ESC European Society of Cardiology

European Journal of Preventive Cardiology (2023) **30**, 583–591
https://doi.org/10.1093/eurjpc/zwad020

FULL RESEARCH PAPER
CVD risk factors

Proteomic profiling identifies novel independent relationships between inflammatory proteins and myocardial infarction

Elsa Valdes-Marquez[1], Robert Clarke [1]*, Michael Hill[1], Hugh Watkins[2,3], and Jemma C. Hopewell[1]*; on behalf of the PROCARDIS Consortium

[1]Clinical Trial Service Unit and Epidemiological Studies Unit, Nuffield Department of Population Health, University of Oxford, Roosevelt Drive, Oxford OX3 7LF, UK; [2]The Wellcome Centre for Human Genetics, University of Oxford, Roosevelt Dr, Headington, Oxford OX3 7BN, UK; and [3]Radcliffe Department of Medicine, Division of Cardiovascular Medicine, University of Oxford, Level 4, Academic Block, John Radcliffe Hospital, Headington, Oxford OX3 9DU, UK

Received 31 August 2022; revised 4 January 2023; accepted 23 January 2023; online publish-ahead-of-print 25 January 2023

See the editorial comment for this article 'Proteomic profiling for investigating the pathophysiology of myocardial infarction', by C. Hage and G. Savarese, https://doi.org/10.1093/eurjpc/zwad019.

Background Inflammation has been implicated in the pathogenesis of coronary heart disease, but the relevance and independence of individual inflammatory proteins is uncertain.

Objective To examine the relationships between a spectrum of inflammatory proteins and myocardial infarction (MI).

Methods and results A panel of 92 inflammatory proteins was assessed using an OLINK multiplex immunoassay among 432 MI cases (diagnosed < 66 years) and 323 controls. Logistic regression was used to estimate associations between individual proteins and MI, after adjustment for established cardiovascular risk factors and medication use, and stepwise regression to identify proteins with independent effects. Machine learning techniques (Boruta analysis and LASSO regression) and bioinformatic resources were used to examine the concordance of results with those obtained by conventional methods and explore the underlying biological processes to inform the validity of the associations. Among the 92 proteins studied, 62 (67%) had plasma concentrations above the lower limit of detection in at least 50% of samples. Of these, 15 individual proteins were significantly associated with MI after covariate adjustment and correction for multiple testing. Five of these 15 proteins (CDCP1, CD6, IL1–8R1, IL-6, and CXCL1) were independently associated with MI, with up to three-fold higher risks of MI per doubling in plasma concentrations. Findings were further validated using machine learning techniques and biologically focused analyses.

Conclusions This study, demonstrating independent relationships between five inflammatory proteins and MI, provides important novel insights into the inflammatory hypothesis of MI and the potential utility of proteomic analyses in precision medicine.

Lay Summary

- The PROCARDIS study conducted a hypothesis-free proteomic study using a panel of 92 inflammatory proteins in cases with early onset myocardial infarction (MI) and healthy controls and identified 15 proteins that were significantly associated with MI, including five proteins that independently contributed to risk of MI.
- The study used state-of-the-art analytical methods including conventional statistical analysis and machine learning approaches to characterize the proteomic associations with MI. It also integrated bioinformatic and genomic data to consider the biological relevance of the proteins independently associated with MI.
- The findings provide novel insights into the 'inflammatory basis' of MI and provide support for prioritizing a wider array of inflammatory proteins for further study than have been previously considered in order to discover if therapeutic modification could be used for treatment and prevention of MI.

* Corresponding author. Tel: +0044 1865 743743, Fax: +0044 1865 743985, Email: robert.clarke@ndph.ox.ac.uk (R.C.); jemma.hopewell@ndph.ox.ac.uk (J.H.)

Introduction

Coronary heart disease (CHD) is the leading cause of death worldwide. Observational and genetic studies have implicated inflammation in the pathogenesis of CHD,[1–3] but the role of individual inflammatory proteins for risk of CHD is not fully understood. The initial observational studies of the inflammatory hypothesis of CHD examined the associations of C-reactive protein (CRP) and fibrinogen ('downstream' markers) with CHD.[4,5] Subsequently, genetic studies using Mendelian randomization approaches refuted the causal relevance of both CRP and fibrinogen for CHD,[6–8] but highlighted the higher risks of CHD associated with inflammatory cytokines ('upstream' markers) that control the production of 'downstream' inflammatory markers, including interleukin-6 (IL-6).[9–11] The CANTOS trial reported that inhibition of interleukin-1β (IL-1β) with Canakinumab, which lowers plasma levels of IL-6 and risk of major vascular events, and provided 'proof of concept' for the inflammatory hypothesis of CHD and its reversibility by targeted drug treatment in high-risk individuals, independent of LDL-C lowering therapy.[12] Further studies have implicated additional inflammatory cytokines for risk of CHD.[13–15]

Advances in multiplex proteomic assays have enabled automated measurements of large numbers of inflammatory proteins using small volumes of plasma.[16] Recognition of the importance of hypothesis-free testing of large panels of proteins for enhanced discovery of novel pathways for risk of CHD and improved targeted approaches to patient care, have prompted the need for more rigorous epidemiological approaches for the analysis of such data in observational studies of cardiovascular disease (CVD).[17]

This hypothesis-free study measured plasma concentrations of 92 inflammatory proteins, using an OLINK multiplex immunoassay, in order to examine associations with risk of early onset myocardial infarction (MI) in the PROCARDIS case-control study. The aim of the present study was to consider the use of both conventional and machine learning approaches to: (i) assess the associations of individual inflammatory proteins with risk of MI after adjustment for established CHD risk factors and medication use; and (ii) assess the independent relevance of the inflammatory proteins with risk of MI.

Methods

Participants

In the PROCARDIS case-control study of CHD, participants were recruited from four European countries (United Kingdom, Italy, Sweden, and Germany) between 2004 and 2008. Cases had an MI before age 66 years as detailed by at least two of three documented sources (e.g. typical ischaemic chest pain and pathological development of Q-waves or elevated plasma cardiac biomarkers). Blood samples were collected on average between 5 and 10 years after onset of their reported coronary artery disease event. Controls were recruited from the same population who had no personal or sibling history of CHD before age 66 years, and it was planned to recruit one control of the same sex, ethnicity, and within 5 years of age of cases.[18,19] A subset of 900 unrelated individuals (450 MI cases and 450 controls) with complete data on relevant covariates [age, sex, smoking status (based on reported smoking habits 10 years prior to enrolment), hypertension status, diabetic status, body mass index, LDL-C, HDL-C, triglycerides, and CRP, as well as medication use including statin and aspirin use], were selected for the present proteomic study.

Proteomic assays

A total of 92 inflammation-related proteins (see Supplementary material online, *Table S1*) were measured using a OLINK immunoassay panel that used a proximity extension assay (OLINK, Uppsala, Sweden).[20] The OLINK assay uses oligonucleotide antibody pairs to bind selected proteins and form unique DNA constructs which are amplified and subsequently quantified using a real-time polymerase chain reaction. These data were pre-processed to generate normalized protein expression (NPX) values, which are presented on a doubling scale (i.e. log base 2), with high NPX values representing high plasma protein levels.[21] Additional details of OLINK measurements, including lower limits of detection and coefficients of variation for the inflammatory panel proteins, are provided at https://www.olink.com.[21]

Of the 900 participants selected for the study, the measurements from blood samples of 106 participants that were mailed to the laboratory were excluded based on assay information subsequently available suggesting possible artefacts due to delayed separation of plasma from red cells on protein concentrations (see Supplementary material online, *Figure S1*). In addition, 39 samples that failed OLINK quality control were excluded (a sample plate median value is calculated each of two internal controls. For each sample, the results of these internal controls are allowed to deviate ±0.3 NPX from the plate median. If the sample deviates more 0.3 NPX, the sample will fail quality control).[22]

Of 92 measured proteins, 24 proteins were excluded because >50% participants had values below the lower limits of detection. In addition, quality control analyses identified six proteins with large deviations between countries of sample collection that may reflect handling errors (see Supplementary material online, *Figure S2*) and were thus excluded. Overall, the main statistical analyses were undertaken on 62 proteins in 755 participants in whom plasma samples were separated within 4 h of blood collection.

Statistical methods

Logistic regression was used to estimate the associations of individual inflammatory proteins with MI risk after adjustment for established CHD risk factors (sex, age, smoking status, hypertension, diabetes, LDL-C, HDL-C, triglycerides, and BMI) and medication use (statins and aspirin). To account for multiple testing, $P < 8.06 \times 10^{-4}$ (Bonferroni correction for 62 tests at $P = 0.05$) was used as a conservative threshold for defining a significant association between individual proteins and MI risk. A Boruta analysis was also used to identify inflammatory proteins associated with MI risk. Boruta is a machine learning classification technique designed to select variables (or features) from multi-dimensional datasets that show associations not consistent with random chance. The Boruta analysis involved multiple random forest runs ($n = 1000$) in which permuted copies of the 62 proteins (referred to as shadow variables), representing proteins with the same distributions as the original proteins but no correlation with MI, were examined. Proteins were classified as 'confirmed' if better, or 'rejected' if not better than the shadow variables. Those not classified at the maximum number of runs were labelled as 'tentative' (see Supplementary material online, *Methods* for additional details).

Stepwise logistic regression was used to identify which of the proteins significantly associated with MI were independent (with all models adjusted for established CHD risk factors and medication use), using a *P*-value threshold for entering and staying in the model of 0.05. To examine the stability of the proteins selected by stepwise logistic regression (i.e. to ascertain that the protein selection was not biased by one or more influential observations), this was repeated for each of 1000 resamples (with replacement) to assess the frequency of the selected proteins.

In addition, stepwise logistic regression with a *P*-value threshold of 0.157 [based on less than 100 events-per-variable (EPV). EPV is the ratio between sample size and number of variables; and quantifies the balance between the amount of information provided by the data and number of estimates],[23] and LASSO logistic regression were used to assess concordance of the selected proteins by the primary conventional stepwise regression analyses (see Supplementary material online, *Methods*). All analyses were performed using SAS 9.4 and R 3.6.2.

Results

Baseline characteristics

Selected characteristics of the 432 MI cases and 323 controls are shown in *Table 1*. The mean (SD) age at blood collection was 62 (7) years in MI cases and 58 (7) in controls. MI cases had higher proportions of hypertension, diabetes and heart failure than controls, in addition to higher mean levels of BMI. MI cases also had higher proportions of statin,

Table 1 Baseline characteristics of myocardial infarction cases and controls

		Controls ($n = 323$)	Cases ($n = 432$)
Baseline Characteristics			
	Age, Years	57.8 (7.1)	62.2 (6.7)
	Sex, Female,%	87 (26.9%)	83 (19.2%)
	Smoker,%	89 (27.6%)	191 (44.2%)
	Hypertension, %	79 (24.5%)	206 (47.7%)
	Diabetes, %	14 (4.3%)	54 (12.5%)
	Heart Failure, %	2 (0.6%)	62 (14.4%)
	Body mass index, kg/m^2	26.2 (3.6)	28.0 (4.1)
	Statin use, %	37 (11.5%)	213 (49.3%)
	Aspirin use, %	19 (5.9%)	274 (63.4%)
	Antihypertensive use, %	58 (18.0%)	333 (77.1%)
Lipids and biomarkers			
	LDL-C, mmol/L	3.2 (0.8)	2.9 (0.8)
	HDL-C, mmol/L	1.4 (0.4)	1.1 (0.3)
	Triglycerides, mmol/L[a]	1.2 (1.1–1.2)	1.6 (1.5–1.7)
	Fibrinogen, mg/L	3.6 (0.8)	4.0 (0.9)
	CRP, mg/L[a]	1.2 (1.0–1.3)	2.2 (1.9–2.4)

[a]Geometric means and 95% confidence intervals.
CRP, C-reactive protein; LDL-C, low-density lipoprotein; HDL-C, high-density lipoprotein. Smoking status based on reported smoking habits 10 years prior to enrolment into the PROCARDIS study.

aspirin, and blood pressure-lowering medication use than controls (49% vs. 11%; 63% vs. 6%; 77% vs. 18% for use of statin, aspirin, and blood-pressure-lowering medications, respectively) at the time of blood collection. Consequently, plasma concentrations of LDL-C were lower in MI cases than in controls. In contrast, the mean levels of CRP (2.20 vs. 1.2 mg/L) were higher in MI cases than in controls (2.20 vs. 1.2 mg/L).

Correlation between established CHD risk factors and individual proteins

The pairwise Spearman correlations between established CHD risk factors (age, LDL-C, HDL-C, triglycerides, BMI, fibrinogen, CRP) and proteins concentrations in controls varied between −0.28 and 0.53 (see Supplementary material online, *Figure S3*), with the strongest correlation between CRP and IL-6 ($r = 0.53$). Age was correlated with many of the inflammatory proteins, varying between a correlation of −0.18 to 0.37, and therefore age-adjusted pairwise correlations between inflammatory proteins are shown in Supplementary material online, *Figure S4*. Of the 1891 age-adjusted pairwise correlations, which varied between −0.23 and 0.74, 97% showed positive correlations and 2% had a correlation over $|r|=0.5$. The strongest of these correlations was between IL-7 and CXCL1 ($r = 0.74$), which were substantially greater than the well-documented correlation observed between CRP and IL-6, illustrating the interdependency of these inflammatory pathways.

Associations of individual proteins with MI

Plasma protein concentrations in MI cases and controls typically had relatively few outliers, with extreme observations chiefly found at high, rather than low plasma protein concentrations (see Supplementary material online, *Figure S5*). After adjusting for established CHD risk factors and medication use, 15 proteins were significantly associated with MI (*Figure 1A* and Supplementary material online, *Figure S6*). LIFR and uPA were most strongly associated with MI, with a six- to eight-fold higher risk of MI per doubling in plasma concentrations (OR: 8.16; 95%CI 3.00 to 22.20; OR: 6.10; 95%CI 2.55 to 14.62; respectively). CSF1, HGF, CD40, CD5, IL-18R1, and CDCP1 were each associated with about four- to five-fold higher risks of MI. The remaining proteins were associated with 1.5 to 2-fold higher risks of MI per doubling in plasma protein concentrations (*Figure 1A*). Of the 32 proteins selected by the Boruta analysis (which did not take into consideration covariate adjustments), 10 overlapped with the 15 proteins identified by conventional approaches (*Figure 1B*, Supplementary material online, *Figure S6*). The distribution of the 15 proteins and their corresponding ORs for MI by fifths of plasma concentrations indicated approximately linear positive associations of MI with plasma protein concentrations (see Supplementary material online, *Figure S7*).

Independent associations of proteins with risk of MI

Stepwise logistic regression analyses (adjusted for CHD risk factors and medication use) indicated that of the 15 proteins associated with MI in PROCARDIS, five proteins (CDCP1, CD6, IL-18R1, IL-6, and CXCL1) were independently associated with MI (*Figure 2*). In the final joint model (including CHD risk factors, medication use and the five independent proteins), CDCP1 was associated with a three-fold higher risk of MI per doubling in plasma concentrations (OR: 2.83; 95%CI 1.65 to 4.85), while IL-6 was associated with a 1.5-fold higher risk of MI (OR: 1.52; 95% CI: 1.09 to 2.21). Assessment of the stability of the five independent proteins identified in the primary analysis (see Supplementary material online, *Methods*) indicated all proteins were selected in >50% of resamples (considered the threshold for confirmation), with the exception of IL-18R1 that was selected in slightly fewer (46%) resamples (*Figure 2* and Supplementary material online, *Table S2*).

Sensitivity analyses yielded broadly concordant results to those evaluated in the primary analyses with an additional six proteins being identified by the different selection approaches considered. Stepwise regression (with a *P*-value threshold based on EPV) identified CCL20

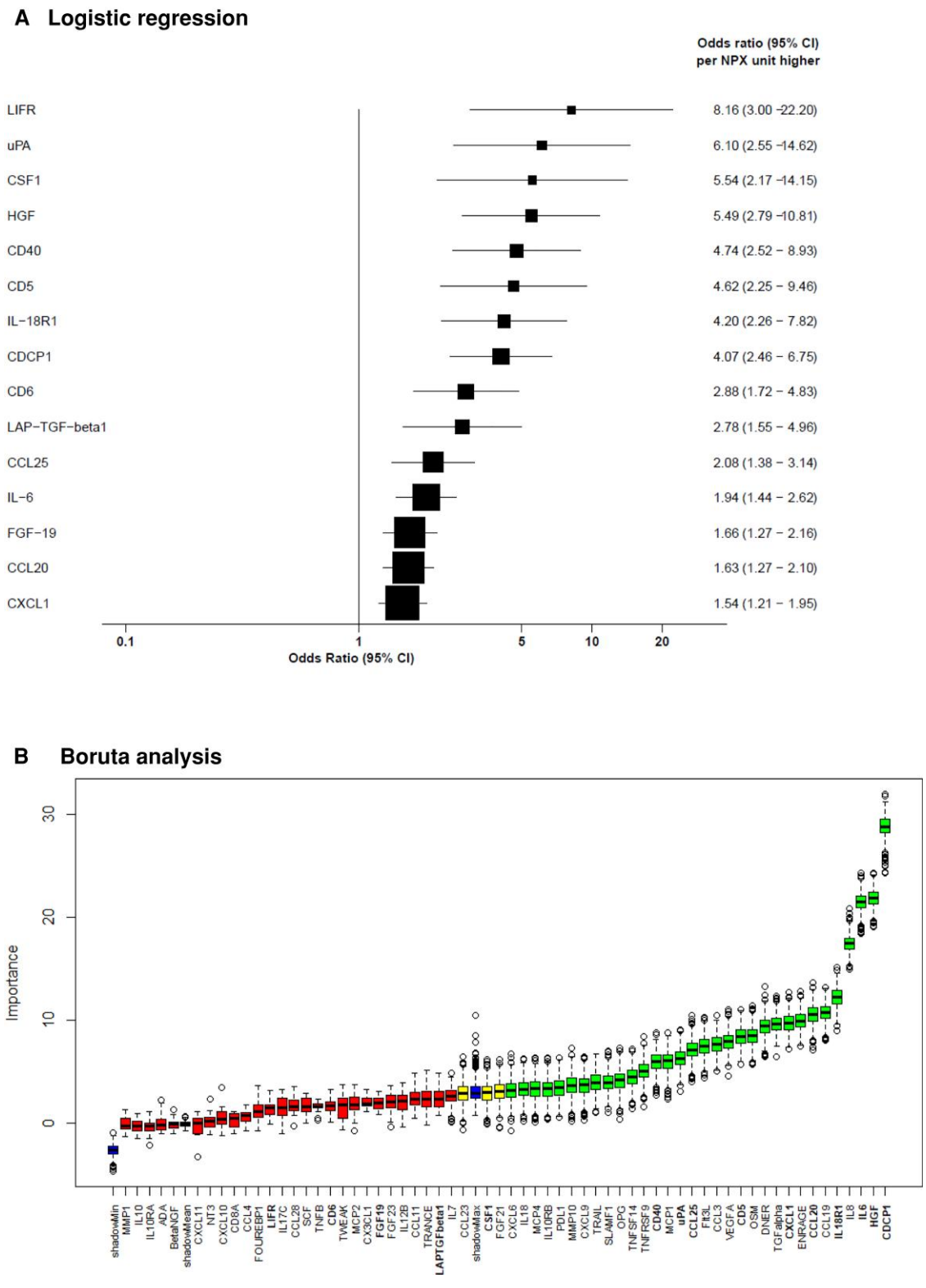

Figure 1 Associations of individual inflammatory proteins with MI using (*A*) logistic regression and (*B*) Boruta analysis. (*A*) Logistic regression. Estimated odds ratios (ORs) of proteins with risk of MI are shown per NPX unit higher after adjustment for age, sex, smoking, hypertension, diabetes status, plasma lipids (LDL-C, HDL-C, and triglycerides), body mass index and medication use (statin and aspirin). A *P*-value of 8.06×10^{-4} (Bonferroni correction for multiple testing, 0.05/62) was used as a threshold for defining significant associations with MI risk. (*B*) Boruta analysis indicates protein 'importance' for MI risk.

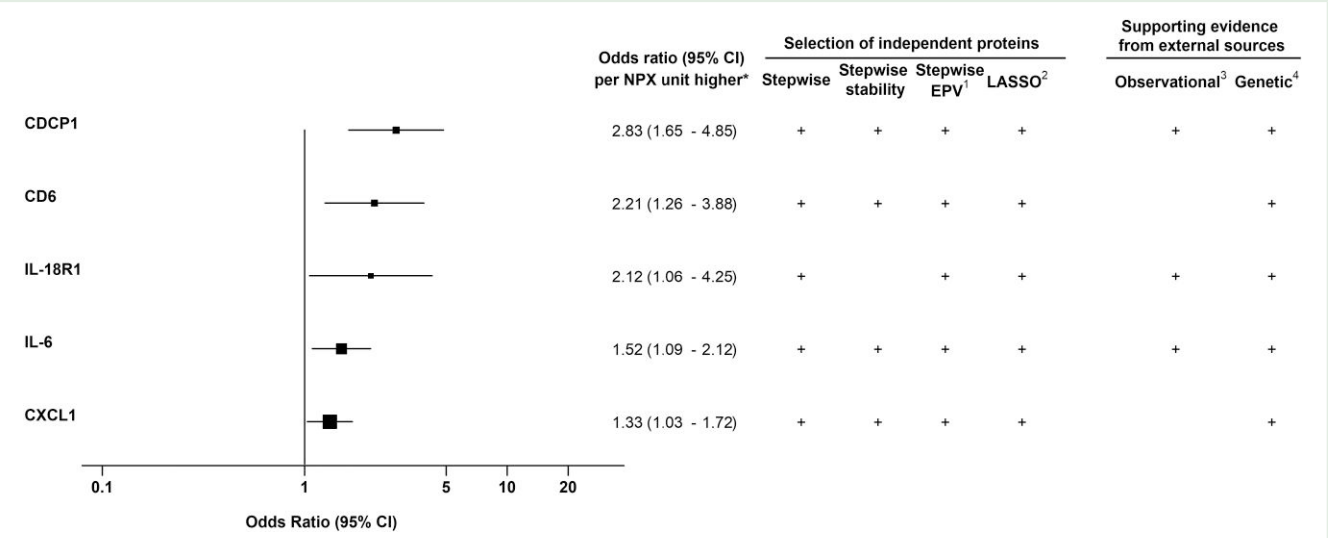

Figure 2 Associations of inflammatory proteins independently associated with MI. Estimated ORs of proteins with risk of MI are shown per NPX unit higher. Notes: *Estimates are based on the joint model adjusted for CHD risk factors, medication use (statin and aspirin), and all other proteins shown in the figure.1 Two additional proteins (CCL20 and FGF-19) were flagged by stepwise EPV.2 Six additional proteins were flagged by LASSO logistic regression (CCL20, CCL25, FGF-19, CD40, uPA, and LIFR).3 Observational evidence obtained from Ferkingstad *et al*. Nat Genet 2021[26] and Wallentin *et al*.[27] PloS Med 2021.4 Genetic evidence was derived from various sources. Phenoscanner (http://www.phenoscanner.medschl.cam.ac.uk/),[24] UK Biobank atlas (http://geneatlas.roslin.ed.ac.uk/phewas/; and CARDIoGRAMplusC4D consortium summary statistic.[25] Genetic variants within the protein coding gene ± 250 kb and with a $P \leq 0.01$ for their association with CHD/MI were considered as genetic evidence. For IL-6 and IL-18R1, two additional genes were considered (IL6R and IL18) besides the gene with the same name as the proteins. For the rest of the proteins, only one gene was considered (gene with the same name as proteins). Additional information can be found in Supplementary material online, *Table S3*.

and FGF-19 in addition to the five proteins from the primary analysis, and all seven proteins were confirmed in corresponding stability analyses (i.e. 50% of resamples), with the exception of CCL20 which was selected in only 49% (*Figure 2* and Supplementary material online, *Table S2*). Finally, LASSO logistic regression selected 11 proteins independently associated with MI, including the five proteins selected in the primary analysis and the two additional proteins observed in the stepwise EPV analysis, as well as CCL25, CD40, uPA, and LIFR (*Figure 2*).

External observational and genetic support for proteins associated with MI

In order to assess the validity of the primary analyses and their biological basis, we screened the findings of external observational and genetic studies, and bioinformatic resources (*Figure 2*). We identified evidence supporting each of the individual proteins that had been shown to be independently associated with MI in this study (i.e. CDCP1, IL-18R1, IL-6, and CXCL1). Observational evidence provided support (i.e. directionally consistent significant associations based on individual study defined statistical significance: Supplementary material online, *Table S3*) with either CVD mortality or CHD/MI for associations for CDCP1, IL-18R1, and IL-6.[26,27] Genetic support was shown for IL-6 through IL6R, while weaker evidence (genetic variants, within the protein coding gene ± 250 kb, and with $P \leq 0.01$ for associations with CHD) was identified for the other four proteins (see Supplementary material online, *Table S3*).

Biological relevance of proteins associated with MI

FUMA (Functional Mapping and Annotation) gene-set enrichment analysis based on Gene Ontology terms and including the 15 proteins individually associated with MI in the present study suggested that these proteins were over-represented in 298 biological processes (adjusted *P*-value less than 5% false discovery rate threshold). The top five biological processes were inflammation response; cytokine mediated signalling pathway; regulation of signalling receptor activity; response to cytokine and positive regulation of intracellular signal transduction (see Supplementary material online, *Figure S8*). Likewise, FUMA gene-set enrichment analysis limited to the five proteins independently associated with MI in PROCARDIS also impacted 49 biological processes. The top five biological processes included inflammatory response, positive regulation of T-cell cytokine production, cellular response to biotic stimulus, CD4 positive alpha beta T-cell cytokine production, and regulation of T-cell cytokine production (*Figure 3*). IL-6 was represented in all 49 biological processes; IL-18R1, CD6, CXCL1 were represented in between 28% and 72% of the processes; however, CDCP1 was not represented in any of the processes.

Discussion

The present study, involving 432 cases and 323 controls, demonstrated that higher plasma concentrations of 15 inflammatory biomarkers were associated with higher risks of MI after adjustment for established CHD risk factors and medication use. Among these 15 inflammatory proteins, five proteins (CDCP1, CD6, IL-18R1, IL-6, and CXCL1) were independently associated with MI. The high degree of concordance of the individual proteins identified using both conventional and machine learning approaches, in addition to support of such associations in external observational and genetic data, provides further support for the validity of the findings of the present study. The 15 proteins associated with MI in the present study included interleukins ($n = 2$), chemokines ($n = 3$), fibroblast growth factors ($n = 1$), and other proteins. The relevance and biology of the individual proteins associated with CHD are considered below.

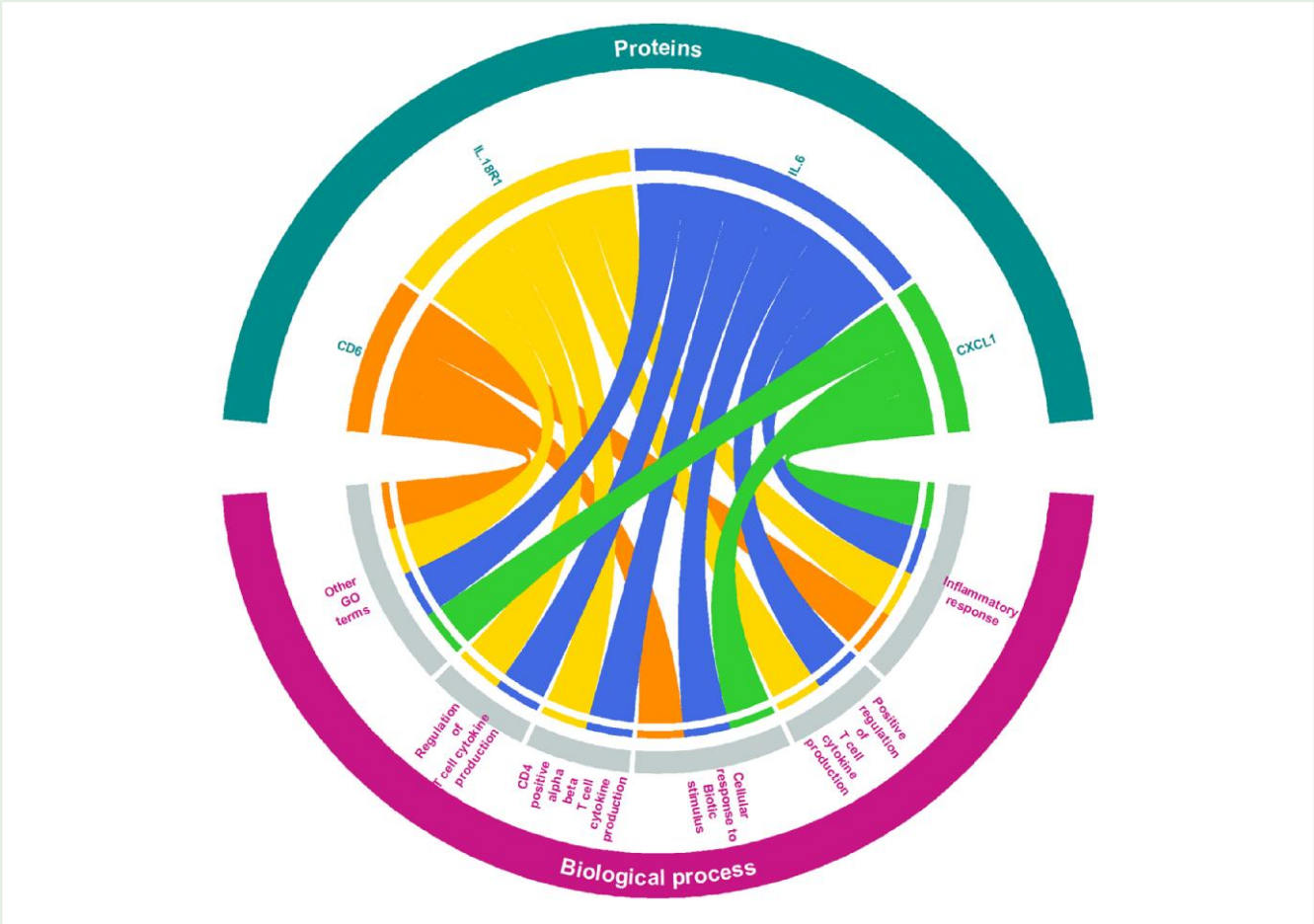

Figure 3 Biological processes most relevant to the five inflammatory proteins independently associated with MI.

Interleukins

Interleukins, produced by leukocytes, and expressed on endothelial cell and smooth muscle cells are typically classified by their amino acid sequence or homology of their receptors.[28] The interleukins significantly associated with MI in the present study were derived from two different families (IL-6 and IL-1), which have been previously linked with atherosclerotic disease.[28,29] This study confirmed the associations of high plasma concentrations of IL-6 and IL-18R1 with risk of MI, consistent with previous studies.[9–11,26] IL-18R1 selectively binds IL-18, which has also been associated with higher risks of CVD.[14,30] However, the present study also demonstrated the independent relevance of IL-6 and IL-18R1 for risk of MI, of each other and other relevant inflammatory proteins.

Chemokines

Chemokines direct leukocytes to local sites of inflammation and play a role in atherosclerosis and CVD.[31] In the present study, two of the most commonly occurring classes of chemokines (C-C and C-X-C) were associated with MI. CXCL1 (C-X-C motif ligand 1) is a chemokine that has chemotactic activity for neutrophils. Higher plasma levels of CXCL1 were associated with higher risks of MI in this study, whilst previous reports have been somewhat inconsistent.[26] In addition to production by immune cells, CXCL1 expression can be induced indirectly by IL-1, TNF-α, IL-17 and is mainly triggered by pathways involved in inflammation. Furthermore, *CXCLC1* gene has a number of interacting proteins including CXCL6 (which has also been linked with higher risks of CHD)[27] as well as other chemokines and interleukins, and is also co-expressed with IL-1β.[32]

CCL25 and CCL20 levels were each associated with MI in PROCARDIS, albeit neither were independent of the other proteins associated with MI. CCL25 (C-C motif chemokine 25) plays a role in the development of T-cells and CCR9 is its unique receptor, which has also been linked with MI through NF-κB and MAP signalling.[33] CCL20 (C-C motif chemokine 20) is a chemokine ligand for C-C chemokine receptor CCR6.[34] Both CCL25 and CCL20 have also been positively associated with higher risks of CVD in previous studies.[27,35]

Fibroblast growth factor

FGF-19 (fibroblast growth factor 19) was associated with MI in the present study. FGF-19 is a hormone and growth factor that regulates bile acid synthesis and has multiple roles in glucose and lipid metabolism.[36] Previous studies have reported that higher levels of FGF-19 were associated with higher risks of CVD.[37]

Other proteins

CDCP1 (CUB domain-containing protein 1) is a transmembrane glycoprotein involved in cell adhesion, autoimmune diseases and cancer. This protein was independently associated with MI risk in the present study

and was positively associated with CHD and CVD mortality in single protein analyses.[26,27] CD6 (T-cell surface glycoprotein CD6 isoform) is a cell adhesion molecule that mediates cell–cell contacts and regulates T-cell response via its interaction with ALCAM/CD166. The *CD6* gene interacts with various proteins such as *CD5, ALCAM, LGALS1, and LGALS3*,[32] which have been positively linked with CVD risk factors, and CVD.[26,27] Support for associations of the remaining proteins individually associated with MI in the present study (LIFR, uPA, CSF1, HGF, CD40, CD5, and LAP TGF-Beta1) has been provided for CVD mortality in another study.[27]

Potential strengths and limitations of present study

This study examined associations with a large number of individual proteins, in addition to assessing their independent relationships, and used external observational studies, in addition to genetic and bioinformatic resources to assess their biological relevance. However, the PROCARDIS study also had several relevant limitations. The blood samples in cases were collected about 5–10 years after a diagnosis of MI, and proteomic assays were conducted in blood samples that had been stored in liquid nitrogen for about 10 years after blood collection, which may impact the stability of the measured proteins.[38] Plasma levels of CDCP1 and FGF-19 are influenced by storage, albeit this only accounted for about 5% of the total variation in plasma concentrations.[39] The case-control design of the PROCARDIS study could also contribute the high proportion of protein concentrations being below the lower limit of detection. Moreover, plasma levels of several proteins could have been influenced by differences in sample handling (e.g. delays prior to centrifugation, freeze–thaw cycles, or sample transportation).[38] The PROCARDIS study did not systematically record data on history of other inflammatory diseases (such as rheumatoid arthritis or psoriasis), presence of CHD symptoms, recurrent CVD events, or detailed use of medications. Hence, the study was unable to account for possible effects of concomitant inflammatory diseases (albeit the prevalence is likely to be low in this population), CHD symptom severity, time between disease onset and blood collection, or use of medications, each of which may influence measured blood levels of inflammatory proteins.

Measurement of a large number of inflammatory proteins using a comprehensive immunoassay panel in a case-control study is an efficient and cost-effective approach to evaluate associations of a large number of inflammatory proteins with risk of MI. However, observational studies cannot fully exclude the effects of confounding or reverse causality, despite rigorous approaches adopted to avoid such biases in the present study. While the proteins identified as associated with MI were replicated by multiple analytical approaches, and supported by external studies, further replication of these associations is still required in prospective studies to further reduce the risk of reverse causality and residual confounding and to assess their relevance for risk prediction. Future research integrating genomic and proteomic data in a wide range of populations is required to assess the causal relevance of these associations.

Conclusions

Overall, the findings of the present study demonstrated that 15 individual inflammatory proteins on the OLINK inflammatory panel were significantly associated with MI risk, of which five contributed independently to the risk of MI. Future Mendelian randomization studies will be necessary to explore the causal relevance of these proteins for MI, and their potential significance as biological and therapeutic targets in atherosclerotic disease. Further analyses of the biological pathways in which these proteins are involved should be informative to distinguish canonical pathways and potential upstream regulators associated with MI. Understanding the role of inflammatory proteins and their related pathways in the pathophysiology of MI could inform drug target prioritization that may lead to discovery of novel treatments for CHD and enhance precision medicine approaches to patient care.

Author contributions

E.V.M., R.C., and J.C.H. conceptualized the study. E.V.M. conducted the analysis, and wrote the first draft of the report. R.C. and J.C.H. have responsibility for the data integrity and accuracy of the data analyses. All authors had full access to all of the study findings, contributed to interpretation of the results and revision of the manuscript, and responsibility for the decision to submit the report for publication.

Supplementary material

Supplementary material is available at *European Journal of Preventive Cardiology* online.

Funding

The PROCARDIS study was funded by 6th Framework Programme of the European Union (LSH-2005-2.1.1.1), British Heart Foundation and Astra Zeneca. R.C., H.W., and J.C.H. acknowledge support by the British Heart Foundation Centre for Research Excellence, Oxford. J.C.H. was funded by the British Heart Foundation (FS/14/55/30806) and also acknowledges support from National Institute for Health Research Oxford Biomedical Research Centre. We thank Federico Murgia for conducting bioinformatic analyses of the biological processes underlying these associations.

Conflict of interest: All authors declare that they have no conflicts of interest in relation to this report.

Data availability

PROCARDIS data are available to bone fide researchers on application to the corresponding authors of this report.

References

1. Hansson GK. Inflammation, atherosclerosis, and coronary artery disease. *N Engl J Med* 2005;**352**:1685–1695.
2. Libby P. Inflammation in atherosclerosis. *Arterioscler Thromb Vasc Biol* 2012;**32**: 2045–2051.
3. Libby P, Ridker PM, Hansson GK. Progress and challenges in translating the biology of atherosclerosis. *Nature* 2011;**473**:317–325.
4. Fibrinogen Studies Collaboration, Danesh J, Lewington S, Thompson SG, Lowe GD, Collins R, Kostis JB, Wilson AC, Folsom AR, Wu K, Benderly M, Goldbourt U, Willeit J, Kiechl S, Yarnell JW, Sweetnam PM, Elwood PC, Cushman M, Psaty BM, Tracy RP, Tybjaerg-Hansen A, Haverkate F, de Maat MP, Fowkes FG, Lee AJ, Smith FB, Salomaa V, Harald K, Rasi R, Vahtera E, Jousilahti P, Pekkanen J, D'Agostino R, Kannel WB, Wilson PW, Tofler G, Arocha-Piñango CL, Rodriguez-Larralde A, Nagy E, Mijares M, Espinosa R, Rodriquez-Roa E, Ryder E, Diez-Ewald MP, Campos G, Fernandez V, Torres E, Marchioli R, Valagussa F, Rosengren A, Wilhelmsen L, Lappas G, Eriksson H, Cremer P, Nagel D, Curb JD, Rodriguez B, Yano K, Salonen JT, Nyyssönen K, Tuomainen TP, Hedblad B, Lind P, Loewel H, Koenig W, Meade TW, Cooper JA, De Stavola B, Knottenbelt C, Miller GJ, Cooper JA, Bauer KA, Rosenberg RD, Sato S, Kitamura A, Naito Y, Palosuo T, Ducimetiere P, Amouyel P, Arveiler D, Evans AE, Ferrieres J, Juhan-Vague I, Bingham A, Schulte H, Assmann G, Cantin B, Lamarche B, Després JP, Dagenais GR, Tunstall-Pedoe H, Woodward M, Ben-Shlomo Y, Davey Smith G, Palmieri V, Yeh JL, Rudnicka A, Ridker P, Rodeghiero F, Tosetto A, Shepherd J, Ford I, Robertson M, Brunner E, Shipley M, Feskens EJ, Kromhout D, Dickinson A, Ireland B, Juzwishin K, Kaptoge S, Lewington S, Memon A, Sarwar N, Walker M, Wheeler J, White I, Wood A. Plasma fibrinogen level and the risk of major cardiovascular diseases and nonvascular mortality: an individual participant meta-analysis. *JAMA* 2005;**294**:1799–1809.

5. Emerging Risk Factors Collaboration, Kaptoge S, Di Angelantonio E, Lowe G, Pepys MB, Thompson SG, Collins R, Danesh J. C-reactive protein concentration and risk of coronary heart disease, stroke, and mortality: an individual participant meta-analysis. *Lancet* 2010;**375**:132–140.
6. Elliott P, Chambers JC, Zhang W,Clarke R, Hopewell JC, Peden JF, Erdmann J, Braund P, Engert JC, Bennett D, Coin L, Ashby D, Tzoulaki I, Brown IJ, Mt-Isa S, McCarthy MI, Peltonen L, Freimer NB, Farrall M, Ruokonen A, Hamsten A, Lim N, Froguel P, Waterworth DM, Vollenweider P, Waeber G, Jarvelin MR, Mooser V, Scott J, Hall AS, Schunkert H, Anand SS, Collins R, Samani NJ, Watkins H, Kooner JS. Genetic loci associated with C-reactive protein levels and risk of coronary heart disease. *JAMA* 2009;**302**:37–48.
7. Keavney B, Danesh J, Parish S, Palmer A, Clark S, Youngman L, Delépine M, Lathrop M, Peto R, Collins R. Fibrinogen and coronary heart disease: test of causality by 'Mendelian randomization'. *Int J Epidemiol* 2006;**35**:935–943.
8. C Reactive Protein Coronary Heart Disease Genetics Collaboration (CCGC), Wensley F, Gao P, Burgess S, Kaptoge S, Di Angelantonio E, Shah T, Engert JC, Clarke R, Davey-Smith G, Nordestgaard BG, Saleheen D, Samani NJ, Sandhu M, Anand S, Pepys MB, Smeeth L, Whittaker J, Casas JP, Thompson SG, Hingorani AD, Danesh J. Association between C reactive protein and coronary heart disease: Mendelian randomisation analysis based on individual participant data. *BMJ* 2011;**342**:d548.
9. Danesh J, Kaptoge S, Mann AG, Sarwar N, Wood A, Angleman SB, Wensley F, Higgins JP, Lennon L, Eiriksdottir G, Rumley A, Whincup PH, Lowe GD, Gudnason V. Long-term interleukin-6 levels and subsequent risk of coronary heart disease: two new prospective studies and a systematic review. *PLoS Med* 2008;**5**:e78.
10. IL6R Genetics Consortium Emerging Risk Factors Collaboration, Sarwar N, Butterworth AS, Freitag DF, Gregson J, Willeit P, Gorman DN, Gao P, Saleheen D, Rendon A, Nelson CP, Braund PS, Hall AS, Chasman DI, Tybjærg-Hansen A, Chambers JC, Benjamin EJ, Franks PW, Clarke R, Wilde AA, Trip MD, Steri M, Witteman JC, Qi L, van der Schoot CE, de Faire U, Erdmann J, Stringham HM, Koenig W, Rader DJ, Melzer D, Reich D, Psaty BM, Kleber ME, Panagiotakos DB, Willeit J, Wennberg P, Woodward M, Adamovic S, Rimm EB, Meade TW, Gillum RF, Shaffer JA, Hofman A, Onat A, Sundström J, Wassertheil-Smoller S, Mellström D, Gallacher J, Cushman M, Tracy RP, Kauhanen J, Karlsson M, Salonen JT, Wilhelmsen L, Amouyel P, Cantin B, Best LG, Ben-Shlomo Y, Manson JE, Davey-Smith G, de Bakker PI, O'Donnell CJ, Wilson JF, Wilson AG, Assimes TL, Jansson JO, Ohlsson C, Tivesten Å, Ljunggren Ö, Reilly MP, Hamsten A, Ingelsson E, Cambien F, Hung J, Thomas GN, Boehnke M, Schunkert H, Asselbergs FW, Kastelein JJ, Gudnason V, Salomaa V, Harris TB, Kooner JS, Allin KH, Nordestgaard BG, Hopewell JC, Goodall AH, Ridker PM, Hólm H, Watkins H, Ouwehand WH, Samani NJ, Kaptoge S, Di Angelantonio E, Harari O, Danesh J. Interleukin-6 receptor pathways in coronary heart disease: a collaborative meta-analysis of 82 studies. *Lancet* 2012;**379**:1205–1213.
11. Interleukin-6 Receptor Mendelian Randomisation Analysis (IL6R MR) Consortium, Swerdlow DI, Holmes MV, Kuchenbaecker KB, Engmann JE, Shah T, Sofat R, Guo Y, Chung C, Peasey A, Pfister R, Mooijaart SP, Ireland HA, Leusink M, Langenberg C, Li KW, Palmen J, Howard P, Cooper JA, Drenos F, Hardy J, Nalls MA, Li YR, Lowe G, Stewart M, Bielinski SJ, Peto J, Timpson NJ, Gallacher J, Dunlop M, Houlston R, Tomlinson I, Tzoulaki I, Luan J, Boer JM, Forouhi NG, Onland-Moret NC, van der Schouw YT, Schnabel RB, Hubacek JA, Kubinova R, Baceviciene M, Tamosiunas A, Pajak A, Topor-Madry R, Malyutina S, Baldassarre D, Sennblad B, Tremoli E, de Faire U, Ferrucci L, Bandenelli S, Tanaka T, Meschia JF, Singleton A, Navis G, Mateo Leach I, Bakker SJ, Gansevoort RT, Ford I, Epstein SE, Burnett MS, Devaney JM, Jukema JW, Westendorp RG, Jan de Borst G, van der Graaf Y, de Jong PA, Mailand-van der Zee AH, Klungel OH, de Boer A, Doevendans PA, Stephens JW, Eaton CB, Robinson JG, Manson JE, Fowkes FG, Frayling TM, Price JF, Whincup PH, Morris RW, Lawlor DA, Smith GD, Ben-Shlomo Y, Redline S, Lange LA, Kumari M, Wareham NJ, Verschuren WM, Benjamin EJ, Whittaker JC, Hamsten A, Dudbridge F, Delaney JA, Wong A, Kuh D, Hardy R, Castillo BA, Connolly JJ, van der Harst P, Brunner EJ, Marmot MG, Wassel CL, Humphries SE, Talmud PJ, Kivimaki M, Asselbergs FW, Voevoda M, Bobak M, Pikhart H, Wilson JG, Hakonarson H, Reiner AP, Keating BJ, Sattar N, Hingorani AD, Casas JP. The interleukin-6 receptor as a target for prevention of coronary heart disease: a Mendelian randomisation analysis. *Lancet* 2012;**379**:1214–1224.
12. Ridker PM, Everett BM, Thuren T, MacFadyen JG, Chang WH, Ballantyne C, Fonseca F, Nicolau J, Koenig W, Anker SD, Kastelein JJP, Cornel JH, Pais P, Pella D, Genest J, Cifkova R, Lorenzatti A, Forster T, Kobalava Z, Vida-Simiti L, Flather M, Shimokawa H, Ogawa H, Dellborg M, Rossi PRF, Troquay RPT, Libby P, Glynn RJ, CANTOS Trial Group. Antiinflammatory therapy with canakinumab for atherosclerotic disease. *N Engl J Med* 2017;**377**:1119–1131.
13. Clarke R, Valdes-Marquez E, Hill M, Gordon J, Farrall M, Hamsten A, Watkins H, Hopewell JC. Plasma cytokines and risk of coronary heart disease in the PROCARDIS study. *Open Heart* 2018;**5**:e000807.
14. Kaptoge S, Seshasai SR, Gao P, *Freitag DF, Butterworth AS, Borglykke A, Di Angelantonio E, Gudnason V, Rumley A, Lowe GD, Jørgensen T, Danesh J*. Inflammatory cytokines and risk of coronary heart disease: new prospective study and updated meta-analysis. *Eur Heart J* 2014;**35**:578–589.
15. Lind L, Zanetti D, Ingelsson M, Gustafsson S, Ärnlöv J, Assimes TL. Large-scale plasma protein profiling of incident myocardial infarction, ischemic stroke, and heart failure. *J Am Heart Assoc* 2021;**10**:e023330.
16. Lam MP, Ping P, Murphy E. Proteomics research in cardiovascular medicine and biomarker discovery. *J Am Coll Cardiol* 2016;**68**:2819–2830.
17. Folkersen L, Gustafsson S, Wang Q, Hansen DH, Hedman ÅK, Schork A, Page K, Zhernakova DV, Wu Y, Peters J, Eriksson N, Bergen SE, Boutin TS, Bretherick AD, Enroth S, Kalnapenkis A, Gådin JR, Suur BE, Chen Y, Matic L, Gale JD, Lee J, Zhang W, Quazi A, Ala-Korpela M, Choi SH, Claringbould A, Danesh J, Davey Smith G, de Masi F, Elmståhl S, Engström G, Fauman E, Fernandez C, Franke L, Franks PW, Giedraitis V, Haley C, Hamsten A, Ingason A, Johansson Å, Joshi PK, Lind L, Lindgren CM, Lubitz S, Palmer T, Macdonald-Dunlop E, Magnusson M, Melander O, Michaelsson K, Morris AP, Mägi R, Nagle MW, Nilsson PM, Nilsson J, Orho-Melander M, Polasek O, Prins B, Pålsson E, Qi T, Sjögren M, Sundström J, Surendran P, Võsa U, Werge T, Wernersson R, Westra HJ, Yang J, Zhernakova A, Ärnlöv J, Fu J, Smith JG, Esko T, Hayward C, Gyllensten U, Landen M, Siegbahn A, Wilson JF, Wallentin L, Butterworth AS, Holmes MV, Ingelsson E, Mälarstig A. Genomic and drug target evaluation of 90 cardiovascular proteins in 30,931 individuals. *Nat Metab* 2020;**2**:1135–1148.
18. Clarke R, Peden JF, Hopewell JC, Kyriakou T, Goel A, Heath SC, Parish S, Barlera S, Franzosi MG, Rust S, Bennett D, Silveira A, Malarstig A, Green FR, Lathrop M, Gigante B, Leander K, de Faire U, Seedorf U, Hamsten A, Collins R, Watkins H, Farrall M, PROCARDIS Consortium. Genetic variants associated with Lp(a) lipoprotein level and coronary disease. *N Engl J Med* 2009;**361**:2518–2528.
19. Farrall M, Green FR, Peden JF, Olsson PG, Clarke R, Hellenius ML, Rust S, Lagercrantz J, Franzosi MG, Schulte H, Carey A, Olsson G, Assmann G, Tognoni G, Collins R, Hamsten A, Watkins H. Genome-wide mapping of susceptibility to coronary artery disease identifies a novel replicated locus on chromosome 17. *PLoS Genet* 2006;**2**:e72.
20. OLINK. Inflammation panel. https://www.olink.com/products/target/inflammation/ (accessed 15 June 2022).
21. OLINK. Validation documentation for inflammation panel (article number 95302). https://www.olink.com/resources-support/document-download-center/ (accessed 15 June 2022).
22. OLINK. FAQ: How is quality control performed? https://olink.com/faq/how-is-quality-control-of-the-data-performed/ (accessed 15 June 2022).
23. Heinze G, Wallisch C, Dunkler D. Variable selection—A review and recommendations for the practicing statistician. *Biom J* 2018;**60**:431–449.
24. Kamat MA, Blackshaw JA, Young R, Surendran P, Burgess S, Danesh J, Butterworth AS, Staley JR. Phenoscanner V2: an expanded tool for searching human genotype-phenotype associations. *Bioinformatics* 2019;**35**:4851–4853.
25. Nikpay M, Goel A, Won HH, Hall LM, Willenborg C, Kanoni S, Saleheen D, Kyriakou T, Nelson CP, Hopewell JC, Webb TR, Zeng L, Dehghan A, Alver M, Armasu SM, Auro K, Bjonnes A, Chasman DI, Chen S, Ford I, Franceschini N, Gieger C, Grace C, Gustafsson S, Huang J, Hwang SJ, Kim YK, Kleber ME, Lau KW, Lu X, Lu Y, Lyytikäinen LP, Mihailov E, Morrison AC, Pervjakova N, Qu L, Rose LM, Salfati E, Saxena R, Scholz M, Smith AV, Tikkanen E, Uitterlinden A, Yang X, Zhang W, Zhao W, de Andrade M, de Vries PS, van Zuydam NR, Anand SS, Bertram L, Beutner F, Dedoussis G, Frossard P, Gauguier D, Goodall AH, Gottesman O, Haber M, Han BG, Huang J, Jalilzadeh S, Kessler T, König IR, Lannfelt L, Lieb W, Lind L, Lindgren CM, Lokki ML, Magnusson PK, Mallick NH, Mehra N, Meitinger T, Memon FU, Morris AP, Nieminen MS, Pedersen NL, Peters A, Rallidis LS, Rasheed A, Samuel M, Shah SH, Sinisalo J, Stirrups KE, Trompet S, Wang L, Zaman KS, Ardissino D, Boerwinkle E, Borecki IB, Bottinger EP, Buring JE, Chambers JC, Collins R, Cupples LA, Danesh J, Demuth I, Elosua R, Epstein SE, Esko T, Feitosa MF, Franco OH, Franzosi MG, Granger CB, Gu D, Gudnason V, Hall AS, Hamsten A, Harris TB, Hazen SL, Hengstenberg C, Hofman A, Ingelsson E, Iribarren C, Jukema JW, Karhunen PJ, Kim BJ, Kooner JS, Kullo IJ, Lehtimäki T, Loos RJF, Melander O, Metspalu A, März W, Palmer CN, Perola M, Quertermous T, Rader DJ, Ridker PM, Ripatti S, Roberts R, Salomaa V, Sanghera DK, Schwartz SM, Seedorf U, Stewart AF, Stott DJ, Thiery J, Zalloua PA, O'Donnell CJ, Reilly MP, Assimes TL, Thompson JR, Erdmann J, Clarke R, Watkins H, Kathiresan S, McPherson R, Deloukas P, Schunkert H, Samani NJ, Farrall M. A comprehensive 1,000 genomes-based genome-wide association meta-analysis of coronary artery disease. *Nat Genet* 2015;**47**:1121–1130.
26. Ferkingstad E, Sulem P, Atlason BA, Sveinbjornsson G, Magnusson MI, Styrmisdottir EL, Gunnarsdottir K, Helgason A, Oddsson A, Halldorsson BV, Jensson BO, Zink F, Halldorsson GH, Masson G, Arnadottir GA, Katrinardottir H, Juliusson K, Magnusson MK, Magnusson OT, Fridriksdottir R, Saevarsdottir S, Gudjonsson SA, Stacey SN, Rognvaldsson S, Eiriksdottir T, Olafsdottir TA, Steinthorsdottir V, Tragante V, Ulfarsson MO, Stefansson H, Jonsdottir I, Holm H, Rafnar T, Melsted P, Saemundsdottir J,

Norddahl GL, Lund SH, Gudbjartsson DF, Thorsteinsdottir U, Stefansson K. Large-scale integration of the plasma proteome with genetics and disease. *Nat Genet* 2021;**53**:1712–1721.
27. Wallentin L, Eriksson N, Olszowka M, Grammer TB, Hagström E, Held C, Kleber ME, Koenig W, März W, Stewart RAH, White HD, Åberg M, Siegbahn A. Plasma proteins associated with cardiovascular death in patients with chronic coronary heart disease: a retrospective study. *PLoS Med* 2021;**18**:e1003513.
28. von der Thüsen JH, Kuiper J, van Berkel TJ, Biessen EA. Interleukins in atherosclerosis: molecular pathways and therapeutic potential. *Pharmacol Rev* 2003;**55**:133–166.
29. Bartekova M, Radosinska J, Jelemensky M, Dhalla NS. Role of cytokines and inflammation in heart function during health and disease. *Heart Fail Rev* 2018;**23**:733–758.
30. Ridker PM, MacFadyen JG, Thuren T, Libby P. Residual inflammatory risk associated with interleukin-18 and interleukin-6 after successful interleukin-1β inhibition with canakinumab: further rationale for the development of targeted anti-cytokine therapies for the treatment of atherothrombosis. *Eur Heart J* 2020;**41**:2153–2163.
31. Noels H, Weber C, Koenen RR. Chemokines as therapeutic targets in cardiovascular disease. *Arterioscler Thromb Vasc Biol* 2019;**39**:583–592.
32. GeneCards. The Human Gene database. https://www.genecards.org/ (accessed 03/January/2023).
33. Huang Y, Wang D, Wang X, Zhang Y, Liu T, Chen Y, Tang Y, Wang T, Hu D, Huang C. Abrogation of CC chemokine receptor 9 ameliorates ventricular remodeling in mice after myocardial infarction. *Sci Rep* 2016;**6**:32660.
34. Manthey HD, Cochain C, Barnsteiner S, Karshovska E, Pelisek J, Koch M, Chaudhari SM, Busch M, Eckstein HH, Weber C, Koenen RR, Zernecke A. CCR6 Selectively promotes monocyte mediated inflammation and atherogenesis in mice. *Thromb Haemost* 2013;**110**:1267–1277.
35. Feldreich T, Nowak C, Carlsson AC, Östgren CJ, Nyström FH, Sundström J, Carrero-Roig JJ, Leppert J, Hedberg P, Giedraitis V, Lind L, Cordeiro A, Ärnlöv J. The association between plasma proteomics and incident cardiovascular disease identifies MMP-12 as a promising cardiovascular risk marker in patients with chronic kidney disease. *Atherosclerosis* 2020;**307**:11–15.
36. Domouzoglou EM, Naka KK, Vlahos AP, Papafaklis MI, Michalis LK, Tsatsoulis A, Maratos-Flier E. Fibroblast growth factors in cardiovascular disease: the emerging role of FGF21. *Am J Physiol Heart Circ Physiol* 2015;**309**:H1029–H1038.
37. Wong YK, Cheung CYY, Tang CS, Au KW, Hai JSH, Lee CH, Lau KK, Cheung BMY, Sham PC, Xu A, Lam KSL, Tse HF. Age-biomarkers-clinical risk factors for prediction of cardiovascular events in patients with coronary artery disease. *Arterioscler Thromb Vasc Biol* 2018;**38**:2519–2527.
38. OLINK. Pre-analytical variation in protein biomarker research. https://www.olink.com/resources-support/white-papers-from-olink/ (accessed 15 June 2022).
39. Enroth S, Hallmans G, Grankvist K, *Gyllensten U.* Effects of long-term storage time and original sampling month on biobank plasma protein concentrations. *EBioMedicine* 2016;**12**:309–314.