Characterization of caffeine response regulatory variants in vascular endothelial cells

Carly Boye [1], Cynthia Kalita [1], Anthony Findley [1], Adnan Alazizi [1], Julong Wei[1], Xiaoquan Wen[2], Roger Pique-Regi [1,3], Francesca Luca [1,3,4]

[1]Center for Molecular Medicine and Genetics, [2]Department of Biostatistics, University of Michigan, Ann Arbor, MI, [3]Department of Obstetrics and Gynecology, Wayne State University, Detroit, MI, [4]Department of Biology, University of Rome Tor Vergata, Rome, Italy

Abstract

Genetic variants in gene regulatory sequences can modify gene expression and mediate the molecular response to environmental stimuli. In addition, genotype-environment interactions (GxE) contribute to complex traits such as cardiovascular disease. Caffeine is the most widely consumed stimulant and is known to produce a vascular response. To investigate GxE for caffeine, we treated vascular endothelial cells with caffeine and used a massively parallel reporter assay to measure allelic effects on gene regulation for over 43,000 genetic variants. We identified 665 variants with allelic effects on gene regulation and 29 variants that regulate the gene expression response to caffeine (GxE, FDR<10%). When overlapping our GxE results with eQTLs colocalized with CAD and hypertension, we dissected their regulatory mechanisms and showed a modulatory role for caffeine. Our results demonstrate that massively parallel reporter assay is a powerful approach to identify and molecularly characterize GxE in the specific context of caffeine consumption.

Introduction

Caffeine is the most widely consumed stimulant in the world(Planning Committee for a Workshop on Potential Health Hazards Associated with Consumption of Caffeine in Food and Dietary Supplements et al., 2014). Caffeine produces a vascular response in the endothelium, causing vaso-dilation. The vascular endothelium, the innermost layer of arteries, is involved in several important functions, including regulation of blood flow, angiogenesis, thrombosis, and coagulation(Hadi et al., 2005; Krüger-Genge et al., 2019). Endothelial dysfunction occurs in diseases such as atherosclerosis and hypertension(Xu et al., 2021), eventually leading to CAD(Matsuzawa and Lerman, 2014) Multiple studies have investigated the role of caffeine in cardiovascular disease (CVD), and more broadly, vascular health in general, with conflicting results(Chieng et al., 2022; Ding et al., 2014; Turnbull et al., 2017) on the role of caffeine in CVD risk. Ding *et al.* (2014)

meta-analyzed 36 studies and found no association between heavy coffee consumption and increased risk of CVD(Ding et al., 2014). Similarly, Turnbull *et al.* (2017) observed that moderate caffeine consumption was not associated with an increased risk of CVD, or other cardiovascular events such as heart failure(Turnbull et al., 2017). Multiple studies suggested that caffeine may be beneficial in reducing the risk of coronary artery disease (CAD)(Choi et al., 2015; Miranda et al., 2018; Voskoboinik et al., 2019), while others provided evidence that caffeine may reduce risk of heart failure, but had no significant effect on the risk of coronary heart disease or CVD(Stevens et al., 2021). Most recently, Chieng *et al.* (2022) found that decaffeinated, ground, and instant coffee significantly decreased CVD risk and mortality(Chieng et al., 2022). The conflicting results from these epidemiological studies may have several causes, including potential interactions between caffeine consumption and other environmental and genetic risk factors. Recent molecular studies investigated the consequences of caffeine exposure on chromatin accessibility and gene expression in vascular endothelial cells(Findley et al., 2019). This study identified response factor motifs for caffeine, defined as transcription factor motifs that are enriched in differentially accessible regions, and demonstrated that caffeine can induce changes in gene regulation in endothelial cells.

Analyzing the changes in gene expression upon exposure to environmental stimuli is a powerful approach to discover genotype-environment interactions (GxE). These molecular GxE result in a different response depending on genotype(Knowles et al., 2018, 2017; Moyerbrailean et al., 2016b), potentially through allele-specific effects on response factor binding or other environmental-specific gene regulatory mechanisms. Yet regulatory sequences that are differentially bound in response to environmental perturbations are poorly annotated. Single nucleotide polymorphisms (SNPs) within caffeine response factor binding sites were enriched for artery expression quantitative trait loci (eQTLs) colocalized with CAD risk variants(Findley et al., 2019). The results of this study thus suggested that SNPs within regulatory elements active in the presence of caffeine may play a role in CAD risk and pointed to gene-environment interactions in gene regulation as a potential mechanism underlying caffeine modulation of genetic risk for CAD. However, only a limited number of molecular GxE for caffeine have been studied so far, thus the transcription factors and regulatory sequences involved in caffeine GxE remain uncharacterized. Furthermore, it is important to study GxE in the relevant cell type; i.e. endothelial cells which constitute the vascular endothelium. For these reasons, it is crucial to investigate and validate the mechanisms behind caffeine GxE in vascular endothelial cells.

Massively parallel reporter assays (MPRA) have allowed studies of non-coding genetic variants and their role in gene regulation, at unprecedented scale(Arnold et al., 2013; Gordon et al., 2020; Kalita et al., 2018; Melnikov et al., 2012; Patwardhan et al., 2012; Tewhey et al., 2016; Ulirsch et al., 2016; Vockley et al., 2015; Wang et al., 2018). Originally developed to study the gene regulatory potential of promoters and enhancer sequences, MPRA protocols have been further developed to study regulatory genetic variation and fine map association signals(Kalita et al., 2018; Tewhey et al., 2016; Ulirsch et al., 2016; Vockley et al., 2015). MPRAs with synthetic regulatory sequences can test allelic activity for candidate regulatory variants independently of their allele frequency in the population(Kalita et al., 2018; Tewhey et al., 2016; Ulirsch et al., 2016; Vockley et al., 2015). In MPRAs, DNA sequences containing each allele are transfected into cells and RNA-Seq is used to quantify the transcripts for each allele. To directly test allelic effects of tens of thousands of candidate regulatory sequences predicted to affect transcription factor binding (CentiSNPs(Moyerbrailean et al., 2016a)), we used an MPRA called Biallelic Targeted STARR-Seq (BiT-STARR-Seq)(Kalita et al., 2018). Only two previous studies have used MPRAs to

investigate DNA sequences that regulate the transcriptional response to treatments(Johnson et al., 2018; Shlyueva et al., 2014). One study utilized STARR-Seq to characterize enhancer activity in Drosophila cells upon treatment with the hormone ecdysone, however it did not investigate genotype-environment interactions (GxE)(Shlyueva et al., 2014). The other study utilized STARR-Seq to investigate the response to glucocorticoid treatment in the human cell line A549(Johnson et al., 2018). Although this study investigated GxE interactions, only a small number of variants were tested as this study was limited to pre-existing variation within the samples (as opposed to designed target sequences), and only 2 variants had significant GxE(Johnson et al., 2018). We aim to identify and validate the DNA sequences that regulate the transcriptional response to caffeine in the vascular endothelium, and how genetic variation present in these regulatory elements may affect the transcriptional response to caffeine (figure 1 A).

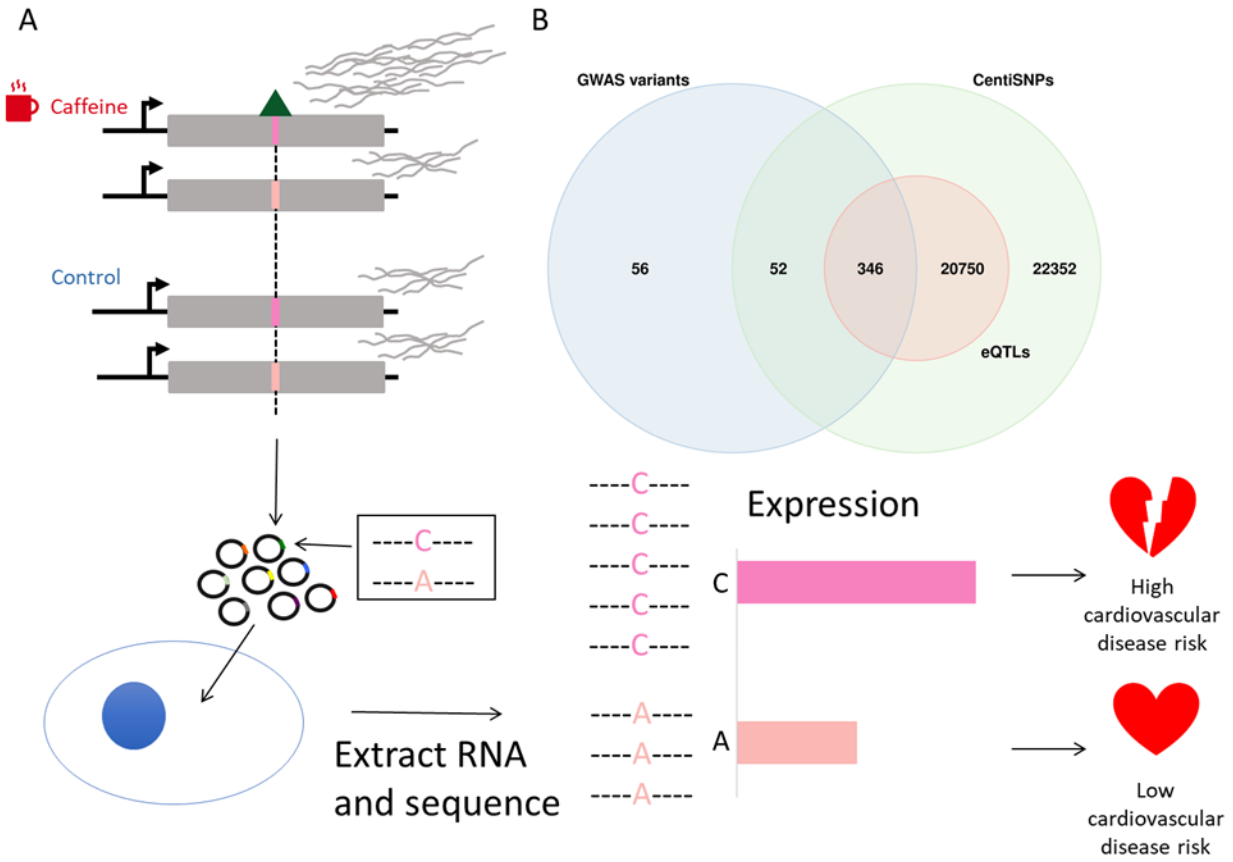

Figure 1. Study design. (A) Genetic variants modulate transcriptional response dependent on environmental conditions. The pink bars represent different alleles present in our targets, and the green triangle represents a bound transcription factor. These target sequences are transfected into cells, RNA is extracted and sequenced, and then activity is measured for targets for both alleles. (B) Library composition based on annotation category: SNPs predicted to alter transcription factor binding using a combination of sequence information and experimental data (centiSNPs) (Moyerbrailean et al., 2016b), SNPs associated with complex traits (GWAS), and eQTLs in GTEx.

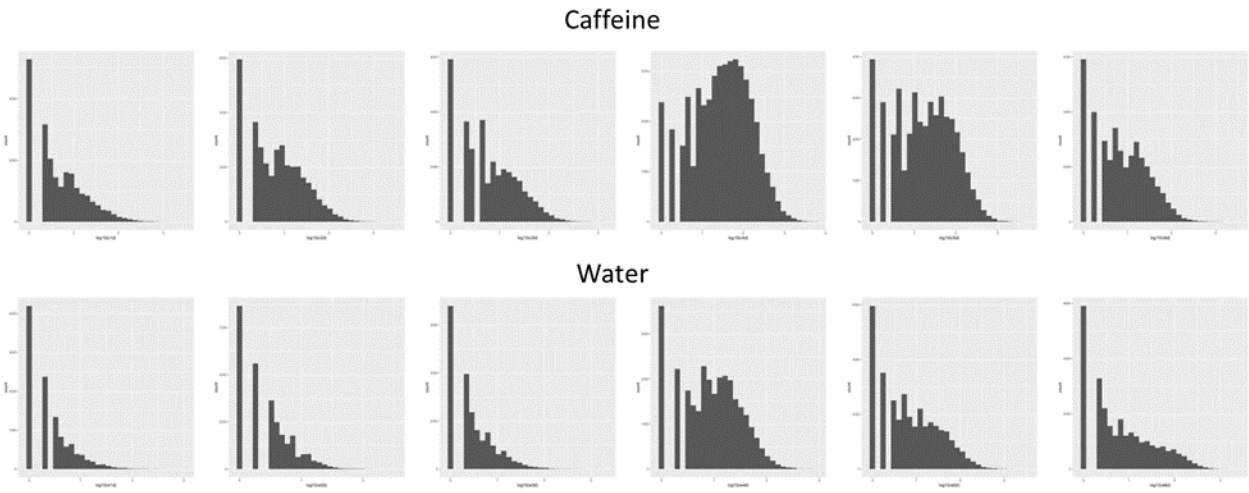

Figure 1- figure supplement 1. Coverage histograms for all 12 libraries used in this study. Caffeine libraries are on top, and water libraries are on bottom.

Results

Active Regulatory Regions in Response to Caffeine

For this study we used a library of targets that was designed to capture a large number of predicted gene regulatory variants in motifs for hundreds of transcription factor binding sites (figure 1 B). The targets in our library consisted of self-transcribing enhancer regions containing a minimal promoter, a reporter gene, and the oligonucleotide containing the candidate regulatory SNP. These candidate regulatory SNPs belong to multiple categories, including SNPs predicted to alter transcription factor binding using a combination of sequence information and experimental data (centiSNPs) (Moyerbrailean et al., 2016b), SNPs associated with complex traits (GWAS), and eQTLs in GTEx. In addition, we included 1,676 negative control sequences. To test if these putative regulatory sequences mediate the response to caffeine, we used DESeq2 to test for differential activity of the targets in cells treated with caffeine compared to cells in the control group (see materials and methods section for the specific model). The library contained motifs in both the forward and reverse orientations within separate targets. Since these motifs may induce direction-specific effects, we performed the differential activity analysis per each direction separately (see Figure 2- figure supplement 1, see also Figure 2- figure supplement 2 and Figure 2- figure supplement 3), and considered any target with FDR<10% in either direction as significant. We observed 772 significantly differentially active targets: 546 upregulated targets and 226 downregulated targets (figure 2 A, supplemental file 1), showing that caffeine overall increases the activity of the regulatory elements.

We then focused on differentially active targets containing a known caffeine response factor as determined based on chromatin accessibility data from endothelial cells treated with caffeine(Findley et al., 2019). We observed that these targets had lower p-values, as observed in the QQ plot in figure 2 A (inset). To identify any additional transcription factors that may be important for the response to caffeine, we conducted a motif scan for 838 known transcription factor binding motifs using the JASPAR CORE Vertebrates 2022 database(Castro-Mondragon et

al., 2022) (supplemental file 2). We found 19 motifs that were enriched for being within differentially active targets (figure 2 B, supplemental file 3). We found the motif for ZNF423, one of the caffeine response factors, was enriched within the differentially active targets. The 3 most enriched motifs were NFATC1, NFATC4, and NFATC2. The NFAT transcription factor family is known for their involvement in the Ca^{2+}/NFAT pathway. This signaling pathway plays an important role in maintaining the homeostasis of vascular endothelial cells(Wang et al., 2020) and contributes to the mediation of proliferation and migration(Johnson et al., 2003; Wang et al., 2020). Thus, improper signaling of the Ca^{2+}/NFAT pathway can induce endothelial dysfunction(Garcia-Vaz et al., 2020; Wang et al., 2020). In diabetic mice, NFAT expression exacerbated atherosclerosis(Blanco et al., 2018; Zetterqvist et al., 2014) and increased foam cell formation(Du et al., 2021). In human coronary artery smooth muscle cells, NFAT signaling mediates vascular calcification(Goettsch et al., 2011). To better understand the regulatory response to caffeine, we then investigated which motifs were enriched for being within upregulated or downregulated targets separately (supplemental file 3). We observed 19 motifs enriched for being within upregulated targets (figure 2 C), and 23 motifs enriched for being within downregulated targets (figure 2 D). Motifs enriched for being within upregulated targets include the previously mentioned NFATC1 and ZNF423. Motifs enriched for being within downregulated targets include the previously mentioned NFATC2 and ZNF423. We also observed that the motif for SREBF2, also called SREBP2, is enriched for being within downregulated targets. In hepatocytes, caffeine is known to suppress SREBF2 activity, which reduces PCSK9 expression, and thus increases LDLR expression, which could be protective against cardiovascular disease(Lebeau et al., 2022). The corresponding transcription factors for these motifs could also play a role in mediating the response to caffeine in vascular endothelial cells. The motif for TEAD4 was also identified as enriched for being within downregulated targets. Interestingly, a CAD GWAS risk variant disrupts binding of TEAD4 in smooth muscle cells, causing lower expression of p16, which could potentially contribute to the risk identified at this locus(Almontashiri et al., 2015).

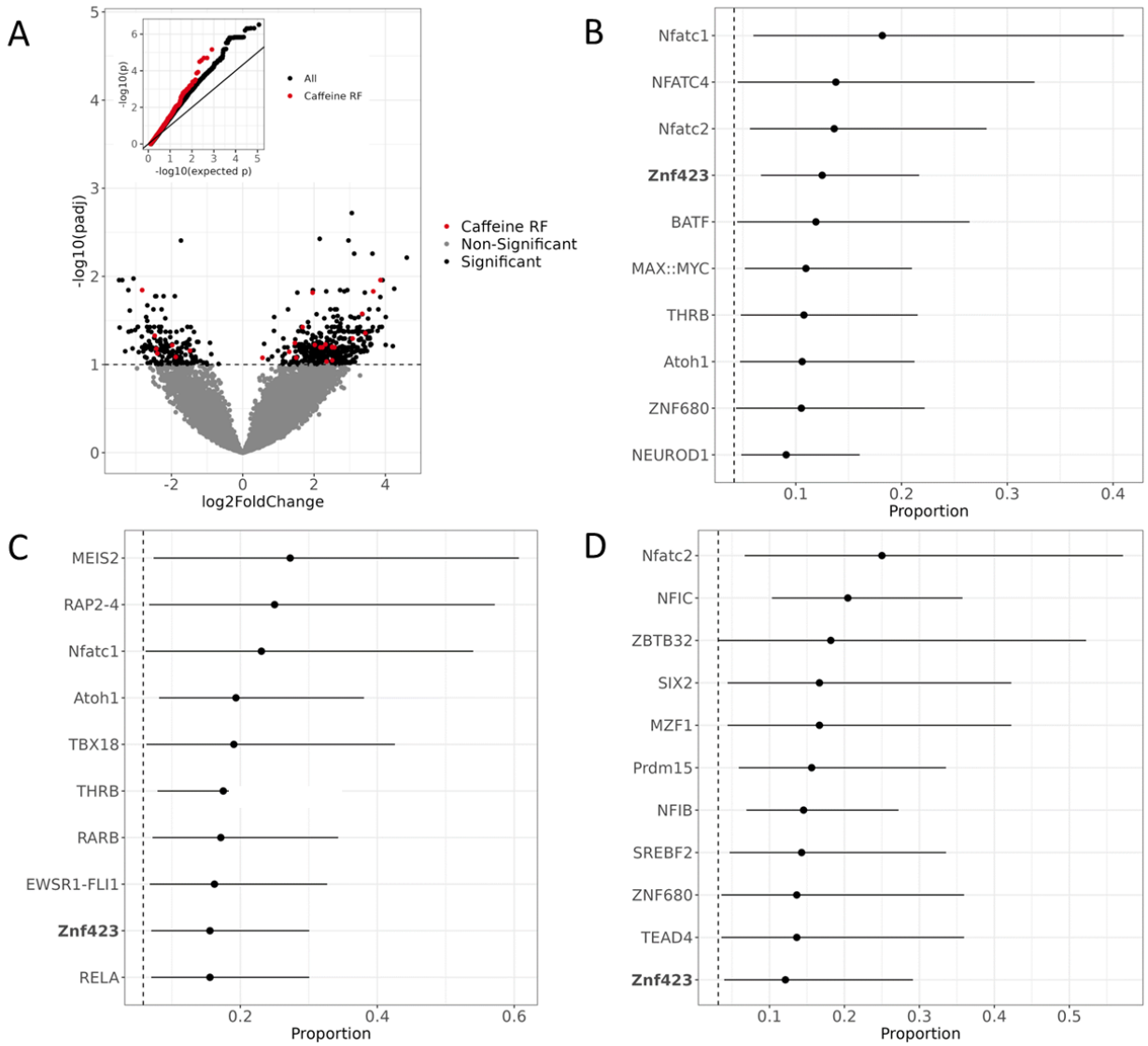

Figure 2. Active regulatory regions in caffeine response in vascular endothelial cells.(A) Volcano plot for DESeq2 results showing targets differentially active in caffeine. The light red points are significant (FDR < 10%) targets containing caffeine response factor binding sites, the black points are significant targets not containing a caffeine response factor binding site and the gray points are non-significant targets. The inset contains a QQ plot for targets containing a caffeine response factor binding site (red), or no caffeine response factor binding site (black). (B) Motifs enriched via test of proportions ($p<0.05$) within differentially active targets. Names of caffeine response factors are bolded. (C) Motifs enriched via test of proportions within upregulated targets ($p<0.05$). (D) Motifs enriched via test of proportions within downregulated targets ($p<0.05$).

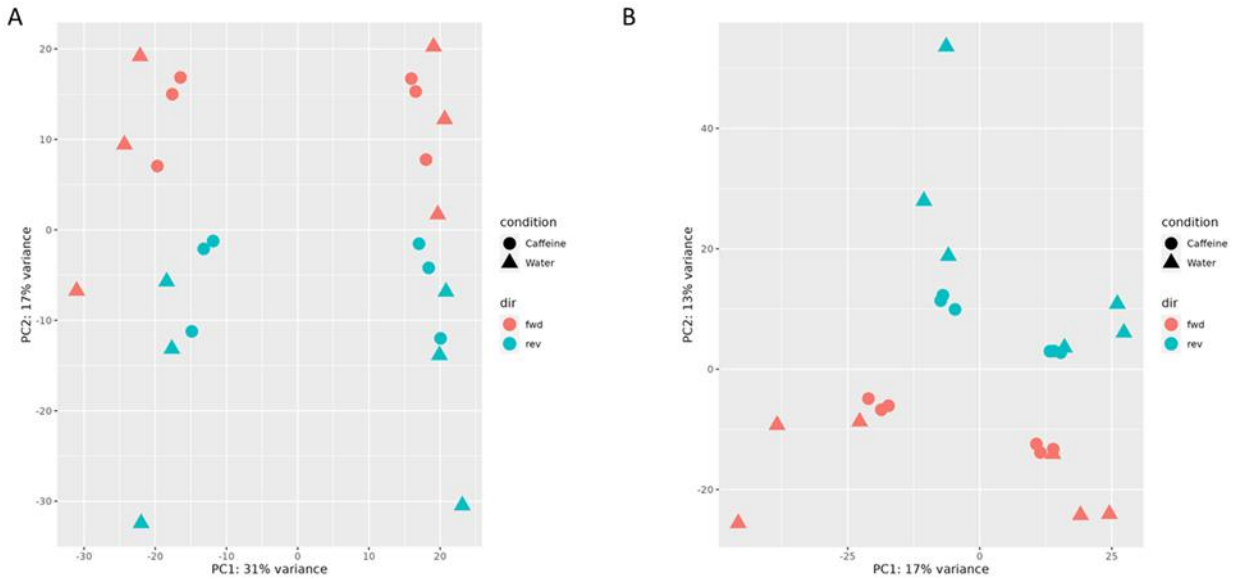

Figure 2- figure supplement 1: PCA plots from read count data in the first (A) and second (B) batch (experiment), annotated by direction.

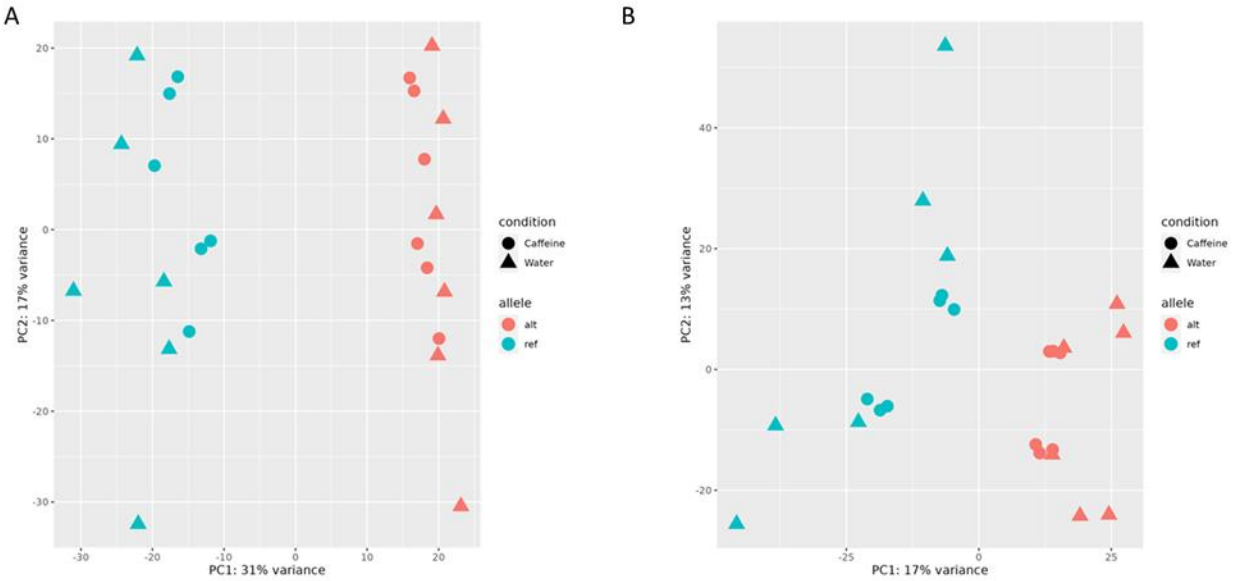

Figure 2- figure supplement 2: PCA plots from read count data in the first (A) and second (B) batch (experiment), annotated by allele.

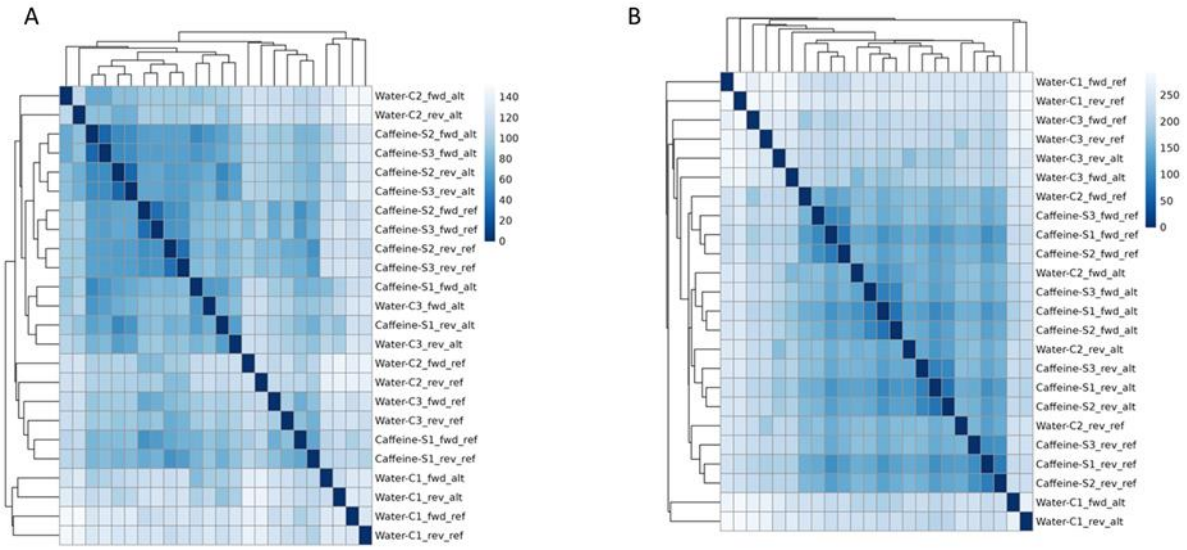

Figure 2- figure supplement 3: Heatmap plots from read count data in the first (A) and second (B) batch (experiment).

Allelic Effects on Gene Regulation Within Conditions and in Response to Caffeine

To investigate how genetic variation affects regulatory sequences and their function in cells treated with caffeine and in the control samples, we tested for allele-specific effects (ASE). Since the library contained the same sequence in both the forward and reverse orientations in independent targets and the regulatory effect may be direction-dependent, we tested for ASE in each SNP/direction pair separately (supplemental file 4, Figure 3- figure supplement 1, see "BiT-STARR-Seq Library Design" section of methods for a more detailed description of terminology used). We observed 689 SNP/direction pairs (corresponding to 665 SNPs) with significant ASE out of 50,914 SNP/direction pairs (30,680 SNPs) tested (2.2%, FDR<10%) (figure 3 A). Additionally, our library contained negative control sequences that were predicted to not have an allelic regulatory effect. These negative control sequences tend to have higher p-values than other sequences in our library, as predicted (figure 3 A). These results demonstrate that genetic variation within regulatory sequences within our library can modulate gene expression levels in vascular endothelial cells.

To directly test for GxE in the molecular response to caffeine, we tested for conditional allele-specific effects (cASE), where ASE is only significant in one condition, or significantly different between the two conditions. When testing for cASE, we observed 29 significant SNP/direction pairs (corresponding to 29 SNPs) out of 23,814 (15,927 SNPs) tested (0.2%, FDR<10%) (figure 3 B, supplemental file 4, Figure 3- figure supplement 2). Additionally, we annotated which of these cASE targets contains a known caffeine response factor as defined in Findley *et al.* 2019, based on ATAC-seq data. There is an enrichment for these targets containing caffeine response factors, as expected (figure 3 B). The size of this enrichment may be underestimated due to the relatively small amount of caffeine response factor motifs (4) present in the designed library. Importantly, these variants contribute to inter-individual variation in response to caffeine. Thus, we sought to

further characterize these variants.

To investigate the regulatory architecture underlying these genetic effects on gene expression, we asked whether ASE (FDR<10%) and cASE variants (here defined at a nominal p<0.01, N= 569) were enriched in open chromatin regions as annotated in Findley *et al.* (2019)(Findley et al., 2019). ASE variants were 2.4-fold enriched within open chromatin regions (p<2.2e-16, figure 3C, green), while a more moderate trend was observed for cASE (1.2-fold, p= 0.088, figure 3 C, purple). This difference in enrichment could be due to the difference between the native chromatin context versus the reporter assay context. Environmental effects on gene regulatory sequences may have a more complex regulatory architecture influenced by the chromatin context that may explain the difference in the enrichment results between ASE and cASE (supplemental file 3).

Genetic regulation of gene expression can be context-dependent, with factors such as cell type(Donovan et al., 2020; Kim-Hellmuth et al., 2020), developmental states(Cuomo et al., 2020; Strober et al., 2019), and environmental stimuli all contributing to gene-environment interactions (GxE-eQTLs, also known as response eQTLs, dynamic eQTLs, context-eQTLs) (see for example(Alasoo et al., 2019; Barreiro et al., 2012; Çalışkan et al., 2015; Findley et al., 2021; Kim-Hellmuth et al., 2017; Maranville et al., 2011; Moyerbrailean et al., 2016b). These context-specific effects can be captured without large cohorts if the appropriate experimental design is applied(Findley et al., 2021). Allele-specific expression experiments in two different conditions can detect GxE in small sample sizes compared to eQTL studies(Moyerbrailean et al., 2016b). To investigate the abundance of GxE missing from large databases such as GTEx(GTEx Consortium, 2020), we tested if cASE variants were enriched for artery eQTLs. Using variants within the open chromatin regions described above, we conducted a Fisher's exact test and found that ASE variants were 1.38 times more likely to be artery eQTLs (p=0.01, figure 3 C, green) compared to variants that did not show ASE. We observed a more moderate trend for cASE variants (1.2 times more likely to be artery eQTLs,p=0.26, figure 3 C, purple, supplemental file 3).

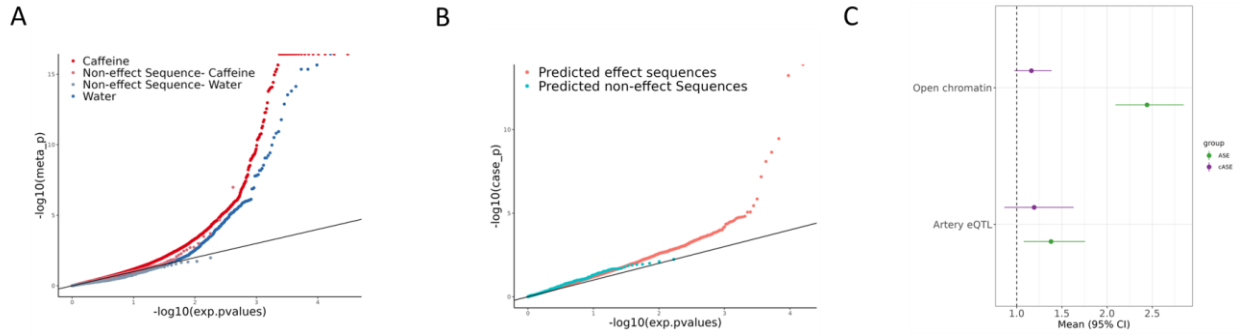

Figure 3. Allelic effects on gene regulation within conditions and in response to caffeine. (A) QQ plot depicting the p-values for ASE, with negative control sequences labeled in lighter red (caffeine) or lighter blue (control) (B) QQ plot depicting the p-values for cASE, with targets containing caffeine response factor motifs annotated in pink and all other sequences in blue. (C) Enrichment via Fisher's exact test of ASE (green) and cASE (purple) variants in open chromatin regions and targets containing artery eQTLs.

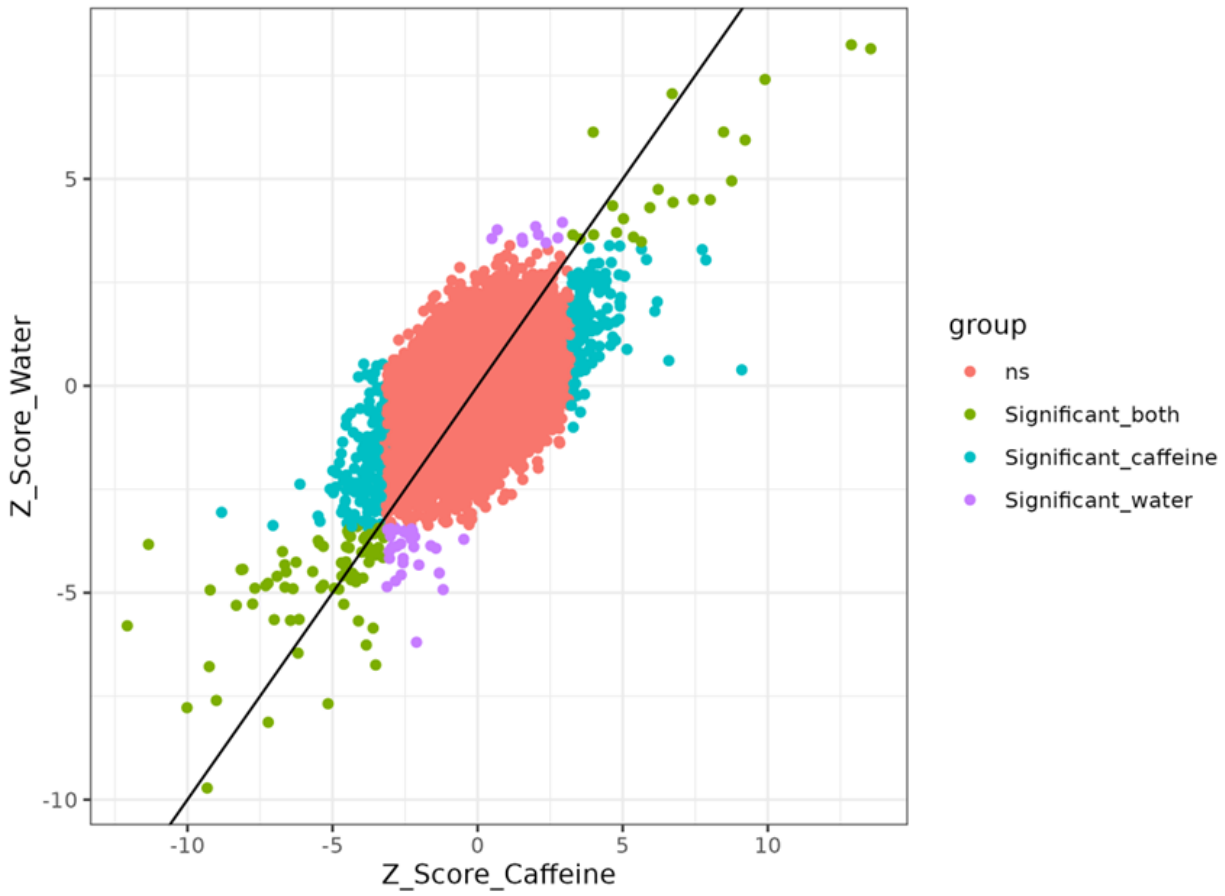

Figure 3- figure supplement 1: Z-score scatter plot from ASE analysis comparing caffeine and water.

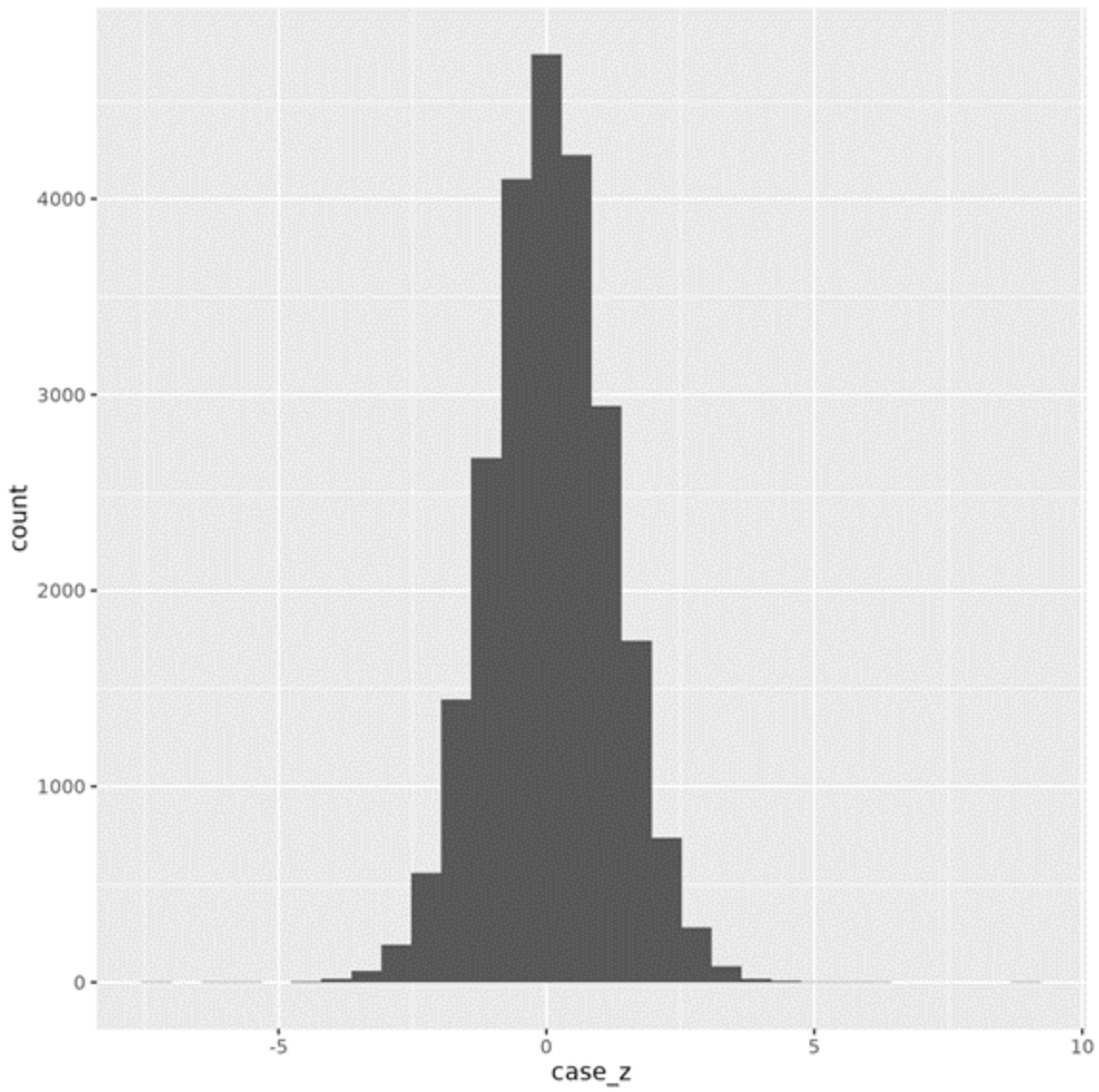

Figure 3- figure supplement 2: Distribution of delta Z scores from cASE analysis.

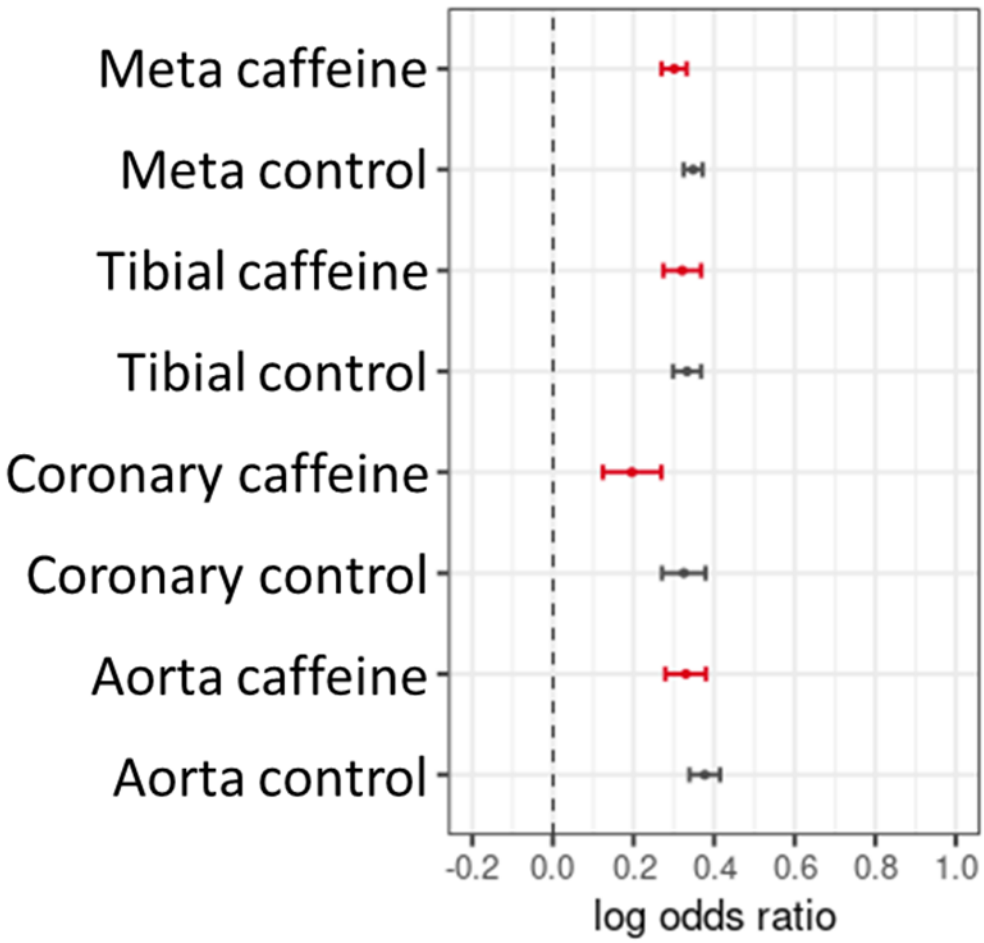

Figure 3- figure supplement 3: Artery eQTL enrichment for footprint SNPs (control, black) or caffeine response factors (caffeine, red). "Meta" is the combination of all 3 artery tissue types in GTEx.

Characterizing Allele-Specific Effects and Conditional Allele-Specific Effects Across Transcription Factor Motifs

We hypothesized that the regulatory context defined by the transcription factor motifs present in each target determines the effect of a genetic variant on expression in caffeine treated cells. We conducted a motif scan of the library of targets for 838 known transcription factor binding motifs from JASPAR(Castro-Mondragon et al., 2022) (supplemental file 2). We then used a test of proportions to identify any motifs that were disproportionately within targets containing significant ASE or cASE variants. For targets containing ASE variants, we observed 44 enriched motifs (figure 4). For targets containing cASE variants, we observed 18 enriched motifs (figure 4, supplemental file 3). Factors of interest for cardiovascular function include NRF1, enriched for targets containing cASE and ASE variants, which is known to regulate lipid metabolism(Hirotsu et al., 2012; Huss and Kelly, 2004), and is annotated as part of the lipid metabolism pathway in

Reactome(Fabregat et al., 2018). KLF15 and KLF14 are also enriched in targets with cASE and ASE. KLF15 is involved in cardiac lipid metabolism (Prosdocimo et al., 2014, 2015), and KLF14, has previously been associated with cardiovascular disease (Chen et al., 2012; Hu et al., 2018). Lastly, SREBF2, which was identified as enriched within our differential activity results, also has implications for disease as discussed previously. This implies that SREBF2 are important both in the interindividual response to caffeine as well as disease state, linking our identified GxE with atherosclerotic disease.

We also wanted to investigate if ASE and cASE variants were disproportionately present in caffeine response factor binding sites, which may indicate that caffeine response factors' regulatory function may be modified by genetic variation. For this analysis, we used annotations from Findley *et al.* (2019)(Findley et al., 2019), which defines caffeine response factors as transcription factors with motifs that were significantly enriched or depleted in differentially accessible chromatin after treatment with caffeine. Factors INSM1 and PLAG1 are caffeine response factors as defined in Findley et al. (2019), confirming that genetic variation may modulate the response to caffeine by increasing binding activity of these transcription factors.

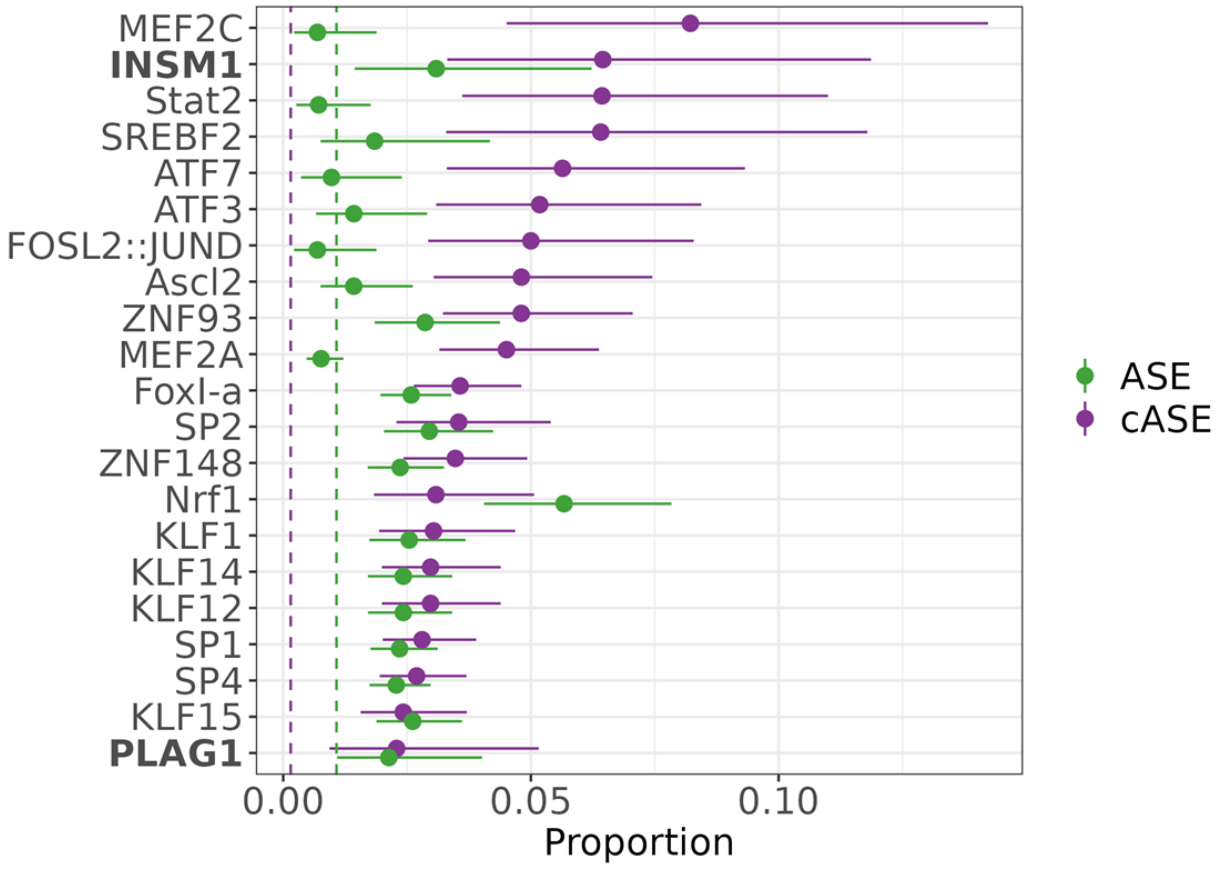

Figure 4. Transcription factors contributing to ASE and cASE. Motifs enriched via test of proportions for significant ASE (green), or cASE (purple). The dotted lines represent the baseline proportion (mean number of significant variants within any motif) for ASE (green) and cASE (purple). Bolded factors are caffeine response factors as defined in Findley *et al.* (2019).

Validation of the Regulatory Mechanism for Fine-Mapped Artery eQTLs

Computational fine-mapping is a commonly used method to identify causal variants, often for complex traits, however further functional validation is usually needed to confirm the regulatory mechanism underlying fine-mapped causal variants. We previously showed that artery eQTLs are enriched in caffeine response factor motifs (also see Figure 3- figure supplement 3)(Findley et al., 2019). We now leverage this finding to fine-map artery eQTLs using DAP-G and the caffeine response factor annotation from Findley et al. (Findley et al., 2019). In our library we tested 187 fine-mapped variants. We identified significant ASE for 6 SNPs (6 SNP/direction pairs), thus validating the regulatory function of these fine mapped causal eQTLs (supplemental file 5). We also identified 2 fine-mapped artery eQTLs with significant cASE ($p<0.01$; 2 SNP/direction; supplemental file 5), which may represent hidden GxE in GTEx.

We then investigated if genes shown to be linked to CAD and hypertension risk using both TWAS and colocalization analysis (INTACT, Okamoto et al., 2023) can be further modulated by GxE with caffeine. INTACT combines TWAS and colocalization approaches, thus overcoming the TWAS inherent linkage disequilibrium limitations. We identified 8 cASE variants that regulate the expression of genes associated with CAD and hypertension (<10% FDR for INTACT analysis, $p<0.01$ for cASE analysis, supplemental file 6). We used available experimental data and computational predictions to dissect the regulatory mechanisms underlying context-specific allelic effects for two of these SNPs and their impact on CAD risk (Figures 5 and 6). We considered allelic effects on transcription factor binding, expression of the transcription factors, and disease risk. Specifically, we considered i)predicted effects on transcription factor binding from CentiSNP, which uses a combination of sequence information and chromatin accessibility data to predict alleles that increase binding of specific transcription factors(Moyerbrailean et al., 2016a); ii)changes in the expression of the genes encoding for the relevant transcription factors and changes in chromatin accessibility (Findley et al., 2019); iii) and allelic effect on the target genes (this study and GTEx).

rs4938344 is an eQTL regulating the long non-coding RNA *AP000892.6*. The reference allele at this locus, G, results in decreased expression of *AP000892.6*. (as measured in GTEx and in the caffeine condition of our assay, Figure 5D and 5E respectively). INTACT associated high expression of *AP000892.6* with decreased risk of hypertension and CAD (Figure 5F). This SNP is predicted to modulate binding of GABP (a known repressor of transcription(Genuario and Perry, 1996)) and ETV1 at this site (Figure 5C). These transcription factors are upregulated in caffeine exposed endothelial cells (Findley et al., 2019)(Figure 5B). This increase in expression uncovers allelic differences in gene regulation which are not detected in the absence of caffeine, likely because of the low expression of the repressor.. The allelic differences in binding of these factors should lead to allelic differences in expression of *AP000892.6*. Accordingly, the reference allele for this variant exhibited lower activity in response to caffeine in our BiT-STARR-seq experiments

(Figure 5E). This effect is consistent with the GTEx artery eQTL for *AP000892.6* (Figure 5D). In summary, caffeine induces higher expression of the ETV1 and GABP transcription factors, which then bind preferentially to the reference allele at rs4938344, this results in lower expression of *AP000892.6* and increased risk for CAD and hypertension (Figure 5A). *AP000892.6* interacts with the *RB1CC1* RNA (Gong et al., 2018), which may play a role in atherosclerosis via its function in forming the autophagosome (Chen et al., 2021).

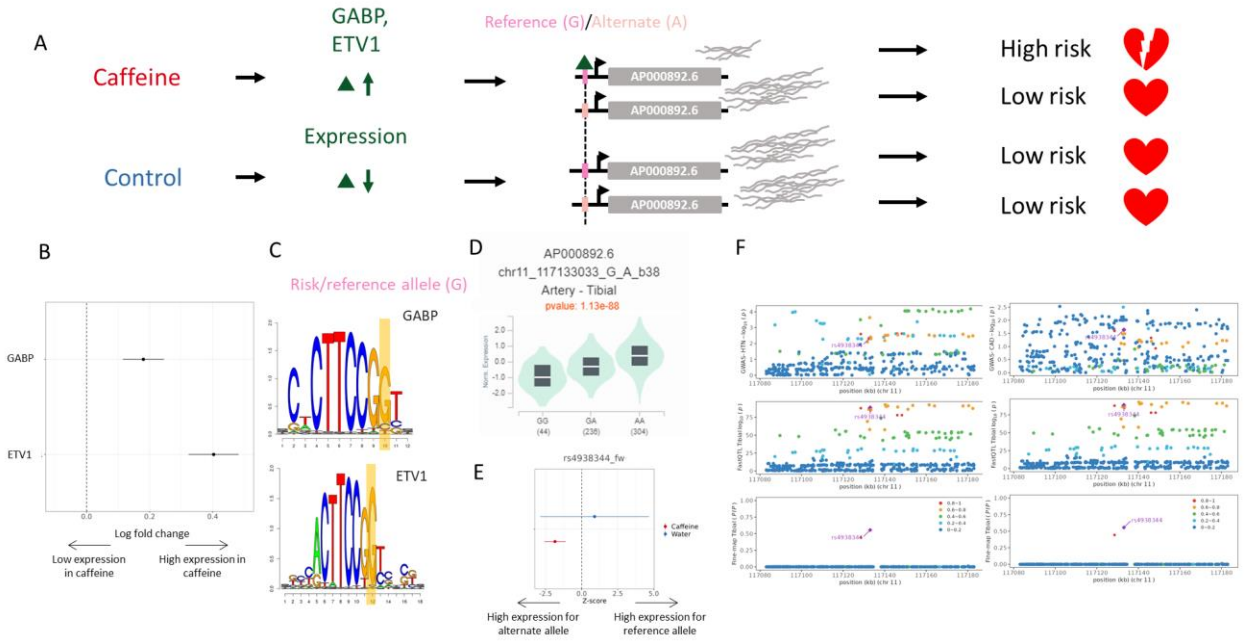

Figure 5: (A) Potential mechanism for rs4938344. (B) Transcription factors containing rs4938344 in a binding site are upregulated upon caffeine exposure. (C) Logos of transcription factor motifs with rs4938344 highlighted. (D) GTEx violin plot for AP000892.6. (E) Effect size from the BiT-STARR-Seq assay for this SNP within each condition. (F) Locus zoom plots showing GWAS and eQTL data for hypertension (left) and CAD (right) in tibial artery tissue.

rs4527034 is an eQTL regulating the *KAT8* gene. The reference allele at this locus, A, results in decreased expression of *KAT8* (as measured in GTEx and in the control condition of our assay, Figure 6D and 6E respectively). INTACT associated high expression of *KAT8* with increased risk of hypertension (Figure 6F). This SNP is predicted to modulate binding of the TERF2IP transcription factor at this site (Figure 6C). TERF2IP is upregulated in caffeine exposed endothelial cells (Findley et al., 2019)(Figure 6B). This increase in expression may saturate all binding sites in the caffeine condition, while the transcription factor may only bind to the preferential allele in the control condition. The allelic differences in binding of these factors should lead to allelic differences in expression of *KAT8* in the control condition, which is what we observe both in our BiT-STARR-seq experiments (Figure 6E) and in GTEx artery eQTL for *KAT8* (Figure 6D). In summary, in the absence of caffeine, TERF2IP binds preferentially to the reference allele at rs4527034, this results in lower expression of *KAT8* and reduced risk for hypertension. In the presence of caffeine, TERF2IP is upregulated, resulting in increased binding and lower expression

of KAT8, independently of the genotype, with an expected overall protective effect on hypertension. Confirming this potential mechanism for disease risk, TERF2IP expression levels were found to affect plaque formation in a mouse model (Kotla et al., 2019). High expression of KAT8, a histone acetyltransferase, also coincides with atherosclerotic progression, and histone acetylation increases in plaques within vascular endothelial cells (Greißel et al., 2016; Zhang et al., 2018).

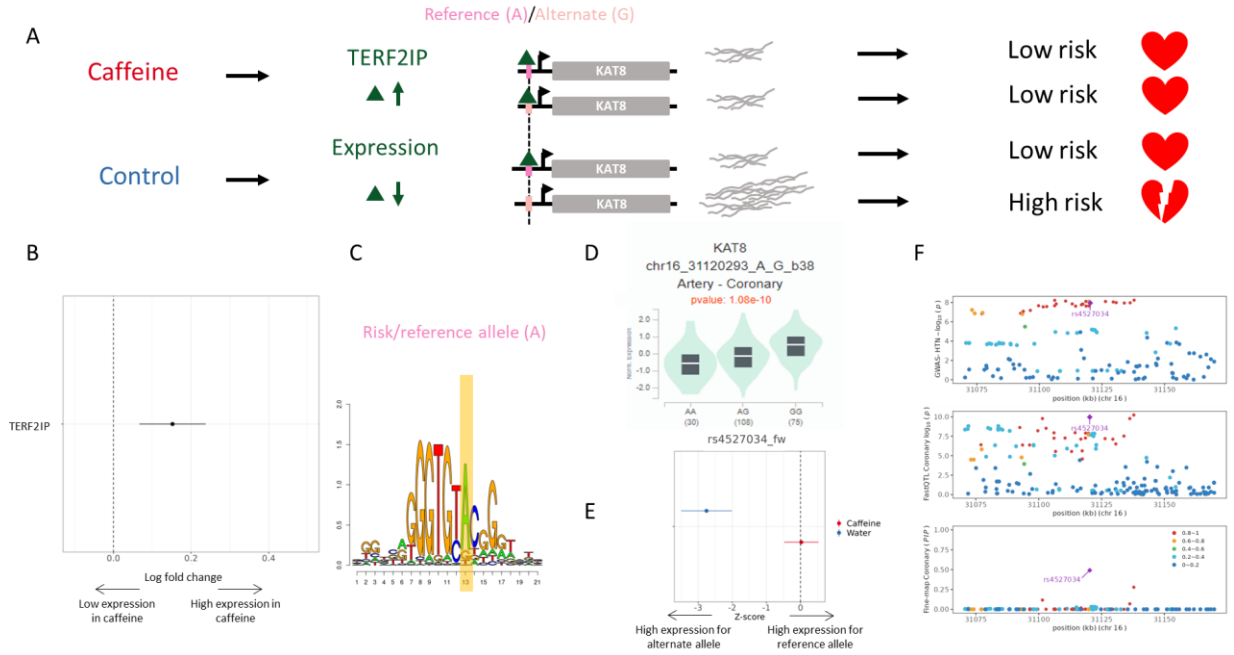

Figure 6: (A) Potential mechanism for rs4527034. (B) TERF2IP is upregulated upon caffeine exposure. (C) Logos of TERF2IP motif with rs4527034 highlighted. (D) GTEx violin plot for *KAT8*. (E) Effect size from the BiT-STARR-Seq assay for this SNP within each condition. (F) Locus zoom plots showing GWAS and eQTL data for hypertension in coronary artery tissue.

Discussion

This study utilized the MPRA BiT-STARR-Seq to identify gene regulatory activity in vascular endothelial cells exposed to caffeine. By utilizing BiT-STARR-Seq, we were able to identify a molecular response to caffeine, allele-specific effects, and conditional allele-specific effects (cASE). By combining our results with pre-existing annotations, we were able to characterize variants exhibiting cASE and identify potential mechanisms for some of these variants.

Heart disease is one of the leading causes of death in the United States according to the CDC (http://wonder.cdc.gov/ucd-icd10.html). The most common type of heart disease is coronary artery disease (CAD), which affects over 18 million adults over 20(Fryar et al., 2012). Common risk factors of CAD include hypertension, high cholesterol levels, and family history(Brown et al., 2022; Hajar, 2017). CAD occurs when plaques form in the arteries (atherosclerosis), causing a

narrowing of the artery, which reduces blood flow to the heart. The innermost layer of the artery is composed of endothelial cells. The endothelium is involved in several important functions, including regulation of blood flow, angiogenesis, thrombosis, and coagulation(Hadi et al., 2005; Krüger-Genge et al., 2019). Endothelial dysfunction occurs in diseases such as atherosclerosis and hypertension(Xu et al., 2021), eventually leading to CAD(Matsuzawa and Lerman, 2014). The molecular mechanisms behind endothelial dysfunction and the resulting diseases are largely unknown. Characterizing these molecular mechanisms is crucial in order to gain a more complete understanding of these disease phenotypes. Additionally, although caffeine is known to produce a vascular response, the current literature does not come to a consensus on the role of caffeine in CAD risk. Here we characterized the regulatory response of non-coding variants to caffeine in vascular endothelial cells using a massively parallel reporter assay (MPRA).

BiT-STARR-Seq, the MPRA used in this study, has several advantages over other methods used to detect genotype-environment interactions (GxE). One common method of detecting GxE is response eQTL mapping, which includes collecting samples from large cohorts and exposing those cells to environmental perturbations (see for example(Alasoo et al., 2019, 2018; Barreiro et al., 2012; Çalışkan et al., 2015; Fairfax et al., 2014; Huang et al., 2020; Kim-Hellmuth et al., 2017; Knowles et al., 2018; Lee et al., 2014; Mangravite et al., 2013; Manry et al., 2017; Maranville et al., 2011; Nédélec et al., 2016; Quach et al., 2016). This method has several disadvantages, as it cannot easily interrogate rare variants, relies on variation existing in a cohort (instead of investigating variants of interest), and requires larger cohort sizes to have enough power to detect GxE. In contrast, because our method uses a designed library of targets(Kalita et al., 2018), we are able to interrogate rare variants easily, as our targets are synthesized. Similarly, we can design a library of specific variants to investigate (such as candidate regulatory variants) instead of relying on variation within a cohort(Kalita et al., 2018). BiT-STARR-Seq also allows us to directly compare 2 alleles within the same sequence context without requiring a large cohort. Despite the advantages of BiT-STARR-Seq, unlike response eQTL mapping, we are not interrogating these SNPs in their native chromatin context. Future work may include using genome-editing tools such as CRISPR to directly insert the desired variants in their endogenous locations in the genome. For our study, we determined BiT-STARR-Seq to be the ideal assay to determine GxE for a large number of SNPs.

We observed a regulatory response to caffeine treatment, consistent with previous studies in the same cell type(Findley et al., 2019; Moyerbrailean et al., 2016b). These results suggest that caffeine exposure significantly changes the regulatory activity of vascular endothelial cells, which may have important implications regarding the impact of lifestyle in CAD. As caffeine may modulate gene regulatory activity, the resulting impact on gene expression may increase or decrease CAD risk. In addition, we identified novel transcription factors contributing to the regulatory response to caffeine including several NFAT transcription factors, and SREBF2. NFAT transcription factors are largely known for their role in the Ca^{2+}/NFAT signaling pathway, where Ca^{2+} binds to calmodulin, stimulating calcineurin, which then causes NFAT factors to localize in the nucleus(Crabtree and Olson, 2002; Klee et al., 1998). Caffeine is known to cause an increase in Ca^{2+} in human aortic endothelial cells(Corda et al., 1995), so it is understandable that we find these factors enriched for being within targets that respond to caffeine exposure. SREBF2, also known as SREBP2, is involved in sterol homeostasis(Horton et al., 2003). This result implies these novel transcription factors important for the regulatory response to caffeine may also contribute to understanding the role of caffeine in CAD risk. This coincides with findings that caffeine exposure

can alter expression of genes, including those for transcription factors in mouse cardiomyocytes(Fang et al., 2014). Another study aimed to uncover mechanisms relevant to cardiovascular disease upon caffeine exposure and found that caffeine inhibits the transcription factor SREBP2, which causes an overall protective effect against cardiovascular disease(Lebeau et al., 2022). These results coincide with our findings.

Non-coding regions of the genome contain regulatory variants that modulate gene expression. In this study we identify and characterize over 600 variants exhibiting allele-specific effects. Numerous non-coding variants have been implicated in CAD risk via GWAS(Hartmann et al., 2022; Kessler and Schunkert, 2021; Koyama et al., 2020; Nikpay et al., 2015; Temprano-Sagrera et al., 2022; van der Harst and Verweij, 2018), but they are generally uncharacterized. Few lead non-coding variants have been thoroughly investigated. One of these led the authors to propose and validate a molecular mechanism connecting expression of the EDN1 gene to the phenotypic outcome(Gupta et al., 2017; Wang and Musunuru, 2018). The specific mechanisms that detail how non-coding variants contribute to CAD will be critical in understanding CAD risk and ultimately developing clinical treatments. While understanding various genetic risk factors for CAD is important, genotype-environment interactions for these variants also have an impact on phenotype and have not been widely studied.

Since previous studies have shown that GxE-eQTL can modulate complex disease risk, we expect that GxE detected in our assay may be relevant to CAD(Alasoo et al., 2019, 2018; Barreiro et al., 2012; Çalışkan et al., 2015; Fairfax et al., 2014; Findley et al., 2021; Huang et al., 2020; Kim-Hellmuth et al., 2017; Knowles et al., 2018; Lee et al., 2014; Mangravite et al., 2013; Manry et al., 2017; Maranville et al., 2011; Nédélec et al., 2016; Quach et al., 2016). We tested for conditional allele-specific effects (cASE), which occurs when allele-specific effects are only significant in one condition, or significantly different between the two conditions. This analysis identifies genotype-environment interactions (GxE) which are important in understanding disease risk while accounting for genetic and environmental context. We identified 29 variants that regulate the gene expression response to caffeine and demonstrated that context-aware MPRAs can be used to dissect molecular mechanisms underlying cardiovascular health.

By fine-mapping artery eQTLs and combining the data collected from our assay with pre-existing annotations, we produced potential mechanisms for 2 cASE variants through altered transcription factor expression and binding in response to caffeine. Our results indicated that both genetic and environmental factors are important in determining risk, and that the interaction between these factors can be informative to mechanisms and phenotypic consequences. Importantly, by utilizing multiple functional annotations, we are able to identify variants that may be relevant to disease but did not reach genome-wide significance in GWAS, possibly because of their context-specific effects. By studying different environmental contexts, we can identify that, in these instances, the presence of caffeine can impact risk of poor cardiovascular health outcomes. If environmental context was not considered and this work was conducted solely in the control condition, the caffeine modulatory effect on risk would have been missed.

Although we investigate GxE for caffeine in vascular endothelial cells, our experimental approach can be applied to various different complex diseases and their relevant cell types and treatments. To further validate our work, genome editing tools could be used to investigate the effect of these variants in their native chromatin context. Additional validation could include allele-specific and

condition-specific transcription factor binding assays (such as electrophoretic mobility shift assays) for the fine-mapped variants.

Our study demonstrates the importance of considering environmental contexts when investigating gene regulatory activity, as we identify several thousand instances of GxE in our library of candidate regulatory variants. Our data, combined with pre-existing annotations, allowed us to identify transcription factors involved in GxE in caffeine, and describe specific potential molecular mechanisms for some of these GxE. Our results provide important insights into the molecular regulatory effect of caffeine exposure and GxE for caffeine in vascular endothelial cells.

Materials and Methods

Cell Culture

Human umbilical vein endothelial cells (HUVECs) were obtained from Lonza (catalog #: CC-2517-0000315288). Cells were cultured at 37°C with 5% CO2 and seeded at 5,000 cells/cm^2. EGM-2 growth medium (Lonza) was used to culture the cells.

Treatment

Treatment concentration was the same as used in previous studies(Findley et al., 2019; Moyerbrailean et al., 2016b). We used a caffeine concentration of $1.16\times10{-}3$ M. In addition, water was used as a vehicle control, as that was the solvent used to prepare the caffeine treatment.

BiT-STARR-Seq Library Design

We designed 43,556 target regulatory regions each containing a SNP in the middle and with a total length of 200 nucleotides. This set of targets corresponds to 87,112 constructs each containing only one of two alleles at the test SNP. Additionally, each construct can be integrated in the forward or reverse orientation, leading to a maximum of 174,224 constructs in either direction. Please also see below for a description of how we use library-related terms throughout the paper. The library used is the same as reported in Kalita *et al.* (Kalita et al., 2018). Briefly, the library of target regulatory sequences consisted of several categories of regulatory variants, including eQTLs (Innocenti et al., 2011; Wen et al., 2015), SNPs predicted to disrupt transcription factor binding (centiSNPs)(Moyerbrailean et al., 2016a), and SNPs associated with complex traits in GWAS(Pickrell, 2014). Negative controls that were not predicted to have a regulatory effect were also included in the library(Moyerbrailean et al., 2016a). It is important to note that these negative controls are only predicted not to have a regulatory effect via computational annotation (Moyerbrailean et al., 2016a), so they may not be representative of true negative controls. This is why we largely do not utilize these SNPs as negative controls within our analyses. Our predictions of regulatory activity also did not account for environmental context, thus these sequences are also not suited to annotate our cASE results.

SNP (n = 43,556): Refers to a genetic variant tested for allelic effects on gene regulation.

Target (n = 43,556): 200 nucleotide-long oligonucleotide sequence that contains the test SNP in

the middle of the target.

Construct (n = 87,112): Synthesized 200 nucleotide-long oligonucleotide sequence that contains only one of the two possible alleles at the test SNP. Each target corresponds to two constructs. .

Direction: Constructs can integrate in either the forward or reverse direction relative to the direction of transcription in the BiT-STARR-seq plasmid. Therefore two directions are possible for each construct.

SNP/direction pair (n = 87,112): A SNP tested for allelic effects on gene regulation contrasting the expression of two constructs that are integrated in the same direction. All statistical tests are performed at this level, testing in each direction separately.

BiT-STARR-Seq Experiments

Oligonucleotides were synthesized and used to create a DNA plasmid library, which was sequenced and used as a subsequent input for the ASE analysis. The DNA library was transfected into HUVECs using the Lonza Nucleofector X platform. Cells were electroporated using the DS-120 setting with primary cell solution P5. Caffeine was added at 1.16×10^{-3} M after transfection. Cells were incubated for 24 hours and lysed. We completed 6 replicates per treatment condition (caffeine and the water vehicle control), or 12 replicates in total.

Library Preparation and Sequencing

RNA was extracted using the RNeasy Plus Mini kit (Qiagen, catalog #: 74136). A cDNA library was prepared using the Superscript III First-Strand Synthesis kit (Invitrogen, catalog #: 18080-400). Sequencing was completed using the Illumina Nextseq 500 to generate 125 cycles for read 1, 30 cycles for read 2, 8 cycles for the fixed multiplexing index 2 and 10 cycles for index 1 (variable barcode). Average sequencing depth per library was 39,235,611 reads, for a total of 470,827,333 reads (Figure 1- figure supplement 1, supplemental file 7).

Processing Sequence Data

To analyze the RNA-seq data, we began by demultiplexing our data using the bcl2fastq software to create demultiplexed FASTQ files. We then aligned to hg19 using HISAT2. Afterwards, we applied a filter to ensure the UMIs present match the expected UMI pattern (RDHBVDHBVD). Reads with short UMIs or those that do not match the expected sequence were removed. The resulting BAM files were then deduplicated using UMItools. We ran samtools mpileup followed by bcftools query to output read counts per each allele/direction combination.

Differential Activity Analysis

To test for a molecular response to caffeine, we used the R package DESeq2(Love et al., 2014). To determine which model would best test for a molecular response to caffeine, we completed principal component analysis to identify major sources of variation. We identified that the first PC clearly represented allelic effects (Figure 2- figure supplement 2, also see Figure 3- figure supplement 1- figure supplement 3), thus we included allele (reference or alternate) as part of our model. Our model tested the effect of treatment, correcting for allele (~allele + treatment), as we

observed a strong allelic effect.We ran DESeq2 for each direction (see Figure 2- figure supplement 1), as the library contained motifs in both the forward and reverse orientations within separate target sequences. We considered targets as significant with Benjamini-Hochberg FDR < 10%.

ASE Analysis

To test for ASE, we utilized the R package quantitative allele-specific analysis of reads (QuASAR-MPRA) (Kalita et al., 2017). QuASAR-MPRA is an extension of the QuASAR package which allows for analysis of barcoded MPRA data. QuASAR-MPRA uses a beta-binomial model and accounts for uneven initial allelic proportions present in the DNA library. We used the fitQuasarMpra() function to test for ASE in each experiment separately, estimating the ASE effect and its standard error. For each SNP/direction pair, we meta-analyzed the effect size using a weighted mean utilizing inverse-variance weighting for each condition separately. The z-score for each SNP-direction pair is subsequently calculated as the meta-analyzed effect size minus the DNA proportion, divided by the meta-analyzed standard error of the effect size. We then required each identifier to be within 4 or greater replicates (out of the 6 total replicates), and performed multiple test correction using the Benjamini-Hochberg procedure. Significant ASE was then defined as having an FDR<10%.

cASE Analysis

To test for cASE, we used a method previously developed by our lab called differential allele-specific test, or ΔAST. The calculation for this parameter ΔZ, is provided in (Moyerbrailean et al., 2016a) as well as below. The QuASAR-MPRA package outputs betas for the treatment (β_T) and the control (β_C), as well as the standard error (se) for both groups, which are used to calculate ΔZ. From this ΔZ parameter, we used the pnorm() function in R to calculate a p-value for each variant, and then used the p.adjust() function to perform multiple test correction using the Benjamini-Hochberg procedure.

$$\Delta Z = \frac{\beta_T - \beta_C}{\sqrt{se_T^2 + se_C^2}}$$

Motif Enrichment Analysis

The motif scan was completed using the PWMScan tool, using all PWMs within the JASPAR(Castro-Mondragon et al., 2022) 2022 CORE database (838 motifs). A threshold of -t 10 (base 2) was used for the motif scan, which was restricted to the regions of our designed targets. Once the scan was complete, motifs that were present less than 100 times in the library were removed. For differentially active targets in response to caffeine, 222 motifs passed this filter. For motifs containing ASE variants, 359 motifs passed this filter. For motifs containing cASE variants, 417 motifs passed the filter. A test of proportion (prop.test() in R) was performed per each transcription factor, where the null proportion was the total number of significant targets/variants containing/within motifs (differentially active target, ASE or cASE) divided by the total number of non-significant targets/variants containing/within motifs. The test was done per motif, where the proportion being tested is the same as the null proportion, but conducted per motif rather than across all motifs. The related calculations are shown below:

Expected null proportion = n_a/n_b
Observed motif proportion = n_c/n_d
n_a = Number of targets/variants of interest containing/within any motif
n_b = Number of targets/variants containing/within any motif
n_c = Number of targets/variants of interest containing/within specific motif
n_d = Number of targets/variants containing/within specific motif

ASE and cASE <10% FDR were considered significant. For follow up analyses of cASE features, we relaxed the significance threshold to nominal p-value <0.01.

Open Chromatin Region Enrichment Analysis

To test if certain variants were within open chromatin regions, we obtained the list of differentially accessible regions tested in (Findley et al., 2019). We considered any accessible region (annotated as differentially accessible or not). Bedtools bed intersect tool was used to complete the overlap with the appropriate datasets (DESeq2, ASE, or cASE results), resulting in a list of targets or SNPs that were within open chromatin regions. This annotation was then used to complete the Fisher's exact test. Enrichments for significant ASE or cASE variants were performed separately.

Artery eQTL Enrichment Analysis

To determine if certain variants were artery eQTLs, we obtained GTEx v8(Castro-Mondragon et al., 2022)data for aorta, coronary, and tibial artery tissues. We then intersected the list of variants of interest with the list of artery eQTLs. To determine if artery eQTLs were within differentially active targets, bedtools intersect was used, resulting in a list of targets that contained artery eQTLs. We further subset this list to variants within open chromatin regions (see previous section). This annotation was then used to complete the Fisher's exact test. Enrichments for significant ASE or cASE variants were performed separately.

Fine-mapping Analysis with DAP-G

Based on a previous study(Findley et al., 2019), we define caffeine response factors as transcription factors with motifs that were significantly enriched or depleted in differentially accessible chromatin after treatment with caffeine. We annotated genetic variants into two categories: (1) genetic variants in motifs for response factors (2) genetic variants in motifs for transcription factors that are not caffeine response factors. By integration of these genetic variants annotation, we estimated the probability of each SNP regulating gene expression in a Bayesian hierarchical model using TORUS(Wen, 2016). These probabilities are then used in DAP-G(Zhang et al., 2020) to fine-map eQTLs from all 3 artery tissues in GTEx V8. 364,427,888 eQTLs were fine-mapped across 3 artery tissues. We filtered for a posterior inclusion probability of greater than 0.9.

Identification of putative causal genes

To better understand how cASE variants may impact traits related to cardiovascular health, we performed colocalization analysis of the causal GWAS variants with eQTLs using fastENLOC (Wen et al., 2017) by integration of fine-mapped eQTLs in the three artery tissues (see above section) and fine-mapped GWAS signals for CAD and hypertension. We fine-mapped GWAS

using DAP (Wen, 2016). From fastENLOC we obtained the gene locus-level colocalization probability for each gene which is used to evaluate how the gene is associated with complex traits or diseases. Intuitively, colocalization analysis identified the overlap of causal eQTLs and GWAS hits. However, it lacked the sensitivity due to the failure in distinguishing between the vertical pleiotropy (genetic effects on traits mediated by gene expression) and horizontal pleiotropy (independent effects on gene expression and traits)(Okamoto et al., 2023). To overcome the limitations in a single approach, we combined the evidence from colocalization and TWAS to estimate the probability of putative causal genes using the R package INTACT(Okamoto et al., 2023). Here, the TWAS data we utilized were from PTWAS (Zhang et al., 2020). We determined putative causal genes with FDR<10%. We then identified variants regulating these putative causal genes using the DAP-G fine-mapping results, requiring a SNP-level colocalization probability >0.5, with no threshold on PIP (supplemental file 6).

Acknowledgements

We thank the members of the Luca and Pique-Regi laboratories. This work was supported by the National Institute of General Medical Sciences of the National Institutes of Health (R01GM109215).

Supplemental Materials

Supplemental files are available at Zenodo: https://zenodo.org/doi/10.5281/zenodo.7327508

Supplemental file 1: Differential Activity Results. Output from DESeq2. Column "V1" is the genomic position (hg19) of the SNP within the construct being tested, "baseMean" refers to normalized expression, "log2FoldChange" refers to the effect size, "lfcSE" is the standard error of the effect size, "stat" is the log2FoldChange divided by lfcSE, "pvalue" is the nominal p-value, and "padj" is the adjusted p-value. Please refer to DESeq2 documentation for additional details.

Supplemental file 2: PWMScan Results. Includes all variants in the designed library that are within a motif from the JASPAR 2022 CORE Vertebrates database determined by PWMScan. Columns 1-3 are the genomic position of the test SNP, column 4 is the rsID for the test SNP, and column 5 is the ID of the motif in JASPAR.

Supplemental file 3: Characterization of Regulatory Regions and Variants. Includes contingency tables for all Fisher's exact tests reported in this study. Please see the legend provided within the file. Briefly, A1-C5 is the contingency table and summary of results for ASE variants within open chromatin regions, A6-C10 for ASE variants overlapping with artery eQTLs, A11-C15 for cASE variants within open chromatin, and A16-C20 for cASE variants overlapping with artery

eQTLs.

Supplemental file 4: Allele-Specific Effects and Conditional Allele-Specific Effects Results. Output from ASE/cASE analysis. The "identifier" column is the SNP/direction pair, "meta_estimate" is the effect size (of the ASE), "meta_se" is the standard error of the effect size, "n" is the number of replicates containing non-zero read counts, "DNA_prop" is the proportion of reads (reference/alternate) present in the DNA library, "meta_z" is the z-score, "meta_p" is the nominal ASE p-value, "meta_padj" is the adjusted ASE p-value before filtering for SNP/direction pairs where n>3, "new_padj" is the adjusted ASE p-value after filtering for SNP/direction pairs where n>3, "group" refers to the applicable condition for the statistic (.x suffix for caffeine, .y suffix for control), "case_z" is the z-score (ΔZ statistic), "case_p" is the nominal p-value for cASE, and "case_padj" is the adjusted p-value for cASE.

Supplemental file 5: Fine-Mapped Artery eQTLs with Significant ASE or cASE. The "chr" column contains the chromosome information, "pos" is the genomic position (0-based, hg19), "pos1" is the genomic position (1-based, hg19),"identifier" is the SNP/direction pair, "rsID" refers to rsID, "padj_min" refers to the minimum p-adjusted value for ASE (across conditions), "ID" is an ID to identify the genomic position in the format of chr_pos1, "PIP" is the posterior inclusion probability as calculated by DAP-G, "gene" refers to the Ensembl gene ID, "tissue" refers to the GTEx artery tissues (tibial, aorta, or coronary), "group" identifies if the variant has significant ASE or significant cASE, "case_p" refers to the cASE nominal p-value, and "case_padj" refers to the cASE adjusted p-value.

Supplemental file 6: Variants that Regulate Putatively Casual Genes as Identified via INTACT. The "Identifier" column is the SNP/direction pair, "ID" is an ID to identify the genomic position in the format of chr_pos1 (see above), "Gene ID" refers to the Ensembl gene ID, "Gene symbol" refers to the gene symbol, "Trait" is CAD or hypertension (HTN), "Tissue" refers to the GTEx artery tissues (tibial, aorta, or coronary),"DAP-G PIP" is the posterior inclusion probability as calculated by DAP-G, "GWAS z-score" is the z-score for the GWAS of the corresponding trait, "GWAS pvalue" is the nominal p-value for the GWAS of the corresponding trait, "PTWAS-INTACT PCG" is the probability of being a putative causal gene as calculated by INTACT, "PTWAS-INTACT FDR" is the FDR statistic as calculated by INTACT, "INTACT z-score" is the z-score statistic as calculated by INTACT, "cASE_p" is the cASE nominal p-value, "cASE_padj_(original)" is the cASE adjusted p-value, "cASE_z" is the cASE z-score (ΔZ statistic), "DE_construct?" indicates if the construct is significantly differentially active as defined by our assay, "Caffeine_ASE_z" is the z-score for ASE (caffeine condition), "Water_ASE_z" is the z-score for ASE (control condition), "centiSNP motifs" refers to motifs the SNP is within as identified by centiSNP, and "centiSNP category" refers to if the motif is a footprintSNP, effectSNP, or switchSNP as identified by centiSNP.

Supplemental file 7: Read Counts. Read counts for all libraries at different steps of the data processing pipeline. "Library Name" is an identifier used for the sample, "Treatment" is the condition (caffeine or control), "Dedup" refers to the number of reads after deduplication, and ">1 read either direction" describes the number of reads that have >1 read for each allele in either direction.

References

Alasoo K, Rodrigues J, Danesh J, Freitag DF, Paul DS, Gaffney DJ. 2019. Genetic effects on promoter usage are highly context-specific and contribute to complex traits. *Elife* **8**. doi:10.7554/eLife.41673

Alasoo K, Rodrigues J, Mukhopadhyay S, Knights AJ, Mann AL, Kundu K, HIPSCI Consortium, Hale C, Dougan G, Gaffney DJ. 2018. Shared genetic effects on chromatin and gene expression indicate a role for enhancer priming in immune response. *Nat Genet*. doi:10.1038/s41588-018-0046-7

Almontashiri NAM, Antoine D, Zhou X, Vilmundarson RO, Zhang SX, Hao KN, Chen H-H, Stewart AFR. 2015. 9p21.3 Coronary Artery Disease Risk Variants Disrupt TEAD Transcription Factor-Dependent Transforming Growth Factor β Regulation of p16 Expression in Human Aortic Smooth Muscle Cells. *Circulation* **132**:1969–1978.

Altshuler D, Hirschhorn JN. 2005. MEF2A sequence variants and coronary artery disease: a change of heart? *J Clin Invest*.

Arkenbout EK, van Bragt M, Eldering E, van Bree C, Grimbergen JM, Quax PHA, Pannekoek H, de Vries CJM. 2003. TR3 orphan receptor is expressed in vascular endothelial cells and mediates cell cycle arrest. *Arterioscler Thromb Vasc Biol* **23**:1535–1540.

Arnold CD, Gerlach D, Stelzer C, Boryń ŁM, Rath M, Stark A. 2013. Genome-wide quantitative enhancer activity maps identified by STARR-seq. *Science* **339**:1074–1077.

Barreiro LB, Tailleux L, Pai AA, Gicquel B, Marioni JC, Gilad Y. 2012. Deciphering the genetic architecture of variation in the immune response to Mycobacterium tuberculosis infection. *Proc Natl Acad Sci U S A* **109**:1204–1209.

Blanco F, Heinonen SE, Gurzeler E, Berglund LM, Dutius Andersson A-M, Kotova O, Jönsson-Rylander A-C, Ylä-Herttuala S, Gomez MF. 2018. In vivo inhibition of nuclear factor of activated T-cells leads to atherosclerotic plaque regression in IGF-II/LDLR-/-ApoB100/100 mice. *Diab Vasc Dis Res* **15**:302–313.

Brown JC, Gerhardt TE, Kwon E. 2022. Risk Factors For Coronary Artery DiseaseStatPearls. Treasure Island (FL): StatPearls Publishing.

Çalışkan M, Baker SW, Gilad Y, Ober C. 2015. Host genetic variation influences gene expression response to rhinovirus infection. *PLoS Genet* **11**:e1005111.

Castro-Mondragon JA, Riudavets-Puig R, Rauluseviciute I, Lemma RB, Turchi L, Blanc-Mathieu R, Lucas J, Boddie P, Khan A, Manosalva Pérez N, Fornes O, Leung TY, Aguirre A, Hammal F, Schmelter D, Baranasic D, Ballester B, Sandelin A, Lenhard B, Vandepoele K, Wasserman WW, Parcy F, Mathelier A. 2022. JASPAR 2022: the 9th release of the open-access database of transcription factor binding profiles. *Nucleic Acids Res* **50**:D165–D173.

Chen Y, Zeng A, He S, He S, Li C, Mei W, Lu Q. 2021. Autophagy-Related Genes in Atherosclerosis. *J Healthc Eng* **2021**:6402206.

Chieng D, Canovas R, Segan L, Sugumar H, Voskoboinik A, Prabhu S, Ling LH, Lee G, Morton JB, Kaye DM, Kalman JM, Kistler PM. 2022. The impact of coffee subtypes on incident cardiovascular disease, arrhythmias, and mortality: long-term outcomes from the UK Biobank. *Eur J Prev Cardiol*. doi:10.1093/eurjpc/zwac189

Choi Y, Chang Y, Ryu S, Cho J, Rampal S, Zhang Y, Ahn J, Lima JAC, Shin H, Guallar E. 2015. Coffee consumption and coronary artery calcium in young and middle-aged

asymptomatic adults. *Heart* **101**:686–691.
Corda S, Spurgeon HA, Lakatta EG, Capogrossi MC, Ziegelstein RC. 1995. Endoplasmic reticulum Ca2+ depletion unmasks a caffeine-induced Ca2+ influx in human aortic endothelial cells. *Circ Res* **77**:927–935.
Crabtree GR, Olson EN. 2002. NFAT signaling: choreographing the social lives of cells. *Cell* **109 Suppl**:S67–79.
Cuomo ASE, Seaton DD, McCarthy DJ, Martinez I, Bonder MJ, Garcia-Bernardo J, Amatya S, Madrigal P, Isaacson A, Buettner F, Knights A, Natarajan KN, HipSci Consortium, Vallier L, Marioni JC, Chhatriwala M, Stegle O. 2020. Single-cell RNA-sequencing of differentiating iPS cells reveals dynamic genetic effects on gene expression. *Nat Commun* **11**:810.
Ding M, Bhupathiraju SN, Satija A, van Dam RM, Hu FB. 2014. Long-term coffee consumption and risk of cardiovascular disease: a systematic review and a dose-response meta-analysis of prospective cohort studies. *Circulation* **129**:643–659.
Donovan MKR, D'Antonio-Chronowska A, D'Antonio M, Frazer KA. 2020. Cellular deconvolution of GTEx tissues powers discovery of disease and cell-type associated regulatory variants. *Nat Commun* **11**:955.
Du M, Yang L, Liu B, Yang L, Mao X, Liang M, Huang K. 2021. Inhibition of NFAT suppresses foam cell formation and the development of diet-induced atherosclerosis. *FASEB J* **35**:e21951.
Fabregat A, Sidiropoulos K, Viteri G, Marin-Garcia P, Ping P, Stein L, D'Eustachio P, Hermjakob H. 2018. Reactome diagram viewer: data structures and strategies to boost performance. *Bioinformatics* **34**:1208–1214.
Fairfax BP, Humburg P, Makino S, Naranbhai V, Wong D, Lau E, Jostins L, Plant K, Andrews R, McGee C, Knight JC. 2014. Innate immune activity conditions the effect of regulatory variants upon monocyte gene expression. *Science* **343**:1246949.
Fang X, Mei W, Barbazuk WB, Rivkees SA, Wendler CC. 2014. Caffeine exposure alters cardiac gene expression in embryonic cardiomyocytes. *Am J Physiol Regul Integr Comp Physiol* **307**:R1471–87.
Findley AS, Monziani A, Richards AL, Rhodes K, Ward MC, Kalita CA, Alazizi A, Pazokitoroudi A, Sankararaman S, Wen X, Lanfear DE, Pique-Regi R, Gilad Y, Luca F. 2021. Functional dynamic genetic effects on gene regulation are specific to particular cell types and environmental conditions. *Elife* **10**. doi:10.7554/eLife.67077
Findley AS, Richards AL, Petrini C, Alazizi A, Doman E, Shanku AG, Davis GO, Hauff N, Sorokin Y, Wen X, Pique-Regi R, Luca F. 2019. Interpreting Coronary Artery Disease Risk Through Gene-Environment Interactions in Gene Regulation. *Genetics* **213**:651–663.
Fryar CD, Chen T-C, Li X. 2012. Prevalence of uncontrolled risk factors for cardiovascular disease: United States, 1999-2010. *NCHS Data Brief* 1–8.
Garcia-Vaz E, McNeilly AD, Berglund LM, Ahmad A, Gallagher JR, Dutius Andersson A-M, McCrimmon RJ, Zetterqvist AV, Gomez MF, Khan F. 2020. Inhibition of NFAT Signaling Restores Microvascular Endothelial Function in Diabetic Mice. *Diabetes* **69**:424–435.
Genuario RR, Perry RP. 1996. The GA-binding protein can serve as both an activator and repressor of ribosomal protein gene transcription. *J Biol Chem* **271**:4388–4395.
Goettsch C, Rauner M, Hamann C, Sinningen K, Hempel U, Bornstein SR, Hofbauer LC. 2011. Nuclear factor of activated T cells mediates oxidised LDL-induced calcification of vascular smooth muscle cells. *Diabetologia* **54**:2690–2701.

Gong J, Shao D, Xu K, Lu Z, Lu ZJ, Yang YT, Zhang QC. 2018. RISE: a database of RNA interactome from sequencing experiments. *Nucleic Acids Res* **46**:D194–D201.

Gordon MG, Inoue F, Martin B, Schubach M, Agarwal V, Whalen S, Feng S, Zhao J, Ashuach T, Ziffra R, Kreimer A, Georgakopoulos-Soares I, Yosef N, Ye CJ, Pollard KS, Shendure J, Kircher M, Ahituv N. 2020. lentiMPRA and MPRAflow for high-throughput functional characterization of gene regulatory elements. *Nat Protoc* **15**:2387–2412.

Greißel A, Culmes M, Burgkart R, Zimmermann A, Eckstein H-H, Zernecke A, Pelisek J. 2016. Histone acetylation and methylation significantly change with severity of atherosclerosis in human carotid plaques. *Cardiovasc Pathol* **25**:79–86.

GTEx Consortium. 2020. The GTEx Consortium atlas of genetic regulatory effects across human tissues. *Science* **369**:1318–1330.

Gupta RM, Hadaya J, Trehan A, Zekavat SM, Roselli C, Klarin D, Emdin CA, Hilvering CRE, Bianchi V, Mueller C, Khera AV, Ryan RJH, Engreitz JM, Issner R, Shoresh N, Epstein CB, de Laat W, Brown JD, Schnabel RB, Bernstein BE, Kathiresan S. 2017. A Genetic Variant Associated with Five Vascular Diseases Is a Distal Regulator of Endothelin-1 Gene Expression. *Cell* **170**:522–533.e15.

Hadi HAR, Carr CS, Al Suwaidi J. 2005. Endothelial dysfunction: cardiovascular risk factors, therapy, and outcome. *Vasc Health Risk Manag* **1**:183–198.

Hajar R. 2017. Risk Factors for Coronary Artery Disease: Historical Perspectives. *Heart Views* **18**:109–114.

Hartmann K, Seweryn M, Sadee W. 2022. Interpreting coronary artery disease GWAS results: A functional genomics approach assessing biological significance. *PLoS One* **17**:e0244904.

Hirotsu Y, Hataya N, Katsuoka F, Yamamoto M. 2012. NF-E2-related factor 1 (Nrf1) serves as a novel regulator of hepatic lipid metabolism through regulation of the Lipin1 and PGC-1β genes. *Mol Cell Biol* **32**:2760–2770.

Horton JD, Shah NA, Warrington JA, Anderson NN, Park SW, Brown MS, Goldstein JL. 2003. Combined analysis of oligonucleotide microarray data from transgenic and knockout mice identifies direct SREBP target genes. *Proc Natl Acad Sci U S A* **100**:12027–12032.

Huang QQ, Tang HHF, Teo SM, Mok D, Ritchie SC, Nath AP, Brozynska M, Salim A, Bakshi A, Holt BJ, Khor CC, Sly PD, Holt PG, Holt KE, Inouye M. 2020. Neonatal genetics of gene expression reveal potential origins of autoimmune and allergic disease risk. *Nat Commun* **11**:3761.

Huss JM, Kelly DP. 2004. Nuclear receptor signaling and cardiac energetics. *Circ Res* **95**:568–578.

Innocenti F, Cooper GM, Stanaway IB, Gamazon ER, Smith JD, Mirkov S, Ramirez J, Liu W, Lin YS, Moloney C, Aldred SF, Trinklein ND, Schuetz E, Nickerson DA, Thummel KE, Rieder MJ, Rettie AE, Ratain MJ, Cox NJ, Brown CD. 2011. Identification, replication, and functional fine-mapping of expression quantitative trait loci in primary human liver tissue. *PLoS Genet* **7**:e1002078.

Johnson EN, Lee YM, Sander TL, Rabkin E, Schoen FJ, Kaushal S, Bischoff J. 2003. NFATc1 mediates vascular endothelial growth factor-induced proliferation of human pulmonary valve endothelial cells. *J Biol Chem* **278**:1686–1692.

Johnson GD, Barrera A, McDowell IC, D'Ippolito AM, Majoros WH, Vockley CM, Wang X, Allen AS, Reddy TE. 2018. Human genome-wide measurement of drug-responsive regulatory activity. *Nat Commun* **9**:5317.

Kalita CA, Brown CD, Freiman A, Isherwood J, Wen X, Pique-Regi R, Luca F. 2018. High

throughput characterization of genetic effects on DNA:protein binding and gene transcription. *Genome Res*. doi:10.1101/gr.237354.118

Kalita CA, Moyerbrailean GA, Brown C, Wen X, Luca F, Pique-Regi R. 2017. QuASAR-MPRA: Accurate allele-specific analysis for massively parallel reporter assays. *Bioinformatics*. doi:10.1093/bioinformatics/btx598

Kessler T, Schunkert H. 2021. Coronary Artery Disease Genetics Enlightened by Genome-Wide Association Studies. *JACC Basic Transl Sci* **6**:610–623.

Kim-Hellmuth S, Aguet F, Oliva M, Muñoz-Aguirre M, Kasela S, Wucher V, Castel SE, Hamel AR, Viñuela A, Roberts AL, Mangul S, Wen X, Wang G, Barbeira AN, Garrido-Martín D, Nadel BB, Zou Y, Bonazzola R, Quan J, Brown A, Martinez-Perez A, Soria JM, GTEx Consortium, Getz G, Dermitzakis ET, Small KS, Stephens M, Xi HS, Im HK, Guigó R, Segrè AV, Stranger BE, Ardlie KG, Lappalainen T. 2020. Cell type-specific genetic regulation of gene expression across human tissues. *Science* **369**. doi:10.1126/science.aaz8528

Kim-Hellmuth S, Bechheim M, Pütz B, Mohammadi P, Nédélec Y, Giangreco N, Becker J, Kaiser V, Fricker N, Beier E, Boor P, Castel SE, Nöthen MM, Barreiro LB, Pickrell JK, Müller-Myhsok B, Lappalainen T, Schumacher J, Hornung V. 2017. Genetic regulatory effects modified by immune activation contribute to autoimmune disease associations. *Nat Commun* **8**:266.

Klee CB, Ren H, Wang X. 1998. Regulation of the calmodulin-stimulated protein phosphatase, calcineurin. *J Biol Chem* **273**:13367–13370.

Knowles DA, Burrows CK, Blischak JD, Patterson KM, Serie DJ, Norton N, Ober C, Pritchard JK, Gilad Y. 2018. Determining the genetic basis of anthracycline-cardiotoxicity by molecular response QTL mapping in induced cardiomyocytes. *Elife* **7**. doi:10.7554/eLife.33480

Knowles DA, Davis JR, Edgington H, Raj A, Favé M-J, Zhu X, Potash JB, Weissman MM, Shi J, Levinson DF, Awadalla P, Mostafavi S, Montgomery SB, Battle A. 2017. Allele-specific expression reveals interactions between genetic variation and environment. *Nat Methods*. doi:10.1038/nmeth.4298

Kotla S, Vu HT, Ko KA, Wang Y, Imanishi M, Heo K-S, Fujii Y, Thomas TN, Gi YJ, Mazhar H, Paez-Mayorga J, Shin J-H, Tao Y, Giancursio CJ, Medina JL, Taunton J, Lusis AJ, Cooke JP, Fujiwara K, Le N-T, Abe J-I. 2019. Endothelial senescence is induced by phosphorylation and nuclear export of telomeric repeat binding factor 2-interacting protein. *JCI Insight* **4**. doi:10.1172/jci.insight.124867

Koyama S, Ito K, Terao C, Akiyama M, Horikoshi M, Momozawa Y, Matsunaga H, Ieki H, Ozaki K, Onouchi Y, Takahashi A, Nomura S, Morita H, Akazawa H, Kim C, Seo J-S, Higasa K, Iwasaki M, Yamaji T, Sawada N, Tsugane S, Koyama T, Ikezaki H, Takashima N, Tanaka K, Arisawa K, Kuriki K, Naito M, Wakai K, Suna S, Sakata Y, Sato H, Hori M, Sakata Y, Matsuda K, Murakami Y, Aburatani H, Kubo M, Matsuda F, Kamatani Y, Komuro I. 2020. Population-specific and trans-ancestry genome-wide analyses identify distinct and shared genetic risk loci for coronary artery disease. *Nat Genet* **52**:1169–1177.

Krüger-Genge A, Blocki A, Franke R-P, Jung F. 2019. Vascular Endothelial Cell Biology: An Update. *Int J Mol Sci* **20**. doi:10.3390/ijms20184411

Lebeau PF, Byun JH, Platko K, Saliba P, Sguazzin M, MacDonald ME, Paré G, Steinberg GR, Janssen LJ, Igdoura SA, Tarnopolsky MA, Wayne Chen SR, Seidah NG, Magolan J, Austin RC. 2022. Caffeine blocks SREBP2-induced hepatic PCSK9 expression to enhance LDLR-

mediated cholesterol clearance. *Nat Commun* **13**:770.
Lee MN, Ye C, Villani A-C, Raj T, Li W, Eisenhaure TM, Imboywa SH, Chipendo PI, Ran FA, Slowikowski K, Ward LD, Raddassi K, McCabe C, Lee MH, Frohlich IY, Hafler DA, Kellis M, Raychaudhuri S, Zhang F, Stranger BE, Benoist CO, De Jager PL, Regev A, Hacohen N. 2014. Common genetic variants modulate pathogen-sensing responses in human dendritic cells. *Science* **343**:1246980.
Li H, Li Q, Zhang Y, Liu W, Gu B, Narumi T, Siu KL, Youn JY, Liu P, Yang X, Cai H. 2019. Novel Treatment of Hypertension by Specifically Targeting E2F for Restoration of Endothelial Dihydrofolate Reductase and eNOS Function Under Oxidative Stress. *Hypertension* **73**:179–189.
Love MI, Huber W, Anders S. 2014. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol* **15**:550.
Mangravite LM, Engelhardt BE, Medina MW, Smith JD, Brown CD, Chasman DI, Mecham BH, Howie B, Shim H, Naidoo D, Feng Q, Rieder MJ, Chen Y-DI, Rotter JI, Ridker PM, Hopewell JC, Parish S, Armitage J, Collins R, Wilke RA, Nickerson DA, Stephens M, Krauss RM. 2013. A statin-dependent QTL for GATM expression is associated with statin-induced myopathy. *Nature* **502**:377–380.
Manry J, Nédélec Y, Fava VM, Cobat A, Orlova M, Van Thuc N, Thai VH, Laval G, Barreiro LB, Schurr E. 2017. Deciphering the genetic control of gene expression following Mycobacterium leprae antigen stimulation. *PLoS Genet* **13**:e1006952.
Maranville JC, Luca F, Richards AL, Wen X, Witonsky DB, Baxter S, Stephens M, Rienzo A. 2011. Interactions between glucocorticoid treatment and Cis-regulatory polymorphisms contribute to cellular response phenotypes. *PLoS Genet* **7**. doi:10.1371/journal.pgen.1002162
Matsuzawa Y, Lerman A. 2014. Endothelial dysfunction and coronary artery disease: assessment, prognosis, and treatment. *Coron Artery Dis* **25**:713–724.
Medina-Leyte DJ, Zepeda-García O, Domínguez-Pérez M, González-Garrido A, Villarreal-Molina T, Jacobo-Albavera L. 2021. Endothelial Dysfunction, Inflammation and Coronary Artery Disease: Potential Biomarkers and Promising Therapeutical Approaches. *Int J Mol Sci* **22**. doi:10.3390/ijms22083850
Melnikov A, Murugan A, Zhang X, Tesileanu T, Wang L, Rogov P, Feizi S, Gnirke A, Callan CG Jr, Kinney JB, Kellis M, Lander ES, Mikkelsen TS. 2012. Systematic dissection and optimization of inducible enhancers in human cells using a massively parallel reporter assay. *Nat Biotechnol* **30**:271–277.
Miranda AM, Steluti J, Goulart AC, Benseñor IM, Lotufo PA, Marchioni DM. 2018. Coffee Consumption and Coronary Artery Calcium Score: Cross-Sectional Results of ELSA-Brasil (Brazilian Longitudinal Study of Adult Health). *J Am Heart Assoc* **7**. doi:10.1161/JAHA.117.007155
Moyerbrailean GA, Kalita CA, Harvey CT, Wen X, Luca F, Pique-Regi R. 2016a. Which Genetics Variants in DNase-Seq Footprints Are More Likely to Alter Binding? *PLoS Genet* **12**:e1005875.
Moyerbrailean GA, Richards AL, Kurtz D, Kalita CA, Davis GO, Harvey CT, Alazizi A, Watza D, Sorokin Y, Hauff N, Zhou X, Wen X, Pique-Regi R, Luca F. 2016b. High-throughput allele-specific expression across 250 environmental conditions. *Genome Res* **26**:1627–1638.
Nédélec Y, Sanz J, Baharian G, Szpiech ZA, Pacis A, Dumaine A, Grenier J-C, Freiman A, Sams AJ, Hebert S, Pagé Sabourin A, Luca F, Blekhman R, Hernandez RD, Pique-Regi R,

Tung J, Yotova V, Barreiro LB. 2016. Genetic Ancestry and Natural Selection Drive Population Differences in Immune Responses to Pathogens. *Cell* **167**:657–669.e21.

Nikpay M, Goel A, Won H-H, Hall LM, Willenborg C, Kanoni S, Saleheen D, Kyriakou T, Nelson CP, Hopewell JC, Webb TR, Zeng L, Dehghan A, Alver M, Armasu SM, Auro K, Bjonnes A, Chasman DI, Chen S, Ford I, Franceschini N, Gieger C, Grace C, Gustafsson S, Huang J, Hwang S-J, Kim YK, Kleber ME, Lau KW, Lu X, Lu Y, Lyytikäinen L-P, Mihailov E, Morrison AC, Pervjakova N, Qu L, Rose LM, Salfati E, Saxena R, Scholz M, Smith AV, Tikkanen E, Uitterlinden A, Yang X, Zhang W, Zhao W, de Andrade M, de Vries PS, van Zuydam NR, Anand SS, Bertram L, Beutner F, Dedoussis G, Frossard P, Gauguier D, Goodall AH, Gottesman O, Haber M, Han B-G, Huang J, Jalilzadeh S, Kessler T, König IR, Lannfelt L, Lieb W, Lind L, Lindgren CM, Lokki M-L, Magnusson PK, Mallick NH, Mehra N, Meitinger T, Memon F-U-R, Morris AP, Nieminen MS, Pedersen NL, Peters A, Rallidis LS, Rasheed A, Samuel M, Shah SH, Sinisalo J, Stirrups KE, Trompet S, Wang L, Zaman KS, Ardissino D, Boerwinkle E, Borecki IB, Bottinger EP, Buring JE, Chambers JC, Collins R, Cupples LA, Danesh J, Demuth I, Elosua R, Epstein SE, Esko T, Feitosa MF, Franco OH, Franzosi MG, Granger CB, Gu D, Gudnason V, Hall AS, Hamsten A, Harris TB, Hazen SL, Hengstenberg C, Hofman A, Ingelsson E, Iribarren C, Jukema JW, Karhunen PJ, Kim B-J, Kooner JS, Kullo IJ, Lehtimäki T, Loos RJF, Melander O, Metspalu A, März W, Palmer CN, Perola M, Quertermous T, Rader DJ, Ridker PM, Ripatti S, Roberts R, Salomaa V, Sanghera DK, Schwartz SM, Seedorf U, Stewart AF, Stott DJ, Thiery J, Zalloua PA, O'Donnell CJ, Reilly MP, Assimes TL, Thompson JR, Erdmann J, Clarke R, Watkins H, Kathiresan S, McPherson R, Deloukas P, Schunkert H, Samani NJ, Farrall M. 2015. A comprehensive 1,000 Genomes-based genome-wide association meta-analysis of coronary artery disease. *Nat Genet* **47**:1121–1130.

Okamoto J, Wang L, Yin X, Luca F, Pique-Regi R, Helms A, Im HK, Morrison J, Wen X. 2023. Probabilistic integration of transcriptome-wide association studies and colocalization analysis identifies key molecular pathways of complex traits. *Am J Hum Genet* **110**:44–57.

Patwardhan RP, Hiatt JB, Witten DM, Kim MJ, Smith RP, May D, Lee C, Andrie JM, Lee S-I, Cooper GM, Ahituv N, Pennacchio LA, Shendure J. 2012. Massively parallel functional dissection of mammalian enhancers in vivo. *Nat Biotechnol* **30**:265–270.

Pickrell JK. 2014. Joint analysis of functional genomic data and genome-wide association studies of 18 human traits. *Am J Hum Genet* **94**:559–573.

Planning Committee for a Workshop on Potential Health Hazards Associated with Consumption of Caffeine in Food and Dietary Supplements, Food and Nutrition Board, Board on Health Sciences Policy, Institute of Medicine. 2014. Caffeine in Food and Dietary Supplements: Examining Safety: Workshop Summary. Washington (DC): National Academies Press (US).

Quach H, Rotival M, Pothlichet J, Loh Y-HE, Dannemann M, Zidane N, Laval G, Patin E, Harmant C, Lopez M, Deschamps M, Naffakh N, Duffy D, Coen A, Leroux-Roels G, Clément F, Boland A, Deleuze J-F, Kelso J, Albert ML, Quintana-Murci L. 2016. Genetic Adaptation and Neandertal Admixture Shaped the Immune System of Human Populations. *Cell* **167**:643–656.e17.

Shlyueva D, Stelzer C, Gerlach D, Yáñez-Cuna JO, Rath M, Boryń ŁM, Arnold CD, Stark A. 2014. Hormone-responsive enhancer-activity maps reveal predictive motifs, indirect repression, and targeting of closed chromatin. *Mol Cell* **54**:180–192.

Stevens LM, Linstead E, Hall JL, Kao DP. 2021. Association Between Coffee Intake and Incident Heart Failure Risk: A Machine Learning Analysis of the FHS, the ARIC Study, and the CHS. *Circ Heart Fail* **14**:e006799.

Strober BJ, Elorbany R, Rhodes K, Krishnan N, Tayeb K, Battle A, Gilad Y. 2019. Dynamic genetic regulation of gene expression during cellular differentiation. *Science* **364**:1287–1290.

Temprano-Sagrera G, Sitlani CM, Bone WP, Martin-Bornez M, Voight BF, Morrison AC, Damrauer SM, de Vries PS, Smith NL, Sabater-Lleal M. 2022. Multi-phenotype analyses of hemostatic traits with cardiovascular events reveal novel genetic associations. *J Thromb Haemost* **20**:1331–1349.

Tewhey R, Kotliar D, Park DS, Liu B, Winnicki S, Reilly SK, Andersen KG, Mikkelsen TS, Lander ES, Schaffner SF, Sabeti PC. 2016. Direct Identification of Hundreds of Expression-Modulating Variants using a Multiplexed Reporter Assay. *Cell* **165**:1519–1529.

Turnbull D, Rodricks JV, Mariano GF, Chowdhury F. 2017. Caffeine and cardiovascular health. *Regul Toxicol Pharmacol* **89**:165–185.

Ulirsch JC, Nandakumar SK, Wang L, Giani FC, Zhang X, Rogov P, Melnikov A, McDonel P, Do R, Mikkelsen TS, Sankaran VG. 2016. Systematic Functional Dissection of Common Genetic Variation Affecting Red Blood Cell Traits. *Cell* **165**:1530–1545.

van der Harst P, Verweij N. 2018. Identification of 64 Novel Genetic Loci Provides an Expanded View on the Genetic Architecture of Coronary Artery Disease. *Circ Res* **122**:433–443.

Vockley CM, Guo C, Majoros WH, Nodzenski M, Scholtens DM, Hayes MG, Lowe WL Jr, Reddy TE. 2015. Massively parallel quantification of the regulatory effects of noncoding genetic variation in a human cohort. *Genome Res* **25**:1206–1214.

Voskoboinik A, Koh Y, Kistler PM. 2019. Cardiovascular effects of caffeinated beverages. *Trends Cardiovasc Med* **29**:345–350.

Wang L, Fan C, Topol SE, Topol EJ, Wang Q. 2003. Mutation of MEF2A in an inherited disorder with features of coronary artery disease. *Science* **302**:1578–1581.

Wang X, He L, Goggin SM, Saadat A, Wang L, Sinnott-Armstrong N, Claussnitzer M, Kellis M. 2018. High-resolution genome-wide functional dissection of transcriptional regulatory regions and nucleotides in human. *Nat Commun* **9**:5380.

Wang X, Musunuru K. 2018. Confirmation of Causal rs9349379- PHACTR1 Expression Quantitative Trait Locus in Human-Induced Pluripotent Stem Cell Endothelial Cells. *Circ Genom Precis Med* **11**:e002327.

Wang Y, Hu J, Liu J, Geng Z, Tao Y, Zheng F, Wang Y, Fu S, Wang W, Xie C, Zhang Y, Gong F. 2020. The role of Ca2+/NFAT in Dysfunction and Inflammation of Human Coronary Endothelial Cells induced by Sera from patients with Kawasaki disease. *Sci Rep* **10**:4706.

Weng L, Kavaslar N, Ustaszewska A, Doelle H, Schackwitz W, Hébert S, Cohen JC, McPherson R, Pennacchio LA. 2005. Lack of MEF2A mutations in coronary artery disease. *J Clin Invest* **115**:1016–1020.

Wen X. 2016. Molecular QTL discovery incorporating genomic annotations using Bayesian false discovery rate control. *aoas* **10**:1619–1638.

Wen X, Luca F, Pique-Regi R. 2015. Cross-population joint analysis of eQTLs: fine mapping and functional annotation. *PLoS Genet* **11**:e1005176.

Wen X, Pique-Regi R, Luca F. 2017. Integrating molecular QTL data into genome-wide genetic association analysis: Probabilistic assessment of enrichment and colocalization. *PLoS Genet* **13**:e1006646.

Xu S, Ilyas I, Little PJ, Li H, Kamato D, Zheng X, Luo S, Li Z, Liu P, Han J, Harding IC, Ebong EE, Cameron SJ, Stewart AG, Weng J. 2021. Endothelial Dysfunction in Atherosclerotic Cardiovascular Diseases and Beyond: From Mechanism to Pharmacotherapies. *Pharmacol Rev* **73**:924–967.

You B, Jiang Y-Y, Chen S, Yan G, Sun J. 2009. The orphan nuclear receptor Nur77 suppresses endothelial cell activation through induction of IkappaBalpha expression. *Circ Res* **104**:742–749.

Zetterqvist AV, Berglund LM, Blanco F, Garcia-Vaz E, Wigren M, Dunér P, Andersson A-MD, To F, Spegel P, Nilsson J, Bengtsson E, Gomez MF. 2014. Inhibition of nuclear factor of activated T-cells (NFAT) suppresses accelerated atherosclerosis in diabetic mice. *PLoS One* **8**:e65020.

Zhang W, Song M, Qu J, Liu G-H. 2018. Epigenetic Modifications in Cardiovascular Aging and Diseases. *Circ Res* **123**:773–786.

Zhang Y, Quick C, Yu K, Barbeira A, GTEx Consortium, Luca F, Pique-Regi R, Kyung Im H, Wen X. 2020. PTWAS: investigating tissue-relevant causal molecular mechanisms of complex traits using probabilistic TWAS analysis. *Genome Biol* **21**:232.