

## Metabolic reaction fluxes as amplifiers and buffers of risk alleles for coronary artery disease

Carles Foguet<sup>1,2,3\*</sup>, Xilin Jiang<sup>1,2,3,4</sup>, Scott C. Ritchie<sup>1,2,3,5,6,7</sup>, Elodie Persyn<sup>1,2,3</sup>, Yu Xu<sup>1,2,3</sup>, Chief Ben-Eghan<sup>1,2,3</sup>, Emanuele Di Angelantonio<sup>2,3,5,6,8,9</sup>, John Danesh<sup>2,3,5,6,8,10</sup>, Adam S. Butterworth<sup>2,3,5,6,8</sup>, Samuel A. Lambert<sup>1,2,3,6,11</sup>, Michael Inouye<sup>1,2,3,5,6,7\*</sup>

1. Cambridge Baker Systems Genomics Initiative, Department of Public Health and Primary Care, University of Cambridge, Cambridge, UK.
2. British Heart Foundation Cardiovascular Epidemiology Unit, Department of Public Health and Primary Care, University of Cambridge, Cambridge, UK.
3. Victor Phillip Dahdaleh Heart and Lung Research Institute, University of Cambridge, Cambridge, UK.
4. Department of Epidemiology, Harvard T.H. Chan School of Public Health, Boston, MA, USA.
5. British Heart Foundation Centre of Research Excellence, University of Cambridge, Cambridge, UK.
6. Health Data Research UK Cambridge, Wellcome Genome Campus and University of Cambridge, Cambridge, UK.
7. Cambridge Baker Systems Genomics Initiative, Baker Heart and Diabetes Institute, Melbourne, VIC, Australia.
8. National Institute for Health and Care Research Blood and Transplant Research Unit in Donor Health and Behaviour, University of Cambridge, Cambridge, UK.
9. Health Data Science Research Centre, Fondazione Human Technopole, Milan, Italy.
10. Department of Human Genetics, the Wellcome Trust Sanger Institute, Wellcome Trust Genome Campus, Hinxton, UK.
11. European Molecular Biology Laboratory, European Bioinformatics Institute, Wellcome Genome Campus, Hinxton, Cambridge, UK.

\* Corresponding authors: [cf545@medschl.cam.ac.uk](mailto:cf545@medschl.cam.ac.uk) (C.F.) & [mi336@medschl.cam.ac.uk](mailto:mi336@medschl.cam.ac.uk) (M.I.)

### **Abstract**

Genome-wide association studies have identified thousands of variants associated with disease risk but the mechanism by which such variants contribute to disease remains largely unknown. Indeed, a major challenge is that variants do not act in isolation but rather in the framework of highly complex biological networks, such as the human metabolic network, which can amplify or buffer the effect of specific risk alleles on disease susceptibility. In our previous work, we established that metabolic models can be leveraged to simulate the emerging metabolic effects of genetically driven variation in transcript levels and estimate personalized metabolic reaction fluxes. Here we use genetically predicted reaction fluxes to perform a systematic search for metabolic fluxes acting as buffers or amplifiers of coronary artery disease (CAD) risk alleles. Our analysis identifies 30 risk locus - reaction flux pairs with significant interaction on CAD susceptibility involving 18 individual reaction fluxes and 8 independent risk loci. Notably, many of these reactions are linked to processes with putative roles in the disease such as the metabolism of inflammatory mediators and fatty acids. In summary, this work establishes proof of concept that biochemical reaction fluxes can have non-additive effects with risk alleles and provides novel insights into the interplay between metabolism and genetic variation on disease susceptibility.

NOTE: This preprint reports new research that has not been certified by peer review and should not be used to guide clinical practice.

## **Introduction**

Genome-wide association studies (GWAS) have identified tens of thousands of single nucleotide polymorphisms (SNPs) associated with disease risk<sup>1</sup>. However, variant to function (V2F) remains largely unsolved limiting the potential to leverage GWAS results to uncover the underlying disease biology and identify novel therapeutic targets<sup>2,3</sup>. Genetic variants do not act in isolation but within the context of highly complex biological networks that can modulate the effect of specific alleles on disease susceptibility<sup>4,5</sup>. Therefore, unveiling non-additive effects between genetic variants and environment or genetic factors is a powerful approach to understanding the functional relationships of genetic variants and their role in disease<sup>6,7</sup>.

Metabolism is a major biological network comprised of metabolites, enzymes, and transmembrane carriers and underlies many processes in health and disease<sup>8</sup>. One of the most direct manifestations of the metabolic phenotype are metabolic fluxes: the rate at which substrates are converted to products in biochemical reactions or transported across compartments in a metabolic network<sup>9,10</sup>. Dysregulation of metabolic fluxes can play a major role in disease onset and progression<sup>11–14</sup>. For instance, it is well established that alterations in lipid metabolic pathways can disrupt the concentration of lipids in blood and promote the formation of atherosclerotic lesions<sup>15,16</sup>.

Furthermore, metabolic fluxes are particularly attractive as therapeutic targets for perturbation as it has been shown that they can be safely modulated to minimise disease risk and progression. For instance, statins, which are widely prescribed to reduce cardiovascular disease risk, act by reducing the flux of cholesterol synthesis through the inhibition of the enzyme HMG-CoA Reductase<sup>17</sup>. Statins can be combined with inhibitors of bile acid reabsorption, which can further reduce cholesterol levels by increasing the flux of bile acid synthesis<sup>16,18</sup>. Similarly, one of the mechanisms of action of the type 2 diabetes drug metformin, which is also reported to have cardioprotective effects<sup>19</sup>, is the reduction of the oxidative phosphorylation flux<sup>20</sup>.

Notably, intracellular metabolic fluxes cannot be directly measured and instead must be predicted from data integrated into the framework of metabolic networks<sup>9,10</sup>. We have previously demonstrated the feasibility and insights that can be gleaned from using a genotyped population-based biobank to estimate metabolic reaction fluxes and perform a fluxome-wide association analysis for coronary artery disease (CAD)<sup>21</sup>. This analysis was performed using genetically predicted fluxes derived from the integration of genetically imputed transcript abundances<sup>22</sup> within the constraints defined by the stoichiometric relationships of enzymes and transmembrane carriers in human genome-scale metabolic networks<sup>23</sup>. We found that fluxes through several reactions linked to the metabolism of inflammatory mediators (e.g., histamine and prostaglandins) and polyamines were strongly associated with CAD risk<sup>21</sup>.

Since alterations in metabolic function contribute to disease susceptibility and progression, here we set out to test the hypothesis that metabolic fluxes can also act as buffers or amplifiers of the effects of risk alleles. We performed a systematic search for interactions between genetically personalized organ-specific fluxes and the dosage of risk alleles on CAD susceptibility in UK Biobank (UKB)<sup>24,25</sup>. In doing so, we identified numerous instances where reaction fluxes significantly amplify or buffer the effects of genetic variants on CAD risk.

## **Results**

### **Overview of the methods**

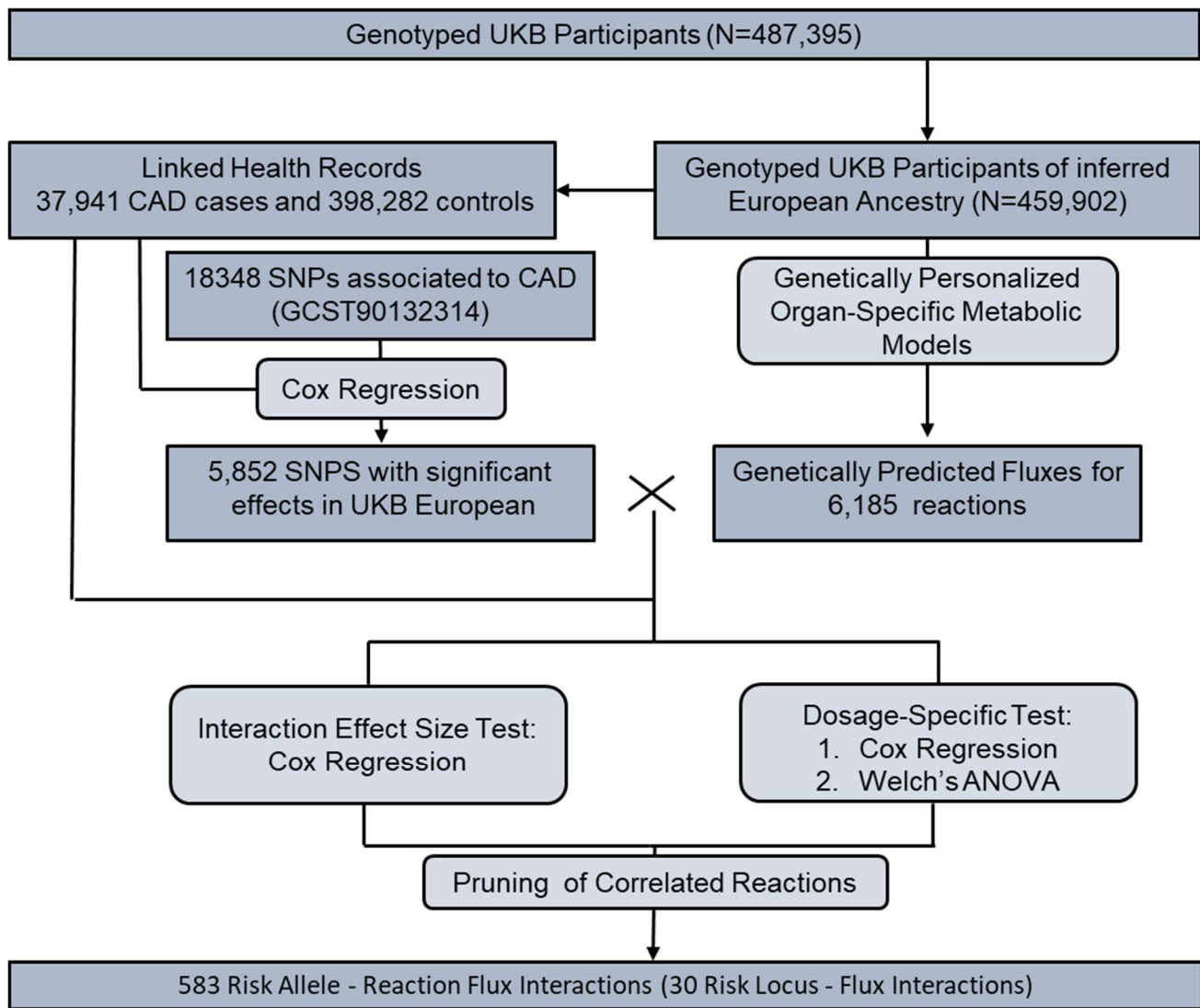
We estimated genetically personalized metabolic reaction fluxes for 6,185 organ-specific reactions in 459,902 UK Biobank participants<sup>24,25</sup> of European genetic ancestries<sup>26</sup> (**Methods**). From hospital episode statistics and cause of death records for these participants, there were 37,941 CAD cases in total (combined prevalent and incident cases) using the PheWAS Catalog definition of coronary atherosclerosis<sup>27</sup>, and 398,282 non-CAD controls (N=398,282). Using a published genome-wide association meta-analysis performed on over one million participants of European ancestry<sup>28</sup>, we extracted a set of 18,348 SNPs associated with CAD risk ( $P < 5 \times 10^{-8}$ ). For downstream analyses, we used the subset of 5,852 SNPs which were found significantly associated with CAD in European UKB participants ( $P < 5 \times 10^{-8}$ ; Cox regression model; **Methods**; **Figure 1**).

We next assessed the extent to which each of the 6,185 reaction fluxes either buffered (i.e. negative interaction) or amplified (i.e. positive interaction) the effects of the 5,852 risk alleles on CAD (**Methods**). To robustly identify such events, we used the intersection of two complementary approaches with different statistical assumptions. In the first approach, we tested for a significant interaction effect size between risk allele dosage and reaction flux value using a Cox proportional-hazards model for CAD events. Before testing for interaction, the effect of the risk allele was regressed from reaction flux to control for potential false positives arising from dependencies between genetically predicted flux values and risk allele dosage<sup>29</sup>. In the second approach, termed dosage-specific test, we tested the differences in reaction flux effect sizes on CAD risk between individuals carrying different risk allele dosages (**Figure S1**; **Methods**). The P-values from the interaction model and the dosage-specific test were then adjusted for multiple testing using the Benjamini-Hochberg method (i.e. FDR). A buffering or amplification effect of a reaction flux on a risk allele had to be significant under both approaches (FDR-adjusted P-value  $< 0.05$ ; **Figure 1**) to be deemed valid.

### **Buffering and amplification of the effect of risk variants by reaction fluxes**

In total, we found 583 pairs of SNP-reaction fluxes which were significant (FDR-adjusted P-value  $< 0.05$ ) in both the interaction effect size test and the dosage-specific test (separately, 669 and 595 SNP-reaction flux pairs were significant, respectively) (**Supplementary Data S1**). We observed strong correlations between interaction effect estimates ( $r = 0.998$ ) and the p-values ( $r = 0.713$ ) of both approaches (**Figure S2**). Indeed, from the 86 pairs significant for the interaction effect size test but not the dosage-specific test, 82 were significant with an FDR-adjusted P-value  $< 0.25$  in the latter. Similarly, 7 out of the 12 pairs significant with the dosage-specific test but not the interaction effect size test were also borderline significant, with the remaining pairs being instances where the interaction might significantly deviate from linearity.

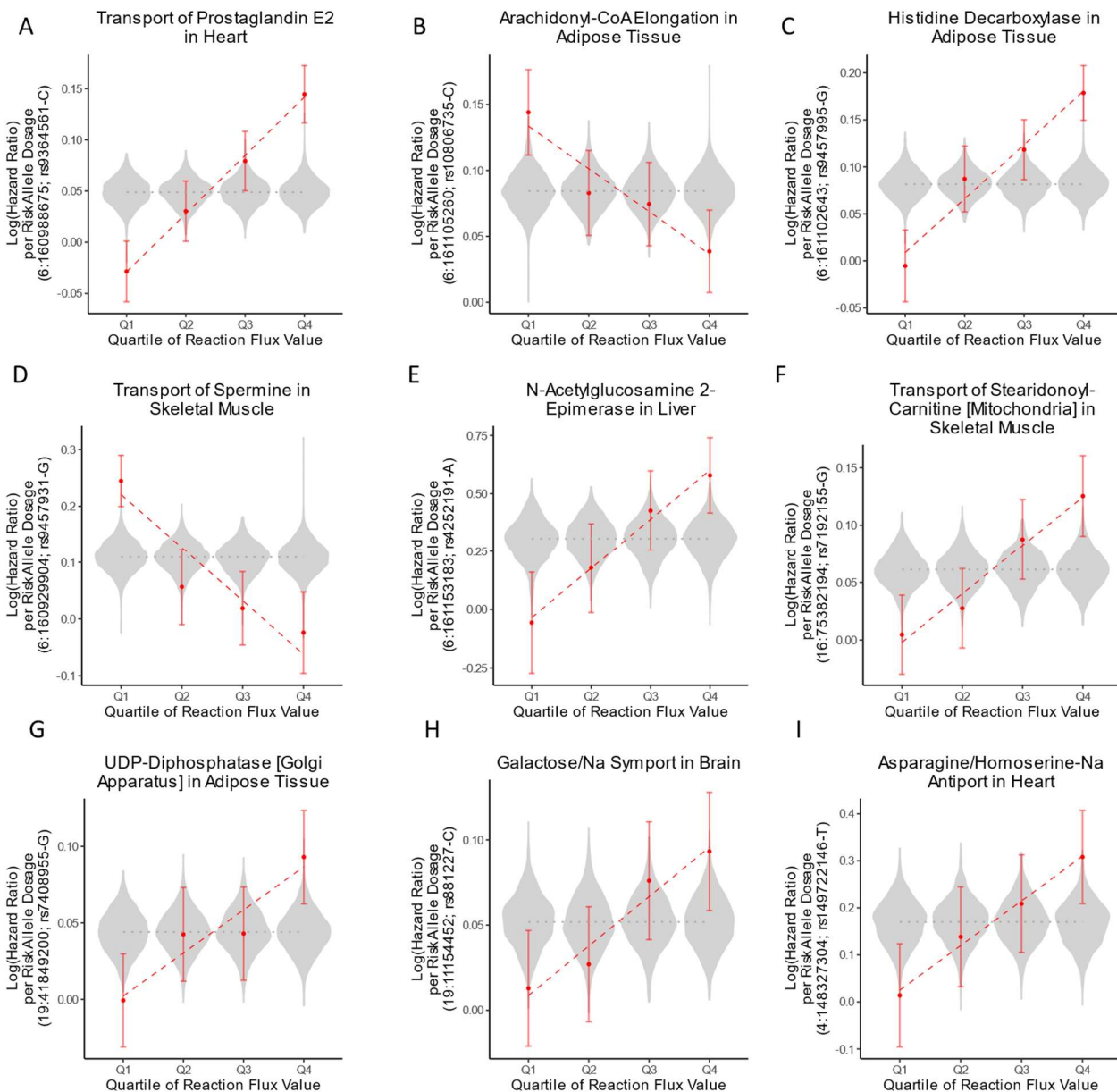
Of the 583 SNP-reaction flux pairs, 353 displayed buffering (i.e. negative interaction effect size) and 230 displayed amplification (i.e. positive interaction effect size) of the effects of the risk allele by reaction fluxes (**Figure 2**). The significant pairs comprised 279 unique SNPs mapped to 8 independent risk loci ( $R^2 < 0.6$ ; **Methods**) leading to 30 risk locus – reaction flux pairs with significant interaction on disease susceptibility. These interactions encompassed 18 unique reaction fluxes. Notably, only five of these fluxes had a significant effect on CAD risk in univariate association analysis, indicating that the majority of the metabolic associations here unveiled have their association with CAD masked unless analysed together with common risk alleles (**Figure S3**).



**Figure 1:** Summary of the methods to detect interactions between reaction fluxes and risk alleles on CAD risk.

### The genomic region encoding Lp(a) and plasminogen is a major site for amplification and buffering

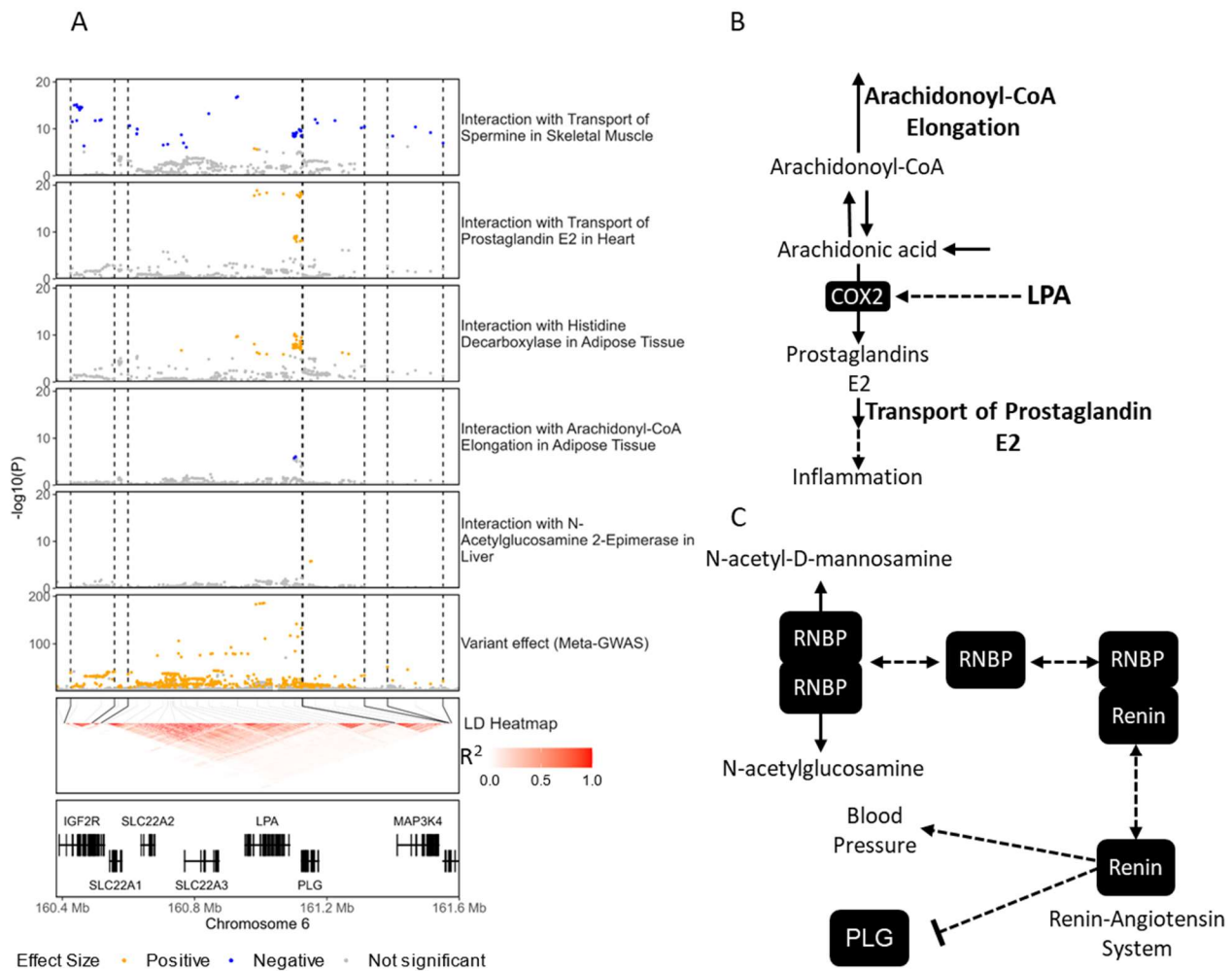
The majority (530 out of 583) of the SNP-reaction flux pairs with significant interaction on CAD risk were mapped to four risk loci in the chromosome 6 genomic region encoding the *LPA* and *PLG* genes (**Figure 2A-E; Figure 3A; Figure S1A-E; Figure S4; Supplementary Data S1**). *LPA* codes for apolipoprotein(a) which is the primary constituent of Lp(a) and has been established to have a causal role in the formation of atherosclerotic lesions by promoting lipid accumulation, inflammation, and calcification in the artery wall<sup>30-32</sup>. *PLG* encodes plasminogen which can contribute to atherosclerosis by modulating, cell migration, extracellular matrix structure, vascular smooth muscle cell (VSMC) function, and inflammation<sup>33-35</sup>.



**Figure 2:** Examples of buffering and amplification of variant effect sizes by reaction fluxes on CAD risk. For each reaction, UKB participants of European genetic ancestries were quartile-binned according to the personalized reaction flux value, and variant effect sizes (Log(Hazard Ratio)) were estimated within each subset using Cox regression. Error bars denote the 95% confidence intervals for variant effect sizes. The dashed lines indicate the linear regression of variant effect sizes per flux quartile. Violin plots indicate the distribution of variant effect size on CAD risk for all other reaction fluxes. The dotted grey line indicates variant effect size in all UKB participants of European genetic ancestries. Genome coordinates correspond to the GRCh37 genome assembly.

The flux of prostaglandin E2 transport in both heart and brain tissues strongly amplifies the effect size of a set of risk variants mapped to the risk locus which includes the *LPA* gene (Figure 2A; Figure 3A; Figure S1A; Figure S4). Prostaglandin E2 is an inflammatory mediator that plays a role in the pathogenesis of atherosclerosis<sup>36</sup>. Interestingly, the flux of elongation of arachidonoyl-CoA in adipose tissue has the opposite effect and buffers the effect size of two SNPs in the same risk locus (Figure 2B; Figure S1B). Arachidonoyl-CoA is the CoA-conjugated form of arachidonic acid which is a precursor for prostaglandin synthesis<sup>23</sup> (Figure 3B). Similarly, variants in the same locus also had their effect size amplified by the flux of histamine synthesis in adipose tissue (Figure 2C; Figure S1C) and a flux of histamine transport into the liver (Figure S4). Histamine is also an inflammatory

mediator that has been linked to atherosclerosis by modulating inflammation, blood lipids, and lipoprotein fractions<sup>37,38</sup>.



**Figure 3:** Examples of amplification and buffering at the *LPA/PLG* risk loci. A) Regional association plots showing the  $-\log_{10}(P\text{-value})$  for interaction and variant effect sizes on CAD risk. P-values for interaction and variant effect sizes were derived from the interaction effect size test and meta-GWAS summary statistics, respectively. Only a representative subset of reactions involved in significant interactions are shown, the complete set of interactions for this region is shown in **Figure S4**. The LD heatmap indicates the pairwise LD for SNPs with genome-wide significant effect size on CAD in UKB participants of European genetic ancestries. Dashed black lines indicate the limits of LD blocks ( $R^2 > 0.6$ ) used to define independent risk loci. To facilitate visualization, only LD blocks with variants involved in significant interactions are highlighted. Only protein-coding genes are shown in the gene plot. Genome coordinates correspond to the GRCh37 genome assembly. B) Potential mechanism of interaction between reactions of prostaglandin metabolism and *LPA*. *LPA* induces the expression of the enzyme *COX2* which catalyses the formation of prostaglandin E2 from arachidonic acid. Conversely, the elongation of arachidonoyl-CoA diverts arachidonic acid away from prostaglandin synthesis. C) Potential mechanism of interaction between N-Acetylglucosamine 2-Epimerase and plasminogen. This reaction is catalysed by a homodimer of RNBP, which can also form a heterodimer with renin inhibiting the renin-angiotensin system. The renin-angiotensin system can suppress plasminogen activation. Solid arrows denote metabolic reactions or transport processes and dashed lines other functional relationships.

We also found that fluxes involving polyamine transport in adipose tissue, heart, and skeletal muscle can have buffering or amplification effects in a variant- and tissue-specific manner across the four adjacent risk loci (**Figure 3A**; **Figure 2D**; **Figure S1D**; **Figure S4**). Polyamines are a family of pleiotropic compounds that regulate cell proliferation, cell differentiation, and protein synthesis and have a generally protective effect against inflammation and oxidative stress<sup>39</sup>. Polyamine-rich diets have been established to protect against cardiovascular disease by countering the age-related myocyte and vascular

endothelial dysfunctions<sup>39</sup> while dysregulation of endogenous polyamine metabolism can also lead to pathologies such as cardiac hypertrophy<sup>40,41</sup>. In adipose tissue, polyamine synthesis is reported to protect against obesity by promoting vascularization and lipolysis<sup>42</sup> while in skeletal muscle it promotes muscle mass growth and regeneration<sup>43</sup>.

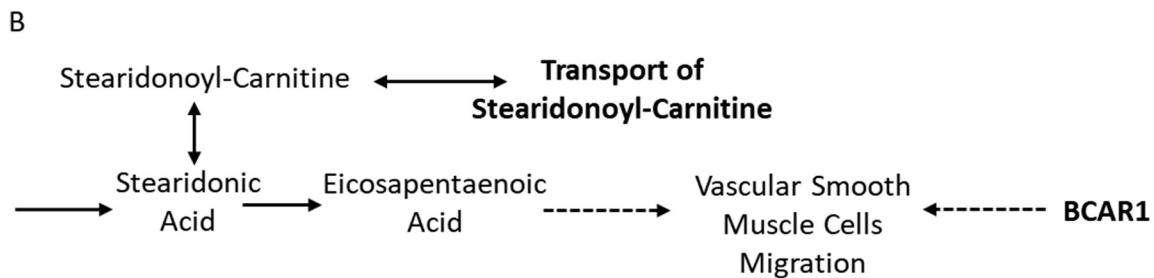
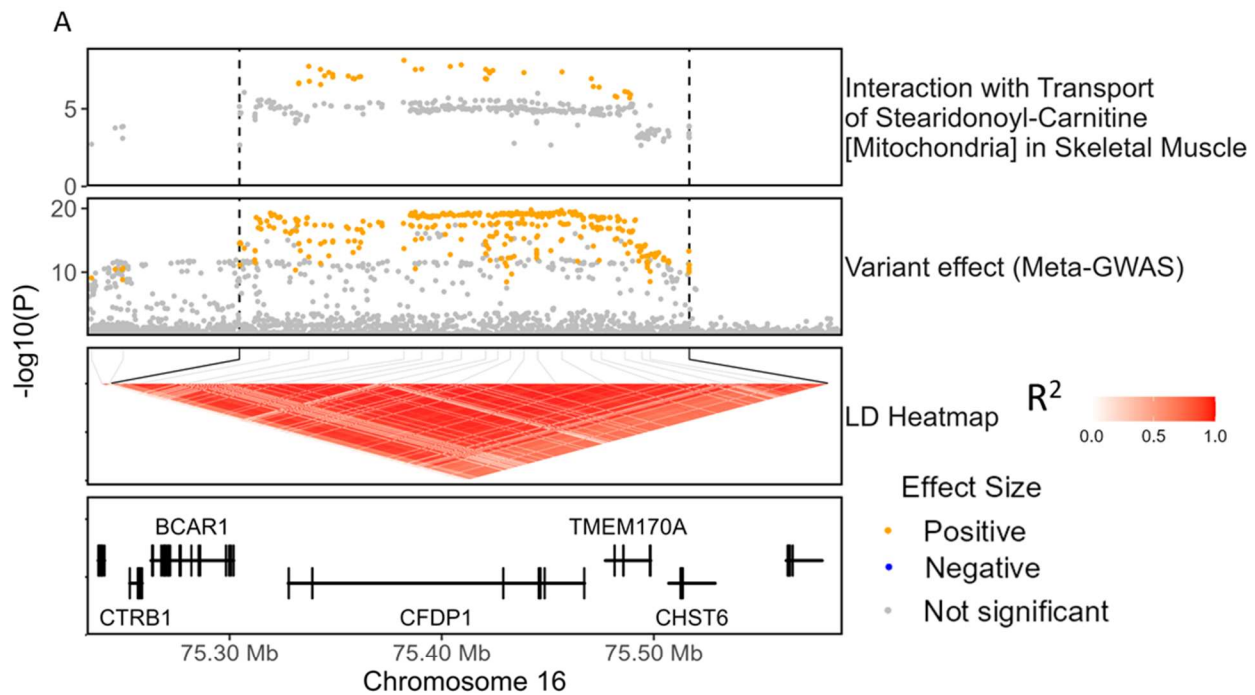
A commonality between prostaglandin and polyamine transport is that they are mediated by transmembrane carriers coded by genes in the *LPA/PLG* genomic region. Namely, prostaglandin transport can be mediated by either *SLC22A1*, *SLC22A2*, or *SLC22A3* and polyamine transport is mediated by *SLC22A1*<sup>23,44</sup>. In this regard, expression quantitative trait loci (eQTL) variants for these transporters had been used as input for computing genetically personalized flux values<sup>21,22</sup> (**Methods**) and are also mapped within this genomic region. In addition to regressing out the effect of the individual risk allele when testing for interaction, we quantified the linkage disequilibrium (LD) between eQTL variants used in flux computation and CAD risk variants, and the contribution of the former to flux values. We found that, while some eQTL variants had a strong contribution to reaction fluxes, they were not in strong LD with CAD risk variants with significant interaction with reaction fluxes (**Figure S5**). Furthermore, the flux through the elongation of arachidonoyl-CoA was not associated with any eQTL variants at the *LPA/PLG* loci, suggesting that prostaglandin metabolism may modulate the effect of risk alleles independently of the activity of the *SLC22A1-3* transporters.

Finally, in the risk locus encoding the *PLG* gene, two intron variants displayed significant interaction on disease risk with the flux through N-Acetylglucosamine 2-Epimerase in the liver (**Figure 3A**; **Figure 2E**; **Figure S1E**). This reaction is catalysed by a homodimer of the renin-binding protein (RNBP)<sup>23,45</sup>. The formation and stabilization of its catalytically active form prevents the formation of a heterocomplex with renin, which sequesters the latter and inhibits the renin-angiotensin system<sup>45</sup> (**Figure 3C**). The renin-angiotensin system increases blood pressure and has been linked to atherosclerosis<sup>46</sup>.

### **Transport of stearidonoyl-carnitine amplifies the effect size of variants mapped to the *BCAR1/CFDP1* risk locus**

In the *BCAR1/CFDP1* risk locus, there was evidence of risk allele amplification by the flux of mitochondrial transport of stearidonoyl-carnitine in skeletal muscle, with 38 risk variants showing significant interaction with this flux (**Figure 2F**; **Figure 4A**; **Figure S1F**; **Supplementary Data S1**). Polymorphisms within this region have been associated with carotid intima-media thickness (i.e. a marker of subclinical atherosclerosis) and CAD risk with *BCAR1* identified as the likely causal gene<sup>47,48</sup>. *BCAR1* regulates cell migration, proliferation and apoptosis and is essential for cardiovascular development in embryogenesis<sup>49</sup>. In particular, *BCAR1* is a major regulator of VSMC function and it has been theorized that *BCAR1*'s contribution to the formation of atherosclerotic lesions arises from its role in VSMC migration<sup>50</sup>.

Stearidonoyl-carnitine is the carnitine-conjugated form of stearidonic acid (SDA), an omega-3 polyunsaturated fatty acid that can be efficiently metabolized to eicosapentaenoic acid (EPA)<sup>51</sup>. EPA has well-established anti-inflammatory and cardioprotective effects<sup>52,53</sup> and has been shown to inhibit the progression of atherosclerotic lesions<sup>54</sup>. Likewise, SDA has also been found to have anti-inflammatory effects in cell and animal models<sup>51,55</sup>. However, SDA levels in the blood have also been associated with an increased risk of cardiovascular pathologies such as hypertension<sup>56</sup> or thrombosis<sup>57</sup> suggesting a context-specific role in disease.



**Figure 4:** Transport of stearidonoyl-carnitine amplifies the effect size of risk variants of the *BCAR1/CFDP1* locus. A) Regional association plots showing the  $-\log_{10}(P)$ -value for interaction and variant effect sizes on CAD risk. P-values for interaction and variant effect sizes were derived from the interaction effect size test and meta-GWAS summary statistics, respectively. B) Potential mechanism of interaction between the transport of stearidonoyl-carnitine and risk variants linked to *BCAR1*. Solid arrows denote metabolic reactions or transport processes and dashed lines other functional relationships, namely activation of the migration of vascular smooth muscle cells.

### Galactose transport into the brain amplifies the effect size of variants mapped to the *SMARCA4* risk locus

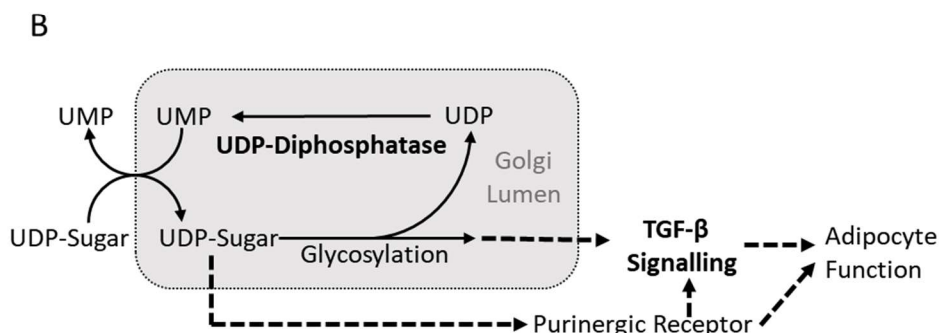
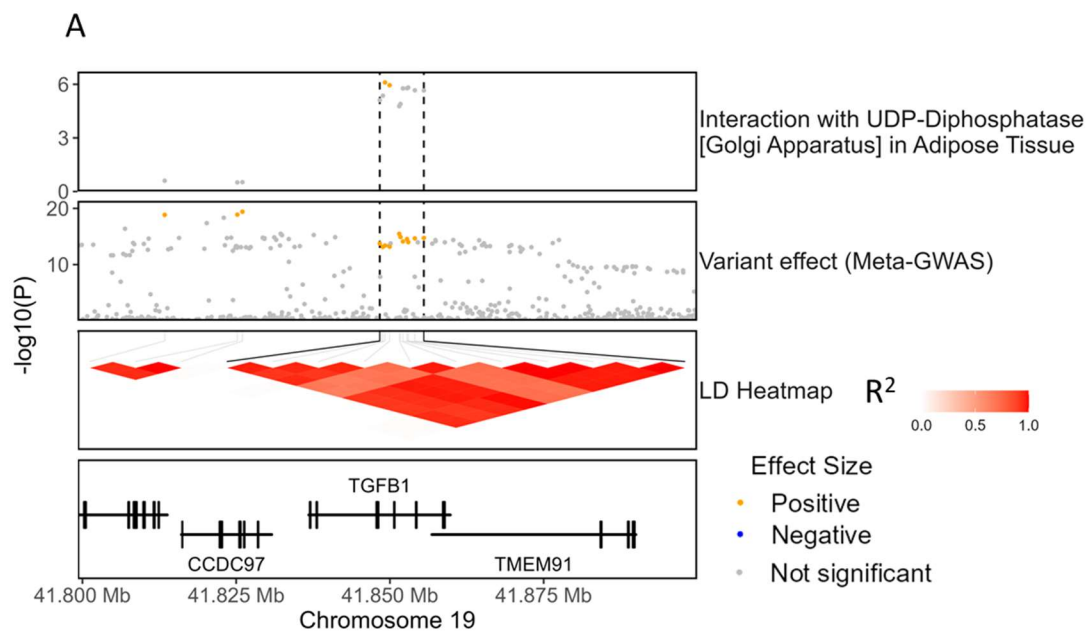
A set of 19 variants mapped to a risk locus encoding the *SMARCA4* gene, were identified as having their effect size amplified by the flux of the sodium-coupled galactose transport in the brain (Figure 2H; Figure S1H; Figure S6; Supplementary Data S1). *SMARCA4* codes for a chromatin-remodelling factor that has a major role in transcriptional regulation, DNA repair, and cell proliferation in a wide range of processes and tissues<sup>58</sup>. The *SMARCA4* gene is adjacent to *LDLR*, a well-established risk locus for CAD, but it has *LDLR*-independent roles in CVD risk<sup>59</sup>. For instance, *SMARCA4* has been reported to mediate vascular calcification<sup>60</sup>, inflammation<sup>59</sup> and myocardial proliferation<sup>61</sup>. Indeed, the variants showing significant interaction with galactose transport are in a different LD block than those linked to *LDLR* (Figure S6).



## UDP-Diphosphatase amplifies the effect size of variants mapped to the TGF- $\beta$ -risk locus

In the TGF- $\beta$  risk locus, there was evidence of risk allele amplification by the flux through UDP-diphosphatase (Golgi apparatus) in adipose tissue, with two intron variants of *TGFB1* showing significant interaction on disease risk with this flux (**Figure 2G**; **Figure 5A**; **Figure S1G**; **Supplementary Data S1**). *TGFB1* codes for a ligand of the Transforming Growth Factor  $\beta$  (TGF- $\beta$ ) family which regulates a wide range of biological processes such as morphogenesis, tissue homeostasis, and inflammation in a context-specific manner<sup>62</sup>. TGF- $\beta$  plays a major role in the cardiovascular system by regulating the proliferation, differentiation and function of endothelial, smooth muscle, and immune cells, and its dysregulation can lead to a wide range of cardiovascular diseases such as atherosclerosis<sup>63</sup>. TGF- $\beta$  signalling is also linked to adiposity through the regulation of adipocyte differentiation and oxidative metabolism<sup>64</sup>, and *TGFB1* is overexpressed in obesity<sup>65,66</sup>.

UDP-diphosphatase plays a major role in the Golgi apparatus, which is the primary location of protein and lipid glycosylation. Nucleotide sugars (such as UDP-sugars) are transported from the cytosol to the Golgi's lumen where they act as donors for a glycosylation process releasing the nucleotide diphosphate (e.g. UDP)<sup>67</sup>. Notably, the transport of UDP-sugars to Golgi involves an antiport with luminal UMP<sup>68,69</sup>. As such, UDP-Diphosphatase contributes to glycosylation by catalysing the hydrolysis of UDP to UMP hence enabling the UMP-dependent transport of nucleotide sugars<sup>67,68</sup>. Additionally, the UMP-dependent transport of nucleotides-sugars to the Golgi's lumen is also reported to contribute to the vesicle-based release of UDP-sugars to the extracellular space<sup>69</sup> (**Figure 5B**).



**Figure 5:** UDP-diphosphatase amplifies the effect size of CAD risk variants within the *TGF- $\beta$*  risk locus. A) Regional association plots showing the  $-\log_{10}(\text{P-value})$  for interaction and variant effect sizes on CAD risk. P-values for interaction and variant effect sizes were derived from the interaction effect size test and meta-GWAS summary statistics, respectively. B) Potential mechanism of interaction between UDP-diphosphatase and risk variants linked to *TGB1*. Solid arrows denote metabolic reactions or transport processes and dashed lines other functional relationships.

## Amino acid transport amplifies the effect of a risk variant upstream of the *EDNRA* gene

A significant positive interaction in CAD risk was detected between a risk variant, mapped upstream of the *EDNRA* gene, and the flux of an amino acid transport process in the heart (i.e. the sodium-coupled exchange of homoserine and asparagine) (**Figure 2I**; **Figure S1I**; **Figure S7**; **Supplementary Data S1**). *EDNRA*, which is expressed in VSMC and cardiomyocytes, codes for the receptor for endothelin-1. Endothelin-1, acting through the activation of *EDNRA*, is a potent vasoconstrictor with a well-established role in cardiovascular diseases<sup>70</sup>. *EDNRA* expression in vascular tissue and endothelin-1 levels in blood are altered in atherosclerosis<sup>71,72</sup> and *EDNRA* antagonists can attenuate the progression of coronary atherosclerosis lesions<sup>73</sup>. The concentration in blood of both asparagine and homoserine have been associated with cardiovascular disease risk<sup>74–77</sup>, however the mechanism of their effect remains largely unknown.

## Buffering and amplification of the risk of myocardial infarction

Myocardial infarction (MI) is caused by a sudden blockage of blood flow to the myocardium, primarily due to coronary atherosclerosis with or without a blood clot<sup>78</sup>. Given the close relationship between MI and CAD and that the metabolic reactions we assess here may not *a priori* be involved in hard outcomes like MI, we investigated the extent to which the identified instances of buffering and amplification of risk allele penetrance on CAD risk could also be relevant for MI risk.

On the UKB participants of European genetic ancestries, we identified 36,007 MI cases and 423,629 controls using an established definition of MI<sup>79</sup>. As expected, there was a substantial overlap of cases and controls between MI and coronary atherosclerosis (**Table S1**). Risk variant and reaction flux effect sizes, estimated using Cox regression, were highly correlated between coronary atherosclerosis and MI (**Figure S8A-D**). However, reaction effect sizes for MI risk were on average lower than the equivalent effect sizes derived for coronary atherosclerosis.

We evaluated buffering and amplification effects between risk variants and risk allele dosage on MI risk. We tested the same SNP-reaction flux pairs that had been evaluated with coronary atherosclerosis to facilitate the comparison of the results. For MI, there were 426 pairs significant with both interaction effect size and dosage-specific tests (FDR-adjusted P-value<0.05) (**Supplementary Data S1**). Taking into account the LD structure of risk SNPs, these represented 26 risk locus – reaction flux pairs with evidence of significant amplification or buffering of risk alleles by reaction fluxes. Out of the 583 significant SNP-flux pairs for coronary atherosclerosis, 360 were also significant in MI (20 risk locus – reaction flux pairs) with 91 additional SNP-flux pairs (4 additional risk locus – reaction flux pairs) being borderline significant (FDR-adjusted P-value<0.25 for both the interaction effect size test and the dosage-specific tests). Overall, interaction effect sizes were also strongly correlated between both MI and coronary atherosclerosis (**Figure S8E-H**).

Notably, there were no significant interactions on MI risk between the transport of stearidonoyl-carnitine and the risk locus of *BCAR1/CFDP1* nor UDP-diphosphatase and the *TGFB1* risk locus. Effect sizes for variants mapped to these loci estimated with Cox

regression were also slightly lower in MI compared to coronary atherosclerosis (**Figure S9**). Conversely, two loci had significant buffering of disease risk alleles by reaction fluxes in MI and not in CAD: effect sizes of 15 risk alleles mapped to the *PDE5A/MAD2L1* locus, and 13 risk alleles mapped to the *MAP1S/FCHO1* locus were buffered by the flux of orotate phosphoribosyltransferase in heart and uracil transport in skeletal muscle, respectively (**Figure S10**). Both reactions are functionally part of pyrimidine metabolism, which has been implicated in MI<sup>80–82</sup>. Notably, PDE5A is also linked to nucleotide metabolism and its inhibition is well-established to have cardioprotective effects in MI<sup>83,84</sup>.

## **Discussion**

Here, we used genetically personalized organ-specific metabolic fluxes to study the interaction between risk variants and biochemical reactions on disease risk using CAD as a case study. Our analysis identified 18 metabolic reaction fluxes that can amplify or buffer risk allele effect sizes at 8 well-established CAD risk loci unveiling a total of 30 risk locus - reaction flux pairs with significant interaction on disease susceptibility. Most of such reactions involved metabolic processes with known roles in atherosclerosis such as inflammation or fatty acid metabolism. Furthermore, the majority of the interactions detected for CAD were also relevant for MI.

We identify that the genomic region of *LPA/PLG*, a well-known locus for CAD risk<sup>28,85</sup>, is a major site of interaction between risk variants and reaction fluxes. For instance, we find that a set of variants in the *LPA* and *PLG* risk loci have their effect size amplified by the flux of reactions involved in the synthesis or transport of histamine and prostaglandin E2, two inflammatory mediators that can contribute to the formation of atherosclerotic lesions<sup>36,38</sup>. In our previous work<sup>21</sup>, we identified that some of those reaction fluxes were associated with CAD risk, here we show that this effect can be further amplified or buffered by the dosage of specific risk alleles within the *LPA/PLG* region (**Figure S1A,C**). Inflammation is one of the mediators of pathogenicity of *LPA* and *PLG*<sup>31–33</sup> and the former has been reported to induce the expression of cyclooxygenase-2, which catalyses the first step of prostaglandin E2 synthesis<sup>31</sup>. Hence, the identified interactions may reflect a mechanism where *LPA* or *PLG* variants that increase inflammation have their effect on CAD risk amplified by a high capacity to transport prostaglandin E2 or histamine across cellular membranes. Conversely, the flux of elongation of arachidonoyl-CoA in adipose tissue would attenuate this effect by diverting arachidonic acid away from prostaglandin synthesis (**Figure 3B**).

Notably, most interactions were also instances where the effect of a biochemical reaction flux on CAD risk only becomes apparent when analysed in conjunction with risk variants. For instance, neither the flux through N-Acetylglucosamine 2-Epimerase in the liver, the transport of stearidonoyl-carnitine in skeletal muscle nor UDP-diphosphatase in adipose tissue have a significant effect on disease risk when analysed in univariate analysis. However, the interaction analysis reveals that such fluxes can significantly amplify the effect size of risk variants mapped to the *PLG*, *BCAR1* and *TGFB1* risk loci, respectively, providing insights into their roles in cardiovascular disease. For instance, the reaction N-Acetylglucosamine 2-Epimerase is mediated by an enzyme that moonlights as an inhibitor of renin<sup>45</sup> and the formation of its catalytically active form prevents it from inhibiting the renin-angiotensin system<sup>46</sup>, a known regulator of plasminogen<sup>86,87</sup>, suggesting a potential mechanism of interaction with variants of the plasminogen risk locus(**Figure 3C**). Concerning the transport of stearidonoyl-carnitine, the transport of acylcarnitines to mitochondria is the limiting step for mitochondrial  $\beta$ -oxidation<sup>88</sup>, and hence this reaction might be affecting CAD risk by modulating the bioavailability of SFA and other polyunsaturated fatty acids. In this regard, one of the cardioprotective actions of omega-3 polyunsaturated fatty acids is the inhibition of VSMC proliferation and migration<sup>52,89</sup>, thus

providing a link to *BCAR1* which is also reported to modulate VSMC function<sup>50</sup> (**Figure 4B**). Similarly, UDP-diphosphatase is reported to play a major role in the transport of nucleotide-sugars to the Golgi apparatus<sup>67,68</sup> and hence can potentially modulate both protein glycosylation and the vesicle-based release of UDP-sugars<sup>69</sup>. The formation of the functional form of TGF- $\beta$ , and many of the other proteins involved in TGF- $\beta$  signalling, requires glycosylation in the Golgi apparatus<sup>90,91</sup> suggesting a potential avenue of interaction. Additionally, the release of UDP-sugars, acting through the activation of purinergic receptors, can modulate adipocyte differentiation, lipolysis, and inflammation within the adipose tissue<sup>92</sup>. Indeed, there is ample evidence of crosstalk between TGF- $\beta$  signalling and purinergic receptors<sup>93–95</sup> suggesting a second potential mechanism of interaction (**Figure 5B**).

In conclusion, this work establishes proof of concept that biochemical reaction fluxes can modulate the effect of disease risk alleles and highlights the importance of considering the burden of risk variants to understand the contribution of metabolism to cardiovascular disease susceptibility. Given that disease-associated metabolic processes represent potential targets against disease, such findings have important implications for personalized medicine as they highlight that the therapeutic efficacy of targeting specific metabolic pathways may depend on each individual's genetic background.

## **Methods**

### **UK Biobank**

UKB is a cohort of approximately 500,000 participants from the general UK population (<https://www.ukbiobank.ac.uk/>). Participants were between age 40 and 69 at recruitment (median 58 years of age; 54% women) and accepted an invitation to attend one of the assessment centres that were established across the United Kingdom between 2006 and 2010<sup>24</sup>. We used the version 3 release of the UK Biobank genotype data<sup>25</sup> (<https://biobank.ndph.ox.ac.uk/showcase/label.cgi?id=263>), which had been imputed to the UK10K/1000 genomes and haplotype reference consortium (HRC)<sup>96</sup> panels.

### **Ancestry inference**

Genotyped UKB participants were assigned a genetic ancestry using KING<sup>26</sup>. Briefly, genotyped UKB samples were projected to the 10 main genetic principal components computed from samples from the 1000 genome project with known superpopulation groups (American, East Asian, European, and South Asian). Using this projection, KING uses a support-vector-machine-based method to infer the most likely ancestral group of each sample<sup>26</sup>. UKB participants were assigned to European Ancestry if the probability of belonging to that group was estimated to be 95% or higher.

### **Genetically personalized organ-specific fluxes**

Genetically personalized organ-specific fluxes were computed as previously described<sup>21</sup>. Briefly, organ-specific metabolic models for adipose tissue, brain, heart, liver and skeletal muscle were extracted from Harvey/Harvetta whole body models<sup>97</sup> and ported to HUMAN1<sup>98</sup>, the latest reconstruction of human metabolism. The GIM3E<sup>99</sup> algorithm was then used to compute an average reaction flux distribution for each organ consistent with average organ-specific transcript abundances obtained from GTEx<sup>100</sup>. In parallel, genotype data was used to impute personalized organ-specific transcript abundances for UKB participants using the elastic net models from PredictDB<sup>22</sup>. Finally, the quadratic metabolic transformation algorithm (qMTA) was used to integrate the individual-level organ-specific transcript abundances and the average reaction flux distribution and compute genetically

personalised organ-specific reaction flux values for each analysed organ. Flux values for each reaction were log<sub>2</sub>-transformed and standardized to zero-mean and unit variance.

### **Definition of coronary artery disease**

Cases of CAD were identified using the definition of coronary atherosclerosis (Phecode 411.4) of the PheWAS Catalog (version 1.2)<sup>27</sup>. Namely, cases were defined by the presence of any of the constitutive ICD9 (411.81, 414.0, 414.01, 414.02, 414.03, 414.04, 414.05, 414.2, 414.3, 414.4, 996.03 or V45.81, V45.82) and ICD10 (I24.0, I25.1, Z95.1 or Z95.5) codes in hospital episode statistics or death records. Additionally, non-cases with any of the constitutive ICD codes of ischemic heart disease (Phecode 410-414.99) were excluded from the controls<sup>27</sup>. We identified 37,941 CAD cases and 398,282 controls in genotyped UKB participants of European ancestry. The earliest coded or reported date for disease was converted to the age of phenotype onset. Controls were censored according to the maximum follow-up of the health linkage data (October 31, 2022) or the date of death.

### **Definition of myocardial infarction**

MI was defined as evidence of a fatal or nonfatal myocardial infarction or major coronary surgery in hospital episode statistics, death records, or self-reported during the verbal interview at UKB enrolment<sup>79</sup>. More in detail, myocardial infarction was defined as the presence of ICD-9 codes 410-412, ICD-10 codes I21–I24 or I25.2 in hospital episode statistics or cause of death records or reporting a heart attack during the verbal interview at UKB enrolment (Self-report field 6150 and Self-report field 20002). Major coronary surgery was defined by ICD-9 code V45.81, ICD-10 code Z95.1, OPCS-3 codes 309.4 or 884, OPSC-4 codes K40–K46 or reporting a coronary angioplasty, coronary artery bypass grafts or triple heart bypass at the verbal interview (Self-report field 20004)<sup>79</sup>. We identified 36,007 cases and 423,629 controls in genotyped UKB participants of European ancestry. The earliest coded or reported date for disease was converted to the age of phenotype onset. Controls were censored according to the maximum follow-up of the health linkage data (October 31, 2022) or the date of death.

### **Risk variant selection**

A set of variants associated with CAD risk was obtained from a published genome-wide association meta-analysis performed on over one million participants of European ancestry<sup>28</sup>. Summary statistics from this study were downloaded from the GWAS catalog (GCST90132314)<sup>1</sup> and they identified 18,348 biallelic SNPs with genome-wide significance ( $P < 5 \times 10^{-8}$ ).

We anticipated that only a subset of these variants might have significant effects in our cohort. As such, we tested the effect of these variants on CAD risk in the European subset of UKB. For this, we used a similar model and assumptions that we would subsequently use for the amplification and buffering analysis (i.e., Cox proportional-hazards model using age as a time scale stratifying by sex and genotyping array and using the first 10 genetic principal components as covariates). We selected the 5,852 variants that had genome-wide significance in our cohort as candidate variants to test for amplification and buffering of disease risk by reaction fluxes.

## Identifying interactions between reaction fluxes and risk variants on disease risk

To robustly identify instances of buffering/amplification between reaction fluxes and risk variants we used two complementary methods to detect interaction: an interaction effect size test and a dosage-specific test.

In the interaction effect size test, for each pair of reaction flux values and risk allele dosages, we test for a significant interaction term using a Cox proportional-hazards model with age as a time scale for CAD risk:

$$h(t) = h_0(t) \cdot \exp(\beta_{SNP} \cdot SNP + \beta_{Flux} \cdot Flux + \beta_{SNP^2} \cdot SNP^2 + \beta_{Flux^2} \cdot Flux^2 + \beta_{SNP:Flux} \cdot SNP \cdot Flux)$$

where,

$h(t)$  is the hazard function defining the risk of CAD at age  $t$

$h_0$  is the baseline hazard

$SNP$  is the dosage of the risk allele

$Flux$  is the adjusted reaction flux value. Before testing for interaction, flux values are adjusted with linear regression to remove any potential effects of  $SNP$  over the reaction flux value.

$\beta_{SNP}$ ,  $\beta_{Flux}$ ,  $\beta_{SNP^2}$ ,  $\beta_{Flux^2}$  are the first and second-order effect sizes for SNP dosage and Flux value

$\beta_{SNP:Flux} \cdot SNP \cdot Flux$  is the interaction effect size between SNP dosage and Flux value

The first ten genetic principal components were also used as covariates but have been omitted from this equation for clarity. Additionally, this analysis was stratified by sex and genotyping array. The model was fitted using the `CoxPHFitter` function from the `lifelines` python package<sup>101</sup>. The significant interaction effect on CAD risk was evaluated with a two-tailed Wald test for the interaction effect size.

In the dosage-specific effect size test, for each pair of reaction flux values and risk alleles, we split the analysed UKB participants based on risk allele dosage (0, 1 or 2) and estimated the effect of reaction flux value on CAD risk within each allele dosage using a Cox proportional hazards model with age as time scale.

$$h_{SNP_i}(t) = h_{0_{SNP_i}}(t) \cdot \exp(\beta_{Flux,SNP_i} \cdot Flux_{SNP_i})$$

where,

$h_{SNP_i}(t)$  is the hazard function defining the risk of developing CAD at age  $t$  in individuals with dosage  $i$  of the risk allele

$h_{0_{SNP_i}}$  is the baseline hazard in individuals with dosage  $i$  of the risk allele

$\beta_{Flux,SNP_i}$  is flux effect size in individuals with dosage  $i$  of the risk allele.

$Flux$  is the reaction flux value. Adjusting flux values to regress out any potential effects of  $SNP$  dosage has no effect in this analysis as the dosage of the risk allele is constant for each test.

As with the interaction model, the first ten genetic principal components are also used as covariates and the analysis is stratified by sex and genotyping array.

The `CoxPHFitter` function from the `lifelines` python package<sup>101</sup> was used to estimate the effect size of reaction flux values and its standard error (SE) for each dosage of the risk allele. A Welch's ANOVA test (i.e. a variant of ANOVA that does not assume homogeneity of variance)<sup>102</sup> was used to evaluate if there were significant differences between reaction effect sizes across risk allele dosages. To facilitate comparing these results with the interaction effect size, effect size variation per dosage of risk allele were computed with a linear regression of flux effect size per dosage weighted by the standard error of effect sizes estimates at each dosage ( $1/SE^2$ ).

## Reaction selection and pruning

In the HUMAN1-derived organ-specific models<sup>21,23</sup>, some reactions lack gene annotation. This can occur because they are spontaneous processes, but more often than not this arises because the gene(s) mediating the reaction are unknown or because they are artificial reactions needed to balance the models (e.g., exchange reactions). As such, we decided to focus the amplification/buffering analysis only on the fluxes of reactions with annotated genes (adipose tissue: 721, brain: 1116, heart: 1314, liver: 1956, skeletal muscle: 1078).

Additionally, in any metabolic network, many reaction fluxes have a significant degree of flux correlation. This can arise from stoichiometric coupling between reactions (e.g., the product of the first reaction is the substrate of the second reaction) and from proteins that mediate multiple reactions (e.g. some transmembrane carriers can transport a wide range of substrates). To account for this and facilitate the interpretation of the results, for each analyzed organ metabolic network, we pruned reaction fluxes with more than 50% correlation from the interaction results. Briefly, Pearson correlation coefficients were computed between all reaction fluxes from each organ. Next, all SNP-reaction fluxes pairs were ranked based on the maximum P-value between the interaction effect size test and the dosage-specific test. Then, starting from the reaction flux in the most significant reaction-SNP pair, reactions with more than 0.5 flux correlation ( $r$ ) to this reaction were identified and all pairs involving these reactions were removed. The process was subsequently repeated for all ranked reaction-SNP pairs until no pairs involving reactions from the same organ-metabolic network with more than 0.5 flux correlation remained. In total 1,670 reactions remained (adipose tissue: 280, brain: 417, heart: 360, liver: 263, skeletal muscle: 350)

The interaction effect size test and the dosage-specific test P-values were adjusted for multiple testing against all remaining SNP-reaction flux pairs using the Benjamini and Hochberg (i.e., FDR) method.

Additionally, the univariate effect of the uncorrelated reactions on disease risk was also evaluated using a Cox model stratifying by sex and genotyping array and using the first 10 genetic principal components as covariates. The resulting p-values were FDR-adjusted for all uncorrelated reactions within each disease definition.

## Linkage disequilibrium estimation and variant annotation

LD between risk SNPs or SNPs used in flux prediction was measured using the genotype data from the UKB subset of inferred European ancestry using Plink 1.9<sup>103</sup>. LD blocks used to delimitate independent risk loci were defined using hierarchical clustering with the single linkage method (“friends of friends” clustering) using  $1-R^2$  between risk variants as a distance metric. The resulting hierarchical tree was cut at a height of 0.4, identifying LD blocks such that each risk SNP had an  $R^2 > 0.6$  with at least one other risk SNP within the block and an  $R^2 < 0.6$  with all risk SNPs outside the block.

Ensembl Variant Effect Predictor was used to annotate the effect of risk variants involved in interactions<sup>104</sup>. Similarly, known eQTL and sQTL variants and their target genes were obtained from GTEx Portal (<https://gtexportal.org/home/>)<sup>100</sup> and the INTERVAL RNA-SEQ Portal (<https://www.intervalrna.org.uk/>)<sup>105</sup> and were used to further annotate the risk variants involved in interactions.

Gene features shown in plots were extracted from the TxDb.Hsapiens.UCSC.hg19.knownGene package with the makeGenesDataFromTxDb function from the karyoploteR package<sup>106</sup>.

## **Conflicts of Interest**

A.S.B. reports institutional grants from AstraZeneca, Bayer, Biogen, BioMarin, Bioverativ, Novartis, Regeneron and Sanofi. J.D. serves on scientific advisory boards for AstraZeneca, Novartis, Our Future Health and UK Biobank, and has received multiple grants from academic, charitable and industry sources outside of the submitted work. M.I. is a trustee of the Public Health Genomics (PHG) Foundation, a member of the Scientific Advisory Board of Open Targets, and has research collaborations with AstraZeneca, Nightingale Health and Pfizer which are unrelated to this study.

## **Acknowledgements**

This research has been conducted using the UK Biobank Resource under Application 7439. This work was performed using resources provided by the Cambridge Service for Data-Driven Discovery (CSD3) operated by the University of Cambridge Research Computing Service ([www.csd3.cam.ac.uk](http://www.csd3.cam.ac.uk)), provided by Dell EMC and Intel using Tier-2 funding from the Engineering and Physical Sciences Research Council (capital grant EP/P020259/1), and DiRAC funding from the Science and Technology Facilities Council ([www.dirac.ac.uk](http://www.dirac.ac.uk)). This work was supported by core funding from the British Heart Foundation (RG/18/13/33946: RG/F/23/110103), NIHR Cambridge Biomedical Research Centre (NIHR203312) [\*], BHF Chair Award (CH/12/2/29428), Cambridge BHF Centre of Research Excellence (RE/18/1/34212), and by Health Data Research UK, which is funded by the UK Medical Research Council, Engineering and Physical Sciences Research Council, Economic and Social Research Council, Department of Health and Social Care (England), Chief Scientist Office of the Scottish Government Health and Social Care Directorates, Health and Social Care Research and Development Division (Welsh Government), Public Health Agency (Northern Ireland), British Heart Foundation and the Wellcome Trust. X.J. was also supported by the Wellcome Trust [227566/Z/23/Z]. M.I. was also supported by the UK Economic and Social Research 878 Council (ES/T013192/1).

\*The views expressed are those of the authors and not necessarily those of the NIHR or the Department of Health and Social Care.

## **References**

1. Sollis, E. *et al.* The NHGRI-EBI GWAS Catalog: knowledgebase and deposition resource. *Nucleic Acids Res.* **51**, D977–D985 (2023).
2. Nandakumar, S. K., Liao, X. & Sankaran, V. G. In The Blood: Connecting Variant to Function In Human Hematopoiesis. *Trends Genet.* **36**, 563–576 (2020).
3. Claussnitzer, M. & Susztak, K. Gaining insight into metabolic diseases from human genetic discoveries. *Trends Genet.* **37**, 1081–1094 (2021).
4. Joshi, A., Rienks, M., Theofilatos, K. & Mayr, M. Systems biology in cardiovascular disease: a multiomics approach. *Nat. Rev. Cardiol.* **18**, 313–330 (2021).
5. Lappalainen, T., Li, Y. I., Ramachandran, S. & Gusev, A. Genetic and molecular architecture of complex traits. *Cell* **187**, 1059–1075 (2024).
6. Niel, C., Sinoquet, C., Dina, C. & Rocheleau, G. A survey about methods dedicated to epistasis detection. *Front. Genet.* **6**, (2015).
7. Wu, H., Eckhardt, C. M. & Baccarelli, A. A. Molecular mechanisms of environmental exposures and human disease. *Nat. Rev. Genet.* **24**, 332–344 (2023).



8. Frayn, K. N. *Metabolic Regulation: A Human Perspective*. (2010).
9. Zamboni, N., Saghatelian, A. & Patti, G. J. Defining the metabolome: size, flux, and regulation. *Mol. Cell* **58**, 699–706 (2015).
10. Niedenfür, S., Wiechert, W. & Nöh, K. How to measure metabolic fluxes: A taxonomic guide for <sup>13</sup>C fluxomics. *Curr. Opin. Biotechnol.* **34**, 82–90 (2015).
11. Oller, J. *et al.* Rewiring Vascular Metabolism Prevents Sudden Death due to Aortic Ruptures-Brief Report. *Arterioscler. Thromb. Vasc. Biol.* **42**, 462–469 (2022).
12. Sims, N. R. & Muyderman, H. Mitochondria, oxidative metabolism and cell death in stroke. *Biochim. Biophys. Acta - Mol. Basis Dis.* **1802**, 80–91 (2010).
13. Tian, Z. & Liang, M. Renal metabolism and hypertension. *Nat. Commun.* **12**, 1–12 (2021).
14. Lopaschuk, G. D., Karwi, Q. G., Tian, R., Wende, A. R. & Abel, E. D. Cardiac Energy Metabolism in Heart Failure. *Circ. Res.* **128**, 1487–1513 (2021).
15. Libby, P. *et al.* Atherosclerosis. *Nat. Rev. Dis. Prim.* **5**, 1–18 (2019).
16. Soppert, J., Lehrke, M., Marx, N., Jankowski, J. & Noels, H. Lipoproteins and lipids in cardiovascular disease: from mechanistic insights to therapeutic targeting. *Adv. Drug Deliv. Rev.* **159**, 4–33 (2020).
17. Ward, N. C., Watts, G. F. & Eckel, R. H. Statin Toxicity. *Circ. Res.* **124**, 328–350 (2019).
18. Mach, F. *et al.* 2019 ESC/EAS Guidelines for the management of dyslipidaemias: lipid modification to reduce cardiovascular risk. *Eur. Heart J.* **41**, 111–188 (2020).
19. Han, Y. *et al.* Effect of metformin on all - cause and cardiovascular mortality in patients with coronary artery diseases : a systematic review and an updated meta - analysis. *Cardiovasc. Diabetol.* 1–16 (2019) doi:10.1186/s12933-019-0900-7.
20. Foretz, M., Guigas, B. & Viollet, B. Metformin : update on mechanisms of action and repurposing potential. **19**, (2023).
21. Foguet, C. *et al.* Genetically personalised organ-specific metabolic models in health and disease. *Nat. Commun.* **13**, 7356 (2022).
22. Gamazon, E. R. *et al.* A gene-based association method for mapping traits using reference transcriptome data. *Nat. Genet.* **47**, 1091–8 (2015).
23. Robinson, J. L. *et al.* An atlas of human metabolism. *Sci. Signal.* **13**, 1–12 (2020).
24. Sudlow, C. *et al.* UK Biobank: An Open Access Resource for Identifying the Causes of a Wide Range of Complex Diseases of Middle and Old Age. *PLoS Med.* **12**, 1–10 (2015).
25. Bycroft, C. *et al.* The UK Biobank resource with deep phenotyping and genomic data. *Nature* **562**, 203–209 (2018).
26. Manichaikul, A. *et al.* Robust relationship inference in genome-wide association studies. *Bioinformatics* **26**, 2867–73 (2010).
27. Wu, P. *et al.* Mapping ICD-10 and ICD-10-CM Codes to Phecodes: Workflow Development and Initial Evaluation. *JMIR Med. informatics* **7**, e14325 (2019).
28. Aragam, K. G. *et al.* Discovery and systematic characterization of risk variants and genes for coronary artery disease in over a million participants. *Nat. Genet.* **54**, 1803–1815 (2022).

29. Hemani, G. *et al.* Phantom epistasis between unlinked loci. *Nature* **596**, E1–E3 (2021).
30. Reyes-Soffer, G. *et al.* Lipoprotein(a): A Genetically Determined, Causal, and Prevalent Risk Factor for Atherosclerotic Cardiovascular Disease: A Scientific Statement From the American Heart Association. *Arterioscler. Thromb. Vasc. Biol.* **42**, e48–e60 (2022).
31. Cho, T., Romagnuolo, R., Scipione, C., Boffa, M. B. & Koschinsky, M. L. Apolipoprotein(a) stimulates nuclear translocation of  $\beta$ -catenin: a novel pathogenic mechanism for lipoprotein(a). *Mol. Biol. Cell* **24**, 210–21 (2013).
32. Van Der Valk, F. M. *et al.* Oxidized phospholipids on Lipoprotein(a) elicit arterial wall inflammation and an inflammatory monocyte response in humans. *Circulation* **134**, 611–624 (2016).
33. Plow, E. F. & Hoover-Plow, J. The functions of plasminogen in cardiovascular disease. *Trends Cardiovasc. Med.* **14**, 180–186 (2004).
34. Rossignol, P. *et al.* Plasminogen activation: a mediator of vascular smooth muscle cell apoptosis in atherosclerotic plaques. *J. Thromb. Haemost.* **4**, 664–70 (2006).
35. Kremen, M. *et al.* Plasminogen mediates the atherogenic effects of macrophage-expressed urokinase and accelerates atherosclerosis in apoE-knockout mice. *Proc. Natl. Acad. Sci. U. S. A.* **105**, 17109–14 (2008).
36. Gomez, I., Foudi, N., Longrois, D. & Norel, X. The role of prostaglandin E2 in human vascular inflammation. *Prostaglandins. Leukot. Essent. Fatty Acids* **89**, 55–63 (2013).
37. Inouye, M. *et al.* Metabonomic, transcriptomic, and genomic variation of a population cohort. *Mol. Syst. Biol.* **6**, 441 (2010).
38. Wang, K. Y. *et al.* Histamine deficiency decreases atherosclerosis and inflammatory response in apolipoprotein e knockout mice independently of serum cholesterol level. *Arterioscler. Thromb. Vasc. Biol.* **31**, 800–807 (2011).
39. Xuan, M. *et al.* Polyamines: their significance for maintaining health and contributing to diseases. *Cell Commun. Signal.* **21**, 1–24 (2023).
40. Meana, C. *et al.* Correlation between endogenous polyamines in human cardiac tissues and clinical parameters in patients with heart failure. *J. Cell. Mol. Med.* **20**, 302–12 (2016).
41. Giordano, E. *et al.* Polyamines in cardiac physiology and disease. *Open Hear. Fail. J.* **3**, 25–30 (2010).
42. Monelli, E. *et al.* Angiocrine polyamine production regulates adiposity. *Nat. Metab.* **4**, 327–343 (2022).
43. Lee, N. K. L. & Maclean, H. E. Polyamines, androgens, and skeletal muscle hypertrophy. *J. Cell. Physiol.* **226**, 1453–1460 (2011).
44. Nigam, S. K. The SLC22 Transporter Family: A Paradigm for the Impact of Drug Transporters on Metabolic Pathways, Signaling, and Disease. *Annu. Rev. Pharmacol. Toxicol.* **58**, 663–687 (2018).
45. Takahashi, S., Hori, K., Ogasawara, H., Hiwatashi, K. & Sugiyama, T. Effects of nucleotides on the interaction of renin with GlcNAc 2-epimerase (renin binding protein, RnBP). *J. Biochem.* **140**, 725–30 (2006).
46. Wu, C.-H. *et al.* Renin-Angiotensin System and Cardiovascular Functions. *Arterioscler. Thromb. Vasc. Biol.* **38**, e108–e116 (2018).

47. Gertow, K. *et al.* Identification of the BCAR1-CFDP1-TMEM170A locus as a determinant of carotid intima-media thickness and coronary artery disease risk. *Circ. Cardiovasc. Genet.* **5**, 656–665 (2012).
48. Boardman-Pretty, F. *et al.* Functional Analysis of a Carotid Intima-Media Thickness Locus Implicates BCAR1 and Suggests a Causal Variant. *Circ. Cardiovasc. Genet.* **8**, 696–706 (2015).
49. Barrett, A., Pellet-Many, C., Zachary, I. C., Evans, I. M. & Frankel, P. p130Cas: a key signalling node in health and disease. *Cell. Signal.* **25**, 766–77 (2013).
50. Camacho Leal, M. del P. *et al.* p130Cas/BCAR1 scaffold protein in tissue homeostasis and pathogenesis. *Gene* **562**, 1–7 (2015).
51. Whelan, J., Gouffon, J. & Zhao, Y. Effects of dietary stearidonic acid on biomarkers of lipid metabolism. *J. Nutr.* **142**, 630S-634S (2012).
52. Hirafuji, M., Machida, T., Hamaue, N. & Minami, M. Cardiovascular protective effects of n-3 polyunsaturated fatty acids with special emphasis on docosahexaenoic acid. *J. Pharmacol. Sci.* **92**, 308–316 (2003).
53. Khan, S. U. *et al.* Effect of omega-3 fatty acids on cardiovascular outcomes: A systematic review and meta-analysis. *eClinicalMedicine* **38**, (2021).
54. Mita, T. *et al.* Eicosapentaenoic acid reduces the progression of carotid intima-media thickness in patients with type 2 diabetes. *Atherosclerosis* **191**, 162–167 (2007).
55. Sung, J., Jeon, H., Kim, I. H., Jeong, H. S. & Lee, J. Anti-Inflammatory Effects of Stearidonic Acid Mediated by Suppression of NF- $\kappa$ B and MAP-Kinase Pathways in Macrophages. *Lipids* **52**, 781–787 (2017).
56. Al Ashmar, S. *et al.* Metabolomic profiling reveals key metabolites associated with hypertension progression. *Front. Cardiovasc. Med.* **11**, 1–11 (2024).
57. Li, S. *et al.* Novel Metabolites Are Associated With Thrombosis After Cardiac Surgery in Children With Congenital Heart Disease. *Arterioscler. Thromb. Vasc. Biol.* **43**, 2231–2239 (2023).
58. Trotter, K. W. & Archer, T. K. The BRG1 transcriptional coregulator. *Nucl. Recept. Signal.* **6**, (2008).
59. Cave, L. *et al.* Hemin-driven chromatin remodelling by atherosclerotic risk gene SMARCA4 switches human blood-derived macrophages from leukocyte disposal to erythrocyte disposal. *bioRxiv* 2023.05.01.538808 (2023) doi:10.1101/2023.05.01.538808.
60. Wang, C. *et al.* Label-free quantitative proteomics identifies Smarca4 is involved in vascular calcification. *Ren. Fail.* **41**, 220–228 (2019).
61. Xiao, C. *et al.* Chromatin-remodelling factor Brg1 regulates myocardial proliferation and regeneration in zebrafish. *Nat. Commun.* **7**, 1–13 (2016).
62. Massagué, J. TGF $\beta$  signalling in context. *Nat. Rev. Mol. Cell Biol.* **13**, 616–30 (2012).
63. Goumans, M.-J. & Ten Dijke, P. TGF- $\beta$  Signaling in Control of Cardiovascular Function. *Cold Spring Harb. Perspect. Biol.* **10**, 1–39 (2018).
64. Zamani, N. & Brown, C. W. Emerging roles for the transforming growth factor- $\beta$  superfamily in regulating adiposity and energy expenditure. *Endocr. Rev.* **32**, 387–403 (2011).
65. Samad, F., Yamamoto, K., Pandey, M. & Loskutoff, D. J. Elevated expression of

- transforming growth factor- $\beta$  in adipose tissue from obese mice. *Mol. Med.* **3**, 37–48 (1997).
66. Fain, J. N., Tichansky, D. S. & Madan, A. K. Transforming growth factor beta1 release by human adipose tissue is enhanced in obesity. *Metabolism.* **54**, 1546–51 (2005).
  67. Stanley, P. Golgi glycosylation. *Cold Spring Harb. Perspect. Biol.* **3**, 1–13 (2011).
  68. Caffaro, C. E. & Hirschberg, C. B. Nucleotide sugar transporters of the Golgi apparatus: From basic science to diseases. *Acc. Chem. Res.* **39**, 805–812 (2006).
  69. Sesma, J. I. *et al.* Endoplasmic reticulum/golgi nucleotide sugar transporters contribute to the cellular release of UDP-sugar signaling molecules. *J. Biol. Chem.* **284**, 12572–12583 (2009).
  70. Karmazyn, M. Chapter 3 - Endothelin-1 as a Cardiac-Derived Autocrine, Paracrine and Intracrine Factor in Heart Health and Disease. in *Endocrinology of the Heart in Health and Disease* (eds. Schisler, J. C., Lang, C. H. & Willis, M. S.) 59–85 (Academic Press, 2017). doi:<https://doi.org/10.1016/B978-0-12-803111-7.00003-8>.
  71. Fan, J., Unoki, H., Iwasa, S. & Watanabe, T. Role of endothelin-1 in atherosclerosis. *Ann. N. Y. Acad. Sci.* **902**, 84–93; discussion 93-4 (2000).
  72. Stölting, M. *et al.* Monitoring endothelin-a receptor expression during the progression of atherosclerosis. *Biomedicines* **8**, 1–15 (2020).
  73. Yoon, M. H. *et al.* Long-term endothelin receptor antagonism attenuates coronary plaque progression in patients with early atherosclerosis. *Int. J. Cardiol.* **168**, 1316–21 (2013).
  74. Hasokawa, M. *et al.* Identification of biomarkers of stent restenosis with serum metabolomic profiling using gas chromatography/mass spectrometry. *Circ. J.* **76**, 1864–1873 (2012).
  75. Ottosson, F., Smith, E., Melander, O. & Fernandez, C. Altered asparagine and glutamate homeostasis precede coronary artery disease and type 2 diabetes. *J. Clin. Endocrinol. Metab.* **103**, 3060–3069 (2018).
  76. Patel, A., Abdelmalek, L., Thompson, A. & Jialal, I. Decreased homoserine levels in metabolic syndrome. *Diabetes Metab. Syndr. Clin. Res. Rev.* **14**, 555–559 (2020).
  77. Aa, N. *et al.* Plasma Metabolites Alert Patients With Chest Pain to Occurrence of Myocardial Infarction. *Front. Cardiovasc. Med.* **8**, (2021).
  78. Ojha, N. & Dhamoon, A. S. Myocardial Infarction. in *StatPearls [Internet]. Treasure Island (FL): StatPearls Publishing* (2024).
  79. Inouye, M. *et al.* Genomic Risk Prediction of Coronary Artery Disease in 480,000 Adults: Implications for Primary Prevention. *J. Am. Coll. Cardiol.* **72**, 1883–1893 (2018).
  80. Li, S. *et al.* Cardiomyocytes disrupt pyrimidine biosynthesis in nonmyocytes to regulate heart repair. *J. Clin. Invest.* **132**, (2022).
  81. Mazzola, A. *et al.* Opposite effects of uracil and adenine nucleotides on the survival of murine cardiomyocytes: In Focus: Heart. *J. Cell. Mol. Med.* **12**, 522–536 (2008).
  82. Yitzhaki, S., Shneyvays, V., Jacobson, K. A. & Shainberg, A. Involvement of uracil nucleotides in protection of cardiomyocytes from hypoxic stress. *Biochem. Pharmacol.* **69**, 1215–23 (2005).
  83. Hutchings, D. C., Anderson, S. G., Caldwell, J. L. & Trafford, A. W.

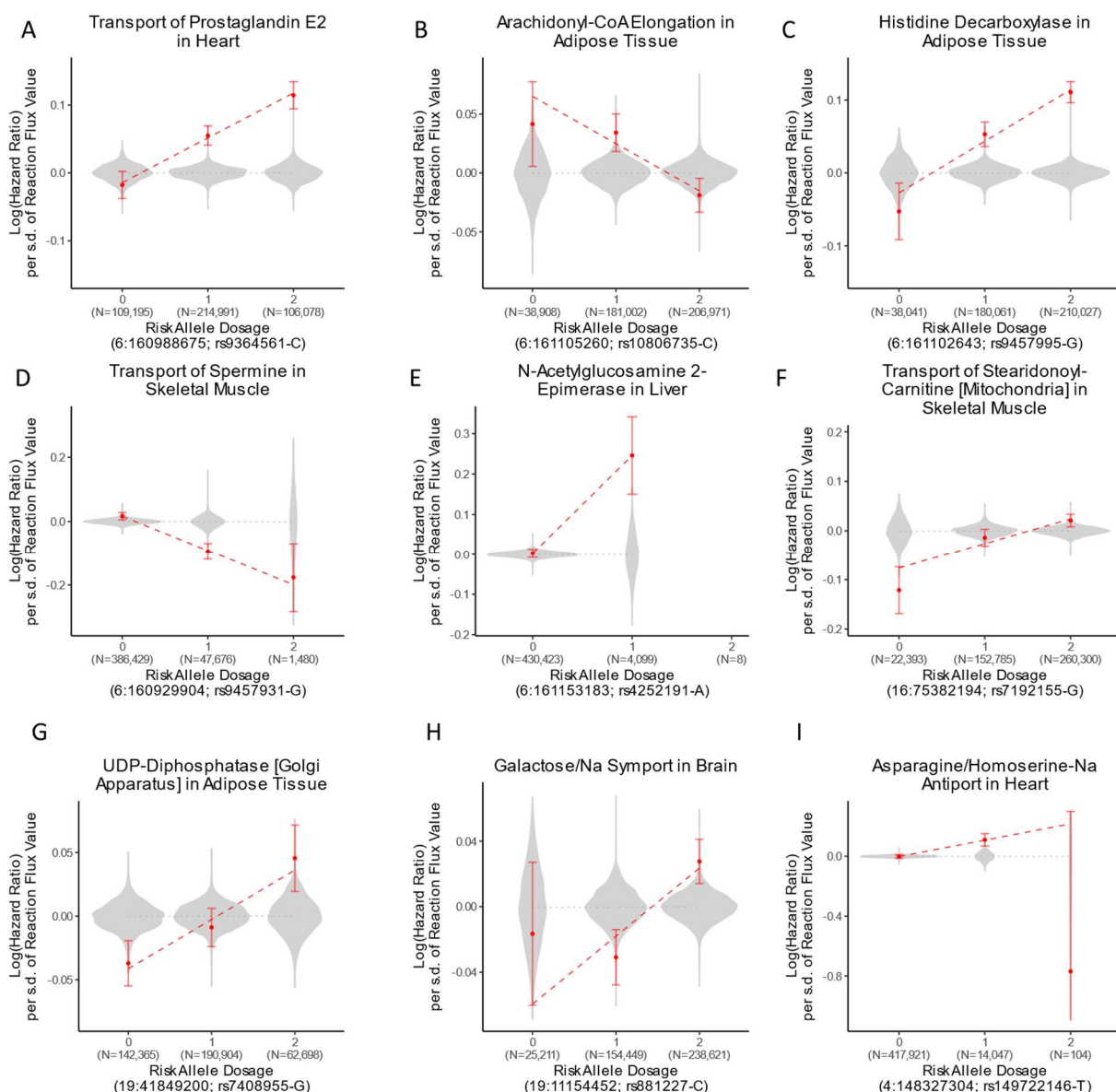
- Phosphodiesterase-5 inhibitors and the heart: Compound cardioprotection? *Heart* **104**, 1244–1250 (2018).
84. Li, S. *et al.* Phosphodiesterase-5a Knock-out Suppresses Inflammation by Down-Regulating Adhesion Molecules in Cardiac Rupture Following Myocardial Infarction. *J. Cardiovasc. Transl. Res.* **14**, 816–823 (2021).
  85. Schunkert, H. *et al.* Large-scale association analysis identifies 13 new susceptibility loci for coronary artery disease. *Nat. Genet.* **43**, 333–8 (2011).
  86. Vaughan, D. E., Lazos, S. A. & Tong, K. Angiotensin II regulates the expression of plasminogen activator inhibitor-1 in cultured endothelial cells. A potential link between the renin-angiotensin system and thrombosis. *J. Clin. Invest.* **95**, 995–1001 (1995).
  87. Brown, N. J., Agirbasli, M. A., Williams, G. H., Litchfield, W. R. & Vaughan, D. E. Effect of activation and inhibition of the renin-angiotensin system on plasma PAI-1. *Hypertens. (Dallas, Tex. 1979)* **32**, 965–71 (1998).
  88. Lehner, R. & Quiroga, A. D. *Fatty Acid Handling in Mammalian Cells. Biochemistry of Lipids, Lipoproteins and Membranes: Sixth Edition* (Elsevier, 2016). doi:10.1016/B978-0-444-63438-2.00005-5.
  89. Yin, X., Xu, C., Xu, Q. & Lang, D. Docosahexaenoic acid inhibits vascular smooth muscle cell migration and proliferation by decreasing microRNA-155 expression levels. *Mol. Med. Rep.* **22**, 3396–3404 (2020).
  90. Nüchel, J. *et al.* TGFB1 is secreted through an unconventional pathway dependent on the autophagic machinery and cytoskeletal regulators. *Autophagy* **14**, 465–486 (2018).
  91. Zhang, J., ten Dijke, P., Wuhrer, M. & Zhang, T. Role of glycosylation in TGF- $\beta$  signaling and epithelial-to-mesenchymal transition in cancer. *Protein and Cell* vol. 12 89–106 at <https://doi.org/10.1007/s13238-020-00741-7> (2021).
  92. Jain, S. & Jacobson, K. A. Adipocyte purinergic receptors activated by uracil nucleotides as obesity and type 2 diabetes targets. *Curr. Opin. Pharmacol.* **63**, 102190 (2022).
  93. Chen, J. *et al.* P2Y12 inhibitor clopidogrel inhibits renal fibrosis by blocking macrophage-to-myofibroblast transition. *Mol. Ther.* **30**, 3017–3033 (2022).
  94. Mederacke, I. *et al.* The purinergic P2Y14 receptor links hepatocyte death to hepatic stellate cell activation and fibrogenesis in the liver. *Sci. Transl. Med.* **14**, eabe5795 (2022).
  95. Borges da Silva, H. *et al.* Sensing of ATP via the Purinergic Receptor P2RX7 Promotes CD8<sup>+</sup> Trm Cell Generation by Enhancing Their Sensitivity to the Cytokine TGF- $\beta$ . *Immunity* **53**, 158-171.e6 (2020).
  96. Loh, P.-R. *et al.* Reference-based phasing using the Haplotype Reference Consortium panel. *Nat. Genet.* **48**, 1443–1448 (2016).
  97. Thiele, I. *et al.* Personalized whole-body models integrate metabolism, physiology, and the gut microbiome. *Mol. Syst. Biol.* (2020) doi:10.15252/msb.20198982.
  98. Pornputtpong, N., Nookaew, I. & Nielsen, J. Human metabolic atlas: an online resource for human metabolism. *Database* **2015**, 1–9 (2015).
  99. Schmidt, B. J. *et al.* GIM3E: condition-specific models of cellular metabolism developed from metabolomics and expression data. *Bioinformatics* **29**, 2900–8 (2013).

100. GTEx Consortium. The Genotype-Tissue Expression (GTEx) project. *Nat. Genet.* **45**, 580–5 (2013).
101. Davidson-Pilon, C. lifelines: survival analysis in Python. *J. Open Source Softw.* **4**, 1317 (2019).
102. Wilcox, R. R. One-way anova. in *Applying Contemporary Statistical Techniques* 285–328 (2003). doi:10.1016/B978-0-12-751541-0.50030-4.
103. Chang, C. C. *et al.* Second-generation PLINK: rising to the challenge of larger and richer datasets. *Gigascience* **4**, 7 (2015).
104. McLaren, W. *et al.* The Ensembl Variant Effect Predictor. *Genome Biol.* **17**, 122 (2016).
105. Tokolyi, A. *et al.* Genetic determinants of blood gene expression and splicing and their contribution to molecular phenotypes and health outcomes. *medRxiv* 1–29 (2023).
106. Gel, B. & Serra, E. karyoploteR: an R/Bioconductor package to plot customizable genomes displaying arbitrary data. *Bioinformatics* **33**, 3088–3090 (2017).

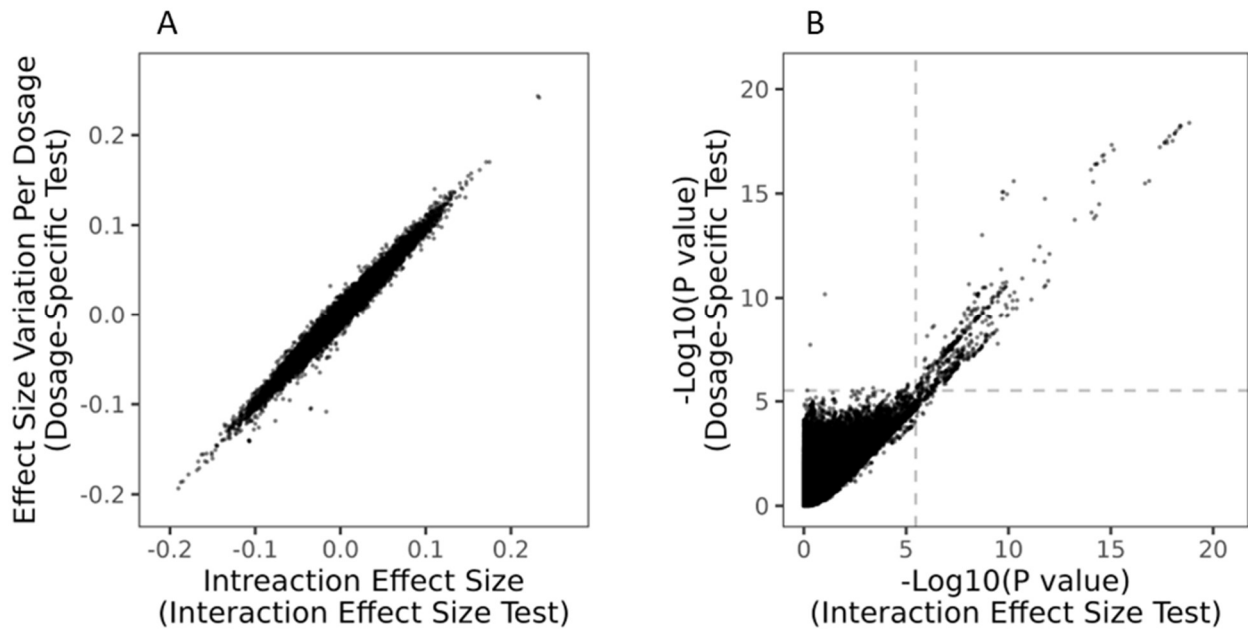
## Supplementary Figures and Tables

**Table S1:** Cases and controls under the definition of coronary atherosclerosis or myocardial infarction in the inferred European ancestry subset of UKB. Under the definition of coronary atherosclerosis, individuals with any of the constitutive ICD codes of ischemic heart disease were excluded from the controls.

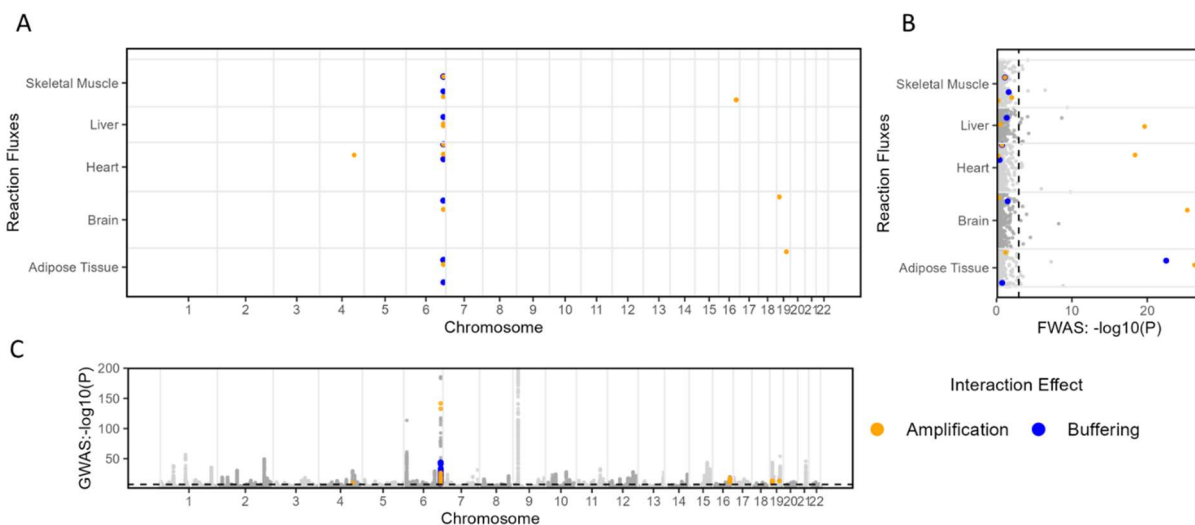
		Myocardial Infarction		
		Control	Case	Excluded
Coronary Atherosclerosis	Control	397254	970	58
	Case	10220	27682	39
	Excluded	16155	7355	0



**Figure S1:** Examples of risk allele dosage-specific reaction effect sizes for a representative set of SNP-reaction flux pairs with significant interaction. Reaction flux effect sizes (i.e.  $\log(\text{Hazard Ratio})$ ) on CAD risk were estimated using Cox regression in European UKB participants carrying different risk allele dosages. Error bars denote the 95% confidence intervals for reaction effect sizes. The dashed line indicates the linear regression of flux effect size per dosage weighted by the standard error of effect size estimates ( $1/\text{SE}^2$ ). Violin plots indicate the distribution of all other risk reaction flux effect sizes for the same risk allele and the dotted line is the linear regression for these effect sizes. Genome coordinates correspond to the GRCh37 genome assembly.

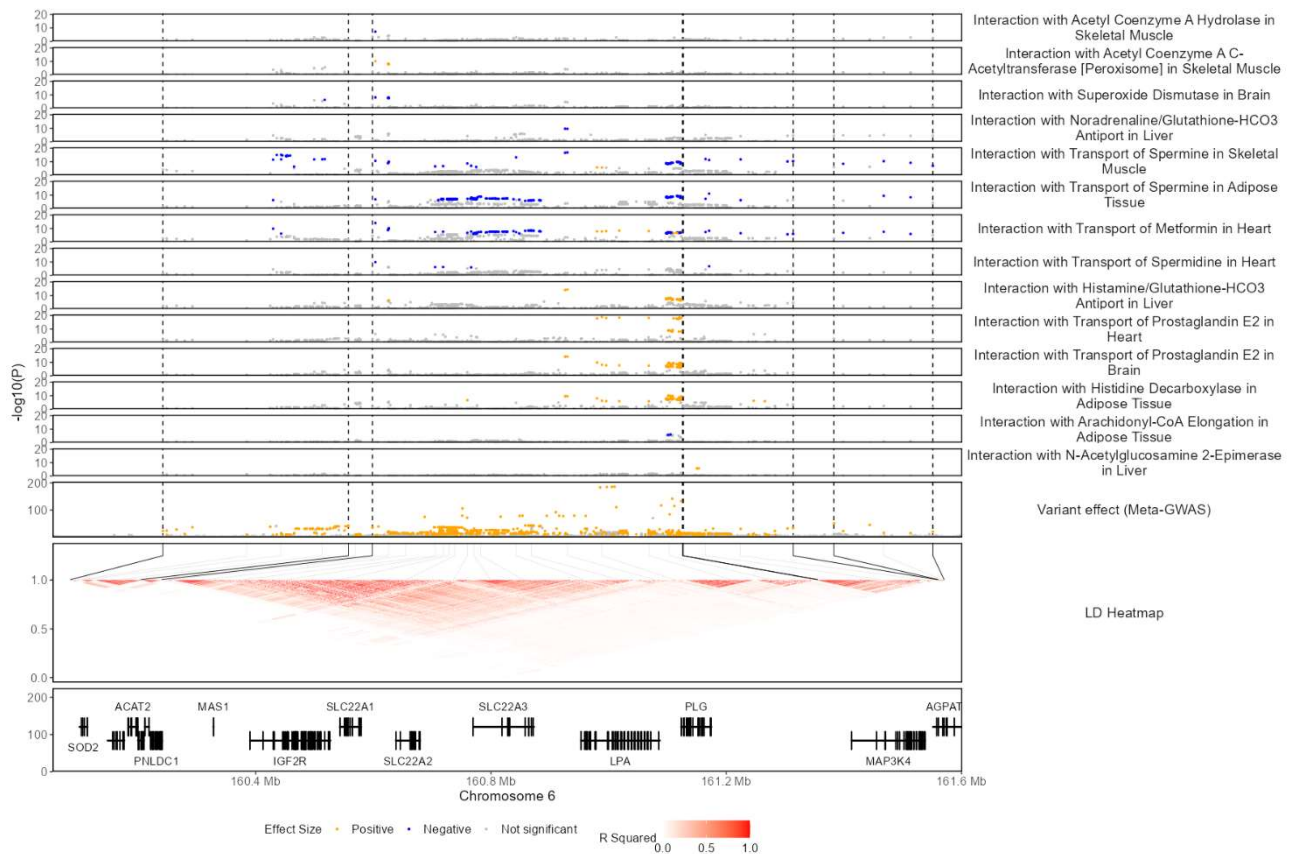


**Figure S2:** Comparison of interaction effect size and dosage-specific tests. A) Comparison of interaction effect sizes estimates. B) Comparison of interaction effect significance ( $-\log_{10}(\text{P-values})$ ). The dashed grey line indicates the nominal P-value significance threshold estimated after Benjamini-Hochberg multiple testing correction (i.e.,  $\text{FDR} < 0.05$ ).

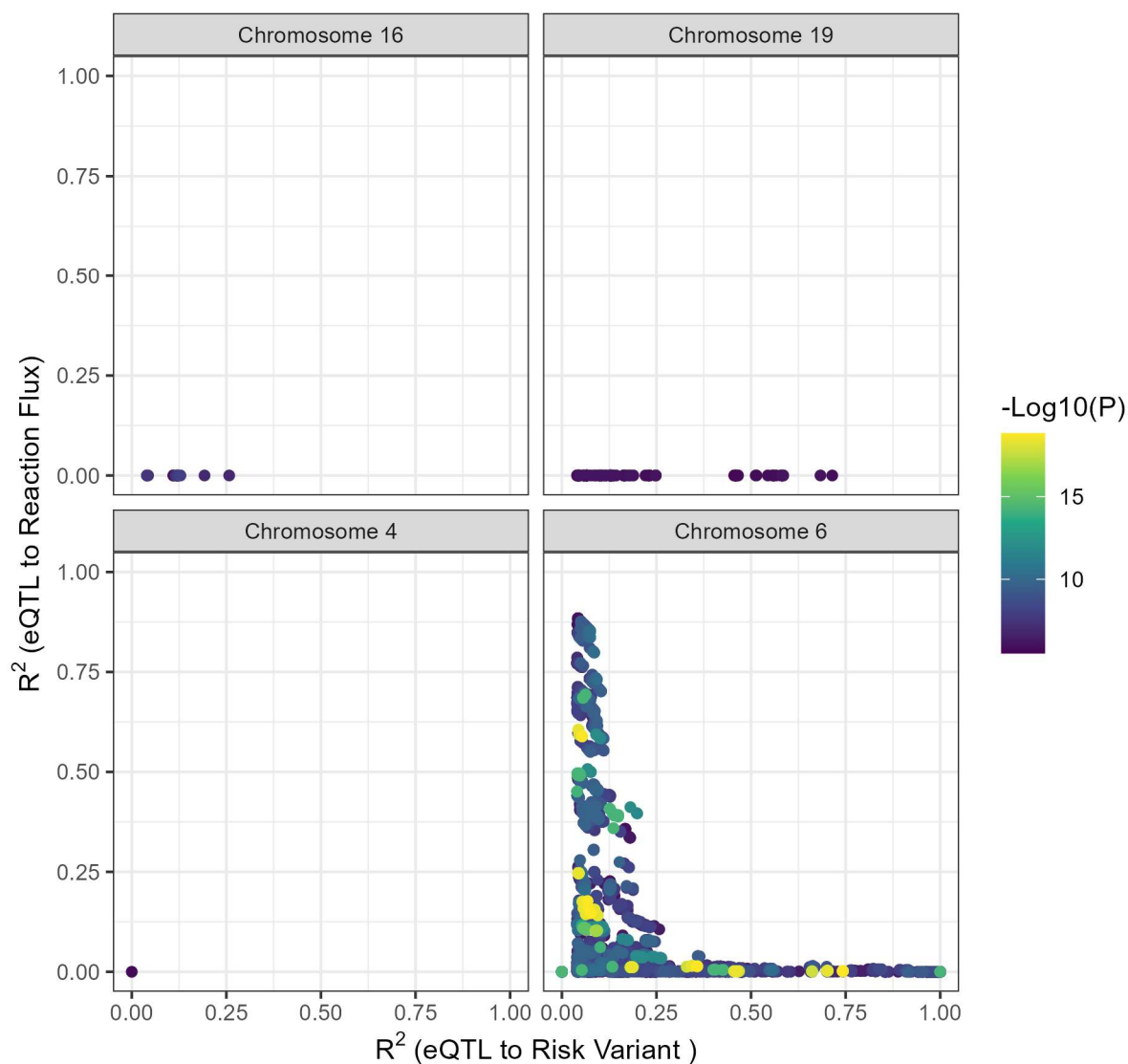


**Figure S3:** Variants and reaction fluxes with significant interaction. A) Organ metabolic network and chromosome of the reaction fluxes and variants, respectively, with significant interaction on CAD risk. B) and C) Manhattan plots of reaction fluxes and variants effects on CAD risk when analysed in univariate analysis. Variant effect sizes P-values were obtained from a published GWAS meta-analysis<sup>28</sup>. P-values for reaction effect size were computed for uncorrelated reaction fluxes with a Cox proportional hazard model in the inferred European ancestry subset of UKB. The dashed grey line indicates the P-value threshold to consider a variant or reaction effect size significant.

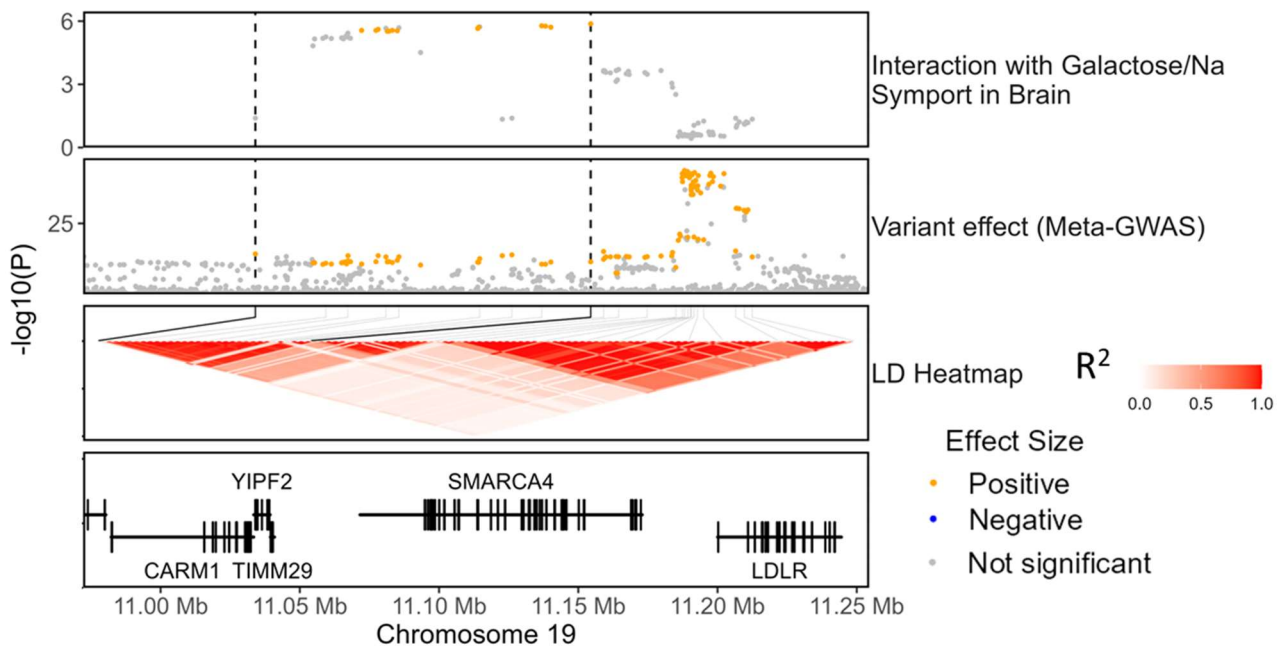




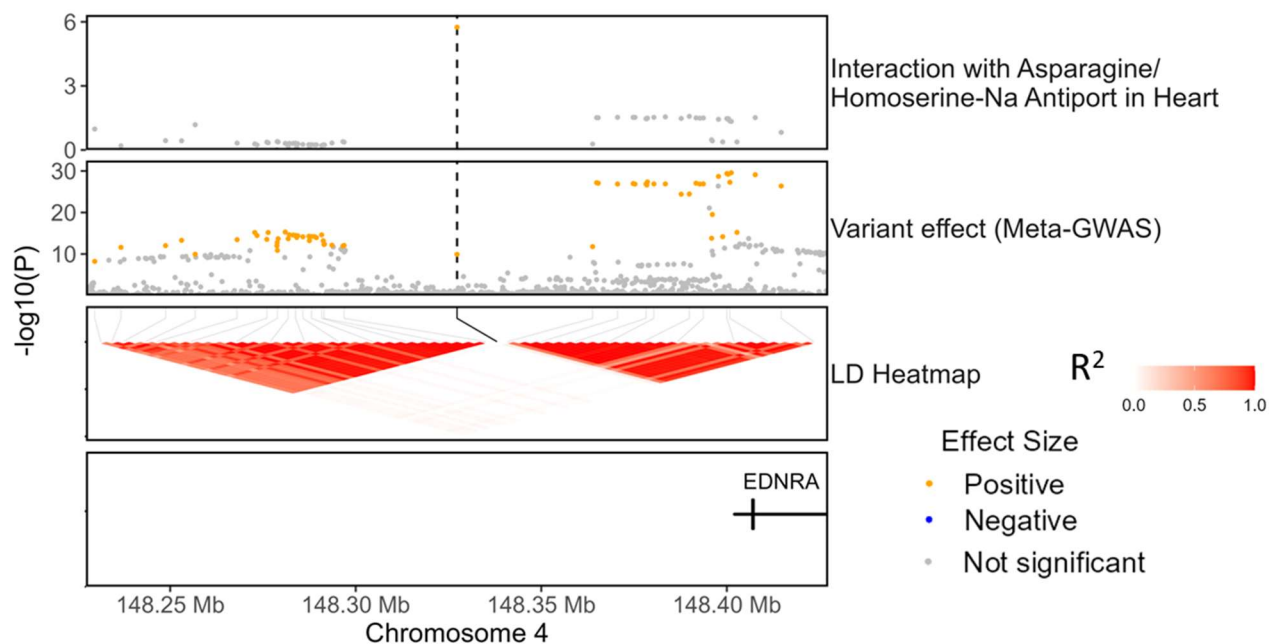
**Figure S4:** Significant SNP-Flux Interactions in the LPA/PLG risk loci. The regional association plots show the  $-\log_{10}(P)$  value for interaction and variant effect sizes on CAD risk. P-values for interaction and variant effect sizes were derived from the interaction effect size test and meta-GWAS summary statistics, respectively. The LD heatmap indicates the pairwise LD for SNPs with genome-wide significant effect size on CAD in UKB participants of European genetic ancestries. Dashed black lines indicate the limits of LD blocks ( $R^2 > 0.6$ ) used to define independent risk loci. To facilitate visualization, only LD blocks with variants involved in significant interactions are highlighted. Only protein-coding genes are shown in the gene plot. Genome coordinates correspond to the GRCh37 genome assembly.



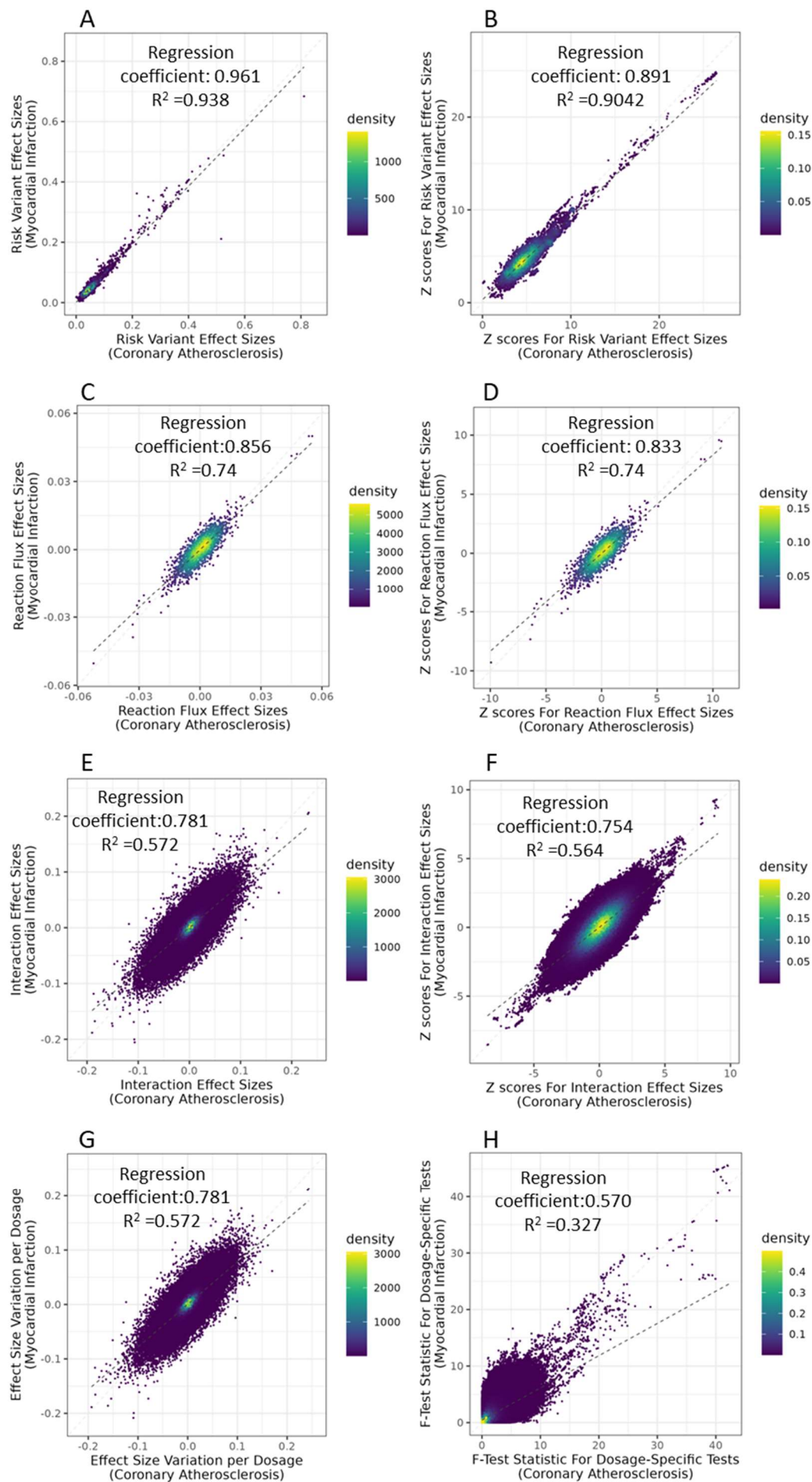
**Figure S5:** Linkage disequilibrium between risk variants and eQTL variants used as input for flux simulation and their correlation to the reaction flux value for each pair of risk-variant and reaction flux with significant interaction in CAD risk. For each pair, the effect of the risk variant has been regressed out of the flux (**Methods**). Linkage disequilibrium under 0.04 are plotted as 0. Data is coloured based on the interaction effect size P-value for each risk-variant reaction flux pair.



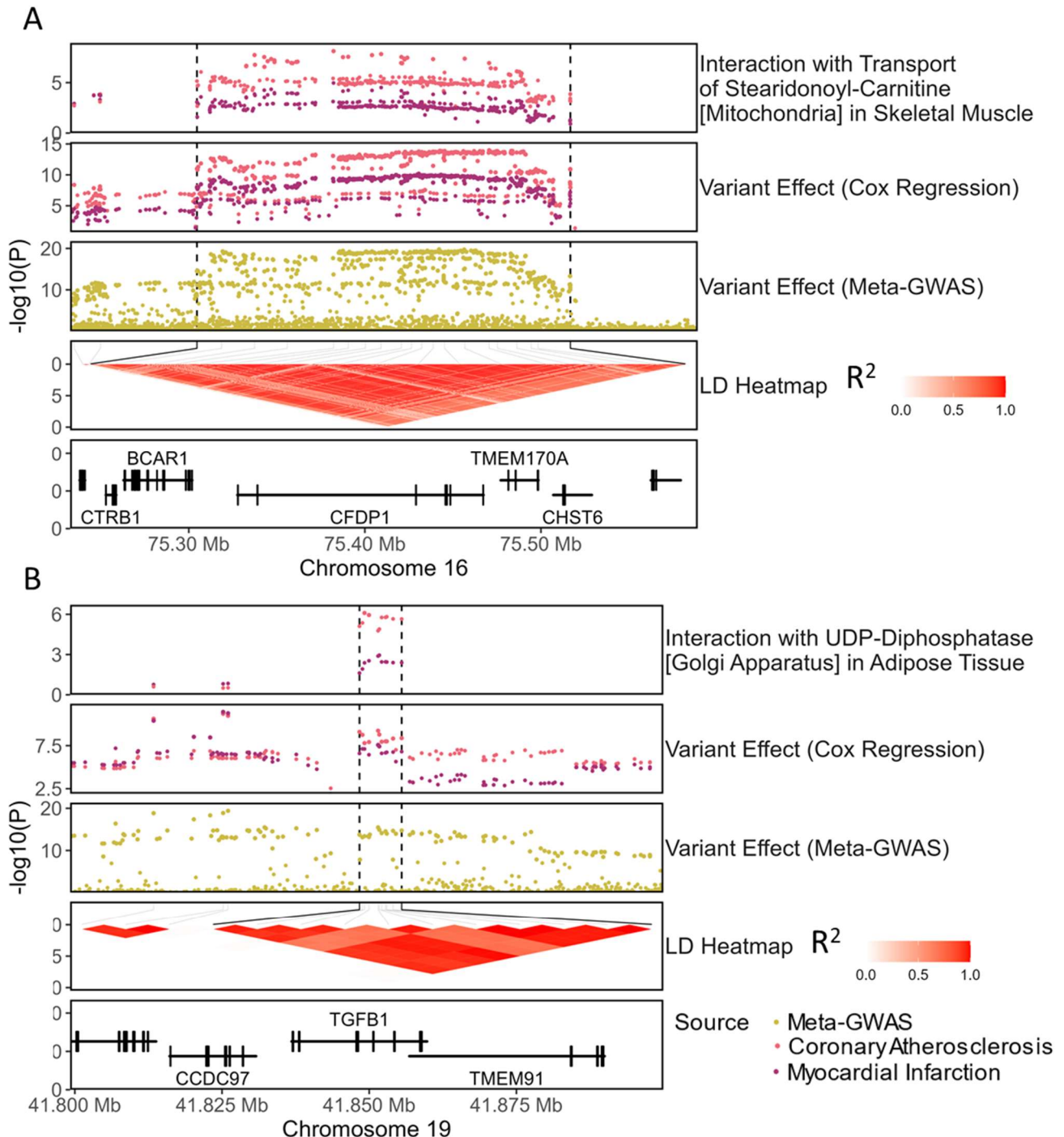
**Figure S6:** Interaction between galactose transport in brain and variants in the *SMARCA4* risk locus. The regional association plots show the  $-\log_{10}(P\text{-value})$  for interaction and variant effect sizes on CAD risk. P-values for interaction and variant effect sizes were derived from the interaction effect size test and meta-GWAS summary statistics, respectively.



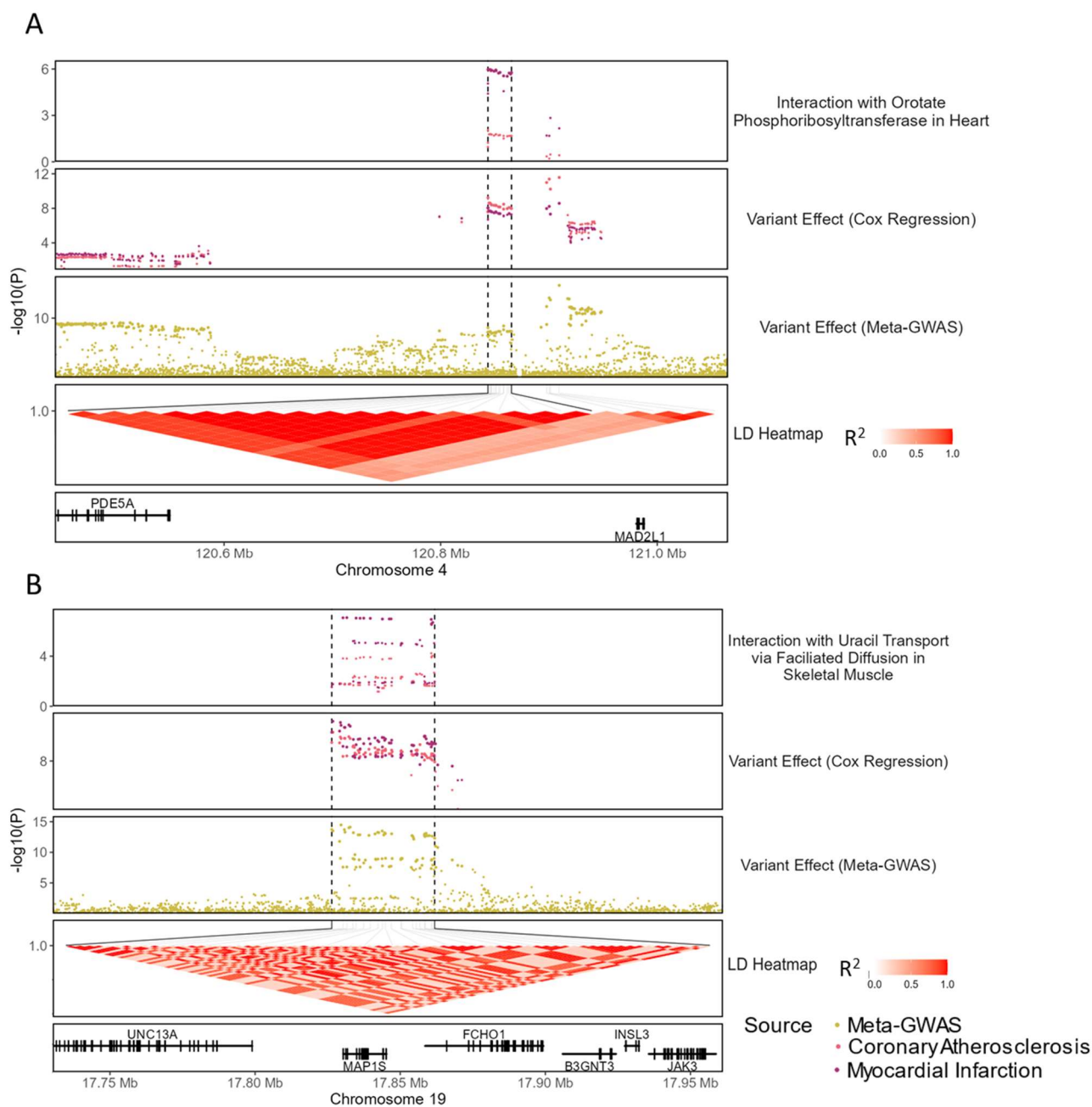
**Figure S7:** Amino acid transport amplifies the effect of a risk variant at the *EDNRA* locus. The regional association plots show the  $-\log_{10}(P\text{-value})$  for interaction and variant effect sizes on CAD risk. P-values for interaction and variant effect sizes were derived from the interaction effect size test and meta-GWAS summary statistics, respectively.



**Figure S8:** Effect sizes for the risk of coronary atherosclerosis and myocardial infarction. Effect sizes and Z scores for effect sizes for risk variants (A,B), reaction fluxes (C,B), interaction between reaction fluxes and risk variants (E,F), and reaction effect size variation and F-statistic for the dosage-specific test (G,H) using two alternative disease definitions: coronary atherosclerosis and myocardial infarction. The set of evaluated risk variants was obtained from a published GWAS meta-analysis<sup>28</sup>. The reaction flux effect size was evaluated for 1,670 uncorrelated reaction fluxes (**Methods**). Effect sizes were evaluated with a Cox proportional hazard model in the inferred European ancestry subset of UKB. The dosage-specific test evaluated the variation of reaction flux effect size per each risk allele dosage in the same UKB subset (**Methods**). The dashed black line indicates the linear regression for the plotted values.



**Figure S9:** Example of interactions that are attenuated in myocardial infarction. The regional association plots show the  $-\log_{10}(P\text{-value})$  for interaction and variant effect sizes on disease risk. P-values for interaction effect sizes and variant effect (Cox Regression) were derived from the interaction and variant effect size test, respectively, in the European subset of UKB for coronary atherosclerosis or myocardial infarction events. As a reference, the variant effects for the meta-GWAS on CAD risk are also plotted.



**Figure S10:** Example of interactions that are specific to myocardial infarction. The regional association plots show the  $-\log_{10}(P\text{-value})$  for interaction and variant effect sizes on disease risk. P-values for interaction effect sizes and Variant Effect(Cox Regression) were derived from the interaction and variant effect size test, respectively, in the European subset of UKB for coronary atherosclerosis or myocardial infarction events. As a reference, the variant effects for the meta-GWAS on CAD risk are also plotted.

**Supplementary Data S1:** Table providing the reaction and variant annotation, and summary stats for pairs of risk alleles and reaction fluxes with significant interaction in either CAD or MI risk. Human1 and Recon3D Reaction IDs indicate the identifiers of the reactions in the HUMAN1 and Recon3D genome-scale models of human metabolism. Reaction organ indicates the organ-specific metabolic network where the flux is located. Genes mapped to reaction indicate the enzymes or transmembrane carriers mediating the reactions as defined in the gene-reaction rules of HUMAN1. eGene(s) and sGene(s) indicate the genes whose expression and splicing, respectively, are associated with the risk alleles. Variant risk locus indicates the LD block ( $R^2 > 0.6$ ) of risk variants where interacting risk alleles are mapped. The interaction between risk alleles and reaction fluxes is measured with two complementary approaches: an interaction effect size test and a dosage-specific test (**Methods**). Genome coordinates correspond to the GRCh37 genome assembly.