Multiomics Data Analysis Identified CpG Sites That Mediate the Impact of Smoking on Cardiometabolic Traits

Majid Nikpay *

Posted Date: 6 May 2023

doi: 10.20944/preprints202305.0409.v1

Keywords: smoking,; DNA methylation; cardiometabolic traits; mendelian randomization

Article

Multiomics Data Analysis Identified CpG Sites that Mediate the Impact of Smoking on Cardiometabolic Traits

Majid Nikpay

Omics and Biomedical Analysis Core Facility, University of Ottawa Heart Institute, Ottawa, K1Y 4W7, Canada; mnikpay@ottawaheart.ca

Abstract: Understanding epigenome paths through which smoking contributes to cardiometabolic traits is important for downstream applications. In this study, I used a SNP-based analytical pipeline to integrate several publicly available datasets and identify CpG sites that mediate the impact of smoking on cardiometabolic traits. After applying stringent statistical criteria, 11 CpG sites were detected that showed significant association ($P<5e^{-8}$) with cardiometabolic traits at both discovery and replication stage. By integrating eQTL data, I identified genes behind a number of these associations. cg05228408 was hypomethylated in smokers and contributed to higher blood pressure by lowering the expression of *CLCN6* gene. cg08639339 was hypermethylated in smokers and lowered metabolic rate by increasing the expression of *RAB29*; furthermore, I noted *TMEM120A* mediated the impact of smoking-cg17325771 on LDL, and *LTBP3* mediated the smoking-cg07029024 effect on heart rate. This study provides a list of CpG sites that mediates the impact of smoking on cardiometabolic traits and a framework to investigate epigenome path through which a lifestyle habit modifies disease risks.

Keywords: smoking; DNA methylation; cardiometabolic traits; mendelian randomization

Introduction

Lifestyle choices could modify the risk of diseases partly through their impacts on the epigenome which are genomic sites that the interaction of genetics and environmental factors happen. Regular tobacco smoking is known to impact a number of phenotype conditions including cardiometabolic traits. The aim of this study was to investigate CpG methylation sites through which tobacco smoking impacts cardiometabolic traits. With the advancement of high-throughput screening methods, previous studies have already identified CpG sites that show differential levels of methylation in smokers as compared to non-smokers. Genome-wide association studies (GWAS) also provided a comprehensive catalogue of biological entities (traits, biomarkers, …) and their underlying SNPs; meanwhile, analytical tools have been developed that can infer the relation between two entities using the knowledge available at SNP level.[1,2] Motivated by these developments, in this study, I devised an analytical pipeline (Figure 1) to integrate the previous findings in order to investigate epigenome paths through which smoking contributes to cardiometabolic traits.

Findings from such studies could have different applications. First, many complex phenotypes including cardiometabolic traits progress gradually over the time until they pass the liability threshold point and become diseases. As such, epigenomic biomarkers can greatly benefit preventive medicine, because they allow the health practitioners to detect the early presence of a disorder and monitor its condition over time. This is notable because epigenomic changes are reversible by changing the lifestyle. Second, at the molecular level, designing medications for every macromolecule (protein, metabolite, …) is not straightforward; however, with the development of CRISPER-based epigenome editing system, targeting the epigenome sites underlying a trait could be a universal

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 6 May 2023 doi:10.20944/preprints202305.0409.v1

therapeutic solution applicable to various diseases. Finally, understanding the molecular path through which a lifestyle habit causes a disease is important for biological insight and downstream research.

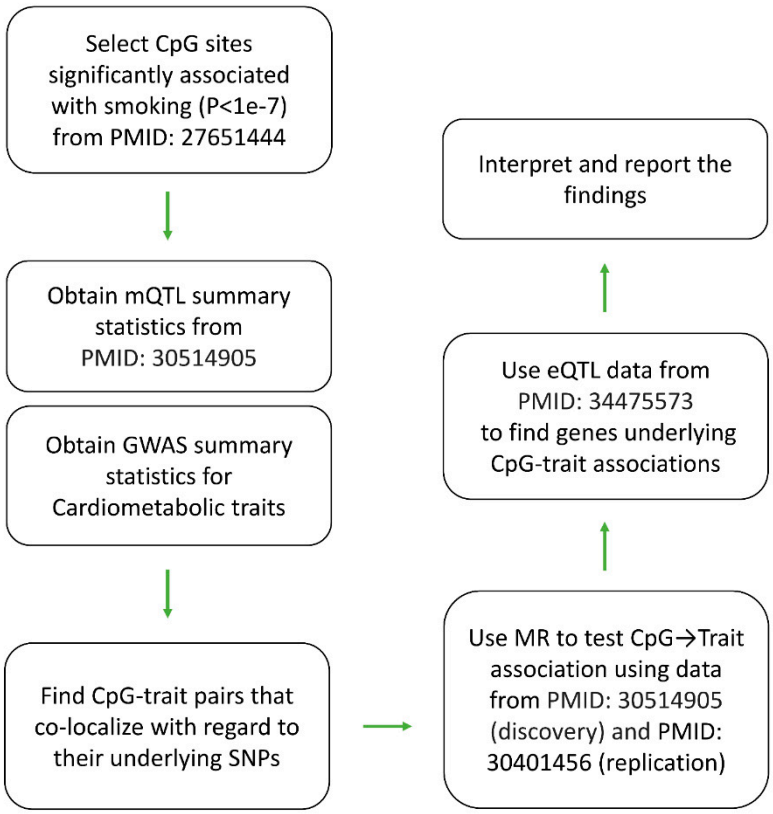

Figure 1. Overview of the analyses undertaken in this study to find CpG sites that mediate the impact of smoking on cardiometabolic traits. This study was conducted using data obtained from previous studies. The list of CpG sites that showed differential levels of methylation between smokers and never-smokers (at Bonferroni correction level, $P<1e^{-8}$) was obtained from PMID: 27651444. Then, a colocalization analysis was performed using GWAS summary statistics to select CpG-trait pairs for MR analysis. We used a two-stages discovery and replication design to identify CpG sites that significantly contribute to a cardiometabolic trait (P<5e-8) through MR analysis. Finally, eQTL data from the eQTLGen consortium were integrated to investigate genes that mediate the effect of a CpG site on a trait.

The nature of association between smoking, epigenome and cardiometabolic traits has been the subject of a number of studies.[3–5] However, such studies normally measure DNA methylation and traits in the same group of subjects. Such a design does not differentiate between causation, and reverse causation. Furthermore, limited sample sizes hinder the power of such studies. Here, an analytical pipeline was used that relies on the concept of Mendelian randomization and allows integrating data from large GWAS consortia. In the Methods section, I detailed the approach.

Methods

Data sources:

Previously Joehanes et al.[6] conducted a meta-analysis of genome-wide DNA methylation using DNA samples derived from the blood of 9,389 participants (2,433 current smokers and 6,956 never smokers). The authors identified 2,623 CpG sites that showed differential levels of methylation between smokers and never-smokers at Bonferroni threshold of $P<1e^{-7}$. In this study, I chose these

 doi:10.20944/preprints202305.0409.v1

sites and examined their contribution to cardiometabolic traits through the analytical pipeline described in Figure 1.

Association with cardiometabolic traits:

Initially, co-localization analysis was performed using the SMR program[1] to identify genomic-regions where GWAS signal for a CpG site and a trait overlap. Following this stage, CpG-trait pairs with colocalization $P<5e^{-8}$ were selected. Next, Mendelian randomization was used to investigate CpG sites that are causally contributing to cardiometabolic traits.

Mendelian randomization (MR) is a statistical method that can investigate the nature of association between two biological entities by comparing their magnitude of associations to the same set of reference SNPs. Because the assortment of SNP alleles in offsprings is a random process (unaffected by confounding environmental factors). Therefore, a set of independent SNPs can be used to investigate the nature of association between two entities. In this study, I used the GSMR algorithm implemented in GCTA software[2] to conduct MR analysis. SNPs that are used for MR analysis must pass a number of criteria to prevent bias in the assessment. First, they must not be in linkage disequilibrium. In this study, SNPs that the extent of LD among them do not exceed, $r2<0.2$ (based on genotype data available from the 1000 genomes on subjects of European ancestry) were included into the analysis. Second, they must not show a pleiotropic effect (i.e., exposure ← SNP → outcome). Such SNPs were excluded from the instrument using the HEIDI test ($P<0.01$) implemented in the GSMR algorithm. Third, they must be significantly associated with the exposure. For this purpose, I chose SNPs that were associated with the exposure (CpG site) at GWAS significance level ($P<5e^{-8}$).

mQTL summary statistics from McRae et al.[7] were used to conduct co-localization analysis and to find CpG sites that are causally contributing to a trait. The authors used the Infinium HumanMethylation 450 BeadChip (Illumina, CA, USA) to measure DNA methylation in blood samples taken from 1980 subjects of European descent. GWAS summary statistics for cardiometabolic traits were obtained from the OpenGWAS database[8] by considering the studies conducted using samples from the European population. This consideration is important to prevent bias due to the population stratification.

CpG sites that were causally contributing to a trait ($P<5e^{-8}$) and did not show evidence of reverse causation ($P<0.05$) were selected and their contributions were assessed once more using mQTL data from Hannon et al.[9,10] Finally, I selected CpG-trait pairs with $P<5e^{-8}$ and integrated eQTL data from the eQTLGen consortium[11] to identify genes that mediate the impact of methylation sites on traits.

Results

By choosing CpG sites that showed differential levels of methylation between smokers and never-smokers in Joehanes et al. study[6], I examined their contribution to cardiometabolic traits through the SNP-based analytical pipeline described in Figure 1. After applying rigorous statistical criteria, 11 CpG sites were identified that co-localized with cardiometabolic traits (Table S1) and showed significant evidence of causal contribution ($P<5e^{-8}$) at both discovery and replication stages (Table 1). The description of CpG sites and their nature of association with smoking is provided in Table S2. By inspecting data from the EWAS atlas[12] which is a repository of trait-epigenome modifications, I found confirmatory evidence from other studies with regard to the association of the identified CpG sites with smoking (Table S3). Next, eQTL data from the eQTLGen consortium[11] were integrated to investigate genes that may mediate the impact of CpG sites on the traits. In the following sections, I review the notable findings:

Table 1. CpG sites that mediate the impact of smoking on cardiometabolic traits.

Probe	Trait		Discovery				Replication			
	Description	OpenGWAS ID	Beta	SE	P-value	NSNPs	Beta	SE	P-value	NSNPs
cg05228408	Hypertension	ukb-a-61	-0.03	0.002875	2.3E-20	6	-0.55	0.055627	3.8E-23	4
cg02998240	Low density lipoprotein	ebi-a-GCST90002412	-0.02	0.002364	1.9E-21	17	-0.29	0.030195	2.2E-21	8
cg01465596	Systolic blood pressure	ukb-a-360	-0.03	0.004722	3.3E-11	6	-0.53	0.084893	6.0E-10	5
cg08639339	Basal metabolic rate	ukb-a-268	-0.02	0.002418	4.1E-11	17	-0.32	0.045082	1.5E-12	15
cg27526649	Pulse rate	ukb-a-363	-0.48	0.059679	4.8E-16	13	-7.93	0.965334	2.1E-16	8
cg10676309	Basal metabolic rate	ukb-a-268	-0.03	0.004693	8.2E-12	3	-0.86	0.1373	3.5E-10	3
cg11105358	Immune reaction	ukb-b-17241	-0.01	0.001132	3.0E-10	13	-0.21	0.033844	3.6E-10	3
cg05789250	Systolic blood pressure	ukb-a-360	-0.03	0.004582	1.5E-09	7	-0.82	0.145834	1.7E-08	3
cg12583553	Basal metabolic rate	ukb-a-268	-0.02	0.003135	4.8E-09	10	-0.31	0.049638	5.7E-10	6
cg12583553	Body fat percentage	ukb-b-8909	-0.02	0.0034	1.6E-10	9	-0.32	0.051055	2.1E-10	4
cg17325771	Low density lipoprotein	ebi-a-GCST90002412	-0.03	0.003576	6.9E-14	7	-0.74	0.092528	1.2E-15	3
cg07029024	Pulse rate	ukb-b-18103	0.03	0.004262	1.5E-09	7	0.39	0.061896	3.0E-10	3

doi:10.20944/preprints202305.0409.v1

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 6 May 2023 doi:10.20944/preprints202305.0409.v1

cg05228408, CLCN6

AGTRAP-PLOD1 is a well-established locus for hypertension.[13,14] Within this locus, I found cg05228408 undergoes hypomethylation as a result of smoking (B=-0.01, P=6.4e^{-10}, Table S1) and consequently increases the risk of hypertension (Table 1). Several genes are located within *AGTRAP-PLOD1* locus that are to varying degrees implicated in hypertension. By integrating the eQTL data, I noted GWAS signals for hypertension and cg05228408 overlap with eQTLs for *CLCN6* (Figure 2). The outcome of MR analysis was also consistent. Namely, higher methylation at cg05228408 site was associated with higher levels of *CLCN6* (B=0.81, P=3.0e^{-42}, Figure 2) and consequently this lowered the risk of hypertension (B=-0.02, P=2.4e^{-18}, Figure 2). *CLCN6* encodes a protein that acts as a voltage-dependent chloride channel. This protein is primarily localized to late endosomes and functions as a chloride/proton antiporter.

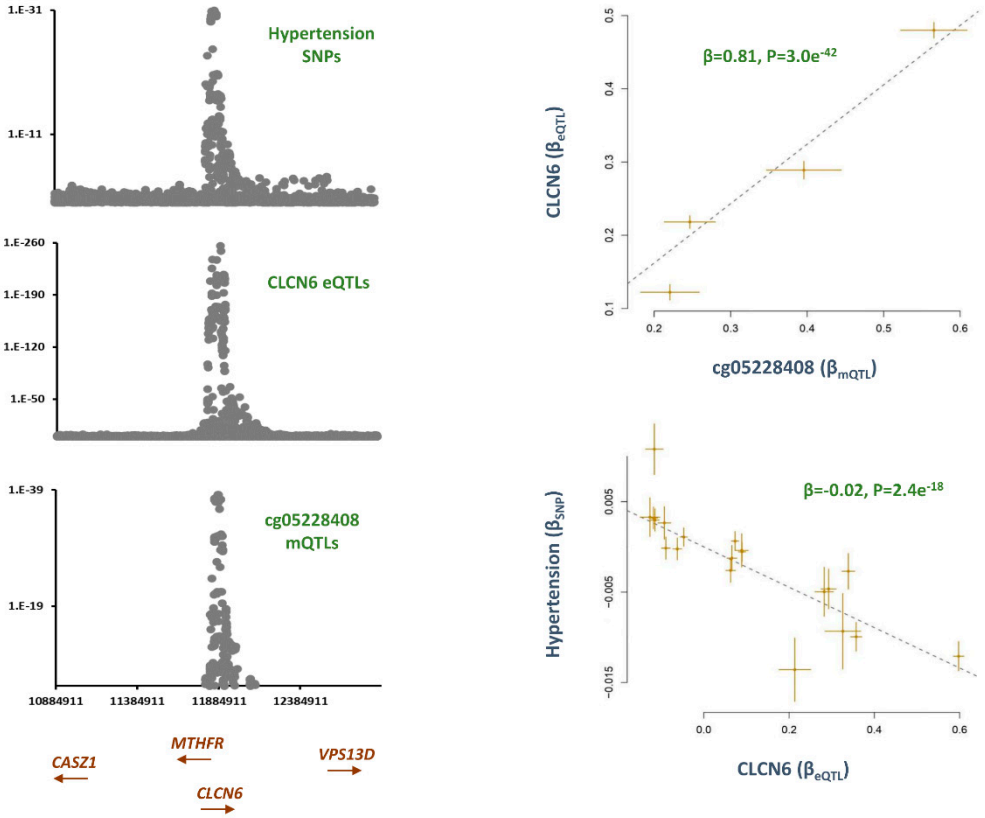

Figure 2. Mechanism whereby cg05228408 mediates the impact of smoking on hypertension. cg05228408 site is reported to be hypomethylated in smokers as compared to never smokers (Tables S2 and S3) (A) We found regional association plots for mQTLs of cg05228408, eQTLs of CLCN6 and risk SNPs of hypertension overlap. (B) MR analysis then revealed as cg05228408 becomes hypermethylated, the expression of CLCN6 increases, and this contributes to lower risk of hypertension in non-smokers. Each point on the MR plots represents a SNP; the x-value of a SNP is its β effect size on the exposure, and the horizontal error bar represents the standard error around the β. The y-value of the SNP is its β effect size on the outcome, and the vertical error bar represents the standard error around its β. The dashed line represents the line of best fit (a line with the intercept of 0 and the slope of β from the MR test).

cg08639339, RAB29

Co-localizaion analysis revealed mQTLs for cg08639339 overlap with SNPs contributing to basal metabolic rate (P_{SMR}=1.1e^{-11}, P_{HEIDI}=0.07, Table S1). The top SNP in this region, rs6673687-T was associated with higher basal metabolic rate (B=0.01, P=3.2e^{-13}) but lower methylation at cg08639339 (B =–0.60, P=1.5e^{-78}, Table S1). Consistently, the MR analysis revealed that higher methylation at this

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 6 May 2023 doi:10.20944/preprints202305.0409.v1

site contributes to lower basal metabolic rate (B=−0.2, P=3.6e^{-10}, Table 1). By investigating the eQTL data, I noted eQTLs for *RAB29* show overlap with mQTLs for cg08639339 and GWAS signals for basal metabolic rate (BMR) (Figure 2). Higher methylation at cg08639339 site contributed to higher expression of *RAB29* (B=0.4, P=1.6e-83) and this lowered BMR (B=-0.03, P=3.1e-13, Figure 3). *RAB29*, formerly known as *RAB29* encodes a protein which is involved in lysosomal trafficking and maintenance.

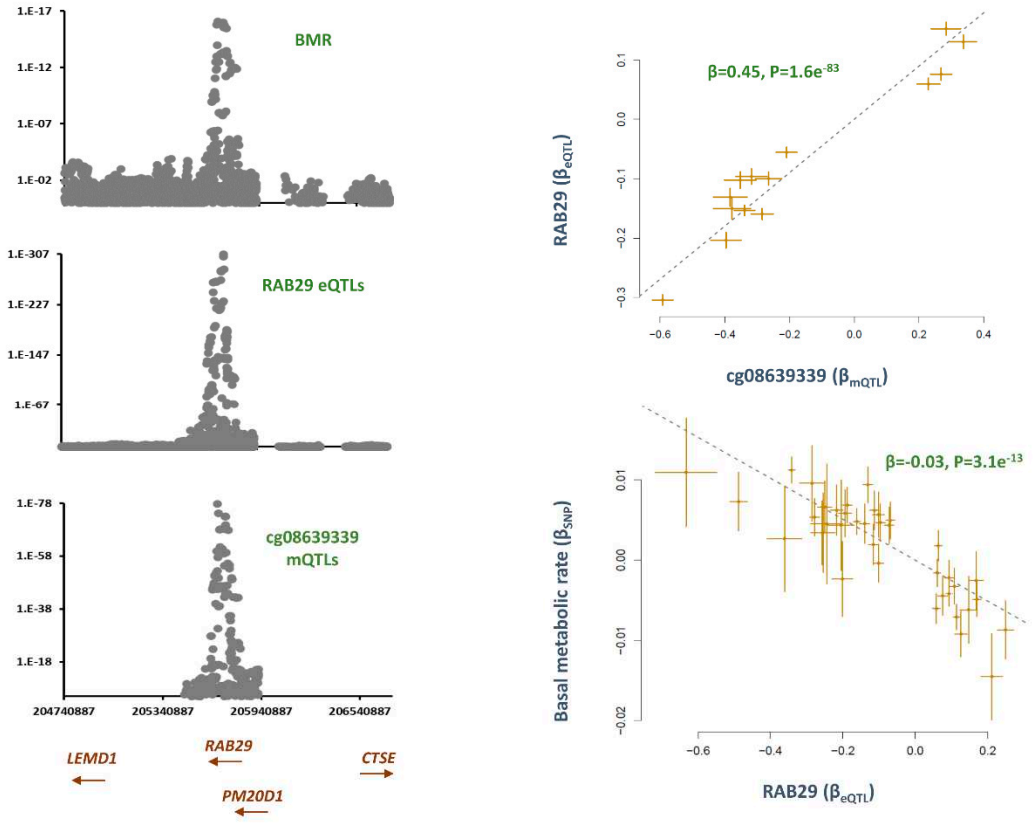

Figure 3. *RAB7L1* mediates the impact of smoking-cg08639339 on basal metabolic rate. We noted cg08639339 is hypermethylated in smokers vs. never-smokers (Tables S2 and S3). MR analysis revealed higher methylation at cg08639339 contributes to higher expression of *RAB7L1*. This consequently contributed to lower basal metabolic rate. Each point on the MR plots represents a SNP; the x-value of a SNP is its β effect size on the exposure, and the horizontal error bar represents the standard error around the β. The y-value of the SNP is its β effect size on the outcome, and the vertical error bar represents the standard error around its β. The dashed line represents the line of best fit (a line with the intercept of 0 and the slope of β from the MR test).

cg17325771, TMEM120A

The methylation site, cg17325771 was hypomethylated (B=-0.01, P=6.5e^{-11}) in smokers as compared to non-smokers. Co-localization analysis revealed mQTLs for cg17325771 overlap with SNPs contributing to LDL (P_{SMR}=3.4e^{-14}, P_{HEIDI}=0.013, Table S1). Subsequently, Mendelian randomization revealed lower methylation at this site contributes to higher LDL levels (B=-0.03, P=6.9e^{-14}, Table 1). By plotting the distribution of eQTLs, I found *TMEM120A* to be the likely gene that mediate the impact of cg1732577 on LDL (Figure 4). The outcome of MR analysis also revealed lower methylation at cg17325771 site is associated with higher expression of *TMEM120A* (B=-0.22, P=1.1e^{-31}) and this consequently contributes to higher LDL level (B=0.09, P=2.3e^{-15}; Figure 4). The protein encoded by *TMEM120A* is a transmembrane protein induced by tumor necrosis factor alpha (TNF-α).

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 6 May 2023 doi:10.20944/preprints202305.0409.v1

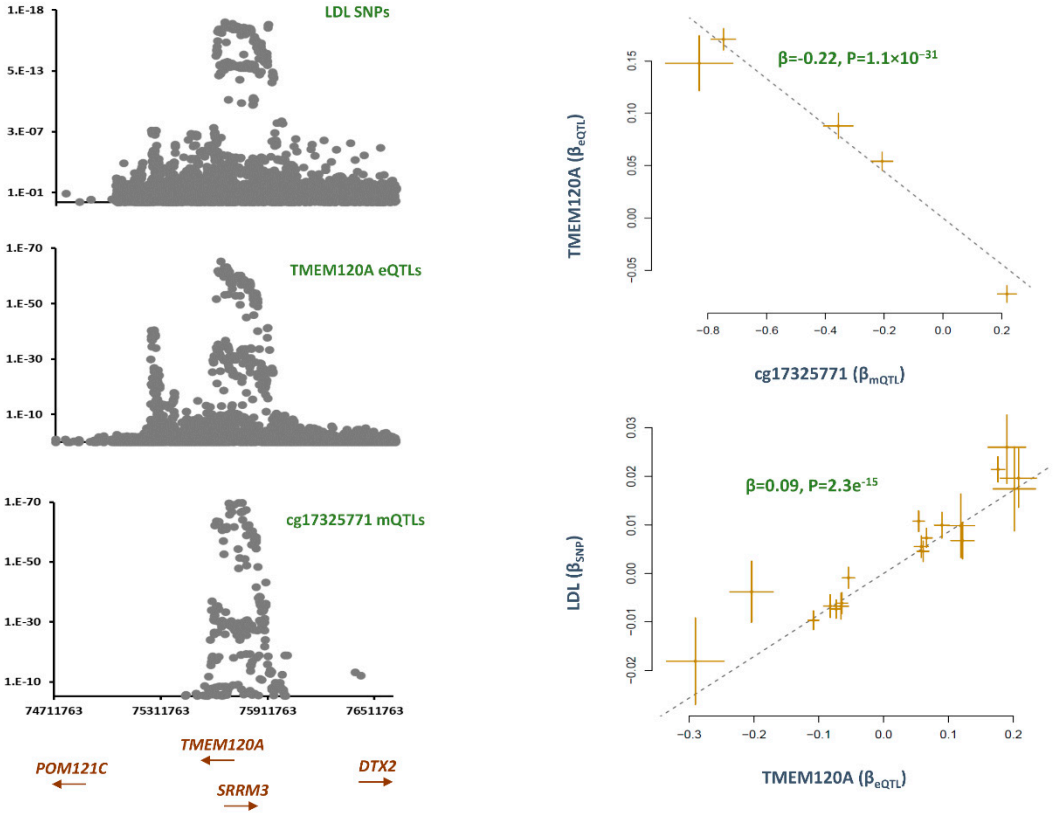

Figure 4. Mechanism whereby cg17325771 mediates the impact of smoking on heart rate. (A) Regional association plots for mQTLs of cg17325771, eQTLs of *TMEM120A* and SNPs for LDL overlapped. (B) MR analysis then revealed as cg17325771 becomes hypomethylated, the expression of *TMEM120A* increases, and this contributes to higher LDL. Each point on the MR plots represents a SNP; the x-value of a SNP is its β effect size on the exposure, and the horizontal error bar represents the standard error around the β. The y-value of the SNP is its β effect size on the outcome, and the vertical error bar represents the standard error around its β. The dashed line represents the line of best fit (a line with the intercept of 0 and the slope of β from the MR test).

cg07029024, LTBP3

Among the identified CpG sites, cg07029024 showed the strongest association with smoking (B=0.01, P=5.5e^{-21}, Table S2). I noted mQTLs for this site, colocalize with SNPs contributing to higher LDL (P_{SMR}=1.3e^{-8}, P_{HEIDI}=0.07, Table S1). Mendelian randomization indicated higher methylation at this site is associated with higher heart rate (B=0.03, P=1.5e^{-9}, Figure S1). The site is located on chromosome band 11q13.1. Among genes in this region, I detected an association between cg07029024 and the expression of *LTBP3*. The outcome of analyses indicated as this site becomes methylated, the expression of *LTBP3* decreases (B=-0.7, P=6.7e^{-21}) and this contributes to higher heart rate (B=-0.04, P=3.4e-14; Figure S1). *LTBP3* encoded protein forms a complex with transforming growth factor beta (TGF-beta) proteins and may be involved in their subcellular localization.

 doi:10.20944/preprints202305.0409.v1

Discussion

Over the past two-decades high throughput studies have provided research community with vast amounts of findings which are continually being added to the databases. In the current time, there are efforts toward joining these data for new insights. This study is another attempt in this direction. Lifestyle habits can predispose or protect us against diseases. At the molecular level, investigating the paths through which such changes happen is important for downstream applications such as early diagnosis and intervention. In the current work, data from several studies were combined in order to investigate epigenome paths through which smoking contributes to cardiometabolic traits.

Using a discovery and replication study design and by setting stringent statistical criteria, I identified 11 CpG sites that mediated the impact of smoking on cardiometabolic traits (Table 1). I found mQTLs for cg05228408, and eQTLs for *CLCN6* show overlap with GWAS signal for hypertension. MR analysis further underlined this finding. I noted as this site becomes hypomethylated (as observed in smokers) the expression of *CLCN6* decreases and this elevates the blood pressure. The role of *CLCN6* in blood pressure regulation is known as previous, GWAS and sequencing studies have linked mutations and variants within this gene to hypertension.[13,14] Recently, Klemens et al. provided functional evidence that *ClCN6* affects blood pressure by regulating golgi calcium reserves which in turn contribute to vascular smooth muscle function.[13] Of note, *CLCN6* is within *AGTRAP-PLOD1* locus which contains several genes implicated in blood pressure regulation such as *MTHFR*, *NPPA*, and *NPPB*; therefore, as underlined earlier[14] further research is required to elucidate the role of this region in blood pressure regulation; however, finding from this study indicates hypermethylating the cg05228408 site could represent a novel therapeutic intervention for lowering the blood pressure. Furthermore, measuring methylation level at this site could represent a biomarker for early diagnosis of hypertension.

The analysis revealed cg08639339 mediates the impact of smoking on basal metabolic rate through *RAB29*. A recent exome-sequencing study found this gene to be associated with cardiometabolic risk in ARIC cohort[15]. The expression of *RAB29* is reported to be upregulated in presence of cholesterol biosynthesis.[16] *RAB29* encoded protein is involved in lysosomal trafficking and maintenance and *RAB29* knock-out mice show lysosomal defects characterized by accumulation of lipids in kidney proximal tubule cells.[17]

TMEM120A is a trans-membrane protein that is known to be expressed in fat tissue and impacts adipogenesis/fat metabolism differentiation.[18,19] *TMEM120A* deficiency is reported to broadly impact lipid metabolism and causes lipodystrophy by altering genome topology.[20] Here, I found the methylation site cg17325771 to mediate the impact of smoking on LDL through this gene.

cg07029024 site showed the strongest association with smoking. By integrating the eQTL data, I found *LTBP3* as the gene that mediates the impact of this site on heart rate. *LTBP3* encodes latent TGF-β binding protein-3 (LTBP-3), which belongs to a family of proteins that regulate TGF-β activity by enabling its secretion, directing it to specific sites in the extracellular matrix, and participating in its activation. It is impact on heart rate could be attributed to its cardiac function. The role of LTBP3-TGF-β signalling in differentiation of cardiac progenitor cells and formation of heart has been researched [21]; besides, *LTBP3* pathogenic variants are reported to predispose individuals to thoracic aortic aneurysms and aortic dissections [22].

The analytical pipeline that was used in this study relies on publicly available data and can be applied to other lifestyle traits. This underlines the value of data sharing by researchers and encourages future studies that aim to catalogue SNPs for less explored functional elements. In the long term, data from such studies will greatly facilitate functional annotations.

In this study, I took a conservative approach to lower the likelihood of false positives. Furthermore, mQTL data came from studies with relatively small sample sizes. Future studies that integrate data from larger consortiums and more dense methylation arrays are expected to provide a more comprehensive picture of epigenomic sites that mediate the impact of lifestyle traits on disease risks. In this regard, reporting trans-regulatory effects are very important, because they appear to be common[23] but often remain unreported by the original QTL studies. In this study, we could not

reveal any functional insight for a number of the identified sites. Therefore, future studies require integrating a more diverse and comprehensive set of data, in order to investigate the mechanism whereby a functional element impacts a phenotype.

In summary, here we combined findings from several studies to identify CpG sites that mediate the impact of smoking on cardiometabolic traits and investigate the underlying genes. This study provides a framework to investigate the molecular paths through which lifestyle habits modify disease risks.

Supplementary Materials: The following supporting information can be downloaded at the website of this paper posted on Preprints.org.

Acknowledgments: This research work was enabled in part by computational resources and support provided by the Compute Ontario and the Compute Canada.

Conflicts of Interest: None

References

1. Zhu, Z., et al., *Integration of summary data from GWAS and eQTL studies predicts complex trait gene targets.* Nature genetics, 2016. **48**(5): p. 481-487.
2. Zhu, Z., et al., *Causal associations between risk factors and common diseases inferred from GWAS summary data.* Nature communications, 2018. **9**(1): p. 224.
3. Ligthart, S., et al., *Tobacco smoking is associated with DNA methylation of diabetes susceptibility genes.* Diabetologia, 2016. **59**(5): p. 998-1006.
4. Maas, S.C., et al., *Smoking-related changes in DNA methylation and gene expression are associated with cardio-metabolic traits.* Clinical Epigenetics, 2020. **12**: p. 1-16.
5. Fragou, D., et al., *Smoking and DNA methylation: Correlation of methylation with smoking behavior and association with diseases and fetus development following prenatal exposure.* Food and Chemical Toxicology, 2019. **129**: p. 312-327.
6. Joehanes, R., et al., *Epigenetic Signatures of Cigarette Smoking.* Circulation: Cardiovascular Genetics, 2016. **9**(5): p. 436-447.
7. McRae, A.F., et al. *Identification of 55,000 Replicated DNA Methylation QTL.* Scientific reports, 2018. **8**, 17605 DOI: 10.1038/s41598-018-35871-w.
8. Elsworth, B., et al., *The MRC IEU OpenGWAS data infrastructure.* bioRxiv, 2020: p. 2020.08.10.244293.
9. Hannon, E., et al., *An integrated genetic-epigenetic analysis of schizophrenia: evidence for co-localization of genetic associations and differential DNA methylation.* Genome Biology, 2016. **17**(1): p. 176.
10. Hannon, E., et al., *Leveraging DNA-Methylation Quantitative-Trait Loci to Characterize the Relationship between Methylomic Variation, Gene Expression, and Complex Traits.* American journal of human genetics, 2018. **103**(5): p. 654-665.
11. Võsa, U., et al., *Large-scale cis- and trans-eQTL analyses identify thousands of genetic loci and polygenic scores that regulate blood gene expression.* Nature genetics, 2021. **53**(9): p. 1300-1310.
12. Li, M., et al., *EWAS Atlas: a curated knowledgebase of epigenome-wide association studies.* Nucleic acids research, 2018. **47**(D1): p. D983-D988.
13. Klemens, C.A., et al., *Loss of Chloride Channel 6 (CLC-6) Affects Vascular Smooth Muscle Contractility and Arterial Stiffness* via *Alterations to Golgi Calcium Stores.* Hypertension, 2021. **77**(2): p. 582-593.
14. Flister, M.J., et al., *Identifying multiple causative genes at a single GWAS locus.* Genome Research, 2013. **23**(12): p. 1996-2002.
15. Feofanova, E.V., et al., *Exome sequence association study of levels and longitudinal change of cardiovascular risk factor phenotypes in European Americans and African Americans from the Atherosclerosis Risk in Communities Study.* Genetic Epidemiology, 2021. **45**(6): p. 651-663.
16. Helip-Wooley, A. and J.G. Thoene, *Sucrose-induced vacuolation results in increased expression of cholesterol biosynthesis and lysosomal genes.* Experimental Cell Research, 2004. **292**(1): p. 89-100.
17. Kuwahara, T., et al., *LRRK2 and RAB7L1 coordinately regulate axonal morphology and lysosome integrity in diverse cellular contexts.* Scientific reports, 2016. **6**(1): p. 29945.
18. Xue, J., et al., *TMEM120A is a coenzyme A-binding membrane protein with structural similarities to ELOVL fatty acid elongase.* eLife, 2021. **10**: p. e71220.

19. Batrakou, D.G., et al., *TMEM120A and B: Nuclear Envelope Transmembrane Proteins Important for Adipocyte Differentiation.* PloS one, 2015. **10**(5): p. e0127712.
20. Czapiewski, R., et al., *Genomic loci mispositioning in Tmem120a knockout mice yields latent lipodystrophy.* Nature communications, 2022. **13**(1): p. 321.
21. Zhou, Y., et al., *Latent TGF-β binding protein 3 identifies a second heart field in zebrafish.* Nature, 2011. **474**(7353): p. 645-648.
22. Guo, D.-c., et al., *LTBP3 Pathogenic Variants Predispose Individuals to Thoracic Aortic Aneurysms and Dissections.* The American Journal of Human Genetics, 2018. **102**(4): p. 706-712.
23. Nikpay, M., S. Ravati, and R. McPherson, *Genome-wide screening identifies DNA methylation sites that regulate the blood proteome.* Epigenomics, 2022. **14**(13): p. 837-848.