

基于慢性脑灌注不足的蛋白质组孟德尔随机化研究揭示白质高信号的血浆靶点

刘小燕¹, 余卓远², 杨 美¹

(1.重庆医科大学基础医学院神经科学研究中心, 重庆 400016; 2.重庆医科大学附属第一医院泌尿外科, 重庆 400016)

【摘要】目的:以白质高信号(white matter hyperintensities, WMH)作为人群层面的主要生物标志物,旨在利用人类遗传学证据筛选并确定与之相关的血浆蛋白。**方法:**在欧洲人群队列中系统性筛选了血浆蛋白定量性状位点(quantitative trait loci, QTL)集,并对每个蛋白质进行了两样本孟德尔随机化(Mendelian randomization, MR)分析。主要结局指标为 WMH($n=37\ 355$)。在血管周围间隙(perivascular space, PVS; $n=38\ 598$)和腔隙性卒中(Lacunar stroke, LS; $n=232\ 596$)中进行了方向性验证。采用贝叶斯共定位分析评估位点共享性,错误发现率控制在 0.05。细胞类型注释信息来源于包含 34 种细胞类型的人类脑单核转录组图谱。因果推断仅限于 WMH;PVS 和 LS 仅用于方向一致性检验。**结果:**共有 13 种蛋白质在多重检验校正后仍与 WMH 相关。对其中 4 个具有顺式(cis)遗传证据的蛋白质进行了深入分析:ACOX1、ART4、MGP 以及以 LAMC1 为锚点的层粘连蛋白测定组。作用方向分析表明,ART4 增加 WMH 风险,而 ACOX1、MGP 和层粘连蛋白具有保护作用。贝叶斯共定位分析进一步支持 ART4 和 MGP 与 WMH 共享相同的因果变异,从而增强了二者关联的可信度;而 ACOX1 倾向于独立的遗传信号,层粘连蛋白的共享证据较弱。其余 9 个仅由反式(trans)信号驱动的蛋白质作为假设生成项予以保留。在 PVS 和 LS 中,这些蛋白质的作用方向与 WMH 一致,但效应有所减弱。**结论:**通过蛋白质组学 MR 筛选,结合共定位分析、跨终点一致性验证和细胞类型定位,将 ART4 和 MGP 确认为 WMH 的高置信度靶点。ACOX1 和层粘连蛋白虽显示保护性 MR 关联,但其定位结果不一致或较弱,需后续进行更精细的定位与功能验证。

【关键词】脑小血管病;脑白质疏松;孟德尔随机化分析;蛋白质组学;数量性状基因座

【中图分类号】R743

【文献标志码】A

Proteome-wide Mendelian randomization identifies plasma targets for white-matter hyperintensities in chronic cerebral hypoperfusion

Liu Xiaoyan¹, Yu Zhuoyuan², Yang Mei¹

(1.Neuroscience Research Center, School of Basic Medical Sciences, Chongqing Medical University;

2.Department of Urology, The First Affiliated Hospital of Chongqing Medical University)

【Abstract】Objective: Chronic cerebral hypoperfusion (CCH) is an upstream driver of cerebral small vessel disease, yet with a lack of druggable pathways and targets. With white matter hyperintensities (WMH) as the primary biomarker at the population level, this article aims to identify related plasma proteins based on human genetic evidence. **Methods:** The full set of plasma protein quantitative trait loci was identified in the European cohort, and two-sample Mendelian randomization (MR) was performed for every protein. WMH was the primary outcome measure ($n=37\ 355$). In addition, directional validation was performed in perivascular spaces (PVS) ($n=38\ 598$) and lacunar stroke (LS) ($n=232\ 596$). Bayesian colocalization was used to assess locus sharing, and the false discovery rate was controlled at 0.05. Cell-type annotation data were collected from a 34-cell-type human brain single-nucleus transcriptome atlas. Causal inference was only performed for WMH, and PVS and LS were only used for directional validation. **Results:** There were 13 proteins associated with WMH after multiple tests. In-depth analysis was performed for 4 proteins with cis-acting genetic evidence, i.e., ACOX1, ART4, MGP, and the laminin assay group anchored at LAMC1. Directionality analysis showed that ART4 increased the risk of WMH, whereas ACOX1, MGP, and laminin had a protective effect. The Bayesian colocalization analysis further supported a shared causal variant between ART4/MGP and WMH, thereby strengthening the confidence for the association between ART4 and MGP, conversely, ACOX1 appeared to have an independent genetic signal, while laminin showed weaker evidence for sharing. The remaining 9

作者简介:刘小燕, Email:2019110058@stu.cqmu.edu.cn,

研究方向:慢性脑灌注不足。

通信作者:杨 美, Email:yangmei503@cqmu.edu.cn。

基金项目:国家自然科学基金资助项目(编号:81971230);重庆市自然科学基金资助项目(编号:CSBT2023NSCQ-MSX0565)。

优先出版:https://link.cnki.net/urlid/50.1046.R.20251125.1326.002

(2025-11-26)

trans-signal-driven proteins were retained as hypothesis-generating candidates. In PVS and LS, the directional effects of these proteins aligned with those observed in WMH, albeit with attenuated magnitudes. **Conclusion**: Proteome-wide MR screening, combined with colocalization analysis, cross-endpoint checks, and cell-type mapping, identifies ART4 and MGP as high-confidence targets for WMH under the CCH framework. Although ACOX1 and laminin show protective MR associations, their inconsistent or weak colocalization results require further fine mapping and functional validation.

[Key words] cerebral small vessel disease; leukoaraiosis; Mendelian randomization analysis; proteomics; quantitative trait loci

慢性脑灌注不足 (chronic cerebral hypoperfusion, CCH) 是老年脑损伤和血管性认知障碍的关键上游病理状态,但目前缺少明确、可操作的药物靶点^[1]。脑白质高信号 (white matter hyperintensities, WMH) 是 CCH 相关白质损伤的一种可量化影像学标志,并且与卒中、痴呆及死亡风险的升高密切相关,这凸显了其重要的临床与流行病学意义^[2]。观察性研究本身难以阐明低灌注、白质损伤以及多种病因并存现象之间的因果联系;因此,需要借助更强大的工具来推断暴露因素与临床结局之间的因果关系^[3]。

孟德尔随机化 (Mendelian randomization, MR) 方法利用与可改变暴露因素强相关的遗传变异作为“自然随机化”的工具变量,能够有效减少混杂倚倚和反向因果的干扰,已被广泛应用于识别潜在的干预通路和药物靶点^[3-4]。在蛋白质组学研究中,优先选择顺式蛋白数量性状位点 (protein quantitative trait loci, pQTL) 作为工具变量,能提高生物学解释的清晰度并减少远端多效性影响,该方法在各种疾病的靶点优先级排序中已展现出其价值^[5-6]。与此同时,贝叶斯共定位分析能够在基因座层面判定暴露因素与结局是否更可能由同一个因果变异驱动,这有助于避免对“显著但位于不同峰值”的遗传信号产生误读,从而为 MR 结果提供正交证据支持^[6]。

与脑脊液 (cerebrospinal fluid, CSF) 相比,血浆蛋白检测具有侵入性小、可扩展性强、成本更低等优势,这使得研究人员能够获得更大的 pQTL 样本量以及统计效力更强的工具变量;此外,血浆对血管-内皮界面来源的信号尤为敏感,这与 CCH 的血管病理本质相契合^[7-8]。因此,选择血浆蛋白作为暴露因素,一方面是因为大规模人群的血浆 pQTL 资源已趋成熟且易于扩展——这为开展两样本 MR 分析提供了强有力的工具;另一方面则是因为血浆能灵敏地反映与 CCH 病理机制密切相关的血管-内皮信号,这对于后续的机制推断和临床转化潜力至关重要^[9-10]。值得注意的是,近期已有研究将脑脊液/血浆 pQTL 与 WMH、血管周围间隙 (perivascular

spaces, PVS) 等磁共振成像 (magnetic resonance imaging, MRI) 白质指标相结合进行 MR 分析,为生物体液中的蛋白质与白质影像学改变之间的因果关联提供了初步证据,也间接支持了本研究方法的可行性^[11]。

基于上述考量和方法学优势,本研究将 WMH 视为 CCH 所致白质损伤的主要特征,开展了双样本蛋白质组 MR 分析,并利用贝叶斯共定位在基因座层面强化了因果关系的解读。为了评估本研究结论在 CCH 相关不同终点间的一致性,课题组进一步利用大样本资源进行了跨终点的一致性验证^[12-13]。在细胞定位方面,本课题组整合了人类蛋白质图谱的人脑单核转录组图谱[涵盖 11 个脑区、34 种细胞类型,提供标准化的每百万转录本数 (normalized Transcripts Per Million, nTPM) 矩阵],将遗传/因果信号投射到内皮细胞、周细胞、血管平滑肌细胞和成纤维样细胞等关键谱系上,以辅助生物学解读,旨在系统性地发现并分层评估与白质损伤相关的候选蛋白质。

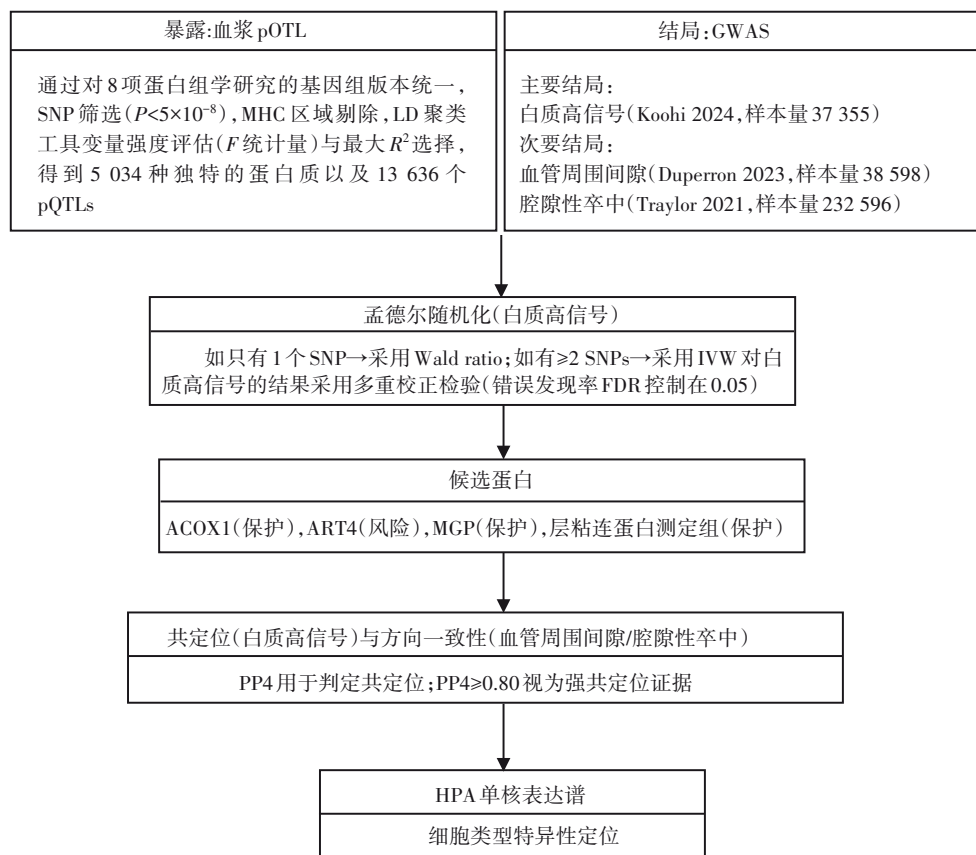
1 资料与方法

1.1 研究设计与终点指标

本研究的工作流程如图 1 所示。本课题组筛选了来自 8 项大型欧洲蛋白质组学研究资源 (Fenland, INTERVAL, UKB-PPP) 的血浆 pQTL 集,并评估了其在全基因组关联研究 (Genome-Wide Association Study, GWAS) 数据集的关系:WMH (主要指标)^[14]、PVS (方向性验证)^[12]及腔隙性卒中 (lacunar stroke, LS) (方向性验证)^[13]。随后,依据以下 4 个证据层面对蛋白靶点进行分级评估:①蛋白质组孟德尔随机化分析;②基因座水平的贝叶斯共定位分析;③跨终点一致性验证;④基于人脑单核转录组图谱的细胞类型定位。

1.2 蛋白质组学工具变量

如前所述^[15],本研究整合了 8 项针对欧洲血统人群开展的蛋白质组学研究:Fenland (10 708 名参与者,4 775 种蛋白质)^[16]、Iceland (35 559 名参与者,4 719 种蛋白质)^[8,17]、INTERVAL (3 301 名参与者,2 995 种蛋白质)^[7]、UKB-PPP (34 557 名参与者,2 922 种蛋白质)^[18]、UKB-PPP 第一阶段



注:本研究全蛋白质组分析流程概览:首先,整合了来自 8 项欧洲研究的血浆蛋白 pQTL 数据,统一至同一基因组版本(build),随后通过全基因组显著性筛选(排除 MHC 区域并进行 LD 聚类),最终获得 5 034 种蛋白质和 13 636 个 pQTLs。以 WMH 为主要结局指标进行 MR 分析,继而针对 WMH 结果进行贝叶斯共定位分析,并在 PVS 和 LS 中进行了方向一致性检验。人脑单核转录组表达数据为结果提供了生物学背景支持。最终,4 种蛋白质被优先筛选出:ACOX1(保护性)、ART4(风险性)、MGP(保护性)和层粘连蛋白复合物(保护性)。分析中,单工具变量采用 Wald 比率法,多工具变量则采用逆方差加权(IVW)法,并为控制多重检验假阳性,对所有结果进行了 FDR 校正

图 1 慢性脑灌注不足框架下的研究流程图

(35 571 名参与者,1 463 种蛋白质)^[19],KORA F4(1 000 名参与者,1 124 种蛋白质)^[20],SCALLOP(30 931 名参与者,90 种蛋白质)^[21],以及 FHS(6 861 名参与者,71 种蛋白质)^[22],所纳入蛋白质组学研究的详细信息见表 1。数据标准化与跨研究合并遵循既往工作流程^[15];最终形成的每种蛋白质工具集部分展示在表 2。

表 1 蛋白质组学研究汇总

研究	平台	人群	蛋白数量	pQTL 数量	样本量
Fenland	SomaLogic	欧洲	4 775	8 328	10 708
Iceland	SomaLogic	欧洲	4 719	18 084	35 559
INTERVAL	SomaLogic	欧洲	2 995	1 927	3 301
UKB-PPP	Olink	欧洲	2 922	14 287	34 557
UKB-PPP phase I	Olink	欧洲	1 463	10 248	35 571
KORA F4	SomaLogic	欧洲	1 124	539	1 000
SCALLOP	Olink	欧洲	90	451	30 931
FHS	xMAP	欧洲	71	16 602	6 861

关于层粘连蛋白测量的说明:在 INTERVAL 研究(使用 SomaScan 技术平台)中,部分适配体可结合层粘连蛋白不同亚基上的共同表位;该平台采用蛋白质组注释,因此将该检测标识为“LAMB1/LAMC1/LAMA1”。鉴于此,本课题组将该检测作为 1 个整体指标进行分析,但仅提取了 LAMC1 基因座±500 kb 范围内的顺式(cis)变异作为遗传工具变量,原因是该检测所有显著的顺式 pQTL 均锚定在 LAMC1 基因座(在 LAMB1 或 LAMA1 基因座±500 kb 范围内未发现独立的、达到显著性阈值的顺式工具变量)。因此,本研究的 MR 估计值在统计学上等效于对 LAMC1(γ1 链)的遗传代理效应,而非跨基因的聚合效应。

1.3 结局数据处理与基因组版本对齐

WMH、PVS 和 LS 的汇总统计数据均对齐至同一基因组版本(hg19/GRCh37),并经标准化处理,确保效应量在暴露变量与结局变量间均指向相同效应等位基因^[23]。

1.4 MR 主要分析

若在聚类(clumping)后仅剩 1 个工具变量,本研究采用

表 2 部分蛋白以及对应的 SNPs

蛋白	研究	样本量	SNP	染色体	位置	效应等位基因	非效应等位基因	效应等位基因频率	Beta	标准误	P	R ²	R ² 总和	F	顺式/反式
ACOX1	UKB-PPP phase I	35 571	rs10852766	17	73951864	C	T	0.658	0.07	0.01	2.24E-20	0.003	0.002 5	87.8	cis
ARL1	INTERVAL	3 301	rs429358	19	45411941	C	T	0.153	-0.30	0.03	1.10E-19	0.023	0.023 3	78.8	trans
ART4	Fenland	10 708	rs10772808	12	14990587	A	G	0.400	0.90	0.01	1.00E-200	0.386	0.386 2	6 736.5	cis
CTF1	INTERVAL	3 301	rs429358	19	45411941	C	T	0.153	0.80	0.03	3.50E-148	0.166	0.165 9	656.1	trans
DCK	INTERVAL	3 301	rs429358	19	45411941	C	T	0.153	0.31	0.03	7.40E-21	0.025	0.024 9	84.3	trans
LAMB1_LAM C1_LAMA1	INTERVAL	3 301	rs4129858	1	183004334	G	A	0.431	-0.29	0.02	7.10E-34	0.041	0.083 4	141.9	cis
LAMB1_LAM C1_LAMA1	INTERVAL	3 301	rs62199218	20	60932425	A	G	0.039	0.75	0.06	1.10E-31	0.042	0.083 4	145.2	trans
MGP	Fenland	10 708	rs2430687	12	15047396	T	C	0.600	0.41	0.01	1.00E-200	0.079	0.106 2	919.9	cis
MGP	Fenland	10 708	rs704	17	26694861	A	G	0.470	-0.15	0.01	7.62E-28	0.011	0.106 2	118.1	trans
MGP	Fenland	10 708	rs77542162	17	67081278	A	G	0.980	-0.64	0.05	4.24E-42	0.016	0.106 2	175.8	trans
NEFL	Fenland	10 708	rs429358	19	45411941	T	C	0.850	1.25	0.02	1.00E-200	0.399	0.3991	7 109.9	trans
PDZK11P1	INTERVAL	3 301	rs429358	19	45411941	C	T	0.153	0.35	0.03	3.10E-26	0.032	0.031 7	108.2	trans
RGS7	Fenland	10 708	rs429358	19	45411941	T	C	0.850	-0.16	0.02	1.54E-16	0.006	0.006 4	68.6	trans
SARS	Iceland	35 559	rs429358	19	45411941	C	T	0.174	-0.08	0.01	4.47E-13	0.002	0.001 8	65.6	trans
SNRPF	KORA F4	1 000	rs4420638	19	45422946	G	A	0.170	-0.44	0.06	3.00E-14	0.053	0.053 4	56.3	trans
TBCA	INTERVAL	3 301	rs429358	19	45411941	C	T	0.153	-1.34	0.02	1.00E-200	0.465	0.465 4	2 871.8	trans

Wald 比率法并结合 Delta 法计算标准误(standard error, SE)。若有 2 个或以上独立工具变量可用,则采用逆方差加权(inverse-variance weighted, IVW)固定效应模型来合并各 SNP 的比率估计值。结果以比值比(OR)、95%CI 及双侧 P 值的形式报告。必要时采用 Cochran's Q 统计量描述异质性,并以加权中位数法或 MR-Egger 法作为敏感性分析进行参考,但不将其用于主要结论的推断^[24-26]。错误发现率(false discovery rate, FDR)控制在 0.05^[27]。

1.5 贝叶斯共定位分析

本课题组采用贝叶斯共定位法检验特定基因座上的 MR 暴露与 WMH 关联是否由同一因果变异驱动。针对每种蛋白质,将分析窗口锚定于暴露侧哨兵变异位点,并在 hg19 基因组构建中 ± 1 000 千碱基的对称窗口内比较暴露与结局效应量及标准误。仅当变异在等位基因水平标准化及基础质量检查后同时存在于两数据集时才纳入分析。先验概率设定为:暴露关联 0.000 1,结局关联 0.000 1,共同关联 0.000 01。汇总了 5 项标准假设的后验支持度,重点关注共享因果变异概率(PP4)与不同因果变异概率(PP3)。将 PP4>0.80 视为强支持共享变异,0.50~0.80 视为提示性支持,<0.50 视为支持不足。PP3 接近 1 则表明存在独立变异^[28]。

1.6 跨终点一致性

对于 PVS,本研究采用了 2023 年发表于《自然-医学》(Nature Medicine)的多队列 GWAS/WES/WGS 资源(样本量高达 38 598 名参与者),其通路富集分析结果突出了细胞外基质(extracellular matrix, ECM)/发育过程——这与 WMH 的

生物学背景非常契合,十分适用于方向性重复验证^[12]。对于 LS,采用了 2021 年发表于《柳叶刀-神经病学》(Lancet Neurology)的国际病例对照 GWAS(约 7 338 例病例 vs. 254 798 例对照),并据此生成了相应的荟萃分析汇总统计数据^[13]。针对这 2 个外部终点,采用了与 WMH 分析相同的工具变量选择、数据协调和估计流程,并主要在作用方向层面上与 WMH 的结果进行比对解读^[23]。

1.7 单核脑图谱与细胞类型定位

为将遗传/因果线索投射到特定细胞类型,本研究采用了 HPA 人脑单核转录组图谱。其下载站点提供了 34 种细胞聚类类型 × 基因的 nTPM 矩阵,覆盖 11 个脑区,并附有标准化的字段说明和文档。采用统一尺度的点图(dot plot)来展示 4 种候选蛋白(ACOX1, ART4, MGP 及层粘连蛋白复合物)在这 34 种细胞类型中的绝对表达量和检测比例,避免了仅在部分细胞亚集中进行排名展示可能带来的解释偏倚;该图谱整合了来自公共单核 RNA 测序(snRNA-seq)数据集的超过 250 万个脑细胞核。

1.8 质量控制与稳健性

连锁不平衡(linkage disequilibrium, LD)参考数据和等位基因频率均来自 1000 Genomes Project(第三阶段,欧洲人群),用于在回文 SNP 位点进行聚类和链推断,确保了跨资源数据的一致性^[29]。所有 MR 和共定位分析均使用相同的基因组版本(hg19)并采用双侧检验。通过数量级变化调整聚类参数和共定位先验概率来进行稳定性检验,报告结果以主要分析为准,敏感性分析结果作为支持性证据^[23]。

1.9 伦理声明

本研究仅使用公开的去识别化 GWAS 汇总统计数据及在线资源;原始研究均获得适当伦理审批及知情同意程序。未涉及新受试者招募或干预措施。

2 结果

2.1 筛查结果与后续分析范围

在以 WMH 为主要终点的全蛋白质组 MR 筛查中,经多重检验校正 (FDR<0.05) 后,共有 13 种蛋白质与 WMH 存在关联,它们分别是: ACOX1、ARL1、CTF1、DCK、NEFL、PDZK1IP1、RGS7、SARS、TBCA、SNRPF、ART4、MGP 及 LAMB1_LAMC1_LAMA1 (层粘连蛋白测定组)。其中 9 个信号仅由反式 pQTL 驱动 (ARL1、CTF1、DCK、NEFL、PDZK1IP1、RGS7、SARS、TBCA、SNRPF), 鉴于其生物学特异性有限, 本文未对其进行深入讨论, 仅作为探索性发现予以报告。因此, 本课题组随后的分析重点聚焦于另外 4 个拥有顺式 (cis) pQTL 证据的蛋白质: ACOX1、ART4、MGP 和层粘连蛋白复合物, 并进一步评估了它们的作用方向一致性和细胞定位。

2.2 4 个具有顺式证据的蛋白

经错误发现校正后, 4 种蛋白均对白质高信号呈统计学显著关联。ACOX1 (单一变异位点, 采用 Wald 比值法评估) 显示保护性关联, $OR=0.433$ (95%CI=0.355~0.527)。ART4 (同样采用单变异位点 Wald 分析) 显示轻微风险关联 $OR=1.037$ (1.022~1.053)。采用多重工具与逆方差加权后, MGP 和层粘连蛋白测定组均显示中度保护作用: MGP $OR=0.935$ (0.906~0.965); 层粘连蛋白 $OR=0.934$ (0.905~0.965)。该模式具有一致性: ACOX1 具有最大保护效力; ART4 具有微小但稳定的风险效应; MGP 与层粘连蛋白在多变量聚合分析中显示出相似保护作用。其余 9 项信号由转录因子驱动, 虽

通过多重检验阈值, 但效应较弱。FDR<0.05 的部分孟德尔随机化分析结果见表 3。

2.3 跨终点一致性

为避免 WMH 特异性单表型假阳性, 本课题组在另外 2 个脑小血管病变终点中评估了这些发现的外部一致性。PVS 采用欧洲人群的影像 GWAS 汇总统计数据, 其表型基于 3D T2/T1 结构 MRI 上半自动或自动分割的 PVS 体积/评分得出; 队列主要为基于人群的研究 (以英国生物银行 MRI 子队列为主), 样本量达数万人; 基因型均经过标准质量控制和参照面板的基因型插补。LS 分析则采用了国际病例对照 GWAS 汇总统计数据, 病例经临床或影像学确认 (≤ 10 000 例), 对照组规模庞大 (数十万例); 各项研究均控制了人群结构差异, 并采用固定效应/随机效应荟萃分析合并结果。对于这两个终点, 我们均采用了与 WMH 主要分析相同的数据协调和工具变量选择流程, 并报告了双侧 P 值, 同时注明了名义显著性和 FDR 状态 (表 4)。在此分析框架下, 上述 4 个顺式优先候选蛋白在 PVS 和 LS 中均表现出与 WMH 一致的效应方向: ACOX1、MGP 和层粘连蛋白复合物依然显示保护作用, 而 ART4 则持续显示风险增加作用。其绝对效应量均小于在 WMH 中的效应, 这与在不同表型间进行效应迁移时常见的衰减现象一致; 其统计学显著性则各不相同, 部分达到名义显著 ($P<0.05$), 但多数在经过多重检验校正后变为临界或不再显著, 这与 PVS 和 LS 相对较小的样本量及较大的测量误差 (尤其是 PVS) 相符。方向的一致性——特别是 ACOX1 和 MGP 的保护性信号, 以及 ART4 的微弱风险信号——在 2 个小血管终点中都得到了重现, 这为结论的可移植性提供了证据; 层粘连蛋白复合物在 PVS/LS 中的保护方向也与其基底膜/ECM 的生物学功能相吻合。考虑到统计效能和测量差异, 将跨终点结果视为支持性证据, 并纳入本研究的证据分级体系 (图 2 和表 4)。

表 3 孟德尔随机化结果—WMH (FDR<0.05)

蛋白	方法	SNP 数量	B	SE	P 值	CI 下限	CI 上限	OR	OR (95%CI)		FDR
									上限	下限	
ACOX1	Wald ratio	1	-0.837	0.101	8.72E-17	-1.035	-0.640	0.433	0.355	0.527	4.09E-13
ARL1	Wald ratio	1	-0.158	0.031	3.39E-07	-0.219	-0.097	0.854	0.803	0.907	1.77E-04
CTF1	Wald ratio	1	0.059	0.012	3.39E-07	0.037	0.082	1.061	1.037	1.086	1.77E-04
DCK	Wald ratio	1	0.153	0.030	3.39E-07	0.094	0.212	1.166	1.099	1.236	1.77E-04
NEFL	Wald ratio	1	-0.038	0.007	3.39E-07	-0.053	-0.023	0.963	0.949	0.977	1.77E-04
PDZK1IP1	Wald ratio	1	0.136	0.027	3.39E-07	0.084	0.188	1.145	1.087	1.207	1.77E-04
RGS7	Wald ratio	1	0.297	0.058	3.39E-07	0.183	0.411	1.345	1.200	1.508	1.77E-04
SARS	Wald ratio	1	-0.594	0.116	3.39E-07	-0.822	-0.365	0.552	0.440	0.694	1.77E-04
TBCA	Wald ratio	1	-0.035	0.007	3.39E-07	-0.049	-0.022	0.965	0.952	0.978	1.77E-04
SNRPF	Wald ratio	1	-0.095	0.019	1.04E-06	-0.133	-0.057	0.909	0.875	0.945	4.88E-04
ART4	Wald ratio	1	0.036	0.008	1.57E-06	0.022	0.051	1.037	1.022	1.053	6.69E-04
MGP	Inverse variance weighted	3	-0.068	0.016	2.72E-05	-0.099	-0.036	0.935	0.906	0.965	1.06E-02
LAMB1_LAMC1_LAMA1	Inverse variance weighted	2	-0.068	0.017	4.34E-05	-0.100	-0.035	0.934	0.905	0.965	1.57E-02

2.4 共定位分析

本课题组以暴露侧的哨兵变异为中心,在 ± 1 000 kb 的对称窗口内进行了贝叶斯共定位分析,从而对 4 个拥有顺式支持的蛋白质进行了区分。ART4 和 MGP 显示出强有力的证据,表明其 MR 暴露(蛋白质水平)与 WMH 的关联源于同一个因果变异;两者共享因果变异的后验概率(PP4)均为 0.995,而支持不同变异(PP3)的证据则微乎其微(图 3)。相比之下,ACOX1 虽在 MR 中显示保护性效应,但其基因座更倾向于支持不同的遗传信号:其不同变异的概率(PP3)几乎为 1,而共享变异的概率(PP4)可忽略不计。对于以 LAMC1 为锚点的层粘连蛋白检测组,其支持共享变异的证据较

弱,不同变异的解释仍然较为合理。完整的数值结果报告于表 5。

2.5 单核(single-nucleus)细胞类型定位(HPA)

本课题组将遗传/因果线索投射至 HPA 34 类单核(single-nucleus)细胞图谱。为避免相对缩放伪影,正文仅展示绝对表达量点图(图 4),并在统一尺度下解读全部 34 类细胞;每个基因的前 5 个高表达细胞簇列于补充表 6。总体上,MGP 在内皮、血管平滑肌、周细胞及成纤维样谱系中较高;层粘连蛋白测定组在成纤维与平滑肌等基质/基底膜相关细胞中较高;ACOX1 在内皮富集,并在部分胶质亚型中等;ART4 总体较低,但在内皮达到峰值(图 4)。

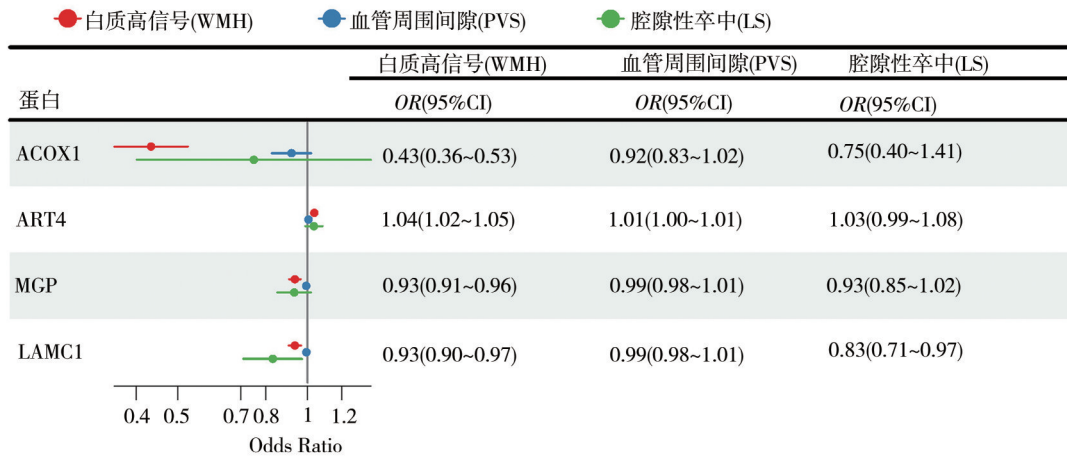


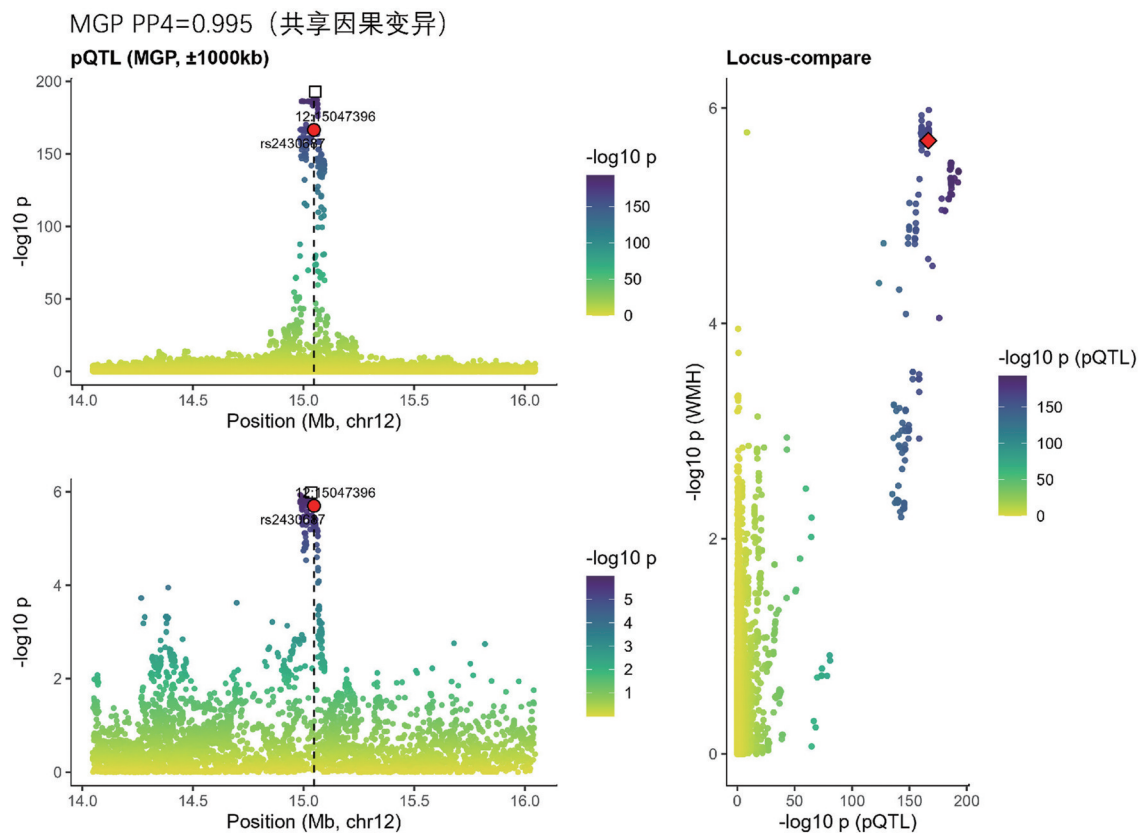
图 2 4 个顺式支持蛋白在不同终点中的效应方向

表 4 方向一致性验证

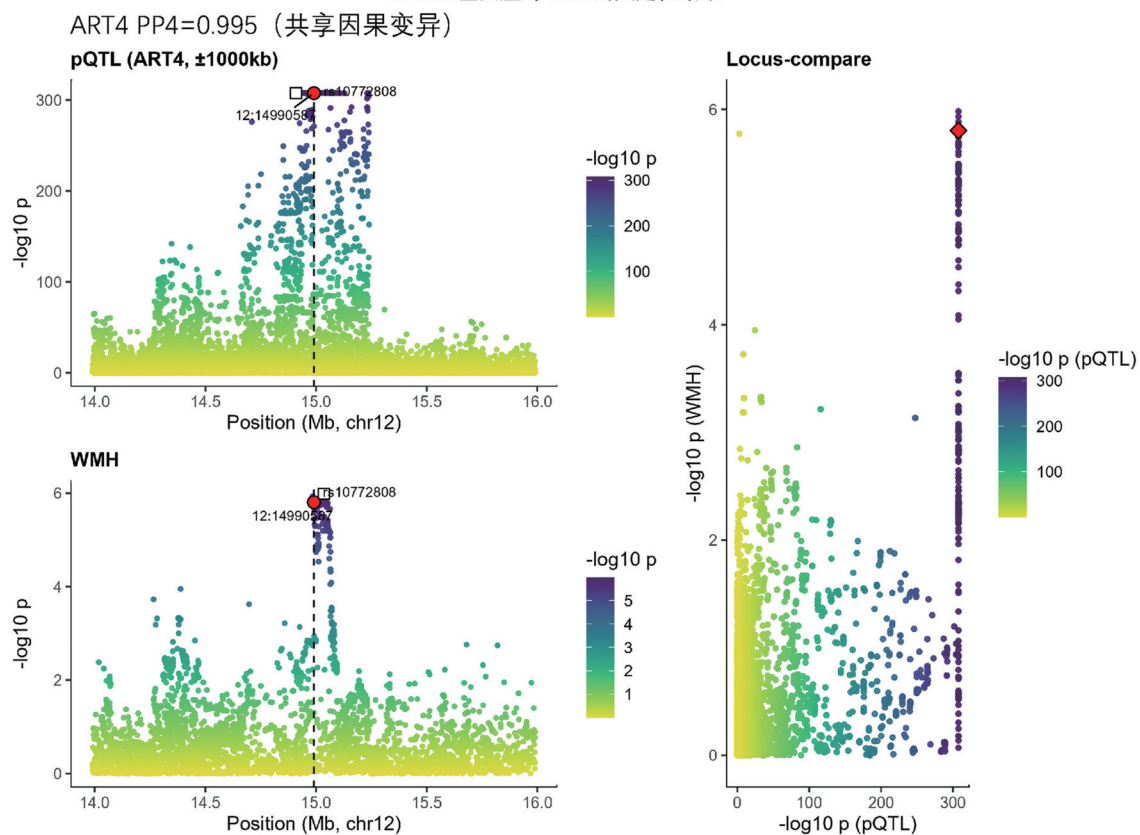
蛋白	方法	SNP 数量	beta	标准误	P	OR	95%CI 上限	95%CI 下限	特征
LAMB1_LAMC1_LAMA1	Inverse variance weighted	2	-0.068	0.017	0.000	0.934	0.905	0.965	WMH
LAMB1_LAMC1_LAMA1	Inverse variance weighted	2	-0.006	0.009	0.511	0.994	0.976	1.012	PVS
LAMB1_LAMC1_LAMA1	Inverse variance weighted	2	-0.186	0.080	0.021	0.830	0.709	0.972	Lacunar
MGP	Inverse variance weighted	3	-0.068	0.016	0.000	0.935	0.906	0.965	WMH
MGP	Inverse variance weighted	3	-0.007	0.008	0.377	0.993	0.978	1.008	PVS
MGP	Inverse variance weighted	3	-0.071	0.046	0.125	0.931	0.850	1.020	Lacunar
ART4	Wald ratio	1	0.036	0.008	0.000	1.037	1.022	1.053	WMH
ART4	Wald ratio	1	0.005	0.004	0.195	1.005	0.997	1.013	PVS
ART4	Wald ratio	1	0.034	0.024	0.153	1.035	0.987	1.084	Lacunar
ACOX1	Wald ratio	1	-0.837	0.101	0.000	0.433	0.355	0.527	WMH
ACOX1	Wald ratio	1	-0.086	0.054	0.114	0.918	0.825	1.021	PVS
ACOX1	Wald ratio	1	-0.287	0.320	0.370	0.750	0.401	1.405	Lacunar

表 5 共定位后验概率

蛋白	SNP	PP.H0.abf	PP.H1.abf	PP.H2.abf	PP.H3.abf	PP.H4.abf
ACOX1	rs10852766	3.37E-48	1.63E-32	2.06E-16	1	1.25E-11
ART4	rs10772808	0	0.005	0	2.82E-06	0.995
LAMC1	rs4129858	1.05E-27	0.593	4.17E-28	0.236	0.171
MGP	rs2430687	2.89E-181	0.005	6.18E-182	2.69E-06	0.995



A. MGP基因座与WMH的共定位结果



B. ART4基因座与WMH的共定位结果

注: 区域共定位分析表明在这 2 个基因座上, 蛋白质与 WMH 的关联均源于同一因果变异

图 3 ART4 和 MGP 基因座与白质高信号 (WMH) 的共定位分析

表 6 基因表达最高的 5 种细胞类型

基因名称	详细细胞类型	细胞类型	表达量(ntpm)
ACOX1	内皮细胞	内皮细胞	70.3
ACOX1	脉络丛上皮细胞	胶质细胞	56.6
ACOX1	成纤维细胞	间充质细胞	54.3
ACOX1	伯格曼胶质细胞	胶质细胞	53.3
ACOX1	星形胶质细胞	胶质细胞	38.9
ART4	内皮细胞	内皮细胞	4.0
ART4	脉络丛上皮细胞	胶质细胞	3.9
ART4	海马 CA4 区神经元	神经元细胞	3.4
ART4	下菱形来源神经元 lower rhombic lip	神经元细胞	2.8
ART4	乳头体神经元	神经元细胞	2.1
LAMB1_LAMC1_LAMA1	成纤维细胞	间充质细胞	327.6
LAMB1_LAMC1_LAMA1	脉络丛上皮细胞	胶质细胞	135.0
LAMB1_LAMC1_LAMA1	星形胶质细胞	胶质细胞	124.4
LAMB1_LAMC1_LAMA1	血管相关平滑肌细胞	肌肉细胞	110.2
LAMB1_LAMC1_LAMA1	中脑来源抑制性神经元	神经元细胞	96.4
MGP	成纤维细胞	间充质细胞	249.1
MGP	内皮细胞	内皮细胞	241.5
MGP	血管相关平滑肌细胞	肌肉细胞	238.8
MGP	周细胞	肌肉细胞	105.7
MGP	室管膜细胞	胶质细胞	7.1

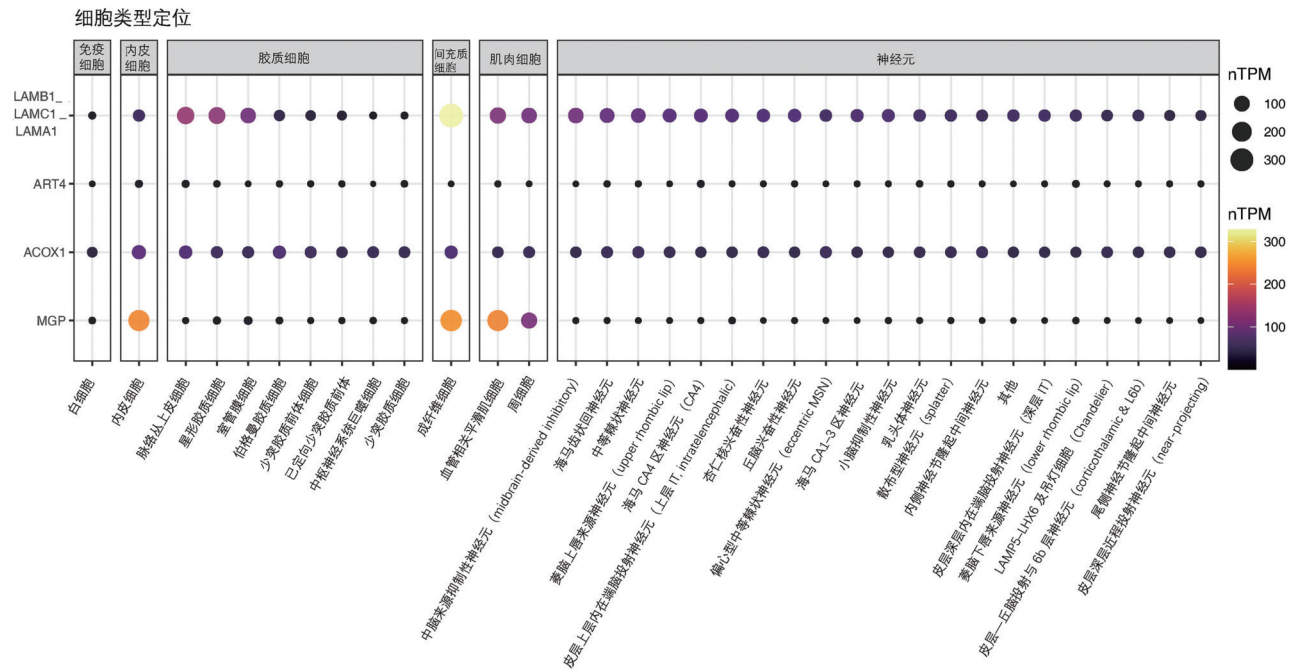


图 4 人脑单核图谱中的细胞类型定位

2.6 证据整合与分级

通过结合 MR 估计值及其 FDR 支持、跨终点一致性、基因座共定位及单核(single-nucleus)细胞类型定位,本课题组对靶点进行分级。ART4 和 MGP 为高可信度靶点:它们通过 MR,且存在强有力的共享因果变异证据,在血管周围间隙和

腔隙性卒中中呈现方向一致的信号,并与合理的细胞类型表达相吻合。ACOX1 与层粘连蛋白测定组为中等可信度:MR 信号与细胞水平模式具有生物学一致性,但共定位较弱或存在不一致,因此优先开展精细定位、强化顺式调控工具及功能验证工作。

3 讨 论

本研究将 CCH 确立为脑小血管损伤的上游病理生理背景,并将 WMH 作为其主要特征。在此框架下,蛋白质 MR 筛查在多重检验校正后,共鉴定出 13 种与 WMH 存在关联的蛋白质;其中 4 种拥有顺式(cis)工具变量的蛋白质(ACOX1、ART4、MGP 及以 LAMC1 为锚点的层粘连蛋白测定组)被纳入后续深入分析。需要明确的是,本研究的设计旨在推断其对 WMH 的因果效应。对 PVS 和 LS 的分析,主要用于检验这些因果效应方向在小血管病变谱系中的一致性,而非宣称其在这些终点上也存在因果效应。这种“估计目标”的区分在 CCH 背景下至关重要,因为各种影像学表型实质上是共享的微血管应激因素的概率性表现,而非单一上游因素的充分必要条件式的结果^[12-13]。因此,在 CCH 背景下解读 MR 结果,应着重关注其生物学一致性和效应方向的普适性,而非跨表型的因果断言。

本研究筛选出的 4 个核心蛋白,为理解 CCH 背景下 WMH 的机制提供了具体切入点。ACOX1 是过氧化物酶体脂肪酸 β -氧化通路的起始与限速酶,参与髓鞘相关脂质稳态与细胞氧化还原平衡^[30-31]。MR 结果显示较高的血浆 ACOX1 与较低的 WMH 风险相关,提示系统性脂质代谢与过氧化体稳态可能对白质具有保护作用。ART4(CD297)为红细胞膜上的 GPI 锚定蛋白,人源 ART4 的细胞外单 ADP-核糖基转移酶活性尚未确证。本研究观察到血浆 ART4 与 WMH 风险的因果关联,可能作为血液学或血管相关状态的间接指示,具体通路仍需实验验证。MGP 为维生素 K 依赖的抗钙化分子,可拮抗 BMP2/4 信号以维持血管弹性与基质稳态^[32]。本研究发现其与 WMH 存在强共定位证据,支持“血管壁重塑—基底膜和细胞外基质稳态失衡”在 WMH 形成中的关键作用。LAMC1 编码层粘连蛋白 γ 1 链,层粘连蛋白是微血管基底膜的核心组成,影响细胞黏附、迁移与存活^[33]。尽管本研究中共定位证据较弱,但遗传层面呈现的负相关提示系统性层粘连蛋

白代谢或与微血管基底膜完整性相关,这与 WMH 所见的血脑屏障受损过程相一致。

共定位分析与细胞类型定位的结果对候选靶点进行了有效分层,并将其映射至慢性低灌注背景下预期相关的生物学过程。ART4-WMH 和 MGP-WMH 关联均显示出强有力的共享变异证据(高 PPH4),这与 MR 结果相互印证,形成了强有力的证据链。ART4 定位于内皮细胞富集的区域,可能捕捉了核苷酸信号传导与内皮细胞的激活状态;而 MGP 则定位于基底膜和细胞外基质谱系,提示其可能参与了血管壁的重塑过程。与之相反,ACOX1 虽显示出保护性的 MR 估计值,但其基因座更倾向于支持不同于 WMH 的独立因果变异(高 PPH3),这与该区域的复杂性(如存在相邻或多个信号)相符;层粘连蛋白测定组则显示出保护性 MR 效应但共定位证据较弱^[28]。结合人脑单核转录组数据,这些发现共同指向了与 CCH 病理密切相关的 2 条核心轴线:内皮细胞代谢与结构耦合功能紊乱,以及基底膜与细胞外基质重塑障碍。前者与低灌注诱导的代谢应激、剪切力降低、氧化还原失衡及对少突胶质细胞的营养支持受损等机制相吻合;后者则与血管周围单元的渐进性改变相一致,这些变化会损害血脑屏障功能并增加白质易损性,这已在临床和实验性低灌注模型中得到观察^[34-36]。遗传因果性、基因座共享性与细胞定位三者之间的高度吻合,极大地增强了我们的信心,表明这些通路并非由混杂因素或反向因果关系所导致的假象。

基于上述分级结果,本研究衍生出潜在的临床与转化医学意义。ART4 和 MGP 作为高置信度候选靶点脱颖而出,其可能作为靶向干预策略的分子,也可能作为血浆药效动力学生物标志物或伴随诊断标志物,用于监测 WMH 的进展或治疗反应。在 CCH 背景下,WMH 进展仍是寻求小血管神经保护作用的临床试验最合适的主要影像学终点,因为它能整合随时间推移的弥漫性白质损伤,并在中等样本量的研究中提供足够的统计效能。PVS 和 LS 则可作为次要或机制探索性终点,这与其虽方向一致

但效应量更小且在不同队列中存在较大测量误差的特性相符^[12-13]。此外,这 2 条轴线也为老药新用或新分子实体开发提供了可检验的假设。在内皮层面,调控过氧化物酶体脂质氧化和核苷酸介导的信号传导,或许能改善内皮细胞的代谢适应性与屏障完整性。在血管壁层面,增强细胞外基质的组装与稳态维持,则可能稳定基底膜、支持周细胞-内皮细胞对话,并最终改善微循环储备功能。这些假设可在分阶段的研究体系中得到验证,例如将富含生物标志物的临床试验与定量 MRI 终点和系列血浆蛋白检测相结合。

当然,本研究的一些局限性和未来研究方向也需要谨慎考量。MR 估计的是与基因终生相关的蛋白质水平变化,其效应大小和时效性可能无法完全模拟药理干预的效果;剂量反应关系和时间窗口需要通过实验来明确。尽管已优先采用顺式 pQTL 并进行共定位分析,但残余多效性、连锁不平衡结构以及单一因果变异的假设仍可能给估计值带来偏倚^[28]。人群祖先背景、蛋白质组学平台和影像处理流程的差异,可能是导致效应在 PVS 和 LS 中出现衰减的原因之一,这也支持了本课题组将这些终点仅用于方向性验证而非因果推断的策略^[12-13]。未来的研究应将 MR 与共定位分析延伸至以 PVS 和 LS 为主要结局的分析中;应用多变量 MR 来控制诸如血压、代谢特征和炎症等共有的血管风险因素;采用多信号共定位分析进行精细定位,以解析 ACOX1 和层粘连蛋白检测组等复杂基因座;并前瞻性地以 WMH 进展为主要结局的观察性队列和早期干预研究中,验证 ART4 和 MGP 作为血浆生物标志物的潜力。总而言之,在 CCH 框架下,汇聚的遗传学与细胞学证据共同揭示了内皮-代谢通路和基底膜-细胞外基质通路在脑小血管性白质损伤中的关键作用,其中 ART4 和 MGP 已成为最具转化前景的领衔候选分子。

利益冲突 所有作者声明不存在利益冲突

作者贡献声明 刘小燕:研究方案设计、数据整理与分析、论文撰写;余卓远:协助完成图表、论文修改;杨美:学术指导、论文修改

参 考 文 献

- [1] Rajeev V, Chai YL, Poh L, et al. Chronic cerebral hypoperfusion: a critical feature in unravelling the etiology of vascular cognitive impairment[J]. *Acta Neuropathol Commun*, 2023, 11(1): 93.
- [2] Guo WJ, Shi J. White matter hyperintensities volume and cognition: A meta-analysis[J]. *Front Aging Neurosci*, 2022, 14: 949763.
- [3] Davey Smith G, Hemani G. Mendelian randomization: genetic anchors for causal inference in epidemiological studies[J]. *Hum Mol Genet*, 2014, 23(R1): R89-R98.
- [4] Daghlas I, Gill D. Mendelian randomization as a tool to inform drug development using human genetics[J]. *Camb Prism Precis Med*, 2023, 1: e16.
- [5] Schmidt AF, Finan C, Gordillo-Marañón M, et al. Genetic drug target validation using Mendelian randomisation[J]. *Nat Commun*, 2020, 11(1): 3255.
- [6] Zhao HL, Rasheed H, Nøst TH, et al. Proteome-wide Mendelian randomization in global biobank meta-analysis reveals multi-ancestry drug targets for common diseases[J]. *Cell Genom*, 2022, 2(11): 100195.
- [7] Sun BB, Maranville JC, Peters JE, et al. Genomic atlas of the human plasma proteome[J]. *Nature*, 2018, 558(7708): 73-79.
- [8] Eldjam GH, Ferkingstad E, Lund SH, et al. Large-scale plasma proteomics comparisons through genetics and disease associations[J]. *Nature*, 2023, 622(7982): 348-358.
- [9] Aschenbrenner AJ, Li Y, Henson RL, et al. Comparison of plasma and CSF biomarkers in predicting cognitive decline[J]. *Ann Clin Transl Neurol*, 2022, 9(11): 1739-1751.
- [10] Dammer EB, Ping LY, Duong DM, et al. Multi-platform proteomic analysis of Alzheimer's disease cerebrospinal fluid and plasma reveals network biomarkers associated with proteostasis and the matrisome[J]. *Alzheimers Res Ther*, 2022, 14(1): 174.
- [11] Dobbie S, Caro I, Western D, et al. Proteogenomics in cerebrospinal fluid and plasma reveals new biological fingerprint of cerebral small vessel disease[J]. *Res Sq*, 2024; rs. 3. rs-rs. 4535534.
- [12] Duperron MG, Knol MJ, Le Grand Q, et al. Genomics of perivascular space burden unravels early mechanisms of cerebral small vessel disease[J]. *Nat Med*, 2023, 29(4): 950-962.
- [13] Traylor M, Persyn E, Tomppa L, et al. Genetic basis of lacunar stroke: a pooled analysis of individual patient data and genome-wide association studies[J]. *Lancet Neurol*, 2021, 20(5): 351-361.
- [14] Koohi F, Harshfield EL, Shatunov A, et al. Does thrombosis play a causal role in lacunar stroke and cerebral small vessel disease? [J]. *Stroke*, 2024, 55(4): 934-942.
- [15] Yu Z, Liu X, Yang X, et al. Identification of Potential Therapeutic

- Targets in Dupuytren's Contracture Through Integrated Multi-Omics Analysis[EB/OL]. Rochester, NY: Social Science Research Network, 2025[2025-09-06]. <https://papers.ssrn.com/abstract=5278478>.
- [16] Pietzner M, Wheeler E, Carrasco-Zanini J, et al. Mapping the proteo-genomic convergence of human diseases[J]. *Science*, 2021, 374(6569): eabj1541.
- [17] Ferkingstad E, Sulem P, Atlason BA, et al. Large-scale integration of the plasma proteome with genetics and disease[J]. *Nat Genet*, 2021, 53(12): 1712-1721.
- [18] Sun BB, Chiou J, Traylor M, et al. Plasma proteomic associations with genetics and health in the UK Biobank[J]. *Nature*, 2023, 622(7982): 329-338.
- [19] Sun BB, Chiou J, Traylor M, et al. Genetic regulation of the human plasma proteome in 54, 306 UK Biobank participants[EB/OL]. bioRxiv, 2022: 2022.06.17.496443[2025-07-18]. <https://www.biorxiv.org/content/10.1101/2022.06.17.496443v1>.
- [20] Suhre K, Arnold M, Bhagwat AM, et al. Connecting genetic risk to disease end points through the human blood plasma proteome[J]. *Nat Commun*, 2017, 8: 14357.
- [21] Folkersen L, Gustafsson S, Wang Q, et al. Genomic and drug target evaluation of 90 cardiovascular proteins in 30, 931 individuals[J]. *Nat Metab*, 2020, 2(10): 1135-1148.
- [22] Yao C, Chen G, Song C, et al. Genome-wide mapping of plasma protein QTLs identifies putatively causal genes and pathways for cardiovascular disease[J]. *Nat Commun*, 2018, 9(1): 3268.
- [23] Hemani G, Zheng J, Elsworth B, et al. The MR-Base platform supports systematic causal inference across the human phenome[J]. *Elife*, 2018, 7: e34408.
- [24] Burgess S, Butterworth A, Thompson SG. Mendelian randomization analysis with multiple genetic variants using summarized data[J]. *Genet Epidemiol*, 2013, 37(7): 658-665.
- [25] Bowden J, Del Greco M F, Minelli C, et al. Improving the accuracy of two-sample summary-data Mendelian randomization: moving beyond the NOME assumption[J]. *Int J Epidemiol*, 2019, 48(3): 728-742.
- [26] Bowden J, Davey Smith G, Haycock PC, et al. Consistent estimation in mendelian randomization with some invalid instruments using a weighted Median estimator[J]. *Genet Epidemiol*, 2016, 40(4): 304-314.
- [27] Benjamini Y, Hochberg Y. Controlling the false discovery rate: a practical and powerful approach to multiple testing[J]. *J R Stat Soc Ser B Methodol*, 1995, 57(1): 289-300.
- [28] Giambartolomei C, Vukcevic D, Schadt EE, et al. Bayesian test for colocalisation between pairs of genetic association studies using summary statistics[J]. *PLoS Genet*, 2014, 10(5): e1004383.
- [29] 1000 Genomes Project Consortium; Auton A, Brooks LD, et al. A global reference for human genetic variation[J]. *Nature*, 2015, 526(7571): 68-74.
- [30] Berger J, Dorninger F, Forss-Petter S, et al. Peroxisomes in brain development and function[J]. *Biochim Biophys Acta*, 2016, 1863(5): 934-955.
- [31] Kassmann CM. Myelin peroxisomes - Essential organelles for the maintenance of white matter in the nervous system[J]. *Biochimie*, 2014, 98: 111-118.
- [32] Yao YC, Shahbazian A, Boström KI. Proline and gamma-carboxylated glutamate residues in matrix Gla protein are critical for binding of bone morphogenetic protein-4[J]. *Circ Res*, 2008, 102(9): 1065-1074.
- [33] Xu LL, Nirwane A, Yao Y. Basement membrane and blood-brain barrier[J]. *Stroke Vasc Neurol*, 2018, 4(2): 78-82.
- [34] Shibata M, Ohtani R, Ihara M, et al. White matter lesions and glial activation in a novel mouse model of chronic cerebral hypoperfusion[J]. *Stroke*, 2004, 35(11): 2598-2603.
- [35] Kakae M, Kawashita A, Onogi H, et al. Bilateral common carotid artery stenosis in mice: a model of chronic cerebral hypoperfusion-induced vascular cognitive impairment[J]. *Bio Protoc*, 2024, 14(13): e5022.
- [36] Xu WQ, Bai QK, Dong Q, et al. Blood-brain barrier dysfunction and the potential mechanisms in chronic cerebral hypoperfusion induced cognitive impairment[J]. *Front Cell Neurosci*, 2022, 16: 870674.
- (收稿:2025-10-28;修回:2025-10-29;录用:2025-11-20)
(责任编辑:唐秋姗)
- 本文引用格式:
刘小燕,余卓远,杨美.基于慢性脑灌注不足的蛋白质组孟德尔随机化研究揭示白质高信号的血浆靶点[J].重庆医科大学学报,2026, 51(1):55-65.