DNA Methylation Markers of Type 2 Diabetes Mellitus Among Male Veterans With or Without Human Immunodeficiency Virus Infection

Raina Mathur,¹ Qin Hui,¹ Yunfeng Huang,¹ Marta Gwinn,¹ Kaku So-Armah,² Matthew S. Freiberg,³ Amy C. Justice,^{4,5} Ke Xu,^{4,6} Vincent C. Marconi,^{7,8,9} and Yan V. Sun^{1,7,10}

Departments of ¹Epidemiology and ⁹Global Health, Emory University Rollins School of Public Health, and ²Division of Infectious Diseases and ¹⁰Department of Biomedical Informatics, Emory University School of Medicine, Atlanta, and ⁷Atlanta Veterans Affairs Medical Center, Georgia, ³Boston University School of Medicine, Massachusetts, ⁴Cardiovascular Medicine Division, Vanderbilt University School of Medicine and Tennessee Valley Healthcare System, Nashville, ⁵Connecticut Veteran Health System, West Haven, and ⁶Department of Psychiatry, ⁵Yale School of Medicine, New Haven, Connecticut

Epigenetic modifications such as DNA methylation are associated with both human immunodeficiency virus (HIV) infection and type 2 diabetes mellitus (T2DM). We investigated epigenetic associations with T2DM according to HIV infection status and assessed interaction effects among 681 male participants of the Veterans Aging Cohort Study. Methylation at previously reported sites, cg1963031 (*TXNIP*), cg18181703 (*SOC3*), and cg09152259 (*PROC*), was significantly associated with T2DM in HIV-infected individuals. We identified 3 novel associations with suggestive statistical significance: cg1231141 (*ADAMTS2*), cg19534769 (*HGFAC*), and cg13163919 (*TLE3*). Suggestive interaction with HIV infection status was found at cg17862404 (*TSC22D1*). The implicated genes are involved in inflammation, pancreatic β -cell function, and T2DM pathogenesis.

Keywords. DNA methylation; epigenome; HIV; Type 2 Diabetes.

As antiretroviral therapy (ART) has become more effective at lengthening the lives of persons with human immunodeficiency virus (HIV) infection, they are at greater risk for age-related diseases, such as type 2 diabetes mellitus (T2DM), cardiovascular disease, kidney disease, and cancer [1]. In a study of prevalent T2DM in HIV-infected and HIV-uninfected veterans from the Veterans Aging Cohort Study (VACS) study, Butt et al found that HIV infection itself was not associated with increased risk of T2DM; however, the effects of increasing age, hepatitis C virus coinfection, and high body mass index (BMI) were

magnified in veterans with HIV infection, whose risk of T2DM increased further with long-term ART [2]. Genetic and molecular mechanisms related to the development of T2DM among persons with HIV infection have not yet been fully investigated.

Epigenetic studies can help elucidate how environmental factors and comorbid conditions may contribute to disease susceptibility. The most widely studied epigenetic modification is DNA methylation of the cytosine residue of a C-G dinucleotide (CpG), although other forms of methylation occur as well. Epigenetic changes have been associated with T2DM and its risk factors, including age, obesity, physical activity, and diet [3]. Previous candidate gene studies with limited sample size and genomic coverage have reported differential methylation within a number of genes involved in pathways underlying T2DM. A large epigenome-wide association study (EWAS) of T2DM recently identified and replicated a CpG site in *TXNIP*, a previously reported T2DM-associated gene, and novel CpG sites within *ABCG1*, *PHOSPHO1*, *SOC3*, and *SREBF1* [4]. HIV infection is also known to be associated with epigenomic profile [5] and epigenetic age [6]. We hypothesized that HIV infection or its treatment could result in epigenetic modification of genomic loci that affect the risk of developing T2DM. We performed an EWAS to examine differential DNA methylation associated with T2DM in HIV-infected and uninfected participants in the VACS cohort.

METHODS

VACS Data Set

VACS is a prospective, observational study of HIV-infected and HIV-uninfected veterans matched for age, race/ethnicity, and sex and enrolled from 8 Veterans Affairs (VA) facilities beginning in 1997. Clinical and laboratory information were collected from each patient's electronic medical record. Total white blood cell counts and CD4⁺ T-cell subsets were analyzed at the time of peripheral blood sample collection. Our study included a subset of 564 HIV-infected and 117 HIV-uninfected male VACS participants, collected initially for an EWAS investigating intravenous drug abuse [7]. The study was approved by the Veteran's Administration Research and Development Committee and the Institutional Review Board of the Atlanta Veteran's Administration. All study participants signed an informed consent. VACS defined T2DM as (1) glucose level ≥ 200 mg/dL on 2 separate occasions or (2) glucose level ≥ 200 mg/dL on 1 occasion plus treatment with an oral hypoglycemic or insulin for ≥ 30 days. We used Student *t* and χ^2 tests to compare continuous and categorical variables in HIV-infected and HIV-uninfected groups.

Received 22 August 2018; editorial decision 7 January 2019; accepted 10 January 2019; published online January 11, 2019.

Correspondence: Y. V. Sun, Department of Epidemiology, Emory University, 1518 Clifton Rd NE, Room 3049, Atlanta, GA 30322 (yvsun@emory.edu).

The Journal of Infectious Diseases® 2019;219:1959–62

© The Author(s) 2019. Published by Oxford University Press for the Infectious Diseases Society of America. All rights reserved. For permissions, e-mail: journals.permissions@oup.com. DOI: 10.1093/infdis/jiz023

DNA Methylation Assessment and Statistical Analysis

DNA methylation was measured in samples of peripheral blood mononuclear cells using the Illumina 450K platform, as described elsewhere [5]. A total of 927 CpG sites were removed from the analysis owing to a high missing rate (>5%). Another 35 605 CpG sites were removed because their probes contained single-nucleotide polymorphisms within 10 base pairs of the CpG site, including 2 reported T2DM-associated sites (cg06500161 and cg11024682) [4]. An additional 24 729 CpG sites were removed because their probes mapped to multiple locations in the genome. After quality control procedures, 412 583 autosomal and 11 232 X chromosome sites from 681 individuals remained for epigenetic association analysis. The data were normalized using subset-quantile within array normalization in the *minfi* package in R software. Heterogeneity of cell type proportions in blood and other tissues is a well-established confounder in epigenetic association studies and was adjusted for in this study using calculated cell type proportions (CD4⁺ T cells, CD8⁺ T cells, natural killer cells, B cells, monocytes, and granulocytes) in the blood [8].

We used a linear mixed effect regression (the nlme package in R) to model DNA methylation as a function of T2DM status, controlling for HIV status, current smoking status, chronological age, BMI, hepatitis C virus infection, hypertension, and cell type proportions. A random intercept was included to correct for the potential correlation of samples processed on the same array (eg, batch effect). We then repeated this analysis, stratified on HIV infection status. We also adjusted for CD4 T-cell counts in the HIV-infected group. Finally, we assessed potential interaction effects between T2DM and HIV infection status. These 3 analyses were also performed using DNA methylation data for the X chromosome, as described elsewhere [9]. For the replication analyses of published T2DM-associated sites, the threshold for statistical significance was set at .05. For the EWAS, a multiple-testing corrected significance threshold ($P < 1 \times 10^{-7}$ or false discovery rate <0.05) was used to control false-positive findings, and a threshold of $P < 1 \times 10^{-5}$ was used to indicate “suggestive” statistical significance, to report additional epigenetic associations for potential replication studies.

DNA methylation age was calculated using the algorithm developed by Horvath, using age-related CpG sites from the

Illumina Infinium platform (450K) [10]. The correlation between the calculated epigenetic age and the self-reported chronological age was assessed. The difference between chronological age and DNA methylation age was calculated to represent accelerated epigenetic aging [6]. We carried out multiple logistic regression analyses to examine whether the age difference was independently associated with T2DM controlled for chronological age.

RESULTS

Descriptive statistics for the study population, stratified by HIV infection status, are summarized in [Supplementary Table 1](#). The median HIV viral load for the HIV-infected men was 75 copies/μL, and 83.2% of them were receiving ART. Neither HIV viral load nor ART status was associated with T2DM among HIV-infected participants. The prevalence of T2DM was higher in the HIV-uninfected group than in the HIV-infected group (35% vs 20%, respectively).

To replicate previously reported T2DM-associated CpG sites, we conducted multiple regression analyses in pooled samples. We observed significant associations for cg1963031 (*TXNIP* gene) and cg18181703 (*SOC3* gene) of inverse association with T2DM, as identified by Chambers et al [4] ([Table 1](#) and [Supplementary Figure 1](#)). Stratified by HIV infection status, the association of cg1963031 (*TXNIP*) was replicated in both HIV-infected and HIV-uninfected groups ([Table 1](#)). CpG sites cg18181703 and cg09152259 (*PROC* gene) were significantly associated with T2DM in the HIV-infected group but not in the HIV-uninfected group ([Table 1](#)). For cg1963031 (*TXNIP*), we observed an approximately 2-fold difference between the β coefficients in the HIV-infected and HIV-uninfected groups ($\beta = -.042$ and $\beta = -.020$, respectively), but this difference was marginally significant ($P = .08$). We found no evidence of interaction for any of the other previously reported CpG sites.

The regional plot for chromosome 1 ([Supplementary Figure 1](#)) shows the base-pair location and $-\log_{10}(P \text{ value})$ of CpG sites mapped to the *TXNIP* gene. In all, we tested 17 CpG sites within the *TXNIP* gene and 136 CpG sites within ± 200 kb of the flanking region in the HIV-infected group. We found no significant associations with any of these CpG sites except cg1963031.

Table 1. Replication of Candidate CpG Sites Associated With T2DM in All Study Participants and Stratified by HIV Status

CpG Site	Chromosome	Gene Name	Position	All Participants		HIV-Infected Participants		HIV-Uninfected Participants	
				Regression Coefficient (SE)	PValue	Regression Coefficient (SE)	PValue	Regression Coefficient (SE)	PValue
cg18181703	17	<i>SOC3</i>	763 54621	−0.013 (0.004)	5.1×10^{-3}	−0.012 (0.005)	.02	−0.016 (0.011)	.13
cg19693031	1	<i>TXNIP</i>	145 441 552	−0.020 (0.005)	2.0×10^{-5}	−0.017 (0.005)	8.4×10^{-4}	−0.042 (0.012)	7.2×10^{-4}
cg02650017	17	<i>PHOSPHO1</i>	47 301 614	−0.003 (0.002)	.20	−0.002 (0.003)	.38	2.8×10^{-4} (0.005)	.95
cg04999691	7	<i>C7orf29</i>	1 500 27050	−0.002 (0.002)	.15	−0.003 (0.002)	.20	−0.002 (0.004)	.59
cg09152259	2	<i>PROC</i>	128 156 114	−0.010 (0.005)	.03	−0.009 (0.005)	.07	−0.012 (0.013)	.37

Abbreviations: HIV, human immunodeficiency virus; SE, standard error; T2DM, type 2 diabetes mellitus.

In the entire study population (Supplementary Figures 2 and 3), no CpG sites achieved epigenome-wide significance ($P < 10^{-7}$). When we applied a relaxed significance threshold ($P < 10^{-5}$), we found suggestive associations with 5 novel CpG sites, including cg01920980 (*ST18*) and cg03575666 (*SFRP1*) (Table 2). Our EWAS analyses restricted to the HIV-infected (Supplementary Figures 4 and 5) and HIV-uninfected groups (Supplementary Figures 6 and 7), respectively, found no epigenome-wide significant associations. However, in the HIV-infected group, 9 CpG sites met the threshold for suggestive association (Table 2).

Our EWAS analysis of interaction between T2DM status and HIV infection status produced no statistically significant results; however, we observed an interesting effect for cg17862404 (*TSC22D1*). Associations with this site differed between the HIV-infected and uninfected groups in both magnitude and direction (Supplementary Table 2).

The Manhattan and Q-Q plots for all 3 EWAS analyses supported the lack of epigenome-wide significant findings. The Q-Q plots for the EWAS analyses of the entire study population and the HIV-infected group showed moderate inflation (inflation factors of 1.03 and 1.12, respectively), which did not require further correction for global inflation. The Q-Q plot for the EWAS interaction analysis showed deflation (inflation factor, 0.87).

An initial analysis of the data showed a strong correlation between chronological age and DNA methylation age ($r = 0.85$; Supplementary Figure 8). Although the median chronological age was the same in the HIV-infected and HIV-uninfected groups, the median DNA methylation age in the HIV-infected group was 52.9 years, compared with 51.8 years in the uninfected group. This difference in DNA methylation age was statistically significant ($P = 4.0 \times 10^{-3}$). DNA methylation age was

significantly associated with T2DM status ($P = 1.5 \times 10^{-4}$), but the association with T2DM was not independent after adjustment for chronological age ($P = .14$).

DISCUSSION

Our study replicated the previously reported T2DM-associated site cg19693031 (*TXNIP* gene), which had a stronger association with T2DM in the HIV-uninfected than in the HIV-infected group. *TXNIP* encodes thioredoxin-interacting protein, which is a regulator of cellular redox signaling that protects cells from oxidative stress. Expression of *TXNIP* is inhibited by insulin and stimulated by hyperglycemic conditions within the body [11]. In prediabetes and diabetes, consistently elevated *TXNIP* expression in muscle cells is inversely associated with glucose uptake [11].

Our study also replicated T2DM-associated sites cg18181703 (*Socs3*) and cg09152259 (*PROC*) in the HIV-infected group. *Socs3*-knockout mice have comparatively more normal adiposity and energy expenditure and are protected against the development of hyperinsulinemia and insulin resistance [12]. *PROC* provides instructions for protein C, an important factor for controlling blood clotting and inflammation. Activated protein C, a derivative of protein C, functions in reducing inflammation and apoptosis and prevents pancreatic β -cell death [13].

The pooled EWAS analyses combining both HIV-infected and HIV-uninfected groups found only suggestive association with cg01920980 and cg03575666, which map to the *ST18* and *SFRP1* genes, respectively. Increased *ST18* expression, which stimulates pancreatic β -cell apoptosis, has been found in the pancreatic islet cells of obese animals. Increased *ST18* expression has also been found to decrease insulin secretion in the pancreas [14]. Our EWAS in the HIV-infected group identified suggestive evidence for associations of cg1231141 (*ADAMTS2*), cg19534769

Table 2. CpG Sites With Suggestive Significance in Association With T2DM, Including All Study Participants and HIV-Infected Participants^a

Participants by CpG Site	Chromosome	Gene Name	Regression Coefficient (SE)	PValue
All participants				
cg08069675	12	<i>NOC4L</i>	-0.009 (0.002)	1.4×10^{-6}
cg14721531	6	<i>CDK19</i>	-0.007 (0.001)	3.9×10^{-6}
cg19534769	4	<i>HGFAC</i>	0.009 (0.002)	5.5×10^{-6}
cg03575666	8	<i>SFRP1</i>	0.021 (0.005)	7.4×10^{-6}
HIV-infected participants only				
cg19534769	4	<i>HGFAC</i>	0.011 (0.002)	4.7×10^{-7}
cg01231141	5	<i>ADAMTS2</i>	0.053 (0.011)	1.6×10^{-6}
cg01439753	16	<i>CMIP</i>	-0.005 (9.6×10^{-4})	2.7×10^{-6}
cg03732014	21	<i>BACH1</i>	0.013 (0.003)	4.0×10^{-6}
cg13163919	15	<i>TLE3</i>	-0.015 (0.003)	6.3×10^{-6}
cg19848732	17	NA	-0.005 (0.001)	7.1×10^{-6}
cg20480097	13	<i>MBNL2</i>	-0.004 (9.0×10^{-4})	8.3×10^{-6}
cg14836313	11	<i>MRPL23</i>	0.008 (0.002)	8.6×10^{-6}
cg12214399	4	NA	0.147 (0.033)	9.7×10^{-6}

Abbreviations: HIV, human immunodeficiency virus; NA, not available; SE, standard error; T2DM, type 2 diabetes mellitus.

^aThe suggestive significance threshold is $P < 1 \times 10^{-5}$.

(*HGFAC* gene), and cg13163919 (*TLE3*) with T2DM. The *HGFAC* gene regulates pancreatic β -cell mass and β -cell function; its expression is regulated by HNF1 α , a transcription factor that has been implicated in forms of monogenic diabetes [15].

Our study had several limitations, owing in part to VACS population characteristics and study procedures. Using results of a random plasma glucose test to diagnose T2DM could have missed some persons with fasting glucose levels >7 mmol/L. Because the data were not collected prospectively, we could not examine the temporal direction of the association between T2DM and DNA methylation. The VA population is not representative of the US population, including a higher proportion of African Americans, and our analysis was limited to male patients. The VA patient population has a high burden of chronic diseases, which could help explain why the prevalence of T2DM was approximately 1.7 times higher in the HIV-uninfected group than in the HIV-infected group.

We were unable to evaluate 2 of the reported candidate CpG sites because they did not meet quality control standards. Because of these limitations, this study provides only preliminary evidence for novel CpG candidate sites that may contribute to T2DM risk in HIV-infected populations. The epigenetic associations with T2DM that we identified in the cohort of persons with HIV infection need to be further replicated in independent samples and validated by other technologies (eg, pyrosequencing) in future studies.

Using rich phenotypic data from the VACS, we examined epigenetic association with T2DM in an HIV-infected population, controlling for confounders including race/ethnicity, BMI, chronological age, current smoking status, comorbid conditions, and cell type proportions. We identified several T2DM-associated CpG sites for further investigation in persons with HIV infection. Our analyses showed a few differences in epigenetic associations between the HIV-infected and HIV-uninfected groups. Although some of these differences may have arisen owing to unbalanced sample sizes between the HIV-infected and HIV-uninfected groups, these CpG sites are candidates for further investigation of T2DM in relation to HIV infection.

Supplementary Data

Supplementary materials are available at *The Journal of Infectious Diseases* online. Consisting of data provided by the authors to benefit the reader, the posted materials are not copyedited and are the sole responsibility of the authors, so questions or comments should be addressed to the corresponding author.

Notes

Financial support. This work was supported the National Institutes of Health Center for AIDS Research at Emory University (developmental grant P30 AI050409 to V. C. M. and Y. V. S.) and the National Institutes of Health (grant K01HL134147 to K. S. and grants R03DA039745 and K12DA000167 to K. X.).

Potential conflict of interest. All authors: No reported conflicts. All authors have submitted the ICMJE Form for Disclosure of Potential Conflicts of Interest. Conflicts that the editors consider relevant to the content of the manuscript have been disclosed.

References

1. Gallant J, Hsue PY, Shreay S, Meyer N. Comorbidities among US patients with prevalent HIV infection-A trend analysis. *J Infect Dis* **2017**; 216:1525–33.
2. Butt AA, McGinnis K, Rodriguez-Barradas MC, et al; Veterans Aging Cohort Study. HIV infection and the risk of diabetes mellitus. *AIDS* **2009**; 23:1227–34.
3. Kwak SH, Park KS. Recent progress in genetic and epigenetic research on type 2 diabetes. *Exp Mol Med* **2016**; 48:e220.
4. 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.
5. Zhang X, Justice AC, Hu Y, et al. Epigenome-wide differential DNA methylation between HIV-infected and uninfected individuals. *Epigenetics* **2016**; 11:750–60.
6. Nelson KN, Hui Q, Rimland D, et al. Identification of HIV infection-related DNA methylation sites and advanced epigenetic aging in HIV-positive, treatment-naive U.S. veterans. *AIDS* **2017**; 31:571–5.
7. Zhang X, Hu Y, Justice AC, et al. DNA methylation signatures of illicit drug injection and hepatitis C are associated with HIV frailty. *Nat Commun* **2017**; 8:2243.
8. Houseman EA, Accomando WP, Koestler DC, et al. DNA methylation arrays as surrogate measures of cell mixture distribution. *BMC Bioinformatics* **2012**; 13:86.
9. Klebaner D, Huang Y, Hui Q, et al. X chromosome-wide analysis identifies DNA methylation sites influenced by cigarette smoking. *Clin Epigenetics* **2016**; 8:20.
10. Horvath S. DNA methylation age of human tissues and cell types. *Genome Biol* **2013**; 14:R115.
11. Parikh H, Carlsson E, Chutkow WA, et al. TXNIP regulates peripheral glucose metabolism in humans. *PLoS Med* **2007**; 4:e158.
12. Jorgensen SB, O'Neill HM, Sylow L, et al. Deletion of skeletal muscle SOCS3 prevents insulin resistance in obesity. *Diabetes* **2013**; 62:56–64.
13. Xue M, Jackson CJ. Activated protein C and its potential applications in prevention of islet β -cell damage and diabetes. *Vitam Horm* **2014**; 95:323–63.
14. Henry C, Close AF, Buteau J. A critical role for the neural zinc factor ST18 in pancreatic β -cell apoptosis. *J Biol Chem* **2014**; 289:8413–9.
15. Ohki T, Sato Y, Yoshizawa T, Yamamura K, Yamada K, Yamagata K. Identification of hepatocyte growth factor activator (*Hgf*) gene as a target of HNF1 α in mouse β -cells. *Biochem Biophys Res Commun* **2012**; 425:619–24.