Genetic associations with carotid intima-media thickness link to atherosclerosis biology with sex-differences in sub-Saharan Africans

Palwende Boua (romyboua@gmail.com)
Clinical Research Unit of Nanoro , Institut de Recherche en Sciences de la Santé https://orcid.org/0000-0001-8325-2665

Jean-Tristan Brandenburg
Sydney Brenner Institute for Molecular Bioscience

Ananyo Choudhury
University of the Witwatersrand https://orcid.org/0000-0001-8225-9531

Hermann Sorgho
Clinical Research Unit of Nanoro , Institut de Recherche en Sciences de la Santé

Engelbert Nonterah
Navrongo Health Research Centre, Ghana Health Service

Godfred Agongo
NavrongoHealth Research Centre , Navrongo , Ghana

Gershim Asiki
African Population and Health Research Center, Nairobi, Kenya https://orcid.org/0000-0002-9966-1153

Lisa Micklesfield
MRC/Wits Developmental Pathways for Health Research Unit

Solomon Choma
Department of Pathology and Medical Sciences

Xavier Gómez-Olivé
8MRC/Wits Rural Public Health and Health Transitions Research Unit (Agincourt), School of Public Health, Faculty of Health Sciences , University of the Witwatersrand

Scott Hazelhurst
Sydney Brenner Institute for Molecular Bioscience, Faculty of Health Sciences & School of Electrical & Information Engineering, University of the Witwatersrand https://orcid.org/0000-0002-0581-149X

Halidou Tinto
Institute de Recherche en Sciences de la Santé

Nigel Crowther
11Department of Chemical Pathology, National Health Laboratory Service

Christopher Mathew
King's College London School of Medicine https://orcid.org/0000-0003-4178-1838

Michele Ramsay
University of the Witwatersrand https://orcid.org/0000-0002-4156-4801

Article

Keywords: Atherosclerosis, cardiovascular diseases, carotid intima-media thickness

Posted Date: January 22nd, 2021

DOI: https://doi.org/10.21203/rs.3.rs-147838/v1

Version of Record: A version of this preprint was published at Nature Communications on February 14th, 2022. See the published version at https://doi.org/10.1038/s41467-022-28276-x.

Abstract

Atherosclerosis precedes the onset of many clinical manifestations of cardiovascular diseases (CVDs). We used carotid intima-media thickness (cIMT) to investigate genetic susceptibility to atherosclerosis in 7894 unrelated adults (3963 women, 3931 men) aged 40 to 60 years resident in four sub-Saharan African countries. cIMT was measured by ultrasound and genotyping was performed on the H3Africa SNP Array. Two new African-specific genome-wide significant loci, SIRPA (p=4.7E-08), and FBXL17 (p=2.5E-08), were identified in the combined dataset. Sex-stratified analysis revealed associations with two male-specific loci, SNX29 (p=6.3E-09) and MAP3K7 (p=5.3E-08), and two female-specific loci, LARP6 (p=2.4E-09) and PROK1 (p=1.0E-08). Regional associations were replicated with known risk loci for atherosclerosis and CVDs with different lead SNPs than in Europeans and significant enrichment for oestrogen response genes for female-specific signals were identified. The genes identified showed biological relevance to atherosclerosis and/or CVDs, as well as sex-differences and transferability of signals from non-African studies.

Introduction

Atherosclerosis is a complex multifactorial trait with an enigmatic genetic aetiology. Despite discoveries from genome-wide association studies (GWAS), little is known about the genetic contributions to atherosclerosis. Meanwhile, the worldwide epidemic of cardiovascular diseases (CVDs), including clinical manifestation of atherosclerosis, is growing and has become the leading cause of deaths worldwide (Fuster, 2014; Roth et al., 2017). Moreover, the health and demographic transition in sub-Saharan Africa (SSA) has shifted the major causes of death from communicable and nutritional diseases to non-communicable diseases (NCDs).

Atherosclerosis results from injury to the arterial endothelium, resulting in an inflammatory response in the vessel wall. The location and morphology of the atherosclerotic lesions predict the nature of the resulting vascular disease. Whereas family and twins studies provided evidence of high heritability of cIMT (20–65%) (Fox et al., 2003; Sacco et al., 2009; Fagnani et al., 2013; Medda et al., 2014), the GWAS studies reported associations that account for only 1.1% of the variance of cIMT (Bis et al., 2012).

The genetic diversity of African populations and their deep evolutionary roots represent opportunities for novel genetic discoveries. Because haplotypes blocks are shorter in Africans compared to other populations (average haplotype block ~ 8.8 kb in Africans, ~ 20.7 kb in Europeans, and ~ 25.2 kb in Han Chinese), identification of causal variants is facilitated (Lonjou et al., 2003; Hinds et al., 2005). The role of ancestry in atherosclerosis risk has been established from studies in multi-ethnic settings and admixture studies for atherosclerosis (Shendre, Irvin, *et al.*, 2017; Shendre, Wiener, *et al.*, 2017). African ancestry was reported to be associated with a higher risk of atherosclerosis compared to Europeans, Hispanics and Asians.

Since phenotypic differences between men and women are a pervasive feature of several quantitative traits, studies of sex interactions for complex human traits may shed light on the molecular mechanisms that lead to biological differences between men and women. Sex has been found to play a role in variations between gene expression and genotype across a range of human complex traits (Rawlik et al., 2016). Sex-differences in the transcriptomes of cells involved in the atherosclerotic process have been previously reported (Franconi et al., 2017) and are supported by sex-stratified GWAS analyses (Dong et al., 2015; Lin et al., 2015). Sex provides two different environmental contexts determined by the hormonal milieu and differential gene expression between the sexes.

Several genetic association studies of cIMT have been performed in the major world populations and provided insights into genes and tissue-specific regulatory mechanisms linking atherosclerosis both to its functional genomic origins and its clinical consequences in humans. To date, 76 SNPs have been found to be robustly associated with cIMT (GWAS Catalog) (Buniello et al., 2019), but none of the studies focused on sub-Saharan African populations.

The Africa Wits-INDEPTH Partnership for Genomic Studies cohort (AWI-Gen) was developed to examine genetic and environmental contributions to cardiometabolic diseases in Africans. It has over 12 000 participants from four sub-Saharan African countries, Burkina Faso, Ghana, Kenya and South Africa, and the distributions and associated risk factors for cIMT have been described (Ramsay et al., 2016; Ali et al., 2018; Nonterah et al., 2018). This study aimed to investigate genetic susceptibility

to atherosclerosis in sub-Saharan Africans in the AWI-Gen cohort. cIMT was used as an endophenotype, with further investigation of sex-differences.

Results

Genetic Association with cIMT

Analyses were performed using the imputed dataset of 13.9M SNPs in 7894 participants from the AWI-Gen study and association with mean-max-cIMT. Despite the population sub-structure demonstrated by principal component analysis in the study sample (**Supplementary Figure 1**), our results did not show evidence of genomic inflation (l = 0.997). The genome-wide association results for the combined dataset are illustrated in the Manhattan plot and the genomic inflation by the QQ-plot (**Figure 1a, 1b**). In the combined dataset, we identified two new genome-wide significant loci in *SIRPA* on chromosome 20 (rs6045318, p = 4.7E-08) and *FBXL17* on chromosome 5 (rs552690895, p = 2.5E-08). These two SNPs are African specific and have not been observed in European or Asian populations. Other suggestive association signals had lead variants located in an intergenic region on chromosome 8 (rs11781274, p = 1.8E-07), an intronic region in *SORCS1* (rs11193156, p = 2.1E-07), an intronic region in *ANKK1* (rs11214599, p = 5.4E-07), an exonic region in *CTBP2* (rs3781409, p = 6.6E-07) and an intronic region in *SMARCA2* (rs1324201, p = 8.6E-07) (**Table 1, Supplementary Table 1**).

Sex-specific analyses revealed four significant loci (as illustrated in the Miami plots, **Figure 2**): two male-specific loci led by intronic variants in *SNX29* (rs190770959, p = 6.3E-9) and *MAP3K7* (rs284509, p = 5.3E-8 (near significance)), and two female-specific loci one located in an intergenic region between *UACA* and *LRP6* (a downstream variant located in the promoter flanking region) (rs150840489, p = 2.4E-09) and a variant in a transcription factor binding site near *PROK1/ CYMP* (rs115473055, p = 1.0E-08). Loci with suggestive associations (p < 1E-06) included variants in *AK2P2*, *RBMS3*, *FIP1L1:LNX1*, *CDH18*, *FLT4*, *FOXK1*, *CDH17*, and *C16orf45* in female-specific analysis (**Table 2, Supplementary Table 1**), whereas for male-specific analysis suggestive associations were identified in *U3*, *OR2T35*, *FBXL17*, *SORCS1*, and *USP12* (**Table 3, Supplementary Table 1**).

Sex-differences that were limited to men or women were assessed (**Supplementary Figure 2**). We found suggestive signals in intergenic region of *IGFBPL/ FAM95C* (rs12350396, p = 3.4E-07), *UACA/ LARP6* (rs150840489, p = 4.4E-07), *MYOD1/ KCNC1* (rs150481830, p = 4.9E-07), and in an intronic region of *CROCC* (rs11585710, p = 4.8E-07). These sex-difference are shown in **Table 4 and Supplementary Table 1**. Regional plots of significant loci are shown in **Figure 3**.

Analysis of sex-dimorphism requires both a significant SNP association with cIMT in at least in one sex and a significant sex-difference for the SNP association (P-value testing for difference in sex-specific effect estimates). Several scenarios can describe sexual dimorphism for SNP associations: (i) concordant effect direction (CED); (ii) single sex effect (SSE); or (iii) opposite effect direction (OED) (Randall *et al*., 2013). In our study, we identified all three types of sexual dimorphism: the *LARP6* locus was a case of a single sex effect (rs150840489: p-female = 2.4E-09, beta- female = -0.051; p- male = 0.17, beta- male = 0.012); the *CROCC* locus showed opposite effect direction (rs11585710: p-female = 490E-03, beta- female = 0.007; p-male = 2.7E-05, beta-male = -0.012), and the *FBXL17* variant showed a concordant effect direction (rs547840497, p-female = 0,037, beta-female = -0.022; p-male = 1.8E-07, beta-male= -0.062). In total 177 SNPs showed CED, 89 SNPs had OED and 3213 SNPs showed SSE.

Replication with GWAS Catalog

In view of the limited number of genome-wide significant SNPs for cIMT previously reported, our replication analysis also included screening for phenotypes similar to cIMT (coronary artery calcification (CAC), abdominal aortic aneurysm (AAA)). Our study replicated (see criteria in Methods section) the locus for association with cIMT in the *CBFA2T3* region with a SNP that is 18979 bp from and associated SNP reported in Europeans: rs9934287 (p = 6.6E-06) was suggestively associated in our study in the *CBFA2T3* region. The rs844396 (p = 6.00E-09) that was previously reported by Franceschini and colleagues in a European ancestry population (Franceschini *et al.*, 2018) and later replicated by a UK Biobank analysis of cIMT (Strawbridge *et al.*, 2020), did not replicate in our study (p = 0.85). The rs9934287 SNP (MAF=0.047 in our study) is monomorphic in populations of European ancestry from the 1000 Genomes Project.

A previously reported locus for association with carotid plaque in European populations (Pott *et al.*, 2017) at *GEM* (rs72672639, p = 4.0E-06) was suggestively associated in our female-specific subset with two SNPs (rs78571209, rs76489670, p = 7.8E-05) located approximatively 2200 bp from the SNP reported for plaque in Europeans. Similarly, the association with the *MRPL37* locus (rs11206301, p= 8.00E-06) for plaque in European populations was suggestively associated in our male-specific analysis for cIMT (rs13374450, p = 3.0E-05). The two SNPs in the *MRPL37* locus were not in LD despite their proximity (201 bp). The suggestive variant in our study rs4773141 (p = 4.7E-05, in the combined dataset), located in *COL4A1*, was previously reported for CAD (p = 4.0E-17) in European populations (Van der Harst and Verweij, 2017).

In our combined analysis, a total of 10 SNPs replicated for CAC (a surrogate marker of atherosclerosis as cIMT) (Inouye *et al.* 2012; O'Donnell *et al.* 2011) and for CAC in Type 2 Diabetes African patients (Divers *et al.*, 2017). Fourteen SNPs replicated for coronary heart disease and coronary artery disease (**Supplementary Table 2-3**).

In female-specific GWAS, 13 SNPs were suggestive of replication for CAC (Inouye *et al.*, 2012; Wojczynski *et al.*, 2013; Divers *et al.*, 2017), and 16 SNPs replicated loci for coronary heart diseases, coronary artery disease, coronary aneurism and coronary atherosclerosis. Five variants (p < 3.1E-05) replicated a locus on *MIR100HG* reported for association with stroke in African Americans (Carty *et al.*, 2015), and rs114299344 (p = 2.2E-05, beta = -0.043) replicated the *ADAMTS2* locus in paediatric stroke (rs469568, p = 8.0E-06) (Arning *et al.*, 2012) (**Supplementary Table 2-3**).

For male-specific analyses, 18 SNPs replicated 6 loci for CAC (Inouye *et al.*, 2012; Wojczynski *et al.*, 2013; Divers *et al.*, 2017). For coronary heart disease, coronary artery disease, and coronary atherosclerosis, 13 SNPs replicated loci (*ZNF652*, *ZFPM2-AS1/ZFPM2*, *PKD2L1*, *CFDP1*, *AC027506.1/AC007948.1*, *AC096558.2/ARHGAP15/AC096558.1*, *C1GALT1*, *SLC22A3*, *LPAL2*, *LPA*). rs75601989 (p = 2.0E-05) replicated *RN7SL363P/FURIN* locus for stroke (Malik *et al.*, 2018) (**Supplementary Table 2-3**).

Functional annotation

Annotation of the genic positions of the 467, 515 and 581 SNPs respectively from combined, female-specific and male-specific analyses with significant and suggestive associations (p<1E-05) showed that these were mostly intronic or intergenic. 50 SNPs displayed a CADD score above 12.37 suggestive of being potentially deleterious (19 in the combined; 18 in female-specific; 13 in male-specific datasets) (**Supplementary Table 4a, 4b, 4c**). In the female-specific sample, the lead SNP in *CYMP* (rs115473055) had a Regulome DB score of 2a suggesting the variant was likely affecting a transcription binding site (CTCF). Positional mapping, eQTL mapping (matched cis-eQTL SNPs) and chromatin interaction mapping (on the basis of 3D DNA–DNA interactions) is reported (**Supplementary Table 5a, 5b, 5c**). We found that rs78172571, in high LD with rs150840489 (the top SNP

associated in our female-specific), was involved in HiC type chromatin interactions in multiple tissues including aorta, in which the variant acts as an enhancer of *THAP10* (FDR = 2.03E-17).

Gene-based and gene-set analysis

In a gene-based analysis (using MAGMA threshold of p< 2.6E-06) of the combined dataset analysis there was a significant association with *CALD1* (p = 5.9E-07) (**Supplementary Figure 3A**) with Mean-Max cIMT, whereas in female-specific analysis *FLT4* (p = 4.3E-07) was significantly associated (**Supplementary Figure 3B**). The results from gene-set analysis in the combined dataset showed significant enrichment for “Chemical and Genetic perturbation” gene-set (adjP = 3.9E-05). The female-specific analysis revealed significant enrichment of gene-sets (**Supplementary Table 6a, 6b, 6c**), with among them “Hallmark gene-sets for Oestrogen response”, with “Early Oestrogen response” and “Late Oestrogen response” both being significant (2.2E-6).

Discussion

In this African population GWAS for cIMT as the outcome variable, and as a proxy for atherosclerosis, we identified two new loci associated with cIMT in the full dataset, two new loci specific to the female only analysis and two loci associated in the male only analysis (p < 5E-08). We replicated regional associations with known loci associated with atherosclerosis and CVDs, but with different lead SNPs at almost all loci. In total, 54 loci were replicated with SNPs at p<1E-04 within 25 kb of previously reported genome-wide significantly associated variants. Those loci were associated with the following atherosclerosis phenotypes (cIMT, carotid plaque, coronary artery calcification, and abdominal aortic aneurism) and outcomes (coronary artery disease, coronary heart disease, myocardial infarction, and stroke).

Measurements of cIMT are used clinically to assess vascular pathophysiology and to reflect the atherosclerosis process. Our study identified cIMT-associated loci relevant to genes related to macrophage activity and polarisation (*SIRPA*), to vascular smooth muscle cells (*MAP3K7, CALD1*), to vascular endothelial growth (*PROK1, FLT4*), to collagen synthesis and plaque stability (*LARP6*), and a pathway of blood vessel occlusion (*SNX29*) (**Figure 4**).

The associated loci are discussed with regard to their potential functions and biological evidence from previous studies and previous reports from GWAS (**Extended discussion in Supplementary Note**).

FBXL17 (lead SNP:rs552690895; p = 2.5E-08) in the combined data set, is linked to cardiovascular physiology through its involvement in protein degradation where it plays a central role in cardiovascular physiology and disease: from endothelial function, the cell cycle, atherosclerosis, myocardial ischaemia, cardiac hypertrophy, inherited cardiomyopathies, and heart failure. A GWAS in Lithuanian families found that variants in *FBXL17* were associated with coronary heart diseases (Domarkiene *et al.*, 2013).

Signal regulatory protein alpha (*SIRPA*) (lead SNP:rs6045318; p = 4.7E-08 in the combined data set) has a role in the mediation of phagocytosis and polarization of macrophages which is important in the pathophysiology of atherosclerosis (Chen *et al.*, 2019). There is evidence that *SIRPA* is involved in discrete stages of cardiovascular cell lineage differentiation (Skelton *et al.*, 2014) and that defects in the gene (knock out) reduces atherosclerosis in mice (Szilagyi *et al.*, 2014). *SIRPA* expression has been found as a signature of inflamed atherosclerotic plaque (Puig *et al.*, 2011).

On the chromosome 16, rs147978408 (p = 6.3E-09) was the top cIMT associated variant in *SNX29* for the male-specific analysis. The sorting nexin (*SNX*) family genes are associated with CVDs, and dysfunction of the *SNX* pathway is involved in several forms of cardiovascular disease (CVD) (Yang *et al.*, 2019). In a study of genes that regulate smooth muscle cell differentiation and disease risk, *SNX29* was involved in pathways for occlusion of blood vessels and atherosclerosis (Iyer *et al.*, 2018). Ito and colleagues identified sex-dependent differentially methylated regions close to *SNX29* in mouse liver and found that this methylation status was influenced by testosterone and contributed to sex-dimorphic chromatin decondensation (Ito *et al.*, 2015). This might explain the sex-specific effect observed in our study. Because of the previous link between *SNX29* and hypertension, we ran further GWAS analysis stratified by hypertensive status and found that the association was driven by the hypertensive group (effect three times higher in hypertensives compared to the non-hypertensives), therefore demonstrating that the association of *SNX29* with cIMT might be mediated by the vascular remodeling caused by hypertension.

In the male-specific analysis, rs284509 (p=5.3E-08) in *MAP3K7* region on chromosome 6 to was associated with cIMT. Mitogen-activated protein kinase kinase kinase 7 (*MAP3K7*) is known to play a role in growth inhibition in vascular smooth muscle cells. The sex-specific association observed might be related to the fact that *MAP3K7IP3* (located on the X chromosome), which is known to form a ternary complex with *MAP3K7* in response to inflammatory stimuli, has shown sex-differential expression in ischemic stroke (Stamova *et al.*, 2012; Rocha *et al.*, 2016). In a study on expression of androgen-modulated micro-RNAs, it was reported that *MAP3K7* was a target of mmu-miR-467h and mmu-miR-669i in the angiogenesis and transforming growth factor beta receptor signalling pathways (Bouhaddioui, Provost and Tremblay, 2016). Our study is the first to report *MAP3K7* association with a CVD phenotype.

LARP6 (La-related protein 6) is a ribonucleoprotein domain family member 6 with a role in collagen regulation by targeting mRNA encoding Type I collagen (Zhang and Stefanovic, 2016; Stefanovic *et al.*, 2019). Collagen is a hallmark of atherosclerotic plaque stability, thus alteration of the collagen balance may lead to an instability of atherosclerotic lesions, and therefore promote plaque formation and rupture (Puig *et al.*, 2011; Higashi *et al.*, 2016). In the Taiwanese population, the *LARP6* locus was found to be associated with coronary artery disease (Assimes *et al.*, 2016). Myocardial gene expression in non-ischemic human heart failure found that *LARP6* was differentially expressed between men and women (1.36 fold) (Fermin *et al.*, 2008). The female-specific effect of this locus in our study may be explained by the enhancer function of rs78172571 in high LD with rs150840489 (the top SNP associated in our female-specific) on *THAP10* gene (FDR = 2.03E-17), known to be regulated by oestrogen.

Our study is the first to report prokineticin 1 (*PROK1)* for any trait in a GWAS. It was associated with cIMT in the female-specific analysis (lead SNP:rs115473055, p=1.00E-08). *PROK1* is a specific placental angiogenic factor that plays a role in the control of normal (e.g. endometrial decidualization) and pathological placental angiogenesis (Hoffmann *et al.*, 2006). The gene is known to be predominantly expressed in the steroidogenic glands, such as ovary, testis, and adrenal cortex, and is often complementary to the expression of vascular endothelial growth factor (*VEGF*), suggesting that these molecules function in a coordinated manner. The function and particular pattern of this gene's activity might explain why we identified the locus only in our female-specific analysis.

Our gene-based analysis identified caldesmon 1 (*CALD1*) as significantly associated with cIMT in our combined set led by rs7781307 (p = 2.1E-06) on 7q33. This gene plays a major role in the regulation of smooth muscle contraction, cell migration and cell invasion (Mayanagi and Sobue, 2011). *CALD1* was linked to advanced coronary atherosclerosis (Tan *et al.*, 2017) and abdominal aortic aneurysm (Wan *et al.*, 2018). Under expression of *CALD1* was found to be a key feature of calcification of vascular smooth muscle cells from atherosclerotic plaque (Puig *et al.*, 2011; Goikuria *et al.*, 2018; Trillhaase *et al.*, 2018).

Additionally, studies on epigenetic modifications reported *CALD1* to exhibit differential methylation in atherosclerosis (Zaina *et al.*, 2014; Nazarenko *et al.*, 2015; Fernández-Sanlés *et al.*, 2017).

Our study is the first to report and association of *FLT4* (also known as vascular endothelial growth factor receptor 3 (*VEGFR3*)) (rs112967731; p = 5.7E-07, female-specific) with cIMT or any cardiovascular phenotype in GWAS studies. *FLT4* is a major signalling protein involved in angiogenesis, vasculogenesis and maintenance of the endothelium. Defect and/or downregulation of *VEGR3* was found to lead to cardiovascular failure in embryonic stage and to higher mortality after myocardial infarction in mouse models (Dumont *et al.*, 1998; Vuorio *et al.*, 2018). Biological studies have highlighted the role of *FLT4* in atherosclerosis in major pathological processes. The gene has been reported to be involved in plaque instability by two process: the mediation of monocytes/macrophages apoptosis and consequently alteration plaque stability (Schmeisser *et al.*, 2006); and the modulation of vascular remodelling and shear stress resulting in plaques haemorrhages and calcification in carotid arteries (Baeyens *et al.*, 2015; Tuenter *et al.*, 2016; Camaré *et al.*, 2017).

The SNP rs116517341, which leads the association with the *CCDC71L* locus in our male-specific analysis (p=6.30E-05), is located over 100 kb from the lead-SNP found in European, but our lead SNP was closer to *CCDC71L* than that found by the study of cIMT in Europeans. Therefore, different variants may influence the association of the *CCDC71L* gene in cIMT. When analysing the variants in *CBFA2T3*, our lead-SNP (rs9934287) was located 18,979 bp away from the SNP reported by Franceschini et al. (2018). The regional plot showed more dense signals with numerous SNPs in LD with the lead-SNP in European, whereas our lead- SNP had fewer SNPs in high LD (**Figure 5 C-D**). The LD structure using 1000 Genomes Project Europen (CEU) compared to African populations (YRI-LWK-GWD-MSL-ESN) showed that LD blocks were smaller in Africans (**Supplementary Figure 4**), providing opportunity for extended fine-mapping and reducing the credible set toward identifying causal variants.

Our sex-specific analyses revealed loci that support the hypothesis that sex differences may be due to sex-specific epigenetic modification, independent of sex hormone levels. When analysing sex-specific or gene-sex interactions, it is important to keep in mind that they also reflect the influences of non-genetic factors such as behaviour, as evidence by the previously reported gene-smoking interactions (Boua *et al.*, 2020). Hence, environmental exposure, anatomical differences, and sex hormone environment, which create systemic differences between males and females for trait expression, affect disease risk and heritability (Gilks, Abbott and Morrow, 2014).

Our study identified significant enrichment of oestrogen pathway genes in our female-specific analysis. Oestrogen-dependent regulation of vascular gene expression and vascular physiology encompasses complex processes involving both nuclear and membrane-associated oestrogen signalling pathways. In recent years we have witnessed major progress in understanding how these regulatory processes contribute to the atheroprotective effects exerted by oestrogens. Animal models of atherosclerosis provided compelling evidence that physiological oestrogen levels potently attenuate both early and advanced stages of atherosclerosis lesion development in females, and suggested similar protective effects in males. The effect of oestrogens on atherosclerosis can target metabolism (lipid, glucose), macrophage function or smooth muscle cells. Nonetheless, hormone replacement therapy during menopause has not been shown to conclusively reduce atherosclerosis risk, suggesting that more studies are needed to fully decipher the biological mechanisms.

Strengths and limitations

Our study is the first population-based study to investigate the genetic architecture of cIMT in sub-Saharan African populations. In addition to our sex-specific analyses, we tested for sex-difference between the two strata using a minimal model (adjustment for age and PCs) to avoid the “collider bias”. We used an analysis framework allowing us to identify genetic effects that point in opposite directions in men and women and to detect genetic effects that are only (or more pronounced) in one stratum, a method that has been shown to have better power to identify qualitative gene-sex interactions (Winkler *et al.*, 2017). The use of a new SNP genotyping array with better representation of common African variants and imputation reference panels from African participants has improved the SNP coverage in ethnically diverse African populations.

The lack of an ethnically matched replication cohort is a limitation in our study, and it will be important to replicate these findings in additional suitable cohorts. We identified African-specific variants in new loci and replicated previously reported loci, revealing opportunities for trans-ancestry fine-mapping.

We found evidence of gene set enrichment for biological processes. Our study is the first GWAS to report significant enrichment of genes in the oestrogen pathway for cIMT in our female-specific analysis. The findings from our study support the notion that genomics studies in Africa are likely to contribute to the understanding of complex traits, such as atherosclerosis.

Materials And Methods

Study population and phenotype assessments

This is a cross-sectional study that investigated populations from six sub-Saharan African sites in West Africa (Burkina Faso (Nanoro) and Ghana (Navrongo)), East Africa (Kenya (Nairobi)) and South Africa (Agincourt, Dikgale and Soweto) as part of the AWI-Gen study (Richter *et al.*, 2007; Derra *et al.*, 2012; Kahn *et al.*, 2012; Oduro *et al.*, 2012; Alberts *et al.*, 2015; Beguy *et al.*, 2015; Ramsay *et al.*, 2016; Ali *et al.*, 2018). The participants for this study include 10,703 black African men and women from two urban settings (Nairobi and Soweto) and four rural settings (Agincourt, Dikgale, Nanoro and Navrongo), aged 40 to 60 years. Participants completed a questionnaire requesting information on demography, health history and behaviour. Anthropometric measurements were taken and blood collected for genotyping (H3Africa SNP array) and phenotyping (biomarkers) (Ali *et al.*, 2018). Ultrasound scans were performed to assess cIMT of the right and left carotid arteries. No cIMT data was collected for female participants from Soweto because they were drawn from the Study of Women Entering and Endocrine Transition (SWEET) study, and they were therefore not included in the subsequent GWAS. This study received approval from the Human Research Ethics Committee (Medical), University of the Witwatersrand, South Africa (M121029, M1706110). All the participants provided written informed consent prior to enrolment and participation in the study.

cIMT Measurement

cIMT was measured using Dual B-mode ultrasound images of the carotid tree showing a typical double line for the arterial wall. Details of the method for measurement are provided in Ali et al. 2018 (Ali *et al.*, 2018). The cIMT values were QCed according to the Mannheim Consensus defining the use of cIMT in population-based studies. The Mean Max cIMT was generated as the average of the maximum cIMT from the left and right, and this value was used for the GWAS analyses.

Genotyping and Imputation

The H3Africa genotyping array (https://chipinfo.h3abionet.org), designed as an African-common-variant-enriched GWAS array (Illumina) with ~2.3 million SNPs, was used to genotype genomic DNA using the Illumina FastTrack Sequencing Service (https://www.illumina.com/services/sequencing-services.html). The following pre-imputation QC steps were applied to the entire AWI-Gen genotype data set. Individuals with a missing SNP calling rate greater than 0.05 were removed. SNPs with a genotype missingness greater than 0.05, MAF less than 0.01 and Hardy-Weinberg equilibrium (HWE) P-value less than 0.0001 were removed. Non-autosomal and mitochondrial SNPs, and ambiguous SNPs that did not match the GRCh37 reference alleles

or strands were also removed. Imputation was performed on the cleaned dataset (with 1,729,661 SNPs and 10,903 individuals) using the Sanger Imputation Server and the African Genome Resources as reference panel. We selected EAGLE2 (Loh *et al.*, 2016) for pre-phasing and the default PBWT algorithm was used for imputation. After imputation, poorly imputed SNPs with info scores less than 0.6, MAF less 0.01, and HWE P-value less than 0.00001 were excluded. The final QC-ed imputed data had 13.98 M SNPs, and only participants with both good quality cIMT and genotyping data (n = 7894) were used for the GWAS analyses.

Genome-wide association analysis

Linear regression of Mean Max cIMT was performed with covariates in R (https://www.R-project.org/). Residuals were extracted from the linear regression analyses and used for the GWAS analysis. We used as covariates age, sex and 8 principal components (PCs) computed on genetics data. In our sex stratified analysis (3963 women, 3931 men), the covariates were age and 5 PCs. The number of PCs to include in each model were determined using a stepwise regression. We performed all association testing with the residuals in BOLT-LMM, which implement testing using a Linear Mixed Model (LMM). To run efficiently, BOLT-LMM required three components: the (imputed) genotypic data for association testing; a reference panel of LD scores per SNP, calculated using 1000 Genomes Project African samples; and genotype data used to approximate a genetic relationship matrix (GRM) (Using a subset of the SNP Array genotypes with LD filtering). This method is expected to account for all forms of relatedness, ancestral heterogeneity in the samples and other (potentially hidden) structure in the data. The analyses were run on the automated workflow of H3abionet/H3agwas (http://github.com/h3abionet/h3agwas/) (Baichoo et al., 2018). We screened the output for a genome-wide significance threshold (p-values < 5.E-08). To assess genomic inflation, we compared our observed distribution of −log10(P) values to that expected in the absence of association (Lambda) and illustrated the results in QQ plots. The same process was applied for sex-stratified analyses.

We used EasyStrata (Winkler *et al.*, 2015) to test for the joint effect calculated from sex strata results (Aschard *et al.*, 2010) and to test for the difference between two strata results as a means to test for sex effects (Randall *et al.*, 2013). The joint and stratified frameworks were found to be the most efficient way to test for gene-environment interactions (Sung *et al.*, 2016). Power calculations were performed with Quanto (Version 1.2.4) (http://biostats.usc.edu/Quanto.html).

Replication from the GWAS catalog

The GWAS Catalog database was downloaded (https://www.ebi.ac.uk/gwas/, accessed on 12 Jan 2019) and a subset of the data generated using the following key words relevant to our study: coronary artery disease, carotid atherosclerosis, cIMT, coronary artery calcification and abdominal artery aneurism. The marker co-ordinates from the GWAS Catalog are given in build 38. Since our dataset was in build 37, we performed lift-over of GWAS Catalog to build 37 in order to allow accurate comparison. In order to look whether our study was replicating previous findings, we searched for the same marker or any markers within 25 kb (considering the highest mean size of LD blocks). We then searched for SNPs in a 25 kb region of all SNPs with suggestive associations (p-value < 1E-04) found in our study. We further defined loci by grouping SNPs with p-value < 1E-04 within 250 kb of each other. These loci were used for regional replication and transferability analyses.

Functional analysis

The FUMA online platform (http://fuma.ctglab.nl/) (Watanabe *et al.*, 2017) was used to annotate, prioritize, visualize and interpret GWAS results. GWAS summary statistics (p<1E-05) from out study was used as the input. FUMA provided extensive functional annotation for all SNPs in genomic areas identified by lead SNPs. From the list of gene IDs (as identified by SNP2GENE option in FUMA) FUMA annotated genes in a biological context (Watanabe *et al.*, 2017). We selected all candidate SNPs in the associated genomic region having $r^2 \geq 0.6$ (with 1000 Genome Project African references) with one of the independently significant SNPs, with a suggestive P-value (p < 1E-05) and MAF > 0.01 for annotation. Predicted functional consequences for these SNPs were obtained by matching the SNP's chromosome base-pair position, and reference and alternate alleles, to databases containing known functional annotations, including ANNOVAR (Wang, Li, and Hakonarson 2010), combined annotation-dependent depletion (CADD) scores (Kircher *et al.*, 2014), and RegulomeDB (RDB) (Boyle *et al.* 2012) scores. Additionally, eQTLs scans (GTex Consortium *et al.*, 2017) were performed.

Functional annotation of mapped genes

Genes implicated by mapping of significant GWAS SNPs were further investigated using the GENE2FUNC procedure in FUMA (Watanabe *et al.*, 2017), which provides hypergeometric tests of enrichment of the list of mapped genes in 53 GTEx tissue-specific gene expression sets (GTex Consortium *et al.*, 2017), 7,246 MSigDB gene sets (http://software.broadinstitute.org/gsea/msigdb), and chromatin states (Consortium Roadmap Epigenomics *et al.*, 2015).

MAGMA Gene-based and gene-sets analysis

Multi-marker analysis of genomic annotation (MAGMA, v1.6) gene analysis were performed using summary statistics of our association results as input. Gene-based analysis enabled summarization of SNPs associations at the gene level and association of the set of genes to biological pathways. MAGMA employs multiple linear regression to obtain gene-based p-values (de Leeuw *et al.*, 2015; Watanabe *et al.*, 2017). The window for gene annotation was set for 25kb and genome-wide significance was set at 0.05/number of tested genes. MAGMA gene-set analysis used a competitive testing framework, with gene-sets from MsigDB (v6.2, 10678 gene sets (curated gene sets: 4761, GO terms: 5917)) (Liberzon *et al.*, 2015). MAGMA analysis was implemented within FUMA.

Declarations

Data Availability

AWI-Gen phenotype dataset is available at study number EGA00001002482 [https://ega459 archive.org/datasets/EGAD00001006425]. AWI-Gen genotype dataset accession id: EGAD00010001996 [https://ega-archive.org/datasets/EGAD00010001996].

Author's contribution

PRB, HS, HT, AC, CM, and MR designed the study. PRB and J-TB performed the analysis. PRB wrote the manuscript. All authors critically reviewed and approved the manuscript.

Acknowledgements

This study would not have been possible without the generosity of the participants who spent many hours responding to questionnaires, being measured and having samples taken. We wish to acknowledge the sterling contributions of our field workers, phlebotomists, laboratory scientists, administrators, data personnel, and all other staff who contributed to the data and sample collections, processing, storage, and shipping. Investigators responsible for the conception and design of the AWI-Gen study include the following: MR (PI, Wits), Osman Sankoh (co-PI, INDEPTH), Stephen Tollman, and Kathleen Kahn (Agincourt PI), Marianne Alberts (Dikgale PI), Catherine Kyobutungi (Nairobi PI), HT (Nanoro PI), Abraham Oduro (NavrongoPI), Shane Norris (Soweto PI), and SH, Nigel Crowther, Himla Soodyall, and Zane Lombard (Wits). We would like to acknowledge each of the following investigators for their significant contributions to this research, mentioned according to affiliation: Wits AWI-Gen Collaborative Centre: Stuart Ali, AC, SH, Freedom Mukomana, Cassandra Soo; Soweto (DPHRU): Nomses Baloyi, Yusuf Guman.

Funding

This study was funded by the National Institutes of Health (NIH) through the H3Africa AWI-Gen project (NIH grant number U54HG006938) and the Wits Non-Communicable Disease Research Leadership Programme (NIH Fogarty International Centre grant number D43TW008330). AWI-Gen is supported by the National Human Genome Research Institute (NHGRI), Eunice

Kennedy Shriver National Institute of Child Health & Human Development (NICHD), Office of the Director (OD) at the National Institutes of Health. PRB was funded by the National Research Foundation/The World Academy of Sciences “African Renaissance Doctoral Fellowship” (Grant no. 100004).

References

Alberts, M. *et al.* (2015) ‘Health & Demographic Surveillance System Profile: The Dikgale Health and Demographic Surveillance System’, *International Journal of Epidemiology*, 44(5), pp. 1565–1571. doi: 10.1093/ije/dyv157.

Ali, S. A. *et al.* (2018) ‘Genomic and environmental risk factors for cardiometabolic diseases in Africa: methods used for Phase 1 of the AWI-Gen population cross-sectional study’, *Global Health Action*. Taylor & Francis, 11(sup2). doi: 10.1080/16549716.2018.1507133.

Arning, A. *et al.* (2012) ‘Agenome-wide association study identifies a gene network of ADAMTS genes in the predisposition to pediatric stroke’, *Blood*, 120(26), pp. 5231–5236. doi: 10.1182/blood-2012-07-442038.

Aschard, H. *et al.* (2010) ‘Genome-wide meta-analysis of joint tests for genetic and gene-environment interaction effects’, *Human Heredity*, 70(4), pp. 292–300. doi: 10.1159/000323318.

Assimes, T. L. *et al.* (2016) ‘Genetics of Coronary Artery Disease in Taiwan: A Cardiometabochip Study by the Taichi Consortium’, *PloS one*, 11(3), p. e0138014. doi: 10.1371/journal.pone.0138014.

Baeyens, N. *et al.* (2015) ‘Vascular remodeling is governed by a VEGFR3-dependent fluid shear stress set point’, *eLife*, 4, pp. 1–16. doi: 10.7554/elife.04645.

Baichoo, S. *et al.* (2018) ‘Developing reproducible bioinformatics analysis workflows for heterogeneous computing environments to support African genomics’, *BMC Bioinformatics*. BMC Bioinformatics, 19(1), p. 457. doi: 10.1186/s12859-018-2446-1.

Beguy, D. *et al.* (2015) ‘Health & Demographic Surveillance System Profile: The Nairobi Urban Health and Demographic Surveillance System (NUHDSS) Donatien’, *International Journal of Epidemiology*, 44(5), pp. 1565–1571. doi: 10.1093/ije/dyv157.

Bis, J. C. *et al.* (2012) ‘Meta-analysis of genome-wide association studies from the CHARGE consortium identifies common variants associated with carotid intima media thickness and plaque’, *Nat Genet*, 43(10), pp. 940–947. doi: 10.1038/ng.920.Meta-analysis.

Boua, P. R. *et al.* (2020) ‘Novel and known gene-smoking interactions with cIMT identified as potential drivers for atherosclerosis risk in West-African populations of the AWI-Gen Study’, *Frontiers in Genetics*, 10(1), p. 100004. doi: 10.3389/fgene.2019.01354.

Bouhaddioui, W., Provost, P. R. and Tremblay, Y. (2016) ‘Expression profile of androgen-modulated microRNAs in the fetal murine lung’, *Biology of Sex Differences*. Biology of Sex Differences, 7(1), pp. 1–13. doi: 10.1186/s13293-016-0072-z.

Boyle, A. P. *et al.* (2012) ‘Annotation of functional variation in personal genomes using RegulomeDB’, *Genome Research*, 22(9), pp. 1790–1797. doi: 10.1101/gr.137323.112.

Buniello, A. *et al.* (2019) ‘The NHGRI-EBI GWAS Catalog of published genome-wide association studies, targeted arrays and summary statistics 2019’, *Nucleic Acids Research*. Oxford University Press, 47(D1), pp. D1005–D1012. doi: 10.1093/nar/gky1120.

Camaré, C. *et al.* (2017) ‘Angiogenesis in the atherosclerotic plaque’, *Redox Biology*. Elsevier, 12(January), pp. 18–34. doi: 10.1016/j.redox.2017.01.007.

Carty, C. L. *et al.* (2015) ‘Meta-Analysis of Genome-Wide Association Studies Identifies Genetic Risk Factors for Stroke in African Americans’, *Stroke; a journal of cerebral circulation*, 46(8), pp. 2063–2068. doi: 10.1161/STROKEAHA.115.009044.

Chen, W. *et al.* (2019) 'miR-378a Modulates Macrophage Phagocytosis and Differentiation through Targeting CD47-SIRPα Axis in Atherosclerosis', *Scandinavian Journal of Immunology*, 90(e12766), pp. 1–10. doi: https://doi.org/10.1111/sji.12766.

Consortium Roadmap Epigenomics *et al.* (2015) 'Integrative analysis of 111 reference human epigenomes', *Nature*, 518(7539), pp. 317–330. doi: 10.1038/nature14248.

Derra, K. *et al.* (2012) 'Profile: Nanoro health and demographic surveillance system', *International Journal of Epidemiology*, 41(5), pp. 1293–1301. doi: 10.1093/ije/dys159.

Dimas, A. S. *et al.* (2012) 'Sex-biased genetic effects on gene regulation in humans', *Genome Research*, 22(12), pp. 2368–2375. doi: 10.1101/gr.134981.111.

Divers, J. *et al.* (2017) 'Genome-wide association study of coronary artery calcified atherosclerotic plaque in African Americans with type 2 diabetes', *BMC Genetics*. BMC Genetics, 18(1), pp. 1–13. doi: 10.1186/s12863-017-0572-9.

Domarkiene, I. *et al.* (2013) 'RTN4 and FBXL17 genes are associated with coronary heart disease in genome-wide association analysis of Lithuanian families', *Balkan Journal of Medical Genetics*, 16(2), pp. 17–22. doi: 10.2478/bjmg-2013-0026.

Dong, C. *et al.* (2015) 'Genetic variants in LEKR1 and GALNT10 modulate sex-difference in carotid intima-media thickness: A genome-wide interaction study', *Atherosclerosis*. Elsevier Ltd, 240(2), pp. 462–467. doi: 10.1016/j.atherosclerosis.2015.04.019.

Dumont, D. J. *et al.* (1998) 'Cardiovascular Failure in Mouse Embryos Deficient in VEGF Receptor-3', *Science*, 282(October), pp. 946–950.

Fagnani, C. *et al.* (2013) 'Genetic and environmental components of carotid artery elasticity: An Italian twin study', *European Journal of Internal Medicine*. European Federation of Internal Medicine., 24(4), pp. e53–e54. doi: 10.1016/j.ejim.2013.03.001.

Fasanaro, P., Capogrossi, M. C. and Martelli, F. (2010) 'Regulation of the endothelial cell cycle by the ubiquitin-proteasome system', *Cardiovascular Research*, 85(2), pp. 272–280. doi: 10.1093/cvr/cvp244.

Fermin, D. R. *et al.* (2008) 'Sex and age dimorphism of myocardial gene expression in nonischemic human heart failure.', *Circulation. Cardiovascular genetics*, 1(2), pp. 117–125. doi: 10.1161/CIRCGENETICS.108.802652.

Fernández-Sanlés, A. *et al.* (2017) 'Association between DNA methylation and coronary heart disease or other atherosclerotic events: a systematic review', *Atherosclerosis*, 263, pp. 325–333. doi: 10.1111/ijlh.12426.

Fox, C. S. *et al.* (2003) 'Genetic and environmental contributions to atherosclerosis phenotypes in men and women: Heritability of carotid intima-media thickness in the Framingham heart study', *Stroke*, 34(2), pp. 397–401. doi: 10.1161/01.STR.0000048214.56981.6F.

Franceschini, N. *et al.* (2018) 'GWAS and colocalization analyses implicate carotid intima-media thickness and carotid plaque loci in cardiovascular outcomes', *Nature Communications*, 9(1), p. 5141. doi: 10.1038/s41467-018-07340-5.

Franconi, F. *et al.* (2017) 'Human cells involved in atherosclerosis have a sex', *International Journal of Cardiology*. Elsevier Ireland Ltd, 228, pp. 983–1001. doi: 10.1016/j.ijcard.2016.11.118.

Fuster, V. (2014) 'Global burden of cardiovascular disease: Time to implement feasible strategies and to monitor results', *Journal of the American College of Cardiology*, 64(5), pp. 520–522. doi: 10.1016/j.jacc.2014.06.1151.

Gershoni, M. and Pietrokovski, S. (2017) 'The landscape of sex-differential transcriptome and its consequent selection in human adults', *BMC Biology*. BMC Biology, 15(1), pp. 1–15. doi: 10.1186/s12915-017-0352-z.

Gilks, W. P., Abbott, J. K. and Morrow, E. H. (2014) 'Sex differences in disease genetics: Evidence, evolution, and detection', *Trends in Genetics*. Elsevier Ltd, 30(10), pp. 453–463. doi: 10.1016/j.tig.2014.08.006.

Goikuria, H. *et al.* (2018) 'Characterization of Carotid Smooth Muscle Cells during Phenotypic Transition', *Cells*, 7(3), p. 23. doi: 10.3390/cells7030023.

GTex Consortium *et al.* (2017) 'Genetic effects on gene expression across human tissues', *Nature*, 550(7675), pp. 204–213. doi: 10.1038/nature24277.

Van der Harst, P. and Verweij, N. (2017) 'Identification of 64 Novel Genetic Loci Provides an Expanded View on the Genetic Architecture of Coronary Artery Disease', *Circulation Research*, (122), pp. 433–443. doi: 10.17632/2zdd47c94h.1.

Herrmann, J., Lerman, L. O. and Lerman, A. (2010) 'On to the road to degradation: Atherosclerosis and the proteasome', *Cardiovascular Research*, 85(2), pp. 291–302. doi: 10.1093/cvr/cvp333.

Higashi, Y. *et al.* (2016) 'Abstract 13854: Micro-RNA Regulation of Collagen Production by Vascular Smooth Muscle Cells Mediated by La Ribonucleoprotein Domain Family Member 6: Potential Mechanisms Underlying Stable Phenotype of Atherosclerotic Plaque by Insulin-like Growth Factor 1', *Circulation*. American Heart Association, 134(suppl_1), pp. A13854–A13854. doi: 10.1161/circ.134.suppl_1.13854.

Hinds, D. A. *et al.* (2005) 'Whole-Genome Patterns of Common DNA Variation in Three Human Populations', *Science*, 307(5712), pp. 1072–1079. doi: 10.1126/science.1105436.

Hoffmann, P., Feige, J. J. and Alfaidy, N. (2006) 'Expression and oxygen regulation of endocrine gland-derived vascular endothelial growth factor/prokineticin-1 and its receptors in human placenta during early pregnancy', *Endocrinology*, 147(4), pp. 1675–1684. doi: 10.1210/en.2005-0912.

Inouye, M. *et al.* (2012) 'Novel Loci for Metabolic Networks and Multi-Tissue Expression Studies Reveal Genes for Atherosclerosis', *PLoS Genetics*, 8(8). doi: 10.1371/journal.pgen.1002907.

Ito, S. *et al.* (2015) 'Novel sex-dependent differentially methylated regions are demethylated in adult male mouse livers', *Biochemical and Biophysical Research Communications*. Elsevier Ltd, 462(4), pp. 332–338. doi: 10.1016/j.bbrc.2015.04.137.

Iyer, D. *et al.* (2018) 'Coronary artery disease genes SMAD3 and TCF21 promote opposing interactive genetic programs that regulate smooth muscle cell differentiation and disease risk', *PLoS Genetics*, 14(10), pp. 1–28. doi: 10.1371/journal.pgen.1007681.

Kahn, K. *et al.* (2012) 'Profile: Agincourt health and socio-demographic surveillance system', *International Journal of Epidemiology*, 41(4), pp. 988–1001. doi: 10.1093/ije/dys115.

Kircher, M. *et al.* (2014) 'A general framework for estimating the relative pathogenicity of human genetic variants', *Nature Genetics*. Nature Publishing Group, 46(3), pp. 310–315. doi: 10.1038/ng.2892.

de Leeuw, C. A. *et al.* (2015) 'MAGMA: Generalized Gene-Set Analysis of GWAS Data', *PLoS Computational Biology*, 11(4), pp. 1–19. doi: 10.1371/journal.pcbi.1004219.

Liberzon, A. *et al.* (2015) 'The Molecular Signatures Database Hallmark Gene Set Collection', *Cell Systems*, 1(6), pp. 417–425. doi: 10.1016/j.cels.2015.12.004.

Lin, H.-F. *et al.* (2015) 'Age and sex differences in the effect of parental stroke on the progression of carotid intima-media thickness.', *Atherosclerosis*. Elsevier Ltd, 241(1), pp. 229–233. doi: 10.1016/j.atherosclerosis.2015.02.025.

Loh, P. *et al.* (2016) 'Reference-based phasing using the Haplotype Reference Consortium panel', *Nature Genetics*, 48(11), pp. 1443–1448. doi: 10.1038/ng.3679.

Lonjou, C. *et al.* (2003) 'Linkage disequilibrium in human populations', *Proceedings of the National Academy of Sciences*, 100(10), pp. 6069–6074.

Malik, R. *et al.* (2018) 'Multiancestry genome-wide association study of 520,000 subjects identifies 32 loci associated with stroke and stroke subtypes', *Nature Genetics*, 50(4), pp. 524–537. doi: 10.1038/s41588-018-0058-3.

Mayanagi, T. and Sobue, K. (2011) 'Diversification of caldesmon-linked actin cytoskeleton in cell motility', *Cell Adhesion and Migration*, 5(2), pp. 150–159. doi: 10.4161/cam.5.214398.

Medda, E. *et al.* (2014) 'Heritability of arterial stiffness and carotid intima-media thickness: An Italian twin study', *Nutrition, Metabolism and Cardiovascular Diseases*. Elsevier Ltd, 24(5), pp. 511–517. doi: 10.1016/j.numecd.2013.10.031.

Nazarenko, M. S. *et al.* (2015) 'A comparison of genome-wide DNA methylation patterns between different vascular tissues from patients with coronary heart disease', *PLoS ONE*, 10(4), pp. 1–15. doi: 10.1371/journal.pone.0122601.

Nonterah, E. A. *et al.* (2018) 'Subclinical Atherosclerosis In Sub-Saharan African Countries Is Driven By Prevailing Socio-Economic Conditions, HIV and Classical Cardiovascular Risk Factors: Findings From the Africa-Wits-INDEPTH Partnership For Genomic Studies', *Global Heart*. Elsevier Inc, 13(4), p. 380. doi: 10.1016/j.gheart.2018.09.024.

O'Donnell, C. J. *et al.* (2011) 'Genome-Wide Association Study for Coronary Artery Calcification With Follow-Up in Myocardial Infarction', *Circulation*, 124(25), pp. 2855–2864. doi: 10.1161/circulationaha.110.974899.

Oduro, A. R. *et al.* (2012) 'Profile of the Navrongo health and demographic surveillance system', *International Journal of Epidemiology*, 41(4), pp. 968–976. doi: 10.1093/ije/dys111.

Pott, J. *et al.* (2017) 'Genome-wide meta-analysis identifies novel loci of plaque burden in carotid artery', *Atherosclerosis*, 259, pp. 32–40. doi: 10.1016/j.atherosclerosis.2017.02.018.

Powell, S. R. and Divald, A. (2010) 'The ubiquitin-proteasome system in myocardial ischaemia and preconditioning', *Cardiovascular Research*, 85(2), pp. 303–311. doi: 10.1093/cvr/cvp321.

Puig, O. *et al.* (2011) 'A gene expression signature that classifies human atherosclerotic plaque by relative inflammation status', *Circulation: Cardiovascular Genetics*, 4(6), pp. 595–604. doi: 10.1161/CIRCGENETICS.111.960773.

Ramsay, M. *et al.* (2016) 'H3Africa AWI-Gen Collaborative Centre: a resource to study the interplay between genomic and environmental risk factors for cardiometabolic diseases in four sub-Saharan African countries', *Global Health, Epidemiology and Genomics*, 1(e20). doi: 10.1017/gheg.2016.17.

Randall, J. C. *et al.* (2013) 'Sex-stratified Genome-wide Association Studies Including 270,000 Individuals Show Sexual Dimorphism in Genetic Loci for Anthropometric Traits', *PLoS Genetics*. Edited by G. Gibson, 9(6), p. e1003500. doi: 10.1371/journal.pgen.1003500.

Rawlik, K. *et al.* (2016) 'Evidence for sex-specific genetic architectures across a spectrum of human complex traits', *Genome Biology*. Genome Biology, 17(1), p. 166. doi: 10.1186/s13059-016-1025-x.

Richter, L. *et al.* (2007) 'Cohort profile: Mandela's children: The 1990 birth to twenty study in South Africa', *International Journal of Epidemiology*, 36(3), pp. 504–511. doi: 10.1093/ije/dym016.

Rocha, D. M. *et al.* (2016) 'Saturated fatty acids trigger TLR4-mediated inflammatory response', *Atherosclerosis*. Elsevier Ltd, 244, pp. 211–215. doi: 10.1016/j.atherosclerosis.2015.11.015.

Roth, G. A. *et al.* (2017) 'Global, Regional, and National Burden of Cardiovascular Diseases for 10 Causes, 1990 to 2015.', *Journal of the American College of Cardiology*, 70(1), pp. 1–25. doi: 10.1016/j.jacc.2017.04.052.

Sacco, R. L. *et al.* (2009) 'Heritability and Linkage Analysis for Carotid Intima-Media Thickness: The Family Study of Stroke Risk and Carotid Atherosclerosis', *Stroke*, 40(7), pp. 2307–2312. doi: 10.1161/STROKEAHA.109.554121.

Schmeisser, A. *et al.* (2006) 'Apoptosis of human macrophages by Flt-4 signaling: Implications for atherosclerotic plaque pathology', *Cardiovascular Research*, 71(4), pp. 774–784. doi: 10.1016/j.cardiores.2006.06.012.

Shendre, A., Wiener, H. W., *et al.* (2017) 'Genome-wide admixture and association study of subclinical atherosclerosis in the Women's Interagency HIV Study (WIHS)', *PLoS ONE*, 12(12), pp. 1–23. doi: 10.1371/journal.pone.0188725.

Shendre, A., Irvin, M. R., *et al.* (2017) 'Local Ancestry and Clinical Cardiovascular Events Among African Americans From the Atherosclerosis Risk in Communities Study', *Journal of the American Heart Association*, 6(4). doi: 10.1161/JAHA.116.004739.

Skelton, R. J. P. *et al.* (2014) 'SIRPA, VCAM1 and CD34 identify discrete lineages during early human cardiovascular development', *Stem Cell Research*. Elsevier B.V., 13(1), pp. 172–179. doi: 10.1016/j.scr.2014.04.016.

Stamova, B. *et al.* (2012) 'The X-chromosome has a different pattern of gene expression in women compared with men with ischemic stroke', *Stroke*, 43(2), pp. 326–334. doi: 10.1161/STROKEAHA.111.629337.

Stangl, K. and Stangl, V. (2010) 'The ubiquitin-proteasome pathway and endothelial (dys)function', *Cardiovascular Research*, 85(2), pp. 281–290. doi: 10.1093/cvr/cvp315.

Stefanovic, B. *et al.* (2019) 'Discovery and evaluation of inhibitor of LARP6 as specific antifibrotic compound', *Scientific Reports*, 9(1), pp. 1–15. doi: 10.1038/s41598-018-36841-y.

Strawbridge, R. J. *et al.* (2020) 'Carotid intima-media thickness novel loci, sex-specific effects, and genetic correlations with obesity and glucometabolic traits in UK Biobank', *Arteriosclerosis, Thrombosis, and Vascular Biology*, (February), pp. 446–461. doi: 10.1161/ATVBAHA.119.313226.

Sung, Y. J. *et al.* (2016) 'An Empirical Comparison of Joint and Stratified Frameworks for Studying GxE Interactions: Systolic Blood Pressure and Smoking in the CHARGE Gene-Lifestyle Interactions Working Group', *Genetic Epidemiology*, 40(5), pp. 404–415. doi: 10.1002/gepi.21978.

Szilagyi, K. *et al.* (2014) 'Defective signal regulatory protein alpha (sirpa) signaling reduces atherosclerosis in mice', *Atherosclerosis*. Elsevier Ltd, 235(2), p. e91. doi: 10.1016/j.atherosclerosis.2014.05.242.

Tan, X. *et al.* (2017) 'Identification of Key Pathways and Genes in Advanced Coronary Atherosclerosis Using Bioinformatics Analysis', *BioMed Research International*. Hindawi, 2017, pp. 1–12. doi: 10.1155/2017/4323496.

Trillhaase, A. *et al.* (2018) 'Di ff erentiation of human iPSCs into VSMCs and generation of VSMC-derived calcifying vascular cells', *Stem Cell Research*. Elsevier, 31(July), pp. 62–70. doi: 10.1016/j.scr.2018.07.008.

Tuenter, A. *et al.* (2016) 'High shear stress relates to intraplaque haemorrhage in asymptomatic carotid plaques', *Atherosclerosis*. Elsevier Ltd, 251, pp. 348–354. doi: 10.1016/j.atherosclerosis.2016.05.018.

Vuorio, T. *et al.* (2018) 'Downregulation of VEGFR3 signaling alters cardiac lymphatic vessel organization and leads to a higher mortality after acute myocardial infarction', *Scientific Reports*, 8(1), pp. 1–13. doi: 10.1038/s41598-018-34770-4.

Wan, L. *et al.* (2018) 'Screening key genes for abdominal aortic aneurysm based on gene expression omnibus dataset'. BMC Cardiovascular Disorders, pp. 1–13.

Wang, K., Li, M. and Hakonarson, H. (2010) 'ANNOVAR : functional annotation of genetic variants from high-throughput sequencing data', *Nucleic Acids Research*, 38(16), pp. 1–7. doi: 10.1093/nar/gkq603.

Watanabe, K. *et al.* (2017) 'Functional mapping and annotation of genetic associations with FUMA', *Nature Communications*. Springer US, (8), pp. 1–10. doi: 10.1038/s41467-017-01261-5.

Winkler, T. W. *et al.* (2015) 'EasyStrata: Evaluation and visualization of stratified genome-wide association meta-Analysis data', *Bioinformatics*, 31(2), pp. 259–261. doi: 10.1093/bioinformatics/btu621.

Winkler, T. W. *et al.* (2017) 'Approaches to detect genetic effects that differ between two strata in genome-wide meta-analyses: Recommendations based on a systematic evaluation', *PLoS ONE*, 12(7). doi: 10.1371/journal.pone.0181038.

Wojczynski, M. K. *et al.* (2013) 'Genetics of coronary artery calcification among African Americans, a meta-analysis', *BMC Medical Genetics*, 14:75. Available at: http://www.embase.com/search/results?subaction=viewrecord&from=export&id=L52698119%5Cnhttp://www.biomedcentral.com/1471-2350/14/75%5Cnhttp://dx.doi.org/10.1186/1471-2350-14-75%5Cnhttp://sfx.library.uu.nl/utrecht?sid=EMBASE&issn=14712350&id=doi:10.1186%2F1.

Yang, J. *et al.* (2019) 'The emerging role of sorting nexins in cardiovascular diseases', *Clinical Science*, 133(5), pp. 723–737. doi: 10.1042/cs20190034.

Yu, X. and Kem, D. C. (2010) 'Proteasome inhibition during myocardial infarction', *Cardiovascular Research*, 85(2), pp. 312–320. doi: 10.1093/cvr/cvp309.

Zaina, S. *et al.* (2014) 'DNA methylation map of human atherosclerosis', *Circulation: Cardiovascular Genetics*, 7(5), pp. 692–700. doi: 10.1161/CIRCGENETICS.113.000441.

Zhang, Y. and Stefanovic, B. (2016) 'LARP6 meets collagen mRNA: Specific regulation of type I collagen expression', *International Journal of Molecular Sciences*, 17(3). doi: 10.3390/ijms17030419.

Tables

Table 1: Selected SNPs (p<1E-06) associated with Mean Max cIMT for the combined AWI-Gen dataset (n=7894)

rsID	chr	pos	non effect allele	effect allele	MAF	gwasP	Beta	SE	IndSigSNP	Nearest Gene	Genic position
rs12048810	1	22690547	A	G	0,083	6,30E-07	-0,015	0,003	rs12048810	*RP11-415K20.1*	intergenic
rs547840497	5	107469179	C	T	0,014	2,90E-07	-0,041	0,008	rs552690895	*FBXL17*	intronic
rs552690895	**5**	**107570359**	**A**	**G**	**0,013**	**2,50E-08**	**-0,043**	**0,008**	**rs552690895**	***FBXL17***	**intronic**
rs116216559	8	129492156	A	G	0,026	7,70E-07	-0,026	0,005	rs116216559	*RP11-89M16.1*	ncRNA_intronic
rs28731472	8	138092782	C	A	0,090	3,90E-07	-0,016	0,003	rs11781274	*RNU6-144P*	intergenic
rs11781274	8	138097368	A	G	0,082	1,80E-07	-0,017	0,003	rs11781274	*RNU6-144P*	intergenic
rs1324201	9	2183330	C	T	0,256	8,60E-07	-0,010	0,002	rs1324201	*SMARCA2*	intronic
rs11193156	10	108759228	A	T	0,134	2,10E-07	-0,016	0,003	rs11193156	*SORCS1*	intronic
rs10444192	10	126714166	T	C	0,100	7,50E-07	-0,017	0,003	rs10444192	*CTBP2*	intronic
rs3781409	10	126715629	T	C	0,100	6,60E-07	-0,017	0,003	rs3781409	*CTBP2*	exonic
rs11214599	11	113271360	T	C	0,020	5,40E-07	-0,039	0,008	rs11214599	*ANKK1*	downstream
rs2511520	11	113274771	T	C	0,020	5,90E-07	-0,039	0,008	rs11214599	*RP11-159N11.3*	intergenic
rs2242591	11	113279921	T	C	0,020	8,10E-07	-0,039	0,008	rs11214599	*RP11-159N11.3*	ncRNA_intronic
rs6278	11	113280724	A	C	0,020	8,60E-07	-0,038	0,008	rs11214599	*RP11-159N11.3:DRD2*	ncRNA_intronic
rs1124491	11	113282090	A	G	0,020	6,40E-07	-0,039	0,008	rs11214599	*RP11-159N11.3:DRD2*	ncRNA_intronic
rs685058	18	2146395	G	A	0,143	7,10E-07	0,014	0,003	rs685058	*RP11-161I6.2*	ncRNA_intronic
rs6045318	**20**	**1883451**	**G**	**A**	**0,024**	**4,70E-08**	**-0,031**	**0,006**	**rs6045318**	***SIRPA***	**intronic**

chr: Chromosome; pos: position; MAF: Minor allele frequency of effect allele; gwasP: GWAS p-value; SE: Standard Error; IndSigSNP: Index significant SNP.

Table 2: Selected SNPs (p<1E-06) for the Female-stratified analysis

sID	chr	pos	Non effect allele	Effect allele	MAF	gwasP	beta	se	IndSigSNP	Nearest Gene	Genic position
)672087	1	106514615	G	A	0,036	8,20E-07	0,053	0,011	rs116672087	*RP11-24P14.1*	ncRNA_intronic
)9398	1	111006986	G	A	0,154	8,90E-07	-0,018	0,004	rs4839398	*PROK1:RP11-470L19.5*	intergenic
)473055	**1**	**111020548**	**T**	**C**	**0,080**	**1,00E-08**	**-0,026**	**0,005**	**rs115473055**	***CYMP***	**intergenic**
)337389	1	111033882	A	G	0,061	9,50E-08	-0,037	0,007	rs140337389	*CYMP*	ncRNA_exonic
)454995	2	32062127	C	A	0,014	3,20E-07	-0,060	0,012	rs139454995	*AK2P2*	intergenic
)0575	3	29486586	G	A	0,026	5,80E-07	-0,041	0,008	rs9830575	*RBMS3*	intronic
)13301	3	29569688	G	C	0,022	8,40E-07	-0,040	0,008	rs9830575	*RBMS3*	intronic
)28487	4	54487185	C	T	0,022	7,10E-08	-0,050	0,009	rs79028487	*FIP1L1:LNX1*	intronic
)752200	4	54488817	G	A	0,022	6,50E-08	-0,051	0,010	rs145752200	*FIP1L1:LNX1*	intronic
)178248	4	54495071	A	T	0,022	9,40E-08	-0,050	0,010	rs139178248	*FIP1L1:LNX1*	intronic
)15295	4	54508985	G	C	0,023	1,60E-07	-0,049	0,010	rs78315295	*FIP1L1:LNX1*	intronic
)055133	4	54511431	T	C	0,022	4,50E-07	-0,048	0,010	rs184055133	*FIP1L1:LNX1*	intronic
)744905	5	20536559	C	T	0,012	7,50E-07	-0,049	0,010	rs190082809	*CDH18*	intronic
)793722	5	20542546	T	C	0,012	4,50E-07	-0,052	0,010	rs190082809	*CDH18*	intronic
)849737	5	20544251	G	C	0,012	4,10E-07	-0,052	0,010	rs190082809	*CDH18*	intronic
)340920	5	20549408	G	A	0,012	5,20E-07	-0,051	0,010	rs190082809	*CDH18*	intronic
)082809	5	20552942	A	G	0,013	2,80E-07	-0,053	0,010	rs190082809	*CDH18*	intronic
)59032	5	165602212	T	C	0,033	3,30E-07	-0,048	0,009	rs11959032	*AC122720.1*	intergenic
)35228	5	165603117	A	C	0,033	9,00E-07	-0,047	0,010	rs11959032	*AC122720.1*	intergenic
)437926	5	165614455	A	T	0,036	6,50E-07	-0,046	0,009	rs11959032	*AC122720.1*	intergenic
)967731	5	180083479	A	G	0,090	5,70E-07	-0,023	0,005	rs112967731	*FLT4*	intergenic
)285788	7	4779605	T	A	0,016	3,40E-07	0,056	0,011	rs140285788	*FOXK1*	intronic
3153	8	95092061	C	A	0,165	6,50E-07	-0,018	0,004	rs114684581	*KB-1184D12.1*	intergenic
)684581	8	95100302	C	T	0,166	3,40E-07	-0,018	0,004	rs114684581	*KB-1184D12.1*	intergenic
)79158	8	95174434	T	C	0,073	6,90E-07	-0,024	0,005	rs80279158	*CDH17*	intronic
)60551	9	12260408	G	C	0,274	1,20E-07	0,015	0,003	rs10960551	*RP11-71E22.1*	intergenic
)0521	9	12261388	A	G	0,274	2,70E-07	0,015	0,003	rs10960551	*RP11-71E22.1*	intergenic
6577	9	12264458	C	G	0,291	9,50E-07	0,014	0,003	rs10960551	*RP11-71E22.1*	intergenic
)44940	9	93865035	A	G	0,023	7,40E-07	-0,051	0,010	rs74744940	*RP11-305L7.1*	intergenic
5324	11	115446096	A	G	0,487	6,90E-07	-0,013	0,003	rs922205	*RP11-136I14.4*	intergenic
5326	11	115446460	A	T	0,487	7,20E-07	-0,013	0,003	rs922205	*RP11-136I14.4*	intergenic
)6890	11	115447446	G	A	0,488	8,10E-07	-0,013	0,003	rs922205	*RP11-136I14.4*	intergenic
)205	11	115448115	A	T	0,440	4,60E-07	-0,013	0,003	rs922205	*RP11-136I14.4*	upstream
)840489	**15**	**71088277**	**A**	**G**	**0,037**	**2,40E-09**	**-0,051**	**0,009**	**rs150840489**	***RPL29P30***	**downstream**

sID	chr	pos	Non effect allele	Effect allele	MAF	gwasP	beta	se	IndSigSNP	Nearest Gene	Genic position
98235	15	80852328	T	G	0,030	6,40E-08	-0,044	0,008	rs77298235	*ARNT2:RP11-379K22.2*	ncRNA_intronic
17996	15	80852338	C	T	0,064	5,30E-07	-0,032	0,006	rs76017996	*ARNT2:RP11-379K22.2*	ncRNA_intronic
882243	15	80882112	T	C	0,019	8,40E-07	-0,046	0,009	rs141882243	*ARNT2*	intronic
132984	16	15690120	C	T	0,088	3,80E-07	-0,026	0,005	rs116132984	*C16orf45:KIAA0430*	UTR3

chr: Chromosome; pos: position; MAF: Minor allele frequency of effect allele; gwasP: GWAS p-value; SE: Standard Error; IndSigSNP: Index significant SNP;

Table 3: Selected SNPs (p<1E-06) for the Male-stratified analysis

rsID	chr	pos	Non effect allele	Effect allele	MAF	gwasP	beta	se	IndSigSNP	Nearest Gene	Genic position
114119990	1	163865550	C	A	0,043	8,20E-07	0,032	0,007	rs114119990	*U3*	intergenic
67416193	1	222666445	C	G	0,322	8,20E-07	0,015	0,003	rs67416193	*RNU6-791P*	intergenic
146962840	1	248797787	G	C	0,047	8,00E-07	-0,049	0,010	rs146962840	*OR2T35*	intergenic
145287839	4	169571145	G	T	0,025	9,20E-07	-0,052	0,011	rs145287839	*PALLD*	intronic
547840497	5	107469179	C	T	0,014	1,80E-07	-0,062	0,012	rs547840497	*FBXL17*	intronic
806282	6	91197382	T	G	0,264	4,80E-07	0,015	0,003	rs284509	*MAP3K7*	intergenic
284509	**6**	**91208376**	**G**	**T**	**0,231**	**5,30E-08**	**0,017**	**0,003**	**rs284509**	***MAP3K7***	**intergenic**
284511	6	91208542	C	T	0,237	6,80E-08	0,017	0,003	rs284509	*MAP3K7*	intergenic
182259	6	91209111	T	C	0,231	1,70E-07	0,016	0,003	rs284509	*MAP3K7*	intergenic
11193156	10	108759228	A	T	0,134	1,30E-07	-0,024	0,005	rs11193156	*SORCS1*	intronic
112120989	13	27636336	T	C	0,023	1,10E-07	-0,041	0,008	rs56940748	*USP12*	intergenic
542629237	13	27661417	C	T	0,030	3,50E-07	-0,038	0,007	rs56940748	*USP12*	intronic
56940748	13	27662100	A	G	0,023	8,40E-08	-0,045	0,008	rs56940748	*USP12*	intronic
73497361	13	27663613	C	T	0,023	8,70E-08	-0,045	0,008	rs56940748	*USP12*	intronic
190770959	**16**	**12158574**	**T**	**C**	**0,015**	**6,30E-09**	**-0,056**	**0,010**	**rs190770959**	***SNX29***	**intronic**
147978408	**16**	**12171475**	**C**	**T**	**0,016**	**6,60E-09**	**-0,055**	**0,010**	**rs190770959**	***SNX29***	**intronic**
4239212	17	26309966	G	A	0,092	3,70E-07	-0,025	0,005	rs4239212	*RP11-218F4.1*	intergenic
6045318	20	1883451	G	A	0,024	7,10E-07	-0,040	0,008	rs6045318	*SIRPA*	intronic

chr: Chromosome; pos: position; MAF: Minor allele frequency of effect allele; gwasP: GWAS p-value; SE: Standard Error; IndSigSNP: Index significant SNP

Table 4: Selected SNPs (p<1E-06) for the sex-difference test

P	CHR	BP	A1	A0	A1FREQ W	BETA W	SE W	P W	A1FREQ M	BETA M	SE MEN	P M	p Sex diff	Genic position	Nearest Gene
0396	9	38487113	T	C	0,860	0,009	0,004	1,30E-02	0,852	-0,018	0,004	8,20E-06	3,46E-07	intergenic	*IGFBPL1;FAM95C*
40489	15	71088277	G	A	0,975	-0,051	0,009	2,40E-09	0,977	0,012	0,009	1,70E-01	4,38E-07	intergenic	*UACA;LARP6*
5710	1	17276223	C	T	0,617	0,007	0,003	4,90E-03	0,612	-0,012	0,003	2,70E-05	4,84E-07	intronic	*CROCC*
6910	1	17276357	T	C	0,617	0,007	0,003	4,90E-03	0,612	-0,012	0,003	2,70E-05	4,84E-07	intronic	*CROCC*
81830	11	17754454	T	C	0,458	0,012	0,003	7,70E-05	0,469	-0,009	0,003	3,30E-03	4,90E-07	intergenic	*MYOD1;KCNC1*

SNP: Single Nucleotide Polymorphism; CHR: chromosome; A1: alternative allele; A0: reference Allele; A1FREQW: alternative allele frequency women; A1FREQM: alternative allele frequency men; PW: p-value women; PM: p-value men; BETA W: beta women; BETA W: beta men; SE W: standard error women; SE M: standard error men; p Sex diff: p-value sex-difference

Figures

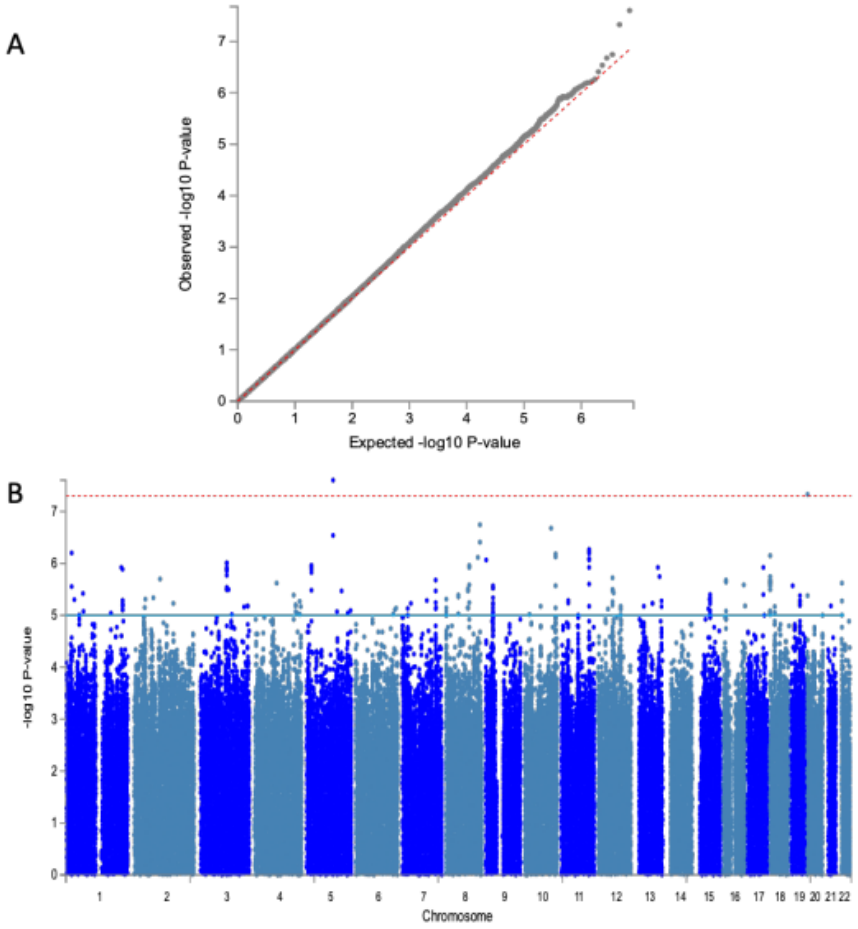

Figure 1

QQ and Manhattan plots for cIMT association results in AWI-Gen study (7894 participants).

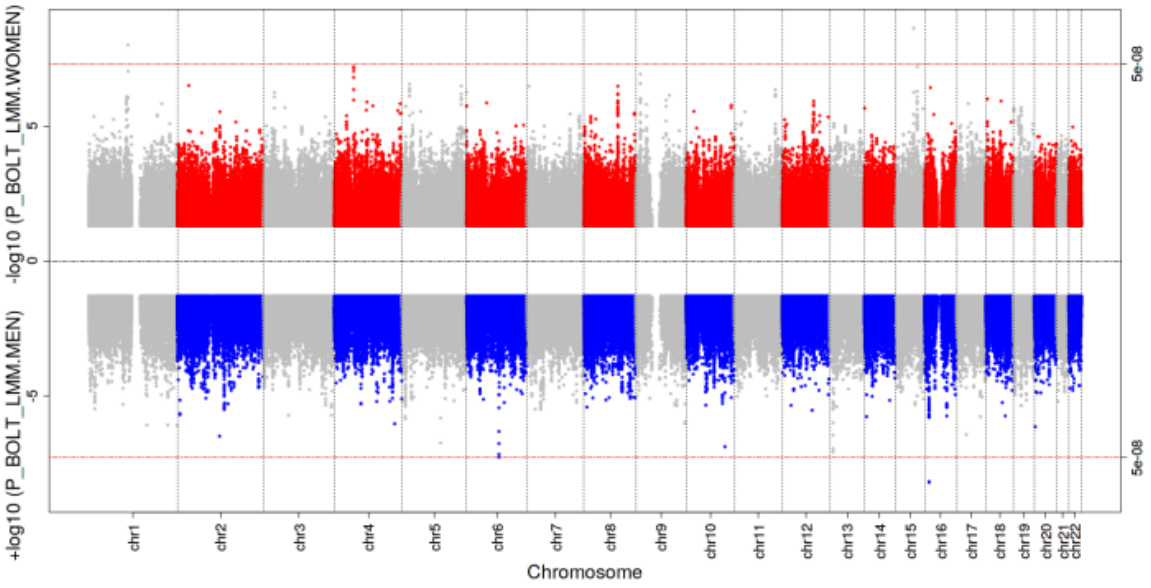

Figure 2

Miami plot showing female and male-specific associated p-values for mean max cIMT.

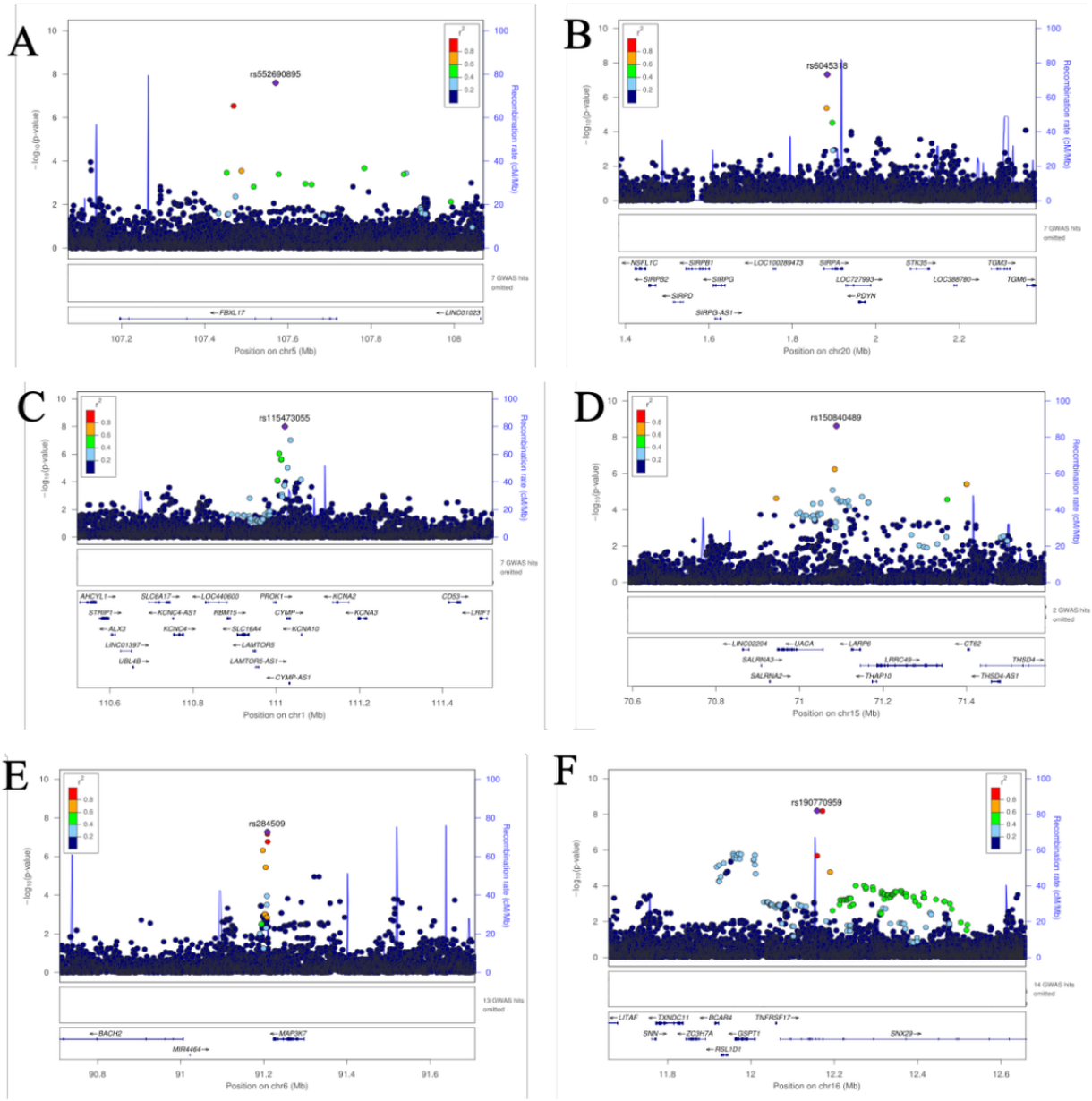

Figure 3

Regional association plots for the selected top SNPs showing genetic associations with meam max cIMT.

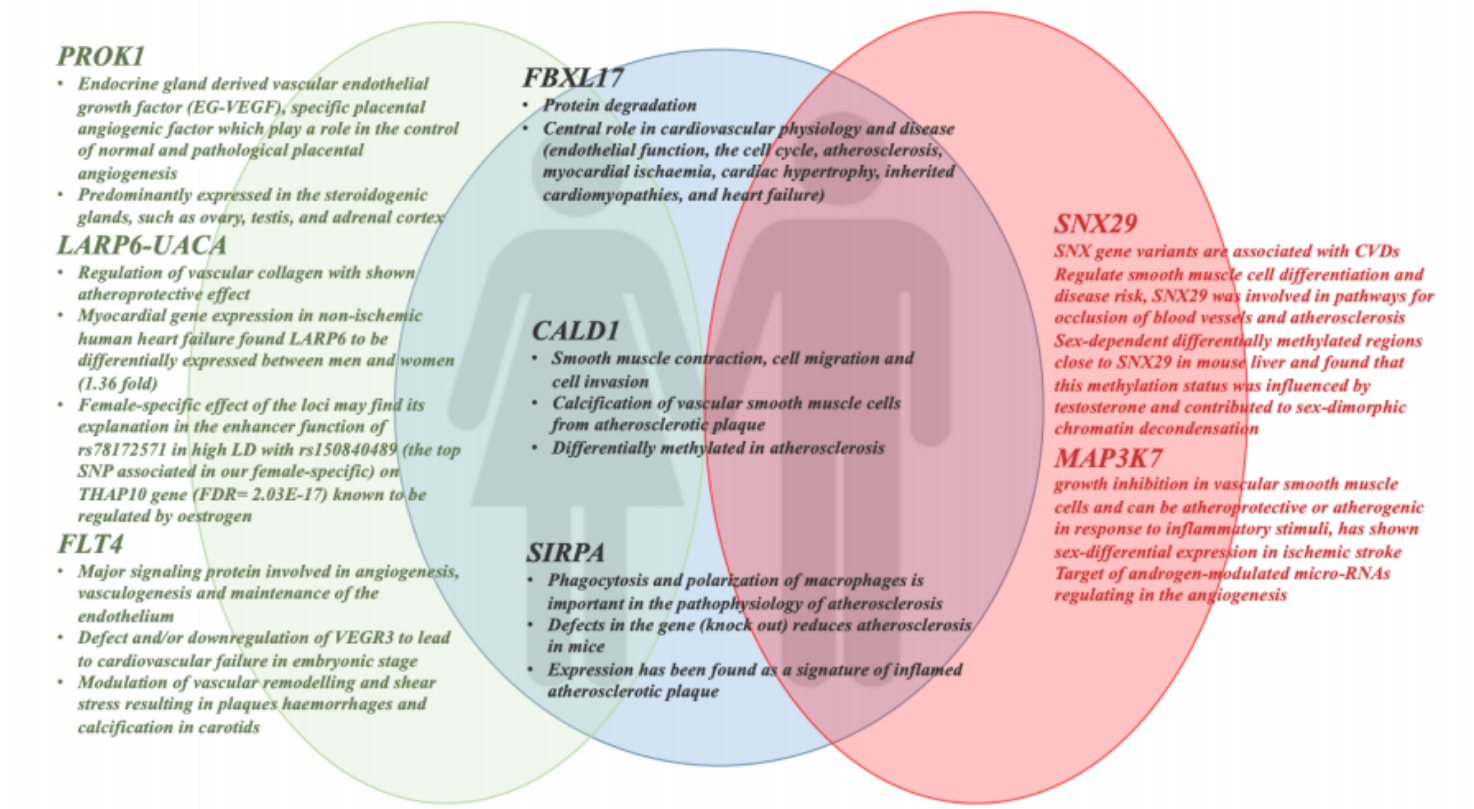

Figure 4

Biological relevance of identified loci

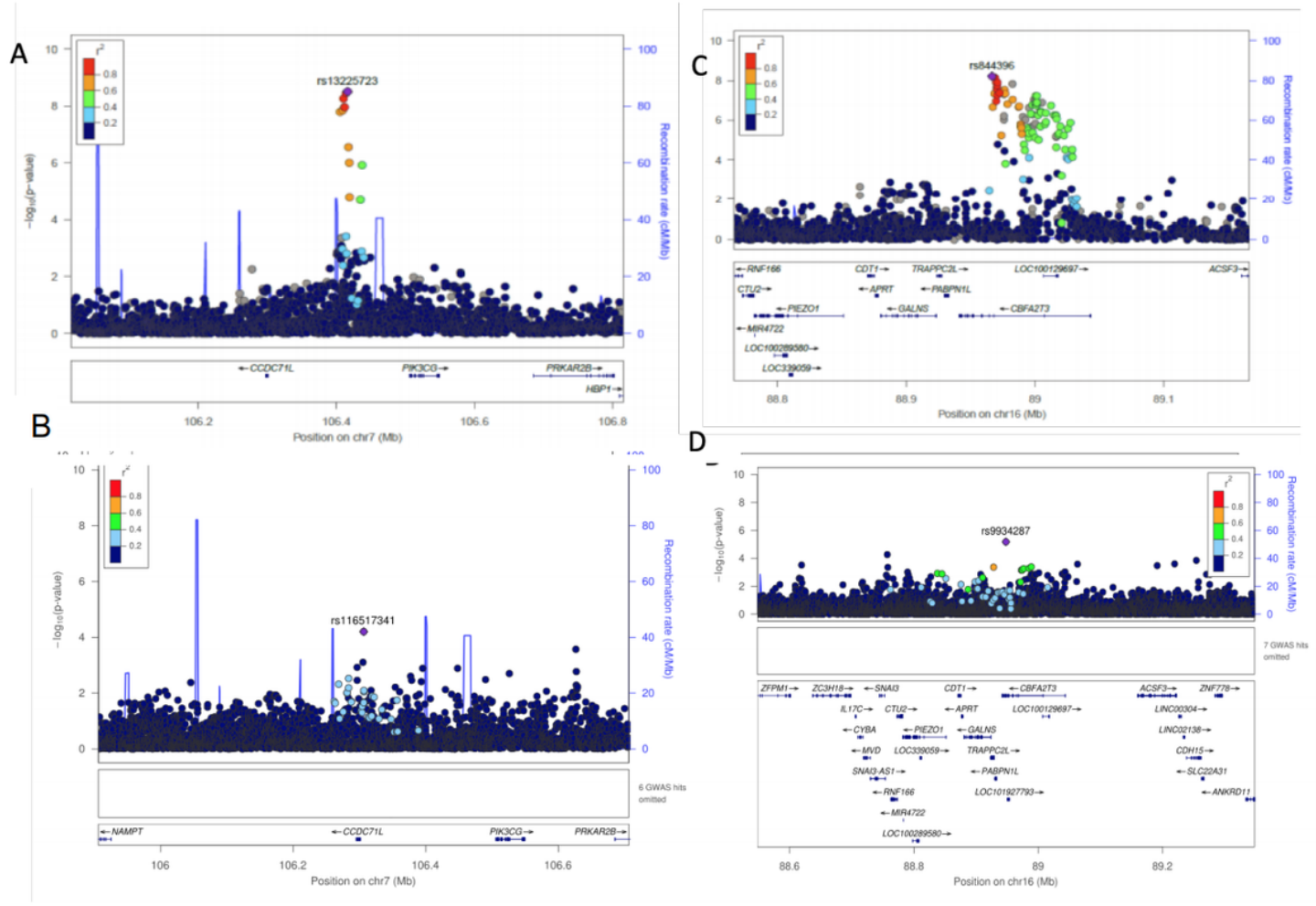

Figure 5

Regional plots for genetic associations with cIMT, focusing on signals close to two genes (CCDC71L and CBFA2T3) previously identified in European GWASs and replicated in our study.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- BouaGWAScIMTSuppTextFigures.pdf
- SuppTable123.xlsx
- SuppTable4aFunctionalAnnotationAll.xlsx
- SuppTable4bFunctionalAnnotationFemale.xlsx
- SuppTable4cFunctionalAnnotationMale.xlsx
- SuppTable5aGeneMapAll.xlsx
- SuppTable5bGeneMapFemale.xlsx
- SuppTable5cGeneMapMale.xlsx
- SuppTable6aGeneSetAll.xlsx
- SuppTable6bGeneSetFemale.xlsx
- SuppTable6cGeneSetMale.xlsx