

RESEARCH

Open Access



Epigenome-wide association study in Chinese monozygotic twins identifies DNA methylation loci associated with blood pressure

Weijing Wang¹, Jie Yao^{1,2}, Weilong Li³, Yili Wu¹, Haiping Duan⁴, Chunsheng Xu⁴, Xiaocao Tian⁴, Shuxia Li⁵, Qihua Tan⁵ and Dongfeng Zhang^{1*}

Abstract

Background Hypertension is a crucial risk factor for developing cardiovascular disease and reducing life expectancy. We aimed to detect DNA methylation (DNAm) variants potentially related to systolic blood pressure (SBP) and diastolic blood pressure (DBP) by conducting epigenome-wide association studies in 60 and 59 Chinese monozygotic twin pairs, respectively.

Methods Genome-wide DNA methylation profiling in whole blood of twins was performed using Reduced Representation Bisulfite Sequencing, yielding 551,447 raw CpGs. Association between DNAm of single CpG and blood pressure was tested by applying generalized estimation equation. Differentially methylated regions (DMRs) were identified by *comb-P* approach. Inference about Causation through Examination of Familial Confounding was utilized to perform the causal inference. Ontology enrichment analysis was performed using Genomic Regions Enrichment of Annotations Tool. Candidate CpGs were quantified using Sequenom MassARRAY platform in a community population. Weighted gene co-expression network analysis (WGCNA) was conducted using gene expression data.

Results The median age of twins was 52 years (95% range 40, 66). For SBP, 31 top CpGs ($p < 1 \times 10^{-4}$) and 8 DMRs were identified, with several DMRs within *NFATC1*, *CADM2*, *IRX1*, *COL5A1*, and *LRAT*. For DBP, 43 top CpGs ($p < 1 \times 10^{-4}$) and 12 DMRs were identified, with several DMRs within *WNT3A*, *CNOT10*, and *DAB2IP*. Important pathways, such as Notch signaling pathway, p53 pathway by glucose deprivation, and Wnt signaling pathway, were significantly enriched for SBP and DBP. Causal inference analysis suggested that DNAm at top CpGs within *NDE1*, *MYH11*, *SRRM1P2*, and *SMPD4* influenced SBP, while SBP influenced DNAm at CpGs within *TNK2*. DNAm at top CpGs within *WNT3A* influenced DBP, while DBP influenced DNAm at CpGs within *GNA14*. Three CpGs mapped to *WNT3A* and one CpG mapped to *COL5A1* were validated in a community population, with a hypermethylated and hypomethylated direction in hypertension cases, respectively. Gene expression analysis by WGCNA further identified some common genes and enrichment terms.

Conclusion We detect many DNAm variants that may be associated with blood pressure in whole blood, particularly the loci within *WNT3A* and *COL5A1*. Our findings provide new clues to the epigenetic modification underlying hypertension pathogenesis.

Keywords Blood pressure, Causality, DNA methylation, Monozygotic twins

*Correspondence:

Dongfeng Zhang

zhangdf1961@126.com

Full list of author information is available at the end of the article



© The Author(s) 2023. **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>. The Creative Commons Public Domain Dedication waiver (<http://creativecommons.org/publicdomain/zero/1.0/>) applies to the data made available in this article, unless otherwise stated in a credit line to the data.

Introduction

Hypertension is a chronic disease condition, and approximately 90% of the cases are considered as essential hypertension without a definitive cause. The prevalence of hypertension is 28.9%, and the rates of treatment and control of hypertension are only 35.3% and 13.4% in China. Hypertension has been a considerable risk factor for developing cardiovascular disease and reducing life expectancy and has become a huge burden on the nationwide health system [1, 2].

As a complex phenotype, hypertension is controlled by both genetic and environmental factors through the interface of epigenetics. At present, the magnitude of genetic sources of variance in hypertension has been extensively explored. Wu et al. found that blood pressure had a moderate heritability with 0.61 for systolic blood pressure (SBP) and 0.58 for diastolic blood pressure (DBP) in Chinese population [3]. Additionally, genome-wide association studies (GWASs) have reported some genetic variants that were responsible for susceptibility to blood pressure variation, such as the genetic loci within *ADRB1*, *ATP2B1*, *SOX6*, *CHIC2*, *IGFBP3*, and *KCNK3* [4–8]. However, the previous reported genetic variants only partially contributed to the pathogenesis of hypertension.

In recent years, increasingly strong evidence has supported the significant role of epigenetic mechanisms with altered gene expression in the increased susceptibility to diseases. Currently, a large number of epigenome-wide association studies (EWASs) have been conducted to explore the underlying association between genomic DNA methylation (DNAm) variants and complex traits, such as heart failure [9] and permanent atrial fibrillation [10]. Meanwhile, accumulating evidence has also demonstrated a functional role of DNAm variants in the regulation of blood pressure or the development of hypertension [11, 12]. However, to date, very few studies have investigated the blood pressure or hypertension-related DNAm loci by applying an EWAS approach [13–16], and few results are replicated. In addition, the causal nature of the association, i.e., if DNAm exerts a causal effect on blood pressure or vice versa, is unknown. Hence, it is essential to further perform EWAS as well as causal inference analysis to investigate the association and causal relationship between DNAm and blood pressure.

Furthermore, the previous EWASs were most performed using samples from unrelated individuals, where the confounding effects from different genetic backgrounds were not well controlled for. Nowadays, the trait or disease-discordant twin design has been a popular and powerful tool for EWAS while controlling for individual genetic make-up [17, 18]. In this study based on a sample

of blood pressure-discordant Chinese monozygotic twins, we conducted an EWAS to explore the association between the DNAm at CpGs and blood pressure as well as their causality and validated the candidate CpGs in a community population. Additionally, we further integrated the differentially methylated results with gene expression data.

Materials and methods

The primary materials and methods of this study were in accord with our previously published studies [19–22].

Participants

The sample collection was carried out through the Qingdao Twin Registry [23], and details of study recruitment have been previously described [24]. Participants who were pregnant and breastfeeding, who suffered from cardiovascular disease, stroke and/or tumor, and who were regularly taking any medications within one month before participation were excluded. Meanwhile, participants who were unconscious, unable, or unwilling to cooperate were also dropped. Considering that we used trait-discordant monozygotic twin design, the twins with intra-pair blood pressure difference ≥ 2 mmHg for SBP or intra-pair blood pressure difference ≥ 1 mmHg for DBP were separately chosen. A total of 60 SBP-discordant monozygotic twin pairs and 59 DBP-discordant twin pairs were included in the methylation analysis. The median of absolute values of intra-pair blood pressure difference was 18 mmHg (95% range 2, 55) for SBP and 10 mmHg (95% range 2, 28) for DBP, respectively. Additionally, a subsample of 12 monozygotic twin pairs were included in the gene expression analysis. All co-twin pairs completed a questionnaire and undertook a health examination after a 10–12-h overnight fast.

This study was approved by the Regional Ethics Committee of the Qingdao CDC Institutional Review Boards. The ethical principles of the Helsinki Declaration were also followed. Prior written informed consent was achieved.

Zygoty determination

We first identified potential monozygotic and dizygotic twins through sex and ABO blood types. Twins with opposite sex and/or different blood types were classified as dizygotic twins. Then, the zygoty of twins with same sex and blood types was further determined by DNA testing using 16 short tandem repeat markers [23, 25, 26].

Measurement of blood pressure

Blood pressure was measured in a sitting position following standard procedure using a mercurial table stand model sphygmomanometer. SBP was measured

as Korotkoff phase I (appearance of sound) and DBP as Korotkoff phase V (disappearance of sound). Each subject took three repeated measurements, with at least one-minute interval. The mean value of these three measurements was calculated and used in subsequent analysis. All measurements greater than three standard deviations above or below the means were assigned as missing values.

Reduced representation bisulfite sequencing (RRBS) data preparation

The total DNA extracted from whole blood was used in RRBS experiment. Briefly, genomic DNA was first digested to generate short fragments. Then the CpG-rich DNA fragments was bisulfite-converted. Finally, the cDNA library was obtained and sequenced. The raw methylation data covered 551,447 CpGs across the genome of each individual. We mapped the raw sequencing data to the human GRCh37 by *Bismark* [27] and then imported data to *BiSeq* to smooth the methylation level [28]. We controlled the coverage to 90% quantile and dropped CpGs with average methylation β -values less than 0.01 or more than 10 missing observations. After quality control, a total of 248,262 CpGs for SBP and 248,955 CpGs for DBP remained for subsequent analyses. The methylation β -value was transformed to M -value by applying \log_2 transformation.

Since total DNA was extracted from whole blood, different methylation profiles of distinct cell-types may lead to false discoveries [29]. In our analysis, we applied *ReFACTor* method, a reference-free method to account for cell-type heterogeneity, and we used the top five components to correct for the cell-type composition effect on DNAm [30].

Gene expression data preparation

Briefly, the total mRNA was first extracted from whole blood. Subsequently, the RNA-Seq library was constructed and sequenced to obtain the sequenced data, which was then mapped to the human genome by *TopHat2* [31]. The gene expression level was evaluated by FPKM value through *Cufflinks* software [32].

Epigenome-wide association analysis

The association between the DNAm M -value at each CpG and SBP or DBP was tested by using generalized estimating equation (GEE) approach through the *geeglm* function in R-package *geepack*, adjusting for age, sex, and cell-type composition. Moreover, in order to address the paired structure of the twin data, we included a vector which identified the clusters of twins within a pair into the GEE model. To correct for multiple testing, we calculated false discovery rate (FDR) [33] and defined

FDR < 0.05 as genome-wide significance. For CpGs with FDR ≥ 0.05 , we defined $p < 1 \times 10^{-6}$ as suggestive significance and $1 \times 10^{-6} \leq p < 1 \times 10^{-5}$ as weaker-than-suggestive significance [34]. The CpGs with $p < 1 \times 10^{-4}$ were reported as top CpGs of this EWAS [35]. The identified genomic CpGs ($p < 0.05$) were annotated to the nearest genes by using R-package *biomaRt* [36, 37].

Causal inference analysis

For the top CpGs ($p < 1 \times 10^{-4}$), the causal relationship with blood pressure was investigated by the Inference about Causation through Examination of Familial Confounding (ICE FALCON) method which was a regression based method for causal inference in association studies using twins or family data [38, 39]. In this method, ‘familial’ meant both genetic and shared environmental factors in twins, which was essential to make explicit causal inference. The GEE model was applied for parameter estimation with correction for twin pairing. Estimations of β_{self} , $\beta_{\text{co-twin}}$ as well as β'_{self} and $\beta'_{\text{co-twin}}$ were calculated, where β_{self} was the estimation of overall correlation including the causal proportion and family confounding proportion, $\beta_{\text{co-twin}}$ estimated only the family confounding proportion of the correlation, and β'_{self} and $\beta'_{\text{co-twin}}$ was the estimation of full model. If $|\beta_{\text{co-twin}} - \beta'_{\text{co-twin}}|$ was similar to $|\beta_{\text{self}} - \beta'_{\text{self}}|$, then the association was due to family confounding; and if $|\beta_{\text{co-twin}} - \beta'_{\text{co-twin}}|$ was much larger than $|\beta_{\text{self}} - \beta'_{\text{self}}|$ (ratio > 1.5), then it indicated a causal effect.

Region-based analysis

We applied the *comb-p* approach to detect the blood pressure-associated differentially methylated regions (DMRs) [40]. The significant enriched DMRs were determined by Stouffer–Liptak–Kechris (*slk*) corrected $p < 0.05$.

Ontology enrichments analysis

We submitted the identified CpGs ($p < 0.05$) to the Genomic Regions Enrichment of Annotations Tool (GREAT) online to analyze the ontology enrichments [41]. Annotation was based on the human GRCh37, and the default ‘‘basal plus extension’’ association rule was used. The false discovery rate (FDR) < 0.05 was considered as statistically significant in ontology enrichments analysis.

EWAS power estimation

We have recently published a computer simulation study on the power of EWAS using twin design [17]. According to this study, if one trait/disease had a heritability (h^2) of 0.6 and there was a low correlation between environmental factors and DNAm ($R^2_{M,E} = 0.1$), the sample size

required for statistical power to exceed 80% in trait or disease-discordant twin design ranged from 22 (when the correlation within twin pair due to either shared genetic background or common environment, denoted as $\rho_\varepsilon=0.8$) to 63 (when $\rho_\varepsilon=0.1$) pairs, which was an immense improvement over the ordinary case-control design. Hence, we speculated that our study based on nearly 60 twin pairs would get a statistical power of about 80%.

We also estimated the correlation between environmental factors (i.e., blood pressure) and DNAm based on the top CpGs identified in this EWAS. We tested the correlation between intra-pair blood pressure difference and intra-pair DNAm difference of each top CpG in EWAS by using partial correlation analysis model, adjusting for age and sex. The median of absolute values of partial correlation coefficients was 0.34 (range 0.03, 0.47) for SBP and 0.27 (range 0.04, 0.46) for DBP (Additional file 1: Table S1), indicating that the $R^2_{M,E}$ of our study was likely to be greater than 0.1 and close to 0.3. The heritability of SBP and DBP was about 0.60 in the same twin population as our study [3]. According to the computer simulation study [17], for SBP and DBP with $h^2=0.6$ and $R^2_{M,E}=0.3$, the sample size required for statistical power to exceed 80% in our twin design would range from 17 (when $\rho_\varepsilon=0.8$) to 25 (when $\rho_\varepsilon=0.1$) pairs, which were much less than 60 pairs. Hence, our study based on nearly 60 twin pairs would get an enough statistical power.

Quantitative methylation analysis of COL5A1 and WNT3A

We randomly recruited 118 hypertension cases and 149 health controls from the community to validate the CpGs mapped to *COL5A1* and *WNT3A* in EWAS. The cases were defined as those with SBP ≥ 140 mmHg and DBP ≥ 90 mmHg. The subjects with a history of diabetes, obesity, cancer, stroke, and cardiovascular disease were excluded. The participants were interviewed when blood samples were taken and stored under -80°C for DNA methylation analysis. We designed the primers for *COL5A1* and *WNT3A* gene to cover the region with the most CpGs ($p < 0.05$) in EWAS. The mass spectra of cleavage products were collected using the MALDI-TOF mass spectrometry based on the MassARRAY System (Bio Miao Biological Technology, Beijing, China), and the spectra's methylation ratio was generated by MassARRAY EpiTYPER software (Agena Bioscience, San Diego, California). The DNAm of CpGs between the two independent groups was compared by Wilcoxon rank-sum test. Binary logistic regression model was applied to evaluate the association of each CpG with hypertension while adjusting for BMI, triglyceride (TG), and fasting blood glucose (FBG). The $p < 0.05$ was set as statistically significant.

Weighted gene co-expression network analysis (WGCNA)

We conducted the weighted gene co-expression network analysis (WGCNA) [42, 43] to identify the specific modules and genes potentially associated with blood pressure. Briefly, a weighted adjacency matrix using gene expression profile data was established, and then, a topological overlap matrix was constructed and used as input for hierarchical clustering analysis. Gene modules were detected by the dynamic tree cutting algorithm, and module eigengenes were correlated with SBP or DBP to detect the module of interest. Enrichment analysis was conducted for the genes clustered in the specific module by DAVID tool [44, 45]. The significant enriched terms were identified with $p < 0.05$ from a modified Fisher's exact test.

Results

Epigenome-wide association analysis

A total of 60 twin pairs with a median value of 134.00 mmHg (95% range 102.05, 184.90) for SBP and 59 twin pairs with a median value of 80.00 mmHg (95% range 62.00, 105.03) for DBP were included in the methylation analysis. The median age of twins was 52 years (95% range 40, 66). The other clinical indicators, i.e., BMI, serum uric acid, FBG, high-density lipoprotein cholesterol (HDL), low-density lipoprotein cholesterol (LDL), and TG, showed statistically intra-pair correlated, indicating the co-twin design beneficial (Additional file 2: Table S2).

The Manhattan plot of EWAS on SBP is shown in Additional file 3: Fig. S1 (a), and we identified 31 SBP-related top CpGs with $p < 1 \times 10^{-4}$ (Table 1). After correcting for multiple testing, no CpG reached genome-wide significance as defined by FDR < 0.05 . The four strongest associations ($\beta = -0.01$, $p = 5.76 \times 10^{-6}$ – 9.58×10^{-6}) were detected for the CpGs (chr3: 84,330,415–84,330,441 bp) located at *SRRMIP2*, showing weaker-than-suggestive significance ($1 \times 10^{-6} \leq p < 1 \times 10^{-5}$). All these top CpGs were located at/near 15 genes, including *SRRMIP2*, *COL5A1*, *MIR1268A*, *NFATC1*, *NDE1*, *MYH11*, *SMPD4*, *TXNL1P1*, *MIR3147*, *PIP5K1C*, *TNK2*, *CACHD1*, *SLC47A1*, etc.

The association between DNAm of 43 top CpGs and DBP reached $p < 1 \times 10^{-4}$ level (Additional file 3: Fig. S1 (b) and Table 2). There were four CpGs (chr1: 228,195,260–228,195,292 bp) within *WNT3A* and one CpG (chr1: 2,391,479 bp) within *PLCH2* detected as showing genome-wide significance (FDR < 0.05). Seven CpGs within *SIMI*, *PLCH2*, *ATXN7L3B*, and *LOC646588* showed weaker-than-suggestive significance with $1 \times 10^{-6} \leq p < 1 \times 10^{-5}$. All the top CpGs were located at/near 16 genes, and there was more than one CpG located

Table 1 The results of epigenome-wide association study on systolic blood pressure ($p < 1 \times 10^{-4}$)

Chromosome	Position (bp)	Coefficient	p-value	FDR	Ensembl gene ID	HGNC symbol
chr3	84,330,432	-0.009	5.756E-06	0.161	ENSG00000242195	<i>SRRM1P2</i>
chr3	84,330,437	-0.009	7.677E-06	0.161	ENSG00000242195	<i>SRRM1P2</i>
chr3	84,330,415	-0.009	8.172E-06	0.161	ENSG00000242195	<i>SRRM1P2</i>
chr3	84,330,441	-0.009	9.579E-06	0.161	ENSG00000242195	<i>SRRM1P2</i>
chr3	84,330,448	-0.009	1.248E-05	0.161	ENSG00000242195	<i>SRRM1P2</i>
chr17	58,216,280	-0.008	1.285E-05	0.161	ENSG00000267095	NA
chr7	57,472,878	0.013	1.662E-05	0.196	ENSG00000266168	<i>MIR3147</i>
chr8*	9,260,932	-0.076	2.635E-05	0.267	ENSG00000254235	NA
					ENSG00000254237	NA
chr17	58,216,262	-0.008	2.687E-05	0.267	ENSG00000267095	NA
chr3	84,330,462	-0.008	2.873E-05	0.272	ENSG00000242195	<i>SRRM1P2</i>
chr9	137,673,895	-0.009	2.971E-05	0.272	ENSG00000130635	<i>COL5A1</i>
chr9	137,673,907	-0.009	3.066E-05	0.272	ENSG00000130635	<i>COL5A1</i>
chr16*	15,814,807	-0.048	3.425E-05	0.286	ENSG00000072864	<i>NDE1</i>
					ENSG00000133392	<i>MYH11</i>
chr13	87,444,790	-0.011	3.545E-05	0.286	ENSG00000231879	<i>TXNL1P1</i>
chr18	77,269,485	0.011	4.124E-05	0.320	ENSG00000131196	<i>NFATC1</i>
chr9	137,673,888	-0.009	4.699E-05	0.341	ENSG00000130635	<i>COL5A1</i>
chr13	87,444,783	-0.011	5.913E-05	0.366	ENSG00000231879	<i>TXNL1P1</i>
chr15	22,545,461	0.007	5.953E-05	0.366	ENSG00000221641	<i>MIR1268A</i>
chr2	130,937,909	0.059	6.043E-05	0.366	ENSG00000136699	<i>SMPD4</i>
chr15	22,545,464	0.007	6.663E-05	0.384	ENSG00000221641	<i>MIR1268A</i>
chr8*	9,260,942	-0.071	6.811E-05	0.384	ENSG00000254235	NA
					ENSG00000254237	NA
chr19	3,670,396	0.010	7.317E-05	0.399	ENSG00000186111	<i>PIP5K1C</i>
chr3	195,609,985	0.050	7.510E-05	0.399	ENSG00000061938	<i>TNK2</i>
chr2	130,937,907	0.055	7.555E-05	0.399	ENSG00000136699	<i>SMPD4</i>
chr1	64,880,619	-0.108	7.798E-05	0.403	ENSG00000158966	<i>CACHD1</i>
chr17	19,436,923	-0.028	8.193E-05	0.415	ENSG00000142494	<i>SLC47A1</i>
chr18	77,269,508	0.012	8.697E-05	0.428	ENSG00000131196	<i>NFATC1</i>
chr16*	15,814,759	-0.056	8.814E-05	0.428	ENSG00000072864	<i>NDE1</i>
					ENSG00000133392	<i>MYH11</i>
chr18	77,269,476	0.011	8.970E-05	0.428	ENSG00000131196	<i>NFATC1</i>
chr15	22,545,472	0.008	9.480E-05	0.444	ENSG00000221641	<i>MIR1268A</i>
chr17	62,775,172	0.012	9.845E-05	0.450	ENSG00000215769	<i>ARHGAP27P1-BPTFP1-KPNA2P3</i>

NA, not available; FDR, false discovery rate

* The CpG sites annotated to two genes

at/near genes *ATXN7L3B*, *DAB2IP*, *WNT3A*, *GNA14*, *EYS*, *KCNT1*, *LOC646588*, *MGEA5*, *PGR*, *PLCH2*, *SAE1*, and *SIM1*.

We found 21 common CpGs ($p < 1 \times 10^{-3}$) between SBP and DBP, and these CpGs were annotated at genes *CACNA1B*, *LARP4B*, *CSNK1G2*, *LOC646588*, *HES4*, *PPIAP45*, *GPX1*, *METRNL*, *ROBO3*, and *LINC00943*.

We also compared previously reported significant blood pressure or hypertension-associated differentially

methylated genes in EWASs [13–15] with our results. We defined our genes where CpGs with $p < 0.05$ were located as supportive to the reported results. The genes *CDC42BPB*, *ALDH3B2*, *DAB2IP*, *SLC7A5*, *VPS25*, *SLC43A1*, *SKOR2*, *ATXN1*, *ZMIZ1*, and *CPT1A* for SBP and *MAN2A2*, *CFLAR*, *CPT1A*, *DAB2IP*, *SLC7A5*, *PHGDH*, *SKOR2*, and *ZMIZ1* for DBP could be replicated (Additional file 4: Table S3).

Table 2 The results of epigenome-wide association study on diastolic blood pressure ($p < 1 \times 10^{-4}$)

Chromosome	Position (bp)	Coefficient	p-value	FDR	Ensembl gene ID	HGNC symbol
chr1	228,195,277	0.028	5.764E-08	0.010	ENSG00000154342	WNT3A
chr1	228,195,289	0.029	1.291E-07	0.010	ENSG00000154342	WNT3A
chr1	2,391,479	-0.020	1.540E-07	0.010	ENSG00000149527	PLCH2
chr1	228,195,292	0.029	1.633E-07	0.010	ENSG00000154342	WNT3A
chr1	228,195,260	0.029	2.857E-07	0.014	ENSG00000154342	WNT3A
chr6	100,909,431	0.026	2.450E-06	0.090	ENSG00000112246	SIM1
chr1	2,391,466	-0.018	2.541E-06	0.090	ENSG00000149527	PLCH2
chr12	74,797,036	-0.047	4.032E-06	0.125	ENSG00000253719	ATXN7L3B
chr6	100,909,425	0.025	5.889E-06	0.163	ENSG00000112246	SIM1
chr12	74,797,017	-0.044	7.783E-06	0.185	ENSG00000253719	ATXN7L3B
chr7	25,898,451	0.020	8.166E-06	0.185	ENSG00000223561	LOC646588
chr12	74,797,049	-0.037	9.188E-06	0.191	ENSG00000253719	ATXN7L3B
chr7	25,898,447	0.020	1.231E-05	0.236	ENSG00000223561	LOC646588
chr17	38,088,968	0.185	1.593E-05	0.257	ENSG00000264968	NA
chr6	66,373,850	0.055	1.624E-05	0.257	ENSG00000188107	EYS
chr12	74,797,053	-0.035	1.720E-05	0.257	ENSG00000253719	ATXN7L3B
chr9	80,272,835	-0.068	1.752E-05	0.257	ENSG00000156049	GNA14
chr1	228,195,243	0.029	1.988E-05	0.273	ENSG00000154342	WNT3A
chr9	80,272,842	-0.066	2.086E-05	0.273	ENSG00000156049	GNA14
chr9	80,272,845	-0.066	2.338E-05	0.288	ENSG00000156049	GNA14
chr6	66,373,857	0.053	2.469E-05	0.288	ENSG00000188107	EYS
chr9	80,272,847	-0.065	2.663E-05	0.288	ENSG00000156049	GNA14
chr12	74,797,056	-0.034	2.825E-05	0.293	ENSG00000253719	ATXN7L3B
chr9	138,637,356	-0.013	4.158E-05	0.400	ENSG00000107147	KCNT1
chr19	35,324,068	0.168	4.340E-05	0.400	ENSG00000267767	LINC01801
chr19	47,635,288	-0.021	4.779E-05	0.425	ENSG00000142230	SAE1
chr9	138,637,337	-0.011	5.604E-05	0.462	ENSG00000107147	KCNT1
chr12	74,796,990	-0.040	5.764E-05	0.462	ENSG00000253719	ATXN7L3B
chr17	38,088,944	0.171	5.910E-05	0.462	ENSG00000264968	NA
chr9	124,308,134	0.014	6.588E-05	0.462	ENSG00000136848	DAB2IP
chr9	124,308,131	0.014	6.979E-05	0.462	ENSG00000136848	DAB2IP
chr19	47,635,313	-0.017	7.074E-05	0.462	ENSG00000142230	SAE1
chr5	30,864,593	0.145	7.349E-05	0.462	ENSG00000241668	RPL19P11
chr9	124,308,128	0.014	7.539E-05	0.462	ENSG00000136848	DAB2IP
chr11	100,999,098	0.026	8.179E-05	0.462	ENSG00000082175	PGR
chr16	8,619,841	0.014	8.388E-05	0.462	ENSG00000232258	TMEM114
chr9	124,308,155	0.013	8.503E-05	0.462	ENSG00000136848	DAB2IP
chr10	103,551,798	0.150	8.742E-05	0.462	ENSG00000198408	OGA
chr17	38,088,933	0.169	8.800E-05	0.462	ENSG00000264968	NA
chr11	100,999,104	0.025	8.815E-05	0.462	ENSG00000082175	PGR
chr9	124,308,115	0.014	9.133E-05	0.462	ENSG00000136848	DAB2IP
chr9	124,308,162	0.012	9.381E-05	0.462	ENSG00000136848	DAB2IP
chr10	103,551,806	0.148	9.390E-05	0.462	ENSG00000198408	OGA

NA, not available; FDR, false discovery rate

Table 3 Results of causal inference analysis for systolic blood pressure and diastolic blood pressure

CpG No.	Chromosome	Position	HGNC symbol	Methylation to blood pressure				Blood pressure to methylation					
				β_{self_change}	P_{self_change}	$\beta_{co-twin_change}$	$P_{co-twin_change}$	Absolute value of ratio	β_{self_change}	P_{self_change}	$\beta_{co-twin_change}$	$P_{co-twin_change}$	Absolute value of ratio
<i>SBP</i>													
CpG1	chr3	84,330,462	SRRM1P2	0.640	0.247	-2.869	0.046	4.483	-0.001	0.307	0.001	0.198	-
CpG2	chr16	15,814,807	NDE1, MYH11	0.086	0.287	-0.355	<0.001	4.136	-0.004	0.405	0.008	0.210	-
CpG3	chr13	87,444,790	TXNLP1	0.628	0.212	-2.764	0.001	4.403	-0.001	0.208	0.002	0.035	1.746
CpG4	chr13	87,444,783	TXNLP1	0.540	0.288	-2.704	0.002	5.006	-0.001	0.240	0.002	0.045	1.820
CpG5	chr2	130,937,909	SMPD4	0.080	0.041	0.290	<0.001	3.650	0.004	0.456	-0.017	0.006	4.545
CpG6	chr3	195,609,985	TNK2	0.209	0.041	0.322	0.006	1.538	-0.003	0.602	-0.012	0.062	-
CpG7	chr2	130,937,907	SMPD4	0.132	0.005	0.359	<0.001	2.728	0.001	0.789	-0.017	0.002	12.877
CpG8	chr1	64,880,619	CACHD1	-0.091	0.019	-0.300	0.015	3.309	-0.005	0.533	0.022	0.080	-
CpG9	chr16	15,814,759	NDE1, MYH11	0.082	0.262	-0.421	<0.001	5.111	-0.006	0.284	0.010	0.128	-
<i>DBP</i>													
CpG1	chr1	228,195,277	WNT3A	-0.745	0.102	-2.312	<0.001	3.102	-0.001	0.538	-0.003	0.353	-
CpG2	chr1	228,195,289	WNT3A	-0.785	0.142	-2.759	<0.001	3.515	0.000	0.819	-0.001	0.737	-
CpG3	chr1	228,195,292	WNT3A	-0.787	0.157	-2.875	<0.001	3.655	0.000	0.936	0.000	0.907	-
CpG4	chr1	228,195,260	WNT3A	-0.451	0.133	-1.083	0.015	2.404	-0.002	0.294	-0.005	0.103	-
CpG5	chr7	25,898,451	LOC646588	-0.111	0.864	-1.332	0.040	11.959	0.001	0.623	0.003	0.217	-
CpG6	chr7	25,898,447	LOC646588	-0.143	0.820	-1.301	0.038	9.112	0.001	0.634	0.002	0.263	-
CpG7	chr6	66,373,850	EYS	0.143	0.700	-0.402	0.317	-	0.012	0.094	0.038	<0.001	3.234
CpG8	chr9	80,272,835	GNA14	-0.078	0.761	0.121	0.670	-	0.011	0.062	-0.021	0.010	1.877
CpG9	chr9	80,272,842	GNA14	-0.098	0.724	0.156	0.608	-	0.011	0.070	-0.021	0.007	2.013

DBP, diastolic blood pressure; SBP, systolic blood pressure

Causal inference analysis

The results of causal inference on the top CpGs ($p < 1 \times 10^{-4}$) with SBP and DBP are provided in Table 3. Interestingly, a causal effect of DNAm to SBP was clearly supported for 9 CpGs located at/near *NDE1* and *MYH11*, *TXNL1P1*, *SMPD4*, *SRRM1P2*, *TNK2*, and *CACHD1*, respectively. Out of these 9 CpGs, a causal

effect of SBP to DNAm of 4 CpGs located at/near *TXNL1P1* and *SMPD4* was also observed.

As for DBP, the causal effect of DNAm to DBP was obviously found for 6 CpGs, with 4 at *WNT3A* and 2 at *LOC646588*. A causal effect of DBP influencing DNAm was also observed for another 8 CpGs, with 4 CpGs at

Table 4 Results of annotation to differentially methylated regions for systolic blood pressure and diastolic blood pressure

ID	Chromosome	Start (bp)	End (bp)	Length	Stouffer–Liptak–Kechris (slk) corrected p -value	Ensembl ID	Gene symbol
<i>SBP</i>							
A	chr18	77,269,147	77,269,528	20	<0.001	ENSG00000131196	<i>NFATC1</i>
B	chr3	84,330,387	84,330,523	11	0.001	ENSG00000175161	<i>CADM2</i>
C	chr4	155,665,297	155,665,627	16	0.007	ENSG00000121207	<i>LRAT</i>
D	chr13	19,173,908	19,174,405	29	0.009	ENSG00000198033	<i>TUBA3C</i>
E	chr12	74,564,341	74,564,858	20	0.010	ENSG00000251138	<i>LOC100507377</i>
F	chr1	50,881,821	50,882,443	18	0.025	ENSG00000142700	<i>DMRTA2</i>
G	chr16	32,857,318	32,857,950	31	0.031	ENSG00000214617	<i>SLC6A10P</i>
H	chr5	3,605,630	3,606,797	44	0.031	ENSG00000170549	<i>IRX1</i>
<i>DBP</i>							
A	chr1	228,195,226	228,195,293	6	<0.001	ENSG00000154342	<i>WNT3A</i>
B	chr20	21,376,425	21,376,894	28	0.003	ENSG00000125816	<i>NKX2-4</i>
C	chr16	8,619,759	8,619,952	10	0.006	ENSG00000232258	<i>TMEM114</i>
D	chr14	64,965,186	64,965,446	11	0.009	ENSG00000089775	<i>ZBTB25</i>
E	chr3	32,822,274	32,822,412	13	0.012	ENSG00000182973	<i>CNOT10</i>
F	chr17	38,088,678	38,088,969	11	0.018	ENSG00000172057	<i>ORMDL3</i>
G	chr19	47,933,149	47,933,251	4	0.027	ENSG00000118160	<i>SLC8A2</i>
H	chr19	1,465,543	1,467,185	80	0.029	ENSG00000115266	<i>APC2</i>
I	chr7	25,898,313	25,898,710	24	0.030	ENSG00000050344	<i>NFE2L3</i>
J	chr1	1,872,273	1,872,775	18	0.030	ENSG00000142609	<i>CFAP74</i>
K	chr9	124,308,098	124,308,286	11	0.031	ENSG00000136848	<i>DAB2IP</i>
L	chr18	14,999,329	15,000,083	47	0.040	ENSG00000180777	<i>ANKRD30B</i>

DBP, diastolic blood pressure; SBP, systolic blood pressure

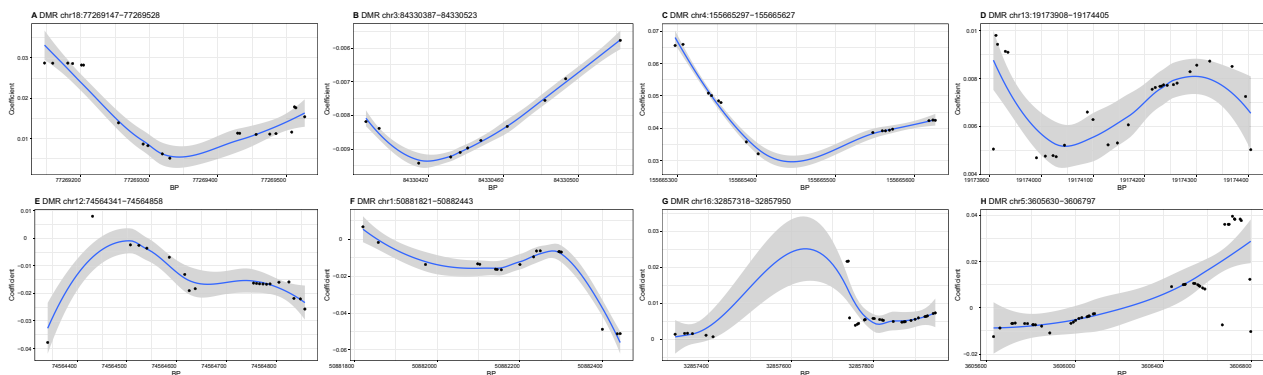


Fig. 1 Differential methylation patterns from the identified differentially methylated regions for systolic blood pressure. The dots represent the CpGs. The x-axis shows the position of CpGs on chromosome and the y-axis shows regression coefficients. BP, base pair; DMR, differentially methylated region

GNA14, 2 CpGs at *EYS*, 1 CpG at *SAE1*, and 1 CpG at *TMEM114*, respectively.

Region-based analysis

A total of 8 DMRs were identified for SBP (Table 4). As illustrated in Fig. 1, the methylation levels of 4 DMRs (A, C, D, and G) at/near *NFATC1*, *LRAT*, *TUBA3C*, and *SLC6A10P* were positively and 3 DMRs (B, E, and F) at/near *CADM2*, *LOC100507377*, and *DMRTA2* negatively correlated with SBP, whereas the trend of association between one DMR (H) at *IRX1* and SBP was uncertain.

Out of the 12 DMRs identified for DBP (Fig. 2; Table 4), the methylation level of 6 DMRs (A, C, E, F, K, and L) showed positive associations and two DMRs (G and H) showed negative associations with DBP. But it was difficult to determine the trend of association between 4 DMRs (B, D, I, and J) and DBP. These DMRs were annotated within 12 genes, such as *WNT3A*, *CNOT10*, and *DAB2IP*.

Ontology enrichments analysis

Lots of important ontology enrichments potentially associated with SBP were found, such as nicotinic acetylcholine receptor signaling pathway, p53 pathway by glucose deprivation, notch signaling pathway, Hedgehog signaling pathway, and PI3 kinase pathway (Table 5). For DBP, the ontology enrichments mainly highlighted inflammation mediated by chemokine and cytokine signaling pathway, notch signaling pathway, angiogenesis, Wnt signaling pathway, TGF-beta signaling pathway, etc. (Table 6).

Many common ontology enrichments for SBP and DBP were observed, such as nicotinic acetylcholine receptor signaling pathway, p53 pathway by glucose deprivation, Notch signaling pathway, Hedgehog signaling pathway, and PI3 kinase pathway (Additional file 5: Table S4).

We found that 2 pathways (PKA-mediated phosphorylation of CREB, regulation of insulin secretion) for SBP and 2 pathways (NCAM1 interactions, dorso-ventral axis formation) for DBP were also enriched in our previous GWAS of blood pressure in twins [8].

Quantitative methylation analysis of *COL5A1* and *WNT3A*

Eight CpGs ($p < 0.05$) mapped to *COL5A1* in EWAS of SBP were quantified using the Sequenom MassARRAY platform. As shown in Additional file 6: Table S5, just one CpG (Chr9: 137,673,907) was validated to be hypomethylated ($\beta = -0.439$, $p = 0.048$) in hypertension cases, and this CpG was also regarded as top signal as in Table 1.

Among the 5 CpGs ($p < 0.05$) mapped to *WNT3A* in EWAS of DBP, 3 were quantified using the Sequenom MassARRAY platform. As shown in Additional file 7: Table S6, all of the 3 CpGs were validated in the same direction as in EWAS and also regarded as top signal as in Table 2. Overall, the validation analysis showed clear consistency of hypermethylation in 3 CpGs within *WNT3A* associated with DBP in a community population.

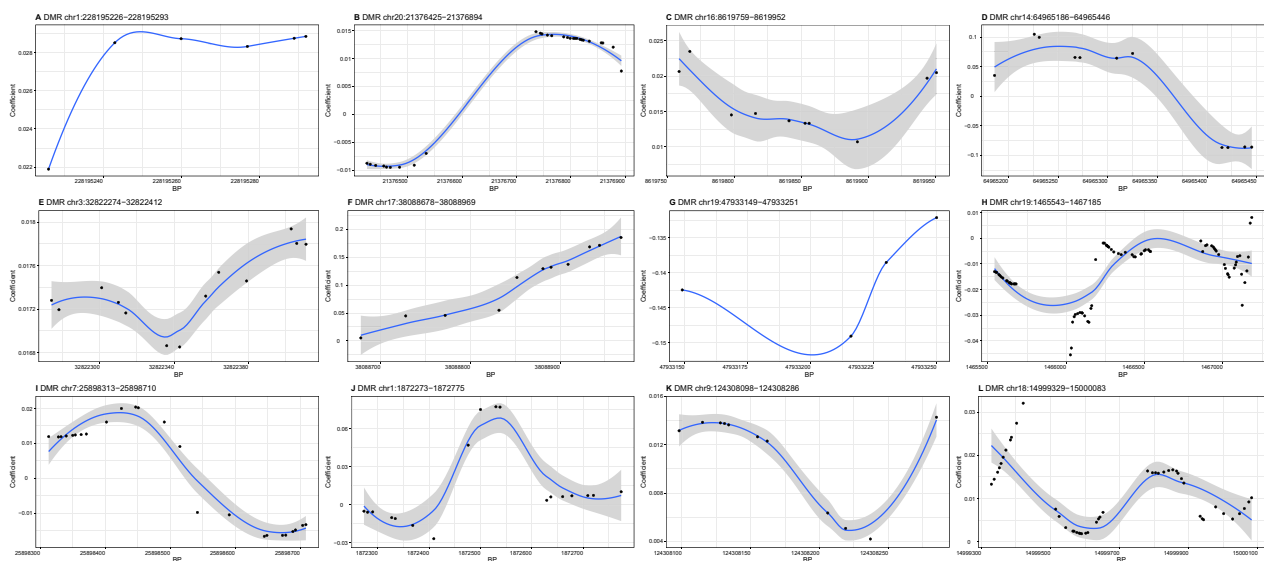


Fig. 2 Differential methylation patterns from the identified differentially methylated regions for diastolic blood pressure. The dots represent the CpGs. The x-axis shows the position of CpGs on chromosome and the y-axis shows regression coefficients. BP, base pair; DMR, differentially methylated region

Table 5 The top GREAT ontology enrichments for regions potentially related to systolic blood pressure

Ontology database	Term name	Binom FDR Q-value	Binom region fold enrichment
PANTHER pathway	Cytoskeletal regulation by Rho GTPase	1.29E-17	2.19
PANTHER pathway	Nicotinic acetylcholine receptor signaling pathway	8.79E-14	1.83
PANTHER pathway	Metabotropic glutamate receptor group II pathway	5.05E-11	2.01
PANTHER pathway	GABA-B receptor II signaling	2.71E-10	1.97
PANTHER pathway	p53 pathway by glucose deprivation	6.02E-09	2.57
PANTHER pathway	Angiogenesis	1.50E-07	1.35
PANTHER pathway	Inflammation mediated by chemokine and cytokine signaling pathway	2.66E-07	1.37
PANTHER pathway	Endogenous cannabinoid signaling	2.98E-07	2.05
PANTHER pathway	Notch signaling pathway	6.81E-07	1.78
PANTHER pathway	Hedgehog signaling pathway	4.77E-06	1.94
PANTHER pathway	Thyrotropin-releasing hormone receptor signaling pathway	7.80E-06	1.54
PANTHER pathway	Gamma-aminobutyric acid synthesis	8.53E-06	3.62
PANTHER pathway	Nicotine pharmacodynamics pathway	1.43E-05	1.82
PANTHER pathway	Heterotrimeric G-protein signaling pathway-rod outer segment phototransduction	5.11E-05	1.74
PANTHER pathway	Heterotrimeric G-protein signaling pathway-Gq alpha and Go alpha-mediated pathway	7.86E-05	1.29
PANTHER pathway	Corticotropin-releasing factor receptor signaling pathway	7.85E-04	1.68
PANTHER pathway	Adrenaline and noradrenaline biosynthesis	2.65E-03	1.67
PANTHER pathway	Insulin/IGF pathway-mitogen activated protein kinase kinase/MAP kinase cascade	5.49E-03	1.42
PANTHER pathway	PI3 kinase pathway	1.32E-02	1.34
PANTHER pathway	Histamine H1 receptor-mediated signaling pathway	1.97E-02	1.33
MSigDB pathway	Focal adhesion	1.43E-17	1.51
MSigDB pathway	Type II diabetes mellitus	2.42E-14	1.95
MSigDB pathway	Taurine and hypotaurine metabolism	8.37E-13	3.96
MSigDB pathway	RAC1 signaling pathway	1.00E-12	2.05
MSigDB pathway	Insulin signaling pathway	8.51E-10	1.49
MSigDB pathway	Regulation of RhoA activity	6.87E-09	1.82
MSigDB pathway	Arachidonic acid metabolism	7.32E-08	2.01
MSigDB pathway	T cell receptor signaling pathway	8.83E-06	1.39
MSigDB pathway	mTOR signaling pathway	7.31E-05	1.53
MSigDB pathway	VEGF signaling pathway	5.69E-04	1.42

Weighted gene co-expression network analysis (WGCNA) and gene expression analysis

We included 12 twin pairs (including 7 male pairs) with a median age of 53 years (95% range 43–65), a median SBP of 126 mmHg (95% range 94–195), and a median DBP of 81 mmHg (95% range 64–100) in the analyses.

Additional file 8: Fig. S2 illustrates the genes clustered in mediumpurple3 module (including 4,380 genes) were both negatively correlated with SBP ($r = -0.45$, $p = 0.03$) and DBP ($r = -0.45$, $p = 0.03$). Among the genes where the top CpGs ($p < 1 \times 10^{-4}$) and DMRs were annotated in methylation analysis, 3 genes (*MYH11*, *NEATC1*, and *PIP5K1C*) for SBP and 7 genes (*WNT3A*, *EYS*, *GNAI4*, *SAE1*, *CNOT10*, *APC2*, and *CFAP74*) for DBP were also clustered in mediumpurple3 module in WGCNA.

The genes in methylation analysis and genes clustered in mediumpurple3 module were involved in some

common enrichment terms, such as voltage-gated calcium channel activity, NADH dehydrogenase (ubiquinone) activity, PPAR signaling pathway, and acetylcholine receptor activity (Additional file 9: Table S7).

Discussion

It has been demonstrated that epigenetics plays a crucial part in the development hypertension; hence, looking for the specific DNAm variants potentially associated with blood pressure is still a research hotspot [46]. In this study, we detected multiple CpGs, genes, DMRs, and pathways that could not only elucidate the mechanisms of blood pressure variation but also have important implications for the intervention and treatment of hypertension.

In our EWAS on SBP, many genes where the top CpGs and DMRs were located, such as *SRRMIP2*, *COL5A1*,

Table 6 The top GREAT ontology enrichments for regions potentially related to diastolic blood pressure

Ontology database	Term name	Binom FDR Q-value	Binom region fold enrichment
PANTHER pathway	Inflammation mediated by chemokine and cytokine signaling pathway	1.42E-12	1.50
PANTHER pathway	Thyrotropin-releasing hormone receptor signaling pathway	9.07E-12	1.82
PANTHER pathway	Nicotine pharmacodynamics pathway	2.13E-11	2.27
PANTHER pathway	Endogenous cannabinoid signaling	2.44E-11	2.35
PANTHER pathway	Cytoskeletal regulation by Rho GTPase	2.68E-10	1.83
PANTHER pathway	Notch signaling pathway	1.52E-09	1.93
PANTHER pathway	Histamine H1 receptor-mediated signaling pathway	2.86E-09	1.82
PANTHER pathway	Muscarinic acetylcholine receptor 1 and 3 signaling pathway	6.39E-09	1.64
PANTHER pathway	Angiogenesis	4.17E-08	1.35
PANTHER pathway	2-Arachidonoylglycerol biosynthesis	1.35E-07	3.31
PANTHER pathway	Nicotinic acetylcholine receptor signaling pathway	7.36E-07	1.52
PANTHER pathway	Corticotropin-releasing factor receptor signaling pathway	1.51E-06	1.94
PANTHER pathway	p53 pathway by glucose deprivation	2.02E-06	2.21
PANTHER pathway	Heterotrimeric G-protein signaling pathway-rod outer segment phototransduction	2.16E-06	1.83
PANTHER pathway	Angiotensin II-stimulated signaling through G proteins and beta-arrestin	6.23E-06	1.79
PANTHER pathway	Wnt signaling pathway	7.41E-05	1.17
PANTHER pathway	GABA-B receptor II signaling	1.80E-04	1.54
PANTHER pathway	Blood coagulation	1.64E-03	1.64
PANTHER pathway	Insulin/IGF pathway-mitogen activated protein kinase kinase/MAP kinase cascade	5.21E-03	1.40
PANTHER pathway	TGF-beta signaling pathway	7.04E-03	1.22
PANTHER pathway	Beta3 adrenergic receptor signaling pathway	2.62E-02	1.46
PANTHER pathway	PI3 kinase pathway	3.04E-02	1.28
PANTHER pathway	FGF signaling pathway	3.54E-02	1.16
PANTHER pathway	Hedgehog signaling pathway	3.99E-02	1.40
PANTHER pathway	Toll receptor signaling pathway	4.29E-02	1.30
MSigDB pathway	Ceramide signaling pathway	1.96E-17	2.27
MSigDB pathway	RhoA signaling pathway	1.56E-13	2.05
MSigDB pathway	p53 pathway	3.68E-09	1.81
MSigDB pathway	VEGF signaling pathway	5.46E-06	1.54
MSigDB pathway	Insulin signaling pathway	3.20E-03	1.23

NFATC1, *NDE1*, *MYH11*, *SMPD4*, *LRAT*, *CADM2*, *IRX1*, and *TNK2*, may play important roles in regulating blood pressure. The SNP rs6794880 (chr3:84,402,361) in *SRRMIP2* was reported to be related to obesity [47], and we suspected that this locus might influence the development of obesity through regulating the DNAm at one CpG (chr3:84,330,462) in *SRRMIP2* we identified. Moreover, the association between obesity and hypertension has clearly been confirmed [48]. It was indicated that the SNPs rs4841895 in *COL5A1* [49], rs4799055 in *NFATC1* (from dbGaP database), rs1449386 in *CADM2* [50], and rs954767 in *IRX1* [51] might play a role in blood pressure regulation, and we suspected that these loci might influence the development of hypertension through regulating the DNAm in these genes. *NDE1* gene was involved in the signaling pathway by Rho GTPases, which could play a critical role in the pathogenesis of hypertension

[52]. The protein encoded by *MYH11* is a smooth muscle myosin in vascular smooth muscle cell (SMC) whose principal functions were contraction and regulation of blood pressure and blood flow distribution. The DNAm variation of *MYH11* might influence the function of SMC and hence took part in the pathogenesis of hypertension [53]. The protein encoded by *SMPD4* was a sphingomyelinase involved in sphingolipid metabolism pathway, and mounting evidence pointed toward that a derangement of this pathway might trigger the precursor clinical conditions of hypertension and hypertension itself [54]. It was found that *LRAT* may be a critical biomarker of vitamin A deficiency in target organs and may regulate blood pressure through affecting renin angiotensin system biomarkers [55]. *TNK2* gene was involved in the oxidative damage response pathway, and it was demonstrated that inflammation and oxidative stress

significantly contributed to the vascular dysfunction and renal damage associated with hypertension [56]. However, the mechanism of other genes, such as *TXNL1P1*, *PIP5K1C*, *MIR3147*, and *SLC47A1*, underlining hypertension requires further investigation.

As for DBP, several interesting genes were also found, including *DAB2IP*, *WNT3A*, *GNAI4*, *KCNT1*, *PGR*, *PLCH2*, *SIMI*, and *CNOT10*. It was previously reported that the SNPs rs35061590 and rs13290547 in *DAB2IP* might be associated with heart rate [57] and hence might influence the pathogenesis of hypertension. *WNT3A* gene was a member of the WNT gene family, and Wnt signaling pathway played an emerging role in regulating blood pressure [58]. The protein encoded by *GNAI4* was involved in the regulation of insulin secretion pathway, and the relationship of insulin, insulin sensitivity, and hypertension had been clearly confirmed [59]. It was reported that the genetic knockout mouse strain lacking K_{Na} channels (*KCNT1* and *KCNT2*) showed a modest hypertensive phenotype [60]. The SNP rs61892344 in *PGR* was previously reported to be associated with DBP [51]. The protein encoded by *PLCH2* was involved in the inositol phosphate metabolism pathway, and the inositol phosphate production in blood vessels differed in normotensive and spontaneously hypertensive rats [61]. An association of *SIMI* variants with early-onset obesity in children was demonstrated [62], but the association of *SIMI* with hypertension was currently unknown. The *CNOT10* gene was probably associated with left ventricular remodeling in hypertension by bioinformatics-based analysis [63]. Up until now, the association of other genes, such as *ATXN7L3B*, *LOC646588*, *EYS*, *MGEA5*, and *SAE1*, with hypertension had not been extensively researched, but they may also serve as candidates to be further verified.

There is a particular challenge regarding the causal effects in observational epidemiological studies using high-dimensional omics data [64]. Our study provides evidence for the causation underlying the blood pressure–DNA methylation association using ICE FALCON method. We found the causal effect that SBP was in response to the DNAm at CpGs located at several genes. *NDE1* and *MYH11* were involved in the Rho GTPase effectors pathway whose important role in the pathogenesis of vasospasm, hypertension, pulmonary hypertension, and heart failure had been demonstrated [65]. *TNK2* was involved in the oxidative damage response pathway that could cause vascular dysfunction and renal damage associated with hypertension [56]. As for DBP, clear causal effect from DNAm to DBP was found for CpGs within *WNT3A* and *LOC646588*. *WNT3A* was involved in Wnt signaling pathway whose role in regulating blood pressure had previously been

reported [58]. However, the mechanism of DNAm variation response to blood pressure changes was currently unclear, and further research was needed.

As additional validation, we quantified candidate CpGs mapped to *WNT3A* and *COL5A1* using Sequenom MassARRAY platform in a community population, and three CpGs mapped to *WNT3A* and one CpG mapped to *COL5A1* were successfully validated. As additional replication, we also compared previously reported results in EWASs with ours. A list of differentially methylated genes could be replicated, especially the well-known hypertension-associated gene *DAB2IP* [57]. We also compared the results from methylation and gene expression analyses and found a list of common genes. For SBP, these genes were involved in various biological pathways, such as nicotinic acetylcholine receptor signaling pathway (*MYH11*), Wnt signaling pathway (*NEFATC1*), and RhoA signaling pathway (*PIP5K1C*). For DBP, these common genes took part in Wnt signaling pathway (*WNT3A*, *APC2*, and *GNAI4*), ubiquitin proteasome pathway (*SAE1*), and RNA degradation pathway (*CNOT10*), etc. Moreover, we also found many common enrichment terms, such as voltage-gated calcium channel activity [66], NADH dehydrogenase (ubiquinone) activity [67], PPAR signaling pathway [68], and acetylcholine receptor activity [69], for which the relationships with hypertension were clear. All of these indicated that the DNAm variants we identified Additional file 8 played a significant role in the development of hypertension.

Several strengths can be noticed in our study. First, the trait or disease-discordant twin design we adopted has been revealed as a powerful tool for detecting the epigenetic variation underlying complex diseases [18]. Second, we also performed causal inference to investigate the causation underlying the cross-sectional epigenetic associations and found that blood pressure changes had a causal effect on the DNAm variants at some CpGs, and vice versa. Third, given the various genetic constitutions, environmental exposures, and a multitude of lifestyles in different ethnic populations worldwide, our findings will specifically help elucidate the underlying pathogenesis of hypertension in the Chinese population.

Nevertheless, the sample size of the present study was relatively limited due to the challenges of recruiting and identifying qualified twins. However, the trait or disease-discordant twin design we adopted had greater statistical power over the traditional cross-sectional or case–control design. For blood pressure with a moderate heritability, this design would allow for large sample size reductions comparing to the traditional designs. According to our previous study [17], this study based on nearly 60 twin pairs would get a statistical power of about 80%.

Conclusions

In summary, we found evidence that in peripheral blood from middle and old-aged Chinese twins, the DNAm at several loci within *WNT3A* and *COL5A1* is associated with blood pressure. Additionally, we also found evidence that blood pressure has a causal effect on peripheral blood DNAm, and vice versa. Our findings provide new clues to the epigenetic modification underlying hypertension pathogenesis.

Abbreviations

ICE FALCON	Causation through examination of familial confounding
DMR	Differentially methylated region
DBP	Diastolic blood pressure
EWAS	Epigenome-wide association study
FDR	False discovery rate
GREAT	Genomic Regions Enrichment of Annotations Tool
GEE	Generalized estimating equation
RRBS	Reduced-representation bisulfite sequencing
SBP	Systolic blood pressure
WGCNA	Weighted gene co-expression network analysis

Supplementary Information

The online version contains supplementary material available at <https://doi.org/10.1186/s13148-023-01457-1>.

Additional file 1: Table S1. The results of partial correlation analysis model between intra-pair blood pressure difference and intra-pair DNA methylation difference of each top CpG in epigenome-wide association analysis

Additional file 2: Table S2. Basic characteristics of the participants

Additional file 3: Fig. S1. Circular Manhattan plot for epigenome-wide association studies of systolic blood pressure (a) and diastolic blood pressure (b). The numbers of chromosome and the $-\log_{10}$ of p-values for statistical significance are shown. The dots represent the observed CpGs.

Additional file 4: Table S3. Comparison between our results and other previously reported blood pressure or hypertension-associated differentially methylated genes

Additional file 5: Table S4. Common ontology enrichments by GREAT tool between systolic blood pressure and diastolic blood pressure

Additional file 6: Table S5. The results of validation analysis for the CpGs mapped to *COL5A1* on systolic blood pressure

Additional file 7: Table S6. The results of validation analysis for the CpGs mapped to *WNT3A* on diastolic blood pressure

Additional file 8: Fig. S2. Relationships of consensus module eigengenes and external trait of blood pressure. Numbers in the table report the correlations with the p-values printed in parentheses. The table is color coded by correlation according to the color legend.

Additional file 9: Table S7. Common enrichment terms for blood pressure between methylation analysis and weighted gene co-expression network analysis

Acknowledgements

Not applicable.

Author contributions

Material preparation, data collection, and analysis were performed by WJW, JY, WLL, and DFZ. Investigation, resource, and data curation were performed by YLW, HPD, CSX, and XCT. The first draft of the manuscript was written by WJW and JY. The draft was revised by SXL, QHT, and DFZ. All authors read and approved the final manuscript.

Funding

This work is funded by the National Natural Science Foundation of China (31741063), Natural Science Foundation of Shandong Province (ZR2020QH304), and China Postdoctoral Science Foundation (2020M682129).

Availability of data and materials

The data used or analyzed during the current study are available from the corresponding author on reasonable request.

Declarations

Ethics approval and consent to participate

This study was approved by the Regional Ethics Committee of the Qingdao CDC Institutional Review Boards. Prior written informed consent was obtained from all participants.

Consent for publication

Not applicable.

Competing interests

The authors declare that they have no competing interests.

Author details

¹Department of Epidemiology and Health Statistics, Public Health College, Qingdao University, No. 308 Ningxia Road, Qingdao 266021, Shandong, China.

²Jiangsu Health Development Research Center, Nanjing, Jiangsu, China.

³Population Research Unit, Faculty of Social Sciences, University of Helsinki, Helsinki, Finland.

⁴Qingdao Municipal Center for Disease Control and Prevention/Qingdao Institute of Preventive Medicine, Qingdao, Shandong, China.

⁵Unit of Human Genetics, Department of Clinical Research, University of Southern Denmark, Odense, Denmark.

Received: 8 August 2022 Accepted: 24 February 2023

Published online: 03 March 2023

References

- Bundy JD, He J. Hypertension and related cardiovascular disease burden in China. *Ann Glob Health*. 2016;82(2):227–33.
- Wang Y, Peng X, Nie X, Chen L, Weldon R, Zhang W, et al. Burden of hypertension in China over the past decades: systematic analysis of prevalence, treatment and control of hypertension. *Eur J Prev Cardiol*. 2016;23(8):792–800.
- Wu Y, Zhang D, Pang Z, Jiang W, Wang S, Li S, et al. Multivariate modeling of body mass index, pulse pressure, systolic and diastolic blood pressure in Chinese twins. *Twin Res Hum Genet*. 2015;18(1):73–8.
- Ganesh SK, Tragante V, Guo W, Guo Y, Lanktree MB, Smith EN, et al. Loci influencing blood pressure identified using a cardiovascular gene-centric array. *Hum Mol Genet*. 2013;22(8):1663–78.
- Franceschini N, Fox E, Zhang Z, Edwards TL, Nalls MA, Sung YJ, et al. Genome-wide association analysis of blood-pressure traits in African-ancestry individuals reveals common associated genes in African and non-African populations. *Am J Hum Genet*. 2013;93(3):545–54.
- Wain LV, Verwoert GC, O'Reilly PF, Shi G, Johnson T, Johnson AD, et al. Genome-wide association study identifies six new loci influencing pulse pressure and mean arterial pressure. *Nat Genet*. 2011;43(10):1005–11.
- Kato N, Loh M, Takeuchi F, Verweij N, Wang X, Zhang W, et al. Trans-ancestry genome-wide association study identifies 12 genetic loci influencing blood pressure and implicates a role for DNA methylation. *Nat Genet*. 2015;47(11):1282–93.
- Chen J, Wang W, Li Z, Xu C, Tian X, Zhang D. Heritability and genome-wide association study of blood pressure in Chinese adult twins. *Mol Genet Genomic Med*. 2021;9(11):e1828-e.
- Li B, Feng ZH, Sun H, Zhao ZH, Yang SB, Yang P. The blood genome-wide DNA methylation analysis reveals novel epigenetic changes in human heart failure. *Eur Rev Med Pharmacol Sci*. 2017;21(8):1828–36.

10. Zhao G, Zhou J, Gao J, Liu Y, Gu S, Zhang X, et al. Genome-wide DNA methylation analysis in permanent atrial fibrillation. *Mol Med Rep*. 2017;16(4):5505–14.
11. Demura M, Saijoh K. The role of DNA methylation in hypertension. *Adv Exp Med Biol*. 2017;956:583–98.
12. Liang M. Epigenetic mechanisms and hypertension. *Hypertension*. 2018;72(6):1244–54.
13. Boström AE, Mwynyi J, Voisin S, Wu W, Schultes B, Zhang K, et al. Longitudinal genome-wide methylation study of Roux-en-Y gastric bypass patients reveals novel CpG sites associated with essential hypertension. *BMC Med Genomics*. 2016;9:20.
14. Huang Y, Ollikainen M, Muniandy M, Zhang T, van Dongen J, Hao G, et al. Identification, heritability, and relation with gene expression of novel DNA methylation loci for blood pressure. *Hypertension*. 2020;76(1):195–205.
15. Richard MA, Huan T, Lighthart S, Gondalia R, Jhun MA, Brody JA, et al. DNA methylation analysis identifies loci for blood pressure regulation. *Am J Hum Genet*. 2017;101(6):888–902.
16. Wang X, Falkner B, Zhu H, Shi H, Su S, Xu X, et al. A genome-wide methylation study on essential hypertension in young African American males. *PLoS ONE*. 2013;8(1): e53938.
17. Li W, Christiansen L, Hjelmborg J, Baumbach J, Tan Q. On the power of epigenome-wide association studies using a disease-discordant twin design. *Bioinformatics*. 2018;34(23):4073–8.
18. Tan Q, Christiansen L, Von Bornemann HJ, Christensen K. Twin methodology in epigenetic studies. *J Exp Biol*. 2015;218(1):134–9.
19. Li W, Zhang D, Wang W, Wu Y, Mohammadnejad A, Lund J, et al. DNA methylome profiling in identical twin pairs discordant for body mass index. *Int J Obes (Lond)*. 2019;43(12):2491–9.
20. Wang W, Li W, Jiang W, Lin H, Wu Y, Wen Y, et al. Genome-wide DNA methylation analysis of cognitive function in middle and old-aged Chinese monozygotic twins. *J Psychiatr Res*. 2021;136:571–80.
21. Wang W, Jiang W, Hou L, Duan H, Wu Y, Xu C, et al. Weighted gene co-expression network analysis of expression data of monozygotic twins identifies specific modules and hub genes related to BMI. *BMC Genomics*. 2017;18(1):872.
22. Wang W, Li W, Wu Y, Tian X, Duan H, Li S, et al. Genome-wide DNA methylation and gene expression analyses in monozygotic twins identify potential biomarkers of depression. *Transl Psychiatry*. 2021;11(1):416.
23. Duan H, Ning F, Zhang D, Wang S, Zhang D, Tan Q, et al. The Qingdao twin registry: a status update. *Twin Res Hum Genet*. 2013;16(1):79–85.
24. Xu C, Zhang D, Tian X, Wu Y, Pang Z, Li S, et al. Genetic and environmental basis in phenotype correlation between physical function and cognition in aging Chinese twins. *Twin Res Hum Genet*. 2017;20(1):60–5.
25. Becker A, Busjahn A, Faulhaber HD, Bähring S, Robertson J, Schuster H, et al. Twin zygosity. Automated determination with microsatellites. *J Reprod Med*. 1997;42(5):260–6.
26. Jackson RW, Snieder H, Davis H, Treiber FA. Determination of twin zygosity: a comparison of DNA with various questionnaire indices. *Twin Res*. 2001;4(1):12–8.
27. Krueger F, Andrews SR. Bismark: a flexible aligner and methylation caller for Bisulfite-Seq applications. *Bioinformatics*. 2011;27(11):1571–2.
28. Hebestreit K, Dugas M, Klein HU. Detection of significantly differentially methylated regions in targeted bisulfite sequencing data. *Bioinformatics*. 2013;29(13):1647–53.
29. Jaffe AE, Irizarry RA. Accounting for cellular heterogeneity is critical in epigenome-wide association studies. *Genome Biol*. 2014;15(2):R31.
30. Rahmani E, Zaitlen N, Baran Y, Eng C, Hu D, Galanter J, et al. Sparse PCA corrects for cell type heterogeneity in epigenome-wide association studies. *Nat Methods*. 2016;13(5):443–5.
31. Kim D, Pertea G, Trapnell C, Pimentel H, Kelley R, Salzberg SL. TopHat2: accurate alignment of transcriptomes in the presence of insertions, deletions and gene fusions. *Genome Biol*. 2013;14(4):R36.
32. Trapnell C, Williams BA, Pertea G, Mortazavi A, Kwan G, van Baren MJ, et al. Transcript assembly and quantification by RNA-Seq reveals unannotated transcripts and isoform switching during cell differentiation. *Nat Biotechnol*. 2010;28(5):511–5.
33. Benjamini Y, Hochberg Y. Controlling the false discovery rate: a practical and powerful approach to multiple testing. *J R Stat Soc Series B Stat Methodol*. 1995;57(1):289–300.
34. Li S, Wang W, Zhang D, Li W, Lund J, Kruse T, et al. Differential regulation of the DNA methylome in adults born during the Great Chinese Famine in 1959–1961. *Genomics*. 2021;113(6):3907–18.
35. Mohammadnejad A, Soerensen M, Baumbach J, Mengel-From J, Li W, Lund J, et al. Novel DNA methylation marker discovery by assumption-free genome-wide association analysis of cognitive function in twins. *Aging Cell*. 2021;20(2): e13293.
36. Durinck S, Moreau Y, Kasprzyk A, Davis S, De Moor B, Brazma A, et al. BioMart and bioconductor: a powerful link between biological databases and microarray data analysis. *Bioinformatics*. 2005;21(16):3439–40.
37. Durinck S, Spellman PT, Birney E, Huber W. Mapping identifiers for the integration of genomic datasets with the R/Bioconductor package biomaRt. *Nat Protoc*. 2009;4(8):1184–91.
38. Li S, Wong EM, Bui M, Nguyen TL, Joo JE, Stone J, et al. Inference about causation between body mass index and DNA methylation in blood from a twin family study. *Int J Obes*. 2019;43(2):243–52.
39. Li W, Baumbach J, Larsen MJ, Mohammadnejad A, Lund J, Christensen K, et al. Differential long noncoding RNA profiling of BMI in twins. *Epigenomics*. 2020;12(17):1531–41.
40. Pedersen BS, Schwartz DA, Yang IV, Kechris KJ. Comb-p: software for combining, analyzing, grouping and correcting spatially correlated P-values. *Bioinformatics*. 2012;28(22):2986–8.
41. McLean CY, Bristor D, Hiller M, Clarke SL, Schaar BT, Lowe CB, et al. GREAT improves functional interpretation of cis-regulatory regions. *Nat Biotechnol*. 2010;28(5):495–501.
42. Langfelder P, Horvath S. Fast R functions for robust correlations and hierarchical clustering. *J Stat Softw*. 2012;46(11):111.
43. Langfelder P, Horvath S. WGCNA: an R package for weighted correlation network analysis. *BMC Bioinformatics*. 2008;9:559.
44. da Huang W, Sherman BT, Lempicki RA. Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nat Protoc*. 2009;4(1):44–57.
45. da Huang W, Sherman BT, Lempicki RA. Bioinformatics enrichment tools: paths toward the comprehensive functional analysis of large gene lists. *Nucleic Acids Res*. 2009;37(1):1–13.
46. Raftopoulos L, Katsi V, Makris T, Tousoulis D, Stefanadis C, Kallikazaros I. Epigenetics, the missing link in hypertension. *Life Sci*. 2015;129:22–6.
47. Justice AE, Winkler TW, Feitosa MF, Graff M, Fisher VA, Young K, et al. Genome-wide meta-analysis of 241,258 adults accounting for smoking behaviour identifies novel loci for obesity traits. *Nat Commun*. 2017;8:14977.
48. DeMarco VG, Aroor AR, Sowers JR. The pathophysiology of hypertension in patients with obesity. *Nat Rev Endocrinol*. 2014;10(6):364–76.
49. Simino J, Shi G, Bis JC, Chasman DI, Ehret GB, Gu X, et al. Gene-age interactions in blood pressure regulation: a large-scale investigation with the CHARGE, Global BPgen, and ICBP Consortia. *Am J Hum Genet*. 2014;95(1):24–38.
50. Giri A, Hellwege JN, Keaton JM, Park J, Qiu C, Warren HR, et al. Trans-ethnic association study of blood pressure determinants in over 750,000 individuals. *Nat Genet*. 2019;51(1):51–62.
51. Evangelou E, Warren HR, Mosen-Ansorena D, Mifsud B, Pazoki R, Gao H, et al. Genetic analysis of over 1 million people identifies 535 new loci associated with blood pressure traits. *Nat Genet*. 2018;50(10):1412–25.
52. Loirand G, Pacaud P. The role of Rho protein signaling in hypertension. *Nat Rev Cardiol*. 2010;7(11):637–47.
53. Gomez D, Swiatlowska P, Owens GK. Epigenetic control of smooth muscle cell identity and lineage memory. *Arterioscler Thromb Vasc Biol*. 2015;35(12):2508–16.
54. Kulkarni H, Mamtani M, Blangero J, Curran JE. Lipidomics in the study of hypertension in metabolic syndrome. *Curr Hypertens Rep*. 2017;19(1):7.
55. Liang X, Chen M, Wang D, Wen J, Chen J. Vitamin A deficiency indicating as low expression of LRAT may be a novel biomarker of primary hypertension. *Clin Exp Hypertens*. 2021;43(2):151–63.
56. Small HY, Migliarino S, Czesnikiewicz-Guzik M, Guzik TJ. Hypertension: focus on autoimmunity and oxidative stress. *Free Radic Biol Med*. 2018;125:104–15.
57. Jeff JM, Ritchie MD, Denny JC, Kho AN, Ramirez AH, Crosslin D, et al. Generalization of variants identified by genome-wide association studies for electrocardiographic traits in African Americans. *Ann Hum Genet*. 2013;77(4):321–32.

58. Abou Ziki MD, Mani A. Wnt signaling, a novel pathway regulating blood pressure? State of the art review *Atherosclerosis*. 2017;262:171–8.
59. Ferrari P, Weidmann P. Insulin, insulin sensitivity and hypertension. *J Hypertens*. 1990;8(6):491–500.
60. Li P, Halabi CM, Stewart R, Butler A, Brown B, Xia X, et al. Sodium-activated potassium channels moderate excitability in vascular smooth muscle. *J Physiol*. 2019;597(20):5093–108.
61. Vila E, Macrae IM, Reid JL. Differences in inositol phosphate production in blood vessels of normotensive and spontaneously hypertensive rats. *Br J Pharmacol*. 1991;104(2):296–300.
62. Zegers D, Beckers S, Hendrickx R, Van Camp JK, de Craemer V, Verrijken A, et al. Mutation screen of the SIM1 gene in pediatric patients with early-onset obesity. *Int J Obes (Lond)*. 2014;38(7):1000–4.
63. Pang B, Hu C, Wu G, Zhang Y, Lin G. Identification of target genes in hypertension and left ventricular remodeling. *Medicine (Baltimore)*. 2020;99(28): e21195.
64. Richmond RC, Hemani G, Tilling K, Davey Smith G, Relton CL. Challenges and novel approaches for investigating molecular mediation. *Hum Mol Genet*. 2016;25(R2):R149–56.
65. Shimokawa H, Sunamura S, Satoh K. RhoA/Rho-kinase in the cardiovascular system. *Circ Res*. 2016;118(2):352–66.
66. Zamponi GW, Striessnig J, Koschak A, Dolphin AC. The physiology, pathology, and pharmacology of voltage-gated calcium channels and their future therapeutic potential. *Pharmacol Rev*. 2015;67(4):821–70.
67. Chen X, He XY, Zhu C, Zhang Y, Li Z, Liu Y, et al. Interaction between mitochondrial NADH dehydrogenase subunit-2 5178 C > A and clinical risk factors on the susceptibility of essential hypertension in Chinese population. *BMC Med Genet*. 2019;20(1):121.
68. Toral M, Romero M, Pérez-Vizcaino F, Duarte J, Jiménez R. Antihypertensive effects of peroxisome proliferator-activated receptor- β/δ activation. *Am J Physiol Heart Circ Physiol*. 2017;312(2):H189–200.
69. Liu C, Su D. Nicotinic acetylcholine receptor $\alpha 7$ subunit: a novel therapeutic target for cardiovascular diseases. *Front Med*. 2012;6(1):35–40.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Ready to submit your research? Choose BMC and benefit from:

- fast, convenient online submission
- thorough peer review by experienced researchers in your field
- rapid publication on acceptance
- support for research data, including large and complex data types
- gold Open Access which fosters wider collaboration and increased citations
- maximum visibility for your research: over 100M website views per year

At BMC, research is always in progress.

Learn more biomedcentral.com/submissions

