SUPPLEMENT

Genetic risk factors underlying white matter hyperintensities and cortical atrophy

Yash Patel[1,2 *], Jean Shin[1,2*], Eeva Sliz[3], Ariana Tang[1,2], Aniket Mishra[4], Rui Xia[5], Edith Hofer[6,7], Hema Sekhar Reddy Rajula[4], Ruiqi Wang[8], Frauke Beyer[4,9], Katrin Horn[10], Max Riedl[10], Jing Yu[11,12], Henry Völzke[13], Robin Bülow[14], Uwe Völker[15], Stefan Frenzel[16], Katharina Wittfeld[16], Sandra Van der Auwera[16,17], Thomas H. Mosley[18], Vincent Bouteloup[4,19], Jean-Charles Lambert[20], Geneviève Chêne[4,21], Carole Dufouil[4], Christophe Tzourio[4,21], Jean-François Mangin[22], Rebecca F. Gottesman[23], Myriam Fornage[5], Reinhold Schmidt[7], Qiong Yang[8], Veronica Witte[9], Markus Scholz[10,], Markus Loeffler[10,24], Gennady V Roshchupkin[11,12], M. Arfan Ikram[11], Hans J. Grabe[15,16], Sudha Seshadri[25], Stephanie Debette[4,26], Tomas Paus[27-30#], and Zdenka Pausova[1,2,27,30,31#]

1. The Hospital for Sick Children, Toronto, Ontario, Canada;
2. Departments of Physiology and Nutritional Sciences, University of Toronto, Toronto, Ontario, Canada;
3. Research Unit of Population Health, Faculty of Medicine, University of Oulu, Oulu, Finland;
4. University of Bordeaux, INSERM, Bordeaux Population Health research center, UMR1219, F-33000, Bordeaux, France;
5. The Brown Foundation Institute of Molecular Medicine, McGovern Medical School, The University of Texas Health Science Center at Houston, Houston, TX, USA;
6. Institut für Medizinische Informatik, Statistik und Dokumentation, Graz, Austria;
7. Division of Neurogeriatrics, Department of Neurology, Medical University of Graz, Graz, Austria;
8. Department of Biostatistics, Boston University School of Public Health, Boston, MA, USA;
9. Max Planck Institute for Human Cognitive and Brain Sciences, Leipzig, Germany;
10. Institute for Medical Informatics, Statistics and Epidemiology; Leipzig University; Leipzig; Germany
11. Department of Epidemiology, Erasmus MC University Medical Center Rotterdam, Rotterdam, The Netherlands;
12. Department of Radiology and Nuclear Medicine, Erasmus MC University Medical Center Rotterdam, Rotterdam, The Netherlands;
13. Institute for Community Medicine, University Medicine Greifswald, Greifswald, Germany;
14. Institute of Diagnostic Radiology and Neuroradiology, University Medicine Greifswald, Greifswald, Germany;
15. Interfaculty Institute of Genetics and Functional Genomics, University Medicine Greifswald, 17475 Greifswald, Germany;
16. Department of Psychiatry and Psychotherapy, University Medicine Greifswald, Greifswald, Germany;
17. German Centre for Neurodegenerative Diseases (DZNE), Site Rostock/Greifswald, Greifswald, Germany;
18. The MIND Center, The University of Mississippi Medical Center, Jackson, MS, USA;
19. CHU Bordeaux, CIC 1401 EC, Pôle Santé Publique, Bordeaux, France
20. U1167-RID-AGE facteurs de risque et déterminants moléculaires des maladies liées au vieillissement, INSERM, CHU Lille, Institut Pasteur de Lille, University of Lille, Lille, France
21. Department of Public Health, CHU de Bordeaux, Bordeaux, France
22. Université Paris-Saclay, CEA, CNRS, Neurospin, Baobab, Saclay, France
23. National Institute of Neurological Disorders and Stroke Intramural Research Program, Bethesda, Maryland, USA.
24. Leipzig Research Centre for Civilization Diseases; Leipzig University; Leipzig; Germany;
25. University of Texas, San Antonio, USA;
26. Bordeaux University Hospital, Department of Neurology, Institute for Neurodegenerative Diseases, F-33000, Bordeaux, France;
27. Centre hospitalier universitaire Sainte-Justine, University of Montreal, Montreal, Canada;
28. Departments of Psychiatry of Neuroscience, Faculty of Medicine, University of Montreal, Montreal, Canada;
29. Departments of Psychology and Psychiatry, University of Toronto, Toronto, Ontario, Canada;
30. ECOGENE-21, Chicoutimi, Canada.
31. Department of Pediatrics, Faculty of Medicine, University of Montreal, Montreal, Canada

[*]These authors contributed equally: Yash Patel, Jean Shin

[#]These authors jointly supervised this work: Tomas Paus, Zdenka Pausova

Supplementary Note 1. Study populations

Atherosclerosis Risk in Communities* (*ARIC)

The Atherosclerosis Risk in Communities (ARIC) study is a prospective epidemiologic study originally designed to investigate the etiology and natural history of atherosclerosis and clinical atherosclerotic diseases[1]. At its inception (1987-1989), 15,792 men and women, including 11,478 white and 4,266 black participants were recruited from four U.S. communities: Suburban Minneapolis, Minnesota; Washington County, Maryland; Forsyth County, North Carolina; and Jackson, Mississippi. In the first 3 communities, the sample reflects the demographic composition of the community. In Jackson, only black residents were enrolled. Participants were between age 45 and 64 years at their baseline examination in 1987-1989 when blood was drawn for DNA extraction and participants consented to genetic testing. Vascular risk factors and outcomes, including transient ischemic attack, stroke and dementia, were determined via standardized questionnaires and/or laboratory procedures. This study included 1,026 White participants who had genotype and brain MRI data. The ARIC Study has been approved by institutional review boards at participating institutions, and all participants provided written informed consent.

Austrian Stroke Prevention Family Study (ASPS-Fam)

The Austrian Stroke Prevention Family (ASPS-Fam) study is a prospective single-center community-based study on the cerebral effects of vascular risk factors in the normal aged population of the city of Graz, Austria[2]. ASPS-Fam represents an extension of the Austrian Stroke Prevention Study (ASPS). Between 2006 and 2013, study participants of the ASPS and their first-grade relatives were invited to enter ASPS-Fam. Inclusion criteria were no history of previous stroke or dementia and a normal neurologic examination. A total of 419 individuals from 176 families of 2 to 6 members, all European Caucasians, were included into the study. All participants underwent a thorough diagnostic workup including clinical history, laboratory evaluation, cognitive testing, and an extended vascular risk factor assessment. The 297 participants who passed genotyping quality control and underwent brain MRI scanning were included in this study. All study protocols were approved by the ethics committee of the Medical University of Graz, Austria, and written informed consent was obtained from all participants.

Three City Dijon (3C-Dijon)

The Three City Dijon (3C-Dijon) study is a population-based cohort of French non-institutionalized individuals aged 65 years and older[3]. The study protocol was approved by the Ethical Committee of the University Hospital of Kremlin-Bicêtre and each participant signed an informed consent. The overall design of the 3C-Dijon study is detailed elsewhere[3–5]. The Three City Dijon (3C-Dijon) study is a population-based cohort of 4931 French non-institutionalized individuals aged 65 years and older (PMID: 14598854). This study involved 436 3C-Dijon participants with genotype and brain MRI data.

Framingham Heart Study (FHS)

The Framingham Heart Study is a community-based, longitudinal cohort study that was started in 1948 when the National Heart Institute chose the town of Framingham, Massachusetts, USA, to conduct an epidemiological study[6]. The primary aim of the study was to identify determinants of cardiovascular disease to guide public health prevention[6]. The Original Cohort constitutes of 5,209 individuals most of whom were selected based on random sampling with an additional group of volunteers[7]. In 1971 and in 2002, respectively, the Offspring Cohort (FHS-GEN2) and the Third Generation (FHS-GEN3) were initiated: the participants of GEN2 (N=5,124) are children of the Original Cohort or spouses of these children, and the participants of GEN3 (N=4,095) have at least one parent in the GEN2 cohort[6]. In addition, two OMNI cohorts have begun (in years 1994 and 2003) to include individuals with diverse ethnic backgrounds[6]. As part of multiple large ancillary studies on brain structure and cognitive function starting in March 1999, FHS participants were recruited to undergo MRI of the brain[8]. Written informed consent is obtained on every visit from all participants[6]. The institutional review board of Boston University Medical Center has approved the study protocol. This study included 4,084 White participants, from the Original, GEN2, GEN3, New Offspring Spouse, and OMNI1 cohorts, who had genotyping data and underwent brain MRI.

LIFE-Adult

LIFE-Adult is a population-based study and a part of the large-scale research project LIFE (Leipzig Research Center for Civilization Diseases). 10,000 residents (40 – 79 years old) from the district of Leipzig in Germany were recruited and extensively phenotyped[9]. All subjects gave written informed consent. The procedures were conducted according to the Declaration of Helsinki and approved by the University of Leipzig's ethics committee (registration-number: 263-2009-14122009). This study included 2,146 individuals with both genotype and brain MRI data.

MEMENTO

The MEMENTO cohort is a French memory clinic-based prospective study of 2,323 participants, recruited between April 2011 and June 2014, with cognitive complaints but no dementia at baseline[10]. The recruitment took place within the French national network of university-based memory clinics (Centres de Mémoires de Ressources et de Recherche [CMRR]). The study aim is to improve the understanding of Alzheimer's disease and related dementias' natural history and identify new phenotypes of participants who will develop dementia over time. This study was performed in accordance with the guidelines of the Declaration of Helsinki. The study protocol has been approved by the local ethics committee ("Comité de Protection des Personnes Sud-Ouest et Outre Mer III"; approval number 2010-A01394-35). All participants provided written informed consent. This study included 1,950 individuals with both genotype and brain MRI data.

Rotterdam Study (RS)

The Rotterdam Study (RS) is a prospective cohort study initiated in 1990 in the city of Rotterdam, the Netherlands[11]. It was designed to study risk factors of diseases in the elderly, including cardiovascular, neurological, ophthalmological and endocrine diseases[11]. The initial cohort (RS-I) was among 7,983 persons aged 55-106 years living in the Ommoord district of Rotterdam[11]. The pilot phase of the study took place in 1989 and cohort recruitment in years 1990-1993 (RS-I). Since year 2000, new cohorts have been initiated (RS-II, RS-III, RS-IV), and up to 2008, nearly 15,000 participants had been recruited[12]. All participants had DNA extracted at their first visit. In 1995-1996, 563 non-demented persons of the 7,983 participants from the Rotterdam Study I were randomly selected in strata of age and sex to undergo cranial MRI scanning. From 2005 onwards, cranial MRI scanning, including assessment of cerebral white matter lesion burden, was added to the core protocol[13]. The study has been approved by the institutional board (Medical Ethics Committee) of the Erasmus Medical Center and by the review board of The Netherlands Ministry of Health, Welfare and Sports, and the approval has been renewed every 5 years and with the introduction of major new elements in the study[12]. Analyses of this study included 6,872 participants with both genotype and brain MRI data in RS-I, RS-II and RS-III cohorts.

Study of Health in Pomerania (SHIP)

Study of Health in Pomerania (SHIP) is a population-based study with an overall objective to assess prevalence and incidence of common risk factors and diseases as well as to investigate the complex associations among risk factors and health outcomes[14–16]. Two independent cohorts, SHIP-START and SHIP-TREND, were selected in the same region of West Pomerania in the north-east of Germany[15]. From the total of 213,057 inhabitants in the region in 1996, a two-stage stratified cluster sample of adults between 20-79 years of age were drawn, with a net sample of 6,265 eligible individuals[14–16]. The second follow up examinations of the first SHIP cohort (SHIP-START-2) and baseline examinations of the second one (SHIP-TREND-0) took place between 2009 and 2012[14–16]. This study included 1,109 participants from SHIP-START-2 and 2,063 participants from SHIP-TREND-0 who had both genotype and brain MRI data. SHIP and SHIP
TREND were approved by the local ethics committee.

UK Biobank

The UK Biobank is a prospective cohort study with deep genetic, physical and health data collected on ~500K individuals across the United Kingdom from 2006-2010[17]. The purpose of the study is to allow detailed investigations of the genetic and nongenetic determinants of many important diseases of middle and old age, such as dementia, ultimately aiming at improving the prevention, diagnosis and treatment of a wide range of serious and life-threatening illnesses. Participants aged 40-69 years were recruited between 2006 and 2010 by letter and answered detailed questions about their health and lifestyle, had body measures taken and donated blood, urine and saliva that allow various type of assay (e.g., genetic, biochemical, metabolomic, proteomic and hematologic). An imaging (including MRI of the brain) extension to the existing UK Biobank study was funded in 2016 to scan 100,000 subjects from the existing cohort[18]. As of February 2023, ~52K participants had undergone brain MRI, and ~5K of them had repeat scans within 2 years after their initial scan. All participants provided informed consent at baseline assessment. The UK Biobank obtained approval from the National Information Governance Board for Health and Social Care and the National Health Service Northwest Multicentre Research Ethics Committee (reference # 11/NW/0382). This study used data accessed under UK Biobank project ID 43688 and included 31,082 unrelated White-British individuals with genotype data and brain MRI data acquired during their first imaging visit. It was approved by the SickKids Research Ethics Board (#1000073323).

FinnGen

FinnGen (https://www.finngen.fi/en), initiated in 2017, is a collaborative public-private research initiative integrating genome data with digital healthcare records from approximately 500,000 Finns. It is a nationwide project that aims to generate new medical and therapeutic insights into human diseases. FinnGen is a pre-competitive alliance involving Finnish biobanks and their affiliated institutions (universities and hospitals), international pharmaceutical companies, and the

Finnish biobank cooperative (FINBB). A full list of FinnGen partners can be found here: https://www.finngen.fi/en/partners. In this study, we utilized data from the FinnGen data freeze 12, which includes 3,624 cases with vascular dementia (FinnGen endpoint F5_VASCDEM), 21,257 cases with all-cause dementia (FinnGen endpoint KRA_PSY_DEMENTIA), and 9,690 cases with late-onset Alzheimer's disease (FinnGen endpoint AD_LO), covering a total sample of 500,348 individuals. Case-control statuses were defined using International Classification of Diseases (ICD) codes as defined below:

Phenotype	FinnGen endpoint	Case definition	Control definition
Vascular dementia	F5_VASCDEM	ICD-10: F01 or ICD-9: 4378	Those not listed as cases and not having an entry of ICD-10: F00-F09; ICD-9: 290, 3310, 4378A; or ICD-8: 290.
All-cause dementia	KRA_PSY_DEMENTIA	ICD-10: F00-F09, F05.1, or G30; ICD-9: 290, 2912A, 2828C, 2941A, 3310A, 3311A, or 4378A; ICD-8: 290	Those not listed as cases.
Late-onset Alzheimer's disease	AD_LO	ICD-10: F00.1*, F00.10*, F00.10*G30.1, G30.1, G30.1+F00.10	Those not listed as cases and not having an entry of ICD-10: G30 or ICD-9: 3310.

FinnGen samples were genotyped with Illumina and Affymetrix arrays (Illumina Inc., San Diego, and Thermo Fisher Scientific, Santa Clara, CA, USA), and genotype calls were made with the GenCall or zCall (for Illumina) and the AxiomGT1 algorithm for Affymetrix data. Individuals with ambiguous sex, high genotype missingness (>5%), excess heterozygosity (+-4SD), and non-Finnish ancestry were excluded, as well as all variants with high missingness (>2%), low Hardy–Weinberg equilibrium p-value ($<1e^{-6}$) and minor allele count < 3. Array data pre-phasing was carried out with Eagle 2.3.527[19], and the number of conditioning haplotypes was set to 20,000. Genotype imputation was done with Beagle 4.128[20] (as described in https://doi.org/10.17504/protocols.io.xbgfijw) by using the SISu v3 population-specific reference panel developed from high-quality data for 3,775 high-coverage (25-30x) whole-genome sequences in Finns.

FinnGen participants provided written informed consent for biobank research based on the Finnish Biobank Act. Alternatively, separate research cohorts, collected before the Finnish Biobank Act came into effect (in September 2013) and the start of FinnGen (August 2017), were collected based on study-specific consents and later transferred to the Finnish biobanks after approval by Fimea (Finnish Medicines Agency), the National Supervisory Authority for Welfare and Health. Recruitment protocols followed the biobank protocols approved by Fimea. The Coordinating Ethics Committee of the Hospital District of Helsinki and Uusimaa (HUS) statement number for the FinnGen study is Nr HUS/990/2017. The FinnGen study is approved by Finnish Institute for Health and Welfare (permit numbers: THL/2031/6.02.00/2017, THL/1101/5.05.00/2017, THL/341/6.02.00/2018, THL/2222/6.02.00/2018, THL/283/6.02.00/2019, THL/1721/5.05.00/2019 and THL/1524/5.05.00/2020), Digital and population data service agency (permit numbers: VRK43431/2017-3, VRK/6909/2018-3, VRK/4415/2019-3), the Social Insurance Institution (permit numbers: KELA 58/522/2017, KELA 131/522/2018, KELA 70/522/2019, KELA 98/522/2019, KELA 134/522/2019, KELA 138/522/2019, KELA 2/522/2020, KELA 16/522/2020), Findata permit numbers THL/2364/14.02/2020, THL/4055/14.06.00/2020, THL/3433/14.06.00/2020, THL/4432/14.06/2020, THL/5189/14.06/2020, THL/5894/14.06.00/2020, THL/6619/14.06.00/2020, THL/209/14.06.00/2021, THL/688/14.06.00/2021, THL/1284/14.06.00/2021, THL/1965/14.06.00/2021, THL/5546/14.02.00/2020, THL/2658/14.06.00/2021, THL/4235/14.06.00/2021, Statistics Finland (permit numbers: TK-53-1041-17 and TK/143/07.03.00/2020 (earlier TK-53-90-20) TK/1735/07.03.00/2021, TK/3112/07.03.00/2021) and Finnish Registry for Kidney Diseases permission/extract from the meeting minutes on 4th July 2019. The Biobank Access Decisions for FinnGen samples and data utilized in FinnGen Data Freeze 11 include: THL Biobank BB2017_55, BB2017_111, BB2018_19, BB_2018_34, BB_2018_67, BB2018_71, BB2019_7, BB2019_8, BB2019_26, BB2020_1, BB2021_65, Finnish Red Cross Blood Service Biobank 7.12.2017, Helsinki Biobank HUS/359/2017, HUS/248/2020, HUS/430/2021 §28, §29, HUS/150/2022 §12, §13, §14, §15, §16, §17, §18, §23, §58, §59, HUS/128/2023 §18, Auria Biobank AB17-5154 and amendment #1 (August 17 2020) and amendments BB_2021-0140, BB_2021-0156 (August 26 2021, Feb 2 2022), BB_2021-0169, BB_2021-0179, BB_2021-0161, AB20-5926 and amendment #1 (April 23 2020) and it´s modifications (Sep 22 2021), BB_2022-0262, BB_2022-0256, Biobank Borealis of

Northern Finland_2017_1013, 2021_5010, 2021_5010 Amendment, 2021_5018, 2021_5018 Amendment, 2021_5015, 2021_5015 Amendment, 2021_5015 Amendment_2, 2021_5023, 2021_5023 Amendment, 2021_5023 Amendment_2, 2021_5017, 2021_5017 Amendment, 2022_6001, 2022_6001 Amendment, 2022_6006 Amendment, 2022_6006 Amendment, 2022_6006 Amendment_2, BB22-0067, 2022_0262, 2022_0262 Amendment, Biobank of Eastern Finland 1186/2018 and amendment 22§/2020, 53§/2021, 13§/2022, 14§/2022, 15§/2022, 27§/2022, 28§/2022, 29§/2022, 33§/2022, 35§/2022, 36§/2022, 37§/2022, 39§/2022, 7§/2023, 32§/2023, 33§/2023, 34§/2023, 35§/2023, 36§/2023, 37§/2023, 38§/2023, 39§/2023, 40§/2023, 41§/2023, Finnish Clinical Biobank Tampere MH0004 and amendments (21.02.2020 & 06.10.2020), BB2021-0140 8§/2021, 9§/2021, §9/2022, §10/2022, §12/2022, 13§/2022, §20/2022, §21/2022, §22/2022, §23/2022, 28§/2022, 29§/2022, 30§/2022, 31§/2022, 32§/2022, 38§/2022, 40§/2022, 42§/2022, 1§/2023, Central Finland Biobank 1-2017, BB_2021-0161, BB_2021-0169, BB_2021-0179, BB_2021-0170, BB_2022-0256, BB_2022-0262, BB22-0067, Decision allowing to continue data processing until 31st Aug 2024 for projects: BB_2021-0179, BB22-0067,BB_2022-0262, BB_2021-0170, BB_2021-0164, BB_2021-0161, and BB_2021-0169, and Terveystalo Biobank STB 2018001 and amendment 25th Aug 2020, Finnish Hematological Registry and Clinical Biobank decision 18th June 2021, Arctic biobank P0844: ARC_2021_1001.

Supplementary Table 1. Cohort characteristics

Study Cohort	N (51,065)	Male (%)	Age (years)	Age range (years)	BMI (kg/m^2)	Current smoking (%)	Hypertension (%)	Type 2 diabetes (%)
ARIC	1,026	42.0	76.7 (5.2)	67 - 90	27.9 (5.1)	5.2	70.8	26.1
ASPS-Fam	297	39.1	64.6 (10.5)	38 - 84	26.6 (4.7)	16.2	69.0	7.1
Dijon_3C	436	42.0	72.6 (4.0)	65 - 84	25.8 (3.6)	8.0	75.9	7.6
FHS	4,084	46.3	56.8 (13.1)	25 - 100	27.9 (5.4)	8.5	34.5	7.7
LIFE-Adult	2,146	52.6	57.6 (15.3)	19 - 82	26.9 (4.3)	14.7	49.6	8.3
MEMENTO	1,950	37.6	70.9 (8.6)	33 - 93	25.5 (4.3)	6.5	61.2	8.7
RS	6,872	45.6	62.4 (8.8)	45 - 95	27.3 (4.1)	18.5	59.0	9.5
SHIP-START-2	1,109	48.3	55.6 (12.8)	31 - 90	27.6 (4.4)	19.4	47.6	8.6
SHIP-TREND-0	2,063	48.0	51.5 (14.0)	21 - 83	27.6 (4.5)	23.7	43.1	8.4
UK Biobank	31,082	47.3	63.8 (7.5)	45 - 81	26.5 (4.4)	3.2	50.0	2.4

Supplementary Table 2. Meta-GWAS of PC1 (WMH and insular CT) and their previous GWAS associations with WMH or CT

			Previous GWAS-significant loci*		References**	
PC1 locus #	rsID	chr:pos (hg19)	WMH	CT	WMH	CT
4	rs7454868, rs190945449	6:26799828, 6:26828359	no	no		
8	rs62477728	7:75132471	no	no		
9	rs11191163, rs11191268	10:103733624, 10:104115262	no	no		
14	rs11075976	16:51498626	no	no		
19	rs112783265	18:32358907	no	no		
7	rs798528	7:2772431	no	yes		9-11
12	rs3765066	15:75140854	no	yes		10,12
20	rs2072859	22:38322350	no	yes		9,10,12
5	rs13208741	6:45461253	no	yes		9-12
10	rs4630220	10:105459116	yes	no	2-5	
11	rs11838776	13:111040681	yes	no	2-4	
1	rs3762515	2:56150864	yes	no	2-5,7,8	
17	rs563065735, rs12950988	17:43129103, 17:43127708	yes	no	2-4	
18	rs3744027	17:73888743	yes	no	1-8	
6	rs4272224	6:151035800	yes	yes	2-4,6,8	9,10,12
2	rs147100405, rs72932753	2:203720774, 2:203670122	yes	yes	1-4	9,10
3	rs79934840, rs11711420	3:183403240, 3:183349010	yes	yes	2,4	9,10
13	rs17616633	16:51451683	yes	yes	2,4,6	9-13
15	rs9308343	16:87224857	yes	yes	2-6	9-14
16	rs1472932	17:19220666	yes	yes	2	10

*yes (no): Top SNP or it's LD-proxy with r^2>0.2 was (not) associated with WMH or CT at the genome-wide significance level of 5e-08
GWAS catalogue search was done on May 28, 2024.

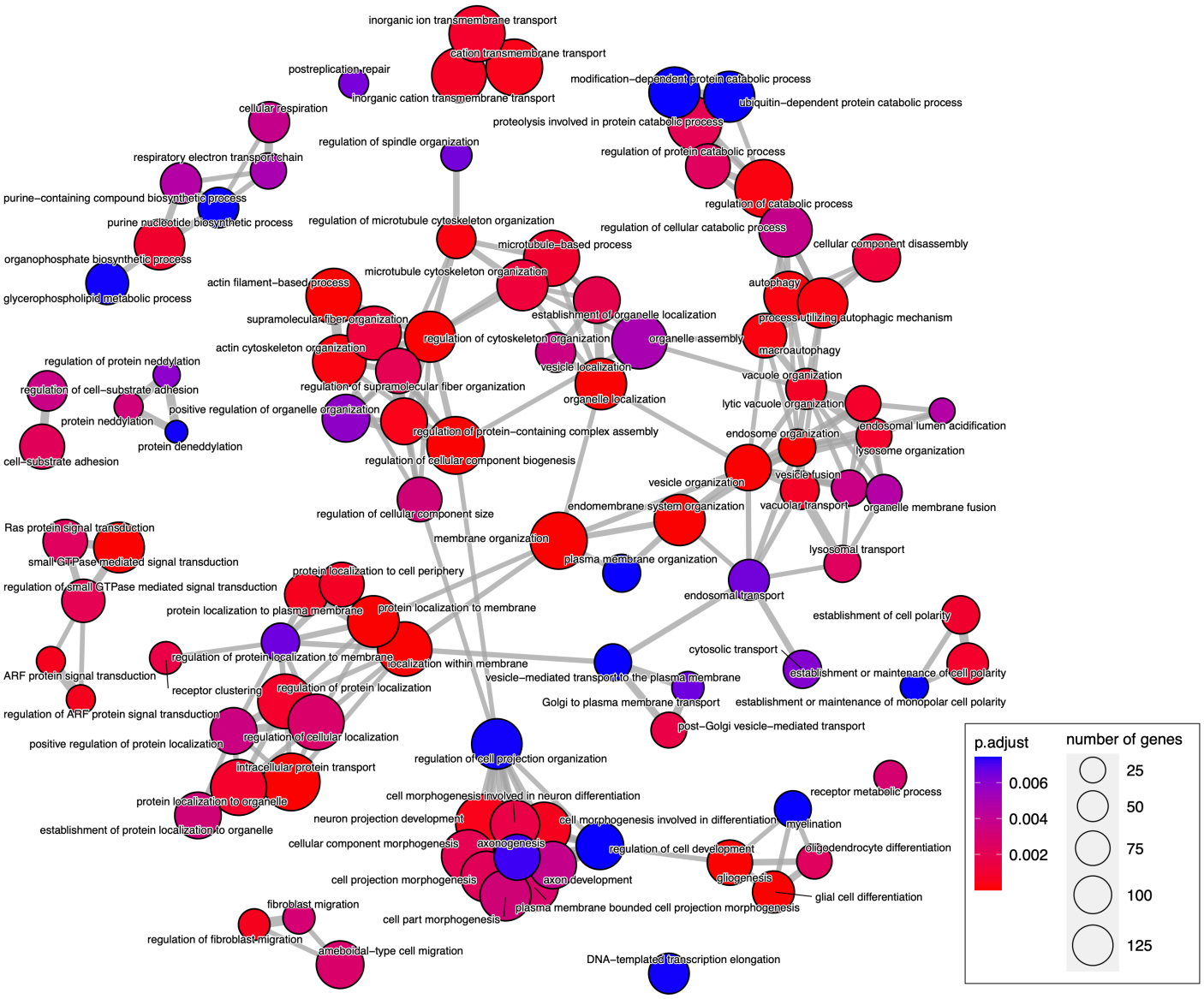
**References

1. Jian, Xueqiu et al. "Exome Chip Analysis Identifies Low-Frequency and Rare Variants in MRPL38 for White Matter Hyperintensities on Brain Magnetic Resonance Imaging." Stroke vol. 49,8 (2018): 1812-1819. doi:10.1161/STROKEAHA.118.020689
2. Persyn, Elodie et al. "Genome-wide association study of MRI markers of cerebral small vessel disease in 42,310 participants." Nature Communications vol. 11,1 2175. 1 May. 2020, doi:10.1038/s41467-020-15932-3
3. Armstrong, Nicola J et al. "Common Genetic Variation Indicates Separate Causes for Periventricular and Deep White Matter Hyperintensities." Stroke vol. 51,7 (2020): 2111-2121. doi:10.1161/STROKEAHA.119.027544
4. Sargurupremraj, Muralidharan et al. "Cerebral small vessel disease genomics and its implications across the lifespan." Nature Communications vol. 11,1 6285. 8 Dec. 2020, doi:10.1038/s41467-020-19111-2
5. Verhaaren, Benjamin F J et al. "Multiethnic genome-wide association study of cerebral white matter hyperintensities on MRI." Circulation. Cardiovascular genetics vol. 8,2 (2015): 398-409. doi:10.1161/CIRCGENETICS.114.000858
6. Smith, Stephen M et al. "An expanded set of genome-wide association studies of brain imaging phenotypes in UK Biobank." Nature Neuroscience vol. 24,5 (2021): 737-745. doi:10.1038/s41593-021-00826-4
7. Rutten-Jacobs, Loes C A et al. "Genetic Study of White Matter Integrity in UK Biobank (N=8448) and the Overlap With Stroke, Depression, and Dementia." Stroke vol. 49,6 (2018): 1340-1347. doi:10.1161/STROKEAHA.118.020811
8. Traylor, Matthew et al. "Genetic variation in PLEKHG1 is associated with white matter hyperintensities (n = 11,226)." Neurology vol. 92,8 (2019): e749-e757. doi:10.1212/WNL.0000000000006952
9. Shadrin, Alexey A et al. "Vertex-wise multivariate genome-wide association study identifies 780 unique genetic loci associated with cortical morphology." NeuroImage vol. 244 (2021): 118603. doi:10.1016/j.neuroimage.2021.118603
10. van der Meer, Dennis et al. "The genetic architecture of human cortical folding." Science advances vol. 7,51 (2021): eabj9446. doi:10.1126/sciadv.abj9446
11. van der Meer, Dennis et al. "Understanding the genetic determinants of the brain with MOSTest." Nature communications vol. 11,1 3512. 14 Jul. 2020, doi:10.1038/s41467-020-17368-1
12. Makowski, Carolina et al. "Larger cerebral cortex is genetically correlated with greater frontal area and dorsal thickness." Proceedings of the National Academy of Sciences of the United States of America vol. 120,11 (2023): e2214834120. doi:10.1073/pnas.2214834120
13. Hofer, Edith et al. "Genetic correlations and genome-wide associations of cortical structure in general population samples of 22,824 adults." Nature communications vol. 11,1 4796. 22 Sep. 2020, doi:10.1038/s41467-020-18367-y
14. Grasby, Katrina L et al. "The genetic architecture of the human cerebral cortex." Science (New York, N.Y.) vol. 367,6484 (2020): eaay6690. doi:10.1126/science.aay6690

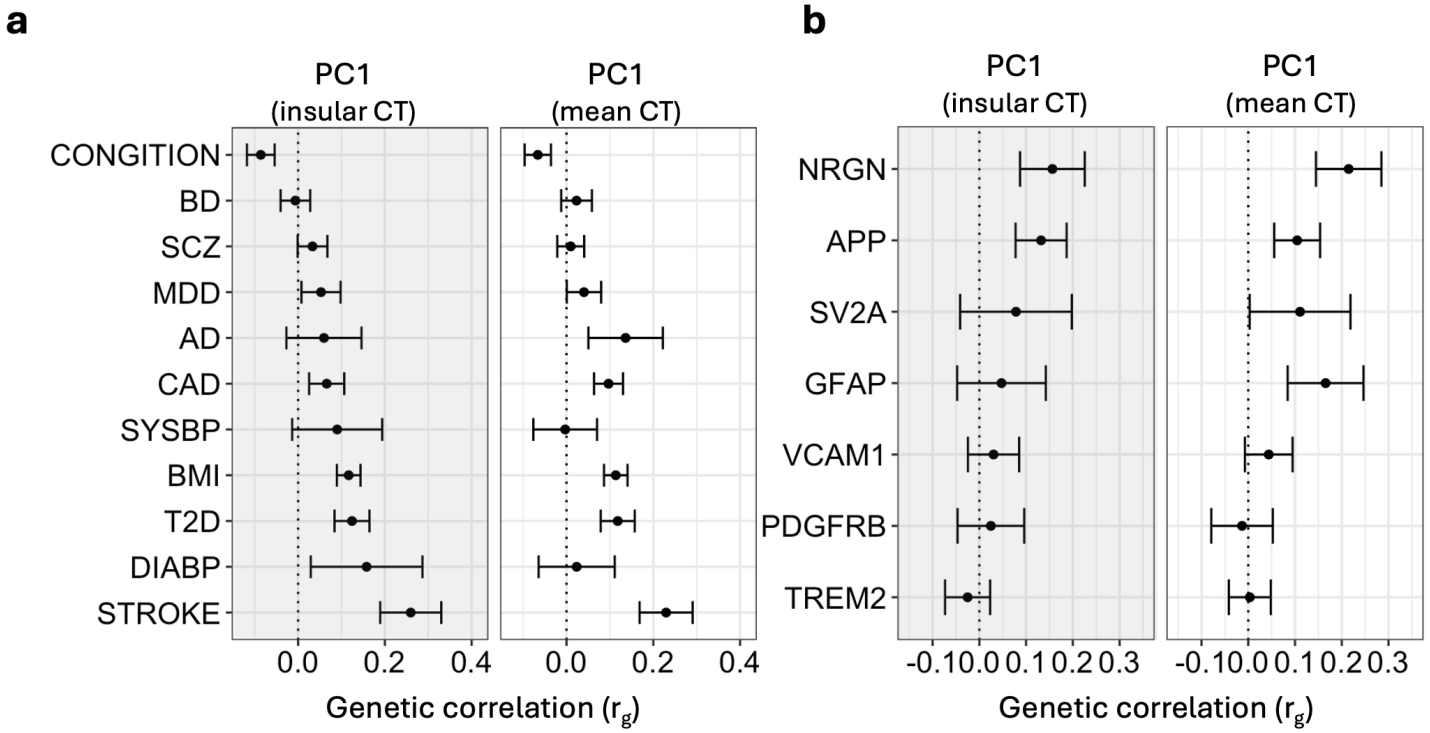
Supplementary Table 3. Associations between polygenic risk scores (PRS) of WMH and insular or mean cortical thickness (with adjustment for sample-overlap)

		Insular cortical thickness				Mean cortical thickness			
P-value Threshold	**Number of SNPs**	**Coefficient**	**SE**	**R-squared**	**P**	**Coefficient**	**SE**	**R-squared**	**P**
5.00E-08	1	0.0016	0.0025	1.29E-05	5.26E-01	-0.0023	0.0024	3.06E-05	3.30E-01
1.00E-06	3	-0.0091	0.0051	1.04E-04	7.27E-02	-0.0062	0.0047	5.68E-05	1.84E-01
1.00E-05	7	-0.0177	0.0090	1.24E-04	4.93E-02	-0.0043	0.0084	8.65E-06	6.04E-01
0.0001	28	-0.0343	0.0216	8.12E-05	1.12E-01	0.0040	0.0201	1.31E-06	8.40E-01
0.001	249	-0.2455	0.0769	3.28E-04	1.41E-03	0.0410	0.0715	1.06E-05	5.66E-01
0.01	2235	-0.6417	0.2759	1.74E-04	2.00E-02	-0.0285	0.2565	3.99E-07	9.11E-01
0.05	9992	-2.7940	0.6377	6.17E-04	1.18E-05	-1.9955	0.5930	3.64E-04	7.66E-04
0.1	18416	-3.7471	0.9038	5.53E-04	3.39E-05	-3.4368	0.8402	5.38E-04	4.32E-05
0.5	63748	-6.7320	2.1298	3.21E-04	1.57E-03	-7.9867	1.9798	5.24E-04	5.50E-05
1	88930	-9.1529	2.9065	3.19E-04	1.64E-03	-10.3110	2.7019	4.69E-04	1.36E-04
PRS_{PCA}-WMH*		-0.0109	0.0026	5.54E-04	3.33E-05	-0.0079	0.0024	3.39E-04	1.17E-03

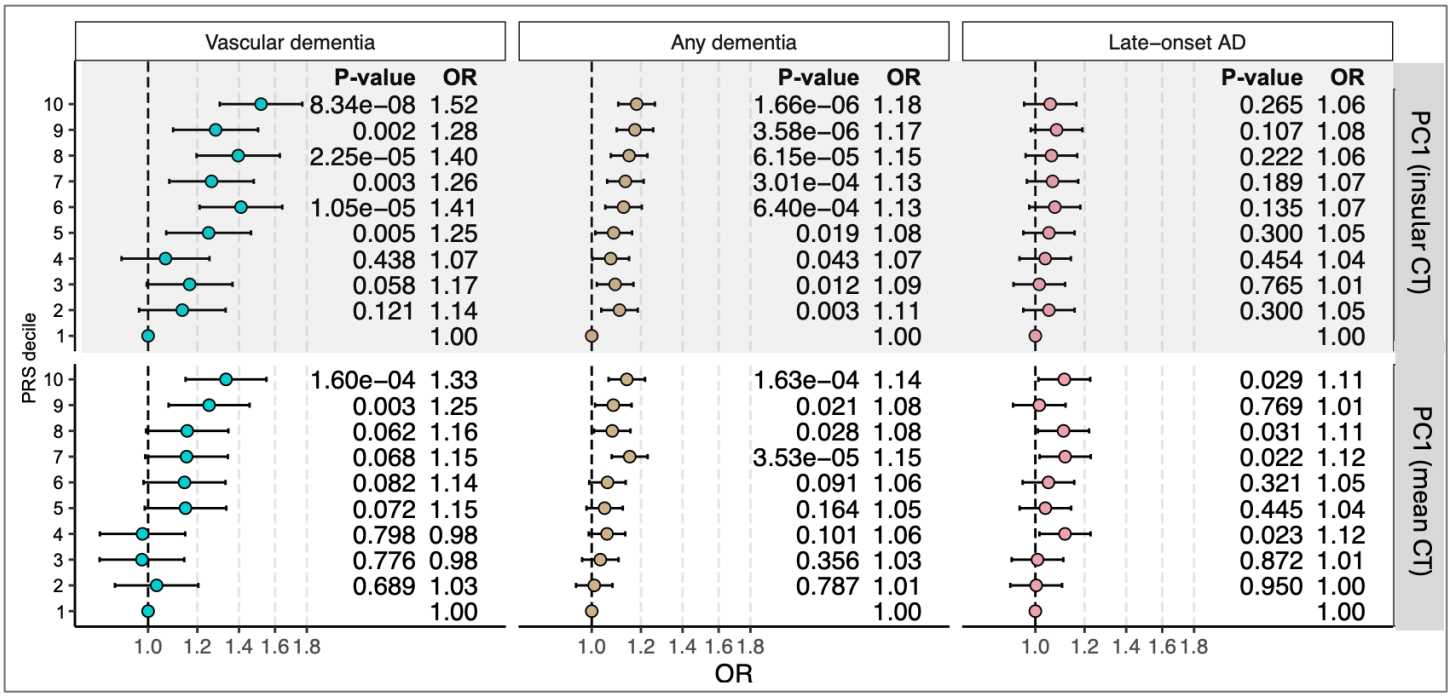
*PCA approach[31] was applied to the polygenic scores across the 10 p-value thresholds and the first principal component (PRS_{PCA}-WMH) was tested for the association with the covariate-adjusted insular and mean cortical thickness. PRSPCA-WMH was positively associated with WMH (p=1.2e-09) and explains 0.12% of the total variance of WMH and 0.44% of total variance of the polygenic risk scores across the 10 p-value thresholds in the UK Biobank participants.



Supplementary Figure 1. Gene ontology enrichment analysis of genes co-expressed with eQTL-regulated genes by genome-wide significant variants from the GWAS of PC1. Each node represents a significant biological process. Clusters of terms with high similarity are linked by edges.



Supplementary Figure 2. Genetic correlations between PC1 (derived from WMH and insular or mean CT) and (a) cognitive, vascular-risk, neurodegenerative and neuropsychiatric traits and (b) plasma proteins related to neurodegenerative biological processes. DIAB: Diastolic Blood Pressure, T2D: Type 2 Diabetes, BMI: Body Mass Index, SYSBP: Systolic Blood Pressure, CAD: Coronary Artery Disease, AD: Alzheimer's Disease, MDD: Major Depressive Disorder, SCZ: Schizophrenia, BD Bipolar Disorder. NRGN: neurogranin, APP: amyloid precursor protein, SV2A: Synaptic vesicle protein 2, GFAP: Glial fibrillary acidic protein, VCAM1: Vascular cell adhesion protein 1, PDGFRB: platelet derived growth factor receptor beta, TREM2: Triggering receptor expressed on myeloid cells 2. These genes were *a priori* selected as related to biomarkers of neurodegeneration[32], and for which there was available plasma proteomic GWAS summary statistics from the UKBB-PPP initiative[33]. Error bars represent standard error of the estimate.



Supplementary Figure 3. Association between polygenic risk score (PRS) of WMH and insular CT or mean CT-derived PC1 and the risk of vascular dementia, all-cause dementia, and Alzheimer's disease. The odds ratios were calculated in FinnGen (n=500,348) by comparing each of the top nine PRS deciles to the lowest decile and adjusting for age, sex, the first 10 genetic principal components and genotyping arrays. Error bars represent 95% confidence intervals.

Supplementary References

1. Wright, J. D. *et al.* The ARIC (Atherosclerosis Risk In Communities) Study: JACC Focus Seminar 3/8. *J Am Coll Cardiol* **77**, 2939–2959 (2021).
2. Ghadery, C. *et al.* R2* mapping for brain iron: associations with cognition in normal aging. *Neurobiology of Aging* **36**, 925–932 (2015).
3. Antoniak, M. *et al.* Vascular factors and risk of dementia: design of the Three-City Study and baseline characteristics of the study population. *Neuroepidemiology* **22**, 316–325 (2003).
4. Godin, O. *et al.* White matter lesions as a predictor of depression in the elderly: the 3C-Dijon study. *Biological psychiatry* **63**, 663–669 (2008).
5. Soumaré, A. *et al.* White matter lesions volume and motor performances in the elderly. *Annals of Neurology: Official Journal of the American Neurological Association and the Child Neurology Society* **65**, 706–715 (2009).
6. Tsao, C. W. & Vasan, R. S. Cohort Profile: The Framingham Heart Study (FHS): Overview of milestones in cardiovascular epidemiology. *International Journal of Epidemiology* **44**, 1800–1813 (2015).
7. Dawber, T. R., Meadors, G. F. & Moore, F. E. Epidemiological approaches to heart disease: the Framingham Study. *American journal of public health* **41**, 279–281 (1951).
8. Massaro, J. M. *et al.* Managing and analysing data from a large-scale study on Framingham Offspring relating brain structure to cognitive function. *Stat Med* **23**, 351–367 (2004).
9. Loeffler, M. *et al.* The LIFE-Adult-Study: objectives and design of a population-based cohort study with 10,000 deeply phenotyped adults in Germany. *BMC public health* **15**, 1–14 (2015).
10. Dufouil, C. *et al.* Cognitive and imaging markers in non-demented subjects attending a memory clinic: study design and baseline findings of the MEMENTO cohort. *Alzheimer's Research & Therapy* **9**, 1–13 (2017).
11. Hofman, A. *et al.* The Rotterdam Study: Objectives and design update. *European Journal of Epidemiology* **22**, 819–829 (2007).

12. Ikram, M. A. *et al. Objectives, Design and Main Findings until 2020 from the Rotterdam Study*. *European Journal of Epidemiology* vol. 35 (Springer Netherlands, 2020).
13. Ikram, M. A. *et al.* The Rotterdam Scan Study: design update 2016 and main findings. *European journal of epidemiology* **30**, 1299–1315 (2015).
14. John, U. *et al.* Study of Health in Pomerania (SHIP): A health examination survey in an east German region: Objectives and design. *Sozial- und Praventivmedizin* **46**, 186–194 (2001).
15. Völzke, H. *et al.* Cohort profile: The study of health in Pomerania. *International Journal of Epidemiology* **40**, 294–307 (2011).
16. Völzke, H. *et al.* Cohort profile update: the study of health in Pomerania (SHIP). *International Journal of Epidemiology* **51**, e372–e383 (2022).
17. Sudlow, C. *et al.* UK biobank: an open access resource for identifying the causes of a wide range of complex diseases of middle and old age. *PLoS medicine* **12**, e1001779 (2015).
18. Miller, K. L. *et al.* Multimodal population brain imaging in the UK Biobank prospective epidemiological study. *Nature neuroscience* **19**, 1523–1536 (2016).
19. Loh, P.-R. *et al.* Reference-based phasing using the Haplotype Reference Consortium panel. *Nat Genet* **48**, 1443–1448 (2016).
20. Browning, B. L. & Browning, S. R. Genotype Imputation with Millions of Reference Samples. *Am J Hum Genet* **98**, 116–126 (2016).
21. Howard, G. *et al.* Cigarette smoking and other risk factors for silent cerebral infarction in the general population. *Stroke* **29**, 913–917 (1998).
22. Schmidt, R., Fazekas, F., Kapeller, P., Schmidt, H. & Hartung, H.-P. MRI white matter hyperintensities: three-year follow-up of the Austrian Stroke Prevention Study. *Neurology* **53**, 132–132 (1999).
23. Maillard, P. *et al.* An automated procedure for the assessment of white matter hyperintensities by multispectral (T1, T2, PD) MRI and an evaluation of its between-centre reproducibility based on two large community databases. *Neuroradiology* **50**, 31–42 (2008).

24. Lampe, L. *et al.* Visceral obesity relates to deep white matter hyperintensities via inflammation. *Annals of neurology* **85**, 194–203 (2019).
25. Regy, M. *et al.* Association of APOE ε4 with cerebral gray matter volumes in non-demented older adults: the MEMENTO cohort study. *NeuroImage* **250**, 118966 (2022).
26. Samaille, T. *et al.* Contrast-based fully automatic segmentation of white matter hyperintensities: method and validation. *PloS one* **7**, e48953 (2012).
27. Schmidt, P. *et al.* An automated tool for detection of FLAIR-hyperintense white-matter lesions in multiple sclerosis. *Neuroimage* **59**, 3774–3783 (2012).
28. 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).
29. Han, B. & Eskin, E. Random-effects model aimed at discovering associations in meta-analysis of genome-wide association studies. *The American Journal of Human Genetics* **88**, 586–598 (2011).
30. Minelli, C. *et al.* The use of two-sample methods for Mendelian randomization analyses on single large datasets. *Int J Epidemiol* **50**, 1651–1659 (2021).
31. Coombes, B. J., Ploner, A., Bergen, S. E. & Biernacka, J. M. A principal component approach to improve association testing with polygenic risk scores. *Genet Epidemiol* **44**, 676–686 (2020).
32. Hansson, O. Biomarkers for neurodegenerative diseases. *Nature medicine* **27**, 954–963 (2021).
33. Sun, B. B. *et al.* Plasma proteomic associations with genetics and health in the UK Biobank. *Nature* **622**, 329–338 (2023).