*Journal of Developmental Origins of Health and Disease*

**www.cambridge.org/doh**

# Associations between epigenome-wide DNA methylation and height-related traits among Sub-Saharan Africans: the RODAM study

Galatea Swart[1], Karlijn Meeks[2,4,5], Felix Chilunga[2], Andrea Venema[1], Charles Agyemang[2,5], Eva van der Linden[2,3] and Peter Henneman[1]

[1]Department of Human Genetics, Department of Human Genetics, Genome Diagnostic Laboratory, Amsterdam Reproduction and Development, Research Institute, Amsterdam University Medical Centers, Amsterdam, The Netherlands; [2]Department of Public and Occupational Health, Amsterdam University Medical Centers, University of Amsterdam, Amsterdam Public Health Research Institute, Amsterdam, The Netherlands; [3]Department of Vascular Medicine, Amsterdam University Medical Centers, University of Amsterdam, Amsterdam Cardiovascular Sciences, Amsterdam, The Netherlands; [4]Center for Research on Genomics and Global Health, National Human Genome Research Institute, National Institute of Health, Bethesda, MD, USA and [5]Department of Medicine, Division of Endocrinology, Diabetes and Metabolism, The John Hopkins University School of Medicine, Baltimore, MD, USA

**Original Article**

**Cite this article:** Swart G, Meeks K, Chilunga F, Venema A, Agyemang C, van der Linden E, and Henneman P. (2023) Associations between epigenome-wide DNA methylation and height-related traits among Sub-Saharan Africans: the RODAM study. *Journal of Developmental Origins of Health and Disease* **14**: 658–669. doi: 10.1017/S2040174423000034X

Received: 4 July 2023
Revised: 19 September 2023
Accepted: 29 October 2023
First published online: 4 December 2023

**Keywords:**
DNA methylation; epigenetic wide association study; anthropometrics; RODAM study

**Corresponding author:**
G. Swart; Email: galateaswart@gmail.com

## Abstract

Human height and related traits are highly complex, and extensively research has shown that these traits are determined by both genetic and environmental factors. Such factors may partially affect these traits through epigenetic programing. Epigenetic programing is dynamic and plays an important role in controlling gene expression and cell differentiation during (early) development. DNA methylation (DNAm) is the most commonly studied epigenetic feature. In this study we conducted an epigenome-wide DNAm association analysis on height-related traits in a Sub-Saharan African population, in order to detect DNAm biomarkers across four height-related traits. DNAm profiles were acquired in whole blood samples of 704 Ghanaians, sourced from the Research on Obesity and Diabetes among African Migrants study, using the Illumina Infinium HumanMethylation450 BeadChip. Linear models were fitted to detect differentially methylated positions (DMPs) and regions (DMRs) associated with height, leg-to-height ratio (LHR), leg length, and sitting height. No epigenome-wide significant DMPs were recorded. However we did observe among our top DMPs five informative probes associated with the height-related traits: cg26905768 (leg length), cg13268132 (leg length), cg19776793 (height), cg23072383 (LHR), and cg24625894 (sitting height). All five DMPs are annotated to genes whose functions were linked to bone cell regulation and development. DMR analysis identified overlapping DMRs within the gene body of *HLA-DPB1* gene, and the *HOXA* gene cluster. In this first epigenome-wide association studies of these traits, our findings suggest DNAm associations with height-related heights, and might influence development and maintenance of these traits. Further studies are needed to replicate our findings, and to elucidate the molecular mechanism underlying human height-related traits.

## Introduction

Human height and height-related traits, such as leg-to-height ratio (LHR), leg length, and sitting height, are complex traits that are determined by both genetic and environmental factors. Moreover, these traits have been shown to have low correlations with each other and therefore are assumed as independent variables when utilized by epidemiologists to measure growth factors among populations at risk of prolonged malnutrition.[1] The longitudinal effects of environmental and genetic factors on height have been recorded since the first published Developmental Origin of Disease Hypothesis (DODH) by Baker. Investigating gene–environment interaction in the form of DNA methylation (DNAm) as part of DODH is of potential relevance, as only the use of genome-wide association studies (GWAS) loci scores could only explain 18.5–19.8% of inter-individual variation in height heritability.[2] Additionally, GWAS have identified 3290 loci associated with height and height-related traits, however those loci only represent 24.6% of the variance in height.[3,4] Environmental factors, such as nutritional shortage, are considered important factors determining height.[5–7] It has been shown that during periods of nutritional shortage populations generally have shorter statures, whereas during nutritional surplus, average population length goes up.[8–10] In the light of low explained variance of genetic factors and the important effect of environmental factors, studying epigenetic modifications associated with height and height-related traits is of interest, as they can reflect gene–environment interaction (Fig. 1).[1,11] Epigenetics is the study of heritable, yet reversible molecular modifications of DNA without altering the DNA sequence[12] and comprises histone

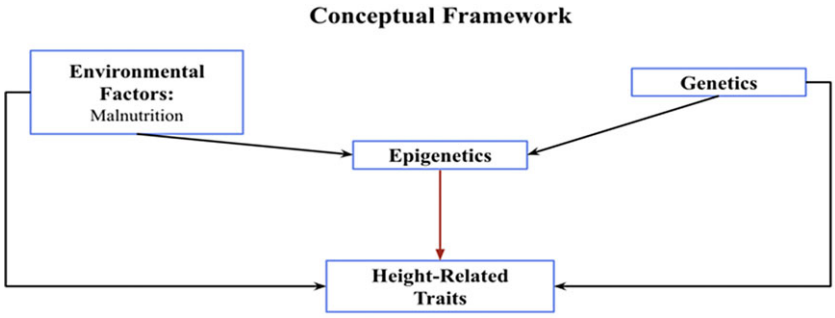


**Figure 1.** Conceptual framework diagram. The diagram demonstrates this study's exploration of DNA methylation profiles associated with height-related traits due to these traits being influenced by both genetics and environmental factors.

modifications, microRNA, and DNAm of which the latter is the most commonly studied. Epigenetic processes influence the phenotypic development, through regulation of gene expression.[12] Previous studies have found genes relating to bone cell development and maintenance to be under the influence of DNAm[13,14] but epigenome-wide association studies (EWAS) on height and height-related traits have not been performed. The aim of this study was to assess the relation between DNAm and height and height-related traits (leg length, sitting height, and LHR), in a population of Ghanaians, in order to increase our physiological understanding of growth regulation. This of particular relevance to a West African population, among whom an association between height and height-related with cardiometabolic conditions later in life was demonstrated.[15]

## Methods

### Study population

This study population of adults was derived from the Research on Obesity and Diabetes among African Migrants (RODAM) study. The data collection procedures for the RODAM study have been published previously.[16] The study was conducted from 2012 and 2015 with 6385 Ghanaians resident sampled from across in five geographical locations. Ghanaian individuals were randomly sampled from various urban and rural areas within Ghana and from migrant communities in Amsterdam, Berlin, and London[16] with the mean age for participants estimated to 51.2 (± 9.73 standard deviations (SD)). Ethical committees from participating institutions in Ghana, the Netherlands, Germany, and the UK approved the study before the start of data collection. Written informed consent was obtained from each study participant. All participants for the DNAm profiling were selected from a case–control design in which ~ 300 individuals were diabetic cases, ~300 were deemed diabetic control cases, and ~ 135 were obese control cases).[17] The EWAS was performed using the DNAm profiles of 736 participants of the RODAM cohort,[18] of which 713 samples passed quality control. Quality control procedures were based on those described by Meeks *et al.*[18] and Chen *et al.*[19] Individuals missing one or more height-related traits and those whose measurements were erroneously recorded were omitted from inclusion in the final sample for this study ($n = 704$).

### DNA isolation methylation profiling, processing, and quality control

High molecular DNAm was extracted from whole blood samples. The blood samples were processed manually, aliquoted, and stored at −20°C at local research locations. Samples were transported to central laboratories in each country to be registered and stored again −80°C.

The DNA extraction methylation profiling, processing, and quality control protocols of the RODAM samples have been described in detail previously.[20] In brief, bisulfite conversion of DNA and DNA extraction protocols were conducted in the Source BioScience laboratories, Nottingham, UK, using the Zymo EZ DNAm™ kit. Infinium® HumanMethylation450 BeadChip amplified and hybridized the converted DNA, thereby quantifying DNAm levels for ~ 485,000 CpG sites. Methylation levels were determined from the methylated or unmethylated intensities for each individual CpG site present within the array. Methylation levels were quantified via Beta values, with zero representing an unmethylated probe and one representing a fully methylated probe. M values subsequently were calculated and were applied in further analyses. Using R statistical software (version 4.2.0), quality control was performed using the *MethylAid* package (version 1.30.0). The raw 450k data was normalized using the *Minfi* package (version 1.42). Additionally probes annotated to the X and Y chromosomes were removed as well as cross-hybridised CpGs, and probes known to contain single nucleotide polymorphisms with a minor allele frequency of > 0.05.[19] These procedures reduced the data available for analysis to approximately 429,449 CpG sites per participant. Lastly, relative blood cell type distribution (CD8+, T lymphocytes, CD4+ T lymphocytes, natural killer cells, B cell, monocytes, and granulocytes) was estimated according to the methods of Houseman *et al.*[21]

### Physical measurements

All height and height-related measurements were standardized across all sampling locations.[1,22] Height was determined from a portable stadiometer (SECA 217) to the nearest 0.1 cm without participants wearing shoes. Sitting height was derived from measuring participants as they sat upright on a flat seat, and with their heads level, feet on the floor, and with the thighs unsupported. Then the sitting height distance (cm) from the floor to the top of the head was recorded. The leg length (cm) was calculated by subtracting an individual's sitting height from their total height. By dividing the calculated leg length by skeletal height, the LHR was estimated.[1,22]

### Differentially methylated positions

Differentially methylated positions (DMPs) were derived from multivariate linear regressions between height-related traits (independent variable) and DNAm M values (dependent variable) using the *Limma* package (version 3.52). Methylation M values were used to ensure normal distribution for the statistical analysis, while Beta values were used for interpretation

https://doi.org/10.1017/S204017442300034X Published online by Cambridge University Press

and visualization of data.[23] Models were adjusted for sex, age, estimated blood cell type proportions, and technical covariates (hybridisation batch and array position), because of correlation with DNAm in the principal components analysis. QQ plots were used to assess best model fit for each height-related trait (Fig. 2). False discovery rate (FDR) p-value adjustments were applied to reduce type I error through multiple testing. An FDR of < 0.05 was assumed epigenome-wide significance, while an FDR within a range of 0.05–0.5 was eligible for differentially methylated region (DMR) analysis. Since all traits analyzed in this study are to a greater or lesser extent correlated to each other, we assumed that traits were not fully independent and multiple test penalty was only applied per analyzed trait. CpG probes and genes were annotated using the Human Genome build 37 Illumina platform via the *IlluminaHumanMethylation450kanno.ilmn12.hg19* R package (Version 0.6, UCSC build). Lastly, in order to detect and remove undocumented genetic variation we ran a post hoc analysis on the top 25 DMPs using the Bioconductor package *gaphunter*,[24] applying the delta difference cluster (>2) threshold of 0.05.[24,25]

### Differentially methylated regions

DMRs were identified using two different R packages: *bumphunter* (version 1.38) and *DMRcate* (version 3.15). Both procedures generated results based on the lowest *p*-values and FDRs using fitted models similar to the DMP analysis.[20] For *bumphunter*, the M value cutoff determining the effect size for the DMR analysis was set at 0.0025, corresponding to an effect size of 0.25% M value difference per cm for the corresponding height-related trait. Here we applied 500 bootstrap permutations on the bumphunter models. DMRs were defined to compromise ≥ 3 CpGs. A family-wise error rate (FWER) < 0.2 was considered statistically significant. With the *DMRcate* analysis, FDR threshold cutoffs varied for every height-related trait, i.e. this threshold was relaxed to each traits' lowest observed FDR value minus 0.01. We assumed that DMRs with a Stouffer coefficient and a smoothed FDR of < 0.05 were epigenome-wide significant. A smoothed FDR is a method of refining the multiple-hypothesis test by implementing a weighted distribution. *DMRcate* uses a Gaussian kernel bandwidth for the smoothed-function estimation.[27]

### Biological relevance

Biological functions of DMPs and DMRs were assessed through the systematic search of multiple academic sources.[28–31] All probes and gene functions were required to be consistent across at least three independent sources for the function of the probe, or gene, to be considered informative and relevant to this study. The sources used to verify the function of identified DMPs, and DMRs include EWAS Atlas Database,[28] iMethyl,[29] and UCSC Genome.[30] The function of identified genes annotated to our top DMPs and DMRs were verified using UCSC Genome,[30] the National Library of Medicine's Gene Database,[31] as well as peer-reviewed article sourced from the PubMed database (https://pubmed.ncbi.nlm.nih.gov/). The purpose of these reviews was to confirm a probe/genes' citation in publications exploring metabolic conditions, growth factors, adverse environmental influenced development, or general probe/gene function.

### Pathway analysis

Enrichment pathway analysis identifies molecular and genetic mechanisms associated with specific DMPs. The MissMethyl[32] an enrichment procedure was conducted with the first 100 significant probes per anthropometric trait based on p-value and lowest FDR. Enrichment results were generated using MissMethyl package in R; generating results from databases are the Gene Ontology[33] (GO) and Kyoto Encyclopedia of Genes and Genomes[34] (KEGG). The purpose of the enrichment pathway analysis was to confirm if there existed a correlation connection between the molecular mechanism associated with our top 100 DMPs and embryonic development, growth factors, or height-related traits.

## Results

### Participants characteristics

The subset of the RODAM study presented, included 704 participants after quality control protocols (Table 1). Among this population, the mean age was 51.2 (± 9.73 SD), while average height was calculated at 164.04 cm (± 8.33), LHR was 0.50 (± 0.02), leg length 82.65 cm (± 5.21), and sitting height measuring 81.45 (± 4.72) (Table 1). While the various blood cells were observed to distribute at similar levels: CD8T+ lymphocytes, CD4T+ lymphocytes, natural killer cells (NK), B cell, monocytes, and granulocytes. When the subset was stratified between males and females, we observed that male participants were older and taller. The subset had an average age of 52 (± 9 SD) among males and 51 (± 10 SD) among females. Demographic analysis did not demonstrate a correlation between sex and a difference in blood cell type distributions, or habitual smoking (Table 1).

### Differentially methylated positions

We detected no epigenome-wide significant DMPs among the four height-related traits. Although not statistically significant, we evaluated the top 25 DMPs per trait. Among the leg length DMPs, we identified cg26905768 annotated to the body of *BMPER* and cg13268132 annotated to the promoter region of *TNFRSF11B,* which were both hypomethylated (Table 2). Top DMPs among height included hypomethylated cg19776793 annotated to the body of *SLC38A10,* and for LHR, hypomethylated cg23072383 annotated to the TSS1500 of *SLC35E4*. For sitting height, we associated a hypomethylated probe cg24625894 annotated to the promoter region of *SLC39A4*. Note that the latter three DMPs are all members of the *SLC30* gene family (Supplementary Table S1, S2, and S4). The possible informative nature of these probes will be discussed at length in the Discussion section. *Post hoc* analyses applying *gaphunter,* did not return any additional genetic variation among the 25 DMPs with the smallest p-values, per trait. Additionally, sex-stratified analysis did not reveal statistically significant DMPs associated with any of the traits.

### Differentially methylated regions

Next we aimed to detect DMRs by applying *DMRcate* for all traits. From the *DMRcate* analysis, height was associated with 1291 DMRs according to a FDR threshold cutoff of 0.12 of which 110 demonstrated significance (Supplementary Table S5). Sixteen significant DMRs were detected for sitting height out of a total of 2053 according to an FDR threshold cutoff of 0.27 (Supplementary Table S7). Leg length generated 1506 DMRs, however none of the DMRs were epigenome-wide significantly represented (Supplementary Table S6). Several DMRs were annotated within the same gene; these repetitive DMRs were observed in one or more height-related traits. These recurring DMRs were located within the body of the *HOXA* gene cluster (covering 23, and 26

https://doi.org/10.1017/S204017442300034X Published online by Cambridge University Press

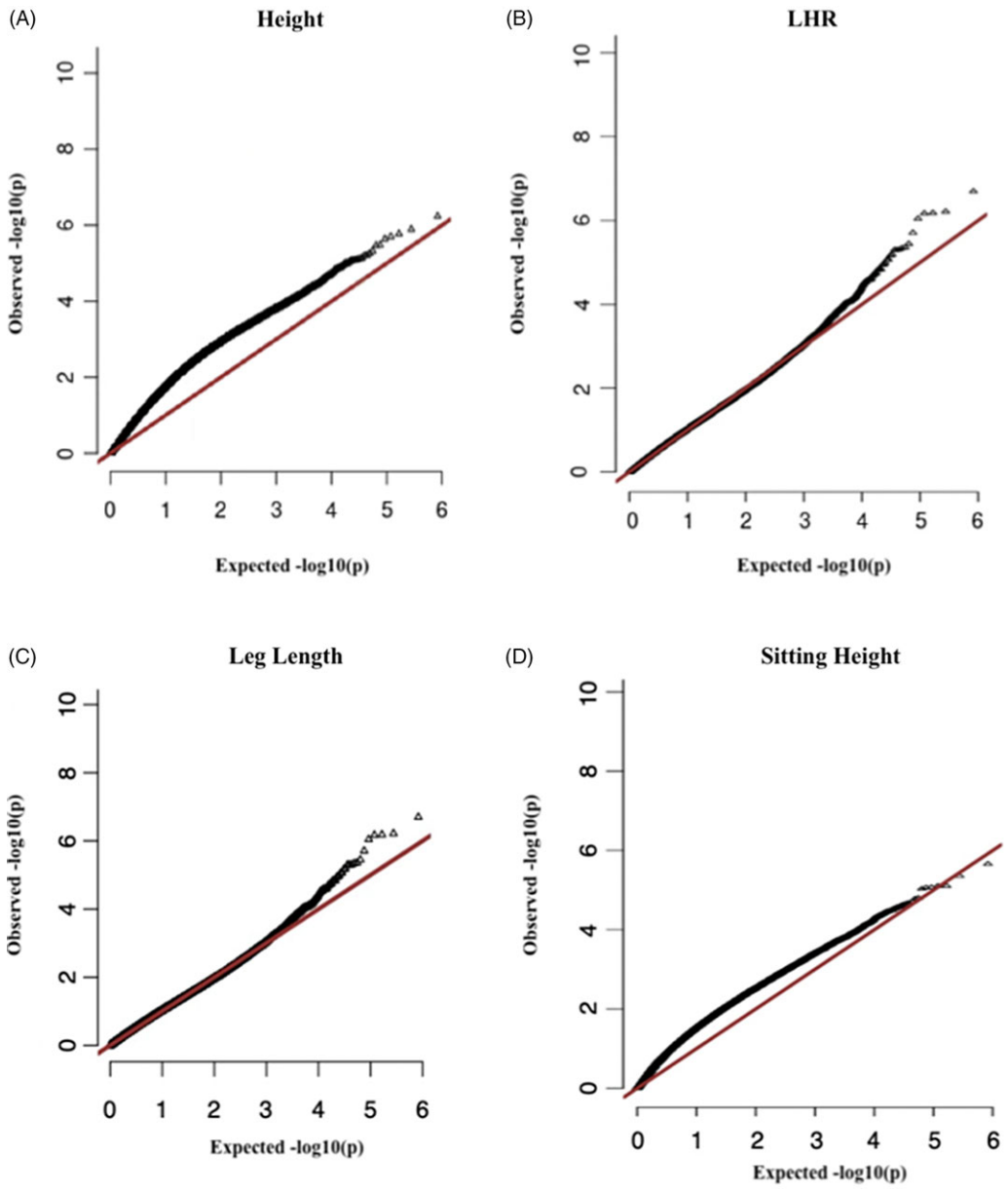


**Figure 2.** QQ plots visualizing the fit of the linear regression models used in DMP analysis for height (A), LHR (B), leg length (C), and sitting height (D).

DMPs for height and sitting height, respectively), *HLA-DPB1* (covering 23, and 24 DMPs for height and sitting height, respectively), and *HIC1* (covering 50, and 53 DMPs for height and sitting height, respectively), which were all three hypomethylated (Table 2, Figs. 3–5, Supplementary Table S7). DMR analyses applying *Bumphunter*, using an effect size cutoff of 0.0025, produced one significant DMR. This DMR was detected for both leg length and sitting height (covering 9, and 9 DMPs, respectively) and were located in the *HLA-DPB1* gene. Neither height nor LHR generated DMRs. Noteworthy, the DMRs observed within the *HLA-DPB1* gene associated with sitting height, were detected in both the *DMRcate* and *bumphunter* procedures, associated with sitting height.

### *Pathway analysis*

The pathway analysis produced no results which could be used to add or subtract credibility from our hypothesis. Several p-value significant molecular mechanisms relating to bone development and growth factors were observed, however all mechanisms were stipulated to have an FDR = 1. Zinc ion homeostasis was identified with sitting height using both GO ($n = 40$, DE = 3, $p = 0.001$, FDR = 1) and KEGG databases ($n = 8$, DE = 2, $p = 0.001$, FDR = 1). Leg length yielded pathways as well, however the most noteworthy was osteoclast differentiation ($n = 120$, DE = 4, $p = 0.004$, FDR = 1) which was gleaned from the KEGG database.

## Discussion

### *Key findings*

We conducted an exploratory association study to investigate associations between DNAm profiles and height-related traits using data of the RODAM study. Even though not epigenome-wide significant, we did identify potentially informative DMPs and DMRs, located in genes previously linked to the growth regulation. To our knowledge, none of the observed DMPs or DMRs have ever been associated with these height-related traits before. The relevant

https://doi.org/10.1017/S204017442300034X Published online by Cambridge University Press

**Table 1.** Characteristics of RODAM study including participants demographic information, lifestyle factors, height-related traits, and distribution of the cell types observed

| Demographics (standard deviation (SD)) | Participant characteristics | All participants | Male | Female | *P* value | Missing (%) |
|---|---|---|---|---|---|---|
| | *n* | 704 | 301 | 403 | | 0 |
| | Age, years (SD) | 51.2 (9.73) | 51.86 (9.08) | 50.53 (10.38) | 0.077 | 0 |
| | Location % (SD) | | | | **<0.001** | |
| | Amsterdam | 150 (21.3) | 89 (29.6) | 61 (15.1) | | 0 |
| | Berlin | 81 (11.5) | 56 (18.6) | 25 (6.2) | | 0 |
| | London | 130 (18.46) | 52 (17.3) | 78 (19.4) | | 0 |
| | Urban Ghana | 241 (34.23) | 72 (23.9) | 169 (41.9) | | 0 |
| | Rural Ghana | 102 (14.49) | 32 (10.6) | 70 (17.4) | | 0 |
| Lifestyle factors, % (SD) | | | | | | |
| | Habitual smoking | 2.07 (0.33) | 2.13 | 2.03 | **<0.001** | 2.8 |
| height-related traits, cm, (SD) | | | | | | |
| | Height | 164.04 (8.33) | 170.38 (6.74) | 159 (5.90) | **<0.001** | 0 |
| | Leg-height ratio (LHR) | 0.50 (0.02) | 0.50 (0.02) | 0.50 (0.02) | 0.784 | 0 |
| | Leg length | 82.65 (5.21) | 85.88 (4.64) | 80.24 (4.22) | **<0.001** | 0 |
| | Sitting height | 81.45 (4.72) | 85.3 (3.63)) | 79.07 (3.98) | **<0.001** | 0 |
| Cell types distribution % (SD) | | | | | | |
| | CD8+ T lymphocytes | 0.11 (0.05) | 0.10 (0.05) | 0.12 (0.04) | **<0.001** | 0 |
| | CD4+T lymphocytes | 0.18 (0.06) | 0.18 (0.06) | 0.18 (0.06) | 0.2 | 0 |
| | Natural killer | 0.11 (0.06) | 0.12 (0.06) | 0.10 (0.05) | **<0.001** | 0 |
| | B cell | 0.11 (0.03) | 0.00 (0.00) | 0.00 (0.00) | 0.0 | 0 |
| | Monocytes | 0.08 (0.02) | 0.08 (0.03) | 0.08 (0.02) | 0.79 | 0 |
| | Granulocytes | 0.46 (0.09) | 0.46 (0.10) | 0.45 (0.09) | 0.78 | 0 |

DMPs include multiple probes annotated to the *SLCA* gene family, across both height and sitting height. Leg length identified two CpGs within genes associated with bone cell regulation: cg26905768 annotated to *BMPER* and cg13268132 annotated to *TNFRSF11B*. *BMPER* specifically encodes a secreted protein that limits bone morphogenetic protein (*BMP*) function, inhibits BMP2- and BMP4-dependent osteoblast differentiation[35], and modulates BMP-dependent differentiation among endothelial cells.[32] The observed DMP was annotated to the body of the *BMPER* gene, and hypomethylated in this gene region generally associated with lower expression of the gene. In the context of regulation of leg length, this would be in line with our expectation that expression of bone cell genes should decrease upon aging. *TNFRSF11B*, encodes osteoprotegerin, a protein that regulates osteoclastogenesis inhibitory factors.[36,37] *TNFRSF11B's* role in bone cell resorption has been hypothesized as contributing to osteoporosis as well as other conditions caused by decreased bone density.[36,37] As hypomethylation of promoter regions generally is associated with more expression of the gene, it thus seems that *TNFRSF11B* is more expressed. As osteoclasts play a role in age-related decrease in bone mineral density, activity of this gene might be relevant given the average age of the study population.[36,37] The alignment of multiple SLCA genes has potential importance as this gene family assists in transport of protein, zinc, and iron throughout the body during development.[38] Specifically SLC39A4 gene (annotated to sitting height) is correlated with bone cell maintenance.[39] Unfortunately, pathway analysis returned no significant results. In summary, our observed DMPs associated with height-related traits have previously been associated with growth factors or bone cell regulation, and the observed methylation levels, in the respective locations, are mostly in accordance with what we would expect in height-related traits.

We identified several DMRs that were annotated to a hypomethylated *HLA-DPB1* gene, as well as hypomethylation among the *HOXA* genes cluster for at least two height-related traits. In both DMRs we observed hypomethylation in the body of each gene, which asserts less gene expression. The *HLA* gene family is mostly associated with chronic autoimmune diseases[31,40] and inflammatory conditions. *HLA-DPB1* is a class II *HLA* gene associated with the body's defense against infection.[31,40] This gene, however, has never been studied in the context of growth regulation. We therefore cannot assert a physiological rationale for multiple DMRs being detected amongst only two of our height-related traits. For both sitting height and LHR, we found DMRs annotated to the *HOXA* cluster. The homeobox, or *HOX*, gene family is highly influential during embryonic development in most species.[30,40] Researchers have demonstrated *HOXA* genes are associated with various forms of development, including skeletal regulation.[40] Therefore, hypomethylation of the body of the *HOXA* genes, the implication of less expression, is in line with our expected findings as *HOXA* gene misregulation or mutations leading to changes in skeletal development are correlated with

https://doi.org/10.1017/S204017442300034X Published online by Cambridge University Press

**Table 2.** Differentially methylated regions (DMRs) for height, sitting height, and leg length identified using two methods (*DMRcate* and *bumphunter*). DMRs with FWER > 0.2 or Stouffer coefficient > 0.05

| | DMR analysis | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| *DMRcate* | Trait | Chr | Start | End | Width | No. CpGs | Smoothed FDR | Stouffer | Direction of affect | UCSC Reference | Genes |
| | **Height** | | | | | | | | | | |
| | | 6 | 33047944 | 33049505 | 1562 | 23 | 1.86E-11 | 9.31E-05 | −1.85E-03 | Body | *HLA-DPB1, HLA-DPA1* |
| | | 7 | 27142427 | 27144302 | 1876 | 23 | 2.75E-21 | 1.66E-04 | −1.23E-03 | Body | *HOXA2* |
| | | 6 | 126068919 | 126070802 | 1884 | 15 | 1.21E-07 | 2.71E-04 | 1.46E-04 | Body | *HEY2, RP11-624M8.1* |
| | | 8 | 37962386 | 37963526 | 1141 | 15 | 4.12E-11 | 3.02E-04 | 8.34E-05 | 1st Intron | *ASH2L* |
| | | 2 | 190305524 | 190306491 | 968 | 16 | 1.62E-09 | 4.12E-04 | 1.08E-04 | Body | *WDR75* |
| | | 13 | 99852409 | 99853209 | 801 | 14 | 1.82E-08 | 4.90E-04 | 9.41E-05 | Body | *UBAC2, UBAC2-AS1* |
| | | 11 | 94277826 | 94279068 | 1243 | 11 | 2.15E-13 | 5.39E-04 | −8.72E-04 | 3'UTR | *PIWIL4, FUT4* |
| | | 2 | 33824170 | 33825072 | 903 | 13 | 1.62E-08 | 7.10E-04 | 8.96E-05 | | *FAM98A* |
| | | 20 | 60639404 | 60641183 | 1780 | 9 | 1.29E-06 | 8.07E-04 | −1.35E-04 | Body | *hsa-mir-3195, TAF4* |
| | | 14 | 81999503 | 82000858 | 1356 | 14 | 6.05E-09 | 8.41E-04 | 1.91E-04 | | *SEL1L* |
| | | 12 | 110841169 | 110842031 | 863 | 16 | 9.19E-09 | 9.81E-04 | 1.53E-04 | Promoter | *ANAPC7, RP11-478C19.2* |
| | | 4 | 119199352 | 119200372 | 1021 | 13 | 2.60E-09 | 1.01E-03 | 2.40E-04 | | *SNHG8* |
| | | 17 | 1956668 | 1959066 | 2399 | 50 | 3.98E-16 | 1.02E-03 | 1.07E-04 | Body | *HIC1* |
| | | 6 | 32811181 | 32813715 | 2535 | 57 | 2.46E-12 | 1.08E-03 | 1.26E-04 | 3'UTR | *TAPSAR1, PSMB9, PSMB8, TAP1* |
| | | 18 | 12947517 | 12948653 | 1137 | 13 | 2.74E-08 | 1.17E-03 | 1.52E-04 | Promoter | *SEH1L* |
| | | 6 | 33290281 | 33292029 | 1749 | 49 | 4.53E-13 | 1.17E-03 | 1.35E-04 | 3'UTR | *DAXX* |
| | | 12 | 123459152 | 123460194 | 1043 | 15 | 6.44E-08 | 1.34E-03 | 1.17E-04 | Body | *OGFOD2, RP11-197N18.2, ABCB9* |
| | | 15 | 72766139 | 72767565 | 1427 | 12 | 3.16E-08 | 1.50E-03 | 8.65E-05 | Body | *ARIH1, RP11-1007O24.3* |
| | | 4 | 56212072 | 56212685 | 614 | 8 | 2.90E-08 | 1.60E-03 | 1.65E-04 | Body | *SRD5A3* |
| | | 3 | 38035098 | 38036060 | 963 | 8 | 1.27E-07 | 1.61E-03 | 1.76E-04 | | *VILL* |
| | | 10 | 72141924 | 72142824 | 901 | 11 | 3.24E-10 | 1.62E-03 | 1.03E-04 | Body | *LRRC20* |
| | | 12 | 51664245 | 51664655 | 411 | 9 | 3.45E-08 | 1.70E-03 | 1.56E-04 | Body | *DAZAP2, SMAGP* |
| | | 16 | 11349023 | 11350746 | 1724 | 16 | 5.67E-09 | 1.74E-03 | 1.39E-04 | Promoter; Body | *RMI2, SOCS1* |
| | | 15 | 65822346 | 65823539 | 1194 | 15 | 2.28E-10 | 1.84E-03 | 3.31E-05 | | *PTPLAD1* |
| | | 3 | 122283003 | 122283684 | 682 | 14 | 7.49E-08 | 1.85E-03 | 8.27E-05 | Body | *DTX3L, PARP9* |
| | **Sitting height** | | | | | | | | | | |
| | | 7 | 27169208 | 27171528 | 2321 | 26 | 2.94E-11 | 4.00E-03 | −1.70E-03 | Promoter; Body | *HOXA-AS2, HOXA-AS3, HOXA3* |
| | | 3 | 45077070 | 45078075 | 1006 | 9 | 2.13E-08 | 1.20E-02 | −2.02E-03 | 3'UTR | *EXOSC7, CLEC3B* |
| | | 16 | 53406901 | 53407808 | 908 | 8 | 2.85E-08 | 2.60E-02 | 3.95E-03 | | *RP11-44F14.2, RP11-44F14.1* |

(*Continued*)

**Table 2.** (*Continued*)

| DMR analysis | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| *DMRcate* | Trait | Chr | Start | End | Width | No. CpGs | Smoothed FDR | Stouffer | Direction of affect | UCSC Reference | Genes |
| | | 8 | 145637525 | 145639181 | 1657 | 8 | 1.54E-06 | 3.00E-02 | 1.44E-03 | | *SLC39A4* |
| | | 13 | 88328009 | 88329407 | 1399 | 8 | 3.53E-05 | 3.30E-02 | −2.10E-03 | | *SLITRK5* |
| | | 17 | 7311030 | 7312081 | 1052 | 9 | 2.64E-05 | 3.70E-02 | −1.28E-03 | Body | *NLGN2* |
| | | 13 | 20989142 | 20990243 | 1102 | 6 | 2.93E-06 | 3.80E-02 | −1.79E-03 | Body | *CRYL1* |
| | | 2 | 43327937 | 43328657 | 721 | 7 | 1.10E-05 | 4.00E-02 | −1.93E-03 | Body | *AC093609.1* |
| | | 11 | 64108241 | 64110178 | 1938 | 10 | 2.19E-04 | 4.10E-02 | −7.96E-04 | Body | *CCDC88B* |
| | | 3 | 113160071 | 113160821 | 751 | 10 | 5.05E-07 | 4.20E-02 | −1.34E-03 | Body; 1st Intron | *WDR52* |
| | | 1 | 949392 | 949893 | 502 | 5 | 4.47E-07 | 4.40E-02 | −2.16E-03 | Body | *ISG15* |
| | | 6 | 33047944 | 33050124 | 2181 | 24 | 4.24E-08 | 4.40E-02 | −3.03E-03 | Body | *HLA-DPB1, HLA-DPA1* |
| | | 1 | 164545553 | 164546143 | 591 | 5 | 6.63E-05 | 4.50E-02 | −8.47E-04 | | *PBX1* |
| | | 19 | 8590567 | 8591776 | 1210 | 5 | 2.58E-04 | 4.60E-02 | −1.15E-03 | Body | *MYO1F* |
| | | 1 | 27675934 | 27677240 | 1307 | 6 | 4.53E-04 | 4.70E-02 | −1.37E-03 | | *SYTL1* |
| | | 6 | 26225258 | 26226256 | 999 | 7 | 4.71E-04 | 5.00E-02 | −1.05E-03 | Promoter; Body | *HIST1H3E* |

| Bumphunter | Trait | Chr | Start | End | P value | No.cpgs | Fwer | Affect size | UCSC reference | Gene |
|---|---|---|---|---|---|---|---|---|---|---|
| | **Leg length** | | | | | | | | | |
| | | 6 | 33048529 | 33048875 | 3.89E-04 | 9 | 0.198 | 0.0025 | Body | *HLA-DPB1* |
| | **Sitting height** | | | | | | | | | |
| | | 6 | 33048529 | 33048875 | 3.50E-04 | 9 | 0.178 | 0.0025 | Body | *HLA-DPB1* |

Are listed in *Supplemental Table X*. Chr = Chromosome, UCSC Reference: the gene feature based on the UCSC genome browser, Gene: the annotation performed via. Illumina R package version 0.06, UCSC build HG37.

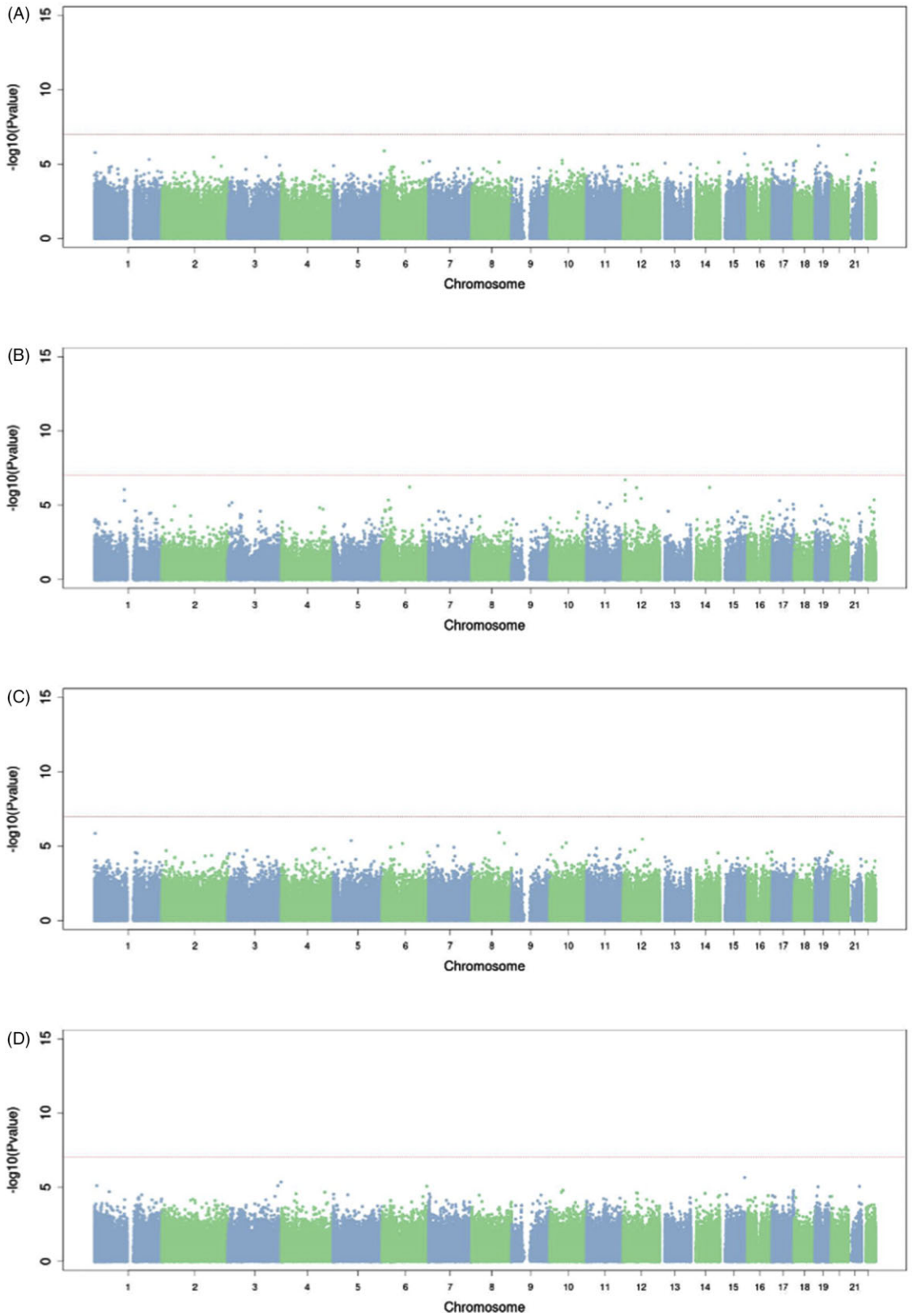


**Figure 3.** Manhattan plot of epigenome-wide p-values for height (A), LHR (B), leg length (C), and sitting height (D).

https://doi.org/10.1017/S204017442300034X Published online by Cambridge University Press

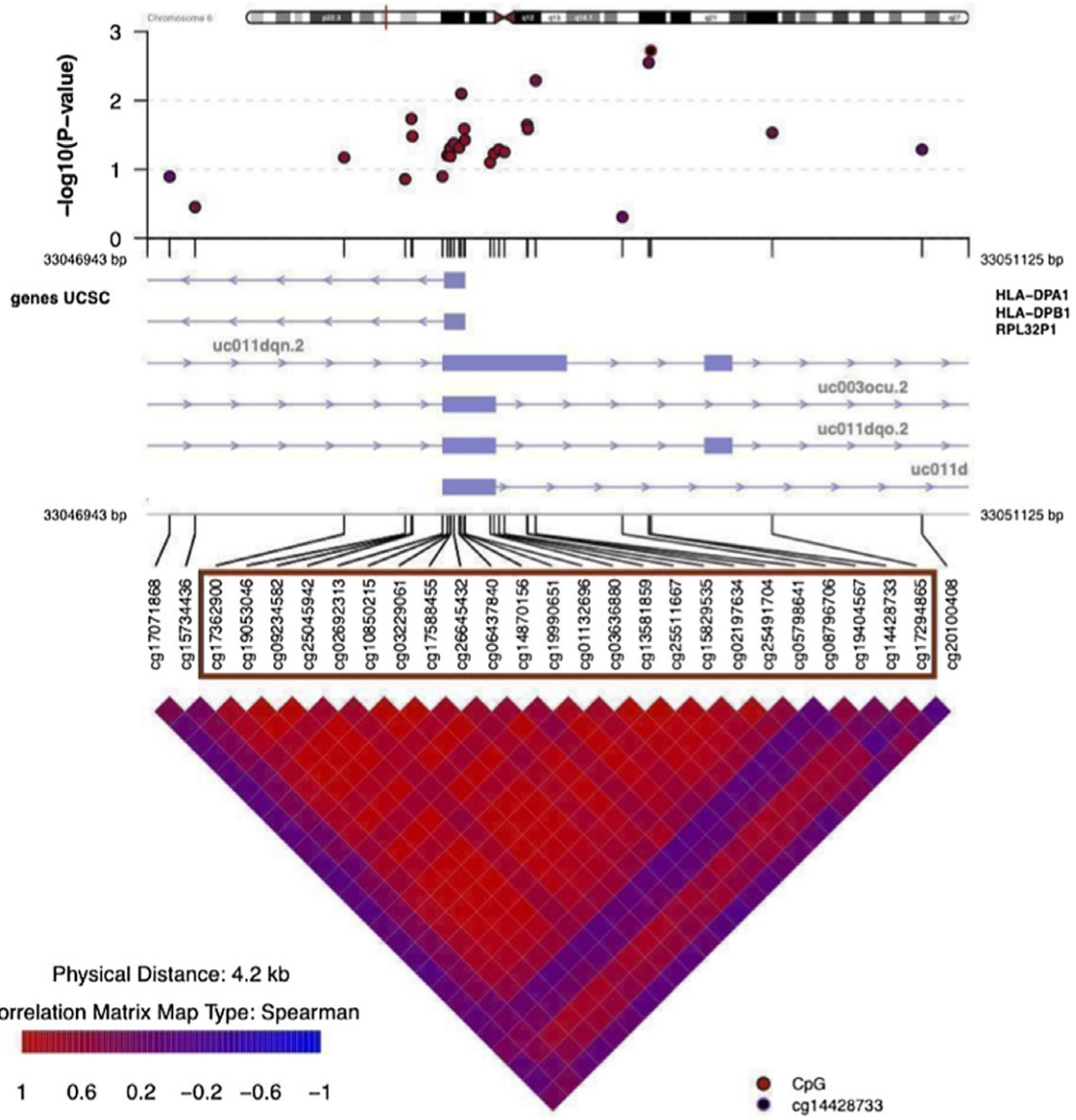


**Figure 4.** Comet plot of the differentially methylated region identified *HLA-DPB1* gene based on the sitting height trait results. Plot show a differentially methylated region in chromosome 6, 33.646943 and 33.051125 megabases (mb) obtained via *DMRcate*. The red box highlights the probes annotated to the DMR. Correlations are measured using the spearman rambling coefficient.

early development,[40,42] not during middle life, which is the average age of the of our RODAM study subset. Our findings could serve as a starting point for further research to assess the role of the *HOXA* gene in human skeletal development.

Although there is no explicit information about exposure to malnutrition in early life, the average age of the RODAM cohort allows for speculation on exposure to food shortage and the impact on height. In Ghana, two periods of famine occurred in the latter half of the 20$^{th}$ century. In the era between 1960 and 1974 there were widespread food shortages in rural areas throughout the country.[43,44] At the time, an estimated 70% of Ghanaian citizens were estimated to reside in rural regions.[41] More widely known is the 1981–1983 famine leading to a nationwide occurrence of malnutrition and inaccessibility to cash crops.[43] Based on the mean age of RODAM participants, most would have been born or experienced infancy in Ghana between 1960 and 1974. This early-life exposure to famine could have impacted DNAm patterns, thereby affecting height and related traits.

## Strengths and limitations

A major benefit of using the RODAM study population is its relative genetic homogeneity, meaning that all individuals stemmed from one region in Ghana, the Ashanti region, and the majority of participants identified as originating from one ethnic groups, the Akan.[16] Additionally, the prevalence of confounding factors in DNAm studies like smoking and alcohol consumption were very low and therefore we assume that our results were not impacted.

Our study has several limitations. The RODAM cohort was designed to investigate metabolic disease and immigration-related health concerns among Sub-Saharan African populations. It was not the primary purpose of the RODAM cohort to explore height-related traits, or the confounding factors that contribute toward height development. This difference between our use of the RODAM cohort and its original epidemiological purpose could contribute to our overall lack of statistically significant results.

https://doi.org/10.1017/S204017442300034X Published online by Cambridge University Press

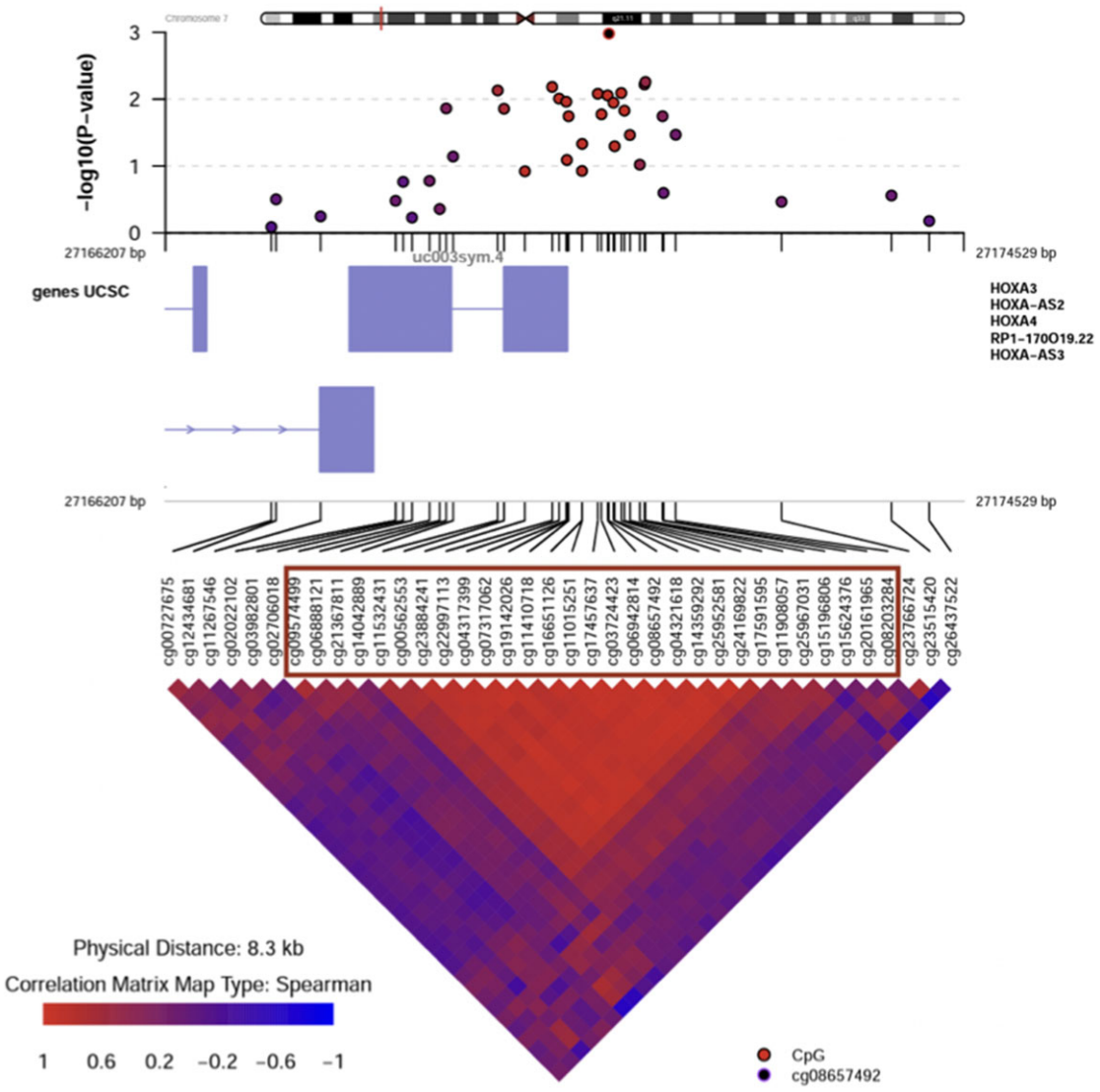


**Figure 5.** Comet plot of the differentially methylated region identified *HOXA* cluster gene based on the sitting height trait results. Plot show a differentially methylated region in chromosome 7, 27.167207 and 27.173529 megabases (mb) obtained via *DMRcate*. The red box highlights the probes annotated to the DMR. Correlations are measured using the spearman rambling coefficient.

Note the additional co-factors of RODAM representing a mean age of 50 for its participants and that height-related traits have an ~ 80% heritability rate. This does pose the possibility that epigenetic signaling was diminished due to environmental factors over the course of participants' lifetimes. We suggest the use of younger cohorts in future. Additionally due to the relatively small sample size, as well as the effect size, our study was limited in statistical power to detect epigenome-wide significant DMPs. The small number of participants in our study subset meant that a sex-stratified analysis was not possible, but differences based on sex were not expected as correlation between sex and height-related traits were shown to be low. We relaxed significance thresholds for both DMP and DMR protocols, potentially resulting in false positive findings. Moreover, we cannot definitively state whether or not our height-related traits were affected by any nutritional deprivation during early-life development, as we did not apply a longitudinal design. This limits our capacities to make statements on causality. Additionally, we assumed that DNAm patterns of height-related genes remain stable in adult, however, this assumption cannot be verified in this study. Differential epigenetic aberrations might echo in later life without having a current functional effect, but would be involved in other traits, biological mechanisms, or co-morbidity such as cardiovascular disease.[18] Lastly, this study was conducted utilizing epigenetic profiles based on whole bloods samples, though we focused on the height-related traits determined by bone development. As DNAm is tissue-specific, and DNAm profiles derived from blood might therefore not be representative of processes occurring in bone tissue.[45] Epigenetic profiles derived from bone-related tissues would help to validate our findings.

## Conclusion

In this proof-of-principle study, we found several potential DNAm markers for height, and height-related traits annotated to genes involved in skeletal and early development in humans. These findings can serve as a starting point to further elucidate the role of DNAm in skeletal development. Future research including a larger

https://doi.org/10.1017/S204017442300034X Published online by Cambridge University Press

sample size, information on early-life factors, DNAm profiles derived from bone tissue and translational research will all help to gain better physiological understanding of growth regulation in human development.

**Supplementary material.** The supplementary material for this article can be found at https://doi.org/10.1017/S204017442300034X.

**Acknowledgments.** The authors are grateful to Prof. Karien Stronks (Department of Public Health, Academic Medical Center, The Netherlands) and Prof. Ama de-Graft Aikins (Regional Institute for Populations Studies, University of Ghana, Ghana) for their conceptualization and contributions to the RODAM study. We thank Nadine Huckauf and Candy Kalischke (Department of Endocrinology and Metabolism, Charite University Medicine, Berlin, Germany) for their excellent support in biochemical characterization of the participants and data handling. Jan van Straalen from the Academic Medical Center should be acknowledged for his valuable support with standardization of the laboratory procedures, and the AMC Biobank for support in biobank management and storage of collected samples. We also extend our personal gratitude to Alejandro Conde-Carrillo and Maika Conde-Carrillo. Lastly, the authors are grateful to the advisory board members for their valuable support in shaping the methods, to the research assistants, interviewers, and other staff of the five research locations who have taken part in gathering the data and, most of all, to the Ghanaian volunteers participating in this project.

**Author contribution.** Eva van der Linden and Peter Henneman are contributed equally to this work.

**Financial support.** The RODAM study was supported by the European Commission under the Seventh Framework Program (Grant Number: 278901). Additional support was funded by the Intramural Research Program of the National Institutes of Health in the Center for Research on Genomics and Global Health (CRGGH). The CRGGH is supported by the National Human Genome Research Institute, the National Institute of Diabetes and Digestive and Kidney Diseases, the Center for Information Technology, and the Office of the Director at the National Institutes of Health (1ZIAHG200362). The study sponsor was not involved in the design of the study; the collection, analysis and interpretation of data; writing the report; nor the decision to submit the report for publication.

**Competing interests.** None.

**Ethical standards.** The RODAM study was conducted according to the guidelines laid down in the Declaration of Helsinki. All procedures involving human subjects were reviewed and approved by the respective ethics committees in Ghana, the Netherlands, the UK, and Germany. Written informed consent was obtained from all participants.

## References

1. Bogin B, Varela-Silva MI. Leg length, body proportion, and health: a review with a note on beauty. *Int J Environ Res Public Health*. 2010; 7(3), 1047–1075.
2. Shah S, Bonder MJ, Marioni RE, *et al.* Improving phenotypic prediction by combining genetic and epigenetic associations. *Am J Hum Genet*. 2015; 97(1), 75–85.
3. Yengo L, Sidorenko J, Kemper KE, *et al.* Meta-analysis of genome-wide association studies for height and body mass index in ~700000 individuals of European ancestry. *Hum Mol Genet*. 2018; 27(20), 3641–3649.
4. Manolio TA, Collins FS, Cox NJ, *et al.* Finding the missing heritability of complex diseases. *Nature*. 2009; 461(7265), 747–753.
5. Jantz LM, Jantz RL. Secular change in long bone length and proportion in the United States, 1800-1970. *Am J Phys Anthropol*. 1999; 110(1), 57–67.
6. Johnston LW, Harris SB, Retnakaran R, *et al.* Short leg length, a marker of early childhood deprivation, is associated with metabolic disorders underlying type 2 diabetes: the PROMISE cohort study. *Diabetes Care*. 2013; 36(11), 3599–3606.
7. Li L, Dangour AD, Power C. Early life influences on adult leg and trunk length in the 1958 British birth cohort. *Am J Hum Biol*. 2007; 19(6), 836–843.
8. van Abeelen AF, Elias SG, Bossuyt PM, *et al.* Famine exposure in the young and the risk of type 2 diabetes in adulthood. *Diabetes*. 2012; 61(9), 2255–2260.
9. Eriksson JG. Early growth, and coronary heart disease and type 2 diabetes: experiences from the Helsinki birth cohort studies. *Int J Obes (Lond)*. 2006; 30(S4), S18–22.
10. Barker DJ, Martyn CN. The maternal and fetal origins of cardiovascular disease. *J Epidemiol Community Health*. 1992; 46(1), 8–11.
11. Varela-Silva MI, Frisancho AR, Bogin B, *et al.* Behavioral, environmental, metabolic and intergenerational components of early life undernutrition leading to later obesity in developing nations and in minority groups in the U.S.A. *Coll Antropol*. 2007; 31(1), 39–46.
12. Hamilton JP. Epigenetics: principles and practice. *Dig Dis*. 2011; 29(2), 130–135.
13. Liu Y, Zhang XL, Chen L, *et al.* Epigenetic mechanisms of bone regeneration and homeostasis. *Prog Biophys Mol Biol*. 2016; 122(2), 85–92.
14. White TB B, Folkens P. *Human Osteology*. 3 edn. 2011. Academic Press, New York, pp. 688.
15. van der Heijden TGW, Chilunga FP, Meeks KAC, *et al.* The magnitude and directions of the associations between early life factors and metabolic syndrome differ across geographical locations among migrant and non-migrant Ghanaians-the RODAM study. *Int J Environ Res Public Health*. 2021; 18(22), 11996.
16. Agyemang C, Beune E, Meeks K, *et al.* Rationale and cross-sectional study design of the research on obesity and type 2 Diabetes among African migrants: the RODAM study. *BMJ Open*. 2014; 4(3), e004877.
17. Chilunga FP, Henneman P, Venema A, *et al.* DNA methylation as the link between migration and the major noncommunicable diseases: the RODAM study. *Epigenomics*. 2021; 13(9), 653–666.
18. Krzyzewska I.M. LP, Mul AN, van der Laan L, Yim AYFL, Cobben JM. Expression quantitative trait methylation analysis identifies whole blood molecular footprint in fetal alcohol spectrum disorder (FASD). *Int J Mol Sci*. 2023; 24(7), 6601.
19. Chen YA, Lemire M, Choufani S, *et al.* Discovery of cross-reactive probes and polymorphic CpGs in the illumina infinium HumanMethylation450 microarray. *Epigenetics*. 2013; 8(2), 203–209.
20. Meeks KAC, Henneman P, Venema A, *et al.* Epigenome-wide association study in whole blood on type 2 diabetes among sub-saharan African individuals: findings from the RODAM study. *Int J Epidemiol*. 2019; 48(1), 58–70.
21. Houseman EA, Accomando WP, Koestler DC, *et al.* DNA methylation arrays as surrogate measures of cell mixture distribution. *BMC Bioinformatics*. 2012; 13(1), 86.
22. Boateng D, Danquah I, Said-Mohamed R, *et al.* Early-life exposures and cardiovascular disease risk among Ghanaian migrant and home populations: the RODAM study. *J Dev Orig Health Dis*. 2020; 11(3), 250–263.
23. Du P, Zhang X, Huang CC, *et al.* Comparison of beta-value and M-value methods for quantifying methylation levels by microarray analysis. *BMC Bioinformatics*. 2010; 11(1), 587.
24. Fortin JP, Triche TJ Jr., Hansen KD. Preprocessing, normalization and integration of the illumina humanMethylationEPIC array with minfi. *Bioinformatics*. 2017; 33(4), 558–560.
25. Andrews SV, Ladd-Acosta C, Feinberg AP, Hansen KD, Fallin MD. Gap hunting' to characterize clustered probe signals in illumina methylation array data. *Epigenetics Chromatin*. 2016; 9(1), 56.
26. Peters TJ, Buckley MJ, Chen Y, Smyth GK, Goodnow CC, Clark SJ. Calling differentially methylated regions from whole genome bisulphite sequencing with DMRcate. *Nucleic Acids Res*. 2021; 49(19), e109–e109.
27. Li M, Zou D, Li Z, *et al.* EWAS atlas: a curated knowledgebase of epigenome-wide association studies. *Nucleic Acids Res*. 2019; 47(D1), D983–D8.
28. Hachiya T, Furukawa R, Shiwa Y, *et al.* Genome-wide identification of inter-individually variable DNA methylation sites improves the efficacy of epigenetic association studies. *NPJ Genom Med*. 2017; 2(1), 11.
29. Kent WJ, Sugnet CW, Furey TS, *et al.* The human genome browser at UCSC. *Genome Res*. 2002; 12(6), 996–1006.

30. Brown GR, Hem V, Katz KS, *et al.* Gene: a gene-centered information resource at NCBI. *Nucleic Acids Res*. 2015; 43(D1), D36–D42.
31. Florencio-Silva R, Sasso GR, Sasso-Cerri E, Simoes MJ, Cerri PS. Biology of bone tissue: structure, function, and factors that influence bone cells. *Biomed Res Int*. 2015; 2015, 421746–17.
32. Phipson B, Oshlack A. DiffVar: a new method for detecting differential variability with application to methylation in cancer and aging. *Genome Biol*. 2014; 15(9), 2–17.
33. Harris MA, Clark J, Ireland A, *et al.* The gene ontology (GO) database and informatics resource. *Nucleic Acids Res*. 2004; 32(Database issue), D258–61.
34. Kanehisa M. The KEGG database. *Novartis Found Symp*. 2002; 247, 91–101.
35. Yang P, Troncone L, Augur ZM, Kim SSJ, McNeil ME, Yu PB. The role of bone morphogenetic protein signaling in vascular calcification. *Bone*. 2020; 141, 115542.
36. Richards JB, Rivadeneira F, Inouye M, *et al.* Bone mineral density, osteoporosis, and osteoporotic fractures: a genome-wide association study. *Lancet*. 2008; 371(9623), 1505–1512.
37. Boronova I, Bernasovska J, Macekova S, *et al.* TNFRSF11B gene polymorphisms, bone mineral density, and fractures in slovak postmenopausal women. *J Appl Genet*. 2015; 56(1), 57–63.
38. Demontiero O, Vidal C, Duque G. Aging and bone loss: new insights for the clinician. *Ther Adv Musculoskelet Dis*. 2011; 4(2), 61–76.
39. Jeong J, Eide DJ. The SLC39 family of zinc transporters. *Mol Aspects Med*. 2013; 34(2-3), 612–619.
40. Liu B, Shao Y, Fu R. Current research status of HLA in immune-related diseases. *Immun Inflamm Dis*. 2021; 9(2), 340–350.
41. Mallo M. Reassessing the role of hox genes during vertebrate development and evolution. *Trends Genet*. 2018; 34(3), 209–217.
42. Rux DR, Wellik DM. Hox genes in the adult skeleton: novel functions beyond embryonic development. *Dev Dyn*. 2017; 246(4), 310–317.
43. Campbell MON. *The Political Ecology of Agricultural History in Ghana*, 2013. Nova Science Publishers Inc, New York, pp. 106.
44. Nott JD. *Between Famine and Malnutrition: Spatial Aspects of Nutritional Health During Ghana's Long Twentieth Century, c. 1896-2000*, 2016. University of Leeds, University of Leeds.
45. Hernandez Cordero AI, Yang CX, Li X, *et al.* The blood DNA methylation clock grimAge is a robust surrogate for airway epithelia aging. *Biomedicines*. 2022; 10(12), 3094.

https://doi.org/10.1017/S204017442300034X Published online by Cambridge University Press