CXCL12 regulates coronary artery dominance in diverse populations and links development to disease

Pamela E Rios Coronado Ph.D[1], Daniela Zanetti Ph.D.[2,3,4], Jiayan Zhou Ph.D.[2,3], Jeffrey A Naftaly B.A.[1], Pratima Prabala B.S.[1], Pik Fang Kho Ph.D.[2,3], Azalia M Martínez Jaimes B.S. & B.A.[5,1], Austin T Hilliard Ph.D.[6], Saiju Pyarajan Ph.D.[7], Daniel Dochtermann[7], Million Veteran Program[8], Kyong-Mi Chang M.D.[9,10], Virginia D Winn MD, PhD[11], Anca M Pașca MD[12], Mary E Plomondon Ph.D.[13,14], Stephen W Waldo M.D.[13,14,15], Philip S Tsao Ph.D.[6,16,17], Shoa L. Clarke MD, PhD[18,3,2], Kristy Red-Horse Ph.D[1,19,20], Themistocles L. Assimes MD, PhD[2,3,21,17]

[1]Department of Biology, Stanford University, Stanford, CA, USA, [2]Department of Medicine, Division of Cardiovascular Medicine, Stanford University School of Medicine, Stanford, CA, USA, [3]VA Palo Alto Healthcare System, Palo Alto, CA, USA, [4]Institute of Genetic and Biomedical Research, National Research Council, Cagliari, Sardinia, Italy, [5]Department of Developmental Biology, Stanford University School of Medicine, Stanford, CA, USA, [6]VA Palo Alto Health Care System, Palo Alto, CA, USA, [7]Center for Data and Computational Sciences, VA Boston Healthcare System, Boston, MA, USA, [8]A list of consortium members can be found in the Supplementary File, [9]Corporal Michael J. Crescenz VA Medical Center, Philadelphia, PA, USA, [10]Department of Medicine, Division of Gastroenterology and Hepatology, University of Pennsylvania Perelman School of Medicine, Philadelphia, PA, USA, [11]Department of Obstetrics and Gynecology, Stanford University School of Medicine, Stanford, CA, USA, [12]Department of Pediatrics, Neonatology, Stanford University School of Medicine, Stanford, CA, USA, [13]Department of Medicine, Rocky Mountain Regional VA Medical Center, Aurora, CO, USA, [14]CART Program, VHA Office of Quality and Patient Safety, Washington, DC, USA, [15]Division of Cardiology, University of Colorado School of Medicine, Aurora, CO, USA, [16]Department of Medicine, Stanford University School of Medicine, Stanford, CA, USA, [17]Cardiovascular Institute, Stanford University School of Medicine, Stanford, CA, USA, [18]Department of Medicine, Stanford Prevention Research Center, Stanford University School of Medicine, Stanford, CA, USA, [19]Institute for Stem Cell Biology and Regenerative Medicine, Stanford University School of Medicine, Stanford, CA, USA, [20]Howard Hughes Medical Institute, Chevy Chase, MD, USA, [21]Department of Epidemiology and Population Health, Stanford University School of Medicine, Stanford, CA, USA

Pamela E Rios Coronado, Daniela Zanetti, and Jiayan Zhou contributed as co-lead junior authors.
Shoa L. Clarke, Kristy Red-Horse, and Themistocles L. Assimes contributed as co-supervising senior authors.

Corresponding Authors:
Kristy Red-Horse, PhD
Department of Biology, Stanford University, Stanford, CA 94305
Howard Hughes Medical Institute, Stanford, CA
E-mail: kredhors@stanford.edu

Themistocles L. Assimes, MD PhD FAHA
Department of Medicine, Stanford University School of Medicine, Stanford, CA 94305
Palo Alto VA Health Care System, Palo Alto, CA 94304
E-mail: tassimes@stanford.edu

Abstract

Mammalian cardiac muscle is supplied with blood by right and left coronary arteries that form branches covering both ventricles of the heart. Whether branches of the right or left coronary arteries wrap around to the inferior side of the left ventricle is variable in humans and termed right or left dominance. Coronary dominance is likely a heritable trait, but its genetic architecture has never been explored. Here, we present the first large-scale multi-ancestry genome-wide association study of dominance in 61,043 participants of the VA Million Veteran Program, including over 10,300 Africans and 4,400 Admixed Americans. Dominance was moderately heritable with ten loci reaching genome wide significance. The most significant mapped to the chemokine *CXCL12* in both Europeans and Africans. Whole-organ imaging of human fetal hearts revealed that dominance is established during development in locations where *CXCL12* is expressed. In mice, dominance involved the septal coronary artery, and its patterning was altered with Cxcl12 deficiency. Finally, we linked human dominance patterns with coronary artery disease through colocalization, genome-wide genetic correlation and Mendelian Randomization analyses. Together, our data supports *CXCL12* as a primary determinant of coronary artery dominance in humans of diverse backgrounds and suggests that developmental patterning of arteries may influence one's susceptibility to ischemic heart disease.

Introduction

Ischemic heart disease is the number one cause of death in both men and women[1], and it occurs when myocardial tissue is chronically or acutely under perfused, leading to angina, myocardial infarction, heart failure, and arrhythmias with a high mortality rate[2]. Perfusion of the heart is provided by a branching system of coronary arteries, and pathologies of these arteries are the driver of ischemic heart disease[2,3]. Discovering the mechanisms guiding coronary artery formation during heart development could identify targetable molecular pathways that repair or regenerate diseased arteries[4-6]. However, a roadblock to progress towards this goal is the paucity of information on how this process occurs in humans[7].

Some variability in human coronary artery branch structure exists, and this variability is likely heritable as are most human traits[8]. Exploring the genetic basis of this variability thus could identify genes regulating coronary development. Most of the major branches of the coronary arterial tree follow a stereotypic pattern. The left main coronary artery arises from the left coronary cusp of the aorta and gives rise to the left anterior descending (LAD) and left circumflex (LCx), while the right coronary artery (RCA) arises from the right coronary cusp. However, the origin of the posterior descending artery (PDA) is variable across humans[2,3] as it can stem from branches of the right or left coronary arteries. In approximately ~80% of the population, the PDA arises from the RCA (right dominant), but in ~10% it arises from the LCx (left dominant)[2,3]. In the remaining ~10%, a more balanced anatomy exists termed mixed or co-dominance2,3. Importantly, the LAD, LCx, and PDA collectively perfuse the entire left ventricle, which is primarily responsible for maintaining the systemic circulation2,3. Several studies have suggested that left dominance is associated with worse outcomes after a myocardial infarction9 or after coronary procedures10, but there is conflicting data on whether it is associated with higher11, lower12, or no difference13 in the risk of cardiovascular events. Furthermore, how, and when dominance is determined and whether it has an impact on the propensity to develop coronary artery disease is not known.

We address these questions through large-scale human genetics analyses coupled with examination of human fetal hearts and *in vivo* mouse experiments. Since the genetic architecture of coronary artery dominance had not previously been explored, we use more than 61,000 invasive coronary angiograms in the Million Veteran Program (MVP) to estimate genetic heritability and conduct the first ever genome-wide association study (GWAS) of coronary dominance. We thus identify

CXCL12, a gene first implicated in coronary artery disease (CAD) over sixteen years ago[14], as a candidate regulator of coronary artery dominance, and we support this hypothesis through *in vivo* mouse experiments. Finally, we use Mendelian randomization to show that coronary artery dominance may impact risk for CAD, thus linking a fetal developmental process to an acquired disease that manifests in mid to late adulthood.

Results

A large-scale, genetically diverse genome-wide association study of coronary artery dominance

We leveraged the MVP biobank study combined with the Veterans Health Administration (VA) Clinical Assessment, Reporting, and Tracking (CART) Program, a national quality and safety organization for invasive cardiac procedures, to conduct a large-scale, multi-ancestry GWAS of dominance (**Figure 1**). Among 658,164 participants with genotype data available imputed to a multi-ancestry TOPMed reference panel, we identified 61,043 participants who had undergone one or more coronary angiogram that included an assessment of dominance (**Supplementary Table 1)**. We used Scalable and Accurate Implementation of GEneralized mixed model (SAIGE)[15] to conduct GWAS in the three largest genetically inferred superpopulations, white/European (EUR), black/African (AFR), and admixed-American (AMR), followed by all subjects combined irrespective of ancestral assignment. We ran two GWAS for each population comparing first left and mixed dominance (co-dominance) combined to right dominance (reference) (RvsLM) and the second comparing left to right (reference) dominance excluding mixed (RvsL) as a form of extreme phenotyping approach to discovery.

The quantile-quantile (QQ) plots of p-values for genetic variants tested for association with dominance in the combined GWAS and in the GWAS of EUR and AFR demonstrated a strong genetic signal with no evidence of population stratification based on the calculated genomic control lambdas for common variants with minor allele frequency >0.01 (**Figure 2A**). These observations were bolstered by a point estimate of total SNP-based narrow-sense heritability (h^2) of 0.277 (95% confidence intervals, 0.150-0.404) among EUR using a linkage disequilibrium (LD) and MAF stratified multicomponent approach implemented in GCTA, GREML-LDMS-I, and assuming the observed prevalence of non-right dominance of 17% (**Supplementary Table 2, Supplementary Figure 1A,B**). A large fraction of the additive genetic variance is derived from low frequency and rare variants (**Supplementary Figure 1C**).

The genetic signal was weak to non-existent among AMR, likely a reflection of restricted power of discovery given the small number of subjects with non-right dominance in this subgroup.

Ten loci, as defined by an algorithm implemented in FUMA-GWAS[16], reached GWS in one or more of the GWAS performed (**Table 1, Figure 2, Supplementary Table 3 and Figure 2**). Of these, seven involved common variants observed in two or more of our GWAS. Two additional loci were defined by either one or a small number of rare variants in high LD with MAF<0.005, and one locus on chromosome three was defined by multiple low frequency variants reaching genome wide significance only in AMR despite higher frequencies and better power to implicate the same region in the other two ancestral groups and in the combined analysis. Furthermore, the saddlepoint approximation (SPA) for p-values of SNPs found in this region for AMR did not converge in the other populations.

The most significant locus with the highest effect size was located just downstream of *CXCL12* which, remarkably, reached GWS in both EUR and AFR (**Figure 2B**). Regional association and LD plots for GWS loci are included in **Supplementary Figure 3**. Statistical fine mapping with SusieX for the seven loci involving common variants observed in multiple GWAS identified a single credible set of causal SNPs within each of five loci--two credible sets within the *NKX2-5* locus on chromosome 5 and four credible sets within the *CXCL12* locus on chromosome 10. **Supplementary Table 4** lists the SNPs within each credible set with the highest posterior probability for being independently causal. **Supplementary Tables 5** provides full SAIGE output for all lead and independently significant variants identified by FUMA, and all SUSIEx lead genetic variants. **Supplementary Table 6** is a list of mapped genes using GTex v8 and chromatin folding maps as implemented in FUMA. Lastly, **Supplementary Table 7** provides annotation hyperlinks for the lead SNPs in Table 1 to five online variant-based portals.

Lead genetic variants in both European and African Americans are linked to Cxcl12 expression

Although any, or all, novel loci identified could causally impact coronary dominance, we focused on *CXCL12* due to its GWS in two genetically distant populations and previous studies demonstrating a role during coronary artery development in model organisms[17-20].

Multiple genetic analyses suggested that dominance-associated variants downstream of *CXCL12* influence gene expression levels, rather than protein function. First, a subset of significant intergenic variants in the locus are predicted to be functionally important by CADD score, a tool for scoring the deleteriousness of single nucleotide variants (SNPs) as well as insertion/deletion variants in the human genome[21,22], in both EUR and AFR (**Figure 3A,E**). Among EUR, several can be identified as expression

quantitative trait loci (eQTL) for up to six genes within the locus, including *CXCL12* (**Figure 3A**), and in several tissues, including ovary, testis, atrial appendage, and tibial artery (**Figure 3B**). Importantly, eQTLs observed among AFR implicate only one gene, *CXCL12*, and further identify the coronary artery as a relevant tissue (**Figure 3E-F**). Assuming the arterial tissues are most pertinent mechanistically, our association results combined with the GTEx eQTL data suggest the allele associated with more *CXCL12* expression is more often observed with right dominance (**Figure 3B,C,F,G**). Lastly, databases utilizing HiC data revealed that dominance-associated eQTLs interacted with the *CXC12* and *TMEM72* promoters among EUR, but only with *CXCL12* among AFR (**Figure 3D,H**). Thus, *CXCL12* is the strongest candidate for a causal gene regulating coronary artery patterning.

Coronary dominance is apparent at fetal stages when CXCL12 is expressed

When dominance is determined in unknown; we hypothesized it would be established during heart development with the formation of the coronary arteries. Since angiograms and non-invasive imaging procedures are not performed during healthy pregnancies, we investigated whether dominance is presence during fetal stages by performing whole-organ immunofluorescence that allowed visualization of the entire artery tree in 3D (**Figure 4A and B**) [23,24]. Tracing primary and secondary branches of the right and left coronary arteries of each heart revealed a pattern reminiscent of the adult anatomy. Left and right primary branches stemmed from the aorta and took a course typical of the LAD, LCx, and RCA in all hearts (**Supplementary Figure 4A-C**).

As in adults, there was variability in the extent to which the right and left branches covered the posterior and inferior aspects of the ventricles at gestational weeks (GW) 14-22. 6 of the 8 hearts (75%) were clearly right dominant in that the interventricular septum had a branch from the right where the adult posterior descending artery (PDA) would be located, and additional right branches extended onto the inferior left ventricle (**Figure 4C**). We also traced a number of lower order arteries and arterioles (**Figure 4D**) and found that the septum was perfused by branches from the right coronary on the inferior side and the left anterior descending on the anterior side (**Figure 4E**). 2 of 8 hearts (25%) had a different coverage that we scored as co-dominance because both right and left arteries branched into the inferior wall and septum (**Figure 4F-H** and **Supplementary Figure 4D-F**). For example, one heart showed an artery branching onto the interventricular groove from the right, but it only extended halfway down while the rest of the inferior aspect contained branches from the left circumflex and left anterior descending (**Figure 4F**). This resulted in the inferior septum being irrigated by both right and left

(**Figure 4H**). The observation that dominance can be observed as early as GW 14 and is at ratios reminiscent of those in adults (**Figure 4I**), supports a model where it is established during fetal development.

If dominance is established during fetal development and impacted by *CXCL12*, this chemokine should be expressed near developing coronary arteries during this time. Fluorescence *in situ* hybridization on fetal heart ventricles from GWs 13, 18, and 20 showed expression in coronary arteries, and, at the earliest stage, trabecular myocardium (**Figure 4J** and **Supplementary Figure 4G**). Thus, *CXCL12* is expressed at the right time and place to influence human coronary artery dominance.

CXCL12 levels regulate septal artery dominance in mice

To provide evidence that *CXC12* impacts major anatomical variation in the coronary tree, we investigated coronary artery patterning in mice expressing different levels of *Cxcl12* (**Figure 1***). Whole mouse hearts were subjected to a whole-organ immunofluorescence[23,24] (**Figure 5A**). Artery tracing using Imaris software and measuring coronary artery branching revealed a lack of variance at the inferior wall in both CD1 and C57BL/6J strains. All hearts were left dominant in this location (**Supplemental Figure 5A and B)**. Thus, coronary artery dominance is not identical between human and mice.

We next explored whether coronary branching varies in other regions of the heart and found that the mouse septal artery (SpA) displayed features of dominance. Unlike the human heart, the mouse ventricular septum does not rely on end branches of the left and right coronary arteries but instead receives blood flow from a dedicated SpA (**Figure 5B**)[25-28]. Vascular filling experiments have demonstrated that the SpA has been reported to display variation in where it originates. It can arise from the proximal branch of either the left, right, or both coronary arteries; in rare instances, it originates directly from the aorta[26,27]. We confirmed this using our imaging and tracing methods on neonatal mouse hearts (**Figure 5C**), and, consistent with genetic regulation of SpA patterning, different mouse strains displayed different origin ratios (**Figure 5D**). Ratios were similar in neonatal and adult hearts, supporting the notion that septal dominance is established early in life, similar to human observations (**Supplementary Figure 5C**)[26,27]. Therefore, coronary artery dominance in mice manifests differently than in humans, and we refer to it as SpA dominance.

We next utilized these SpA branching variations to investigate CA dominance in *Cxcl12* haploinsufficient mice. In the *Cxcl12*^{DsRed} mouse strain, the *Cxcl12* coding region is replaced with a red

fluorescent protein, DsRed, so that it is both a deletion line and a gene expression reporter[29]. Wild type (*Cxcl12*$^{+/+}$) and heterozygous (*Cxcl12*$^{DsRed/+}$) littermates were compared rather than using homozygous knockouts for two reasons: (1) We wanted to better model the likely effects of dominance-associated variants, which are in non-coding regions predicted to modify gene expression levels (**see Figure 3**), and (2) Homozygous knockout animals are lethal between embryonic days 15-18[19]. QPCR assessing *Cxcl12* mRNA levels in hearts from *Cxcl12*$^{+/+}$, *Cxcl12*$^{DsRed/+}$, and *Cxcl12*$^{DsRed/DsRed}$ mice demonstrated that expression in heterozygous mice is approximately 50% lower than wild type (**Figure 5E**), making them a good model for assessing SpA dominance in hearts with attenuated *Cxcl12* expression. Because the SpA connection was in a prominent proximal location, we could also assess dominance using an *in vivo* labeling method that labeled all coronary vasculature with fluorescently tagged Isolectin GS-IB$_4$ (**Supplementary Figure 5D-F**). This facilitated analyses of 126 neonatal mouse hearts obtained from crosses between *Cxcl12*$^{DsRed/+}$ male mice and wild type C57Bl/6J females. Analyzing SpA dominance showed that *Cxcl12*$^{DsRed/+}$ mice had shifted SpA dominance ratios when compared with wild type littermates (**Figure 5F**). They had a lower incidence of right dominance and a higher incidence of co-dominance and aorta, but there was no change in the percentage of left dominance. Therefore, *Cxcl12* heterozygosity changed mouse SpA dominance patterns. These data demonstrate that *in vivo* mouse experiments can aid in the validation of human susceptibility loci related to coronary artery patterning phenotypes.

To begin understanding how *Cxcl12* levels might influence SpA development, we investigated when SpA dominance was established and where *Cxcl12* was expressed during the process. Immunolabeling embryonic hearts with the artery endothelial cell marker Jagged1 revealed that positive vessels appeared in the septum around embryonic days (E)14.5 and 15.5 as a network of interconnected, immature artery precursors that were attached to both the right and left main branches (**Figure 5G and Supplementary Figure 6B** and **C**)[30]. Over the next two days, remodeling resulted in progressive development of an artery within the septum, and, here, the majority were located close to the right ventricular lumen (**Supplementary Figure 6D and E)**. Related to dominance, we frequently observed that one of the original connections, either to the left or right coronary artery, had grown larger than the other side at these time points (**Supplementary Figure 6D)**. At E17.5 (approximately one day before birth), we began to observe SpAs that had disconnected from one side leading to the appearance of a mature form that suggested left, right, or co-dominance[30](**Figure 5H** and **Supplementary Figure 6E**). This developmental trajectory occurred in an area with abundant *Cxcl12* expression, including in the

outflow tract, right ventricle trabecular myocardium, and in the artery endothelial cells themselves (**Figure 5I**). Together, these data are consistent with SpA dominance being regulated by Cxcl12 during embryonic development.

Colocalization, genetic correlation, and Mendelian Randomization analyses with coronary artery disease

The *CXCL12* region identified in our dominance GWAS highly overlaps with the widely replicated signal reported with the initial GWASs for CAD in 2007[14]. We performed colocalization analysis[31] to identify potentially shared dominance and CAD causal variants at this locus. We then quantified the similarity in the genetic architecture of dominance and CAD through genome-wide genetic correlation analyses[32] within and external to MVP and assessed causality between dominance and CAD using a two-sample Mendelian randomization (MR) approach.

Colocalization analyses supported the existence of a shared causal variant within the *CXCL12* locus for coronary dominance and CAD[33] (**Figure 6A**). We also detected a modest negative genome-wide genetic correlation (r_g) using LD score regression between coronary dominance and CAD within MVP (r_g=-0.193, p=0.004) as well as CAD from the CardiogramplusC4D data[34] (r_g=-0.155, p=0.02). To determine if the genetic correlation between dominance and CAD is driven solely by the *CXCL12* locus, we next calculated r_g after excluding a 900 kb region in chromosome 10 (from 43910000 to 44810000 base pairs) that encompasses the *CXCL12* gene signal. Even with exclusion of the *CXCL12* region, we observed similar genetic correlations (r_g=-0.180, p=0.003 and r_g=-0.117, p=0.06, respectively). These findings suggest that coronary dominance and CAD share a modest overlap in genetic architecture beyond *CXCL12*. We tested the hypothesis that dominance influences CAD susceptibility using two-sample Mendelian randomization. Genetic instrument variables for the exposure of non-right dominance were derived from our EUR GWAS of right versus left and right versus left+mixed, which provided 6 independent genome-wide significant SNPs, including 2 at the *CXCL12* locus and 4 at other loci. For the CAD outcome, we used summary statistics from the largest available CAD GWAS from a non-overlapping sample of predominantly European ancestry subjects[34]. We observed evidence for a causal association between non-right dominance and protection against CAD (**Figure 6B-D, Supplementary Tables 8-9**). This finding was consistent in sensitivity analyses using multiple methods for Mendelian randomization (**Figure 6C**), and we did not detect evidence of horizontal pleiotropy or influential

instruments. Leave-one-out analysis shows that this causal relationship is not driven by any single SNP (**Figure 6D**).

Discussion

Here, we present a combined and stratified multi-ancestry GWAS of coronary artery dominance using genotypes from MVP participants who had undergone angiograms, including large numbers of participants of European, African, or Admixed American descent. Ten loci reached GWS with the most statistically significant SNPs located downstream of *CXCL12*. Whole-organ immunofluorescence and spatial transcriptomics showed that dominance is apparent in humans during fetal development when *CXCL12* is expressed in the heart. Evaluation of heterozygous *Cxcl12* knockout mice indicated that *Cxcl12* expression likely influences left versus right patterning during development of the major coronary vessels. Taken together, these data support the notion that *CXCL12* expression directly influences the patterning of human coronary dominance during development. Finally, given that the *CXCL12* locus has been previously implicated in CAD, we explored the genetic relationship between dominance and CAD in humans, finding evidence for shared genetic architecture even outside of the *CXCL12* locus. Mendelian randomization further implicated dominance as a possible causal factor in CAD susceptibility.

Why study the determinants of coronary dominance? Even since the turn of the 20th century when this anatomical variation was categorized in postmortem hearts, cardiologists have been fascinated by whether it impacts heart function or disease[35]. Several studies have attempted to correlate dominance with cardiovascular events, but no consistent data has emerged[9-13]. Our findings show that studying the genetics of human coronary artery development can reveal pathways for cardiac regeneration and repair. Indeed, our top hit for dominance was *CXCL12*, which in animal models revascularizes the infarcted heart through inducing collateral artery growth[36]. This provides a strong rationale to investigate the genes associated with additional loci for their ability to stimulate coronary artery growth during cardiac injury or whether they might synergize with CXCL12 to enhance the function of induced collateral arteries[28].

Key strengths of our human population genetics study include the large number of participants from multiple ancestral groups with a reliable classification of dominance captured through a nationally mandated quality assurance registry, combined with state-of-the-art genotyping, imputation, and linear mixed-model analyses of genetic data. We note that prior initial discovery GWAS for various human

traits has rarely led to the identification of a susceptibility locus in both EUR and AFR. This observation is most likely due to the genetic distance between the populations as well as the imbalance of sample size that usually accompanies such studies due to the longstanding EUR bias in GWAS. The observation at the *CXCL12* locus in both populations despite this imbalance underscores the importance of this locus in determining dominance patterns. We expect to extend discovery in the future when genotyping of additional MVP participants is made available, and through the accrual of additional initial procedures over time among participants who have not yet had an angiogram. A weakness of our study is the very low fraction of female participants with an even lower fraction having undergone an angiogram. This imbalance can only be addressed primarily through the study of dominance in other biobanks with a more balanced number of males and females. Nevertheless, we expect the core biological mechanisms of dominance patterns to be the same in males and females, and this expectation is supported by epidemiologic population genetic data demonstrating no differences in the distribution of dominance patterns between sexes[9,10,37] as well as an overall genetic correlation of one between sexes for a vast majority of human traits[38].

Three of our ten loci will require further replication and scrutiny to increase our confidence that the initial genetic association findings will persist with further study. Two of these loci on chromosome three and five, respectively, involve lead variants that are rare with MAF ≤ 0.5%. Despite these variants being well imputed, the probability of a false positive association increases with rare variants. Additional genotyping in MVP including pending whole genome sequencing will help support or refute these two discoveries. Furthermore, our heritability analyses suggest low-frequency and rare variants contribute a large fraction of the overall heritability of dominance increasing our confidence that these loci will retain their significance over time. The third locus on chromosome three was identified only in AMR despite a substantially higher power to detect the same association in EUR and AFR. To help explain the discrepant findings between ancestral groups at this locus, a comprehensive review of the reliability of the calling of the genotypes in this region by ancestry, combined with a local ancestry analysis, will be required once a larger number of genotyped and sequenced AMRs individuals is available.

Our genetic correlation and instrumental variable analyses exposed a somewhat unanticipated causal relationship of a lower risk of CAD with non-right dominance. Importantly, this finding was not driven solely by the *CXCL12* locus. To reconcile our findings with the literature, we considered research publications on the prognosis after a myocardial infarction separate to publications focused on the

overall burden of coronary atherosclerosis observed by dominance status. While multiple reports suggest left dominance is associated with a poorer prognosis after a myocardial infarction or revascularization procedures[9,10,37,39-42], the clinical relationship between dominance and coronary plaque burden or incident myocardial infarction is less clear[43-45]. A poorer prognosis with a left dominant circulation conditional on a plaque rupture is consistent with the intuition that an individual with a less balanced coronary circulation to the left ventricle will suffer greater damage to it when a plaque rupture occurs within a major branch especially if the rupture and thrombosis is high up in the coronary arterial tree, but this observation does not rule out the possibility that a left dominant or co-dominant circulation otherwise protects against the development of coronary atherosclerosis. It is well established that artery regions exhibiting disturbed blood flow are prone to plaque buildup, e.g., at curved regions or branch points[46]. One might hypothesize that hearts with left- or co-dominance would exhibit hemodynamic differences compared to those with right-dominance, making one or the other experience more non-optimal flow, leading to higher susceptibility to the developments of coronary atherosclerosis. Alternatively, if dominance does not influence the development of atherosclerosis, then it may be that coronary patterning impacts the likelihood of atherosclerosis clinically manifesting as ischemia or infarction. To answer these questions more definitively, large scale studies examining the relationship between dominance and an unbiased documentation of burden need to be conducted. Still, our evidence that human dominance is established during fetal development coupled with the genetic correlation between it and CAD supports an intriguing possibility of a developmental component to adult disease susceptibility.

Methods

Population genetic studies

Design

The design of the MVP has been previously described[57]. Briefly, active users of the Veterans Health Administration (VA) of any age have been recruited from more than 60 VA Medical Centers nationwide since 2011 with current enrollment at >980,000. Informed consent is obtained from all participants to provide blood for genomic analysis and access to their full EHR within the VA prior to and after enrollment. We linked MVP participants to the Veterans Affairs Clinical Assessment, Reporting, and Tracking (CART) Program, a national quality and safety organization for invasive cardiac procedures, to

identify participants who had undergone at least one coronary angiogram by June 2021[63]. Data were available retrospectively starting in 2004 in select sites and from all sites by 2010[64]. The study received ethical and study protocol approval from the VA Central Institutional Review Board.

Genetic Data and Quality Control

A total of 658,164 participants enrolled in MVP between 2011 and 2019 were genotyped with a customized Affymetrix Axiom array in three batches. A genotyping quality control (QC) procedure described in detail elsewhere[58] was applied to all three batches of data and further adapted to remove markers out of Hardy-Weinberg Equilibrium or with differential missingness between sexes. A total of 590,511 autosomal and 12,693 nonPAR X-chromosome markers passing QC were then phased using SHAPEIT4 v4.2.0[47]. Using Minimac4 (v1.0.2)[61], phased haplotypes were then imputed to a TOPMed panel (GRCh38 reference genome) that included ~194,000 phased haplotypes.

Assignment of coronary dominance

Coronary dominance is collected as a structured variable in the EHR template report used to document the results of coronary angiograms for the CART registry and is typically entered by a member of the team who performed the procedure. One of three dominance options can be selected including right, mixed, and left. We first excluded all coronary angiogram reports from individuals with a history of a heart transplant if they were performed on the day of or after the date of their cardiac transplant. For individuals with a single procedure, we assigned dominance based on that single procedure. Most individuals with more than one procedure in CART were fully concordant in their dominance assignments across all procedures. For the remaining, we required that at least 80% of the assignments were concordant to assign a dominance, otherwise the participant was excluded. A total of 61,044 MVP participants were found to have at least one coronary angiogram evaluation with a documentation of dominance.

Statistical Analysis

Genetically inferred ancestry assignment

Population membership to all genotyped participants was assigned using a reference dataset of unrelated individuals from the 1000 Genomes Project (1KGP). This assignment was performed centrally as part of a core MVP project and made available as a core resource for all MVP investigators[33]. Specifically, the smartpca module in the EIGENSOFT package[48] was used to project the principal components (PCs)

loadings from the 1KGP to genotyped MVP participants. To define the genetically similar population, we trained a random forest classifier using cross-population meta-data based on the top 10 PCs from the reference training data. We then used a random forest classifier trained on the predicted MVP PCA data to assign individuals to one of five 1000 Genomes superpopulation groups including White/Europeans (EUR), non-Hispanic Black/African American (AFR), Admixed American (AMR), East Asians (EAS), and South Asian (SAS). Individuals with random forest probability over 50% for a population were assigned to that population. Those who could not be assigned to a population were assigned to 'other'. PCs were then generated for all members within each of the five superpopulations.

Heritability analyses

We used GREML-LDMS as implemented in Genome-wide Complex Trait Analysis (GCTA) v1.94.1 Linux to estimate the multicomponent narrow sense heritability (h^2) of dominance. GREML-LDMS is one of the most accurate heritability estimation methods when considering MAF and LD structures that may bias such estimates[49]. Analyses were restricted to GIA groups with at least > 4,000 case subjects with non-right dominance when applying GREML-LDMS to binary traits to ensure an estimate of heritability with an acceptable standard error[50]. Restricted by computing memory requirements, we selected all non-right dominance subjects and two random right dominance subjects per to run through GREML-LDMS[51,52]. First, SNPs with allele dosage information used for all GWAS were converted to hard-call genotypes using the default settings in PLINK v2.00a3LM. SNPs that were multi-allelic, had MAC < 3, or genotype call-rate < 95% were removed. Since dominance status is a binary trait, SNPs with $p < 0.05$ for Hardy-Weinberg equilibrium or differential missingness in cases (non-right) vs controls (right) were also removed[51,52]. LD scores were calculated separately for each autosome using default GCTA parameters with an r^2 cutoff of 0.01, and the genome-wide LD score distribution was utilized to allocate SNPs to 1 of 4 LD quartile groups, where groups 1-4 represented SNPs exhibiting increasing LD scores. Within each LD group, SNPs were further stratified into 6 MAF bins ([0.001, 0.01], [0.01, 0.1], [0.1, 0.2], [0.2, 0.3], [0.3, 0.4], [0.4, 0.5]) and a genetic relatedness matrix (GRM) was constructed from each bin, ultimately creating 24 GRMs. Finally, reml function in GCTA was used to fit a model of case status (non-right vs right) based on the 24 GRMs, with age and sex as covariates. Total observed heritability estimates were transformed to estimate disease liability across a range of presumed non-right dominance prevalence estimates in the general population including the observed prevalence in MVP.

Genome-wide association study in MVP

We used SAIGE v1.1.6.2[15] which uses the saddlepoint approximation (SPA) to calibrate unbalanced case-control ratios in score tests based on logistic mixed models and allows for model fitting with either full or sparse genetic relationship matrix (GRM) to conduct the coronary dominance GWASs. Each GWAS was adjusted for sex and 10 principal components where the PCs used for all subjects combined analysis were global and generated using genotype data and PLINK while the PCs used for the analyses stratified by superpopulation were generated within each population alone. All summary statistics underwent quality control using EasyQC[53]. Variants were excluded if minor allele count was less than or equal to 6, imputation quality was less than 0.7, or a variant was monomorphic. Sex chromosomes analyses were also excluded.

A locus was considered GWS if at least one lead genetic variant within it reached a $P<5\times10^{-8}$ in any of the individual population GWAS or in the combined multi-population GWAS. Since no previous large-scale GWAS of coronary dominance has been reported, all loci reaching GWS were considered novel. We used Functional Mapping and Annotation of Genome-Wide Association Studies (FUMA-GWAS)[76] to define genomic risk loci including independent, lead, and candidate variants. First, independent genetic variants were identified as variants with a *P* below a specific threshold and not in substantial linkage disequilibrium (LD) with each other ($r2 < 0.6$). Second, variants in LD ($r2 \geq 0.6$) with an independent variant and with $p < 0.05$ were retained as candidate variants to form an LD block. Third, LD blocks within 500kb of each other were merged into one locus. Lastly, a second clumping of the independent variants was performed to identify the subset of lead SNPs with LD $r^2 < 0.1$ within each locus. For our meta-analyses of Whites alone, we used a UK Biobank release 2b EUR reference panel of genotype data imputed to the UK10K/1000G SNPs created by FUMA including ~17 million SNPs. This panel includes a random subset of 10,000 unrelated subjects among all subjects with genotype data mapped to the 1000G populations based on the minimum Mahalanobis distance. We used the 1000G AFR reference panel of 661 subjects with ~43.7 million SNPs for our Blacks, and the AMR reference panel of 347 subjects with ~29.5 million SNPs for our Hispanics. We looked up select independently significant variants identified by FUMA in the GTExPortal to further characterize their expression profile across tissues and determine the directionality of expression in relation to the dominance status. Lastly, we annotated the lead SNP(s) at each novel locus by creating URL hyperlinks to five variant-base portals: OpenTargets, QTLbase, Common Metabolic Disease Knowledge, Open GWAS, and PhenoScanner.

Statistical fine mapping

Statistical fine-mapping was performed using SuSiEx[54], an accurate and computationally efficient method for cross-population fine-mapping using GWAS summary statistics, which builds on the single-population fine-mapping framework, Sum of Single Effects (SuSiE)[55]. SuSiEx improves the power and resolution of fine-mapping while producing well-calibrated false positive rates and retaining the ability to identify population specific causal variants[54]. We used the 1000 Genomes population phase 3 as reference panel to prioritize SNPs or sets of SNPs (credible set) driving each association signal, accounting for multiple causal variants in a genomic region. We calculated 95% credible sets for each identified signal representing the fewest number of variants whose posterior inclusion probabilities (PIP) for the signal summed to ≥ 0.95. We discarded credible sets in which the variants had a minimum $P>1\times10^{-5}$ for the population group in which the signal was mapped.

Colocalization

We assessed the presence of colocalization of genetic association signals between the *CXCL12* for coronary dominance and associations in analogous regions for CAD using COLOC[31]. For these analyses, we used the summary statistics from the Cardiogram*plus*C4D Consortium GWAS[34] and our GWAS of coronary dominance, analyzing the genomic window where the *CXCL12* locus is located (from 44210000 bp to 44810000 bp). Evidence of colocalization at a locus with the same causal variant shared between CAD and coronary dominance was defined as a posterior probability Bayesian factor H4 (PP.H4.abf) > 0.7 while evidence of colocalization with a different variant was defined as defined as a posterior probability Bayesian factor H3 (PP.H3.abf) > 0.7.

Genetic correlation

We applied LD-score regression (LDSC)[32] using available GWAS summary statistics for CAD from a recently completed phenome-wide GWAS across 2,068 traits in 449,042 MVP EUR participants[33], and separately, a recently completed Cardiogram*plus*C4D consortium meta-analysis of summary statistics for CAD in largely EUR populations that did not include MVP[56] to evaluate pairwise genome-wide genetic correlations (r_g) between coronary dominance and other traits. We used pre-calculated European LD scores and restricted the analysis to SNPs found in HapMap Phase 3[57] removing the human leukocyte antigen (HLA) region due to its unusual LD structure and genetic architecture.

Mendelian randomization

We used genome-wide significant independent hits (after LD clumping using a window of 10 Mb and r2 cutoff = 0.001) associated with the coronary dominance GWAS as instrument variables (IVs) for associations with Cardiogram*plus*C4D Consortium[34]. We performed two-sample MR using four separate methods to estimate causal effects: the standard inverse-variance weighted (IVW) regression with and without MR-PRESSO (to minimize the risk of horizontal pleiotropy)[58]; as well as two robust regression methods, the weighted median-based method, and Egger regression[59] using the R package TwoSampleMR[60]. The MR-PRESSO analysis attempts to reduce heterogeneity in the estimate of the causal effect by removing SNPs that contribute to the heterogeneity disproportionately more than expected. The number of distributions in MR-PRESSO analysis was set to 1000. In addition, we calculated the intercept term in MR Egger regression, as a useful indication of whether directional horizontal pleiotropy is driving the results of a MR analysis. Additionally, to identify potentially influential SNPs, we performed a "leave-one-out" sensitivity analysis to where the MR is performed again but leaving out each SNP in turn.

Animal Studies

Animals

All mouse colonies were housed and bred in the animal facility at Stanford University in accordance with institutional animal care and use committee (IACUC) guidance on 12 h/12 h day and night cycle with water and food ad libitum.

The following mouse strains were used: C57BL/6J, (Jackson Laboratory, strain code: 000664), CD1 (Charles River, strain 022) and *Cxcl12*^{DsRed} (Jackson Laboratory, strain code: 022458). All experiments were conducted in accordance with protocols approved by the Institutional Animal Care and Use (IACUC) Committee of Stanford University.

Cxcl12^{DsRed} mice were maintained on a C57BL/6J background. To assess SpA dominance, *Cxcl12*$^{DsRed/+}$ males were mated with wild type C57BL/6J females to produce litters with wild type (*Cxcl12*$^{+/+}$) and heterozygous (*Cxcl12*$^{DsRed/+}$) mice. All neonatal studies used mice between postnatal days (P)0-P6. For all embryonic studies, timed pregnancies were determined by defining the day on which a vaginal plug was found as E0.5.

Assessing SpA dominance

Dominance was assigned from whole-organ images of coronary arteries imaged using light sheet microscopy and obtained by one of two experimental methods, whole-organ immunofluorescence or *in vivo* labeling (see details for each below).

SpA dominance was determined in 3D using Imaris. First, the ventricular septum was identified using the Oblique & Ortho Slicer and Scissors Tools. Then, it was confirmed that the septum was principally irrigated by one or two SpAs. Lastly, using the Oblique & Ortho Slicer and Scissors tools, the SpA was followed until it merged with the most proximal RCA or LCA segments. On rare occasions, the SpA(s) were followed to an independent connection point (ostium) on the aorta. Based on these observations, SpA dominance was scored as right-, left-, or co-dominant, or as aorta. When scoring SpA dominance in *Cxcl12* mice, researcher was blinded to genotype.

Whole-organ immunolabeling

Whole heart immunofluorescence was performed following the modified iDISCO+ protocol previously described[24,61] For all following steps, tissue was always agitated unless noted otherwise. Briefly, animals were perfused with PBS through the dorsal vein, and fixed in 4% paraformaldehyde (Electron Microscopy Science 15714) at 4°C for 1hr (embryonic and neonatal hearts) or 2hr (adult hearts), washed 3X in PBS and stored in PBS with 0.01% sodium azide (w/v, Sigma-Aldrich S8032) until ready to process. Hearts were dehydrated in increasing series of methanol/ddH-$_2$O dilutions (20%, 40%, 60%, 80%,100% 2X) for 1hr each, followed by overnight incubation in 66% dichloromethane (DCM, Sigma-Aldrich 34856) and 33% methanol. Next, tissue was washed 2X in 100% methanol for 4hrs and bleached overnight at 4°C in 5% hydrogen peroxide (Sigma-Aldrich 216763) in methanol. Next, the hearts are rehydrated in methanol/ddH-$_2$O dilutions (80%, 60%, 40%, 20%) for 1hr each, followed by PBS, 0.2% Triton X-100 PBS (2X) and incubated in 20% dimethyl sulfoxide (DMSO), 2.3% Glycine (w/v, Sigma G7126), and 0.2% Triton X-100 PBS at 37°C for 2 days.

For immunofluorescence, hearts were blocked in 10% DMSO, 6% Normal Donkey Serum (NDS, Jackson ImmunoResearch 017-000-121) in 0.2% Triton X-100 for 2 days at 37C. Primary antibodies were prepared in PBS with 5%DMSO, 3% NDS in 0.2% Tween-20 and 0.1% Heparin (w/v, Sigma-Aldrich H3393) and incubated at 37°C for 4-14 days. Antibodies included Cy3 conjugated-αSMA-Cy3 (1:300, Sigma C6198), Podocalyxin (1:1000, R&D Systems MAB1556), RFP (1:1000, Rockland 600-401-379), Jagged1 (Novus, AF599). Secondary antibodies conjugated to either Alexa 555

or 647 (Jackson ImmunoResearch) were matched 1:1 in concentration to their primary target and in prepared in PBS with 3% NDS in 0.2% Tween-20, 0.1% Heparin for the same primary incubation days at 37°C. Washes after each antibody incubation in PBS with 0.2% Tween-20, 0.1% in Heparin were performed in 30min increment until the end of the day, followed by an overnight wash. Before clearing, samples were embedded in 1% low-melting agarose (Sigma-Aldrich A9414) in PBS and dehydrated in methanol/ddH-$_2$O dilutions (20%, 40%, 60%, 80%,100% 2X) for 1hr each and 100% overnight. Next, hearts were incubated in 66% DCM and 33% Methanol for 2.5hrs, followed 2X 30min 100% DCM. Finally, samples were cleared in ethyl cinnamate (ECi, Sigma Aldrich 112372), manually inverted a few times, and kept at RT in the dark until imaging.

In vivo vascular labeling

Neonate pups were gently wrapped in gauze and cooled on ice for 6 minutes to induce hypothermic circulatory arrest. Mice received an intravenous 10 µl retro-orbital injection of Isolectin GS-IB$_4$ DyLight™ 649 (1:5, Invitrogen, I32450, in sterile 1X PBS). Neonates were then allowed to recover at 37°C on a warm plate and, when conscious, returned to their mother's care for 30 min before euthanasia. Hearts were then dissected and fixed in 4% PFA (Electron Microscopy Science, 15714) at 4°C for 1 hour. Subsequently, hearts were washed three times in 1X PBS and stored in cold PBS with 0.01% sodium azide (w/v, Sigma-Aldrich, S8032) and covered from light until ready to process.

Hearts were then cleared following the protocol previously described[62]. For all following steps, tissue was always agitated unless noted otherwise. Before clearing, samples were embedded in 1% low-melting agarose (Sigma-Aldrich, A9414) in PBS and dehydrated in methanol/ddH-$_2$O dilutions (20%, 40%, 60%, 80%,100% 2X) for 1hr each and 100% overnight. Next, hearts were incubated in 66% DCM and 33% Methanol overnight, followed next day by 2X 100% DCM washes the next day. Finally, samples were cleared in ethyl cinnamate (ECi, Sigma Aldrich, 112372), manually inverted a few times, and kept at RT in the dark until and after imaging.

Light sheet imaging

Samples were imaged with Imspector Pro 7.0.98 software and LaVision BioTec Ultramicroscope II Light sheet microscope in a quartz cuvette filled with ECi. For imaging, we used a MVX10 zoom body (Olympus) with a 2x objective (pixel size of 3.25 µm / x,y) at magnification from 0.63x up to 1.6x. Up to 1400 images were taken for each heart using a z-step size of 3.5µm z step size, and light sheet numerical aperture to 0.111 NA. Band-pass emission filters (mean nm / spread) were used, depending on

the excited fluorophores: 525/50 for autofluorescence; 595/40 for Cy3; and 680/30 for AF647. Exposure time was 10ms for single channel and 25ms for multichannel acquisition.

Image processing and visualization

Maximum projections were performed using NIH Fiji ImageJ 1.53 software. Representative images in figures were in some cases processed with Despeckle and/or Unsharp Mask features. Raw images are included in Supplement. Imaris 9.5.0 software (Bitplane) was used for 3D rendering. Pixel dimensions were updated from the non-reduced 16-bit image metadata. The Filament Object Tracer module in Imaris was used to generate a simple 3D outline of the main coronary arteries. A root point was manually placed where the RCA and LCA originate from the aorta, i.e. right and left ostium, and filaments were then created and drawn until the farthest length possible using the Autopath and AutoDepth features.

Quantitative PCR

Whole hearts were collected from E16.5 mice and lysed in 1 mL of TRIzol (Invitrogen 15596026) in tubes containing Lysing Matrix G (MP Biomedicals: 116916050) with homogenization (3X 30 sec at 30 Hz) in a Qiagen TissueLyser II followed by isolation of the aqueous phase according to the TRIzol reagent user guide. RNA was then extracted and isolated with a NucleoSpin RNA isolation kit (Takara 740955.50) according to the manufacturers' protocols. 100 ng of total RNA was then reverse transcribed with the SensiFAST cDNA synthesis kit (Meridian Bioscience BIO-65053). Total cDNA was then diluted 1:10 and qPCR was performed with the SensiFAST SYBR® No-ROX Kit (Meridian Bioscience BIO-98005) with 10 µL of reactants in a 384 well plate (ThermoFisher 4483285).

Primer pairs were designed such that they span exon/exon junctions to preclude amplification of genomic DNA and were specific for murine *Cxcl12*. Each well of the plate received 5µL of SensiFAST SYBR® No-ROX Mix, .04 µL of 100mM forward primer, .04 µL of 100mM reverse primer, 2.92µL of water and 2µL of cDNA. After plate preparation with gene specific primers combined in a well with a sample specific cDNA the plates were sealed with MicroAmp™ Optical Adhesive Film (ThermoFisher 4311971) and centrifuged for 5 min at 350g. Plates were run on a Quantstudio Flex 7 (ThermoFisher) with the following parameters: Initial dissociation (2 min at 50°C followed by 10 min at 95°C), 40x PCR stage with SYBR detection (15 sec 95°C, 1 min 60°C) with the final stage being a dissociation step to generate a melt curve for the PCR products to ensure a single product was amplified. During analysis the linear phase of the fluorescence was used to generate Ct (cycle threshold) values. The ddCt (delta-

delta Ct) method was used to analyze the results. *Gapdh*, a housekeeping gene, was used as an internal control to normalize between samples which were then compared to the expression of the wild type to quantify *Cxcl12* expression between genotypes. For each well there were 3 technical replicates as well as a water negative control.

Statistical analysis

Statistical analysis and graphs were generated using Prism 9 (GraphPad). Graphs represent mean values obtained from multiple experiments and error bars represent standard deviation. For qPCR, two-tailed unpaired Student's t test was used to compare groups within an experiment and the level of significance were assigned to statistics in accordance with their p values (0.05 flagged as $*$, 0.01 flagged as $**$, less than 0.001 flagged as $***$, less than 0.0001 flagged as $****$). When analyzing SpA dominance in *Cxcl12* mice, a two-tailed Chi-square without Yates-correction test was performed in a 2x2 contingency table (right & non-right x wild type & heterozygotes).

Human Fetal Heart studies

Under IRB approved protocols, human fetal hearts were obtained for developmental analysis[63]. Gestational age was determined by standard dating criteria by last menstrual period and ultrasound[64]. Pregnancies complicated by multiple gestations and known fetal or chromosomal anomalies were excluded. Tissues were processed within 1hr following procedure at which time it was rinsed with cold, sterile PBS. For whole-organ immunolabeling, tissue was fixed in 4% PFA for 4hrs at 4°C before whole-organ immunolabeling. For *in situ* hybridization, hearts were fixed in 4% PFA for 24–48 hr at 4°C, followed by three 15 min washes in PBS.

Dominance analysis at fetal stages

3D reconstructions of fetal hearts subjected to whole-organ immunolabeling were performed by the method mentioned above under the heading Image processing and visualization. Briefly, root points were manually placed at the right and left ostia and filaments were then created and drawn until the farthest length possible using the Autopath and AutoDepth features in the Filament Object Tracer module in Imaris. Arterial landmarks and large diameter vessels were prioritized and traced first, followed by vessels extending into the septum on the inferior side of the heart.

Right dominance was assigned when the filament trace of the PDA was connected to the RCA filament trace. Additionally, when looking at a central cross-section, the arteries dominating the volume

were traced from the RCA and its ostium. Co-dominance was assigned when, looking at a central cross-section, arteries in this volume were equally occupied by those traced from the RCA and LCA to their ostia.

In situ hybridization

Hearts were sequentially dehydrated in 30%, 50%, 70%, 80%, 90%, and 100% ethanol, washed 3X for 30 min in xylene, washed several times in paraffin, and finally embedded in paraffin which was allowed to harden into a block. For each heart, the whole ventricle was cut into 10-µm-thick sections and captured on glass slides.

CXCL12 mRNA and α-SMA protein were simultaneously localized on sections using RNAscope® per manufacturers' instructions. Briefly, slides were baked in 60°C for 1hr in a dry oven. Next, slides were incubated 2X in xylene with slight agitation for 5 min in RT under fume hood. Immediately after, slides were incubated 2X in 100% ethanol for 2 min under the same conditions. Slides were then dried for 5 min in the drying oven at 60°C. Afterwards, ~5 drops of hydrogen peroxide were added to the deparaffinized slides and incubated at RT for 10 min. Slides were washed 2X with fresh distilled water. In a food steamer, target retrieval was performed for 15 minutes with RNAscope Retrieval Buffer. The tissue was pretreated in mild conditions, with RNAscope Protease III for 15 min. *In situ* hybridization was continued with the RNAscope® Multiplex Fluorescent Reagent Kit v2 (ACD 323100) using the following probes: human *CXCL12* (ACD 422991), positive control (ACD 320861), and negative control (ACD 320871). The ACD RNAscope protocol was followed in Channel 1 for all probes with Opal 650 fluorophore (Akoya FP1496001KT).

Prior to mounting, slides were fixed in 4% PFA for 30 min, followed by 3X 15 min washes in PBS-0.1% Tween. The slides were then immunostained with Cy3 conjugated-α-SMA-Cy3 (1:300, Sigma C6198) for 1hr at RT. Then, slides were washed for 10 min 3X in PBS-0.1% Tween. To remove blood cells and reduce autofluorescence, slides were washed in $CuSO_4$ for 30 min at RT in a nutating mixer. Slides were then washed with distilled water before adding DAPI and mounting in ProLong Gold Antifade reagent (Invitrogen P36934). Finally, slides were imaged within a week on a confocal microscope (Zeiss LSM 980).

URLs

GCTA, http://cnsgenomics.com/software/gcta/#Overview;

SAIGE, https://saigegit.github.io/SAIGE-doc/

PLINK: https://www.cog-genomics.org/plink/;

EasyQC, https://www.uni-regensburg.de/medizin/epidemiologie-praeventivmedizin/genetische-epidemiologie/software/;

R statistical software, www.R-project.org;

FUMA, http://fuma.ctglab.nl/;

SUSIEx, https://github.com/getian107/SuSiEx

Locus Zoom, https://my.locuszoom.org/

GTExPortal, https://www.gtexportal.org/home/

OpenTargets, https://genetics.opentargets.org/;

QTLbase, http://www.mulinlab.org/qtlbase;

Common Metabolic Disease Knowledge Portal, https://hugeamp.org/;

Open GWAS, https://gwas.mrcieu.ac.uk/;

PhenoScanner, http://www.phenoscanner.medschl.cam.ac.uk/;

CADD, https://cadd.gs.washington.edu/

COLOC, https://chr1swallace.github.io/coloc/articles/a01_intro.html

LDSC, https://github.com/bulik/ldsc;

CARDIoGRAMplusC4D, http://www.cardiogramplusc4d.org;

TwoSampleMR, https://mrcieu.github.io/TwoSampleMR/index.html

MR-PRESSO, https://github.com/rondolab/MR-PRESSO

Data availability

The full summary level association statistics from each genetically inferred ancestry population association analyses in MVP from this report will be available through dbGaP, with accession number phs001672 at the time of publication in a peer reviewed journal.

Funding/support

This research is based on data from the Million Veteran Program, Office of Research and Development, Veterans Health Administration, and was supported by Veterans Administration awards I01-01BX003362. The content of this manuscript does not represent the views of the Department of Veterans Affairs or the United States Government. P.E.R.C was supported by NIH-TIMBS-T32HL094274. E.N.F was supported by NIH-T32HL007444. J.A.N was supported by NIH-

T32GM007276. S.L.C was supported by NIH-CTSA KL2TR003143. K.R-H was supported by NIH-R01HL12850307 and the Howard Hughes Medical Institute.

Acknowledgements

Support for title page creation and format was provided by AuthorArranger, a tool developed at the National Cancer Institute. The Genotype-Tissue Expression (GTEx) Project was supported by the Common Fund of the Office of the Director of the National Institutes of Health, and by NCI, NHGRI, NHLBI, NIDA, NIMH, and NINDS. The data/figures used for the analyses described in this manuscript were obtained from the GTExPortal on 10/18/21.

Consortia

The VA Million Veteran Program. Members of the consortium can be found in the Supplementary Text File.

Author Contributions

Concept and design: P.E.R.C., D.Z., S.L.C, K.R-H., T.L.A.
Acquisition, analysis, or interpretation of data: P.E.R.C., D.Z., J.Z., J.A.N, P.P., P.K., A.M.M., A.T.H., S.P., D.D., K.C., V.D.W., A.M.P., M.E.P., S.W.W., P.S.T, S.L.C, K.R-H., T.L.A.
Drafting of the manuscript: P.E.R.C, D.Z., J.Z., S.L.C., K.R-H., T.L.A.
Critical revision of the manuscript for important intellectual content: P.E.R.C, D.Z., S.L.C., K.R-H., T.L.A.

Ethics Declarations

None.

Table 1. Lead genetic variants, effect sizes, allele frequencies, and mapped genes at loci reaching genome wide significance for coronary artery dominance status

Genomic Locus No	GIA	GWAS Analysis	chr	snp rsID	bp hg38/GRCh38	bp hg19/GRCh37	Allele1	Allele2	AF Allele2	BETA	SE	OR (95%CIs)	p.value	AF case	AF ctrl	N_case	N_ctrl	N	Mapped Genes
1	AFR	RvsL	3	rs544905040	2757792	2841160	A	G	0.005	1.33	0.25	3.8 (2.33-6.2)	4.71E-08	0.013	0.004	1167	8280	9447	CNTN6
2	AMR	RvsLM	3	rs9873160	197417938	197144809	G	C	0.073	0.56	0.10	1.75 (1.43-2.14)	1.82E-08	0.108	0.066	738	3717	4455	ZNF32
3	EUR	RvsLM	4	rs56410855	173452142	174373293	G	T	0.094	0.17	0.03	1.19 (1.12-1.26)	2.32E-09	0.107	0.091	7708	37613	45321	ATP6V0E1, BNIP1, NKX2-5
3	ALL	RvsLM	4	rs112133513	173487900	174409051	G	A	0.081	0.17	0.03	1.19 (1.12-1.25)	3.40E-10	0.090	0.079	10638	50405	61043	ATP6V0E1, BNIP1, NKX2-5
3	EUR	RvsL	4	rs56410855	173452142	174373293	G	T	0.094	0.21	0.04	1.24 (1.15-1.33)	1.03E-09	0.111	0.091	4944	37613	42557	SCRG1, HMGB2, FBXO8, CEP44
3	ALL	RvsL	4	rs78510568/rs748824055	173475861	174397012	C	T	0.078	0.21	0.03	1.24 (1.16-1.32)	1.95E-10	0.092	0.076	6640	50405	57045	ATP6V0E1, BNIP1, NKX2-5
4	ALL	RvsLM	5	rs567007743	99585888	98921592	A	G	0.002	0.67	0.15	1.95 (1.45-2.63)	5.05E-06	0.003	0.002	10638	50405	61043	ELL2, PCSK1, ERAP1, NA, ERAP2, LNPEP, NA, LIX1, RIOK2, RGMB, CHD1, FAM174A, ST8SIA4, SLCO4C1
5	ALL	RvsLM	5	rs954635	173220844	172647847	A	T	0.362	0.11	0.02	1.12 (1.08-1.16)	3.03E-11	0.389	0.357	10638	50405	61043	INSC, SOX6, C11orf58
5	EUR	RvsLM	5	rs954635	173220844	172647847	A	T	0.282	0.15	0.02	1.16 (1.11-1.2)	1.46E-13	0.306	0.277	7708	37613	45321	INSC, SOX6, C11orf58
5	EUR	RvsL	5	rs62385088	173222031	172649034	G	A	0.164	0.21	0.03	1.23 (1.17-1.3)	1.09E-13	0.190	0.161	4944	37613	42557	INSC, SOX6, C11orf58
5	ALL	RvsL	5	rs55893552	173223617	172650620	T	A	0.158	0.17	0.03	1.18 (1.12-1.24)	3.79E-11	0.175	0.156	6640	50405	57045	INSC, SOX6, C11orf58
6	ALL	RvsL	10	rs2902339	44018658	44514106	G	A	0.784	-0.24	0.02	0.79 (0.75-0.82)	5.14E-28	0.746	0.788	6640	50405	57045	CXCL12
6	EUR	RvsL	10	rs7917534	44041975	44537423	G	C	0.798	-0.25	0.03	0.78 (0.74-0.82)	1.29E-21	0.762	0.803	4944	37613	42557	CXCL12
6	EUR	RvsLM	10	rs589655	44249867	44745315	C	G	0.100	0.31	0.03	1.37 (1.3-1.45)	4.37E-29	0.126	0.095	7708	37613	45321	CXCL12
6	AFR	RvsL	10	rs2576354	44254945	44750393	G	A	0.259	0.32	0.05	1.37 (1.25-1.51)	9.21E-11	0.317	0.251	1167	8280	9447	CXCL12
6	AFR	RvsLM	10	rs2576354	44254945	44750393	G	A	0.262	0.27	0.04	1.3 (1.21-1.41)	1.17E-11	0.305	0.251	2040	8280	10320	CXCL12
6	ALL	RvsLM	10	rs606314	44255031	44750479	C	G	0.101	0.30	0.02	1.36 (1.29-1.42)	6.58E-37	0.126	0.096	10638	50405	61043	CXCL12
7	ALL	RvsL	10	rs1194743	52452837	54212597	T	C	0.779	0.15	0.02	1.17 (1.11-1.22)	2.44E-10	0.802	0.776	6640	50405	57045	ASAH2, SGMS1, ASAH2B, A1CF, PRKG1, CSTF2T, DKK1, MBL2
7	EUR	RvsL	10	rs1194743	52452837	54212597	T	C	0.747	0.15	0.03	1.16 (1.1-1.23)	7.51E-09	0.770	0.744	4944	37613	42557	A1CF, PRKG1, DKK1
8	ALL	RvsL	11	rs35342212/rs67428213	15874288	15895834	G	A	0.229	-0.13	0.02	0.87 (0.83-0.92)	1.05E-08	0.206	0.232	6640	50405	57045	INSC, SOX6, C11orf58
8	ALL	RvsLM	11	rs35342212/rs67428213	15874288	15895834	G	A	0.228	-0.11	0.02	0.9 (0.86-0.93)	7.43E-09	0.208	0.232	10638	50405	61043	INSC, SOX6, C11orf58
9	ALL	RvsL	15	rs58473469	174592722	174885063	C	T	0.200	-0.16	0.03	0.85 (0.81-0.9)	9.75E-10	0.183	0.202	6640	50405	57045	SEMA7A, UBL7, ARID3B, CLK3, CYP1A1, LMAN1L, SCAMP2, SCAMP5, SNUPN
10	ALL	RvsLM	17	rs851058	43760351	41837719	G	A	0.391	0.09	0.02	1.09 (1.06-1.13)	1.40E-08	0.405	0.388	10638	50405	61043	DUSP3, C17orf105, MPP3
10	EUR	RvsLM	17	rs851058	43760351	41837719	G	A	0.417	0.11	0.02	1.11 (1.07-1.15)	4.54E-09	0.438	0.413	7708	37613	45321	DUSP3, C17orf105, MPP3

Genomic locus No: Index of genomic rick loci, GIA: Genetically inferred ancestry (ALL; all ancestral groups combined, EUR; genetically inferred European subjects, AFR; genetically inferred African subjects, AMR; genetically inferred Admixed-Americans), GWAS: Genome-wide assocociation analysis comparison (RvsLM; right dominance as reference versus left and mixed subjects combined, RvsL; right dominance (reference) versus left dominance subjects excluding subjects with mixed dominance, chr: chromosome, SNP rsID: variant ID, bp hg38/GRCh38: base position on human build 38, bp hg19/GRCh37: base position on human build 37, Allele1: allele 1, Allele2: allele 2, AF_Allele2: allele frequency of allele 2, BETA: effect size of allele 2, SE: standard error of BETA, OR (95%CIs): Odds ratio and 95% confidence intervals of non-right dominance to right dominance for allele 2 compared to allele1, p.value: p value with SaddlePoint Approximation applied, AF_case: allele frequency of allele 2 in cases, AF_ctrl: allele frequency of allele 2 in controls, N_case: sample size in cases, N_ctrl: sample size in controls, N: sample size cases and controls combined

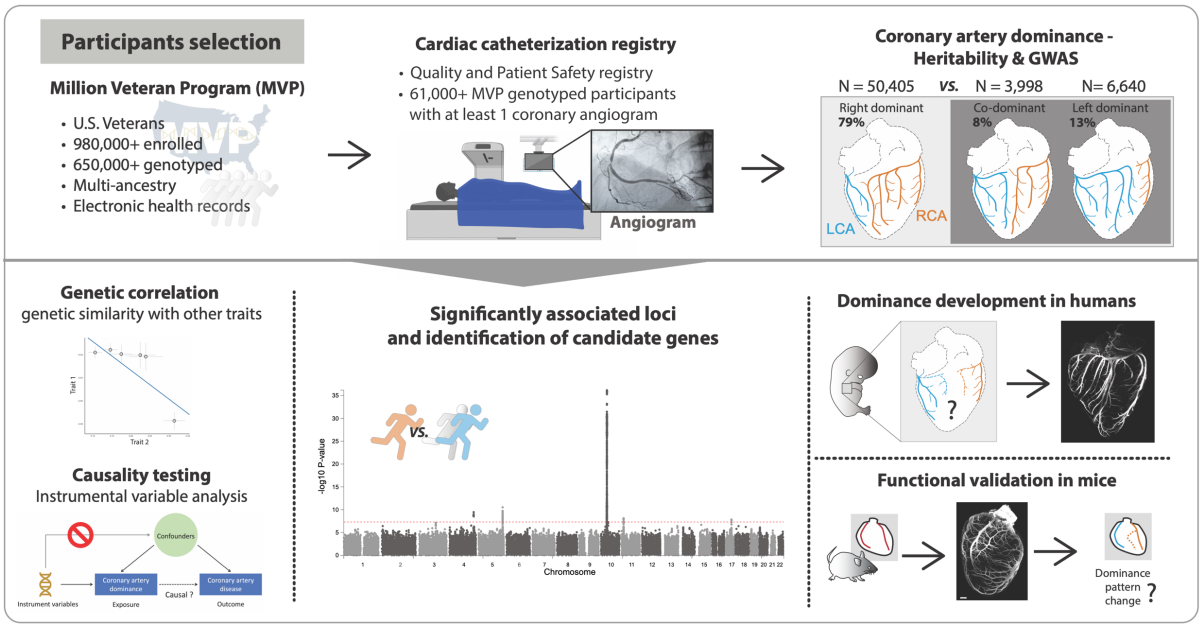

Figure 1: Design of a study to identify and understand the genetic determinants of coronary artery dominance. A multi-ancestry Genome-wide Association Study (GWAS) of coronary artery dominance was conducted using genotype data imputed to the TOPMed panel in 61,043 participants in the Million Veteran Program (MVP) with at least one determination of dominance status through a coronary angiogram. Analyses were conducted in all participants combined and stratified by genetically inferred ancestry. In one set of analyses, right dominance was compared to left and mixed dominance combined, and in the second set, right was compared to left alone after excluding mixed dominance. In addition to the identification of susceptibility loci, an estimate of heritability, genetic correlation, and instrumental variable analyses was pursued in European participants. Whole heart fetal imaging and mouse validation studies were also pursued to shed light on the mechanism of association for the most significant findings in the MVP.

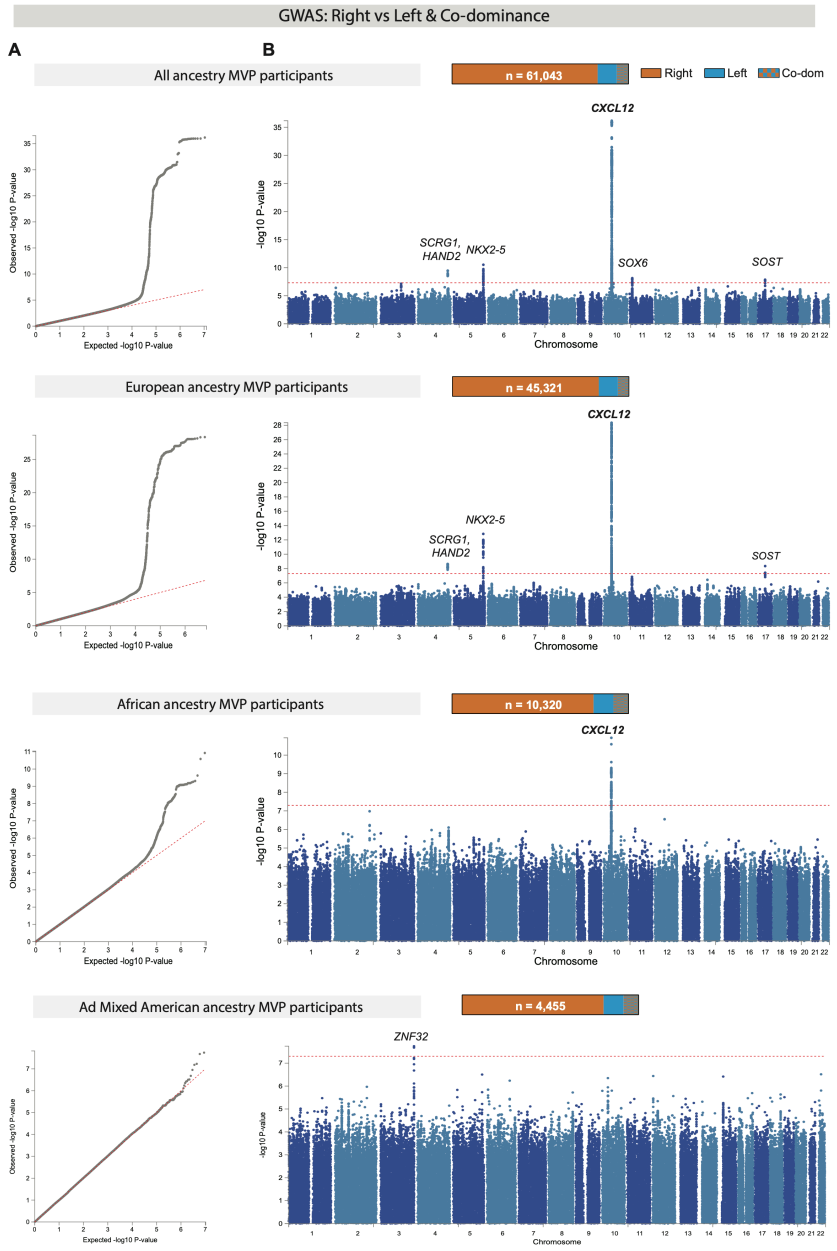

Figure 2: Genome-wide Association Study (GWAS) of right versus left and mixed dominance combined in the Million Veteran Program. A GWAS was conducted in all Veteran participants combined and in each of the three largest genetically inferred ancestry subgroups: Europeans (EUR), Africans (AFR) and Admixed Americans (AMR). (**A**) Evidence of a strong genetic signal in the combined, EUR, and AFR analyses with deviation from the null of the p-values of association for all genetic variants on quantile-quantile (q-q) plots without inflation. The genetic signal is weakest in the smallest AMR group. (**B**) Manhattan plots for the corresponding q-q plots with p-values of genetic variants represented in order by chromosome and position on the chromosome. Gene(s) mapped to regions reaching genome wide significance are labeled. Q-Q and Manhattan plots for the right versus left only dominance can be found in the Supplementary Figures file.

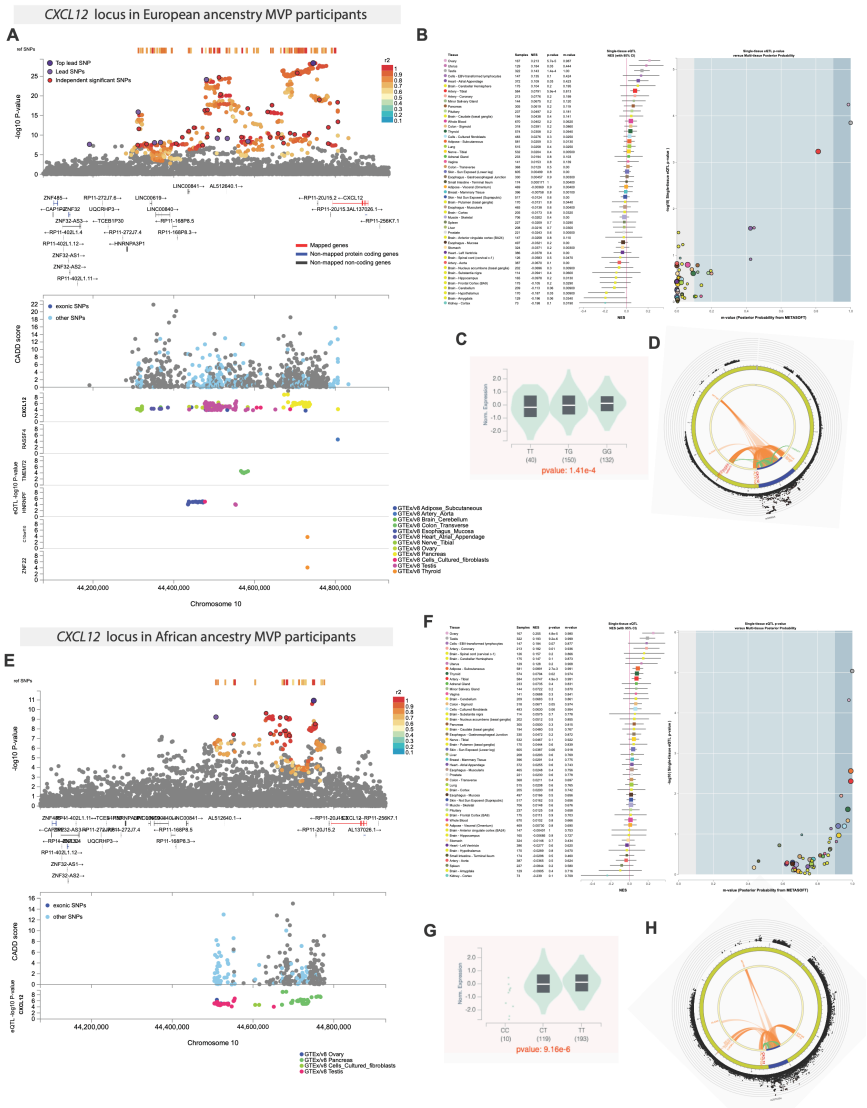

Figure 3: Annotation of CXCL12 susceptibility locus for coronary dominance in European and African participants in the Million Veteran Program. (**A**, **E** top) Analogous locus zoom plots derived by FUMA-GWAS for the dominance locus immediately downstream of *CXCL12* in the European (EUR) and African (AFR) participant GWAS, respectively. The plots demonstrate substantial overlap of the signal between the two ancestral groups. (**A**, **E** bottom) show CADD scores of the independently significant genetic variants. Variants with scores >10 are predicted to affect the expression of genes nearby most likely by affecting binding to regulatory sites. The same genetic variants are also identified as expression quantitative trait loci (eQTL) in multiple tissues in the Genotype-Tissue Expression (GTEx) project database version 8 (v8). Notably only CXCL12 is implicated by eQTLs in the GWAS data from AFR compared to CXCL12 and four other genes in EUR. (**B, F**) eQTL multi-tissue plots from the GTExPortal for two genetic variants found to be independently significant for dominance in EUR and AFR, respectively. While the strongest effects of tissue expression for these SNPs are in ovary, testis, and the uterus, the genetic variant in EUR is also an eQTL for the tibial artery and the variant in AFR is an eQTL for both tibial and coronary arteries. (**C, G**) eQTL violin plots of gene expression levels by genotype in testis for same two variants, respectively, with a direction of effect on expression levels that is concordant with the arterial tissues. When combined with the GWAS, plots show that increased expression of CXCL12 is associated with an increased prevalence of right dominance. (**D, H**) Circos plots of chromatin interactions and eQTLs. Only mapped genes by either chromatin interaction and/or eQTLs conditional on user defined parameters in FUMA-GWAS are displayed. When gene is mapped only by chromatin interactions or only by eQTLs, it is colored orange or green, respectively. When it is mapped by both, it is colored red. Of note, only *CXCL12* is mapped in AFR by both eQTLs and chromatin.

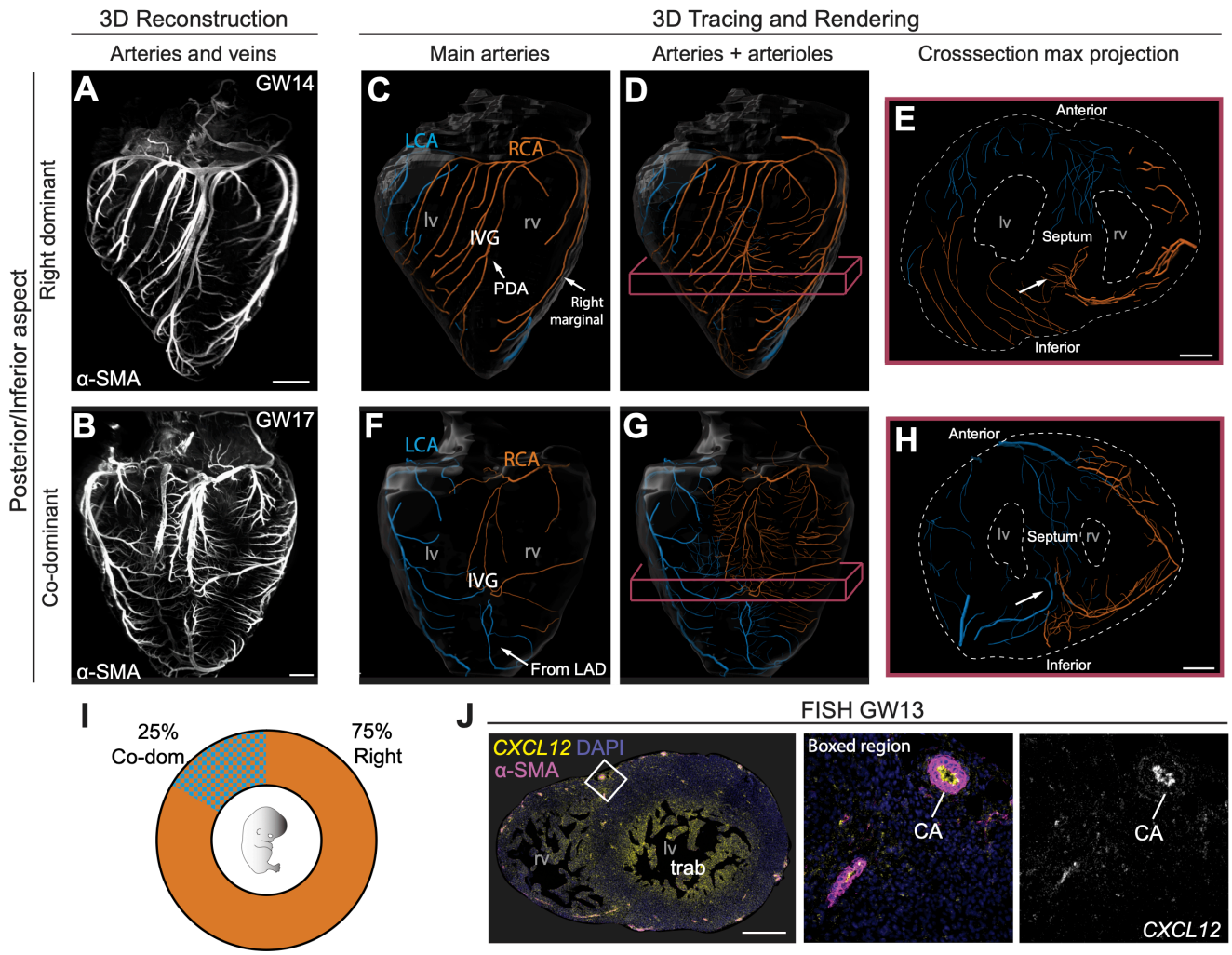

Figure 4. Coronary dominance is apparent during fetal development when *CXCL12* is expressed. (**A** and **B**) Three dimensional (3D) reconstructions of hearts from the indicated gestational weeks (GW) subjected to whole-organ immunolabeling for alpha-smooth muscle actin (α-SMA). (**C-H**) 3D reconstructions of representative tracings of α-SMA+ arteries. Traces highlighted branches of the main arteries (**C, F**) or main with lower order arteries (**D, E, G, H**) that originated from the right (RCA, orange) or left (LCA, blue) coronary ostia where they meet the aorta. (**C-E**) Right dominance was indicated when the artery extending along the interventricular groove (IVG), i.e., the posterior descending artery (PDA), originated from the RCA (**C** and **D**) and when the inferior septum was primarily occupied by right branches (arrow in **E**). (**F-H**) Co-dominance was indicated when arteries within the IVG stemmed from both the RCA and LCA (**F** and **G**) and when the inferior septum was occupied equally by RCA or LCA branches (arrow in **H**). (**I**) Graphic representation of the percentages of right- and co-dominance in our cohort. N=8 hearts. (**J**) *CXCL12* fluorescence *in situ* hybridization (FISH) and α-SMA immunolabeling on a transverse section through a GW13 heart showed expression in coronary arteries (CA) and trabecular myocardium (trab). DAPI labeled nuclei. Scale bars: A, B, E, and J, 500µm; H, 1mm.

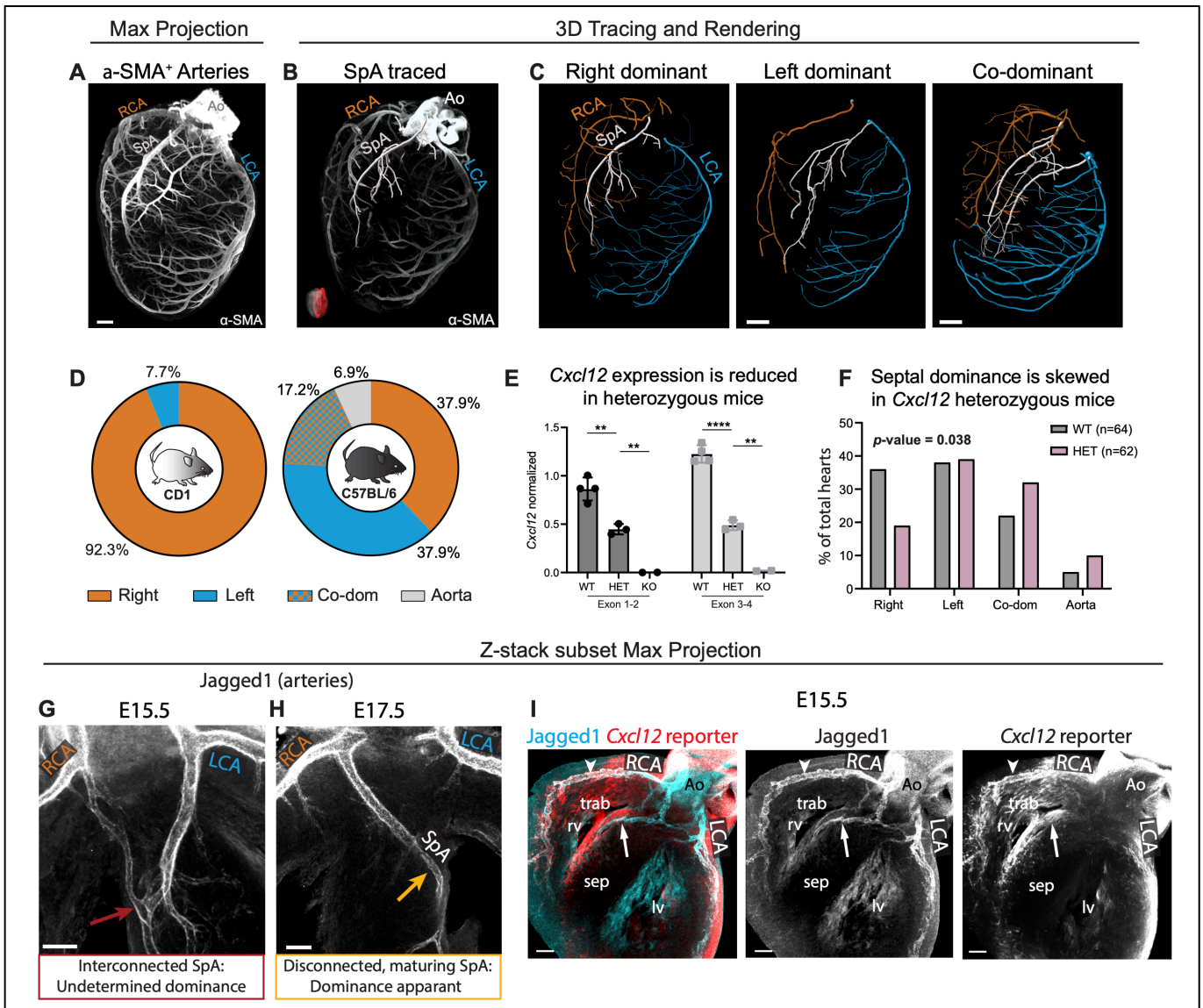

Figure 5. *Cxcl12* influences septal artery dominance in mice. (**A**) Whole-organ immunolabeling of coronary artery smooth muscle (α-SMA) in a postnatal day (P) 6 mouse heart. (**B**) Tracing of the septal artery (SpA) in the same heart overlayed onto a Z-stack subset max projection showing internal coronary arteries (see inset 3D navigator, red region indicates Z-stack). (**C**) Tracings of arteries in whole-organ images demonstrating the various SpA connections (left panel is same heart as **A** and **B**). (**D**) Quantifications of SpA dominance in neonatal hearts from the indicated wild type mouse strains. N=13 CD1, N=29 C56BL/6J. (**E**) QPCR measuring *Cxcl12* mRNA levels in whole hearts from embryonic day (E) 16.5 mice of the indicated genotypes. WT, $Cxcl12^{+/+}$; HET, $Cxcl12^{DsRed/+}$; KO, $Cxcl12^{DsRed/DsRed}$. Each dot represents an individual heart. Error bars are mean □ st dev: **, p=0.0027, p=0.0016, p=0.0011; ****, p<0.0001 by two-sided Student's t-test. (**F**) Septal dominance in wild type ($Cxcl12^{+/+}$) and heterozygous ($Cxcl12^{DsRed/+}$) neonatal mice. p-value calculated by two-tailed Chi-square test. (**G** and **H**) Representative images from whole-organ immunolabeling of wild type embryonic mouse hearts demonstrates stages of SpA development. Earlier in embryonic development, an immature artery network is attached to both sides, which later matures into a disconnected SpA with dominance shortly before birth. (**I**) Representative E15.5 heart heterozygous for the $Cxcl12^{DsRed}$ allele shows that the SpA (arrow) develops in regions of *Cxcl12* expression. Arrowhead indicates expression in artery endothelial cells. Ao, aorta; LCA, left coronary artery; lv, left ventricle; RCA, right coronary artery; rv, right ventricle; sep, septum; trab, trabecular myocardium. Scale bars: A, C 400µm; G - I, 100µm

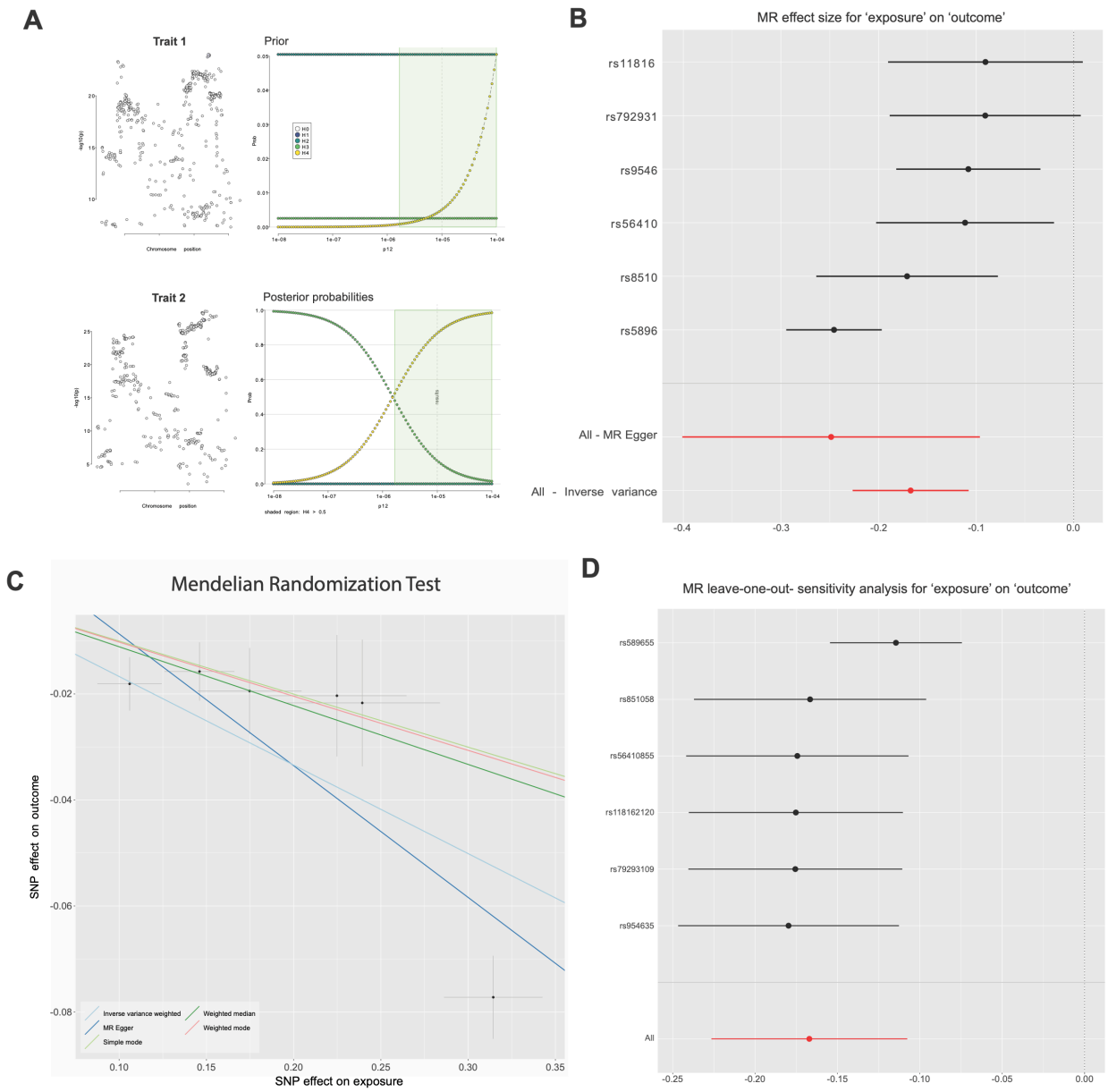

Figure 6: Colocalization of the *CXCL12* dominance locus with coronary artery disease and two sample Mendelian Randomization (MR) study. (**A**) Colocalization plot for the *CXCL12* for coronary artery disease (trait 1) and dominance (trait 2) showing a high probability of the same genetic variant being causal in both traits with a posterior probability Bayesian factor H4 =0.87. (**B**) Two sample MR forest plot using six genetic instruments from four dominance loci suggesting a causal association between left dominance and lower risk of CAD using standard inverse variance weighted meta-analysis. Results remained significant using MR-Egger regression sensitivity analysis. (**C**) Scatter plot performed across all MR methods tested in our study which were all significant. Estimated effect sizes of the genetic variants on both the exposure and outcome are displayed (black dots) (**D**) Leave one out sensitivity MR analysis demonstrating persistent significance.

References

1. Global Health Estimates 2019: Deaths by Cause, Age, Sex, by Country and by Region, 2000-2019. *World Health Organization* (2020).
2. Libby, P.*, et al. Braunwald's heart disease : a textbook of cardiovascular medicine*, (Elsevier, Philadelphia, 2021).
3. Baroldi, G., Scomazzoni, G. & United States. Surgeon-General's, O. *Coronary circulation in the normal and the pathologic heart*, (Office of the Surgeon General, Dept. of the Army; for sale by the Superintendent of Documents, U.S. Govt. Print. Off., Washington, 1967).
4. Yuan, R.*, et al.* Vascular endothelial growth factor gene transfer therapy for coronary artery disease: A systematic review and meta-analysis. *Cardiovasc Ther* **36**, e12461 (2018).
5. Rasanen, M.*, et al.* VEGF-B Promotes Endocardium-Derived Coronary Vessel Development and Cardiac Regeneration. *Circulation* **143**, 65-77 (2021).
6. Raftrey, B.*, et al.* Dach1 Extends Artery Networks and Protects Against Cardiac Injury. *Circ Res* **129**, 702-716 (2021).
7. Yla-Herttuala, S., Bridges, C., Katz, M.G. & Korpisalo, P. Angiogenic gene therapy in cardiovascular diseases: dream or vision? *Eur Heart J* **38**, 1365-1371 (2017).
8. Polderman, T.J.*, et al.* Meta-analysis of the heritability of human traits based on fifty years of twin studies. *Nature genetics* (2015).
9. Veltman, C.E.*, et al.* Influence of coronary vessel dominance on short- and long-term outcome in patients after ST-segment elevation myocardial infarction. *Eur Heart J* **36**, 1023-1030 (2015).
10. Parikh, N.I.*, et al.* Left and codominant coronary artery circulations are associated with higher in-hospital mortality among patients undergoing percutaneous coronary intervention for acute coronary syndromes: report From the National Cardiovascular Database Cath Percutaneous Coronary Intervention (CathPCI) Registry. *Circulation. Cardiovascular quality and outcomes* **5**, 775-782 (2012).
11. Veltman, C.E.*, et al.* Prognostic value of coronary vessel dominance in relation to significant coronary artery disease determined with non-invasive computed tomography coronary angiography. *Eur Heart J* **33**, 1367-1377 (2012).
12. Wang, L.*, et al.* Association between coronary dominance and acute inferior myocardial infarction: a matched, case-control study. *BMC cardiovascular disorders* **19**, 35 (2019).
13. Gebhard, C.*, et al.* Coronary dominance and prognosis in patients undergoing coronary computed tomographic angiography: results from the CONFIRM (COronary CT Angiography EvaluatioN For Clinical Outcomes: An InteRnational Multicenter) registry. *Eur Heart J Cardiovasc Imaging* **16**, 853-862 (2015).
14. Samani, N.J.*, et al.* Genomewide association analysis of coronary artery disease. *The New England journal of medicine* **357**, 443-453 (2007).
15. Zhou, W.*, et al.* Efficiently controlling for case-control imbalance and sample relatedness in large-scale genetic association studies. *Nature genetics* **50**, 1335-1341 (2018).
16. Watanabe, K., Taskesen, E., van Bochoven, A. & Posthuma, D. Functional mapping and annotation of genetic associations with FUMA. *Nature communications* **8**, 1826 (2017).

17. Harrison, M.R., *et al.* Chemokine-guided angiogenesis directs coronary vasculature formation in zebrafish. *Developmental cell* **33**, 442-454 (2015).
18. Cavallero, S., *et al.* CXCL12 Signaling Is Essential for Maturation of the Ventricular Coronary Endothelial Plexus and Establishment of Functional Coronary Circulation. *Developmental cell* **33**, 469-477 (2015).
19. Ivins, S., *et al.* The CXCL12/CXCR4 Axis Plays a Critical Role in Coronary Artery Development. *Developmental cell* **33**, 455-468 (2015).
20. D'Amato, G., *et al.* Endocardium-to-coronary artery differentiation during heart development and regeneration involves sequential roles of Bmp2 and Cxcl12/Cxcr4. *Developmental cell* **57**, 2517-2532 e2516 (2022).
21. Kircher, M., *et al.* A general framework for estimating the relative pathogenicity of human genetic variants. *Nature genetics* **46**, 310-315 (2014).
22. Rentzsch, P., Witten, D., Cooper, G.M., Shendure, J. & Kircher, M. CADD: predicting the deleteriousness of variants throughout the human genome. *Nucleic Acids Res* **47**, D886-D894 (2019).
23. Renier, N., *et al.* iDISCO: a simple, rapid method to immunolabel large tissue samples for volume imaging. *Cell* **159**, 896-910 (2014).
24. Rios Coronado, P.E. & Red-Horse, K. Enhancing cardiovascular research with whole-organ imaging. *Curr Opin Hematol* **28**, 214-220 (2021).
25. Icardo, J.M. & Colvee, E. Origin and course of the coronary arteries in normal mice and in iv/iv mice. *J Anat* **199**, 473-482 (2001).
26. Fernandez, B., *et al.* The coronary arteries of the C57BL/6 mouse strains: implications for comparison with mutant models. *J Anat* **212**, 12-18 (2008).
27. Lopez-Garcia, A., *et al.* Unusual anatomical origins of the coronary arteries in C57BL/6 mice. Are they strain-specific? *J Anat* **229**, 703-709 (2016).
28. Anbazhakan, S., *et al.* Blood flow modeling reveals improved collateral artery performance during the regenerative period in mammalian hearts. *Nat Cardiovasc Res* **1**, 775-790 (2022).
29. Ding, L. & Morrison, S.J. Haematopoietic stem cells and early lymphoid progenitors occupy distinct bone marrow niches. *Nature* **495**, 231-235 (2013).
30. Kolesova, H., Bartos, M., Hsieh, W.C., Olejnickova, V. & Sedmera, D. Novel approaches to study coronary vasculature development in mice. *Dev Dyn* **247**, 1018-1027 (2018).
31. Giambartolomei, C., *et al.* Bayesian test for colocalisation between pairs of genetic association studies using summary statistics. *PLoS Genet* **10**, e1004383 (2014).
32. Bulik-Sullivan, B.K., *et al.* LD Score regression distinguishes confounding from polygenicity in genome-wide association studies. *Nature genetics* **47**, 291-295 (2015).
33. Verma, A., *et al.* Diversity and Scale: Genetic Architecture of 2,068 Traits in the VA Million Veteran Program. (Cold Spring Harbor Laboratory, 2023).
34. Aragam, K.G., *et al.* Discovery and systematic characterization of risk variants and genes for coronary artery disease in over a million participants. *Nature genetics* **54**, 1803-1815 (2022).
35. Thiene, G., Frescura, C., Padalino, M., Basso, C. & Rizzo, S. Coronary Arteries: Normal Anatomy With Historical Notes and Embryology of Main Stems. *Front Cardiovasc Med* **8**, 649855 (2021).
36. Das, S., *et al.* A Unique Collateral Artery Development Program Promotes Neonatal Heart Regeneration. *Cell* **176**, 1128-1142 e1118 (2019).
37. Goldberg, A., *et al.* Coronary dominance and prognosis of patients with acute coronary syndrome. *Am Heart J* **154**, 1116-1122 (2007).

38. Bernabeu, E., *et al.* Sex differences in genetic architecture in the UK Biobank. *Nature genetics* **53**, 1283-1289 (2021).
39. Kuno, T., *et al.* Impact of coronary dominance on in-hospital outcomes after percutaneous coronary intervention in patients with acute coronary syndrome. *PLoS One* **8**, e72672 (2013).
40. Veltman, C.E., *et al.* Relation between coronary arterial dominance and left ventricular ejection fraction after ST-segment elevation acute myocardial infarction in patients having percutaneous coronary intervention. *Am J Cardiol* **114**, 1646-1650 (2014).
41. Mikaeilvand, A., *et al.* Association of coronary artery dominance and mortality rate and complications in patients with ST-segment elevation myocardial infarction treated with primary percutaneous coronary intervention. *J Res Med Sci* **25**, 107 (2020).
42. Selcuk, E., Cevirme, D. & Bugra, O. Prognostic Value of Coronary Dominance in Patients Undergoing Elective Coronary Artery Bypass Surgery. *Braz J Cardiovasc Surg* **35**, 452-458 (2020).
43. Vasheghani-Farahani, A., *et al.* The association between coronary arterial dominancy and extent of coronary artery disease in angiography and paraclinical studies. *Clin Anat* **21**, 519-523 (2008).
44. Yan, B., *et al.* Association of coronary dominance with the severity of coronary artery disease: a cross-sectional study in Shaanxi Province, China. *BMJ Open* **8**, e021292 (2018).
45. Azour, L., *et al.* Influence of coronary dominance on coronary artery calcification burden. *Clin Imaging* **77**, 283-286 (2021).
46. Tamargo, I.A., Baek, K.I., Kim, Y., Park, C. & Jo, H. Flow-induced reprogramming of endothelial cells in atherosclerosis. *Nat Rev Cardiol* **20**, 738-753 (2023).
47. Delaneau, O., Zagury, J.F., Robinson, M.R., Marchini, J.L. & Dermitzakis, E.T. Accurate, scalable and integrative haplotype estimation. *Nature communications* **10**, 5436 (2019).
48. Patterson, N., Price, A.L. & Reich, D. Population structure and eigenanalysis. *PLoS Genet* **2**, e190 (2006).
49. Evans, L.M., *et al.* Comparison of methods that use whole genome data to estimate the heritability and genetic architecture of complex traits. *Nature genetics* **50**, 737-745 (2018).
50. Visscher, P.M., *et al.* Statistical power to detect genetic (co)variance of complex traits using SNP data in unrelated samples. *PLoS Genet* **10**, e1004269 (2014).
51. Lee, S.H., Wray, N.R., Goddard, M.E. & Visscher, P.M. Estimating missing heritability for disease from genome-wide association studies. *Am J Hum Genet* **88**, 294-305 (2011).
52. Lee, S.H., *et al.* Estimating the proportion of variation in susceptibility to schizophrenia captured by common SNPs. *Nature genetics* **44**, 247-250 (2012).
53. Winkler, T.W., *et al.* Quality control and conduct of genome-wide association meta-analyses. *Nature protocols* **9**, 1192-1212 (2014).
54. Yuan, K., *et al.* Fine-mapping across diverse ancestries drives the discovery of putative causal variants underlying human complex traits and diseases. (Cold Spring Harbor Laboratory, 2023).
55. Wang, G., Sarkar, A., Carbonetto, P. & Stephens, M. A simple new approach to variable selection in regression, with application to genetic fine mapping. *Journal of the Royal Statistical Society Series B-Statistical Methodology* **82**, 1273-1300 (2020).
56. Nikpay, M., *et al.* A comprehensive 1,000 Genomes-based genome-wide association meta-analysis of coronary artery disease. *Nature genetics* **47**, 1121-1130 (2015).
57. Altshuler, D.M., *et al.* Integrating common and rare genetic variation in diverse human populations. *Nature* **467**, 52-58 (2010).

58. Verbanck, M., Chen, C.Y., Neale, B. & Do, R. Detection of widespread horizontal pleiotropy in causal relationships inferred from Mendelian randomization between complex traits and diseases. *Nature genetics* **50**, 693-698 (2018).
59. Burgess, S., Bowden, J., Fall, T., Ingelsson, E. & Thompson, S.G. Sensitivity Analyses for Robust Causal Inference from Mendelian Randomization Analyses with Multiple Genetic Variants. *Epidemiology* **28**, 30-42 (2017).
60. Hemani, G.*, et al.* The MR-Base platform supports systematic causal inference across the human phenome. *eLife* **7**, 29 (2018).
61. Renier, N.*, et al.* Mapping of Brain Activity by Automated Volume Analysis of Immediate Early Genes. *Cell* **165**, 1789-1802 (2016).
62. Roostalu, U.*, et al.* Effect of captopril on post-infarction remodelling visualized by light sheet microscopy and echocardiography. *Scientific reports* **11**, 5241 (2021).
63. Cunningham, F.G.*, et al.* Abortion. in *Williams Obstetrics, 25e* (McGraw-Hill Education, New York, NY, 2018).
64. Cunningham, F.G.*, et al.* Prenatal Care. in *Williams Obstetrics, 25e* (McGraw-Hill Education, New York, NY, 2018).