·论著·一次研究·

跨组织转录组关联研究揭示年龄相关性白内障新型易感基因及候选化合物

李 芳[1]，周鑫阳[2]，黄国鑫[2,3]

1. 武汉大学中南医院眼科（武汉 430071）
2. 湖北医药学院附属襄阳市第一人民医院循证医学中心（湖北襄阳 441000）
3. 青岛大学青岛医学院药学院（山东青岛 266023）

【摘要】目的 探索年龄相关性白内障的易感基因及候选化合物。**方法** 整合FinnGen R7年龄相关性白内障GWAS数据与GTEx V8多组织eQTL数据，采用跨组织TWAS、单组织TWAS、COJO及MAGMA筛选候选基因，并通过孟德尔随机化和多组织共定位分析验证；选择最优基因开展虚拟筛选和分子对接。**结果** 跨组织TWAS、单组织TWAS联合COJO及MAGMA分别筛选出38、120和680个显著基因，获得13个交集基因。孟德尔随机化分析显示，其中9个基因与疾病风险可能存在因果关系。多组织共定位分析显示，*IHH*在4种组织中存在边际共定位信号（PP.H4=0.803~0.835），但未达到严格标准。以*IHH*为靶点进行虚拟筛选和分子对接，获得71种候选化合物，其中Silibinin B的Docking得分和MM/GBSA ΔG Bind得分分别为−6.60和−41.87 kcal/mol，综合排名最优。**结论** 本研究筛选出13个候选易感基因，其中*IHH*获得孟德尔随机化和边际共定位证据支持。虚拟筛选提示Silibinin B具有较高的IHH计算机模拟结合优先级，可作为后续药物重定位研究的候选化合物，但仍需进一步实验验证。

【关键词】年龄相关性白内障；转录组关联研究；孟德尔随机化；共定位分析；分子对接

【中图分类号】R776.1 **【文献标志码】**A

A cross-tissue transcriptome-wide association study identifies a novel susceptibility gene for age-related cataract and candidate compounds

LI Fang[1], ZHOU Xinyang[2], HUANG Guoxin[2,3]

1. Department of Ophthalmology, Zhongnan Hospital of Wuhan University, Wuhan 430071, China
2. Evidence-Based Medicine Center, Xiangyang No.1 People's Hospital, Hubei University of Medicine, Xiangyang 441000, Hubei Province, China
3. School of Pharmacy, Qingdao University, Qingdao 266023, Shandong Province, China
Corresponding author: HUANG Guoxin, Email: xzyxhgx@163.com

【Abstract】Objective To explore susceptibility genes for age-related cataract and screen candidate compounds. **Methods** Genome-wide association study data on age-related cataract from FinnGen R7 were integrated with multi-tissue expression quantitative trait locus data from GTEx

DOI: 10.12173/j.issn.1005-0698.202608069
基金项目：湖北省“323”攻坚行动襄阳市第一人民医院重点专项科研基金（XYY2022-323）
通信作者：黄国鑫，主治医师，Email：xzyxhgx@163.com

V8. Cross-tissue transcriptome-wide association study, single-tissue transcriptome-wide association study, conditional and joint analysis, and gene-level analysis were used to screen candidate genes, which were further evaluated using Mendelian randomization and multi-tissue colocalization analysis. The optimal gene was selected as the target for virtual screening and molecular docking. **Results** Cross-tissue transcriptome-wide association study, single-tissue transcriptome-wide association study combined with conditional and joint analysis, and gene-level analysis identified 38, 120, and 680 significant genes, respectively, yielding 13 overlapping genes. Mendelian randomization analysis showed that 9 of these genes might have causal associations with disease risk. Multi-tissue colocalization analysis showed marginal colocalization signals for *IHH* in four tissues, with PP. H4 values ranging from 0.803 to 0.835, although none reached the stringent colocalization threshold. Virtual screening and molecular docking targeting IHH yielded 71 candidate compounds. Silibinin B ranked highest overall, with a docking score of −6.60 and an MM/GBSA ΔG Bind free energy of −41.87 kcal/mol. **Conclusion** This study identified 13 candidate susceptibility genes, among which *IHH* was supported by Mendelian randomization and marginal colocalization evidence. Virtual screening indicated that Silibinin B had a relatively high computational binding priority for *IHH*, suggesting its potential as a candidate compound for subsequent drug repurposing studies. However, further experimental validation is still required.

【Keywords】Age-related cataract; Transcriptome-wide association study; Mendelian randomization; Colocalization analysis; Molecular docking

年龄相关性白内障是由晶状体蛋白变性、聚集及晶状体透明度进行性丧失所致的常见眼病，是全球首位可逆性致盲及视力损害病因[1]。全球疾病负担研究显示，年龄相关性白内障相关的中重度视力损害患者数量随人口老龄化持续增长，尤其在中低收入地区，因手术资源可及性受限，其造成的致盲负担尤为突出[2]。尽管年龄相关性白内障摘除联合人工晶状体植入术已成为成熟且有效的治疗手段，但手术本身存在感染、术后并发症及医疗资源分布不均等局限[3-4]。因此，从病因学层面理解年龄相关性白内障的发生机制、寻找可用于早期干预或延缓疾病进展的药物靶点，仍具有重要的临床和公共卫生意义[5]。

基于双生子及家族聚集性研究，遗传因素在年龄相关性白内障易感性中发挥重要作用。双生子研究显示，年龄相关性核性白内障的遗传力约为48%，而皮质性白内障的遗传力约为53%~58%，提示遗传变异在疾病发生发展中扮演着不可忽视的角色[6]。近年来，大规模全基因组关联研究（genome-wide association studies，GWAS）已成功识别出数十个与年龄相关性白内障显著相关的风险位点，其中涵盖了晶状体蛋白基因、间隙连接蛋白基因以及膜内在蛋白基因等经典候选区域，极大地增进了对遗传变异影响晶状体透明性的认识[7]。然而，与多数复杂性状疾病类似，GWAS鉴定的风险变异绝大多数位于基因间区或内含子等非编码区域，其下游真正的效应基因及调控机制难以通过单纯的统计关联加以明确[8-9]。此外，复杂连锁不平衡（linkage disequilibrium，LD）结构的存在，使得从被动的统计信号中精确甄别出驱动疾病发生的因果变异变得尤为困难，严重阻碍了遗传发现向临床转化和药物靶点开发的推进[10]。

近年来，跨组织全转录组关联研究（Transcriptome-Wide Association Study，TWAS）分析已被广泛应用于多种复杂性状及疾病易感基因的鉴定，包括类风湿关节炎、肺癌、自闭症谱系障碍等多系统疾病，并在眼科及神经系统疾病领域取得进展[11-12]：既有研究通过跨组织TWAS联合单组织连接基因表达与疾病的桥梁（Functional Summarybased Imputation，FUSION）验证、基因组注释的多标记分析（Multi-marker Analysis of Genomic Annotation，MAGMA）、孟德尔随机化（mendelian randomization，MR）及共定位分析，先后鉴定出偏头痛的*REV1*和*SREBF2*[13]、原发性开角型青光眼的*AFAP1*和*FMNL2*[14]、感音神经性耳聋的*SPTBN1*等多个候选易感基因[15]，并进一步在重度抑郁症研究中将

该策略拓展至候选靶点的分子对接及成药性评估[16]。这些研究提示，通过多方法交叉验证的跨组织TWAS策略，能够在缺乏目标组织直接表达数据的情形下，仍从全身多组织的遗传调控信号中捕获与疾病风险相关的因果候选基因，为复杂疾病的分子机制解析和药物重定位提供了可行路径[17]。

然而，目前尚缺乏针对年龄相关性白内障的系统性跨组织TWAS研究。鉴于晶状体组织表达数据的缺失是限制年龄相关性白内障遗传学研究向机制解析转化的关键瓶颈，本研究拟整合年龄相关性白内障GWAS汇总数据与GTEx V8多组织表达数量性状位点（expression Quantitative Trait Locus，eQTL）数据，采用跨组织TWAS分析，并通过FUSION单组织验证、MAGMA基因水平分析及多组织共定位分析（colocalization analysis，COLOC）联合分析进一步筛选候选易感基因；在此基础上，利用MR及贝叶斯共定位分析验证候选基因与年龄相关性白内障风险的因果关联。在遗传学分析基础上，选择具有较高优先级的候选基因开展虚拟筛选与分子对接，从美国食品药品监督管理局（Food and Drug Administration，FDA）批准药物库中筛选潜在结合化合物。本研究旨在为阐明年龄相关性白内障的遗传学基础提供新的候选基因和分子层面的解释，并为后续药物重定位及实验验证提供候选靶点和化合物依据。

1 资料和方法

本研究设计策略见流程图1。整合年龄相关性白内障GWAS汇总统计数据与GTEx V8多组织eQTL数据，通过跨组织TWAS、单组织TWAS、MAGMA、条件与联合分析（Conditional and Joint Analysis，COJO）及MR等多层次分析，筛选出获得多种方法共同支持的优先候选基因。然后针对候选基因为候选靶点开展虚拟筛选和分子对接。

1.1 数据来源

1.1.1 年龄相关性白内障数据

本研究年龄相关性白内障的GWAS汇总数据来源于FinnGen R7数据库，疾病终点编号为H7_CATARACTSENILE。该表型依据芬兰国家医疗登记系统中的ICD-10 H25、ICD-9 3661及ICD-8 37402等诊断编码定义，研究对象主要为芬兰人群。该GWAS共纳入42 843例年龄相关性白内障病例和262 698名对照，其中女性病例23 208例、男性病例19 635例。病例的平均首次发生年龄为71.49岁，其中女性为71.13岁，男性为71.90岁[18]。

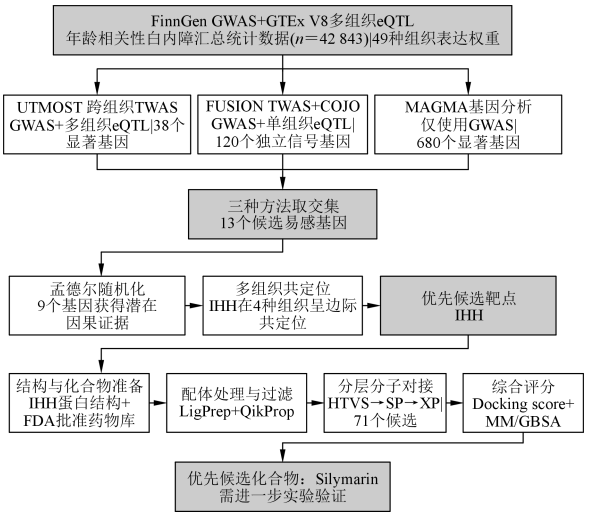


图1 研究流程图

Figure 1. Research flowchart

1.1.2 表达数量性状位点数据

表达数量性状位点数据来自GTEx V8数据集（https://www.gtexportal.org/home/）。该数据系统性地研究人类不同组织中，遗传变异如何影响基因表达，V8版本是目前最大的人类基因表达与基因型数据图谱之一。数据来自838例死后捐献者，检查了来自49种组织，涵盖了人体主要器官和系统，包括神经系统、消化系统、内分泌系统等[19]。

1.2 统计方法

1.2.1 跨组织转录组关联分析

跨组织全转录组关联分析是基于公开GTEx多组织表达权重数据库，通过整合多组织eQTL信息，提升基因-性状关联检验的统计效能。该方法构建跨组织基因表达预测模型，估算遗传调控基因表达量与目标性状的关联，并采用广义Berk-Jones检验整合各组织统计量，校正组织间表达相关性，提升检验效能。所有分析使用UTMOST标准流程完成，对基因水平关联P值采用Benjamini-Hochberg方法进行FDR校正，以FDR<0.05作为统计学显著性标准，筛选候选易感基因。

1.2.2 单组织转录组关联分析

单组织转录组关联研究分析是基于GTEx v8参考面板的cis-eQTL数据，构建基因表达水平的

线性预测模型，并将其应用于GWAS汇总统计量，从而评估基因预测表达水平与目标性状之间的关联。该方法以基因转录起始位点上下游1 Mb为cis窗口，筛选该区间内所有SNP；然后采用五种预测模型（BLUP、BSLMM、LASSO、Elastic Net及top eQTL）分别拟合SNP基因型与基因表达量之间的关系，通过五折交叉验证选取预测性能最优的模型；再利用1000 Genomes Project欧洲参考人群估算SNP间的LD矩阵，将最优模型的权重系数与GWAS SNP效应Z-score进行线性组合，计算基因水平的TWAS Z-score及其显著性*P*值；最后对每个基因在不同组织中的关联结果分别进行FDR校正，以FDR<0.05为统计学显著性阈值。为避免局部LD造成的假阳性，进一步采用COJO分析识别条件独立的多基因信号。

1.2.3 MAGMA

MAGMA是将GWAS汇总统计量中SNP的关联信号，在充分考虑LD结构的前提下，聚合至基因层面，从而计算每个基因与表型的关联显著性。该方法基于1000 Genomes Project Phase 3欧洲参考人群估算LD矩阵；然后将SNP注释至其所在基因的转录区域（±0 kb）；采用基于多元回归的基因水平关联模型，计算每个基因的关联Z-score及显著性*P*值；最后应用FDR法对多重检验进行校正，以FDR<0.05为关联达到统计学显著性。

1.2.4 MR分析

MR是利用遗传变异作为工具变量（instrumental variables，IVs），通过模拟随机对照试验的分配机制，推断暴露因素（基因表达）与结局（疾病）之间的因果效应[20]。该方法首先提取各候选基因转录起始位点上下游1 Mb范围内的顺式表达数量性状位点作为IVs；然后筛选全基因组显著水平的工具变量（$P<5\times10^{-8}$），并进行LD聚类以获取独立SNP（$r^2<0.001$）；再采用TwoSampleMR R包进行MR分析，当仅有一个独立IV时使用Wald比值法（Wald ratio）估计因果效应，当有两个及以上IV时采用逆方差加权法（inverse variance weighted，IVW）作为主要分析方法；最后以*P*<0.05为关联达到统计学显著性，判断基因表达与疾病之间可能存在提示性因果关联。

1.2.5 COLOC

COLOC基于近似贝叶斯因子框架，批量匹配GWAS与GTEx多组织cis-eQTL信号，区分组织特异性、跨组织共享致病变异；通过5类假设后验概率量化共定位证据，规避LD造成的假关联，明确基因表达调控在多组织中与疾病的共享遗传位点。该方法首先提取目标基因组区域内GWAS及各组织eQTL的汇总统计信息，然后将多组织QTL数据与GWAS数据按SNP进行匹配与一致化；再根据研究目标选择适用的多组织共定位算法，执行共定位计算，输出各组织或性状组合的后验概率；最后以PP.H4>0.9、PP.H4/PP.H3>5为严格共定位标准，PP.H4>0.75为边际阈值，筛选具有跨组织共享遗传机制的目标基因。

1.2.6 虚拟筛选与分子对接

根据COLOC的结果，去探索最优目标基因的候选化合物。目标蛋白的三维结构来源于AlphaFold蛋白质结构数据库（https://alphafold.ebi.ac.uk/）。采用Schrodinger软件套件（2023-1版，Schrodinger，LLC）开展虚拟筛选和分子对接，并将获取的蛋白质三维结构作为分子对接受体。按照软件默认设置对受体结构进行预处理，随后采用SiteMap模块对蛋白质表面的潜在结合空腔进行扫描。口袋筛选时将最小口袋体积设置为100 Å^3，并启用能量网格计算，以评价候选口袋的疏水作用势、氢键作用特征等性质，探针半径设置为1.0 Å。综合SiteScore和Dscore对候选结合口袋进行评价，并根据综合评价结果确定分子对接位点。以候选结合位点为中心构建尺寸为20 Å×20 Å×20 Å的立方体对接网格，用于后续分子对接分析。

用于虚拟筛选的小分子化合物来源于陶术FDA批准药物库（L1300-FDA-approved-Drug-Library）。采用Schrodinger软件中的LigPrep模块对配体结构进行标准化和构象处理，相关参数均采用软件默认设置，并过滤不合理的分子结构或构象。采用QikProp模块对处理后的化合物进行吸收、分布、代谢、排泄和毒性相关性质预测，同时开展类药性评价及化学空间过滤，以构建用于后续虚拟筛选的候选化合物库。

将预处理后的受体结构和小分子结构提交至Schrodinger虚拟筛选流程进行分层分子对接。首先采用高通量虚拟筛选模式（high-throughput virtual screening，HTVS）进行初步筛选，并保留对接评分排名前15%的化合物；随后采用标准精

度模式（standard precision，SP）进行进一步筛选，并保留该阶段对接评分排名前30%的化合物；最后采用超精度模式（extra precision，XP）进行精细分子对接，并按照对接评分排序保留排名前95%的化合物。完成分层筛选后，对候选化合物进行同分异构体去除处理，并计算结合自由能（MM/GBSA ΔG Bind）得分和对接（Docking）得分，获得用于后续分析的候选化合物。

2 结果

2.1 跨组织转录组关联分析结果

在跨组织转录组关联分析中共筛选出333个基因*P*值小于0.05，经过FDR调整后，剩下38个显著基因。38个基因主要集中在2号染色体，基因宽度的中位数（四分位数）为18 731（2 699，66 273），29个基因为蛋白编码，见表1。

2.2 单组织转录组关联分析及COJO分析结果

在单组织转录组关联分析中共筛选出38 002个基因*P*值小于0.05，经过FDR调整后，剩下3 923个显著基因。3 923个基因主要集中在17、1、8号染色体，基因宽度的中位数（四分位数）为21 118（7 107，59 307），2 851个基因为蛋白编码。采用COJO分析消除连锁不平衡导致的假阳性结果，共筛选出120个显著基因，见附件表1。

2.3 MAGMA结果

在MAGMA中，经过FDR调整后，共筛选出680个显著基因。680个基因主要集中在6、17、2号染色体，基因宽度的中位数（四分位数）为26 001.00（9 106.00，89 845.00），基因碱基窗口内SNP数量的中位数（四分位数）为112.50（35.00，363.00），详细结果见附件表2。

2.4 交集基因

综合跨组织转录组关联分析结果、单组织转录组关联分析及COJO分析结果、MAGMA结果，共筛选出13个基因与年龄相关性白内障显著相关，分别为*ANKRD49P1*、*FAM117B*、*CGREF1*、*PLCL1*、*KHK*、*NHEJ1*、*RETREG2*、*RFTN2*、*IHH*、*OTOF*、*ABCB6*、*ZFAND2B*、*CCDC117*。见图2。

2.5 MR结果

对13个基因进行MR分析，最终只有9个基因可进行匹配。9个基因分别为*ABCB6*、*CGREF1*、*FAM117B*、*IHH*、*KHK*、*NHEJ1*、*PLCL1*、*RFTN2*、*ZFAND2B*。其中基因*KHK*与29个组织之间可能存在提示性因果关联；基因*CGREF1*、*IHH*、*NHEJ1*、*ZFAND2B*与5个组织之间可能存在提示性因果关联；基因*PLCL1*、*RFTN2*与4个组织之

表1 跨组织转录组关联分析结果

Table 1. Multi-tissue transcriptome-wide association study results

基因编号	基因名	染色体号	FDR	基因编号	基因名	染色体号	FDR
ENSG00000228446	ANKRD49P1	2	0.00	ENSG00000163501	IHH	2	0.03
ENSG00000171174	RBKS	2	0.01	ENSG00000171885	AQP4	18	0.03
ENSG00000248494	LNX1-AS2	4	0.01	ENSG00000213901	SLC23A3	2	0.03
ENSG00000133477	FAM83F	22	0.01	ENSG00000211611	IGKV6-21	2	0.04
ENSG00000138439	FAM117B	2	0.01	ENSG00000143869	GDF7	2	0.04
ENSG00000163521	GLB1L	2	0.01	ENSG00000182944	EWSR1	22	0.04
ENSG00000227557	MTND3P9	2	0.01	ENSG00000204640	NMS	2	0.04
ENSG00000138028	CGREF1	2	0.02	ENSG00000225774	SIRPAP1	22	0.04
ENSG00000115896	PLCL1	2	0.03	ENSG00000264707	L3MBTL4-AS1	18	0.04
ENSG00000152056	AP1S3	2	0.03	ENSG00000115155	OTOF	2	0.04
ENSG00000181378	CFAP65	2	0.03	ENSG00000115657	ABCB6	2	0.04
ENSG00000226261	PIMREGP1	2	0.03	ENSG00000158552	ZFAND2B	2	0.04
ENSG00000138030	KHK	2	0.03	ENSG00000159873	CCDC117	22	0.04
ENSG00000187736	NHEJ1	2	0.03	ENSG00000171132	PRKCE	2	0.04
ENSG00000127824	TUBA4A	2	0.03	ENSG00000264012	GATA6	18	0.04
ENSG00000143921	ABCG8	2	0.03	ENSG00000171951	SCG2	2	0.04
ENSG00000115020	PIKFYVE	2	0.03	ENSG00000135900	MRPL44	2	<0.05
ENSG00000144567	RETREG2	2	0.03	ENSG00000163798	SLC4A1AP	2	<0.05
ENSG00000162944	RFTN2	2	0.03	ENSG00000237350	CDC42P6	4	<0.05

间可能存在提示性因果关联，见图3。

2.6 COLOC结果

在COLOC中以PP.H4>0.9、PP.H4/PP.H3>5为严格共定位标准，PP.H4>0.75为边际阈值。全部位点均未达到严格共定位标准，仅*IHH*基因区域在4种组织存在边际信号：胫神经（PP.H4=0.835）、小脑半球（PP.H4=0.826）、下肢光暴露皮肤（PP.H4=0.818）、非暴露耻骨皮肤（PP.H4=0.803），上述区域PP.H4/PP.H3均>9，提示*IHH*区域可能存在跨组织共享的遗传调控信号。其余8个基因在全部组织PP.H4均<0.75，PP.H3占比更高，说明GWAS与eQTL多为独立遗传信号。整体提示多数白内障致病变异不通过外周组织cis表达调控发挥作用。

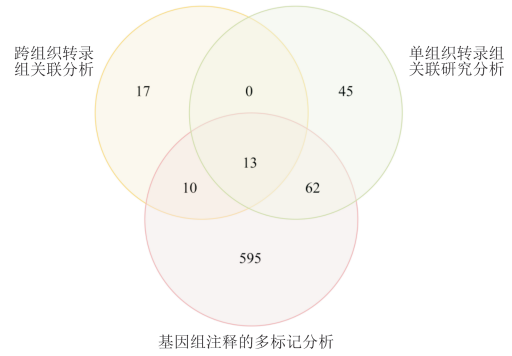


图2 交集基因韦恩图

Figure 2. Venn diagram of overlapping genes

注：跨组织TWAS、单组织TWAS联合COJO以及MAGMA分析分别获得不同数量的显著基因，三种方法共有13个交集基因。

2.7 虚拟筛选与分子对接结果

根据筛选出的最优基因IH*H*进行虚拟筛选和分子对接。AlphaFold蛋白质结构数据库中*IHH*的ID为AF-Q14623-F1-model_v6。经LigPrep模块进行配体结构处理后，共生成6 122种小分子结构。进一步通过QikProp进行ADME/T性质预测、类药性评价及化学空间过滤后，共有3 912种小分子结构符合筛选要求并进入分层分子对接流程。在HTVS中，按照对接评分保留排名前15%的化合物，共获得586种候选化合物；随后采用SP进行进一步筛选，保留对接评分排名前30%的化合物，共获得175种候选化合物。最后采用XP进行精细分子对接，按对接评分保留排名前95%的化合物，并进一步去除同分异构体，最终获得71种候选化合物，计算MMGBSA dG Bind得分和Docking得分，两种得分各前10名情况见表2。综合Docking得分和MM/GBSA ΔG Bind得分的排序结果，Silymarin在两种评价指标中均具有较高排名，因此被确定为具有进一步研究价值的优先候选化合物。数据库中Silymarin的实际结果为Silibinin B（Silybin B，PubChem CID：1548994，CAS：142797-34-0，分子式$C_{25}H_{22}O_{10}$），并进一步绘制IHH蛋白与Silibinin B的分子对接相互作用模式图，见图4。

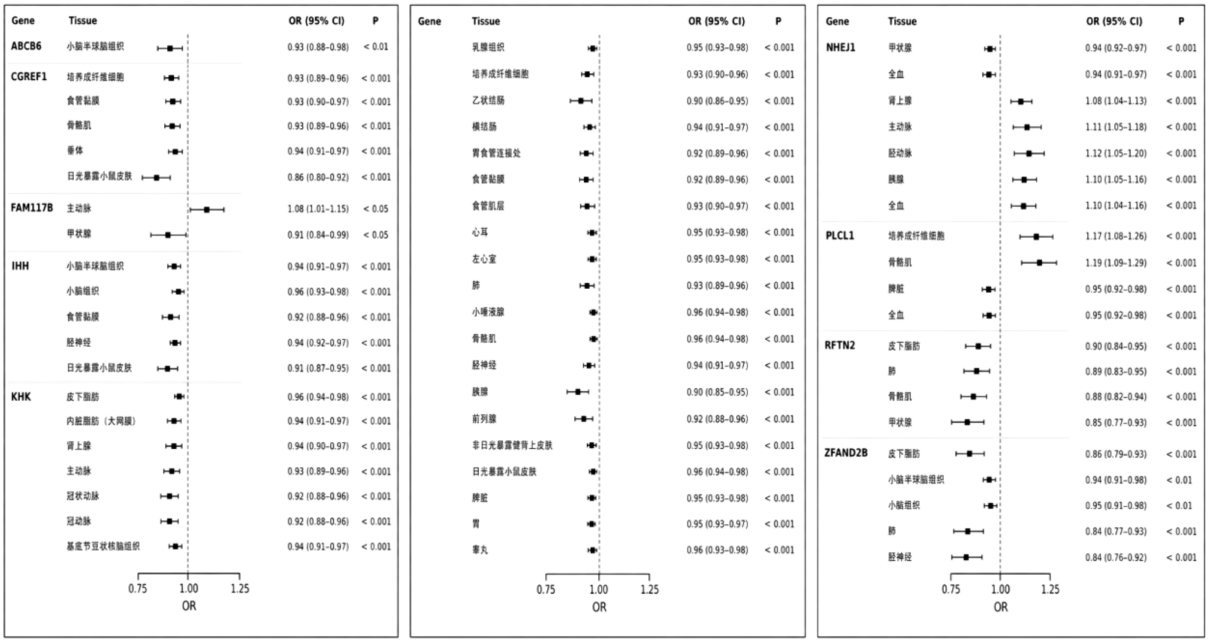


图3 MR分析结果

Figure 3. Results of Mendelian randomization analysis

表2 结合自由能得分和对接得分前十药物

Table 2. Top ten compounds ranked by MM/GBSA ΔG Bind score and docking score

MM/GBSA ΔG Bind得分前十药物			Docking得分前十药物		
药物名称	MM/GBSA ΔG Bind	Docking	药物名称	MM/GBSA ΔG Bind	Docking
氢溴酸依来曲坦	−41.16	−4.52	氧化多巴胺	−28.11	−6.54
水飞蓟素	−41.87	−6.60	盐酸异丙肾上腺素	−32.72	−6.54
二丙酸咪多卡	−43.78	−5.34	水飞蓟素	−41.87	−6.60
磷酸卡维地洛	−44.74	−6.36	二水硫酸异丙肾上腺素	−29.31	−6.76
盐酸地芬尼多	−44.75	−4.50	沙丁胺醇	−35.84	−6.86
盐酸莱克多巴胺	−46.26	−5.68	刃天青钠	−21.83	−6.88
西洛多辛	−46.46	−4.94	甲磺酸异他林	−29.92	−6.88
甲磺酸阿帕替尼	−47.74	−4.85	酒石酸肾上腺素	−24.24	−7.08
达可替尼	−48.24	−4.59	盐酸肾上腺素	−24.25	−7.08
沙美特罗	−50.63	−5.47	盐酸去甲肾上腺素	−26.73	−7.09

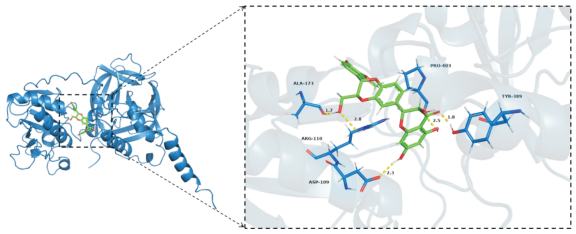


图4 IHH蛋白与Silibinin B的分子对接分析

Figure 4. Molecular docking analysis of IHH protein and Silibinin B

注：Silibinin B在预测IHH结合口袋中的三维结合构象及空间匹配情况。

3 讨论

本研究通过整合年龄相关性白内障GWAS汇总数据与GTEx V8多组织eQTL数据，采用UTMOST跨组织TWAS分析、FUSION单组织验证、COJO条件联合分析及MAGMA基因水平分析等多层次验证策略，系统筛选出13个基因同时获得多种基因水平分析方法支持。进一步通过MR分析确认其中9个基因（*ABCB6*、*CGREF1*、*FAM117B*、*IHH*、*KHK*、*NHEJ1*、*PLCL1*、*RFTN2*、*ZFAND2B*）与年龄相关性白内障存在潜在因果关联。COLOC提示*IHH*基因区域在胫神经、小脑半球及皮肤组织中存在边际共定位信号（PP.H4=0.803~0.835），表明白内障致病变异可能通过神经或皮肤组织的基因表达调控间接参与疾病发生，最后本研究以遗传学证据优先级较高的*IHH*为候选靶点，通过虚拟筛选和分子对接对FDA批准化合物库进行探索性筛选，结果提示Silibinin B具有相对较高的计算机模拟结合优先级。

跨组织TWAS分析已被广泛应用于多种复杂疾病易感基因的鉴定。与既往偏头痛研究通过跨组织TWAS联合FUSION、MAGMA、MR及共定位分析鉴定出*REV1*和*SREBF2*[13]，以及感音神经性耳聋研究鉴定出*SPTBN1*、*EML6*等8个候选基因的策略一致[15]，本研究采用的多层次交叉验证方法有效降低了假阳性率，提高了候选基因的可信度。同时，本研究所识别的13个候选基因中有9个通过了MR因果关联验证，这一比例高于部分既往研究，可能与年龄相关性白内障较高的遗传力（约35%~50%）以及FinnGen数据库较大的样本量有关。

本研究发现的候选基因可能参与多个与白内障发生密切相关的生物学过程。其中，*KHK*（ketohexokinase）编码果糖代谢关键酶，参与糖代谢调节[21]。近年来研究发现，异常糖代谢和氧化应激是促进晶状体蛋白损伤和聚集的重要因素，高糖环境可诱导活性氧产生增加，导致晶状体蛋白发生氧化修饰、糖基化及结构改变，最终降低晶状体透明度[22-23]。因此，*KHK*表达异常可能通过影响细胞能量代谢和氧化还原平衡参与年龄相关性白内障发生。此外，*IHH*作为Hedgehog信号通路的重要成员，在细胞增殖、分化及组织稳态维持中发挥关键作用[24]。已有研究表明，Hedgehog信号通路参与眼组织发育以及晶状体形成过程，其异常调控可能影响晶状体上皮细胞功能及细胞稳态。本研究MR分析显示*IHH*与年龄相关性白内障风险存在潜在因果关系，提示该基因可能成为连接遗传调控与晶状体退行性改变的重要节点[25-26]。本研究的COLOC进一步显示，*IHH*基因区域在胫神经、小脑半球及暴露/非暴露皮肤组织中存在边际共定位信号（PP.H4=0.803~0.835，且PP.H4/PP.H3均>9），提示该位点的遗

传变异可能通过神经外胚层来源组织中的顺式表达调控参与白内障相关的病理过程，但其在晶状体或眼周组织中的直接作用仍有待验证。*ABCB6*、*NHEJ1*、PLC*L1*及*ZFAND2B*等候选基因同样具有潜在生物学意义。ABCB6属于ATP结合盒转运蛋白家族，参与卟啉和血红素相关代谢物的跨膜运输，并可能参与细胞氧化应激防御。线粒体功能异常和氧化还原稳态破坏被认为与年龄相关性白内障的发生密切相关。随着年龄增长，晶状体长期受到紫外线及其他氧化因素影响，线粒体损伤和活性氧累积可能导致晶状体上皮细胞功能异常，并促进晶状体蛋白氧化修饰、交联与聚集[27]。因此，*ABCB6*表达异常可能通过影响代谢物转运和细胞氧化应激耐受性参与白内障发生。另一方面，*NHEJ1*编码非同源末端连接因子XLF，能够与XRCC4－DNA连接酶Ⅳ复合物协同参与DNA双链断裂修复，对维持基因组稳定性具有重要作用[28]。既往研究发现，年龄相关性白内障患者晶状体上皮细胞中的DNA损伤显著增加。因此，*NHEJ1*表达或功能异常可能削弱DNA损伤修复能力，促进晶状体细胞衰老及透明性丧失，但这一推测仍需功能实验验证[26]。

本研究以IHH蛋白为受体开展虚拟筛选，Silibinin B的Docking得分和MM/GBSA ΔG Bind得分分别为−6.60和−41.87 kcal/mol，且均进入前10位，提示其可能与IHH形成相对稳定的结合。既往研究表明，Silymarin可改善糖尿病大鼠晶状体的氧化应激及多元醇通路异常[29]；其主要活性成分silibinin还能抑制糖基化、晶状体蛋白聚集和醛糖还原酶活性，并减弱TGF−β2诱导的晶状体上皮细胞迁移及上皮−间质转化[30-31]。鉴于氧化损伤、蛋白糖基化和晶状体上皮细胞功能障碍均参与年龄相关性白内障发生，上述证据为Silymarin的潜在晶状体保护作用提供了生物学依据。此外，silibinin还能抑制TGF−β2诱导的人晶状体上皮细胞迁移及上皮－间质转化，该作用可能与PI3K－Akt信号通路受到抑制有关[29]。结合本研究的遗传学结果，推测Silymarin除发挥抗氧化和抗糖基化作用外，还可能通过调节IHH/Hedgehog信号影响晶状体稳态。然而，分子对接仅反映理论结合可能性，尚不能证明Silymarin能够直接调控IHH或治疗白内障。此外，Silymarin属于多成分混合物，其具体活性单体及眼内生物利用度仍不明确，需通过分子动力学、结合实验、晶状体上皮细胞及动物模型验证其作用机制与治疗效果。

从遗传学或计算数据库筛选获得候选靶点，到最终形成可用于临床的治疗药物，仍存在较大的转化距离。目前尚缺乏统一数据能够准确估计数据库筛选候选化合物的最终临床转化率。已有研究显示，即使进入临床开发阶段，药物最终获批的总体比例仍约为10%；但具有明确人类遗传学支持的药物机制，临床成功概率约为缺乏遗传学支持者的2.6倍[32]。另有研究显示，2013—2022年FDA批准的新药中约63%的靶点或相关蛋白具有相应的人类遗传学证据[33]。因此，TWAS、MR及虚拟筛选等方法的主要价值在于提高候选靶点和化合物的优先级。

本研究仍存在一些局限。首先，本研究GWAS数据主要来源于欧洲人群，因此研究结果是否适用于亚洲及其他种族群体仍需进一步验证。其次，由于目前缺乏大规模人类晶状体组织eQTL数据库，本研究主要依赖GTEx多组织表达数据进行推断，可能无法完全反映晶状体特异性调控机制。未来随着单细胞转录组和晶状体组织遗传资源的发展，可进一步验证候选基因在不同晶状体细胞类型中的功能。此外，所有MR关联在FDR校正后均未达到统计学显著性，因此目前结果仅能作为探索性和假设生成性证据，同时MR分析显示IHH预测表达与白内障风险可能存在因果关系，而共定位分析仅观察到边际共享信号，尚不能排除由不同连锁变异驱动关联的可能，仍需要细胞实验和动物模型进一步验证候选基因的具体作用机制。并且，虚拟筛选、分子对接及MM/GBSA计算仅反映候选化合物与IHH蛋白在特定构象下的理论结合可能性，不能直接证明其真实结合、靶点调控作用及对白内障的治疗效果。此外，由于GTEx数据库缺乏晶状体组织eQTL数据，本研究的跨组织共定位分析主要基于脑、神经、皮肤等非晶状体组织，因此其结果仅能提供间接遗传学线索，不能替代晶状体组织特异性证据。未来应结合晶状体上皮细胞、晶状体组织及单细胞转录组数据，并通过功能实验进一步验证候选基因的组织特异性作用。

综上所述，本研究通过整合多组织转录组关联分析及多种遗传学方法，筛选出13个与年龄相

关性白内障相关的候选基因，其中 *IHH* 获得 MR 和边际共定位证据支持。进一步的虚拟筛选与分子对接提示，Silibinin B 可能是作用于 IHH 的候选化合物。上述结果为年龄相关性白内障的遗传机制研究及药物重定位提供了新的线索，但仍需通过细胞和动物实验进一步验证。

附件见《药物流行病学杂志》官网附录（https://ywlxbx.whuznhmedj.com/futureApi/storage/appendix/202608069.xlsx）

利益冲突声明：作者声明本研究不存在任何经济或非经济利益冲突。

参考文献

1 中国康复医学会视觉康复专业委员会. 白内障健康管理流程专家共识(2024 版)[J]. 中华预防医学杂志, 2025, 59(3): 263-270. [Visual Rehabilitation Professional Committee of the Chinese Association of Rehabilitation Medicine. Expert consensus on the health management process of cataracts (2024)[J]. Chinese Journal of Preventive Medicine, 2025, 59(3): 263-270.] DOI: 10.3760/cma.j.cn112150-20241223-01029.

2 Wan Z, Bai J, Wang W, et al. Global, regional, and national burden of cataract among older adults from 1990 to 2021: a comprehensive analysis based on the Global Burden of Disease Study 2021[J]. Front Med (Lausanne), 2025, 12: 1679828. DOI: 10.3389/fmed.2025.1679828.

3 de Diego-García L, Rejas-González R, Latre IC, et al. Pharmacological strategies for cataract management: from molecular targets to clinical translation[J]. Int J Mol Sci, 2025, 26(12): 5658. DOI: 10.3390/ijms26125658.

4 柳玉林, 曲华, 杜金库, 等. 白内障摘除手术感染防控措施的最佳证据总结[J]. 上海护理, 2023, 23(1): 44-49. [Liu YL, Qu H, Du JK, et al. Summary of best evidence for infection prevention and control measures after cataract extraction surgery[J]. Shanghai Nursing, 2023, 23(1): 44-49.] DOI: 10.3969/j.issn.1009-8399.2023.01.011.

5 卫佳奇, 张琇雯, 黄滔敏. 白内障的药物治疗及研究进展[J]. 眼科学报, 2025, 40(7): 565-573. [Wei JQ, Zhang XW, Huang TM. Pharmacological treatment and research progress ofcataracts[J]. Eye Science, 2025, 40(7): 565-573.] DOI: 10.12419/2041507.

6 Hammond CJ, Snieder H, Spector TD, et al. Genetic and environmental factors in age-related nuclear cataracts in monozygotic and dizygotic twins[J]. N Engl J Med, 2000, 342(24): 1786-1790. DOI: 10.1056/NEJM200006153422404.

7 Shiels A, Hejtmancik JF. Molecular Genetics of Cataract[J]. Prog Mol Biol Transl Sci, 2015, 134: 203-218. DOI: 10.1016/bs.pmbts.2015.05.004.

8 Liao J, Su X, Chen P, et al. Meta-analysis of genome-wide association studies in multiethnic Asians identifies two loci for age-related nuclear cataract[J]. Hum Mol Genet, 2014, 23(22): 6119-6128. DOI: 10.1093/hmg/ddu315.

9 Choquet H, Melles RB, Anand D, et al. A large multiethnic GWAS meta-analysis of cataract identifies new risk loci and sex-specific effects[J]. Nat Commun, 2021, 12(1): 3595. DOI: 10.1038/s41467-021-23873-8.

10 Diaz-Torres S, Lee SS, García-Marín LM, et al. Uncovering genetic loci and biological pathways associated with age-related cataracts through GWAS meta-analysis[J]. Nat Commun, 2024, 15(1): 9116. DOI: 10.1038/s41467-024-53212-6.

11 Hu Y, Li M, Lu Q, et al. A statistical framework for cross-tissue transcriptome-wide association analysis[J]. Nat Genet, 2019, 51(3): 568-576. DOI: 10.1038/s41588-019-0345-7.

12 Ni J, Wang P, Yin K, et al. Novel insight into the aetiology of rheumatoid arthritis gained by a cross-tissue transcriptome-wide association study[J]. RMD Open, 2022, 8(2): e002529. DOI: 10.1136/rmdopen-2022-002529.

13 Gui J, Yang X, Tan C, et al. A cross-tissue transcriptome-wide association study reveals novel susceptibility genes for migraine[J]. J Headache Pain, 2024, 25(1): 94. DOI: 10.1186/s10194-024-01802-6.

14 Chen J, Zhuo X, Li Y, et al. Cross-tissue transcriptome-wide association study identifies novel genes associated with POAG[J]. Invest Ophthalmol Vis Sci, 2025, 66(6): 7. DOI: 10.1167/iovs.66.6.7.

15 孙红, 卜凤娇, 闫晶超, 等. 感音神经性耳聋的潜在药物靶点：一项跨组织转录组关联研究[J]. 复旦学报(医学版), 2026, 53(3): 350-359. [Sun H, Bu FJ, Yan JC, et al. Potential drug targets for sensorineural hearing loss: a cross-tissue transcriptome-wide association study[J]. Fudan University Journal of Medical Sciences, 2026, 53(3): 350-359.] DOI: 10.3969/j.issn.1672-8467.2026.03.006.

16 商然, 蒙秋晓, 史卫忠, 等. 跨转录组关联研究识别重度抑郁症潜在药物靶点[J]. 临床药物治疗杂志, 2026, 24(6): 40-46. [Shang R, Meng QX, Shi WZ, et al. Cross-tissue transcriptome-wide association study identifies novel susceptibility genes for major depressive disorder[J]. Clinical Medication Journal, 2026, 24(6): 40-46.] DOI: 10.3969/j.issn.1672-3384.2026.06.006.

17 方堃, 李国壮, 王霖霆, 等. 从基因表达到全转录组关联研究的方法学发展与比较[J]. 协和医学杂志, 2026, 17(1): 223-229. [Fang K, Li GZ, Wang LT, et al. From gene expression to transcriptome-wide association study: development and comparison of methodolgy[J]. Medical Journal of Peking Union Medical College Hospital, 2026, 17(1): 223-229.] DOI: 10.12290/xhyxzz.2024-0920.

18 Kurki MI, Karjalainen J, Palta P, et al. FinnGen provides genetic insights from a well-phenotyped isolated population[J]. Nature, 2023, 613(7944): 508-518. DOI: 10.1038/s41586-022-05473-8.

19 Consortium G. The GTEx Consortium atlas of genetic regulatory effects across human tissues[J]. Science, 2020, 369(6509): 1318-1330. DOI: 10.1126/science.aaz1776.

20 王莉娜, Zhang Zuofeng. 孟德尔随机化法在因果推断中的应用[J]. 中华流行病学杂志, 2017, 38(4): 547-552. [Wang LN, Zhang ZF. Mendelian randomization approach, used for causal inferences[J]. Chinese Journal of Epidemiology. 2017, 38(4): 547-552.] DOI: 10.3760/cma.j.issn.0254-6450.2017.04.027.

21 Wishart TFL, Flokis M, Shu DY, et al. Hallmarks of lens aging and cataractogenesis[J]. Exp Eye Res, 2021, 210: 108709. DOI: 10.1016/j.exer.2021.108709.

22 Mirtschink P, Jang C, Arany Z, et al. Fructose metabolism, cardiometabolic risk, and the epidemic of coronary artery disease[J]. Eur Heart J, 2018, 39(26): 2497-2505. DOI: 10.1093/eurheartj/ehx518.

23 Fan X, Monnier VM. Protein posttranslational modification (PTM) by glycation: Role in lens aging and age-related cataractogenesis[J]. Exp Eye Res, 2021, 210: 108705. DOI: 10.1016/j.exer.2021.108705.

24 St-Jacques B, Hammerschmidt M, McMahon AP. Indian hedgehog signaling regulates proliferation and differentiation of chondrocytes and is essential for bone formation[J]. Genes Dev, 1999, 13(16): 2072-2086. DOI: 10.1101/gad.13.16.2072.

25 Burnett JB, Lupu FI, Eggenschwiler JT. Proper ciliary assembly is critical for restricting Hedgehog signaling during early eye development in mice[J]. Dev Biol, 2017, 430(1): 32-40. DOI: 10.1016/j.ydbio.2017.07.012.

26 Kerr CL, Huang J, Williams T, et al. Activation of the hedgehog signaling pathway in the developing lens stimulates ectopic FoxE3 expression and disruption in fiber cell differentiation[J]. Invest Ophthalmol Vis Sci, 2012, 53(7): 3316-3330. DOI: 10.1167/iovs.12-9595.

27 Ulrich DL, Lynch J, Wang Y, et al. ATP-dependent mitochondrial porphyrin importer ABCB6 protects against phenylhydrazine toxicity[J]. J Biol Chem, 2012, 287(16): 12679-12690. DOI: 10.1074/jbc.M111.336180.

28 Ahnesorg P, Smith P, Jackson SP. XLF interacts with the XRCC4-DNA ligase IV complex to promote DNA nonhomologous end-joining[J]. Cell, 2006, 124(2): 301-313. DOI: 10.1016/j.cell.2005.12.031.

29 Malaviya PB, Shukal DK, Sharma T, et al. Silibinin suppresses TGFβ2-induced lens epithelial cell migration and epithelial-mesenchymal transition[J]. J Biosci, 2022,47: 51. DOI: 10.1007/s12038-022-00286-3.

30 Patil KK, Gacche RN. Inhibition of glycation and aldose reductase activity using dietary flavonoids: A lens organ culture studies[J]. Int J Biol Macromol, 2017, 98: 730-738. DOI: 10.1016/j.ijbiomac.2017.01.129.

31 Patil KK, Meshram RJ, Barage SH, et al. Dietary flavonoids inhibit the glycation of lens proteins: implications in the management of diabetic cataract[J]. 3 Biotech, 2019, 9(2): 47. DOI: 10.1007/s13205-019-1581-3.

32 Minikel EV, Painter JL, Dong CC, et al. Refining the impact of genetic evidence on clinical success[J]. Nature, 2024, 629(8012): 624-629. DOI: 10.1038/s41586-024-07316-0.

33 Rusina PV, Falaguera MJ, Romero JMR, et al. Genetic support for FDA-approved drugs over the past decade[J]. Nat Rev Drug Discov, 2023, 22(11): 864. DOI: 10.1038/d41573-023-00158-x.

收稿日期：2026 年 08 月 12 日 修回日期：2026 年 09 月 04 日
本文编辑：李绪辉 桂裕亮