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:

_____ 04/22/2019

Zhenyu Gan Date

Metabolomic Profiling of Carotid Intima-Media Thickness in Middle-Aged Males: A Twin Study

By

Zhenyu Gan

Master of Science in Public Health

Epidemiology

_____

Yan V. Sun

Committee Chair

Metabolomic Profiling of Carotid Intima-Media Thickness in Middle-Aged Males: A Twin Study

By

Zhenyu Gan

Bachelor of Engineering

Beijing University of Chemical Technology

2017

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 Science in Public Health

in Epidemiology

2019

Abstract

Metabolomic Profiling of Carotid Intima-Media Thickness in Middle-Aged Males:

A Twin Study

By Zhenyu Gan

Background: Atherosclerotic cardiovascular disease (ASCVD) is a leading cause of death worldwide. Carotid intima-media thickness (IMT) is a reliable measurement of subclinical atherosclerosis and predictor of future cardiovascular events. High throughput metabolomic studies possess increased ability to detect changes of metabolic levels and exploring novel associations. This study aims to investigate the association between metabolites in plasma and carotid IMT among middle-aged male twins.

Methods: We performed an untargeted metabolome-wide association study (MWAS) including 12,527 negative ionized features and 7,508 positive ionized features using data from 92 twin pairs and 3 singletons in the Emory Twin Study. Metabolomics data were generated using fasting blood samples, and the relative abundance of metabolites was evaluated using high-performance liquid chromatography–mass spectrometry (HPLC-MS). Carotid IMT was measured using high-resolution B-mode ultrasonography. Other cardiovascular factors were measured at the same time. Mixed effect linear regression models were used to examine within-pair and between-pair effects to explain the unshared environment effects and general association. Metabolic pathway analysis was performed using Mummichog, and metabolic features were annotated via the HMDB database.

Results: According to the MWAS analyze results, among 92 twin pairs and 3 singletons, we did not identify any metabolic feature statistically significantly associated with carotid IMT. Pathway analysis further suggests 17 significant metabolism pathways, where glutathione metabolism is associated with carotid IMT regardless of the twin effects. 2 lipid metabolism pathways, including saturated fatty acids metabolism and arachidonic acid metabolism, were significantly associated between co-twins. The unshared environment effect may contribute to the difference in lipid metabolism level, which further associated with carotid IMT and related cardiovascular events.

Conclusion: No metabolite alone was statistically associated with the changes of carotid IMT in this twin study. Pathway analysis suggests the glutathione metabolism pathway is associated with carotid IMT and potential ASCVD events. The differences in unshared environmental factors between the co-twins may explain the association between lipid metabolism pathways and carotid IMT.

Metabolomic Profiling of Carotid Intima-Media Thickness in Middle-Aged Males: A Twin Study

By

Zhenyu Gan

Bachelor of Engineering

Beijing University of Chemical Technology

2017

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 Science in Public Health

in Epidemiology

2019

Table of Contents

Background

Cardiovascular diseases (CVD) are the leading cause of death worldwide. Atherosclerotic Cardiovascular diseases (ASCVD) are functional disorders of heart and blood vessels and commonly caused by accumulation of fatty deposits on blood vessel walls that supply blood to heart and brain (1). Atherosclerosis is the most common precursor and process underlying most CVD, and the most commonly occurring diseases are coronary artery disease, carotid artery disease, myocardial infarction, and stroke (2, 3). It is a condition that artery wall becomes thicken and harden due to inflammatory plaque formation and lipoproteins accumulation. Atherosclerosis is a multistage, multifocal metabolic disorder, and it starts early in life and may take decades to develop into clinical CVD (4). Atherosclerosis can lead to coronary artery disease and carotid artery disease when the arteries that carry oxygen-rich blood to heart and brain tissue is blocked (5). Inflammatory response from the leukocytes is a significant player in atherosclerotic plaque formation and disease progression. (2). Cholesterol derivatives are well-established atherosclerosis risk factors, and their accumulation is major contribution of atherosclerotic plaques (6, 7).

The prevalence of metabolic disorders is significantly increased among the elderly population and further lead to increased cardiovascular morbidity and mortality (8). Established risk factors of ASCVD, including smoking, blood pressure, and lipid level contributes to current cardiovascular risk assessment, but these characteristics are insufficient to fully explain cardiovascular risks. Therefore, novel biomarkers for screening, predicting, and identifying ASCVD has invoked substantial interest (4, 9).

Carotid intima-media thickness (IMT) is a subclinical measurement of atherosclerosis and a strong predictor of cardiovascular risk (3, 10-12). It is the observable distance between the lumen-intima and the media-adventitia interface of carotid artery measured by B-mode ultrasound (13). It is the most reliable and well-established measurement of the potential development of atherosclerosis and to predict future cardiovascular events. In the last two decades, the carotid IMT measurement for risk estimation of ASCVD has significantly increased (13). Carotid IMT could not only identifying and quantifying atherosclerosis development much earlier in subjects with significant CVD risk factors including diabetes, hypertension, and high cholesterol level, and also independent of other traditional risk factors such as age, gender, and race (14). The change of carotid IMT overtime is often used to measure the progression of atherosclerosis. Carotid IMT test is suitable to be used in studies with large population because this test is simple, affordable, non-invasive, methodologically standardized, and reproducible (11, 14). The IMT measurement is usually made either on the wall nearest or furthest to the ultrasound transducer, which are near wall and far wall measurements respectively (15).

Metabolic profiling is an analytical method for investigating small molecular metabolites with unique chemical fingerprints which are the final representation of the organism's phenotype (16). The concept of "metabolic profiling" was introduced in 1971 by Horning, et al. after they used gas chromatography-mass spectrometry (GC-MS) to measure metabolites in human urine (17). Metabolomic is an emerging field of system biology and commonly defined as a comprehensive measurement of all low-molecular-weight metabolic molecules in the human specimen (18). Metabolomics researches usually based

on mass spectrometry (MS) and nuclear magnetic resonance (NMR), which are highly sensitive instrumentations producing reliable quantitative measurement of the metabolites (19). The identification of novel biomarkers associated with atherosclerosis is important to help early diagnosis and prevention, even the mechanism of the metabolites may not identically lead to atherosclerosis or other CVD (16). Because the biochemical pathway has been a focused research area over a century, the changes in metabolism are more observable, and the interpretation of these metabolism changes are more precise and applicable (20). The enormous potential of biomarker identification and evaluation has made metabolomic profiling is one of the most powerful tools to investigate downstream genetic production and identify novel metabolomic biomarkers for various diseases.

Molecular level studies on atherosclerosis have provided valuable information on the pathology and identify novel biomarkers for the disease prevention and diagnosis. One study investigated the association of endothelial inflammation as a risk factor of atherosclerosis and increased plasma phospholipid polyunsaturated fatty acids levels, while obesity modifies this association (21). Another large population study aiming to improve the diagnosis of atherosclerosis was conducted in 2012. Metabolomic profiles in the serum samples revealed that docosahexaenoic acid, glutamine, and tyrosine levels may be potential biomarker for atherosclerosis (4). Chen et al. performed an untargeted metabolomics study using Gas Chromatography-Mass Spectrometry (GC-MS) suggested that plasma levels of palmitate, stearate, and 1-monolinoleoylglycerol may constitute biomarkers of this disease (22). One study was conducted to investigate the metabolome composition from atherosclerotic plaque tissue and intima tissue. The authors compared

the metabolites in the tissue samples from patients who underwent carotid or femoral endarterectomy. They observed the metabolism level changes of acylcarnitines in atherosclerosis plaque tissues and suggested phosphatidylethanoleamine-ceramides may serve as a biomarker of atherosclerosis pathogenesis (7).

Twin studies have a unique value in learning about the pathology of diseases for separating genetic factors from environment effects. Molecular profiling among twins is a powerful tool to investigate the molecular association with complex traits (23). Twins are naturally matched pairs who share similar biological and behavioral features such as maternal effects or age. Therefore, some potential confounding effect would be removed when making comparison between twins who share them. The association derived from twin study can be explained by general association among twin pairs treated as single subjects, or individual (unshared environment) influences between the two twins in a pair. The differences in phenotype between co-twins imply the contribution due to unshared environmental factors like diet, smoking status, mental health (24). In Carlin's paper, the authors discussed the regression model and the interpretation of coefficients in the twin studies (25). The within-twin coefficient represents the expected change of the outcome explained by unshared effect between co-twins, which is the one-unit change in the difference between individual exposure value and the average exposure value in the twin pairs. The between-twin coefficient gives the expected changes in the outcome for a one-unit change in the average value of exposure in the twin pair, which is explained by the shared environmental or genetic factors.

Metabolomics study to identify potential biomarkers or metabolic pathways of atherosclerosis in a population of twin pairs possess unique value in learning about how shared and unshared environment would contribute to this disease between twins. The goal of this study, therefore, is taking the advantage of twin pairs to investigate the association of metabolites in plasma samples with Carotid IMT measurement among 307 male twins. The most valuable benefit of the study is utilizing a mixed effect model to consider metabolic profiling changes within twin pairs, which emphasize environmental factors and exposures ruling out the shared effects (such as maternal effect, nutrition supply in early life), and the between-twin effects that indicate general associations.

Methods

Study Population

The twins included in the study were selected from the Vietnam Era Twin (VET) Registry, which contained 7,369 male twin pairs who were born between 1946 and 1956. All twins served in the United States military during the period of the Vietnam War. The VET registry is one of the largest twin registries in the United States. Our project utilized data from the Emory Twin Study (ETS), which recruited 307 monozygotic and dizygotic twin pairs from the VET Registry to investigate behavioral, biological, and pathological risk factors of cardiovascular diseases among twins. The Emory University General Clinical Research Center examined and updated the medical information on all twins between March 2002 and March 2006 (3). We measured metabolomic profiles of 187 subjects including 92 twin pairs and 3 singletons from the ETS. The twin pairs were selected based on the availability of metabolic data and valid Carotid IMT of the twin pairs. All twin pairs in this study are white male Caucasians. DNA analysis were performed to determine the zygosity information of the twins. Details of the Registry construction and composition are well documented (26). All participants signed informed consent, and the Emory Institutional Review Board approved the study protocol.

Measurements

All measurements were performed in the morning after an overnight fast and instructed to restrain smoking. All medication uses were held for 24 hours before testing. Measurements for both twin pairs were performed at the same time. A medical history and physical exam were obtained from all twins. Body weight and height were measured when participants

wore light clothes and no shoes. BMI was calculated as weight in kilograms divided by height in meters squared. Systolic blood pressure (SBP) and diastolic blood pressure (DBP) were measured twice using a mercury sphygmomanometer on the right arm in the sitting position after a 10-minute rest. Smoking status was classified as current smoker and past or never smoker. Physical activity was assessed using the Baecke Questionnaire of Habitual Physical Activity, which is a 16-question measurement of physical activity levels. The total physical activity score was used in the analysis. Venous blood samples were drawn to collect plasma and evaluate relative abundance of metabolites. Low-density lipoprotein (LDL) and high-density lipoprotein (HDL) were obtained using homogeneous assays (Equal Diagnostics, Exton, PA) (3). Carotid IMT was measured on both the near and front wall of the carotid artery using high-resolution B-mode ultrasonography. The sonographer measured four different segments on the carotid artery. Mean of the four measurements were taken as carotid IMT value for each subjects in this study. The difference of carotid IMT value of 2 subjects in a twin pair is considered as within-twin effect, and the mean of carotid IMT value is considered as between-twin effect. For 3 singletons, to avoid missing measurement, we duplicate the IMT value to calculate the IMT difference and mean. Therefore, the IMT mean is IMT value itself and IMT difference is 0 for these 3 singletons.

Metabolomics Measurements

Metabolomics data were collected from fasting blood samples. The high-resolution metabolomics (HRM) profiling of twins were performed using high-performance liquid chromatography-mass spectrometry (HPLC-MS). Each plasma sample was analyzed for

three replicates using a dual C_{18} chromatography (C_{18} Higgins Analytical Inc) (27). Blood plasma samples were treated with acetonitrile acid to precipitate protein before being centrifuged on 13,000 rpm at 4°C. MS data were collected by using both positive and negative electrospray ionization from high-resolution mass spectrometer (Q-Exactive Thermo Fisher Scientific, San Diego, CA). Mass-to-charge ratio (m/z), intention time, and ion intensity were collected. Adaptive processing software apLCMS was applied to extract LC-MS raw data and perform quantification of metabolites (28). Data filtering, normalization, merging, and summarization were performed using xMSanalyzer (29).

Metabolome-wide Association Study (MWAS) and Statistical Methods

HRM collected relative abundance of 20,035 metabolic features in plasma samples, including 12,527 negative ionized features and 7508 positive ionized features. All relative abundance All relative abundance values were $\log_2$ transformed to increase normality and missing values were imputed from half of the minimum detected value. Previous studies suggested that age, BMI, cholesterol level, can be confounders of the associations of metabolite levels with increasing IMT that should be controlled for. Smoking status was also considered as a major confounder of the association between this association. Linear mixed effect models that considering both within-twin effect and between-twin effect were applied to investigate the association between carotid IMT and metabolomics features among twins, while controlling for age, BMI, total cholesterol level, and smoking status.

In this study, within-twin coefficient and between-twin coefficient estimates were considered as random effect and fixed effect, respectively (25). The within-twin coefficient represents the expected change of the outcome for a one-unit change in the difference between individual exposure value and the average exposure value in the twin pairs. This estimate provided information about the difference of carotid IMT measurement explained by the two-time difference in metabolites level between co-twins, which further represents the contribution of unshared environmental factors such as mental health or diet. The between-twin coefficient gave the expected changes in the outcome for a one-unit change in the average value of exposure in the twin pair, controlling for individual variance constant. This estimate demonstrated the difference in carotid IMT measurements explained on the population level effect of shared maternal, genetic, and environmental effects.

Pathway Analysis and Annotation

Pathway analysis was performed using Mummichog, which is a software that computes all possible metabolites matched with the input significant features and find the metabolic pathways and network that could possibly formed by those features (30). All features with a raw significant association with IMT ($p<0.05$) were input to Mummichog as significant features, and other features are input as reference features. After excluding features with missing p-value and selecting features met the threshold, we had 433 and 422 positive ionized significant features for within-twin effect and between-twin effect, and 649 and 686 negative ionized significant features for within-twin effect and between-twin effect to input to Mummichog. The MS mode was selected for each batch of input features, number

of permutations to estimate null distributions is set to 1000, and all remaining options are set as default. Identified statistically significant pathways were further classified into typical metabolism classes based on KEGG Pathway. In order to improve the reliability of metabolism pathway analysis result, metabolism pathways with more than 3 overlapped significant metabolites were included. Other pathways that do not meet this criterion were not included even the p-value is lower than 0.05.

Potential metabolite identification and annotation are based on the Human Metabolomics Database (HMDB). We compare annotation from HMDB to compare and supplement results from Mummichog that based on database Kyoto Encyclopedia of Genes and Genomes (KEGG) (30).

All statistical analyses were performed in the R statistical environment version 3.4.4 (http://www.r-project.org/). Implementation of linear mixed effect models utilized R package nlme.

Results

The baseline characteristics of the study population were divided into two groups based on their zygotic information and summarized in Table 1. The study population included 135 individuals of monozygotic twins and 52 individuals of dizygotic twins. Among the study population, 30% of the participants were current smokers, and 62% were alcohol consumers. Monozygotic twins were more likely to be smokers and less likely to be alcohol consumers than dizygotic twins. The age of the twins ranged from 47 to 59 years old with an average age of 55 years old. Carotid IMT measurements and other cardiovascular factors such as cholesterol level were similar between the two groups; BMI and measured blood pressure were slightly higher in zygotic twins.

Quantile-quantile plots and Manhattan plots were created for positive ionized features and negative ionized features considering both within-twin effect and between-twin effect (Figure 1. – Figure 4.). The inflation factors range from 0.973 to 1.081, indicating moderate global inflation of low p-values. Absence of abnormal tail deviance and no large inflation supported the validity of the linear mixed effect model used in this study. Manhattan plot show no single metabolic feature is statistically significantly associated with carotid IMT after adjusting for multiple tests.

After Bonferroni correction of multiple tests (12,527 for negative ionized features and 7508 for positive ionized features), no significant metabolic feature in both positive and negative ionization mode was found to be associated with increasing carotid IMT measurement,

considering both within-twin effects and between-twin effects and adjusting for age, BMI, and cholesterol level. Among 855 positive ionization mode features with raw p-value < 0.05, 86 features were identified as certain metabolites. Among 1,335 negative features with raw p-value < 0.05, 135 metabolic features were identified. The top 5 annotated metabolites associated with carotid IMT of between-twin effect were listed in Table 2, and top 5 annotated metabolites associated with carotid IMT of within-twin effect were listed in Table 3. Both tables included identified negative and positive ionized features with lowest p-value. All identifications and annotations were based on HMDB database. Features with multiple possible annotation were not included because we lack further information to select the correct annotation or address the false annotation in this study. The top 5 identified metabolites that associated with carotid IMT between twin pairs were ribose-1-arsenate, nitroxoline, 8-dimethoxy-2,6-dimethyl-2-octanol, 5-dimethyl-1H-pyrrole, and 5-phosphonooxy-L-lysine. All metabolic features were not significantly associated with carotid IMT. Top 5 identified metabolites associated with carotid IMT between twin pairs were 5-acetylamino-6-amino-3-methyluracil, austalide G, (S)-3-sulfonatolactate, diethanolamine, and histidinal. Similarly, none of these metabolites were associated with carotid IMT between twin pairs.

Result of pathway network analysis performed by Mummichog was listed in Table 4. We identified 17 metabolism pathways statistically significantly associated with carotid IMT and risk of having atherosclerosis, under the threshold of raw p-value lower than 0.05. Among 17 metabolism pathways, 8 pathways classified as amino acid related metabolism pathways, 3 were classified as metabolism of cofactors and vitamins, 2 lipid metabolisms,

2 carbohydrate metabolisms, 1 energy metabolism, and 1 nucleotide metabolism pathway. Among those, 4 pathways were statistically significant between twin pairs including cholecalciferol metabolism, glutathione metabolism, ascorbate and aldarate metabolism, and galactose metabolism. Glutathione metabolism, which further links to the glutamate metabolism, is the only pathway that significantly associated with carotid IMT considering both within-twin effects and between-twin effects. According to the results, alanine and aspartate metabolism, a branch of amino acid metabolism, was the most significant pathway to explain the association between plasma metabolites and carotid IMT on a population level. The significant amino acid metabolism pathways supported the fact that amino acids and their downstream metabolites might play important roles in atherosclerosis pathology, and some typical metabolites within these pathways might serve as potential biomarkers of higher risk of having atherosclerosis and other related cardiovascular events. Lipid metabolism pathways were only significant between the co-twins, which may be explained by unshared environment effect within a pair.

All pathway identification and classification were based on KEGG Pathway database (30). The significant level of input metabolite features was raw association p-value < 0.05. We only included metabolism pathways with more than 3 overlapped significant metabolites in order to improve reliability of pathway analysis.

Discussion

Our research of metabolomics study on carotid IMT did not identify any significant metabolites associated with atherosclerosis in a sample of twins. The pathway analysis of metabolic features with p-value lower than 0.05, we did not identify any top metabolic associations (Table 2 and Table 3) as a member of any significant metabolism pathway. Ribose-1-arsenate is an intermediate in arsenate detoxification I pathway. The ribose can conjugate to arsenate by the enzyme purine nucleoside phosphorylase (PNP), and this process belongs to nicotinate and nicotinamide metabolism. Histidinal is involved in the histidine biosynthesis I pathway, which is produced by the reaction between histidinol and NAD+. Some of the metabolites are food and drug intake related chemical compounds, such as nitroxoline, 2,5-dimethyl-1H-pyrrole, and 8,8-dimethoxy-2,6-dimethyl-2-octanol. Nitroxoline is a urinary antibacterial agent that against susceptible organisms commonly found in urinary tract infections. It is only found in individuals that have used this drug and do not belong to any metabolism pathway. 8,8-dimethoxy-2,6-dimethyl-2-octanol is a flavoring ingredient. The appearance of this metabolites is possibly come from drug use and food intake. Even the participants were required to held medicine uses and took an overnight fasting, the ingredients from drug and food may retain in blood for a relatively long time. Similarly, we also found 5-acetylamino-6-amino-3-methyluracil (AAMU), one of caffeine major metabolites, is associated with carotid IMT. However, according to the results of pathway network analysis, caffeine metabolism is not a significant or plausible pathway that associated with carotid IMT.

Inaccurate annotation is also a plausible reason to explain the identified metabolites with lowest p-value were not metabolic members in the significant metabolism pathways. In this study, we excluded the multiple annotated metabolites due to the uncertainty of specific features. Some features with same m/z ratio and retention time may be annotated to multiple metabolites or different molecular structural form. For example, one feature of m/z 105.07 and retention time of 82.89 could be identified as 3-methyl-2-butanethiol, 2-methyl-1-butanethiol, or 1-pentanethiol. Differentiating the metabolites requires the utilization of further identification technologies such as ^{1}H NMR, which was not performed in this study.

Studies have shown that amino acid metabolism is strongly associated with carotid IMT. Glutamate is metabolized into glutamine in cells and glutamine is an important precursor of glucose during fast. One study found glutamine could be a marker of preclinical atherosclerosis, and the level of its related metabolites are also possibly associated with coronary artery disease (4). After a series of metabolic activity, glutamate metabolism will lead to alanine metabolic activity. Studies also illustrated elevated alanine aminotransferase level is significantly associated with atherosclerosis, and may be a biomarker for preclinical diagnosis or medical intervention (31, 32). A medical experiment performed on New Zealand rabbits suggests L-aspartate and L-glutamate has protective effect on the atherosclerosis disease (33). There is a general agreement that lipid metabolism was an important character in atherosclerosis progression (34). For example, cholesterol is a well-established risk factor of atherosclerosis. The metabolism level unsaturated fatty acid is associated with metabolic oxidation condition. Oxidative stress

will lead to inflammation, which is a very important composition of atherosclerosis progression (7). Studies have suggested the causal connection between lipid metabolism and atherosclerosis according to the observation that patients with myocardial infraction are in general have higher cholesterol level than normal group (19). Moreover, the lipolysis could also produce fatty acid. Studies has demonstrated that long-chain saturated fatty acid are cytotoxic, thus may introduce cell death and initiating inflammation, which are key attributes and necessary process of atherogenesis (35). The accumulation of fatty acid deposit, inflammation plaque, and death cell on the carotid would increase the thickness artery wall and lead to potential ASCVD events. The result from pathway analysis suggests atherosclerosis is strongly associated with amino acid metabolism and lipid metabolism. This association has evoked great research interest, and further study on the related metabolism pathway will hopefully provide us more novel biomarkers for preclinical diagnosis and early therapeutic intervention of atherosclerosis and ASCVD.

There are several strengths in this study we should consider. We utilized a cross-sectional analysis to study the association of metabolites level in plasma with carotid IMT, considering both between-twin effects and within-twin effects. Thus, genetic factors and early environmental exposures were addressed since twin pairs share similar biological characteristics and early life environment. In addition, participants in this study were mostly white and all were males, and thus race and sex, which may confound the association of metabolites with carotid IMT, were already removed. However, limitations do exist in this study. First, our study population is derived from military veterans, they tend to be stronger, physically healthier, while less mentally health due to higher proportion

of Post-traumatic Stress Disorder (PTSD). Therefore, the generalizability to general population is unknown. Second, the Exclusion of twins with missing metabolomics data may be a potential selection bias. Third, the cross-sectional study design lack of ability to detect casual effect between carotid IMT and metabolites. Therefore, the temporal relationship of metabolite concentrations and the presence of carotid IMT could not be addressed. Fourth, this study have small sample size. Thus we have limited power to detect the metabolic association with carotid IMT.

Conclusion

In conclusion, our findings from the twin study, taken in context with those from previous studies, suggest that no single metabolic feature alone can explain the carotid IMT changes. However, pathway analysis indicates glutathione metabolism pathway is associated with carotid IMT among middle-aged males. The unshared environmental effect between co-twins may explain the association between different level of lipid metabolism and changes in carotid IMT. Further study with larger sample size and prospective study design would be benefitted to identify potential biomarkers for early diagnosis and prevention of carotid IMT and the potential event of atherosclerosis.

References

1. Ellulu MS, Patimah I, Khaza'ai H, et al. Atherosclerotic cardiovascular disease: a review of initiators and protective factors. *Inflammopharmacology* 2016;24(1):1-10.
2. Falk E. Pathogenesis of atherosclerosis. *J Am Coll Cardiol* 2006;47(8 Suppl):C7-12.
3. Zhao J, Cheema FA, Bremner JD, et al. Heritability of carotid intima-media thickness: a twin study. *Atherosclerosis* 2008;197(2):814-20.
4. Wurtz P, Raiko JR, Magnussen CG, et al. High-throughput quantification of circulating metabolites improves prediction of subclinical atherosclerosis. *Eur Heart J* 2012;33(18):2307-16.
5. Jashari F, Ibrahimi P, Nicoll R, et al. Coronary and carotid atherosclerosis: similarities and differences. *Atherosclerosis* 2013;227(2):193-200.
6. Radovic B, Aflaki E, Kratky D. Adipose triglyceride lipase in immune response, inflammation, and atherosclerosis. *Biol Chem* 2012;393(9):1005-11.
7. Vorkas PA, Shalhoub J, Isaac G, et al. Metabolic phenotyping of atherosclerotic plaques reveals latent associations between free cholesterol and ceramide metabolism in atherogenesis. *J Proteome Res* 2015;14(3):1389-99.
8. Costantino S, Paneni F, Cosentino F. Ageing, metabolism and cardiovascular disease. *J Physiol* 2016;594(8):2061-73.
9. Wang TJ, Gona P, Larson MG, et al. Multiple biomarkers for the prediction of first major cardiovascular events and death. *N Engl J Med* 2006;355(25):2631-9.
10. O'Leary DH, Polak JF. Intima-media thickness: a tool for atherosclerosis imaging and event prediction. *Am J Cardiol* 2002;90(10C):18L-21L.
11. Simon A, Gariepy J, Chironi G, et al. Intima-media thickness: a new tool for diagnosis and treatment of cardiovascular risk. *J Hypertens* 2002;20(2):159-69.

12. Lorenz MW, von Kegler S, Steinmetz H, et al. Carotid intima-media thickening indicates a higher vascular risk across a wide age range: prospective data from the Carotid Atherosclerosis Progression Study (CAPS). *Stroke* 2006;37(1):87-92.

13. Bauer M, Caviezel S, Teynor A, et al. Carotid intima-media thickness as a biomarker of subclinical atherosclerosis. *Swiss Med Wkly* 2012;142:w13705.

14. Onut R, Balanescu AP, Constantinescu D, et al. Imaging Atherosclerosis by Carotid Intima-media Thickness in vivo: How to, Where and in Whom ? *Maedica (Buchar)* 2012;7(2):153-62.

15. Polak JF, Szklo M, O'Leary DH. Associations of Coronary Heart Disease with Common Carotid Artery Near and Far Wall Intima-Media Thickness: The Multi-Ethnic Study of Atherosclerosis. *J Am Soc Echocardiogr* 2015;28(9):1114-21.

16. Kordalewska M, Markuszewski MJ. Metabolomics in cardiovascular diseases. *J Pharm Biomed Anal* 2015;113:121-36.

17. Horning EC, Horning MG. Metabolic profiles: gas-phase methods for analysis of metabolites. *Clin Chem* 1971;17(8):802-9.

18. Clish CB. Metabolomics: an emerging but powerful tool for precision medicine. *Cold Spring Harb Mol Case Stud* 2015;1(1):a000588.

19. Goonewardena SN, Prevette LE, Desai AA. Metabolomics and atherosclerosis. *Curr Atheroscler Rep* 2010;12(4):267-72.

20. German JB, Hammock BD, Watkins SM. Metabolomics: building on a century of biochemistry to guide human health. *Metabolomics* 2005;1(1):3-9.

21. Steffen BT, Steffen LM, Tracy R, et al. Obesity modifies the association between plasma phospholipid polyunsaturated fatty acids and markers of inflammation: the Multi-Ethnic Study of Atherosclerosis. *Int J Obes (Lond)* 2012;36(6):797-804.

22. Chen X, Liu L, Palacios G, et al. Plasma metabolomics reveals biomarkers of the atherosclerosis. *J Sep Sci* 2010;33(17-18):2776-83.

23. van Dongen J, Slagboom PE, Draisma HH, et al. The continuing value of twin studies in the omics era. *Nat Rev Genet* 2012;13(9):640-53.

24. Kaminsky ZA, Tang T, Wang SC, et al. DNA methylation profiles in monozygotic and dizygotic twins. *Nat Genet* 2009;41(2):240-5.

25. Carlin JB, Gurrin LC, Sterne JA, et al. Regression models for twin studies: a critical review. *Int J Epidemiol* 2005;34(5):1089-99.

26. Goldberg J, Curran B, Vitek ME, et al. The Vietnam Era Twin Registry. *Twin Res* 2002;5(5):476-81.

27. Huang Y, Hui Q, Walker DI, et al. Untargeted metabolomics reveals multiple metabolites influencing smoking-related DNA methylation. *Epigenomics* 2018;10(4):379-93.

28. Yu T, Park Y, Johnson JM, et al. apLCMS--adaptive processing of high-resolution LC/MS data. *Bioinformatics* 2009;25(15):1930-6.

29. Uppal K, Soltow QA, Strobel FH, et al. xMSanalyzer: automated pipeline for improved feature detection and downstream analysis of large-scale, non-targeted metabolomics data. *BMC Bioinformatics* 2013;14:15.

30. Li S, Park Y, Duraisingham S, et al. Predicting network activity from high throughput metabolomics. *PLoS Comput Biol* 2013;9(7):e1003123.

31. Yamada J, Tomiyama H, Yambe M, et al. Elevated serum levels of alanine aminotransferase and gamma glutamyltransferase are markers of inflammation and oxidative stress independent of the metabolic syndrome. *Atherosclerosis* 2006;189(1):198-205.

32. Adibi P, Sadeghi M, Mahsa M, et al. Prediction of coronary atherosclerotic disease with liver transaminase level. *Liver Int* 2007;27(7):895-900.

33. Yanni AE, Agrogiannis G, Nomikos T, et al. Oral supplementation with L-aspartate and L-glutamate inhibits atherogenesis and fatty liver disease in cholesterol-fed rabbit. *Amino Acids* 2010;38(5):1323-31.

34. Gould RG. Lipid metabolism and atherosclerosis. *Am J Med* 1951;11(2):209-27.

35. Linton MRF, Yancey PG, Davies SS, et al. The Role of Lipids and Lipoproteins in Atherosclerosis. In: Feingold KR, Anawalt B, Boyce A, et al., eds. *Endotext*. South Dartmouth (MA), 2000.

Tables and Figures

Table 1. Demographic and Cardiovascular Characteristics of Twins, Emory Twin Study (ETS)

Variables	Overall (n = 187)	MZ Twins (n = 135)	DZ Twins (n = 52)
Demographics			
Age, years (s.d.)	55.8 (3.3)	55.7 (3.3)	56.3 (3.0)
Current smoker, n (%)	59 (31.6)	46 (34.1)	13 (25.0)
Current alcohol consumer, n (%)	116 (62.0)	84 (62.2)	32 (61.5)
Physical activity[a] (s.d.)	7.2 (1.7)	7.2 (1.7)	7.4 (1.6)
Cardiovascular factors			
BMI, kg/m^2 (s.d.)	29.5 (4.7)	29.3 (4.8)	30.0 (4.4)
LDL-cholesterol, mg/dL (s.d.)	119.6 (38.1)	120.6 (40.4)	117.3 (31.8)
HDL-cholesterol, mg/dL (s.d.)	38.8 (11.8)	38.3 (12.0)	40.1 (11.5)
Total cholesterol, mg/dL (s.d.)	185.5 (42.1)	183.9 (43.2)	189.7 (39.2)
Systolic blood pressure, mmHg (s.d.)	131.0 (16.1)	130.9 (16.4)	131.3 (15.7)
Diastolic blood pressure, mmHg (s.d.)	81.4 (10.6)	81.0 (10.6)	82.8 (10.6)
Hypertension, n (%)	63 (33.7)	43 (31.9)	20 (38.5)
IMT, mm (s.d.)	0.763 (0.126)	0.767 (0.130)	0.753 (0.116)

Abbreviations: MZ: monozygotic; DZ: dizygotic; BMI, body mass index; HDL, high-density lipoprotein; LDL, low-density lipoprotein; IMT, intima-media thickness; s.d., standard deviation.
[a] Physical activity level was measured based on Baecke score.

Table 2. Summary of Top 5 Annotated Metabolites from Between-twin MWAS Analysis

Metabolites	m/z	Retention time	Between-twin Estimate	Within-twin Estimate	Between-twin p-value	Within-twin p-value
Ribose-1-arsenate	-272.9572	165.7528	-1.3691	-0.0540	1.71E-04	0.923
Nitroxoline	-189.0289	73.0188	1.4628	1.0480	5.38E-04	0.130
5-phosphonooxy-L-lysine	-241.0590	219.4582	-1.1078	0.4306	3.36E-03	0.490
8,8-dimethoxy-2,6-dimethyl-2-octanol	+219.1953	231.1526	-2.0205	0.0260	1.61E-03	0.982
2,5-dimethyl-1H-pyrrole	+96.0812	80.7621	0.9859	-0.1552	2.14E-03	0.711

Abbreviations: MWAS, Metabolome-Wide Association Study; M/Z, mass-to-charge ratio.

Table 3. Summary of Top 5 Annotated Metabolites from Within-twin MWAS Analysis

Metabolites	m/z	Retention time	Between-twin Estimate	Within-twin Estimate	Between-twin p-value	Within-twin p-value
5-acetylamino-6-amino-3-methyluracil	-197.0670	115.0579	-0.4327	3.5314	0.760	2.87E-04
Austalide G	-517.2418	142.2371	-1.4096	-6.8214	0.433	1.01E-03
(S)-3-sulfonatolactate	-168.9799	438.1638	0.4161	-7.9353	0.742	1.03E-03
Diethanolamine	+106.0867	548.9763	-0.7332	2.9753	0.267	2.28E-03
Histidinal	-138.0660	39.5965	-0.6954	2.3751	0.189	3.25E-03

Abbreviations: MWAS, Metabolome-Wide Association Study; M/Z, mass-to-charge ratio.

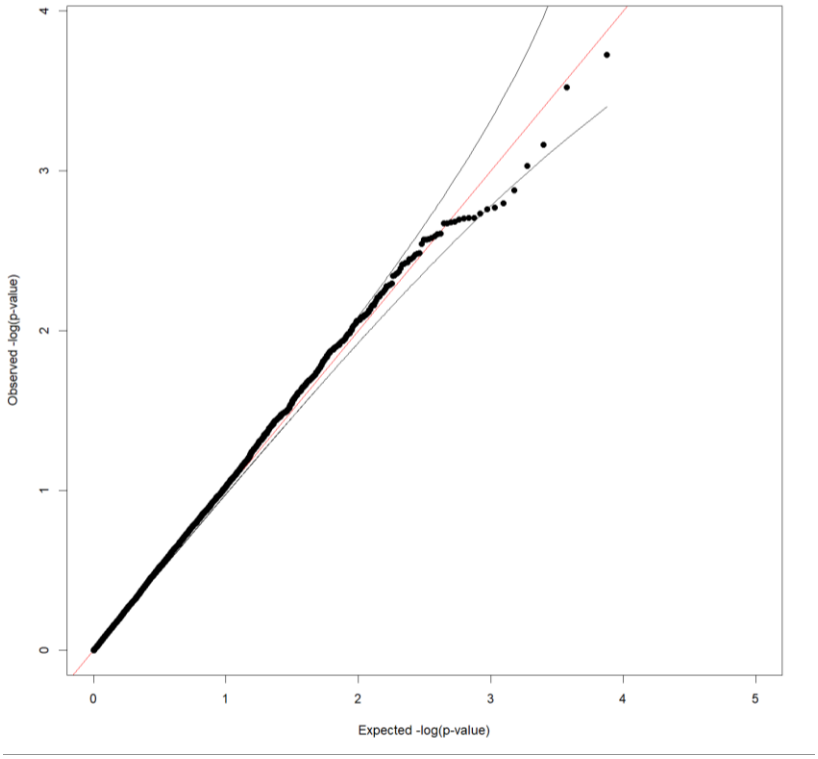
Table 4. Summary of Significant Pathways from Mummichog Metabolic Network Analysis

Pathway	Significant Metabolites[a]	p-value[b]	Metabolism Class
Significant Between-twin Metabolism Pathways			
Vitamin D3 (cholecalciferol) metabolism	5	9.20E-04	Metabolism of cofactors and vitamins
Glutathione metabolism	3	8.23E-03	Metabolism of other amino acids
Ascorbate and aldarate metabolism	4	3.64E-02	Carbohydrate metabolism
Galactose metabolism	3	4.20E-02	Carbohydrate metabolism
Significant Within-twin Metabolism Pathways			
Alanine and aspartate metabolism	7	1.60E-04	Amino acid metabolism
Nicotinate and nicotinamide metabolism	5	1.10E-03	Metabolism of cofactors and vitamins
Arginine and proline metabolism	8	1.44E-03	Amino acid metabolism
Saturated fatty acids metabolism	4	1.55E-03	Lipid metabolism
Nitrogen metabolism	4	2.08E-03	Energy metabolism
Vitamin B9 (folate) metabolism	4	3.16E-03	Metabolism of cofactors and vitamins
Lysine metabolism	7	3.70E-03	Amino acid metabolism
Arachidonic acid metabolism	6	3.86E-03	Lipid metabolism
Aspartate and asparagine metabolism	10	5.42E-03	Amino acid metabolism
Glutathione metabolism	3	8.23E-03	Metabolism of other amino acids
Beta-alanine metabolism	3	1.93E-02	Metabolism of other amino acids
Purine metabolism	5	3.14E-02	Nucleotide metabolism
Glutamate metabolism	3	3.49E-02	Amino acid metabolism

[a, b]Significant threshold is raw p-value < 0.05.

Figure 1. Metabolome-wide Association with Carotid Intima-media Thickness of Positive Ionized Metabolic Features Between Twin Pairs, Adjusted for Age, BMI, Smoking Status and Cholesterol Level; (A) Quantile-Quantile Plot (Red Straight Line: y=x, Black Curves: 95% Confidence Interval for the Global Null Hypothesis, Inflation Factor = 1.023); (B) Manhattan Plot (No Significant Feature Identified)

A)



B)

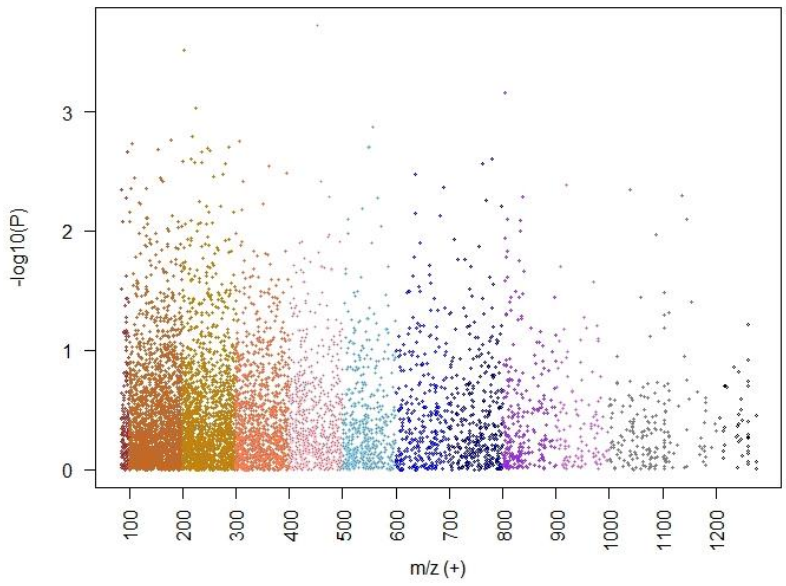
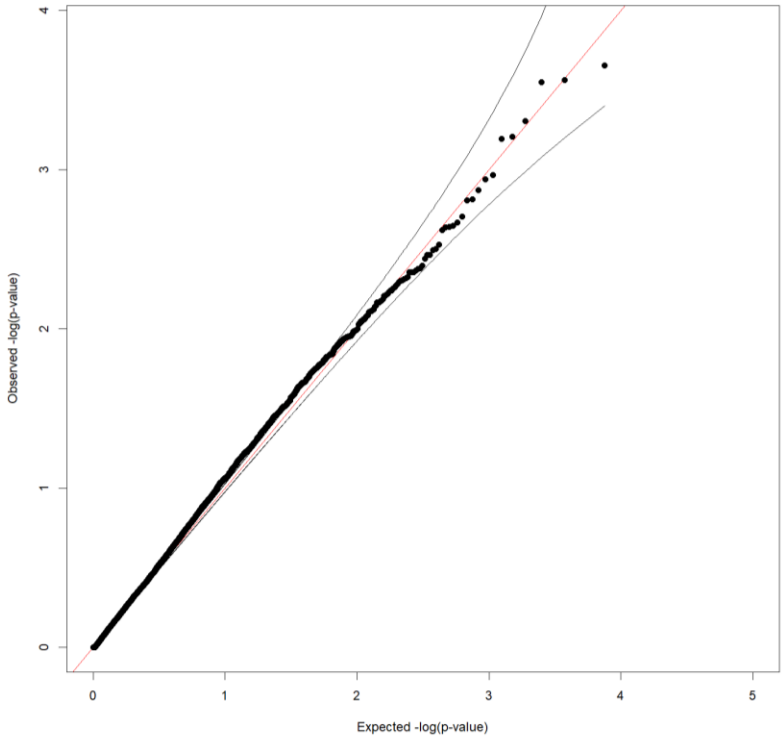


Figure 2. Metabolome-wide Association with Carotid Intima-media Thickness of Positive Ionized Metabolic Features Within Co-twins, Adjusted for Age, BMI, Smoking Status and Cholesterol Level; (A) Quantile-Quantile Plot (Red Straight Line: y=x, Black Curves: 95% Confidence Interval for the Global Null Hypothesis, Inflation Factor = 1.033); (B) Manhattan Plot (No Significant Feature Identified)

A)



B)

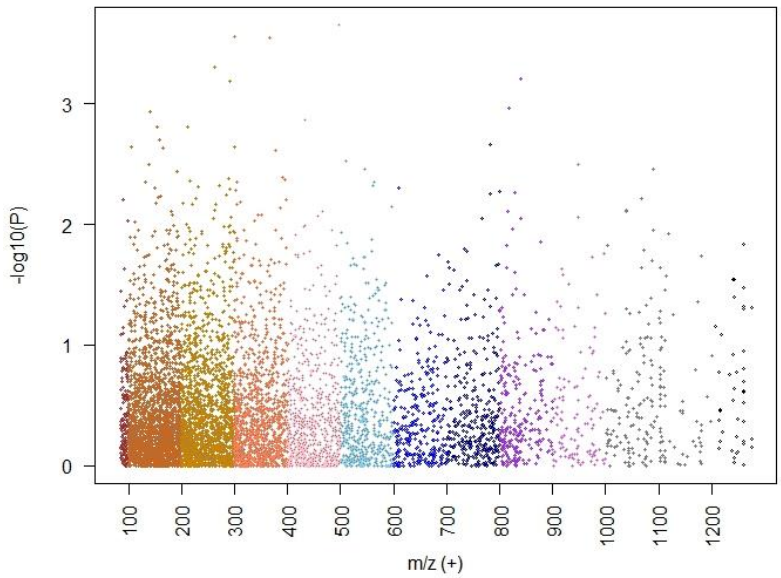
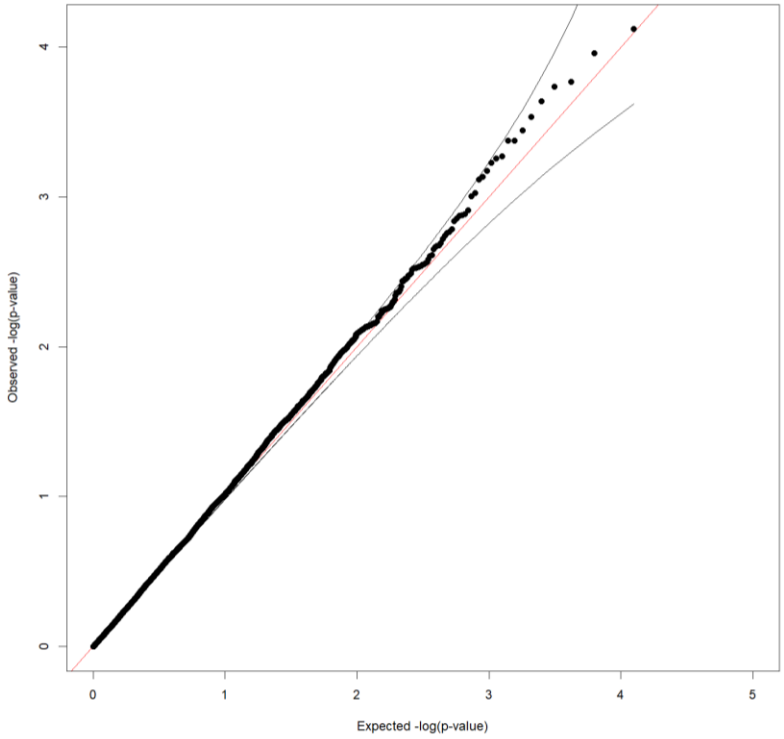


Figure 3. Metabolome-wide Association with Carotid Intima-media Thickness of Negative Ionized Metabolic Features Between Twin Pairs, Adjusted for Age, BMI, Smoking Status and Cholesterol Level; (A) Quantile-Quantile Plot (Red Straight Line: y=x, Black Curves: 95% Confidence Interval for the Global Null Hypothesis, Inflation Factor = 1.004); (B) Manhattan Plot (No Significant Feature Identified)

A)



B)

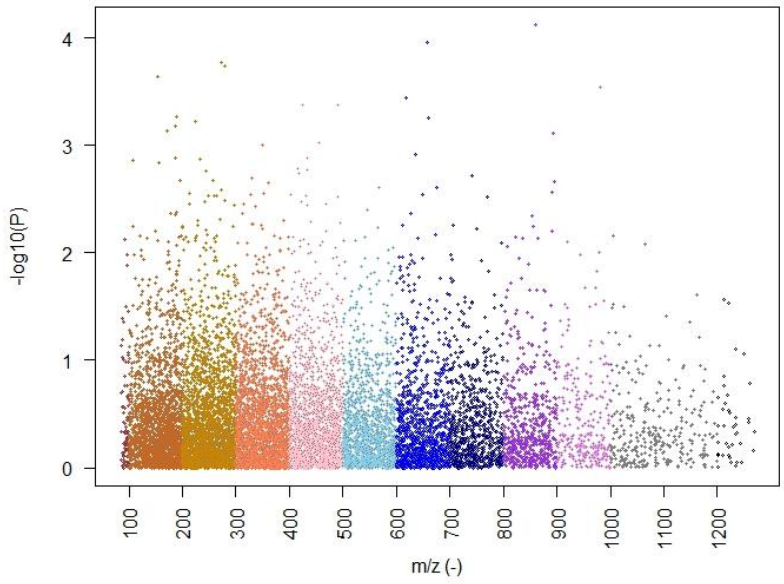
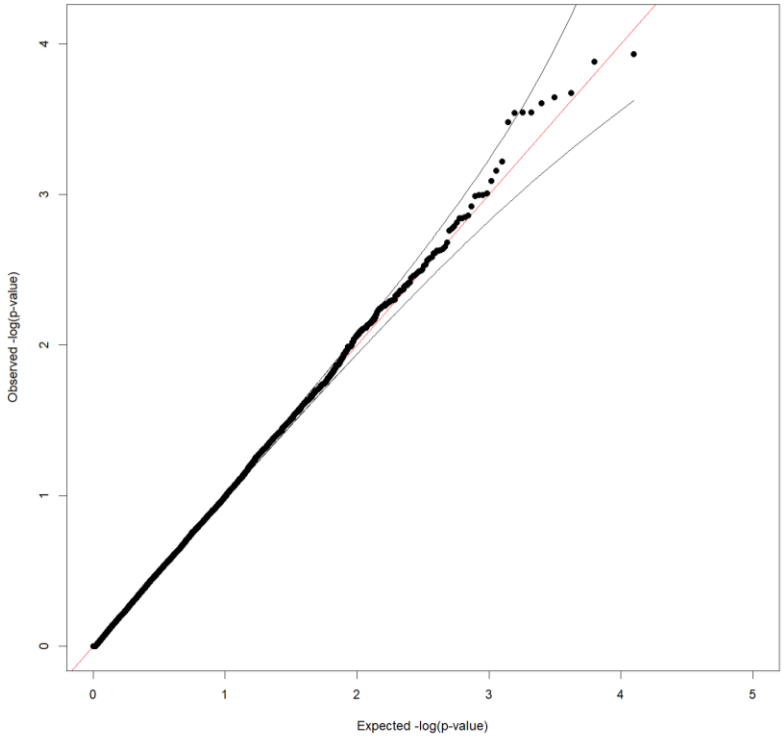


Figure 4. Metabolome-wide Association with Carotid Intima-media Thickness of Negative Ionized Metabolic Features Within Co-twins, Adjusted for Age, BMI, Smoking Status and Cholesterol Level; (A) Quantile-Quantile Plot (Red Straight Line: y=x, Black Curves: 95% Confidence Interval for the Global Null Hypothesis, Inflation Factor = 0.9734); (B) Manhattan Plot (No Significant Feature Identified)

A)



B)