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:

04/25/2020

Lijin Wang Date

Epigenome-wide Association of HbA1c among the Non-diabetic Twins

By

Lijin Wang

MPH

Epidemiology

Dr. Yan Sun

Faculty Thesis Advisor

Epigenome-wide Association of HbA1c among the Non-diabetic Twins

By

Lijin Wang

MD Public Health

Sichuan University

2018

Faculty Thesis Advisor: Dr. Yan Sun, PhD, MS

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

2020

Abstract

Epigenome-wide Association of HbA1c among the Non-diabetic Twins

By Lijin Wang

Aim: Conducted an epigenome-wide association study to identify novel HbA1c-associated CpG sites in whole blood DNA.

Patients & methods: We performed an epigenome-wide association study of hyperglycemia among 258 non-diabetic male twins using linear mixed effect models. DNA Methylation at each cytosine-phosphate-guanine (CpG) sites as a function of HbA1c, adjusting for age, smoking, BMI and calculated proportion of peripheral blood leukocytes subtypes.

Results: Regression analysis showed that identified three CpG sites significantly associated with HbA1c according to within-twin effect after multiple testing correction (FDR less than 0.05), including cg00725722 (CCDC74B, p-value = 2.74×10^{-7}), cg19142411(RAD52, p-value = 9.54×10^{-8}), cg00773483 (RNF150; p-value = 2.04×10^{-7}). We also identified 25 CpG sites significantly associated (FDR<0.05) with HbA1C via between-twin effect (unshared association between co-twins).

Conclusion: This epigenome-wide study discovered significant association (FDR <0.05) between HbA1c and DNA methylation level in within-twin and between-twin effects. There were no DNA methylation sites significantly associated with HbA1C in both between-twin effect and within-twin effect after multiple testing correction, which suggested that for different HbA1C-associated DNA methylation sites, the epigenetic associations were mostly driven by either shared or unshared factors, but not both.

Epigenome-wide Association of HbA1c among the Non-diabetic Twins

By

Lijin Wang

MD Public Health

Sichuan University

2018

Faculty Thesis Advisor: Dr. Yan Sun, PhD, MS

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

2020

Acknowledgements

I would like to express my deepest appreciation to my thesis advisor, Dr. Yan Sun, for his encouragement, guidance, and patience. Dr. Sun's enthusiasm for epigenetic is utterly infectious, and he has opened my eyes to a realm of scientific study that reconciles my passions for epidemiology.

I would like to acknowledge and thank Zeyuan Wang, Yijin Xiang, Radhika Asrani, Dr. Qin and Dr. Huang for serving as my colleagues, mentors, and friends. My journey into a world described by learning, research, and teaching began in their presence.

Finally, I would also like to acknowledge and thank my mother, Dongmei Li, my father, Xiaoping Wang, my uncle, Dr. Dongxiao Li, my grandfather, Peiqing Li, and my grandmother, Youjun Fu, for their unconditional love and support.

Table of Contents

Introduction

Hyperglycemia remains a major global health problem. The global prevalence of hyperglycemia is increasing markedly, along with subsequent diabetes mellitus (Danaei et al., 2011). Hemoglobin A1c (HbA1c), which is produced by the reaction between hemoglobin and blood glucose, indicates the average plasma glucose level in an individual over the past 2 to 3 months (Nathan et al., 2008). Hyperglycemia is associated with complications such as arterial cardiovascular disease (CVD) (Selvin et al., 2010), fibrinolysis damage (Seljeflot, Larsen, Dahl-Jorgensen, Hanssen, & Arnesen, 2006), which may induce higher healthcare costs and reduce the qualities of their lives.

The role of genetics in hyperglycemia is reflected in the context of diabetes risk based on family history. Genetic susceptibility is an important risk factor of type 2 diabetes, since heritability of this disease estimates ranging from 20% to 80% (Walaszczyk et al., 2018). Although nuclear genetic factors contribute to the development of type 2 diabetes, a concordance rate for diabetes in monozygotic (MZ) twins of 24% (Poulsen, Kyvik, Vaag, & Beck-Nielsen, 1999), indicates that considerable contribution of unique environmental risk factors and/or stochastic epigenetic to type 2 diabetes. Therefore, epigenetic modifications, heritable alterations that are not due to changes in DNA sequences, may play a critical role in the pathogenesis of type 2 diabetes and hyperglycemia. Epigenome-wide association (EWAS) studies is an important method to examine a genome-wide set of quantifiable epigenetic marks, such as DNA methylation (Joshi & Shrestha, 2010; Ling & Groop, 2009). It provides platforms to identify regions of the genome-harboring DNA methylation variation associated with disease phenotypes (Joshi & Shrestha, 2010; Ling & Groop, 2009).

Several EWAS study were interested in demonstrated an association between hba1c and DNA methylation at CpG site upstream (TSS1500) of ANKRD11 and (cg19693031—TXNIP) (Meeks et al., 2019; Ronn et al., 2015), but some studies indicated no association between hba1c and DNA (Walaszczyk et al., 2018). There were more EWAS studies have explained the association between type 2 diabetes and DNA methylation at certain CpG sites, such as ABCG1, PHOSPHO1, SOCS3, SREBF1 and TXNIP (Chambers et al., 2015; Florath et al., 2016; Ling & Ronn, 2019; Rakyan, Down, Balding, & Beck, 2011). DNA methylation at cg06500161 (annotated to ABCG1) was associated with all the 2-hour glucose, and the other three CpG sites showed an association with fasting insulin only after additional adjustment for body mass index (BMI) (B. Wang et al., 2016). Additionally, disturbances in methylation at the ABCG1 and SREBF1 had been reported in blood, liver and adipose tissue from people with prevalent type 2 diabetes. These findings indicated that alterations in DNA methylation may contribute to the development of hyperglycemia (B. Wang et al., 2016). However, these studies did not disentangle the contribution of genetic factors when measuring the association between DNA methylation and hyperglycemia. Hence, we needed to design a twin study to eliminate the effect of genetic factors.

Therefore, we performed epigenome-wide association studies (EWAS) using whole blood samples aimed to investigate associations between DNA methylation and HbA1c

Methods

Study population

Our peripheral blood-based metabolomic and DNA methylation data were obtained from samples of the Emory Twin Study (ETS), which consisted of 296 middle-aged male monozygotic and dizygotic twin pairs from the Vietnam Era Twin Registry who were born between 1946 and 1956 (Forsberg et al., 2020). There were 258 twin pairs left after dropping subjects with type 2 diabetes or treatment of type 2 diabetes. All twins were examined in pairs at the Emory University General Clinical Research Center between 2002 and 2010. Zygosity information was determined by DNA analysis. All twins involved in this study were male and Caucasians. The ETS followed identical procedures, measurements, and protocols, and it was approved by the Emory Institutional Review Board, and all twins signed informed consent.

Phenotypic measurements

Twins were taken the same food the night before the assessments and required to avoid smoking. All measurements were conducted in the morning after an overnight fast, and both twin pairs were tested at the same time. All medications were stopped for about 24 h prior to testing. All assays for each twin pair were performed in the same analytical turn. A medical history and a physical examination were obtained from all twin pairs. BMI was calculated as weight in kilograms divided by height in meters squared. Cigarette smoking was classified into current smokers (any number of cigarettes) versus never or past smokers. Venous blood samples were drawn from the collection of plasma and peripheral blood leukocytes (PBL). HbA1c values were determined using high performance liquid chromatography (HPLC) (Rathmann et al., 2012).

Plasma and PBL samples were stored at -80°C before biomedical assays analysis(Huang et al., 2018; Zhao, Goldberg, Bremner, & Vaccarino, 2012).

DNA methylation measurement

Genomic DNA was isolated from PBL samples. The total amount of 0.5 μg genomic DNA was bisulfite modified using the Illumina EPIC BeadChip (850 K) in two batches, respectively (Barcelona de Mendoza, Huang, Crusto, Sun, & Taylor, 2018). DNA was firstly bisulfite converted, then whole-genome amplified, enzymatically fragmented and purified. Samples were then hybridized to each BeadChip (Huang et al., 2018). Quantile normalization of β values for each DNA methylation site was performed. Methylation sites were excluded from analyses if they overlapped with SNPs, had a missing rate greater than 10% or were not uniquely mapped to the reference genome.

Epigenome-wide association analysis

To identify the association between diabetes-related metabolites and DNA methylation, we performed statistical analyses using linear mixed models. The random effect of twin pair number and chip number was added to address the heterogeneity between twins and batch effects, respectively. The fixed effect of twin-pair average methylation level at each CpG site and individual deviation from the twin-pair average methylation level at each CpG site was included in the model. MZ twins and DZ twins are treated the same while we use the individual methylation level at each CpG site to represent the twin-pair average for singletons(Carlin, Gurrin, Sterne, Morley, & Dwyer, 2005). Age, current smoking (yes/no), BMI, and calculated proportion of PBL subtypes were included in the model as potential confounders. A reference-

based method was implemented to infer PBL subtypes including B cells, granulocytes, monocytes, natural killer cells, $CD4^+$and $CD8^+$ T cells (Houseman et al., 2012). DNA methylation sites were associated with HbA1c if they reached epigenome-wide significance overall (p-value <0.05).

For initial CpG site discovery analyses, we used a false discovery rate (FDR) of 0.05 to account for multiple testing. Manhattan plots of the epigenome-wide results were created using an FDR-adjusted threshold of 0.05 for significant sites. We also produced normalized quantile-quantile plots comparing observed to expected p-values. All statistical analyses were performed in the R statistical environment version 3.1.2. R package nlme was used to implement a linear mixed effect model.

Results

After exclusion of low-quality EWAS samples, and merge with phenotypic data, a total of 258 non-diabetic male twins were included in the present study. The mean age of participants was 56.1 years old (SD of 3.24). The characteristics of other phenotypes such as diabetes-related biomarkers, smoking status, drinking status, education and body mass index (BMI) were summarized in Table1.

We conducted a twin specific epigenome-wide association analysis of HbA1c among non-diabetic male participants. Quantile-quantile plots for between-pair effects and with-pair effects comparing the observed to the expected p-values from the EWAS showed moderate inflation (IF = 1.74) (Figure 1). The Manhattan plot for between-pair effects and within-pair effects depicted the notable significance level of the multiple selected CpG sites at the FDR-adjusted threshold of

0.05 (Figure 2). We identified three CpG sites significantly associated with HbA1c (Table 2) according to within-twin effect after multiple testing correction (FDR less than 0.05), including cg00725722 (CCDC74B, p-value = 2.74×10^{-7}), cg19142411(RAD52, p-value = 9.54×10^{-8}), cg00773483 (RNF150; p-value = 2.04×10^{-7}). We also identified 25 CpG sites significantly associated (FDR less than 0.05) with HbA1c via the between-twin effect (unshared association between co-twins). However, the direction of the association between DNA methylation level and HbA1c was not the same (Table 3).

Using our twin specific association results, we examined the significant DNA methylation sites reported in the KORA F4 study (Kriebel et al., 2016). Structural equation modeling (SEM) was used in the KORA F4 study, the effect size of our study was not directly comparable (Kriebel et al., 2016). Among a total of 711 DNA methylation sites, 31 sites were marginally significantly associated with HbA1c (Table 4) according to the between-twin effect, and 69 sites were significantly associated with HbA1C according to within-twin effect (Table 5). But none was significant after correcting for B-H adjusted tests.

We took forward several DNA methylation sites identified in relation to type 2 diabetes and DNA methylation sites from genes ABCG1, PHOSPHO1, SOCS3 and SREBF1 in the previous studies. But only one DNA methylation site was replicated, and it was not significant after FDR correlation. Direction and P-value for the type 2 diabetes-related probe was provided in Table 6.

Discussion

Though 3 CpG sites were significantly associated with HbA1c (Table 2) according to within-twin effect and 25 CpG sites were associated with HbA1c for between-twin comparison after

FDR correlation. Most of the DNA methylation sites were not reported by comparing to previously reported DNA methylation sites, which needed further investigation.

There were some interesting results from the replication. One of the important CpG methylation loci that was replicated in KORA F4 study was cg04116983, which located at the dentin matrix protein (DMP) -1-mediated gene. Some previous experimental findings that came from clinical statistical analysis showed that DMP-1 was associated with glucose-related diseases, like diabetes kidney disease and Osteocyte (Kalaitzoglou, Popescu, Bunn, Fowlkes, & Thrailkill, 2016; Valk, Bruijn, & Bajema, 2011). Though specific mechanisms of DMP-1 were unclear, DMP-1 can decrease reactive oxygen species (ROS) such as H_2O_2 and increase renal function, which may have effective protection to diabetes-related diseases (Wang, Li, & Kong, 2016). DMP-1 supplement had been used to help type 2 diabetes patients to control their blood glucose level (Du et al., 2017).

We were also interested in the difference between DNA methylation significant for HbA1c and type 2 diabetes. Previous studies show a significant association of type 2 studies with methylation variation at the ABCG1, PHOSPHO1, SOCS3, SREBF1, and TXNIP gene (Chambers et al., 2015; Dayeh et al., 2016). They play roles in key pathways underlying type 2 diabetes and related metabolic defects, which may also affect the blood level of HbA1c. ABCG1 is a key component involved in cholesterol and phospholipid transport and in β-cells. It is reported that hepatic ABCG1 impaired glucose tolerance and insulin secretion in vivo, most likely through its role in the release of HDL in plasma and its interaction with β-cell ABCG1 (Kruit et al., 2012). The suppressor of cytokine signaling 3 (SOCS3), induced by pro-inflammatory cytokines, such as TNFα and IL-6, is involved in inflammation-mediated insulin resistance in the liver and adipocytes. SOCS3 is a mediator of insulin resistance in the liver and

lack of SOCS3 in the liver induced systemic insulin resistance due to chronic inflammation (Emanuelli et al., 2000). SREBF1, a transcriptional regulator, is essential in the hepatic lipogenesis (Sekiya et al., 2007). SREBF1 occurs in obesity, insulin resistance, and type 2 diabetes by medicating dyslipidemia and hepatic steatosis (Sekiya et al., 2007). But we did not replicate DNA methylation sites from these genes, which suggested that the type 2 diabetes-related DNA methylation sites may not directly be involved in the pathological process of hyperglycemia, or such associations were not linear among the non-diabetic subjects. Thioredoxin-interacting protein (TXNIP) has now emerged as a mediator of glucotoxic β-cell death, whereas TXNIP down regulation protects against type 1 and type 2 diabetes by promoting β-cell survival and preserving β-cell mass (Minn, Hafele, & Shalev, 2005). TXNIP might also introduce regulation of adiposity and energy expenditure through hypothalamic pathways. Most recently, TXNIP has been discovered to control β-cell microRNA expression, β-cell function, and insulin production (Shalev, 2014). PHOSPHO1, a bone-specific phosphatase plays a role in the initiation of bone mineralization. Cardiovascular calcification is a common consequence of aging, diabetes, and hypercholesterolemia (Dayeh et al., 2016). PHOSPHO1 can be considered as remarkable protection against obesity and diabetes in mice (Dayeh et al., 2016). One of the DNA methylation sites from PHOSPHO1 in our study and another significant DNA methylation from TXNIP in the previous study showed that HbA1c may associate with these genes (Meeks et al., 2019).

Strengths and limitations

There were some limitations in our study. First, DNA methylation was measured using DNA extracted from blood, which may not be the ideal tissue to investigate the etiology and

mechanism of hyperglycemia. Epigenetic studies of adipose and pancreatic tissues may produce more functional links between DNA methylation and type 2 diabetes mechanism at molecular and population levels (Pidsley & Mill, 2011). Since the pathogenesis of type 2 diabetes is close to hyperglycemia, adipose and pancreatic tissues can also apply to hyperglycemia. Moreover, the association between HbA1c and DNA methylation levels may vary in different subtypes of peripheral blood leukocytes. Although we had adjusted all the cell type proportions as a potential confounder, there may still exist residual confounding, which may impact the association between HbA1c and DNA methylation level (Houseman et al., 2012). In addition, the sample size in this twin cohort was relatively small that limited our ability to identify the association between HbA1c and DNA methylation of moderate effect size. Finally, all subjects included in the present study were older males, which can be limited to represent the epigenetic association with HbA1c in females and other demographic groups. Future epigenomic studies can benefit from larger samples with a wider range of ethnicity and demography.

We also had some strengths in this study. First, this study used the Illumina Infinium Methylation EPIC (850 K) BeadChip, which provided the ability to measure DNA methylation at over 850,000 CpG dinucleotides. This technique helped us identify novel HbA1c-associated CpG sites in whole blood DNA with a wider range than usual, which may lead to better understand the epigenetic modification on hyperglycemia and other type 2 diabetes-related diseases. Also, the study design of the twin study helped us to evaluate the association between HbA1c and DNA methylation, with separating genetic effects from environmental effects.

Conclusion

Using well-matched twins from the ETS, we examined the epigenetic association with HbA1c among non-diabetic twins, controlling for age, current smoking (yes/no), BMI, and calculated proportion of PBL subtypes. We identified several novel HbA1c-associated CpG sites for further investigation. There were no DNA methylation sites significantly associated with HbA1c in both between-twin effect and within-twin effect after multiple testing correction, which suggests that for different HbA1c-associated DNA methylation sites, the epigenetic associations were mostly driven by either shared or unshared factors, but not both.

Reference

Barcelona de Mendoza, V., Huang, Y., Crusto, C. A., Sun, Y. V., & Taylor, J. Y. (2018). Perceived Racial Discrimination and DNA Methylation Among African American Women in the InterGEN Study. *Biol Res Nurs, 20*(2), 145-152. doi:10.1177/1099800417748759

Carlin, J. B., Gurrin, L. C., Sterne, J. A., Morley, R., & Dwyer, T. (2005). Regression models for twin studies: a critical review. *Int J Epidemiol, 34*(5), 1089-1099. doi:10.1093/ije/dyi153

Chambers, J. C., Loh, M., Lehne, B., Drong, A., Kriebel, J., Motta, V., . . . Kooner, J. S. (2015). 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, 3*(7), 526-534. doi:10.1016/s2213-8587(15)00127-8

Danaei, G., Finucane, M. M., Lu, Y., Singh, G. M., Cowan, M. J., Paciorek, C. J., . . . Ezzati, M. (2011). 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. *Lancet, 378*(9785), 31-40. doi:10.1016/s0140-6736(11)60679-x

Dayeh, T., Tuomi, T., Almgren, P., Perfilyev, A., Jansson, P. A., de Mello, V. D., . . . Ling, C. (2016). DNA methylation of loci within ABCG1 and PHOSPHO1 in blood DNA is associated with future type 2 diabetes risk. *Epigenetics, 11*(7), 482-488. doi:10.1080/15592294.2016.1178418

Du, N., Liu, S., Cui, C., Zhang, M., Jia, J., & Cao, X. (2017). DMP-1 attenuates oxidative stress and inhibits TGF-beta activation in rats with diabetic kidney disease. *Ren Fail, 39*(1), 229-235. doi:10.1080/0886022x.2016.1256319

Emanuelli, B., Peraldi, P., Filloux, C., Sawka-Verhelle, D., Hilton, D., & Van Obberghen, E. (2000). SOCS-3 is an insulin-induced negative regulator of insulin signaling. *J Biol Chem, 275*(21), 15985-15991. doi:10.1074/jbc.275.21.15985

Florath, I., Butterbach, K., Heiss, J., Bewerunge-Hudler, M., Zhang, Y., Schottker, B., & Brenner, H. (2016). Type 2 diabetes and leucocyte DNA methylation: an epigenome-wide association study in over 1,500 older adults. *Diabetologia, 59*(1), 130-138. doi:10.1007/s00125-015-3773-7

Forsberg, C., Liu, C., Mori, A., Tsai, M., Sporleder, J., Moore, K., . . . Smith, N. (2020). Cohort Profile: The Vietnam Era Twin Registry (VET Registry). *Int J Epidemiol, 49*(1), 22-22d. doi:10.1093/ije/dyz217

Houseman, E. A., Accomando, W. P., Koestler, D. C., Christensen, B. C., Marsit, C. J., Nelson, H. H., . . . Kelsey, K. T. (2012). DNA methylation arrays as surrogate measures of cell mixture distribution. *BMC Bioinformatics, 13*(1), 86. doi:10.1186/1471-2105-13-86

Huang, Y., Hui, Q., Walker, D. I., Uppal, K., Goldberg, J., Jones, D. P., . . . Sun, Y. V. (2018). Untargeted metabolomics reveals multiple metabolites influencing smoking-related DNA methylation. *Epigenomics, 10*(4), 379-393. doi:10.2217/epi-2017-0101

Joshi, S. K., & Shrestha, S. (2010). Diabetes mellitus: a review of its associations with different environmental factors. *Kathmandu Univ Med J (KUMJ), 8*(29), 109-115. doi:10.3126/kumj.v8i1.3233

Kalaitzoglou, E., Popescu, I., Bunn, R. C., Fowlkes, J. L., & Thrailkill, K. M. (2016). Effects of Type 1 Diabetes on Osteoblasts, Osteocytes, and Osteoclasts. *Curr Osteoporos Rep, 14*(6), 310-319. doi:10.1007/s11914-016-0329-9

Kriebel, J., Herder, C., Rathmann, W., Wahl, S., Kunze, S., Molnos, S., . . . Grallert, H. (2016). Association between DNA Methylation in Whole Blood and Measures of Glucose Metabolism: KORA F4 Study. *PLoS One, 11*(3), e0152314. doi:10.1371/journal.pone.0152314

Kruit, J. K., Wijesekara, N., Westwell-Roper, C., Vanmierlo, T., de Haan, W., Bhattacharjee, A., . . . Hayden, M. R. (2012). Loss of both ABCA1 and ABCG1 results in increased disturbances in islet sterol homeostasis, inflammation, and impaired beta-cell function. *Diabetes, 61*(3), 659-664. doi:10.2337/db11-1341

Ling, C., & Groop, L. (2009). Epigenetics: a molecular link between environmental factors and type 2 diabetes. *Diabetes, 58*(12), 2718-2725. doi:10.2337/db09-1003

Ling, C., & Ronn, T. (2019). Epigenetics in Human Obesity and Type 2 Diabetes. *Cell Metab, 29*(5), 1028-1044. doi:10.1016/j.cmet.2019.03.009

Meeks, K. A. C., Henneman, P., Venema, A., Addo, J., Bahendeka, S., Burr, T., . . . Agyemang, C. (2019). Epigenome-wide association study in whole blood on type 2 diabetes among sub-Saharan African individuals: findings from the RODAM study. *Int J Epidemiol, 48*(1), 58-70. doi:10.1093/ije/dyy171

Minn, A., Hafele, C., & Shalev, A. (2005). Thioredoxin-Interacting Protein Is Stimulated by Glucose through a Carbohydrate Response Element and Induces β-Cell Apoptosis. *Endocrinology, 146*, 2397-2405. doi:10.1210/en.2004-1378

Nathan, D. M., Kuenen, J., Borg, R., Zheng, H., Schoenfeld, D., & Heine, R. J. (2008). Translating the A1C assay into estimated average glucose values. *Diabetes Care, 31*(8), 1473-1478. doi:10.2337/dc08-0545

Pidsley, R., & Mill, J. (2011). Epigenetic studies of psychosis: current findings, methodological approaches, and implications for postmortem research. *Biol Psychiatry, 69*(2), 146-156. doi:10.1016/j.biopsych.2010.03.029

Poulsen, P., Kyvik, K. O., Vaag, A., & Beck-Nielsen, H. (1999). Heritability of type II (non-insulin-dependent) diabetes mellitus and abnormal glucose tolerance--a population-based twin study. *Diabetologia, 42*(2), 139-145. doi:10.1007/s001250051131

Rakyan, V. K., Down, T. A., Balding, D. J., & Beck, S. (2011). Epigenome-wide association studies for common human diseases. *Nat Rev Genet, 12*(8), 529-541. doi:10.1038/nrg3000

Rathmann, W., Kowall, B., Tamayo, T., Giani, G., Holle, R., Thorand, B., . . . Meisinger, C. (2012). Hemoglobin A1c and glucose criteria identify different subjects as having type 2 diabetes in middle-aged and older populations: the KORA S4/F4 Study. *Ann Med, 44*(2), 170-177. doi:10.3109/07853890.2010.531759

Ronn, T., Volkov, P., Gillberg, L., Kokosar, M., Perfilyev, A., Jacobsen, A. L., . . . Ling, C. (2015). Impact of age, BMI and HbA1c levels on the genome-wide DNA methylation and mRNA expression patterns in human adipose tissue and identification of epigenetic biomarkers in blood. *Hum Mol Genet, 24*(13), 3792-3813. doi:10.1093/hmg/ddv124

Sekiya, M., Yahagi, N., Matsuzaka, T., Takeuchi, Y., Nakagawa, Y., Takahashi, H., . . . Shimano, H. (2007). SREBP-1-independent regulation of lipogenic gene expression in adipocytes. *J Lipid Res, 48*(7), 1581-1591. doi:10.1194/jlr.M700033-JLR200

Seljeflot, I., Larsen, J. R., Dahl-Jorgensen, K., Hanssen, K. F., & Arnesen, H. (2006). Fibrinolytic activity is highly influenced by long-term glycemic control in Type 1 diabetic patients. *J Thromb Haemost, 4*(3), 686-688. doi:10.1111/j.1538-7836.2006.01803.x

Selvin, E., Steffes, M. W., Zhu, H., Matsushita, K., Wagenknecht, L., Pankow, J., . . . Brancati, F. L. (2010). Glycated hemoglobin, diabetes, and cardiovascular risk in nondiabetic adults. *N Engl J Med, 362*(9), 800-811. doi:10.1056/NEJMoa0908359

Shalev, A. (2014). Minireview: Thioredoxin-interacting protein: regulation and function in the pancreatic beta-cell. *Mol Endocrinol, 28*(8), 1211-1220. doi:10.1210/me.2014-1095

Valk, E. J., Bruijn, J. A., & Bajema, I. M. (2011). Diabetic nephropathy in humans: pathologic diversity. *Curr Opin Nephrol Hypertens, 20*(3), 285-289. doi:10.1097/MNH.0b013e328345bc1c

Walaszczyk, E., Luijten, M., Spijkerman, A. M. W., Bonder, M. J., Lutgers, H. L., Snieder, H., . . . van Vliet-Ostaptchouk, J. V. (2018). DNA methylation markers associated with type 2 diabetes, fasting glucose and HbA1c levels: a systematic review and replication in a case-control sample of the Lifelines study. *Diabetologia, 61*(2), 354-368. doi:10.1007/s00125-017-4497-7

Wang, B., Gao, W., Li, J., Yu, C., Cao, W., Lv, J., . . . Li, L. (2016). Methylation loci associated with body mass index, waist circumference, and waist-to-hip ratio in Chinese adults: an epigenome-wide analysis. *The Lancet, 388*, S21. doi:10.1016/S0140-6736(16)31948-1

Wang, X., Li, W., & Kong, D. (2016). Cyclocarya paliurus extract alleviates diabetic nephropathy by inhibiting oxidative stress and aldose reductase. *Ren Fail, 38*(5), 678-685. doi:10.3109/0886022x.2016.1155394

Zhao, J., Goldberg, J., Bremner, J. D., & Vaccarino, V. (2012). Global DNA methylation is associated with insulin resistance: a monozygotic twin study. *Diabetes, 61*(2), 542-546. doi:10.2337/db11-1048

Tables

Table 1. Descriptive Statistics of Participants, Stratified by Types of Twins, Emory Twin Study, 2002-2010

	Singleton(n=28)	MZ twins(n=153)	DZ twins(n=77)
Age (years)			
Mean (SD)	54.2 (2.76)	56.4 (3.26)	56.0 (3.16)
HbA1c (%)			
Mean (SD)	5.27 (0.51)	5.59 (0.51)	5.55 (0.45)
Insulin (mIU/L)			
Mean (SD)	7.18 (4.14)	8.47 (6.69)	8.50 (7.12)
Blood Glucose Levels (mg/dL)			
Mean (SD)	94.0 (18.4)	99.6 (12.5)	102 (14.2)
Hypertension			
Yes	12 (42.9%)	52 (34.0%)	19 (24.7%)
BMI (kg/m^2)			
Mean (SD)	28.9 (4.76)	29.4 (4.45)	30.1 (6.05)
Current Smoking Status			
Yes	10 (35.7%)	39 (25.7%)	23 (29.9%)
Number of alcoholic beverages in a typical week			
Mean (SD)	6.93 (12.7)	5.24 (11.8)	6.71 (10.1)

*MZ: Monozygotic Twins DZ: Dizygotic Twins

Table 2. Result of within-twin HbA1c -DNA methylation association in the Emory Twin study. The betas represent the slope of the regression model indicating the rate of change in the HbA1c as independent variable (deviation value of individual methylation beta from twin-pair average methylation beta) changes.

Probe	Chr	Position	Gene symbol	Beta	SE	p value	FDR
cg00725722	2	130902131	CCDC74B	-10.145	1.728	2.74E-07	0.0386
cg19142411	12	1059047	RAD52	51.793	8.412	9.54E-08	0.0386
cg00773483	4	141805605	RNF150	3.071	0.515	2.04E-07	0.0386

Table 3. Result of between-twin HbA1c -DNA methylation association in the Emory Twin study. The betas represent the slope of the regression model indicating the rate of change in the HbA1c as independent variable (deviation value of individual methylation beta from twin-pair average methylation beta) changes.

Probe	Chr	Position	Gene symbol	Beta	SE	p value	FDR
cg24808105	5	148339194	NA	-5.133	0.741	7.51E-11	6.35E-05
cg24575376	7	17707138	NA	-3.187	0.492	8.67E-10	0.0002
cg04712207	10	64893676	NRBF2	34.611	5.470	1.92E-09	0.0003
cg16478852	2	166789623	TTC21B;TTC21B-AS1	2.866	0.499	4.04E-08	0.0043
cg05080976	11	3010620	NAP1L4	-4.874	0.910	2.56E-07	0.0114
cg03445143	17	30810170	PSMD11	-4.492	0.851	3.69E-07	0.0133
cg15603959	3	52719735	PBRM1;GNL3	14.836	2.824	4.18E-07	0.0142
cg22960635	16	1543151	TELO2	11.072	2.112	4.41E-07	0.0143
cg06069280	10	111815726	ADD3	2.712	0.518	4.65E-07	0.0146
cg05906620	10	69523794	NA	-49.333	9.639	7.87E-07	0.0217
cg02527472	11	27743348	BDNF	8.693	1.699	7.93E-07	0.0217
cg00419061	14	25135098	NA	-17.942	3.529	9.22E-07	0.0237
cg23259973	19	14196803	C19orf67	-4.187	0.833	1.19E-06	0.0262
cg15715045	13	49139035	NA	8.568	1.705	1.20E-06	0.0262
cg00934312	7	149570931	ATP6V0E2;ATP6V0E2-AS1	17.823	3.547	1.21E-06	0.0262
cg00562976	12	95927545	USP44	4.182	0.841	1.52E-06	0.0293
cg20604447	14	32492063	NA	-3.522	0.713	1.76E-06	0.0310
cg23519397	4	17629225	NA	-8.998	1.822	1.80E-06	0.0310
cg13598219	3	141899907	GK5	6.967	1.429	2.39E-06	0.0372
cg13799227	1	228652581	NA	15.673	3.216	2.39E-06	0.0372
cg26125384	2	70314274	PCBP1;PCBP1-AS1	6.100	1.252	2.42E-06	0.0372
cg01698752	8	91042601	DECR1	7.345	1.518	2.80E-06	0.0408
cg02592615	6	12012261	HIVEP1	9.962	2.060	2.83E-06	0.0408
cg06473182	5	172571413	BNIP1	18.831	3.898	2.89E-06	0.0408
cg03085150	8	676080	ERICH1	-3.121	0.650	3.31E-06	0.0430

Table 4. Top CpG methylation sites associated with HbA1c for the within-twin effect in ETS (n=258) and KORA (n=1,814) at the level of significance ($P < 0.05$) in the discovery phase.

Probe	Gene	Emory Cohort			KORA F4 Study		
		Beta	SE	p value	Beta	SE	p value
cg07807219	NA	5.440	1.222	4.31E-05	-0.029	0.007	2.70E-05
cg18075930	NA	2.409	0.551	0.0001	-0.031	0.007	2.40E-05
cg27313403	POLDIP3;RNU12	-24.710	6.620	0.0005	-0.028	0.007	4.70E-05
cg08573180	TOP2B	-11.843	3.442	0.0011	-0.026	0.006	2.60E-05
cg22043667	PECI;C6orf201	2.115	0.624	0.0013	-0.038	0.008	3.20E-06
cg20724862	NA	3.115	0.933	0.0015	-0.049	0.011	2.40E-05
cg01000408	LTBP4	-8.920	2.674	0.0015	-0.024	0.005	1.70E-05
cg03376794	HS3ST4	-11.422	3.434	0.0016	-0.034	0.007	5.20E-06
cg08146372	NA	-12.756	3.852	0.0017	-0.029	0.005	8.00E-08
cg19378376	TMEM120B	-8.561	2.607	0.0018	-0.019	0.004	2.60E-05
cg25054890	NA	3.383	1.055	0.0023	-0.032	0.006	3.20E-06
cg04116983	DMP1	3.112	0.990	0.0027	-0.024	0.006	5.00E-05
cg17820591	ENO3	-21.436	7.101	0.0039	-0.044	0.010	3.30E-05
cg02990672	ENHO	-8.626	2.895	0.0043	-0.028	0.006	1.80E-05
cg09725013	NCKAP5L;BCDIN3D-AS1	-15.158	5.104	0.0044	-0.029	0.007	6.00E-05
cg15743657	TAS1R2	3.256	1.105	0.0047	-0.024	0.006	4.70E-05
cg00797837	NA	-13.938	4.789	0.0052	-0.035	0.006	1.50E-07
cg07588128	JUB	-4.635	1.615	0.0058	-0.022	0.005	6.30E-05
cg14869845	C5orf13	-22.997	8.131	0.0066	-0.024	0.006	6.60E-05
cg15652577	RASSF3	-11.829	4.205	0.0068	-0.025	0.006	5.80E-05
cg00219956	ZNF629	3.618	1.299	0.0074	-0.014	0.003	5.60E-05
cg26246744	NA	7.132	2.566	0.0075	0.032	0.007	3.70E-05
cg16884295	NA	2.515	0.926	0.0088	-0.027	0.006	5.60E-05
cg11987455	ERMAP	2.840	1.053	0.0093	-0.037	0.009	5.00E-05
cg05861030	NA	2.654	0.989	0.0096	-0.042	0.009	1.80E-05
cg03749777	NUAK1	-5.356	2.008	0.0101	-0.023	0.004	1.20E-06
cg18065874	NA	-2.883	1.084	0.0103	-0.040	0.010	7.10E-05
cg20081364	ISM1	-5.452	2.057	0.0105	-0.025	0.006	5.30E-05
cg23141096	TM9SF3	-8.078	3.052	0.0106	-0.030	0.007	2.60E-05
cg07955105	NA	3.764	1.437	0.0114	-0.063	0.014	2.60E-05
cg04534926	CBLN1	-8.634	3.362	0.013	-0.035	0.007	2.50E-06
cg05160563	USH2A	-2.131	0.837	0.0137	-0.044	0.009	4.90E-06
cg10691848	EIF4G1	-8.859	3.543	0.0155	-0.038	0.007	1.40E-06
cg26009803	LSM14A	-6.244	2.505	0.0158	-0.037	0.008	3.20E-05
cg23913350	LOX	-4.039	1.634	0.0166	-0.024	0.005	6.40E-06
cg23309825	MYL6B	-5.099	2.064	0.0167	-0.033	0.008	2.60E-05
cg02556954	ETF1	-3.848	1.561	0.0169	-0.026	0.006	4.80E-05
cg07781995	NA	2.412	0.991	0.0183	-0.038	0.008	9.80E-06
cg09941176	LOC100286844;NCKAP5L	-10.461	4.320	0.0188	-0.031	0.007	1.50E-05
cg00751937	NDUFA12	7.354	3.090	0.0209	-0.027	0.006	4.10E-05
cg23036405	NA	-2.659	1.123	0.0215	-0.051	0.012	7.40E-05

cg00469341	ETV1	-8.907	3.776	0.022	-0.029	0.007	6.80E-05
cg10857427	ERP27	1.541	0.656	0.0225	-0.035	0.007	9.00E-06
cg20765522	PBX2	-1.331	0.577	0.025	-0.023	0.005	1.70E-05
cg24306277	TUBB2A	-11.923	5.232	0.0266	-0.035	0.007	2.10E-06
cg07907534	SIN3A	-7.126	3.146	0.0275	-0.021	0.005	6.10E-05
cg14751503	EIF3F	-6.179	2.729	0.0276	-0.032	0.007	4.00E-05
cg09331206	YPEL3	2.286	1.012	0.028	-0.027	0.006	1.30E-05
cg15556863	TMEM62	1.620	0.718	0.0282	-0.048	0.011	5.20E-05
cg17695831	NA	2.124	0.955	0.0303	-0.056	0.012	1.40E-05
cg19100169	CCNJL	2.623	1.182	0.0307	-0.046	0.011	3.20E-05
cg08349573	C1QTNF1	10.373	4.676	0.0307	0.024	0.005	1.50E-05
cg23527468	NA	-2.260	1.020	0.031	-0.031	0.007	2.70E-05
cg25108325	UBE2U	1.185	0.544	0.0338	-0.072	0.015	9.90E-06
cg02381853	NA	-1.961	0.912	0.036	-0.038	0.008	1.90E-05
cg06656260	C12orf62	-8.717	4.053	0.036	0.033	0.007	2.50E-05
cg20097440	CYP26B1	3.460	1.624	0.0376	-0.031	0.007	3.70E-05
cg16282910	NA	1.469	0.691	0.0381	-0.033	0.008	2.70E-05
cg12914530	ANKRD11	1.732	0.817	0.0385	-0.044	0.010	2.40E-05
cg04240950	GTF3C1	2.675	1.268	0.0395	-0.028	0.006	2.30E-05
cg19702802	NANOS2	-1.276	0.618	0.0439	-0.057	0.012	1.40E-05
cg11429664	BLM	-7.901	3.840	0.0445	-0.026	0.005	6.80E-06
cg27331851	SOX2OT;SOX2	-6.490	3.193	0.047	-0.029	0.007	4.00E-05
cg27452691	SFTPA1	2.294	1.130	0.0472	-0.025	0.005	2.10E-05
cg25866059	UBE2G2	-5.669	2.794	0.0474	-0.025	0.006	2.90E-05
cg09344986	NA	-6.114	3.021	0.0479	0.025	0.006	7.20E-05
cg19102271	C9orf40	-11.470	5.682	0.0485	-0.026	0.005	7.30E-06
cg12115882	NA	2.261	1.123	0.0491	-0.031	0.008	7.00E-05
cg07709210	PGC	1.381	0.688	0.0496	-0.034	0.008	3.20E-05

Table 5. Top CpG methylation sites associated with HbA1c for the between-twin effect in ETS (n=258) and KORA (n=1,814) at the level of significance (P < 0.05) in the discovery phase.

Probe	Gene	Emory Cohort			KORA F4 Study		
		Beta	SE	p value	Beta	SE	p value
cg12914530	ANKRD11	-3.479	1.062	0.0013	-0.044	0.010	2.40E-05
cg12891498	HRK	-9.362	3.147	0.0033	-0.024	0.005	3.40E-05
cg14918391	CAPN5	24.428	8.310	0.0037	-0.035	0.008	1.30E-05
cg14124818	NA	-3.685	1.332	0.0062	-0.038	0.007	1.70E-06
cg23527468	NA	-3.083	1.142	0.0076	-0.031	0.007	2.70E-05
cg11255394	HOXC13	-3.614	1.403	0.0108	0.028	0.007	5.70E-05
cg16993435	CENPH	-12.614	4.915	0.0111	-0.027	0.006	3.40E-05
cg01689892	BET1L;RIC8A	-14.134	5.621	0.0128	-0.030	0.007	6.50E-05
cg22678977	LRRC6	2.033	0.812	0.0132	0.053	0.012	1.70E-05
cg09941176	LOC100286844;NCKAP5L	-10.253	4.158	0.0146	-0.031	0.007	1.50E-05
cg08084454	TBL1XR1	12.439	5.179	0.0173	-0.030	0.007	3.40E-05
cg11387705	NA	5.761	2.416	0.0181	-0.025	0.006	5.80E-05
cg06894585	DAB2	1.738	0.742	0.0203	0.046	0.010	2.30E-05
cg05200715	KCNC2	1.385	0.613	0.0251	-0.041	0.010	4.60E-05
cg13142700	SLC6A1	5.056	2.284	0.0281	-0.032	0.008	5.30E-05
cg07907534	SIN3A	-6.438	2.946	0.0301	-0.021	0.005	6.10E-05
cg19973121	PFKL	1.785	0.822	0.0312	-0.047	0.011	7.80E-05
cg05366401	TM9SF1	-21.424	9.869	0.0313	-0.019	0.005	6.50E-05
cg18102738	SMTNL2	1.136	0.525	0.0317	0.058	0.013	2.50E-05
cg11030575	AGK	-9.413	4.352	0.0319	-0.027	0.006	3.70E-05
cg09304395	NA	1.993	0.929	0.0333	-0.021	0.005	6.60E-05
cg02558362	MIR548F5;NBEA	-4.291	2.005	0.0337	-0.022	0.005	5.20E-05
cg25608370	SCD5	-4.924	2.330	0.0359	-0.022	0.005	8.20E-06
cg05532669	SRI	-11.331	5.388	0.0368	-0.021	0.005	1.70E-05
cg07690290	LLPH	15.167	7.259	0.0381	-0.030	0.006	5.80E-06
cg26182406	NA	-2.346	1.130	0.0394	-0.035	0.007	1.20E-05
cg06938356	COL23A1	2.628	1.266	0.0394	-0.030	0.007	3.10E-05
cg09183102	MT1F	7.695	3.797	0.0442	-0.028	0.007	7.60E-05
cg02401978	ZNF746	1.617	0.804	0.0457	-0.037	0.009	6.90E-05
cg03017480	RUNX1	-3.185	1.598	0.0477	-0.027	0.006	4.10E-05
cg22177707	C4orf33;SCLT1	-10.395	5.265	0.0498	-0.025	0.006	2.60E-05

Table 6. Result of replicated methylation sites associated with type 2 diabetes for within-twin effect in ETS (n=258).

Probe	Chr	Position	Gene symbol	Beta	SE	p value	FDR
cg02650017	17	47301614	PHOSPHO1	-9.381	2.495	0.0004	0.1192

Figures and Figure Legends

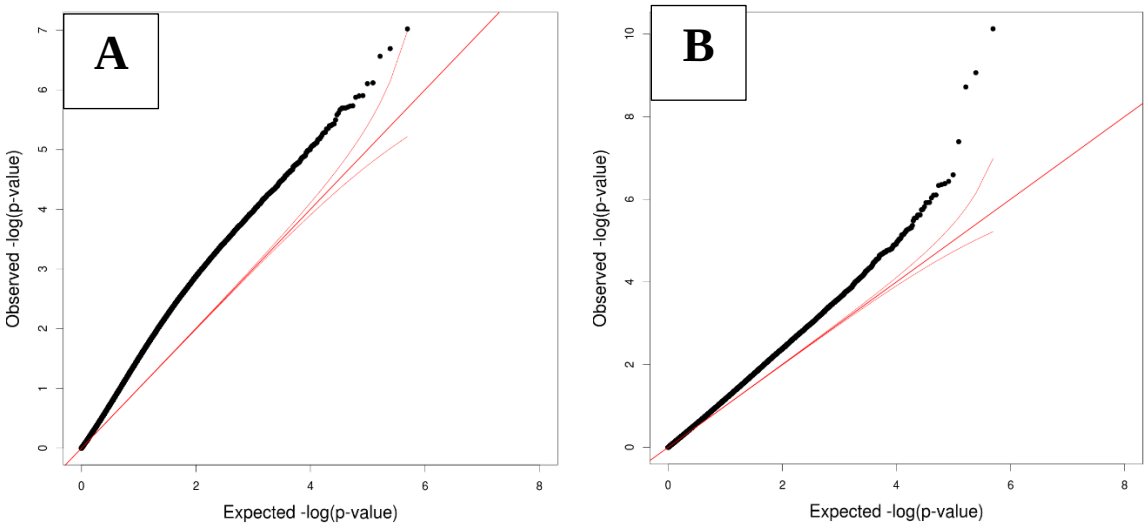
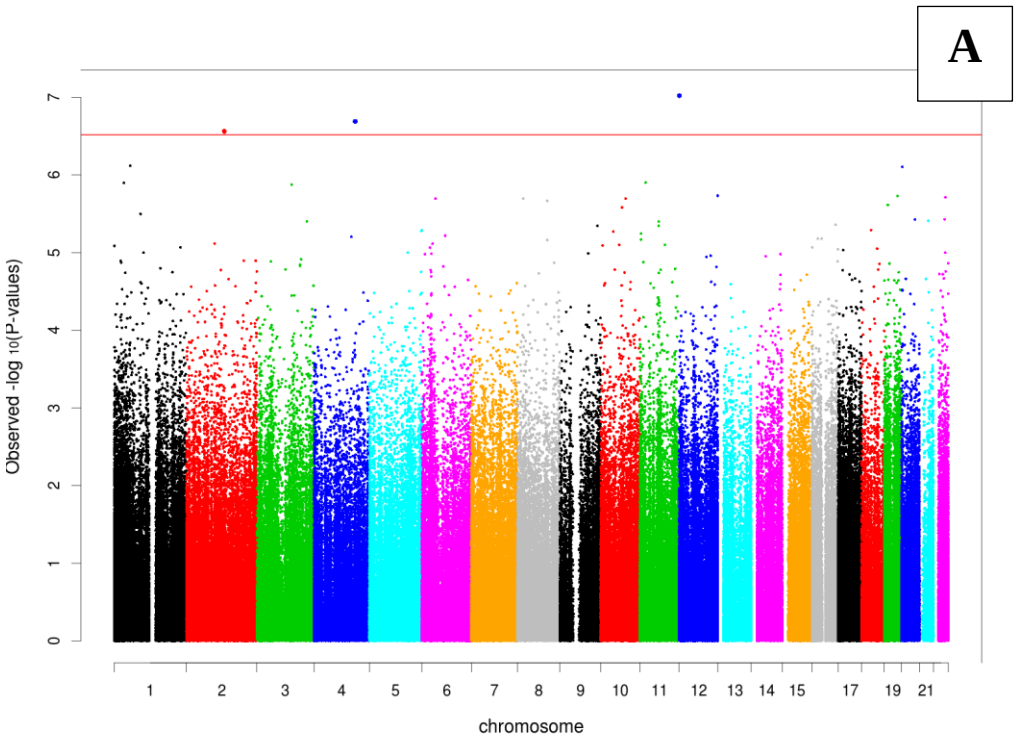


Figure 1. Quantile–quantile (QQ) plots of multivariate EWAS of HbA1c among the non-diabetic twins. **A,** QQ plot for within-twin associations, IF = 1.74; **B,** QQ plot for between-twin associations, IF = 1.74.



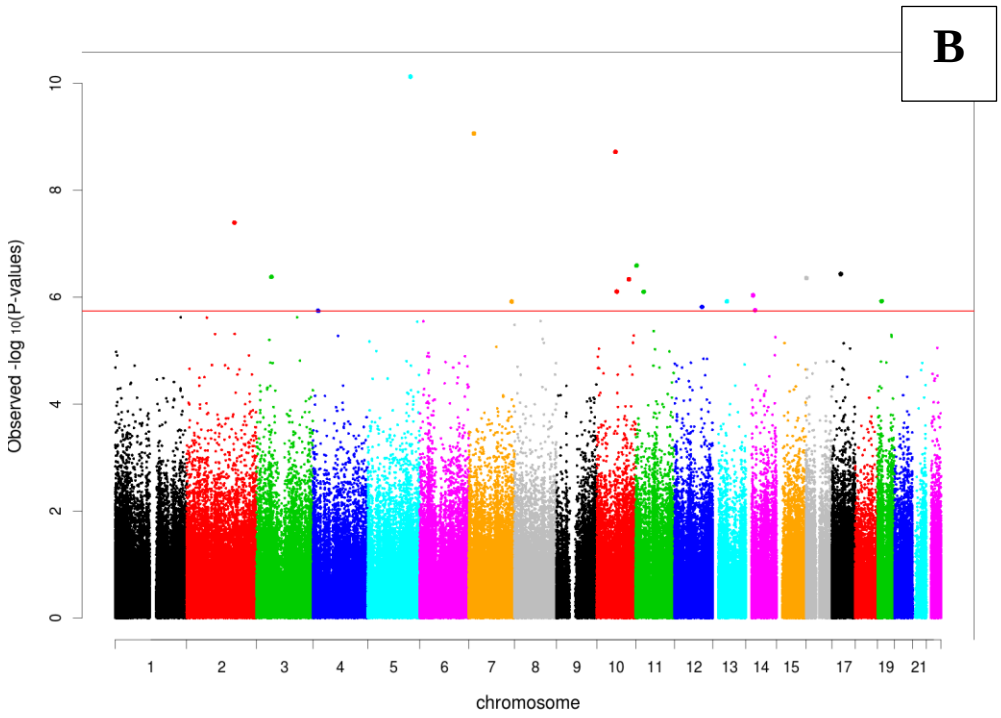


Figure 2. Manhattan plots for EWAS of HbA1c among the non-diabetic twins. Red horizontal line: Multiple testing correction threshold (FDR < 0.05). **A,** Manhattan plot for within-twin associations; **B,** Manhattan plot for between-twin associations.