

ПЛАТФОРМА СИСТЕМИ І ТЕХНОЛОГІЇ ШТУЧНОГО ІНТЕЛЕКТУ. РОБОТОТЕХНІКА. ГНУЧКІ СИСТЕМИ.

Mykhailo Krolevets^{1, 2, 3}

Vincent ten Cate^{1, 4, 5}

Jürgen H. Prochaska^{1, 4, 5}

Andreas Schulz¹

Steffen Rapp^{1, 4}

Stefan Tenzer⁸

Miguel A. Andrade-Navarro⁷

Steve Horvath⁶

Christof Niehrs²

Philipp S. Wild^{1, 3, 4, 5}

^{1.} Preventive Cardiology and Preventive Medicine, Department of Cardiology, University Medical Center of the Johannes Gutenberg University Mainz, Langenbeckstr. 1, 55131 Mainz, Germany

^{2.} Institute of Molecular Biology (IMB), 55128 Mainz, Germany and Division of Molecular Embryology, DKFZ-ZMBH Alliance, 69120 Heidelberg, Germany

^{3.} Systems medicine, Institute of Molecular Biology (IMB), Ackermannweg 4, 55128 Mainz, Germany

^{4.} Clinical Epidemiology and Systems Medicine, Center for Thrombosis and Hemostasis (CTH), Mainz, Germany

^{5.} German Center for Cardiovascular Research (DZHK), Partner Site Rhine Main, University Medical Center of the Johannes Gutenberg University Mainz, Mainz, Germany

^{6.} Altos Labs, San Diego, USA

^{7.} Institute for Immunology, University Medical Center of the Johannes Gutenberg-University Mainz, Mainz, Germany

^{8.} Institute of Organismic and Molecular Evolution, Johannes Gutenberg-University Mainz, Mainz, Germany

A SYSTEMATIC REVIEW AND DATABASE OF KNOWN CPG METHYLATION SITES

Introduction

Cardiovascular disease (CVD) is the leading cause of death globally, accounting for approximately 30% of all deaths worldwide [1]. This puts an enormous burden on healthcare systems across the globe and has prompted clinicians and researchers to explore potential causes of this widespread disease. The World Health Organization recognizes CVD as one of the most environmentally driven diseases with a comparatively small genetic component [2]. It is likely that epigenetic changes mediate, at least in part, the environmental risk for developing or progressing CVD. One prominent factor that is thought to play a role is DNA methylation, an epigenetic mark that can modify gene expression. DNA methylation usually refers to the methylation of the fifth carbon of cytosine residues found throughout the genome. Methylation is most commonly observed at so-called CpG islands [3], which are long repeats of cytosine-guanine nucleotides. As technologies to study genome-wide DNA methylation continue to develop, there is growing evidence that there is a strong link between DNA methylation and CVD [4–6]. However, there is still little to no evidence of the directionality of this relationship, and clear findings on relevant CpG sites or genes have not yet emerged from the extensive research in this area. The main aim of this systematic review was to synthesize results of studies that have investigated the relationship between DNA methylation and CVD and to create an easily accessible and searchable database from the results of these studies. Based on this database, we identified overlapping differentially methylated CpG sites and neighboring genes across studies and performed functional enrichment and interaction network analyses. Based on these results, we highlighted pathways that may be

involved in the development or progression of CVD through a mechanism that is associated with DNA methylation changes.

Results

Identified studies on DNA methylation in cardiovascular disease An overview on the workflow of the systematic review is provided in Fig. 1.

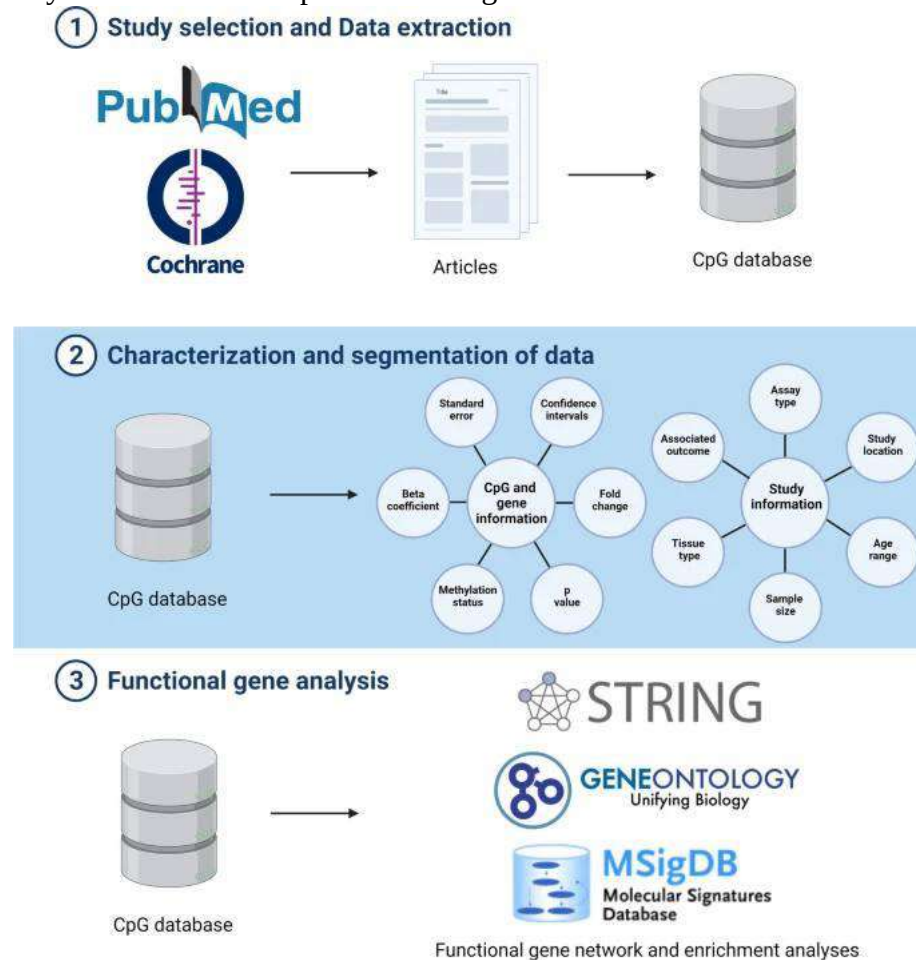


Fig.1 Workflow for the systematic review

Using the a priori defined search queries, a total of 5327 articles was identified in PubMed and 236 articles in CENTRAL. After irrelevant articles were removed, a total of 207 studies were assessed for relevance by reading the full text. A total of 99 articles [7] were considered eligible for systematic review (Fig. 2).

After grading for risk of bias, the majority of publications (n=93) was assigned a study quality metric of “fair” using the NHLBI quality assessment tools, with two studies graded “poor” and four graded “good.” Eighty-five articles provided direct data comparing CpG sites and genes (beta-estimates of methylation level, p-value, standard errors, etc.), nine measured global DNA methylation (i.e., total hyper- or hypomethylation), and eight studies used a methylation-based clock without directly reporting summary statistics at the CpG site level.

A total of 87,827 individuals was included in the 99 selected studies. Across studies that reported the sex distribution of participants, the overall proportion of females was 54% and that of males was 46%. DNA methylation measurements were performed using a variety of assays including the Illumina Infinium® HumanMethylation 450 BeadChip (Illumina, California, USA), Infinium® HumanMethylation 850 BeadChip (Illumina, California, USA), Pyrosequencing, whole genome bisulfite sequencing, and others. Study cohorts were located in four continents: Europe, Asia, North and South America, and Australia. Samples for DNA methylation measurements were extracted from a variety of tissues, with whole blood used in the majority of cases. Since the type of outcome varied widely across the studies, they were grouped into the following five main categories for analysis: ‘CVD’, ‘heart’, ‘vasculature’, ‘cardiovascular risk factor (CVRf)’ and ‘others’. Study

designs were grouped as follows: ‘investigation of epigenome-wide methylation’, ‘global genome methylation’, ‘specific CpG site methylation’, ‘specific gene methylation’ and ‘epigenetic clock methylation’.

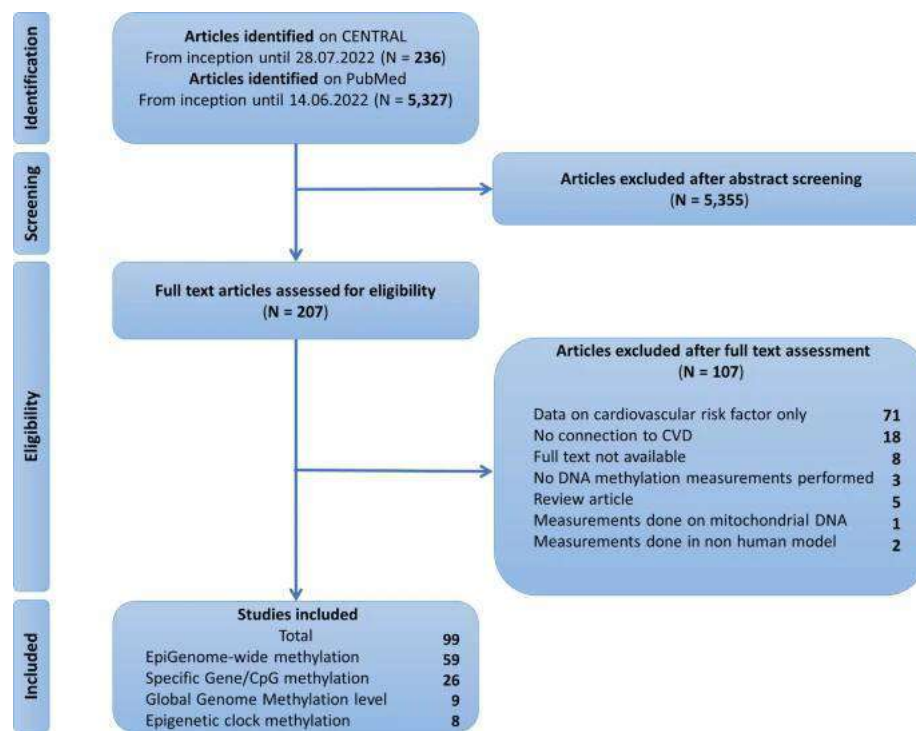


Fig.2 PRISMA flowchart. CVD—cardiovascular disease

A meta-analysis of the collected data could not be performed since the endpoints, methods and reported estimates of the individual studies could not be sufficiently reconciled.

Database of CpG sites associated with CVD

A CpG database was created by aggregating all CpG and gene-related information from the collected studies together with three methylation clocks: Horvath [8], Hannum [9] and GrimAge [10]. The database includes 74,580 unique CpG entries (73,550 entries without the methylation clock “GrimAge”) and 19,127 gene entries (18,374 entries without the methylation clock “GrimAge”). For each CpG/gene, there is information on the outcome, follow-up, cohort, measurement method, location of study, sample size, tissue analyzed, age range and related publication. For each entry, additional quantitative information such as methylation beta value, p-value, z-score, standard error, false discovery rate, etc. is provided if the information was reported in the publication.

Two sites were referenced in ≥ 6 publications: cg01656216 (near ZNF438) related to vascular disease and epigenetic age, and cg03636183 (near F2RL3) related to coronary heart disease, myocardial infarction, smoking and air pollution. Of 19,127 mapped genes, 5,807 were reported in ≥ 2 studies. Most frequently reported were TEAD1 (TEA Domain Transcription Factor 1) and PTPRN2 (Protein Tyrosine Phosphatase Receptor Type N2) in association with outcomes ranging from vascular to cardiac disease. Gene set enrichment analysis of 4,532 overlapping genes revealed enrichment for Gene Ontology molecular function “DNA-binding transcription activator activity” ($q = 1.65 \times 10^{-11}$) and biological processes “skeletal system development” ($q = 1.89 \times 10^{-23}$). Gene enrichment demonstrated that general CVD-related terms are shared, while “heart” and “vasculature” specific genes have more disease-specific terms as PR interval for “heart” or platelet distribution width for “vasculature.” STRING analysis revealed significant protein–protein interactions between the products of the differentially methylated genes ($p = 0.003$) suggesting that dysregulation of the protein interaction network could contribute to CVD. Overlaps with curated gene sets from the Molecular Signatures Database showed enrichment of genes in hemostasis ($p = 2.9 \times 10^{-6}$) and atherosclerosis ($p = 4.9 \times 10^{-4}$).

Conclusion

This review highlights the current state of knowledge on significant relationship between DNA methylation and CVD in humans. An open-access database has been compiled of reported CpG methylation sites, genes and pathways that may play an important role in this relationship.

Availability of data and materials

All data collected, generated, or analyzed as part of this study, except for data from the GrimAge methylation clock, are included in this published article, or available as additional files.

Abbreviations

AMI: Acute myocardial infarction

ACS: Acute coronary syndrome

CpG: Cytosine-phosphate-guanine

CVD: Cardiovascular disease

CVRF: CVD risk factor

DNA: Deoxyribonucleic acid

DNAm: Deoxyribonucleic acid methylation

GO: Gene ontology

GOBP: Gene ontology biological process

HP: Human phenotype

PRISMA: Preferred reporting items for systematic reviews and meta-analyses

References

1. Bloom DE, Jané-Llopis E, Abrahams-Gessel S, Bloom LR, Fathima S, et al. Global economic burden non communicable diseases. Geneva: World Economic Forum.
2. Bhatnagar A. Environmental determinants of cardiovascular disease. *Circ Res.* 2017;121(2):162–80.
3. Jang H, Shin W, Lee J, Do J. CpG and non-CpG methylation in epigenetic gene regulation and brain function. *Genes.* 2017;8(6):148.
4. Chilunga FP, Henneman P, Venema A, Meeks KAC, Requena-Méndez A, Beune E, et al. Genome-wide DNA methylation analysis on C-reactive protein among Ghanaians suggests molecular links to the emerging risk of cardiovascular diseases. *Npj Genom Med.* 2021;6(1):46.
5. Ramos RB, Fabris V, Lecke SB, Maturana MA, Spritzer PM. Association between global leukocyte DNA methylation and cardiovascular risk in postmenopausal women. *BMC Med Genet.* 2016;17(1):71.
6. Kazmi N, Elliott HR, Burrows K, Tillin T, Hughes AD, Chaturvedi N, et al. Associations between high blood pressure and DNA methylation. *PLoS ONE.* 2020;15(1): e0227728.
7. Krolevets, M., Cate, V.t., Prochaska, J.H. et al. DNA methylation and cardiovascular disease in humans: a systematic review and database of known CpG methylation sites. *Clin Epigenet* 15, 56 (2023). <https://doi.org/10.1186/s13148-023-01468-y>
8. Horvath S. DNA methylation age of human tissues and cell types. *Genome Biol.* 2013;14(10):R115.
9. Hannum G, Guinney J, Zhao L, Zhang L, Hughes G, Sadda S, et al. Genome-wide methylation profiles reveal quantitative views of human aging rates. *Mol Cell.* 2013;49(2):359–67.
10. Lu AT, Quach A, Wilson JG, Reiner AP, Aviv A, Raj K, et al. DNA methylation GrimAge strongly predicts lifespan and healthspan. *Aging.* 2019;11(2):303–27.