Genetic Effects on the Correlation Structure of CVD Risk Factors

Exome-Wide Data From a Ghanaian Population

Nuri Kodaman[†,‡,§,*], Rafal S. Sobota[†,§,*], Folkert W. Asselbergs[‖,¶,#], Matthew T. Oetjens[**], Jason H. Moore[§,††], Nancy J. Brown[‡‡], Melinda C. Aldrich[§§], Scott M. Williams[‡,§]

Nashville, TN, USA; Cleveland, OH, USA; Hanover, NH, USA: Utrecht, the Netherlands; London, United Kingdom; Ann Arbor, MI, USA; and Philadelphia, PA, USA

ABSTRACT

Plasma concentration of plasminogen activator inhibitor-1 (PAI-1) is highly correlated with several cardiovascular disease (CVD) risk factors. It also plays a direct role in CVD, including myocardial infarction and stroke, by impeding the dissolution of thrombi in the blood. Insofar as PAI-1 links CVD's risk factors to its endpoints, genetic variants modulating the relationship between PAI-1 and risk factors may be of particular clinical and biological interest. The high heritability of PAI-1, which has not been explained by genetic association studies, may also, in large part, be due to this relationship with CVD risk factors. Using exome-wide data from 1,032 Ghanaian study participants, we tested for heterogeneity of correlation by genotype between PAI-1 and 4 CVD risk factors (body mass index, triglycerides, mean arterial pressure, and fasting glucose) under the hypothesis that loci involved in the relationship between PAI-1 and other risk factors will also modify their correlational structure. We found more significant heterogeneities of correlation by genotype than we found marginal effects, with no evidence of type I inflation. The most significant result among all univariate and multivariate tests performed in this study was the heterogeneity of correlation between PAI-1 and mean arterial pressure at rs10738554, near *SLC24A2*, a gene previously associated with high blood pressure in African Americans.

Cardiovascular disease (CVD) is responsible for almost one-half of all noncommunicable disease—related deaths worldwide [1]. It comprises multiple disorders of the circulatory system, among which venous and arterial thrombotic disorders are the most common [2]. The enzyme plasminogen activator inhibitor-1 (PAI-1) plays a major role in the etiology of thrombosis by impeding fibrinolysis, or clot breakdown [3]. Elevated plasma PAI-1 is accordingly a major risk factor for thrombotic events, such as deep vein thrombosis, myocardial infarction, and stroke [4].

Plasma PAI-1 concentration has been considered a promising endophenotype for CVD, because it is linked etiologically to correlated clinical endpoints. Endophenotypes are likely to have simpler genetic architectures than the complex diseases with which they associate. They can also be defined unambiguously and measured precisely, making them potentially valuable targets for genome-wide association studies (GWAS) [5]. These advantages were recently demonstrated by a GWAS on serum transferrin (a biomarker for iron deficiency) that identified 2 loci explaining 40% of the genetic variation in this protein [6]. Similar studies on PAI-1, however, have not been nearly as successful. A recent meta-analysis identified only 3 genome-wide significant loci, which together explained <3% of the genetic variance [7]. The inability to identify any major genetic factors beyond the well-documented 4G/5G variant in the *PAI-1* promoter is rather puzzling [8], particularly because the heritability of PAI-1 has been estimated to be as high as 0.83 [9]. Furthermore, the small number of variants that are associated with PAI-1 do not appear to be associated with CVD-related outcomes [10], despite the fact that high PAI-1 levels are.

One possible explanation for this paradox is that indirect genetic effects are responsible for the high heritability of PAI-1. For example, even if the genes directly involved in PAI-1 production were devoid of any variation, PAI-1 would still be a heritable trait, because many conditions that associate with PAI-1, such as obesity, hypertriglyceridemia, hypertension, and even dietary habits, are themselves heritable [11-13], and PAI-1 concentration increases steadily across the entire distribution of these and most other cardiometabolic risk factors [14]. Whereas increasing GWAS sample sizes can improve the likelihood of detecting such indirect associations, the purpose of detecting them would not be entirely clear. In fact, variables such as body mass index [BMI] and triglycerides are adjusted for as covariates in most PAI-1 association studies [7,15,16].

There are, however, potentially more important ways in which the positive association between cardiovascular risk factors and PAI-1 can help explain its missing heritability. In particular, there may be genetic variants that have

The authors report no relationships that could be construed as a conflict of interest.
This work was supported by National Institutes of Health Grants HL65234 and LM010098 (to F.W. Asselbergs, J.H. Moore, and S.M. Williams), K07 CA172294 to (M. C, Aldrich), and by a Dekker scholarship-Junior Staff Member 2014T001—Netherlands Heart Foundation and University College London Hospitals, NIHR (National Institute of Health Research) Biomedical Research Centre (to F.W. Asselbergs).
R. S. Sobota was supported by Public Health Service Award T32 GM07347 from the National Institute of General Medical Studies for the Vanderbilt Medical-Scientist Training Program.
R. S. Sobota, N. Kodaman, and S.M. Williams were partially supported by National Institutes of Health Grant P20 GM103534.
From the †Vanderbilt Genetics Institute, Vanderbilt University Medical School, Nashville, TN, USA; ‡Department of Epidemiology and Biostatistics, Case Western Reserve University, Cleveland, OH, USA; §Department of Genetics, Geisel School of Medicine, Dartmouth College, Hanover, NH, USA; ‖Department of Cardiology, Division Heart and Lungs, UMC (University Medical Center) Utrecht, Utrecht, the Netherlands; ¶Durrer Center for Cardiogenetic Research, Netherlands Heart Institute, Utrecht, the Netherlands; #Institute of Cardiovascular Science, Faculty of Population Health Sciences, University College London, London, United Kingdom; **Department of Genetics, University of Michigan, Ann

Arbor, MI, USA; ††Department of Biostatistics and Epidemiology, The Perelman School of Medicine, University of Pennsylvania, Philadelphia, PA, USA; ‡‡Department of Medicine, Vanderbilt University Medical School, Nashville, TN, USA; and the §§Department of Thoracic Surgery and Division of Epidemiology, Vanderbilt University Medical School, Nashville, TN, USA.
Correspondence: S. M. Williams (smw154@case.edu).
* N. Kodaman and R. S. Sobota contributed equally to this work.

GLOBAL HEART

VOL. ■, NO. ■, 2017
ISSN 2211-8160/$36.00.
http://dx.doi.org/10.1016/j.gheart.2017.01.013

direct but conditional effects on the concentration of PAI-1, dependent on other specific risk factors. For example, a (hypothetical) variant may directly increase *PAI-1* expression only when adiposity exceeds a certain threshold; such a variant would be difficult to detect at a genome-wide level of significance, owing to its purely conditional effect, despite its likely contribution to the heritability of PAI-1. Another kind of context-dependent variant may be involved in a (hypothetical) homeostatic pathway that either raises or lowers PAI-1 as adiposity increases or decreases. Genetic effects of this type would be difficult if not impossible to detect in conventional GWAS.

If a large portion of the missing heritability of PAI-1 is due to indirect genetic effects and/or to context-dependent genetic effects, we would expect to observe small effect sizes in association studies, unpredictable attempts at replication, and highly variable heritability estimates, as we do [17,18]. Whereas increasing sample size can improve the power to detect variants whether they have indirect or context-dependent effects, culling the biologically meaningful results from the scores of trivially indirect associations in GWAS will become increasingly difficult [19]. However, the context-dependent variants, many of which likely remain to be found [20,21], may be particularly important from a clinical perspective [22].

With the potential importance of context-dependency in mind, we chose to explicitly test whether the correlation between PAI-1 and associated risk factors differs by genotype, complementing conventional tests for differences in mean alone. The motivation for our analysis is the hypothesis that genetic variants that increase PAI-1 in response to another risk factor (and vice versa) will also modify the correlational structure between the two. Few studies have directly sought to identify genetic variants associated with changes in correlation. Two decades ago, Reilly et al. [23] found that the correlation structure between various apolipoproteins varied with apoliporotein E (ApoE) genotype, and in a sex-specific manner. This ability of ApoE to modulate lipid trait relationships was again demonstrated in a 2013 study, which concluded that the ApoE isoform genotype not only influenced the correlation between triglycerides and total cholesterol, but changed the relationship between both those traits and incident coronary heart disease, but in a population-specific manner [24]. However, to our knowledge, no high-throughput study of genes influencing the covariance among traits has been performed.

METHODS

Study population

The study population has been previously described [25]. Briefly, unrelated participants were identified from Sunyani, the capital of the Brong Ahafo region of Ghana, population ~250,000 as of the 2012 census. Recruitment for the study began in May 2002 and ended in November 2006. Participants learned about the study at public venues, including local churches and markets. Individuals were excluded from analyses if they had signs of acute illness (e.g., malarial infection), were under 18 years of age, or were a first- or second-degree relative of someone already enrolled in the study. Participants provided information via questionnaire regarding their previous medical histories and other demographic and socioeconomic variables, including age, sex, education, smoking status, alcohol consumption, and current medications. All participants provided informed consent. Institutional review boards at Vanderbilt University, Dartmouth College, and Regional Hospital, Sunyani, approved all protocols.

Anthropometric measurements and biochemical analysis. Standing height and weight were measured to calculate BMI. Blood pressure was measured twice; the means for both systolic blood pressure (SBP) and diastolic blood pressure (DBP) were used in subsequent statistical analyses [26]. Mean arterial pressure (MAP) was calculated using the formula: $MAP = DBP + [(SBP-DBP)/3]$, which approximates the average arterial pressure during a single cardiac cycle. Blood was drawn between the hours of 8:00 AM and 10:00 AM, after an 8-h minimum fast. These samples were used to assess fasting glucose, lipid, and PAI-1 levels. Fasting glucose levels were measured using a handheld Sure Step glucose monitor by LifeScan (Milpitas, CA, USA), using blood drops from the blood draw needles. Plasma samples were stored in liquid nitrogen prior to shipment to Vanderbilt University, where concentrations of the PAI-1 antigen were measured using a commercially available enzyme-linked immunoassay (Biopool AB, Umea, Sweden).

Genotyping

A subset of 1,105 urban participants from the Ghanaian cohort was selected for genotyping. DNA was genotyped using the Illumina Infinium HumanExome BeadChip platform (Illumina Inc., San Diego, CA, USA). This platform interrogates strictly exonic variants, covering ~240,000 markers.

Quality control

All single nucleotide polymorphisms (SNP) with a genotyping call rate <95% were removed. Individuals for whom <95% of variants were called were removed from analyses. Variants with a minor allele frequency <20% were also removed, as were variants that failed the test for Hardy-Weinberg equilibrium ($p < 0.001$). Cryptic relatedness was assessed in the data, and 1 participant in each pair of related individuals (pi-hat > 0.2) was randomly removed. Following quality control, 1,032 of the 1,105 participants and 15,890 variants remained for analyses. All quality control procedures were performed in PLINK (version 1.07 [http://pngu.mgh.harvard.edu/purcell/plink/]) [27].