DOI: 10.16016/j.1000-5404.202101018

腹主动脉瘤 MTHFR-MMP2 基因-基因交互作用研究

王斯悦[1],秦雪英[1],隗英琦[2],左尚维[3],陈泓伯[1],王梦莹[1],吴 瑶[1],吴俊慧[1],王小文[1],王紫荆[1],王伽婷[1],于 欢[1],武轶群[1],吴 涛[1],胡永华[1] 100191 北京,北京大学公共卫生学院流行病与卫生统计学系[1]; 100000 北京,北京市疾病预防控制中心[2];100853 北京,中国人民解放军总医院血管外科[3]

[摘要] **目的** 探索 MTHFR-MMP2 基因-基因交互作用是否影响患腹主动脉瘤(abdominal aortic aneurysm, AAA)的风险。**方法** 基于 1∶2匹配病例对照研究设计,以 2011-2014 年解放军总医院血管外科的 AAA 患者、同期同一医院及北京市房山区两地性别、年龄匹配的非 AAA 患者为研究对象;采用标准化问卷收集其基本人口学资料、病史及行为危险因素;开展体格检查、腹主动脉超声,并采集 5 mL 外周静脉血液标本对基因分型。采用传统及机器学习方法,包括条件 Logistic 回归、广义多因子降维(generalized multifactor dimension reduction, GMDR)、非监督式稀疏因子分析的机器学习方法(EPIstasis sparse factor analysis, EPISFA),分别在相加交互作用、相乘交互作用尺度上,探索 MTHFR 与 MMP2 基因-基因交互作用是否影响患腹主动脉瘤风险,并采用交互作用超额危险度(relative excess risk of interaction, RERI)、交互作用归因比(attributable proportion of interaction, API)、交互作用指数(the synergy index, S)3 种指标评价 MTHFR 与 MMP2 间的基因-基因交互作用的方向和大小。**结果** 2011-2014 年共纳入 155 组 1∶2 的匹配病例对照(共 465 人)。在调整协变量后,同时在乘法和加法尺度下发现 MTHFR 基因与 MMP2 基因上多个位点(rs1132896, rs1477017, rs1992116, rs243847)可能存在拮抗交互作用($P<0.05$),影响 AAA 风险,并在 3 种分析方法中彼此验证;在高维交互作用上,GMDR 和 EPISFA 提示 MTHFR-MMP2 间可能存在高维交互作用,具有较高的交叉验证一致性,且结果具有统计学差异($P<0.05$)。**结论** MTHFR-MMP2 基因间可能存在拮抗基因-基因交互作用影响患 AAA 风险。

[关键词] 腹主动脉瘤;基因-基因交互作用;匹配病例对照;交互作用指数;相乘模型;相加模型
[中图法分类号] R181.3;R394.3;R654.3 [文献标志码] A

MTHFR-MMP2 gene-gene interactions may increase the risk of abdominal aortic aneurysm

WANG Siyue[1], QIN Xueying[1], WEI Yingqi[2], ZUO Shangwei[3], CHEN Hongbo[1], WANG Mengying[1], WU Yao[1], WU Junhui[1], WANG Xiaowen[1], WANG Zijing[1], WANG Jiating[1], YU Huan[1], WU Yiqun[1], WU Tao[1], HU Yonghua[1]
[1]Department of Epidemiology and Biostatistics, School of Public Health, Peking University, Beijing, 100191; [2]Beijing Center for Disease Control and Prevention, Beijing, 100000; [3]Department of Vascular Surgery, Chinese PLA General Hospital, Beijing, 100853, China

[**Abstract**] **Objective** To explore the role of genetic interactions of MTHFR-MMP2 in risks for abdominal aortic aneurysm (AAA). **Methods** A 1∶2 matched case-control study was performed on the AAA patients admitted in the vascular surgery department of Chinese PLA General Hospital between 2011 and 2014, and on the non-AAA patients (control group) with matched gender and age from both the above hospital and a community of Fangshan district during the same period. Standardized questionnaires were used to collect information such as demographic characteristics, medical history and behavioral risk factors.

[通信作者] 胡永华,E-mail:yhhu@bjmu.edu.cn

1000-5404/Copyright © 2021, 陆军军医大学(第三军医大学)。这篇开放获取文章遵循 CC BY 许可协议(https://creativecommons.org/licenses/by/4.0/)

Physical and abdominal aortic ultrasound examinations were carried out, and 5 mL peripheral venous blood samples were collected for genotyping. The interactions between MTHFR and MMP2 genes were investigated using both traditional and machine learning methods, including conditional Logistic regression model, generalized multifactor dimension reduction (GMDR) and EPI stasis sparse factor analysis (EPISFA). Whether the genetic interactions affects the risk of AAA was analyzed and determined, on additive and multiplicative scales, respectively. In addition, relative excess risk of interaction (RERI), attributable proportion of interaction (API) and synergy index (S) were adopted to assess the direction and magnitude of the genetic interactions between MTHFR and MMP2. **Results** A total of 465 individuals, including 155 AAA patients and 310 non-AAA subjects at a matched ratio of 1∶2 were enrolled in this study during 2011 and 2014. After adjustment of covariates, antagonistic interactions between MTHFR gene and multiple loci of MMP2 gene (rs1132896, rs1477017, rs1992116, rs243847) were found on both additive and multiplicative scales simultaneously ($P<0.05$), which could affect AAA risks. And the results were verified by the above 3 methods. GMDR and EPISFA results suggested that there might be high-dimensional interactions between MTHFR and MMP2, with high cross-validation consistency and significant difference ($P<0.05$). **Conclusion** There may be antagonistic gene-gene interactions between MTHFR and MMP2, and the interactions may potentially influence the risks for AAA.

[**Key words**] abdominal aortic aneurysm; gene-gene interaction; matched case-control; synergy index; multiplicative model; additive model

Corresponding author: HU Yonghua, E-mail: yhhu@bjmu.edu.cn

腹主动脉瘤(abdominal aortic aneurysm, AAA)是血管外科常见的大动脉疾病,其定义是由于动脉壁的损伤或病变,引起动脉管壁永久性局限性扩张>3.0 cm或超过正常血管直径的50%[1]。AAA严重威胁患者生命安全,患者腹主动脉瘤破裂发生后病死率可达90%[2-3],是心血管疾病的重要死因。腹主动脉瘤是受遗传和环境共同作用影响的复杂疾病[4-5];据双生子研究估计其遗传度可达70%[6]。截至目前,多项全基因组关联研究(genome-wide association studies, GWASs)发现了24个与AAA风险有关的遗传位点或区域[7-10],但这些位点仅能解释极少部分遗传度[11]。对于这部分"缺失的遗传度",开展基因-基因交互作用有利于更好地解释复杂疾病的遗传风险,且有利于深入探索复杂疾病遗传病因的机制[12]。

MTHFR与MMP2基因是影响腹主动脉瘤风险的重要候选基因[13-15,18]。然而,既往研究多采用实验研究探索两基因在AAA中的作用,或仅关注两基因的独立主效应探索,而基于人群研究的MTHFR-MMP2基因-基因交互作用证据较少[11]。因此,基于人群研究探索MTHFR-MMP2基因-基因交互作用与AAA风险的关联具有重要价值。近年来候选基因研究证据提示亚甲基四氢叶酸还原酶基因(methylenetetrahydrofolate reductase, MTHFR)单核苷酸多态性(single nucleotide polymorphism, SNP)与AAA的疾病风险有关[13-15],但不同人群研究间证据仍存在不一致[16-17]。由于MTHFR C677T在AAA发生及进展中多参与信号通路的上调或下调,与其他基因存在着协同或拮抗的交互作用,因此探索MTHFR C677T的基因-基因交互作用为深入了解AAA疾病病因具有重要意义。

MMP家族基因多态性参与调节基质金属蛋白酶,其作为生物活性分子,在细胞增殖,迁移和凋亡中具有重要作用[18],其中以MMP2基因多态性与AAA风险的关联研究证据较为充分[19]。实验研究提示MMP家族基因与MTHFR C677T可能协同参与了Hcy水平的调控[20]。

近年来基因-基因交互作用的方法不断发展,除Logistic回归外,广义多因子降维(generalized multifactor dimension reduction, GMDR)[21]、非监督式稀疏因子分析的机器学习方法(EPIstasis sparse factor analysis, EPISFA)[12]等机器学习方法的发展有效补充了Logistic回归模型在计算高维交互作用时的效能不足,同时为探索二维及高维复杂遗传交互作用和不同分析方法结果的验证提供了基础。

因此,本研究采用匹配病例-对照研究设计,采用多种遗传交互作用分析方法(条件Logistic回归、GMDR、EPISFA),分别在相加交互作用、相乘交互作用尺度上,探索MTHFR与MMP2间是否存在二维及高维基因-基因交互作用影响腹主动脉瘤风险,并基于多种交互作用指标评价MTHFR与MMP2间的基因-基因交互作用的方向和大小。

1 资料与方法

1.1 研究设计

本研究设计为1:2个体匹配的病例对照研究,以性别和年龄作为匹配因素。本研究于2011年7月至2014年12月开展,以解放军总医院为基础序贯纳入AAA病例组与医院对照组研究对象,同时以北京市房山区某社区为基础,纳入社区对照组研究对象。根据性别和年龄匹配的原则,每例AAA患者与来自医院和社区的各一名对照组研究对象匹配形成一个对子。更详细的研究设计请参见左尚维等[1]的研究。所有研究对象包括患者和对照者均签署知情同意书。项目已通过中国人民解放军总医院生物医学伦理委员会审查批准(伦理审批批号:2011013)。

1.2 研究对象

研究病例组、医院对照组及社区对照组的纳入排除标准分别如下。病例组序贯纳入解放军总医院血管外科AAA患者,入选标准:(1)年龄大于40岁;(2)经腹主动脉超声或腹部计算机断层扫描(CT)检查确诊为肾下腹主动脉瘤;(3)调查资料可完整得到并签署知情同意书。排除标准:(1)诊断为感染性腹主动脉瘤;(2)诊断为继发性腹主动脉瘤(如外伤所致);(3)有严重精神疾病,难以配合调查和体检;(4)有严重感染性或自身免疫性疾病;(5)孕妇、恶性肿瘤患者。

医院对照组(对照组1)与社区对照组(对照组2)研究对象采用统一的纳入排除标准,入选标准:(1)与病例组患者性别相匹配,年龄相差3岁以内;(2)经腹主动脉超声或腹部CT检查,未患任何部位、任何原因引起的腹主动脉瘤;(3)调查资料可完整得到并签署知情同意书。排除标准:(1)患有主动脉夹层、胸主动脉瘤等主动脉疾病;(2)患有症状性外周动 脉闭塞;(3)严重感染性或自身免疫性疾病患者;(4)身体或精神有重大缺陷,不能配合问卷和抽血检测者;(5)孕妇、恶性肿瘤患者。

1.2 研究方法

1.2.1 问卷调查及一般体格检查 研究对象纳入后收集如下资料及血样,包括:问卷调查信息、一般体格检查、影像学动脉直径测量、血清检测并采集5 mL外周静脉血用于基因分型。各项资料均由解放军总医院临床专业人员经过统一培训后组织收集。数据经过严格的质控后纳入分析。问卷包括一般人口学特征、慢性病史及治疗史、行为因素等。吸烟定义为总吸烟量超过100支;饮酒定义为规律饮酒至少半年且每次酒精摄入量大于50 mL白酒酒精含量。标准化体格检查主要包括测量身高、体质量、腰臀围、血压、腹主动脉超声或腹部CT等;清晨空腹采集外周静脉血用以进行血清学检测,指标包括血清葡萄糖(serum glucose,GLU)、甘油三酯(triglyce- rides,TG)、总胆固醇(total cholesterol,TC)、高密度脂蛋白胆固醇(high-density lipoprotein cholesterol ester,HDL-C)、低密度脂蛋白胆固醇(low-density lipo- protein cholesterol ester,LDL-C)、载脂蛋白A1(apoli- poprotein A1,ApoA1)、载脂蛋白B(apolipoprotein B,ApoB)、超敏C反应蛋白(high-sensitivity C-reactive protein,hsCRP)、同型半胱氨酸(homocysteine,Hcy)。

1.2.2 基因型检测及遗传位点选择 采用血凝块标本应用蛋白质沉淀法提取基因组DNA。简要操作步骤为:捣碎血凝块并去除红细胞碎片;裂解白细胞;在蛋白酶K的作用下将染色体蛋白质分离并沉淀;应用异丙醇析出DNA;冷无水乙醇清洗DNA及并将DNA溶解于200 μL乙二胺四乙酸二钠钙溶液制成DNA原液。采用紫外分光光度法测定DNA原液浓度和纯度。取2 μL DNA原液,用UV7501型紫外分光光度仪测定样品在260 nm和280 nm的光密度值并读取浓度值。由毅新兴业(北京)科技有限公司进行基因位点分型,SNP基因位点分型采用飞行时间质谱法,检测平台为Sequenom MassARRAY ©飞行时间质谱基因分型系统。本研究采用基质辅助激光解吸电离-飞行时间质谱的解离方式,通过检测延伸产物与未延伸引物间的分子量差异确定碱基。

基于既往文献报道,本研究分别对位于MTHFR C677T(rs1801133),MMP2(rs1132896,rs1477017,rs1992116,rs243847,rs243865,rs243866)的单核苷酸多态性进行了分型。依据遗传数据质量控制标准,排除未达到质量控制标准的位点:(1)基因型信息缺失>10%;(2)对照组中弱势等位基因频率低(MAF)<0.01;(3)对照组中Hardy-Weinberg平衡的显著偏差$P<0.05$。

1.3 统计学分析

采用Epidata3.1建立数据库,并进行平行双录入核查。基于R软件(Version 3.5.2)进行统计学分析。对于符合正态分布的连续型变量分布采用$\bar{x}\pm s$表示,对于非正态的连续型变量分布采用中位数±四分位数间距(median ± IQR)表示。组间比较采用非参数Wilcoxon成对检验;分类资料的分布使用频数及率表示。采用条件Logistic回归模型控制协变量的影响,探索两基因的乘法交互项是否有统计学意义。以乘法交互项的P值是否达到统计学显著性推断是否存在乘法交互作用;以乘法交互项的OR值及其95%可信区

间推断乘法交互作用的方向及大小。在加法交互作用探索框架中,采用交互作用超额危险度(relative excess risk of interaction, RERI)、交互作用归因比(attributable proportion of interaction, API)、交互作用指数(the synergy index, S)推断加法交互作用的方向,采用Assmann等的Boostrap Percentile Method[28],基于Bootstrap方法,以匹配编号为单元,进行1000次等样本量的重采样,以2.5%及97.5%分位数为指标的95%可信区间,估计3种交互作用评价指标的可信区间,进而推断是否存在统计学显著的加法交互作用。此外,同时采用广义多因子降维(GMDR)、非监督式稀疏因子分析的机器学习方法(EPISFA)等在控制协变量的同时验证交互作用结果。统计学显著性水平设置为0.05。

2 结果

2.1 研究对象基本特征

研究共募集155组1∶2匹配的病例/对照,共计465人。病例组男性138人(89.0%),女性17人(11.1%),年龄(69.2±10.0)岁,医院对照组与社区对照组研究对象年龄分别为(69.6±10.9)和(69.5±9.0)岁。吸烟和饮酒行为在病例与对照组间的分布差异有统计学意义($P<0.05$,表1)。吸烟者患AAA的风险是从不吸烟者的7.72倍(95%*CI*: 3.95~15.06, $P<0.01$);饮酒也是AAA的危险因素,饮酒患者AAA风险为从不饮酒者的1.61倍(95%*CI*:1.07~2.43, $P=0.02$)。

经过基因数据质量控制,共纳入MTHFR C677T(rs1801133),MMP2(rs1132896, rs1477017, rs1992116, rs243847)共5个SNPs。对照组中,这5个SNPs均符合Hardy-Weinberg平衡($P>0.05$)。这5个SNPs的不同基因型在病例组和对照组的分布如表1所示。

本研究分别对MTHFR C677T和MMP2基因上的5个SNPs进行多因素条件Logistic回归分析。每个SNP的弱势等位基因以加性算法纳入模型,在控制了其他混杂因素的情况下,MTHFR C677T和MMP2基因上的5个SNPs的独立主效应均达到了统计学阈值($P<0.05$)。

2.2 腹主动脉瘤的二维MTHFR-MMP2基因-基因交互作用

在乘法交互作用研究框架下,条件Logistic回归模型在控制协变量后显示,MTHFR C677T rs1801133与位于MMP2基因上的4个SNP的二维交互作用均达到统计学差异($P<0.05$),乘法交互项的OR值均<1,提示可能两基因间存在拮抗交互作用,影响罹患腹主动脉瘤的风险,见表2。MTHFR和MMP2基因上的弱势等位基因在各自单独存在时,均可独立升高AAA风险;当两基因上弱势等位基因同时存在时,两者则拮抗削弱了其各自单独作用时对AAA风险的影响。

表1 研究对象基本特征

特征	病例组($n=155$)	对照组($n=310$)	*P*值
性别			
男	138(89.0)	276(89.0)	-
女	17(11.1)	34(11.1)	-
年龄[a]	69.2±10.0	69.6±9.6	-
环境暴露			
吸烟	132(85.16)	167(53.87)	<0.001
饮酒	79(52.67)	131(42.26)	<0.001
生化指标			
Hcy[b]	19.30±10.90	18.42±8.98	<0.001
HDL-C[a]	1.02±0.39	1.20±0.28	<0.001
LDL-C[a]	2.77±1.25	2.69±0.70	0.095
hsCRP[b]	0.28±0.92	0.21±0.13	0.009
GLU[b]	4.93±0.81	5.90±0.80	<0.001
SBP[b]	131.00±25.00	138.00±31.34	0.003
DBP[b]	77.00±12.00	80.50±11.96	<0.001
合并症			
冠心病	53(34.19)	80(25.81)	0.059
高血压	108(69.68)	143(46.13)	<0.001
糖尿病	18(11.61)	46(14.84)	0.341
缺血性脑卒中	29(18.71)	31(10.00)	0.008
下肢静脉栓塞	51(32.90)	51(16.45)	<0.001
基因型			
MTHFR C677T			
rs1801133			
TT	45(30.00)	117(38.87)	0.089
TC	76(50.67)	121(40.20)	
CC	29(19.33)	63(20.93)	
MMP2			
rs1132896			
GG	94(61.44)	204(66.23)	0.251
GC	50(32.68)	95(30.85)	
CC	9(5.88)	9(2.92)	
rs1477017			
AA	74(48.36)	146(47.40)	0.778
AG	63(41.18)	135(43.83)	
GG	16(10.46)	27(8.77)	
rs1992116			
CC	75(49.02)	155(50.00)	0.230
CT	60(39.22)	133(42.90)	
TT	18(11.76)	22(7.10)	
rs243847			
TT	59(38.56)	97(31.60)	0.166
TC	66(43.14)	161(52.44)	
CC	28(18.20)	49(15.96)	

a: $\bar{x}\pm s$;b:M±IQR

此外,在加法交互作用探索框架下,本研究人群中MTHFR-MMP2基因-基因交互作用归因比API<0(最大值-0.837,最小值-0.465),即全部腹主动脉瘤病例中可归因于MTHFR-MMP2基因-基因交互作用的比例均<0;与交互作用指数S<1(即两基因同时作用时的效应与两基因独立作用时的效应之和的比值<1,亦即两基因同时作用时的效应<两基因独立作用时的效应之和)及交互作用超额危险度RERI<0(两基因同时作用时的效应与两基因独立作用时的效应之和的差值为负)结果一致,且3种指标经Bootstrap Percentile Method计算的95%可信区间均未跨越无效假设线(RERI与API的无效假设线为0,S的无效假设线为1)。结果提示MTHFR-MMP2间可能存在拮抗交互作用,影响人群中患腹主动脉瘤的风险。

本研究基于两种机器学习算法GMDR及EPISFA算法同时验证了MTHFR C677T rs1801133与MMP2基因的交互作用,两种机器学习模型交叉验证一致性较高,且结果均达到了统计学差异,见表3,进一步验证了本研究人群中发现的MTHFR-MMP2基因-基因交互作用。

表2 腹主动脉瘤的二维MTHFR-MMP2基因-基因交互作用

MTHFR-MMP2	β	*OR*	95%*CI*	*P*	RERI[a],95%*CI*[b]	API[c],95%*CI*	S[d],95%*CI*
rs1801133, rs1132896							
rs1801133,“C”	0.478	1.608	1.123~2.302	0.009	-1.472 (-2.084~-0.511)	-0.837 (-1.351~-0.668)	0.340 (0.292~0.547)
rs1132896,“C”	0.964	2.623	1.380~4.988	0.003			
rs1801133,“C”-rs1132896,“C”	-0.877	0.416	0.231~0.749	0.003			
rs1801133, rs1477017							
rs1801133,“C”	0.455	1.576	1.064~2.333	0.023	-0.659 (-1.394~-0.429)	-0.465 (-1.008~-0.241)	0.387 (0.299~0.602)
rs1477017,“C”	0.405	1.499	0.946~2.377	0.085			
rs1801133,“C”- rs1477017,“C”	-0.512	0.599	0.387~0.928	0.022			
rs1801133, rs1992116							
rs1801133,“C”	0.591	1.806	1.206~2.707	0.004	-1.434 (-2.933~-1.076)	-0.797 (-1.183~-0.549)	0.358 (0.195~0.581)
rs1992116,“T”	0.887	2.428	1.263~4.669	0.008			
rs1801133,“C”- rs1992116,“T”	-0.890	0.411	0.229~0.736	0.003			
rs1801133, rs243847							
rs1801133,“C”	1.076	2.931	0.949~9.056	0.062	-1.419 (-2.054~-1.197)	-0.643 (-0.889~-0.424)	0.460 (0.277~0.697)
rs243847,“T”	0.529	1.696	0.594~4.847	0.324			
rs1801133,“C”- rs243847,“T”	-0.813	0.443	0.271~0.725	0.001			

表3 机器学习方法验证的腹主动脉瘤的二维MTHFR-MMP2基因-基因交互作用

MTHFR-MMP2	Logistic	GMDR		EPISFA	
	$P_{\text{for interaction}}$	$P_{\text{for interaction}}$	交叉验证一致性	$P_{\text{for interaction}}$	交叉验证一致性
rs1801133, rs1132896	0.003	0.009	9/10	0.009	10/10
rs1801133, rs1477017	0.022	-	-	0.028	9/10
rs1801133, rs1992116	0.003	-	-	0.035	8/10
rs1801133, rs243847	0.012	-	-	0.014	9/10

2.3 腹主动脉瘤的高维 MTHFR-MMP2 基因-基因交互作用

控制协变量后,本研究同时基于 GMDR 和 EPISFA 两种机器学习模型发现 MTHFR-MMP2 基因间可能存在高维基因-基因交互作用。GMDR 筛选出的 MTHFR-MMP2 最优模型为三维基因-基因交互模型(rs1801133, rs1132896, rs1992116),该 3 位点组合在 EPISFA 模型中也得到了验证。两种机器学习模型在该三维基因-基因交互作用的交叉验证一致性较高(均为 7/10),且均达到了统计学差异(P<0.05),见表 4。MTHFR-MMP2 基因组合构成的 4 位点模型(rs1801133, rs1132896, rs1992116, rs243847)在 GMDR 和 EPISFA 模型下则未达到统计学差异。

表 4 腹主动脉瘤的高维 MTHF -MMP2 基因-基因交互作用

MTHFR-MMP2 交互	维度	分析方法	$P_{\text{for interaction}}$	交叉验证一致性
rs1801133, rs1132896, rs1992116	3	GMDR	0.017	7/10
rs1801133, rs1132896, rs1992116, rs243847	4	GMDR	0.054	5/10
rs1801133, rs1132896, rs1992116	3	EPISFA	0.044	7/10
rs1801133, rs1132896, rs1992116, rs243847	4	EPISFA	0.072	6/10

3 讨论

腹主动脉瘤是严重威胁患者生命健康的疾病。腹主动脉瘤破裂患者的病死率高,是高于 65 岁老年男性的首要猝死原因之一,造成严重的家庭和社会负担。因此探索其病因具有重要意义。MTHFR 与 MMP2 基因是影响腹主动脉瘤风险的重要候选基因[13-15,18]。然而,既往研究大多关注其独立主效应,或采用实验探索两基因及其交互作用在 AAA 中的作用,基于人群研究的 MTHFR-MMP2 基因-基因交互作用证据较少[11]。本研究采用 1:2 匹配的病例-对照研究设计(共 465 人),同时在乘法和加法尺度下发现 MTHFR 基因与 MMP2 基因可能存在拮抗交互作用,影响 AAA 罹患风险,并在 3 种分析方法中彼此验证;此外,GMDR 和 EPISFA 提示了 MTHFR-MMP2 间可能存在高维交互作用,具有较高的交叉验证一致性,且结果具有统计学差异。本研究结果为深入探索复杂疾病遗传病因的机制[12]、预测腹主动脉瘤的发生、发展及制定人群防制策略提供研究线索。

本研究在相加、相乘两种交互作用识别尺度下探索 MTHFR-MMP2 间基因-基因交互作用是否影响患腹主动脉瘤风险,并基于交互作用评价指标描述交互作用的方向和大小。基因-基因交互作用表现为两个或以上基因对表型的联合作用不同于其独立作用[22]。ROTHMAN 等[23]提出交互作用分为统计学交互作用,生物学交互作用和公共卫生学交互作用。统计学交互作用常基于相加、相乘两种尺度;其中,由于相乘模型检验较为方便,是目前应用较为广泛的尺度[24],但有研究者认为其对生物学交互作用的提示作用较为有限[24-25]。相反,相加模型尽管在模型拟合及指标评价等计算较为复杂,但其更适于提示生物学交互线索[24-25]。在生物学领域,基因-基因交互作用按照具体的作用机制可以分为协同交互(synergistic interaction)和拮抗交互(antagonistic interaction)[26]。ROTHMAN 等认为,在专业生物学知识解释前提下,可通过交互作用指标的评价等角度,从统计学交互证据提示生物学交互的线索[23]。常见的交互作用评价指标包括:交互作用超额危险度、交互作用归因比、交互作用指数等[23,27];具有正向相加交互作用的两因素可以指示具有生物学意义的协同作用;具有负向相加交互作用的两因素可以指示具有生物学意义的拮抗作用[23]。本文采用匹配病例对照设计,采用多种统计方法,同时分析相乘和相加交互尺度下 MTHFR-MMP2 基因-基因交互作用是否存在,及其存在的方向及大小,从多角度评价 MTHFR-MMP2 基因-基因交互在腹主动脉瘤风险中的作用。在乘法交互作用研究框架下,条件 Logistic 回归模型提示 MTHFR C677T rs1801133 与位于 MMP2 基因上的 4 个 SNP 的二维交互作用均为负向乘法交互作用(OR 值均<1),且具有统计学差异(P<0.05),提示可能两基因间存在拮抗交互作用,影响腹主动脉瘤的风险。本研究又在加法尺度下,验证了该拮抗交互作用的方向。在本研究人群中,MTHFR-MMP2 基因-基因交互作用归因比 API<0,即全部腹主动脉瘤病例中可归因于 MTHFR-MMP2 基因-基因交互作用的比例均<0,交互作用指数 S<1 且交互作用超额危险度<0,3 种评价指标均显示该交互作用方向为拮抗交互。为得到交互作用指标的可信区间,本研究采用 Bootstrap Percentile Method 对数据进行 1 000 次重采样。ASSMANN 等[27-28]的模拟研究提示,经 Bootstrap Percentile Method 估计的可信区间在交互作用的评价上更具有可靠性和稳定性。3 种指标经 Bootstrap Percentile Method 计算的 95%可信区间均未跨越无效假设线(RERI 与 API 的无效假设线为 0,S 的无效假

设线为 1),即本研究同时在相加、相乘尺度上发现 MTHFR-MMP2 基因-基因间可能存在拮抗交互作用,影响人群罹患腹主动脉瘤的风险。

本研究同时采用经典条件 Logistic 回归与新型机器学习(GMDR 广义多因子降维、EPISFA)探索与腹主动脉瘤患病风险有关的基因-基因交互作用,在多种数据分析策略下得到的结果彼此间可相互验证,具有较高的一致性。通常在病例对照研究中,基因-基因交互作用效应假设检验的统计方法主要可以分为三类[29]。第一类统计量称为基于 Logit 模型的统计量,基因-基因交互作用可以表示为 Logistic 回归中的交互作用项,是目前大多数基因-基因交互作用的识别方法。本研究在此框架下采用了条件 Logistic 回归探索 MTHFR-MMP2 是否存在基因-基因交互作用。第二类统计量是基于连锁不平衡原理。病例组和对照组中连锁不平衡系数的差值则为基因-基因交互作用效应得到估计量[30]。第三类统计量是单纯病例统计量,其定义基因型之间的 Pearson 相关系数的大小为基因-基因交互作用的统计量。本研究采用的 EPISFA 模型即是根据此统计量展开基因-基因交互作用探索[12],其以单纯病例研究的性质为基础,基于非监督式机器学习框架识别基因-基因交互作用。与监督式方法不同,非监督式机器学习方法不直接利用多态性位点与表型之间的关联信息,而利用患病人群中交互位点之间的相关性性质。此外,LOU 等[21]提出的 GMDR 将广义线性模型的概念引人到多因子降维中,可在估计交互作用的同时纳入协变量,并可探索高维交互作用,有效补充了 Logistic 回归模型在计算高维交互作用时的效能不足。目前 GMDR 已经较为广泛应用于心血管疾病、精神疾病等复杂疾病的遗传研究中[31-32]。EPISFA 方法较新:XIANG 等[12]应用 EPISFA 发现了 6 对可能影响缺血性脑卒中的基因-基因交互作用,并在多种模型下得到了交互作用证据的验证;但在腹主动脉瘤领域尚无此类方法的应用,具有应用此类新方法探索的价值。在二维交互中,本研究发现 MTHFR 基因与 MMP2 基因上多个位点间的交互作用拮抗影响了 AAA 患病风险($P<0.05$),并在 3 种分析方法中彼此验证,且 GMDR 和 EPISFA 的交叉验证结果均显示交互作用证据有较高的一致性;此外,GMDR 和 EPISFA 提示了 MTHFR-MMP2 rs1801133, rs1132896, rs1992116 位点组合可能存在高维交互作用,存在较高的交叉检验一致性,且具有统计学差异;本研究所采取的 3 种分析模型从不同的数据分析框架出发,得到了一致性较高的研究结果,提示 MTHFR 基因与 MMP2 间可能存在基因-基因交互作用拮抗影响 AAA 患病风险。

近年来,多项研究证据提示 MTHFR 基因与 MMP2 基因多态性与 AAA 的发展有关[13-15],但对于两者交互作用展开探索的人群研究证据较少[20]。本研究采用 1∶2匹配病例对照设计发现 MTHFR 基因与 MMP2 基因上多个位点间的交互作用拮抗影响了 AAA 患病风险($P<0.05$)。这可能与两个基因共同参与了 Hcy 水平调控有关[20,33],而高 Hcy 水平已经多项研究发现为 AAA 的独立危险因素。SMALLWOOD 等[20]基于组织水平实验观察到在 miR-516a-5p 过表达后,MTHFR C677T 与 MMP 家族基因同时改变,并参与了 Hcy 异常代谢。范义川等[34]在动物实验模型中也发现诱导大鼠高 Hcy 后,引起 MMP2 的活性改变,进而影响 AAA 的风险。

本研究采用匹配病例-对照研究设计共纳入 465 人,同时在加法和乘法交互作用识别尺度下,采用多种分析方法对 MTHFR 与 MMP2 基因-基因交互作用展开探索和相互验证。本研究同时在乘法和加法尺度下,发现 MTHFR 基因与 MMP2 基因可能存在拮抗交互作用影响 AAA 患病风险,有待未来更多本土人群研究及功能实验对结果的验证。

参考文献:

[1] 左尚维, 隗瑛琦, 陈峰, 等. 腹主动脉瘤危险因素的病例对照研究[J]. 北京大学学报(医学版), 2014, 46(3): 412-416. DOI: 10.3969/j.issn.1671-167X.2014.03.014.
ZUO S W, WEI Y Q, CHEN F, *et al*. A matched case-control study of risk factors in abdominal aortic aneurysm[J]. J Peking Univ Heal Sci, 2014, 46(3): 412-416. DOI: 10.3969/j.issn.1671-167X.2014.03.014.

[2] KEISLER B, CARTER C. Abdominal aortic aneurysm[J]. N Engl J Med, 1993, 329(17): 1275-1277. DOI: 10.1056/nejm199310213291715.

[3] SCHMITZ-RIXEN T, KEESE M, HAKIMI M, *et al*. Ruptured abdominal aortic aneurysm—epidemiology, predisposing factors, and biology[J]. Langenbecks Arch Surg, 2016, 401(3): 275-288. DOI: 10.1007/s00423-016-1401-8.

[4] SIDLOFF D, STATHER P, DATTANI N. Aneurysm global epidemiology study: public health measures can further reduce abdominal aortic aneurysm mortality[J]. J Vasc Surg, 2014, 59(5): 1471. DOI: 10.1016/j.jvs.2014.03.266.

[5] PINARD A, JONES G T, MILEWICZ D M. Genetics of thoracic and abdominal aortic diseases[J]. Circ Res, 2019, 124(4): 588-606. DOI: 10.1161/circresaha.118.312436.

[6] WAHLGREN C M, LARSSON E, MAGNUSSON P K, *et al*. Genetic and environmental contributions to abdominal aortic aneurysm development in a twin population[J]. J Vasc Surg, 2010, 51(1): 3-7; discussion7. DOI: 10.1016/j.jvs.2009.08.036.

[7] GRETARSDOTTIR S, BAAS A F, THORLEIFSSON G, *et al*. Genome-wide association study identifies a sequence variant within the DAB2IP gene conferring susceptibility to abdominal aortic aneurysm[J]. Nat Genet, 2010, 42(8): 692-697.

[8] BOWN M J, JONES G T, HARRISON S C, *et al*. Abdominal aortic aneurysm is associated with a variant in low-density lipoprotein receptor-related protein 1[J]. Am J Hum Genet, 2011, 89(5): 619-627. DOI: 10.1016/j.ajhg.2011.10.002.

[9] BRADLEY D T, HUGHES A E, BADGER S A, *et al*. A variant in LDLR is associated with abdominal aortic aneurysm[J]. Circ Cardiovasc Genet, 2013,6(5): 498-504. DOI: 10.1161/CIRCGENETICS.113.000165.

[10] JONES G T, TROMP G, KUIVANIEMI H, *et al*. Meta-analysis of genome-wide association studies for abdominal aortic aneurysm identifies four new disease-specific risk loci[J]. Circ Res, 2017, 120(2): 341-353. DOI: 10.1161/CIRCRESAHA.116.308765.

[11] KLARIN D, VERMA S S, JUDY R, *et al*. Genetic architecture of abdominal aortic aneurysm in the Million Veteran Program[J]. Circulation, 2020, 142(17): 1633-1646. DOI: 10.1161/CIRCULATIONAHA.120.047544.

[12] XIANG X, WANG S, LIU T, *et al*. Exploring gene-gene interaction in family-based data with an unsupervised machine learning method: EPISFA[J]. Genet Epidemiol, 2020, 44(8): 811-824. DOI: 10.1002/gepi.22342.

[13] SAKALIHASAN N, DEFRAIGNE J O, KERSTENNE M A, *et al*. Family members of patients with abdominal aortic aneurysms are at increased risk for aneurysms: analysis of 618 probands and their families from the Liège AAA Family Study[J]. Ann Vasc Surg, 2014, 28(4): 787-797. DOI: 10.1016/j.avsg.2013.11.005.

[14] HINTERSEHER I, TROMP G, KUIVANIEMI H. Genes and abdominal aortic aneurysm[J]. Ann Vasc Surg, 2011, 25(3): 388-412. DOI: 10.1016/j.avsg.2010.09.004.

[15] CAO H, HU X H, ZHANG Q, *et al*. Hyperhomocysteinaemia, low folate concentrations and MTHFR C677T mutation in abdominal aortic aneurysm[J]. Vasa, 2014, 43(3): 181-188. DOI: 10.1024/0301-1526/a000347.

[16] STRAUSS E, WALISZEWSKI K, GABRIEL M, *et al*. Increased risk of the abdominal aortic aneurysm in carriers of the MTHFR 677T allele[J]. J Appl Genet, 2003, 44(1): 85-93.

[17] BRUNELLI T, PRISCO D, FEDI S, *et al*. High prevalence of mild hyperhomocysteinemia in patients with abdominal aortic aneurysm[J]. J Vasc Surg, 2000, 32(3): 531-536. DOI: 10.1067/mva.2000.107563.

[18] WANG X, KHALIL R A. Matrix metalloproteinases, vascular remodeling, and vascular disease[J]. Adv Pharmacol, 2018, 81: 241-330. DOI: 10.1016/bs.apha.2017.08.002.

[19] HINTERSEHER I, BERGERT H, KUHLISCH E, *et al*. Matrix metalloproteinase 2 polymorphisms in a Caucasian population with abdominal aortic aneurysm[J]. J Surg Res, 2006, 133(2): 121-128. DOI: 10.1016/j.jss.2005.12.001.

[20] SMALLWOOD L, ALLCOCK R, VAN BOCKXMEER F, *et al*. Polymorphisms of the matrix metalloproteinase 9 gene and abdominal aortic aneurysm[J]. Br J Surg, 2008, 95(10): 1239-1244. DOI: 10.1002/bjs.6345.

[21] LOU X Y, CHEN G B, YAN L, *et al*. A generalized combinatorial approach for detecting gene-by-gene and gene-by-environment interactions with application to nicotine dependence[J]. Am J Hum Genet, 2007, 80(6): 1125-1137. DOI: 10.1086/518312.

[22] CORDELL H J. Epistasis: what it means, what it doesn't mean, and statistical methods to detect it in humans[J]. Hum Mol Genet, 2002, 11(20): 2463-2468. DOI: 10.1093/hmg/11.20.2463.

[23] ROTHMAN K J, GREENLAND S. Modern epidemiology[M]. 3rd ed. Philadelphia: Lippincot Williams & Wilkins, 2008: 94-113.

[24] 袁悦, 李楠, 任爱国, 等. 流行病学研究中相加和相乘尺度交互作用的分析[J]. 现代预防医学, 2015, 42(6): 961-965,975.
YUAN Y, LI N, REN A G, *et al*. Analysis of the application of the additive model and the multiplicative statistical model in biologial interaction[J]. Mod Prev Med, 2015, 42(6): 961-965,975.

[25] KNOL M J, VAN DER TWEEL I, GROBBEE D E, *et al*. Estimating interaction on an additive scale between continuous determinants in a logistic regression model[J]. Int J Epidemiol, 2007, 36(5): 1111-1118. DOI: 10.1093/ije/dym157.

[26] LI W, REICH J. A complete enumeration and classification of two-locus disease models[J]. Hum Hered, 2000, 50(6): 334-349. DOI: 10.1159/000022939.

[27] 邱宏, 余德新, 王晓蓉, 等. logistic 回归模型中交互作用的分析及评价[J]. 中华流行病学杂志, 2008, 29(9): 934-937. DOI: 10.3321/j.issn: 0254-6450.2008.09.019.
QIU H, IGNATIUS T S, WANG X R, *et al*. Study on the

interaction under logistic regression modeling[J]. Chin J Epidemiol, 2008, 29(9): 934-937. DOI: 10.3321/j.issn: 0254-6450.2008.09.019.

[28] ASSMANN S F, HOSMER D W, LEMESHOW S, *et al*. Confidence intervals for measures of interaction[J]. Epidemiology, 1996, 7(3): 286-290. DOI: 10.1097/00001648-199605000-00012.

[29] HU J K, WANG X, WANG P. Testing gene-gene interactions in genome wide association studies[J]. Genet Epidemiol, 2014, 38(2): 123-134. DOI: 10.1002/gepi.21786.

[30] YANG Y, HE C, OTT J. Testing association with interactions by partitioning Chi-squares[J]. Ann Hum Genet, 2009, 73(1): 109-117. DOI: 10.1111/j.1469-1809.2008.00480.x.

[31] 吴同, 冉柳毅, 靳佳佳, 等. 脑源性神经营养因子基因和神经营养性酪氨酸激酶受体Ⅱ型基因交互作用与青年自杀未遂行为的关系[J]. 第三军医大学学报, 2019, 41(11): 1065-1069. DOI: 10.16016/j.1000-5404.201811215.

WU T, RAN L Y, JIN J J, *et al*. Association of interaction between brain-derived neurotrophic factor gene and neurotrophic tyrosine kinase receptor type 2 gene with suicide attempt in middle and later adolescents[J]. J Third Mil Med Univ, 2019, 41(11): 1065-1069. DOI: 10.16016/j.1000-5404.201811215.

[32] YI X, LIN J, LUO H, *et al*. Interactions among variants in TXA2R, P2Y12 and GPIIIa are associated with carotid plaque vulnerability in Chinese population[J]. Oncotarget, 2018, 9(25): 17597-17607. DOI: 10.18632/oncotarget.24801.

[33] HALAZUN K J, BOFKIN K A, ASTHANA S, *et al*. Hyperhomocysteinaemia is associated with the rate of abdominal aortic aneurysm expansion[J]. Eur J Vasc Endovascular Surg, 2007, 33(4): 391-394. DOI: 10.1016/j.ejvs.2006.10.022.

[34] 范义川. 甲硫氨酸诱导高同型半胱氨酸血症在腹主动脉瘤发病中的作用及分子机制的实验性研究[D]. 沈阳: 中国医科大学, 2019.

FAN Y C. Role of methionine-induced hyperhomocysteinemia in the pathogenesis of abdominal aortic aneurysm and its molecular mechanisms[D]. Shenyang: China Medical University, 2019.

(**收稿**:2021-01-31;**修回**:2021-02-16)

(编辑 张 维)