Title

Gene-Environment Interactions for Cardiovascular Disease.

Journal

Current Atherosclerosis Reports, 23(12)

Authors

Hartiala, Jaana
Hilser, James
Biswas, Subarna
et al.

Publication Date

2021-10-14

DOI

10.1007/s11883-021-00974-9

Peer reviewed

HHS Public Access

Author manuscript

Curr Atheroscler Rep. Author manuscript; available in PMC 2022 October 14.

Published in final edited form as:

Curr Atheroscler Rep. ; 23(12): 75. doi:10.1007/s11883-021-00974-9.

Gene-Environment Interactions for Cardiovascular Disease

Jaana A. Hartiala, PhD[1], **James R. Hilser, MPH**[1,2], **Subarna Biswas, MD**[1,3], **Aldons J. Lusis, PhD**[4,5,6], **Hooman Allayee, PhD**[1,2]

[1]Department of Population and Public Health Sciences, Keck School of Medicine, University of Southern California, Los Angeles, CA 90033

[2]Department of Biochemistry and Molecular Medicine, Keck School of Medicine, University of Southern California, Los Angeles, CA 90033

[3]Department of Surgery, Keck School of Medicine, University of Southern California, Los Angeles, CA 90033

[4]Department of Medicine, David Geffen School of Medicine of UCLA, Los Angeles, CA 90095.

[5]Department of Microbiology, Immunology, and Molecular Genetics, David Geffen School of Medicine of UCLA, Los Angeles, CA 90095.

[6]Department of Human Genetics, David Geffen School of Medicine of UCLA, Los Angeles, CA 90095.

Abstract

Purpose of review: We provide an overview of recent findings with respect to gene-environment (GxE) interactions for cardiovascular disease (CVD) risk and discuss future opportunities for advancing the field.

Recent findings: Over the last several years, GxE interactions for CVD have mostly been identified for smoking and coronary artery disease (CAD) or related risk factors. By comparison, there is more limited evidence for GxE interactions between CVD outcomes and other exposures, such as physical activity, air pollution, diet, and sex. The establishment of large consortia and population-based cohorts, in combination with new computational tools and mouse genetics platforms, can potentially overcome some of the limitations that have hindered human GxE interaction studies and reveal additional association signals for CVD-related traits.

Summary: The identification of novel GxE interactions is likely to provide a better understanding of the pathogenesis and genetic liability of CVD, with significant implications for healthy lifestyles and therapeutic strategies.

Address correspondence and reprint requests to: Hooman Allayee, PhD, Department of Population and Public Health Sciences, Keck School of Medicine of USC, 2250 Alcazar Street, CSC202, Los Angeles, CA 90033, Phone: (323) 442-1736, Fax: (332) 442-2764, hallayee@usc.edu.

Conflicts of Interest

The authors declare that they have no conflict of interest.

Human and Animal Rights and Informed Consent

This article does not contain any studies with human or animal subjects performed by any of the authors.

Keywords

cardiovascular disease; genetics; environmental factors; gene-environment interactions; genome-wide association study

Introduction.

Coronary artery disease (CAD), myocardial infarction (MI), peripheral artery disease (PAD), and stroke (collectively referred to as atherosclerotic cardiovascular disease [CVD]) are the leading causes of death in Western societies [1], even in the contemporary era of high-potency lipid-lowering therapy [2]. The pathophysiological basis of CVD is atherosclerosis, a process that develops slowly over decades without necessarily having overt manifestations. Thus, individuals with CAD or PAD are typically asymptomatic, with the first clinical indication often being an adverse event (i.e. MI, stroke, or death) [3]. To date, elevated plasma lipids, diabetes mellitus, systemic arterial hypertension, inflammation, and cigarette smoking are established causal pathways for atherothrombotic diseases [4], but strategies to lower these traditional risk factors only partially mitigates risk and >50% of patients with an acute cardiovascular event do not exhibit elevated levels of these risk factors. These observations suggest the involvement of additional unrecognized causal mechanisms.

Recent large-scale genetic studies and decades of epidemiological research have established the contribution of heritable susceptibility factors and environmental exposures to risk of CVD. Despite the recognition that gene-environment (GxE) interactions are also likely to play important roles in development of CVD, their identification has been hampered by insufficient power due to small sample sizes and weak genetic effects, imprecise measurement of relevant exposures, and computational limitations [5]. However, the establishment of collaborative consortia comprised of multiple cohorts that in aggregate can now include over 1 million subjects, the availability of population-based biobanks and datasets that have collected environmental, genetic, and clinical data in the same individuals, and newly developed statistical methods will help overcome the challenges of carrying out GxE interaction studies on a large and unbiased scale. In this review, we discuss these advances in the field and highlight recent discoveries that illustrate how successful identification of GxE interactions for CVD and related traits can be achieved.

Role of the Environment in Modulating Risk of CVD.

It is generally accepted that risk of CVD is characterized by heritable susceptibility factors in the context of lifetime exposure to the type of atherogenic environment that has become so prevalent in industrialized societies. This concept is supported by decades of epidemiological research and even in migration studies in which the adverse effect of a Western lifestyle is evident. For example, risk of CVD was shown to progressively increase in Japanese men as function of migrating from Japan to Hawaii and California [6]. While a detailed discussion of the role of environmental factors in mediating CVD risk is beyond the scope of this review, we briefly highlight those components that are relevant to GxE interactions (Table 1) and refer the reader to other recent publications for a more thorough analysis [7].

The best understood environmental risk factors for CVD are perhaps diet, physical activity, and smoking (Table 1). It is well-known that a sedentary lifestyle or consumption of high levels of cholesterol and fats, particularly saturated and *trans* fats, increase risk of CVD whereas regular physical activity and diets rich in fruits and vegetables and non-red meat lean protein are protective [8-10]. The environmental risk factor with the strongest adverse effect on the development of CVD is smoking, which doubles risk even in light smokers [11]. CVD has also been estimated to account for over half of the ~500,000 annual premature deaths that are attributed to smoking [1]. In addition to poor diet and tobacco use, a large body of evidence has also shown consistent associations between CVD and short- or long-term exposure to components of traffic-related air pollution (TRAP), such as coarse, fine, and ultrafine particulate matter (PM) or nitrogen oxides (NO_x) [12, 13]. Notably, after smoking and dietary risk factors, more deaths can be attributed to air pollution than low-density lipoprotein (LDL) cholesterol levels and obesity [14], highlighting the importance of this environmental factor. Lastly, sex can also be considered another variable that can interact with genetic factors since it is well known that men are at higher risk of CVD than women, at least until menopause due to the protective effect of estrogen.

Current State of CVD Genetics.

CVD has a strong genetic component based on heritability estimates that range between ~40-60% [15-18]. These results are corroborated by genome-wide association studies (GWAS), where in the 10 years after the initial identification of the CAD locus on *9p21*, over 100 additional genomic regions were identified for CVD outcomes [19]. However, just in the last four years alone, the number of susceptibility loci for CAD, MI, PAD, and stroke has more than doubled to over 250 [•20-27]. These discoveries have been made possible, in part, through the establishment of collaborative consortia that have harnessed the cumulative power of meta-analyzing data across dozens of case-control cohorts as well as large population-based cohorts, such as the UK Biobank [28], FinnGen [29], and Biobank Japan [30]. The results of these large GWAS provide further evidence that loci influence risk of CVD via perturbations of lipid metabolism, blood pressure regulation, inflammation, and coagulation but an obvious underlying biological mechanism for half or more of the association signals is not readily apparent. Furthermore, the risk alleles at nearly all identified loci are common in the population and only explain a fraction of the overall heritability for CVD outcomes [26]. Thus, it has been postulated that there are other unrecognized contributions from additional common variants, rare, highly penetrant susceptibility alleles, and/or GxE interactions.

If the results of the last few years are any indication, additional common risk alleles for CVD will undoubtedly be revealed by even larger GWAS that are already starting to exceed one million individuals. However, the effect sizes of such alleles will be even weaker than those already identified and therefore not likely to explain a significant fraction of the remaining so-called missing heritability. In this regard, rare variants have also been previously associated with CVD risk through exome chip or sequencing analysis, with some exhibiting large effect sizes. Even so, there have still been far fewer rare variants identified for CVD than common single nucleotide polymorphisms (SNPs) from GWAS [31]. Even with increased sample sizes in more recent rare variant studies [32-35], the number of

novel genes identified for CVD or related risk factors has not appreciably increased. These observations suggest that rare variants may not explain a significant fraction of the genetic risk for CVD, although it is still possible that expansion of rare variant analysis to the entirety of population-based biobanks or more subjects of non-European ancestry could increase the evidence that this class of variants still has unrecognized significant contributions to mediating CVD risk.

Overview of GxE Interactions.

From a statistical perspective, GxE interactions are defined as those where the combined effect of genotype and exposure differs significantly from the additive effects of genotype and exposure (Figure 1). A straightforward example for such synergy would be where the effect of smoking on CAD is significantly greater among carriers of a risk allele than the effect of smoking observed in non-carriers (Figure 1). To test for such an interaction, a multiplicative interaction term (GxE) can be included in a standard statistical model along with the main effects of genotype (G) and exposure (E). Alternatively, genetic analyses can be carried out in exposed (e.g. current or ever smokers) and non-exposed (e.g. never smokers) groups separately, followed by interaction tests to determine whether effect sizes are significantly different between the two exposure groups. As described below, this latter approach has been used in several studies since it is relatively straightforward to implement in multiple cohorts individually and then meta-analyzed together to test for GxE interactions.

GxE Interactions with Smoking.

For the present discussion on GxE interactions for CVD outcomes, we focus primarily on recent studies carried out by large multi-cohort consortia or in population-based cohorts since the sample sizes used in these analyses increase the likelihood of identifying true interactions. Of the various environmental exposures highlighted above (Table 1), several recent GxE interactions have been reported for smoking using both candidate gene and whole-genome approaches (Table 2). In one large study with ~140,000 cases and controls, stratified analyses with known CAD loci revealed a significant GxE interaction where the protective association of a variant (rs7178051) upstream of *ADAMTS7* with CAD was attenuated by 60% in smokers (5% lower risk) compared to non-smokers (12% lower) [••36]. Notably, the basis for the overall reduced risk of CAD observed with rs7178051 was likely due to the protective T allele decreasing *ADAMTS7* mRNA levels based on expression data in human aortic endothelial cells and lymphoblastoid cells lines [••36]. Another important aspect of this study was in providing directionally consistent functional evidence for the observed GxE interaction since cigarette smoke extract was shown to upregulate *ADAMTS7* mRNA levels in vascular wall cells [••36]. More recently, a moderately linked *ADAMTS7* variant (rs3825807) was also shown to affect CAD risk through a gene-smoking interaction [37]. Importantly, rs3825807 encodes a Ser214Pro substitution that decreases ADAMTS7 function [38] and was associated with a greater reduction in CAD risk among never smokers compared to ever smokers [37]. Thus, the GxE interaction association with rs3825807 on CAD risk is directionally consistent

with that observed with rs7178051 and supports the notion that smoking attenuates the cardioprotective effects of genetically decreased ADAMTS7 expression/function.

In another candidate gene approach, stratified analyses were also used to evaluate a polygenic risk score (PRS) constructed with known susceptibility loci for association with CAD in smokers and non-smokers [39]. These analyses identified a GxE interaction where cumulative genetic burden of the 50 variants examined modified the association of smoking with CAD, which also happened to be stronger in men than women (Table 2) [39]. While not for CVD outcomes directly, unbiased, genome-wide studies have sought to identify gene-smoking interactions for CVD risk factors as well [40-42]. These meta-analysis-based studies incorporated data from approximately half a million subjects and identified ~20 new loci associated with LDL cholesterol, high-density lipoprotein (HDL) cholesterol, and triglyceride levels, or systolic and diastolic blood pressure (SBP and DBP), pulse pressure (PP), and mean arterial pressure (MAP) through GxE interactions (Table 2) [40-42]. However, in some cases, the GxE loci were only observed in subjects of African ancestry, were not replicated in independent datasets, and were supported only by limited functional data that did not point to obvious underlying molecular mechanisms [40-42].

GxE Interactions with Diet and Physical Activity.

In comparison to smoking, fewer recent GxE studies have been reported for other exposures relevant to CVD outcomes. Part of this may be due to the difficulty in obtaining precise estimates of exposure in large numbers of subjects, such as for diet or physical activity which are usually based on questionnaire data. In addition, diet and physical activity are typically categorized for GxE interaction analyses based on percentile cutoffs, and thus may not best reflect the biological effects of such environmental factors. Like other GxE interaction studies, most prior analyses with diet have been based on testing candidate genes [43]. This approach was recently applied to a genetic risk score (GRS)-based analysis with known lipid loci, which revealed that the cumulative genetic association of LDL-raising alleles with increased risk of stroke was attenuated by a high quality diet (Table 2) [•44]. In another study, Francis et al. [45] carried out a GWAS analysis with fish oil supplementation as the dietary exposure in ~81,000 subjects from the UK Biobank and ARIC cohorts to identify four novel loci associated with LDL cholesterol (*ABCG6*, *MLXIPL*), HDL cholesterol (*SLC12A3*), and triglyceride (*GJB6-GJB2-GJA3*) levels through synergistic or antagonistic interactions (Figure 1; Table 2). A similar GWAS study design with larger numbers of subjects also identified novel loci for lipid levels through interactions with physical activity, defined as a dichotomized exposure variable [46]. The results indicated that the HDL-raising effects of three novel loci were enhanced as a function of being physically active whereas the LDL-raising effects at another locus were attenuated by physical activity (Table 2).

GxE Interactions with Air Pollution.

Numerous previous studies have evaluated GxE interactions with various air pollutants, including PM, nitrogen or sulfur containing molecules, and ozone [47]. Most of these studies involved candidate gene analyses with small sample sizes [47] compared to the

more recent GxE interaction studies described above for other exposures. However, two studies used a GWAS approach for gene-air pollution interaction studies. Using distance to nearest major road as the exposure, one study identified two loci (*BMP8A* and *BMP2*) that were associated with PAD at the genome-wide significance level [48] and a suggestive locus (*PIGR-FCAMR*) associated with coronary atherosclerosis [49]. Lastly, another recent GWAS investigated GxE interactions for cardiac function traits and identified a variant downstream of *CXCL12* that was associated with greater QT prolongation in subjects exposed to high PM than those in the low exposure group [50]. Notably, the *CXCL12* locus was previously known to influence risk of CVD through main effects since it was identified in the first series of GWAS for CAD and MI over 10 years ago [51, 52].

Gene-sex Interactions.

Surprisingly, few studies have investigated interactions with respect to sex for CVD even though it is known that men are at higher risk than women, at least until menopause. In this regard, variation at *CPS1* was shown to exhibit sexually dimorphic pleiotropic associations with several choline-related and urea cycle metabolites, which translated into decreased risk of CAD in women but not men (Table 2) [•53]. This finding remains one of the few gene-sex interaction loci reported in the literature for CAD. More recently, Huang and colleagues used over 300,000 subjects from the UK Biobank to demonstrate that association of a genome-wide PRS and a 161-locus GRS was more strongly associated with CAD in men than women (Table 2) [54]. In addition, studies have also examined variation on the X chromosome for association with CAD as another approach to exploring gene-sex interactions. One such study found no associations on the X chromosome [55, 56]whereas a more recent study identified a locus that reduced CAD risk through strong lipid lowering effects [55, 56]. However, these effects were equivalent in men and women [55, 56]. Thus, it does not appear that variation at the X chromosome plays a major role in mediating sex differences with respect to CVD.

Recent Advances for Investigating GxE Interactions.

The analytical frameworks typically used for investigating GxE interactions described above poses certain challenges that may explain why the results of candidate gene GxE interaction studies are often not replicated or why unbiased GxE interaction studies have, until recently, not been largely pursued. For example, detecting GxE interactions have been estimated to require ~4-fold larger sample sizes than detecting main effects of exposures or SNPs alone [57]. This is compounded by the imprecise measures of many CVD-relevant exposures. Notably, the availability of large population-based cohorts, such as the UK Biobank [28], FinnGen [29], and Biobank Japan [30] may overcome such sample size limitations, as has been shown recently for cardiometabolic traits [58]. Moreover, such large datasets can also facilitate replication of the GxE interactions since the associations highlighted in Table 2 have only been identified very recently and still require validation in independent populations prior to considering them as true GxE signals.

Since subjects from large biobanks could be related to varying degrees, genetic analysis of such datasets requires the use of liner mixed models (LMMs), which can control for

potential confounding due to cryptic relatedness and/or population stratification. Although genetic analyses with LMMs can be computationally intensive, especially with large datasets, they have been successfully adapted in GWAS to detect main effects for CVD [20-27]. However, efficient LMM methods for carrying out unbiased GxE interaction studies have also recently been developed. For example, a novel method termed Linear Environment Mixed Model Analysis (LEMMA) has been introduced for investigating GxE interactions on a genome-wide scale [••59]. Importantly, LEMMA differs from previous approaches in that it implements a Bayesian whole-genome regression model that allows joint modeling of both main SNP effects and GxE interactions and can be used with biobank-level data [••59]. Since it is also biologically plausible that multiple exposures can jointly affect CVD risk and that genetic variants can modify this association, another advantage of LEMMA is that it allows linear combinations of environmental variables to be tested jointly for GxE interactions [••59]. For example, rather than testing only one component of air pollution for GxE interactions, LEMMA would allow the inclusion of all types of PM or NO_x into the model. In addition to LEMMA, a non-linear regression and an even more computationally efficient method, termed Gaussian Prior Linear Environment Mixed Model Analysis (GPLEMMA), has also been developed for GxE interaction studies [••60]. Mixed model approaches have also been developed for assessing GxE interactions with whole exome or genome sequencing data where sets of variants at the gene level can be tested in aggregate [61]. Lastly, other statistical approaches, such as machine learning and Bayesian analyses, have also been proposed to address the GxE problem [62]. Taken together, the availability of these methods will foster additional research on GxE interactions and allow such studies to be carried in large numbers of subjects, which will increase the likelihood of identifying additional true GxE signals.

As an alternative to human studies, mouse models provide an excellent complementary approach to investigating GxE interactions since many experimental variables, such as age, sex, and exposures of interest can be tightly controlled. Animal models also have additional advantages as compared to human studies, including access to tissues, defined genetic backgrounds, and experimental validation. This is illustrated by studies using ~100 strains from the Hybrid Mouse Diversity Panel [63], which has led to identification of robust GxE interactions for diet-induced obesity [64] and responses to air pollution-related exposures [65], among other traits [63].

Conclusions.

In summary, efforts to elucidate GxE interactions for CVD are becoming increasingly successful due, in part, to collaborative studies with large population-based cohorts and collections of case-control datasets. Coupled with advances in statistical methodology, these resources will allow even more sophisticated types of GxE analyses to be carried out, which will increase the likelihood of identifying additional novel interactions with exposures that have important contributions to CVD risk. Taken together with recent discoveries, these efforts are likely to provide us with a better understanding of the pathogenesis of CVD and the role of GxE interactions in mediating its genetic liability. However, elucidating the biological mechanisms underlying GxE interactions and determining their translational applications will remain a challenge that will require extensive follow up studies.

Acknowledgments

Work in the author's laboratories is supported, in part, by the USC Center for Genetic Epidemiology Pilot Project Program and NIH grants R01ES025786, R01MD010358, R01HL133169, R01HL148110, R01DK117850, R01HL147883, and R01HL144651. The funders had no role in preparation, review, or approval of the manuscript.

References

1. Virani SS, Alonso A, Benjamin EJ, Bittencourt MS, Callaway CW, Carson AP, et al. Heart Disease and Stroke Statistics-2020 Update: A Report From the American Heart Association. Circulation. 2020;141(9):e139–e596. [PubMed: 31992061]
2. Ridker PM, Danielson E, Fonseca FA, Genest J, Gotto AM Jr., Kastelein JJ, et al. Rosuvastatin to prevent vascular events in men and women with elevated C-reactive protein. N Engl J Med. 2008;359(21):2195–207. [PubMed: 18997196]
3. Hansson GK. Inflammation, atherosclerosis, and coronary artery disease. N Engl J Med. 2005;352(16):1685–95. [PubMed: 15843671]
4. Grundy SM, Stone NJ, Bailey AL, Beam C, Birtcher KK, Blumenthal RS, et al. 2018 AHA/ACC/AACVPR/AAPA/ABC/ACPM/ADA/AGS/APhA/ASPC/NLA/PCNA Guideline on the Management of Blood Cholesterol: Executive Summary: A Report of the American College of Cardiology/American Heart Association Task Force on Clinical Practice Guidelines. Circulation. 2019;139(25):e1046–e81. [PubMed: 30565953]
5. Sackton TB, Hartl DL. Genotypic context and epistasis in individuals and populations. Cell. 2016;166(2):279–87. [PubMed: 27419868]
6. Worth RM, Kato H, Rhoads GG, Kagan K, Syme SL. Epidemiologic studies of coronary heart disease and stroke in Japanese men living in Japan, Hawaii and California: mortality. Am J Epidemiol. 1975;102(6):481–90. [PubMed: 1202950]
7. Bhatnagar A Environmental Determinants of Cardiovascular Disease. Circ Res. 2017;121(2):162–80. [PubMed: 28684622]
8. Hu FB, Stampfer MJ, Manson JE, Rimm E, Colditz GA, Rosner BA, et al. Dietary fat intake and the risk of coronary heart disease in women. N Engl J Med. 1997;337(21):1491–9. [PubMed: 9366580]
9. Hu FB, Rimm EB, Stampfer MJ, Ascherio A, Spiegelman D, Willett WC. Prospective study of major dietary patterns and risk of coronary heart disease in men. Am J Clin Nutr. 2000;72(4):912–21. [PubMed: 11010931]
10. Mozaffarian D, Katan MB, Ascherio A, Stampfer MJ, Willett WC. Trans fatty acids and cardiovascular disease. N Engl J Med. 2006;354(15):1601–13. [PubMed: 16611951]
11. Schane RE, Ling PM, Glantz SA. Health effects of light and intermittent smoking: a review. Circulation. 2010;121(13):1518–22. [PubMed: 20368531]
12. Pope CA 3rd, Turner MC, Burnett RT, Jerrett M, Gapstur SM, Diver WR, et al. Relationships between fine particulate air pollution, cardiometabolic disorders, and cardiovascular mortality. Circ Res. 2015;116(1):108–15. [PubMed: 25348167]
13. Cosselman KE, Navas-Acien A, Kaufman JD. Environmental factors in cardiovascular disease. Nat Rev Cardiol. 2015;12(11):627–42. [PubMed: 26461967]
14. Brauer M, Casadei B, Harrington RA, Kovacs R, Sliwa K, Group WHFAPE. Taking a Stand Against Air Pollution-The Impact on Cardiovascular Disease: A Joint Opinion From the World Heart Federation, American College of Cardiology, American Heart Association, and the European Society of Cardiology. Circulation. 2021;143(14):e800–e4. [PubMed: 33506685]
15. Wienke A, Holm NV, Skytthe A, Yashin AI. The heritability of mortality due to heart diseases: a correlated frailty model applied to Danish twins. Twin Res. 2001;4(4):266–74. [PubMed: 11665307]
16. Zdravkovic S, Wienke A, Pedersen NL, Marenberg ME, Yashin AI, De Faire U. Heritability of death from coronary heart disease: a 36-year follow-up of 20 966 Swedish twins. J Intern Med. 2002;252(3):247–54. [PubMed: 12270005]

17. Bevan S, Traylor M, Adib-Samii P, Malik R, Paul NL, Jackson C, et al. Genetic heritability of ischemic stroke and the contribution of previously reported candidate gene and genomewide associations. Stroke. 2012;43(12):3161–7. [PubMed: 23042660]

18. Vinkhuyzen AA, Wray NR, Yang J, Goddard ME, Visscher PM. Estimation and partition of heritability in human populations using whole-genome analysis methods. Annu Rev Genet. 2013;47:75–95. [PubMed: 23988118]

19. Nikpay M, Goel A, Won HH, Hall LM, Willenborg C, Kanoni S, et al. A comprehensive 1,000 Genomes-based genome-wide association meta-analysis of coronary artery disease. Nat Genet. 2015;47(10):1121–30. [PubMed: 26343387]

•20. Howson JMM, Zhao W, Barnes DR, Ho WK, Young R, Paul DS, et al. Fifteen new risk loci for coronary artery disease highlight arterial-wall-specific mechanisms. Nat Genet. 2017;49(7):1113–9. [PubMed: 28530674]

•21. Nelson CP, Goel A, Butterworth AS, Kanoni S, Webb TR, Marouli E, et al. Association analyses based on false discovery rate implicate new loci for coronary artery disease. Nat Genet. 2017;49(9):1385–91. [PubMed: 28714975]

•22. Klarin D, Zhu QM, Emdin CA, Chaffin M, Horner S, McMillan BJ, et al. Genetic analysis in UK Biobank links insulin resistance and transendothelial migration pathways to coronary artery disease. Nat Genet. 2017;49(9):1392–7. [PubMed: 28714974]

•23. van der Harst P, Verweij N. Identification of 64 novel genetic loci provides an expanded view on the genetic architecture of coronary artery disease. Circ Res. 2018;122(3):433–43. [PubMed: 29212778]

•24. Malik R, Chauhan G, Traylor M, Sargurupremraj M, Okada Y, Mishra A, et al. Multiancestry genome-wide association study of 520,000 subjects identifies 32 loci associated with stroke and stroke subtypes. Nat Genet. 2018;50(4):524–37. [PubMed: 29531354]

•25. Klarin D, Lynch J, Aragam K, Chaffin M, Assimes TL, Huang J, et al. Genome-wide association study of peripheral artery disease in the Million Veteran Program. Nat Med. 2019;25(8):1274–9. [PubMed: 31285632]

•26. Koyama S, Ito K, Terao C, Akiyama M, Horikoshi M, Momozawa Y, et al. Population-specific and trans-ancestry genome-wide analyses identify distinct and shared genetic risk loci for coronary artery disease. Nat Genet. 2020;52(11):1169–77. [PubMed: 33020668]

•27. Hartiala JA, Han Y, Jia Q, Hilser JR, Huang P, Gukasyan J, et al. Genome-wide analysis identifies novel susceptibility loci for myocardial infarction. Eur Heart J. 2021;42(9):919–33. [PubMed: 33532862] In just the last few years, these eight studies have collectively identified over 100 additional susceptibility loci for CVD outcomes.

28. Bycroft C, Freeman C, Petkova D, Band G, Elliott LT, Sharp K, et al. The UK Biobank resource with deep phenotyping and genomic data. Nature. 2018;562(7726):203–9. [PubMed: 30305743]

29. Tabassum R, Ramo JT, Ripatti P, Koskela JT, Kurki M, Karjalainen J, et al. Genetic architecture of human plasma lipidome and its link to cardiovascular disease. Nat Commun. 2019;10(1):4329. [PubMed: 31551469]

30. Nagai A, Hirata M, Kamatani Y, Muto K, Matsuda K, Kiyohara Y, et al. Overview of the BioBank Japan Project: Study design and profile. J Epidemiol. 2017;27(3S):S2–S8. [PubMed: 28189464]

31. Hartiala J, Schwartzman WS, Gabbay J, Ghazalpour A, Bennett BJ, Allayee H. The genetic architecture of coronary artery disease: current knowledge and future opportunities. Curr Atheroscler Rep. 2017;19(2):6. [PubMed: 28130654]

32. Eicher JD, Chen MH, Pitsillides AN, Lin H, Veeraraghavan N, Brody JA, et al. Whole exome sequencing in the Framingham Heart Study identifies rare variation in HYAL2 that influences platelet aggregation. Thromb Haemost. 2017;117(6):1083–92. [PubMed: 28300864]

33. Jian X, Satizabal CL, Smith AV, Wittfeld K, Bis JC, Smith JA, et al. Exome chip analysis identifies low-frequency and rare variants in mrpl38 for white matter hyperintensities on brain magnetic resonance imaging. Stroke. 2018;49(8):1812–9. [PubMed: 30002152]

34. Paraboschi EM, Khera AV, Merlini PA, Gigante L, Peyvandi F, Chaffin M, et al. Rare variants lowering the levels of coagulation factor X are protective against ischemic heart disease. Haematologica. 2020;105(7):e365–e9. [PubMed: 31699787]

35. Pedicino D, Vinci R, Giglio AF, Pisano E, Porto I, Vergallo R, et al. Alterations of hyaluronan metabolism in acute coronary syndrome: implications for plaque erosion. J Am Coll Cardiol. 2018;72(13):1490–503. [PubMed: 30236312]

••36. Saleheen D, Zhao W, Young R, Nelson CP, Ho W, Ferguson JF, et al. Cardioprotective effects at the ADAMTS7 locus as a result of gene-smoking interactions. Circulation. 2017;135(24):2336–53. [PubMed: 28461624] Of 50 known CAD loci evaluated, this study identified a significant GxE interaction with the ADAMTS7 locus where the protective association of the effect allele was attenated by 60% in smokers. Functional data support this gene-smoking interaction since expression of ADAMTS7 was significantly upregulated in vascular cells exposed to cigaratte smoke extract.

37. Yu Z, Zekavat SM, Haidermota S, Bernardo R, Libby P, Finucane H, et al. Genome-wide pleiotropy analysis of coronary artery disease and pneumonia identifies shared immune pathways. medRxiv. 2021:2021.07.05.21260028.

38. Pu X, Xiao Q, Kiechl S, Chan K, Ng FL, Gor S, et al. ADAMTS7 cleavage and vascular smooth muscle cell migration is affected by a coronary-artery-disease-associated variant. Am J Hum Genet. 2013;92(3):366–74. [PubMed: 23415669]

39. Hindy G, Wiberg F, Almgren P, Melander O, Orho-Melander M. Polygenic risk score for coronary heart disease modifies the elevated risk by cigarette smoking for disease incidence. Circ Genom Precis Med. 2018;11(1):e001856. [PubMed: 29874179]

40. Bentley AR, Sung YJ, Brown MR, Winkler TW, Kraja AT, Ntalla I, et al. Multi-ancestry genome-wide gene-smoking interaction study of 387,272 individuals identifies new loci associated with serum lipids. Nat Genet. 2019;51(4):636–48. [PubMed: 30926973]

41. Sung YJ, de Las Fuentes L, Winkler TW, Chasman DI, Bentley AR, Kraja AT, et al. A multi-ancestry genome-wide study incorporating gene-smoking interactions identifies multiple new loci for pulse pressure and mean arterial pressure. Hum Mol Genet. 2019;28(15):2615–33. [PubMed: 31127295]

42. Sung YJ, Winkler TW, de Las Fuentes L, Bentley AR, Brown MR, Kraja AT, et al. A large-scale multi-ancestry genome-wide study accounting for smoking behavior identifies multiple significant loci for blood pressure. Am J Hum Genet. 2018;102(3):375–400. [PubMed: 29455858]

43. Voruganti VS. Nutritional genomics of cardiovascular disease. Curr Genet Med Rep. 2018;6(2):98–106. [PubMed: 30581653]

•44. Hellstrand S, Ericson U, Schulz CA, Drake I, Gullberg B, Hedblad B, et al. Genetic susceptibility to dyslipidemia and incidence of cardiovascular disease depending on a diet quality index in the Malmo Diet and Cancer cohort. Genes Nutr. 2016;11:20. [PubMed: 27551321] This relatively large study provided evidence that the adverse association of cummulative genetic burden for increased LDL cholesterol levels with risk of stroke could be mitigated by a high quality diet.

45. Francis M, Li C, Sun Y, Zhou J, Li X, Brenna JT, et al. Genome-wide association study of fish oil supplementation on lipid traits in 81,246 individuals reveals new gene-diet interaction loci. PLoS Genet. 2021;17(3):e1009431. [PubMed: 33760818]

46. Kilpelainen TO, Bentley AR, Noordam R, Sung YJ, Schwander K, Winkler TW, et al. Multi-ancestry study of blood lipid levels identifies four loci interacting with physical activity. Nat Commun. 2019;10(1):376. [PubMed: 30670697]

47. Ward-Caviness CK. A review of gene-by-air pollution interactions for cardiovascular disease, risk factors, and biomarkers. Hum Genet. 2019;138(6):547–61. [PubMed: 30968250]

48. Ward-Caviness CK, Neas LM, Blach C, Haynes CS, LaRocque-Abramson K, Grass E, et al. Genetic variants in the bone morphogenic protein gene family modify the association between residential exposure to traffic and peripheral arterial disease. PLoS One. 2016;11(4):e0152670. [PubMed: 27082954]

49. Ward-Caviness CK, Neas LM, Blach C, Haynes CS, LaRocque-Abramson K, Grass E, et al. A genome-wide trans-ethnic interaction study links the PIGR-FCAMR locus to coronary atherosclerosis via interactions between genetic variants and residential exposure to traffic. PLoS One. 2017;12(3):e0173880. [PubMed: 28355232]

50. Gondalia R, Avery CL, Napier MD, Mendez-Giraldez R, Stewart JD, Sitlani CM, et al. Genome-wide association study of susceptibility to particulate matter-associated QT prolongation. Environ Health Perspect. 2017;125(6):067002. [PubMed: 28749367]

51. Samani NJ, Erdmann J, Hall AS, Hengstenberg C, Mangino M, Mayer B, et al. Genomewide association analysis of coronary artery disease. N Engl J Med. 2007;357(5):443–53. [PubMed: 17634449]

52. Myocardial Infarction Genetics C, Kathiresan S, Voight BF, Purcell S, Musunuru K, Ardissino D, et al. Genome-wide association of early-onset myocardial infarction with single nucleotide polymorphisms and copy number variants. Nat Genet. 2009;41(3):334–41. [PubMed: 19198609]

•53. Hartiala JA, Tang WH, Wang Z, Crow AL, Stewart AF, Roberts R, et al. Genome-wide association study and targeted metabolomics identifies sex-specific association of CPS1 with coronary artery disease. Nat Commun. 2016;7:10558. [PubMed: 26822151] This study integrated targeted metabolomics with genetic analyses to identify one of the only sexually dimorphic assocations known for CAD where a variant in CPS1 reduces risk in women but not men.

54. Huang Y, Hui Q, Gwinn M, Hu YJ, Quyyumi AA, Vaccarino V, et al. Sexual differences in genetic predisposition of coronary artery disease. Circ Genom Precis Med. 2021;14(1):e003147. [PubMed: 33332181]

55. Loley C, Alver M, Assimes TL, Bjonnes A, Goel A, Gustafsson S, et al. No Association of Coronary Artery Disease with X-Chromosomal Variants in Comprehensive International Meta-Analysis. Sci Rep. 2016;6:35278. [PubMed: 27731410]

56. Natarajan P, Pampana A, Graham SE, Ruotsalainen SE, Perry JA, de Vries PS, et al. Chromosome Xq23 is associated with lower atherogenic lipid concentrations and favorable cardiometabolic indices. Nat Commun. 2021;12(1):2182. [PubMed: 33846329]

57. Smith PG, Day NE. The design of case-control studies: the influence of confounding and interaction effects. Int J Epidemiol. 1984;13(3):356–65. [PubMed: 6386716]

58. Wang H, Zhang F, Zeng J, Wu Y, Kemper KE, Xue A, et al. Genotype-by-environment interactions inferred from genetic effects on phenotypic variability in the UK Biobank. Sci Adv. 2019;5(8):eaaw3538. [PubMed: 31453325]

••59. Kerin M, Marchini J. Gene-by-environment interactions with a bayesian whole-genome regression model. Am J Hum Genet. 2020;107(4):698–713. [PubMed: 32888427]

••60. Kerin M, Marchini J. A non-linear regression method for estimation of gene-environment heritability. Bioinformatics. 2020;36(4):5632–9. These two papers describe new computationally efficient linear mixed model algorithms for large-scale GxE interaction analyses. The major advance of these methods is they allows joint modeling of both main SNP effects and GxE interactions and can be used with large biobank-level data.

61. Wang X, Lim E, Liu CT, Sung YJ, Rao DC, Morrison AC, et al. Efficient gene-environment interaction tests for large biobank-scale sequencing studies. Genet Epidemiol. 2020;44(8):908–23. [PubMed: 32864785]

62. Li J, Li X, Zhang S, Snyder M. Gene-environment interaction in the era of precision medicine. Cell. 2019;177(1):38–44. [PubMed: 30901546]

63. Lusis AJ, Seldin MM, Allayee H, Bennett BJ, Civelek M, Davis RC, et al. The Hybrid Mouse Diversity Panel: a resource for systems genetics analyses of metabolic and cardiovascular traits. J Lipid Res. 2016;57(6):925–42. [PubMed: 27099397]

64. Parks BW, Nam E, Org E, Kostem E, Norheim F, Hui ST, et al. Genetic control of obesity and gut microbiota composition in response to high-fat, high-sucrose diet in mice. Cell Metab. 2013;17(1):141–52. [PubMed: 23312289]

65. Maazi H, Hartiala JA, Suzuki Y, Crow AL, Shafiei Jahani P, Lam J, et al. A GWAS approach identifies Dapp1 as a determinant of air pollution-induced airway hyperreactivity. PLoS Genet. 2019;15(12):e1008528. [PubMed: 31869344]

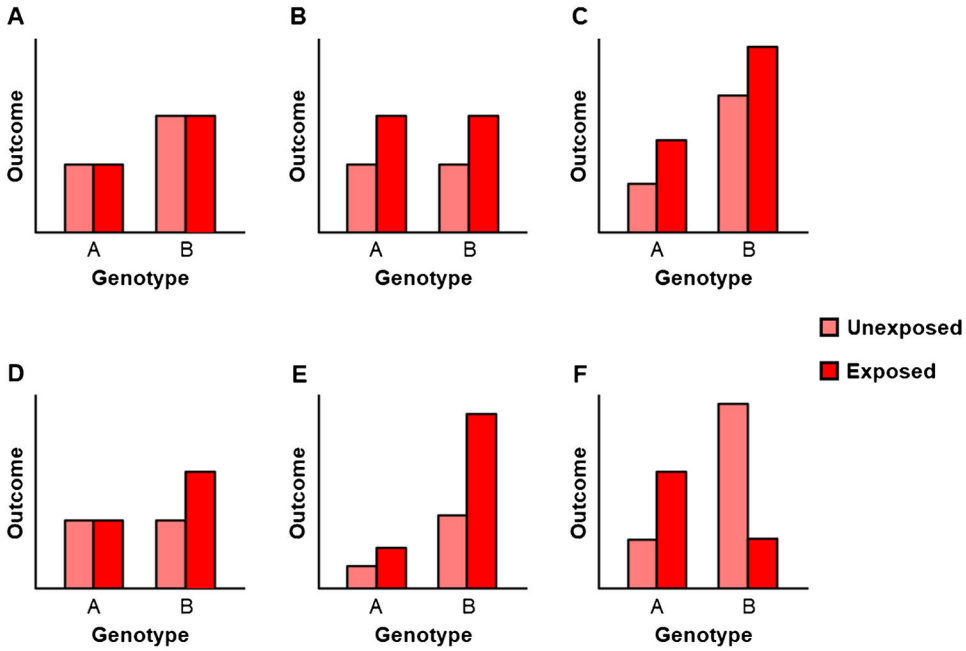

Figure 1. Conceptual Illustration of Genetic, Environmental, and GxE Interaction Effects.
Differences in an outcome (i.e. CAD risk) can be due to main effects of genetic (A) or environmental (B) factors or the additive effect of both genetic and environmental factors (C). Interactions are conceptualized as effects that manifest (D), synergize (E), or are antagonized (F) by an exposure.

Table 1.

Selected Environmental Variables Associated with CVD Outcomes.

Exposure	Direction of Association with CVD
Diet	Positively associated with saturated and *trans* fat and cholesterol content and inversely associated with vegetable, fruit, whole grain, lean protein, and polyunsaturated fat content. Certain species of gut bacteria are positively associated through their nutrients, such as choline, L-carnitine, and phenylalanine.
Physical activity	Negatively associated with increased levels
Smoking	Positively associated with increased levels and accounts for ~250,000 annual premature deaths due to CVD in the United States alone.
Air pollution	Positively associated with increased exposure levels of particulate matter, aerosols, and volatile organic chemicals.
Sex	Men are more susceptible than women until menopause at which point risk becomes equivalent.

Table 2.

Recent GxE Interactions Identified for CVD Outcomes and Risk Factors.

Exposure	Approach(s) used	Variants tested	Sample Size	Outcome	Nature of Interaction	Ref.
Smoking	Candidate gene	50 known CAD-associated variants	~140,000 subjects of European ancestry	CAD	Protective association of *ADAMTS7* variant diminished by smoking, which induced expression of *ADAMTS7*.	[••36, 37]
	Candidate gene	PRS with 50 known CAD-associated variants	~24,000 subjects of European ancestry	CAD	CAD risk was higher in male smokers with the lowest cumulative genetic burden.	[39]
	GWAS meta-analysis	Genome-wide SNPs	~400,000 subjects of European, African, Asian, Hispanic, and Brazilian ancestries	Lipid levels	13 novel loci identified for LDL cholesterol, HDL cholesterol, and triglyceride levels, most of which were only significant in subjects of African ancestry.	[40]
	GWAS meta-analysis	Genome-wide SNPs	~610,000 subjects of European, African, Asian, and Hispanic ancestries	SBP, DBP, MAP, PP	18 loci exhibited higher blood pressure-raising effects in smokers than non-smokers and most loci were only identified in subjects of African ancestry.	[41, 42]
Diet	Candidate gene	GRS with 91 lipid-associated SNPs	~25,000 subjects of European ancestry	Lipid levels, CAD, and stroke	Association of GRS constructed for LDL cholesterol levels with stroke was attenuated by high quality diet.	[•44]
	GWAS	Genome-wide SNPs	~81,000 subjects of European ancestry	Lipid levels	Four novel loci identified for LDL, HDL, or triglyceride levels that exhibited synergistic or antagonist interactions with fish oil.	[45]
Physical activity	GWAS meta-analysis	Genome-wide SNPs	~252,000 subjects of European, African, Asian, Hispanic, and Brazilian ancestries	Lipid levels	Physical activity attenuated LDL cholesterol-raising effects at *CNTNAP2* locus and increased HDL-raising effects at *CLASP1*, *LHX1*, and *SNTA1* loci.	[46]
Air pollution	GWAS	Genome-wide SNPs	~2200 subjects of European and African ancestries	PAD	Association of minor allele at *BMP8A* and *BMP2* loci with risk of PAD increased by residential proximity to traffic.	[48]
	GWAS	Genome-wide SNPs	~2100 subjects of European and African ancestries	CAD	Only suggestive associations observed at locus harboring *PIGR-FCAMR* in subjects of African ancestry.	[49]
	GWAS meta-analysis	Genome-wide SNPs	~22,000 subjects of European, African, and Hispanic ancestries	QT interval	Compared to the low exposure group, variant downstream of *CXCL12* was associated with greater QT prolongation in subjects exposed to high PM.	[50]
Sex	Candidate gene	Single variant	~54,000 subjects of European ancestry	CAD	Variant of *CPS1* associated with decreased risk of CAD in women but not men.	[•53]
	Genome-wide and candidate gene	Genome-wide PRS and GRS with 161 CAD SNPs	~318,000 individuals of European ancestry	CAD	Effect of cumulative burden of genome-wide PRS and candidate GRS stronger in men.	[54]

CAD, coronary artery disease; PRS, polygenic risk score; GRS, genetic risk score; SBP, systolic blood pressure; diastolic blood pressure; MAP, mean arterial pressure; PP, pulse pressure; PAD, peripheral artery disease.