medRxiv preprint doi: https://doi.org/10.1101/2024.06.10.24308599; this version posted June 10, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

Identification of distinct and shared biomarker panels in different manifestations of cerebral small vessel disease through proteomic profiling

Ines Hristovska[1]; Alexa Pichet Binette[1]; Atul Kumar[1]; Chris Gaiteri[2,3]; Linda Karlsson[1]; Olof Strandberg[1]; Shorena Janelidze[1]; Danielle van Westen[4,5]; Erik Stomrud[1,6]; Sebastian Palmqvist[1,6]; Rik Ossenkoppele[1,7,8]; Niklas Mattsson-Carlgren[1,9,10]; Jacob W. Vogel[1,11]; Oskar Hansson[1,6]

Affiliations:

[1]Clinical Memory Research Unit, Department of Clinical Sciences Malmö, Lund University, Lund, Sweden

[2]Department of Psychiatry, SUNY Upstate Medical University, Syracuse, NY, USA

[3]Rush University Alzheimer's Disease Center, Rush University, Chicago IL, USA

[4]Diagnostic Radiology, Department of Clinical Sciences Lund, Lund University

[5]Imaging and Function, Skåne University Hospital, Lund, Sweden

[6]Memory Clinic, Skåne University Hospital, Malmö, Sweden

[7]Alzheimer Center Amsterdam, Neurology, Vrije Universiteit Amsterdam, Amsterdam UMC location VUmc, Amsterdam, The Netherlands

[8]Amsterdam Neuroscience, Neurodegeneration, Amsterdam, The Netherlands

[9]Department of Neurology, Skåne University Hospital, Lund, Sweden

[10]Wallenberg Center for Molecular Medicine, Lund University, Lund, Sweden

[11]Department of Clinical Sciences, Malmö, SciLifeLab, Lund University, Lund, Sweden

Correspondence: oskar.hansson@med.lu.se or ines.hristovska@med.lu.se

ABSTRACT

The pathophysiology underlying various manifestations of cerebral small vessel disease (cSVD) remains obscure. Using cerebrospinal fluid proximity extension assays and co-expression network analysis of 2,943 proteins, we found common and distinct proteomic signatures between white matter lesions (WML), microbleeds and infarcts measured in 856 living patients, and validated WML-associated proteins in three additional datasets. Proteins indicative of extracellular matrix dysregulation and vascular remodeling, including ELN, POSTN, CCN2 and MMP12 were elevated across all cSVD manifestations, with MMP12 emerging as an early cSVD indicator. cSVD-associated proteins formed a co-abundance network linked to metabolism and enriched in endothelial and arterial smooth muscle cells, showing elevated levels at early disease manifestations. Later disease stages involved changes in microglial proteins, associated with longitudinal WML progression, and changes in neuronal proteins mediating WML-associated cognitive decline. These findings provide an atlas of novel cSVD biomarkers and a promising roadmap for the next generation of cSVD therapeutics.

NOTE: This preprint reports new research that has not been certified by peer review and should not be used to guide clinical practice.

medRxiv preprint doi: https://doi.org/10.1101/2024.06.10.24308599; this version posted June 10, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

INTRODUCTION

Manifestations of cerebral small vessel disease (cSVD) have emerged as the leading vascular contributors to cognitive decline and subsequent dementia[1], and as major risk factors for stroke[2]. cSVD can act as both a primary etiology and as a secondary pathology in numerous neurodegenerative conditions. These manifestations stem from a variety of pathological processes that affect the brain's small arteries, arterioles, venules, and capillaries in the brain, leading to structural and functional abnormalities within the brain's microvascular network[3]. Although lifestyle changes and controlling vascular risk factors can reduce the risk for cSVD[4], the development of more targeted preventative and therapeutic strategies is impeded due to the relatively limited understanding of the molecular pathophysiology of cSVD.

At present, clinical detection of cSVD relies strongly on magnetic resonance imaging (MRI), which identifies markers such as white matter lesions (WML), subcortical infarcts and microbleeds[5]. Neuropathological studies and model systems have helped to characterize isolated pathways contributing to the etiology of cSVD[6,7,8], but the larger landscape and timeline of altered pathways in living humans has been elusive. While existing literature offers insights into pathophysiological mechanisms for cSVD, including endothelial dysfunction, BBB breakdown and inflammation, these findings have not been consistently replicated, requiring further validation across larger cohorts with a broader spectrum of markers[9,10]. Associative evidence exists linking certain lesions together and with other diseases, but ultimately, we have little insight into their underlying etiology[11]. Additionally, MRI lesions related to cSVD pathology are likely indicative of later stages in the pathological processes, underscoring the necessity of investigating the early changes that may precede these visible abnormalities. Such gaps in our knowledge emphasize the need for investigation using richer biomarker panels and larger cohorts to provide deeper insight in the pathophysiological processes involved in cSVD. Discovering robust markers and defining molecules related to earliest stages of disease will be crucial to the development and monitoring of therapeutic interventions.

Recent progress in highly sensitive and high throughput protein measurement techniques now allow for a comprehensive large-scale proteomic profiling in cerebrospinal fluid (CSF) and blood[12–14]. These advances enable identification of novel biomarkers and delineate molecular mechanisms associated with disease pathogenesis *in vivo*[15,16]. In our study, we utilized proximity extension assay (Olink) proteomics technology to extensively assess pathophysiological pathways concurrent with cSVD manifestations in the CSF of several large population and patient cohorts. Our principal aims were i) to identify both common and unique CSF proteome signatures associated with WML, microbleeds and infarcts; ii) to provide in-depth characterization of the biological underpinnings of those signatures; iii) to identify proteins implicated in WML progression and mediating WML-associated cognitive decline, and iv) to validate novel biomarkers in independent cohorts.

medRxiv preprint doi: https://doi.org/10.1101/2024.06.10.24308599; this version posted June 10, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

RESULTS

We analyzed cerebrospinal fluid (CSF) protein levels for 2,943 proteins in 856 participants from the Swedish BioFINDER-2 cohort, ranging from cognitively unimpaired individuals to patients with mild cognitive impairment and dementia (Extended Data Fig. 1). Participants were stratified into groups reflecting the presence or absence of cSVD pathology, namely WML, microbleeds or infarcts. We distinguished 571 participants without WML pathology and 285 with WML pathology (cut-off for WML (% of icv) = 0.508, Extended Data Fig. 2). Similarly, individuals were classified based on the presence of microbleeds (675 without and 143 with ≥1 microbleed(s)) and cerebral infarcts (790 without and 66 with ≥1 infarct(s)). Extended Data Table 1 indicates associations between the presence of each cSVD pathology and various demographic and clinical measures. As expected, participants with cSVD manifestations tended to be older, more likely male, more likely demented and were more likely diagnosed with vascular risk factors.

Existence of shared and distinct proteomic signatures in different cSVD manifestations

We first identified differentially abundant proteins (DAPs) in individuals exhibiting WML pathology (n=424), microbleeds (n=45), or infarcts (n=41, pFDR<0.05, Fig. 1a). Almost all associations remained significant after adjusting for co-morbid Alzheimer's or Parkinson's disease (Extended Data Fig. 3). Since we aimed to identify and study proteins specifically associated with each cSVD manifestation, proteins were retained for further analysis only if their association with WML pathology, microbleeds, or infarcts, respectively, persisted when adjusting for the other two manifestations of cSVD (e.g., proteins associated with WML when adjusting for presence of microbleeds and infarcts). Most proteins remained significant after adjusting for other cSVDs (Fig. 1b). Relative to microbleeds and infarcts, a large part of the proteome exhibited alterations in the presence of WML. Summary statistics from the differential expression analyses can be found in Supplementary Table 1.

Several proteins, including MMP12, PDLIM7, POSTN, ELN and CCN2, showed a significant *increase* in abundance across all three examined manifestations of cSVD independently (Fig. 1c). However, the expression of many proteins was found to be specific to a certain cSVD manifestation, notably in WML, but also in microbleeds. For instance, several matrix metalloproteinases, including MMP1 and MMP8, were exclusively elevated in WML. KYNU and ITGAM, which are linked to immune response, were found to be increased only with microbleeds. PMCH and TNFRSF10A, associated with stress response and apoptosis respectively, were upregulated only in infarcts. This differential protein expression pattern with little shared upregulation between microbleeds and infarcts is in agreement with the hypothesis that these manifestations may have distinct underlying etiologies. On the other hand, some proteins were found to be increased in WML as well as in either microbleeds or infarcts, as illustrated in Fig. 1d. A similar pattern was observed for proteins that were *downregulated*, but no proteins were consistently decreased across all forms of cSVD. Compared to microbleeds and infarcts, many proteins were downregulated in the context of WML, with CBLN4 and LRTM2 showing substantial negative associations (Fig. 1e).

We next sought to investigate how key proteins fluctuated in response to continuous (as opposed to binary) measures of WML (Fig. 1f) and microbleed (Fig. 1g) loads, respectively.

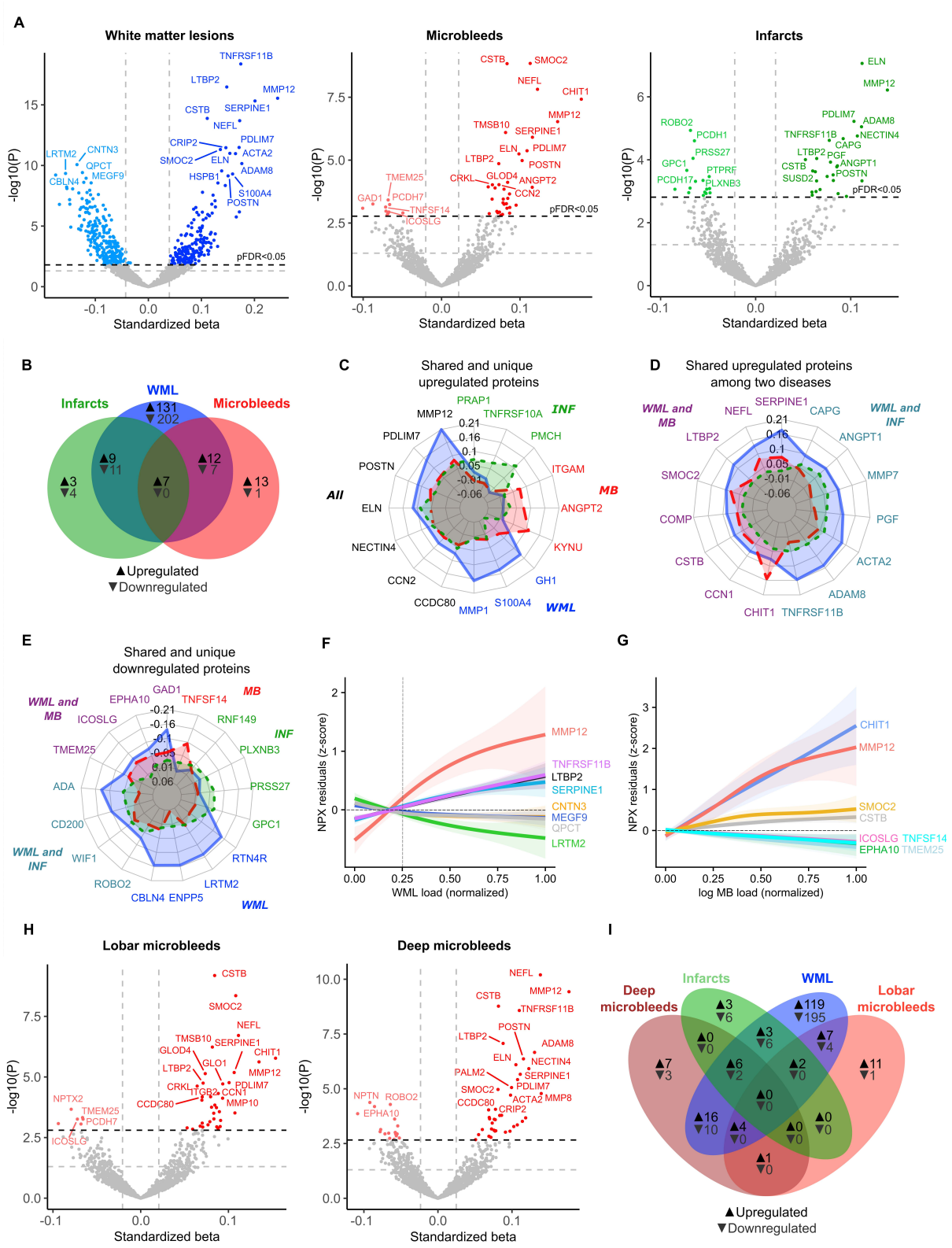
medRxiv preprint doi: https://doi.org/10.1101/2024.06.10.24308599; this version posted June 10, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

Among the proteins that showed a gradual decrease, LRTM2 exhibited the most notable reduction as the severity of WML increased. MMP12 levels rose sharply in both conditions as pathology became more pronounced, with CHIT1 levels following the same trend specifically with increased microbleed load.

Given that the distinct topography of microbleeds is considered to reflect their underlying etiology, we explored whether variations in their locations were mirrored by specific protein expression profiles (Fig. 1h). Our analysis identified proteins predominantly associated with either deep or lobar microbleeds (Fig. 1i). While most proteins were unique to one type, a few such as KYNU were associated with both, while others like CSTB and SMOC2 were also linked to WML. ITGB2, ITGAM and TMSB10 were unique to lobar microbleeds, whereas TNFRSF12A and FABP2 were exclusive to deep microbleeds. Interestingly, DAPs related to deep microbleeds frequently overlapped with those found in WML, and to a lesser extent, infarcts – a pattern not common for lobar microbleeds.

medRxiv preprint doi: https://doi.org/10.1101/2024.06.10.24308599; this version posted June 10, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

Figure 1. Differential protein expression in subjects with WML, microbleeds, and infarcts



Legend: **(a)** Volcano plots showing differentially abundant proteins (DAPs) when comparing individuals with and without specific cSVD pathology: WML (N=571 vs N=285), microbleeds (N=675 vs N=143), and infarcts (N=790 vs N=66). The models are adjusted for age, sex, and average protein level. The dashed lines represent significance threshold at α=0.05 before (gray) and after (black) FDR correction. Proteins below the p[FDR]<0.05 threshold were considered significant. For clarity, only the

medRxiv preprint doi: https://doi.org/10.1101/2024.06.10.24308599; this version posted June 10, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

top 20 proteins are labeled. **(b)** Venn diagram illustrating the shared and exclusive proteins associated with WML, microbleeds, and infarcts, with proteins of increased abundance indicated in black and those with decreased abundance in gray. **c,e,** Radar charts depicting standardized beta coefficients for top proteins, in relation to their association to disease: **(c)** shared and disease-specific upregulated proteins; **(d)** proteins commonly upregulated between two conditions; and **(e)** shared and unique downregulated proteins. Each axis represents a different protein, with axis length indicating the standardized beta coefficients magnitude for each condition. **f,g,** Proteins of interest from panel **(a)**, selected based on disease relevance or expression pattern, plotted against normalized WML load **(f)** and microbleed burden **(g)**. Shaded areas correspond to the 95% confidence interval. The vertical line indicates the cutoff for WML positivity. **(h)** Volcano plots showing differentially abundant proteins (DAPs) when comparing individuals with and without lobar and deep microbleeds. The models are adjusted for age, sex, and average protein level. **(i)** Venn diagram illustrating the shared and exclusive proteins with increased abundance associated with WML, deep microbleeds, lobar microbleeds, and infarcts.

cSVD is associated with greater neurovascular and lesser neuronal cellular responses

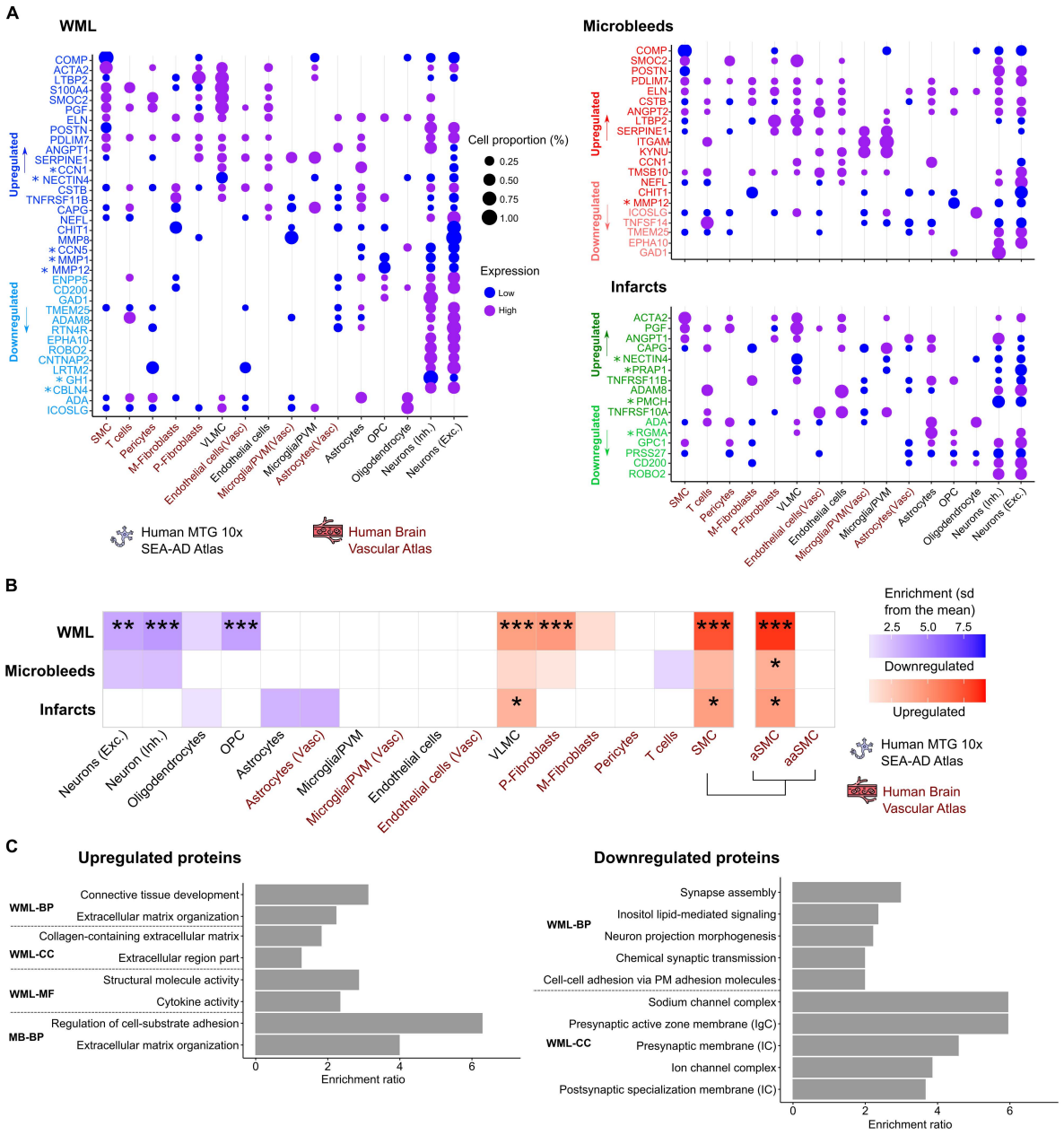
To better understand the specific cellular mechanisms underlying the different cSVD manifestations, we sought to uncover the cell-type specificity of DAPs and their biological roles. For cell-type specificity, we used Expression Weighted Cell Type Enrichment (EWCE) on single-cell transcriptome data from the SEA-AD atlas and the Human Brain Vascular Atlas, the latter providing greater granularity due to the vascular underpinnings of cSVD.

Proteins *upregulated* across all three manifestations of cSVD predominantly originated from vascular-associated cells, particularly vascular leptomeningeal cells (VLMC) and smooth muscle cells (SMC) (Fig. 2a, 2b). Further dissection of DAPs in smooth muscle cell populations showed a significant enrichment specifically in arterial (as opposed to arteriolar) SMC associated with all cSVD manifestations. WML-associated proteins were uniquely enriched in perivascular fibroblasts. Conversely, *downregulated* proteins in WML were predominantly associated with neuronal cells – both excitatory and inhibitory – as well as oligodendrocyte precursor cells (OPCs), indicating a distinct cellular response within the neuronal and oligodendrocyte populations.

Building on these cellular insights, we sought to understand the biological significance of DAPs using functional enrichment analysis with the Gene Ontology (GO) database. All significant GO terms are listed in Supplementary Table 2. Upregulated proteins in WML and microbleeds, previously linked to vascular cells, were involved in extracellular matrix organization, with microbleed-specific proteins further enriched in cell-substrate adhesion (Fig 2c). For WML, the analysis pointed to an extracellular localization of enriched proteins, suggesting their contribution to maintaining structural integrity and mediating cytokine signaling. In contrast, downregulated proteins in WML were related to synaptic functions and transmission. The proteins associated with infarcts were not enriched for any GO terms.

medRxiv preprint doi: https://doi.org/10.1101/2024.06.10.24308599; this version posted June 10, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

Figure 2. Cell-type and functional enrichment characterization of DAPs



Legend: **(a)** Proportion of mRNA expression by cell type from single-cell transcriptomics data from the Human MTG 10x SEA-AD and the Human Brain Vascular Atlas for a selection of significant proteins for WML, microbleeds and infarcts. The proteins that are denoted with an asterisk do not have information about mRNA expression in the Human Brain Vascular Atlas. Gene expression was classified based on Gaussian Mixture Model (GMM)-derived cutoffs for each cell type, with expression above the cutoff marked as high, while those below are indicated as low. **(b)** Cell-type enrichment analyses for upregulated and downregulated proteins based on single-cell transcriptomics data from the Human MTG 10x SEA-AD and the Human Brain Vascular Atlas for different DAP categories using the Expression Weighted Cell Type Enrichment (EWCE) package. The enrichment for upregulated proteins is shown in a palette of red, while the enrichment for downregulated proteins is shown in a palette of blue. **(c)** Top 5 summary terms from functional enrichment analyses using Gene Ontology (GO) databases (Biological Processes, Cellular Components and Molecular Functions) for different DAP categories. For cell-type and functional enrichment analyses, the 1331 Olink proteins were used as background.

*corresponds to $p_{FDR}<0.05$, **corresponds to $p_{FDR}<0.01$, ***corresponds to $p_{FDR}<0.001$.

medRxiv preprint doi: https://doi.org/10.1101/2024.06.10.24308599; this version posted June 10, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

Abbreviations: OPC = oligodendrocyte precursor cells, PVM = perivascular macrophages, VLMC = vascular leptomeningeal cell, P-Fibroblasts=perivascular fibroblasts, M-Fibroblasts = meningeal fibroblasts, SMC = smooth muscle cell, aSMC = arterial smooth muscle cell, aaSMC=arteriolar smooth muscle cell, BP = Biological Processes, CC = Cellular Components, MF = Molecular Functions, IgC = Integral Component, IC = Intrinsic Component, PM = Plasma Membrane

cSVD is associated with alterations in prominent metabolic, immune and neuronal processes along the cSVD continuum

Changes in protein abundance often occur in concert with changes in other proteins as part of complex molecular responses. Therefore, we sought to establish whether the changes in protein abundance that we observed in association with cSVD pathology occurred in isolation or were embedded in broader molecular networks. To identify such networks, we applied the SpeakEasy2[17] network-based clustering algorithm to the available CSF proteomics data and used this information to further annotate our cSVD-associated proteins. We focused on the first 10 modules of co-abundant proteins (Fig. 3a), and the assignment of proteins to each module is found in Supplementary Table 1.

Modules consisted of varying number of proteins, ranging from 16 to 222. Modules 1-5 and 10 were characterized by proteins that showed an increase in abundance in cSVD, while modules 6-9 were enriched in proteins exhibiting mostly decreased abundance (Fig. 3b, 3c).

To characterize the protein co-abundance modules, we first assessed cell-type specificity and involvement in biological processes through comprehensive enrichment analysis (Fig. 3d, Supplementary Table 3). Subsequently, to contextualize these modules within the disease framework, we examined their associations with cSVD manifestations (Fig. 3e). Finally, we charted the average protein levels from each significant module against WML and microbleed severity to deepen our understanding of the modules' dynamics across cSVD pathology load (Fig. 3f-3g). Modules 5 and 10 showed robust *positive* correlations across all three cSVD manifestations. Module 5 exhibited a notable enrichment of proteins predominantly located in endothelial and arterial smooth muscle cells, participating in the metabolism of nitrogenous catabolic compounds and heterocyclic molecules, including several of our key proteins, like MMP12, PDLIM7 and SERPINE1. In contrast, Module 10's profile remained undefined due to non-significant enrichment after multiple comparison correction, but included proteins showing strong association with cSVD, such as SMOC2. A marked increase in protein levels in these two modules were observed even at lower levels of WML and microbleed accumulation (Fig. 3f-3g).

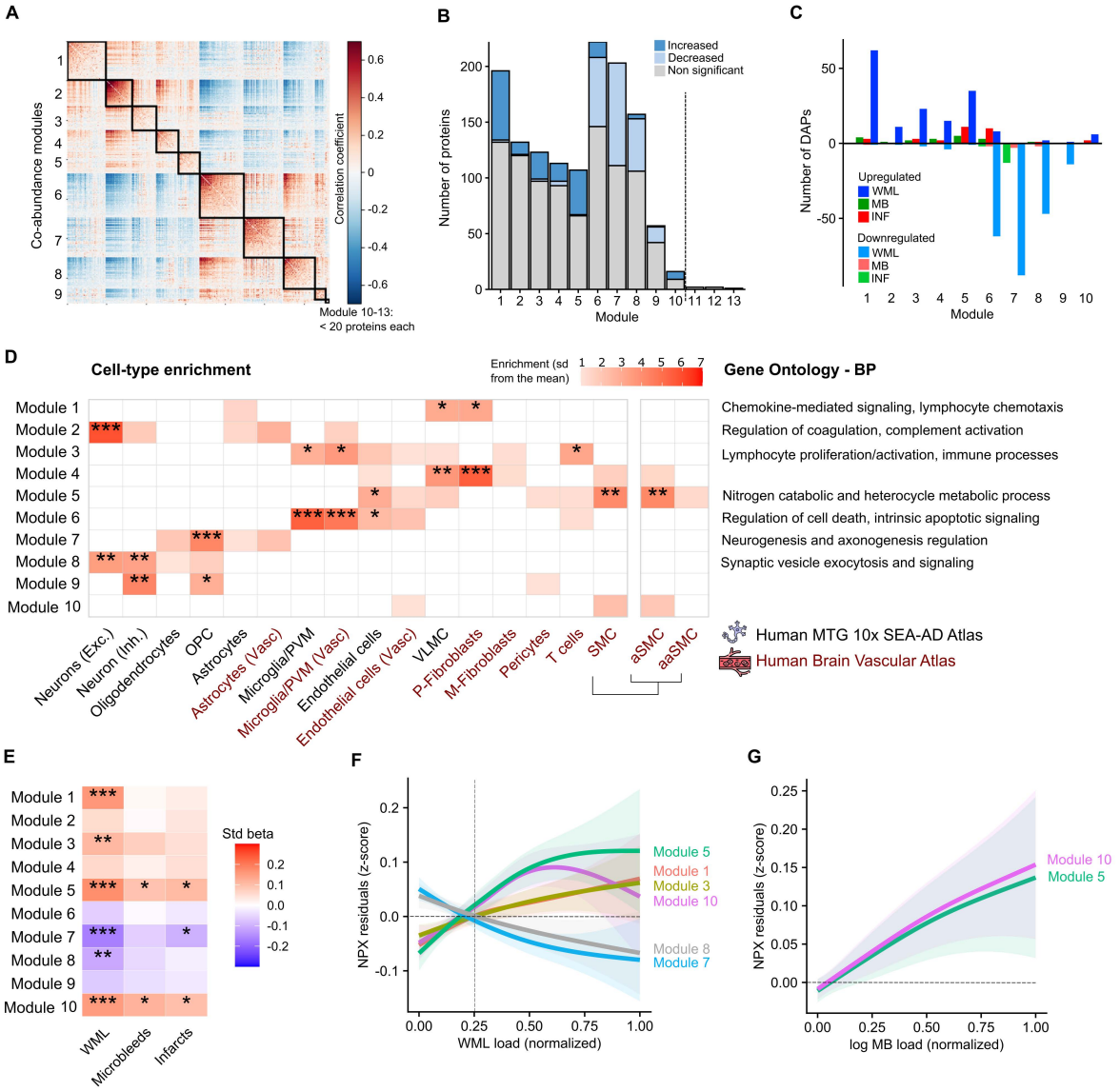
Both WML and infarcts showed a *negative* association with Module 7, enriched in proteins found in OPCs, and involved in the regulation of neurogenesis and axonogenesis. White matter lesions were also negatively associated with Module 8, enriched in proteins involved in synaptic vesicle exocytosis and signaling in excitatory and inhibitory neurons. For both module 7 and 8, we observed a progressive lowering of protein levels with increasing WML pathology.

Alongside Module 8, Modules 1 and 3 showed associations exclusively with WML. The proteins in Module 1 were predominantly enriched in vascular leptomeningeal cells and perivascular fibroblasts, in line with our previous findings as to their significant association with WML and were involved in chemokine-mediated signaling pathways and lymphocyte chemotaxis. We also uncovered an association with Module 3, containing proteins enriched in

medRxiv preprint doi: https://doi.org/10.1101/2024.06.10.24308599; this version posted June 10, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

microglia/perivascular macrophages and T cells, involved in immune processes, including the regulation of lymphocyte activation and proliferation. Module 1 and 3 followed a similar pattern, showing a steady rise without reaching a plateau, even at the highest levels of white matter lesion load. In summary, we identified a complex interplay between molecular networks and cSVD pathology, shedding light on the dysregulations in metabolic, vascular, immune, and neuronal processes within cSVD.

Figure 3. Clustering algorithm-derived modules of protein co-abundance and their association with cSVD



Legend: **(a)** Heatmap depicting protein-protein correlation matrix within the Olink data. Black outlined boxes indicate co-abundance modules, with modules 10-13 comprising fewer than 20 proteins each. **(b)** Distribution of proteins across co-abundance modules. Bar heights represent the number of proteins in each module, categorized by changes in abundance in all three cSVD manifestations combined: increased (dark blue), decreased (pale blue), and non-significant (gray). Modules 11 to 13, demarcated by a dashed line, contain few proteins, and are not used in subsequent analysis. **(c)** Bar chart illustrating the module-wise distribution of DAPs across various cSVD conditions. Upregulated proteins are represented above the horizontal axis, and downregulated proteins are shown below the axis in corresponding colors for each condition. **(d)** Heatmap displaying cell-type enrichment across protein co-abundance modules, using single-cell transcriptomics data from the Human MTG 10x SEA-AD and

medRxiv preprint doi: https://doi.org/10.1101/2024.06.10.24308599; this version posted June 10, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

the Human Brain Vascular Atlas. Adjacent to the heatmap, key biological processes identified through Gene Ontology (GO) – Biological Processes (BP) enrichment analysis are listed for each module. **(e)** Heatmap of standardized beta coefficients illustrating associations between biological modules and WML, microbleeds and infarcts. **f,g,** Trajectories of average protein expression within significant modules as a function of normalized **(f)** WML load and **(g)** microbleed load, showcasing the relationship between protein abundance and disease burden. Shaded areas correspond to the 95% confidence interval. The vertical line indicates the cutoff for WML positivity.
*corresponds to $p_{FDR}<0.05$, **corresponds to $p_{FDR}<0.01$, ***corresponds to $p_{FDR}<0.001$

Crucial contribution of immune-related proteins in the progression of WML over time

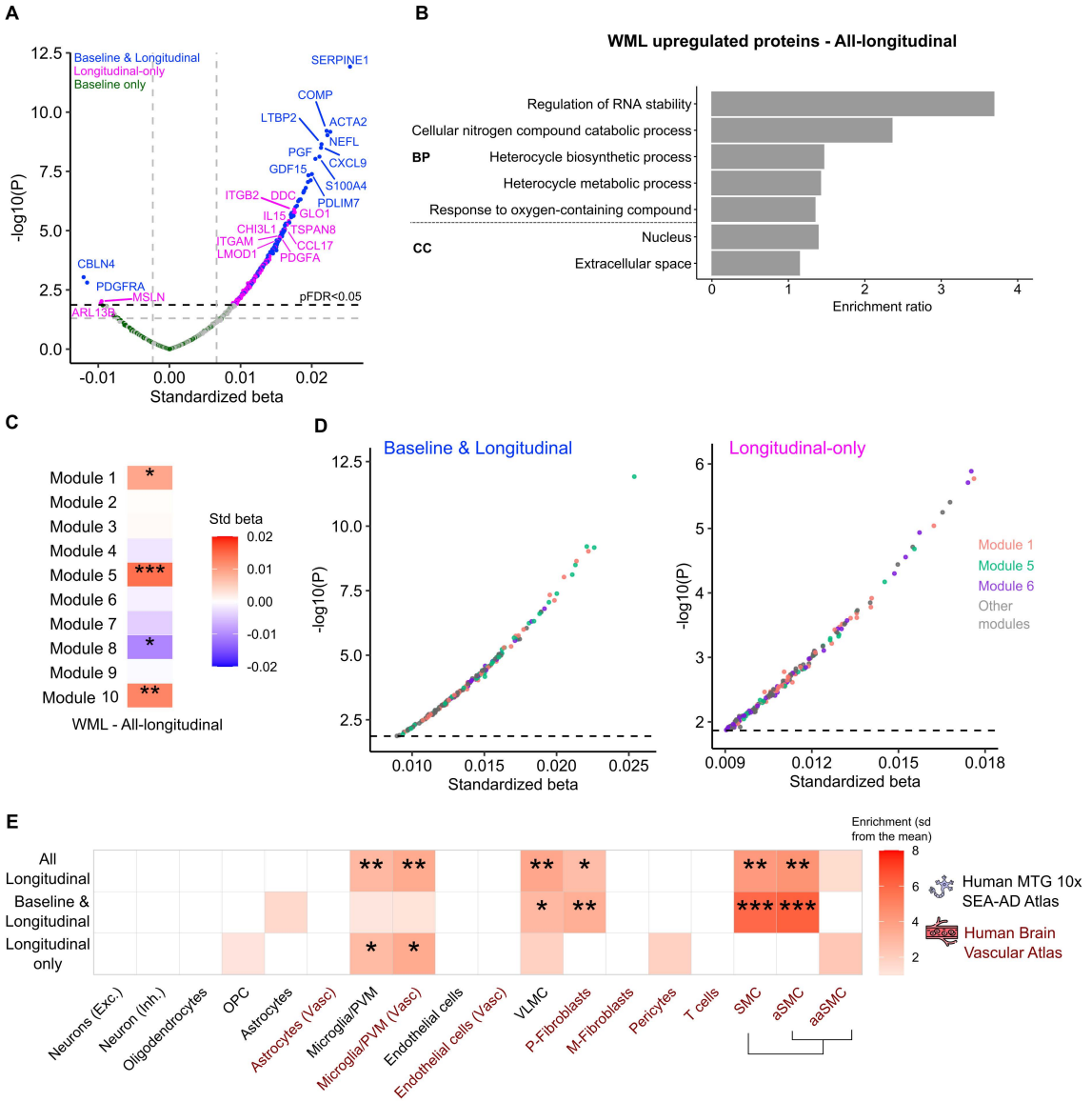
Our data expectedly showed an overall increase in WML volume over time in subjects with longitudinal follow-up (n=571 subjects, average follow-up time=2.67±0.93, Extended Data Fig. 4a). We proceeded in identifying the proteins at baseline that were associated with the rate of longitudinal increase in WML burden by applying linear mixed-effect models. This analysis revealed that abundance of several proteins at baseline was associated with faster increase in WML progression over time (Fig. 4a, Supplementary Table 4). To gain a deeper insight into the role these proteins play in WML progression, we divided them into two categories: associated with both at baseline and longitudinal WML (shared, in blue), and those associated with longitudinal WML only (in magenta). Approximately half of the longitudinal WML-associated proteins were also associated with WML at baseline, indicating sustained upregulation, with SERPINE1, LTBP2 and NEFL being prominent among these proteins. In contrast, immune-related proteins, such as ITGB2, IL15 and CHI3L1 showed strong associations only with the rate of longitudinal WML progression, but not WML at baseline.

Functional enrichment analysis using all longitudinal-associated proteins emphasized the metabolic breakdown of cellular nitrogen compounds, as well as the biosynthesis and metabolism of heterocyclic structures, aligning closely with Module 5 (Fig. 4b). Subsequently, we found a strong association with Module 5 and WML progression (Fig. 4c). To better understand the relationship between the modules and the two groups of DAPs associated with WML progression, shared (left panel) and longitudinal-only (right panel) cSVD-associated proteins were color-coded based on their respective co-abundance module, in particular Module 1, 5 and 6 (Fig. 4d, all modules color-coding at Extended Data Fig. 4b). We found a distinct pattern: proteins with the most prominent effects in the shared group tended to cluster within Module 5, while longitudinal-only proteins frequently clustered into Module 6. The Module 6 proteins were strongly associated with microglia and perivascular macrophages and the regulation of apoptotic signaling pathways. A two-proportion z-test confirmed a significant disparity between the shared and longitudinal-only groups in the presence of Module 5 and Module 6 proteins. Furthermore, when distinguishing between the shared and longitudinal-only, we found that the shared proteins were significantly enriched in VLMC (per the SEA-AD Atlas) and perivascular fibroblasts and smooth muscle cells (per the Human Vascular Atlas) (Fig. 4e). On the other hand, longitudinal-only proteins were significantly enriched in microglial and macrophage populations.

In parallel, we aimed to identify proteins linked to progression in microbleed pathology, but no proteins exhibited a significant association (Extended Data Fig. 5). This lack of significant associations may be attributed to the modest cohort size of subjects exhibiting an increase in the microbleed counts over the observed period (n=49).

medRxiv preprint doi: https://doi.org/10.1101/2024.06.10.24308599; this version posted June 10, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

Figure 4. Differential abundance, cell-type and functional enrichment, and module characterization in WML progression



Legend: **(a)** Volcano plots showing differentially abundant proteins (DAPs) associated at baseline with WML progression over time. Linear mixed effect models were adjusted for age, sex, average protein level, and intracranial volume, with a protein*time interaction included in the model. The dashed lines represent p<0.05 (gray) and pFDR<0.05 (black) significance thresholds, and proteins below the pFDR<0.05 threshold were considered significant. Protein color-coding indicates significance relative to WML at baseline: proteins associated with WML both at baseline and longitudinally are highlighted in blue; those exclusively associated with longitudinal WML are in magenta, and proteins associated solely with WML at baseline are shown in dark green. For clarity, only the top upregulated and downregulated proteins from the first two groups are labeled. **(b)** Summary terms from functional enrichment analyses using Biological Processes and Cellular Components from Gene Ontology (GO) database for all upregulated DAPs. **(c)** Heatmap of standardized beta coefficients from a linear mixed-effects model showing the associations between biological modules and WML change over time. **(d)** Significant shared (e.g. associated with cross-section and longitudinal WML) proteins (left panel) and longitudinal-only (right panel) were color-coded according to the modules they were assigned to in Figure 4. **(e)** Cell-type enrichment analyses based on single-cell transcriptomics data from the Human MTG 10x SEA-AD and the Human Brain Vascular Atlas for different DAP categories using the

medRxiv preprint doi: https://doi.org/10.1101/2024.06.10.24308599; this version posted June 10, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

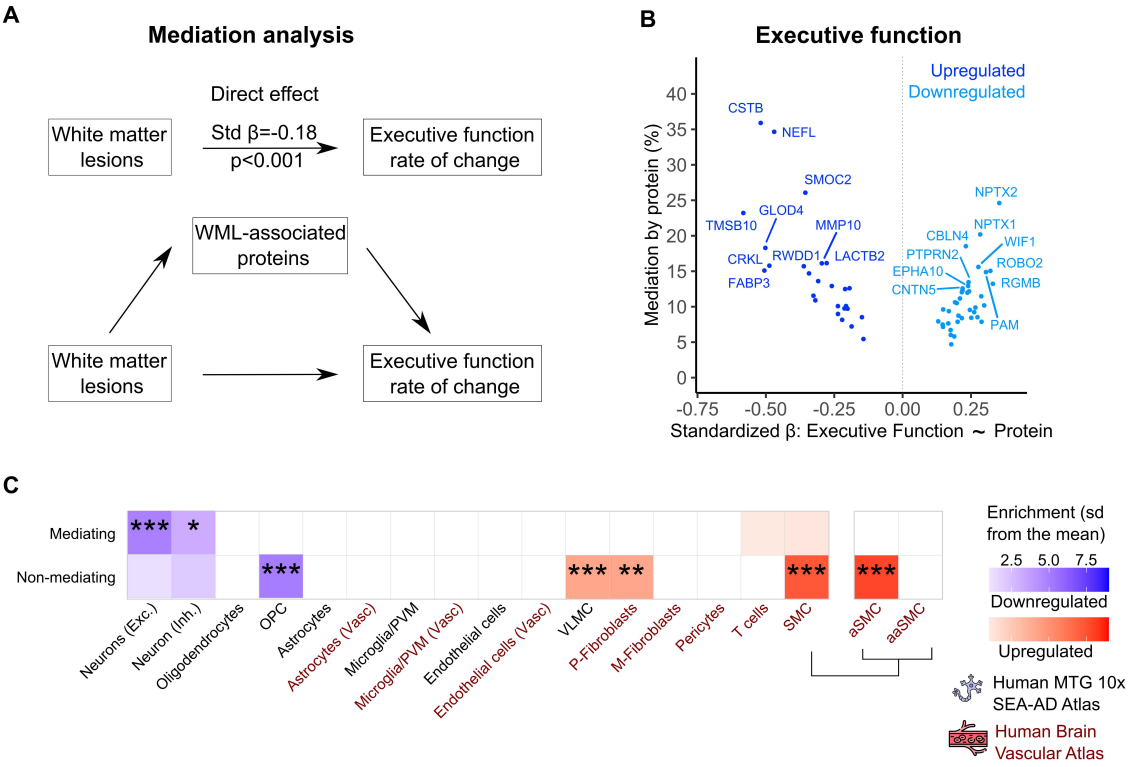
Expression Weighted Cell Type Enrichment (EWCE) package. For cell-type and functional enrichment analyses the 1331 Olink proteins were used as background.
*corresponds to $p_{FDR}<0.05$, **corresponds to $p_{FDR}<0.01$, ***corresponds to $p_{FDR}<0.001$.
Abbreviations: OPC = oligodendrocyte precursor cells, PVM = perivascular macrophages, VLMC = vascular leptomeningeal cell, P-Fibroblasts=perivascular fibroblasts, M-Fibroblasts = meningeal fibroblasts, SMC = smooth muscle cell, aSMC = arterial smooth muscle cell, aaSMC=arteriolar smooth muscle cell, BP = Biological Processes, CC = Cellular Components

Neuronal proteins partly mediate the association between WML and cognitive decline

Since WML were associated with cognitive decline, especially executive functions (Fig. 5a), we next assessed the possible mediation of DAPs on the association between WML at baseline and longitudinal cognitive decline. We studied the impact of these proteins on changes in executive function (Trail Making Test-A (TMT-A), Trail Making Test-B (TMT-B) and the Symbol Digit Modalities Test (SDMT), Fig. 5, Supplementary Table 5) and global cognition (the modified Preclinical Alzheimer Cognitive Composite, mPACC, Extended Data Fig. 6). We found that proteins like CSTB and NEFL, which are upregulated in WML, alongside ECM and cell adhesion-related proteins (SMOC2, MMP10, TMSB10), exhibited a significant partial mediation effect on rate of executive function decline, each contributing over 15% to the mediating effect (Fig. 5b). Similarly, key downregulated proteins associated with synaptic function, such as NPTX2, NPTX1, and CBLN4, were major contributors to this mediation. Cell specificity analysis of proteins that were negatively associated with WML and that mediated cognitive decline were predominantly neuronal, including glutamatergic and GABAergic subtypes (Fig. 5c). In contrast, proteins negatively associated with WML that had no mediating effect were primarily observed in OPCs. While proteins positively associated with WML and mediating cognitive decline did not exhibit significant cell-type enrichment, the non-mediating proteins were enriched in VLMCs, perivascular fibroblasts, and arterial smooth muscle cells. These results were largely reproduced when looking at change in global cognition instead of executive function (Extended Data Fig. 6).

medRxiv preprint doi: https://doi.org/10.1101/2024.06.10.24308599; this version posted June 10, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

Figure 5. Characterization of WML-associated proteins mediating the relationship between WML and cognitive decline as measured by executive function



Legend: **(a)** Path diagram of the mediation analysis models. **(b)** Volcano plots showing the extent to which WML-associated proteins mediate the association between WML and executive function rate of change by the standardized beta values derived from the association between proteins and the executive function rate of change. The top 10 upregulated and top 10 downregulated proteins with highest mediation effect are labeled. **(c)** Comparison for cell-type enrichment between WML-associated proteins that mediate and those that do not mediate the interaction between WML and executive function rate of change based on single cell transcriptomics data from the Human MTG 10x SEA-AD and the Human Brain Vascular Atlas using the EWCE package. For this analysis, the 1331 Olink proteins were used as background.
*corresponds to $p_{FDR}<0.05$, **corresponds to $p_{FDR}<0.01$, ***corresponds to $p_{FDR}<0.001$.
Abbreviations: OPC = oligodendrocyte precursor cells, PVM = perivascular macrophages, VLMC = vascular leptomeningeal cell, P-Fibroblasts=perivascular fibroblasts, M-Fibroblasts = meningeal fibroblasts, SMC = smooth muscle cell, aSMC = arterial smooth muscle cell, aaSMC=arteriolar smooth muscle cell

Validation of DAPs in independent cohorts

To ensure robustness and generalizability of our findings regarding WML-associated DAPs, we sought validation in two additional independent cohorts with CSF proteomics data (using WML volume as a continuous variable, Extended Data Table 2, Extended Data Fig. 7). The first included 383 participants from the independent Swedish BioFINDER-1 study, all of whom had both CSF proteomics and MRI data. Our analysis focused on a subset of CSF proteins (N=202) measured with Olink-based technology that were overlapping with those examined in our BioFINDER-2 study. The second validation cohort was derived from the Alzheimer's Disease Neuroimaging Initiative (ADNI), which included 729 participants, and used the aptamer-based SOMAscan 7K (v4.1) platform. Linear regression models revealed a high degree of concordance in the differential protein abundance between the BioFINDER-1 and

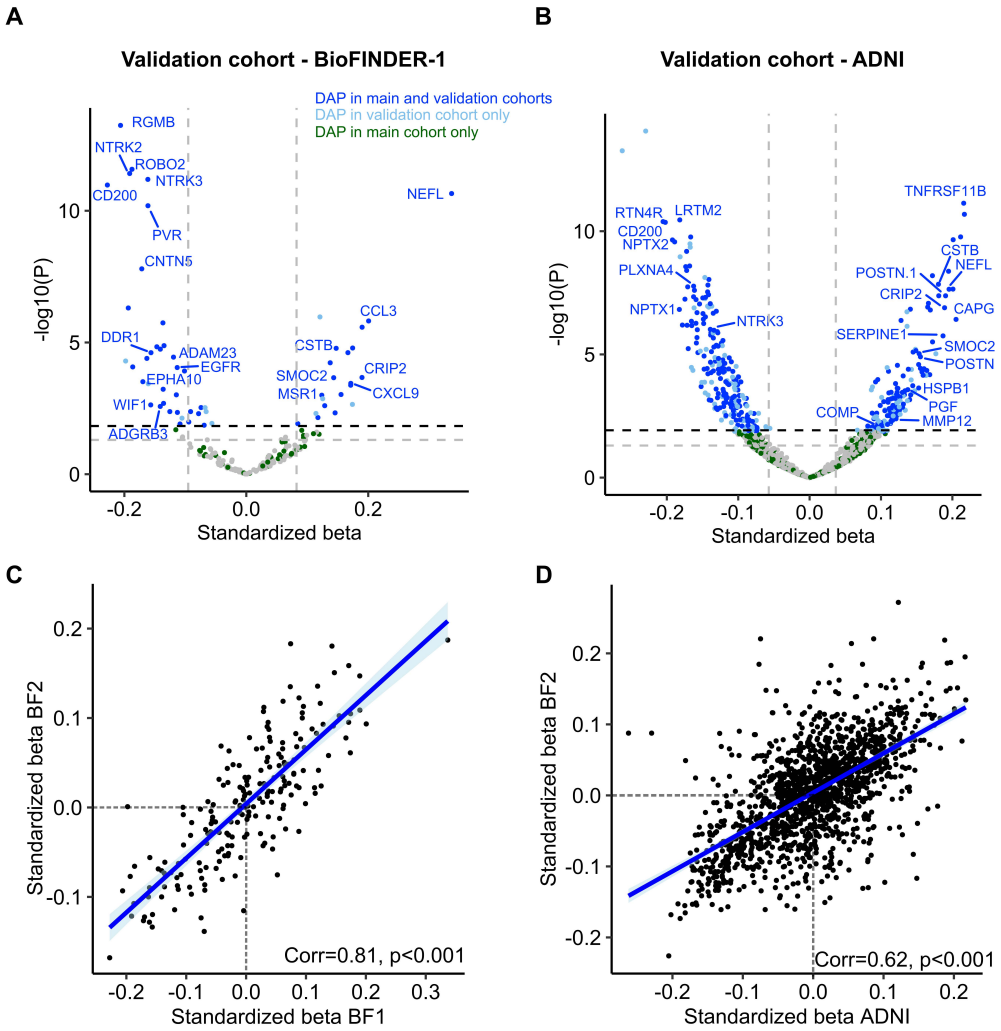
medRxiv preprint doi: https://doi.org/10.1101/2024.06.10.24308599; this version posted June 10, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

BioFINDER-2 cohorts, with a 78% agreement (Fig. 6a, Supplementary Table 6). Similarly, a 71% concordance was observed between the ADNI cohort and BioFINDER-2 (Fig. 6b, Supplementary Table 7). The correlation of standardized beta values revealed robust associations, with a strong correlation coefficient of 0.81 between BioFINDER-1 and BioFINDER-2 (Fig. 6c). Additionally, a correlation of 0.62 between the ADNI cohort and BioFINDER-2 (Fig. 6d) indicated a high level of consistency in DAP patterns even in this cross-platform context.

We also investigated the association of WML with DAPs in *plasma*, using the extensive UK Biobank dataset where plasma samples (n=5712) had been analyzed using Olink-based technology (Extended Data Table 2, Extended Data Fig. 8). This cohort exhibited a younger demographic profile and a longer gap between proteomic measurement and neuroimaging. While the results showed less concordance compared to CSF cohorts (57% agreement, correlation = 0.18), we still validated several key proteins. Specifically, NEFL, CSTB, and PGF were found to be significantly upregulated in plasma of individuals showing WML pathology, while LRTM2 and CD200 were significantly downregulated (Supplementary Table 8). Notably, some proteins, such as BST2, showed contrasting trends, being significantly increased in UK Biobank plasma but significantly decreased in CSF from the BioFINDER-2 cohort. Fewer proteins, such as FAP, showed the opposite trend.

medRxiv preprint doi: https://doi.org/10.1101/2024.06.10.24308599; this version posted June 10, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

Figure 6. Validation of differentially abundant proteins in two independent cohorts



Legend: **a,b,** Volcano plots showing differentially abundant proteins (DAPs) from the **(a)** BioFINDER-1 validation cohort or **(b)** Alzheimer's Disease Neuroimaging Initiative (ADNI) validation cohort, analyzed with WML as a continuous variable, adjusting for age, sex, intracranial volume, and average protein level. Color-coding reflects significance across cohorts: proteins significant in both the main and validation cohorts are shown in blue, those significant only in the validation cohort are depicted in pale blue, and proteins significant solely in the main cohort are shown in green. Labels highlight the top 20 differentially abundant proteins in BioFINDER-2. **c,d,** Scatter plots showing the correlations between standardized beta values from linear regression analysis for proteins overlapping between **(c)** BioFINDER-1 or **(d)** ADNI and BioFINDER-2.

medRxiv preprint doi: https://doi.org/10.1101/2024.06.10.24308599; this version posted June 10, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

DISCUSSION

This study represents a thorough investigation of CSF proteomic alterations concomitant to various cSVD manifestations, including WML, microbleeds and infarcts in a large cohort. Earlier work focused primarily on isolated cSVD manifestations, often centered on WML, with limited protein panels assessed mostly in blood and fewer studies in CSF. These studies identified significant associations of cSVD pathology with endothelial dysfunction (ICAM-1[18], E-selectin[19], PGF[20,21]), inflammation (IL-6[22]), clotting pathways (vWF, CSTB, LTBR, TNFRSF11B)[23] and alteration of neuronal structural integrity (NEFL[24]). In addition to validating these changes in CSF, our proteomic study introduces novel candidate biomarkers for the different cSVD manifestations and provides an enriched expression profile that deepens the understanding of the pathophysiological framework of cSVD. We have identified a spectrum of dysregulated proteins, uncovering both unique and shared patterns across different cSVD conditions, with MMP12 emerging as an early and robust indicator of diverse cSVD manifestations. Furthermore, our findings highlight a common mechanism in cSVD pathophysiology involving extracellular matrix dysregulation and a critical role for arterial smooth muscle cells. Through comprehensive protein co-abundance network analysis, we also shed light on the dynamics between disrupted metabolic, vascular, immune, and neuronal mechanisms in cSVD, especially related to white matter lesions. In summary, we have identified proteins and biological processes present in cSVD that hold promise as novel therapeutic targets and biomarkers for these diseases.

Several proteins showed elevated levels across all three cSVD manifestations. In addition to the extracellular matrix-related proteins such as MMP12 and ELN (further detailed below), matricellular proteins crucial in modulating cell-matrix interaction and signaling, like POSTN and CCN2, were also altered in the presence of multiple cSVD pathologies. CCN2 is thought to act as a 'matricellular organizer', and is involved in extracellular matrix (ECM) remodeling, angiogenesis, wound repair and cell adhesion[25]. In mice models, CCN2 was found to be an important mediator of arterial remodeling in vascular disease[26]. On the other hand, POSTN was found to be increased in vascular injury, with its deletion protecting against atherosclerosis by altering ECM remodeling and inflammation[27,28]. The upregulation of PDLIM7, CCDC80, and NECTIN4 alongside these proteins suggest shared mechanisms in cSVD pathology related to extracellular matrix remodeling, vascular repair, and altered cell-matrix communication. Taken together, these observations suggest a common cluster of upregulated proteins across all cSVD manifestations, pointing to a critical involvement in vascular integrity and remodeling processes to cSVD pathophysiology.

Among these shared proteins, MMP12 seems to be a key early player in the onset of cSVD, marked by an early and substantial increase in abundance across cSVD manifestations, and associated with WML progression. While previous research has associated MMP2, MMP3 and MMP9 with vascular cognitive impairment[29,30], our analysis did not reveal a significant association with MMP3, and technical constraints prevented the evaluation of MMP2 (not present in the Olink Explore 3072 panel) and MMP9 (above LOD in <70% of subjects). Interestingly, we found a significant association of other MMPs with WML, such as MMP1, MMP7, MMP8 and MMP10. MMPs participate in many processes relevant to cSVD, such as proteolyzation of cerebrovascular basement membranes and alteration of tight junctions, as well as activation of bioactive molecules, including clotting factors and cell-cell signaling proteins[31]. MMP12 plays a key role in breaking down elastin and other extracellular matrix

medRxiv preprint doi: https://doi.org/10.1101/2024.06.10.24308599; this version posted June 10, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

components[32], which in pathological conditions may lead to compromised vessel wall integrity and elasticity. This could result in arterial stiffening – a condition linked to induction of SMC MMP12 expression in animal models and human aortic samples[33]. This aligns with our observations of increased levels of elastin across the three studied cSVD pathologies, suggesting the breakdown of elastic fibers in the vessel walls, and elevated levels of various collagen types, especially with white matter lesions. Arterial stiffness can increase endothelial permeability, promote macrophage adhesion, and can lead to smooth muscle proliferation and vessel remodeling[34]. Additionally, MMP12 plays a role in modulating neuroinflammation through its regulatory effects on macrophages and neutrophils[35,36]. These findings suggest that the upregulation of MMP12 may be a direct and early reaction to vascular injury, further initiating a cascade of responses.

MMP12 protein in the CSF was co-abundant with a network of proteins that was generally elevated in the presence of all three manifestations of cSVD (Module 5). This module was comprised of proteins enriched in aSMC and endothelial cells and involved in the metabolism of heterocyclic molecules and nitrogenous catabolic compounds. An early increase in the protein levels observed during the initial stages of WML and microbleed pathology underscores the critical role of endothelial and smooth muscle cells' dysfunction at the onset of cSVD pathogenesis. Indeed, endothelial dysfunction is often considered as the initial event in the development of cSVD[2], but its mechanisms are not fully understood. A leading hypothesis is that vascular risk factors, like aging and hypertension, are associated with vascular oxidative stress, reducing the bioavailability of nitric oxide (NO)[37,38]. NO promotes vasodilation and blood vessel integrity, both of which are essential for proper cerebral blood flow and pressure regulation[39]. Disruption in NO production can lead to vascular resistance, potentially initiating aSMC-related changes. Our findings revealed that many proteins associated with cSVD were enriched in aSMC and may participate in vascular remodeling. Among other processes, these may include transition to a synthetic phenotype with decreased contractility and increased proliferation, as suggested in an animal model of hypertension[40], and/or progressive degeneration and loss of smooth muscle cells[41]. A more detailed investigation of the biological pathways specific to aSMC-associated proteins could provide deeper insights. Nevertheless, the presence of cSVD-associated proteins such as HSPB1, HSPB6, MMP12, PDLIM7, SERPINE1 and CSTB within this module suggests mechanisms through which oxidative stress, endothelial dysfunction and vascular remodeling may contribute to the pathogenesis and progression of cSVD.

Another module (Module 10) was significantly associated with all cSVD, yet it displayed no notable cell-type or functional enrichment, possibly because it comprised only a few proteins. Nonetheless, this module demonstrated an early increase in protein levels in WML and microbleeds. This module included cSVD-associated proteins such as SMOC2 and VWF. SMOC2 has been involved in ECM degradation, proliferation and migration of vascular SMC, as well as microgliosis[42,43]. On the other hand, VWF acts as a marker of endothelial dysfunction and is emerging as a mediator of vascular inflammation[44,45]. This may suggest a synergistic impact of these two modules, with metabolic and vascular dysfunction contributing to vascular damage and further facilitating the entry of potentially harmful toxins and immune cells in the brain[46].

medRxiv preprint doi: https://doi.org/10.1101/2024.06.10.24308599; this version posted June 10, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

Even though there are commonalities between the different cSVD manifestations, our study further reveals unique protein landscapes of microbleeds and infarcts, which exhibit less overlap (compared to the overlap with WML), indicating at least a few distinct disease mechanisms. Microbleeds seem to be characterized with a more pronounced immune response, as evidenced by proteins such as ITGAM and KYNU[47,48]. Differential abundance of angiogenic factors also differs between these conditions: ANGPT1 and PGF in infarcts may promote the maturation of new blood vessels for ischemic tissue recovery[49,50], whereas ANGPT2 increase in microbleeds may lead to vascular remodeling and destabilization by antagonizing ANGPT1 activity[51]. Further investigation into the localization of microbleeds reveals a compelling separation of molecular pathways. Deep microbleeds, commonly associated with hypertensive arteriosclerosis[52], are associated with a set of proteins that overlap with those implicated in broader cSVD, including WML and infarcts. This shared proteomic profile, including molecules such as MMP12, TNFRSF11B and SERPINE1, underscores a common pathway related to arterial stiffness and atherosclerosis pathogenesis[53–56]. In contrast, lobar microbleeds are linked with cerebral amyloid angiopathy (CAA) – the accumulation of β-amyloid in the vessel walls of the cortex and leptomeninges. Proteins predominantly linked to lobar microbleeds, such as ITGB2, TMSB10, and GLOD4, are not linked to WML and infarcts, but align with amyloid pathology, as corroborated by recent literature[57]. Nevertheless, certain proteins associated both with lobar microbleeds and WML, including CHIT1, CCN1 and MMP10, suggest an intersection of pathological processes implicated in vascular injury and tissue repair. These distinctions enhance our understanding of the etiologies presumed to underlie different types of microbleeds – hypertension in the case of deep microbleeds, and CAA, associated with AD, for lobar microbleeds. This evidence supports a model where deep microbleeds share molecular pathways with other SVD manifestations, while lobar microbleeds parallel more closely amyloid pathways.

Compared to infarcts and microbleeds, our findings suggest that overall, the CSF proteome is much more perturbed in the presence of WML. These perturbations are likely due to factors including disruptions in CSF flow or breaches in the BBB integrity caused by vascular disease, and the pronounced CSF response to WML may reflect a highly conserved biological reaction to vascular pathology. Moreover, the detection of changes in the CSF could be facilitated by the specific location or characteristics of WML-related responses, such as effects predominantly occurring in the extracellular matrix or involving cells that are more peripherally located and, therefore, more readily affected by or contributing to changes in the CSF composition.

The link between inflammation and vascular damage is well established, but the direction of causality remains debated, with existing research offering conflicting insights[58]. Our study contributes to this discussion by detecting a pattern associated with WML proliferation, in addition to initial metabolic and vascular dysfunction. In individuals with a faster rate of WML progression, we noted an increase in immune-related proteins, including chemokines and cytokines such as CXCL and CCL, as well as TNFSF/TNFRSF members, indicating a complex immune response. The response driven by perivascular fibroblasts and VLMC seems to involve chemokine-mediated signaling and immune cells chemotaxis, including lymphocytes, monocytes, and neutrophils. This process potentially guides immune cells to areas of vascular damage, intensifying the local inflammatory response. Even though the role of CNS fibroblasts is largely unknown, in the periphery, they can recruit immune cells to sites of injury, influence

medRxiv preprint doi: https://doi.org/10.1101/2024.06.10.24308599; this version posted June 10, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

leukocyte migration and promote immune cell survival[59]. Microglial activation, a well-established driver of CNS inflammation, may be driven by blood proteins, such as fibrin, within the neurovascular unit[60]. This may lead to a persistent inflammatory state, attracting peripheral inflammatory cells, and facilitating the interaction of activated lymphocytes with CNS antigens.

Building on this foundation, our study reveals a crucial involvement of immune cells in WML progression. The strong association of WML progression with microglia and perivascular macrophage-associated proteins, such as ITGAM, ITGB2, CHI3L1 and IL15, suggests a strong pro-inflammatory response, which might accelerate cell death, and the development and expansion of white matter lesions. Indeed, increased levels of serum CHI3L1 have recently been associated with cSVD and with the destruction of white matter macroscopic and microstructure, further associated with cognitive deficits[61]. In multiple sclerosis, CHI3L1 was associated with neuronal deterioration among other processes, affirming its contribution to the progression of the disease[62]. In a rodent model, reducing microglia activation by minocycline administration, known for its anti-inflammatory properties, was associated with a considerable reduction in white matter damage[63]. Currently, minocycline effectiveness on microglial activation and BBB permeability in cSVD patients is being evaluated in a phase-2 clinical trial[64]. Future research should aim to define the underlying microglial states and the possible disease stage-specific contributions of microglia to cSVD.

In the context of WML, distinct from microbleeds and infarcts, we observe a widespread *downregulation* of proteins related to neurons and OPCs. This phenomenon appears to have different impacts on cognitive outcomes. Specifically, proteins abundant in OPCs are not linked to cognitive decline and may be involved in the myelination process and serve broader roles including proliferation, migration and differentiation of OPCs (PDGFRA)[65–67] and extracellular matrix organization (TIMP4)[68]. While these proteins are vital for maintaining efficient neuronal signaling, their direct impact on cognitive functions may not be immediate, and the brain may partially compensate for reductions in myelination efficiency. An important factor contributing to OPC dysfunction may include endothelial cell dysfunction, leading to a disruption of the release of essential factors for OPCs, compromising their survival and proliferation and increasing their susceptibility to damage[69]. Additionally, angiogenesis in cSVD may hinder OPC migration along blood vessels, reducing the repair mechanisms[70].

Conversely, downregulated proteins enriched in neurons appear to directly mediate cognitive decline, likely due to their integral roles in synaptic function and neuronal plasticity, with disruptions tending to precipitate cognitive impairment. Key mediating proteins included CBLN4, involved in synaptic organization and plasticity[71], and NPTX1 and NPTX2, markers of synaptic activity and potential CSF biomarkers for synaptic dysfunction in neurodegeneration[72]. Additionally, Eph receptor family members, EphA10 and EphB6, both pseudokinases, stand out as significant, yet underexplored mediators of cognitive decline. While the Eph family is recognized for its involvement in synaptogenesis and synaptic plasticity, contributing to neurological diseases characterized by memory impairments[73], EphA10 and EphB6 are not well characterized. Recent studies suggested that EphB6 is present in both excitatory and inhibitory synapses[74], indicating its potential relevance in synaptic communication and increasing the need for more studies.

medRxiv preprint doi: https://doi.org/10.1101/2024.06.10.24308599; this version posted June 10, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

Several proteins that have been identified as *upregulated* in WML and mediating WML-associated cognitive decline also exhibit heightened levels in an AD context[75,76]. Among these, cystatin B (CSTB) and NEFL stand out for their notable mediation effects. CSTB, a lysosomal cysteine protease, plays a crucial role by inhibiting cathepsins, whose dysfunction was found to be critical to the pathogenesis of neurodegenerative disorders[77]. NEFL, a significant component of the neuronal cytoskeleton, reflects neuronal damage and is associated with neurodegenerative diseases[78]. Moreover, proteins such as SMOC2, MMP10, TMSB10, which are associated with cognitive decline and play roles in modulating the extracellular matrix and cell adhesion[42,79,80], highlight the interplay between structural neural integrity and cognitive function.

In the context of microbleed pathology, CHIT1 has emerged as a protein of interest due to its significant increase already at low levels of microbleed pathology and continuing progressive increase, even at higher pathology loads. CHIT1 has been recognized as a promising novel marker of acute or chronic inflammation in several neurological disorders, and its expression has been mostly ascribed to microglia and infiltrating macrophages[81]. CHIT1 levels have been found to be increased in CSF in both AD[82,83] and cerebrovascular dementia[84], but the association with microbleeds remained after correcting for AD pathology. Interestingly, in multiple sclerosis, CSF CHIT1 levels at baseline reflect microglial activation early in the disease and correlated with subsequent disease progression[85]. Since microbleed pathology seems to be associated with a heightened inflammatory state, the significant early increase in CHIT1 suggests that CHIT1 might act as an early marker in the inflammatory response in microbleeds. This underscores the need for further exploration into the role of CHIT1 in immune responses and cSVD, and further characterizing its specificity.

The cohort's recruitment from southern Sweden, primarily composed of individuals of European descent narrows ethnic and racial representation. This points to the need for inclusion of more diverse populations in future studies, especially considering recent evidence of ancestry-specific protein co-abundance modules[86]. The validation of many DAPs in two separate datasets, The Swedish BioFINDER-1 study and the American ADNI study—with ADNI employing the SOMAscan proteomics platform—demonstrates the robustness of our findings despite varied methodologies. The concordance between the SOMAscan and Olink platforms supports the credibility of these protein signatures as biomarkers for cSVD, suggesting their applicability across diverse patient groups and detection technologies. To gain further insight into the relationship between CSF markers and cSVD progression, it is imperative to pursue longitudinal studies that track protein level changes over time. This study primarily focused on CSF biomarkers, and an investigation into plasma protein variations among participants with corresponding CSF data could offer critical insights into the utility of these markers across different biological fluids. Nevertheless, a subset of these proteins was validated in plasma samples from the UK Biobank cohort. Despite this cohort's comparatively younger demographic, such validation underscores the potential for these biomarkers to be effective across both plasma and CSF contexts.

In conclusion, our study has leveraged extensive proteomic profiling to deepen our understanding of CSF protein dysregulation across various manifestations of cSVD. We have demonstrated that specific imaging markers are associated with diverse molecular signatures, distinguishing shared and distinct pathophysiological profiles. These findings compel a reevaluation of the traditional view of 'cerebral small vessel disease' as an umbrella term

medRxiv preprint doi: https://doi.org/10.1101/2024.06.10.24308599; this version posted June 10, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

towards a more nuanced classification into distinct 'cerebral small vessel diseases', each characterized by unique biomarker profiles. This refined insight not only paves the way for a tailored approach to disease management but also highlights the critical need for further research, aiming to uncover further genetic underpinnings, etiological and pathophysiological mechanisms to facilitate the development of targeted therapies and substantially improve clinical outcomes.

METHODS

Participants – BioFINDER-2 cohort

We included participants from the ongoing, prospective Swedish BioFINDER-2 cohort (NCT03174938). All participants were recruited from the south of Sweden at Skåne University Hospital or the Hospital of Ängelholm, Sweden, and the data were acquired between May 2017 and December 2023. The main inclusion and exclusion criteria has been described in detail elsewhere[87]. All participants included in this study had complete CSF proteomic data and MRI measurements available at baseline, with the final dataset including 856 subjects. The patients ranged from adults with intact cognition or subjective cognitive decline (both included in the CU group), mild cognitive impairment (MCI), and dementia. Causes of cognitive impairment included AD and non-AD neurodegenerative diseases. The group of non-AD neurodegenerative diseases included Parkinson's disease (PD) and vascular dementia (VaD) patients. The underlying etiology was established based on primary criteria for each condition, either at baseline or during a follow-up visit. Written informed consent was obtained from all participants prior to entering the study. Ethical approval was obtained from the Regional Ethical Committee in Lund, Sweden.

MRI as measure of cerebrovascular pathology

MRI was performed using a 3 Tesla Siemens MAGNETOM Prisma scanner, equipped with a 64-channel receiver coil array (Siemens Healthcare, Erlangen, Germany). Imaging sequences included a whole-brain T1-weighted anatomical magnetization-prepared rapid gradient echo sequence (MPRAGE; TR = 1900 ms, TE = 2.54 ms, voxel size = 1 × 1 × 1 mm3), a T2-weighted fluid-attenuated inversion recovery sequence (FLAIR; TR = 5000 ms, TE = 393 ms, matching resolution with the T1-weighted image), and a 3D, multi-gradient-echo pulse sequence (TR = 24 ms; TEs = 5.00, 8.80, 12.60, 16.40, and 20.20 ms with monopolar/fly-back readout gradients, flip angle = 15°; voxel size = 0.8 × 0.8 × 1.4 mm3).

Following established criteria, microbleeds and infarcts were identified and counted via visual inspection by an experienced neuroradiologist (D.W.), from the T2* and the FLAIR images, respectively[5,88]. We categorized microbleeds and infarcts as either 0 (absent) and ≥1 (present). When necessary, microbleed counts were treated as a continuous variable and log-transformed for analysis. Additionally, microbleeds were categorized by their localization: lobar, deep and infratentorial. White matter lesions (WML) were segmented using the longitudinal Sequence Adaptive Multimodal SEGmentation (SAMSEG) tool from FreeSurfer (version 7.1) using the T1 and FLAIR images as input and quantified[89,90]. When treating WML volume as a continuous variable, considering the values are in cubic millimeters (mm^3), we applied min-max scaling to rescale the range to fall between 0 and 1. In the main analysis, WML was treated as a binary variable to maintain uniformity in comparing different cSVD. This binary classification was

medRxiv preprint doi: https://doi.org/10.1101/2024.06.10.24308599; this version posted June 10, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

achieved by first adjusting WML volumes with the intracranial volume and separating these adjusted values into tertiles. Participants in the lower two tertiles were classified as not exhibiting WML pathology, whereas individuals in the top tertile were identified as having WML pathology (cut-off for WML (% of icv) = 0.508). Baseline data on microbleeds were missing for 38 participants. MRI follow-ups were conducted 1, 2, and 4 years after the initial baseline measurement, for an average of 2.67±0.93 years of follow-up.

CSF proteomic measures

CSF samples were collected during baseline visits and were subjected to analysis using the Olink proteomic assay, a highly sensitive and specific antibody-based Proximity Extension Assay (PEA™ technology)[91]. This method involves the use of multiplexed oligonucleotide labeled antibody probe pairs that bind to their respective target protein. Protein expression levels were reported as normalized protein expression (NPX) values on a log2 scale, providing a relative quantification unit. We used the Olink Explore 3072 panel, allowing for quantification of 2943 proteins within eight multiplex protein panels, each containing 367-369 proteins (Oncology I and II, Neurology I and II, Cardiometabolic I and II, and Inflammation I and II). To ensure rigorous quality control (QC) in line with Olink technology standards, samples were randomly distributed across plates, and each plate included relevant controls. Samples failing to meet stringent QC criteria were labeled with a warning. A lower limit of detection (LOD) was defined as three standard deviations above the control-based background for each assay. Our analysis was focused on 1331 proteins demonstrating strong expression, with levels exceeding the limit of detection (LOD) in ≥70% of the participants. A minor portion of the data (approximately 0.2% of the dataset, involving 46 to 64 measurements across 37 proteins), received assay warning and was excluded from the analysis. After exclusion, we retained all data points below the LOD, following Olink recommendations[15,92]. Building on our prior study that identified the existence of inter-individual variability in average CSF protein levels[93], we included a measure of average overall protein level as a covariate in our analysis. This was computed as the mean z scored NPX value for proteins with high detection rates (n=1157, including proteins exceeding the LOD in over 90% of participants).

Cell-type and functional enrichment analysis

For cell-type enrichment, we used two sources with single-nucleus RNA-sequencing data: 1) single-nucleus transcriptomes from 166,868 nuclei in the middle temporal gyrus (MTG) obtained from the SEA-AD (Human MTG 10x SEA-AD 2022, https://portal.brain-map.org/atlases-and-data/rnaseq/human-mtg-10x_sea-ad), from five post-mortem human brain samples and 2) the Human Brain Vascular Atlas[94], which profiles major vascular and perivascular cell types of the human brain using 143,793 single-nucleus transcriptomes from the hippocampus and cortex from eight post-mortem samples. We acquired the Seurat objects for both atlases and used the function AverageExpression (R package Seurat version 4.3.0) to determine cell type expression levels. By leveraging the class and subclass annotations from the Human MTG 10x SEA-AD data, we computed the average expression from both non-neuronal cells (microglia, astrocytes, oligodendrocytes, oligodendrocyte precursor cells, vascular leptomeningeal cells, and endothelial cells) and neuronal cells (GABAergic neurons and glutamatergic neurons), after removing the annotations labeled as “None”. Following the same procedure, we calculated the average expression in major vascular and perivascular cell

medRxiv preprint doi: https://doi.org/10.1101/2024.06.10.24308599; this version posted June 10, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

types from the Human Brain Vascular Atlas, including endothelial cells, smooth muscle cells (distinguishing between arterial and arteriolar smooth muscle cells in a subsequent analysis), pericytes, astrocytes, macrophages, T cells, perivascular and meningeal fibroblasts. The percentage expression across the cell types was then calculated based on their average expression levels. For assessing cell-type enrichment statistically, we used the EWCE[95] (Expression Weighted Cell Type Enrichment) package in R (v1.6.0) to determine whether a specific set of proteins (DAPs or modules) showed significant enrichment within certain cell types beyond what would be expected by chance. This was done against a background of the 1331 Olink proteins, using 10,000 bootstrap iterations for analysis.

Functional enrichment analysis on relevant set of proteins (DAPs and modules) was performed using WebGestalt[96], with the 1331 Olink proteins as background set. We performed Over-Representation Analysis (ORA) with ontology sources from Gene Ontology (GO) Biological Processes (BP), Cellular Component (CC) and Molecular Function (MF). We report terms with pFDR<0.05 (Benjamini–Hochberg method), after having applied affinity propagation for redundance reduction of the enrichment categories. All pFDR<0.05 terms are reported in the Supplementary table.

Co-abundance modules analysis

To identify modules of co-expressed proteins, we used the SpeakEasy2 consensus clustering algorithm, which detects robust protein communities by applying both top-down and bottom-up approaches to clustering simultaneously[17]. To achieve this, we initially normalized protein levels by regressing out age, sex, and average overall protein levels. Using these adjusted levels, we constructed a protein-by-protein Spearman correlation matrix of the 1331 proteins, serving as the basis for the clustering process. We obtained 13 modules, but modules 11-13 were excluded due to them containing only a low number of outlying proteins. For each participant, we calculated the average level of proteins within each module, which was then used in our analysis.

Validation of key DAPs in independent datasets

Three separate cohorts were used to validate key proteins identified in the BioFINDER-2 cohort: BioFINDER-1 and Alzheimer's disease Neuroimaging Initiative (ADNI), both of which used CSF for analysis; and the UK Biobank cohort (UKBB), which used plasma proteomics data. The BioFINDER-1 study (registered under NCT01208675) is comprised of cognitively unimpaired (CU) individuals and individuals with mild cognitive impairment (MCI). The methodologies for participant selection and the criteria for inclusion or exclusion have been described previously[97]. In BioFINDER-1, the proteomic analysis was conducted on a limited set of proteins from different panels, including Neuro-exploratory, Neurology, Inflammation, and Cardiovascular-III, resulting in an overlap of 202 proteins with the main BioFINDER-2 cohort. These overlapping proteins were used for further analysis. From this cohort, we selected participants who had both proteomic data and MRI data, specifically those for whom data on white matter lesion volumes was available, segmented using SAMSEG. This subset included 383 subjects.

medRxiv preprint doi: https://doi.org/10.1101/2024.06.10.24308599; this version posted June 10, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

To validate key proteins externally, we used the dataset from Alzheimer's disease Neuroimaging Initiative (ADNI), an ongoing, multisite observation study aimed at understanding the progression of Alzheimer's disease (https://adni.loni.usc.edu/). We selected participants from ADNI1, ADNI-GO and ADNI2 who had proteomic and MRI data available, including information on white matter lesions, resulting in a total of 729 individuals. Using SomaScan 7K (v4.1), an aptamer-based proteomics technology, protein levels of ~7000 proteins were measured, and analyses were performed on 1438 overlapping aptamers with BioFINDER-2. White matter lesion volume was taken from the "ST128SV" variable, which was quantified using FreeSurfer from T1-weighted images. The total intracranial volume was taken from the variable "ST10CV".

Finally, we aimed to verify whether key proteins can be validated in plasma by using data from the UK Biobank dataset[98]. This large-scale, multicenter prospective cohort study includes approximately 500,000 participants aged between 40 and 69, who were enrolled from 2006 to 2010. More information about the study's methodology and data collection can be found in the UKBB online protocol (www.ukbiobank.ac.uk). In this cohort, a total of 2,922 plasma proteins were analyzed using the Olink Explore 3072 assay. Applying the same standards for protein retention as used in BioFINDER-2, 1020 proteins were overlapping between the two cohorts and used for analysis. We selected participants who had both Olink proteomics data and white matter lesions information from MRI scans, amounting to a total of 5712 individuals. White matter lesions volumes were obtained from the variable "Total volume of white matter hyperintensities from T1 and T2 FLAIR images", with lesion segmentation performed using the BIANCA tool[99]. The total intracranial volume was taken from the variable "Volume of EstimatedTotalIntraCranial (whole brain)".

Statistical analysis

All analyses were performed in R version 4.2.1 or Python version 3.9.2. Linear models were used to assess DAPs for each cSVD manifestation independently (WML+ vs WML-, Microbleeds+ vs Microbleeds-, Infarcts+ vs Infarcts-), including covariates such as age, sex, and average overall protein levels. Further analysis distinguished between lobar microbleeds (Lobar microbleeds+ vs Lobar microbleeds-) and deep microbleeds (Deep microbleeds+ vs Deep microbleeds-). Subsequently, significant DAPs for each pathology were analyzed in a linear model including all cSVD as independent variables, to determine if the association persisted after adjusting for other cSVD. Only proteins that retained significance for each disease independently were considered in subsequent analyses, including cell-type and functional enrichment. In certain analyses, models were adjusted for presence of clinical AD and PD diagnosis. In the validation cohorts, given the variability in pathology severity, WML volumes were analyzed as continuous variables. Within the CSF validation cohorts, BioFINDER-1 and ADNI, we included age, gender, intracranial volume and mean overall protein levels as covariates. However, for the UK Biobank cohort, the average overall protein level was not included as a covariate due to the absence of reported similar observations in plasma. Additionally, the time interval between MRI scans and CSF collection was included in the model for the UK Biobank cohort due to the notable time gap.

Linear-mixed effect models including random slope and intercept were used to test the effects of CSF baseline protein levels on WML volume changes and microbleed number variation over

medRxiv preprint doi: https://doi.org/10.1101/2024.06.10.24308599; this version posted June 10, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

time, using protein*time interaction in the model. Age, gender, mean overall protein levels and intracranial volume (for WML only) were used as covariates. To investigate the relationship between protein modules and cSVD, we used linear models to analyze the association between average protein levels of each module and cSVD at baseline. For longitudinal assessments, linear-mixed effect models with random slope and intercept, along with module*time interaction were used to examine the effect of average protein levels in each module on the changes of WML volume over time. To compare the trajectories of key proteins and modules with increasing WML and microbleed pathology, we plotted average protein levels and average protein modules against a continuous measure of WML and microbleeds using Generalized Additive Model (GAM). P values for all analyses were adjusted for multiple comparisons using the false discovery rate (FDR) according to the Benjamini-Hochberg procedure, pFDR<0.05.

Mediation analysis

For the mediation analyses, we investigated the mediation of WML-associated proteins on the association between WML volume (corrected for icv) and executive function or mPACC rate of change. Mediation analyses were conducted using the mediation R package v4.5.0. The significance of the mediating effect was assessed using 1000 bootstrapping iterations and all paths of the mediation model were controlled for age, sex, mean overall protein levels and education. The mPACC composite score was calculated using a mean of z-scores of the Alzheimer's Disease Assessment Scale (ADAS) delayed recall (weighted double), animal fluency, Mini-Mental State Examination (MMSE), and Symbol Digit Modalities Test (SDMT). The executive function composite score was calculated as the mean of z-scores from the Trail Making Test-A (TMT-A), Trail Making Test-B (TMT-B) and the SDMT, following methodology described previously[100,101]. To calculate the rate of change, linear mixed effect models with random slope and intercept were fitted, with executive function or mPACC as the dependent variable and time in years from the baseline scan as the independent variable. To assess the association between cognitive function (rate of change in executive function or mPACC) and WML-associated proteins, we used linear models adjusted for age, sex, mean overall protein levels and education.

Data availability

Pseudonymized data will be shared by request from a qualified academic investigator for the sole purpose of replicating procedures and results presented in the article and as long as data transfer is in agreement with EU legislation on the general data protection regulation and decisions by the Swedish Ethical Review Authority and Region Skåne, which should be regulated in a material transfer agreement.

Code availability

All code was written using existing packages in R and python and can be accessed by contacting the authors.

medRxiv preprint doi: https://doi.org/10.1101/2024.06.10.24308599; this version posted June 10, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

ACKNOWLEDGEMENTS

We would like to acknowledge all the BioFINDER team members as well as participants in the study and their family members for their dedication. The BioFINDER study was supported by the National Institute of Aging (R01AG083740), European Research Council (ADG-101096455), Alzheimer's Association (ZEN24-1069572, SG-23-1061717), GHR Foundation, Swedish Research Council (2022-00775, 2021-02219, 2018-02052), ERA PerMed (ERAPERMED2021-184), Knut and Alice Wallenberg foundation (2022-0231), Strategic Research Area MultiPark (Multidisciplinary Research in Parkinson's disease) at Lund University, Swedish Alzheimer Foundation (AF-980907, AF-980832, AF-993465, AF-939981, AF-AF-994229), Swedish Brain Foundation (FO2021-0293, FO2023-0163), Parkinson foundation of Sweden (1412/22), Cure Alzheimer's fund, WASP and DDLS Joint call for research projects (WASP/DDLS22-066), EU Joint Programme Neurodegenerative Diseases (2019-03401), Rönström Family Foundation (FRS-0003), Konung Gustaf V:s och Drottning Victorias Frimurarestiftelse, Skåne University Hospital Foundation (2020-O000028), Regionalt Forskningsstöd (2022-1259) and Swedish federal government under the ALF agreement (2022-Projekt0080, 2022-Projekt0107). Author JWV is funded through the SciLifeLab & Wallenberg Data Driven Life Science Program (KAW 2020.0239), the Swedish Alzheimer Foundation (AF-994626), and the Crafoord Foundation (20230790). The funding sources had no role in the design and conduct of the study; in the collection, analysis, interpretation of the data; or in the preparation, review, or approval of the manuscript.

AUTHOR CONTRIBUTIONS

IH, JWV and OH designed the study. IH had full access to the raw data and carried out the final statistical analyses. IH, JWV and OH wrote the manuscript and IH had the final responsibility to submit for publication. SP, ES and OH were responsible for clinical assessment and data collection. CG performed the co-expression module analysis and provided transcriptomics experience. All authors contributed to the interpretation of the results and critically reviewed the manuscript.

COMPETING INTERESTS

OH has acquired research support (for the institution) from AVID Radiopharmaceuticals, Biogen, C2N Diagnostics, Eli Lilly, Eisai, Fujirebio, GE Healthcare, and Roche. In the past 2 years, he has received consultancy/speaker fees from AC Immune, Alzpath, BioArctic, Biogen, Bristol Meyer Squibb, Cerveau, Eisai, Eli Lilly, Fujirebio, Merck, Novartis, Novo Nordisk, Roche, Sanofi and Siemens. SP has acquired research support (for the institution) from ki elements/ADDF. In the past 2 years, he has received consultancy/speaker fees from BioArtic, Biogen, Lilly and Roche. The remaining others declare no competing interests.

medRxiv preprint doi: https://doi.org/10.1101/2024.06.10.24308599; this version posted June 10, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

REFERENCES

1. Gorelick, P. B. *et al.* Vascular contributions to cognitive impairment and dementia: a statement for healthcare professionals from the american heart association/american stroke association. *Stroke* **42**, 2672–2713 (2011).

2. Wardlaw, J. M., Smith, C. & Dichgans, M. Small vessel disease: mechanisms and clinical implications. *Lancet Neurol.* **18**, 684–696 (2019).

3. Wardlaw, J. M., Smith, C. & Dichgans, M. Mechanisms of sporadic cerebral small vessel disease: insights from neuroimaging. *Lancet Neurol.* **12**, 483–497 (2013).

4. Wardlaw, J. M. *et al.* ESO Guideline on covert cerebral small vessel disease. *Eur. Stroke J.* **6**, CXI–CLXII (2021).

5. Wardlaw, J. M. *et al.* Neuroimaging standards for research into small vessel disease and its contribution to ageing and neurodegeneration. *Lancet Neurol.* **12**, 822–838 (2013).

6. Hosoki, S. *et al.* Molecular biomarkers for vascular cognitive impairment and dementia. *Nat. Rev. Neurol.* **19**, 737–753 (2023).

7. Yaghi, S. *et al.* Lacunar stroke: mechanisms and therapeutic implications. *J. Neurol. Neurosurg. Psychiatry* **92**, 823–830 (2021).

8. Ungvari, Z., Tarantini, S., Kirkpatrick, A. C., Csiszar, A. & Prodan, C. I. Advances in Cardiovascular Geroscience: Cerebral microhemorrhages: mechanisms, consequences, and prevention. *Am. J. Physiol. - Heart Circ. Physiol.* **312**, H1128 (2017).

9. Liu, X., Sun, P., Yang, J. & Fan, Y. Biomarkers involved in the pathogenesis of cerebral small-vessel disease. *Front. Neurol.* **13**, 969185 (2022).

10. Wu, L.-Y. *et al.* Blood-based biomarkers of cerebral small vessel disease. *Ageing Res. Rev.* **95**, 102247 (2024).

11. Inoue, Y., Shue, F., Bu, G. & Kanekiyo, T. Pathophysiology and probable etiology of cerebral small vessel disease in vascular dementia and Alzheimer's disease. *Mol. Neurodegener.* **18**, 46 (2023).

12. Gold, L. *et al.* Aptamer-Based Multiplexed Proteomic Technology for Biomarker Discovery. *PLoS ONE* **5**, e15004 (2010).

13. Lundberg, M., Eriksson, A., Tran, B., Assarsson, E. & Fredriksson, S. Homogeneous antibody-based proximity extension assays provide sensitive and specific detection of low-abundant proteins in human blood. *Nucleic Acids Res.* **39**, e102–e102 (2011).

14. Assarsson, E. *et al.* Homogenous 96-Plex PEA Immunoassay Exhibiting High Sensitivity, Specificity, and Excellent Scalability. *PLoS ONE* **9**, e95192 (2014).

15. del Campo, M. *et al.* CSF proteome profiling across the Alzheimer's disease spectrum reflects the multifactorial nature of the disease and identifies specific biomarker panels. *Nat. Aging* **2**, 1040–1053 (2022).

medRxiv preprint doi: https://doi.org/10.1101/2024.06.10.24308599; this version posted June 10, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

16. Del Campo, M. *et al.* CSF proteome profiling reveals biomarkers to discriminate dementia with Lewy bodies from Alzheimer´s disease. *Nat. Commun.* **14**, 5635 (2023).

17. Gaiteri, C. *et al.* Robust, scalable, and informative clustering for diverse biological networks. *Genome Biol.* **24**, 228 (2023).

18. Shoamanesh, A. *et al.* Inflammatory biomarkers, cerebral microbleeds, and small vessel disease: Framingham Heart Study. *Neurology* **84**, 825–832 (2015).

19. Zuliani, G. *et al.* Markers of endothelial dysfunction in older subjects with late onset Alzheimer's disease or vascular dementia. *J. Neurol. Sci.* **272**, 164–170 (2008).

20. Hinman, J. D. *et al.* Placental growth factor as a sensitive biomarker for vascular cognitive impairment. *Alzheimers Dement.* **19**, 3519–3527 (2023).

21. Gertje, E. C. *et al.* Associations Between CSF Markers of Inflammation, White Matter Lesions, and Cognitive Decline in Individuals Without Dementia. *Neurology* **100**, e1812–e1824 (2023).

22. Gu, Y. *et al.* Circulating inflammatory biomarkers are related to cerebrovascular disease in older adults. *Neurol. Neuroimmunol. Neuroinflammation* **6**, e521 (2019).

23. Kuipers, S. *et al.* A cluster of blood-based protein biomarkers reflecting coagulation relates to the burden of cerebral small vessel disease. *J. Cereb. Blood Flow Metab.* **42**, 1282–1293 (2022).

24. Peters, N. *et al.* Serum Neurofilament Light Chain Is Associated with Incident Lacunes in Progressive Cerebral Small Vessel Disease. *J. Stroke* **22**, 369–376 (2020).

25. Ponticos, M. Connective tissue growth factor (CCN2) in blood vessels. *Vascul. Pharmacol.* **58**, 189–193 (2013).

26. Brewster, L. P. *et al.* Role of Cellular Communication Network Factor 2 and Male Sex on Flow-mediated Arterial Remodeling and Atherosclerotic Plaque Formation. *JVS-Vasc. Sci.* **3**, 410 (2022).

27. Schwanekamp, J. A., Lorts, A., Vagnozzi, R. J., Vanhoutte, D. & Molkentin, J. D. Deletion of Periostin Protects Against Atherosclerosis in Mice by Altering Inflammation and Extracellular Matrix Remodeling. *Arterioscler. Thromb. Vasc. Biol.* **36**, 60–68 (2016).

28. Shimazaki, M. *et al.* Periostin is essential for cardiac healingafter acute myocardial infarction. *J. Exp. Med.* **205**, 295–303 (2008).

29. Adair, J. C. *et al.* Measurement of gelatinase B (MMP-9) in the cerebrospinal fluid of patients with vascular dementia and Alzheimer disease. *Stroke* **35**, e159-162 (2004).

30. Candelario-Jalil, E. *et al.* Matrix metalloproteinases are associated with increased blood-brain barrier opening in vascular cognitive impairment. *Stroke* **42**, 1345–1350 (2011).

31. Brkic, M., Balusu, S., Libert, C. & Vandenbroucke, R. E. Friends or Foes: Matrix Metalloproteinases and Their Multifaceted Roles in Neurodegenerative Diseases. *Mediators Inflamm.* **2015**, e620581 (2015).

medRxiv preprint doi: https://doi.org/10.1101/2024.06.10.24308599; this version posted June 10, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

32. Chandler, S., Cossins, J., Lury, J. & Wells, G. Macrophage metalloelastase degrades matrix and myelin proteins and processes a tumour necrosis factor-alpha fusion protein. *Biochem. Biophys. Res. Commun.* **228**, 421–429 (1996).

33. Liu, S.-L. *et al.* Matrix metalloproteinase-12 is an essential mediator of acute and chronic arterial stiffening. *Sci. Rep.* **5**, 17189 (2015).

34. Hooglugt, A., Klatt, O. & Huveneers, S. Vascular stiffening and endothelial dysfunction in atherosclerosis. *Curr. Opin. Lipidol.* **33**, 353–363 (2022).

35. Guan, C. *et al.* MMP-12 regulates proliferation of mouse macrophages via the ERK/P38 MAPK pathways during inflammation. *Exp. Cell Res.* **378**, 182–190 (2019).

36. Chalise, U. *et al.* MMP-12 polarizes neutrophil signalome towards an apoptotic signature. *J. Proteomics* **264**, 104636 (2022).

37. De Silva, T. M. & Faraci, F. M. Contributions of Aging to Cerebral Small Vessel Disease. *Annu. Rev. Physiol.* **82**, 275–295 (2020).

38. Dharmashankar, K. & Widlansky, M. E. Vascular Endothelial Function and Hypertension: Insights and Directions. *Curr. Hypertens. Rep.* **12**, 448–455 (2010).

39. Faraci, F. M. & Brian, J. E. Nitric oxide and the cerebral circulation. *Stroke* **25**, 692–703 (1994).

40. Liu, N. *et al.* The dynamic change of phenotypic markers of smooth muscle cells in an animal model of cerebral small vessel disease. *Microvasc. Res.* **133**, 104061 (2021).

41. Frösen, J. & Joutel, A. Smooth muscle cells of intracranial vessels: from development to disease. *Cardiovasc. Res.* **114**, 501–512 (2018).

42. Wang, X. *et al.* SMOC2 promoted vascular smooth muscle cell proliferation, migration, and extracellular matrix degradation by activating BMP/TGF-β1 signaling pathway. *J. Clin. Biochem. Nutr.* **73**, 116–123 (2023).

43. Lloyd-Burton, S. M., York, E. M., Anwar, M. A., Vincent, A. J. & Roskams, A. J. SPARC regulates microgliosis and functional recovery following cortical ischemia. *J. Neurosci. Off. J. Soc. Neurosci.* **33**, 4468–4481 (2013).

44. Horvath, B. *et al.* Measurement of von Willebrand factor as the marker of endothelial dysfunction in vascular diseases. *Exp. Clin. Cardiol.* **9**, 31–34 (2004).

45. Gragnano, F. *et al.* The Role of von Willebrand Factor in Vascular Inflammation: From Pathogenesis to Targeted Therapy. *Mediators Inflamm.* **2017**, 5620314 (2017).

46. Sweeney, M. D., Sagare, A. P. & Zlokovic, B. V. Blood–brain barrier breakdown in Alzheimer disease and other neurodegenerative disorders. *Nat. Rev. Neurol.* **14**, 133–150 (2018).

47. Kishimoto, T. K. & Rothlein, R. Integrins, ICAMs, and Selectins: Role and Regulation of Adhesion Molecules in Neutrophil Recruitment to Inflammatory Sites. in *Advances in Pharmacology* (eds. August, J. T., Anders, M. W. & Murad, F.) vol. 25 117–169 (Academic Press, 1994).

medRxiv preprint doi: https://doi.org/10.1101/2024.06.10.24308599; this version posted June 10, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

48. Sorgdrager, F. J. H., Naudé, P. J. W., Kema, I. P., Nollen, E. A. & Deyn, P. P. D. Tryptophan Metabolism in Inflammaging: From Biomarker to Therapeutic Target. *Front. Immunol.* **10**, (2019).

49. Jeansson, M. *et al.* Angiopoietin-1 is essential in mouse vasculature during development and in response to injury. *J. Clin. Invest.* **121**, 2278–2289 (2011).

50. De Falco, S. The discovery of placenta growth factor and its biological activity. *Exp. Mol. Med.* **44**, 1–9 (2012).

51. Maisonpierre, P. C. *et al.* Angiopoietin-2, a natural antagonist for Tie2 that disrupts in vivo angiogenesis. *Science* **277**, 55–60 (1997).

52. Vernooij, M. W. *et al.* Prevalence and risk factors of cerebral microbleeds: the Rotterdam Scan Study. *Neurology* **70**, 1208–1214 (2008).

53. Hosbond, S. E. *et al.* Osteoprotegerin as a marker of atherosclerosis: a systematic update. *Scand. Cardiovasc. J. SCJ* **46**, 203–211 (2012).

54. Tousoulis, D. *et al.* Serum osteoprotegerin and osteopontin levels are associated with arterial stiffness and the presence and severity of coronary artery disease. *Int. J. Cardiol.* **167**, 1924–1928 (2013).

55. Schneiderman, J. *et al.* Increased type 1 plasminogen activator inhibitor gene expression in atherosclerotic human arteries. *Proc. Natl. Acad. Sci. U. S. A.* **89**, 6998–7002 (1992).

56. Tofler, G. H. *et al.* Plasminogen activator inhibitor and the risk of cardiovascular disease: The Framingham Heart Study. *Thromb. Res.* **140**, 30–35 (2016).

57. Pichet Binette A. Cerebrospinal fluid proteomics reveal new associations with Aβ plaque and tau tangle pathologies across the Alzheimer's disease continuum. *Nat Neurosci* In Press.

58. Varatharaj, A. & Galea, I. The blood-brain barrier in systemic inflammation. *Brain. Behav. Immun.* **60**, 1–12 (2017).

59. Van Linthout, S., Miteva, K. & Tschöpe, C. Crosstalk between fibroblasts and inflammatory cells. *Cardiovasc. Res.* **102**, 258–269 (2014).

60. Adams, R. A. *et al.* The fibrin-derived γ377-395 peptide inhibits microglia activation and suppresses relapsing paralysis in central nervous system autoimmune disease. *J. Exp. Med.* **204**, 571–582 (2007).

61. Zhang, W. *et al.* YKL-40 as a novel biomarker related to white matter damage and cognitive impairment in patients with cerebral small vessel disease. *Brain Res.* **1807**, 148318 (2023).

62. Ahmad, I., Wergeland, S., Oveland, E. & Bø, L. An Association of Chitinase-3 Like-Protein-1 With Neuronal Deterioration in Multiple Sclerosis. *ASN Neuro* **15**, 17590914231198980 (2023).

medRxiv preprint doi: https://doi.org/10.1101/2024.06.10.24308599; this version posted June 10, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

63. Jalal, F. Y., Yang, Y., Thompson, J. F., Roitbak, T. & Rosenberg, G. A. Hypoxia-induced neuroinflammatory white-matter injury reduced by minocycline in SHR/SP. *J. Cereb. Blood Flow Metab. Off. J. Int. Soc. Cereb. Blood Flow Metab.* **35**, 1145–1153 (2015).

64. Brown, R. B. *et al.* MINocyclinE to Reduce inflammation and blood brain barrier leakage in small Vessel diseAse (MINERVA) trial study protocol. *Eur. Stroke J.* **7**, 323–330 (2022).

65. Fruttiger, M. *et al.* Defective oligodendrocyte development and severe hypomyelination in PDGF-A knockout mice. *Dev. Camb. Engl.* **126**, 457–467 (1999).

66. van Heyningen, P., Calver, A. R. & Richardson, W. D. Control of progenitor cell number by mitogen supply and demand. *Curr. Biol. CB* **11**, 232–241 (2001).

67. Zhu, Q. *et al.* Genetic evidence that Nkx2.2 and Pdgfra are major determinants of the timing of oligodendrocyte differentiation in the developing CNS. *Dev. Camb. Engl.* **141**, 548–555 (2014).

68. Arpino, V., Brock, M. & Gill, S. E. The role of TIMPs in regulation of extracellular matrix proteolysis. *Matrix Biol.* **44–46**, 247–254 (2015).

69. Arai, K. & Lo, E. H. An Oligovascular Niche: Cerebral Endothelial Cells Promote the Survival and Proliferation of Oligodendrocyte Precursor Cells. *J. Neurosci.* **29**, 4351–4355 (2009).

70. Tsai, H.-H. *et al.* Oligodendrocyte precursors migrate along vasculature in the developing nervous system. *Science* **351**, 379–384 (2016).

71. Liakath-Ali, K., Polepalli, J. S., Lee, S.-J., Cloutier, J.-F. & Südhof, T. C. Transsynaptic cerebellin 4–neogenin 1 signaling mediates LTP in the mouse dentate gyrus. *Proc. Natl. Acad. Sci.* **119**, e2123421119 (2022).

72. Gómez de San José, N. *et al.* Neuronal pentraxins as biomarkers of synaptic activity: from physiological functions to pathological changes in neurodegeneration. *J. Neural Transm.* **129**, 207–230 (2022).

73. Dines, M. & Lamprecht, R. The Role of Ephs and Ephrins in Memory Formation. *Int. J. Neuropsychopharmacol.* **19**, pyv106 (2015).

74. Bartolini, G. *et al.* Neuregulin 3 Mediates Cortical Plate Invasion and Laminar Allocation of GABAergic Interneurons. *Cell Rep.* **18**, 1157–1170 (2017).

75. Kamalian, A. *et al.* Exploratory Assessment of Proteomic Network Changes in Cerebrospinal Fluid of Mild Cognitive Impairment Patients: A Pilot Study. *Biomolecules* **13**, 1094 (2023).

76. van der Ende, E. L. *et al.* CSF proteomics in autosomal dominant Alzheimer's disease highlights parallels with sporadic disease. *Brain* **146**, 4495–4507 (2023).

77. Drobny, A. *et al.* The role of lysosomal cathepsins in neurodegeneration: Mechanistic insights, diagnostic potential and therapeutic approaches. *Biochim. Biophys. Acta BBA - Mol. Cell Res.* **1869**, 119243 (2022).

medRxiv preprint doi: https://doi.org/10.1101/2024.06.10.24308599; this version posted June 10, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

78. Gaetani, L. *et al.* Neurofilament light chain as a biomarker in neurological disorders. *J. Neurol. Neurosurg. Psychiatry* **90**, 870–881 (2019).

79. Martino Adami, P. V. *et al.* Matrix metalloproteinase 10 is linked to the risk of progression to dementia of the Alzheimer's type. *Brain* **145**, 2507–2517 (2022).

80. Li, W. *et al.* The Clinical Relevance and Functional Implications of Thymosin Beta-10 in Glioma. *Genet. Res.* **2023**, e5517445 (2023).

81. Pinteac, R., Montalban, X. & Comabella, M. Chitinases and chitinase-like proteins as biomarkers in neurologic disorders. *Neurol. Neuroimmunol. Neuroinflammation* **8**, e921 (2021).

82. Mattsson, N. *et al.* Cerebrospinal fluid microglial markers in Alzheimer's disease: elevated chitotriosidase activity but lack of diagnostic utility. *Neuromolecular Med.* **13**, 151–159 (2011).

83. Whelan, C. D. *et al.* Multiplex proteomics identifies novel CSF and plasma biomarkers of early Alzheimer's disease. *Acta Neuropathol. Commun.* **7**, 169 (2019).

84. Di Rosa, M. *et al.* Chitotriosidase and inflammatory mediator levels in Alzheimer's disease and cerebrovascular dementia. *Eur. J. Neurosci.* **23**, 2648–2656 (2006).

85. Oldoni, E. *et al.* CHIT1 at Diagnosis Reflects Long-Term Multiple Sclerosis Disease Activity. *Ann. Neurol.* **87**, 633–645 (2020).

86. Modeste, E. S. *et al.* Quantitative proteomics of cerebrospinal fluid from African Americans and Caucasians reveals shared and divergent changes in Alzheimer's disease. *Mol. Neurodegener.* **18**, 48 (2023).

87. Palmqvist, S. *et al.* Discriminative Accuracy of Plasma Phospho-tau217 for Alzheimer Disease vs Other Neurodegenerative Disorders. *JAMA* **324**, 772–781 (2020).

88. Gregoire, S. M. *et al.* The Microbleed Anatomical Rating Scale (MARS): reliability of a tool to map brain microbleeds. *Neurology* **73**, 1759–1766 (2009).

89. Puonti, O., Iglesias, J. E. & Van Leemput, K. Fast and sequence-adaptive whole-brain segmentation using parametric Bayesian modeling. *NeuroImage* **143**, 235–249 (2016).

90. Cerri, S. *et al.* A contrast-adaptive method for simultaneous whole-brain and lesion segmentation in multiple sclerosis. *NeuroImage* **225**, 117471 (2021).

91. Wik, L. *et al.* Proximity Extension Assay in Combination with Next-Generation Sequencing for High-throughput Proteome-wide Analysis. *Mol. Cell. Proteomics MCP* **20**, 100168 (2021).

92. Sun, B. B. *et al.* Plasma proteomic associations with genetics and health in the UK Biobank. *Nature* **622**, 329–338 (2023).

93. Karlsson, L. *et al.* Cerebrospinal fluid reference proteins increase accuracy and interpretability of biomarkers for brain diseases. *Nat. Commun.* **15**, 3676 (2024).

94. Yang, A. C. *et al.* A human brain vascular atlas reveals diverse mediators of Alzheimer's risk. *Nature* **603**, 885–892 (2022).

medRxiv preprint doi: https://doi.org/10.1101/2024.06.10.24308599; this version posted June 10, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

95. Skene, N. G. & Grant, S. G. N. Identification of Vulnerable Cell Types in Major Brain Disorders Using Single Cell Transcriptomes and Expression Weighted Cell Type Enrichment. *Front. Neurosci.* **10**, (2016).

96. Wang, J., Vasaikar, S., Shi, Z., Greer, M. & Zhang, B. WebGestalt 2017: a more comprehensive, powerful, flexible and interactive gene set enrichment analysis toolkit. *Nucleic Acids Res.* **45**, W130–W137 (2017).

97. Palmqvist, S. *et al.* An accurate fully automated panel of plasma biomarkers for Alzheimer's disease. *Alzheimers Dement. J. Alzheimers Assoc.* **19**, 1204–1215 (2023).

98. Miller, K. L. *et al.* Multimodal population brain imaging in the UK Biobank prospective epidemiological study. *Nat. Neurosci.* **19**, 1523–1536 (2016).

99. Griffanti, L. *et al.* BIANCA (Brain Intensity AbNormality Classification Algorithm): A new tool for automated segmentation of white matter hyperintensities. *NeuroImage* **141**, 191–205 (2016).

100. Palmqvist, S. *et al.* Cognitive effects of Lewy body pathology in clinically unimpaired individuals. *Nat. Med.* **29**, 1971–1978 (2023).

101. Mattsson-Carlgren, N. *et al.* Longitudinal plasma p-tau217 is increased in early stages of Alzheimer's disease. *Brain J. Neurol.* **143**, 3234–3241 (2020).