

运用孟德尔随机化探索肠道微生物与非酒精性脂肪性肝病的因果关联

吴 闯¹,安祯祥^{1,2},何远利^{1,3},何 松^{1,2},刘 闯^{1,2},孙 凯^{1,2}

摘要 目的 采用孟德尔随机化分析方法探讨肠道微生物与非酒精性脂肪性肝病 (NAFLD) 之间的因果关联。方法 肠道微生物群的遗传工具变量是从 18 340 例参与者的全基因组关联研究确定的,NAFLD 的汇总统计数据来自 FinnGen 数据库,该数据库提供了 894 例 NAFLD 病例的数据和 217 898 例对照,使用 IVW 方法作为主要分析,为了检验结果的稳健性,还采用了 MR-Egger 法、WM 法、Simple Mode 法、Weighted Mode 法进行了孟德尔随机化分析,并进行异质性检验、敏感性分析、多效性分析。结果 丙型变形菌纲 IVW 结果显示 ($OR = 0.621$, $95\% CI = 0.412 \sim 0.934$, $P = 0.022$); 肠杆菌科 IVW 结果显示 ($OR = 1.481$, $95\% CI = 1.069 \sim 2.053$, $P = 0.018$); 毛螺菌科 IVW 结果显示 ($OR = 1.405$, $95\% CI = 1.036 \sim 1.904$, $P = 0.029$); 普氏菌属 7 IVW 结果显示 ($OR = 0.834$, $95\% CI = 0.714 \sim 0.974$, $P = 0.021$); 普氏菌属 9 IVW 结果显示 ($OR = 1.251$, $95\% CI = 1.025 \sim 1.527$, $P = 0.027$); 脱硫弧菌目 IVW 结果显示 ($OR = 0.714$, $95\% CI = 0.519 \sim 0.982$, $P = 0.038$); 肠杆菌目 IVW 结果显示 ($OR = 1.481$, $95\% CI = 1.069 \sim 2.053$, $P = 0.018$)。并且异质性检验不存在异质性,敏感性分析均显示稳健,且未显示多效性。结论 丙型变形菌纲、普氏菌属 7、脱硫弧菌目、肠杆菌科、毛螺菌科、普氏菌属 9、肠杆菌目种类的肠道微生物与 NAFLD 之间可能存在因果关系。

关键词 肠道微生物;非酒精性脂肪性肝病;孟德尔随机化
中图分类号 R 575.5

文献标志码 A **文章编号** 1000-1492(2024)02-0236-07
doi:10.19405/j.cnki.issn1000-1492.2024.02.009

非酒精性脂肪性肝病 (non-alcoholic fatty liver disease, NAFLD) 是临床慢性肝病常见的病因,在没

有过量饮酒或其他慢性肝病的情况下,有 5% 以上肝细胞中有脂肪变性的存在,与代谢风险相关^[1],并且是肝硬化和肝细胞癌的主要原因。NAFLD 给患者带来了沉重的负担,损害了与健康相关的生活质量和工作效率。目前 NAFLD 的发病机制尚不清楚,但肠道微生物组与 NAFLD 发病有着重要关系。当下还没有明确 NAFLD 的治疗方法,有研究^[2]表明,肠道组织可以通过肠-肝轴途径来参与 NAFLD 疾病的发生发展,通过调控肠道微生物改善 NAFLD 患者肝脏表型。尽管肠道微生物群与 NAFLD 密切相关,但是其因果关系难以捉摸。孟德尔随机化 (Mendelian randomization, MR) 是一种统计分析方法,主要采用遗传变异为工具变量,观察关联结果推断潜在的因果关系^[3]。该研究采用孟德尔随机化的思想,探讨肠道微生物不同菌群之间与 NAFLD 风险的潜在因果关系。

1 材料与方法

1.1 研究设计 采用两样本孟德尔随机化分析肠道微生物与 NAFLD 之间的因果关系。以肠道微生物为暴露因素,包含 211 种肠道细菌性状,结局变量为 NAFLD。

1.2 数据来源 使用的肠道微生物数据来自于一项多个国家的全基因组关联研究 (genome-wide association study, GWAS), 该研究纳入了 24 个队列, 18 340 例受试者,分别来自美国、加拿大、以色列、荷兰、比利时、瑞典、韩国、德国、丹麦、芬兰和英国,对他们进行基因测序谱和基因分型数据,通过调整年龄、性别、技术及遗传主成分等探索肠道微生物与人类染色体遗传变异之间的关系,课题组在 5 个水平 (门、纲、目、科和属) 评估了 211 个肠道微生物群分类群。该研究基于 16S rRNA 基因的 3 个不同可变区 (V4、V3-V4 和 V1-V2) 分析了肠道微生物群的组成,并使用微生物群数量性状基因座 (microbiome quantitative trait loci, mbQTL) 作图来鉴定影响微生物分类群相对丰度的遗传变体^[4]。NAFLD 基因结果关联的数据来自 FinnGen 数据库,包括 894 例

2023-12-29 接收

基金项目:国家自然科学基金 (编号:82160893); 国家中医药管理局第五批全国中医临床优秀人才研修项目 (编号:国中医药人教函[2022]1 号); 贵州中医药大学中医脾胃病科技创新人才团队建设 (编号:贵中医 TD[2022]005 号)

作者单位:¹ 贵州中医药大学第一临床医学院,贵阳 550001

贵州中医药大学第一附属医院² 消化内科、³ 老年病科,贵阳 550001

作者简介:吴 闯,女,博士研究生;

安祯祥,男,博士,主任医师,教授,博士生导师,责任作者,
E-mail: 407206115@qq.com

表 1 全基因组关联研究的细节与分析中使用的数据集

暴露或结局	样本量	人口	数据下载地址
肠道微生物	18 340 例参与	混合	https://mibiogen.gcc.rug.nl
非酒精性脂肪性肝病	894 例阳性, 217 898 例对照	欧洲	https://storage.googleapis.com/finngen-public-data-r9/summary_stats/finngen_R9_NAFLD.gz

NAFLD 病例, 217 898 例对照。使用数据集来源详见表 1。

1.3 两样本孟德尔随机化的假设 为了保证暴露因素与结局指标之间因果关系的真实性和准确性, MR 必须符合 3 个重要假设: ① 工具变量 (single nucleotide polymorphism, SNP) 与暴露因素肠道菌群强相关联。② 工具变量 SNP 与结局变量 NAFLD 之间和其他混杂因素没有关联, 即排除如肥胖、体重指数 (body mass index, BMI)、自身免疫性疾病等混杂因素。③ 工具变量 SNP 仅通过影响暴露因素肠道微生物对结局变量 NAFLD 产生影响, 不能通过其他途径对 NAFLD 造成影响。F 统计量一般用于评估工具变量与暴露之间的相关性强度。F 统计量的计算公式为:

$$F = \frac{R^2(n-k-1)}{k(1-R^2)}$$

F 应 > 10, 如果 F < 10, 则说明工具变量与暴露之间的关联较弱。

1.4 工具变量 选择与肠道微生物组明显相关的 SNPs 作为工具变量, 参照千个全基因组项目的全基因组信息, 将全基因组统计显著性阈值参数设置为 $P < 5 \times 10^{-8}$ 。将连锁不平衡参数设置为 $r^2 < 0.001$, 聚集距离 = 10 000 kb, 从肠道微生物数据中挑选出无连锁效应的 SNP, 再从已选出的 SNP 中剔除与 NAFLD 有意义的 SNP ($P < 0.05$), 系统收集 SNPs 的主要信息, 包括效应等位基因、其他等位基因、 β 、SE 和 P 值, 用于进一步分析^[5]。

1.5 孟德尔随机化分析 孟德尔随机化方法是一种基于遗传变异的随机化方法, 用于评估因果关系。它利用自然界中存在的基因变异, 将个体分为不同的基因型, 从而模拟随机化试验的效果, 以评估某个因素对某种疾病的影响是否具有因果性。这种方法可以避免传统流行病学研究中的混杂因素和反向因果关系, 从而提高因果推断的可信度。常使用包括逆方差加权 (inverse-variance weighted, IVW)、MR-Egger、加权中位数法 (weighted median, WM)、Simple Mode 法、Weighted Mode 法在内的高效统计方法来推断人类肠道微生物组对 NAFLD 风险是否存在因果效应。使用 IVW 方法作为主要分析, 因为它提

供了最精确的效应估计, 几乎所有 MR 分析都将其作为主要分析。采用 MR-Egger 回归检验潜在的水平多效性, 如果截距的 $P < 0.05$, 则可能存在 SNP 的水平多效性。MR-PRESSO 检测多效性偏差的异常值并纠正水平多效性。此外, Cochran 的 Q 统计量用于量化所选 SNP 之间的异质性。为确定是否存在潜在的强影响 SNP, 采用留一敏感度分析法验证因果效应估计的可靠性和稳定性。使用 R 软件 (版本 4.0.2, TwoSampleMR、MR-PRESSO 包) 进行统计分析。

2 结果

2.1 孟德尔随机化分析结果 IVW 分析结果表明, 丙型变形菌纲 ($OR = 0.621, 95\% CI = 0.412 \sim 0.934, P = 0.022$)、普氏菌属 7 ($OR = 0.834, 95\% CI = 0.714 \sim 0.974, P = 0.021$)、脱硫弧菌目 ($OR = 0.714, 95\% CI = 0.519 \sim 0.982, P = 0.038$) 3 种肠道细菌性状与 NAFLD 风险负相关。肠杆菌科 ($OR = 1.481, 95\% CI = 1.069 \sim 2.053, P = 0.018$)、毛螺菌科 ($OR = 1.405, 95\% CI = 1.036 \sim 1.904, P = 0.029$)、普氏菌属 9 ($OR = 1.251, 95\% CI = 1.025 \sim 1.527, P = 0.027$)、肠杆菌目 ($OR = 1.481, 95\% CI = 1.069 \sim 2.053, P = 0.018$) 4 种肠道细菌性状与 NAFLD 风险正相关。并且筛选出来 SNP 的 F 均 > 10, 说明工具变量与暴露之间存在较强的关联。考虑到 NAFLD 易受生活方式、饮酒的影响, 本实验使用 PhenoScanner 查询了阳性结果的 SNP, 并排除了与上述混杂因素相关的 SNP。详见表 2。IVW 结果森林图见图 1。

2.2 敏感性分析 Cochran Q 检验结果显示 7 种肠道细菌性状 Cochran Q 检验 P 均 > 0.05, 提示 SNP 间不存在异质性; 多效性分析 MR-Egger 结果显示 7 种肠道细菌性状的 P 值均 > 0.05, 因此, 以上肠道细菌性状不存在多效性水平的情况。本研究的敏感性分析采用 Leave-one-out 检验并绘制图, 留一图表明在去除其中任何一个 SNP, 剩下的 SNPs 均在无效线的同一侧, 去掉任何一个 SNP 后对结果产生的影响均较小, 则提示该 MR 研究结果比较稳健。散点图则提示 7 种肠道菌群不同的统计方法的趋势基本

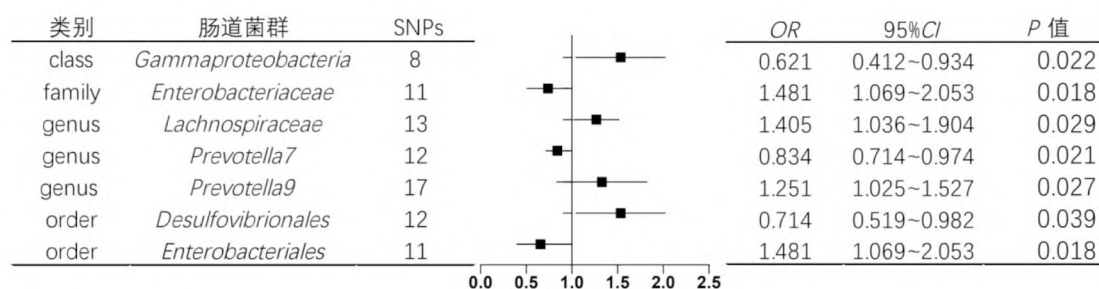


图1 IVW 结果森林图

表2 孟德尔随机化分析主要结果

暴露	SNPs	方法	OR (95% CI)	P 值
class <i>ammaproteobacteria</i>	8	IVW	0.621 (0.412 ~ 0.934)	0.022
		MR Egger	0.648 (0.173 ~ 2.431)	0.544
		Simple Mode	0.731 (0.37 ~ 1.446)	0.398
		Weighted Median	0.71 (0.421 ~ 1.199)	0.200
		Weighted Mode	0.731 (0.358 ~ 1.494)	0.419
family <i>Enterobacteriaceae</i>	11	IVW	1.418 (1.069 ~ 2.053)	0.018
		MR Egger	2.036 (0.456 ~ 9.09)	0.376
		Simple Mode	1.092 (0.503 ~ 2.37)	0.829
		Weighted Median	1.343 (0.825 ~ 2.116)	0.204
		Weighted Mode	1.058 (0.495 ~ 2.263)	0.888
genus <i>Lachnospiraceae</i>	13	IVW	1.405 (1.036 ~ 1.904)	0.029
		MR Egger	2.677 (0.776 ~ 9.23)	0.147
		Simple Mode	1.935 (0.826 ~ 4.531)	0.154
		Weighted Median	1.578 (1.044 ~ 2.383)	0.030
		Weighted Mode	1.935 (0.906 ~ 4.135)	0.114
genus <i>Prevotella7</i>	12	IVW	0.834 (0.714 ~ 0.974)	0.021
		MR Egger	0.607 (0.237 ~ 1.554)	0.322
		Simple Mode	0.827 (0.567 ~ 1.207)	0.346
		Weighted Median	0.848 (0.688 ~ 1.045)	0.122
		Weighted Mode	0.833 (0.593 ~ 1.169)	0.313
genus <i>Prevotella9</i>	17	IVW	1.251 (1.025 ~ 1.527)	0.027
		MR Egger	1.271 (0.731 ~ 2.21)	0.410
		Simple Mode	1.526 (0.906 ~ 2.571)	0.132
		Weighted Median	1.238 (0.941 ~ 1.628)	0.127
		Weighted Mode	1.436 (0.879 ~ 2.345)	0.168
order <i>Desulfovibrionales</i>	12	IVW	0.714 (0.519 ~ 0.982)	0.039
		MR Egger	0.412 (0.101 ~ 1.687)	0.246
		Simple Mode	0.741 (0.352 ~ 1.557)	0.445
		Weighted Median	0.751 (0.488 ~ 1.155)	0.192
		Weighted Mode	0.774 (0.386 ~ 1.552)	0.485
order <i>Enterobacteriales</i>	11	IVW	1.481 (1.069 ~ 2.053)	0.018
		MR Egger	2.036 (0.456 ~ 9.09)	0.376
		Simple Mode	1.092 (0.521 ~ 2.286)	0.821
		Weighted Median	1.343 (0.861 ~ 2.095)	0.194
		Weighted Mode	1.058 (0.54 ~ 2.071)	0.873

一致,说明其偏差较小。漏斗图显示 IVW 左右两边的点大致对称,说明发病偏移较小。敏感性分析见表3,留一图见图2,散点图见图3,漏斗图见图4。

3 讨论

本研究采用两样本孟德尔随机化方法对 211 种

表3 敏感性分析

暴露 id	Q_P 值	MR-Egger_P 值	MRPRESSO	肠道微生物群
ebi-a-GCST90016917	0.918 956	0.948 292	0.937	class Gammaproteobacteria
ebi-a-GCST90016936	0.529 943	0.679 311	0.541	family Enterobacteriaceae
ebi-a-GCST90017026	0.444 157	0.314 666	0.454	genus Lachnospiraceae
ebi-a-GCST90017044	0.491 933	0.517 036	0.538	genus Prevotella7
ebi-a-GCST90017045	0.542 337	0.954 148	0.575	genus Prevotella9
ebi-a-GCST90017097	0.743 414	0.450 961	0.755	order Desulfovibrionales
ebi-a-GCST90017098	0.529 943	0.679 311	0.569	order Enterobacteriales

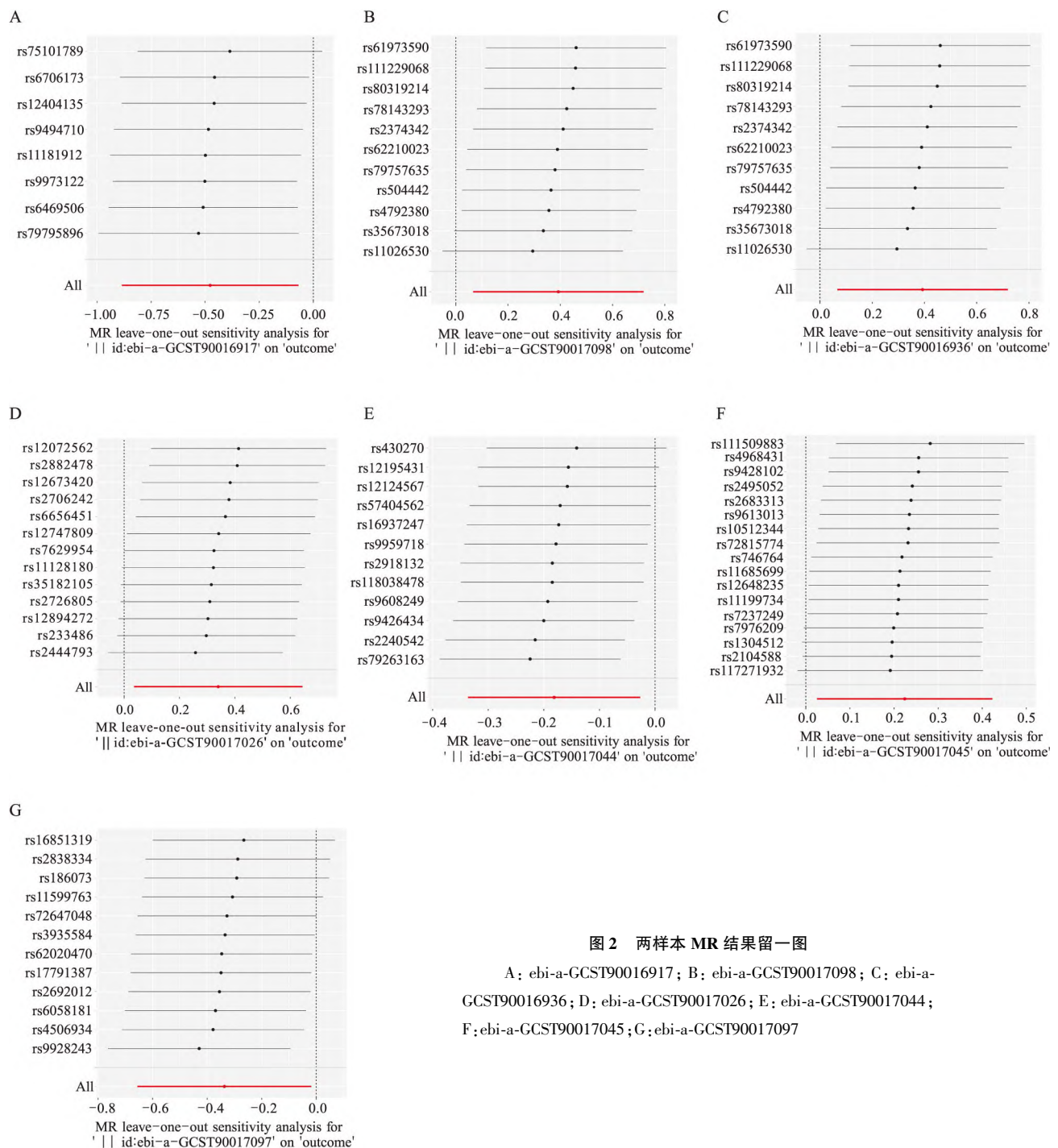


图2 两样本 MR 结果留一图

A: ebi-a-GCST90016917; B: ebi-a-GCST90017098; C: ebi-a-GCST90016936; D: ebi-a-GCST90017026; E: ebi-a-GCST90017044; F: ebi-a-GCST90017045; G: ebi-a-GCST90017097

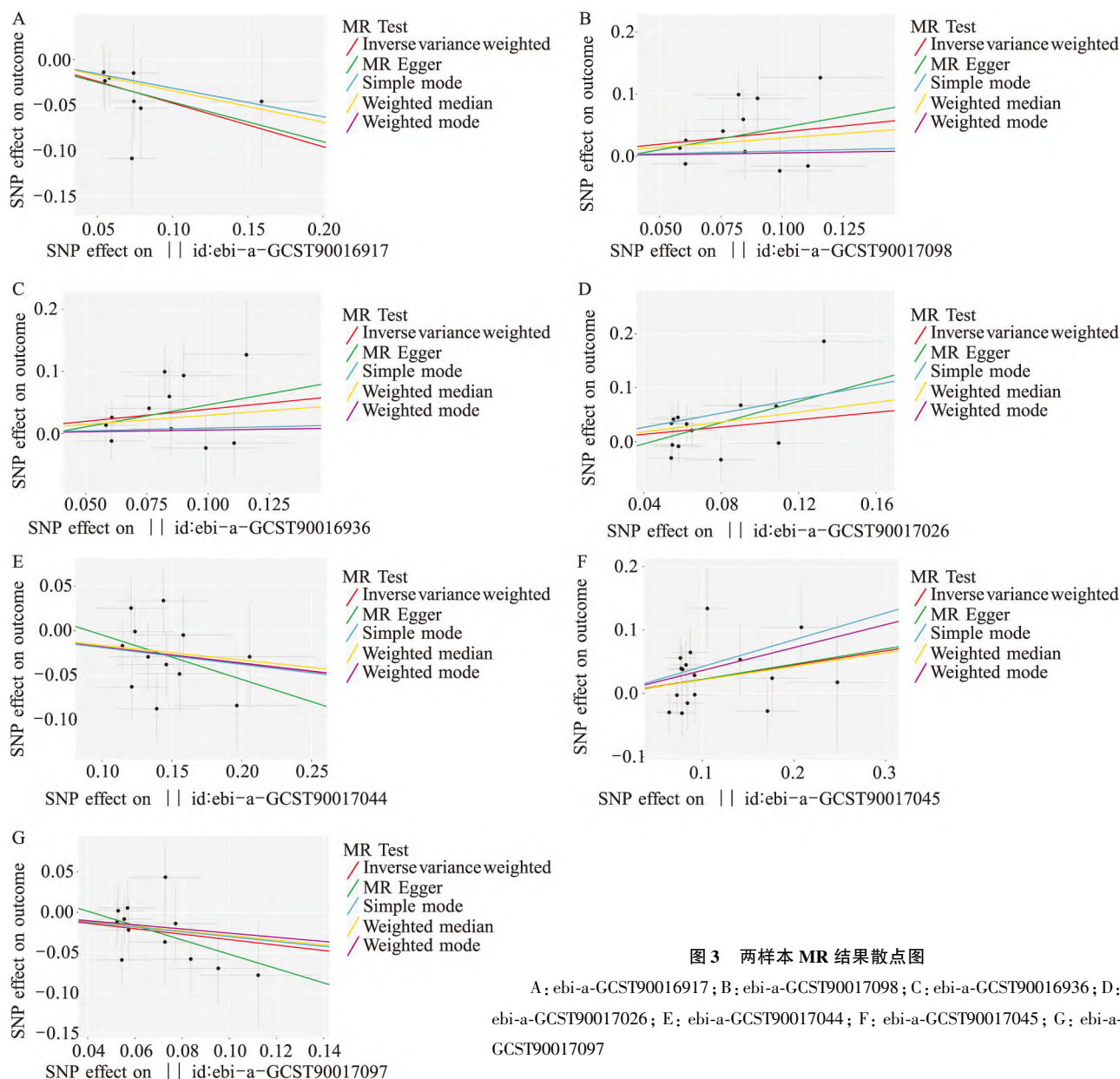


图3 两样本MR结果散点图

A: ebi-a-GCST90016917; B: ebi-a-GCST90017098; C: ebi-a-GCST90016936; D: ebi-a-GCST90017026; E: ebi-a-GCST90017044; F: ebi-a-GCST90017045; G: ebi-a-GCST90017097

肠道细菌性状与NAFLD的GWAS数据进行因果关联分析,共确定了7种细菌性状与NAFLD存在因果关系,其中丙型变形菌纲、普氏菌属7、脱硫弧菌目3种细菌性状与NAFLD的风险成负相关,肠杆菌科、毛螺菌科、普氏菌属9、肠杆菌目4种肠道细菌性状是NAFLD发病的风险因素。

人类身体的胃肠道黏膜表面被数以万计的肠道微生物所定植,这些肠道微生物组参与宿主免疫系统的调节和维持。因肠道微生物群的差异,涉及了免疫、能量、脂质和葡萄糖代谢等多种途径,并参与了全身多种疾病的发生。近年来肠道微生物组在NAFLD中得到了广泛的研究,有证据^[6]表明,肠道微生物组-肝脏轴在NAFLD中发生作用,尤其是在

纤维化和进展到晚期的疾病阶段,并且该研究证明肠道微生物组中肠杆菌科的改变会影响NAFLD疾病的发展。

丙型变形菌纲是肠道微生物主要类之一,丙型变形菌纲与NAFLD等疾病有关^[7],并且在NAFLD中丙型变形菌纲丰度增加。普氏菌属是常见且丰富的专性厌氧菌,该属的成员属于胃肠道和呼吸道的微生物群落,并与糖尿病、心血管疾病和NAFLD等疾病高度相关,其共同特征是低度全身炎症^[8]。在非糖尿病队列中^[9],普氏菌属丰度增加可能与胰岛素抵抗有关。有研究^[7]指出,患有NAFLD的儿童携带的普氏菌属丰度较健康儿童高。并且使用富含普氏菌属的生态失调微生物群从具有缺陷的炎症小

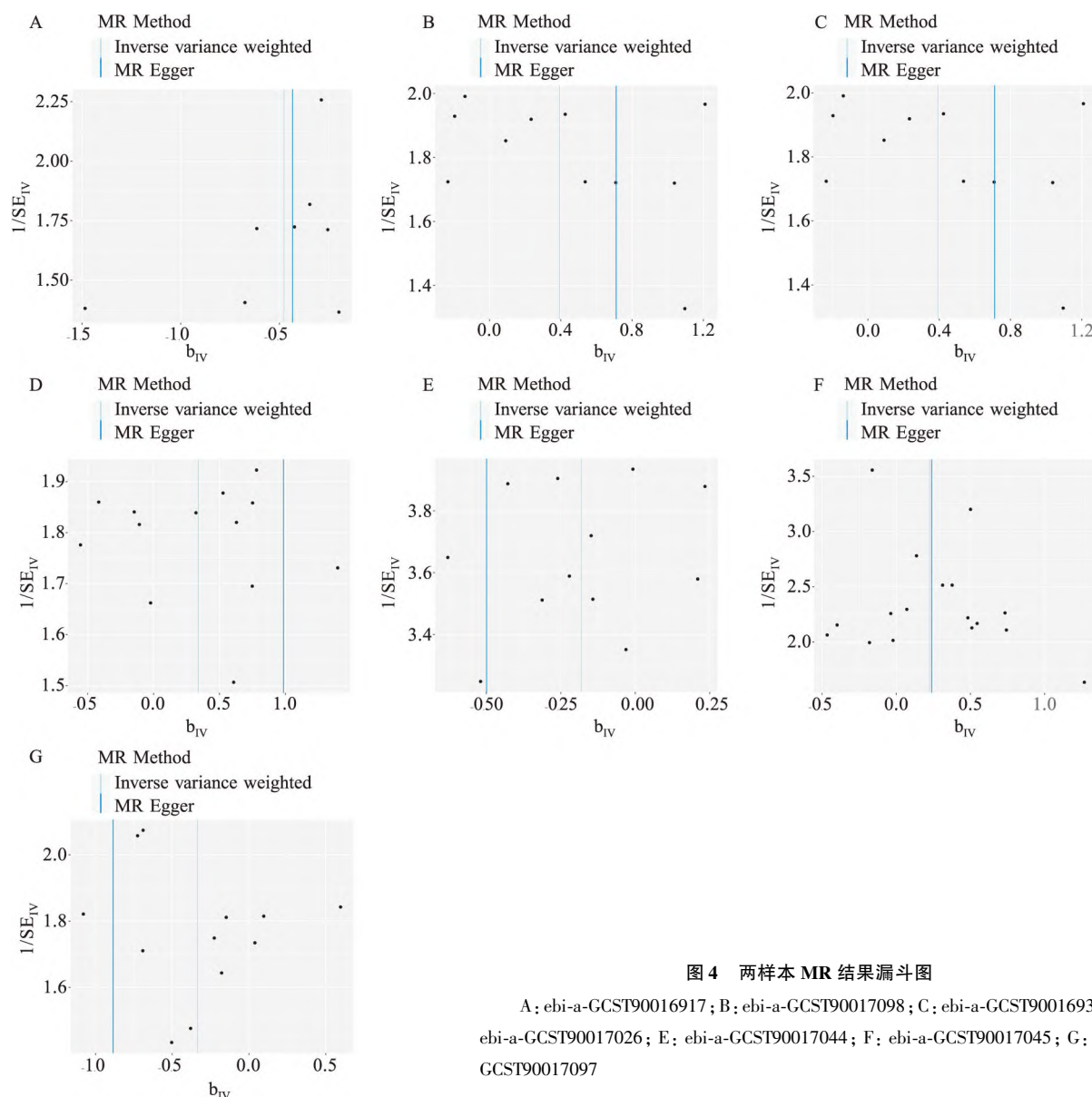


图4 两样本 MR 结果漏斗图

A: ebi-a-GCST90016917; B: ebi-a-GCST90017098; C: ebi-a-GCST90016936; D: ebi-a-GCST90017026; E: ebi-a-GCST90017044; F: ebi-a-GCST90017045; G: ebi-a-GCST90017097

体途径的小鼠转移到野生型小鼠中,研究富含普氏菌属的生态失调在 NAFLD 和肥胖中的作用,且该菌还加剧了蛋氨酸胆碱缺乏型饮食诱导的非酒精性脂肪性肝炎,使肝脏脂肪变性和炎症增加以及肝功异常^[10]。而另一项研究^[11]则提示普氏菌属丰度水平的降低可能对患有 NAFLD 的成年人有害,但毛螺菌科水平升高可能是 NAFLD 疾病进展的主要因素。肠道微生物中肠杆菌科是变形杆菌门的重要成员,该菌群包括许多已知在哺乳动物的大肠和小肠中定植,并且在 NAFLD 及结直肠癌中该细菌水平升高^[12]。脱硫弧菌目菌多与神经类疾病^[13]、消化系统疾病密切相关^[14]。而在 NAFLD 疾病中脱硫弧菌目菌与瘦的 NAFLD 患者成负相关,与肥胖 NAFLD 患者呈正相关,两者表现出相反的趋势,但与总

NAFLD 呈负相关^[15]。这说明肠道微生物群与疾病的发生密切相关,尤其是肠内外疾病,但其对肠内外疾病产生有益或有害作用的机制仍有待进一步研究。

虽然本研究分析了肠道微生物群,但肠道微生物群太大且异质,因此数据有限;其次,本研究虽然单独分析了肠道微生物群与 NAFLD 之间的遗传因果关系,但未进一步探索肠道微生物与 NAFLD 之间的生物学关系。

参考文献

- [1] Powell E E, Wong V W, Rinella M. Non-alcoholic fatty liver disease [J]. Lancet. 2021, 397(10290): 2212-24.
- [2] 邵 丽, 宋 育, 施军平. 肠-肝免疫轴在非酒精性脂肪性肝炎发病中的作用[J]. 中华肝脏病杂志, 2021, 29(6): 505

- 9.
- [3] 刘明, 高亚, 杨珂璐, 等. 孟德尔随机化研究的报告规范 (STROBE-MR) 解读[J]. 中国循证医学杂志, 2022, 22(8): 978-87.
- [4] Kurilshikov A, Medina-Gomez C, Bacigalupe R, et al. Large-scale association analyses identify host factors influencing human gut microbiome composition [J]. *Nat Genet*, 2021, 53(2): 156-65.
- [5] Cheng Q, Yang Y, Shi X, et al. MR-LDP: a two-sample Mendelian randomization for GWAS summary statistics accounting for linkage disequilibrium and horizontal pleiotropy [J]. *NAR Genom Bioinform*, 2020, 2(2): lqaa028.
- [6] Loomba R, Seguritan V, Li W, et al. Gut microbiome-based metagenomic signature for non-invasive detection of advanced fibrosis in human nonalcoholic fatty liver disease [J]. *Cell Metab*, 2019, 30(3): 607.
- [7] Michail S, Lin M, Frey M R, et al. Altered gut microbial energy and metabolism in children with non-alcoholic fatty liver disease [J]. *FEMS Microbiol Ecol*, 2015, 91(2): 1-9.
- [8] Larsen J M. The immune response to Prevotella bacteria in chronic inflammatory disease [J]. *Immunology*, 2017, 151(4): 363-74.
- [9] Pedersen H K, Gudmundsdottir V, Nielsen H B, et al. Human gut microbes impact host serum metabolome and insulin sensitivity [J]. *Nature*, 2016, 535(7612): 376-81.
- [10] Henao-Mejia J, Elinav E, Jin C, et al. Inflammasome-mediated dysbiosis regulates progression of NAFLD and obesity [J]. *Nature*, 2012, 482(7384): 179-85.
- [11] Shen F, Zheng R D, Sun X Q, et al. Gut microbiota dysbiosis in patients with non-alcoholic fatty liver disease [J]. *Hepatobiliary Pancreat Dis Int*, 2017, 16(4): 375-81.
- [12] Liu C J, Zhang Y L, Shang Y, et al. Intestinal bacteria detected in cancer and adjacent tissue from patients with colorectal cancer [J]. *Oncol Lett*, 2019, 17(1): 1115-27.
- [13] Murros K E, Huynh V A, Takala T M, et al. *Desulfovibrio* bacteria are associated with parkinson's disease [J]. *Front Cell Infect Microbiol*, 2021, 11: 652617.
- [14] Xiong Y, Wu L, Shao L, et al. Dynamic alterations of the gut microbial pyrimidine and purine metabolism in the development of liver cirrhosis [J]. *Front Mol Biosci*, 2022, 8: 811399.
- [15] Yun Y, Kim H N, Lee E J, et al. Fecal and blood microbiota profiles and presence of nonalcoholic fatty liver disease in obese versus lean subjects [J]. *PLoS One*, 2019, 14(3): e0213692.

Using Mendelian randomization to explore the causal association between gut microbiota and nonalcoholic fatty liver disease

Wu Min¹, An Zhenxiang^{1,2}, He Yuanli^{1,3}, He Song^{1,2}, Liu Chuang^{1,2}, Sun Kai^{1,2}

(¹The First Clinical Medical College of Guizhou University of Traditional Chinese Medicine, Guiyang 550001; ²Dept of Gastroenterology, ³Dept of Geriatrics, The First Affiliated Hospital of Guizhou University of Traditional Chinese Medicine, Guiyang 550001)

Abstract *Objective* To explore the causal association between gut microbes and non-alcoholic fatty liver disease (NAFLD) by Mendelian randomisation analysis. *Methods* Genetic instrumental variables for gut microbiota were identified from a gene-wide association study of 18 340 participants, and summary statistics for NAFLD were obtained from the FinnGen database, which provided data on 894 NAFLD cases and 217 898 controls using the IVW method as the primary analysis. In order to test the robustness of the results, MR-Egger method, WM method, Simple Mode method, Weighted Mode method were used for Mendelian randomisation analysis, and heterogeneity test, sensitivity analysis, and multiplicity analysis were performed. *Results* class Gammaproteobacteria IVW results showed ($OR = 0.621$, 95% $CI = 0.412 \sim 0.934$, $P = 0.022$); family Enterobacteriaceae IVW results showed ($OR = 1.481$, 95% $CI = 1.069 \sim 2.053$, $P = 0.018$); genus Lachnospiraceae IVW results showed ($OR = 1.405$, 95% $CI = 1.036 \sim 1.904$, $P = 0.029$); genus Prevotella7 IVW results showed ($OR = 0.834$, 95% $CI = 0.714 \sim 0.974$, $P = 0.021$); genus Prevotella9 IVW results showed ($OR = 1.251$, 95% $CI = 1.025 \sim 1.527$, $P = 0.027$); order Desulfovibrionales IVW results showed ($OR = 0.714$, 95% $CI = 0.519 \sim 0.982$, $P = 0.038$); order Enterobacteriales IVW results showed ($OR = 1.481$, 95% $CI = 1.069 \sim 2.053$, $P = 0.018$). And there was no heterogeneity in the heterogeneity test, and the sensitivity analyses all showed robustness and no pleiotropy was found. *Conclusion* This study implicates class Gammaproteobacteria, family Enterobacteriaceae, genus Lachnospiraceae, genus Prevotella7, genus Prevotella9, order Desulfovibrionales, order Enterobacteriales seven species of gut microorganisms have a causal relationship with NAFLD.

Key words gut microbiota; non-alcoholic fatty liver disease; Mendelian randomization