

Article

Variation in Genotype and DNA Methylation Patterns Based on Alcohol Use and CVD in the Korean Genome and Epidemiology Study (KoGES)

Myoungjee Jung ¹, Yeon-Soon Ahn ², Sei-Jin Chang ¹ , Chun-Bae Kim ¹ , Kyoung Sook Jeong ³ , Sang-Baek Koh ^{1,4,*} and Jeong-An Gim ^{5,*}

¹ Department of Preventive Medicine, Yonsei University Wonju College of Medicine, 20 Ilsan-ro, Wonju 26426, Korea; laurie810305@naver.com (M.J.); chang0343@yonsei.ac.kr (S.-J.C.); kimcb@yonsei.ac.kr (C.-B.K.)

² Department of Preventive Medicine and Genomic Cohort Institute, Yonsei University Wonju College of Medicine, 20 Ilsan-ro, Wonju 26426, Korea; ysahn1203@gmail.com

³ Department of Occupational and Environmental Medicine, Yonsei University Wonju College of Medicine, 20 Ilsan-ro, Wonju 26426, Korea; bandyoem@naver.com

⁴ Institute of Genomic Cohort, Yonsei University Wonju College of Medicine, 20 Ilsan-ro, Wonju 26426, Korea

⁵ Medical Science Research Center, College of Medicine, Korea University, 8 Gamasan-ro 20-gil, Guro-gu, Seoul 08308, Korea

* Correspondence: kohhj@yonsei.ac.kr (S.-B.K.); vitastar@korea.ac.kr (J.-A.G.); Tel.: +82-33-741-0345 (S.-B.K.); +82-2-2626-2362 (J.-A.G.)



Citation: Jung, M.; Ahn, Y.-S.; Chang, S.-J.; Kim, C.-B.; Jeong, K.S.; Koh, S.-B.; Gim, J.-A. Variation in Genotype and DNA Methylation Patterns Based on Alcohol Use and CVD in the Korean Genome and Epidemiology Study (KoGES). *Genes* **2022**, *13*, 172. <https://doi.org/10.3390/genes13020172>

Academic Editors: Donato Gemmati and Robert Philibert

Received: 14 November 2021

Accepted: 14 January 2022

Published: 19 January 2022

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

Abstract: Alcohol consumption can increase the risk of chronic diseases, such as myocardial infarction, coronary artery disease, hyperlipidemia, and hypertension. We aimed to assess the association between genotype, DNA methylation patterns, alcohol consumption, and chronic diseases in Korean population. We analyzed 8840 subjects for genotypes and 446 for DNA methylation among the 9351 subjects from the Korean Genome and Epidemiology Study (KoGES). We further divided both groups into two sub-groups according to the presence/absence of chronic diseases. We selected genes whose methylation varied significantly with alcohol consumption, and visualized genotype and DNA methylation patterns specific to each group. Genome-wide association study (GWAS) revealed single nucleotide polymorphisms (SNPs) *rs2074356* and *rs11066280* in HECT domain E3 ubiquitin protein ligase 4 (*HECTD4*) to be significantly associated with alcohol consumption in both the presence. The *rs12229654* genotype also displayed significantly different patterns with alcohol consumption. Furthermore, we retrieved differentially methylated regions (DMRs) from four groups based on sex and chronic diseases and compared them by drinking status. In genotype analysis, cardiovascular diseases (CVDs) showed a higher proportion in drinker than in non-drinker, but not in DMR analysis. Additionally, we analyzed the enriched Gene Ontology terms and Kyoto Gene and Genome Encyclopedia (KEGG) pathways and visualized the network, heatmap, and upset plot. We show that the pattern of DNA methylation associated with CVD is strongly influenced by alcoholism. Overall, this study identified genetic and epigenetic variants influenced by alcohol consumption and chronic diseases.

Keywords: alcohol consumption; chronic diseases; DNA methylation; genotype

1. Introduction

Alcohol use disorder or alcoholism ranks sixth among factors that increase the risk of disease-related mortality and disability [1]. Moreover, alcoholism is a prevalent lifestyle factor relevant to chronic cardiovascular diseases (CVDs) such as myocardial infarction, coronary artery, hyperlipidemia, and hypertension [2]. The correlation between alcohol consumption and chronic diseases is well known, and restriction of alcohol intake is commonly recommended for the management of chronic diseases. To date, many studies

have analyzed the association of alcohol-induced chronic diseases in preventive medicine and epidemiologic studies. However, a new approach based on omics data and data integration is required.

CVD is the leading cause of global morbidity and mortality in both men and women [3–6]. Monogenic conditions can also lead to severe premature CVD and early death, if unrecognized and untreated [7]. CVD also inflicts a significant economic burden on countries [6,8] as well individuals. Although alcohol-dependent people may have diverse medical conditions or diseases, the major cause of their death is CVD [9,10]. The association between alcohol intake and CVD is explained as the J-shaped curve. Since alcohol exerts complex effects on the cardiovascular system, elaborating the major pathways involved is challenging. Epidemiology-based, genetic, and epigenetic approaches have been employed to explain the relationship between alcoholism and cardiovascular diseases, and an integrated analysis is relevant.

A mendelian randomization study provides an alternative approach to establish the causal role of moderate alcohol use in a suitable population and reveals that a genetic variant affects alcohol metabolism and thereby alcohol use. Genetic variants of the several genes affect alcohol metabolism [11]. Cook and colleagues (2021) identified subgroups among Asian Americans at high risks of CVD-related conditions associated with excessive alcohol consumption—females of Chinese, Japanese, Korean, and Vietnamese ethnicities in higher *ALDH2*2* groups. Overall, members of the higher *ALDH2*2* ethnic groups were associated with lower risks of CVD-related conditions [12]. The genotype of the alcohol metabolism genes is related to alcohol-related CVD, and the relationship requires verification under other conditions.

DNA methylation, a major type of epigenetic modification, is associated with various complex human diseases and traits [13] and is potentially an important mechanism to regulate gene expression [14]. DNA methylation patterns are modulated by several diseases, including CVD [15]. DNA methylation affects the cytosine within CpG sites in the entire genome [7], and different diseases are associated with varying DNA methylation patterns. Previous studies have reported a genome-wide increase in the levels of DNA methylation in patients diagnosed with CVDs [16], as well as confirmed coronary artery disease (CAD) [17]. DNA methylation status is also affected by the environment [18], and the variation is associated with alcohol consumption [19]. Alcohol modulates the DNA methylation machinery [20] as well as the metabolism [21]. Hence, CpGs can serve as useful biomarkers for this disease, and the identification of a subset of CpGs involved in the disease can provide insights into the etiology of this disease [22]. In addition, identification of differentially methylated regions (DMRs) is necessary to understand the relationship between DNA methylation and its function in organisms. Epigenome-wide association studies (EWAS) have been performed to reveal the effects of DMRs on various diseases [21,22].

The genetic and behavioral differences between males and females need to be considered before analysis. From an epidemiological perspective, the prevalence and prognosis of chronic diseases based on large cohorts are sex dependent [23]. In addition, drinking habits differ according to sex. Epigenetically, female-specific hypermethylation is observed as a result of X-chromosome inactivation [24]. Therefore, in order to circumvent the sex bias, male and female populations should be analyzed separately. Alternatively, genes located on the sex chromosomes X and Y should be excluded before analysis. In this study, we separated the subjects into four groups based on the incidence of CVDs and sex.

So far, the associations have been studied between drinking and CVD, and the genotype has been reported to be correlated with alcoholism. However, very few studies have been undertaken to screen and analyze the genotypes and epi-genotypes in Koreans. In this study, we investigated the correlation between the occurrence of CVD and genotypes in alcoholic and non-alcoholic subjects. Our goals are to identify the SNPs associated with CVDs induced by alcohol. We also tried to find DMRs from four groups based on sex and CVDs by drinking status.

2. Materials and Methods

2.1. Study Subjects and Data Source

The epidemiological and genomic data sets of the present study were collected from a survey during 2001–2002 and were obtained from the Ansan and Ansong cohort of the Korean Genome and Epidemiology Study (KoGES) released by the Korea Center for Disease Control & Prevention [25]. Among the baseline participants ($n = 10,030$, Ansong: 5018; Ansan: 5012) of the Ansan and Ansong cohort, the DNA samples of 10,004 of the participants were available and genotyped. After the data were filtered using standard quality control procedures, the genotype data of 8840 participants were considered for further analysis [26].

Participants were 40–69 years of age and belonged to the Ansan and Ansong communities located in Gyeonggi province, South Korea. All participants signed an informed consent form. This study was approved by the Institutional Review Board (IRB) of Yonsei Wonju Severance Christian Hospital (Approval Number: CR320355) and Korea University (Approval Number: KUIRB-2020-0191-01) and performed in accordance with the Declaration of Helsinki.

2.2. Study Design

This study investigated the genetic and epigenetic risk factors of CVDs [27]. We divided the subjects into two groups based on the incidence of CVDs (myocardial infarction, coronary artery, hyperlipidemia, and hypertension) and an absence of CVDs. A flowchart of how subjects for genotype and methylation analyses were selected from 9351 participants is presented in Figure 1. As a case–control study, we compared drinkers and non-drinkers in the two groups, CVD and non-CVD groups. We then identified the frequently detected genotypes and epigenotypes via association studies.

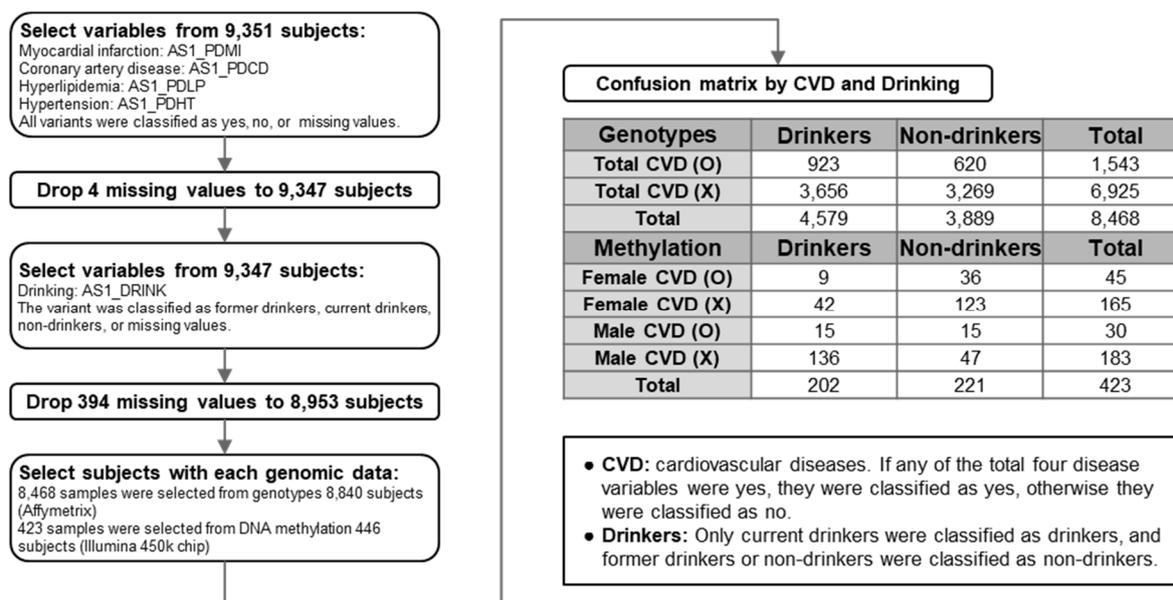


Figure 1. Flowchart for selection of the study population. The study population included 8840 genotypes and 446 methylation analysis results among 9351 subjects. Chronic cardiovascular diseases (CVDs) were considered as present when there was at least one of the following four chronic diseases: myocardial infarction, coronary artery, hyperlipidemia, and hypertension. “AS1_PDMI”, “AS1_PDCD”, “AS1_PDLP”, “AS1_PDHT”, and “AS1_DRINK” indicate the variable codes of myocardial infarction, coronary artery, hyperlipidemia, hypertension, and drinking status, respectively. Analysis was performed on 8468 subjects for genotyping and 423 subjects for methylation, which correspond to the intersection of the four aforementioned CVD-affected and drinking sample groups.

Among 8468 people in the genotype analysis, 4579 were drinkers and 3889 were non-drinkers. A confusion matrix was developed for 1543 CVD subjects and 6925 non-CVD subjects (Figure 1). For the presence or absence of CVD, a genome-wide association study (GWAS) between drinkers and non-drinkers was performed.

To avoid sex-biased DNA methylation differences, we also divided the groups by two sexes. Of the 423 patients divided into the aforementioned groups, 202 were drinkers and 221 were non-drinkers. The female group with CVDs consisted of 45 individuals: nine drinkers and 36 non-drinkers. The female group with no CVDs consisted of 165 subjects: 42 drinkers and 123 non-drinkers. Among 30 males with CVDs, there were 15 drinkers and 15 non-drinkers (Figure 1). One hundred and eighty-three subjects with non-CVDs were identified among men, of which, 136 were classified as drinkers and 47 as non-drinkers.

2.3. Definition and Measurement of Lifestyle Factors

The incidence of alcohol consumption and CVD were assessed using questionnaires designed by well-trained investigators. Questionnaires about alcohol consumption specifically investigated the presence or absence of alcohol, and the question was precisely framed as “Can’t you drink alcohol or don’t you drink from the beginning (for religious reasons, etc.)?” The subjects were required to respond in either “Yes, I do not drink alcohol” or “No (currently drinking)”.

CVD was considered as the dependent variable, and the CVDs were investigated in patients with myocardial infarction, coronary artery, hyperlipidemia, or hypertension. Questionnaires about the presence and absence of each condition were composed as “Have you ever been diagnosed with the following by a clinician?”, with the possible responses “Yes” or “No”. In all categorical questionnaire items, “Yes” was coded as 2 and “No” was coded as 1. If there was no answer, it was marked as a missing value, and “NA” was indicated when coding.

In the questionnaires of exercise intensity, the time spent per day was written by the subjects in minutes. Questions were separated by intensity, and the levels were stable, sedentary, light, moderate, and vigorous.

2.4. SNP Genotypes

The genotypes were obtained using Affymetrix Genome-Wide Human SNP array 5.0 (Affymetrix Inc., Santa Clara, CA, USA), which contains 500,568 SNPs constructed by the Korea Biobank Array. To examine the relationship between genetic polymorphism and CVD, we selected the significant SNPs by *p*-values.

2.5. DNA Methylation Analysis

In this study, genomic DNA extracted from the peripheral blood of the subjects was used for the evaluation of DNA methylation levels. Total 500 ng genomic DNA of each sample was modified by treatment with sodium bisulfite provided in the EZ DNA methylation kit (Zymo Research, Irvine, CA, USA), according to the manufacturer’s manual. Genome-wide DNA methylation was profiled using Illumina Infinium Human Methylation 450 k BeadChip (Illumina, San Diego, CA, USA), which contains over 485,000 CpG probes covering 99% of the RefSeq genes. Each CpG probe has a β value ranging between 0 and 1; higher methylation of CpG gives rise to a value closer to 1. We analyzed 364,050 CpG probes from 423 subjects, excluding the missing values. Furthermore, the gene symbol and genomic location corresponding to each CpG probe of the Illumina 450k methylation chip were obtained using the “get Annotation” function of the IlluminaHumanMethylationEPI-Canno.ilm10b2.hg19 library provided by Bioconductor (<https://www.bioconductor.org>, accessed on 6 December 2021).

2.6. Statistical Analysis of Genotypes and Methylation

Statistical analyses were performed using the PLINK software (version 1.9) (<https://www.cog-genomics.org/plink2>, accessed on 6 December 2021) and R software 4.0.5.

The PLINK software was used to filter irrelevant SNPs, including those on the X and Y chromosomes and the mitochondrial genome. During PLINK analysis, genotyping call rate (0.05) was used as the SNP filtering parameter to remove missing genotypes of the SNPs [28]. Male and female subjects were analyzed in stratification. The association between SNPs and CVD disease related to alcohol consumption is represented by a Manhattan plot. The R software was used to generate the Manhattan plot using qqman package version 0.1.8 for genome-wide association studies (<http://cran.r-project.org/web/packages/qqman/>, accessed on 6 December 2021).

The DMRs were identified using a *t*-test performed between drinkers and non-drinkers in each of the four groups. The filtered DMRs were visualized as volcano plots and heatmaps. Enrichment analyses were performed using DMRs and visualized as KEGG and GO dot plots and gene-concept networks and heatmaps. To design the classification models, two machine learning methods, that is, decision trees and random forests were applied using the “rpart” and “randomForest” packages, respectively. Enrichment analysis was performed using the “pathfindR” package, and enrichment terms were subsequently retrieved as an upset plot.

3. Results

3.1. Study Processes

The process followed in this study for selection and classification of the study subjects has been represented as a flowchart in Figure 1. We recruited 9351 subjects. Five variables were used to distinguish between patients with CVDs and drinkers, and subjects with missing values were excluded. Therefore, 8953 subjects were finally followed. The GWAS was performed on 8468 subjects, which is the intersection of 8840 subjects whose genotype was analyzed and 8953 subjects with no missing values in the five variables. DNA methylation analysis was performed on 423 subjects, which is the intersection of 446 subjects whose methylation patterns was analyzed and 8953 subjects with no missing values in the five variables. A confusion matrix classifying drinkers and non-drinkers in the GWAS and methylation analyses is presented at the end of the flowchart.

3.2. Identification of Genotype Patterns between the Four Selected Conditions

The GWAS results between drinking status were summarized as top three most significant SNPs in two separated groups (Table 1) and visualized as two Manhattan plots (Figure 2). CVD had the lowest odds ratio (OR) of 0.2402 in *rs2074356* belonging to *HECTD4*, which was the most significant genotype. On the contrary, without CVD, *rs2074356* was also the most statistically significant genotype. Therefore, *HECTD4* showed statistically significant values in both classes. The top three genotypes were in the same order by OR (Table 1). In the Manhattan plot of the CVD group, a region with $-\log_{10}(p\text{-value}) > 8$ was found on chromosome 12, and a region with $\log_{10}(p\text{-value}) > 5$ was found on chromosome X (Figure 2, top). In the Manhattan plot of non-CVD group, a region with $-\log_{10}(p\text{-value}) > 8$ was found on chromosomes 12 and X, and a region with $\log_{10}(p\text{-value}) > 5$ was found on chromosome 4 (Figure 2, bottom). The two groups showed a similar pattern but a more significant statistical difference because more subjects were included in the non-CVD group.

Table 1. The most significant SNPs associated with chronic cardiovascular diseases (CVDs) and without CVD between drinking status.

Class	Location (hg38)	dbSNP ID	Gene	Odds Ratio	<i>p</i> -Value
With CVD	chr12:112207597	<i>rs2074356</i>	<i>HECTD4</i> : Intron Variant	0.2402	1.49×10^{-26}
With CVD	chr12:112379979	<i>rs11066280</i>	<i>HECTD4</i> : Intron Variant	0.2919	1.15×10^{-24}
With CVD	chr12:110976657	<i>rs12229654</i>	None	0.308	2.01×10^{-20}
Without CVD	chr12:112207597	<i>rs2074356</i>	<i>HECTD4</i> : Intron Variant	0.3013	1.41×10^{-107}
Without CVD	chr12:112379979	<i>rs11066280</i>	<i>HECTD4</i> : Intron Variant	0.3326	1.95×10^{-106}
Without CVD	chr12:110976657	<i>rs12229654</i>	None	0.3689	1.43×10^{-75}

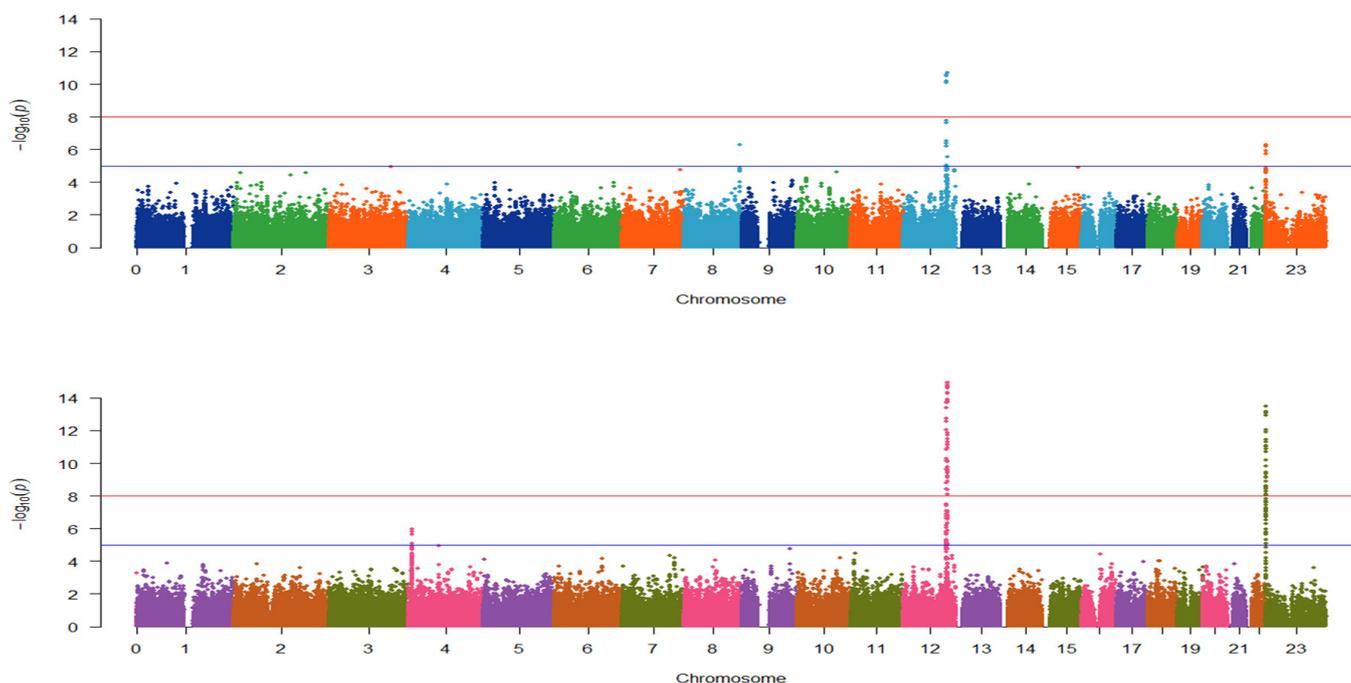


Figure 2. Manhattan plot of the association between drinkers and non-drinkers. (Top) Among 8468 subjects, 1543 subjects had at least chronic cardiovascular diseases (CVDs). In the CVD group, 923 were drinkers and 620 were non-drinkers. (Bottom) In the non-CVD group, 3656 were drinkers and 3269 were non-drinkers. The plot shows $-\log_{10} p$ value for each SNP against the chromosomal location.

3.3. DNA Methylation Analysis of Various Groups

We examined demographic variables to classify the subjects based on the incidence of CVD and drinking habits for methylation analysis, as shown in Table 2. Total 14 demographic and sociological variables, such as sex and age, associated with CVDs caused by drinking were analyzed. Among the remaining 423 participants, 75 (17.7%) were identified as CVD patients. Among female subjects, the 45 participants with CVD were classified into 9 drinkers (20.0%) and 36 non-drinkers. Among male subjects, the 30 participants with CVD were classified into 15 drinkers (50.0%) and 15 non-drinkers. Furthermore, among the 165 female participants with non-CVD, 42 were classified as drinkers (25.5%) and 123 as non-drinkers (74.5%). Among male subjects, 183 participants with non-CVD were classified as 136 drinkers (74.3%) and 47 non-drinkers.

Table 2. Basal characteristics of participants classified according to drinking and chronic cardiovascular diseases (CVDs) for methylation analysis.

Variable	CVDs (n = 75)			Non-CVDs (n = 348)		
	Never Drinkers	Drinkers	p-Value	Never Drinkers	Drinkers	p-Value
Sex	15/36/29.41	15/9/62.5	0.011	47/123/27.65	136/42/76.4	<0.001
Age	56.75 ± 7.63	53.62 ± 6.97	<0.001	53.12 ± 8.37	49.57 ± 7.95	<0.001
BMI	26.58 ± 3.05	25.5 ± 3.27	<0.001	23.91 ± 3.21	24.53 ± 3.39	<0.001
Area	34/17/66.67	14/10/58.33	0.607	102/68/60.00	72/106/40.45	0.001
Education	37/13/74.00	11/13/45.83	0.022	112/58/65.88	68/108/38.64	<0.001
Income	33/16/67.35	8/16/33.33	0.011	97/69/58.43	70/108/39.33	0.001
Exercise1	33/17/66.00	13/11/54.17	0.443	114/55/67.46	100/77/56.50	0.046
Exercise2	12/37/24.49	7/17/29.17	0.778	43/125/25.60	35/141/19.89	0.246
Exercise3	20/29/40.82	9/15/37.50	1.000	55/113/32.74	46/129/26.29	0.195

Table 2. Cont.

Variable	CVDs (n = 75)			Non-CVDs (n = 348)		
	Never Drinkers	Drinkers	p-Value	Never Drinkers	Drinkers	p-Value
Exercise4	37/10/78.72	19/5/79.17	1.000	120/47/71.86	127/48/72.57	0.904
Exercise5	32/16/66.67	18/6/75	0.591	111/58/65.68	125/52/70.62	0.356
Myocardial infarction	1/50/1.96	0/24/0	<0.001	0/170/0	0/178/0	1.000
Coronary artery disease	3/48/5.88	3/21/12.50	0.377	0/170/0	0/178/0	1.000
Hyperlipidemia	3/48/5.88	4/20/16.67	0.201	0/170/0	0/178/0	1.000
High blood pressure	45/6/88.24	20/4/83.33	0.717	0/170/0	0/178/0	1.000

Continuous variables are indicated as “average \pm standard deviation”, and categorical variables are indicated as “Counted samples in group 1/Counted samples in group 2/Ratio of group 1 in total samples”. Sex: Male = 1, Female = 2; Area: Ansong = 1, Ansan = 2; Education: Below middle school = 1, Over high school = 2; Income: Under 1.5 million won = 1, over 1.5 million won = 2; Exercise status: Under 60 min/day = 1, over 60 min/day = 2. Exercise levels are indicated as 1, 2, 3, 4, and 5 corresponding to stable, sedentary, light, moderate, and vigorous, respectively; Four CVD diseases: disease group = 1, normal group = 2.

The mean and standard deviation of age and BMI according to the drinking status in the CVD and non-CVD groups are presented. Exercising for more than 60 min per day by area, education, income, and intensity is presented. The most common chronic disease in the CVD group was high blood pressure, and coronary artery disease and hyperlipidemia were high in the drinkers group, but there was no statistically significant difference. Only one patient in the CVD group had myocardial infarction.

3.4. Identification of DMRs

Owing to the differences in DNA methylation patterns between female and male subjects, the groups were divided according to sex to avoid bias. Volcano plots (Figure 3) depict DNA methylation matrices of the four groups (classified by sexes and CVDs). Points that meet the conditions are marked in pink (highly methylated in drink conditions) or sky-blue (highly expressed in no-drink conditions). In each plot, the X-axis represents the fold-changes (FC) and the Y-axis represents the p-value on a $-\log_{10}$ scale. The area that satisfies the threshold of FC and p-value is represented with two vertical lines and one horizontal line. Based on these criteria, the differences in methylation levels were confirmed across the female, CVD and male, CVD groups (Figure 3A,B); and the female, non-CVD and male, non-CVD groups (Figure 3C,D). Four pairs of DMRs were finally screened from each group based on the methylation patterns and were represented as a heatmap (Figure 4). Heatmaps were generated for the female, CVD and male, CVD groups (Figure 4A,B) as well as the female, non-CVD and male, non-CVD groups (Figure 4C,D). The heatmap shows the variation in DNA methylation between the four groups. Column annotation bars indicate the two parameters in the samples; i.e., alcohol consumption (drink and no-drink), and sex (male and female). The pink and sky-blue bars indicate sex and drinking, respectively. The row annotation bars represent p-value and FC, and a clustering pattern is observed. In the FC bar, pink indicates a pattern indicating high methylation under drink conditions, whereas sky-blue indicates high methylation under non-drinking conditions. The PV bar indicates the p-value, and the intensity of gray bars indicates the significance of the difference between two groups.

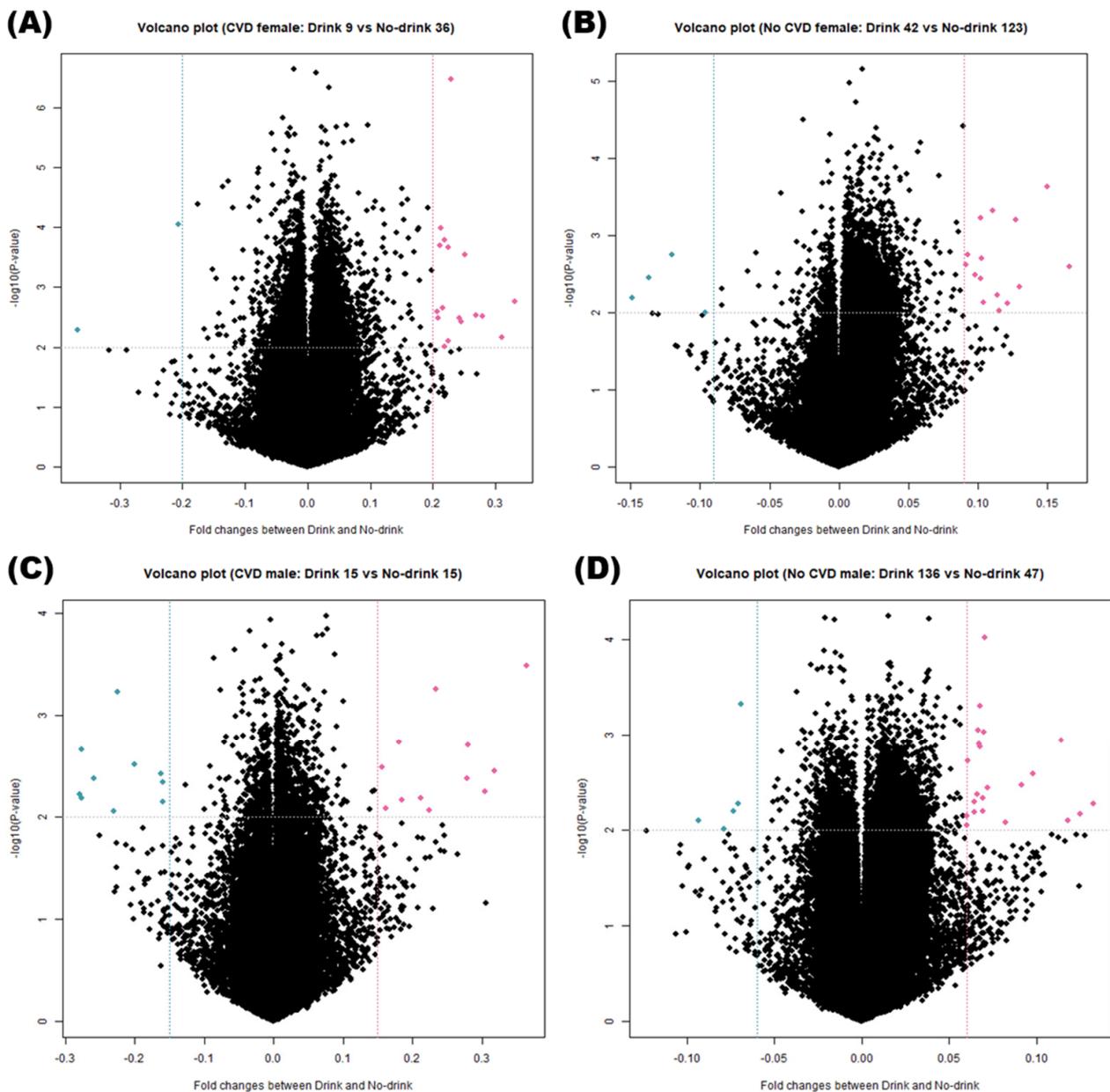


Figure 3. Volcano plot of differentially methylated regions (DMRs) indicating difference between four pairs of four groups. Total 423 subjects were divided into four groups based on the cardiovascular diseases (CVDs) (at least one of myocardial infarction, coronary artery, hyperlipidemia, and hypertension) or non-CVDs, and sex (male or female). **(A)** In females with CVD, *t*-test was performed on nine drinkers versus 36 non-drinkers, and 19 DMRs were found to satisfy the criteria $|\text{fold change}| > 0.2$ and $p\text{-value} < 0.01$. **(B)** In males with CVD, *t*-test was performed on 15 drinkers versus 15 non-drinkers, and 22 DMRs were found to satisfy the criteria $|\text{fold change}| > 0.15$ and $p\text{-value} < 0.01$. **(C)** In females with non-CVD, 42 drinkers versus 123 non-drinkers and 19 DMRs satisfied the criteria $|\text{fold change}| > 0.09$ and $p\text{-value} < 0.01$. **(D)** In females with CVD, the pairs of 136 drinkers versus 47 non-drinkers were subjected to *t*-test, and 27 DMRs satisfied the criteria $|\text{fold change}| > 0.06$ and $p\text{-value} < 0.01$. In the volcano plot, higher methylated regions between drinkers and non-drinkers of four groups are indicated as pink dots.

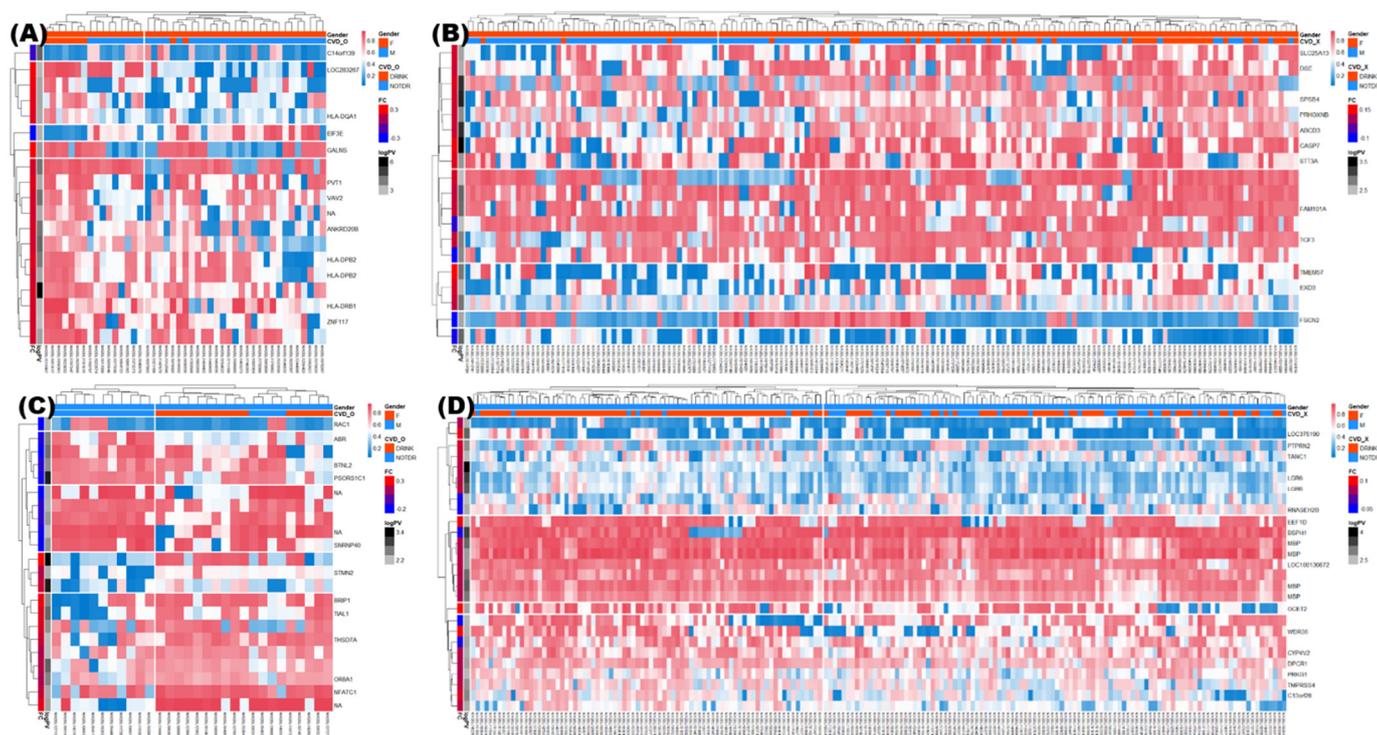


Figure 4. Heatmap of DMRs indicating the difference between four pairs of the four groups. Total 423 subjects were divided on the basis of presence of chronic cardiovascular diseases (CVDs); at least one of myocardial infarction, coronary artery, hyperlipidemia, and hypertension) or non-CVDs, and sexes (male or female). You may consider reporting this data in the body of results, rather than in the figure legends. This would help in avoiding the repetition between Figures 3 and 4. (A) In females with CVD, *t*-test was performed on 9 drinkers versus 36 non-drinkers, and 19 DMRs were found to satisfy the criteria $|\text{fold change}| > 0.2$ and $p\text{-value} < 0.01$. (B) In males with CVD, *t*-test was performed on 15 drinkers versus 15 non-drinkers, and 22 DMRs were identified as satisfying the criteria $|\text{fold change}| > 0.15$ and $p\text{-value} < 0.01$. (C) In females with non-CVD, 42 drinkers versus 123 non-drinkers, and 19 DMRs were identified as satisfying the criteria $|\text{fold change}| > 0.09$ and $p\text{-value} < 0.01$. (D) In females with CVD, the pairs of 136 drinkers versus 47 non-drinkers were subjected to *t*-test, and 27 DMRs were identified as satisfying the criteria $|\text{fold change}| > 0.06$ and $p\text{-value} < 0.01$. In the heatmaps, higher methylated regions between drinkers and non-drinkers of the four groups are indicated as red annotation bars in drinkers. A higher significance deduced from a lower p -value is indicated as darker gray annotation bar. Sexes and CVD are indicated as column annotation bars.

3.5. Enrichment Analysis of DMRs

We obtained a list of results from the KEGG pathway and GO analyses for the four groups (Figure 5). The KEGG dot plots of female, CVD and male, CVD groups are represented in Figure 5A,B, while those of female, non-CVD and male, non-CVD groups are presented in Figure 5C,D. The KEGG pathways with a difference among four pairs of the four groups were significantly enriched by the DMRs. In the CVD group, the neuroactive ligand-receptor interaction was the most enriched in females and was statistically significant. In the same group, males were most enriched with multiple diseases associated with neurogeneration pathways, which was statistically significant. In the non-CVD group, the MAPK signaling pathway showed the common top gene ratio values in the two sexes. Pathway-related genes were enriched in the non-CVD group. Changes in the pathways of genes related to DMR between drinkers and non-drinkers were determined by the presence or absence of CVD rather than differences between the two sexes. Interestingly,

disease-related pathways have been presented in the male and female CVD groups, and will be applied to study each disease, alcoholism, and CVD.

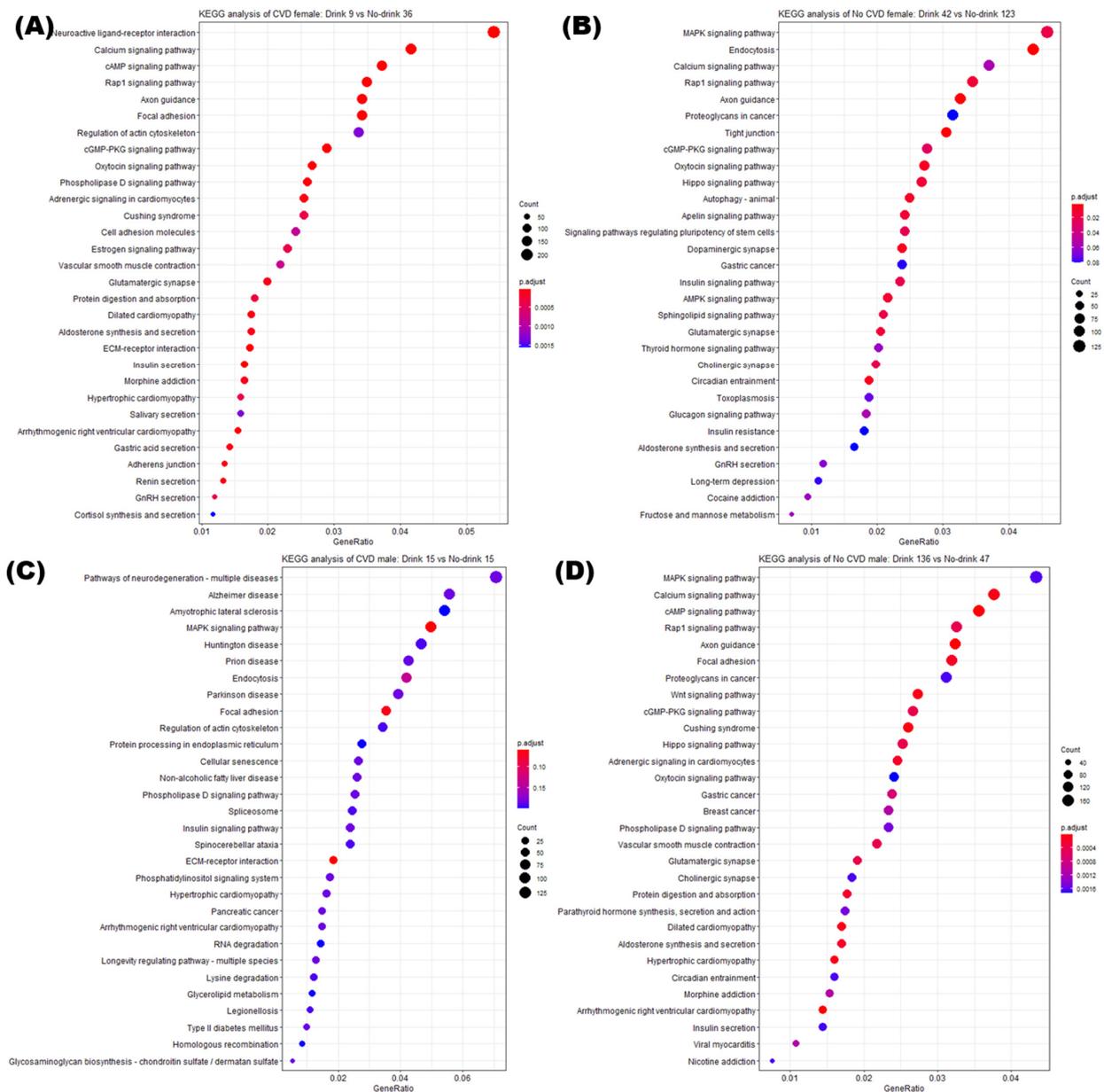


Figure 5. Enriched KEGG pathways of the DMRs indicating the difference between four pairs of the four groups. The top 30 enrichment KEGG pathways for DMRs satisfying the criteria p -value < 0.05 in each of the four pairs are presented. Color of the plot represents the p -value, and the size represents the count of genes. (A) In females with chronic cardiovascular diseases (CVDs), the pairs of 9 drinkers versus 36 non-drinkers were compared. (B) In males with CVD, the pairs of 15 drinkers versus 15 non-drinkers were compared. (C) In females with non-CVD, the pairs of 42 drinkers versus 123 non-drinkers were compared. (D) In females with CVD, the pairs of 136 drinkers versus 47 non-drinkers were compared.

3.6. Network Analysis of DMRs

Based on the network analysis, gene-concept heatmaps were generated (Figure 6). In each heatmap, x and y axes indicate the genes and included terms, respectively. In addition, the highly expressed genes in the drinking group are indicated in red, while the genes showing low methylation are indicated in blue. In female groups, 17 and 11 terms

were obtained, and their related genes were identified along with their methylation levels (Figure 6A,C). In male groups, only three terms were revealed in the gene-concept heatmaps (Figure 6B,D). Furthermore, a difference in terms was recorded between female and male groups. Two GO terms in male group were found to contain numerous genes (systemic arterial pressure and blood pressure finding).

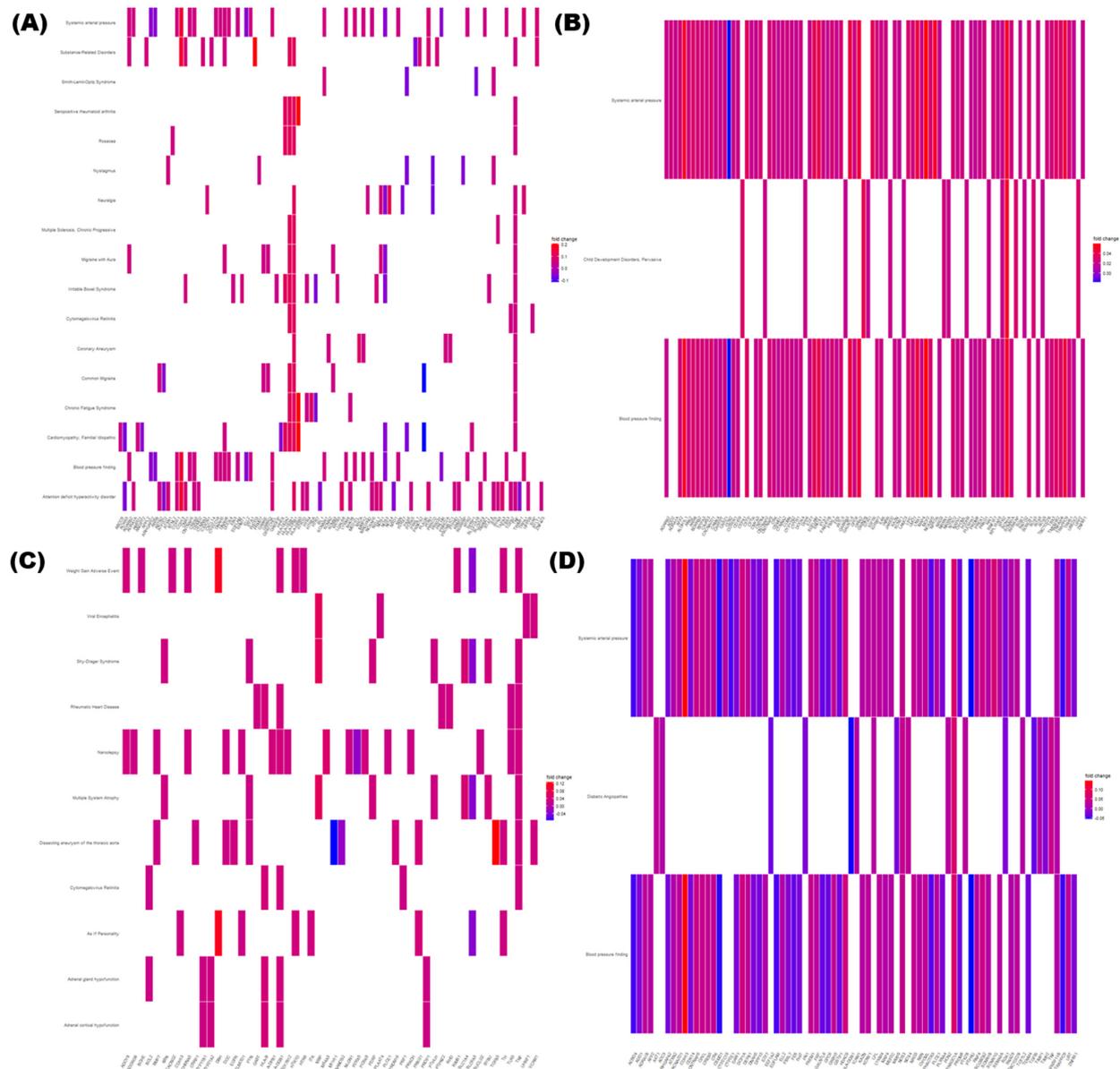


Figure 6. Gene-concept heatmap of the DMRs indicating the difference between four pairs of the four groups. The X-axis represents each gene, and the Y-axis represents the terms corresponding to each gene. Color ranges from red (high methylated in drinkers) to blue (high methylated in non-drinkers). (A) In females with chronic cardiovascular diseases (CVDs), the pairs of 9 drinkers versus 36 non-drinkers were included in the enrichment test, and each term satisfying the criteria $|\text{fold change}| > 0.08$ and $p\text{-value} < 0.05$ was retrieved. (B) In males with CVD, the pairs of 15 drinkers versus 15 non-drinkers were included in the enrichment test, and each term satisfying the criteria $|\text{fold change}| > 0.08$ and $p\text{-value} < 0.05$ was retrieved. (C) In females with non-CVD, the pairs of 42 drinkers versus 123 non-drinkers were subjected to the enrichment test, and each term satisfying the criteria $|\text{fold change}| > 0.04$ and $p\text{-value} < 0.05$ was retrieved. (D) In males with non-CVD, the pairs of 136 drinkers versus 47 non-drinkers were subjected to the enrichment test, and each term satisfying the criteria $|\text{fold change}| > 0.04$ and $p\text{-value} < 0.05$ was retrieved.

3.7. Machine Learning Approaches for Analyzing DMRs

Random forest analysis was performed to describe classification criteria such as decision tree (Figure 7). Two plots were obtained from the random forest model. In the first plot, the number of trees grown for random forest models of four groups was ranked by permutation accuracy importance. In the plot, the x axis indicates the number of grown trees ($n = 500$), and the y axis represents the out-of-bag classification error. Four random forest models were designed, and the number of pre-training tree models was 500. The second plot was a variable importance plot. The top 30 genes retrieved after a random forest analysis were ranked as per the classification parameters. X axis indicates “mean decrease in Gini coefficients”. Gini means inequity, and Gini importance was calculated by the average gain of purity from the given genes. Similar to a high Gini importance, high-weighted genes indicate that the impurity at the decision node is lower at the bottom node.

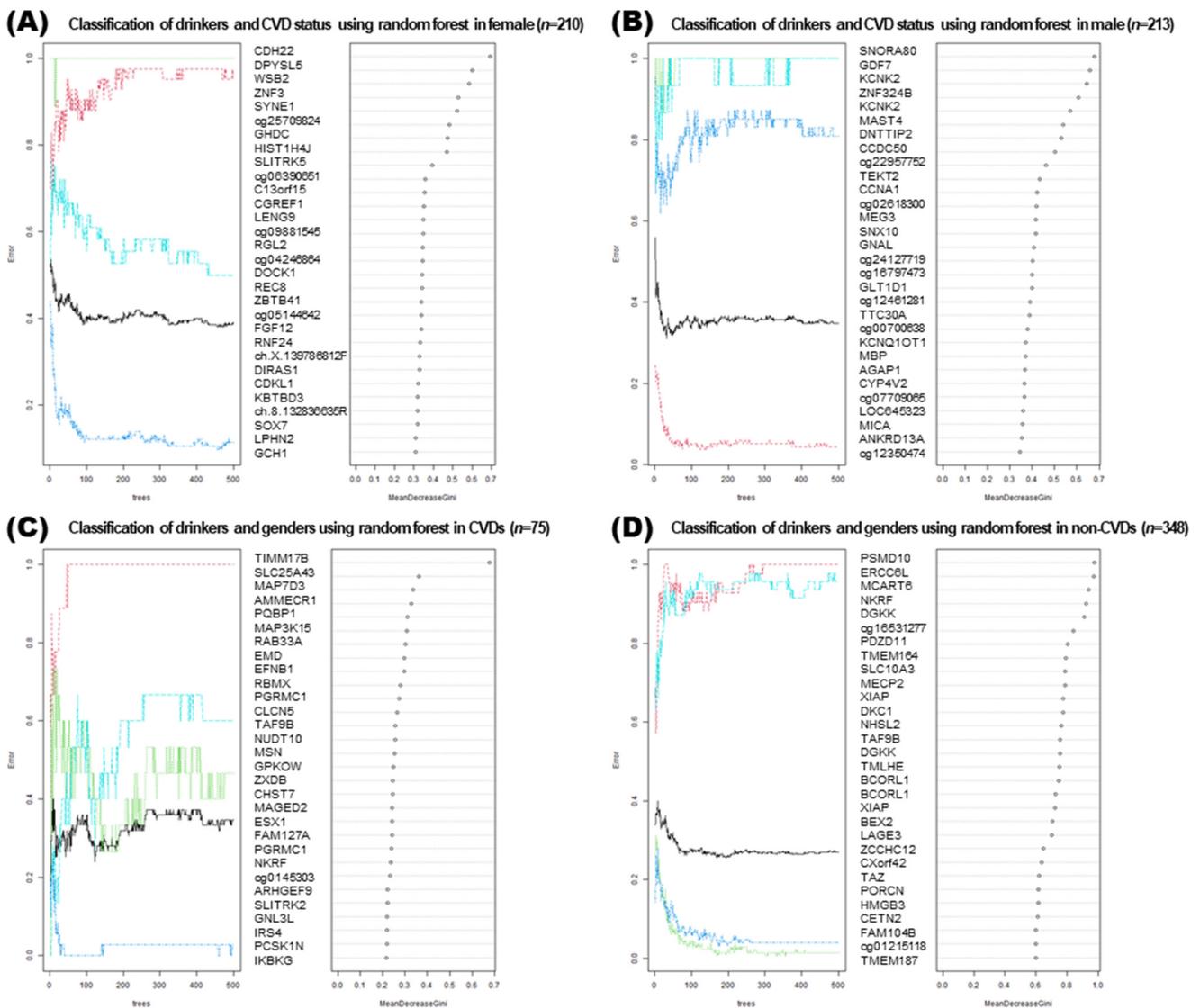


Figure 7. Random forest analysis results classifying the four groups from 423 samples. Each of the four random forest outputs was provided as an error plot and variable importance plot. Three conditions (sex, drinking, and chronic cardiovascular diseases; CVDs) were considered as criteria for classification. Subjects afflicted with at least one of the diseases, including myocardial infarction, coronary artery, hyperlipidemia, and hypertension, were included as subjects with CVDs. Variable importance plot of the random forest analysis resulted from the integration of a large number of

models built by classification of each of the four groups. The variables are ordered top-to-bottom as most-to-least important in classifying the four groups. The ranked list of variables indicates the importance of each variable in classifying the data. The figure shows the top 30 variables important in the classification of the four groups. Four groups were divided based on the drinking and CVD status (A) from 210 female samples and (B) from 213 male samples. Similarly, four groups each were divided based on the drinking status and sex (C) from 75 CVD samples and (D) from 348 non-CVD samples.

In the 210 female subjects, the presence of drinking and CVD were used as factors to classify four groups by random forest. *CDH22*, *DPYSL5*, and *WSB2* were identified as the top three high weighted genes. In the random forest analysis, DMRs that classify four categories based on two variables were extracted. In 348 subjects belonging to the non-CVD group, five genes with a high mean decrease in the Gini coefficient were observed in the genes that suitably explained drinkers and the two sexes (*PSMD10*, *ERCC6L*, *MCART6*, *NKRF*, and *DGKK*).

4. Discussion

4.1. DNA Methylation in Male vs. Female

EWAS are a powerful way to understand the relationships between epigenetic variation and human diseases. In this study, we investigated the epigenetic differences associated with drinking status, and provided DMRs related to alcoholism in two sexes [29,30]. Differential methylation patterns of the two X chromosomes lead to distinction between the inactive X and active X chromosomes [30]. Since the males have one Y and one X chromosomes, the overall methylation patterns of the X chromosomes of both sexes are distinct. A few studies have attempted to compensate for these X chromosome differences [31,32]. In addition, besides sex chromosomes, thousands of CpG sites on autosomes also show dissimilar DNA methylation patterns between males and females [33,34]. Hence, the two sexes are considered an important covariate while undertaking methylation and phenotype association studies. In this study, we attempted to separate the two sexes during classification of the groups to eliminate the bias arising from using information from both sexes.

When extracting DMR from each of the four groups, similar numbers were observed when fold change and *p*-value criteria were similar. This can be visualized in Figures 3 and 4, and the genes included in the DMR, presented in Figure 4, will be helpful for future research related to alcohol consumption and CVD. In the KEGG pathway enrichment analysis, terms with a similar pattern were enriched for both sexes of the CVD and non-CVD groups, respectively (Figure 5).

4.2. Alcohol Consumption and DNA Methylation

DNA methylation may play a role in the progression from normative to problematic drinking and is responsible for several adverse health outcomes associated with alcohol misuse [32,33]. Several studies have attempted EWAS of alcohol consumption by examining alcohol-related DNA methylation patterns among young adults at a heightened risk of alcohol use and associated health problems [19,34,35]. Previous research also showed that the process of alcoholic cirrhosis may directly influence DNA methylation [36]. Furthermore, alcohol cirrhosis has been associated with a multitude of comorbidities, including cardiovascular disease [37].

We divided two groups according to the presence or absence of CVDs, and then obtained genotypes and DMRs that were significantly different according to alcohol consumption in each group. By dividing each group according to sex when identifying DMR, sex-specific bias was removed and each sex-specific pattern was presented. The enriched KEGG pathways and gene-concept heatmap show that the presence of CVD reflects the enriched terms better than the sex (Figures 4 and 6). In four random forest analyses, it was confirmed that the four groups could be classified with similar grades (Gini coefficient) using DMRs (Figure 7). A high mean decrease in the Gini coefficient was observed in the non-CVD group, because the non-CVD group had the highest number of subjects, that is,

348. Although there were 75 subjects with CVD, DMRs were presented to explain male and female, and drinker status. Among them, *TIMM17B* showed the highest Gini coefficient. For genes such as *CDH22*, *DPYSL5*, *WSB2*, *SNORA80*, *GDK7*, and *KCNK2*, which are located in the top of the list, follow-up studies related to alcohol use and CVD are likely necessary.

Therefore, if we combine our results with previous study results, the presence or absence of CVD will be an important relevant factor when studying DNA methylation for alcohol use.

4.3. SNP and DNA Methylation

The GWAS and EWAS conducted to date have discovered genotypes and DMRs that explain alcohol use and CVD. Techniques such as methylation quantitative trait loci (mQTL) for correlating SNPs and DMRs have been introduced and applied in several studies [38,39]. In particular, studies that integrate SNPs and DMRs have been conducted in many cases related to mental health. As it is difficult to explain mental health as a complex phenomenon with a single omics data, several integrated studies have been performed.

Although this study did not progress to mQTL, there was no commonality between the genotypes and DMRs discovered in our classification group. If the factors related to alcohol use and CVD in our study are identified in future studies and mQTL analysis is performed, a stronger explanation can be provided.

4.4. Limitations of This Study

A strength of our study was the discovery of genotypes and DMRs according to alcohol use and CVD in Koreans using various methods. However, our study has the following limitations, which should be further investigated.

First, when classifying drinkers and non-drinkers, former drinkers were included as non-drinkers. In a previous study, DMRs were classified into three groups, but statistically significant results were not obtained due to the limitation of the number of samples. Currently, the KoGES cohort is being expanded, and methylation analysis will be additionally performed. In the future, DMRs for the three groups will be extracted and compared with the results of this study.

Second, there were no common results between genotypes and DMRs, and mQTL analysis was not performed. This will be further analyzed and verified in future extended cohorts. Finally, there were more drinkers than non-drinkers in non-CVD males. In addition, drinkers showed relatively high education and income among males. There was a bias in cohort collection, and it is expected to be corrected with more cohorts in the future.

5. Conclusions

In summary, we show that genetic variation in *HECTD4* is associated with CVD risk in the presence and absence of alcoholism and that alcoholism significantly influences the pattern of DMR methylation associated with CVD in DNA prepared from whole blood. DMRs from four groups based on sex and CVDs were identified and compared by drinking status. Statistical methods (*t*-test), KEGG enrichment analysis, and random forests were used to obtain DMRs for classification. In CVD, enriched terms were found to be related to neuron and MAPK pathway in non-CVD. This study is expected to provide important clues regarding the relationship of GWAS and EWAS with alcohol use and CVD.

Author Contributions: M.J., S.-B.K. and J.-A.G.; methodology, M.J., S.-B.K. and J.-A.G.; software, J.-A.G.; validation, M.J., Y.-S.A. and S.-J.C.; formal analysis, C.-B.K. and K.S.J.; investigation, M.J. and S.-B.K.; resources, M.J. and S.-B.K.; data curation, M.J. and Y.-S.A.; writing—original draft preparation, M.J.; writing—review and editing, S.-B.K. and J.-A.G.; visualization, M.J. and J.-A.G.; supervision, S.-B.K.; project administration, M.J. and S.-B.K.; funding acquisition, S.-B.K. and J.-A.G. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the Basic Science Research Program through the National Research Foundation (NRF) funded by the Ministry of Education, grant number NRF-2020R1I1A1A01052701.

Institutional Review Board Statement: This study was approved by the Institutional Review Board (IRB) of Yonsei Wonju Severance Christian Hospital (Approval Numbers: CR320355) and Korea University (Approval Numbers: KUIRB-2020-0191-01) by accordance with the Declaration of Helsinki.

Informed Consent Statement: Informed consent was obtained from all subjects involved in the study.

Data Availability Statement: KoGES dataset information and data shearing processes can be obtained from the site provided by the National Research Institute of Health, the Korea Disease Control and Prevention Agency, Ministry for Health and Welfare, Korea (<https://nih.go.kr/contents.es?mid=a50401010100>, accessed on 6 December 2021). R source codes will be provided upon request.

Acknowledgments: Data used in this study were from KoGES, the National Research Institute of Health, the Korea Disease Control and Prevention Agency, Ministry for Health and Welfare, Republic of Korea.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. GBD 2015 Risk Factors Collaborators. Global, regional, and national comparative risk assessment of 79 behavioural, environmental and occupational, and metabolic risks or clusters of risks, 1990–2015: A systematic analysis for the Global Burden of Disease Study 2015. *Lancet* **2016**, *388*, 1659–1724. [[CrossRef](#)]
2. Alkerwi, A.; Boutsen, M.; Vaillant, M.; Barre, J.; Lair, M.L.; Albert, A.; Guillaume, M.; Dramaix, M. Alcohol Consumption and the Prevalence of Metabolic Syndrome: A Meta-Analysis of Observational Studies. *Atherosclerosis* **2009**, *204*, 624–635. [[CrossRef](#)] [[PubMed](#)]
3. Murray, C.J.L.; Lopez, A.D. *The Global Burden of Disease: A Comprehensive Assessment of Mortality and Disability from Disease Injuries and Risk Factors in 1990 and Projected to 2002*; Harvard School of Public Health: Boston, MA, USA, 1996.
4. Yusuf, S.; Reddy, S.; Ounpuu, S.; Anand, S. Global burden of cardiovascular diseases: Part I: General considerations, the epidemiologic transition, risk factors, and impact of urbanization. *Circulation* **2001**, *104*, 2746–2753. [[CrossRef](#)]
5. Mozaffarian, D.; Benjamin, E.J.; Go, A.S.; Arnett, D.K.; Blaha, M.J.; Cushman, M.; de Ferranti, S.; Despres, J.P.; Fullerton, H.J.; Howard, W.J.; et al. Heart disease and stroke statistics—2015 update: A report from the American Heart Association. *Circulation* **2015**, *131*, e29–e322. [[CrossRef](#)]
6. Sipos, B.; Jirak, P.; Paar, V.; Rezar, R.; Mirna, M.; Kopp, K.; Hoppe, U.C.; Berezin, A.E.; Lichtenauer, M. Promising Novel Biomarkers in Cardiovascular Diseases. *Appl. Sci.* **2021**, *11*, 3654. [[CrossRef](#)]
7. Vrablik, M.; Dlouha, D.; Todorovova, V.; Stefler, D.; Hubacek, J.A. Genetics of Cardiovascular Disease: How Far Are We from Personalized CVD Risk Prediction and Management? *Int. J. Mol. Sci.* **2021**, *22*, 4182. [[CrossRef](#)] [[PubMed](#)]
8. Gheorghe, A.; Griffiths, U.; Murphy, A.; Legido-Quigley, H.; Lamptey, P.; Perel, P. The economic burden of cardiovascular disease and hypertension in low- and middle-income countries: A systematic review. *BMC Public Health* **2018**, *18*, 975. [[CrossRef](#)]
9. Hurt, R.D.; Offord, K.P.; Croghan, I.T.; Gomez-Dahl, L.; Kottke, T.E.; Morse, R.M.; Melton, L.J., III. Mortality following inpatient addictions treatment. *JAMA* **1996**, *275*, 1097–1103. [[CrossRef](#)]
10. Jarvis, C.M.; Hayman, L.L.; Braun, L.T.; Schwertz, D.W.; Ferrans, C.E.; Piano, M.R. Cardiovascular Risk Factors and Metabolic Syndrome in Alcohol- and Nicotine-Dependent Men and Women. *J. Cardiovasc. Nurs.* **2007**, *22*, 429–435. [[CrossRef](#)]
11. Harada, S.; Agarwal, D.P.; Goedde, H.W.; Tagaki, S.; Ishikawa, B. Possible Protective Role Against Alcoholism for Aldehyde Dehydrogenase Isozyme Deficiency in Japan. *Lancet* **1982**, *2*, 827. [[CrossRef](#)]
12. Cook, W.K.; Tam, C.C.; Luczak, S.E.; Kerr, W.C.; Mulia, N.; Lui, C.; Li, L. Alcohol Consumption, Cardiovascular-Related Conditions, and ALDH2*2 Ethnic Group Prevalence in Asian Americans. *Alcohol. Clin. Exp. Res.* **2021**, *45*, 418–428. [[CrossRef](#)]
13. Portela, A.; Esteller, M. Epigenetic modifications and human disease. *Nat. Biotechnol.* **2010**, *28*, 1057–1068. [[CrossRef](#)] [[PubMed](#)]
14. Zhong, J.; Agha, G.; Baccarelli, A.A. The Role of DNA Methylation in Cardiovascular Risk and Disease: Methodological Aspects, Study Design, and Data Analysis for Epidemiological Studies. *Circ. Res.* **2016**, *118*, 119–131. [[CrossRef](#)]
15. Prasher, D.; Greenway, S.C.; Singh, R.B. The impact of epigenetics on cardiovascular disease. *Biochem. Cell. Biol.* **2020**, *98*, 12–22. [[CrossRef](#)]
16. Kim, M.; Long, T.I.; Arakawa, K.; Wang, R.; Yu, M.C.; Laird, P.W. DNA methylation as a biomarker for cardiovascular disease risk. *PLoS ONE* **2010**, *5*, e9692. [[CrossRef](#)]
17. Sharma, P.; Kumar, J.; Garg, G.; Kumar, A.; Patowary, A.; Karthikeyan, G.; Ramakrishnan, L.; Brahmachari, V.; Sengupta, S. Detection of altered global DNA methylation in coronary artery disease patients. *DNA Cell Biol.* **2008**, *27*, 357–365. [[CrossRef](#)]
18. Kaminsky, Z.A.; Tang, T.; Wang, S.C.; Ptak, C.; Oh, G.H.; Wong, A.H.; Feldcamp, L.A.; Virtanen, C.; Halfvarson, J.; Tysk, C.; et al. DNA methylation profiles in monozygotic and dizygotic twins. *Nat. Genet.* **2009**, *41*, 240–245. [[CrossRef](#)] [[PubMed](#)]
19. Liu, C.; Marioni, R.E.; Hedman, Å.K.; Pfeiffer, L.; Tsai, P.-C.; Reynolds, L.M.; Just, A.C.; Duan, Q.; Boer, C.G.; Tanaka, T.; et al. A DNA methylation biomarker of alcohol consumption. *Mol. Psychiatry* **2018**, *23*, 422–433. [[CrossRef](#)] [[PubMed](#)]
20. Bönsch, D.; Lenz, B.; Fiszer, R.; Frieling, H.; Kornhuber, J.; Bleich, S. Lowered DNA methyltransferase (DNMT-3b) mRNA expression is associated with genomic DNA hypermethylation in patients with chronic alcoholism. *J. Neural Transm.* **2006**, *113*, 1299–1304. [[CrossRef](#)] [[PubMed](#)]

21. Bönsch, D.; Lenz, B.; Reulbach, U.; Kornhuber, J.; Bleich, S. Homocysteine associated genomic DNA hypermethylation in patients with chronic alcoholism. *J. Neural Transm.* **2004**, *111*, 1611–1616. [[CrossRef](#)]
22. Huan, T.; Joehanes, R.; Song, C.; Peng, F.; Guo, Y.; Mendelson, M.; Yao, C.; Liu, C.; Ma, J.; Richard, M.; et al. Genome-wide identification of DNA methylation QTLs in whole blood highlights pathways for cardiovascular disease. *Nat. Commun.* **2019**, *10*, 4267. [[CrossRef](#)]
23. Monti, M.C.; Lonsdale, J.T.; Montomoli, C.; Montross, R.; Schlag, E.; Greenberg, D.A. Familial risk factors for microvascular complications and differential male-female risk in a large cohort of American families with type 1 diabetes. *J. Clin. Endocrinol. Metab.* **2007**, *92*, 4650–4655. [[CrossRef](#)]
24. Cotton, A.M.; Price, E.M.; Jones, M.J.; Balaton, B.P.; Kobor, M.S.; Brown, C.J. Landscape of DNA methylation on the X chromosome reflects CpG density, functional chromatin state and X-chromosome inactivation. *Hum. Mol. Genet.* **2015**, *24*, 1528–1539. [[CrossRef](#)]
25. Kim, Y.; Han, B.G.; KoGES Group. Cohort profile: The Korean genome and epidemiology study (KoGES) consortium. *Int. J. Epidemiol.* **2017**, *46*, e20. [[CrossRef](#)] [[PubMed](#)]
26. Kim, Y.; Kim, J.; Lim, J.E.; Oh, B.; Won, S.; Kim, M.K. Genome Wide Interaction Study of Single-Nucleotide Polymorphisms and Alcohol Consumption on Blood Pressure the Ansan and Ansong Study of the Korean Genome and Epidemiology Study (KoGES). *Genet. Epidemiol.* **2020**, *44*, 300–310. [[CrossRef](#)]
27. Song, Y.; Choi, J.E.; Kwon, Y.J.; Chang, H.J.; Kim, J.O.; Park, D.H.; Park, J.M.; Kim, S.J.; Lee, J.W.; Hong, K.W. Identification of Susceptibility Loci for Cardiovascular Disease in Adults with Hypertension, Diabetes, and Dyslipidemia. *J. Transl. Med.* **2021**, *19*, 85. [[CrossRef](#)]
28. Azman, N.F.; Abdullah, W.Z.; Hanafi, S.; Diana, R.; Bahar, R.; Johan, M.F.; Zilfalil, B.A.; Hassan, R. Genetic polymorphisms of HbE/beta thalassemia related to clinical presentation: Implications for clinical diversity. *Ann. Hematol.* **2020**, *99*, 729–735. [[CrossRef](#)]
29. Rakyán, V.K.; Down, T.A.; Balding, D.J.; Beck, S. Epigenome-wide association studies for common human diseases. *Nat. Rev. Genet.* **2011**, *12*, 529–541. [[CrossRef](#)] [[PubMed](#)]
30. Sharp, A.J.; Stathaki, E.; Migliavacca, E.; Brahmachary, M.; Montgomery, S.B.; Dupre, Y.; Antonarakis, S.E. DNA methylation profiles of human active and inactive X chromosomes. *Genome Res.* **2011**, *21*, 1592–1600. [[CrossRef](#)]
31. Wang, Y.; Zhang, Y.; Zhang, J.; Tang, X.; Qian, Y.; Gao, P.; Zhu, D. Association of a functional single-nucleotide polymorphism in the ALDH2 gene with essential hypertension depends on drinking behavior in a Chinese Han population. *J. Hum. Hypertens.* **2013**, *27*, 181–186. [[CrossRef](#)] [[PubMed](#)]
32. Duncan, C.G.; Grimm, S.A.; Morgan, D.L.; Bushel, P.R.; Bennett, B.D.; NISC Comparative Sequencing Program; Roberts, J.D.; Tyson, F.L.; Merrick, B.A.; Wade, P.A. Dosage compensation and DNA methylation landscape of the X chromosome in mouse liver. *Sci. Rep.* **2018**, *8*, 10138. [[CrossRef](#)]
33. Liu, J.; Morgan, M.; Hutchison, K.; Calhoun, V.D. A study of the influence of sex on genome wide methylation. *PLoS ONE* **2010**, *5*, e10028. [[CrossRef](#)] [[PubMed](#)]
34. Yousefi, P.; Huen, K.; Davé, V.; Barcellos, L.; Eskenazi, B.; Holland, N. Sex differences in DNA methylation assessed by 450K BeadChip in newborns. *BMC Genom.* **2015**, *16*, 911. [[CrossRef](#)]
35. Ward-Caviness, C.K.; Agha, G.; Chen, B.H.; Pfeiffer, L.; Wilcon, R.; Wolf, P.; Gieger, C.; Schwartz, J.; Vokonas, P.S.; Hou, L.; et al. Analysis of repeated leukocyte DNA methylation assessments reveals persistent epigenetic alterations after an incident myocardial infarction. *Clin. Epigenet.* **2018**, *10*, 161. [[CrossRef](#)]
36. Stephenson, M.; Bollepalli, S.; Cazaly, E.; Salvatore, J.E.; Barr, P.; Rose, R.J.; Dick, D.; Kaprio, J.; Ollikainen, M. Associations of Alcohol Consumption with Epigenome-Wide DNA Methylation and Epigenetic Age Acceleration: Individual-Level and Co-twin Comparison Analyses. *Alcohol Clin. Exp. Res.* **2021**, *45*, 318–328. [[CrossRef](#)]
37. Rosen, A.D.; Robertson, K.D.; Hlady, R.A.; Muench, C.; Lee, J.; Philibert, R.; Horvath, S.; Kaminsky, Z.A.; Lohoff, F.W. DNA methylation age is accelerated in alcohol dependence. *Transl. Psychiatry* **2018**, *8*, 182. [[CrossRef](#)] [[PubMed](#)]
38. Zhao, T.; Hu, Y.; Zang, T.; Wang, Y. Integrate GWAS, eQTL, and mQTL data to identify Alzheimer’s disease-related genes. *Front. Genet.* **2019**, *10*, 1021. [[CrossRef](#)] [[PubMed](#)]
39. Ciuculete, D.M.; Boström, A.E.; Voisin, S.; Philipps, H.; Titova, O.E.; Bandstein, M.; Nikontovic, L.; Williams, M.J.; Mwinyi, J.; Schiöth, H.B. A methylome-wide mQTL analysis reveals associations of methylation sites with GAD1 and HDAC3 SNPs and a general psychiatric risk score. *Transl. Psychiatry* **2017**, *7*, e1002. [[CrossRef](#)] [[PubMed](#)]