



Admixture mapping and fine-mapping of type 2 diabetes susceptibility loci in African American women

José M. Uribe-Salazar¹ · Julie R. Palmer² · Stephen A. Haddad² · Lynn Rosenberg² · Edward A. Ruiz-Narváez³

Received: 25 May 2018 / Revised: 18 July 2018 / Accepted: 13 August 2018 / Published online: 22 August 2018
© The Author(s) under exclusive licence to The Japan Society of Human Genetics 2018

Abstract

African American women are disproportionately affected by type 2 diabetes. Genetic factors may explain part of the excess risk. More than 100 genetic variants have been associated with risk of type 2 diabetes, but most studies have been conducted in white populations. Two genome-wide association studies (GWAS) in African Americans have identified three novel genetic variants only. We conducted admixture mapping using 2918 ancestral informative markers in 2632 cases of type 2 diabetes, and 2596 controls nested in the ongoing Black Women's Health Study cohort, with the goal of identifying genomic loci with local African ancestry associated with type 2 diabetes. In addition, we performed replication analysis of 71 previously identified index SNPs, and fine-mapped those genetic loci to identify better or new genetic variants associated with type 2 diabetes in African Americans. We found that individual African ancestry was associated with higher risk of type 2 diabetes. In addition, we identified two genomic regions, 3q26 and 12q23, with excess of African ancestry associated with higher risk of type 2 diabetes. Lastly, we replicated 8 out of 71 index SNPs from previous GWAS, including, for the first time in African Americans, the X-linked rs5945326 SNP near the *DUSP9* gene. In addition, our fine-mapping efforts suggest independent signals at five loci. Our detailed analysis identified two genomic regions associated with risk of type 2 diabetes, and showed that many genetic risk variants are shared across ancestries.

Introduction

Type 2 diabetes is a grave public health problem that affects about 30 million individuals in the United States, or more than 9% of the US adult population [1]. African American women are disproportionately affected by type 2 diabetes, with a prevalence that is more than twice as high as among non-Hispanic white women [2]. Differences in traditional major risk factors, such as obesity, do not fully explain the

racial disparity in the risk of type 2 diabetes [3, 4]. Genetic factors may contribute to the excess risk observed in African Americans.

The list of loci involved in the pathogenesis of type 2 diabetes has grown substantially as a result of recent genome-wide association studies (GWAS). The majority of GWAS of type 2 diabetes susceptibility have been performed in populations of European or East Asian ancestry [5–12]. These studies have discovered more than 100 single-nucleotide polymorphisms (SNPs) associated with type 2 diabetes. GWAS in African ancestry populations have discovered only three novel type 2 diabetes risk variants [13, 14].

To add to the knowledge of the genetic basis of type 2 diabetes in African Americans, we used samples and data from the Black Women's Health Study (BWHS) to (1) examine whether individual African ancestry is associated with higher risk of type 2 diabetes and whether such association is mediated by body weight; (2) conduct genome-wide admixture mapping to identify genomic loci with local African ancestry associated with risk of type 2 diabetes; (3) conduct replication analysis of 71 index SNPs previously found associated with type 2 diabetes in European ancestry

Electronic supplementary material The online version of this article (<https://doi.org/10.1038/s10038-018-0503-2>) contains supplementary material, which is available to authorized users.

✉ Edward A. Ruiz-Narváez
eruiznar@umich.edu

¹ Escuela de Biología, Universidad de Costa Rica, San Pedro, Costa Rica

² Slone Epidemiology Center at Boston University, Boston, MA, USA

³ Department of Nutritional Sciences, University of Michigan School of Public Health, Ann Arbor, MI, USA

populations; and (4) fine-map those genetic loci to identify better and/or new type 2 diabetes genetic variants in African Americans.

Materials and methods

Study population

Subjects in the present study are participants in the BWHS, an ongoing prospective cohort study initiated in 1995 when 59,000 African American women 21–69 years of age from across the US completed a postal health questionnaire [15]. Information on demographic factors, health history, and anthropometric measurements, among others, was collected at baseline. Every 2 years, participants complete follow-up questionnaires to update outcome and covariates data. DNA samples were obtained from BWHS participants by the mouthwash-swish method [16]. Saliva samples were provided by ~50% of BWHS participants. The BWHS was approved by the institutional review board of Boston University (Boston, MA). All study subjects provided written informed consent for use of their saliva samples.

Cases of diabetes for the present study were randomly selected among those with a self-report of newly diagnosed type 2 diabetes in any of the biennial follow-up questionnaires up to 2013 and who had provided a saliva sample. Accuracy of self-reported diabetes was evaluated with 293 women who gave permission to contact their physicians. For 229 women whose physician provided information on the diagnosis; diagnosis of type 2 diabetes was confirmed in 96% of the cases. We randomly selected one control per case matched on year of birth (± 1 year), having completed the same questionnaire as the last questionnaire completed by the case before her diagnosis (index date) of type 2 diabetes, and geographic region of residence (Northeast, South, Midwest, and West) at the matched questionnaire. Controls were selected among those without a self-report of type 2 diabetes who had provided a saliva sample. The prevalence of undiagnosed diabetes in the BWHS was previously assessed using data from collected blood samples [17]. Among 1873 BWHS participants who provided a blood sample in the first year of blood collection and had never reported diabetes, 120 (6.4%) had HbA1c levels of 6.5% (47.5 mmol/mol) or higher, meeting criteria for diabetes [18]. Thus, the prevalence of undiagnosed diabetes in the BWHS is similar to what has been reported for African Americans in the general US population [1].

Assessment of covariates

Age, weight, geographic region of residence, and neighborhood socioeconomic status (SES) information was taken

from the last questionnaire completed before diagnosis for cases, and from the matched questionnaire for controls. BMI (kg/m^2) was calculated from self-reported adult height at baseline and current weight. In a validation study of anthropometric measures conducted in 115 BWHS participants, Spearman correlations for self-reported vs. technician-measured weight and height were 0.97 and 0.93, respectively [19]. Neighborhood SES was measured by linking participants' current address to 2000 US Census block groups using geocoding (Mapping Analytics, Rochester, NY, USA) [20]. We used factor analysis to calculate a score for neighborhood SES that included six group census variables: median household income; median housing value; percentage of households receiving interest, dividend, or net rental income; percentage of adults aged 25 years or older who have completed college; percentage of employed persons aged 16 years or older who are in occupations classified as managerial, executive, or professional; and percentage of families with children that are not headed by a single female. Years of education were self-reported in 1995 and 2003. Data on years of education were taken from the closest 1995 or 2003 questionnaire preceding the year of diagnosis for cases, or index year for controls.

SNP selection

Ancestral informative markers (AIMs)

A list of 2918 AIMs was extracted from the Affymetrix Axiom Genomic Database. These AIMs have large frequency differences between African and European ancestry populations from the 1000 Genomes Project.

Index SNPs and fine-mapping of type 2 diabetes loci

We included 71 index SNPs (Supplementary Table 1), representing 70 genomic regions that have been reported associated with type 2 diabetes in one or more GWAS. Tagging SNPs for fine-mapping were selected for each of these genomic regions (± 100 kb around each index SNP) in order to capture (at $r^2 \geq 0.9$) all SNPs with minor allele frequency (MAF) $\geq 5\%$, based on the African populations in 1000 genomes.

Genotyping and QC

Samples were genotyped on an Affymetrix Axiom 45K custom array (Affymetrix, Santa Clara, CA, USA), designed to include genes and SNPs related to type 2 diabetes, as well as AIMs. Genotyping was carried out at the Affymetrix laboratory (Santa Clara, CA, USA). The Axiom array data underwent extensive QC procedures carried out by

Affymetrix and the study investigators. We removed about 13% of samples due to high missing call rates (>5%) and poor reproducibility. In addition, we excluded about 17% of SNPs because of poor cluster properties, high missing call rates (>10%), deviation from Hardy–Weinberg proportions ($p < 1 \times 10^{-5}$ in controls), or discordance between duplicate samples. After these exclusions, the final dataset included 5228 subjects (2632 cases and 2596 controls) and 38,008 SNPs, including 22,038 SNPs selected for the current analysis.

Imputation

Once genotypes were obtained, we performed SNP imputation using the imputation server from the University of Michigan (<https://www.imputationserver.sph.umich.edu>). We used 1000 Genomes Project phase 3 African (AFR) reference panel that includes 1322 haplotypes (661 subjects). Imputation resulted in 85,375 SNPs with MAF $\geq 1\%$ and info score ≥ 0.5 for analysis in the 70 type 2 diabetes loci.

Data analysis

Ancestry analysis

We used ADMIXMAP software version 3.8.3103 [21] and genotypes of 2918 AIMs to estimate individual percentage of African ancestry. We categorized individual African ancestry in quintiles for statistical analysis, using the first (lowest) quintile as reference group. We used logistic regression to estimate odds ratios (ORs) and 95% confidence intervals (CIs) of the association of individual African ancestry (quintiles) and type 2 diabetes. Our basic model was adjusted for age, geographic region of residence (Northeast, South, Midwest, and West), and genotyping batch. We further adjusted for score of neighborhood SES (quintiles), and years of education (≤ 12 , 13–15, ≥ 16 years). We also modeled individual African ancestry as a continuous variable to estimate ORs and 95% CI of type 2 diabetes per 10% increase of individual African ancestry.

We used several approaches to examine whether individual African ancestry affects risk of type 2 diabetes in part through an effect on BMI. First, we ran models without and with adjustment for BMI (continuous). If β_1 is the estimate of the association between African ancestry and type 2 diabetes without adjustment for BMI, and β_2 is estimate of the association between African ancestry and type 2 diabetes with adjustment for BMI then, the proportion of the association between African ancestry and type 2 diabetes that is explained by BMI is given by $\left(\frac{\beta_1 - \beta_2}{\beta_1}\right)$. Ancestry associations are in logarithmic scale. Second, we conducted BMI-stratified analyses (non-obese women, BMI $< 30 \text{ kg/m}^2$, and

obese women, BMI $\geq 30 \text{ kg/m}^2$). Finally, we evaluated the association of individual African ancestry (continuous) with BMI among controls using linear regression adjusting for age, geographic region of residence (Northeast, South, Midwest, and West), genotyping batch, quintiles of the neighborhood SES score, and years of education (≤ 12 , 13–15, ≥ 16 years).

Admixture mapping

We used ADMIXMAP software version 3.8.3103 [21] to estimate African locus-specific ancestry and identify genomic regions with local ancestry associated with risk of type 2 diabetes. Admixture mapping can be done using case-only and/or case-control analyses. The case-only approach compares locus-specific ancestry at each chromosomal position with average genome-wide ancestry. The case-control method tests for differences of locus-specific ancestry between cases and controls adjusting for individual admixture and other covariates. Case-only analysis has greater statistical power than case-control analysis, assuming that deviation in ancestry between cases and controls is not due to population stratification. We therefore first performed a case-only admixture scan and then a case-control admixture analysis to confirm local ancestry associations from the case-only results. Case-control analyses were adjusted for global individual African ancestry, age, geographic region of residence, and genotyping batch. Statistical significance was assessed using Z statistics, with a threshold of $|Z| > 4.0$ considered genome-wide statistically significant [22]. A positive Z-score indicates that higher African ancestry at that particular locus is associated with higher risk of type 2 diabetes; a negative Z-score indicates that lower African ancestry is associated with higher risk of type 2 diabetes.

Fine-mapping of type 2 diabetes loci

We used logistic regressions to estimate OR and 95% CIs of the association between genetic variants and type 2 diabetes. ORs were adjusted for age at the last completed questionnaire for cases and at the matched questionnaire for controls, current geographic region of residence, individual African ancestry percentage, and genotyping batch. Within each type 2 diabetes locus we corrected for multiple testing using the number of SNPs that were genotyped [23].

Results

Table 1 shows the characteristics of BWHS participants in the present study. Cases had higher BMI and individual African ancestry compared to controls. Cases also had less

Table 1 General characteristics of type 2 diabetes cases and controls in the Black Women's Health Study^a

Characteristic ^b	Cases (n = 2632)	Controls (n = 2596)
Age, years (mean)	50.1	47.7
BMI, kg/m ² (mean)	34.7	29.4
African ancestry, % (mean)	80.0	78.0
Education ≥ 16 years, %	45.6	56.7
Neighborhood SES, %		
First quintile (poorest)	23.0	17.5
Fifth quintile (richest)	15.9	22.8
Region of residence, %		
Northeast	25.5	25.5
South	33.7	33.2
Midwest	24.3	24.3
West	16.5	17.0
Health insurance coverage at baseline, %	92.7	93.2

^aCharacteristic values were taken from the last completed questionnaire before diagnosis for cases, and from the matched questionnaire for controls

^bAll characteristics, except region of residence and health insurance coverage, were significantly different between cases and controls ($P < 0.05$)

years of education, and lived more frequently in neighborhoods within the lowest quintile of SES relative to controls. There were no differences regarding geographic region of residence and health insurance coverage.

Table 2 shows results of the association analyses between individual African ancestry and risk of type 2 diabetes. We found, in the basic model adjusting for age and geographic region of residence, that women in the highest quintile of African ancestry had 75% higher risk of type 2 diabetes compared to women in the lowest quintile (OR = 1.75; 95% CI = 1.46–2.09). The association was attenuated, although it remained highly significant, after successive adjustment for SES (OR = 1.57; 95% CI = 1.31–1.89), and BMI (OR = 1.41; 95% CI = 1.16–1.71). We estimated that each 10% increase of African ancestry was associated with 12% higher risk of type 2 diabetes in the model adjusted for basic covariates plus SES and BMI (OR = 1.12; 95% CI = 1.06–1.18). In BMI-stratified analysis, a significant association between African ancestry and type 2 diabetes was observed only among non-obese women (OR per 10% increase in African ancestry = 1.20; 95% CI = 1.10–1.30). To further explore whether African ancestry may affect risk of type 2 diabetes through an increase in BMI, we assessed the relation of African ancestry with BMI in controls in a multivariate model adjusting for age, geographic region of residence, genotyping batch, and SES. We found that each 10% increase of African ancestry was associated with an

increase of BMI = 0.59 kg/m²; 95% CI = 0.37–0.81 kg/m² ($p < 1 \times 10^{-4}$).

Table 3 shows results of our genome-wide admixture mapping of type 2 diabetes. A total of 2918 autosomal AIMs were included (Supplementary Table 2). We found two genome-wide significant loci, 3q26 and 12q23, with excess of African ancestry associated with higher risk of type 2 diabetes. Each African allele at 3q26 was associated with 23% higher risk of type 2 diabetes (OR = 1.23; 95% CI = 1.09–1.39). ORs for the association were 1.29 (95% CI = 1.10–1.53) among obese women and 1.19 (95% CI = 0.98–1.44) among non-obese women. At 12q23, each African allele was associated with 13% higher risk of type 2 diabetes (OR = 1.13; 95% CI = 1.00–1.29). In BMI stratified analysis, the association was observed among non-obese women only (OR = 1.33; 95% CI = 1.09–1.62). Regarding their relation with BMI, we found that the high-risk African allele at 3q26 was associated with lower BMI in controls (Beta = -0.55 kg/m²; 95% CI = -1.05, -0.038 kg/m²), and the high-risk African allele at 12q23 was associated with higher BMI in controls (Beta = 0.66 kg/m²; 95% CI = 0.13, 1.18 kg/m²).

Table 4 shows the 10 index SNPs, out of 71 SNPs, that were associated with type 2 diabetes at a nominal $P < 0.05$ in the present study. Out of these 10 nominally significant variants, all but two, rs1535500 at *KCNK16* and rs6960043 near *DGKB*, had associations with type 2 diabetes in the same direction as reported in previous GWAS. In six out of the eight SNPs with the same GWAS-reported association direction, the high-risk allele was more frequent in BWHS and African-ancestry populations from 1000 Genomes than in European-ancestry populations. Results for all 71 index SNPs are shown in Supplementary Table 1.

Table 5 shows the five genomic loci with region-wide significant fine-mapping results. At 1q32, we identified a new signal rs12091447 (OR = 1.55, $p = 1.1 \times 10^{-4}$) that was not correlated with the index SNP ($r^2 < 0.01$ in European, EUR, and African, AFR, populations from 1000 Genomes Project). We did not find a significant association of the index SNP (rs2075423) with type 2 diabetes in the present study.

At 5q11, the index SNP rs459193 was nominally associated with type 2 diabetes (OR = 1.09, $p = 0.037$). We found a new signal rs73127858 (OR = 1.29, $p = 7.1 \times 10^{-5}$) uncorrelated to the index SNP both in AFR and EUR populations ($r^2 = 0$). The effect T-allele has frequency of 11% in the BWHS, 13% in AFR populations, and it is almost absent (0.1% frequency) in EUR populations.

At 9p24, we did not find a significant association of the index SNP rs7041847 and type 2 diabetes (OR = 1.10, $p = 0.11$). We found a new signal rs114560781 associated with type 2 diabetes in BWHS (OR = 1.80, $p = 6.6 \times 10^{-5}$) that was uncorrelated with the index SNP in AFR populations

Table 2 Odds ratio (OR) according to quintiles of percentage of African ancestry in the Black Women Health's Study

	Quintiles of percentage of African ancestry (range)					<i>P</i> -trend	Per 10% increase of African ancestry
	Q1 (≤69.3)	Q2 (69.3–77.5)	Q3 (77.6–82.5)	Q4 (82.6–86.7)	Q5 (>86.7)		
<i>All women</i>							
Cases/controls	377/517	492/525	549/515	564/508	650/531		
Basic ^a	1.00 (reference)	1.35 (1.12–1.62)	1.56 (1.30–1.88)	1.63 (1.35–1.95)	1.75 (1.46–2.09)	<0.0001	1.20 (1.14–1.26)
+SES ^b	1.00 (reference)	1.28 (1.06–1.54)	1.44 (1.20–1.74)	1.49 (1.23–1.79)	1.57 (1.31–1.89)	<0.0001	1.16 (1.10–1.22)
+BMI ^c	1.00 (reference)	1.22 (1.00–1.49)	1.38 (1.14–1.69)	1.33 (1.09–1.62)	1.41 (1.16–1.71)	<0.0001	1.12 (1.06–1.18)
<i>Non-obese^d</i>							
Cases/controls	121/367	154/328	175/318	173/303	191/306		
Multivariate ^e	1.00 (reference)	1.44 (1.06–1.95)	1.75 (1.30–2.36)	1.55 (1.15–2.10)	1.67 (1.24–2.25)	<0.0001	1.20 (1.10–1.30)
<i>Obese^d</i>							
Cases/controls	256/150	338/197	374/197	391/205	459/225		
Multivariate ^e	1.00 (reference)	1.01 (0.77–1.33)	1.11 (0.84–1.46)	1.11 (0.84–1.45)	1.15 (0.88–1.50)	0.23	1.05 (0.97–1.13)

^aAdjusted for age and geographic region of residence^bQuintiles of neighborhood socioeconomic status, and years of education^cBody mass index as a continuous variable^dNon-obese, BMI < 30 kg/m²; obese, BMI ≥ 30 kg/m²^eAdjusted for age, geographic region of residence, neighborhood socioeconomic status, years of education, and BMI (continuous)**Table 3** Genome-wide significant regions of local African ancestry associated with type 2 diabetes in the Black Women's Health Study

	Locus	
	3q26.33	12q23.1
<i>Z</i> _{case-only} , <i>P</i> -value	4.01, 6.0 × 10 ⁻⁵	4.42, 9.8 × 10 ⁻⁶
<i>Z</i> _{case-control} , <i>P</i> -value	2.79, 5.3 × 10 ⁻³	2.70, 7.0 × 10 ⁻³
<i>Per African allele OR (95% CI)^a, cases/controls</i>		
All subjects, 2632/2596	1.23 (1.09–1.39)	1.13 (1.00–1.29)
Among non-obese ^b , 814/1622	1.19 (0.98–1.44)	1.34 (1.10–1.63)
Among obese ² , 1818/974	1.29 (1.10–1.53)	0.99 (0.83–1.18)

^aAssociation of locus-specific African ancestry with type 2 diabetes adjusting for individual African ancestry, age, geographic region of residence, neighborhood socioeconomic status, years of education, and BMI (continuous)^bNon-obese, BMI < 30 kg/m²; obese, BMI ≥ 30 kg/m²

($r^2 = 0$). The effect C-allele has frequency of 2% in BWHS, 4% in AFR populations, and is absent in EUR populations.

At 10q25, the index SNP rs7903146 in the *TCF7L2* gene was associated with type 2 diabetes (OR = 1.20, $p = 2.4 \times 10^{-5}$). We identified a new signal rs114770437 associated with type 2 diabetes (OR = 1.37, $p = 6.4 \times 10^{-5}$) and uncorrelated with the index SNP in AFR populations ($r^2 = 0.03$). The effect G-allele has frequency of 93% in BWHS, 91% in AFR populations, and 100% in EUR populations.

At 12q14, the index SNP rs1531343 was not associated with type 2 diabetes (OR = 0.99, $p = 0.74$). We found a stronger marker, the rs2583943 SNP correlated with the index SNP in EUR populations ($r^2 = 0.67$) and associated with type 2 diabetes (OR = 1.25, $p = 5.7 \times 10^{-5}$). In addition, we identified a new signal rs116333053 associated with type 2 diabetes (OR = 1.44, $p = 1.8 \times 10^{-4}$),

uncorrelated with the index SNP in both EUR and AFR populations ($r^2 < 0.01$ in both populations).

Discussion

There were several findings of interest from our admixture mapping and fine-mapping work.

African ancestry and type 2 diabetes

We found that higher individual African ancestry was associated with higher risk of diabetes. Even after adjustment for neighborhood SES and years of education, we observed a clear linear relationship between individual African ancestry and risk of diabetes with 16% higher risk of diabetes per each 10% increase of individual African ancestry. African ancestry was previously found to be associated with higher risk of type 2 diabetes in combined data from the Atherosclerosis Risk in Communities (ARIC) Study and Jackson Heart Study (JHS), even after adjustment for measures of SES [24]. In that analysis, subjects in the upper tertile of African ancestry had 37% higher risk of type 2 diabetes compared to subjects in the lower tertile [24]. As comparison, we found in the BWHS that women in the upper quintile of African ancestry had 57% higher risk of type 2 diabetes relative to the first quintile after adjustment for SES. These results suggest that differences in SES do not completely explain the excess risk due to African ancestry, and genetic factors may account in part for the higher burden of type 2 diabetes in African Americans.

BMI may be playing a partial mediating role between African ancestry and type 2 diabetes as shown by our

Table 4 Type 2 diabetes GWAS index SNPs with $P < 0.05$ in the Black Women's Health Study

Locus	Chromosome	SNP ^a	GWAS OR	Alleles ^b	EAF ^c			BWHS results	
					BWHS	AFR	EUR	OR ^d (95% CI)	P-value
<i>GCKR</i>	2p23	rs780094	1.06	C/T	0.82	0.87	0.59	1.11 (1.00–1.23)	0.049
<i>ANKRD55</i>	5q11	rs459193	1.08	G/A	0.58	0.53	0.71	1.09 (1.01–1.18)	0.037
<i>KCNK16</i>	6p21	rs1535500	1.13	T/G	0.83	0.90	0.49	0.88 (0.79–0.98)	0.021
<i>DGKB</i>	7p21	rs6960043	1.06	C/T	0.24	0.17	0.47	0.89 (0.81–0.98)	0.016
<i>JAZF1</i>	7p15	rs864745	1.10	T/C	0.74	0.79	0.48	1.14 (1.04–1.25)	0.0039
<i>TP53INP1</i>	8q22	rs896854	1.06	T/C	0.69	0.74	0.46	1.18 (1.08–1.29)	1.8×10^{-4}
<i>SLC30A8</i>	8q24	rs13266634	1.12	C/T	0.90	0.93	0.72	1.27 (1.11–1.45)	4.7×10^{-4}
<i>TCF7L2</i>	10q25	rs7903146	1.40	T/C	0.30	0.26	0.32	1.20 (1.10–1.31)	2.4×10^{-5}
<i>KCNQ1</i>	11p15	rs231362	1.08	G/A	0.77	0.80	0.50	1.14 (1.04–1.26)	0.0052
<i>DUSP9</i>	Xq26	rs5945326	1.27	A/G	0.79	0.82	0.75	1.11 (1.01–1.22)	0.038

^aAll index SNPs were directly genotyped in the Affymetrix Axiom 45K custom array

^bEffect allele/reference allele

^cEffect allele frequency; *BWHS* Black Women's Health Study, *AFR* African population from 1KGP, *EUR* European population from 1KGP

^dOdds ratios adjusted for % African ancestry, age, and geographic region of residence

analysis. The association of African ancestry with type 2 diabetes was mostly observed among non-obese subjects, and African ancestry was associated with higher BMI among controls. We estimated that BMI explained about 24% of the observed linear association of African ancestry with risk of type 2 diabetes. Cheng et al. [24] also reported a positive association of African ancestry with BMI, although they did not present BMI-stratified results of the association of African ancestry with type 2 diabetes.

Admixture mapping

Our whole-genome admixture mapping identified two genomic loci, 3q26.33 and 12q23.1, with excess of local African ancestry in cases and confirmed in the case-control analysis.

The 3q26.33 region contains GWAS-identified SNPs associated with type 2 diabetes in Chinese, Malays, and Indians from Singapore [25], acute insulin response in Mexican Americans [26], and total and LDL cholesterol changes in response to fenofibrate in patients with type 2 diabetes [27]. The association of local African ancestry and risk of type 2 diabetes seems to be similar among non-obese and obese subjects, suggesting that the observed association is not modified by body weight.

The admixture signal at 12q23.1 is located about 5 Mb from the insulin-like growth factor 1 (*IGF1*) gene, which harbors genetic variants associated with fasting insulin and insulin resistance in subjects of European ancestry [6, 28]. The association of local African ancestry and risk of type 2 diabetes was observed only among non-obese individuals, suggesting that body weight may be a modifier of the observed association.

Replication and fine-mapping

Results of our replication efforts showed that some risk SNPs are shared across different ancestry groups. We were able to nominally replicate associations for 8 out of 71, or about 11%, of the examined index SNPs. Our results are similar to findings reported from the meta-analysis of type 2 diabetes in African Americans (MEDIA) Consortium, which replicated associations for 17 out of 104 (or about 16%) index SNPs [14]. In addition, the present study is the first one to report an association of the X-linked rs5945326 SNP near the *DUSP9* gene in African Americans, as the previous MEDIA study did not include X-chromosome data.

Our fine-mapping results suggested evidence of independent signals at five loci (*PROX1* at 1q32, *ANKRD55* at 5q11, *GLIS3* at 9p24, *TCF7L2* at 10q25, and *HMGA2* at 12q14), and evidence of a stronger signal than the European index SNP at *HMGA2* at 12q14. Our new SNPs at 1q32 (rs12091447) and 5q11 (rs7317858) were not correlated ($r^2 < 0.01$) with the SNPs identified by the MEDIA consortium at these same loci: rs7548778 and rs459193, respectively [14]. In addition, our new SNP rs114560781 at 9p24 was not correlated ($r^2 < 0.01$) with any of 10 SNPs with locus-wide significant results that were reported by the MEDIA consortium at this same locus [14]. Thus, our findings at 1q32, 5q11, and 9p24 may be false positive results.

We had previously reported about the presence of a new signal at *TCF7L2* tagged by SNP rs114770437 that is independent of the index SNP rs7903146 [29]. SNP rs114770437 is monomorphic in 1000 Genomes European, East Asian, and South Asian populations. This may explain

Table 5 New significant signals in type 2 diabetes loci in the Black Women's Health Study

Locus	Index SNP		Best SNP ^a		r^2 with index EUR/AFR		
	SNP, alleles ^b	OR ^d (95% CI), <i>P</i> -value	EAF ^c				
			BWHS	EUR			
<i>PROX1</i> 1q32	rs2075423, G/T	1.03 (0.95–1.12), 0.49	0.63	0.97	0.98	1.55 (1.24–1.95), 1.1×10^{-4}	<0.01/<0.01
<i>ANKRD55</i> 5q11	rs459193, G/A	1.09 (1.01–1.18), 0.037	0.71	0.11	0.001	1.29 (1.14–1.47), 7.1×10^{-5}	0.00/0.0
<i>GLIS3</i> 9p24	rs7041847, A/G	1.10 (0.98–1.24), 0.11	0.53	0.02	0.0	1.80 (1.35–2.39), 6.6×10^{-5}	0.00/0.0
<i>TCF7L2</i> 10q25	rs7903146, T/C	1.20 (1.10–1.31), 2.4×10^{-5}	0.32	0.93	1.0	1.37 (1.18–1.61), 6.4×10^{-5}	0.00/0.03
<i>HMGA2</i> 12q14	rs1531343, C/G	0.99 (0.91–1.07), 0.74	0.13	0.81	0.16	1.25 (1.12–1.39), 5.7×10^{-5}	0.67/0.03
				0.05	0.001	1.44 (1.19–1.74), 1.8×10^{-4}	<0.01/<0.01

^aMost significant SNP with locus-wide significant fine-mapping results^bEffect allele/reference allele^cEffect allele frequency in the current study (BWHS), and European (EUR) ancestry populations from 1000 Genomes Project^dOdds ratios adjusted for % African ancestry, age, and geographic region of residence^eImputed SNPs: rs12091447 info = 0.93; rs114560781 info = 0.95; rs116333053 info = 0.94

why an independent signal at *TCF7L2* has not been found in previous GWAS of European and Asian ancestry subjects. The MEDIA consortium [14] found a group of SNPs—uncorrelated with the index SNP rs7903146 ($r^2 < 0.05$)—with locus-wide significant associations with type 2 diabetes, in support of our finding of a new independent signal at *TCF7L2*. Three of these SNPs (rs7896811, rs11196199, and rs11196203) have moderate correlations ($r^2 = 0.3$) with rs114770437, suggesting that all of these SNPs may be tagging the same causal SNP. It is noteworthy that we had previously postulated the existence of more than one causal variant at *TCF7L2* [30], based on the presence of several enhancers with both in vitro and in vivo activity in the *TCF7L2* gene [31, 32]. The rs114770437 SNP is located inside one of the enhancer elements identified by Savic et al. [31, 32].

Our strongest association at *HMGA2* SNP rs2583943 may represent a better tag of the causal variant in African ancestry populations. Because the rs2583943 SNP has weak correlations ($r^2 = 0.16$) with two SNPs (rs11175944 and rs1480475) identified by the MEDIA consortium at the *HMGA2* locus, future expanded fine-mapping efforts would help to identify the causal variant(s) at this locus in African American populations.

Our study has several limitations. First, our sample size of 2632 cases and 2596 controls gives only 40% power to detect an OR of 1.10 for a risk allele with a frequency of 0.10. This low power may explain why we failed to replicate association of many of the index SNPs. Second, we used self-reported type 2 diabetes. However, our validation study showed that 96% of self-reported diabetes cases are confirmed in the medical record. Third, our custom genotyping array did not include many of the known type 2 diabetes loci as it was designed before publication of the latest GWAS. Fourth, prevalence of undiagnosed cases of diabetes among controls, although small and comparable to population estimates, may have resulted in underestimation of the true genetic associations. Fifth, although the threshold of call rate > 90% that we used in the QC of the Affymetrix Axiom 45K custom array is not the standard cutoff, we compared association results from genotyped data (shown in the present analysis) and imputed data of the index SNPs. Results were basically the same. Lastly, we cannot completely rule out the presence of unmeasured confounding of the association between genetic ancestry and risk of type 2 diabetes due to socio-economic factors that were not completely captured by either years of education or the neighborhood SES score.

In summary, the present study has several important findings. First, we found that African ancestry was associated with higher risk of type 2 diabetes. This association was mediated in part by BMI, and was not completely explained by differences in SES. Second, our genome-

wide admixture mapping identified two genomic loci, 3q26.33 and 12q23.1, which may harbor genetic risk variants that explain in part the higher risk of type 2 diabetes observed in African American subjects. Third, we were able to replicate several of the index SNPs from European ancestry GWAS. In particular, ours is the first study to replicate the association of the X-linked rs5945326 SNP in African American subjects. Finally, our fine-mapping efforts suggest the presence of new independent signals in at least five loci. Although in four of the loci our strongest SNPs did not correlate with reported SNPs by the MEDIA consortium, findings from the same consortium support the existence of new genetic risk variants in those loci. Future work should help to identify these new causal variants.

Acknowledgements We thank the Black Women's Health Study participants for their continuing participation in this research effort.

Funding This work was supported by grants R01MD007015 (to EARN) from the National Institute on Minority Health and Health Disparities (<http://www.nimhd.nih.gov/>); R01CA058420 (to L.R.), R01CA098663 (to J.R.P.), and UM1CA164974 (to L.R.) from the National Cancer Institute (<http://www.cancer.gov/>); and 11SDG7390014 (to E.A.R.N.) from the American Heart Association (<http://www.heart.org/HEARTORG/>).

Disclaimer The content of this article is solely the responsibility of the authors and does not necessarily represent the official views of the National Institute on Minority Health and Health Disparities, the National Cancer Institute, the National Institutes of Health, or the American Heart Association. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Compliance with ethical standards

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

Ethical approval The BWHS was approved by the institutional review board of Boston University (Boston, MA).

Informed consent All study subjects provided written informed consent for use of their saliva samples.

References

- Centers for Disease Control and Prevention. National diabetes statistics report, 2017. Atlanta, GA: Centers for Disease Control and Prevention, U.S. Department of Health and Human Services; 2017.
- Cowie CC, Rust KF, Byrd-Holt DD, Eberhardt MS, Flegal KM, Engelgau MM, et al. Prevalence of diabetes and impaired fasting glucose in adults in the U.S. population: National Health And Nutrition Examination Survey 1999–2002. *Diabetes Care*. 2006;29:1263–8.
- Lipton RB, Liao Y, Cao G, Cooper RS, McGee D. Determinants of incident non-insulin-dependent diabetes mellitus among blacks and whites in a national sample. The NHANES I Epidemiologic Follow-up Study. *Am J Epidemiol*. 1993;138:826–39.
- Signorello LB, Schlundt DG, Cohen SS, Steinwandl MD, Buchowski MS, McLaughlin JK, et al. Comparing diabetes prevalence between African Americans and Whites of similar socioeconomic status. *Am J Public Health*. 2007;97:2260–7.
- Mahajan A, Go MJ, Zhang W, Below JE, Gaulton KJ, Ferreira T, et al. Genome-wide trans-ancestry meta-analysis provides insight into the genetic architecture of type 2 diabetes susceptibility. *Nat Genet*. 2014;46:234–44.
- Dupuis J, Langenberg C, Prokopenko I, Saxena R, Soranzo N, Jackson AU, et al. New genetic loci implicated in fasting glucose homeostasis and their impact on type 2 diabetes risk. *Nat Genet*. 2010;42:105–16.
- Qi L, Cornelis MC, Kraft P, Stanya KJ, Linda Kao WH, Pankow JS, et al. Genetic variants at 2q24 are associated with susceptibility to type 2 diabetes. *Hum Mol Genet*. 2010;19:2706–15.
- Voight BF, Scott LJ, Steinthorsdottir V, Morris AP, Dina C, Welch RP, et al. Twelve type 2 diabetes susceptibility loci identified through large-scale association analysis. *Nat Genet*. 2010;42:579–89.
- Morris AP, Voight BF, Teslovich TM, Ferreira T, Segre AV, Steinthorsdottir V, et al. Large-scale association analysis provides insights into the genetic architecture and pathophysiology of type 2 diabetes. *Nat Genet*. 2012;44:981–90.
- Scott RA, Scott LJ, Magi R, Marullo L, Gaulton KJ, Kaakinen M, et al. An expanded genome-wide association study of Type 2 diabetes in Europeans. *Diabetes*. 2017;66:2888–902.
- Tsai FJ, Yang CF, Chen CC, Chuang LM, Lu CH, Chang CT, et al. A genome-wide association study identifies susceptibility variants for type 2 diabetes in Han Chinese. *PLoS Genet*. 2010;6:e1000847.
- Cho YS, Chen CH, Hu C, Long J, Ong RT, Sim X, et al. Meta-analysis of genome-wide association studies identifies eight new loci for type 2 diabetes in east Asians. *Nat Genet*. 2011;44:67–72.
- Palmer ND, McDonough CW, Hicks PJ, Roh BH, Wing MR, An SS, et al. A genome-wide association search for type 2 diabetes genes in African Americans. *PLoS One*. 2012;7:e29202.
- Ng MC, Shriner D, Chen BH, Li J, Chen WM, Guo X, et al. Meta-analysis of genome-wide association studies in african americans provides insights into the genetic architecture of type 2 diabetes. *PLoS Genet*. 2014;10:e1004517.
- Rosenberg L, Adams-Campbell L, Palmer JR. The Black Women's Health Study: a follow-up study for causes and preventions of illness. *J Am Med Women's Assoc*. 1995;50:56–8.
- Cozier YC, Palmer JR, Rosenberg L. Comparison of methods for collection of DNA samples by mail in the Black Women's Health Study. *Ann Epidemiol*. 2004;14:117–22.
- Vimalananda VG, Palmer JR, Gerlovin H, Wise LA, Rosenzweig JL, Rosenberg L, et al. Night-shift work and incident diabetes among African-American women. *Diabetologia*. 2015;58:699–706.
- Gillett MJ. International Expert Committee report on the role of the A1c assay in the diagnosis of diabetes: *Diabetes Care* 2009; 32 (7): 1327–1334. *Clin Biochem Rev*. 2009;30:197–200.
- Wise LA, Palmer JR, Spiegelman D, Harlow BL, Stewart EA, Adams-Campbell LL, et al. Influence of body size and body fat distribution on risk of uterine leiomyomata in U.S. black women. *Epidemiology*. 2005;16:346–54.
- Coogan PF, Cozier YC, Krishnan S, Wise LA, Adams-Campbell LL, Rosenberg L, et al. Neighborhood socioeconomic status in relation to 10-year weight gain in the Black Women's Health Study. *Obesity (Silver Spring)*. 2010;18:2064–5.
- Hoggart CJ, Shriver MD, Kittles RA, Clayton DG, McKeigue PM. Design and analysis of admixture mapping studies. *Am J Hum Genet*. 2004;74:965–78.

22. Montana G, Pritchard JK. Statistical tests for admixture mapping with case-control and cases-only data. *Am J Hum Genet.* 2004;75:771–89.
23. Palmer ND, Divers J, Lu L, Register TC, Carr JJ, Hicks PJ, et al. Admixture mapping of serum vitamin D and parathyroid hormone concentrations in the African American—Diabetes Heart Study. *Bone.* 2016;87:71–7.
24. Cheng CY, Reich D, Haiman CA, Tandon A, Patterson N, Selvin E, et al. African ancestry and its correlation to type 2 diabetes in African Americans: a genetic admixture analysis in three U.S. population cohorts. *PLoS ONE.* 2012;7:e32840.
25. Sim X, Ong RT, Suo C, Tay WT, Liu J, Ng DP, et al. Transferability of type 2 diabetes implicated loci in multi-ethnic cohorts from Southeast Asia. *PLoS Genet.* 2011;7:e1001363.
26. Palmer ND, Goodarzi MO, Langefeld CD, Wang N, Guo X, Taylor KD, et al. Genetic variants associated with quantitative glucose homeostasis traits translate to type 2 diabetes in Mexican Americans: the GUARDIAN (Genetics Underlying Diabetes in Hispanics) Consortium. *Diabetes.* 2015;64:1853–66.
27. Rotroff DM, Pijut SS, Marvel SW, Jack JR, Havener TM, Pujol A, et al. Genetic variants in HSD17B3, SMAD3, and IPO11 impact circulating lipids in response to fenofibrate in individuals with type 2 diabetes. *Clin Pharmacol Ther.* 2018;103:712–21.
28. Manning AK, Hivert MF, Scott RA, Grimsby JL, Bouatia-Naji N, Chen H, et al. A genome-wide approach accounting for body mass index identifies genetic variants influencing fasting glycemic traits and insulin resistance. *Nat Genet.* 2012;44:659–69.
29. Haddad SA, Palmer JR, Lunetta KL, Ng MC, Ruiz-Narvaez EA. A novel TCF7L2 type 2 diabetes SNP identified from fine mapping in African American women. *PLoS ONE.* 2017;12:e0172577.
30. Ruiz-Narvaez EA. Redundant enhancers and causal variants in the TCF7L2 gene. *Eur J Hum Genet.* 2014;22:1243–6.
31. Savic D, Bell GI, Nobrega MA. An in vivo cis-regulatory screen at the type 2 diabetes associated TCF7L2 locus identifies multiple tissue-specific enhancers. *PLoS ONE.* 2012;7:e36501.
32. Savic D, Park SY, Bailey KA, Bell GI, Nobrega MA. In vitro scan for enhancers at the TCF7L2 locus. *Diabetologia.* 2013;56:121–5.