NSF (BCS-1534682 to B.A.S) and NIH (R35GM133531 to B.M.H). We would like to thank the Himba for their participation and hospitality. We thank the reviewers and editors for a close reading of the manuscript, which was greatly improved by their comments.

26927691, 0, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/ajpa.24660, Wiley Online Library on [24/11/2022]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

DATA AVAILABILITY STATEMENT

The genetic data that support the findings of this study are available in dbGaP with accession number phs001995.v1.p1 at https://www.ncbi.nlm.nih.gov/gap/advanced_search/?TERM=phs001995.v1.p1. The nongenetic data and R code are also publicly available on GitHub at https://github.com/hennlab/HimbaHeight.

ORCID

Natalie Swinford https://orcid.org/0000-0001-6414-8555
Brooke Scelza https://orcid.org/0000-0001-5875-8875

REFERENCES

Bittles, A. H. (2002). Endogamy, consanguinity and community genetics. *Journal of Genetics*, *81*(3), 91–98. https://doi.org/10.1007/BF02715905

Bollig, M. (1997). Risk and risk minimisation among himba pastoralists in northwestern Namibia. *Nomadic Peoples*, *1*(1), 66–89.

Ceballos, F. C., Hazelhurst, S., Clark, D. W., Agongo, G., Asiki, G., Boua, P. R., Xavier Gómez-Olivé, F., Mashinya, F., Norris, S., Wilson, J. F., & Ramsay, M. (2020). Autozygosity influences cardiometabolic disease-associated traits in the AWI-gen sub-Saharan African study. *Nature Communications*, *11*(1), 5754. https://doi.org/10.1038/s41467-020-19595-y

Ceballos, F. C., Joshi, P. K., Clark, D. W., Ramsay, M., & Wilson, J. F. (2018). Runs of homozygosity: Windows into population history and trait architecture. *Nature Reviews Genetics*, *19*(4), 220–234. https://doi.org/10.1038/nrg.2017.109

Gerbault, P., Liebert, A., Itan, Y., Powell, A., Currat, M., Burger, J., Swallow, D. M., & Thomas, M. G. (2011). Evolution of lactase persistence: An example of human niche construction. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *366*(1566), 863–877. https://doi.org/10.1098/rstb.2010.0268

Harpending, H., & Pennington, R. (1990). Herero households. *Human Ecology*, *18*(4), 417–439. https://doi.org/10.1007/BF00889466

Heckerman, D., Gurdasani, D., Kadie, C., Pomilla, C., Carstensen, T., Martin, H., Ekoru, K., Nsubuga, R. N., Ssenyomo, G., Kamali, A., Kaleebu, P., Widmer, C., & Sandhu, M. S. (2016). Linear mixed model for heritability estimation that explicitly addresses environmental variation. *Proceedings of the National Academy of Sciences of the United States of America*, *113*(27), 7377–7382. https://doi.org/10.1073/pnas.1510497113

Hoddinott, J., & Kinsey, B. (2001). Child growth in the time of drought. *Oxford Bulletin of Economics and Statistics*, *63*(4), 409–436. https://doi.org/10.1111/1468-0084.t01-1-00227

Hollfelder, N., Babiker, H., Granehäll, L., Schlebusch, C. M., & Jakobsson, M. (2021). The genetic variation of lactase persistence alleles in Sudan and South Sudan. *Genome biology and evolution*, *13*(5), evab065. https://doi.org/10.1093/gbe/evab065

Joshi, P. K., Esko, T., Mattsson, H., Eklund, N., Gandin, I., Nutile, T., Jackson, A. U., Schurmann, C., Smith, A. V., Zhang, W., Okada, Y., Stančáková, A., Faul, J. D., Zhao, W., Bartz, T. M., Concas, M. P., Franceschini, N., Enroth, S., Vitart, V., … Wilson, J. F. (2015). Directional dominance on stature and cognition in diverse human populations. *Nature*, *523*(7561), 459–462. https://doi.org/10.1038/nature14618

Marouli, E., Graff, M., Medina-Gomez, C., Lo, K. S., Wood, A. R., Kjaer, T. R., Fine, R. S., Lu, Y., Schurmann, C., Highland, H. M., Rüeger, S., Thorleifsson, G., Justice, A. E., Lamparter, D., Stirrups, K. E., Turcot, V., Young, K. L., Winkler, T. W., Esko, T., … Lettre, G. (2017). Rare and low-frequency coding variants alter human adult height. *Nature*, *542*(7640), 186–190. https://doi.org/10.1038/nature21039

McQuillan, R., Eklund, N., Pirastu, N., Kuningas, M., McEvoy, B. P., Esko, T., Corre, T., Davies, G., Kaakinen, M., Lyytikäinen, L.-P., Kristiansson, K., Havulinna, A. S., Gögele, M., Vitart, V., Tenesa, A., Aulchenko, Y., Hayward, C., Johansson, A., Boban, M., … ROHgen Consortium. (2012). Evidence of inbreeding depression on human height. *PLoS Genetics*, *8*, e1002655. https://doi.org/10.1371/journal.pgen.1002655

Mooney, J. A., Huber, C. D., Service, S., Sul, J. H., Marsden, C. D., Zhang, Z., Sabatti, C., Ruiz-Linares, A., Bedoya, G., Costa Rica/Colombia Consortium for Genetic Investigation of Bipolar Endophenotypes, Freimer, N., & Lohmueller, K. E. (2018). Understanding the hidden complexity of Latin American population isolates. *American Journal of Human Genetics*, *103*(5), 707–726. https://doi.org/10.1016/j.ajhg.2018.09.013

Nabulsi, M. M., Tamim, H., Sabbagh, M., Obeid, M. Y., Yunis, K. A., & Bitar, F. F. (2003). Parental consanguinity and congenital heart malformations in a developing country. *American Journal of Medical Genetics Part A*, *116A*(4), 342–347. https://doi.org/10.1002/ajmg.a.10020

Pemberton, T. J., Absher, D., Feldman, M. W., Myers, R. M., Rosenberg, N. A., & Li, J. Z. (2012). Genomic patterns of homozygosity in worldwide human populations. *The American Journal of Human Genetics*, *91*(2), 275–292. https://doi.org/10.1016/j.ajhg.2012.06.014

Prall, S. P., & Scelza, B. A. (2017). Child fosterage and sex-biased nutritional outcomes among Namibian pastoralists. *American Journal of Human Biology*, *29*(6), e23058. https://doi.org/10.1002/ajhb.23058

Rotwein, P. (2020). Revisiting the population genetics of human height. *Journal of the endocrine Society*, *4*(4), bvaa025. https://doi.org/10.1210/jendso/bvaa025

Scelza, B., Prall, S., & Starkweather, K. (2021). The role of spousal separation on norms related to gender and sexuality among himba pastoralists. *Social Sciences*, *10*(5), 174. https://doi.org/10.3390/socsci10050174

Scelza, B. A., Prall, S. P., Swinford, N. A., Gopalan, S., Atkinson, E. G., McElreath, R., Sheehama, J., & Henn, B. M. (2020). High rate of extrapair paternity in a human population demonstrates diversity in human reproductive strategies. *Science Advances*, *6*(8), eaay6195. https://doi.org/10.1126/sciadv.aay6195

Scelza, B. A. (2011). Female choice and extrapair paternity in a traditional human population. *Biology Letters*, *7*(6), 889–891. https://doi.org/10.1098/rsbl.2011.0478

Scelza, B. A., Atkinson, E. G., Prall, S., McElreath, R., Sheehama, J., & Henn, B. M. (2020). The ethics and logistics of field-based genetic paternity studies. *Evolutionary Human Sciences*, *2*, e22. https://doi.org/10.1017/ehs.2020.23

Scelza, B. A., & Silk, J. B. (2014). Fosterage as a system of dispersed cooperative breeding: Evidence from the Himba. *Human Nature*, *25*(4), 448–464. https://doi.org/10.1007/s12110-014-9211-6

Silventoinen, K., Sammalisto, S., Perola, M., Boomsma, D. I., Cornes, B. K., Davis, C., Dunkel, L., de Lange, M., Harris, J. R., Hjelmborg, J. V. B., Luciano, M., Martin, N. G., Mortensen, J., Nisticò, L., Pedersen, N. L., Skytthe, A., Spector, T. D., Stazi, M. A., Willemsen, G., & Kaprio, J. (2003). Heritability of adult body height: A comparative study of twin cohorts in eight countries. *Twin Research*, *6*(5), 399–408. https://doi.org/10.1375/136905203770326402

Swinford, N., Prall, S., Williams, C., Sheehama, J., Scelza, B., & Henn, B. (2022). Increased homozygosity due to endogamy results in fitness consequences in a human population. *BioRxiv*. https://doi.org/10.1101/2022.07.25.501261

Szpiech, Z. A., Blant, A., & Pemberton, T. J. (2017). GARLIC: Genomic Autozygosity regions likelihood-based inference and classification. *Bioinformatics*, *33*(13), 2059–2062. https://doi.org/10.1093/bioinformatics/btx102

Szpiech, Z. A., Xu, J., Pemberton, T. J., Peng, W., Zöllner, S., Rosenberg, N. A., & Li, J. Z. (2013). Long runs of homozygosity are enriched for deleterious variation. *The American Journal of Human Genetics*, *93*(1), 90–102. https://doi.org/10.1016/j.ajhg.2013.05.003

26927691, 0, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/ajpa.24660, Wiley Online Library on [24/11/2022]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

Visscher, P. M., Hill, W. G., & Wray, N. R. (2008). Heritability in the genomics era — Concepts and misconceptions. *Nature Reviews Genetics*, *9*(4), 255–266. https://doi.org/10.1038/nrg2322

Wainschtein, P., Jain, D., Zheng, Z., TOPMed Anthropometry Working Group, NHLBI Trans-Omics for Precision Medicine (TOPMed) Consortium, Cupples, L. A., Shadyab, A. H., McKnight, B., Shoemaker, B. M., Mitchell, B. D., Psaty, B. M., Kooperberg, C., Liu, C. T., Albert, C. M., Roden, D., Chasman, D. I., Darbar, D., Lloyd-Jones, D. M., Arnett, D. K., Regan, E. A., … Visscher, P. M. (2022). Assessing the contribution of rare variants to complex trait heritability from whole-genome sequence data. *Nature genetics*, *54*(3), 263–273. https://doi.org/10.1038/s41588-021-00997-7

Williams, C. M., Scelza, B. A., Daya, M., Lange, E. M., Gignoux, C. R., & Henn, B. M. (2020). A rapid, accurate approach to inferring pedigrees in endogamous populations. *BioRxiv*, *2020*, 1–24. https://doi.org/10.1101/2020.02.25.965376

Wood, A. R., Esko, T., Yang, J., Vedantam, S., Pers, T. H., Gustafsson, S., Chu, A. Y., Estrada, K., Luan, J.'a., Kutalik, Z., Amin, N., Buchkovich, M. L., Croteau-Chonka, D. C., Day, F. R., Duan, Y., Fall, T., Fehrmann, R., Ferreira, T., Jackson, A. U., … Madden, P. A. F. (2014). Defining the role of common variation in the genomic and biological architecture of adult human height. *Nature Genetics*, *46*(11), 1173–1186. https://doi.org/10.1038/ng.3097

Yang, J., Benyamin, B., McEvoy, B. P., Gordon, S., Henders, A. K., Nyholt, D. R., Madden, P. A., Heath, A. C., Martin, N. G., Montgomery, G. W., Goddard, M. E., & Visscher, P. M. (2010). Common SNPs explain a large proportion of the heritability for human height. *Nature Genetics*, *42*(7), 565–569. https://doi.org/10.1038/ng.608

Yengo, L., Wray, N. R., & Visscher, P. M. (2019). Extreme inbreeding in a European ancestry sample from the contemporary UK population. *Nature Communications*, *10*(1), 1–11. https://doi.org/10.1038/s41467-019-11724-6

Zaitlen, N., Kraft, P., Patterson, N., Pasaniuc, B., Bhatia, G., Pollack, S., & Price, A. L. (2013). Using extended genealogy to estimate components of heritability for 23 quantitative and dichotomous traits. *PLoS Genetics*, *9*(5), e1003520. https://doi.org/10.1371/journal.pgen.1003520

How to cite this article: Swinford, N., Gallagher, B., Sheehama, J., Lin, M., Prall, S., Scelza, B., & Henn, B. M. (2022). Examination of runs of homozygosity in relation to height in an endogamous Namibian population. *American Journal of Biological Anthropology*, 1–9. https://doi.org/10.1002/ajpa.24660

26927691, 0, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/ajpa.24660, Wiley Online Library on [24/11/2022]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License