**Distribution Agreement**

In presenting this thesis or dissertation as a partial fulfillment of the requirements for an advanced degree from Emory University, I hereby grant to Emory University and its agents the non-exclusive license to archive, make accessible, and display my thesis or dissertation in whole or in part in all forms of media, now or hereafter known, including display on the world wide web. I understand that I may select some access restrictions as part of the online submission of this thesis or dissertation. I retain all ownership rights to the copyright of the thesis or dissertation. I also retain the right to use in future works (such as articles or books) all or part of this thesis or dissertation.

Signature:

_____ _____

Yutong Yao Date

The Role of BMI-related DNA Methylations on Type II Diabetes Mellitus Among Male Veterans with HIV

By

Yutong Yao
Master of Public Health

Epidemiology

_____ [Chair's signature]
Yan V. Sun, PhD
Committee Chair

# The Role of BMI-related DNA Methylations on Type II Diabetes Mellitus Among Male Veterans with HIV

By

Yutong Yao

B.S.
University of Notre Dame
2019

Thesis Committee Chair: Yan V. Sun, PhD

An abstract of
A thesis submitted to the Faculty of the
Rollins School of Public Health of Emory University
in partial fulfillment of the requirements for the degree of
Master of Public Health
in Epidemiology
2021

# Abstract

The Role of BMI-related DNA Methylations on Type II Diabetes Mellitus
Among Male Veterans with HIV
By Yutong Yao

**Background:** People with HIV (PWH) are facing a significant burden and risk of non-communicable diseases. Increasing development of T2DM and related mortality among the HIV-infected cohort suggests a need for elucidating the association between diabetes and HIV. High body mass index (BMI) is identified as a common risk factor for T2DM among PWH in cohort studies; epigenetic studies have identified BMI-associated DNA methylation (DNAm) sites, which may lead to T2DM in general populations. However, few studies have examined the association between BMI-related methylations and T2DM among people living with HIV.

**Method:** We examined the association between BMI-related methylation sites and T2DM status among male veterans with HIV from the Veterans Aging Cohort Study (VACS). BMI-related methylation sites were identified from previous BMI EWAS in general population and replicated among people with HIV adjusted for multiple testing. The associations were estimated with and without the adjustment by phenotypic BMI.

**Results:** Seven previously reported BMI-related DNAm were replicated among the cohort of PWH. cg11024682 (*SREBF1*) was significantly associated with T2DM development among the HIV-infected cohort (*p*-value: $4.67\times10^{-5}$, Bonferroni-corrected threshold: $7.14\times10^{-3}$). After adjusting for phenotypic BMI in the model, the association was still significant with cg11024682 (*p*-value: $3.59\times10^{-3}$, Bonferroni-corrected threshold: $7.14\times10^{-3}$).

**Conclusion:** Some BMI-related DNA methylation sites are associated with T2DM among PWH. The epigenetic association with T2DM may be independent from the effect of phenotypic BMI.

# The Role of BMI-related DNA Methylations on Type II Diabetes Mellitus Among Male Veterans with HIV

By

Yutong Yao

B.S.
University of Notre Dame
2019

Thesis Committee Chair: Yan V. Sun, PhD

A thesis submitted to the Faculty of the
Rollins School of Public Health of Emory University
in partial fulfillment of the requirements for the degree of
Master of Public Health
in Epidemiology
2021

Acknowledgements

Everything during the COVID pandemic has become different and challenging. Despite all the difficulties and challenges, many people have relentlessly supported and helped me through this. I would like to express my genuine gratitude to them who supported me through finishing my thesis.

First of all, I want to thank my thesis advisor Dr. Yan Sun for his thoughtful suggestions and tremendous patience throughout my thesis work. His expertise in research methodology and continuous support has kept me on the correct track of my thesis. I am also grateful for all the computation and data support from Dr. Qin Hui and Zeyuan Wang. Coding and data are essential to my research project. Without their technical support, I would have no idea to start the work. I also would like to thank every member in the lab, who have kept an eye on my mistakes and provided insightful suggestions. I would also like to express my gratitude to Julia Sobolik from the Academic Resource Center, who read through my draft and helped me organize my draft.

I am also grateful for all the support from my roommates, friends, and my family who have been there providing all kinds of support and encouragement throughout my work.

# Table of Contents

# I. Background

## Summary of Diseases

Type II Diabetes Mellitus (T2DM), increasing and spreading over the world since the last few decades, is a major non-communicable disease caused by insulin resistance. Characterized by insufficient insulin secretion by pancreatic islet ß cells and decreased insulin sensitivity, T2DM is categorized as a metabolic disorder that prevents the body from processing glucose properly [1]. Main negative outcomes of T2DM include kidney diseases, lower extremity amputations and retinopathy with blindness [2].

HIV, or the human immunodeficiency virus can be spread of blood fluids. Infections can put individuals at risk of weakened immune system with loss of CD4 T cells [3]. At the end of 2019, 38 million people including adults and children were estimated wot be infected and living with HIV globally. Among this HIV-infected population, 67% were taking antiretroviral therapy, which is estimated to have saved 15.3 million lives between 2000 and 2019 [4]. The daily-based treatment is called antiretroviral therapy (ART), combined antiretroviral therapy (cART) or highly active antiretroviral therapy (HAART), which decreases the virus load in the body and prevent individuals from progressing to the severe stage of HIV infection, acquired immunodeficiency syndrome (AIDS) [5].

## T2DM Burden Among HIV-infected Cohort

People living with HIV are facing a significant burden and risk of non-communicable diseases including Type 2 Diabetes Mellitus (T2DM). In recent years, Cross sectional studies have found a high prevalence of diabetes and hypertension among HIV-infected cohort, 19.6%, 26.6% and 15.1% for Ethiopia, Malawi and London respectively [6-8]. Furthermore, the risk of T2DM

seems to elevate in PWH in some cases. Duncan *et al*. and Ye *et al.* found that the risk of T2DM for PWH was two times that for the general population [8, 9]. The outcome can also be worse for the HIV infected cohort than the common public when developing T2DM comorbidity. A three-time higher mortality rate was identified in a retrospective cohort study data collected in the United States between 2006 and 2015 [10].

## Risk Factors of T2DM among PWH

Despite reported risks of T2DM higher in PWH, whether HIV status itself is associated with T2DM remains questionable. A meta-analysis on seven studies conducted in African population between 2008 and 2016 showed no significant differences in T2DM prevalence between HIV-infected and HIV-naive cohorts [11]. Rather than HIV status, ART and other non-HIV related factors have been found to be associated with T2DM among PWH.

### *Antiretroviral Treatment*

Drug treatment may also play a role in elevating T2D risks among PWH. In a large, matched cohort study including 13, 632 individuals from the South Carolina Medicaid system, protease inhibitor-based ART was associated with higher risk of diabetes among HIV-infected people with adjusted relative risk of 1.35 [12]. NNRTI efavirenz was more likely to cause incident diabetes compared with nevirapine with a hazard ratio of 1.33; NRTI zidovudine and stavudine were also reported to be associated with elevated risk of DM [13]. Other drugs found to play a role include protease inhibitors indinavir, saquinavir and NRTI didanosine [14]. Together with a high-fat diet, ART may exacerbate obesity and induce dysregulation of glucose metabolic pathways.

### *Non-HIV Related Factors*

The most common risk factors include obesity or overweight [8, 9, 12, 14-16], gender [12, 15], baseline hypertension [8, 12]; non-white ethnicity, dyslipidaemia, Hepatitis C infection, older

age [12], lipoatrophy [14], statin use over 6 months [9], and hepatic steatosis [8] were factors associated with DM risk among PWH. Age and BMI, however, may play distinct roles in T2DM development among PWH. Isa *et al.* found that patients>40 years of age were associated with T2DM risk at baseline among PWH, and incident diabetes was associated with BMI≥25 [16]. Dimela *et al.* found higher BMI and hypertension were associated with elevated DM risk instead of HAART after adjusting for BMI-defined overweight, hypertension, age, sex, smoking and family history [17].

## T2DM-Related Risk Factors

Obesity is one of the most well-established factors contributing to T2DM risks [18-23]. In a Finnish diabetes prevention study, every kilogram of weight loss was found to be associated with 16% reduction in T2DM development [24]. Biological pathways have also been studied to elucidate the association between BMI/obesity and T2DM. Pathways proposed include that intramyocellular triglyceride in bone, proinflammatory cytokines and adipose tissue macrophages induced by obesity [25-28]. Other factors contributing to T2DM include physical inactivity, smoking behavior, and alcohol consumption [29-31].

Genetic variation plays an essential role in T2DM development as well. Single nucleotide polymorphisms related to T2DM development have been identified using genome-wide association study (GWAS) and used to predict individual risks of T2DM [32-35]. In the latest GWAS meta-analysis with multiple ethnicities, 568 genetic associations were identified including autosomal and X chromosomal loci [35]. Functional enrichment analysis identified the involved biological pathways in AKT2 subnetwork, lung cancer, the GA1 signalosome, protein kinase binding, signal transduction and EGFR signaling [35]. Traits and diseases associated with T2D-

correlated genes included waist circumference, BMI, hypertension, coronary artery disease, dyslipidemia, alcohol intake and smoking [35].

## Epigenome-Wide Association Study (EWAS)

Mainly controlled by DNA methylation (DNAm), histone modification and microRNA, epigenetic regulation is subject to environmental factors and results in differential gene expression of DNA sequences. Common environmental modifiers of epigenetics include cigarette smoking, obesity, physical activity and diet [39]. While GWAS can be used to identify single nucleotide polymorphisms (SNPs) that are causal to diseases and traits, epigenome-wide association study (EWAS) can elucidate inter-individual and individual-environmental variations contributing to variation in phenotypes [36]. The most studied epigenetic modification is DNA methylation, measured by the methylation level of CpG sites on DNA. DNA methylation sites located on genes *TXNIP, ABCG1, PHOSPHO1, SOCS3* and *SREBF1* have been identified in multiple ethnicities as associated with T2DM through EWAS [37-39].

A wide range of EWAS have been conducted to determine the relationship between BMI and DNA methylation. In 2014, Dick *et al.* found that the first methylation sites associated with BMI — methylated HIF3A locus was positively associated with increased BMI within the European origin cohort though none of the sites were found to be significant [40]. Using the Framingham Heart Study and the Atherosclerosis Risk in the Communities (ARIC) cohorts, Aslibekyan *et al.* conducted an EWAS and found methylation sites at *CPT1A, PHGDH* and *CD38* were associated with BMI adjusting for age, gender, study sites, T-cell purity, smoking and family structure [41]. Meeks *et al.* conducted a BMI EWAS within African cohort in 2017 and identified 18 DNAm sites significantly associated with BMI and obesity [42]. Additionally, Mendelson *et al.* reported that the BMI-associated CpG site cg11024682 at *SREBF1* region linked with cardio

metabolic disease [43]. However, these early EWAS had limited sample sizes, which may lead to insufficient power to discover more epigenetic associations with BMI. While most EWAS did not determine the causal relationship between BMI and CpGs, the largest EWAS conducted by Wahl *et al.* in 2017 determined that the methylations mostly resulted from adiposity; in addition, the study identified 187 loci associated with BMI and 62 of them also predicted incident T2D [44]. The most significant and consistent epigenetic markers include *SOCS3, LGALS3BP,* and *ABCG1* regions. All these EWAS studies employed Illumina 450K as the methylation platform to measure DNA methylation from peripheral blood samples. The 450K array, though covered majority of 96% of human methylation islands, does not cover as much as the newest Illumina EPIC BeadChip [45], which surveys over 850,000 DNA sites, almost doubles the coverage of the previous 450K array. However, these studies do not address the epigenetic association with BMI specially among PWH cohorts.

Epigenetic studies have been done to study HIV-1 infection to characterize epigenetic changes correlated with HIV infection. Zhang *et al.* identified 5 CpGs located in gene *LPCAT1, NLRC5,* and *CD4* associated with HIV infection [46]. Methylations at *NLRC5* were correlated with HIV viral load, suggesting an epigenetic function of the gene in virus replication [46]. Nelson *et al.* identified DNA methylation at *VPS37B* associated with HIV infection among ART-naïve PWH [47]. PWH receiving AIDS treatment also had DNA methylations that associated with HIV infection status and response to virus, interferon signaling pathway among people receiving AIDS treatments [48]. In addition to correlating with virus load and ART among PWH, HIV infection was also found to be associated with higher DNA methylation ages compared to HIV-negative groups [47, 49].

Profiling DNA methylations related to T2DM among PWH, Mathur el al. identified novel sites at *ADAMTS2, HGFAC* and *TLE3* among PWH in addition to previously reported sites at *TXNIP, SOCS3* and *PROC* among general population with European and Indian Asian ancestry[50]. With few available studies on the epigenetic markers or mechanisms of developing T2DM within HIV, research needs to address the gap of the roles phenotypic factors play in T2DM development specifically within PWH cohort.

## Research Gap and Our Study

The current gap in the research of T2DM risks among PWH is that how BMI plays a role in developing T2DM in HIV infected cohort. While cohort studies can estimate the effect of BMI on T2DM risks among HIV by controlling other variables, epigenetic studies can better estimate the magnitude of T2DM risks through BMI mediation among PWH. However, not many studies have been done to explore the BMI-mediated effects on T2DM within the HIV-infected cohort. To address the gap, our study will use BMI-associated methylation sites, which incorporated the latest results from the EPIC BeadChip, to estimate the epigenetic effect of BMI on T2DM development within HIV-positive cohort.

# II. Manuscript Chapter

## Introduction

The human immunodeficiency virus (HIV) can spread through blood, semen, breast milk, rectal and vaginal fluids and resulted in 38 million infected individuals worldwide in 2019 [4]. Untreated HIV infections can lead to weakened immune system with lower level of CD4 T cells and progressing to the severe stage of infection, acquired immunodeficiency syndrome (AIDS) [3, 5]. However, progression to AIDS is preventable by the anti-retrovirus therapy (ART), which suppresses the virus load in the body. ART is estimated to have saved 15.3 million lives between 2000 and 2019 [4]. Nowadays, the majority of people with HIV are documented to be taking ART (67%) [4]. While ART effectively decreases HIV-associated mortality among the infected population, HIV remains a major global public health burden due to the challenges in treating every HIV-infected patient with ART and preventing HIV transmission at the population level. Furthermore, people with HIV (PWH) also face the increasing burden and risk of non-communicable diseases including type II diabetes mellitus (T2DM), one of the most common non-communicable diseases around the world. Recent studies have found high prevalence of diabetes among HIV-infected cohort [6-8]. Duncan *et al.* and Ye *et al.* found that the risk of T2DM for PWH was two times that for the general population [8, 9]. The mortality rate due to T2DM was three times higher among PWH than the common population according to a retrospective cohort data collected in the US between 2006 and 2015 [10].

The mechanism underlying elevated risk for diabetes among PWH has been an active research area. Both HIV-related and non-HIV related factors have been identified to contribute to the risk of T2DM among PWH. Several antiviral therapies used to treat HIV were found to be

associated with T2DM among PWH, including efavirenz, zidovudine and protease-inhibitor based ARTs [13]. A potential explanation for the association between HIV and T2DM is that ART may exacerbate obesity and induce dysregulation of glucose metabolic pathways with a high-fat diet [48]. One large Danish cohort study found the risk of developing diabetes to be associated with HIV before 1999 (adjusted incidence risk ratio (aIRR): 2.83; 95%CI: 1.57–5.09). This study suggested that this increased risk was related to the use of older ARTs including indinavir, saquinavir, stavudine and didanosine [14]. Further, the risk was dependent on diagnosis of lipoatrophy (aIRR: 2.30, 95%CI: 1.39-3.80), body mass index (BMI) larger than 30 (aIRR: 9.25, 95%CI: 5.37-15.94) and age over 60 years (aIRR: 8.16, 95%CI: 1.91-34.74) [14]. The findings suggest that non-HIV related factors also play a role in the development of T2DM among PWH. Obesity or overweight defined by BMI according to the NIH [51] is the most common risk factor for T2DM identified through cross-sectional and longitudinal studies. Tripathi *et al.* found that obesity was associated with T2DM among PWH with an OR of 3.37 ($p<0.0001$) in South Carolina, United States [12]. Duncan *et al.* found weight gain following antiretroviral therapy associated with dysglycaemia (OR 1.07, 95% CI 1.04-1.11) in a cohort from London, UK [8]. A retrospective cohort study in Zimbabwe found obesity was associated with T2DM among PWH with aHR of 2.26 (95%CI: 1.17-4.36) [15]. These studies together provide support for the association and its related risk factors between T2DM and HIV. However, they do not articulate a molecular pathway explaining the elevated risk of diabetes among PWH.

Another proposed mechanism is that ART induces obesity and glucose intolerance. Pepin *et al.* found that ART together with high-fat diet can activate G-protein coupled receptor in white adipose tissue, interfering with glucose metabolism [52]. DNA methylations of the cytosine residue of a Cytosine-Guanine-dinucleotide (CpG) may explain the risk of T2DM among PWH

[50]. Epigenome-wide association study (EWAS) can elucidate inter-individual and individual-environmental variations contributing to variation in phenotypes [36]. Previous EWAS have found multiple methylation sites including those on genes *TXNIP, ACG1, PHOSPHO1, SOCS3* and *SREBF1* associated with T2DM across ethnicities [37-39]. Using a cohort with both HIV-infected and HIV-uninfected participants, Mathur *et al.* identified DNA methylation sites associated with T2DM specifically among PWH, including novel sites at *ADAMTS2, HGFAC* and *TLE3* and previously reported *TXNIP, SOCS3* and *PROC* [47]. Nevertheless, the role of BMI reported in the cohort studies has not been considered from these EWAS of T2DM. The association between BMI and DNA methylation is also well established by a wide range of EWAS [40-44]. Wahl *et al.* conducted the largest EWAS on BMI with general European and South Asian cohorts. 187 CpGs were identified to associate with BMI and 62 of them were associated with T2D using DNA methylation in peripheral blood [44]. However, the study did not provide any insight of the DNA methylation among PWH. The methylation data of all these studies were based on previous microarray platforms with limited epigenomic coverage. Further analysis involving BMI is necessary to elucidate the effect of BMI on development of T2DM among PWH using the latest Illumina EPIC BeadChip, which covers more methylation sites than all previous technologies.

To address the gaps in the pathway of BMI affecting T2DM development among PWH, our study examined the association between T2DM within PWH cohort using BMI-related DNA methylation loci. Combining summary statistics from the literature, and methylation data from 450K and EPIC BeadChips, our study aims to address: (1) whether BMI-associated DNA methylation sites are individually associated with T2DM; and (2) whether the epigenetic association with T2DM is dependent on BMI.

## Method

*VACS Dataset Overview*

This study incorporated the epigenetic and phenotypic data from the Veterans Aging Cohort Study (VACS), which is prospective observational cohort of veterans with and without HIV infection matched for age, race/ethnicity, and site enrolled from eight Veterans Affairs facilities beginning in 1997. Consented individuals were enrolled to participate and granted access to their administrative data and complete medical records with their HIV status, comorbid conditions including diabetes, Hepatitis B and Hepatitis C infection, and related treatment details including blood pressures. VACS also collected patient characteristics including BMI, ethnicity, physical activity levels, smoking and drinking status. Laboratory data were collected from medical records for virus load, total white blood cells count, $CD4^+$ T-cell counts in peripheral blood samples. The present study included 522 and 542 male veterans with HIV from VACS with the Illumina EPIC and 450K, respectively. The characteristics between two cohorts were compared using t and Chi-square tests for continuous and categorical variables, respectively. The definition of T2DM according to VACS is: (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 [50]. In addition, we defined hypertension as (1) SBP≥130 mmHg or DBP ≥ 80 mmHg, or (2) taking any anti-hypertensive drugs. The study obtained approval from the Veteran's Administration Research and Development Committee and the Institutional Review Board of the Atlanta Veteran's Administration.

*Previously Reported BMI-related sites*

16 EWAS studies were found through PubMed after searching for BMI or obesity-related methylation (Appendix A1). We selected methylation sites from the largest EWAS on BMI by

Wahl *et al.* in 2017 as the reference BMI-associated CpG sites. 187 methylation sites were identified to be associated with BMI with samples from 5,387 participants including the European and Asian Indian population. The reference study used 450K array for methylation data and the association tests were controlled by age, sex, smoking status, physical activity level, alcohol consumption, cell proportions as well as adjusted by probes [44].

*VACS DNA Methylation Data*

Methylation levels of CpG sites were recorded and derived using the HumanMethylation450 BeadChip (450K) and the EPIC BeadChip provided by Illumina, Inc. Quantile normalization was applied to the signal intensities in both methylation datasets to derive the methylation levels. Additionally, a few filters were applied to the 450K dataset: (1) CpG sites with a call rate smaller than 0.95; (2) CpG sites contain SNP within 10 bp; (3) CpG sites that could be mapped to multiple genome location. 412,583 autosomal CpG sites from 450K remained in our study pool after the filtering. 754,428 autosomal sites from EPIC remained in our study pool after applying the call rate filter as 450K and removing missing legacy and non-positive distance to SNP.

*BMI-associated CpGs*

Reported methylation sites found in either EPIC or 450K pool were selected to estimate and confirm their association with BMI within our cohort. We used linear mixed models to estimate the associations between BMI and the methylation sites from 450K and EPIC, controlling age, race, sex, smoking status, physical activity level, alcohol consumption, cell proportions (CD4T, CD8T, B cell, Monocyte, Granulocyte) and adjusted by probes. Cell proportions were calculated based on the method proposed by Houseman *et al.* to adjust various distribution of white blood cells among subpopulations including diabetic cohorts [53]. NK cell

proportions were not included in the model to avoid correlations between variables. Replicating the test standard by Wahl *et al.*, we used 3 categories for smoking status: current, former, and never smoker. Alcohol consumption was divided into 4 categories: not current, non-hazardous, hazardous and abuse. Race/ethnicity included white Caucasian, African American, Hispanic and others. Association estimates based on methylation data from two BeadChips were combined using meta-analysis with METAL [54]. The association coefficients were estimated as change in DNA methylation levels per unit change in BMI. All p-values in the meta-analysis were adjusted for multiple testing using Bonferroni correction. We examined BMI-associated DNA methylation sites from previous studies among PWH from the VACS.

*Statistical Analysis of Association Between T2DM and BMI-related CpGs*

We estimated the association between T2DM and methylation levels of BMI-associated CpG sites with generalized estimating equations using the "geepack" package in R. Two models were used to explore the relationship between T2DM and BMI-associated CpG methylation levels. To account for potential confounders of T2DM, the estimations were controlled for age, race, hazardous alcohol consumption status, current smoking status, virus load, Hepatitis B and C infection as well as cell type proportions. BMI was modeled as an additional covariate to gauge the impacts of phenotypic BMI on the epigenetic associations with T2DM. Both models were adjusted for batch effect. Subset results for both models were combined using meta-analysis and adjusted with multiple testing using Bonferroni correction.

$$(1)\ GEEGLM(T2DM) = \alpha + \beta\, CpG + \gamma_1\, age + \gamma_2\, HTN + \gamma_3\, HepC + \gamma_4\, HepB + \gamma_5\, VL + \gamma_6 ALC + \gamma_7\, SMK + \gamma_8\, AA)$$

$$(2)\ GEEGLM(T2DM) = \alpha + \beta\, CpG + \gamma_0\, BMI + \gamma_1\, age + \gamma_2\, HTN + \gamma_3\, HepC + \gamma_4\, HepB + \gamma_5\, VL + \gamma_6 ALC + \gamma_7\, SMK + \gamma_8\, AA)$$

*CpG: continuous variable indicating the DNA methylation level*

*Age: continuous variable of sample age*

*HTN: binary status of hypertension defined by SBP/DBP or antihypertensive drugs*

*HepC: binary status of Hepatitis C infection*

*HepB: binary status of Hepatitis B infection*

*VL: binary status of HIV virus load < 75*

*ALC: binary status of hazardous or abuse alcohol consumption level*

*SMK binary status of a current smoker*

*AA: binary status of African American ethnicity*

## Results

The baseline phenotypic characteristics of male veterans with HIV are summarized in Table 1 categorized by methylation BeadChip. The EPIC and 450K sub-cohorts had 529 and 555 participants, respectively. The participants had comorbidities including diabetes (EPIC: 18.2%, 450K: 18.8%), hypertension (EPIC: 76.4%, 450K: 83.3%), Hepatitis C (EPIC: 38.8%, 450K: 54.5%) and Hepatitis B (EPIC: 9.41%, 450K: 11.1%). The median BMI for EPIC was 26.2 and 25.6 for 450K. Majority of both cohorts were African American, hypertensive and had a smoking history; the median age for EPIC was 51.5, and 52.6 for 450K. Majority of both subsets were receiving antiretroviral treatment (83.4%, 83.8%). In the sub-cohort with 450K methylation data, Hepatitis C infection and hypertension were more prevalent than the EPIC cohort. Distributions of CD8T, B cell, NK and granulocyte were statistically different between two sub-cohorts.

To examine how previously reported BMI-associated CpGs replicate in our cohort, we assessed the associations between DNA methylation levels and BMI and meta-analyzed the results of two sub-cohorts with METAL. The test results of significant associations were summarized in

Table 2. Out of 187 previously reported methylation sites, 183 passed quality control procedures in either EPIC or 450K BeadChip data. After the Bonferroni correction of 183 tests, 7 of them were estimated to have directional consistency as literature findings with statistical significance. These CpGs were considered BMI-associated methylation sites and further examined with T2DM. The associations between these DNAm loci and BMI were summarized in Table A3 (Appendix). cg09554443 was only included in the EPIC sub-cohort and cg11024682 (*SREBF1*) in the 450K sub-cohort.

One out of seven BMI-related CpG sites were found to be statistically significantly associated with T2DM status (Table 3) after Bonferroni correction for multiple testing: cg11024682 (*SREBF1*) from EPIC, found on chromosome 17. The estimated odds ratio (OR) was 4.74 (95%CI: 3.99-5.49) for every 10% change in DNA methylation levels (*p*-value: $4.67\times10^{-5}$, Bonferroni-corrected threshold: $7.14\times10^{-3}$). After adjusting for phenotypic BMI in the model, the association was still significant with cg11024682 (*p*-value: $3.59\times10^{-3}$, Bonferroni-corrected threshold: $7.14\times10^{-3}$). DNAm sites at *LGALS3BP* and MAD1L1 (chromosome 17 and 7) were also associated with T2DM with ORs 1.35 (95%CI: 1.13–1.658 *p*-value: $7.87\times10^{-3}$, Bonferroni-corrected threshold: $7.14\times10^{-3}$) and 1.38 (95%CI: 1.11–1.65 *p*-value: 0.0190, Bonferroni-corrected threshold: $7.14\times10^{-3}$) respectively for every 10% change in DNA methylation levels at nominal significance level. Neither results were significant after adjusting for BMI. The association between T2DM and the methylation on *IL5RA*, not significant in BMI-independent model, was nominally significant after adjusting for BMI (OR: 1.64 for every 10% change in DNA methylation levels, 95%CI: 1.17–2.11, *p*-value: 0.0376, Bonferroni-corrected threshold: $7.14\times10^{-3}$).

| | **EPIC** **n = 529** | **450K** **n = 555** | |
|---|---|---|---|
| Variable | n (%) / Mean (SD) | n (%) / Mean (SD) | *p*-value |
| Race | | | |
| White | 48 (9.07) | 58 (10.4) | 0.00304 |
| AA | 426 (80.5) | 469 (76.3) | |
| Hispanic | 40 (7.56) | 15 (2.69) | |
| Other | 15 (2.84) | 13 (2.33) | |
| Smoking | | | |
| Never | 122 (23.6) | 112 (20.3) | 0.506 |
| Current | 288 (54.4) | 316 (56.8) | |
| Past | 119 (22.5) | 127 (22.9) | |
| Virus Load, < 75 | 156 (29.5) | 125 (22.5) | 0.103 |
| Age | 51.5 (7.75) | 52.6 (7.82) | 0.835 |
| BMI | 26.2 (5.05) | 25.6 (4.72) | 0.116 |
| eGFR | 100.3 (31.8) | 98.1 (33.7) | 0.178 |
| CD4T proportion | 0.0980 (0.0547) | 0.114 (0.0582) | 0.150 |
| CD8T proportion | 0.265 (0.0919) | 0.106 (0.0709) | <0.0001 |
| NK proportion | 5.49E-03 (0.0120) | 0.130 (0.0555) | <0.0001 |
| B cell proportion | 0.0888 (0.0384) | 0.131 (0.0476) | <0.0001 |
| Monocyte proportion | 0.147 (0.0336) | 0.103 (0.0371) | 0.0215 |
| Granulocyte proportion | 0.454 (0.101) | 0.525 (0.121) | <0.0001 |
| Alcohol | | | |
| Not Current | 105 (33.5) | 120 (35.8) | 0.734 |
| Non-hazardous | 107 (20.2) | 112 (20.1) | |
| Hazardous | 86 (16.3) | 78 (14.0) | |
| Abuse/dependence | 231 (30.1) | 243 (30.0) | |
| Physical Inactivity | 125 (23.9) | 135 (18.9) | 0.766 |
| Antiretroviral Treatment Recipient | 441 (83.4) | 465 (83.8) | 0.852 |
| Duration HIV | 2950 (1569) | 2842 (1460) | 0.0941 |
| Hypertension | 404 (76.4) | 460 (83.3) | 0.00428 |
| Diabetes | 96 (18.2) | 103 (18.8) | 0.861 |
| Hepatitis C | 187 (38.8) | 274 (54.5) | <0.0001 |
| Hepatitis B | 48 (9.41) | 59 (11.1) | 0.367 |

**Table 1**. Phenotype summary of subsets with EPIC and 450K BeadChip methylation data. Cohorts were consisted of HIV-positive male veterans from VACS. Antiretroviral treatment included ART and HAART. Distribution of race/ethnicity, hypertension, Hepatitis C coinfection rate, and cell proportions (CD8T, NK, B cell, monocyte and granulocyte) were different between 450K and EPIC subsets at a significance alpha

level *p*-value of 0.05. The definition of diabetes according to VACS is: (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 [50]. Hypertension is defined as (1) SBP≥130 mmHg or DBP ≥ 80 mmHg, or (2) taking any anti-hypertensive drugs.

| | **Number** |
|---|---|
| Previously reported BMI-associated CpG loci | 187 |
| Tested in the VACS BMI EWAS | 183 |
| Consistent Direction of association with reported BMI EWAS | 96 |
| Nominally significant ($p$<0.05) | 28 |
| Significant after Bonferroni correction | 7 |

**Table 2.** Replication of previously reported BMI-related methylation sites in the meta-analysis of VACS BMI EWAS

| | | | Meta-analysis (n=948) | | EPIC Subset (n = 465) | | 450K Subset (n = 483) | |
|---|---|---|---|---|---|---|---|---|
| CpG | Chr. | Gene | OR (95% CI) | *P* | OR (95% CI) | *P* | OR (95% CI) | *P* |
| cg11024682 | 17 | *SREBF1* | 4.74 (3.99-5.49) | 4.67E-05* | 4.74 (3.99-5.49) | 4.67E-05* | NA | NA |
| cg08857797 | 17 | *VPS25* | 1.16 (0.797-1.53) | 0.419 | 1.22 (0.601-1.84) | 0.528 | 1.13 (0.679-1.59) | 0.589 |
| cg11202345 | 17 | *LGALS3BP* | 1.35 (1.13-1.58) | 7.87E-03 | 1.19 (0.780-1.60) | 0.407 | 1.43 (1.16-1.70) | 8.48E-03 |
| cg09554443 | 1 | *CD247* | 0.862 (0.446-1.28) | 0.486 | NA | NA | 0.862 (0.446-1.28) | 0.486 |
| cg12593793 | 1 | *LMNA* | 0.787 (0.418-1.16) | 0.202 | 0.811 (0.273-1.35) | 0.446 | 0.766 (0.261-1.27) | 0.300 |
| cg23032421 | 3 | *IL5RA* | 1.21(0.743-1.68) | 0.423 | 0.647 (-0.0267-1.32) | 0.206 | 2.17 (1.52-2.83) | 0.0194 |
| cg05095590 | 7 | *MAD1L1* | 1.38 (1.11-1.65) | 0.0190 | 1.37 (0.925-1.82) | 0.166 | 1.38 (1.05-1.72) | 0.0583 |

* Bonferroni corrected *p*-value < 0.05

**Table 3**. Estimates of association between T2DM and BMI-related CpGs (BMI independent). OR, 95% CI and *P* (*p*-value) were derived from meta-analysis and from subsets (EPIC and/or 450K). NA indicates the corresponding CpG was missing in the subset. 465 and 483 samples were included in EPIC and 450K subsets respectively. The ORs (95%CI) of T2DM were adjusted for every 10% change in DNA methylation levels.

| | | | Meta-analysis (n=946) | | EPIC Subset (N = 463) | | 450K Subset (N = 483) | |
|---|---|---|---|---|---|---|---|---|
| CpG | Chr. | Gene | OR (95% CI) | P | OR (95% CI) | P | OR (95% CI) | P |
| cg11024682 | 17 | *SREBF1* | 3.37 (2.55-4.18) | 3.59E-03* | 3.37 (2.55-4.18) | 3.59E-03* | NA | NA |
| cg08857797 | 17 | *VPS25* | 0.919 (0.489-1.35) | 0.700 | 0.944 (0.287-1.60) | 0.865 | 0.900 (0.331-1.47) | 0.717 |
| cg11202345 | 17 | *LGALS3BP* | 1.19 (0.925-1.46) | 0.198 | 1.05 (0.601-1.49) | 0.841 | 1.28 (0.948-1.61) | 0.145 |
| cg09554443 | 1 | *CD247* | 1.04 (0.670-1.41) | 0.837 | NA | NA | 1.04 (0.670-1.41) | 0.837 |
| cg12593793 | 1 | *LMNA* | 0.924 (0.568-1.28) | 0.662 | 0.965 (0.432-1.50) | 0.897 | 0.891 (0.413-1.37) | 0.637 |
| cg23032421 | 3 | *IL5RA* | 1.64 (1.17-2.11) | 0.0376 | 0.895 (0.225-1.56) | 0.746 | 2.94 (2.28-3.59) | 1.26E-03 |
| cg05095590 | 7 | *MAD1L1* | 1.21 (0.956-1.46) | 0.143 | 1.16 (0.747-1.58) | 0.479 | 1.24 (0.917-1.60) | 0.192 |

* Bonferroni corrected *p*-value < 0.05

**Table 4**. Estimates of association between T2DM and BMI-related CpGs (BMI dependent). OR, 95% CI and P (*p*-value) were derived from meta-analysis and from subsets (EPIC and/or 450K). NA indicates the corresponding CpG was missing in the subset. 463 and 483 samples were included in EPIC and 450K subsets respectively. The ORs (95%CI) of T2DM were adjusted for every 10% change in DNA methylation levels.

## Discussion

We examined the roles of BMI in the development of T2DM among PWH by studying the association between BMI-associated DNA methylations in peripheral blood and risk of T2DM among male veterans with HIV in VACS. Our study confirmed that BMI-related methylation locus cg11024682 (*SREBF1*) is associated with elevated risk of T2DM among PWH. By adjusting BMI in the association between T2DM and cg11024682, we found the epigenetic association with T2DM may be independent from the effect of phenotypic BMI. In a paired correlation test, we found that cg11024682 was moderately correlated with other methylation loci in Chromosome 1, 3 ,7 and 17.

In a cohort of 946 multi-ethnic PWH, we replicated seven BMI-associated DNAm markers that were previously reported in a large cohort with 10,261 samples with European and Indian Asian ancestries from Wahl *et al*. within our VACS cohort. In addition, consistent with Karlsson *et al.*, Campanella *et al.*, and Dhana *et al.*, we identified DNAm sites at *LGALS3BP, LMNA, MAD1L1* and *SREBF1* that were also previously recognized markers associated with BMI among African populations [55-57]. Our study successfully replicated the previous reported BMI-associated methylation markers in our HIV-infected male veteran cohort.

A key finding in our study is that we identified the association between T2DM and the methylation of the Sterol Regulatory Element Binding Transcription Factor 1 (*SREBF1*) on chromosome 17 male veterans with HIV. *SREBF1* gene encodes a transcription factor binding to sterol regulatory element 1, which is part of a promoter of genes controlling sterol biosynthesis and LDL receptor genes [58]. Therefore, methylation of *SREBF1* may disrupt lipid synthesis and cell metabolism through regulation of *SREBF1* expression. *SREBF1* can be inhibited by sterol and the methylation of *SREBF1* has been reported as a predictor of future T2DM development as well

as glycemia, insulin resistance as well as obesity [59, 60]. This finding is also consistent with Wahl *et al.,* who similarly found cg11024682 *(SREBF1*) was associated with T2DM as a single marker even after adjusted for BMI [44]. DNA methylation at *SREBF1* was also previously reported to be associated with elevated T2DM risk through EWAS among Indian Asians, Europeans, and Mexican Americans [38, 39]. Our study finding suggests that the BMI-related DNAm cg11024682 may also be associated with T2DM among the HIV-infected cohort. Future studies can investigate the methylation locus among general HIV-infected population.

We did not find significant association between T2DM and other BMI-related DNAm loci. The findings are different from the referenced results of Wahl *et al*., which reported all 7 methylation markers to be associated with T2DM [44]. Except cg09554443, all markers were associated with T2DM even after adjusted for BMI [44]. Explanations for this difference may likely include an insufficient sample size.

There are several strengths to this study. Instead of using methylation data from a single subset, we incorporated methylation data from 450K and EPIC array in our association tests. Doubling measured CpG sites of 450K, the EPIC array provided us with DNA methylation data at better coverage. Previous studies have found inconsistency in a few sites between the two technologies and combined results were suggested to use for population studies [59]. Through meta-analysis of the association tests, we increased the robustness of our results and avoided potential bias caused by using subset data. Our study is also the first to explore the role of BMI-related DNA methylation markers in the association between T2DM among PWH. We provide a potential epigenetic explanation to the elevated risk of T2DM among PWH mediated by BMI.

Our study is subject to a few limitations. Using VACS cohort data with HIV-positive male veterans, our results would be limited to extending for the general HIV-positive population. Both

VACS cohorts were aged, suffered from chronic diseases and had a higher proportion of African Americans. We were unable to estimate the association of cg11024682 (*SREBF1*) in our 450K data even though the locus was previously reported using the 450K array. Quality control steps of our 450K array removed the CpG since it was within the 10 bp distance of known genome. The test standard of the reference was not fully replicated in our study. The association between DNA methylation and BMI was estimated as the change in DNA methylation per unit change in BMI, whereas the association in the reference was per unit change in methylation.

Our study provided an epigenetic explanation of how T2DM may be mediated by BMI effects specifically among PWH. Using phenotypic data from VACS and combined epigenetic information of 450K and EPIC array, we identified BMI-associated DNA methylation sites and explored their association with T2DM with and without the mediation of phenotypic BMI. The analysis replicated the previous findings of the T2DM predictor and shed light on the potential effect of BMI mediating the association among PWH. The results provide new insights on the epigenetic pathways of T2DM development among the HIV-infected cohort. Specifically, this work confirms the role of DNA methylations through BMI mediation in elevated T2DM risk among PWH. Our study suggests that further molecular or epigenetic research may focus on the *SREBF1* gene or cg11024682 DNAm locus to elucidate the mechanism of T2DM development among PWH. Future research can also be conducted and controlled for types of ART and duration of ART received within a larger cohort. Tissue-specific DNA methylations can be investigated to examine the biological pathway of BMI-mediated T2DM development. As we start to understand this mechanism, future investigations can be conducted to develop targeted therapeutics to prevent obesity and methylation on the key gene *SREBF1*.

# III. Appendices

| First author (Year) | (Discovery) Cohort Characteristics | | | | | Study Design | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | Size | Population | Age, Gender | BMI, Mean (SD) | Underlying conditions | Array, Sample | Study design | Validation | Statistical Analysis | Co-variables |
| Karlsson (2020)[55] | 535 | Swedish | 68.2yrs; 58.5% female | 26.4 (4.2) | Smoking, T2D | 450K, EPIC | Cross-sectional, longitudinal | Yes | Linear mixed effects regression | GLU, CHOL, triglycerides, smoking, T2D, age, sex, methylation array |
| He (2019)[61] | 263 | Penn State | 16.7yrs. 44.1%female. | 65.4% (28.5 %) | Tobacco, alcohol, | Illumina HiSeq 2500 | Cross sectional | Yes | Linear regression | Age, race, sex, batch of assay |
| Sun (2019)[62] | 1485 | The Bogalusa Heart Study | 44yrs; 59% female | 28.7-33.3 | Smoking | 450K | Cross-sectional, longitudinal | Yes | GLM | Age, sex, smoking status, estimated WBC |
| Li (2019)[63] | 60 | CN monozygotic twins | 53.53yrs; 50% female | 25.1 (4.33) | NA | Illumina HiSeq X Ten | Cross-sectional, longitudinal | Yes | Linear mixed effects regression | Cell type composition, GLU, CHOL, TG, HDLC, LDLC |
| Campanella (2018)[56] | 1941 | European | NA | NA | Breast cancer, colorectal cancer, MI, b-cell malignancy, smoking | 450K | Meta-analysis | Yes | meta-analysis | Microarray and position, sex, age at blood draw, case control status |
| Dhana (2018)[57] | 1450 | Dutch | 63.7yrs; 55.9%female | 27.7 (4.4) | Smoking, T2D | 450K | Cross-sectional | Yes | Linear mixed effects regression | Sex, age, smoking, leukocyte proportions, array number, position on array |
| Wahl (2017)[44] | 5387 | European, Asian Indian | >50yrs. >50 %female | >26.8 | Alcohol, hypertension, coronary heart disease, T2D, smoking | 450K | Cross-sectional, longitudinal | Yes | Inverse variance meta-analysis | Age, sex, smoking, physical activity, alcohol, probe, estimated WBC proportions |
| Sayols-Baixeras (2017)[64] | 641 | REGICOR, FOS | 63.2yrs, 50.7% female | 27.0 (4.0) | Hypertension, diabetes, smoking, | 450K | Cross-sectional | Yes | Fixed effects meta-analysis of two cohorts | Age, sex, smoking, surrogate var |
| Mendelson (2017)[43] | 3743 | FHS & LBC | 67yrs,55% female | 28.3 (5.4) | Diabetes, coronary artery disease. | 450K | Cross-sectional | Yes | Meta-analysis of two cohorts | Age, sex |

| | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| Meeks (2017)[42] | 547 | African | 50.5yrs, 57.8%female | 26.7 (0.5) | Without T2D | 450K | Cross-sectional | Yes | Linear regression | Age, sex, recruitment site, estimated cell distributions, technical effects, first principal component from genotyping data |
| Geurts (2017)[65] | 5361 | Melbourne cohort | 60yrs, 32%female | NA | Prostate, colorectal, lung or kidney cancer, urothelial carcinoma or mature B neoplasms, smoking | 450K | Cross-sectional, longitudinal | No | Linear mixed effects regression of case & control | Age, sex, smoking status, country of birth, sample type, white blood cell composition |
| Crujeiras (2017)[66] | 55 | European Caucasian | 27.4yrs; 23% female; | NA | Diabetes | 450K; adipose tissue, leukocytes | Case control | No | Wilcoxon rank test | Age |
| Ali (2016)[67] | 192 | Northern European ancestry | 36.2yrs, 55%females | NA | Insulin resistance, hypertriglyceridemia | 450K | bisulfite validation seq | Yes | SOLAR, rank-normal transformation | Sex, age, interactions |
| Demerath (2015)[68] | 2097 | African | 56.2yrs, 64%female | 30.1 (6.1) | Smoking, alcohol | 450K | Cross-sectional | Yes | Linear mixed effects regression | Leukocyte proportions, sex, age, study center, WBC, education, household income, smoking, alcohol, physical activity, array |
| Aslibekyan (2015)[41] | 991 | GOLDN | 49yrs; 52%female | 28 (6) | Smoking | 450K | Cross sectional; bisulfite sequencing | Yes | Linear mixed effects regression | Age, sex, smoking status, T cell purity, study site, family structure |
| Dick (2014)[40] | 479 | European | 43.8yrs, 78%female | 24.2 (na) | MI, smoking, diabetes, | 450K; skin, adipose tissue | Cross-sectional | Yes | Linear mixed effects regression | age, sex, smoking, methylation array batch, MI |

Table A1. Summary of literature review on EWAS of BMI with discovery set cohort characteristics and study design if available. All studies consist of mixed gender. None of the studies include participants with HIV. All samples came from peripheral blood except noted otherwise.

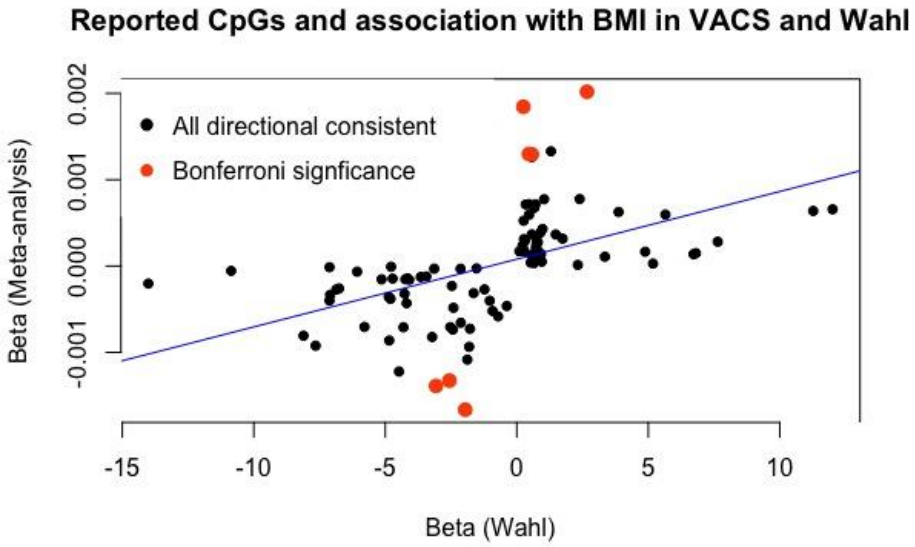


Figure A2. Comparison of beta coefficients between previous reported BMI EWAS (Wahl *et al.*) and the meta-analysis of VACS BMI EWAS. Linear regression was estimated and yielded an adjusted $R^2$ of 0.242. The BMI-related DNA methylation loci demonstrated with directional consistency at Bonferroni significance are shown in red. The association between DNA methylation and BMI (Beta) in meta-analysis was estimated as the change in DNA methylation per unit change in BMI, whereas the beta in the reference was per unit change in methylation.

| EPIC | | CHR 1 | | CHR 3 | CHR 7 | CHR 17 | |
|---|---|---|---|---|---|---|---|
| | | cg12593793 | cg23032421 | cg05095590 | cg08857797 | cg11202345 | cg11024682 |
| CHR 1 | cg12593793 | 1 | | | | | |
| | cg23032421 | -0.0854* | 1 | | | | |
| CHR 3 | cg05095590 | 0.0595 | -0.187* | 1 | | | |
| CHR 7 | cg08857797 | 0.64* | -0.431* | 0.0682 | 1 | | |
| CHR 17 | cg11202345 | 0.042 | 0.208* | 0.191* | 0.0290 | 1 | |
| | cg11024682 | 0.235* | -0.457* | 0.339* | 0.421* | -0.0738 | 1 |

* pair-wise correlation with *p*-value <0.05

Table A3. Pair-wised correlations of DNA methylation sites in EPIC sub cohort. The DNAm locus at cg11024682 has moderate correlation with all other loci except cg11202345 in EPIC array.

| 450K | | CHR 1 | | CHR 3 | CHR 7 | CHR 17 | |
|---|---|---|---|---|---|---|---|
| | | cg12593793 | cg09554443 | cg23032421 | cg05095590 | cg08857797 | cg11202345 |
| CHR 1 | cg12593793 | 1 | | | | | |
| | cg09554443 | -0.304* | 1 | | | | |
| CHR 3 | cg23032421 | -0.226* | 0.534* | 1 | | | |
| CHR 7 | cg05095590 | -0.103* | 0.052 | -0.234* | 1 | | |
| CHR 17 | cg08857797 | 0.658* | -0.526* | -0.541* | 0.0110 | 1 | |
| | cg11202345 | -0.177* | 0.174* | 0.183* | 0.0811 | -0.133* | 1 |

* pair-wise correlation with p-value <0.05

Table A. Pair-wised correlations of DNA methylation sites in 450K sub cohort.. The DNAm locus at cg11024682 has moderate correlation with all other loci except cg05095590 in 450K array.

| | | | Meta-analysis | | EPIC Subset (N =522) | | 450K Subset (N = 542) | |
|---|---|---|---|---|---|---|---|---|
| DNAm | Chr. | Gene | Beta | bonfP | Beta | bonfP | Beta | bonfP |
| cg05095590 | 7 | *MAD1L1* | 2.02E-03 | 6.08E-04 | 1.48E-03 | 1.00 | 2.47E-03 | 5.39E-03 |
| cg08857797 | 17 | *VPS25* | 1.29E-03 | 6.84E-06 | 1.60E-03 | 5.06E-03 | 1.11E-03 | 4.27E-02 |
| cg11024682 | 17 | *SREBF1* | 1.30E-03 | 6.66E-04 | 1.30E-03 | 7.67E-04 | NA | NA |
| cg11202345 | 17 | *LGALS3BP* | 1.84E-03 | 7.65E-04 | 1.39E-03 | 1.00 | 2.50E-03 | 1.27E-02 |
| cg12593793 | 1 | *LMNA* | -1.39E-03 | 1.68E-05 | -1.12E-03 | 0.827 | -1.58E-03 | 8.00E-04 |
| cg23032421 | 3 | *IL5RA* | -1.33E-03 | 1.83E-07 | -1.42E-03 | 5.83E-03 | -1.26E-03 | 1.60E-03 |
| cg09554443 | 1 | *CD247* | -1.66E-03 | 1.97E-03 | NA | NA | -1.66E-03 | 2.28E-03 |

Table A5. Summary of coefficients and p-values of EWAS on BMI. The p-values were adjusted for Bonferroni correction. The p-values for meta-analysis were adjusted for 183 tests; the *p*-values for EPIC cohort and 450K cohort were adjusted for 157 and 175 tests respectively. 522 and 542 samples were included in each subset respectively.

# IV. References


1. Roden, M. and G.I. Shulman, *The integrative biology of type 2 diabetes.* Nature, 2019. **576**(7785): p. 51-60.
2. Health, N.I.o. *Type 2 Diabetes*. Available from: https://www.nih.gov/research-training/accelerating-medicines-partnership-amp/type-2-diabetes#about.
3. Health, N.I.o. *HIV/AIDS: The Basics*. Available from: https://hivinfo.nih.gov/understanding-hiv/fact-sheets/hivaids-basics.
4. Organization, W.H. *HIV/AIDS*. [cited 2021 April]; Available from: https://www.who.int/news-room/fact-sheets/detail/hiv-aids.
5. Health, N.I.o. *HIV Treatment: The Basics*. Available from: https://hivinfo.nih.gov/understanding-hiv/fact-sheets/hiv-treatment-basics.
6. Getahun, Z., et al., *Comorbidity of HIV, hypertension, and diabetes and associated factors among people receiving antiretroviral therapy in Bahir Dar city, Ethiopia.* Journal of Comorbidity, 2020. **10**: p. 2235042X1989931.
7. Divala, O.H., et al., *The burden of hypertension, diabetes mellitus, and cardiovascular risk factors among adult Malawians in HIV care: consequences for integrated services.* BMC Public Health, 2016. **16**(1).
8. Duncan, A.D., L.M. Goff, and B.S. Peters, *Type 2 diabetes prevalence and its risk factors in HIV: A cross-sectional study.* PLOS ONE, 2018. **13**(3): p. e0194199.
9. Ye, Y., et al., *Rates and Correlates of Incident Type 2 Diabetes Mellitus Among Persons Living With HIV-1 Infection.* Frontiers in Endocrinology, 2020. **11**.
10. Park, J., J.A. Zuñiga, and A.A. García, *Diabetes negatively impacts the ten-year survival rates of people living with HIV.* Int J STD AIDS, 2019. **30**(10): p. 991-998.
11. Prioreschi, A., et al., *Incidence and prevalence of type 2 diabetes mellitus with HIV infection in Africa: a systematic review and meta-analysis.* BMJ Open, 2017. **7**(3): p. e013953.
12. Tripathi, A., et al., *Incidence of diabetes mellitus in a population-based cohort of HIV-infected and non-HIV-infected persons: the impact of clinical and therapeutic factors over time.* Diabetic Medicine, 2014. **31**(10): p. 1185-1193.
13. Karamchand, S., et al., *Risk Factors for Incident Diabetes in a Cohort Taking First-Line Nonnucleoside Reverse Transcriptase Inhibitor-Based Antiretroviral Therapy.* Medicine (Baltimore), 2016. **95**(9): p. e2844.
14. Rasmussen, L.D., et al., *Risk of Diabetes Mellitus in Persons with and without HIV: A Danish Nationwide Population-Based Cohort Study.* PLoS ONE, 2012. **7**(9): p. e44575.
15. Chimbetete, C., et al., *New-onset type 2 diabetes mellitus among patients receiving HIV care at Newlands Clinic, Harare, Zimbabwe: retrospective cohort analysis.* Tropical Medicine & International Health, 2017. **22**(7): p. 839-845.
16. Isa, S.E., et al., *Human Immunodeficiency Virus and Risk of Type 2 Diabetes in a Large Adult Cohort in Jos, Nigeria.* Clinical Infectious Diseases, 2016. **63**(6): p. 830-835.
17. Dimala, C.A., et al., *A Comparison of the Diabetes Risk Score in HIV/AIDS Patients on Highly Active Antiretroviral Therapy (HAART) and HAART-Naïve Patients at the Limbe Regional Hospital, Cameroon.* PLOS ONE, 2016. **11**(5): p. e0155560.

18. Danaei, G., et al., *National, regional, and global trends in fasting plasma glucose and diabetes prevalence since 1980: systematic analysis of health examination surveys and epidemiological studies with 370 country-years and 2·7 million participants.* The Lancet, 2011. **378**(9785): p. 31-40.
19. McPhillips, J.B., E. Barrett-Connor, and D.L. Wingard, *CARDIOVASCULAR DISEASE RISK FACTORS PRIOR TO THE DIAGNOSIS OF IMPAIRED GLUCOSE TOLERANCE AND NON-INSULIN-DEPENDENT DIABETES MELLITUS IN A COMMUNITY OF OLDER ADULTS.* American Journal of Epidemiology, 1990. **131**(3): p. 443-453.
20. Cassano, P.A., et al., *Obesity and Body Fat Distribution in Relation to the Incidence of Non-Insulin-dependent Diabetes Mellitus.* American Journal of Epidemiology, 1992. **136**(12): p. 1474-1486.
21. Chan, J.M., et al., *Obesity, Fat Distribution, and Weight Gain as Risk Factors for Clinical Diabetes in Men.* Diabetes Care, 1994. **17**(9): p. 961-969.
22. Kaye, S.A., et al., *Increased incidence of diabetes mellitus in relation to abdominal adiposity in older women.* J Clin Epidemiol, 1991. **44**(3): p. 329-34.
23. Bhargava, S.K., et al., *Relation of Serial Changes in Childhood Body-Mass Index to Impaired Glucose Tolerance in Young Adulthood.* New England Journal of Medicine, 2004. **350**(9): p. 865-875.
24. Lindström, J., et al., *Sustained reduction in the incidence of type 2 diabetes by lifestyle intervention: follow-up of the Finnish Diabetes Prevention Study.* The Lancet, 2006. **368**(9548): p. 1673-1679.
25. Sinha, R., et al., *Assessment of Skeletal Muscle Triglyceride Content by 1H Nuclear Magnetic Resonance Spectroscopy in Lean and Obese Adolescents: Relationships to Insulin Sensitivity, Total Body Fat, and Central Adiposity.* Diabetes, 2002. **51**(4): p. 1022-1027.
26. Eckel, R.H., et al., *Obesity and Type 2 Diabetes: What Can Be Unified and What Needs to Be Individualized?* The Journal of Clinical Endocrinology & Metabolism, 2011. **96**(6): p. 1654-1663.
27. Weisberg, S.P., et al., *Obesity is associated with macrophage accumulation in adipose tissue.* Journal of Clinical Investigation, 2003. **112**(12): p. 1796-1808.
28. Hotamisligil, G.S. and E. Erbay, *Nutrient sensing and inflammation in metabolic diseases.* Nature Reviews Immunology, 2008. **8**(12): p. 923-934.
29. Hu, F.B., et al., *Diet, Lifestyle, and the Risk of Type 2 Diabetes Mellitus in Women.* New England Journal of Medicine, 2001. **345**(11): p. 790-797.
30. Manson, J.E., et al., *A prospective study of cigarette smoking and the incidence of diabetes mellitus among US male physicians.* Am J Med, 2000. **109**(7): p. 538-42.
31. Cullmann, M., A. Hilding, and C.G. Östenson, *Alcohol consumption and risk of pre-diabetes and type 2 diabetes development in a Swedish population.* Diabetic Medicine, 2012. **29**(4): p. 441-452.
32. Wu, Y., et al., *Risk Factors Contributing to Type 2 Diabetes and Recent Advances in the Treatment and Prevention.* International Journal of Medical Sciences, 2014. **11**(11): p. 1185-1200.
33. Zeggini, E., et al., *Replication of Genome-Wide Association Signals in UK Samples Reveals Risk Loci for Type 2 Diabetes.* Science, 2007. **316**(5829): p. 1336-1341.

34. Consortium, T.W.T.C.C., *Genome-wide association study of 14,000 cases of seven common diseases and 3,000 shared controls.* Nature, 2007. **447**(7145): p. 661-678.
35. Vujkovic, M., et al., *Discovery of 318 new risk loci for type 2 diabetes and related vascular outcomes among 1.4 million participants in a multi-ancestry meta-analysis.* Nat Genet, 2020. **52**(7): p. 680-691.
36. Rakyan, V.K., et al., *Epigenome-wide association studies for common human diseases.* Nature Reviews Genetics, 2011. **12**(8): p. 529-541.
37. Florath, I., et al., *Type 2 diabetes and leucocyte DNA methylation: an epigenome-wide association study in over 1,500 older adults.* Diabetologia, 2016. **59**(1): p. 130-138.
38. Kulkarni, H., et al., *Novel epigenetic determinants of type 2 diabetes in Mexican-American families.* Human Molecular Genetics, 2015. **24**(18): p. 5330-5344.
39. Chambers, J.C., 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.* The Lancet Diabetes & Endocrinology, 2015. **3**(7): p. 526-534.
40. Dick, K.J., et al., *DNA methylation and body-mass index: a genome-wide analysis.* The Lancet, 2014. **383**(9933): p. 1990-1998.
41. Aslibekyan, S., et al., *Epigenome-wide study identifies novel methylation loci associated with body mass index and waist circumference.* Obesity, 2015. **23**(7): p. 1493-1501.
42. Meeks, K.A.C., et al., *An epigenome-wide association study in whole blood of measures of adiposity among Ghanaians: the RODAM study.* Clinical Epigenetics, 2017. **9**(1).
43. Mendelson, M.M., et al., *Association of Body Mass Index with DNA Methylation and Gene Expression in Blood Cells and Relations to Cardiometabolic Disease: A Mendelian Randomization Approach.* PLOS Medicine, 2017. **14**(1): p. e1002215.
44. Wahl, S., et al., *Epigenome-wide association study of body mass index, and the adverse outcomes of adiposity.* Nature, 2017. **541**(7635): p. 81-86.
45. Solomon, O., et al., *Comparison of DNA methylation measured by Illumina 450K and EPIC BeadChips in blood of newborns and 14-year-old children.* Epigenetics, 2018. **13**(6): p. 655-664.
46. Zhang, X., et al., *Epigenome-wide differential DNA methylation between HIV-infected and uninfected individuals.* Epigenetics, 2016. **11**(10): p. 750-760.
47. Nelson, K.N., et al., *Identification of HIV infection-related DNA methylation sites and advanced epigenetic aging in HIV-positive, treatment-naive U.S. veterans.* AIDS (London, England), 2017. **31**(4): p. 571-575.
48. Shu, C., et al., *Epigenome-wide association scan identifies methylation sites associated with HIV infection.* Epigenomics, 2020. **12**(21): p. 1917-1927.
49. Rickabaugh, T.M., et al., *Acceleration of age-associated methylation patterns in HIV-1-infected adults.* PLoS One, 2015. **10**(3): p. e0119201.
50. Mathur, R., et al., *DNA Methylation Markers of Type 2 Diabetes Mellitus Among Male Veterans With or Without Human Immunodeficiency Virus Infection.* J Infect Dis, 2019. **219**(12): p. 1959-1962.
51. Health, N.I.o. *Calculate Your Body Mass Index*. [cited 2021 April 29]; Available from: https://www.nhlbi.nih.gov/health/educational/lose_wt/BMI/bmicalc.htm.
52. Pepin, M.E., et al., *Antiretroviral therapy potentiates high-fat diet induced obesity and glucose intolerance.* Mol Metab, 2018. **12**: p. 48-61.

53. Houseman, E.A., et al., *DNA methylation arrays as surrogate measures of cell mixture distribution.* BMC Bioinformatics, 2012. **13**(1): p. 86.
54. Willer, C.J., Y. Li, and G.R. Abecasis, *METAL: fast and efficient meta-analysis of genomewide association scans.* Bioinformatics, 2010. **26**(17): p. 2190-2191.
55. Karlsson, I.K., et al., *Replicating associations between DNA methylation and body mass index in a longitudinal sample of older twins.* International Journal of Obesity, 2020. **44**(6): p. 1397-1405.
56. Campanella, G., et al., *Epigenome-wide association study of adiposity and future risk of obesity-related diseases.* International Journal of Obesity, 2018. **42**(12): p. 2022-2035.
57. Dhana, K., et al., *An Epigenome-Wide Association Study of Obesity-Related Traits.* American Journal of Epidemiology, 2018. **187**(8): p. 1662-1669.
58. Gene, N., *SREBF1 sterol regulatory element binding transcription factor 1 [ Homo sapiens (human) ]*.
59. Eberle, D., et al., *SREBF-1 Gene Polymorphisms Are Associated With Obesity and Type 2 Diabetes in French Obese and Diabetic Cohorts.* Diabetes, 2004. **53**(8): p. 2153-2157.
60. Grarup, N., et al., *Association of Variants in the Sterol Regulatory Element-Binding Factor 1 (SREBF1) Gene With Type 2 Diabetes, Glycemia, and Insulin Resistance: A Study of 15,734 Danish Subjects.* Diabetes, 2008. **57**(4): p. 1136-1142.
61. He, F., et al., *Association between DNA methylation in obesity-related genes and body mass index percentile in adolescents.* Scientific Reports, 2019. **9**(1).
62. Sun, D., et al., *Body Mass Index Drives Changes in DNA Methylation.* Circulation Research, 2019. **125**(9): p. 824-833.
63. Li, W., et al., *DNA methylome profiling in identical twin pairs discordant for body mass index.* International Journal of Obesity, 2019. **43**(12): p. 2491-2499.
64. Sayols-Baixeras, S., et al., *DNA methylation and obesity traits: An epigenome-wide association study. The REGICOR study.* Epigenetics, 2017. **12**(10): p. 909-916.
65. Geurts, Y.M., et al., *Novel associations between blood DNA methylation and body mass index in middle-aged and older adults.* International Journal of Obesity, 2018. **42**(4): p. 887-896.
66. Crujeiras, A.B., et al., *DNA methylation map in circulating leukocytes mirrors subcutaneous adipose tissue methylation pattern: a genome-wide analysis from non-obese and obese patients.* Scientific Reports, 2017. **7**(1): p. 41903.
67. Ali, O., et al., *Methylation of SOCS3 is inversely associated with metabolic syndrome in an epigenome-wide association study of obesity.* Epigenetics, 2016. **11**(9): p. 699-707.
68. Demerath, E.W., et al., *Epigenome-wide association study (EWAS) of BMI, BMI change and waist circumference in African American adults identifies multiple replicated loci.* Human Molecular Genetics, 2015. **24**(15): p. 4464-4479.