Cardiology Research and Cardiovascular Medicine
Lin S, et al. Cardiolog Res Cardiovasc Med 8: 248.
www.doi.org/10.29011/2575-7083.100248
www.gavinpublishers.com

OPEN ACCESS

GAVIN PUBLISHERS

Research Article

Unraveling Novel Therapeutic Targets for Coronary Artery Disease Via Gene Signatures of Plasma Proteome: a Mendelian Randomization Study

Shuyang Lin[1#], Hongming Shang[2#], Ihab Hassanieh[3*]

[1]Division of Hematology, Department of Medicine, Washington University School of Medicine in St Louis, 63110, MO, USA.

[2]Department of Medical Oncology, Dana-Farber Cancer Institute, Harvard Medical School, 02215, MA, USA.

[3]Division of Hospital Medicine, Department of Medicine, Washington University School of Medicine in St Louis, 63110, MO, USA.

[#]contributed equally

***Corresponding author:** Ihab Hassanieh, Division of Hospital Medicine, Department of Medicine, Washington University School of Medicine in St Louis, 63110, MO, USA.

Citation: Lin S, Shang H, Hassanieh I (2024) Unraveling Novel Therapeutic Targets for Coronary Artery Disease Via Gene Signatures of Plasma Proteome: a Mendelian Randomization Study. Cardiol Res Cardio vasc Med 9:248. DOI:https://doi.org/10.29011/2575-7083.100248

Received Date: 08 May, 2024; **Accepted Date:** 14 May, 2024; **Published Date:** 16 May, 2024

Abstract

Background: Coronary Artery Disease (CAD) is a common disease with a significant healthcare burden across the globe. Efforts were made to identify therapeutic targets or prophylaxis for high-risk individuals. **Method:** Here, we used a plasma-proteome-wide Mendelian Randomization approach to estimate the causal effect of plasma proteins on the development of CAD by using the pQTL of the proteins as the exposure, and two large-scale GWAS of CAD as the outcome. **Results:** We identified LPA, PCSK9, PLA2G7, C1S, C1R, ENO2, SNAP25, A1BG, and CA11 as risk proteins casually associated with the onset of CAD. C1S/R and ENO2 were prioritized as the first-tier targets. We further applied pathway enrichment analysis and RNA-seq data to show that aberrant activation of the complement system and glycolysis are casually associated with increased risk for CAD. **Conclusion:** In short, our study revealed novel therapeutic targets for the treatment of CAD, inviting further investigation into these targets and related pathways.

Keywords: Mendelian Randomization, Coronary Artery Disease, Plasma Proteomics, Precision Medicine, Complement System

Introduction

Coronary artery disease (CAD) imposes a significant epidemiological burden as one of the leading causes of mortality globally. Despite advances in pharmacological and procedural intervention, the prevention, diagnosis, and management of CAD remain challenging for healthcare providers. Darapaladib (PLA2G7 inhibitor) and PCI intervention (COURAGE trial) do not significantly reduce cardiovascular events including myocardial infarction and stroke in CAD patients [1,2]. The FOURIER trial demonstrated that PCSK9 inhibitor evolocumab significantly reduced the risk of cardiovascular events but is indicated only in patients with severe risk [3,4]. Thus, there is a pressing need for novel target-based therapies for CAD. Traditional approaches to target discovery have been rely on omic-scale sequencing. While they provide rigorous insights into potential avenues for treatment, they are also associated with difficulties in sample acquisition, preparation, and processing as well as high costs. Here, we proposed an in silico approach that combines plasma proteomics and Mendelian Randomization (MR) to prioritize potential targets for the treatment of CAD. MR is an increasingly popular epidemiological tool that leverages Genome-Wide Association Study (GWAS) data to infer the causal genetic variants underlying the phenotype of interest. The proteome is the central mediator of the etiology and pathophysiology of diseases. Plasma proteome is of particular interest for cardiovascular diseases due to their

involvement in the circulatory system and thus serves as the ideal candidates for omic-scale analyses [5,6]. Here, we used the gene signatures of proteins (protein quantitative trait loci or pQTL) as the instrumental variables to identify potential therapeutic targets for CAD through MR.

Methods

Mendelian Randomization

The main workflow of our study is outlined in (Figure 1A). *TwoSampleMR* package (version 0.5.9) was used to perform the MR analysis by following the standard protocol and as previously described [7,8]. Briefly, we first downloaded the GWAS data from Aragam et al, the largest CAD GWAS to date (discovery cohort), and Sun et al., the newest plasma proteome GWAS from UKB [9,10]. Three assumptions must be satisfied for MR. The relevance assumption assumes that the single-nucleotide polymorphisms (SNPs), which are the instrumental variables used for the MR analysis, are significantly associated with the exposure. To this end, we first extracted the SNPs of pQTLS that have a P-value of being smaller than 5×10^{-6}, this is to ensure that the SNPs are genetically significant. We then clumped the SNPs with the parameter $R^2<0.001$, and kb=10,000. This is to ensure that only one SNP from a defined chromatin region is extracted to ensure that the laws of independent assortment are obeyed. Furthermore, we computed the F-statistics using the formul$F = (\frac{R^2}{1-R^2})(\frac{n-k-1}{k})$, where R^2 Rindicates the exposure variance of SNPs (i.e., the extent to which the exposure as a whole can be accounted for by an individual SNP), n denotes the sample size of SNPs, and k equals the number of IVs included (k = 1 for individual SNP). R^2 was obtained by using the formula $R^2 = \frac{\beta^2}{(\beta^2+SE^2(n)}$,where β is the effect size for the SNP and SE is the standard error for β. SNPs that are not strongly correlated with the exposure (F < 10) were excluded from the study. The second assumption is the exclusion restriction assumption. This is achieved by performing the build-in harmonization function of the *TwoSampleMR* package. Lastly, the independence assumption assumes that the SNPs are independent from pleiotropy. This is achieved by performing the pleiotropy test. A pleiotropy P-value of smaller than 0.05 indicates significant pleiotropy. Odd ratios and P-value were generated from the build-in function of *TwoSampleMR* package using the inverse-weighted variance method. The analysis was replicated in Van der Harst et al (replication cohort) [11]. The Bonferroni correction threshold for the pQTL was $\frac{0.05}{2941} = 1.7x10^{-5}$.

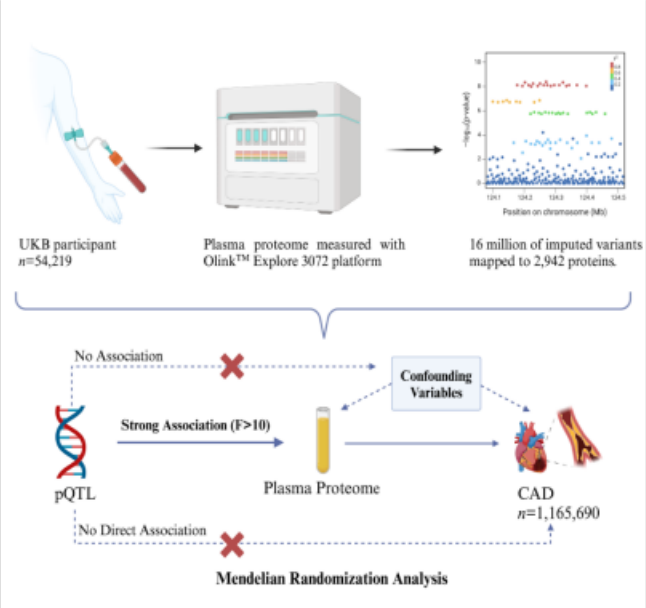

Figure 1 A: Plasma Proteome-based Mendelian Randomization Analysis of CAD Therapeutic Targets) : Overview of the study design(Created with Biorender.com.

Target prioritization

The ties of targets are based on the following algorithm: Tier 1: already existing direct inhibitor that can be repurposed or investigated in clinical trials and pass the pleiotropy test in both cohorts. Tier 2: no direct inhibitors but has the possibility of targeting and passing the pleiotropy test in at least one cohort. Tier 3: high likelihood of off-target effect and does not pass the pleiotropy test in both cohorts. Potential targets: has relevant biological function from cited references.

KEGG Analysis

Clusterprofiler (version 4.3) package and reference to the KEGG pathways were used to annotate the protein list in supplemental [Table 3]. Standard codes of *Clusterprofiler* were used and the vignette can be found in :

https://bioconductor.org/packages/release/bioc/vignettes/clusterProfiler/inst/doc/clusterProfiler.html

GSEA Analysis

GSEA analysis was performed using the GSEA software (version 4.3.2) with normalized gene expression from CAD patients and healthy individuals as the input. Vignette for GSEA software can be found in https://www.gsea-msigdb.org/gsea/index.jsp . The RNA-seq data was downloaded from GSE20680 [12]. PRISM (version 10.2.2) was used to plot the enrichment results.

Results

Plasma-proteomic-wide Mendelian Randomization reveals possible therapeutic targets

Out of the complete 3072 plasma protein panel, we were able to extract sufficient instrumental variables to perform our analyses to estimate the causal effect of 2942 proteins on the Aragam et al GWAS [9]. Over 300 plasma *risk proteins* with OR>1.05 and P < 0.05 were identified as associated with an increased risk of CAD (Figure 1B). To enhance specificity, we performed Bonferroni correction (OR>1.05 and P < 1.7E-5) to reveal a list of *candidate proteins*, consisting of 9 proteins that meet the statistical requirement. The final list of *candidate proteins* is shown in Figure 1C. The GWAS from van der Harst et al was included as a replication cohort [11]. PLA2G7, LPA, and PCSK9 are known to be associated with CAD, substantiating the validity of our analysis.

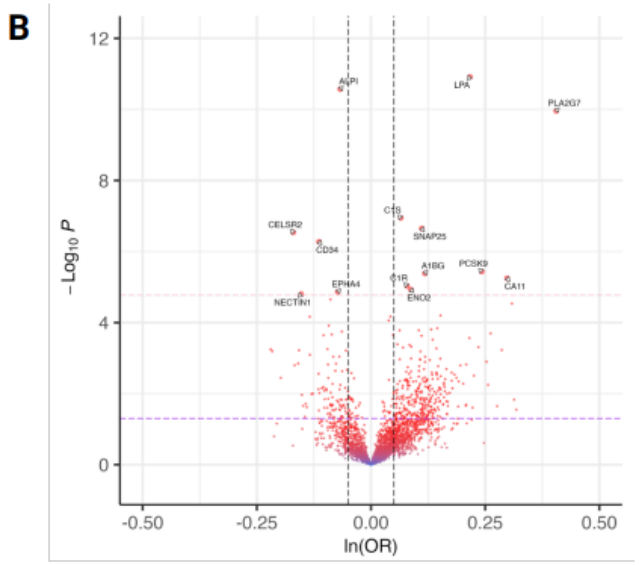

Figure 1 B: Volcano plot showing the MR analysis results. Y-axis: -Log(P-value); X-axis: ln(odd ratios). The purple line indicates the P=0.05 threshold. The pink line indicates the P-value threshold after Bonferroni correction.

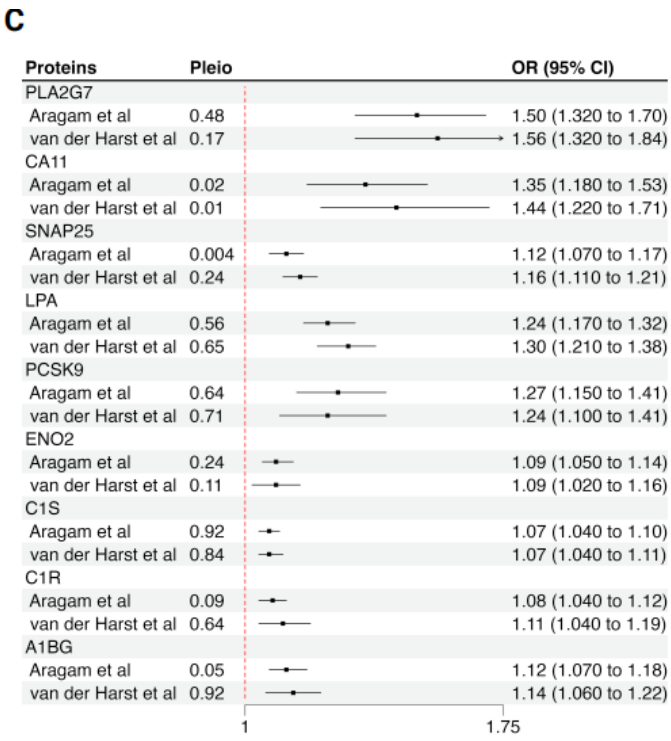

Figure 1 C: Forest plot showing the odd ratio, confidence interval, and Pleitropy test results of selected targets.

Prioritization of therapeutic targets

We prioritized targets based on the availability of direct inhibitors, the annotated function of proteins, and the pleiotropy test results (Figure 1D). C1S [OR= 1.07 (1.04-1.10), Pleiotropy P-value= 0.92], C1R [OR=1.08, Pleiotropy P-value = 0.07] and ENO2 [OR=1.09 (1.05 to 1.14), Pleiotropy P-value=0.24] were determined to be tier 1 target. SNAP25 [OR=1.12 (1.07 to 1.17), Pleiotropy P-value=0.004] and A1BG [OR=1.12 (1.07 to 1.18), Pleiotropy P-value=0.92] were determined to be tier 2 targets. CA11 was determined to be a tier3 target [OR=1.35 (1.18 to 1.53). Pleiotropy P-value=0.02]. The complete analysis results are included in supplemental tables 1 and 2 for the Aragam et al and van der Harst et al cohorts, respectively. The repurposing of C1S

inhibitor sutimlimab and targeting of ENO2 with POMHEX are of particular interest for their availability and mechanistic relevance. C1S is a critical component of the complement system involved in a myriad of biological and pathophysiological processes [13]. ENO2 is an enolase isoenzyme responsible for catalyzing the conversion of 2-phosphoglycerate to phosphoenolpyruvate in the glycolysis pathway [14]. For tier 2 targets, a clinical trial has shown that Botulinum toxin type A (inhibitor of SNAP25) injection to the pericardial fat pad suppressed atrial fibrillation following coronary artery bypass graft (CABG) longitudinally [15]. Interestingly, the involvement of polymorphism of A1BG, a glycoprotein with limited characterization, in CAD was reported [16]. CA11 is a catalytically inactive carbonic anhydrase with difficulty in selective targeting. In addition, it failed the pleiotropy test ($P<0.05$) across the primary and replication cohorts and is thus not prioritized as a therapeutic target.

D

1st Tier Targets		
Protein	*Inhibitor*	*Note*
C1S/C1R	Sutimlimab	Repurposing
ENO2	POMHEX	Investigational
2nd Tier Targets		
Protein	*Inhibitor*	*Note*
SNAP25	Botulinum toxin type A	Investigational
A1BG	N/A	Neutralizing Antibody
3rd Tier Targets		
Protein	*Inhibitor*	*Note*
CA11	N/A	Off-Target Risks
Potential Targets		
Protein	*Inhibitor*	*Note*
LCAT	In development	KO suppresses CAD in mice
SEMA3G	NA	KO reduces obesity in mice
LAG3	Relatlimab	Repurposing
TGFB1	Vactosertib	Investigational

Figure 1 D: List of targets in the descending order of priority.

Risk proteins are involved in immune-related signature and lipid metabolism

To elucidate the functional significance of the *risk proteins* in the pathophysiology of CAD, we selected a list of proteins by overlapping the *risk proteins* identified from the two CAD GWAS (supplemental table 3). Subsequently, we performed KEGG pathway enrichment analysis and showed that these proteins are involved in chemokine signaling, complement activation, chronic myeloid leukemia, and lipid metabolism pathways, among others (Figure 1E, 1F). These pathway analyses provided context into the biological significance of the target identified.

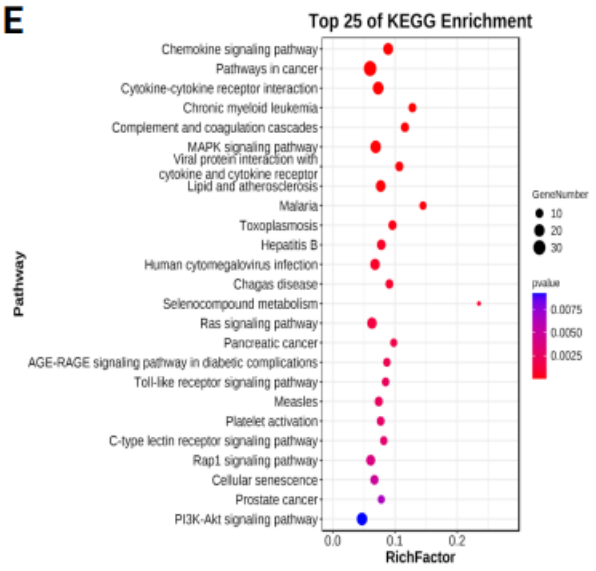

Figure 1 E: KEGG enrichment results from the overlapped proteins identified from MR.

F

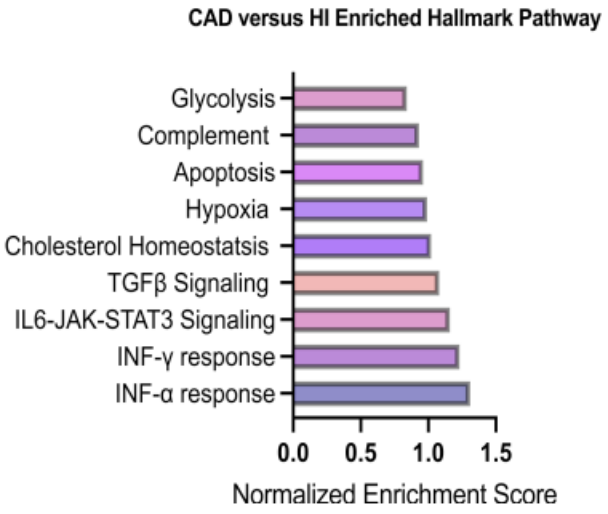

Figure 1 F: GSEA enrichment results from RNA-seq data of CAD patients versus healthy controls.

GSEA analysis shows the enrichment of complement and glycolysis pathways in CAD patients

We further explored RNA-Seq data of CAD patients versus healthy individuals (HI) whole blood from GSE20680. With Gene set enrichment analysis (GSEA), we revealed hallmark pathways similar to the KEGG results such as complement, cholesterol homeostasis, JAK-STAT signaling (involved in leukemia), and

interferon response. These results corroborated our findings on the biological processes of *risk proteins.* Of note, the enrichment in the glycolysis pathway provides additional rationale for targeting ENO2. Based on these findings, we delved into the *risk protein* list and further identified potential targets for CAD (Figure 1D). LCAT, SEM3AG, LAG3, and TGFB1 were considered to be potential targets based on their biological function. LCAT is involved in lipid metabolism, obesity, and cardiovascular events [17]. SEMA3G is a purported adipokine whose KO suppressed high-fat diet (HFD)-induced obesity in mice [18]. LAG3 is involved in T-cell regulation with the recently approved inhibitor relatlimab [19]. Dysregulated TGFb signaling is proposed to be involved in vascular diseases, which is corroborated by our RNA-seq and MR analyses [20].

Discussion

Overall, our study innovatively leveraged the pQTLs of plasma proteome and unraveled an avenue for the treatment of CAD. We have identified C1R/S, two critical components of the complement system, as potential therapeutic targets for the treatment of CAD. Previous studies have highlighted that the complement system plays an important role in atherosclerotic lesions and that suppressing the overactivation of the complement system can have therapeutic value [21,22]. Results from our studies are consistent with these findings while providing evidence from a genetic perspective. Additionally, we also found that ENO2, an isozyme of enolase (ENO), a key component of glycolysis can be a promising target for CAD. Prior studies have reported that the genetic silencing of ENO can suppress apoptosis and mitigate mitochondrial dysfunction by habiting the release of mitochondrial cytochrome C, an important mediator of apoptotic response, into the cytoplasm [23]. Thus, our finding showing that ENO2 is associated with the risk of CAD supported this observation. Furthermore, the complement activation pathway and glycolysis pathway were identified from our RNA-seq analysis to be enriched in CAD patients relative to healthy controls, collaborating with our MR findings. In short, our study has highlighted two potential avenues for the treatment of CAD: 1) suppression of the complement system and 2) glycolysis.

Our study has a few strengths. First, our study is statistically robust as we stratified the candidate proteins after Bonferroni correction into three tiers to prioritize novel targets to ensure that each target identified has genomic significance. Second, most studies applying MR do not further extend their studies by validating MR findings with other modalities, whereas we elucidated the biological function of the *risk proteins* and corroborated our findings with RNA-Seq data. However, our study is not free from limitations. First, PLA2G7 and PCSK9 were both identified. However, the former is not an efficacious target while the latter is indicated only for high-risk CAD patients, highlighting the fallibility of targeted therapy for CAD. Second, while we have endeavored to minimize sample overlaps, some of the reference samples in the outcomes overlap with the UK Biobank cohort, lowering the confidence of our conclusion.

Conclusion

In conclusion, although we have provided new insights into the treatment of CAD and demonstrated the application of pQTLs as valuable instruments in revealing drug targets for diseases, we must proceed with caution as *in silico* analysis cannot fully capture the complexity of disease biology in humans, especially for multifactorial diseases like CAD.

Author Contribution

SL, HS, and IH conceived the project. SL, HS, and IH collected, analyzed and interpreted the data.

SL, HS, and IH wrote the manuscript with input from IH. IH supervised the entire project.

Data Availability

All original GWAS data can be accessed from the cited reference. Detailed analyses results are available upon reasonable request to lshuyang@wustl.edu.

Acknowledgment

We thank the researchers and participants of the GWAS cited in this study. We thank our colleagues, who prefer to remain anonymous, for proofreading the manuscript and providing feedback.

Ethical Considerations: Not applicable because this study used public data.

Conflict of interest: None

References

1. STABILITY Investigators (2014) Darapladib for preventing ischemic events in stable coronary heart disease. The New England journal of medicine. 370: 1702–1711.
2. Boden WE, O'Rourke RA, Teo KK, Hartigan PM, Maron DJ, et al. (2007) Optimal medical therapy with or without PCI for stable coronary disease. The New England Journal of Medicine. 2007, 356:1503–1516.
3. Sabatine MS, Giugliano RP, Keech AC, Honarpour N, Wiviott SD, et al. (2017) Evolocumab and Clinical Outcomes in Patients with Cardiovascular Disease. The New England Journal of Medicine. 2017, 376:1713–1722.
4. Boef AG, Dekkers OM, le Cessie S (2015) Mendelian randomization studies: a review of the approaches used and the quality of reporting. International Journal of Epidemiol. 44:496-511.

5. Zou X, Wang L, Wang S, Zhang Y, Ma J, et al. (2024) Promising therapeutic targets for ischemic stroke identified from plasma and cerebrospinal fluid proteomes: a multicenter Mendelian randomization study. International Journal of Surgery. 110:766-776.

6. Henry A, Gordillo-Marañón M, Finan C, Schmidt AF, Ferreira JP, et al. (2022) Therapeutic Targets for Heart Failure Identified Using Proteomics and Mendelian Randomization. Circulation. 145:1205-1217.

7. Lin S, Gao B, Xu R, Shang H, Xiong Y, et al. (2024) Multiple myeloma, IL6, and risk of schizophrenia: A Mendelian randomization, transcriptome, and Bayesian colocalization study. eJHaem.

8. Burgess S, Thompson SG, CRP CHD Genetics Collaboration (2011) Avoiding bias from weak instruments in Mendelian randomization studies. International Journal of Epidemiology. 40: 755–764.

9. Aragam KG, Jiang T, Goel A, Kanoni S, Wolford BN, et al. (2022) Discovery and systematic characterization of risk variants and genes for coronary artery disease in over a million participants. Nature Genetics. 54:1803–1815.

10. Sun BB, Chiou J, Traylor M, Benner C, Hsu YH, et al. (2023) Plasma proteomic associations with genetics and health in the UK Biobank. Nature. 622: 329–338.

11. van der Harst P, Verweij N (2018) Identification of 64 Novel Genetic Loci Provides an Expanded View on the Genetic Architecture of Coronary Artery Disease. Circulation Research.122:433-443.

12. Elashoff MR, Wingrove JA, Beineke P, Daniels SE, Tingley WG, et al. (2011) Development of a blood-based gene expression algorithm for assessment of obstructive coronary artery disease in non-diabetic patients. BMC Med Genomics. 4:26.

13. Ye J, Yang P, Yang Y, Xia S (2022) Complement C1s as a diagnostic marker and therapeutic target: Progress and prospective. Front Immunol. 13:1015128.

14. Shi J, Miao D, Lv Q, Tan D, Xiong Z, et al. (2023) ENO2 as a Biomarker Regulating Energy Metabolism to Promote Tumor Progression in Clear Cell Renal Cell Carcinoma. Biomedicines. 11: 2499.

15. Brin MF, Burstein R (2023) Botox (onabotulinumtoxinA) mechanism of action. Medicine, 10: e32372.

16. McCaffrey TA, Toma I, Yang Z, Katz R, Reiner J, et al. (2021) RNA sequencing of blood in coronary artery disease: involvement of regulatory T cell imbalance. BMC Medical Genomics. 14: 216.

17. Holleboom AG, Kuivenhoven JA,Vergeer M, Hovingh GK, van Miert GN, et al. (2010) Plasma levels of lecithin: cholesterol acyltransferase and risk of future coronary artery disease in apparently healthy men and women: a prospective case-control analysis nested in the EPIC-Norfolk population study. Journal of lipid research. 51: 416–421.

18. Liu M, Xie S, Liu W, Li J, Li C, et al. (2020) Mechanism of SEMA3G knockdown-mediated attenuation of high-fat diet-induced obesity. J Endocrinol. 244: 223-236.

19. Su J, Fu Y, Cui Z, Abidin Z, Yuan J, et al. (2024) Relatlimab: a novel drug targeting immune checkpoint LAG-3 in melanoma therapy. Frontiers in pharmacology. 14: 1349081.

20. Low EL, Baker AH, Bradshaw AC (2019) TGFβ, smooth muscle cells, and coronary artery disease: a review. Cellular signaling. 53: 90–101.

21. Oksjoki R, Kovanen PT, Meri S, Pentikainen MO (2007) Function and regulation of the complement system in cardiovascular diseases. Front Biosci. 12:4696-4708.

22. Lappegård KT, Garred P, Jonasson L, Espevik T, Aukrust P, et al. (2014) A vital role for complement in heart disease. Mol Immunol. 61: 126-134.

23. Chen S, Zou Y, Song C, Cao K, Cai K,et al. (2023) The role of glycolytic metabolic pathways in cardiovascular disease and potential therapeutic approaches. Basic research in cardiology, 118: 48.