Experimental Research

INTERNATIONAL JOURNAL OF SURGERY

OPEN

Promising therapeutic targets for ischemic stroke identified from plasma and cerebrospinal fluid proteomes: a multicenter Mendelian randomization study

Xuelun Zou, PhD[a], Leiyun Wang, PhD[g], Sai Wang, PhD[a], Yupeng Zhang, PhD[a], Junyi Ma, PhD[a], Lei Chen, PhD[a], Ye Li, PhD[a], Tian-Xing Yao, PhD[a], Huifang Zhou, PhD[a], Lianxu Wu, PhD[a], Qiaoling Tang, PhD[a], Siyuan Ma, PhD[a], Xiangbin Zhang, PhD[a], Rongmei Tang, PhD[a], Yexiang Yi, PhD[a], Ran Liu, PhD[a], Yi Zeng, PhD[c], Le Zhang, PhD[a,b,d,e,f,*]

Background: Ischemic stroke (IS) is more common every year, the condition is serious, and have a poor prognosis. New, efficient, and safe therapeutic targets are desperately needed as early treatment especially prevention and reperfusion is the key to lowering the occurrence of poorer prognosis. Generally circulating proteins are attractive therapeutic targets, this study aims to identify potential pharmacological targets among plasma and cerebrospinal fluid (CSF) proteins for the prevention and treatment of IS using a multicenter Mendelian randomization (MR) approach.

Methods: First, the genetic instruments of 734 plasma and 151 CSF proteins were assessed for causative connections with IS from MEGASTROKE consortium by MR to identify prospective therapeutic targets. Then, for additional validation, plasma proteins from the deCODE consortium and the Fenland consortium, as well as IS GWAS data from the FinnGen cohort, the ISGC consortium and UK biobank, were employed. A thorough evaluation of the aforementioned possible pharmacological targets was carried out using meta-analysis. The robustness of MR results was then confirmed through sensitivity analysis using several techniques, such as bidirectional MR analysis, Steiger filtering, and Bayesian colocalization. Finally, methods like Protein-Protein Interaction (PPI) Networking were utilized to investigate the relationship between putative drug targets and therapeutic agents.

Results: The authors discovered three proteins that may function as promising therapeutic targets for IS and meet the Bonferroni correction ($P < 0.05/885 = 5.65 \times 10^{-5}$). Prekallikrein (OR = 0.41, 95% CI: 0.27–0.63, $P = 3.61 \times 10^{-5}$), a protein found in CSF, has a 10-fold protective impact in IS, while the plasma proteins SWAP70 (OR = 0.85, 95% CI: 0.80–0.91, $P = 1.64 \times 10^{-6}$) and MMP-12 (OR = 0.92, 95% CI: 0.89–0.95, $P = 4.49 \times 10^{-6}$) of each SD play a protective role in IS. Prekallikrein, MMP-12, SWAP70 was replicated in the FinnGen cohort and ISGC database. MMP-12 (OR = 0.93, 95% CI: 0.91–0.94, $P < 0.001$), SWAP70 (OR = 0.92, 95% CI: 0.90–0.94, $P < 0.001$), and prekallikrein (OR = 0.53, 95% CI: 0.33–0.72, $P < 0.001$) may all be viable targets for IS, according to the combined meta-analysis results. Additionally, no evidence of reverse causality was identified, and Bayesian colocalization revealed MMP-12 ($PPH_4 = 0.995$), SWAP70 ($PPH_4 = 0.987$), and prekallikrein ($PPH_4 = 0.894$) shared the same variant with IS, supporting the robustness of the aforementioned causation. Prekallikrein and MMP-12 were associated with the target protein of the current treatment of IS. Among them, Lanadelumab, a new drug whose target protein is a prekallikrein, may be a promising new drug for the treatment of IS.

Conclusion: The prekallikrein, MMP-12, and SWAP70 are causally associated with the risk of IS. Moreover, MMP-12 and prekallikrein may be treated as promising therapeutic targets for medical intervention of IS.

Keywords: ischemic stroke, Mendelian randomization, proteome, therapeutic target

[a]Department of Neurology, Xiangya Hospital, Central South University, [b]Human Brain Disease Biological Resources Platform of Hunan Province, Xiangya Hospital, [c]Department of Geriatrics, Second Xiangya Hospital, [d]National Clinical Research Center for Geriatric Disorders, Xiangya Hospital, [e]Multi-Modal Monitoring Technology for Severe Cerebrovascular Disease of Human Engineering Research Center, Xiangya Hospital, [f]Brain Health Center of Hunan Province, Xiangya Hospital, Central South University, Changsha, Hunan and [g]Department of Pharmacy, Wuhan No.1 Hospital, Wuhan, Hubei, People's Republic of China

Sponsorships or competing interests that may be relevant to content are disclosed at the end of this article.

**Corresponding author. Address: Department of Neurology, Xiangya Hospital, Central South University, Changsha, Hunan 410008, People's Republic of China. Tel.: +86 073 185 295 888; Fax: +86 073 185 533 525. E-mail: zlzdzlzd@csu.edu.cn (L. Zhang).*

International Journal of Surgery (2024) 110:766–776

Received 21 August 2023; Accepted 9 November 2023

Supplemental Digital Content is available for this article. Direct URL citations are provided in the HTML and PDF versions of this article on the journal's website, www.lww.com/international-journal-of-surgery.

Published online 27 November 2023

http://dx.doi.org/10.1097/JS9.0000000000000922

Introduction

The latest data on the global burden of disease shows that stroke is the second leading cause of death after ischemic heart disease and the third leading cause of disability after ischemic heart disease and neonatal disease. Ischemic stroke (IS) is the major subtype of stroke, accounting for 62.4% of new strokes, and is the leading cause of high mortality and disability in stroke. Early pharmacological treatment is the main medical intervention to effectively reduce the high mortality and disability rates of IS[1]. The only thrombolytic drug currently approved by the FDA is recombinant tissue fibrinogen activator[2]. However, the time window is narrow for treating IS, with the complication of hemorrhagic transformation, and less than 10% of stroke patients are currently eligible for this treatment[3,4]. There are still many IS patients without timely treatment available and evolve into patients with severe IS, with no reversal of the fatal or disabling fate. This may be the one of the main reasons why the disease burden of IS is rising each year, with a disability-adjusted life year of 143 (95% CI: 133–153 million) million in 2019[5]. Therefore, it is urgent to address the unmet medical needs for the prevention and treatment of IS.

Clinical drugs can bind to target proteins in the body and alter the activity of the target or downstream proteins to stop the progression of the disease. Previous observations have identified plasma protein targets as potential therapeutic targets in IS[6,7]. However, the high cost and attrition rates of drug development make drug targeting of protein targets in preclinical trial studies IS highly practical. The proportion of drug mechanism-of-action supported by direct genetic evidence increases significantly across drug development trials, and the selection of genetically supported targets can double the success rate of clinical drug development[8].

Mendelian randomization (MR) is an epidemiological approach that can correlate plasma and cerebrospinal fluid (CSF) protein information obtained by high-throughput sequencing with genome-wide association results associated with disease phenotypes. It can be applied for drug target identification and drug repurposing prior to clinical trials. MR assesses the causal relationship between exposure factors (such as proteomics) and outcome (IS) in a genetic context and excludes the influence of confounding factors on causality with the aid of multiple sensitivity analyses[9], and the exclusion of potential confounding factors is one of its advantages over traditional observational studies[10]. Another advantage of MR is the ability to mimic the randomized controlled trial in grouping subjects according to the presence or absence of alleles and to observe the effect of alleles on outcome without the enormous human and material resources of a randomized controlled trial[11]. Although 653 plasma proteins have been used to identify drug targets of IS and its subtypes[12] and 308 plasma proteins were used to explore the association of plasma proteins with stroke and its risk factors[13], no drug target studies have been conducted using CSF proteins for medical intervention of IS and no potential drugs have been discovered. Therefore, we explored potential drug targets for IS in this study and purposes to build a theoretical foundation for the discovery and development of effective therapeutics to address the unmet medical needs of life-threating IS.

HIGHLIGHTS

- Our study explored the causal relationship between 734 plasma proteins and 154 cerebrospinal fluid (CSF) proteins and ischemic stroke (IS) by Mendelian randomization study.
- This study found a causal relationship between the plasma protein matrix metalloproteinase 12, switch-associated protein 70, and the CSF protein prekallikrein and IS risk, which is beneficial for the early prevention and treatment of IS.
- No reverse causal relationship was found between 734 plasma proteins and 154 CSF proteins and IS.
- Matrix metalloproteinase 12 and prekallikrein were associated with the target protein of the current treatment of IS and may be a potential target for new IS drugs.
- Lanadelumab, a new drug whose target protein is a prekallikrein, may be a promising new drug for the treatment of IS.

Methods

Research design

The purpose of this study is to build a groundwork for the later precise prevention and treatment of IS by identifying possible therapeutic targets of IS in the CSF and plasma proteins, as seen in Figure 1. An MR analysis was first conducted to examine the proteins in plasma and CSF that are linked to IS. Data were taken from published plasma and CSF proteomics studies, while IS data were collected from the MEGASTROKE collaboration. Sensitivity analysis using Bayesian colocalization, bidirectional MR, and phenotype scanning were then used to demonstrate that the results were robust. The results were then verified by a multicenter MR study, such as exposure data from Fenland and deCODE database and outcome data from FinnGen cohort, ISGC consortium and UK biobank (Supplemental Figure 1, Supplemental Digital Content 1, http://links.lww.com/JS9/B402). Finally, a comparative and interactive analysis was performed to identify pharmacological targets for IS. There was no need to get informed consent or ethical approval for this study again because all of the data were taken from published sources, and the informed consent and approval were received.

Exposure data acquisition

In the initial phase, the plasma proteins are acquired from a latest publication in *Nature Genetics*[14], which integrates plasma protein quantitative trait loci (pQTL) data obtained from five previously published GWAS studies[15–18]. Three thousand six hundred six pQTLs connected to 2656 proteins were present in this investigation. The CSF pQTL data was acquired from a study published in *Nature Neuroscience*[19], which included a total of 274 pQTLs of 184 CSF proteins. Our requirements for pQTLs found in plasma and CSF were as follows: firstly, they had to be cis-acting pQTLs. Secondly, they had to have a current genome-wide association ($P < 5 \times 10^{-8}$), and thirdly, they had to be found outside the major histocompatibility complex region (Chr6, 26–34 Mb). The effect of linkage disequilibrium (LD) was also eliminated (LD clumping $r^2 < 0.001$). Finally, we included 738 cis-pQTLs in 734 plasma proteins (Supplemental Table 1, Supplemental Digital Content 2, http://links.lww.com/JS9/B403)

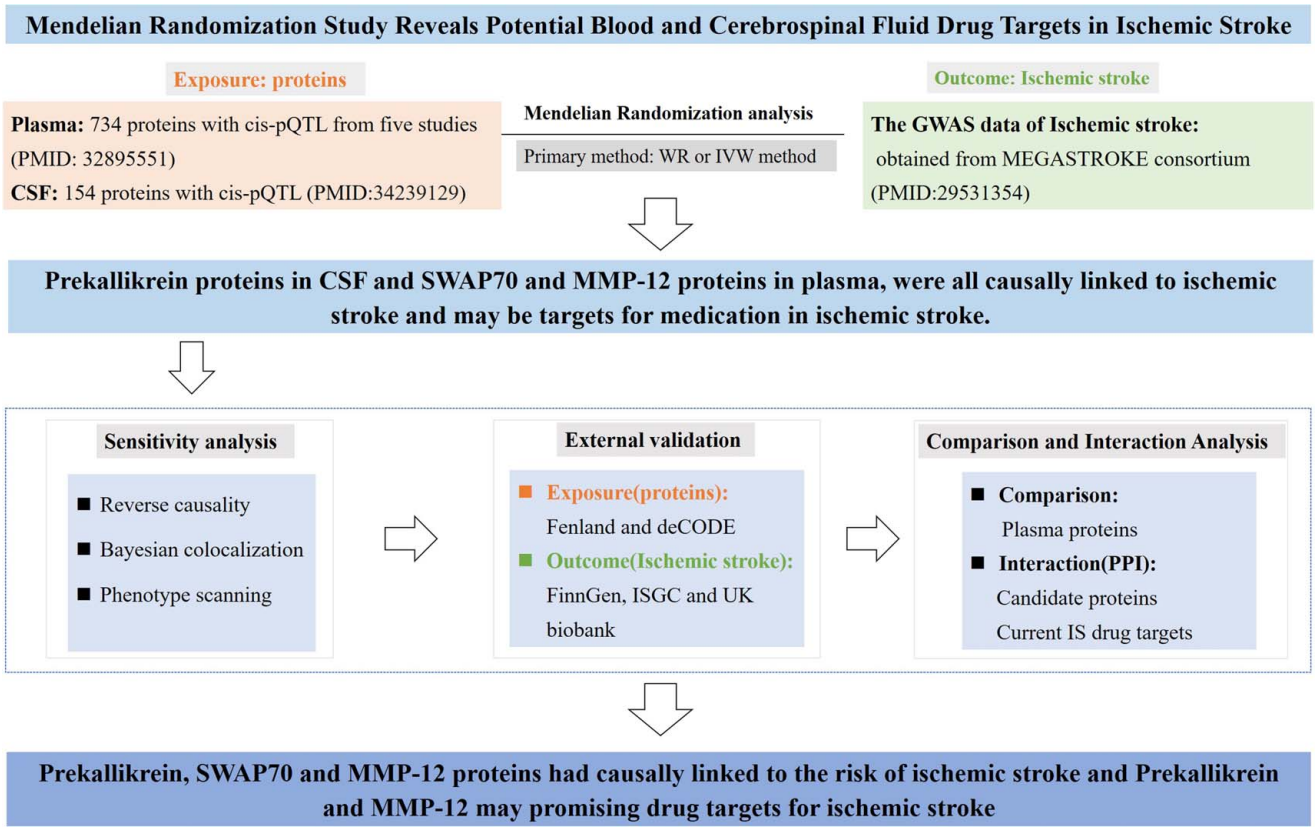

Figure 1. Design of the study and its primary findings. CSF, cerebrospinal fluid; ISGC, International Stroke Genetics Consortium; PPI, Protein-Protein Interaction Networks.

and 154 cis-pQTLs linked to 151 CSF proteins (Supplemental Table 2, Supplemental Digital Content 3, http://links.lww.com/JS9/B404) as exposure ('Primary_', center 1).

For the replication phase, we extracted plasma protein data from two other center studies using the original CSF protein data as validation. The plasma protein pQTL is from a proteomics study conducted by Ferkingstad *et al.*[20] (deCODE database, center 2) in the journal of *Nature Genetics*. This research enumerated the 4907 plasma protein in 35 559 Icelanders' and discovered 18 084 correlations between sequence variations and plasma protein levels[20]. The plasma protein data were obtained from a publication in *Science* that included 4775 proteins (Fenland database, center 3) measured in 10 708 participants[21]. To fill up any gaps in the QTL GWAS summary statistics, such as impact allele frequency, we used the matching human genome construct as a reference.

Sources of outcome data

Data on IS for this study were obtained from a GWAS meta-analysis conducted by the MEGASTROKE consortium ('_Primary', center 4), which included 440 328 participants (34 217 cases and 406 111 controls) and analyzed 7 537 579 SNPs[22]. The investigations by the ISGC consortium[23], which comprised 10 307 cases and 19 326 controls (center 5), and the FinnGen cohort[24], which included 12 632 cases and 206 160 controls, provided the data for the replication phase (center 6). Furthermore, IS GWAS from UK biobank (UKB) has been incorporated (https://www.ukbiobank.ac.uk/) (center 7).

MR analysis

The plasma and CSF proteins as exposure data and IS as outcome data were utilized. In this study, involving two or more instrumental variables, the primary analysis method was the MR study's inverse variance weighting (IVW) method[25]. The Wald ratio approach served as the study's main method of analysis when there was just one instrumental variable[25]. For every additional unit of plasma protein or every additional 10 units of CSF protein, the measured outcome odds ratios represent the odds of a higher stroke risk. MR analysis is performed in the 'TwoSampleMR' package in R version 4.2.3. To reduce the chance of false positives, we conducted a multiple test correction (Bonferroni correction), and the *P*-value of statistical significance in this study was set to less than 0.05/885 (5.65×10^{-5}). During the replication phase, some of the MR analysis results were not in the same direction as the MR analysis results in the initial phase, so we further conducted a meta-analysis with a random-effects model to obtain more robust MR results. The meta-analysis was done in STATA version 14, and the set statistical difference was *P* less than 0.05.

To confirm the accuracy of the data displayed above, we also ran a sensitivity analysis. Reverse MR used IS exposure data from the MEGASTROKE consortium and included a total of 19 eligible SNPs as instrumental factors (Supplemental Table 3,

Supplemental Digital Content 4, http://links.lww.com/JS9/B405). It adhered to the same screening standards as the forward MR study. The plasma[16,20] proteins and CSF[19] proteins, which were utilized as outcome data, was available from three published studies. Reverse causality was principally investigated via the IVW method in addition to MR-Egger regression, weight median (WM), simple mode, and weight mode approaches, which were integrated to analyze causation[10,26]. Steiger filtering was also carried out to establish the directionality of the relationship between proteins and IS[27].

The likelihood that two traits share the same causal variant was evaluated using Bayesian colocalization analysis using the 'coloc' package (https://github.com/chr1swallace/coloc) with default parameters. The gene-based PPH4 greater than 80% was defined that the gene as having evidence of colocalization[10]. We also used phenotypic scanning (http://www.phenoscanner.medschl.cam.ac.uk/)[28], which permits retrieval of potential confounders and pleiotropic linked to the study's instrumental variables and contains published findings from large-scale genetic association studies, to test for the presence of confounding variables and pleiotropic. We initially have to meet the significant genomic connections ($P < 5 \times 10^{-8}$) for instrumental variable inclusion as part of the criterion we defined for the presence of pleiotropy. Second, the GWAS research population of European ancestry was the source of the instrumental factors. Last but not least, the instrumental variables were linked to IS risk factors such as metabolic traits, protein composition, and clinical traits. Additionally, we set the LD value r^2 for pQTLs of priority proteins less to 0.001 to reduce the impact of LD. The LDlink website is then used to explore again whether there is a LD in the results[10].

Protein-protein and protein-drug association analysis

We hypothesize that the connection between plasma pQTL and CSF pQTL is relatively weak because of the blood–brain barrier (BBB). When plasma and CSF pQTL were eventually included in our investigation, we utilized spearman correlation analysis to analyze their correlation. Between – 1.0 and 1.0 is the range of values for the spearman correlation coefficient (r^2). A perfect negative correlation is represented by a correlation near – 1.0, and a perfect positive correlation is represented by a correlation near 1.0. The shifts of the plasma pQTL and CSF pQTL do not have a linear connection when there is a correlation near 0.0. We then established various *P*-value criteria to see if the correlation between the two varied as the degree of significance increased. After analyzing the interrelationship between plasma pQTL and CSF pQTL, we employed protein-protein interaction (PPI) networks to further evaluate the associated protein targets in CSF and plasma proteins that were suggestively ($P < 0.05$) related with the risk of IS in the main MR study. The PPI network is obtained through the STRING website, which is functional protein association networks (https://cn.string-db.org). The minimum required interaction fraction of STRING is 0.4. The protein target pathway is an important approach for the discovery of effective drug compounds. Drugs are able to alter the activity of both target and downstream proteins with the aim of ending disease progression. Therefore, we used the Drugbank database (https://go.drugbank.com/) to explore the relationship between IS protein targets and related genes and approved IS clinical use in recent review[29].

Results

Multicenter MR analysis and meta-analysis revealed that SWAP70, MMP-12 and Prekallikrein were causally associated with IS

Overall, circulating bioproteins from 734 plasma pQTL and 154 CSF pQTL were examined for a potential link to IS (Bonferroni correction of $P < 0.05/885 = 5.65 \times 10^{-5}$). Three proteins, matrix metalloproteinase 12 (MMP-12) and switch-associated protein 70 (SWAP70) in plasma, and prekallikrein (encoded by the KLKB1 gene) in CSF, were discovered to be causally linked to IS under this circumstance (Fig. 2 and Table 1). As to the specific findings, each SD rise in plasma MMP-12 (OR = 0.92, 95% CI: 0.89–0.95, $P = 4.49 \times 10^{-6}$) and SWAP70 (OR = 0.85, 95% CI: 0.80–0.91, $P = 1.64 \times 10^{-6}$) levels would lower the risk of IS, whereas each 10-fold increase in prekallikrein (OR = 0.41, 95% CI: 0.27–0.63, $P = 3.61 \times 10^{-5}$) concentration in CSF would lower the risk of IS. Heterogeneity and pleiotropy were not detected by preliminary study (Supplemental Table 4, Supplemental Digital Content 5, http://links.lww.com/JS9/B406).

To validate the causative association between the aforementioned proteins and IS, we used exposures and outcomes from multicenter datasets. We discovered that SWAP70 was verified to have a causal relationship with IS risk in both the primary plasma protein of SWAP70 and the IS data from multicenter MR study (Exposure_Outcome) of Primary_FinnGen (OR = 0.83, 95% CI: 0.73–0.94, $P = 0.003$), deCODE_FinnGen (OR = 0.89, 95% CI: 0.82–0.96, $P = 0.0033$), Fenland_FinnGen (OR = 0.94, 95% CI: 0.91–0.98, $P = 0.005$), Primary_ISGC (OR = 0.87, 95% CI: 0.77–0.99, $P = 0.037$), deCODE_ISGC (OR = 0.92, 95% CI: 0.85–0.99, $P = 0.0369$), Fenland_ISGC (OR = 0.95, 95% CI: 0.91–0.99, $P = 0.027$), deCODE_Primary (OR = 0.91, 95% CI: 0.87–0.94, $P = 1.6 \times 10^{-6}$), Feland_Primary (OR = 0.96, 95% CI: 0.94–0.99, $P = 0.0121$), Primary_UKB (OR = 0.93, 95% CI: 0.89–0.97, $P = 0.0004$), and deCODE_UKB (OR = 0.97, 95% CI: 0.92–1.01, $P = 0.0004$). Prekallikrein was shown to potentially lower IS risk in the ISGC database (OR = 0.45, 95% CI: 0.21–0.95, $P = 0.035$) but was not validated in the FinnGen cohort and UK biobank. The causal relationship between MMP-12 and IS was confirmed in a partial multicenter (exposure_outcome) MR study of Primary_FinnGen (OR = 0.94, 95% CI: 0.89–0.99, $P = 0.011$), deCODE_FinnGen (OR = 0.93, 95% CI: 0.88–0.98, $P = 0.0114$), Primary_ISGC (OR = 0.90, 95% CI: 0.84–0.96, $P = 0.0008$), deCODE_ISGC (OR = 0.89, 95% CI: 0.83–0.95, $P = 0.0008$), and deCODE_Primary (OR = 0.91, 95% CI: 0.88–0.95, $P = 4.5 \times 10^{-6}$) as shown in Figure 3, despite the causal relationship between MMP-12, prekallikrein and IS risk only being partially confirmed. We further conducted a meta-analysis to confirm the above causality and found that MMP-12 (OR = 0.93, 95% CI: 0.91–0.94, $P < 0.001$), SWAP70 (OR = 0.92, 95% CI: 0.90–0.94, $P < 0.001$) and prekallikrein (OR = 0.53, 95% CI: 0.33–0.72, $P < 0.001$) may reduce the risk of IS (Fig. 4).

Strong and trustworthy causal connection between plasma and CSF proteins with risk of IS confirmed by sensitivity analysis

Sensitivity was divided into three parts, bidirectional MR, colocalization analysis and phenotypic analysis (Table 2). First, The MR Steiger filtering ensured directionality of MMP-12

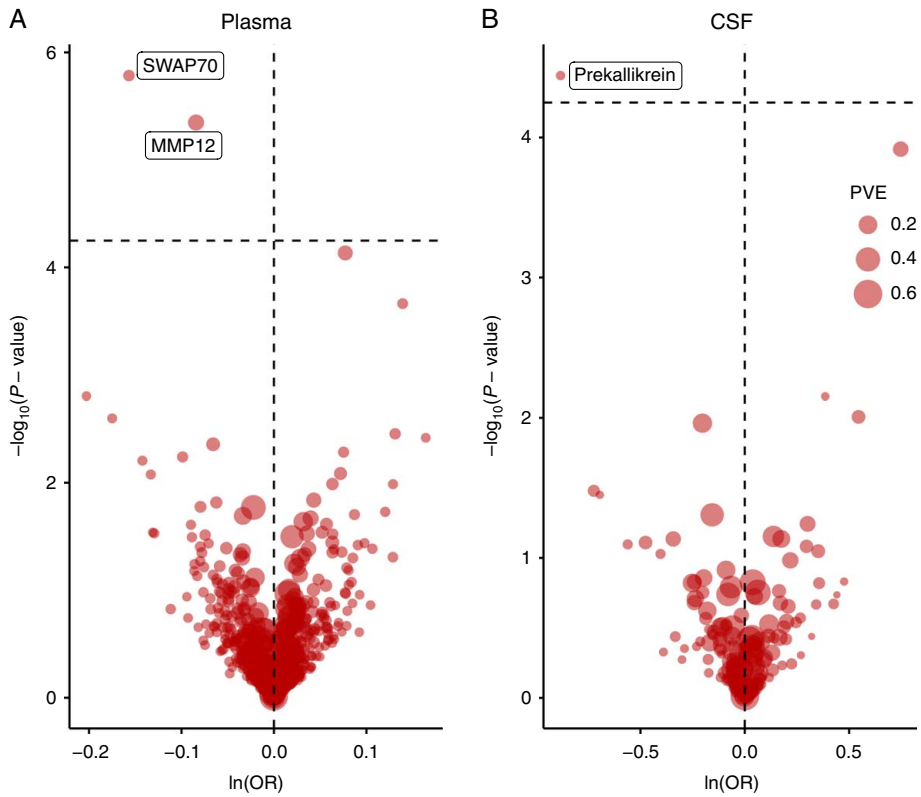

Figure 2. Volcano plot of three proteins identified as potential targets in ischemic stroke from 734 plasma (A) and 151 CSF (B) proteins by Wald ratio or inverse variance weighted method. CSF, cerebrospinal fluid; PVE, proportion of variance explained.

($P = 2.49 \times 10^{-101}$), SWAP70 ($P = 5.21 \times 10^{-23}$), prekallikrein ($P = 8.96 \times 10^{-9}$), and the risk of IS. Bidirectional MR analyses and Steiger filtering were performed, and there is no corresponding result SNP was discovered between IS and prekallikrein, MMP-12 or SWAP70, despite the matching of 27 SNPs between IS and MMP-12 and 20 SNPs between IS and SWAP70 (Supplemental Table 5, Supplemental Digital Content 6, http://links.lww.com/JS9/B407).

In addition, colocalization analysis revealed MMP-12 ($PPH_4 = 0.995$), SWAP70 ($PPH_4 = 0.987$), and prekallikrein ($PPH_4 = 0.894$) shared the same variant with IS, while colocalization results of the *P*-value distributions of the loci 100 KB above and below the pQTL for MMP-12, SWAP70, and prekallikrein and the *P*-value distributions of IS corresponding to the instrumental variables of MMP-12, SWAP70, and prekallikrein could be found in Figure 5. Finally, we performed the phenotypic analysis and found that MMP-12 (rs28381684) was linked to matrix metalloproteinases, postbronchodilator FEV1/FVC ratio, stromelysin-1, and local histogram emphysema pattern, according to a phenotype scan performed using the phenosanner website, while SWAP70 (rs415895) was linked to hematocrit, coronary artery disease, and hypertension. Various proteins, coagulation factors were connected to prekallikrein (rs2304595) in Supplemental Table 6 (Supplemental Digital Content 7, http://links.lww.com/JS9/B408). Prekallikrein is not on the same chromosome as SWAP70 and MMP-12. Therefore, there is no LD. Besides, no LD was found between the pQTLs for MMP-12 and SWAP70 (LD $r^2 = 0.003$), too.

Prekallikrein and MMP-12 may be a potential target for the IS drugs

The results of the plasma and CSF MR tests did not significantly differ at the protein level (Spearman correlation coefficient = −0.005, 95% CI: −0.265–0.223) in Supplemental Figure 2 (Supplemental Digital Content 8, http://links.lww.com/JS9/B409). A negative link; however, persisted and remained

Table 1

Mendelian randomization findings for plasma and cerebrospinal fluid proteins that are substantially linked to ischemic stroke after Bonferroni adjustment.

Tissue	Protein	UniProt ID	SNP	Effect allele	OR (95% CI)	*P*	PVE	F-statistics	Reference
Plasma	SWAP70	B3KUB9; E7EMB1; Q9UH65	rs415895	G	0.85 (0.80–0.91)	1.64E-06	3.16%	107.88	Sun
Plasma	MMP-12	P39900	rs28381684	T	0.92 (0.89–0.95)	4.49E-06	13.18%	500.96	Sun
CSF	Prekallikrein	P03952	rs2304595	G	0.41 (0.27–0.63)	3.61E-05	4.11%	35.76	Yang

Note: Every SNP that was used was cis-acting. The odds ratios for a higher risk of IS were expressed as a function of the SD increase in plasma protein levels and the 10-fold increase in CSF protein levels. CSF, cerebrospinal fluid; PVE, proportion of variance explained; SNP, single nucleotide polymorphism.

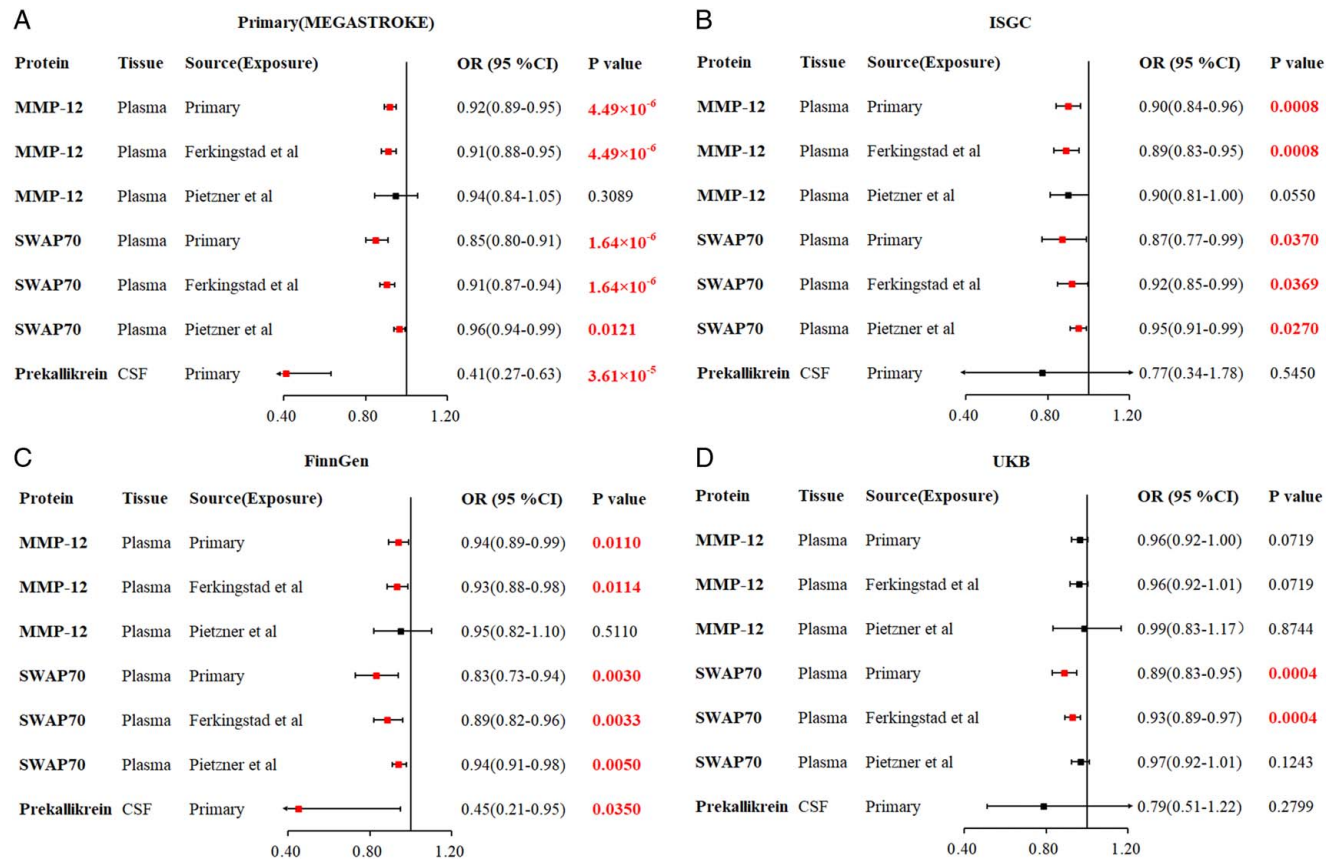

Figure 3. Data from the Primary(MEGASTROKE consortium, A), ISGC consortium (B), FinnGen cohort (C) and UK biobank (D) externally validate the causal link between MMP-12, SWAP70, and prekallikrein and IS. CSF, cerebrospinal fluid. Primary(Exposure): 734 plasma proteins and 151 CSF proteins used in the initial phase. _Primary: IS GWAS from MEGASTROKE consortium. Pietzner *et al*.[21]: Plasma proteins in the Fenland database. Ferkingstad *et al*.[20]: Plasma proteins in the deCODE database. IS: ischemic stroke, Fenland: Plasma proteins in the Fenland database. deCODE: Plasma proteins in the deCODE database. ISGC: IS GWAS data from International Stroke Genetics Consortium, FinnGen: IS GWAS data from FinnGen consortium. UKB: IS GWAS data from UK Biobank. Arrows: 95% upper and lower CI outside the range of the horizontal axis.

insignificant when the number of proteins included in the analysis was limited using various *P*-value thresholds. We explored all IS drugs (review in JAMA) and target-related information (Supplemental Table 7, Supplemental Digital Content 9, http://links.lww.com/JS9/B410). We discovered a connection between target protein of prekallikrein and medications of aspirin and edoxaban, target protein of MMP-12, and medications of alteplase and tenecteplase in the PPI protein network (Fig. 6 and Supplemental Figure 3, Supplemental Digital Content 10, http://links.lww.com/JS9/B411). One can find a strong correlation in prekallikrein with coagulation factor VII (F7), and MMP-12 with fibronectin (FN1) using the STRING website. Among them, STRING also demonstrated between prekallikrein and F7 experimentally determined, co-expression and protein homology, suggesting that these two proteins are related. Drugbank's identification of F7 as a potential therapeutic target for aspirin and edoxaban suggests that prekallikrein may also be a viable target. Furthermore, text mining in FN1 and MMP-12 also reveals intimate interaction between FN1 and MMP-12. A search on the Drugbank website revealed that FN1 are pharmacological targets for Alteplase and Tenecteplased, indicating that MMP-12 may be a possible target for these medications. Additionally, Supplementary Table 8 (Supplemental Digital Content 11, http://links.lww.com/JS9/B412) illustrates the clinical drug and application efficacy of MMP-12 and prekallikrein as pharmacological targets. Among them, lanadelumabis a new drug whose target protein is a prekallikrein, may be a promising new drug for the treatment of IS.

Discussion

This study investigated the possible causative link between plasma and CSF proteins and risk of IS and examined the potential therapeutic targets of IS in plasma and CSF proteins. SWAP70 and MMP-12 in plasma proteins and prekallikrein in CSF proteins, and IS risk were discovered to be causally related. Among them, the causal relationship between SWAP70 and IS was confirmed by FinnGen cohort, ISGC consortium, and UK biobank. While the causal relationship between MMP-12 and prekallikrein and IS is debatable in FinnGen cohort, ISGC consortiums, and UK biobank, further meta-analysis equally confirmed the causal relationship between MMP-12, prekallikrein, and risk of IS. Additionally, it was discovered that prekallikrein in CSF and MMP-12 in plasma may be a potential medication target

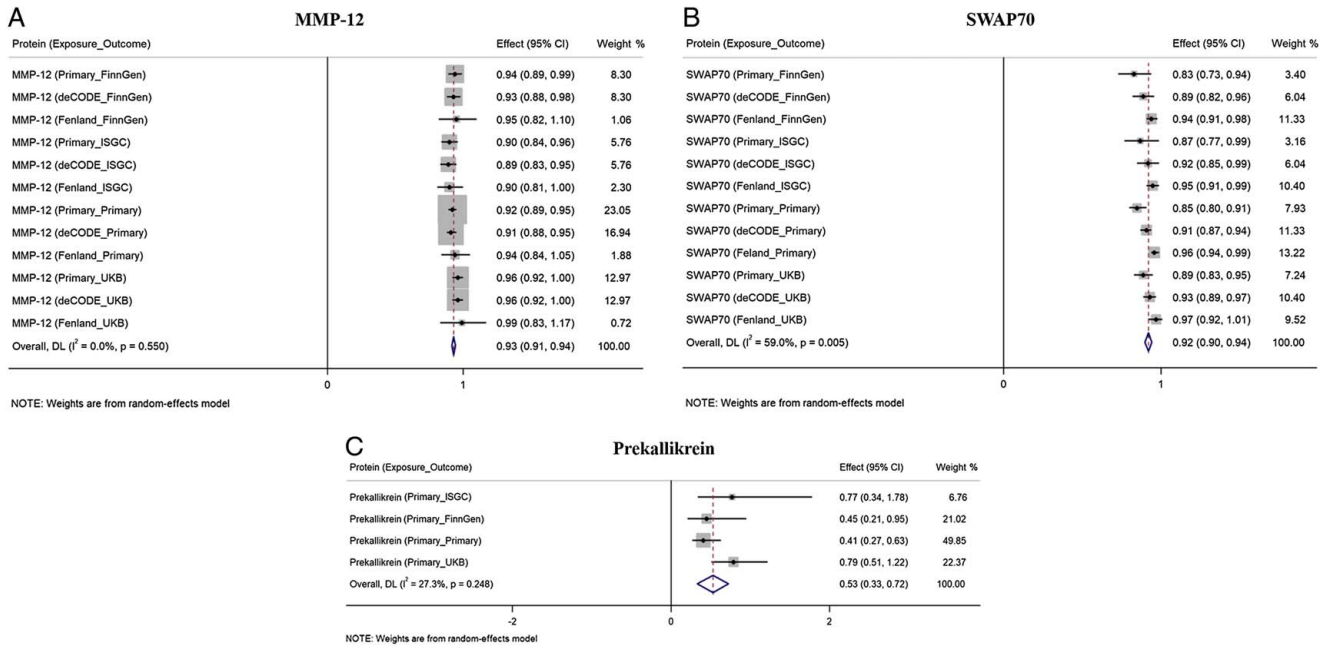

Figure 4. Meta-analysis of the causal relationship between MMP-12 (A), SWAP70 (B), and prekallikrein (C) and IS risk in the initial phase and the replication phase (external validation). CSF, cerebrospinal fluid; IS, ischemic stroke. Primary_: 734 plasma proteins and 151 CSF proteins used in the initial phase, _Primary: IS GWAS data from the MEGASTROKE database, Fenland: Plasma proteins in the Fenland database, deCODE: Plasma proteins in the deCODE database, ISGC: IS GWAS data from the ISGC database, FinnGen: IS GWAS data from the FinnGen cohort, UKB: IS GWAS data from UK Biobank.

for the current IS drugs. These will serve as the cornerstone for further pharmacological therapies and targeted IS prevention.

Drug targets of proteins are important clinical value. We also performed numerous sensitivity studies on the possible target proteins found in this investigation in order to ensure the validity of the study mentioned above. First, we used a combination of MR analysis and colocalization analysis is to establish a causal link between MMP-12, SWAP70, and prekallikrein and risk of IS. However, we conducted a bidirectional MR study and did not detect the presence of reverse causality, excluding the interference of reverse causality, in light of the possibility that this causal relationship could be reverse causality or bias caused by LD. Steiger filtering also supported the results mentioned previously. Although the phenotype scan revealed associations between MMP-12, SWAP70, and prekallikrein and other characteristics, none of these associations could fully explain the connection between the aforementioned proteins and IS. According to phenotype scans, SWAP70 (rs415895) was linked to hematocrit, coronary artery disease, and hypertension whereas MMP-12 (rs28381684) was linked to matrix metalloproteinase, postbronchodilator FEV1/FVC ratio, stromelysin-1, and local histogram emphysema pattern. Prekallikrein (rs2304595) has been connected to a number of proteins and coagulation factors (Supplementary Table 6, Supplemental Digital Content 7, http://links.lww.com/JS9/B408). We cannot entirely rule out the possibility that hypertension and coronary heart disease have a small causal role in the association between IS and SWAP70 because these conditions may be potential causes of IS. Therefore, more research is still required before SWAP70 can be identified as a drug target.

This study discovered a link between the plasma protein MMP-12 and risk of IS, which may lower the chance of having an IS. A previous MR study also suggested a potential causative relationship between MMP-12 (OR = 0.90, 95% CI: 0.86–0.94, $P = 7.43 \times 10^{-5}$) and the likelihood of suffering an IS[30]. Studies on the causative relationship between plasma proteins and risk of IS suggest that MMP-12 (OR = 0.83, 95% CI: 0.77–0.90, $P = 6.56 \times 10^{-6}$) may be causal mediators with large atherosclerotic stroke, a subtype of IS, and MMP-12 may be prospective pharmaceutical targets[12]. This may be connected to the

Table 2

Summary of phenotypic scanning, and reverse causality detection on the candidate causative proteins of MMP-12, SWAP70, and Prekallikrein.

Protein	Tissue	UniProt ID	SNP	Bidirectional MR (MR-IVW)	Steiger filtering	Colocalization PPH4	Previously reported associations
MMP-12	Plasma	P39900	rs28381684	0.73 (0.37-1.42)	Passed (2.49×10^{-101})	0.995	1. FEV1/FVC ratio 2.Stromelysin-1
SWAP70	Plasma	B3KUB9; E7EMB1; Q9UH65	rs415895	1.01 (0.95-1.08)	Passed (5.21×10^{-23})	0.987	blood cells hypertension (SBP/DBP)
Prekallikrein	CSF	P03952	rs2304595	No matching SNPs	Passed (8.96×10^{-9})	0.894	APTT Other proteins

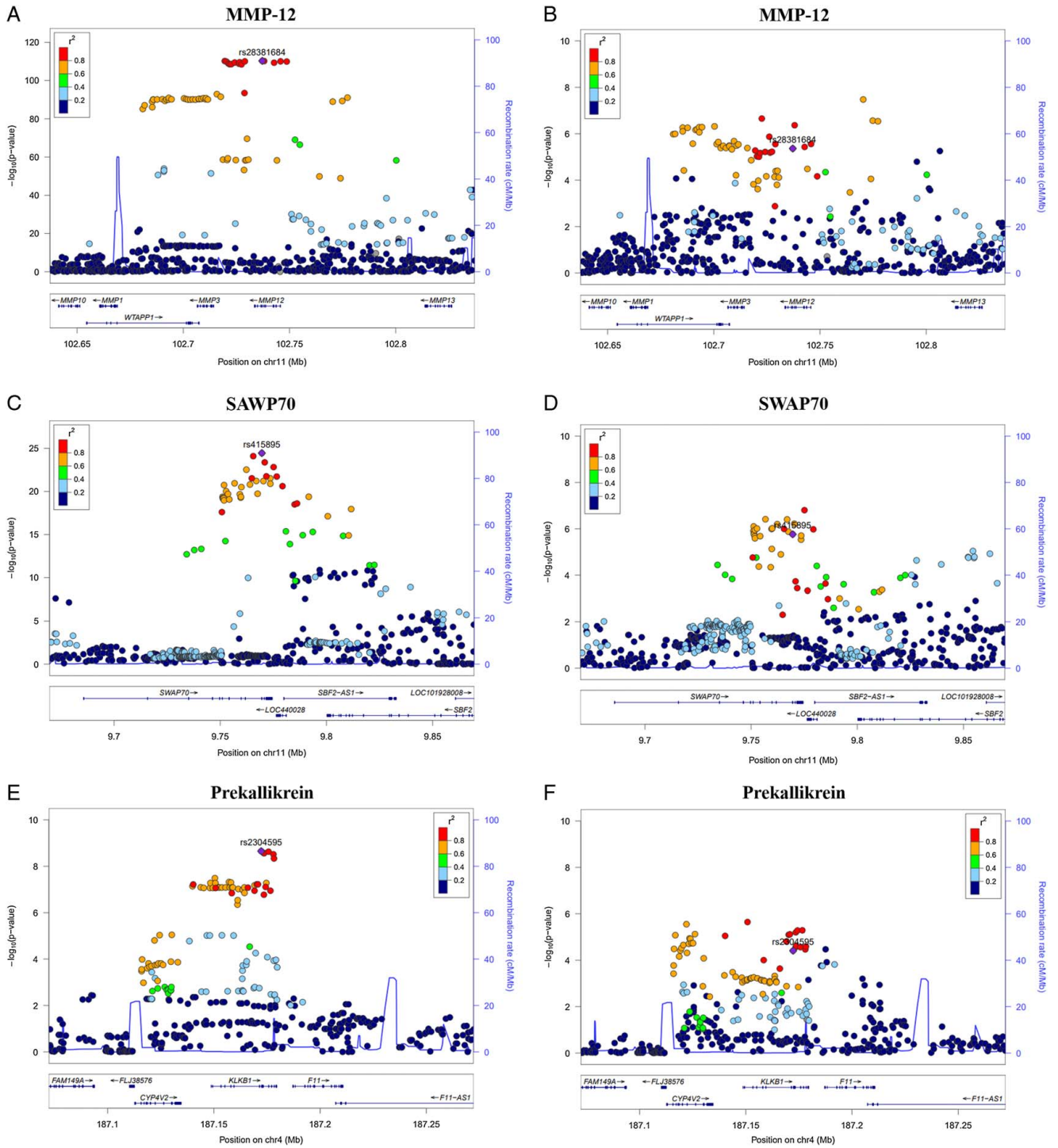

Figure 5. Regional association and matrix plots of colocalization analysis results. (A), (C), and (E) are the *P*-value distributions of the loci 100 KB above and below the pQTL for MMP-12, SWAP70, and Prekallikrein, respectively. (B), (D), and (F) are the *P*-value distributions of IS corresponding to the instrumental variables of MMP-12, SWAP70, and Prekallikrein. In (A–F), each point represents a SNP. Horizontal coordinates indicate the physical location on the chromosome, and the vertical axis indicates the -log *P*-value of GWAS of pQTL or IS.

possibility that MMP-12 protein is connected to the risk of IS through the breakdown of extracellular matrix elements, which in turn affects the stability of atheromatous plaque and impacts the onset of IS[31]. The connection between the medication and its target, MMP-12, is still unknown. In contrast, we discovered that the drugs alteplase, tenecteplase, and atorvastatin may be able to target MMP-12. According to our findings, FN1 may be a mechanism through, which MMP-12 binds to the drugs alteplase, tenecteplase, and atorvastatin. On the other hand, earlier research has demonstrated that MMP-12 has protein hydrolase

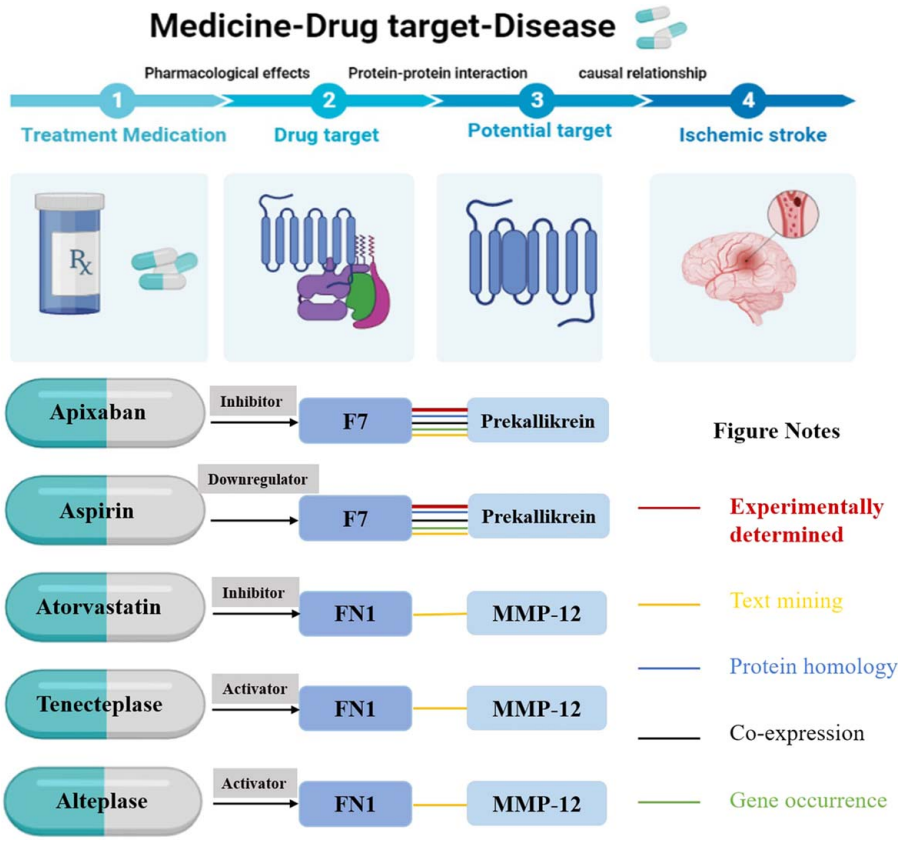

Figure 6. Interactions between identified prospective medication targets and the present IS drug targets.

activity that breaks down elastin, type IV collagen, and FN1 while activating MMP-2 and MMP-3[32]. Pro-MMP-1 and pro-MMP-9 can then be activated by pro-MMP-12, which also activates pro-MMP-2 and pro-MMP-3[33]. As they breakdown various elements of the microvascular basal lamina and BBB tight junction proteins, MMP-2 and MMP-9, further exercising their role as protein hydrolases, are major mediators of BBB disruption[34]. While MMP-12 hydrolysis of FN1 may contribute to thrombus dissolution and post-thrombolytic hemorrhagic transformation, FN1 is an important molecule involved in thrombus formation in vivo[32]. Additionally, MMP-12 and tissue-type fibrinogen activators interact directly or indirectly. By affecting MMP-9 and lowering expression of the proinflammatory M5 phenotype of microglia and macrophages (CD1, iNOS, IL-68, and TNF), tissue-type fibrinogen activator inhibition may attenuate the degradation of the BBB tight junction protein claudin-1, reducing the destruction of the BBB and neuroinflammation, which may then result in a decrease in MMP-12 expression[35]. As a result, the thrombolytic medications alteplase and tenecteplase work by causing FN1 to attach to the protein MMP-12. Additionally, circulating monocytes and neutrophils were both markedly decreased in atorvastatin-treated mice[33]. These changes decreased plaque inflammation, inhibited MMP-12 expression, produced a more stable plaque phenotype, and had a protective effect against IS[35].

In the exogenous coagulation pathway, coagulation factor VII is a key player, while prekallikrein is a molecule that prompts kallikrein to activate coagulation factor XII, which can encourage the endogenous coagulation pathway[36–39]. Prekallikrein and coagulation factor VII can work together to enhance the production of fibrin and the coagulation process[40]. Aspirin inhibits coagulation factor VII, which binds to a possible target of prekallikrein, is the drug's primary target. This allows aspirin to have antiplatelet aggregation effects[41]. A brand-new oral anticoagulant medication called apixaban works to stop the formation of thrombin and thrombosis by coagulation factor Xa, a crucial coagulation factor. Apixaban can downregulate coagulation factor VII and then inhibit coagulation factor X[14]. It may then bind to potential drug targets prebasic proteins and inhibit endogenous and exogenous coagulation pathways to produce anticoagulant effects.

There are some shortcomings in this study that may affect the interpretation and generalization of the findings. First, this investigation discovered that MMP-12 and prekallikrein may be possible therapeutic targets for alteplase, tenecteplase, atorvastatin, aspirin, and apixaban. However, we must recognize that the PPI only offers suggestive findings rather than firm conclusions. Therefore, additional research is required to thoroughly verify the relationships mentioned above. Second, because the plasma proteins employed in this inquiry were merged from five different investigations, there may have been variations in the measuring methods, experimental layout, etc. that were not taken into account. Nevertheless, this has no appreciable effect on how accurate our results are. We used plasma proteins from Fenland, the deCODE collaboration as exposure, and IS data from the ISGC and FinnGen cohort as outcome in order to further validate

the causal link identified in the initial phase. Moreover, we used meta-analysis to confirm the causal link that had been found in the initial phase. The results of the sensitivity test at this level could not be known because the instrumental variables included in this study primarily had only one cis-acting SNP and lacked trans-pQTLs, which would have affected the perform of pleiotropy analysis and heterogeneity test. However, the F-statistics of our so-included SNPs were all greater than 10, indicating that there was almost no weak instrumental variable bias. Moreover, the exposure and outcome data used in this study came from populations with European ancestry, and while we also validated the above three proteins using protein or disease data from British and Finnish populations, the application of the findings of this study to other regions, such as Asia, Africa, and the Americas, needs further research to confirm. Last but not least, despite the causality of MMP-12, SWAP70, prekallikrein and the risk of IS did not validate positive results in few populations—possibly due to ethnic differences, age groups of interest in the research, etc.—these results can still be found to be in a critical state, such as MMP-12_IS in deCODE_UKB (OR = 0.96, 95% CI: 0.92–1.01, $P = 0.072$), which was very close to positive results.

Conclusion

In conclusion, there is a causal link between the plasma proteins MMP-12, SWAP70, and the CSF protein prekallikrein and the risk of IS. MMP-12 and prekallikrein may be prospective drug targets for the IS medications. Follow-up IS clinical drug studies may consider drugs that act on the target proteins MMP-12 and prekallikrein.

Ethics approval and consent to participate

There was no need to get informed consent or ethical approval for this study again because all of the data were taken from published sources, and the informed consent and approval were received.

Consent for publication

There was no need to get informed consent or ethical approval for this study again because all of the data were taken from published sources, and the informed consent and approval were received.

Sources of funding

This study was supported by the National Science & Technology Fundamental Resources Investigation Program of China to L.Z. (No.2018FY100900), The Hunan Provincial Natural Science Foundation of China Grant to Y.Z. (No.2021JJ30923), The Provincial Science and Technology Innovation Leading Talents Project to L.Z. (No.2021RC4014), National Clinical Research Center for Geriatric Disorders (Xiangya Hospital), The Hunan Provincial Natural Science Foundation of China Grant to L.Z. (2023JJ60144) and Y.Z.(No.2021JJ30923), Central South University Research Program of Advanced Interdisciplinary Studies(2023QYJC011), Major Science and Technology Projects in Changsha, China (kq2301008).

Author contribution

X.Z.: conceptualization, methodology, software, validation, formal analysis, data curation, and writing – original draft; L.W.: writing – original draft, investigation, validation, and resources; S.W. and Y.Z.: writing – original draft, data curation, and formal analysis; J.M., L.C., Y.L., T.Y., H.Z., and L.W.: writing – original draft; X.Z., R.T., Y.Y., and R.L.: investigation and validation; Y.Z.: formal analysis, supervision, writing – review and editing; L.Z.: project administration, funding acquisition, supervision, writing – review and editing, and conceptualization.

Conflicts of interests disclosure

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Research registration unique identifying number (UIN)

Not applicable.

Guarantor

Le Zhang.

Provenance and peer review

Not applicable.

Availability of data and materials

The original contributions presented in the study are included in the article. Further inquiries can be directed to the corresponding author.

Acknowledgements

The authors thank the Genetic Investigation of deCODE Consortium, Fenland Consortium, FinnGen cohort, UK biobank, the MEGASTROKE GWAS dataset, and the International Stroke Genetics Consortium and all concerned investigators for sharing GWAS summary statistics on plasma proteins and CSF proteins and stroke.

References

[1] Global, regional, and national burden of stroke and its risk factors, 1990-2019: a systematic analysis for the Global Burden of Disease Study 2019. Lancet Neurol, 2021;20:795–820.
[2] Tao T, Liu M, Chen M, *et al*. Natural medicine in neuroprotection for ischemic stroke: challenges and prospective. Pharmacol Ther 2020;216:107695.
[3] Bansal S, Sangha KS, Khatri P. Drug treatment of acute ischemic stroke. Am J Cardiovasc Drugs 2013;13:57–69.
[4] Bhaskar S, Stanwell P, Cordato D, *et al*. Reperfusion therapy in acute ischemic stroke: dawn of a new era? BMC Neurol 2018;18:8.
[5] Roth GA, Mensah GA, Johnson CO, *et al*. Global burden of cardiovascular diseases and risk factors, 1990-2019: update from the GBD 2019 study. J Am Coll Cardiol 2020;76:2982–3021.

[6] Gawryś K, Turek-Jakubowska A, Gawryś J, *et al*. Platelet-derived drug targets and biomarkers of ischemic stroke-the first dynamic human LC-MS proteomic study. J Clin Med 2022;11.
[7] Turek-Jakubowska A, Dębski J, Jakubowski M, *et al*. New candidates for biomarkers and drug targets of ischemic stroke-a first dynamic LC-MS human serum proteomic study. J Clin Med 2022;11.
[8] Nelson MR, Tipney H, Painter JL, *et al*. The support of human genetic evidence for approved drug indications. Nat Genet 2015;47:856–60.
[9] Davey Smith G, Hemani G. Mendelian randomization: genetic anchors for causal inference in epidemiological studies. Hum Mol Genet 2014;23 (R1):R89–98.
[10] Lin J, Zhou J, Xu Y. Potential drug targets for multiple sclerosis identified through Mendelian randomization analysis. Brain 2023;146:3364–72.
[11] Bennett DA, Holmes MV. Mendelian randomisation in cardiovascular research: an introduction for clinicians. Heart 2017;103:1400–7.
[12] Chong M, Sjaarda J, Pigeyre M, *et al*. Novel drug targets for ischemic stroke identified through Mendelian randomization analysis of the blood proteome. Circulation 2019;140:819–30.
[13] Chen L, Peters JE, Prins B, *et al*. Systematic Mendelian randomization using the human plasma proteome to discover potential therapeutic targets for stroke. Nat Commun 2022;13:6143.
[14] Zheng J, Haberland V, Baird D, *et al*. Phenome-wide Mendelian randomization mapping the influence of the plasma proteome on complex diseases. Nat Genet 2020;52:1122–31.
[15] 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.
[16] Sun BB, Maranville JC, Peters JE, *et al*. Genomic atlas of the human plasma proteome. Nature 2018;558:73–9.
[17] 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;9:3268.
[18] Folkersen L, Fauman E, Sabater-Lleal M, *et al*. Mapping of 79 loci for 83 plasma protein biomarkers in cardiovascular disease. PLoS Genet 2017; 13:e1006706.
[19] Yang C, Farias FHG, Ibanez L, *et al*. Genomic atlas of the proteome from brain, CSF and plasma prioritizes proteins implicated in neurological disorders. Nat Neurosci 2021;24:1302–12.
[20] Ferkingstad E, Sulem P, Atlason BA, *et al*. Large-scale integration of the plasma proteome with genetics and disease. Nat Genet 2021;53:1712–21.
[21] Pietzner M, Wheeler E, Carrasco-Zanini J, *et al*. Mapping the proteo-genomic convergence of human diseases. Science 2021;374:eabj1541.
[22] Malik R, Chauhan G, Traylor M, *et al*. Multiancestry genome-wide association study of 520,000 subjects identifies 32 loci associated with stroke and stroke subtypes. Nat Genet 2018;50:524–37.
[23] Malik R, Traylor M, Pulit SL, *et al*. Low-frequency and common genetic variation in ischemic stroke: The METASTROKE collaboration. Neurology 2016;86:1217–26.
[24] Yuan S, Gill D, Giovannucci EL, *et al*. Obesity, type 2 diabetes, lifestyle factors, and risk of gallstone disease: a mendelian randomization investigation. Clin Gastroenterol Hepatol 2022;20:e529–37.
[25] Deng YT, Ou YN, Wu BS, *et al*. Identifying causal genes for depression via integration of the proteome and transcriptome from brain and blood. Mol Psychiatry 2022;27:2849–57.
[26] Zou X, Wang L, Zeng Y, *et al*. Illuminating the potential causality of serum level of matrix metalloproteinases and the occurrence of cardiovascular and cerebrovascular diseases: a Mendelian randomization study. J Hum Genet 2023.
[27] Hemani G, Tilling K, Davey Smith G. Orienting the causal relationship between imprecisely measured traits using GWAS summary data. PLoS Genet 2017;13:e1007081.
[28] Kamat MA, Blackshaw JA, Young R, *et al*. PhenoScanner V2: an expanded tool for searching human genotype-phenotype associations. Bioinformatics 2019;35:4851–3.
[29] Mendelson SJ, Prabhakaran S. Diagnosis and management of transient ischemic attack and acute ischemic stroke: a review. Jama 2021;325: 1088–98.
[30] Cárcel-Márquez J, Cullell N, Muiño E, *et al*. Causal Effect of MMP-1 (Matrix Metalloproteinase-1), MMP-8, and MMP-12 levels on ischemic stroke: a Mendelian randomization study. Stroke 2021;52:e316–20.
[31] Wang CY, Zhang CP, Li BJ, *et al*. MMP-12 as a potential biomarker to forecast ischemic stroke in obese patients. Med Hypotheses 2020;136: 109524.
[32] Yamamoto J, Ishii I, Okita N, *et al*. The differential involvement of von Willebrand factor, fibrinogen and fibronectin in acute experimental thrombosis in rat cerebral and mesenteric microvessels. Jpn J Physiol 1997;47:431–41.
[33] Matsumoto S, Kobayashi T, Katoh M, *et al*. Expression and localization of matrix metalloproteinase-12 in the aorta of cholesterol-fed rabbits: relationship to lesion development. Am J Pathol 1998;153:109–19.
[34] Challa SR, Nalamolu KR, Fornal CA, *et al*. Therapeutic efficacy of matrix metalloproteinase-12 suppression on neurological recovery after ischemic stroke: optimal treatment timing and duration. Front Neurosci 2022;16:1012812.
[35] Challa SR, Nalamolu KR, Fornal CA, *et al*. The interplay between MMP-12 and t-PA in the brain after ischemic stroke. Neurochem Int 2022;161: 105436.
[36] Schmaier AH. The contact activation and kallikrein/kinin systems: pathophysiologic and physiologic activities. J Thromb Haemost 2016;14: 28–39.
[37] Macfarlane RG. An enzyme cascade in the blood clotting mechanism, and its function as a biochemical amplifier. Nature 1964;202:498–9.
[38] Davie EW, Ratnoff OD. Waterfall sequence for intrinsic blood clotting. Science 1964;145:1310–2.
[39] Wheeler AP, Gailani D. The intrinsic pathway of coagulation as a target for antithrombotic therapy. Hematol Oncol Clin North Am 2016;30: 1099–114.
[40] Seligsohn U, Osterud B, Brown SF, *et al*. Activation of human factor VII in plasma and in purified systems: roles of activated factor IX, kallikrein, and activated factor XII. J Clin Invest 1979;64:1056–65.
[41] Raskob GE, Durica SS, Morrissey JH, *et al*. Effect of treatment with low-dose warfarin-aspirin on activated factor VII. Blood 1995;85:3034–9.