THE ANATOLIAN JOURNAL OF CARDIOLOGY

Official journal of the TURKISH SOCIETY OF CARDIOLOGY

Circulating Protein Biomarkers for Primary Cardiomyopathy: Evidence from Mendelian Randomization and Clinical Validation

ABSTRACT

Background: Primary (non-ischemic) cardiomyopathy poses diagnostic challenges due to phenotypic heterogeneity and non-specific clinical presentation, and commonly used circulating biomarkers have limited disease specificity. The aim was to prioritize mechanism-informed plasma biomarkers and evaluate their diagnostic discrimination in a clinical cohort.

Methods: pQTL data from 8 proteomic studies covering 5,034 proteins were integrated, and Mendelian randomization was performed across 3 cardiomyopathy GWAS datasets. Top MR candidates were clinically validated by enzyme-linked immunosorbent assays quantification in a prospective case–control cohort of 81 subjects (48 cardiomyopathy patients and 33 healthy controls). Echocardiography evaluated cardiac function and remodeling, correlating with plasma biomarker levels.

Results: Proteome-wide MR prioritized 27 proteins linked to cardiomyopathy risk. Clinical validation confirmed significant differences in plasma levels of 5 selected biomarkers (CD163, LGALS3BP, FABP5, FSTL3, VCAM1). CD163 (odds ratio (OR) = 1.019, $P < .0001$) was positively associated with CM, while LGALS3BP (OR = 0.390, $P < .0001$), FABP5, FSTL3, and VCAM1 showed inverse associations. Multivariate analysis revealed CD163 and LGALS3BP as independent CM predictors, strongly correlating with echocardiographic indices of cardiac dysfunction. Receiver operating characteristic analyses showed individual area under the curve (AUC) ranging from 0.632 (VCAM1) to 0.917 (CD163). Given the modest discrimination of VCAM1, 5- versus 4-protein combined models were compared, and a parsimonious 4-protein panel was selected (CD163, LGALS3BP, FABP5, FSTL3). The 4-protein panel achieved an apparent AUC of 0.993, with optimism-corrected and repeated cross-validated AUCs of 0.988 and 0.980.

Conclusion: This integrative genetic-to-clinical approach supports a 4-protein plasma panel with promising diagnostic discrimination for primary cardiomyopathy, warranting external validation. CD163 and LGALS3BP emerged as particularly informative markers for further evaluation.

Keywords: CD163, clinical validation, ELISA, LGALS3BP, mendelian randomization, plasma biomarkers, primary cardiomyopathy

INTRODUCTION

Cardiomyopathy is a major contributor to heart failure, arrhythmias, and sudden cardiac death worldwide, with increasing prevalence and substantial clinical and socioeconomic impact.[1,2] Although echocardiography and cardiac MRI can detect structural and functional abnormalities, diagnostic evaluation can be challenging in routine practice because phenotypic presentations are heterogeneous and symptoms are often non-specific, particularly in earlier or milder stages. In addition, commonly used circulating biomarkers reflect myocardial injury or hemodynamic stress but have limited disease specificity. These gaps highlight the need for complementary circulating markers that can support diagnostic discrimination and patient characterization in primary cardiomyopathy.[3-5]

Mounting evidence suggests that chronic inflammation, myocardial fibrosis, metabolic dysregulation, and extracellular matrix remodeling play central roles in the pathogenesis and progression of cardiomyopathy—even before overt systolic

Copyright@Author(s) - Available online at anatoljcardiol.com.
Content of this journal is licensed under a Creative Commons Attribution-NonCommercial 4.0 International License.

ORIGINAL INVESTIGATION

Jie He[1]
Xiaoyan Liu[2]
Jianhui Tang[1]
Chunrong Tang[1]
Xiaofen Liu[1]
Zhuoyuan Yu[3]

[1]*Department of Cardiology, Jintang County Second People's Hospital, Chengdu, Sichuan, People's Republic of China*
[2]*Department of Military Traffic Injury Prevention and Control, Daping Hospital, Army Medical University, Chongqing, China*
[3]*Department of Urology, The First Affiliated Hospital of Chongqing Medical University, Chongqing, People's Republic of China*

Corresponding author:
Zhuoyuan Yu
✉ dryzy@stu.cqmu.edu.cn

Received: October 23, 2025
Accepted: May 13, 2026
Available Online Date: August 4, 2026

Cite this article as: He J, Liu X, Tang J, Tang C, Liu X, Yu Z. Circulating protein biomarkers for primary cardiomyopathy: evidence from mendelian randomization and clinical validation. *Anatol J Cardiol.* 2026;30(10):684-692.

DOI: 10.14744/AnatolJCardiol.2026.5927

dysfunction develops.[6-8] However, sensitive and mechanism-specific plasma biomarkers that reflect these key processes are lacking in routine clinical practice. Most currently used blood-based markers predominantly reflect end-stage damage rather than upstream, actionable pathophysiology.[9] Recent studies reported oxidative stress and neurohumoral biomarkers associated with heart failure phenotypes, underscoring the need for mechanism-informed plasma markers.[10,11]

Large-scale genetic studies have now cataloged thousands of protein quantitative trait loci (pQTLs), providing a unique opportunity to systematically prioritize circulating proteins with potential causal relevance to disease.[12,13] Mendelian randomization (MR) can leverage these genetic instruments to identify proteins that may not only signal increased risk but could serve as future therapeutic targets. Nonetheless, few studies have moved beyond genetic discovery to real-world clinical translation—prospectively validating MR-nominated candidates in contemporary patient cohorts. Recent reports highlight genetic heterogeneity within primary cardiomyopathies, including β1-adrenergic receptor polymorphisms in hypertrophic cardiomyopathy, supporting precision, genetics-anchored approaches.[14]

In this study, proteome-wide pQTL data were integrated with large-scale cardiomyopathy GWAS, using MR to identify candidate plasma biomarkers, and then validating their diagnostic utility in a prospectively recruited clinical cohort. The clinical validation focused on primary cardiomyopathy as an overall entity: although subtype heterogeneity exists, these phenotypes commonly converge on shared pathways of inflammation, fibrosis, and myocardial remodeling in real-world diagnostic pathways. Therefore, the subtype distribution and key echocardiographic measurements are reported, while standardized echocardiographic assessment and targeted enzyme-linked immunosorbent assays (ELISA) quantification to develop a clinically relevant plasma biomarker panel for primary cardiomyopathy.

METHODS

Study Design

This study followed an integrative design beginning with proteome-wide genetic analyses and culminating in clinical validation (Figure 1). First, summary statistics from large-scale plasma proteomic genome-wide association studies were assembled (pQTL GWAS). Protein quantitative trait loci (pQTLs) across the human proteome were identified and then applied 2-sample MR to test for causal effects of circulating proteins on cardiomyopathy risk. Finally, top MR candidates were examined in a prospective clinical cohort using ELISAs for biomarker validation. The overall workflow from pQTL data integration and MR screening to clinical validation is illustrated schematically in Figure 1.

HIGHLIGHTS

- Proteome-wide MR screens 5,034 proteins for cardiomyopathy risk.
- Prospective enzyme-linked immunosorbent assays validation confirms 4 plasma biomarkers.
- CD163 predicts higher risk; LGALS3BP is protective; both are independent.
- Four-protein panel shows strong discrimination after internal validation.
- Biomarker levels track echocardiographic dysfunction.

Proteomic Protein Quantitative Trait Loci Datasets

pQTL findings from 8 published proteomic studies, encompassing a total of 5 034 unique plasma proteins and 13 636 independent pQTL instruments, were aggregated (2 157 cis-pQTLs and 11 479 trans-pQTLs). The data sources and cohort details are as follows: Fenland study (UK), N = 10 708 participants, 4 775 proteins assayed by SomaScan v4.[15] Iceland study, N = 35 559 Icelanders, 4 719 proteins (Somalogic aptamers).[16,17] INTERVAL study (UK), N = 3 301 blood donors, 2,995 proteins (SomaScan assay).[18] The UK Biobank Pharma Proteomics Project (UKB-PPP), N = 34 557 UK Biobank participants, 2 922 unique proteins (Olink Explore 3072 panels).[19] This is the largest pQTL resource, with ~14,287 significant protein–variant associations (81% novel) reported. The published primary pQTL hits from the full UKB-PPP release were used.[19] Additionally, pQTL data from the initial UKB-PPP Pilot/Phase I (≈35 571 participants) for cross-validation was included.[20] KORA F4 study (Germany), N = 1 000 adults, 1 124 proteins (Somalogic v1).[21] SCALLOP consortium (Europe), N = 30 931, 90 cardiovascular-related proteins (Olink panels).[22] FHS Offspring study (USA), N = 6 861 Framingham Heart Study participants, 71 selected CVD proteins (multiplex immunoassays).[23]

All pQTL coordinates were aligned to the human hg19 reference genome. Uniform quality control and instrument selection criteria were applied across datasets. Only pQTL variants reaching genome-wide significance ($P < 5 \times 10^{-8}$) in their original study were considered. The extended major histocompatibility complex (MHC) region on chr6 (positions 25-34 Mb) were excluded to avoid complex LD. Within each protein's locus, if multiple significant SNPs were in linkage disequilibrium, LD clumping ($r^2 < .001$) were performed to retain a single sentinel SNP per locus. For each protein, one or more independent pQTL instruments were selected that were sufficiently strong (F-statistic > 10, corresponding to an approximate $r^2 > 1\%$ in the discovery cohort). When a protein was measured in multiple studies, the dataset with the highest variance explained (e.g. the pQTL with highest r^2 for that protein) was retained to represent that protein's genetic instrument. Cis-pQTLs were defined as variants located within or near the encoding gene (within ±500 kb of the gene locus), while trans-pQTLs were those mapping outside this cis window. In total, 13 636 SNP–protein instruments were curated (2 157 cis and 11 479 trans) for the MR analysis. Each instrument was oriented to the allele associated with higher protein level, and palindromic SNPs were aligned by effect allele frequency to ensure consistency between exposure and outcome data.

Cardiomyopathy GWAS Datasets

Three complementary genome-wide association datasets for cardiomyopathy outcomes were obtained to serve

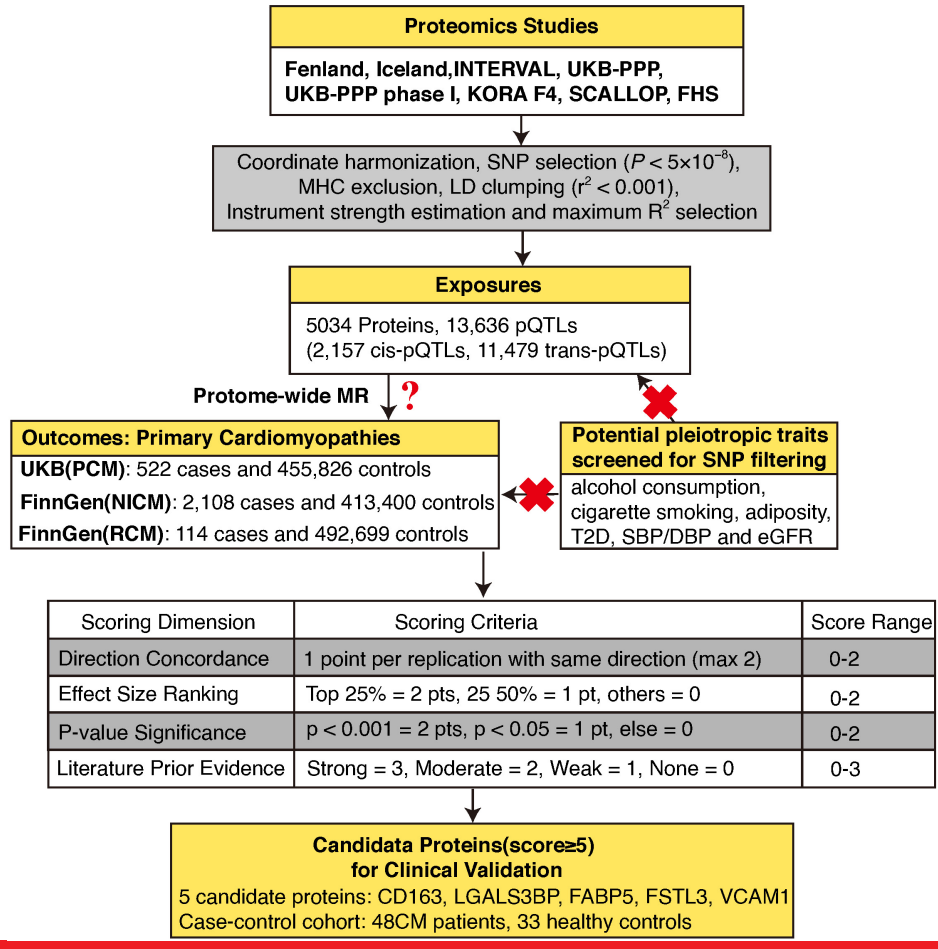


Figure 1. Study design and workflow for biomarker discovery and validation. An overview of the integrative pipeline combining proteome-wide pQTL data from 8 large-scale proteomic studies with 3 cardiomyopathy GWAS datasets (UKB–PCM, FinnGen–NICM, and FinnGen–RCM) using Mendelian randomization (MR). After harmonization, filtering, and instrument selection, 13,636 pQTLs (2,157 cis and 11,479 trans) across 5,034 proteins were evaluated. A multi-dimensional scoring system based on replication consistency, effect size, *P*-value, and literature evidence was used to prioritize candidate proteins for clinical validation. Five proteins (CD163, LGALS3BP, FABP5, FSTL3, and VCAM1) were selected for ELISA testing (48 CM patients, 33 healthy controls); subsequent panel refinement yielded a parsimonious 4-protein panel as the primary combined model. The pleiotropy-screening traits were used for instrument screening to mitigate horizontal pleiotropy, rather than representing a comprehensive list of clinical differential diagnoses.

as MR outcome data. (1) UK Biobank primary cardiomyopathy: Cases of primary (non-ischemic) cardiomyopathy in UK Biobank were defined as participants with any cardiomyopathy diagnosis (e.g. dilated or hypertrophic cardiomyopathy, ICD-10 I42/I43) not attributable to ischemic heart disease. There were 522 cases and 455 826 controls of European ancestry in this analysis. (2) FinnGen Non-ischemic Cardiomyopathy: FinnGen Release 11 results were used for the endpoint "Non-ischemic cardiomyopathy," which had 2 108 cases and 413 400 controls of Finnish ancestry. FinnGen cases were identified via ICD codes (e.g. I42.x excluding ischemic causes); summary statistics were obtained from the FinnGen open data portal. (3) FinnGen restrictive cardiomyopathy: This analysis from FinnGen R11 included 114 cases of restrictive cardiomyopathy (a rare phenotype) and 492 699 controls. In both FinnGen GWAS, logistic models were adjusted for age, sex, genotype batch, and population structure, and individuals with confounding conditions (such as ischemic heart disease, chronic heavy alcohol use, uncontrolled diabetes, obesity, severe hypertension, or chronic kidney disease) were excluded by FinnGen's endpoint definitions to isolate idiopathic cardiomyopathy.

For each protein, its causal effect on cardiomyopathy were estimated in each GWAS dataset using appropriate MR methods: the Wald ratio for single-instrument proteins and inverse-variance weighted (IVW) meta-analysis for proteins with multiple independent pQTLs. Heterogeneity among instruments was evaluated using Cochran's Q test, and a random-effects IVW model was used for sensitivity analysis when significant heterogeneity was detected; however, fixed-effects IVW was used for primary results due to typically small instrument numbers.

In ther analysis, most MR *P*-values were relatively large, and after FDR correction, no protein reached statistical significance. Therefore, the MR results were treated

as hypothesis-generating signals for prioritization rather than definitive discovery. Nominal MR evidence together with a pre-specified multi-dimensional scoring framework were used to select candidates for prospective clinical testing. This approach enables the identification of potentially meaningful biomarkers that might otherwise be overlooked due to statistical conservatism. Finally, MR results across the 3 cardiomyopathy GWAS datasets were synthesized using fixed-effect meta-analysis (inverse-variance weighted) to generate overall estimates and assess consistency.

Protein Prioritization Scoring Scheme

To prioritize candidates for clinical testing, a pre-specified composite scoring framework (Supplementary Table 1) intended as a heuristic ranking tool rather than formal inference. The component weights and cutoffs were chosen pragmatically a priori to reflect translational considerations, and were not derived from an optimized statistical model. The score combines 4 domains capturing direction concordance across datasets, magnitude of MR effects, nominal statistical support, and prior literature evidence. Proteins above a predefined score threshold were shortlisted, and final selection for ELISA testing further considered assay availability and clinical feasibility. This composite scoring framework was defined for the present study and has not been externally validated.

After performing 2-sample MR separately in 3 cardiomyopathy GWAS datasets (UK Biobank primary cardiomyopathy, FinnGen non-ischemic cardiomyopathy, and FinnGen restrictive cardiomyopathy), each protein was prioritized using a 4-dimension composite score (range 0-9):

1. Direction concordance (0–2 points): Proteins received higher scores when the direction of association was consistent across the three cardiomyopathy GWAS datasets.
2. Effect magnitude (0–2 points): Proteins were ranked according to the relative magnitude of the observed MR effect estimates, with larger effects receiving higher scores.
3. Nominal statistical support (0–2 points): Proteins with stronger nominal statistical evidence in the broader cardiomyopathy GWAS analyses received higher scores.
4. Prior literature evidence (0–3 points): Prior evidence was evaluated according to the extent of human, genetic, clinical, or experimental support linking each protein to cardiomyopathy, heart failure, myocardial remodeling, fibrosis, inflammation, or related cardiac phenotypes.

Proteins with a total score ≥5 were shortlisted for potential clinical validation. Ties were broken by smaller meta-analytic *P*-values. This composite scoring framework was defined for the present study and has not been externally validated.

Clinical Cohort

This was a prospective case–control, exploratory clinical validation study. A total of 81 individuals were prospectively recruited for clinical validation of identified biomarkers, including 48 patients with cardiomyopathy (cases) and 33 healthy controls. Cardiomyopathy subtypes (dilated, hypertrophic, and restrictive) were adjudicated by cardiologists based on comprehensive clinical records and echocardiographic findings, in accordance with contemporary diagnostic criteria. Secondary causes were excluded based on documented clinical assessments and available investigations. Baseline assessment and blood sampling were completed at enrollment (morning phlebotomy prior to any cardiac-specific treatment); no longitudinal follow-up was conducted, and therefore no loss to follow-up occurred. The sample size was feasibility-based; no a priori calculation was performed. Effect sizes (odds ratios) with 95% confidence intervals are reported, and findings require external validation in independent cohorts. Patients were enrolled from December 2022 to May 2025 at a single tertiary center. All participants provided written informed consent, and the study protocol was approved by the Institutional Review Board (approval No. 2022-KY-003; April 2022). Inclusion criteria for cases were a diagnosis of cardiomyopathy (predominantly non-ischemic etiologies, confirmed by imaging and clinical evaluation) prior to initiation of therapy. Controls were volunteers with no history of heart failure or structural heart disease, frequency-matched to cases by age and sex. At enrollment, a detailed medical history and physical examination were obtained. Laboratory personnel were blinded to case–control status. Whole blood (5 mL in EDTA) was processed within 30 minutes of collection; plasma was aliquoted and stored at −80 °F until assays.

All participants underwent standard transthoracic echocardiography using a Sonoscape ultrasound system, performed by experienced sonographers. Key echocardiographic parameters included left ventricular end-diastolic diameter (LVEDD), left atrial diameter (LA), interventricular septal thickness (IVS), ejection fraction (LVEF), fractional shortening (FS), and mitral inflow velocities (E, A) for calculation of the E/A ratio. These measurements were used to assess left ventricular size and systolic function (LVEDD, LVEF, FS), wall thickness (IVS), atrial enlargement (LA), and diastolic filling patterns (E/A ratio). All procedures were conducted in accordance with the Declaration of Helsinki and applicable regulations.

All plasma samples were analyzed using commercially available sandwich ELISA kits for 8 proteins: CD163 (Proteintech, KE00190), LGALS3BP (Proteintech, KE00155), FABP5 (CUSABIO, CSB-EL007946HU), FSTL3 (Thermo Fisher, EHFSTL3), VCAM1 (CUSABIO, CSB-E04753h), PICP (Cloud-Clone, SEA573Hu), Galectin-3 (Proteintech, KE00126), and PIIINP (Cloud-Clone, SEA570Hu). All assays were performed in accordance with the manufacturers' protocols. Briefly, samples and standards were loaded in duplicate into 96-well plates pre-coated with target antibodies, incubated, washed, and then treated with detection antibodies and colorimetric substrate according to the kit instructions. Optical density was read at the specified wavelength, and concentrations were calculated using standard curves generated for each assay. All samples were measured blinded to clinical group.

Receiver Operating Characteristic Analysis and Multivariable Biomarker Model

For the single-center case–control cohort (48 cardiomyopathy cases and 33 controls), plasma concentrations of 5

biomarkers (CD163, LGALS3BP, FABP5, FSTL3, VCAM1) were analyzed. Participants with missing biomarker values were excluded from Receiver operating characteristic (ROC) analyses (complete-case analysis). Each biomarker was standardized to z-scores before modeling.

For individual ROC curves, a univariable logistic regression with case/control status as the outcome was fitted for each biomarker, and the model-based predicted probabilities were used to derive the ROC curve and area under the curve (AUC).

For combined prediction, the 4-protein panel (CD163, LGALS3BP, FABP5, and FSTL3) was selected as the primary combined model, given the modest individual discrimination of VCAM1. To quantify the incremental value of VCAM1, a 5-protein model including VCAM1 was additionally fitted, and compared the discrimination of the 4- and 5-protein models using DeLong's test (Supplementary Table 2). As VCAM1 provided negligible incremental contribution, the 4-protein panel was retained as the primary combined model. Predicted probabilities from the multivariable biomarker model were used to construct the combined ROC curve and AUC; this multivariable analysis reflects mutual adjustment among biomarkers rather than adjustment for clinical covariates.

Internal validation was performed using bootstrap optimism correction (2000 resamples) and repeated 10-fold cross-validation (50 repeats); corresponding performance metrics are reported in Supplementary Table 2. To assess incremental value beyond routinely available information, sensitivity analyses were conducted by adding the 4-protein panel to base models including age, sex, systolic blood pressure, and either LVEF or LVEDD (Supplementary Table 3).

Statistical Analysis

Clinical analyses were performed in R (v4.x) and/or Python (standard glm/scikit-learn routines for logistic regression and ROC computation) and GraphPad Prism 9 (for descriptive statistics, graphics, and correlation analysis). Associations between plasma biomarkers and echocardiographic parameters were assessed using Spearman correlation. Diagnostic performance was evaluated as described above. For the clinical validation analyses only, a 2-sided $P < .05$ was considered statistically significant.

RESULTS

Proteome-Wide Data Integration and Mendelian Randomization Screening

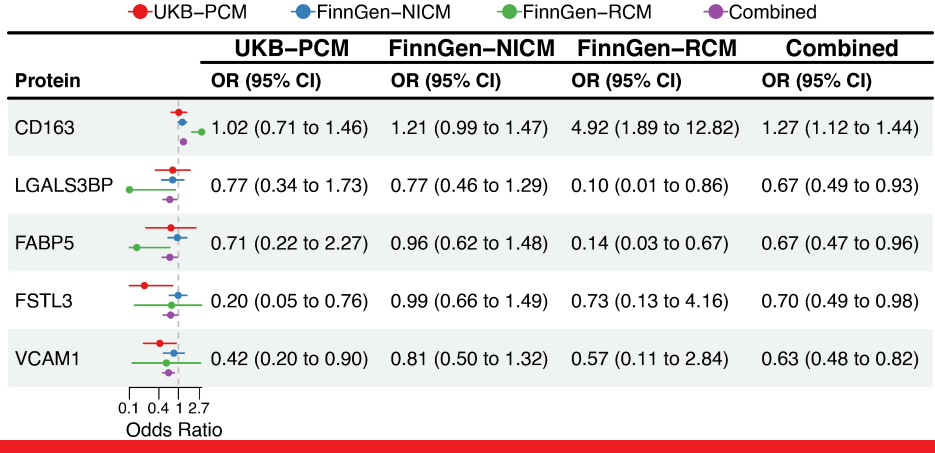
First, pQTL data from 8 large proteomics cohorts were harmonized z(Supplementary Table 4): Fenland (4775 proteins in 10708 participants),[15] Iceland (4719 proteins in 35559 participants),[16,17] UKB-PPP (2922 proteins in 34557 participants),[19] INTERVAL (2995 proteins in 3301 participants),[18] KORA F4 (1124 proteins in 1000 participants),[21] UKB-PPP Phase I (1463 proteins in 35571 participants),[20] SCALLOP (90 proteins in 30931 participants),[22] and FHS (71 proteins in 6861 participants)[23] (Figure 1). Altogether, 5034 proteins (13636 pQTLs: 2157 cis, 11479 trans) were available for downstream MR (Figure 1).

Cardiomyopathy GWAS and Confounder Filtering

Proteome-wide MR was carried out in 3 cardiomyopathy GWAS datasets: primary cardiomyopathy (PCM) in UK Biobank (522 cases, 455826 controls), Non-Ischemic Cardiomyopathy (NICM) in FinnGen R11 (2108 cases, 413400 controls), and restrictive cardiomyopathy (RCM) in FinnGen R11 (114 cases, 492699 controls). Prior to MR, all pQTLs were filtered to remove any variants ($P < 5 \times 10^{-8}$) associated with key confounders—alcohol consumption, smoking status, measures of adiposity, type 2 diabetes, systolic/diastolic blood pressure, and eGFR—thereby minimizing horizontal pleiotropy.

Mendelian Randomization-Derived Protein Associations

Inverse-variance weighted MR (or Wald ratio when only 1 instrument was available) was then performed in each GWAS, and the 3 sets of estimates were combined by fixed-effects meta-analysis.



Protein	UKB–PCM OR (95% CI)	FinnGen–NICM OR (95% CI)	FinnGen–RCM OR (95% CI)	Combined OR (95% CI)
CD163	1.02 (0.71 to 1.46)	1.21 (0.99 to 1.47)	4.92 (1.89 to 12.82)	1.27 (1.12 to 1.44)
LGALS3BP	0.77 (0.34 to 1.73)	0.77 (0.46 to 1.29)	0.10 (0.01 to 0.86)	0.67 (0.49 to 0.93)
FABP5	0.71 (0.22 to 2.27)	0.96 (0.62 to 1.48)	0.14 (0.03 to 0.67)	0.67 (0.47 to 0.96)
FSTL3	0.20 (0.05 to 0.76)	0.99 (0.66 to 1.49)	0.73 (0.13 to 4.16)	0.70 (0.49 to 0.98)
VCAM1	0.42 (0.20 to 0.90)	0.81 (0.50 to 1.32)	0.57 (0.11 to 2.84)	0.63 (0.48 to 0.82)

Figure 2. Forest plots of odds ratios for candidate proteins across 3 cardiomyopathy GWAS datasets. Forest plots showing mendelian randomization-derived odds ratios (ORs) and 95% confidence intervals (CIs) for each candidate protein (CD163, LGALS3BP, FABP5, FSTL3, VCAM1) across UKB–Primary Cardiomyopathy (PCM), FinnGen Non-ischemic Cardiomyopathy (NICM), FinnGen Restrictive Cardiomyopathy (RCM), and the combined meta-analysis. Directionally consistent associations and pooled effects support the potential causal relevance of these proteins to cardiomyopathy risk.

Overall, proteome-wide MR prioritized 27 proteins with scores ≥5 (Supplementary Table 1). From this set, 5 proteins were selected for further validation—CD163, LGALS3BP, FABP5, FSTL3, and VCAM1—based on prior evidence implicating them in cardiomyopathy biology. In fixed-effects meta-analysis of MR estimates across primary cardiomyopathy, non-ischemic cardiomyopathy, and restrictive cardiomyopathy GWAS, the pooled odds ratios per 1-SD increase in genetically predicted protein level were as follows: CD163, 1.27 (95% CI, 1.12-1.44); LGALS3BP, 0.67 (95% CI, 0.49-0.93); FABP5, 0.67 (95% CI, 0.47-0.96); FSTL3, 0.70 (95% CI, 0.49-0.98); and VCAM1, 0.63 (95% CI, 0.48-0.82) (Figure 2).

Clinical Enzyme-Linked Immunosorbent Assays Validation Cohort Characteristics

From December 2022 to May 2025, 81 participants were prospectively recruited at a single tertiary care hospital. All participants provided written informed consent prior to enrollment, and the study protocol was approved by the institutional ethics committee (IRB no. 2022-KY-003). A total of 48 patients diagnosed with primary cardiomyopathy (CM) and 33 age- and sex-matched healthy controls were included. Demographic characteristics such as age, sex, BMI, blood pressure, smoking and drinking status, diabetes prevalence, and eGFR did not significantly differ between groups (Table 1).

Among cardiomyopathy patients (n = 48), 28 were diagnosed with dilated cardiomyopathy, 14 with restrictive cardiomyopathy, and 6 with hypertrophic cardiomyopathy. Key echocardiographic measurements by subtype are summarized in Supplementary Table 5.

All participants underwent transthoracic echocardiography using a Sonoscape system. The following echocardiographic parameters were recorded: LVEDD, LA, IVS, LVEF, FS, and mitral inflow velocities (E and A) for E/A ratio calculation. Compared with controls, CM patients exhibited markedly impaired cardiac structure and function, including increased LVEDD (58.4 ± 10.7 mm vs. 49.6 ± 2.8 mm, $P < .0001$), enlarged LA (42.0 ± 3.6 mm vs. 34.5 ± 2.7 mm, $P < .0001$), thickened IVS (10.6 ± 2.4 mm vs. 9.0 ± 0.9 mm, $P = .0001$), reduced LVEF (50.3 ± 13.1% vs. 64.1 ± 4.5%, $P < .0001$), decreased FS (23.2 ± 7.7% vs. 34.3 ± 3.6%, $P < .0001$), and a non-significant trend toward higher E/A ratio (1.5 ± 1.1 vs. 1.2 ± 0.2, $P = .0553$).

Plasma concentrations of 8 biomarkers were quantified by sandwich ELISA (Table 1). CM patients exhibited significantly higher levels of inflammatory and fibrotic markers—CD163 (339.7 ± 122.2 ng/mL vs. 143.4 ± 77.7 ng/mL; $P < .0001$), PICP (5.4 ± 1.0 ng/mL vs. 3.9 ± 1.0 ng/mL; $P < .0001$), PIIINP (71.6 ± 12.4 ng/mL vs. 41.4 ± 5.5 ng/mL; $P < .0001$), and Gal-3 (1.9 ± 0.4 ng/mL vs. 1.2 ± 0.4 ng/mL; $P < .0001$)—and significantly lower levels of candidate proteins LGALS3BP (3.1 ± 0.5 μg/mL vs. 6.4 ± 1.0 μg/mL; $P < .0001$), FABP5 (3.8 ± 2.0 ng/mL vs. 5.1 ± 2.0

Table 1. Baseline Characteristics, Plasma Biomarkers, and Echocardiographic Parameters of Study Participants

Characteristics	Subjects [Mean ± SD or n (%)] Healthy Controls (n = 33) (%)	Cardiomyopathy Patients (n = 48) (%)	*P*
Age (years)	50.2 ± 7.7	53.0 ± 11.0	.218
Male (%)	18/33 (54.6)	25/48 (52.1)	>.9999
BMI (kg/m^2)	24.6 ± 2.7	26.0 ± 3.6	.0613
SBP (mm Hg)	123.3 ± 9.8	126.9 ± 11.2	.1385
DBP (mm Hg)	79.3 ± 6.7	81.3 ± 8.0	.234
Smoking (%)	9 (27.27)	20 (41.67)	.2402
Drinking (%)	12 (36.36)	16 (33.33)	.8153
Diabetes (%)	2 (6.06)	6 (12.50)	.462
eGFR (mL/min)	92.0 ± 7.7	90.0 ± 13.5	.442
Plasma markers			
CD163 (ng/mL)	143.4 ± 77.7	339.7 ± 122.2	<.0001
LGALS3BP (μg/mL)	6.4 ± 1.0	3.1 ± 0.5	<.0001
FABP5 (ng/mL)	5.1 ± 2.0	3.8 ± 2.0	.0022
FSTL3 (ng/mL)	9.4 ± 1.9	8.0 ± 1.8	.0007
VCAM1 (ng/mL)	649.1 ± 131.7	558.2 ± 232.9	.046
PICP (ng/mL)	3.9 ± 1.0	5.4 ± 1.0	<.0001
PIIINP (ng/mL)	41.4 ± 5.5	71.6 ± 12.4	<.0001
Gal-3 (ng/mL)	1.2 ± 0.4	1.9 ± 0.4	<.0001
Echocardiographic parameters			
LVEF (%)	64.1 ± 4.5	50.3 ± 13.1	<.0001
FS (%)	34.3 ± 3.6	23.2 ± 7.7	<.0001
LVEDD (mm)	49.6 ± 2.8	58.4 ± 10.7	<.0001
LA (mm)	34.5 ± 2.7	42.0 ± 3.6	<.0001
IVS (mm)	9.0 ± 0.9	10.6 ± 2.4	.0001
E/A ratio	1.2 ± 0.2	1.5 ± 1.1	.0553

BMI, body mass index; DBP, diastolic blood pressure; E/A, ratio of early to late mitral inflow velocity; eGFR, estimated glomerular filtration rate; FS, fractional shortening; IVS, interventricular septal thickness at end-diastole measured on routine transthoracic echocardiography; LA, left atrial diameter; LVEDD, left ventricular end-diastolic diameter; LVEF, left ventricular ejection fraction; SBP, systolic blood pressure.

Table 2. Association of Plasma Protein Levels with Cardiomyopathy Risk: Univariate and Multivariate Logistic Regression <.001

Protein	Univariate OR	Univariate 95% CI	Univariate *P*	Multivariate OR	Multivariate 95% CI	Multivariate *P*
CD163	1.019	1.01–1.03	<.001	1.044	1.01–1.08	.012
LGALS3BP	0.390	0.27–0.56	<.001	0.161	0.04–0.63	.009
FABP5	0.716	0.56–0.92	.008	0.488	0.22–1.10	.083
FSTL3	0.645	0.49–0.85	.002	1.700	0.72–4.00	.224
VCAM1	0.998	1.00–1.00	.050	0.994	0.99–1.00	.131

CI, confidence interval; OR, odds ratio.

ng/mL; $P = .0022$), FSTL3 (8.0 ± 1.8 ng/mL vs. 9.4 ± 1.9 ng/mL; $P = .0007$), and VCAM1 (558.2 ± 232.9 ng/mL vs. 649.1 ± 131.7 ng/mL; $P = .0460$).

Clinical Validation of Protein Biomarkers

In the ELISA-validated cohort, univariate logistic regression showed that higher CD163 levels were associated with increased odds of cardiomyopathy [OR (odds ratio) 1.019; 95% CI, 1.010-1.030; $P = 1.49 \times 10^{-6}$], whereas higher LGALS3BP (OR 0.390; 95% CI, 0.270-0.560; $P = 3.92 \times 10^{-7}$), FABP5 (OR 0.716; 95% CI, 0.560-0.920; $P = .0077$), FSTL3 (OR 0.645; 95% CI, 0.490-0.850; $P = .0018$), and VCAM1 (OR 0.998; 95% CI, 1.000-1.000; $P = .0504$) levels were protective (Table 2). When CD163, LGALS3BP, FABP5, FSTL3, and VCAM1 were entered together in a multivariate model, only CD163 (OR 1.044, 95% CI 1.010-1.078; $P = .0116$) and LGALS3BP (OR 0.161, 95% CI 0.040-0.625; $P = .0087$) remained independent predictors of cardiomyopathy (Table 2). Spearman correlation analysis showed that CD163 and LGALS3BP were associated with several echocardiographic indices of cardiac structure and function, whereas FABP5 and VCAM1 showed weaker or limited correlations (Figure 3A). Receiver-operating characteristic analysis further confirmed CD163's superior discrimination (AUC = 0.917) and LGALS3BP's strong performance (AUC = 0.870), with FSTL3 (AUC = 0.716) and FABP5 (AUC = 0.681) showing moderate accuracy and VCAM1 being less discriminative (AUC = 0.632). Figure 3B shows the ROC curves for CD163, LGALS3BP, FABP5, FSTL3, and the combined four-protein panel, which achieved an apparent AUC of 0.993. Given the limited incremental value of VCAM1, the primary 4-protein combined panel (CD163, LGALS3BP, FABP5, FSTL3) were compared with the 5-protein model including VCAM1; the apparent AUCs were similar (0.993 vs 0.991; $P = .523$). Internal validation indicated optimism-corrected and repeated cross-validated AUCs of 0.988 and 0.980 for the 4-protein panel (Supplementary Table 2). In pre-specified sensitivity analyses adjusting for age, sex, systolic blood pressure, and either LVEF or LVEDD, addition of the 4-protein panel improved discrimination compared with covariates alone (Supplementary Table 3).

DISCUSSION

In this study, large-scale proteomic data from 8 cohorts including more than 120 000 participants and 5 034 circulating proteins were harmonized and integrated, followed by rigorous MR analyses across 3 cardiomyopathy GWAS datasets totaling over 460 000 individuals. After stringent filtering of confounding variants, MHC region exclusion, and LD clumping, 27 proteins with genetically informed associations with cardiomyopathy risk were informed. Five proteins (CD163, LGALS3BP, FABP5, FSTL3, VCAM1), selected based on robust biological plausibility and assay availability, underwent detailed clinical validation in the prospectively enrolled cohort of 81 subjects. In panel refinement analyses with internal validation, VCAM1 provided negligible incremental discrimination; therefore, a parsimonious 4-protein panel was adopted as the primary model. The 4-protein panel showed high apparent discrimination (AUC = 0.993), and bootstrap optimism correction and repeated cross-validation

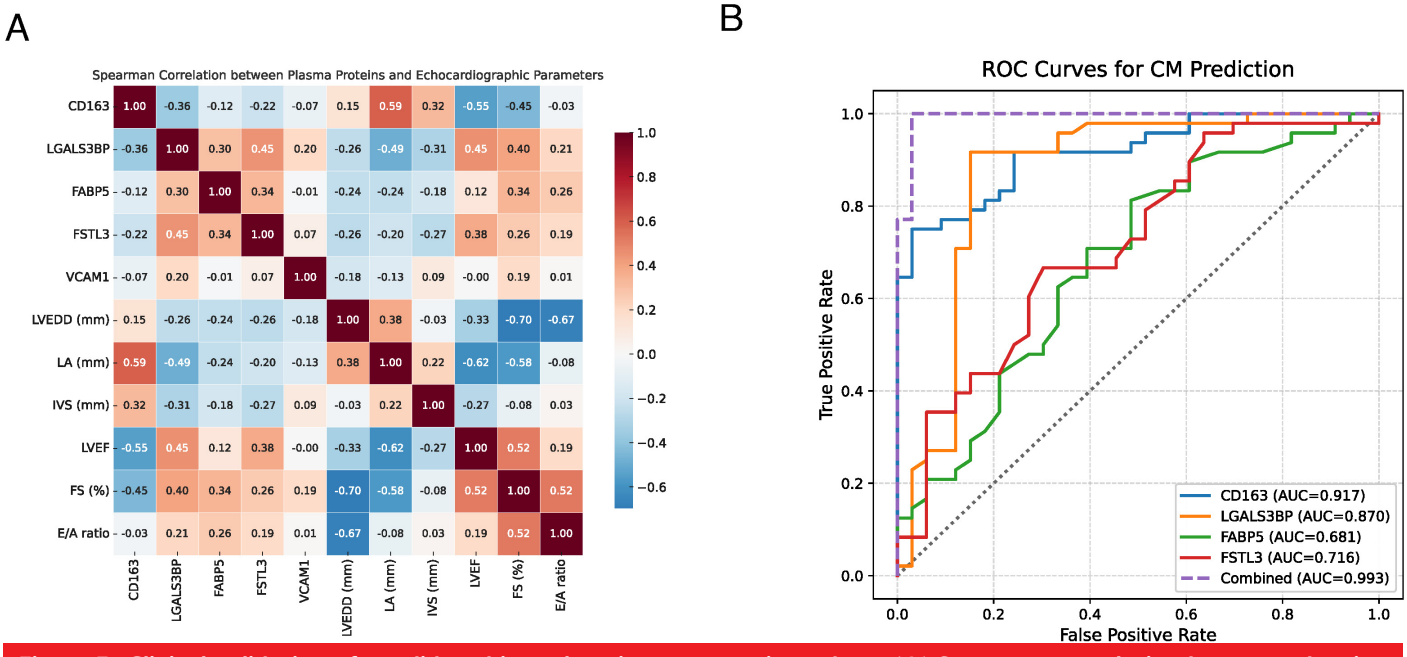


Figure 3. Clinical validation of candidate biomarkers in a prospective cohort. (A) Spearman correlation heatmap showing associations between plasma biomarker levels (CD163, LGALS3BP, FABP5, FSTL3, VCAM1) and echocardiographic parameters (LVEDD, LA, IVS, LVEF, FS, E/A ratio). Red indicates positive correlations; blue indicates negative correlations. (B) Receiver operating characteristic (ROC) curves for the 4 individual proteins (CD163, LGALS3BP, FABP5, and FSTL3) and the combined 4-protein panel. The combined model achieved an apparent (in-sample) AUC of 0.993; given the modest sample size and case–control design, apparent performance may be optimistic. Optimism-corrected and repeated cross-validated AUCs are reported in Supplementary Table 2, and external validation is warranted.

supported its internal performance. However, given the modest sample size and lack of external validation, this result should still be interpreted cautiously.

The study emphasizes CD163 and LGALS3BP as particularly promising biomarkers. CD163, a macrophage-specific scavenger receptor, is widely recognized as a marker of macrophage activation and inflammation, playing a significant role in myocardial remodeling and fibrosis in various cardiovascular conditions.[24] The data, consistent with previous clinical findings, indicate elevated CD163 levels in cardiomyopathy patients correlating strongly with echocardiographic markers of cardiac dysfunction, such as reduced LVEF. Conversely, LGALS3BP has primarily been studied in oncology as a regulator of immune response and extracellular matrix remodeling, but its cardioprotective role in the results, demonstrated by reduced plasma levels and a protective MR effect, suggests an intriguing and previously unrecognized mechanism potentially opposing myocardial fibrosis and remodeling.[25,26]

FABP5 and FSTL3 showed moderate associations with cardiomyopathy in the analyses. FABP5, involved in fatty acid transport and metabolism within cardiomyocytes, has been implicated in metabolic dysregulation contributing to cardiac hypertrophy and dysfunction.[2728] Similarly, FSTL3, known to modulate TGF-β signaling, has been associated with cardiac hypertrophic responses and fibrosis, supporting its relevance as a potential biomarker in cardiac remodeling.[29,30] Although their associations were relatively moderate, both proteins likely reflect distinct but complementary pathogenic pathways contributing to cardiomyopathy. VCAM1 exhibited the weakest association in the study. This adhesion molecule, previously linked with endothelial dysfunction and inflammation in cardiovascular diseases, may reflect broader vascular and inflammatory processes rather than specific myocardial pathology, explaining its relatively lower discriminative capacity.[31-33] Further research in larger and diverse cohorts is warranted to clarify VCAM1's clinical utility and precise role in cardiomyopathy.

Beyond these biological insights, the findings also carry potential clinical implications. In routine practice, establishing an accurate diagnosis of cardiomyopathy can be challenging, and commonly used markers such as NT-proBNP and cardiac troponins are not specific for cardiomyopathy.[34,35] The 4-protein panel—particularly CD163 and LGALS3BP—showed strong discrimination between cardiomyopathy patients and healthy controls in the cohort, supporting its potential as an adjunct to standard diagnostic evaluation. Further validation in larger, independent cohorts will be essential to confirm generalizability and to define its incremental value over routinely available clinical and echocardiographic information.

The MR-driven identification process, integrating extensive proteomic and genetic data with targeted clinical validation, significantly reduces biases and enhances translational potential compared to previous biomarker studies lacking genetic causality support.

Several limitations should be acknowledged. First, the clinical validation was conducted in a relatively small, single-center case–control cohort with a feasibility-based sample size, and no a priori power calculation was performed. Although internal validation using bootstrap optimism correction and repeated cross-validation was performed, the very high apparent AUC of the 4-protein panel should be interpreted with caution, as some degree of model optimism or overfitting cannot be fully excluded in the absence of external validation. Therefore, validation in an independent cohort will be essential to confirm the reproducibility and real-world clinical utility of this biomarker panel. Second, the cohort reflects real-world referrals and may include mild or borderline phenotypes, and cardiomyopathy subtyping relied on clinical records and echocardiography without systematic CMR or genetic testing, introducing potential phenotype misclassification. Third, while extensive instrument filtering was applied, residual horizontal pleiotropy and other sources of bias cannot be fully excluded in MR analyses. In addition, no MR association survived FDR correction, and therefore the proteome-wide MR results should be interpreted as hypothesis-generating, with an increased risk of false-positive prioritization that requires independent replication. Fourth, candidate selection and panel construction were constrained by assay availability, and the cross-sectional design and sample size precluded prognostic assessment and reliable subtype-stratified modeling; therefore, subtype-specific and longitudinal inferences require future studies.

In conclusion, the integrated MR-to-clinical pipeline prioritized circulating protein candidates and identified a 4-protein biomarker panel with promising diagnostic discrimination for primary cardiomyopathy in this exploratory cohort. Given the very high apparent AUC and the lack of external validation, these findings should be considered preliminary and require confirmation in larger, independent cohorts.

AI Disclosure: Large Language Models (ChatGPT) were used only for language polishing of non-substantive text. No AI was used for data analysis or figure generation. All content was verified by the authors.

Ethics Committee Approval: This study was approved by the Ethics Committee of Second People's Hospital of Jintang County, Chengdu, China, (Approval No: 2022-KY-003; Date: April 6, 2022).

Informed Consent: Written informed consent was obtained from all participants before enrollment.

Peer-review: Externally peer-reviewed.

Acknowledgments: The authors thank the Fenland, Iceland, INTERVAL, UKB-PPP, KORA, SCALLOP, and FHS studies for providing proteomics data. Gratitude is extended to all participants, institutions, and staff involved in these studies for their contributions to this research. Appreciation is also extended to all cited studies for sharing data, software, and platforms.

Author Contributions: Concept – Z.Y., J.H.; Design – Z.Y., J.H.; Supervision – Z.Y., J.H.; Resources – J.H., X.Y.L., J.T., C.T., X.F.L.; Materials – J.H., X.Y.L., J.T., C.T., X.F.L.; Data Collection and/or Processing – Z.Y., J.H., X.Y.L., J.T., C.T., X.F.L.; Analysis and/or

Interpretation – J.H., Z.Y., X.Y.L.; Literature Search – J.H., Z.Y., X.Y.L.; Writing – J.H., Z.Y.; Critical Review – J.H., X.Y.L., Z.Y.

Declaration of Interests: None of the authors has any conflicts of interest or serves as a member of the journal's Editorial Board or Advisory Board.

Funding: This work was supported by the Chengdu Municipal Health Commission (Grant No. 2022629).

REFERENCES

1. Kaski JP, Arbelo E. The 2023 ESC guidelines for the management of cardiomyopathies: the 10 commandments. *Eur Heart J*. 2024;45(13):1101. [CrossRef]
2. Karuna N, Tonry C, Ledwidge M, et al. Proteomic-based biomarker discovery reveals panels of diagnostic biomarkers for early identification of heart failure subtypes. *J Transl Med*. 2025;23(1):546. [CrossRef]
3. Nazir A, Nazir A, Afzaal U, et al. Advancements in biomarkers for early detection and risk stratification of cardiovascular diseases-A literature review. *Health Sci Rep*. 2025;8(5):e70878. [CrossRef]
4. Wang Y, Jia H, Song J. Accurate classification of non-ischemic cardiomyopathy. *Curr Cardiol Rep*. 2023;25(10):1299-1317. [CrossRef]
5. Ciarambino T, Menna G, Sansone G, et al. Cardiomyopathies: an overview. *Int J Mol Sci*. 2021;22(14):7722. [CrossRef]
6. Maruyama K, Imanaka-Yoshida K. The pathogenesis of cardiac fibrosis: a review of recent progress. *Int J Mol Sci*. 2022;23(5):2617. [CrossRef]
7. Kong P, Christia P, Frangogiannis NG. The pathogenesis of cardiac fibrosis. *Cell Mol Life Sci*. 2014;71(4):549-574. [CrossRef]
8. Pan K-L, Hsu Y-C, Chang S-T, et al. The role of cardiac fibrosis in diabetic cardiomyopathy: from pathophysiology to clinical diagnostic tools. *Int J Mol Sci*. 2023;24(10):8604. [CrossRef]
9. Stege NM, de Boer RA, van den Berg MP, et al. The time has come to explore plasma biomarkers in genetic cardiomyopathies. *Int J Mol Sci*. 2021;22(6):2955. [CrossRef]
10. Turinay Ertop ZŞ, Aslan AN, Neşelioğlu S, et al. Thiol/disulfide homeostasis: A new oxidative marker in heart failure patients with preserved ejection fraction. *Anatol J Cardiol*. 2024;28(8):406-412. [CrossRef]
11. Örsçelik Ö, Yeşil E, Uyar H, et al. Relationship between urinary angiotensinogen and mortality in patients with heart failure with reduced ejection fraction. *Anatol J Cardiol*. 2023;27(7):417-422. [CrossRef]
12. Si S, Liu H, Xu L, et al. Identification of novel therapeutic targets for chronic kidney disease and kidney function by integrating multi-omics proteome with transcriptome. *Genome Med*. 2024;16(1):84. [CrossRef]
13. Sun J, Zhao J, Jiang F, et al. Identification of novel protein biomarkers and drug targets for colorectal cancer by integrating human plasma proteome with genome. *Genome Med*. 2023;15(1):75. [CrossRef]
14. Raimoglou D, İzgi C, Enar R, et al. Structural and functional impact of Adrenoceptor Beta-1 gene polymorphism in patients with hypertrophic cardiomyopathy and response to beta-blocker therapy. *Anatol J Cardiol*. 2024;28(3):150-157. [CrossRef]
15. Pietzner M, Wheeler E, Carrasco-Zanini J, et al. Mapping the proteo-genomic convergence of human diseases. *Science*. 2021;374(6569):eabj1541. [CrossRef]
16. Eldjarn GH, Ferkingstad E, Lund SH, et al. Large-scale plasma proteomics comparisons through genetics and disease associations. *Nature*. 2023;622(7982):348-358. [CrossRef]
17. Ferkingstad E, Sulem P, Atlason BA, et al. Large-scale integration of the plasma proteome with genetics and disease. *Nat Genet*. 2021;53(12):1712-1721. [CrossRef]
18. Sun BB, Maranville JC, Peters JE, et al. Genomic atlas of the human plasma proteome. *Nature*. 2018;558(7708):73-79. [CrossRef]
19. Sun BB, Chiou J, Traylor M, et al. Plasma proteomic associations with genetics and health in the UK Biobank. *Nature*. 2023;622(7982):329-338. [CrossRef]
20. Sun BB, Chiou J, Traylor M, et al. Genetic regulation of the human plasma proteome in 54,306 UK Biobank participants. *bioRxiv*. 2022:2022.06.17.496443. [CrossRef]
21. Suhre K, Arnold M, Bhagwat AM, et al. Connecting genetic risk to disease end points through the human blood plasma proteome. *Nat Commun*. 2017;8:14357. [CrossRef]
22. Folkersen L, Gustafsson S, Wang Q, et al. Genomic and drug target evaluation of 90 cardiovascular proteins in 30,931 individuals. *Nat Metab*. 2020;2(10):1135-1148. [CrossRef]
23. Yao C, Chen G, Song C, et al. Genome-wide mapping of plasma protein QTLs identifies putatively causal genes and pathways for cardiovascular disease. *Nat Commun*. 2018;9(1):3268. [CrossRef]
24. Skytthe MK, Graversen JH, Moestrup SK. Targeting of CD163+ macrophages in inflammatory and malignant diseases. *Int J Mol Sci*. 2020;21(15):5497. [CrossRef]
25. Capone E, Iacobelli S, Sala G. Role of galectin 3 binding protein in cancer progression: a potential novel therapeutic target. *J Transl Med*. 2021;19(1):405. [CrossRef]
26. Kim D-H, Sung M, Park M-S, et al. Galectin 3-binding protein (LGALS3BP) depletion attenuates hepatic fibrosis by reducing transforming growth factor-β1 (TGF-β1) availability and inhibits hepatocarcinogenesis. *Cancer Commun (Lond)*. 2024;44(10):1106-1129. [CrossRef]
27. Hotamisligil GS, Bernlohr DA. Metabolic functions of FABPs--mechanisms and therapeutic implications. *Nat Rev Endocrinol*. 2015;11(10):592-605. [CrossRef]
28. Storch J, Corsico B. The multifunctional family of mammalian fatty acid-binding proteins. *Annu Rev Nutr*. 2023;43:25-54. [CrossRef]
29. Tian S, Xu X, Yang X, et al. Roles of follistatin-like protein 3 in human non-tumor pathophysiologies and cancers. *Front Cell Dev Biol*. 2022;10:953551. [CrossRef]
30. Ghazal R, Wang M, Liu D, et al. Cardiac fibrosis in the multiomics era: implications for heart failure. *Circ Res*. 2025;136(7):773-802. [CrossRef]
31. Kaur R, Singh V, Kumari P, et al. Novel insights on the role of VCAM-1 and ICAM-1: potential biomarkers for cardiovascular diseases. *Ann Med Surg (Lond)*. 2022;84:104802. [CrossRef]
32. Jager A, van Hinsbergh VW, Kostense PJ, et al. Increased levels of soluble vascular cell adhesion molecule 1 are associated with risk of cardiovascular mortality in type 2 diabetes: the Hoorn study. *Diabetes*. 2000;49(3):485-491. [CrossRef]
33. Gustavsson C, Agardh C-D, Zetterqvist AV, et al. Vascular cellular adhesion molecule-1 (VCAM-1) expression in mice retinal vessels is affected by both hyperglycemia and hyperlipidemia. *PLoS One*. 2010;5(9):e12699. [CrossRef]
34. Spadotto A, Morabito D, Carecci A, et al. The Challenges of Diagnosis and Treatment of arrhythmogenic cardiomyopathy: are We there yet? *Rev Cardiovasc Med*. 2022;23(8):283. [CrossRef]
35. Samanidis G. Current challenges in diagnosis and treatment of cardiovascular disease. *J Pers Med*. 2024;14(8):786. [CrossRef]

Supplementary Table 1. Candidate Protein Prioritization

Protein	Direction-concordance score	Effect-magnitude score	Nominal-evidence score	Prior-evidence score	Total priority score	Selected for ELISA evaluation
CD163	2	2	0	3	7	Yes
LGALS3BP	2	2	0	3	7	Yes
FABP5	2	2	0	2	6	Yes
PTPRR	2	2	2	0	6	No
METTL1	2	2	1	0	5	No
GNRH1	2	2	1	0	5	No
ANXA2	2	0	0	3	5	No
PSAP	2	1	0	2	5	No
SDCBP2	2	2	1	0	5	No
DOS	2	2	1	0	5	No
RCOR1	2	2	1	0	5	No
KRT20	2	2	1	0	5	No
HMBS	2	2	1	0	5	No
HAGH	2	2	1	0	5	No
IL18BP	2	2	1	0	5	No
SH2D1B	2	2	1	0	5	No
PFDN2	2	2	1	0	5	No
DNAJB12	2	2	1	0	5	No
SERPINB4	2	2	1	0	5	No
CCL20	2	2	1	0	5	No
CADM3	2	2	1	0	5	No
CD33	2	2	1	0	5	No
VCAM1	2	2	1	0	5	Yes
ZNF275	1	2	2	0	5	No
FSTL3	2	2	1	0	5	Yes
DDHD2	2	2	1	0	5	No
DNAJB8	2	2	1	0	5	No

Note: The composite score was used as an exploratory heuristic tool for candidate prioritization and not for formal statistical inference. Among the 27 proteins meeting the score threshold, five proteins were selected for ELISA evaluation after additional consideration of biological relevance, complementary pathway representation, commercial assay availability, and clinical feasibility.

Supplementary Table 2. Internal validation of the biomarker panel models

Model	Predictors	Apparent AUC	Optimism-corrected AUC	95% CI	Repeated 10-fold CV AUC	95% range
Four-protein panel (primary)	CD163, LGALS3BP, FABP5, FSTL3	0.993	0.988	0.973–1.000	0.98	0.975–0.985
Five-protein panel (sensitivity)	CD163, LGALS3BP, FABP5, FSTL3, VCAM1	0.991	0.985	0.972–1.000	0.98	0.973–0.985

AUCs were estimated using logistic regression. Internal validation was performed using bootstrap optimism correction (B = 2000 resamples) and repeated stratified 10-fold cross-validation (50 repeats). DeLong test comparing the four- vs five-protein panels: $\Delta AUC = 0.002$, $P = .523$.

Supplementary Table 3. Sensitivity analyses adjusting for age, sex, systolic blood pressure, LVEF and LVEDD

Model	Predictors	Apparent AUC	Optimism-corrected AUC	95% CI	Repeated 10-fold CV AUC	95% range	ΔAUC vs covariates-only	DeLong *P*-value
(A) Sensitivity analyses adjusting for age, sex, systolic blood pressure, and LVEF								
Covariates-only	age, sex, SBP, LVEF	0.878	0.855	0.783–0.934	0.843	0.827–0.855		
Covariates + four-protein panel	age, sex, SBP, LVEF + (CD163, LGALS3BP, FABP5, FSTL3)	0.997	0.993	0.984–1.000	0.988	0.983–0.991	0.119	0.002
(B) Sensitivity analyses adjusting for age, sex, systolic blood pressure, and LVEDD								
Covariates-only	age, sex, SBP, LVEDD	0.826	0.797	0.714–0.889	0.788	0.769–0.803		
Covariates + four-protein panel	age, sex, SBP, LVEDD + (CD163, LGALS3BP, FABP5, FSTL3)	0.996	0.99	0.979–1.000	0.985	0.978–0.990	0.169	0

Models were fitted using logistic regression. Bootstrap optimism correction used B=2000 resamples. Repeated cross-validation used stratified 10-fold CV with 50 repeats; the reported range corresponds to the 2.5th and 97.5th percentiles across repeats. ΔAUC and p-values are based on DeLong tests comparing apparent AUCs.

Supplementary Table 4. Summary Information of Included Proteomics Studies

Study	Platform	Ancestries	Number of proteins	Number of pQTLs	Sample size
Fenland	SomaLogic	European	4,775	8,328	10,708
Iceland	SomaLogic	European	4,719	18,084	35,559
INTERVAL	SomaLogic	European	2,995	1,927	3,301
UKB-PPP	Olink	European	2,922	14,287	34,557
UKB-PPP phase I	Olink	European	1,463	10,248	35,571
KORA F4	SomaLogic	European	1,124	539	1,000
SCALLOP	Olink	European	90	451	30,931
FHS	xMAP	European	71	16,602	6,861

Supplementary Table 5. Key echocardiographic measurements by cardiomyopathy subtype

Subtype	n	LVEF	LVEDD	IVS	LA	E/A
DCM	28	42.4 ± 9.4	66.3 ± 6.0	9.8 ± 1.1	42.2 ± 3.1	0.75 ± 0.16
RCM	14	58.7 ± 9.3	48.4 ± 2.9	9.9 ± 1.0	43.6 ± 3.4	3.02 ± 0.82
HCM	6	67.7 ± 3.0	44.5 ± 2.3	16.2 ± 1.1	37.6 ± 3.4	1.49 ± 0.33

IVS denotes interventricular septal thickness at end-diastole, obtained from routine transthoracic echocardiography.