International Journal of Epidemiology, 2019, 58–70
doi: 10.1093/ije/dyy171
Advance Access Publication Date: 10 August 2018
Original article

Epigenetics & Genetics

Epigenome-wide association study in whole blood on type 2 diabetes among sub-Saharan African individuals: findings from the RODAM study

Karlijn A C Meeks ,[1]* Peter Henneman,[2] Andrea Venema,[2] Juliet Addo,[3] Silver Bahendeka,[4] Tom Burr,[5] Ina Danquah,[6,7] Cecilia Galbete ,[6] Marcel M A M Mannens,[2] Frank P Mockenhaupt,[8] Ellis Owusu-Dabo,[9] Charles N Rotimi,[10] Matthias B Schulze,[6] Liam Smeeth,[3] Joachim Spranger,[11,12,13] Mohammad H Zafarmand,[1,14] Adebowale Adeyemo[10] and Charles Agyemang[1]

[1]Department of Public Health, Amsterdam Public Health Research Institute, Academic Medical Center, Amsterdam UMC, University of Amsterdam, Amsterdam, The Netherlands, [2]Department of Clinical Genetics, Research Institute for Reproduction and Development, Academic Medical Center, Amsterdam UMC, University of Amsterdam, Amsterdam, The Netherlands, [3]Department of Non-communicable Disease Epidemiology, London School of Hygiene and Tropical Medicine, London, UK, [4]Mother Kevin Postgraduate Medical School (MKPGMS), Uganda Martyrs University, Kampala, Uganda, [5]Genomics Department, Source BioScience, Nottingham, UK, [6]Department of Molecular Epidemiology, German Institute of Human Nutrition Potsdam-Rehbruecke, Nuthetal, Germany, [7]Institute for Social Medicine, Epidemiology and Health Economics, Charité – Universitaetsmedizin Berlin, Berlin, Germany, [8]Institute of Tropical Medicine and International Health, Charité – University Medicine Berlin, Berlin, Germany, [9]School of Public Health, Kwame Nkrumah University of Science and Technology, Kumasi, Ghana, [10]Center for Research on Genomics and Global Health, National Human Genome Research Institute, Bethesda, MD, USA, [11]Department of Endocrinology and Metabolism, Charité – University Medicine Berlin, Berlin, Germany, [12]Partner site Berlin, German Centre for Cardiovascular Research (DZHK), Berlin, Germany, [13]Center for Cardiovascular Research (CCR), Charité – University Medicine Berlin, Berlin, Germany and [14]Department of Clinical Epidemiology, Biostatistics and Bioinformatics, Amsterdam Public Health Research Institute, Academic Medical Center, Amsterdam UMC, University of Amsterdam, Amsterdam, The Netherlands

*Corresponding author. Department of Public Health, Amsterdam Public Health Research Institute, Academic Medical Center, Amsterdam UMC, University of Amsterdam, Meibergdreef 9, 1105AZ Amsterdam, The Netherlands. E-mail: k.a.meeks@amc.uva.nl

Editorial decision 17 July 2018; Accepted 25 July 2018

Abstract

Background: Type 2 diabetes (T2D) results from a complex interplay between genetics and the environment. Several epigenome-wide association studies (EWAS) have found DNA methylation loci associated with T2D in European populations. However, data from African

populations are lacking. We undertook the first EWAS for T2D among sub-Saharan Africans, aiming at identifying ubiquitous and novel DNA methylation loci associated with T2D.

Methods: The Illumina 450k DNA-methylation array was used on whole blood samples of 713 Ghanaian participants (256 with T2D, 457 controls) from the cross-sectional Research on Obesity and Diabetes among African Migrants (RODAM) study. Differentially methylated positions (DMPs) for T2D and HbA1c were identified through linear regression analysis adjusted for age, sex, estimated cell counts, hybridization batch, array position and body mass index (BMI). We also did a candidate analysis of previously reported EWAS loci for T2D in non-African populations, identified through a systematic literature search.

Results: Four DMPs [*cg19693031* (*TXNIP*), *cg04816311* (*C7orf50*), *cg00574958* (*CPT1A*), *cg07988171* (*TPM4*)] were associated with T2D after correction for inflation by possible systematic biases. The most strongly associated DMP—*cg19693031*, *TXNIP* ($P = 2.6E\text{-}19$)—showed hypomethylation in T2D cases compared with controls. Two out of the four DMPs [*cg19693031* (*TXNIP*), *cg04816311* (*C7orf50*)] remained associated with T2D after adjustment for BMI, and one locus [*cg07988171* (*TPM4*)] that has not been reported previously.

Conclusions: In this first EWAS for T2D in sub-Saharan Africans, we have identified four DMPs at epigenome-wide level, one of which is novel. These findings provide insight into the epigenetic loci that underlie the burden of T2D in sub-Saharan Africans.

Key words: DNA-methylation, Africans, type 2 diabetes, epigenetic epidemiology, EWAS, RODAM study

Key Messages

- This is the first epigenome-wide study on type 2 diabetes in sub-Saharan Africans.
- A novel DNA methylation locus was associated with type 2 diabetes in a sample of Ghanaians.
- The three DNA methylation loci most strongly associated with type 2 diabetes were also previously reported for other populations, indicating that these are ubiquitous.
- The four epigenetic loci that were genome-wide associated with type 2 diabetes explained 25% of variance in type 2 diabetes.

Background

Type 2 diabetes (T2D) is estimated to affect 415 million adults worldwide, contributing greatly to mortality and health care costs.[1] It is a complex multifactorial disease, meaning that many genetic loci as well as environmental factors play a role.[2,3] The explained heritability of T2D is 5–10%,[4] of an estimated overall heritability of about 10–20%,[4] suggesting that additional genetic factors remain to be identified and that gene-environment interactions may contribute to the onset of T2D.[5]

Epigenetics is the study of mitotically heritable yet reversible molecular modifications to the DNA without altering the DNA sequence.[6] Epigenetic processes play a role in organism development, cell differentiation and genome stability. Furthermore, epigenetics is considered a means of interaction between genes and environment, and one of its main read-outs, DNA methylation, has been previously associated with T2D in European populations.[7,8] Additionally, a great variety of health-related behaviours, such as diet and physical activity, can affect DNA methylation even in a short time span.[9,10] T2D is frequently accompanied by health-related behaviours such as reduced physical activity and unhealthy diet.[11] These health-related behaviours may affect T2D risk through epigenetics.[12]

The prevalence of T2D in sub-Saharan Africa is increasing rapidly, and the International Diabetes Federation estimates that the number of people with T2D in the African region may rise from 14.2 million to 34.2 million in 2040.[1] In

addition, sub-Saharan African migrants in Europe are about three times more likely to have T2D compared with European populations.[13] This increasing T2D burden in sub-Saharan Africa, as well as among migrants, likely reflects increasing obesity rates, diet changes and a reduction in physical activity.[14] Environmental risk factors may be mediated (at least partially) through changes in DNA methylation. Previous epigenome-wide association studies (EWAS) found DNA methylation at multiple loci associated with T2D in blood,[7,8] pancreatic,[15] adipose,[16] liver[17] and muscle tissue[18] from European populations. However, similar data on any tissue from sub-Saharan African populations are lacking. Sub-Saharan Africans are characterized by wide genetic variation, as well as marked differences in genetic variation and linkage disequilibrium, when compared with populations from other continents. The differences in the genetics of sub-Saharan African populations compared with Europeans, as well as differences in environmental exposures and health behaviours, could influence epigenetic loci associated with traits and disorders in sub-Saharan Africans.

In the present study, we conducted an EWAS for T2D in sub-Saharan Africans, aiming at identifying DNA methylation loci in blood associated with T2D using an epigenome-wide association study. Second, we aimed at identifying loci shared between populations of different ancestry (i.e. ubiquitous variants), as well as loci that have not been previously reported in other populations and are therefore potentially unique to sub-Saharan Africans.

Methods

Study population

This study is part of the **R**esearch on **O**besity and **D**iabetes among **A**frican **M**igrants (RODAM) study. Details of the design and data collection of the RODAM study are described elsewhere.[19,20] In brief, the RODAM study collected samples and data between 2012 and 2015 on 6385 Ghanaians resident in five geographical locations. An individual was considered Ghanaian when born in Ghana and at least one of the parents was born in Ghana, or when both parents were born in Ghana. Ethical approval was obtained from the ethical committees of the institutions involved in Ghana, The Netherlands, Germany and the UK. Written informed consent was obtained from all study participants. The EWAS was performed using a subset of 736 participants (265 T2D cases, 471 controls) of the RODAM cohort, of which 713 samples (256 T2D cases, 457 controls) passed quality control (see below). All T2D cases not on glucose-lowering medication were selected with a group of T2D controls on an approximate 1:2 ratio, respectively, which is generally accepted to favour statistical power. The final sample size had over 80% power to detect a 6% difference in methylation between T2D cases and controls with epigenome-wide significance, and over 99% power to detect a 10% difference in methylation between cases and controls.[21]

Phenotypic measurements

Information on demographics and self-reported T2D were collected by self- or interviewer-administered questionnaire. Self-reported T2D comprised an affirmative response to the question 'Have you ever been diagnosed with diabetes by a doctor or health care worker?' Participants were physically examined; height (SECA 217 stadiometer) and weight (SECA 877 scale) were assessed in light clothing, and body mass index (BMI) was calculated (kg/m^2). Participants were instructed to fast from 10 pm on the evening before the physical examination. Fasting venous blood samples were collected, and fasting glucose in mmol/l was measured by the ABX PENTRA 400 and HbA1c in mmol/mol by high-performance liquid chromatography (TOSOH G8 HPLC analyser). T2D was defined according to self-reported diabetes and/or fasting glucose $\geq$7.0 µl/l. In a sensitivity analysis, we additionally applied an HbA1c cut-off of $\geq$6.5% (48 µl/mol) to the T2D definition (Supplementary Material 1, available as Supplementary data at *IJE* online).

DNA methylation profiling and processing

DNA extraction and methylation profiling were performed on whole-blood samples by Source BioScience, Nottingham, UK. Bisulphite DNA treatment was achieved using the Zymo EZ DNA Methylation^{TM} kit, and the quality of the conversion was determined by high-resolution melting analyses.[22] The converted DNA was amplified and hybridized on the Illumina Human Methylation 450 K array, which measures DNA methylation levels of approximately 485 000 CpG sites. The samples were randomly divided over nine bisulphite conversion and hybridization batches. The quality control procedures are described in detail in Supplementary Material 2, available as Supplementary data at *IJE* online. After quality control a sample size of 713 remained for the current analyses.

Statistical analysis

Differentially methylated positions (DMPs)

Linear association analyses were performed in 'R', using the lmFit function from the *Limma* package to identify DMPs. DNA methylation levels were the dependent variable in all analyses. Age, sex, estimated cell counts and technical effects (hybridization batch and array position) were included as covariates both for the epigenome-wide

approach and for the candidate approach. Candidate loci were selected from the literature as described below. Correlation of the included covariates with DNA methylation is shown in Supplementary Material 2, Figures 2 and 3, available as Supplementary data at *IJE* online. Plate position was additionally added as covariate as identified in principal component analysis (Supplementary Material 2, Figure 4, available as Supplementary data at *IJE* online).

Since obesity is a major risk factor for T2D, methylation differences between T2D cases could be confounded by BMI. We therefore added BMI in a secondary model, to identify DMPs associated with T2D independent of BMI. In order to address possible inflation of our test statistics by systematic biases, we applied the recently reported EWAS method by van Iterson *et al.*,[23] using the BACON package (version 1.4.0) in 'R' software. This package includes a Bayesian method that controls biases based on estimation of the empirical null distribution. Model fitting was evaluated using Q-Q plots (Figure 1a–d) . The lambdas for the T2D and HbA1c DMP analyses without and with adjustment for BMI reduced from 1.19, 1.31, 1.40 and 1.51 before implementation of the *BACON* package, to 1.04, 1.10, 1.10 and 1.14 after, respectively. Betas of DMPs identified were subsequently included as independent variables in logistic regression models to assess the association of 1% increase in DNA methylation of respective DMPs with T2D. The Nagelkerke's R squared statistic from the logistic regression models with and without covariates was used to calculate trait variance explained by a locus. The results of the logistic regression analysis were presented as odds ratios with corresponding 95% confidence intervals (CI) and *P*-values.

To confirm our results from the binary T2D analyses, we additionally performed association analyses with the continuous measure of HbA1c in µl/mol as the dependent variable on a subset of 675 samples for which HbA1c measurements were available. HbA1c levels reflect a mean blood glucose over a period of 8–12 weeks[24] and are therefore often used to get an impression of long-term glucose control in both T2D cases and healthy individuals. The same covariates as in the T2D analysis were selected.

For all DMP analyses, M-values were calculated as the $\log_2$ ratio of the intensities of methylated probes versus unmethylated probes. Identification of DMPs was based on M-values instead of beta-values, as M-values are more statistically valid.[25] Corresponding beta-values were

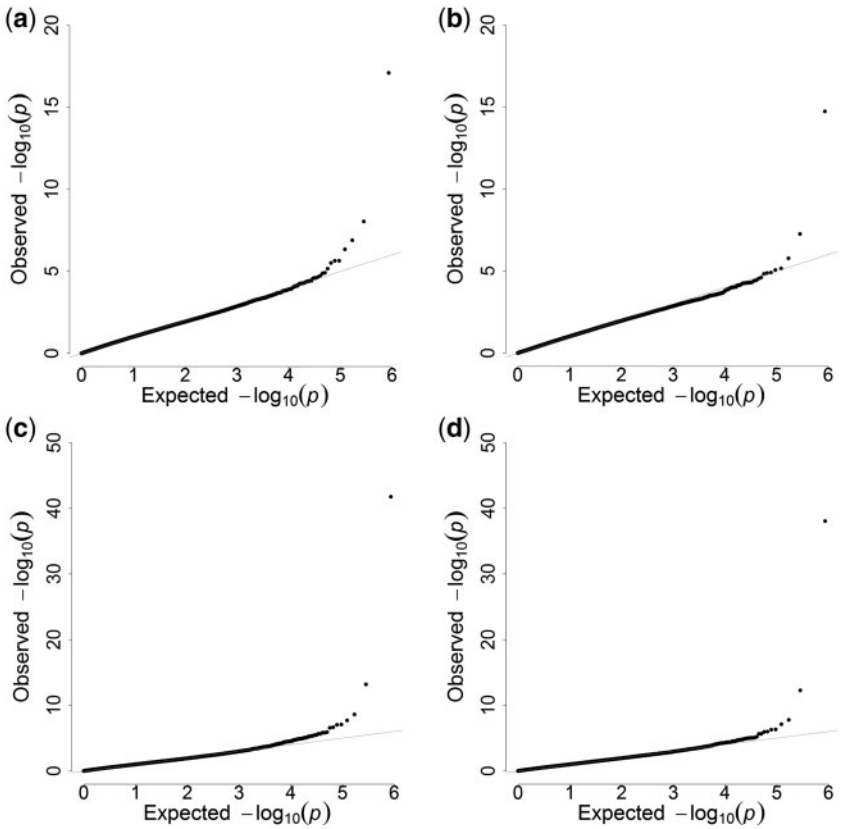

Figure 1. (a) Q-Q plot of EWAS *P*-values for DMP analyses on type 2 diabetes unadjusted for BMI after BACON correction. Lambda = 1.04. (b) Q-Q plot of EWAS *P*-values for DMP analyses on type 2 diabetes adjusted for BMI after BACON correction. Lambda = 1.10. (c) Q-Q plot of EWAS *P*-values for DMP analyses on HbA1c unadjusted for BMI after BACON correction. Lambda = 1.10. (d) Q-Q plot of EWAS *P*-values for DMP analyses on HbA1c adjusted for BMI after BACON correction. Lambda = 1.14.

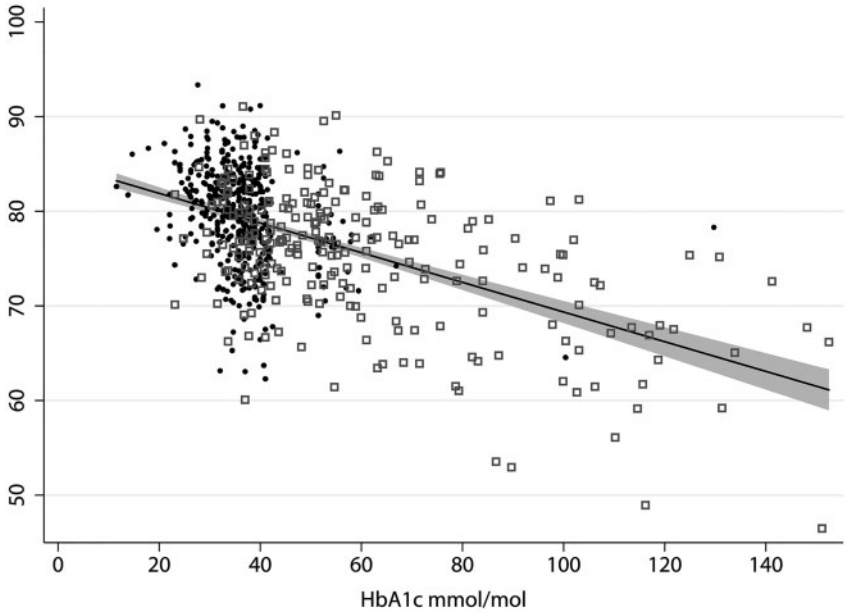

Figure 2. Scatterplot of HbA1c and percentage methylation for differentially methylation position cg19693031 on gene TXNIP by sample. Black dot = T2D control, grey square = T2D case, line = fitted values, grey-shaded area = 95% CI.

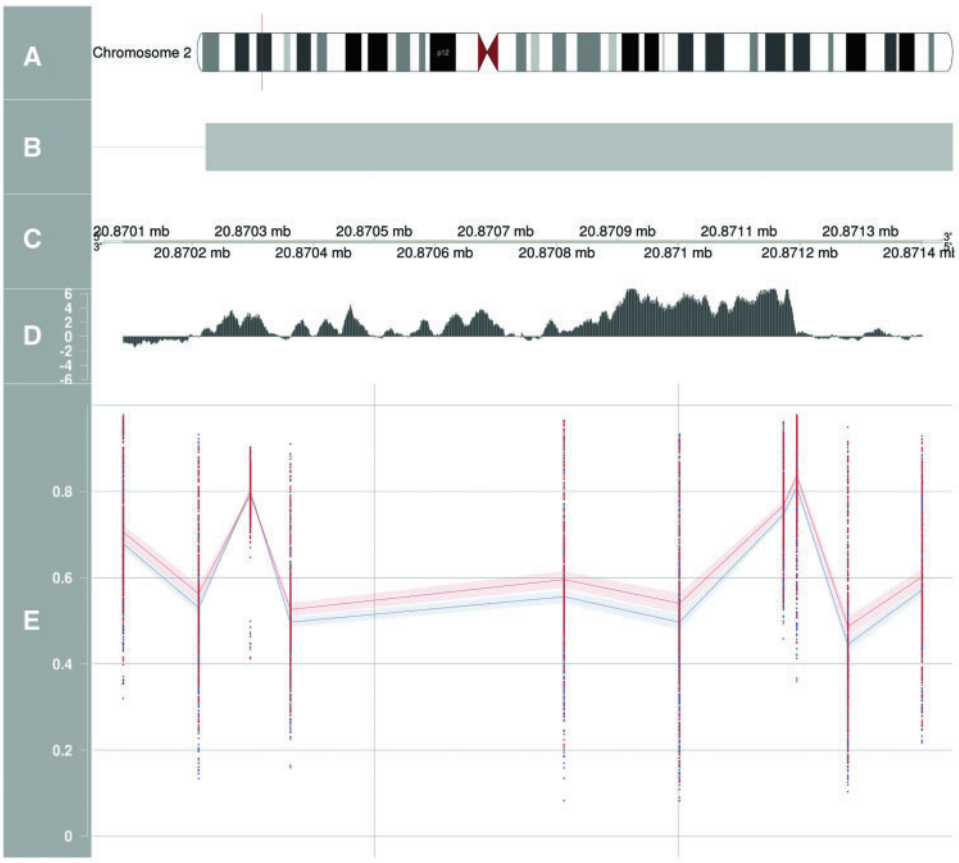

Figure 3. Differentially methylated region located near gene GDF7 associated with type 2 diabetes. A = chromosome, B = reference gene, C = position, D = conservation species track, E = betas 450k DNA methylation + 95% confidence intervals. Blue = diabetes cases, red = diabetes controls.

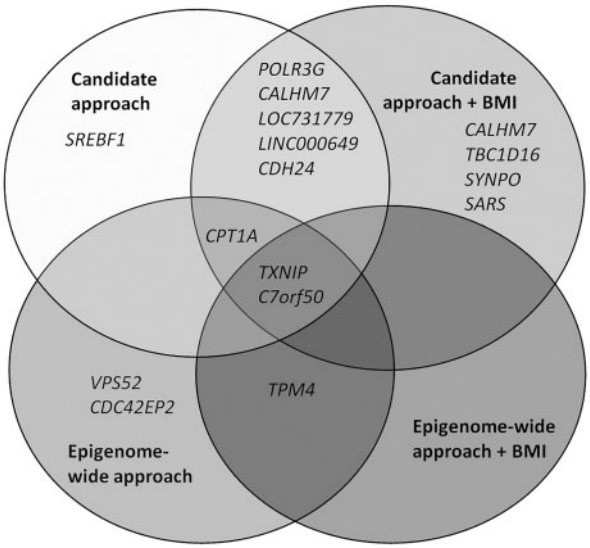

Figure 4. Venn diagram of overlapping genes. Gene names are genes on which identified differentially methylated positions were located in epigenome-wide analyses and in analyses on differentially methylated positions previously reported in other populations (candidate approach).

reported to facilitate biological interpretation.[25] False discovery rate (FDR) adjusted *P*-values corresponding to M-values were used to adjust for multiple testing. FDR was calculated according to the Benjamini–Hochberg procedure.

Differentially methylated regions

To identify differentially methylated regions (DMRs), we fitted models similar to DMP analysis using the *bumphunter* function in the *Minfi* package.[26] We calculated the DMRs with the methylation cut-off of 0.015 (corresponding to 1.5% difference in beta -alues) for T2D and 0.00025 for the continuous measure of HbA1c (corresponding to a 0.025 increase in beta-values per 1-unit increase in HbA1c) using bootstrapping with 500 permutations. In these analyses, beta-values were used instead of M-values as the *bumphunter* function currently does not yet support use of M-values. The DMR methylation cut-off was optimized based on observed effect sizes and significance levels in a volcano plot. We defined a DMR as three or more CpG sites in one cluster. Multiple testing adjustment was performed using family-wise error rate (FWER). We filtered CpG sites with FWER <0.2. Additional DMRs identified with less stringent thresholds are reported in Supplementary Material 3, available as Supplementary data at *IJE* online. Regulation and expression tracks (ENCODE RNA-seq track, ENCODE integrated regulation track and 100 Vertebrate Conservation track) were evaluated for the identified DMRs, using the University of California Santa Cruz (UCSC) human genome browser.[27,28] In order to assess the odds for T2D per 1% increase in DNA methylation of respective DMRs, the average betas for each of the CpGs in a DMR were included in logistic regression models.

Candidate approach of loci reported in non-African populations

Previously reported DMPs and DMRs for T2D were identified through a systematic literature search in PUBMED (Supplementary Material 4, available as Supplementary data at *IJE* online) to use as candidate loci. Multiple testing thresholds using the FDR method were re-calculated for these candidate CpGs ($n = 2876$) due to the lower number of CpGs included. Only one previous EWAS found through the literature search-reported DMRs. Overlap with DMRs identified through *bumphunter* was studied.

Additional post-hoc analyses

We performed additional *posthoc* analyses on identified DMPs of both T2D and HbA1c analyses, including pathway enrichment analyses, evaluation of regulatory regions, evaluation of genetic interaction and gene expression using public databases, and evaluation of methylation levels across recruitment sites. A detailed description of methods for these *post hoc* analyses and their results can be found in Supplementary Material 3, available as Supplementary data at *IJE* online.

Results

Participant characteristics

Out of the 713 samples included in the analysis, 256 were T2D cases and 457 werecontrols (Table 1). Cases and controls were similar in age, gender distribution and sites of data collection. As expected, fasting glucose, HbA1c, BMI and positive family history of T2D were higher in cases than in controls (Table 1). Half of the T2D cases were unaware of their T2D status at the time of examination; the majority of unaware cases resided in Ghana. Among those aware, the average time since diagnosis was 2.4 years. Smoking and alcohol consumption were both low in our study population, with in total only 16 smokers. Both these lifestyle factors were similar between T2D cases and controls. The estimated distribution of cell types inferred from the methylation profile was similar between T2D cases and controls, except for CD4T cell count which was slightly lower in T2D cases (mean 17.3, 95% CI 16.6, 17.9) compared with controls (mean 18.6, 95% CI 18.0, 19.1).

Differentially methylated positions

Six CpG sites (*cg19693031*, *cg04816311*, *cg00574958*, *cg03078690*, *cg07988171*, *cg07477137*) were associated with T2D (FDR <0.05) (Table 2). Implementation of the method to correct for systematic biases resulted in two of these differentially methylated positions (DMPs)

Table 1. Characteristics of type 2 diabetes (T2D) cases and controls[a]

	T2D cases ($n = 256$)	T2D controls ($n = 457$)	Rural Ghana ($n = 104$)	Urban Ghana ($n = 243$)	Europe ($n = 366$)
Age (years)	51.9 (50.6, 53.2)	50.7 (49.8, 51.5)	56.2 (54.4, 57.9)	50.7 (49.5, 51.9)	49.9 (48.8, 50.9)
Sex (% male)	46.5	40.4	30.8	29.6	54.6
Geographical location (%)					
Rural Ghana	16.4	13.6	NA	NA	NA
Urban Ghana	34.4	33.9	NA	NA	NA
Europe	49.2	52.5	NA	NA	NA
BMI (kg/m^2)	27.8 (27.1, 28.4)	26.6 (26.0, 27.1)	22.9 (22.0, 23.7)	26.2 (25.5, 27.0)	28.7 (28.2, 29.2)
Fasting glucose (mmol/L)[b]	9.1 (8.5, 9.6)	5.0 (4.9, 5.0)	6.7 (5.9, 7.5)	7.3 (6.7, 7.8)	5.8 (5.6, 6.0)
HbA1c (%)[b]	7.7 (9.0, 10.1)	5.5 (5.4, 5.5)	5.7 (5.2, 6.1)	6.7 (6.3, 7.0)	6.1 (6.0, 6.2)
HbA1c (mmol/mol)[b]	60.3 (56.8, 63.7)	36.3 (35.4, 37.1)	38.3 (33.6, 42.9)	49.3 (45.9, 52.7)	43.2 (41.6, 44.8)
HOMA-IR[b]	2.74 (3.5)	1.18 (1.04)	1.17 (1.52)	1.69 (2.30)	1.40 (1.69)
HOMA-beta[b]	38.9 (56.8)	74.3 (72.2)	50.1 (59.4)	54.6 (77.7)	67.4 (65.0)
Immediate family with diabetes (%)	37.7 (31.8, 43.9)	17.1 (13.9, 20.9)	17.8 (11.5, 26.6)	29.0 (23.6, 35.1)	23.2 (19.1, 27.9)
Newly detected T2D (%)	47.8 (41.5, 54.1)	NA	22.1 (15.1, 31.2)	19.8 (15.2, 25.3)	12.6 (9.5, 16.4)
Years since diagnosis (years)	2.4 (1.8, 3.0)	NA	2.6 (0.5, 4.6)[c]	2.0 (1.3, 2.8)[c]	2.6 (1.7, 3.5)[c]
Current smokers (%)	2.0 (0.8, 4.8)	2.4 (1.4, 4.4)	0 (NA)	0.4 (0.0, 2.9)	4.2 (2.5, 6.8)
Average alcohol intake (units/week)[d]	1.2 (0.8, 1.6)	1.5 (0.9, 2.1)	1.2 (0.1, 2.4)	0.4 (0.2, 0.6)	2.1 (1.3, 2.9)
Estimated cell counts (%)[e]					
CD8+ T cells	10.7 (10.1, 11.3)	11.0 (10.6, 11.5)	12.0 (11.1, 13.0)	11.9 (11.3, 12.4)	9.9 (9.5, 10.4)
CD4+ T cells	17.3 (16.6, 17.9)	18.6 (18.0, 19.1)	17.5 (16.5, 18.7)	18.0 (17.2, 18.7)	18.4 (17.8, 18.9)
Natural killer cells	10.7 (10.0, 11.3)	10.8 (10.3, 11.3)	13.1 (11.9, 14.3)	11.3 (10.6, 12.0)	9.7 (9.2, 10.2)
B cells	10.6 (10.2, 11.1)	10.6 (10.3, 10.9)	11.5 (10.7, 12.2)	11.0 (10.6, 11.4)	10.1 (9.8, 10.4)
Monocytes	8.0 (7.8, 8.3)	8.2 (7.9, 8.4)	8.2 (7.7, 8.6)	8.2 (7.9, 8.5)	8.0 (7.7, 8.2)
Granulocytes	46.5 (45.4, 47.6)	44.6 (43.8, 45.5)	42.0 (40.0, 44.0)	43.5 (42.4, 44.6)	47.4 (46.6, 48.4)

[a]Numbers are in means or percentages with corresponding 95% confidence intervals; NA, not applicable.

[b]Number of missing results: fasting glucose = 7, HbA1c = 38, HOMA-IR = 6, HOMA-beta = 6. Glycaemic measures were higher in Ghana than in Europe due to exclusion of T2D cases on medication, the proportion of which was higher in Europe.

[c]Years since diagnosed among T2D cases.

[d]Alcohol consumption in units/week, with 500 ml of beer, 250 ml of wine or 80 ml of spirits counted as 1 unit of alcohol.

[e]Cell counts were estimated used the method from Houseman *et al.*[29]

Table 2. Differentially methylated positions associated with type 2 diabetes (T2D) with and without adjustment for body mass index (BMI)

						Base model			Adjusted for BMI		
CpG	CHR	Position	Gene[a]	Feature[b]	Delta β-value[c]	*P*-value[d]	FDR[d]	BACON FDR[e]	*P*-value[d]	FDR[d]	BACON FDR[e]
cg19693031	1	145 441 552	TXNIP	3'UTR	−0.0408	7.35E-18	3.15E-12	3.62E-12	2.45E-17	1.05E-11	8.28E-10
cg04816311	7	1 066 650	C7orf50	Body	0.0198	1.96E-09	4.21E-04	2.03E-03	1.99E-09	4.28E-04	1.22E-02
cg00574958	11	68 607 622	CPT1A	5'UTR	−0.0159	9.26E-08	1.32E-02	1.84E-02	3.09E-06	1.66E-01	8.00E-01
cg07988171	19	16 199 419	TPM4	Body	−0.0176	3.44E-07	3.69E-02	4.98E-02	2.73E-06	3.68E-02	2.47E-01
cg07477137	11	65 083 332	CDC42EP2	5'UTR	0.0044	6.36E-07	4.64E-02	1.66E-01	5.86E-07	6.29E-02	7.62E-01
cg03078690	6	33 235 504	VPS52	Body	0.0089	6.48E-07	4.64E-02	1.66E-01	6.05E-06	1.96E-01	9.56E-01

[a]CpGs are located in the gene if no distance is indicated (genome build Hg19).

[b]Based on manifest feature annotation Illumina.

[c]Beta coefficients were computed from methylation beta-values. Negative beta-values indicate lower DNA methylation (hypomethylation) in cases compared with controls.

[d]*P*-values and FDRs corresponding to M-values.

[e]FDRs corresponding to M-values for the models corrected for inflation by possible systematic biases using the BACON method. Results are sorted on *P*-values for the BACON-adjusted base model.

CHR, Chromosome.

(*cg07988171*, *cg07477137*) no longer passing the threshold for genome-wide significance (FDR = 0.16), but they remained in the top six DMPs. The difference in methylation between cases and controls was most pronounced for *cg19693031* (4.1% less methylated in cases); it is annotated to the *TXNIP* gene (FDR = 3.15E-12) on chromosome 1 in a 3'UTR open sea region. Three out of six DMPs, annotated to genes *TXNIP*, *C7orf50* and *TPM4*, remained associated with T2D after adjustment for BMI (Table 2). The DMPs annotated to *TXNIP* and *TPM4* were hypomethylated in T2D cases, whereas the DMP annotated to *C7orf50* was hypermethylated. The odds for T2D were highest for the DMP annotated to *CDC42EP2* [odds ratio (OR) 1.48, 95% CI 1.27, 1.74] (Table 3). The explained variance in T2D was largest for the DMP annotated to *TXNIP* (14.6%). The combined six DMPs explained 24.6% of variance in T2D.

In the inflation-adjusted models, we found eight DMPs associated with the continuous measure HbA1c unadjusted for BMI and six after adjustment for BMI (FDR <0.05) (Table 4). The DMP with the lowest FDR in association with HbA1c levels (*BACON* FDR = 3.71 x 10^{-33}) was the same as in T2D analysis (*cg19693031—TXNIP*). A 0.1µl/mol (2.2%) increase in HbA1c was associated with 0.15% less methylation of *cg19693031* (Figure 2).

Table 3. Odds ratios for type 2 diabetes differentially methylated positions (DMPs) and differentially methylation regions (DMRs)

					Base model			Adjusted for BMI			Trait variance explained
CpG	CHR	Position	Gene[b]	Feature[c]	OR[d]	95% CI	*P*-value	OR[d]	95% CI	*P*-value	%
cg19693031	1	145 441 552	TXNIP	3'UTR	0.87	0.84, 0.90	<0.001	0.87	0.84, 0.90	<0.001	14.6
cg04816311	7	1 066 650	C7orf50	Body	1.09	1.06, 1.12	<0.001	1.09	1.06, 1.12	<0.001	6.6
cg00574958	11	68 607 622	CPT1A	5'UTR	0.86	0.81, 0.91	<0.001	0.87	0.82, 0.92	<0.001	4.7
cg07988171	19	16 199 419	TPM4	Body	0.89	0.85, 0.93	<0.001	0.89	0.85, 0.93	<0.001	4.6
cg07477137	11	65 083 332	CDC42EP2	5'UTR	1.48	1.27, 1.74	<0.001	1.51	1.28, 1.77	<0.001	4.9
cg03078690	6	33 235 504	VPS52	Body	1.17	1.10, 1.25	<0.001	1.16	1.09, 1.23	<0.001	3.8
DMR[a]	2	20 870 087-20 871 401	GDF7		4.37	1.43, 13.38	0.009	4.48	1.45, 13.84	0.009	1.2

[a]The average methylation level for the nine CpGs in this DMR was taken. There CpGs are: cg09074113, cg05481257, cg07755735, cg14780466, cg10687131, cg12334488, cg10553204, cg07021052, cg04082016.

[b]CpGs are located in the gene if no distance is indicated (genome build Hg19).

[c]Based on manifest feature annotation Illumina.

[d]Odds ratios are per 1% increase in DNA methylation.

CHR, Chromosome.

Table 4. Differentially methylated positions associated with HbA1c with and without adjustment for body mass index

						Base model			Adjusted for BMI		
CpG	CHR	Position	Nearest gene[a]	Feature[b]	Delta β-value[c]	*P*-value[d]	FDR[d]	BACON FDR[e]	*P*-value[d]	FDR[d]	BACON FDR[e]
cg19693031	1	145 441 552	TXNIP	3'UTR	−0.0016	8.26E-45	3.55E-39	7.74E-37	3.18E-44	1.36E-38	3.71E-33
cg04816311	7	10 66 650	C7orf50	Body	0.0008	1.87E-15	4.00E-10	1.37E-08	2.33E-15	5.00E-10	1.23E-07
cg14020176	17	72 764 985	SLC9A3R1	3'UTR	0.0003	2.16E-10	3.09E-05	3.61E-04	4.45E-10	6.20E-05	2.32E-03
cg00574958	11	68 607 622	CPT1A	5'UTR	−0.0004	5.16E-10	5.54E-05	2.16E-03	5.37E-09	4.61E-04	3.98E-02
cg05201300	5	172 443 740	ATP6V0E1	Body	−0.0005	2.79E-09	2.16E-04	6.54E-03	5.77E-10	6.20E-05	8.48E-03
cg08309687	21	35 320 596	LINC00649	IGR	−0.0006	3.01E-09	2.16E-04	6.54E-03	6.82E-09	4.88E-04	4.01E-02
cg03078690	6	33 235 504	VPS52	Body	0.0003	3.46E-08	1.86E-03	1.41E-02	1.77E-07	7.59E-03	9.76E-02
cg12655112	15	42 261 154	EHD4	Body	−0.0005	1.11E-08	6.81E-04	1.49E-02	1.44E-08	8.78E-04	6.00E-02

[a]CpGs are located in the gene if no distance is indicated (genome build Hg19).

[b]Based on manifest feature annotation Illumina; IGR = intergenic region.

[c]Beta coefficients were computed from methylation beta-values. Negative beta-values indicate lower DNA methylation (hypomethylation) for 0.1 µl/mol (2.2%) higher HbA1c.

[d]Numbers are *P*-values and FDR corresponding to M-values.

[e]FDR corresponding to M-values for the models corrected for inflation by possible systematic biases using the BACON method. Results are sorted on *P*-values for the BACON-adjusted base model.

CHR, Chromosome.

We performed *in silico* replication (nominal *P*-value of <0.05) in previously published EWAS for T2D for the DMPs annotated to *VPS52* (*cg03078690*), *TPM4* (*cg07988171*) and *CDC42EP2* (*cg07477137*). None of the 11 EWAS included in the systematic literature search (Supplementary Material 4, available as Supplementary data at *IJE* online) reported *VPS52* (*cg03078690*), *TPM4* (*cg07988171*) or *CDC42EP2* (*cg07477137*).

Differentially methylated regions

We identified one differentially methylated region (DMR) associated with T2D (OR 4.37, 95% CI 1.43, 13.38) annotated to an exon on gene *GDF7* (Figure 3; Table 3). This region was hypomethylated in cases compared with controls and is strongly conserved between species, as is visualized in Figure 3 using data of the Vertebrate Multi Alignment and Conservation track of the UCSC browser.[27] No signals were observed in histone modification track H3K27Ac or ENCODE RNA-seq expression track in muscle tissue-derived cells. One other DMR was associated with HbA1c. This DMR is annotated to chromosome 6 and is closest to gene *ZFP57*.[27,28]

Candidate analysis of DMPs reported in non-African populations

We performed a candidate analysis by fitting the T2D linear regression models to a subset of CpG sites extracted from the literature identified through a systematic search in PUBMED. We could confirm the association with T2D in our dataset for five of the 2876 CpG sites previously reported loci (FDR <0.05) (Table 5). We found a DMP (*cg07988171*) annotated to *TPM4* associated with T2D in the present study, but this was a different DMP from previously reported DMPs in that gene (*cg27377863*,[15] *cg07810884*,[15] *cg06590173*[15]). None of these three CpG sites was associated with T2D in our population. In replication analysis adjusted for BMI, four DMPs were associated with T2D in our study population (Table 5). Only one DMP (*cg11024682*) did not remain associated with T2D after BMI adjustment (Figure 4). The DMPs annotated to *TXNIP* and *C7orf50* were consistently found most strongly associated and with largest effect, that is 4.1% hypomethylation for the DMP annotated to *TXNIP* and 2.0% hypermethylation for the DMP annotated to *C7orf50*.

Replicated DMRs established in non-African populations

Yuan *et al.*[33] reported 20 DMRs identified in relation to T2D in European twins, using MeDIP sequencing and

Table 5. Type 2 diabetes differentially methylated positions in candidate analyses with and without adjustment for BMI

							Base model			Adjusted for BMI		
CpG	Ref[a]	CHR	Position	Gene (distance)[b]	Feature[c]	Delta β-value[d]	*P*-value[e]	FDR[e]	BACON FDR[f]	*P*-value[e]	FDR[e]	BACON FDR[f]
cg19693031	7,30–32	1	145 441 552	TXNIP	3'UTR	−0.0408	7.39E-18	1.97E-14	1.50E-12	2.47E-17	6.57E-14	2.21E-10
cg04816311	30	7	1 066 650	C7orf50	Body	0.0198	1.97E-09	2.62E-06	3.31E-05	2.00E-09	2.66E-06	1.70E-04
cg00574958	30	11	68 607 622	CPT1A	5'UTR	−0.0159	9.28E-08	8.24E-05	7.01E-04	3.10E-06	2.06E-03	3.89E-02
cg19266329	30	1	145 456 128	POLR3GL(-108)	IGR	−0.0150	1.07E-05	7.15E-03	3.33E-02	1.70E-06	1.51E-03	3.27E-02
cg26955383	30	10	105 218 660	CALHM1	TSS200	0.0104	2.00E-05	1.07E-02	4.39E-02	9.64E-06	5.13E-03	5.47E-02

[a]Reference number for article in which candidate CpG was identified.
[b]CpG's are located in the gene if no distance is indicated (genome build Hg19). Distance is expressed in kb.
[c]Based on manifest feature annotation Illumina; IGR = intergenic region.
[d]Beta coefficients were computed from methylation beta-values. Negative beta-values indicate lower DNA methylation (hypomethylation) in cases compared with controls.
[e]Numbers are *P*-values and FDR corresponding to M-values.
[f]FDR corresponding to M-values for the models corrected for inflation by possible systematic biases using the BACON method. Results are sorted on *P*-values for the BACON-adjusted base model.
CHR, Chromosome.

validation on the Illumina 450k array. We did not replicate any of these DMRs. Neither did we identify DMPs associated with T2D within the DMRs reported by Yuan *et al.*

Discussion

Key findings

We conducted an EWAS for T2D among people from Ghana, which to our knowledge, is the first such study in Africa. We identified several CpG sites that were differentially methylated between T2D cases and controls on an epigenome-wide level. Three ubiquitous methylation loci were consistently and strongly associated with T2D in Ghanaians (*cg19693031—TXNIP*, *cg04816311—C7orf50* and *cg00574958—CPT1A*). Three loci (*cg07988171, cg03078690* and *cg07477137)* were associated with T2D on an epigenome-wide level in this study which have not previously been reported in non-African populations. However, two of these no longer passed the epigenome-wide threshold (FDR = 0.16) after correction for possible systematic biases.

Discussion of the key findings

The three DMPs most strongly associated with T2D in our study population had previously been observed in non-African descent populations,[7,30,34] indicating that these loci are associated with T2D across populations (ubiquitous). Out of these three, *cg19693031* (*TXNIP*) has been reported to be associated with T2D in several other populations in blood[7,30–32] and in liver.[7] This DMP was 4.1% less methylated in our T2D cases compared with controls. This is similar to reported effect sizes ranging from -3% to -5%.[7,30–32] Consistent with this finding, DMP *cg19693031* (*TXNIP*) showed a substantial decrease in DNA methylation for higher levels of HbA1c. *TXNIP* is involved in insulin resistance; increased expression negatively affects insulin-mediated glucose uptake and correlates with elevated glucose concentrations.[35] This is consistent with the hypomethylation we found in T2D cases and with the link between hypomethylation and increased HbA1c levels. On the other hand, GWAS in European populations suggest that impaired insulin secretion rather than insulin resistance is considered the main pathogenic cause of T2D.[15] In *post hoc* analyses, we found insulin resistance as assessed by homeostatic modelling to attenuate odds for two of the six T2D DMPs, whereas adjustment for beta-cell dysfunction modelling resulted in hardly any changes (Supplementary Table 2, available as Supplementary data at *IJE* online). Intervention studies are needed to assess whether DNA methylation changes mediate their impact on T2D via insulin resistance or beta-cell dysfunction, in sub-Saharan African populations.

DMP *cg04816311* annotated to *C7orf50* was also associated with T2D in sub-Saharan Africans. This DMP has been previously linked with T2D in Mexican-Americans, Europeans and Arabs, in blood.[30,34] The function of this locus is unknown, although it has been shown to be associated with lipid levels[36] and human longevity.[37]

We also identified three loci—*cg07988171* [*TPM4*], *cg03078690* [*VPS52*] and *cg07477137* [*CDC42EP2*] — not previously reported in non-African populations. Although two (*cg03078690* and *cg07477137)* no longer met the epigenome-wide threshold of FDR <0.05 after correction for statistical inflation by possible systematic biases (FDR = 0.16), they remained in the top six loci and one of them (*cg03078690*) was associated with HbA1c. The DMP *cg07988171* annotated to *TPM4* remained associated with T2D also after correction for possible systematic biases. In our *post hoc* analyses (Supplementary Material 3, available as Supplementary data at *IJE* online), we found that within a 50-bp region around *cg07988171*, there are indications of H3K27Ac histone modifications (i.e. acetylation at the 27th lysine of the histone H3 protein), conservation between species and expression in muscle cells.[27,28] In *post hoc* analyses for genetic interaction (Supplementary Material 3, available as Supplementary data at *IJE* online) around our DMPs, we found that *cg07988171* methylation may interact with gene *RAB8A* which plays a role in glucose homeostasis and is therefore an important target to explore in future research. The DMPs annotated to *VPS52* and *CDC42EP2* showed slight (0.9% and 0.4%) hypermethylation but, whereas *VPS52* is found to be expressed in blood, muscle and adipose tissue, *CDC42EP2* is not.[27,28] Replication studies are needed to confirm the association of these novel loci, in particular *cg07988171*, with T2D in sub-Saharan African populations.

Three DNA methylation loci were associated with T2D at an epigenome-wide level (FDR <0.05) in the model without BMI adjustment, but not after adjustment for BMI, suggesting that they may be mediating their effect on T2D via obesity or that increasing body mass may lead to greater methylation at these sites. It is possible that BMI increases may induce changes in DNA methylation at specific loci, which may in turn influence the risk of T2D. On the other hand, three other DMPs were not confounded by BMI, suggesting that they mediate T2D risk via a mechanism independent of obesity. Enrichment analysis indicated that the identified T2D DMPs might be involved both in muscular function pathways and metabolism pathways (Supplementary Material, available as Supplementary data at *IJE* online).

Our finding that methylation levels of identified DMPs differed by recruitment site should be interpreted with

caution, given the limited sample size upon stratification, but it seems to suggest an environmental influence on the identified DMPs. Three of the six identified DMPs (*cg19693031*, *cg04816311*, *cg00574958*) had differential methylation levels between rural Ghana, urban Ghana and/or Europe (Supplementary Table 8, available as Supplementary data at *IJE* online). In general there was a trend of more hypomethylation in urban Ghana than in rural Ghana and in Europe than in urban Ghana. Possibly, health-related behaviour and environmental changes due to urbanization induce hypomethylation that contributes to a higher T2D prevalence among Ghanaians in Europe and urban Ghana compared with Ghanaians in rural Ghana.[20] The geographical impact on DNA methylation and subsequent T2D requires further exploration.

In the candidate approach, we found five loci associated with T2D out of the 2876 CpGs candidates from previous studies, in either BMI-adjusted or unadjusted analyses. The low yield can be due to population differences, but could also indicate false-positives, particularly when considering the small sample sizes in previous studies (median sample size = 192). We did not find any of the 20 DMRs identified through the systematic literature search to be associated with T2D in our study. It is possible that DMRs associated with T2D differ between Africans and Europeans. However, the lack of overlap could also be attributed to the difference in methods used. Yuan *et al.*[33] validated the DMRs on the Illumina 450k array, but initial detection of DMRs was performed using MeDip-sequencing. Also, differential statistical methods applied could explain the lack of overlap between the studies.

Strengths and limitations

This is the first EWAS for T2D in sub-Saharan Africans. Furthermore, it is one of the largest EWAS on T2D to date. Our sub-Saharan African study population is relatively homogeneous, that is all stem from one region in Ghana. Homogeneity of ethnic groups within Ghana is shown by annotated principal component analysis in Supplementary Figure 4, available as Supplementary data at *IJE* online. Our findings are unlikely to be driven by confounders such as smoking and alcohol intake, as both are low in this study population. Furthermore, smoking DMPs reported in a study among African Americans were not in the vicinity of our T2D DMPs.[38]

A potential limitation is the use of whole-blood samples. As DNA methylation is tissue specific, we ideally would have analysed target tissue in T2D including adipose tissue, skeletal muscle and pancreatic beta cells. However, sampling of these tissues in epidemiological studies is basically impossible, and the actual use of these tissues is very limited. Nevertheless, peripheral blood is thought to give a good reflection and is fairly concordant with other tissue types.[39] More importantly, epigenetic biomarkers derived from a routine blood sample are more feasible for clinical use, for example for the development of biomarkers. A second possible limitation of the present study is the possibility that some of the hits reported are due to genetic [e.g. single nucleotide polymorphisms (SNPs)] rather than epigenetic variation. Unfortunately, we do not yet have GWAS SNP data on the RODAM cohort that would allow us directly test this hypothesis. Existing GWAS data based on non-African populations would not be appropriate for this purpose, due to differences in genetic variation, genetic linkage disequilibrium structure and other factors. Furthermore, we cannot disentangle whether the loci identified were a cause or a consequence of diabetes and its major risk factors, due to the cross-sectional nature of our study. Longitudinal studies assessing new-onset diabetes as an outcome with repeated measurements of phenotype and DNA methylation would be needed to assess whether the epigenetic changes identified occurred before T2D. Intervention studies could be used to study whether modifying DNA methylation affects T2D risk. These future studies should ideally have larger sample sizes and/or combine data in meta-analyses in order to overcome the potential limitation of the present study to detect additional epigenetic loci with small effect sizes (i.e. $<3\%$ methylation difference) due to its sample size.

Conclusions

In summary, this first EWAS on T2D in sub-Saharan Africans has identified methylation loci associated with T2D. We identified both ubiquitous loci and loci not previously reported in non-African populations. Further studies are needed to confirm and extend these findings, including evaluation of the functional relevance of the novel loci and replication in other sub-Saharan African populations. Our findings add to the slowly growing literature on the genetic architecture of T2D in African populations. The datasets created and/or analysed during the current study are available from the corresponding author on reasonable request.

Supplementary Data

Supplementary data are available at *IJE* online.

Funding

This work was supported by the European Commission under the Framework Programme (Grant Number: 278901). C.G. is supported by NutriAct – Competence Cluster Nutrition Research

Berlin-Potsdam funded by the German Federal Ministry of Education and Research (FKZ: 01EA1408A-G). C.N.R. and A.A. are supported 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.

Acknowledgements

The authors are very grateful to Dr Erik Beune (Department of Public Health, Academic Medical Center, The Netherlands) who, with great dedication and eye for detail, coordinated the RODAM study. Special thanks go to Prof. Karien Stronks (Department of Public Health, Academic Medical Center, The Netherlands) for her role in the conceptualization of the RODAM study, and to Prof. Ama de-Graft Aikins (Regional Institute for Populations Studies, University of Ghana, Ghana) for her valuable contribution to the conceptualization of the RODAM study and data collection in Ghana. We thank Nadine Huckauf and Candy Kalischke (Department of Endocrinology and Metabolism, Charité University Medicine, Berlin, Germany) for their excellent support in biochemical characterization of the participants and data handling. We gratefully acknowledge Jan van Straalen from the Academic Medical Center for his valuable support with standardization of the laboratory procedures, and the AMC Biobank for support in biobank management and storage of collected samples. Last but not least, the authors are very 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 Contributions

C.A., F.P.M., E.O.D., L.S., M.B.S., S.B., J.S., I.D., J.A., A.A., P.H., M.M.A.M., K.M., C.N.R. and M.H.Z. conceived and designed the study. C.A., E.O.D., F.P.M., J.A., I.D., C.G., J.S. and K.M. carried out the recruitment and data collection. T.B. and P.H. generated the epigenetics data. A.V. and K.M. performed the statistical analysis. All authors contributed to interpretation of the data. K.M. wrote the manuscript, with the cooperation of all co-authors. All authors read, contributed to and approved the final manuscript. C.A. is responsible for the integrity of the work as a whole.

Conflict of interest: The authors declare that there is no duality of interest associated with this manuscript.

References

1. International Diabetes Federation (IDF). *IDF Diabetes Atlas*, 7th edn. Brussels: International Diabetes Federation, 2015.
2. Patel CJ, Bhattacharya J, Butte AJ. An environment-wide association study (EWAS) on type 2 diabetes mellitus. *PLoS One* 2010;**5**:e10746.
3. Zeggini E, Scott LJ, Saxena R *et al*. Meta-analysis of genome-wide association data and large-scale replication identifies additional susceptibility loci for type 2 diabetes. *Nat Genet* 2008;**40**:638–45.
4. Drong AW, Lindgren CM, McCarthy MI. The genetic and epigenetic basis of type 2 diabetes and obesity. *Clin Pharmacol Ther* 2012;**92**:707–15.
5. Maher B. The case of the missing heritability. *Nature* 2008;**456**: 18–21.
6. Li E. Chromatin modification and epigenetic reprogramming in mammalian development. *Nat Rev Genet* 2002;**3**:662–73.
7. Chambers JC, Loh M, Lehne B *et al*. Epigenome-wide association of DNA methylation markers in peripheral blood from Indian Asians and Europeans with incident type 2 diabetes: a nested case-control study. *Lancet Diabetes Endocrinol* 2015;**3**:526–34.
8. Toperoff G, Aran D, Kark JD *et al*. Genome-wide survey reveals predisposing diabetes type 2-related DNA methylation variations in human peripheral blood. *Hum Mol Genet* 2012;**21**:371–83.
9. Rönn T, Volkov P, Davegårdh C *et al*. A six months exercise intervention influences the genome-wide DNA methylation pattern in human adipose tissue. *PLoS Genet* 2013;**9**:e1003572.
10. Anderson OS, Sant KE, Dolinoy DC. Nutrition and epigenetics: an interplay of dietary methyl donors, one-carbon metabolism and DNA methylation. *J Nutr Biochem* 2012;**23**:853–59.
11. Nolan CJ, Damm P, Prentki M. Type 2 diabetes across generations: from pathophysiology to prevention and management. *Lancet* 2011;**378**:169–81.
12. Ling C, Groop L. Epigenetics: a molecular link between environmental factors and type 2 diabetes. *Diabetes* 2009;**58**:2718–25.
13. Meeks KAC, Freitas-Da-Silva D, Adeyemo A *et al*. Disparities in type 2 diabetes prevalence among ethnic minority groups resident in Europe: a systematic review and meta-analysis. *Intern Emerg Med* 2016;**11**:327–40.
14. Steyn NP, Mchiza ZJ. Obesity and the nutrition transition in Sub-Saharan Africa. *Ann N Y Acad Sci* 2014;**1311**:88–101.
15. Dayeh T, Volkov P, Salö S *et al*. Genome-wide DNA methylation analysis of human pancreatic islets from type 2 diabetic and non-diabetic donors identifies candidate genes that influence insulin secretion. *PLoS Genet* 2014;**10**:e1004160.
16. Ronn T, Volkov P, Gillberg L *et al*. Impact of age, BMI and HbA1c levels on the genome-wide DNA methylation and mRNA expression patterns in human adipose tissue and identification of epigenetic biomarkers in blood. *Hum Mol Genet* 2015; **24**:3792–813.
17. Nilsson E, Matte A, Perfilyev A *et al*. Epigenetic alterations in human liver from subjects with type 2 diabetes in parallel with reduced folate levels. *J Clin Endocrinol Metab* 2015;**100**: E1491–501.
18. Ling C, Poulsen P, Simonsson S *et al*. Genetic and epigenetic factors are associated with expression of respiratory chain component NDUFB6 in human skeletal muscle. *J Clin Invest* 2007;**117**: 3427–35.
19. 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* 2015;**4**:e004877.
20. Agyemang C, Meeks K, Beune E *et al*. Obesity and type 2 diabetes in sub-Saharan Africans—is the burden in today's Africa similar to African migrants in Europe? The RODAM study. *BMC Med* 2016;**14**:166.

21. Tsai P-C, Bell JT. Power and sample size estimation for epigenome-wide association scans to detect differential DNA methylation. *Int J Epidemiol* 2015;**44**:1429–41.
22. Alders M, Bliek J, Vd Lip K, Vd Bogaard R, Mannens M. Determination of KCNQ1OT1 and H19 methylation levels in BWS and SRS patients using methylation-sensitive high-resolution melting analysis. *Eur J Hum Genet* 2009;**17**: 467–73.
23. van Iterson M, van Zwet EW, Heijmans BT. Controlling bias and inflation in epigenome-and transcriptome-wide association studies using the empirical null distribution. *Genome Biol* 2017; **18**:19.
24. Bunn HF, Gabbay KH, Gallop PM. The glycosylation of hemoglobin: relevance to diabetes mellitus. *Science* 1978;**200**: 21–27.
25. Du P, Zhang X, Huang C-C *et al*. Comparison of beta-value and M-value methods for quantifying methylation levels by microarray analysis. *BMC Bioinformatics* 2010;**11**: 587.
26. Jaffe AE, Murakami P, Lee H *et al*. Bump hunting to identify differentially methylated regions in epigenetic epidemiology studies. *Int J Epidemiol* 2012;**41**:200–09.
27. Kent WJ, Sugnet CW, Furey TS *et al*. The human genome browser at UCSC (Human Assembly GRCh37/hg19). *Genome Res* 2002;**12**:996–1006.
28. Rosenbloom KR, Sloan CA, Malladi VS *et al*. ENCODE data in the UCSC Genome Browser: year 5 update. *Nucleic Acids Res* 2012;**41**:D56–63.
29. Houseman EA, Accomando WP, Koestler DC *et al*. DNA methylation arrays as surrogate measures of cell mixture distribution. *BMC Bioinformatics* 2012;**13**:1.
30. Kulkarni H, Kos MZ, Neary J *et al*. Novel epigenetic determinants of type 2 diabetes in Mexican-American families. *Hum Mol Genet* 2015;**24**:5330.
31. Florath I, Butterbach K, Heiss J *et al*. Type 2 diabetes and leucocyte DNA methylation: an epigenome-wide association study in over 1,500 older adults. *Diabetologia* 2016;**59**:130–38.
32. Soriano-Tárraga C, Jiménez-Conde J, Giralt-Steinhauer E *et al*. Epigenome-wide association study identifies TXNIP gene associated with type 2 diabetes mellitus and sustained hyperglycemia. *Hum Mol Genet* 2016;**25**:609–19.
33. Yuan W, Xia Y, Bell CG *et al*. An integrated epigenomic analysis for type 2 diabetes susceptibility loci in monozygotic twins. *Nat Commun* 2014;**5**:1–7.
34. Al Muftah WA, Al-Shafai M, Zaghlool SB *et al*. Epigenetic associations of type 2 diabetes and BMI in an Arab population. *Clin Epigenet* 2016;**8**:1–10.
35. Parikh H, Carlsson E, Chutkow WA *et al*. TXNIP regulates peripheral glucose metabolism in humans. *PLoS Med* 2007;**4**:e158.
36. Ligthart S, Vaez A, Hsu Y-H *et al*. Bivariate genome-wide association study identifies novel pleiotropic loci for lipids and inflammation. *BMC Genomics* 2016;**17**:443.
37. Yashin AI, Wu D, Arbeev KG, Ukraintseva SV. Joint influence of small-effect genetic variants on human longevity. *Aging* 2010;**2**: 612–20.
38. Sun YV, Smith AK, Conneely KN *et al*. Epigenomic association analysis identifies smoking-related DNA methylation sites in African Americans. *Hum Genet* 2013;**132**:1027–37.
39. Kato N, Loh M, Takeuchi F *et al*. Trans-ancestry genome-wide association study identifies 12 genetic loci influencing blood pressure and implicates a role for DNA methylation. *Nat Genet* 2015;**47**:1282–93.