

# 1 **Integrating multi-modal omics to identify therapeutic atherosclerosis**

## 2 **pathways for coronary heart disease**

3

### 4 **Authors and affiliations**

5 Sophie C. de Ruiter, MSc <sup>1</sup>, Marion van Vugt, PhD <sup>2,3,4</sup>, Chris Finan, PhD <sup>3,4,5,6</sup>, Diederick E. Grobbee, PhD <sup>1</sup>,  
6 Dominique P.V. de Kleijn, PhD <sup>7</sup>, Gerard Pasterkamp, PhD <sup>8</sup>, Hester M. den Ruijter, PhD <sup>9</sup>, Ernest Diez  
7 Benavente, PhD <sup>9,\*</sup>, Sanne A.E. Peters, PhD <sup>1,10,\*</sup>, A. Floriaan Schmidt, PhD <sup>2,3,4,5,\*</sup>

8

- 9 1. Julius Center for Health Sciences and Primary Care, University Medical Center Utrecht, Utrecht  
10 University, Utrecht, the Netherlands
- 11 2. Department of Cardiology, Amsterdam Cardiovascular Sciences, Amsterdam University Medical  
12 Centres, University of Amsterdam, Amsterdam, the Netherlands
- 13 3. Institute of Cardiovascular Science, Faculty of Population Health, University College London,  
14 London, United Kingdom
- 15 4. Division Heart and Lungs, Department of Cardiology, University Medical Center Utrecht, Utrecht  
16 University, Utrecht, the Netherlands
- 17 5. UCL British Heart Foundation Research Accelerator, London, United Kingdom
- 18 6. Health Data Research UK and Institute of Health Informatics, University College London, London,  
19 United Kingdom
- 20 7. Department of Vascular Surgery, University Medical Center Utrecht, Utrecht University, Utrecht,  
21 the Netherlands
- 22 8. Central Diagnostic Laboratory, University Medical Center Utrecht, Utrecht University, Utrecht, the  
23 Netherlands
- 24 9. Laboratory of Experimental Cardiology, Division Heart and Lungs, University Medical Center  
25 Utrecht, Utrecht University, Utrecht, the Netherlands
- 26 10. The George Institute for Global Health, School of Public Health, Imperial College London,  
27 London, UK

28 \* joint senior author

29

30 Running title: therapeutic targets and biomarkers for coronary heart disease

31 Word count text: 4036

32 Word count abstract: 245

33 Number of references: 80

34 Number of tables and figures: 6

35 Number of (web) appendices: 3

36

1 **Abstract**

2 **Introduction:** Urinary metabolism breakdown products reflect metabolic changes in atherosclerosis-  
3 relevant tissues and may contain relevant therapeutic leads. We integrated data on urinary metabolism  
4 breakdown products, plasma proteins, atherosclerotic plaque tissue, and single-cell expression to  
5 identify druggable metabolic pathways for coronary heart disease (CHD).

6 **Methods:** Mendelian randomisation was employed to interrogate findings from independent genome-  
7 wide association studies on 954 urinary metabolism breakdown products, 1,562 unique proteins, and  
8 181,522 CHD cases, establishing directionally concordant associations. Using the Athero-Express  
9 Biobank, concordant plasma proteins were linked to plaque vulnerability using protein and mRNA  
10 expression in plaque. Single-cell RNA sequencing data obtained from carotid plaque samples were  
11 used to test for differential expression of concordant proteins across plaque cell types.

12 **Results:** In total, 29 urinary metabolism breakdown products associated with CHD, predominantly  
13 originating from amino acid metabolism (n=12) or unclassified origin (n=9). We identified 113  
14 plasma proteins with directionally concordant associations with these urinary metabolism breakdown  
15 products and CHD. Of the 110 proteins available in plaque, 16 were associated with plaque  
16 vulnerability. This included positive control proteins targeted by drugs indicated for CHD, such as  
17 CAH1 (targeted by aspirin), IL6R (targeted by tocilizumab), and AT1B2 (targeted by digoxin), as  
18 well as two potential repurposing opportunities C1S (targeted by C1-esterase inhibitor and  
19 sutimlimab) and CATH (targeted by bortezomib).

20 **Conclusion:** We have identified amino acid metabolism as an important contributing pathway to  
21 CHD risk and prioritised 16 proteins relevant for CHD with involvement in atherosclerotic plaques,  
22 providing important insights for drug development.

## 1 **Introduction**

2 Atherosclerosis is characterised by the development of lesions accumulating lipoproteins and  
3 inflammatory cells in the arterial wall. As atherosclerosis progresses, plaque accumulates and may  
4 progress into high-risk vulnerable plaques. Such plaques are characterized by a thin fibrous cap, a  
5 large lipid core and high levels of inflammatory cells(1, 2), making them particularly prone to rupture.  
6 Plaque rupture can lead to rapid thrombus formation blocking the blood flow to the heart muscle,  
7 potentially leading to acute clinical events such as coronary heart disease (CHD), which remains a  
8 leading cause of burden and death globally.

9  
10 Metabolic disorders, such as metabolic syndrome (the co-occurrence of hypertension,  
11 hypercholesterolemia, diabetes, and obesity), are associated with a metabolic, pro-inflammatory, and  
12 pro-thrombotic state which are key drivers of atherosclerosis(3), reflecting the close interrelationship  
13 between immune response and metabolic homeostasis, where dysfunction in one system can adversely  
14 affect the other(4). Plasma proteins are markers of disease and targets of most drug compounds.  
15 Previous research has already linked plaque phenotypes to circulating proteins, identifying reduced  
16 plasminogen activator inhibitor levels in patients with an intermediate coronary artery disease risk  
17 profile, for example(5).

18  
19 Urinary metabolism breakdown products, which are the end-products of metabolism excreted in urine,  
20 reflect metabolic changes occurring in atherosclerosis-relevant tissues (e.g., body fat, liver, the arterial  
21 wall). These urinary breakdown products may provide key insights in the metabolic alterations  
22 associated with the progression of atherosclerosis to CHD. While urinary breakdown products are  
23 potentially important markers of disease, they are not directly actionable as they have been excreted  
24 from the body. On the contrary, plasma proteins are actionable targets for therapeutic intervention and  
25 therefore the integration of data on urinary metabolites and plasma proteins identifies relevant  
26 metabolic pathways and potential drug targets, offering a comprehensive approach to understanding  
27 and intervening in the progression of CHD.

1  
2 In the current study, we used Mendelian randomisation (MR) to identify novel metabolic pathways for  
3 CHD by leveraging genetic data on urinary metabolism breakdown products (954 urinary metabolism  
4 breakdown products), plasma proteins (1,562 unique plasma proteins), and CHD (181,522 cases). The  
5 fixed nature of genotypes, established during gametogenesis(6, 7) allows genetic variants to be used  
6 as instrumental variables in MR, which is therefore robust against the presence of confounding bias  
7 and potential reverse causation. In the context of cardiovascular disease, this approach has been  
8 empirically validated(8-11). Plaque involvement was established by using data from carotid  
9 endarterectomy patients from the Athero-Express (AE) Biobank on protein expression, mRNA  
10 expression, and single-cell RNA sequencing. Finally, the potential druggability of these prioritised  
11 proteins was evaluated through linkage to ChEMBL.

12

## 13 **Methods**

### 14 **Genetic data**

15 Genetic associations with 954 urinary metabolism breakdown product values were available from a  
16 genome-wide association study (GWAS) conducted on 1,627 individuals using a non-targeted mass  
17 spectrometry Metabolon assay(12). Metabolism breakdown products were grouped based on their  
18 metabolic origins, which led to nine metabolism classes: amino acid metabolism, carbohydrate  
19 metabolism, cofactor and vitamin metabolism, energy metabolism, lipid metabolism, nucleotide  
20 metabolism, peptide metabolism, xenobiotic metabolism, or unclassified metabolism. Metabolism  
21 breakdown products with an origin in xenobiotic metabolism are breakdown products from substances  
22 not naturally produced by the human body, such as drugs, pollutants, and synthetic compounds.  
23 Unclassified metabolism breakdown products represent potentially novel biomarkers, as they have not  
24 yet been related to an established metabolic origin. These breakdown products were noted for their  
25 recurring chromatographic and spectral properties and have been assigned unique COMP-IDs and/or  
26 CHEM-IDs (**Appendix Table S1**), allowing for identification and comparison in future studies and by  
27 other analysis platforms. We used eight GWAS on genetic associations with plasma protein values,

1 detailed in the **Supplemental Methods**, and a GWAS on genetic associations with incident or  
2 prevalent CHD (181,522 cases among 1,165,690 participants from European ancestry) obtained from  
3 Aragam *et al.*(13).

#### 5 **Mendelian randomization analyses**

6 Three MR analyses were conducted: 1) genome-wide MR to identify urinary metabolism breakdown  
7 products associated with CHD through upstream effects in plasma (step 1 in **Figure 1, Appendix**  
8 **Table S1**), 2) *cis*-MR to identify plasma proteins affecting values of urinary metabolism breakdown  
9 products (step 2 in **Figure 1**), and 3) *cis*-MR to identify plasma proteins associated with CHD (step 3  
10 in **Figure 1**). For the genome-wide MR, genetic variants were selected from across the genome, while  
11 the *cis*-MR utilized genetic variants selected from a 200 kilobase pair window around and within the  
12 protein encoding gene. In all MRs, variants were selected based on an exposure F-statistic of at least  
13 24 and a minimal minor allele frequency of 0.01. The selected variants were clumped to a linkage  
14 disequilibrium (LD) r-squared of 0.30, using a random sample of 5,000 unrelated UK Biobank  
15 participants(14). The *cis*-MR analyses (steps 2-3 in **Figure 1**) were conducted per individual protein  
16 GWAS, where the largest sample size GWAS was used if protein measurements overlapped in  
17 different studies.

18  
19 MR analyses were conducted using generalized least squares (GLS) implementations of the inverse  
20 variance weighted (IVW) estimator, as well as MR-Egger estimator, known for its robustness to  
21 pleiotropy(15). The GLS implementation was used to additionally correct for residual LD, which  
22 optimized estimator precision(16). To reduce the potential for horizontal pleiotropy, variants with  
23 large leverage statistics (larger than three times the mean leverage) and/or outlier statistics (Chi-  
24 square larger than 10.83) were excluded. We discarded analyses with fewer than six variants, ensuring  
25 we had sufficient data to accurately model the exposure effects. To further minimize the potential of  
26 horizontal pleiotropy, we applied a model selection framework identifying the MR model (IVW or  
27 MR-Egger) that is most supported by the available data(17).

28

## 1 **Triangulation of evidence through consideration of concordant CHD effects**

2 We integrated findings from the three distinct MR analyses to triangulate evidence on the effects of  
3 urinary metabolism breakdown products and proteins on CHD. This involved 1) identifying urinary  
4 metabolism breakdown products affecting CHD, 2) identifying proteins associated with these urinary  
5 metabolism breakdown products, and 3) focussing on the subset of proteins with an effect on CHD  
6 that was directionally concordant with the effect of a urinary metabolism breakdown product on CHD  
7 and the protein on the same breakdown product. For example, if a higher value of a urinary  
8 metabolism breakdown product reduced CHD risk, and a protein increased the value of this urinary  
9 breakdown product, the protein effect was considered directionally concordant if higher protein values  
10 reduced CHD risk. We describe concordance in direction as a triangulated association (**Figure 2**).  
11 Given that potential horizontal pleiotropy acts through distinct pathways in each analysis (e.g., pre-  
12 translational pleiotropy in *cis*-MR versus pre- or post-translational horizontal pleiotropy in genome-  
13 wide MR(18)), focusing on results with directional concordance ensures identification of a robust  
14 subset of results with likely limited residual bias due to horizontal pleiotropy.

15

## 16 **Athero-Express Biobank atherosclerotic plaques**

17 Next, the Athero-Express (AE) Biobank data was used to prioritise findings on atherosclerotic plaque  
18 involvement. Patients enrolled in the AE Biobank were at risk of CHD, as well as other  
19 atherosclerosis associated diseases such as ischemic stroke, with carotid plaques from these patients  
20 providing a relevant marker of generalised atherosclerosis (54-58). Specifically, the AE contains  
21 samples of carotid endarterectomy patients, where we identified potential associations between  
22 protein (194 patients) or mRNA (632 patients) expression in plaque and the plaque vulnerability  
23 index. In addition, single-cell RNA sequencing data (4948 cells and 46 patients) were analysed to  
24 study cellular plaque expression. Please see **Supplemental Methods** for the relevant details, and  
25 **Appendix Table S2** for AE patients characteristics. The performed study is in line with the  
26 Declaration of Helsinki, and informed consent was provided by all study participants after approval  
27 for this study by the medical ethical committees of the various hospitals (University Medical Center,  
28 Utrecht, The Netherlands, and St. Antonius Hospital, Nieuwegein, The Netherlands) was obtained.

1

## 2 **Associations with plaque vulnerability**

3 We tested for associations between both protein expression and mRNA expression levels of MR-  
4 prioritised proteins and plaque vulnerability index using a linear model, with protein expression  
5 available for 1,500 proteins and mRNA expression available from 55,105 transcripts, all measured in  
6 atherosclerotic plaques.

7

## 8 **Determining cellular expression using single-cell RNA sequencing**

9 Single-cell RNA sequencing was used to explore whether genes coding for MR-prioritised proteins  
10 where differentially expressed across plaque cell types, providing important information on the  
11 mechanisms driving atherosclerosis. A Wilcoxon rank-sum test was used to compare the expression in  
12 a single cell type to expression in the remaining cell types. The following cell types were considered:  
13 dendritic cells, endothelial cells I (ECs I), endothelial cells II (ECs II), foam cells, inflammatory  
14 macrophages, mast cells, memory B-cells, monocytes, natural killer cells (NK-cells), plasma B-cells,  
15 resident macrophages, smooth muscle cells (SMCs), and T-cells. Additionally, differential expression  
16 across broader clusters of cell types (structural cells, innate immune cells, and adaptive immune cells)  
17 was assessed using the Wilcoxon rank-sum test (**Supplemental Methods**).

18

## 19 **Annotations, effect estimates and multiple testing**

20 Proteins are referred to using their Uniprot label(19) and genes are referred to using the Ensembl label  
21 in italicised font. MR results are presented as mean differences (MD) for urinary metabolism  
22 breakdown products, or odds ratios (OR) for CHD, representing the effect of one standard deviation  
23 (SD) increase of the exposure (i.e., either urinary metabolism breakdown products or plasma protein  
24 value). Associations of the plaque vulnerability index with protein expression levels or mRNA  
25 expression levels are presented as MD in normalized count per unit increase in the plaque  
26 vulnerability index. All effect estimates are accompanied by 95% confidence intervals (CIs) and p-  
27 values. Metabolism breakdown product effects on CHD were filtered for a multiplicity corrected p-  
28 value threshold of  $1.37 \times 10^{-4}$  based on the 365 principal components that were needed to explain

1 90% of the variance in the urinary metabolism breakdown products (**Appendix Figure S1**). Protein  
2 MR effect estimates of the metabolism breakdown product association were evaluated against a  
3 multiplicity corrected p-value threshold of  $1.10 \times 10^{-6}$  based on the number of tested proteins ( $n=1,562$ )  
4 and metabolism breakdown products ( $n=29$ ). The multiplicity corrected p-value threshold was  
5  $7.33 \times 10^{-6}$  for the MRs of protein on CHD based on the number of tested proteins ( $n=682$ ). For  
6 analyses of associations of protein and mRNA expression levels with plaque vulnerability, p-values  
7 were adjusted using the Benjamini-Hochberg method(20) with a false discovery rate threshold of 0.1.  
8 For differential expression testing using single-cell RNA sequencing data, we used a nominal p-value  
9 threshold of 0.05.

10

### 11 **Druggability of prioritised proteins**

12 We identified plasma proteins targeted by approved drugs (drugged proteins), as well as plasma  
13 proteins targeted by a developmental drug (druggable proteins) based on ChEMBL. For the drugged  
14 and druggable proteins, compound indications and side-effects were extracted from ChEMBL and the  
15 British National Formulary (BNF).

16

### 17 **Partial replication of protein associations with cardiac traits**

18 Due to the availability of eight proteomics GWAS, some studies measured the same proteins, which  
19 we used to replicate the associations with CHD risk. Replication was sought using a nominally  
20 significant p-value of 0.05 or smaller and considering the effect direction of the triangulated analysis.  
21 In addition, a more stringent p-value threshold of  $6.94 \times 10^{-4}$  (0.05 divided by the number of proteins  
22 that were available in more than one study) was used for more conservative replication.

23

## 24 **Results**

### 25 **Urinary metabolism breakdown products associating with CHD**

26 Using Mendelian randomisation, we evaluated 954 urinary metabolism breakdown products to  
27 investigate their association with CHD (**Appendix Table S1**). Out of these, 29 were associated with



1 CHD (**Appendix Table S3**, step 1 in **Figure 1, Figure 3**). Twelve of these metabolism breakdown  
2 products originated from amino acid metabolism. For example, higher values of N-acetytyrosine  
3 increased the risk of CHD (OR 1.03 per one standard deviation increase, 95%CI 1.02; 1.04), while  
4 higher values of 3-methylglutaconate decreased the risk of CHD (OR 0.95 per one standard deviation  
5 increase, 95%CI 0.93; 0.97). Other associated urinary metabolism breakdown products originate from  
6 various metabolic processes: one from energy metabolism, four from lipid metabolism, and three from  
7 xenobiotic metabolism. Finally, nine urinary metabolism breakdown products of unclassified  
8 metabolic origin were associated with CHD, representing potential novel metabolic pathways.

9

#### 10 **Triangulated proteins concordant with urinary metabolism breakdown products and CHD**

11 We used Mendelian randomisation to determine associations between plasma proteins and the 29  
12 urinary metabolism breakdown products identified as being associated with CHD. We identified 682  
13 proteins associating with at least one of these 29 urinary metabolism breakdown products (**Appendix**  
14 **Figure S2**). Among these, 113 proteins were identified with a directionally concordant association  
15 with CHD (steps 2-4 in **Figure 1, Appendix Figure S3, Appendix Table S4**). Please see,  
16 **Supplemental Results** and **Appendix Tables S5-S6** for MR analyses replicating 83% of these  
17 findings in smaller independent GWAS.

18

#### 19 **MR-prioritised proteins and plaque vulnerability index**

20 Of the 113 MR-prioritised proteins, 36 were available and detectable in the AE protein expression  
21 data measured in plaque samples (194 patients, **Appendix Table S2** for characteristics) and 10  
22 proteins were associated with increased plaque vulnerability: A2M, AKR7A2, APOF, C1S, CA1,  
23 COMT, CTSD, CTSH, NAGK, and PLTP (**Appendix Figure S4, Appendix Table S7**). Data on  
24 mRNA expression in plaques was available for 110 genes coding for MR-prioritised proteins (632  
25 patients, **Appendix Table S2** for characteristics), identifying an additional 6 proteins where mRNA  
26 expression was associated with plaque vulnerability: FER, IL6RA, ATF6B, AT1B2, SWP70, and  
27 SIG14 (**Figure 4, Appendix Figure S5, Appendix Table S8**).

28

## 1 **Single-cell RNA sequencing of MR-prioritised proteins**

2 To understand the cellular origin of the proteins, we additionally explored which genes encoding  
3 prioritised proteins were differentially expressed across cells in carotid plaques. For this, we used  
4 single-cell RNA sequencing data on over 4,900 cells available for 105 of our MR-prioritised proteins  
5 including all 16 proteins associated with plaque vulnerability. We identified 87 genes that showed to  
6 be enriched in one or more plaque cell types, which included 14 genes encoding for the proteins  
7 associated with plaque vulnerability (**Appendix Table S9, Figure 5**). For example, the gene *ATP1B2*,  
8 coding the protein AT1B2, was higher expressed in smooth muscle cells as compared to other plaque  
9 cell types (**Figure 5, Appendix Figure S6, Appendix Table S9**). The gene *ATF6B*, coding for the  
10 protein ATF6B, was higher expressed in T-cells as compared to other plaque cell types. Next, we used  
11 broader clusters of cell types, namely structural cells, innate immune cells, and adaptive immune  
12 cells. We found three (*IL6RA*, *SIG14*, and *NAGK*) out of 16 plaque vulnerability-associated proteins  
13 to be higher expressed in innate immune cells as compared to the other two clusters of cell types,  
14 while four (*AT1B2*, *C1S*, *A2MG*, and *ARK72*) were higher expressed in structural cells, five (*COMT*,  
15 *CATD*, *CATH*, *PLTP*, and *SWP70*) were higher expressed in both innate immune and structural cells  
16 and two (*FER* and *ATF6B*) were higher expressed in both innate immune cells and adaptive immune  
17 cells (**Figure 4**). Two proteins (*CAH1* and *APOF*) were ubiquitous (i.e., no higher expression in any  
18 of the three clusters).

19

## 20 **Druggability of proteins**

21 The set of 16 proteins associated with plaque vulnerability included eight drugged proteins, four  
22 druggable proteins, and four not yet druggable proteins (step 6 in **Figure 1, Figure 4-5**). These  
23 proteins predominantly associated with amino acid breakdown products and unclassified breakdown  
24 products. Six drugged proteins (*AT1B2*, *CAH1*, *CATD*, *COMT*, *FER*, and *IL6RA*) were targeted by a  
25 drug with a cardiac indication and/or side-effect (**Table 1, Figure 4, Appendix Tables S10-S11**). The  
26 protein *AT1B2* (targeted as part of a protein complex group through digoxin which is contraindicated  
27 in HCM) decreased the risk of CHD (OR 0.97, 95%CI 0.96; 0.99). Higher values of *CAH1* (targeted  
28 by methocarbamol which has registered side-effects including blood pressure disorders, and by

1 dopamine and aspirin indicated for myocardial infarction and coronary artery disease) increased the  
2 risk of CHD (OR 1.31, 95%CI 1.17; 1.48). Higher values of COMT (targeted by the drugs tolcapone  
3 and entacapone belonging to the class of COMT-inhibitors indicated for Parkinson's disease which  
4 have side-effects including chest pain and increased risk of CHD) decreased the risk of CHD (OR  
5 0.75, 95%CI 0.69; 0.81). Higher values of IL6RA, interleukin-6 receptor subunit alpha, decreased the  
6 risk of CHD (OR 0.97, 95%CI 0.97; 0.98). IL6RA is targeted by the interleukin-6 receptor antagonist  
7 tocilizumab indicated for auto-immune diseases, which is also in phase 2 development for CHD(21)  
8 and has registered side-effects of blood pressure disorders. Higher values of CATD (targeted by  
9 chloroquine indicated for malaria and cancer treatment, and in development for atrial fibrillation)  
10 decreased the risk of CHD (OR 0.96, 95%CI 0.94; 0.97). Higher values of FER (targeted by  
11 everolimus indicated in cancer and in development for acute coronary syndrome) decreased the risk of  
12 CHD (OR 0.77, 95%CI 0.72; 0.82).

13

14 KLKB1 and PGFRB are two proteins directionally concordantly associated with unclassified and  
15 amino acid metabolism breakdown products and CHD. Both are targeted by drugs with cardiac  
16 indications or side-effects, but neither is associated with plaque vulnerability. Higher values of  
17 KLKB1, targeted by the inhibitor lanadelumab (in phase 2 development for hypotension during  
18 hemodialysis), decreased the risk of CHD (OR 0.97, 95%CI 0.95; 0.98). Similarly, higher values of  
19 PGFRB, targeted by the inhibiting cancer drugs regorafenib and dasatinib (which have known side-  
20 effects including CHD, QT interval prolongation and blood pressure disorders), decreased the risk of  
21 CHD (OR 0.98, 95%CI 0.98; 0.99).

22

## 23 **Discussion**

24 In this study, we used Mendelian randomisation to assess relationships between 954 urinary  
25 metabolism breakdown products, 1,562 proteins and CHD, and we linked protein and mRNA  
26 expression levels in carotid plaque to plaque vulnerability. This was done to identify potential  
27 biomarkers of altered metabolism informative of CHD risk and plasma proteins which may be used as

1 targets for intervention in these metabolic pathways. Through our multi-layered approach we were  
2 able to prioritise 16 proteins associated with plaque vulnerability and CHD risk, where 83% of these  
3 CHD associations were independently replicated. The majority of these proteins were associated with  
4 urinary metabolism breakdown products from amino acid metabolism or unclassified origins.

5  
6 Altered amino acid metabolism in cardiovascular disease has been observed before(38-41). For  
7 example, the kynurenine pathway has been linked to the pathogenesis of atherosclerosis by  
8 modulating inflammation, oxidative stress, and endothelial function(42). Its breakdown products were  
9 identified as biomarkers of major cardiovascular events(43, 44), with previous studies showing the  
10 therapeutic potential of targeting this pathway to target vascular inflammation and plaque  
11 formation(45). We were able to link N-acetylkynurenine to the druggable proteins SIG14 and A2MG.  
12 SIG14 is member of the sialic acid-binding immunoglobulin-type lectins (Siglecs) family, which is  
13 involved in immune regulation and has been studied as drug target for various diseases including  
14 asthma, cancer, and autoimmune diseases(46). A2MG is a broad-spectrum protease inhibitor with an  
15 anti-inflammatory role and it has a dual role in coagulation(47, 48), which we now link to plaque  
16 vulnerability.

17  
18 The 16 prioritised proteins included two proteins with a known indication or side-effect on CHD:  
19 CAH1 and COMT. CAH1, is targeted by aspirin which is prescribed to reduce the risk of blood  
20 clots(22) in people at a high risk of (subsequent) cardiovascular events. COMT is affected by various  
21 drugs, including the Parkinson's disease drugs entacapone and tolcapone which is known to increase  
22 the risk of myocardial infarction through effects on blood pressure, arteriosclerosis, and aortic  
23 stenosis(23).

24  
25 These positive control findings confirmed our analysis was able to identify known drug targets effect  
26 for CHD. Additionally, we identified four drugged proteins (AT1B2, CATD, FER, and IL6RA) which  
27 are in development for CHD or CHD-related diseases. IL6R inhibiting drugs (e.g., tocilizumab) are  
28 currently being repurposed for treatment of CHD(8, 11), where the phase 3 CANTOS trial evaluating

1 canakinumab, inhibiting IL-1 which is upstream of IL6 and IL6R, showed clinical effectiveness in  
2 preventing CHD(24). This provides additional support for our approach as we rediscovered known  
3 drug target effects for CHD.

4

5 The drugged proteins C1S and CATH were identified as potential repurposing candidates for  
6 treatment of CHD and atherosclerosis. CATH is targeted by Bortezomib which is indicated for  
7 multiple myeloma, and due to side-effects such as thrombocytopenia, and neutropenia, not directly  
8 relevant to consider for CHD management. C1S on the other hand is targeted by inhibiting drugs (e.g.,  
9 berotralstat and human c1-esterase inhibitor) indicated for conditions such as angioedema. C1S  
10 activates the classical pathway of the complement system, contributing to inflammation and  
11 endothelial dysfunction, which promotes atherosclerosis(25, 26). The here identified effects of higher  
12 values of C1S on increased CHD risk, and association with increased plaque vulnerability, along with  
13 higher expression in smooth muscle cells and endothelial cells in plaque, supports drug repurposing  
14 for management of CHD and atherosclerosis.

15

16 We additionally identified four druggable proteins (A2MG, NAGK, PLTP, and SIG14) with  
17 associations with CHD and plaque vulnerability. NAGK is involved in amino sugar metabolism and  
18 plays a role in atherosclerosis through endothelial cell activation and inflammatory responses.  
19 Glucosamine, a substrate of NAGK, has been shown to suppress endothelial cell activation via O-  
20 GlcNAc modification, thereby potentially reducing atherosclerotic inflammation(27). By identifying  
21 an association of higher values of NAGK with a higher CHD risk and plaque vulnerability, the current  
22 study provides further support that inhibiting NAGK may be a viable strategy to treat CHD and  
23 atherosclerosis.

24

25 Four proteins that are not yet druggable (ARK72, APOF, ATF6B, SWP70) were identified as  
26 associated with CHD and plaque vulnerability. Higher values of APOF (apolipoprotein F) were  
27 associated with an increased CHD risk and its expression was higher in plaques with higher  
28 vulnerability. Previous studies have described APOF as a potential target for novel therapeutic

1 approaches(33). APOF inhibits CETP activity and reduces the transfer of cholesteryl esters from high  
2 density lipoprotein to low density lipoprotein (LDL), which is known to reduce LDL cholesterol  
3 plasma concentrations and atherosclerosis risk(34, 35). We have previously shown that despite many  
4 failed attempts to develop sufficiently potent and selective compounds targeting CETP, the drug  
5 target itself remains a viable target for preventing CHD supporting the ongoing phase 3 study of  
6 obicetrapib(10, 36, 37).

7  
8 Despite the robustness of our analyses, our study has a number of potential limitations. Firstly, the  
9 GWAS data we sourced used various high throughput assays, varying in accuracy and analytic  
10 scope(51). These assays do not directly measure concentrations of proteins or metabolism breakdown  
11 products, but evaluate relative values. Therefore, the magnitude of associations is unlikely to be a  
12 strong indicator of the effects of pharmacological perturbation of the target, and smaller MR effects  
13 may translate towards large drug effects and vice versa. Instead, these results predominantly provide  
14 information on the drug mechanisms and whether targets should be inhibited or activated. In the  
15 current study, urinary metabolism breakdown products were used as a proxy to identify more distal  
16 metabolic processes occurring in, or impacting, atherosclerotic relevant tissues. Hence, the  
17 associations between urinary metabolism breakdown products and CHD do not reflect a direct causal  
18 mechanism but rather point to the upstream metabolism origins of these breakdown products, which  
19 may play an important role in the development of atherosclerosis and the risk of CHD. For some of  
20 the MR-prioritised proteins linked with plaque vulnerability, we found an opposite direction of the  
21 association of the protein expression in plaque with the plaque vulnerability as compared to the effect  
22 of the circulating protein on CHD risk. These findings underscore the complexity of tissue-specific  
23 expression, and the molecular pathways involved in atherogenesis. Protein expression patterns can  
24 vary by tissue type and stage of atherosclerosis(52, 53), with certain proteins possibly being present  
25 due to cell death rather than active secretion, which may affect the interpretation of their direction of  
26 effect. Furthermore, MR-prioritised plasma proteins may be involved in CHD through mechanisms in  
27 other tissues, such as arterial constriction, without accumulating in plaques. As a result, these proteins  
28 may not be represented in our AE findings. However, since plaques do not synthesize proteins but

1 contain proteins originating from plasma, the plaque proteins we identified are likely linked to the  
2 plasma proteins prioritised by MR as being related to CHD. Both the GWAS and the AE Biobank  
3 data we used predominantly included European individuals, which may limit the generalizability of  
4 our findings to non-European populations. In addition, the observational nature of the AE and its  
5 high-risk study population introduces a risk of bias due to for example confounding, reverse  
6 causation, and index event bias. However, combining these analyses with MR, which is largely  
7 confounding-resistant, meaningfully enhances the robustness of our findings.

8  
9 In conclusion, we identified urinary biomarkers of altered metabolism, predominantly from amino  
10 acid metabolism and unclassified origins, associating with CHD risk. By integrating these findings  
11 with plasma proteins and plaque vulnerability, we identified 16 plasma proteins as potential drug  
12 targets for CHD, supporting the development of novel and repurposed therapeutic strategies.  
13 Additionally, our study highlights the value of integrating multi-modal evidence to uncover  
14 information potentially relevant for disease diagnosis, prognosis, and etiology.

1 **Funding:** SD and SAEP are supported by a VIDI Fellowship (project number 09150172010050) from  
2 the Dutch Organisation for Health Research and Development (ZonMW) awarded to SAEP. AFS is  
3 supported by BHF grant PG/22/10989, the UCL BHF Research Accelerator AA/18/6/34223,  
4 MR/V033867/1, the National Institute for Health and Care Research University College London  
5 Hospitals Biomedical Research Centre, and the EU Horizon scheme (AI4HF 101080430 and  
6 DataTools4Heart 101057849). MV is supported by a postdoc talent grant from the Amsterdam  
7 Cardiovascular Sciences. This work was funded by UK Research and Innovation (UKRI) under the  
8 UK government’s Horizon Europe funding guarantee EP/Z000211/1. This publication is part of the  
9 project “Computational medicine for cardiac disease” with file number 2023.022 of the research  
10 programme “Computing Time on National Computer Facilities” which is (partly) financed by the  
11 Dutch Research Council (NWO).

12  
13 **Role of funders:** The funding source did not influence the study design, the collection, analysis, and  
14 interpretation of data, the writing of the report, or the decision to submit the manuscript for  
15 publication.

16  
17 **Acknowledgements:** This research has been conducted using the UK Biobank Resource under  
18 Application Number 12113. The authors are grateful to UK Biobank participants.

19  
20 **Declaration of competing interests:** AFS and CF have received funding from New Amsterdam  
21 Pharma for unrelated projects. The other authors declare that they have no conflict of interest.

22  
23 **Author contributions:** SD, EDB, AFS, and SP designed the study. SD and MV accessed the data,  
24 verified the data, and performed analyses. SD drafted the manuscript. All authors provided critical  
25 input on the analysis, as well as the drafted manuscript.

26  
27



1 **Code availability:** Analyses were conducted using Python v3.7.13 (for GNU Linux), Pandas v1.3.5,  
2 Numpy v1.21.6, bio-misc v0.1.4, and matplotlib v3.4.3. Single-cell RNA sequencing data were  
3 analysed in R using Seurat v3(59).  
4  
5 **Data availability:** The GWAS data on CHD used in this study can be accessed via  
6 <https://www.ebi.ac.uk/gwas/studies/GCST90132314> (n=181,522 CHD cases, n=984,168 controls).  
7 The GWAS data on urinary metabolism breakdown product values can be accessed via  
8 <https://www.ebi.ac.uk/gwas/publications/31959995> (n=1,627). The individual GWAS data on plasma  
9 protein values can be accessed as follows: deCODE (n=35,559,  
10 <https://www.decode.com/summarydata/>), SCALLOP (n=30,931,  
11 <https://www.ebi.ac.uk/gwas/publications/33067605>), Ahola-Olli *et al.* (n=8,293,  
12 <https://www.ebi.ac.uk/gwas/publications/27989323>), Framingham (n=6,861,  
13 <https://www.ebi.ac.uk/gwas/publications/21909115>), AGES-Reykjavik (n=5,368,  
14 <https://www.ebi.ac.uk/gwas/publications/35078996>), INTERVAL (n=3,301,  
15 <https://www.ebi.ac.uk/gwas/publications/29875488>), Gilly *et al.* (n=1,328,  
16 <https://www.ebi.ac.uk/gwas/publications/33303764>), and Yang *et al.* (n=636,  
17 <https://www.ebi.ac.uk/gwas/publications/34239129>).

## 1    **References**

2

- 3    1.    Chiorescu RM, Mocan M, Inceu AI, Buda AP, Blendea D, Vlaicu SI. Vulnerable  
4    atherosclerotic plaque: is there a molecular signature? *International Journal of Molecular*  
5    *Sciences*. 2022;23(21):13638.
- 6    2.    Bentzon JF, Otsuka F, Virmani R, Falk E. Mechanisms of plaque formation and  
7    rupture. *Circulation research*. 2014;114(12):1852-66.
- 8    3.    Aboonabi A, Meyer RR, Singh I. The association between metabolic syndrome  
9    components and the development of atherosclerosis. *Journal of human hypertension*.  
10    2019;33(12):844-55.
- 11    4.    Hotamisligil GS. Inflammation and metabolic disorders. *Nature*. 2006;444(7121):860-  
12    7.
- 13    5.    Mokry M, Boltjes A, Slenders L, Bel-Bordes G, Cui K, Brouwer E, et al.  
14    Transcriptomic-based clustering of human atherosclerotic plaques identifies subgroups with  
15    different underlying biology and clinical presentation. *Nature cardiovascular research*.  
16    2022;1(12):1140-55.
- 17    6.    Hingorani A, Humphries S. Nature's randomised trials. *The Lancet*.  
18    2005;366(9501):1906-8.
- 19    7.    de Ruiter SC, Schmidt AF, Grobbee DE, den Ruijter HM, Peters SA. Sex-specific  
20    Mendelian randomisation to assess the causality of sex differences in the effects of risk  
21    factors and treatment: Spotlight on hypertension. *Journal of Human Hypertension*. 2023:1-7.
- 22    8.    Cupido AJ, Asselbergs FW, Natarajan P, Group CIW, Ridker PM, Hovingh GK, et al.  
23    Dissecting the IL-6 pathway in cardiometabolic disease: A Mendelian randomization study  
24    on both IL6 and IL6R. *British Journal of Clinical Pharmacology*. 2022;88(6):2875-84.
- 25    9.    Zheng J, Haberland V, Baird D, Walker V, Haycock PC, Hurle MR, et al. Phenome-  
26    wide Mendelian randomization mapping the influence of the plasma proteome on complex  
27    diseases. *Nature genetics*. 2020;52(10):1122-31.
- 28    10.    Schmidt AF, Hunt NB, Gordillo-Marañón M, Charoen P, Drenos F, Kivimaki M, et  
29    al. Cholesteryl ester transfer protein (CETP) as a drug target for cardiovascular disease.  
30    *Nature communications*. 2021;12(1):5640.
- 31    11.    Consortium I-RMRA. The interleukin-6 receptor as a target for prevention of  
32    coronary heart disease: a mendelian randomisation analysis. *The Lancet*.  
33    2012;379(9822):1214-24.
- 34    12.    Schlosser P, Li Y, Sekula P, Raffler J, Grundner-Culemann F, Pietzner M, et al.  
35    Genetic studies of urinary metabolites illuminate mechanisms of detoxification and excretion  
36    in humans. *Nature genetics*. 2020;52(2):167-76.
- 37    13.    Aragam KG, Jiang T, Goel A, Kanoni S, Wolford BN, Atri DS, et al. Discovery and  
38    systematic characterization of risk variants and genes for coronary artery disease in over a  
39    million participants. *Nature Genetics*. 2022:1-13.
- 40    14.    Sudlow C, Gallacher J, Allen N, Beral V, Burton P, Danesh J, et al. UK biobank: an  
41    open access resource for identifying the causes of a wide range of complex diseases of  
42    middle and old age. *PLoS medicine*. 2015;12(3):e1001779.
- 43    15.    Bowden J, Davey Smith G, Burgess S. Mendelian randomization with invalid  
44    instruments: effect estimation and bias detection through Egger regression. *International*  
45    *journal of epidemiology*. 2015;44(2):512-25.
- 46    16.    Burgess S, Zuber V, Valdes-Marquez E, Sun BB, Hopewell JC. Mendelian  
47    randomization with fine-mapped genetic data: choosing from large numbers of correlated  
48    instrumental variables. *Genetic epidemiology*. 2017;41(8):714-25.

- 1 17. Rücker G, Schwarzer G, Carpenter JR, Binder H, Schumacher M. Treatment-effect  
2 estimates adjusted for small-study effects via a limit meta-analysis. *Biostatistics*.  
3 2011;12(1):122-42.
- 4 18. Schmidt AF, Finan C, Gordillo-Marañón M, Asselbergs FW, Freitag DF, Patel RS, et  
5 al. Genetic drug target validation using Mendelian randomisation. *Nature communications*.  
6 2020;11(1):1-12.
- 7 19. Consortium U. The universal protein resource (UniProt) in 2010. *Nucleic acids*  
8 *research*. 2010;38(suppl\_1):D142-D8.
- 9 20. Benjamini Y, Hochberg Y. Controlling the false discovery rate: a practical and  
10 powerful approach to multiple testing. *Journal of the Royal statistical society: series B*  
11 *(Methodological)*. 1995;57(1):289-300.
- 12 21. Kleveland O, Kunszt G, Bratlie M, Ueland T, Broch K, Holte E, et al. Effect of a  
13 single dose of the interleukin-6 receptor antagonist tocilizumab on inflammation and troponin  
14 T release in patients with non-ST-elevation myocardial infarction: a double-blind,  
15 randomized, placebo-controlled phase 2 trial. *European heart journal*. 2016;37(30):2406-13.
- 16 22. Undas A, Brummel-Ziedins KE, Mann KG. Antithrombotic properties of aspirin and  
17 resistance to aspirin: beyond strictly antiplatelet actions. *Blood*. 2007;109(6):2285-92.
- 18 23. Stocchi F, Rascol O, Kiebertz K, Poewe W, Jankovic J, Tolosa E, et al. Initiating  
19 levodopa/carbidopa therapy with and without entacapone in early Parkinson disease: the  
20 STRIDE- $\square$ PD study. *Annals of neurology*. 2010;68(1):18-27.
- 21 24. Ridker PM, Devalaraja M, Baeres FM, Engelmann MD, Hovingh GK, Ivkovic M, et  
22 al. IL-6 inhibition with ziltivekimab in patients at high atherosclerotic risk (RESCUE): a  
23 double-blind, randomised, placebo-controlled, phase 2 trial. *The Lancet*.  
24 2021;397(10289):2060-9.
- 25 25. Gaboriaud C, Thielens NM, Gregory LA, Rossi V, Fontecilla-Camps JC, Arlaud GJ.  
26 Structure and activation of the C1 complex of complement: unraveling the puzzle. *Trends in*  
27 *immunology*. 2004;25(7):368-73.
- 28 26. Haskard DO, Boyle JJ, Mason JC. The role of complement in atherosclerosis. *Current*  
29 *opinion in lipidology*. 2008;19(5):478-82.
- 30 27. Ju Y, Hua J, Sakamoto K, Ogawa H, Nagaoka I. Modulation of TNF- $\alpha$ -induced  
31 endothelial cell activation by glucosamine, a naturally occurring amino monosaccharide.  
32 *International journal of molecular medicine*. 2008;22(6):809-15.
- 33 28. Champattanachai V, Marchase RB, Chatham JC. Glucosamine protects neonatal  
34 cardiomyocytes from ischemia-reperfusion injury via increased protein-associated O-  
35 GlcNAc. *American Journal of Physiology-Cell Physiology*. 2007;292(1):C178-C87.
- 36 29. van Haperen R, van Tol A, van Gent T, Scheek L, Visser P, van der Kamp A, et al.  
37 Increased risk of atherosclerosis by elevated plasma levels of phospholipid transfer protein.  
38 *Journal of Biological Chemistry*. 2002;277(50):48938-43.
- 39 30. Vergeer M, Dallinga-Thie GM, Dullaart RP, Van Tol A. Evaluation of phospholipid  
40 transfer protein as a therapeutic target. *Clinical Lipidology and Metabolic Disorders*.  
41 2008;3(3):327.
- 42 31. Gill D, Georgakis MK, Koskeridis F, Jiang L, Feng Q, Wei W-Q, et al. Use of genetic  
43 variants related to antihypertensive drugs to inform on efficacy and side effects. *Circulation*.  
44 2019;140(4):270-9.
- 45 32. Jiang X-C, Jin W, Hussain MM. The impact of phospholipid transfer protein (PLTP)  
46 on lipoprotein metabolism. *Nutrition & metabolism*. 2012;9:1-7.
- 47 33. Lagor WR, Fields DW, Bauer RC, Crawford A, Abt MC, Artis D, et al. Genetic  
48 manipulation of the ApoF/Stat2 locus supports an important role for type I interferon  
49 signaling in atherosclerosis. *Atherosclerosis*. 2014;233(1):234-41.

- 1 34. Wang X, Driscoll DM, Morton RE. Molecular cloning and expression of lipid transfer  
2 inhibitor protein reveals its identity with apolipoprotein F. *Journal of Biological Chemistry*.  
3 1999;274(3):1814-20.
- 4 35. Liu Y, Morton RE. Apolipoprotein F: a natural inhibitor of cholesteryl ester transfer  
5 protein and a key regulator of lipoprotein metabolism. *Current opinion in lipidology*.  
6 2020;31(4):194-9.
- 7 36. Cupido AJ, Reeskamp LF, Hingorani AD, Finan C, Asselbergs FW, Hovingh GK, et  
8 al. Joint genetic inhibition of PCSK9 and CETP and the association with coronary artery  
9 disease: a factorial mendelian randomization study. *JAMA cardiology*. 2022;7(9):955-64.
- 10 37. Dunca D, Chopade S, Gordillo-Marañón M, Hingorani AD, Kuchenbaecker K, Finan  
11 C, et al. Comparing the effects of CETP in East Asian and European ancestries: a Mendelian  
12 randomization study. *Nature Communications*. 2024;15(1):5302.
- 13 38. Wang Y, Sun W, Zheng J, Xu C, Wang X, Li T, et al. Urinary metabonomic study of  
14 patients with acute coronary syndrome using UPLC-QTOF/MS. *Journal of Chromatography*  
15 *B*. 2018;1100:122-30.
- 16 39. Huang M, Zhao H, Gao S, Liu Y, Liu Y, Zhang T, et al. Identification of coronary  
17 heart disease biomarkers with different severities of coronary stenosis in human urine using  
18 non-targeted metabolomics based on UPLC-Q-TOF/MS. *Clinica Chimica Acta*. 2019;497:95-  
19 103.
- 20 40. Wei D, Melgarejo JD, Van Aelst L, Vanassche T, Verhamme P, Janssens S, et al.  
21 Prediction of coronary artery disease using urinary proteomics. *European Journal of*  
22 *Preventive Cardiology*. 2023;30(14):1537-46.
- 23 41. van Vugt M, Finan C, Chopade S, Providencia R, Bezzina CR, Asselbergs FW, et al.  
24 Integrating metabolomics and proteomics to identify novel drug targets for heart failure and  
25 atrial fibrillation. *Genome Medicine*. 2024;16(1):120.
- 26 42. Baumgartner R, Forteza MJ, Ketelhuth DF. The interplay between cytokines and the  
27 Kynurenine pathway in inflammation and atherosclerosis. *Cytokine*. 2019;122:154148.
- 28 43. Gáspár R, Halmi D, Demján V, Berkecz R, Pipicz M, Csont T. Kynurenine pathway  
29 metabolites as potential clinical biomarkers in coronary artery disease. *Frontiers in*  
30 *Immunology*. 2022;12:768560.
- 31 44. Pedersen ER, Svingen GFT, Schartum-Hansen H, Ueland PM, Ebbing M, Nordrehaug  
32 JE, et al. Urinary excretion of kynurenine and tryptophan, cardiovascular events, and  
33 mortality after elective coronary angiography. *European heart journal*. 2013;34(34):2689-96.
- 34 45. Polyzos KA, Ovchinnikova O, Berg M, Baumgartner R, Agardh H, Pirault J, et al.  
35 Inhibition of indoleamine 2, 3-dioxygenase promotes vascular inflammation and increases  
36 atherosclerosis in Apoe<sup>-/-</sup> mice. *Cardiovascular research*. 2015;106(2):295-302.
- 37 46. Duan S, Paulson JC. Siglecs as immune cell checkpoints in disease. *Annual review of*  
38 *immunology*. 2020;38(1):365-95.
- 39 47. Beheiri A, Langer C, Düring C, Krümpel A, Thedieck S, Nowak Götzl U. Role of  
40 elevated  $\alpha 2$  macroglobulin revisited: results of a case-control study in children with  
41 symptomatic thromboembolism. *Journal of Thrombosis and Haemostasis*. 2007;5(6):1179-84.
- 42 48. Schulz S, Birkenmeier G, Schagdarsurengin U, Wenzel K, Müller-Werdan U, Rehfeld  
43 D, et al. Role of LDL receptor-related protein (LRP) in coronary atherosclerosis.  
44 *International journal of cardiology*. 2003;92(2-3):137-44.
- 45 49. Li J, Liu X, Xiang Y, Ding X, Wang T, Liu Y, et al. Alpha-2-macroglobulin and  
46 heparin cofactor II and the vulnerability of carotid atherosclerotic plaques: an iTRAQ-based  
47 analysis. *Biochemical and biophysical research communications*. 2017;483(3):964-71.
- 48 50. de Laat-Kremers R, Costanzo S, Yan Q, Di Castelnuovo A, De Curtis A, Cerletti C, et  
49 al. High alpha-2-macroglobulin levels are a risk factor for cardiovascular disease events: A  
50 Moli-sani cohort study. *Thrombosis Research*. 2024;234:94-100.

- 1 51. Katz DH, Robbins JM, Deng S, Tahir UA, Bick AG, Pampana A, et al. Proteomic  
2 profiling platforms head to head: leveraging genetics and clinical traits to compare aptamer-  
3 and antibody-based methods. *Science Advances*. 2022;8(33):eabm5164.
- 4 52. Jeong H, Mason SP, Barabási A-L, Oltvai ZN. Lethality and centrality in protein  
5 networks. *Nature*. 2001;411(6833):41-2.
- 6 53. Larionov S, Dedek O, Birkenmeier G, Thal DR. Expression of  $\alpha$  2-macroglobulin,  
7 neutrophil elastase, and interleukin-1 $\alpha$  differs in early-stage and late-stage atherosclerotic  
8 lesions in the arteries of the circle of Willis. *Acta Neuropathologica*. 2007;113:33-43.
- 9 54. Fernández-Alvarez V, Linares Sánchez M, López Alvarez F, Suárez Nieto C, Mäkitie  
10 AA, Olsen KD, et al. Evaluation of intima-media thickness and arterial stiffness as early  
11 ultrasound biomarkers of carotid artery atherosclerosis. *Cardiology and therapy*.  
12 2022;11(2):231-47.
- 13 55. Bots M. Carotid artery intima-media thickness as an indicator of generalized  
14 atherosclerosis. *Journal of Internal Medicine*. 1994;236:567-73.
- 15 56. Bots ML, Breteler M, Hofman A, Grobbee DE, Van Swieten J, van Gijn J, et al.  
16 Cerebral white matter lesions and atherosclerosis in the Rotterdam Study. *The Lancet*.  
17 1993;341(8855):1232-7.
- 18 57. Bots ML, Hoes AW, Koudstaal PJ, Hofman A, Grobbee DE. Common carotid intima-  
19 media thickness and risk of stroke and myocardial infarction: the Rotterdam Study.  
20 *Circulation*. 1997;96(5):1432-7.
- 21 58. Crouse JR, Raichlen JS, Riley WA, Evans GW, Palmer MK, O'Leary DH, et al.  
22 Effect of rosuvastatin on progression of carotid intima-media thickness in low-risk  
23 individuals with subclinical atherosclerosis: the METEOR Trial. *Jama*. 2007;297(12):1344-  
24 53.
- 25 59. Hao Y, Hao S, Andersen-Nissen E, Mauck WM, Zheng S, Butler A, et al. Integrated  
26 analysis of multimodal single-cell data. *Cell*. 2021;184(13):3573-87. e29.
- 27 60. Hashimshony T, Senderovich N, Avital G, Klochendler A, De Leeuw Y, Anavy L, et  
28 al. CEL-Seq2: sensitive highly-multiplexed single-cell RNA-Seq. *Genome biology*.  
29 2016;17:1-7.
- 30 61. Muraro MJ, Dharmadhikari G, Grün D, Groen N, Dielen T, Jansen E, et al. A single-  
31 cell transcriptome atlas of the human pancreas. *Cell systems*. 2016;3(4):385-94. e3.
- 32 62. Depuydt MA, Prange KH, Slenders L, Örd T, Elbersen D, Boltjes A, et al.  
33 Microanatomy of the human atherosclerotic plaque by single-cell transcriptomics. *Circulation*  
34 *research*. 2020;127(11):1437-55.
- 35 63. Ferkingstad E, Sulem P, Atlason BA, Sveinbjornsson G, Magnusson MI,  
36 Styrismisdottir EL, et al. Large-scale integration of the plasma proteome with genetics and  
37 disease. *Nature Genetics*. 2021;53(12):1712-21.
- 38 64. Folkersen L, Gustafsson S, Wang Q, Hansen DH, Hedman ÅK, Schork A, et al.  
39 Genomic and drug target evaluation of 90 cardiovascular proteins in 30,931 individuals.  
40 *Nature metabolism*. 2020;2(10):1135-48.
- 41 65. Ahola-Olli AV, Würtz P, Havulinna AS, Aalto K, Pitkänen N, Lehtimäki T, et al.  
42 Genome-wide association study identifies 27 loci influencing concentrations of circulating  
43 cytokines and growth factors. *The American Journal of Human Genetics*. 2017;100(1):40-50.
- 44 66. Ehret GB, Munroe PB, Rice KM, Bochud M, Johnson AD, Chasman DI, et al.  
45 Genetic variants in novel pathways influence blood pressure and cardiovascular disease risk.  
46 *Nature*. 2011;478(7367):103-9.
- 47 67. Gudjonsson A, Gudmundsdottir V, Axelsson GT, Gudmundsson EF, Jonsson BG,  
48 Launer LJ, et al. A genome-wide association study of serum proteins reveals shared loci with  
49 common diseases. *Nature communications*. 2022;13(1):1-13.

- 1 68. Sun BB, Maranville JC, Peters JE, Stacey D, Staley JR, Blackshaw J, et al. Genomic  
2 atlas of the human plasma proteome. *Nature*. 2018;558(7708):73-9.
- 3 69. Gilly A, Park Y-C, Png G, Barysenka A, Fischer I, Bjørnland T, et al. Whole-genome  
4 sequencing analysis of the cardiometabolic proteome. *Nature communications*. 2020;11(1):1-  
5 9.
- 6 70. Locke AE, Kahali B, Berndt SI, Justice AE, Pers TH, Day FR, et al. Genetic studies  
7 of body mass index yield new insights for obesity biology. *Nature*. 2015;518(7538):197-206.
- 8 71. Verhoeven BA, Velema E, Schoneveld AH, Vries JPPd, de Bruin P, Seldenrijk CA, et  
9 al. Athero-express: differential atherosclerotic plaque expression of mRNA and protein in  
10 relation to cardiovascular events and patient characteristics. *Rationale and design. European*  
11 *journal of epidemiology*. 2004;19:1127-33.
- 12 72. van Koevorden ID, de Bakker M, Haitjema S, van der Laan SW, de Vries J-PP,  
13 Hoefler IE, et al. Testosterone to oestradiol ratio reflects systemic and plaque inflammation  
14 and predicts future cardiovascular events in men with severe atherosclerosis. *Cardiovascular*  
15 *Research*. 2019;115(2):453-62.
- 16 73. de Bakker M, Timmerman N, van Koevorden ID, de Kleijn DP, de Borst GJ,  
17 Pasterkamp G, et al. The age-and sex-specific composition of atherosclerotic plaques in  
18 vascular surgery patients. *Atherosclerosis*. 2020;310:1-10.
- 19 74. Hellings WE, Pasterkamp G, Vollebregt A, Seldenrijk CA, De Vries J-PP, Velema E,  
20 et al. Intraobserver and interobserver variability and spatial differences in histologic  
21 examination of carotid endarterectomy specimens. *Journal of vascular surgery*.  
22 2007;46(6):1147-54.
- 23 75. Verhoeven B, Hellings WE, Moll FL, de Vries JP, de Kleijn DP, de Bruin P, et al.  
24 Carotid atherosclerotic plaques in patients with transient ischemic attacks and stroke have  
25 unstable characteristics compared with plaques in asymptomatic and amaurosis fugax  
26 patients. *Journal of vascular surgery*. 2005;42(6):1075-81.
- 27 76. Timmerman N, de Kleijn DP, de Borst GJ, den Ruijter HM, Asselbergs FW,  
28 Pasterkamp G, et al. Family history and polygenic risk of cardiovascular disease: Independent  
29 factors associated with secondary cardiovascular events in patients undergoing carotid  
30 endarterectomy. *Atherosclerosis*. 2020;307:121-9.
- 31 77. Davies MJ, Richardson PD, Woolf N, Katz DR, Mann J. Risk of thrombosis in human  
32 atherosclerotic plaques: role of extracellular lipid, macrophage, and smooth muscle cell  
33 content. *Heart*. 1993;69(5):377-81.
- 34 78. Schwartz SM, Galis ZS, Rosenfeld ME, Falk E. Plaque rupture in humans and mice.  
35 *Arteriosclerosis, thrombosis, and vascular biology*. 2007;27(4):705-13.
- 36 79. Lusby RJ, Ferrell LD, Ehrenfeld W, Stoney R, Wylie EJ. Carotid plaque hemorrhage:  
37 its role in production of cerebral ischemia. *Archives of Surgery*. 1982;117(11):1479-88.
- 38 80. Finan C, Gaulton A, Kruger FA, Lumbers RT, Shah T, Engmann J, et al. The  
39 druggable genome and support for target identification and validation in drug development.  
40 *Science translational medicine*. 2017;9(383):eaag1166.
- 41

## Tables

**Table 1.** Positive controls and repurposing opportunities: prioritised proteins related to plaque vulnerability

PROTEIN	GENE NAME	UNIPROT ID	DRUGGABILITY	METABOLISM CLASS(ES)	NO. BREAKDOWN PRODUCTS	BREAKDOWN PRODUCT – CHD EFFECT	PROTEIN-CHD EFFECT	NO. DRUGS	CARDIAC INDICATION AND/OR SIDE-EFFECT	CARDIAC INDICATION(S)	CARDIAC SIDE-EFFECT(S)	DRUG NAME(S)
A2MG	A2M	P01023	Druggable	Unclassified, Amino Acid	2	Increasing	0.89 (0.85; 0.94)	1	No	-	-	Technetium tc 99m succimer
ARK72	AKR7A2	O43488	Not yet druggable	Unclassified	1	Decreasing	1.11 (1.08; 1.15)	0	No	-	-	
APOF	APOF	Q13790	Not yet druggable	Energy	1	Increasing	1.12 (1.07; 1.17)	0	No	-	-	
ATF6B	ATF6B	Q99941	Not yet druggable	Amino Acid	3	Mixed	1.10 (1.08; 1.13)	0	No	-	-	
AT1B2	AT1B2	P14415	Drugged	Lipid	1	Increasing	0.97 (0.96; 0.99)	9	Yes	See Appendix Table S10	Arrhythmias, Cardiac conduction disorder	Deslanoside, Digoxin, Digoxin, Acetyldigitoxin, Istaroxime, Lanatoside c, Omeprazole, Lansoprazole, Pantoprazole
C1S	C1S	P09871	Drugged	Unclassified	1	Decreasing	1.09 (1.07; 1.11)	3	No	-	-	Human c1-esterase inhibitor, Sutinlimab, Berotralstat
CAH1	CA1	P00915	Drugged	Unclassified, Amino Acid	3	Mixed	1.31 (1.17; 1.48)	7	Yes	See Appendix Table S10	Arrhythmias, Blood pressure disorders	See Appendix Table S10
COMT	COMT	P21964	Drugged	Xenobiotics, Unclassified	2	Decreasing	0.75 (0.69; 0.81)	5	Yes	Cardiovascular diseases, Myocardial infarction, Heart failure	Myocardial infarction, Chest pain	Tolcapone, Entacapone, Opicapone, Nebicapone, Dopamine
CATD	CTSD	P07339	Drugged	Amino Acid	1	Increasing	0.96 (0.94; 0.97)	4	Yes	Atrial fibrillation	-	Chloroquine, Amprenavir, Bortezomib, Tipranavir
CATH	CTSH	P09668	Drugged	Xenobiotics	1	Decreasing	0.96 (0.95; 0.97)	1	No	-	-	Bortezomib
FER	FER	P16591	Drugged	Unclassified	1	Increasing	0.77 (0.72; 0.82)	6	Yes	See Appendix Table S10	-	See Appendix Table S10
IL6RA	IL6R	P08887	Drugged	Amino Acid	1	Increasing	0.97 (0.97; 0.98)	7	Yes	Non-st elevated myocardial infarction, Pulmonary hypertension	Blood pressure disorders	Tocilizumab, Sarilumab, Vobasilizumab, Levilimab, Satralizumab, Raloxifene, Bazedoxifene
NAGK	NAGK	Q9UJ70	Druggable	Unclassified	1	Increasing	1.06 (1.04; 1.09)	0	No	-	-	
PLTP	PLTP	P55058	Druggable	Unclassified, Xenobiotics	2	Mixed	1.03 (1.01; 1.04)	0	No	-	-	
SIG14	SIGLEC14	Q08ET2	Druggable	Unclassified, Amino Acid	2	Mixed	0.99 (0.98; 0.99)	0	No	-	-	
SWP70	SWAP70	Q9UH65	Not yet druggable	Amino Acid, Lipid	2	Mixed	0.94 (0.93; 0.96)	0	No	-	-	

N.B. Each row presents a plasma protein associated with coronary heart disease (CHD), and at least one urinary metabolism breakdown product, where the associations between protein, breakdown product(s), and CHD are directionally concordant (**Figure 2**). Associations were identified by Mendelian randomisation.

Protein names are based on their Uniprot synonym. Druggability indicates whether the proteins are targeted by a developmental compound in clinical trials (druggable) or by a compound which has received marketing approval (drugged), based on ChEMBL and the British National Formulary. The metabolism

classes represent the identified metabolism origin of the urinary breakdown product. The breakdown product – CHD effect indicates whether the breakdown product's effect on CHD is risk-increasing or risk-decreasing, based on *cis*-Mendelian randomisation. The protein - CHD effect estimates were obtained from *cis*-Mendelian randomisation and are presented as odds ratios representing the effect of a one standard deviation increase in protein values, along with corresponding confidence intervals. The No. drug column records the number of approved drugs with affinity for the protein as available in ChEMBL. For more details on metabolism breakdown products and full numerical Mendelian randomisation results please refer to **Appendix Table S4**. For a full list of the compounds and their indication and side-effects please refer to **Appendix Table S10**.



## 1 **Figure legends**

2

### 3 **Figure 1. Flowchart of the main steps of this study**

4 N.B. The first three steps are conducted using Mendelian randomisation, based on indicated GWAS  
5 data for CHD, urinary metabolism breakdown products, and plasma proteins. Directionally  
6 concordance is explained in **Figure 2** and the Methods section. Measurements on mRNA and protein  
7 expression are available from carotid plaque samples from the Athero-Express Biobank. Druggability  
8 of plasma proteins is determined using ChEMBL and the British National Formulary. Abbreviations:  
9 CHD = coronary heart disease, MR = Mendelian randomisation, PVI = plaque vulnerability index.

10

### 11 **Figure 2. Example of a triangulated association with directional concordance between a protein** 12 **and its effects on urinary metabolism breakdown product values and CHD**

13 N.B. The plus symbol (in red) represents a positive association. The minus symbol (in blue)  
14 represents a negative association (a risk-decreasing effect). Directional concordance is achieved when  
15 the direction of effects aligns consistently, whether increasing or decreasing.

16

### 17 **Figure 3. Associations of urinary metabolism breakdown products with CHD, presented per** 18 **metabolism class**

19 N.B. Point estimates represent associations between urinary metabolism breakdown products and  
20 coronary heart disease (CHD), obtained from genome-wide Mendelian randomisation. The  
21 metabolism classes (amino acid, energy, lipid, xenobiotics, and unclassified) represent the metabolic  
22 origins of the breakdown product. The right y-axis indicates the number of drugged and druggable  
23 proteins associated with the metabolism breakdown product and CHD, where druggability is sourced  
24 from the British National Formulary and ChEMBL. Genetic associations with 954 urinary metabolism  
25 breakdown products were obtained from Schlosser *et al.* (n=1,627)(12). Genetic associations with  
26 CHD were obtained from Aragam *et al.* (181,522 CHD cases)(13). For more detailed information,

1 please refer to the Methods section and **Appendix Tables S1-S2**. Abbreviations: CI = confidence  
2 interval, OR = odds ratio.

3

4 **Figure 4. Associations of prioritised proteins with CHD and plaque vulnerability in**  
5 **atherosclerotic plaque**

6 N.B. Point estimates represent associations between plasma proteins and coronary heart disease  
7 (CHD), obtained from *cis* Mendelian randomisation. Druggability is based on ChEMBL and the  
8 British National Formulary, distinguishing between proteins targeted by approved drug (drugged  
9 proteins, top panel), proteins targeted by a drug in development (druggable proteins, middle panel)  
10 and not yet druggable proteins (bottom panel). Differential expression of genes coding for prioritised  
11 proteins is determined across cell types based on single-cell RNA sequencing data obtained from  
12 Athero-Express patients (4948 cells, 46 patients). The metabolism classes of the metabolism  
13 breakdown products affected by the protein are indicated on the right y-axis. Genetic associations  
14 with CHD were obtained from Aragam *et al.* (181,522 CHD cases)(13). For a more detailed  
15 description, please refer to the Methods section and the **Supplemental Methods**. Full numerical  
16 results can be found in **Appendix Table S4**. Abbreviations: CI = confidence interval, OR = odds  
17 ratio.

18

19 **Figure 5. Matrix of prioritised proteins associated with CHD and plaque vulnerability in**  
20 **atherosclerotic plaque**

21 N.B. Each column represents a prioritised protein associated with coronary heart disease (CHD) and  
22 plaque vulnerability. The gene names corresponding to each protein are presented in **Table 1**. The  
23 rows present the following: row 1, direction of the protein's association with CHD as obtained by  
24 Mendelian randomisation; row 2-3, association of the gene's mRNA or protein expression in plaque  
25 with plaque vulnerability, obtained from the Athero-Express Biobank; row 4-16, higher expression in  
26 a plaque cell type as compared to expression in all other plaque cell types, obtained from single-cell  
27 RNA sequencing in the Athero-Express Biobank; row 17-21, direction of the protein's association  
28 with metabolism breakdown product(s) presented per metabolism class, based on Mendelian

1 randomisation; row 22, druggability status, distinguishing proteins targeted by approved drug  
2 (drugged proteins) from proteins targeted by a drug in development (druggable proteins) and not yet  
3 druggable proteins, based on ChEMBL and the British National Formulary. Full numerical results can  
4 be found in **Appendix Tables S4, S7-S9**. Abbreviations: CHD = coronary heart disease, ECs =  
5 endothelial cells, NK-Cells = natural killer cells, PVI = plaque vulnerability index, SMCs = smooth  
6 muscle cells.

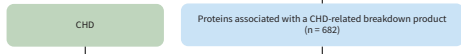
**Step 1**  
*Genome-wide MR*



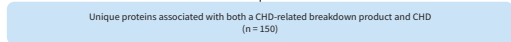
**Step 2**  
*Cis-MR on urinary metabolism breakdown products*



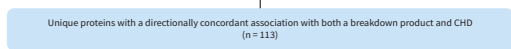
**Step 3**  
*Cis-MR on CHD*



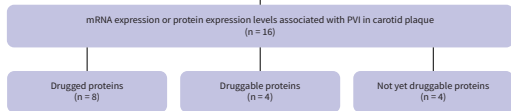
**Step 4**  
*Prioritising on concordance in direction*

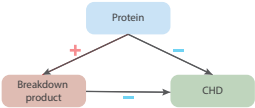


**Step 5**  
*Validation in carotid plaque samples*



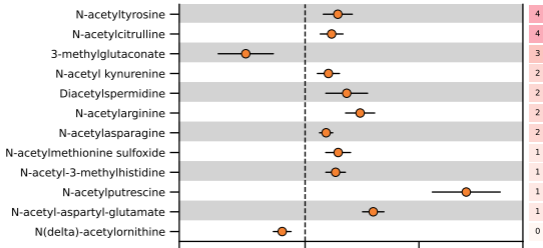
**Step 6**  
*Identification of druggability*





Number of drugged/druggable proteins: 0 1 2 3 4 5 6 7 8 9

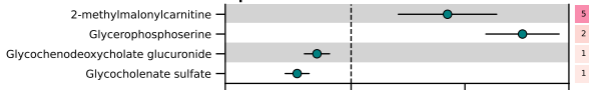
### Amino acid metabolism



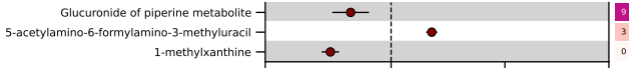
### Energy metabolism



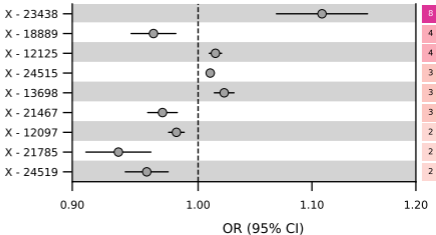
### Lipid metabolism



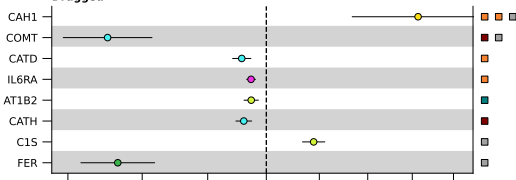
### Xenobiotic metabolism



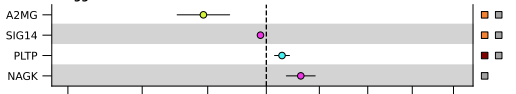
### Unclassified metabolism



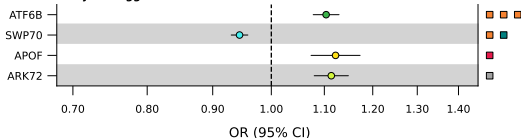
### Drugged



### Druggable



### Not yet druggable

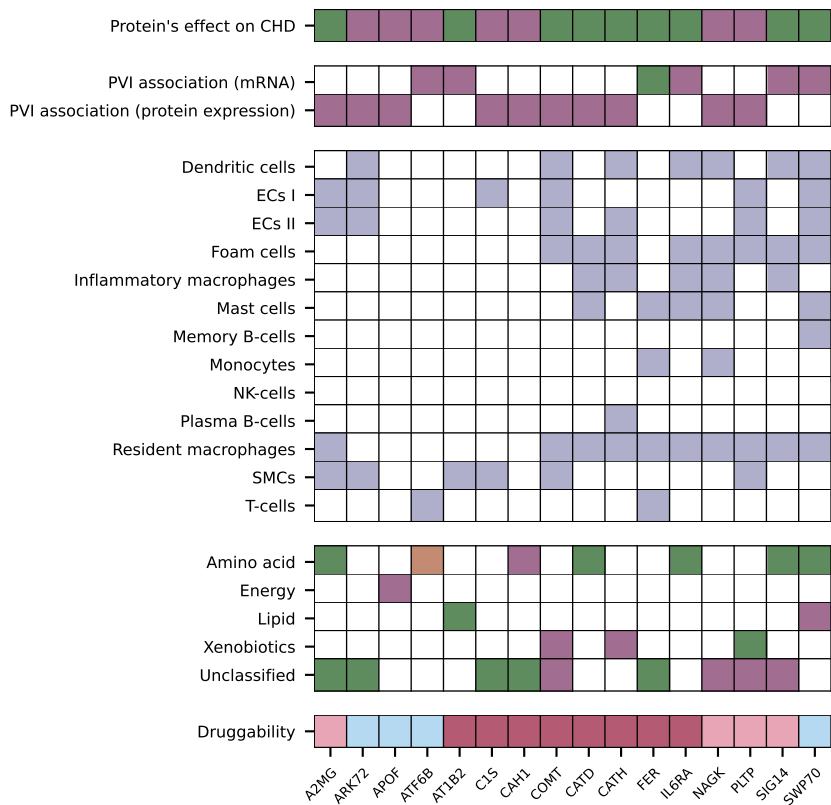


#### Higher expressed in:

- Innate immune cells
- Structural cells
- Innate immune cells and structural cells
- Innate immune cells and adaptive immune cells
- Ubiquitous

#### Metabolism class:

- Amino acid
- Energy
- Lipid
- Xenobiotics
- Unclassified



**Protein's effect on CHD:** Risk decreasing Risk increasing

**Direction of association with PVI:** Negative (lower PVI) Positive (higher PVI) No association

**Higher expression in cell type:** Higher expressed Not higher expressed

**Protein's effect on metabolism breakdown product(s):** Decreasing Increasing Mixed No association

**Druggability:** Drugged Druggable Not yet druggable