Endothelial cells (ECs) mediate several biological functions that are relevant to atherosclerosis and coronary artery disease (CAD), regulating an array of vital processes including vascular tone, wound healing, reactive oxygen species, shear stress response, and inflammation. Although which of these functions is linked causally with CAD development and/or progression is not yet known, genome-wide association studies have implicated more than 400 loci associated with CAD risk, among which several have shown EC-relevant functions. Given the arduous process of mechanistically interrogating single loci to CAD, high-throughput variant characterization methods, including pooled Clustered Regularly Interspaced Short Palindromic Repeats screens, offer exciting potential to rapidly accelerate the discovery of bona fide EC-relevant genetic loci. These discoveries in turn will broaden the therapeutic avenues for CAD beyond lipid lowering and behavioral risk modification to include EC-centric modalities of risk prevention and treatment.

Atherosclerosis is an indolent multicellular disease that culminates over the span of decades. Although plaque rupture and myocardial infarction typically manifest in middle to late adulthood, the clinical manifestations of atherosclerosis are preceded by a long asymptomatic phase of development that begins early in life, even in childhood. A hallmark postmortem study of Korean War trauma victims in their 20s revealed that 77.3% had coronary artery atherosclerosis. Several more recent studies have identified fatty streaks and fibrous plaques in the aorta and coronary arteries of many patients as young as 2 to 15 years of age.

The slow progression of atherosclerosis has confounded long-standing efforts to prevent it. Effective lipid-lowering therapies such as statins and PCSK9 inhibitors have reduced the prevalence of cardiovascular death over the past 30 years, and yet coronary artery disease (CAD)-associated morbidity and mortality has increased over the past decade. Since 2010, mortality attributed to cardiovascular disease in men and women in the United States has increased from 390,000 deaths to 490,000 deaths in 2020. In 2020, there were an estimated 19.05 million deaths from cardiovascular disease globally, which amounted to an 18.71% increase from 2010. These concerning trends suggest that new therapies for CAD are needed in addition to lipid-lowering therapies.

Defining the mechanisms underlying each stage of atherosclerosis will facilitate therapeutic efforts for primary and secondary disease prevention. Each pathologic stage of atherosclerotic plaque development has unique histologic features, and therefore distinct cellular profiles (Figure 1). Early atherosclerosis is characterized by early fatty streak development, and it begins as early as childhood, during which time low-density lipoprotein (LDL) cholesterol particles accumulate in the arterial intima. This accumulation...
triggers further endothelial inflammation, dysfunction, and secretion of chemoattractants for monocyte and lymphocyte recruitment. These fatty streaks disrupt laminar flow through vulnerable vascular regions, causing a turbulent hemodynamic profile at branch sites. Early microscopic changes in the intimal wall of the artery lead to macroscopic lipid accumulation and cellular dysfunction. Numerous studies have identified pathologic roles of macrophages and vascular smooth muscle cells (VSMCs) in this process. Once recruited to the arterial intima, monocytes phagocytose lipids and phenotypically become a foam cell. VSMCs proliferate and migrate into the atherosclerotic plaque, possibly via clonal expansion of distinct VSMCs adjacent to the developing lesion. This proliferation also triggers extracellular matrix accumulation and plaque mineralization. During plaque development the overlying fibrous cap is thick and intact. Over time and with continued cellular apoptosis, however, the fibrous cap thins and increases the likelihood of the final stage of atherosclerosis: plaque rupture. During this phase, the balance between collagen synthesis and degradation shifts to more extracellular matrix degradation. The thin cap can rupture abruptly, triggering thrombosis that leads to the clinical manifestation of macrovascular complications, including myocardial infarction.

Pathologic endothelial cell (EC) phenotypes have been linked to each of these stages of atherosclerotic plaque development. As the point of contact with blood, ECs respond to—and regulate—multiple atherogenic circulatory stimuli such as hyperlipidemia, hypertension, inflammation, and hyperglycemia. Although genetic studies have established a causal role of ECs in atherosclerosis and CAD, the extent of causal contributions of EC dysfunction remains incompletely understood, as does the precise pathologic cascade involving EC phenotypes in relation to other known processes concurrent with atherosclerosis.

This review summarizes the genetic variants that link EC dysfunction to CAD. Genome-wide association studies (GWAS) have implicated more than 400 loci in association with CAD, and many of these loci have EC-relevant functions; however, the biological characterization of these loci remains slow and laborious, requiring single-gene editing for mechanistic interrogation. New and emerging methods for high-throughput functional analysis can accelerate this process to identify causal pathways associated with vascular disease to establish new EC-specific mechanisms for the prevention and treatment of CAD.

The Role of ECs in Atherosclerosis

The vascular endothelium is a continuous layer of cells lined throughout the cardiovascular system and plays a critical role in maintaining normal, or quiescent, vascular function. Endothelial cells mediate vascular reactivity owing to their secretion of three vasoactive peptides: nitric oxide, endothelin-1, and prostacyclin. Abnormalities in the production of these peptides was the first reported evidence of EC dysfunction. Its definition has since grown to encompass all maladaptive changes in EC function that are linked to atherosclerosis, which now includes inflammatory cell recruitment, modification of lipoprotein particles, and the synthesis of extracellular matrix components in the fibrous cap of a plaque. Initial studies of balloon catheter—assisted intimal denudation in animal models supported a hypothesis that ECs are needed only to maintain intimal integrity, and that EC dysfunction is merely a response to mechanical injury. However, subsequent observations of early atherosclerosis in vessels lacking intimal injury have suggested that EC dysfunction is a multifactorial process. For example, atherosclerosis may originate in the adventitia or the vasa vasorum, and EC dysfunction can arise from the resultant inflammation. Although the ability to culture ECs in vitro has expedited research to identify the transcriptomic and biochemical features of EC dysfunction, therapeutic targeting of EC dysfunction constitutes an unmet need in preventive cardiology.

The first major pathologic process ascribed to ECs was its regulation of vascular tone. In 1980, it was discovered that vasodilation induced by acetylcholine requires an intact endothelium and is mediated by an unknown secreted factor, which initially was termed endothelium-derived relaxing factor. Work from several laboratories indicated that one of the first major pathologic process ascribed to ECs was its regulation of vascular tone. In 1980, it was discovered that vasodilation induced by acetylcholine requires an intact endothelium and is mediated by an unknown secreted factor, which initially was termed endothelium-derived relaxing factor. Work from several laboratories indicated that...

**Figure 1** Stages of atherosclerosis. A: Fatty streak. B: Fibrous plaque. C: Rupture and thrombosis. Scale bars = 500 μm.
that, within ECs, L-arginine is metabolized to form nitric oxide (NO) through an enzymatic process, the inhibition of which promotes atherogenesis in animal models.21 The gene responsible for this process is the endothelial isoform of NO synthase (NOS), which is one of three NOS isoforms that catalyze this reaction. Basal production maintains a quiescent, or nonthrombogenic, endothelium. The sole receptor for NO is soluble guanylate cyclase, found enriched within VSMCs, platelets, and leukocytes. The primary effect on VSMCs results in vasodilation, and therefore is a key regulator of systolic blood pressure, supported by the observation that endothelial NOS deficient mice develop refractory hypertension.22 Paradoxically, infusion of the NO stimulant acetylcholine in the coronary arteries of individuals with atherosclerotic disease produces vasoconstriction.20 These studies implicate the NO pathway in the early development of atherosclerosis and were among the first experiments in cultured endothelial cells. Although dysregulation of NO production and signaling initially was considered the sole determinant of EC dysfunction, its definition has since expanded with improved understanding of EC biology (Figure 2).10

The EC-dependent regulation of inflammation represents a second major driver of atherosclerosis. Both innate and adaptive immune responses have been implicated in atherosclerosis and plaque progression. The innate immune response does not require antigenic stimulation and is triggered largely by inflammatory cytokines in the NLR family pyrin domain containing 3 (NLRP3) inflammasome pathway.23 Perhaps in response to oxidized lipoproteins or other pathogen-associated molecular patterns, IL-1 production is augmented. New data show that clonally expanded hematopoietic cells, termed clonal hematopoiesis of indeterminate potential, produce IL-1b.24 ECs also directly secrete IL-1b and other atherosclerosis-inducing cytokines such as monocyte chemoattractant protein-1 (MCP-1) and IL-8. The subsequent inflammatory cascade induces increased expression of endothelial—leukocyte adhesion molecules, such as endothelial—leukocyte adhesion molecule-1 (ELAM-1) and vascular cell adhesion molecule-1 (VCAM-1). ECs that express these ligands actively transport leukocytes across the endothelial barrier, which exacerbates inflammation and plaque progression.

The endothelial lining forms a point of contact with blood, and thus its response to hemodynamic forces is an essential EC function that is dysregulated in the context of atherosclerosis. Exposure of ECs to disturbed flow induces EC turnover and senescence, increases oxidative stress, and disrupts barrier function. Genes important in thrombosis, inflammation, and vascular tone all are regulated by the flow

Figure 2 Genetic risk variants affect different endothelial cell (EC) functions that contribute to atherosclerotic risk. ARHGEF26, Rho guanine nucleotide exchange factor 26; CAD, coronary artery disease; EDNRA, endothelin receptor type A; eNOS, endothelial nitric oxide synthase; FLT1, Fms related receptor tyrosine kinase 1; JCAD, junctional cadherin 5—associated; KLF2/4, Krüppel-like factor 2/4; MFGEB, milk fat globule-EGF factor 8; sGC, soluble guanylyl cyclase; VSMC, vascular smooth muscle cell.
of blood across the ECs.\footnote{25} A master regulator of such an atheroprotective gene expression program is KLF2. KLF2 is among the genes most responsive to hemodynamic stimuli and is up-regulated in atheroresistant regions of human arteries. It stimulates the production of NO and inhibits expression of the potent vasoconstrictive peptide endothelin-1. As the most widely prescribed drugs for the prevention of atherosclerosis, statins up-regulate KLF2 expression in cultured human endothelial cells.\footnote{26} It is now postulated that one component for the large protective effect of statins on CAD and/or myocardial infarction is through the induction of KLF2.

Recent single-cell RNA-sequencing studies have identified EC heterogeneity as an additional variable to account for within the study of atherosclerosis. For example, three distinct EC subtypes were identified in the mouse aorta.\footnote{27} The differentially expressed genes suggest functional specialization of different EC subtypes in atherosclerosis-relevant processes such as lipid handling, angiogenesis, and extracellular matrix gene expression. The distinct mechanism(s) by which these EC subtypes change in response to atherogenic risk factors such as hyperlipidemia and oscillatory shear stress remains unknown. These early single-cell studies suggested that the hallmarks of EC dysfunction may be different in each EC subtype.

### Mendelian Randomization as a Way to Establish a Causal Risk Factor

The importance of ECs in regulating vascular tone, inflammation, and shear stress response are just a few of the functions that implicate its role in atherosclerosis. ECs mediate numerous critical functions including endothelial-to-mesenchymal transition, generation of reactive oxygen species, regulation of platelet activation, and thrombosis/hemostasis. Some of these pathways may be activated by CAD risk factors, and do not contribute causally to the progression of early atherosclerosis. Separating the causal EC-relevant pathways would help identify the most critical steps in the progression of atherosclerosis and introduce new drug targets that act independently from classic lipid-lowering medications.

Human genetics can help link certain EC functions to causal human pathways. Mendelian randomization has become increasingly popular in recent years as a method to infer causality in epidemiologic research. It can be used to investigate the causal relationship between a wide range of modifiable risk factors and disease outcomes, including lifestyle factors, such as smoking and alcohol consumption, and biomarkers, such as blood pressure and cholesterol levels.\footnote{27} The rationale behind Mendelian randomization is that genetic variants are allocated randomly at conception and are not subject to confounding or reverse causation. Therefore, if a genetic variant is associated with both a risk factor and a disease outcome, this risk factor causally may influence the disease outcome.

In a Mendelian randomization study, genetic variants known to be associated with a particular risk factor are used as instrumental variables to estimate the causal effect of the risk factor on the outcome of interest. This approach has been successful for biological pathways with a large number of genetic variants that have been characterized functionally. For instance, LDL-cholesterol is a \textit{bona fide} causal mechanism for atherosclerosis because a large number of genetic variants associated with LDL-cholesterol levels also are associated strongly with CAD/myocardial infarction risk.\footnote{28} Conversely, high-density lipoprotein—cholesterol has a similar number of genetically linked variants, but no association with CAD/myocardial infarction.\footnote{29} Therefore, high-density lipoprotein—cholesterol fails the Mendelian randomization test for biologic causality. Extending Mendelian randomization studies to identify novel risk factors, such as those that implicate EC-relevant functions, would be a major advance in the field. However, this is not feasible because only a few genetic variants are linked to endothelial cell function. The challenge is that most vascular disease—associated genetic variants are in the noncoding genome, so the genes that they regulate are not established. Characterization of individual variants over the past decade has identified several with EC-relevant functions. These variants begin to point to causal pathways for atherosclerosis such as regulation of vascular tone and EC barrier function. With the advent of high-throughput variant characterization methods, the number of functionally annotated genetic variants will increase, dramatically improving the power to link human genetics to causal biological pathways.

### Genetic Association Studies for CAD

The strongest evidence for a disease-relevant gene is when a pathogenic coding mutation is identified. This has been the case for familial hypercholesterolemia and pathogenic mutations in \textit{LDLR}, \textit{PCSK9}, and \textit{APOB}. For each gene there are numerous loss-of-function mutations in the general population that are associated robustly with increased LDL-cholesterol levels. Unfortunately, loss-of-function variants are comparatively rare in EC-relevant genes. This is likely because ECs play a central role in several developmental processes (ie, neural crest development, endocardial development, and vascular patterning), and therefore loss-of-function is incompatible with survival. To address this challenge, rare variant association studies perform sequencing in large numbers of diseased and control individuals. Rare genetic variants are DNA sequence variations that occur at low frequencies in a population. Genes with a greater burden of mutations in the diseased individuals then are implicated in the disease. Burden testing is a popular method for analyzing rare genetic variants because it increases statistical power by aggregating multiple rare variants based on their predicted functional relatedness. It
also reduces the challenge of multiple testing by interrogating a single measure of genetic burden instead of individual variants.

To date, burden testing for rare variants in CAD has implicated lipid genes. The Exome Sequencing Project used burden testing to analyze exome sequencing data from more than 12,000 individuals with and without cardiovascular disease. They identified a statistically significant burden for mutations in LDLR and APOA5, but no novel nonlipid genes. More recently, a larger study of 41,081 CAD cases versus 217,115 controls identified loss-of-function or damaging missense variants in NOS3 in 0.6% of cases. Carriers of these NOS3 loss-of-function mutations have a 2.4-fold higher probability of CAD, as well as higher systolic blood pressure. This suggests that NO synthesis is a key nonlipid driver of CAD risk. Perhaps larger rare variant association studies will implicate more EC-relevant genes for CAD. However, these studies are limited by the low number of pathogenic mutations in nonlipid genes.

**Common Variant Association Studies Identify EC-Relevant Variants for CAD/Myocardial Infarction**

Compared with rare variant association studies, common variant association studies have been more successful at linking human genetic variation to disease: GWAS have implicated more than 350 genomic loci in relation to CAD. In addition, the statistical power gained from meta-analyses has slowly increased the number of associated loci from 3 in 2007 to more than 350 as of 2022 with the inclusion of large biobanks such as the UK Biobank and the Million Veterans Program. These studies now identify an increasing proportion of the heritability for CAD. The recent meta-analysis from the Design of the Coronary Artery Disease Genome-Wide Replication And Meta-Analysis (CARDIOGRAM) consortium identified 15.5% of CAD heritability from the 241 genome-wise significant loci, and 36.1% heritability when including all 897 loci with $P < 2.52 \times 10^{-5}$. One unexpected finding was that 72% of these loci are not associated with circulating LDL or total cholesterol, and 60% are not associated with either cholesterol or hypertension.

Each of the recent GWAS meta-analyses had more than 1 million cases and controls and identified new loci associated with CAD risk. The largest number of novel loci were identified by the multiethnic analysis from the Million Veteran Program. They integrated data from White, Black, and Hispanic individuals with existing studies to identify 95 novel loci from a total population of 243,392 cases and 849,686 controls. The authors then used computational enrichment methods to prioritize relevant cell types. Interestingly, the gene expression profiles of ECs showed the greatest enrichment for CAD genetic risk in multiple analyses and from multiple EC tissues, including fat stromal ECs and heart ECs.

A second large meta-analysis was published recently by the CARDIOGRAM consortium that included 181,522 cases among 1,165,690 total participants of predominantly European ancestry and identified 30 novel loci associated with CAD. Two computational tools were used to identify the biological pathways that explain these genetic associations. First, a machine-learning algorithm called the Polygenic Priority Score was used to identify the Gene Ontology terms that link multiple loci. As expected, this method successfully prioritized known cholesterol and lipoprotein metabolism pathways. However, several other vascular cell functions were also associated significantly with CAD genetic risk, including transforming growth factor β signaling, vascular cell motility, extracellular matrix function, and cytoskeletal regulation. The second method to prioritize biological mechanisms was epigenetic enrichment of CAD-lead variants in candidate cell types. A total of 127 regions had significant enrichment in at least one tissue type. Twenty-two of these variants showed enrichment in human umbilical vein EC epigenetic data, and another 12 showed enrichment in the aorta. These variants likely regulate vascular cell function, and in many cases affect EC biology.

Taken together, these GWAS studies for CAD exemplify that meta-analyses can identify CAD-relevant genomic loci that can be analyzed for prioritized biological pathways to implicate causal cell types. These prioritization methods consistently link CAD-associated genetic variation to EC-relevant functions.

**Evidence of EC-Acting Variants Associated with CAD**

GWAS of complex traits such as CAD have implicated numerous single-nucleotide polymorphisms reaching genome-wide significance ($P < 5 \times 10^{-8}$). However, validating their causal contribution via variant-to-function studies of atherogenesis remains a formidable challenge. In addition, the very large number of candidate studies obscures the unbiased process of discovering true causal variants. Nevertheless, a combination of computational fine-mapping and mechanistic in vitro and in vivo animal studies have revealed genetic variants with bona fide functional relevance in CAD (Table 1).

Among these, the population-wide genetic variation of NOS3, the gene encoding constitutively expressed endothelial NOS, is among the most well-reputed causal genes found using candidate-based approaches owing to its established role in maintaining vascular endothelial homeostasis. Located on chromosome 7p35–36, NOS3 comprises 26 exons and spans 21 kb. Hingorani et al first implicated its role in CAD by uncovering a point mutation in exon 7 of the NOS3 gene (Glu298Asp, G894-T, rs1799983) via single-strand conformation polymorphism analysis that later was found to reduce NOS3 levels and,
consequently, impair both basal and shear stress–dependent NO release. 

Subsequent studies have identified other NOS3 single-nucleotide polymorphisms, which have been incorporated into meta-analyses. However, in 2015 a Flemish-based GWAS finally showed that a promoter-associated -665 C>T (rs3918226) variant predicts cardiovascular mortality independent of other CAD risk factors. 

Additional genome-wide studies of large, multiethnic populations have uncovered similarly strong associations of this rs3918226 variant. 

Other CAD-associated variants also have been found to promote atherogenesis by disrupting NO homeostasis, including the 4q32.1 locus, which was discovered using a meta-analysis of the Han Chinese population. The implicated gene, which encodes the α subunit of soluble guanylate cyclase 1 (GUCY1A3), catalyzes the generation of cyclic guanosine monophosphate (cGMP) pathway signaling within VSMCs and protein kinase G (PKG)—mediated activation of myosin phosphatase leading to VSMC relaxation. Genetic mutations affecting the GUCY1A3 locus lead to accelerated thrombus formation after microcirculatory trauma. Specifically, the risk allele at this locus impairs the binding of transcriptional regulator zinc finger E-box binding homeobox 1 (ZEB1), consequently reducing GUCY1A3 promoter activity and dysregulating both VSMC migration and platelet aggregation.

Table 1  CAD-Associated Genetic Variants With Effects in Endothelial Cells

<table>
<thead>
<tr>
<th>Variant ID</th>
<th>Associated gene(s)</th>
<th>Genomic locus</th>
<th>EC-mediated consequences</th>
<th>Reference</th>
<th>GWAS</th>
</tr>
</thead>
<tbody>
<tr>
<td>rs1799983</td>
<td>NOS3</td>
<td>7q35</td>
<td>Disruption of endothelial nitric oxide—mediated smooth muscle relaxation, increased thrombogenesis</td>
<td>37, 38–42</td>
<td></td>
</tr>
<tr>
<td>rs7692387</td>
<td>GUCY1A3</td>
<td>4q32.1</td>
<td>Disruption of soluble guanylate cyclase in platelets leads to a prothrombotic state</td>
<td>43, 44, 38–41, 45, 46</td>
<td></td>
</tr>
<tr>
<td>rs9349379</td>
<td>PHACTR1 EDN1</td>
<td>6p24</td>
<td>Endothelin 1—mediated vasoconstriction</td>
<td>47, 38, 40–42, 45, 48, 49</td>
<td></td>
</tr>
<tr>
<td>rs74412485</td>
<td>FLT1</td>
<td>9p21</td>
<td>Promotes the expression of inflammatory adhesion molecules</td>
<td>50, 50, 51</td>
<td></td>
</tr>
<tr>
<td>rs1807214</td>
<td>MFGE8</td>
<td>15q26</td>
<td>Integrin-binding glycoprotein that promotes proinflammatory signaling via ABHD2</td>
<td>34, 52, 53, 41</td>
<td></td>
</tr>
<tr>
<td>rs2487928</td>
<td>JCAD</td>
<td>10p11</td>
<td>Gain-of-function mutation causing HIPP0-mediated proinflammatory endothelial signaling, leading to vascular hyper-reactivity and thrombogenesis</td>
<td>54–57, 38, 39, 41, 42, 45</td>
<td></td>
</tr>
<tr>
<td>rs17114036</td>
<td>KLF2</td>
<td>1p32.2</td>
<td>Gain-of-function shear stress—dependent regulation of PLPP3, leading to ICAM1 expression and leukocyte transmigration</td>
<td>58, 59, 38, 45, 46, 49, 60, 61</td>
<td></td>
</tr>
<tr>
<td>rs17114036</td>
<td>PLPP3</td>
<td>1p32.2</td>
<td>Disrupts CEBP-β motif, leading to a proinflammatory phenotype and impaired vascular barrier integrity</td>
<td>62, 63, 38, 45, 46, 60, 61</td>
<td></td>
</tr>
<tr>
<td>rs12493885</td>
<td>ARHGEF26</td>
<td>3q25.2</td>
<td>Gain-of-function mutation drives endothelial dysfunction via VEGF-mediated angiogenesis</td>
<td>64, 40, 42, 49, 60</td>
<td></td>
</tr>
</tbody>
</table>

Biologically relevant causal variants found using fine mapping of the 241 conditionally independent variants associated with CAD by Aragam et al. that reached genome-wise significance in the primary meta-analysis. 

CAD, coronary artery disease; CEBP-β, CCAAT/enhancer-binding protein beta; EC, endothelial cell; GWAS, genome-wide association studies; ICAM1, intercellular adhesion molecule 1; NOS, nitric oxide synthase; PLPP3, phospholipid phosphatase 3; VEGF, vascular endothelial growth factor.
quantitative trait loci or eQTLs) lead to many false positives. Additional variants found within the EDNRA gene also are associated with CAD risk, further implicating endothelin signaling in the pathogenesis of CAD.

Genetic predisposition to CAD has also been linked causally to endothelial proinflammatory signaling. As an example, multietnic population-wide studies have identified several single-nucleotide polymorphisms within the gene encoding vascular endothelial growth factor receptor-1 (FLT1). FLT1 is a high-affinity tyrosine kinase receptor for vascular endothelial growth factor receptor-1 that promotes angiogenesis and vascular smooth muscle proliferation, a loss of which causes increased susceptibility to ischemia reperfusion injury. A GWAS within a Korean population validated two variants within this gene, rs9508025 and rs1333049. A follow-up study in this sample Korean cohort, however, found prognostic association of only rs9319428 with CAD.51 A follow-up GWAS in a Japanese population that validated two variants within this gene, rs9319428 also found rs74412485. In this study, Konta et al discovered that, within its first intron, rs74412485 leads to a gain-of-function effect on FLT1 expression, thereby promoting the expression of inflammatory adhesion molecules likely underlying CAD susceptibility patients harboring this variant.

Another risk locus that has been studied in-depth is the rs1807214 locus, which exists at an intergenic region in 15q26.1. Before polygenic prioritization methods, studies had assigned causal significance of rs1807214 to that of the ABHD2 gene given its proximity at 65 kb downstream of the sentinel variant. However, the rs1807214 variant was subsequently linked to MFGES8, which encodes an integrin-binding glycoprotein located 108 kb upstream. MFGES8 is implicated in VSMC proliferation and proinflammatory signaling, but, more recently, in vitro clustered regularly interspaced short palindromic repeats (CRISPR)—Cas9 disruption of this risk locus was shown to increase MFGES8 expression without altering ABHD2 expression. Therefore, the rs1807214 CAD locus confers CAD risk through disinhibition of ABHD2.

Vascular shear stress—associated signaling is another facet of genetic risk identified via GWAS. Several CAD-associated variants have been identified in association with genetic variants at the 10p11 locus, all of which are within the gene encoding KIAA1462, or, more suitably named, JCAD. Unlike most causal CAD variants, however, the three found within JCAD (rs3739998, rs2505083, and rs2487928) confer relative protection at a genome-wide level of significance. As a key protein involved in vascular integrity via maintenance of endothelial cell—cell junctions, junctional cadherin 5—associated (JCAD) transduces a shear stress—dependent proinflammatory response on the endothelium. Functional analysis of the lead variant rs2487928 in cultured ECs reduces total JCAD levels and suppress the Hippo pathway, which has downstream consequences on a number of downstream phenotypes, including angiogenesis. In vivo studies found that floxed mice harboring an inducible Jcad knock-down backcrossed into an ApoE background were protected against vascular inflammation, vascular hyperreactivity, thrombogenesis, and plaque formation after high-fat feeding.

ECs are activated by disturbed flow in arterial regions prone to atherosclerosis, wherein mechanosensing endothelial transcriptional regulators, including KLF2 and KLF4, have been implicated in genetic predisposition to shear-responsive CAD. The rs17114036 variant is located at the 1p32.2 noncoding region within a hemodynamically regulated endothelial enhancer. CRISPR—Cas9—based disruption has shown that this rs17114036-containing region promotes endothelial quiescence under unidirectional shear stress in part by regulating PLPP3. In vitro experiments have shown that it facilitates leukocyte transmigration via ICAM1 complexing. Klarin et al suggested a gain-of-function phenotype for the causative mutation in humans, which could account for the increased CAD susceptibility in risk allele carriers.

Pooled CRISPR Screens: The Future

Despite the slow progress made by mechanistic single-variant studies, emerging methods that allow high-throughput functional characterization of disease-associated genetic variants will likely accelerate the discovery of novel EC-relevant loci. Each of the examples highlighted in this review demonstrate that successful variant-to-function in ECs required focus on an individual locus, and thus could not identify the shared pathways regulated by many loci. The recent use of gene editing technology to design large-scale, pooled assays for variant characterization finally provide an exciting platform to define the common genetic pathways regulated by multiple CAD risk variants.

Several groups have developed CRISPR-based screens to perturb genes or noncoding regions of the genome systematically to study their function. In a CRISPR perturbation screen, a library of guide RNAs was designed to target specific genes or genomic regions of interest. These guide RNAs are subsequently delivered into cells using viral vectors or other methods. The guide RNAs guide the Cas9 enzyme to cut the DNA at the target site, inducing a DNA repair response that can result in a variety of outcomes, including gene knockout, gene activation, or gene repression.

The effects of these perturbations are measured using high-throughput assays, such as RNA sequencing, DNA sequencing, or cell-based phenotypic screens. By systematically perturbing genes or genomic regions, CRISPR perturbation screens can identify functional relationships between genes and pathways, reveal new therapeutic targets, and provide insights into the underlying biology of complex diseases. The output of these screens is a ranked list of genes that affect cellular function. The first example of a genome-wide CRISPR screen was conducted for the genetic determinants of chemotherapeutic resistance in a melanoma
cell line. Lentiviral delivery of 64,751 guide sequences to A375 melanoma cells identified genes that conferred resistance to vemurafenib.²⁸

CRISPR screens for cancer cell resistance have the advantage of sequencing the cells that survive despite exposure to the drug because these contain the single guide RNA (sgRNA) sequences that conferred resistance. Screening for genes that affect EC function, however, is more complicated because there is no obvious disease-relevant phenotype. To address this challenge, some screens have assayed for specific EC functions and sorting for extremes of the distribution. For example, a screen for novel genes that affect atherosclerosis specifically assayed for genes that regulate low density lipoprotein receptor (LDLR)–independent LDL uptake in ECs.²⁹ With an RNA-interference library targeting approximately 18,000 genes, this screen identified activin A receptor like type 1 (ACVRL1/ALK1) as a key mediator of LDL transcytosis across the endothelium.

CRISPR-mediated knockdown has the potential to expand the capacity and efficacy of these screens for atherosclerosis-relevant genes in ECs. One recent example assessed the effect of 2000 potential causal variants in proximity to 83 CAD GWAS risk loci on six EC phenotypes.³² The study used three kinds of CRISPR perturbations: cutting Cas9 (to achieve gene knockout), CRISPR interference (to reduce gene expression), and CRISPR activation (to increase gene expression). The molecular phenotypes assessed were expression of E-selectin, VCAM1, or ICAM1, and levels of the signaling molecules reactive oxygen species, NO or Ca²⁺. They found variants in 26 loci that affected at least one of these measures, potentially linking these loci to specific vascular and/or inflammatory mechanisms.

Phenotype-specific perturbation screens are limited in that they can only identify genes affecting the specific phenotypes examined. One unbiased approach to phenotyping is Perturb-seq, a pooled CRISPR screen in which the phenotype examined is the entire transcriptome. In this approach, cells engineered to contain CRISPR machinery are transduced with a guide RNA library and subjected to single-cell RNA sequencing in a way that connects the guide each cell received to its transcriptome. Perturb-seq has been used to discover novel aspects of unfolded protein response pathways,³⁰ T-cell activation,³¹ and genetic interaction mapping, on a scale from a few dozen genes to several thousand.³²⁻³⁴

We recently performed Perturb-seq, using CRISPR interference in telomerase immortalized human aortic endothelial cells, measuring the transcriptomic effects of knocking down 2285 target genes (37,637 guides), including all genes within a megabase of the lead variant in 241 CAD GWAS loci.³⁵ Machine learning methods were used, specifically non-negative matrix factorization, to identify 50 transcriptional programs in ECs, including both 37 non–cell-type–specific and 13 EC-specific biological programs, as well as the effects of each CRISPR-based genetic perturbation on these cascades. The 5 programs that had the highest contribution to CAD heritability by intersecting variant-to-gene and Perturb-seq gene-to-program information were prioritized. All five of these CAD-relevant programs are regulated similarly by perturbations to genes in the cerebral cavernous malformation (CCM) signaling pathway. CCM is a disease caused by the loss-of-function mutations in any of the three CCM genes (CCM2, KRIT1, and PDCD10).³⁶ Proteins encoded by these form a complex that integrates EC contact and vascular flow information to modulate downstream signaling, including inhibition of a kinase cascade that activates KLF2, a key flow-sensing transcription factor. The loss of CCM complex function leads to dysregulation of microvascular ECs, which then give rise to brain vascular lesions. The results suggest that common variants that have different effects on CCM function are protective for CAD. These results from the first large-scale, unbiased CRISPR screen in ECs illustrate its power to link CAD genetic variants to a common pathway. Early data suggest that several CAD GWAS loci regulate the same shear stress response pathway already linked to a rare cerebrovascular disease.

Conclusion

Decades of groundbreaking, large-scale, association studies coupled with mechanistic experiments have established a critical role of ECs in vascular disease. Starting with their importance in regulating vascular tone with secreted vasoactive molecules such as NO and endothelin-1, ECs have proven to be far more relevant to CAD physiology than simply lining the blood vessel wall. Variant-to-function studies in ECs have linked the genetic regulation of a few genes to CAD risk. The number of associated genes remains fewer than 10, but some pathways are already showing enrichment. The regulation of vascular tone and EC barrier function seem to be associated causally with CAD risk given the genetic variants in these pathways. With high-throughput variant editing methods, such as CRISPR-based screens, the number of biologically characterized loci will increase quickly. Ultimately, the insights from genetic association studies will have defined the causal pathways for multiple vascular diseases, and guide the development of new therapies that target the cells of the blood vessel wall.

Disclosure Statement

None declared.

References

2. Urbina EM, Williams RV, Alpert BS, Collins RT, Daniels SR, Hayman L, Jacobson M, Mahoney L, Mietus-Snyder M, Rocchini A,


10. 10. G prostaglandins (PGH2) in atherosclerosis in the hypercholesterolemic rabbit thoracic aorta. PGH2 does not contribute to impaired endothelium-dependent relaxation. Arterioscler Thromb 1994, 14:746–752


43. Mauersberger C, Sager HB, Wobst J, Dang TA, Lambrecht L, Erdmann J, Stark K, Esslinger UB, Rumpf PM, Koesling D, de...


