

一种新的调控型遗传变异与中国人群肺癌发病风险的关系:两阶段病例-对照研究

朱小琪¹ 司妮平¹ 付晓宇¹ 程静雯¹ 秦娜² 刘逸辰¹ 田甜¹ 马红霞² 褚敏捷¹

¹南通大学公共卫生学院流行病与卫生统计学系 226019; ²南京医科大学公共卫生学院流行病学系 211166

通信作者:褚敏捷, Email: chuminjie@ntu.edu.cn

【摘要】 目的 调控型数量性状位点(regQTL)理论可以帮助研究者从三维角度评估单核苷酸多态性(SNPs)对重要生物信号的调控作用。本研究拟探讨 regQTL-SNPs 对肺癌易感性的影响。方法 基于 regQTL 理论,利用已知的肺癌 regQTL-SNPs 数据库,筛选出全基因组关联研究(GWAS)报道的肺癌易感区域中发挥 regQTL 功能的 SNPs。并通过两阶段病例-对照研究(初筛阶段:2 331 例肺癌病例和 3 077 例健康对照;验证阶段:626 例肺癌病例和 667 例健康对照),进一步明确上述 regQTL-SNPs 与肺癌易感性的关联。**结果** 在肺癌 GWAS 已报道的易感区域中,共筛选出 8 个 regQTL-SNPs。人群易感性分析的初筛阶段,研究结果显示 3 个 regQTL-SNPs 与肺癌的发病风险存在统计学关联($P < 0.05$),验证阶段结果显示,位于 *ADRA1A* 基因上的 rs6998591 突变等位基因 T 可以显著增加肺癌的发病风险(相加模型: $OR = 1.33, 95\%CI: 1.01 \sim 1.74, P = 0.040$),而位于 *ACTA2* 基因上的 rs11202916 突变等位基因 G 可以明显降低肺癌的发病风险(隐性模型: $OR = 0.71, 95\%CI: 0.52 \sim 0.96, P = 0.026$)。分层分析结果显示,rs6998591 的突变等位基因 T 显著增加肺鳞癌的发病风险(相加模型: $OR = 1.53, 95\%CI: 1.01 \sim 2.32, P = 0.043$),而 rs11202916 的突变等位基因 G 显著降低肺腺癌的发病风险(相加模型: $OR = 0.83, 95\%CI: 0.69 \sim 0.98, P = 0.031$)。基因环境交互作用分析显示携带 rs6998591 突变等位基因 T 且吸烟的个体与不携带 rs6998591 突变等位基因 T 且不吸烟的个体相比,肺癌的发病风险增加 235% ($OR = 3.35, 95\%CI: 2.10 \sim 5.34, P < 0.001$)。**结论** 肺癌 GWAS 已报道的易感区域中存在 2 个发挥 regQTL 功能的 SNPs,并且可以显著影响肺癌的易感性。

【关键词】 肺癌; 调控型数量性状位点; 单核苷酸多态性

基金项目:国家自然科学基金(81703297)

Association between a novel regulatory genetic variants and lung cancer risk in Chinese: a two-stage case-control study

Zhu Xiaoqi¹, Si Niping¹, Fu Xiaoyu¹, Cheng Jingwen¹, Qin Na², Liu Yichen¹, Tian Tian¹, Ma Hongxia², Chu Minjie¹

¹Department of Epidemiology and Biostatistics, School of Public Health, Nantong University, Nantong 226019, China; ²Department of Epidemiology, School of Public Health, Nanjing Medical University, Nanjing 211166, China

Corresponding author: Chu Minjie, Email: chuminjie@ntu.edu.cn

【Abstract】 Objective Regulatory quantitative trait loci (regQTL) theory can help to evaluate the regulation function of single nucleotide polymorphisms (SNPs) on crucial biological signals from a three-dimensional perspective. The aim of this study was to investigate the effect of

DOI: 10.3760/cma.j.cn112338-20210331-00262

收稿日期 2021-03-31 本文编辑 万玉立

引用本文:朱小琪,司妮平,付晓宇,等.一种新的调控型遗传变异与中国人群肺癌发病风险的关系:两阶段病例-对照研究[J].中华流行病学杂志,2021,42(11):2053-2059. DOI: 10.3760/cma.j.cn112338-20210331-00262.



these regQTL-SNPs on the susceptibility of lung cancer. **Methods** Based on the regQTL theory, using the database of identified lung cancer regQTL-SNPs, we screened the SNPs that may function as regQTL in the reported susceptible regions of lung cancer by genome-wide association study (GWAS), and a two-stage case-control study was conducted (screening stage: 2 331 lung cancer cases and 3 077 healthy controls; validation stage: 626 lung cancer cases and 667 healthy controls) to definite the association of related regQTL-SNPs with the susceptibility of lung cancer. **Results** A total of 8 regQTL-SNPs were screened in the reported susceptible regions of lung cancer by GWAS. Among which, 3 SNPs were significantly associated with the risk of lung cancer ($P < 0.05$) in the screening stage. Further validation results indicated that the variant T allele of rs6998591 in *ADRA1A* was significantly associated with increased risk of lung cancer (additive model: $OR = 1.33$, 95%*CI*: 1.01-1.74, $P = 0.040$). In addition, the variant G allele of rs11202916 in *ACTA2* was significantly associated with decreased risk of lung cancer (recessive model: $OR = 0.71$, 95%*CI*: 0.52-0.96, $P = 0.026$). Stratified analysis indicated that the variant T allele of rs6998591 significantly increased lung squamous cell carcinoma risk (additive model: $OR = 1.53$, 95%*CI*: 1.01-2.32, $P = 0.043$), while the variant G allele of rs11202916 significantly decreased lung adenocarcinoma risk (additive model: $OR = 0.83$, 95%*CI*: 0.69-0.98, $P = 0.031$). Gene-environment interaction analysis indicated that the risk of developing lung cancer increased by 235% in smoking individuals carrying rs6998591 variant T allele compared with those non-smoking individuals carrying no rs6998591 variant T allele ($OR = 3.35$, 95%*CI*: 2.10-5.34, $P < 0.001$). **Conclusion** There are two regQTL-SNPs that could significantly affect the susceptibility of lung cancer in the GWAS reported susceptible regions of lung cancer.

【 Key words 】 Lung cancer; Regulatory quantitative trait loci; Single nucleotide polymorphism

Fund program: National Natural Science Foundation of China (81703297)

肺癌是严重危害人类健康的恶性肿瘤之一,也是我国发病率和死亡率最高的恶性肿瘤^[1]。据统计,2020年,我国肺癌新增病例约82万例,死亡病例近71万例。目前,肺癌仍然是危害我国人民群众健康的重要公共卫生问题之一^[2]。肺癌的发生发展是环境和遗传因素共同作用的结果^[3]。尽管吸烟是肺癌最重要的环境危险因素,但是遗传因素在肺癌的发生发展过程中也起着非常关键的作用,有研究表明肺癌的遗传度接近20%^[4]。此外,近期研究表明,与低遗传风险的重度吸烟者相比,存在高遗传风险的轻度吸烟者的肺癌易感性更高^[5],这更加提示遗传因素可能在较大程度上影响了个体对肺癌的易感程度。在肺癌的遗传因素中,研究最多的主要是单核苷酸多态性(single nucleotide polymorphisms, SNPs)^[6]。

研究者通过表达数量性状位点(expression quantitative trait loci, eQTL)分析可以判定一个SNP位点的不同基因型是否能够影响一个或多个基因(mRNA)的表达水平,因此,通过eQTL分析评估SNP的生物学功能是目前公认的一种研究策略^[7]。然而,eQTL分析方法仍然存在一定的不足,例如,它只能从二维的角度探讨SNP与相关mRNA表达之间的关联。近些年,有研究将SNP基因型和mRNA表达数据与微小RNA(miRNA)表达数据相结合,从三维的角度评估某SNP对miRNA以及

mRNA表达水平的调控情况,并将存在统计学意义的SNP定义为调控型QTL(regulatory quantitative trait loci, regQTL)-SNP^[8]。这种regQTL-SNPs的不同基因型对相关miRNA与mRNA表达关联的影响存在明显的差异,例如,在某一SNP野生纯合基因型下,相应的mRNA和miRNA表达水平呈现出较强的正相关性,在其突变杂合基因型下,相应的mRNA和miRNA表达水平无相关性,而在其突变纯合基因型下,相应的mRNA和miRNA表达水平呈现出强烈的负相关性。这种regQTL分析方法的出现为进一步评估相关SNPs的调控功能提供了新的视角。然而,目前与肺癌易感性相关的可以发挥regQTL功能的SNPs却鲜有报道。

本研究依据regQTL调控理论,利用肺癌相关regQTL-SNPs数据库,并结合全基因组关联研究(GWAS)Catalog数据库中的肺癌易感区域,筛选这些区域内可能具有regQTL功能的SNPs,随后,通过两阶段病例-对照研究,进一步明确相关regQTL-SNPs与肺癌易感性的关联。

资料与方法

1. 肺癌相关regQTL-SNPs的筛选:利用已发表的肺癌regQTL-SNPs数据库^[8],得到555 366个regQTL-SNPs($P < 0.05$)。在此基础上,基于两个筛

选标准($P < 10^{-4}$ 和错误发现率 $P < 0.2$),进一步获得 1 723 个满足条件的 regQTL-SNPs (主要分布于 352 个染色体区域)。同时,基于肺癌 GWAS Catalog 数据库(包含 45 篇已发表的肺癌 GWAS 文章,其中研究人群来源于亚洲地区的共有 21 篇),根据 $P < 10^{-5}$ 的标准筛选出 726 个肺癌易感 SNPs (对应于 325 个肺癌易感区域)。将 325 个肺癌易感区域与上文提及的 352 个染色体区域进行匹配,共得到 169 个交集区域(包含 581 个符合条件的 regQTL-SNPs)。这 581 个候选 regQTL-SNPs 主要位于 264 个基因上,利用 GEPIA 数据库(<http://gepia.cancer-pku.cn/index.html>)分析了这些基因在肺癌和癌旁组织间的表达情况,共筛选得到 51 个存在差异表达的基因($P < 0.05$),进一步按照癌和癌旁组织间的差异表达倍数 > 2.0 的标准,最终得到 6 个基因(包含 12 个符合条件的 regQTL-SNPs)。接着,利用千人基因组数据库(<https://www.ncbi.nlm.nih.gov/variation/tools/1000genomes/>),将最小等位基因频率(minor allele frequency) ≤ 0.05 的 SNPs 去除,得到 10 个符合条件的 regQTL-SNPs。最后,通过连锁不平衡分析(linkage disequilibrium, LD)($r^2 < 0.8$),最终筛选得到 8 个候选 regQTL-SNPs。

2. 研究对象:初筛阶段的样本来源于南京医科大学的肺癌 GWAS 数据库^[9]。验证阶段的病例来自 2015–2018 年南通大学附属医院、常熟市第一人民医院和南通市第六人民医院收集的明确诊断的肺癌新发病例(纳入标准:汉族,年龄、性别不限;排除标准:采血前接受过化疗、放疗或有既往肿瘤病史)。验证阶段的对照组选自同时期参加常规体检的健康人群($> 2 000$ 人),与病例组的年龄、性别因素进行频数匹配,随机选择合适的对照人群^[10]。

3. 流行病学调查和临床信息采集:病例和对照采用统一设计的肺癌流行病学调查问卷进行面访调查。内容包括研究对象的一般情况(年龄、性别等)、个人疾病史、肿瘤家族史和生活习惯等。根据统一制定的流行病学调查工作手册培训调查员,采用统一方法和标准进行调查,确保流行病学调查资料的质量。临床资料包括确诊日期、肿瘤病理、TNM 分期、治疗方式和日期等。

4. 生物标本收集:每位研究对象均采用 EDTA 真空抗凝采血管采集静脉血 5 ml,按标准方法在 4 h 内离心,将分离到的血浆、白细胞和红细胞转至 1.5 ml 离心管中(均各保存 2 份), -80 °C 冷冻保存备用。

5. Taqman 基因分型:采用酚氯仿法从外周血中提取基因组 DNA,验证阶段利用南京骥骜生物公司合成的探针、引物和 Master Mix 配成基因分型体系。然后,利用 7500 FAST Real-Time PCR 系统(美国 Applied Biosystems 公司)进行 PCR 扩增并读取荧光信号。根据荧光强度,在 SDS 软件上进行基因分型结果鉴定。

6. 统计学分析:采用 χ^2 检验或 t 检验比较病例组和对照组人口学特征及肺癌发病相关因素的分布情况。采用 logistic 回归分析(共显性模型:突变杂合型 vs. 野生纯合型,突变纯合型 vs. 野生纯合型;显性模型:突变纯合型+突变杂合型 vs. 野生纯合型;隐性模型:突变纯合型 vs. 野生纯合型+突变杂合型;相加模型:疾病发病风险随突变等位基因的增加而改变的概率)计算 OR 值及其 95%CI,估计 regQTL-SNPs 与肺癌发病风险的关联强度。基因与环境间相乘交互作用分析采用 logistic 回归模型。所有检验均为双侧检验,检验水准 $\alpha = 0.05$ 。所有的统计分析均通过 R 3.6.2 软件进行计算。

结 果

1. 基本情况:初筛阶段包含 2 331 例肺癌病例和 3 077 例健康对照。验证阶段包含 626 例肺癌病例和 667 例健康对照,其中,病例组和对照组的年龄和性别分布差异无统计学意义(均 $P > 0.05$)。病例组和对照组的吸烟情况差异有统计学意义($P = 0.008$)。病例组中具有肿瘤家族史的占 22.36%,高于对照组的 16.49%,差异有统计学意义($P = 0.009$)。病例组中肺鳞癌患者有 155 人(24.76%),肺腺癌患者有 471 人(75.24%)。见表 1。

2. 候选 regQTL-SNPs 与肺癌发病风险之间的关联(初筛阶段):候选的 8 个 regQTL-SNPs 位点的基本信息见表 2。经年龄、性别、吸烟状态及人群分层因素调整后,结果显示,位于 ADRA1A 基因上的 rs4548188 (T>A) 和 rs6998591 (C>T),以及 ACTA2 基因上的 rs11202916 (A>G) 多态性位点与肺癌的发病风险存在明显的统计学关联。相加模型结果显示,rs4548188 和 rs6998591 的基因型改变增加肺癌发病风险(rs4548188: OR=1.16, 95%CI: 1.04~1.30, $P = 9.54 \times 10^{-3}$; rs6998591: OR=1.26, 95%CI: 1.11~1.44, $P = 5.60 \times 10^{-4}$), rs11202916 基因型改变则与肺癌发病风险降低有关(OR=0.85, 95%CI: 0.78~0.93, $P = 1.77 \times 10^{-4}$)。其余 5 个

表 1 验证阶段样本的基线信息

变量	病例(n=626)	对照(n=667)	P 值
年龄组(岁, $\bar{x} \pm s$)	62.73±9.23	62.31±6.73	0.348
≤61	267(42.65)	306(45.88)	0.263
>61	359(57.35)	361(54.12)	
性别			0.104
男	388(61.98)	443(66.42)	
女	238(38.02)	224(33.58)	
吸烟状态			0.008
现在吸	254(40.58)	242(36.28)	
过去吸	74(11.82)	55(8.25)	
从不吸	298(47.60)	370(55.47)	
肿瘤家族史			0.009
有	140(22.36)	110(16.49)	
无	486(77.64)	557(83.51)	
病理类型			
鳞癌	155(24.76)		
腺癌	471(75.24)		

注:括号外数据为人数,括号内数据为构成比(%)

regQTL-SNPs 与肺癌发病风险无统计学关联 ($P > 0.05$)。见表 3。

3. 阳性 regQTL-SNPs 与肺癌发病风险之间的关联(验证阶段):对初筛阶段获取的 3 个阳性 regQTL-SNPs (rs4548188、rs6998591、rs11202916)

表 2 8 个 regQTL-SNPs 的基本信息

序号	基因	位点	miRNA	P 值	FDR P 值	MAF
1	TFPI	rs7597815	hsa-mir-106b	2.93×10^{-5}	0.18	0.452
2	GRIAI	rs2926289	hsa-let-7g	2.72×10^{-5}	0.17	0.384
3	FYN	rs7770556	hsa-mir-1247	2.63×10^{-6}	0.08	0.272
4	FYN	rs2026603	hsa-mir-1247	3.74×10^{-6}	0.10	0.320
5	ADRAIA	rs4732646	hsa-mir-1247	2.06×10^{-5}	0.16	0.233
6	ADRAIA	rs4548188	hsa-mir-1247	1.26×10^{-5}	0.14	0.194
7	ADRAIA	rs6998591	hsa-mir-1247	5.20×10^{-7}	0.04	0.141
8	ACTA2	rs11202916	hsa-mir-1306	2.44×10^{-5}	0.17	0.398

注:FDR:错误发现率;MAF:最小等位基因频率

表 3 初筛阶段 8 个 regQTL-SNPs 在病例-对照中的基因型分布及与肺癌易感性的关联

序号	位点	基因	染色体位置	等位基因 ^a	病例 ^b (n=2 331)	对照 ^b (n=3 077)	MAF ^c (病例/对照)	OR 值 ^d (95%CI)	P 值 ^d
1	rs7597815	TFPI	chr2:188436536	T>A	583/1 170/577	747/1 551/777	0.501/0.495	1.02(0.94~1.11)	6.04×10^{-1}
2	rs2926289	GRIAI	chr5:152637000	T>C	463/1 163/704	604/1 544/926	0.448/0.448	1.02(0.94~1.11)	6.78×10^{-1}
3	rs7770556	FYN	chr6:112182111	T>G	235/1 014/1 070	309/1 355/1 381	0.320/0.324	0.99(0.90~1.08)	8.11×10^{-1}
4	rs2026603	FYN	chr6:112215441	T>G	336/1 107/888	442/1 455/1 179	0.382/0.380	1.03(0.94~1.12)	5.54×10^{-1}
5	rs4732646	ADRAIA	chr8:26636585	T>C	110/802/1 418	132/1 022/1 922	0.219/0.209	1.07(0.97~1.19)	1.78×10^{-1}
6	rs4548188	ADRAIA	chr8:26646298	T>A	73/670/1 588	81/807/2 189	0.175/0.158	1.16(1.04~1.30)	9.54×10^{-3}
7	rs6998591	ADRAIA	chr8:26650822	C>T	33/485/1 783	39/522/2 459	0.120/0.099	1.26(1.11~1.44)	5.60×10^{-4}
8	rs11202916	ACTA2	chr10:90736399	A>G	330/1 089/911	546/1 466/1 064	0.375/0.416	0.85(0.78~0.93)	1.77×10^{-4}

注:^a其他等位基因>效应等位基因;^b分别对应突变纯合型、突变杂合型以及野生纯合型的基因型频数;^cMAF:最小等位基因频率;^d调整因素为年龄、性别、吸烟状态和人群分层

进行验证,logistic 回归分析结果显示,在调整年龄、性别和吸烟状态之后,位于 ADRAIA 基因上的 rs6998591(C>T)可以显著增加肺癌的发病风险(相加模型:OR=1.33,95%CI:1.01~1.74,P=0.040)。而位于 ACTA2 基因上的 rs11202916(A>G)可以显著降低肺癌的发病风险(隐性模型:OR=0.71,95%CI:0.52~0.96,P=0.026)。见表 4。

4. 按肺癌病理类型分层分析:考虑到病例组中肺鳞癌与肺腺癌的比例差别较大,且两者具有不完全相同的发病机制。进一步对 2 个阳性 regQTL-SNPs 按肺癌病理类型进行分层分析,结果显示,rs6998591 的突变等位基因 T 增加肺鳞癌的发病风险(相加模型:OR=1.53,95%CI:1.01~2.32,P=0.043),而与肺腺癌的发病风险无显著相关性。相反,rs11202916 的突变等位基因 G 降低肺腺癌的发病风险(相加模型:OR=0.83,95%CI:0.69~0.98,P=0.031),而与肺鳞癌发病风险无显著相关性。见表 5。

5. 基因与环境交互作用分析:基线数据显示,肺癌组的吸烟率显著高于对照组,进一步开展了基因(SNP 位点)与环境(吸烟)的交互作用分析,结果显示,位于 ADRAIA 基因上的 rs6998591(C>T)与吸烟间存在显著的交互作用(OR=1.71,95%CI:1.16~2.52,P=0.007),其中携带 rs6998591 突变等位基因 T 且吸烟的个体与不携带 rs6998591 突变等位基因 T 且不吸烟的个体相比,肺癌的发病风险增加了 235%(OR=3.35,95%CI:2.10~5.34,P<0.001)(图 1)。

讨 论

本研究综合利用肺癌 GWAS Catalog 数据库和肺癌 regQTL-SNPs 数据库,筛选得到 GWAS 已报道的肺癌易感区域中 8 个潜在的发挥 regQTL 功能的

表 4 验证阶段 3 个 regQTL-SNPs 在病例-对照中的基因型分布及与肺癌易感性的关联

位点/基因型	病例 ^a	对照 ^a	aOR 值 ^b (95%CI)	P 值 ^b
rs4548188				
TT	431(69.52)	480(72.51)	1.00	
TA	178(28.71)	164(24.77)	1.19(0.92~1.53)	0.184
AA	11(1.77)	18(2.72)	0.63(0.29~1.35)	0.232
显性模型			1.13(0.88~1.44)	0.331
隐性模型			0.60(0.28~1.28)	0.187
相加模型			1.05(0.85~1.31)	0.635
rs6998591				
CC	488(78.08)	555(83.21)	1.00	
CT	134(21.44)	108(16.19)	1.39(1.05~1.85)	0.023
TT	3(0.48)	4(0.60)	0.84(0.18~3.81)	0.816
显性模型			1.37(1.04~1.82)	0.027
隐性模型			0.79(0.17~3.57)	0.754
相加模型			1.33(1.01~1.74)	0.040
rs11202916				
AA	234(37.56)	230(34.64)	1.00	
AG	303(48.64)	315(47.44)	0.97(0.76~1.23)	0.780
GG	86(13.80)	119(17.92)	0.69(0.50~0.97)	0.032
显性模型			0.89(0.71~1.12)	0.317
隐性模型			0.71(0.52~0.96)	0.026
相加模型			0.86(0.73~1.01)	0.061

注：^a括号外数据为人数，括号内数据为构成比(%)；^b调整因素为年龄、性别和吸烟状态

SNPs。进一步的两阶段病例-对照研究结果提示，位于 *ACTA2* 基因上的 rs11202916 位点 (A>G) 和 *ADRA1A* 基因上的 rs6998591 位点 (C>T) 与肺癌的

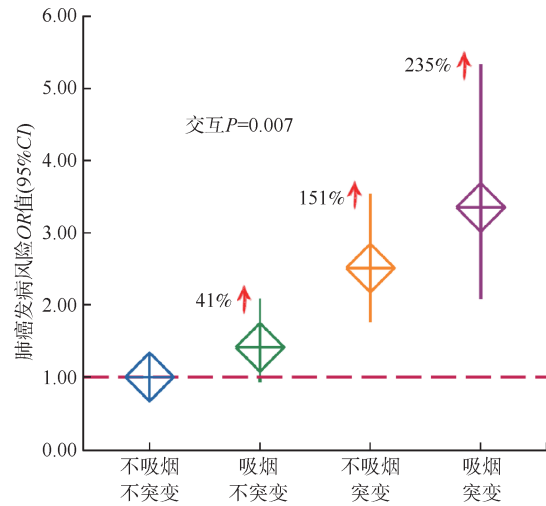


图 1 rs6998591 (C>T) 与吸烟的交互作用分析

发病风险显著相关。基因环境交互作用分析显示，rs6998591 与吸烟在肺癌发病危险中存在显著的交互作用。

ACTA2 又称为 α 平滑肌肌动蛋白，属于肌动蛋白家族。肌动蛋白在细胞中发挥重要的作用^[11]。*ACTA2* 基因的异常表达与多种肿瘤的发生发展过程有关^[12]。GEPIA 数据库显示 *ACTA2* 在肺鳞癌 (LUSC) 和肺腺癌 (LUAD) 癌组织中的表达水平均明显低于癌旁组织，提示 *ACTA2* 可能发挥抑癌基因的作用。研究表明，在肺腺癌细胞中，*ACTA2* 的高

表 5 2 个阳性 regQTL-SNPs 按肺癌病理类型分层分析

位点	病理类型	基因型	病例 ^a	对照 ^a	aOR 值 ^b (95%CI)	P 值 ^b	
rs6998591	鳞癌	CC	116(74.84)	555(83.21)	1.00		
		CT	38(24.52)	108(16.19)	1.65(1.05~2.57)	0.028	
		TT	1(0.65)	4(0.60)	1.03(0.11~9.79)	0.978	
		显性模型			1.62(1.04~2.52)	0.031	
		隐性模型			0.93(0.10~8.84)	0.952	
		相加模型			1.53(1.01~2.32)	0.043	
		腺癌	腺癌	CC	372(79.15)	555(83.21)	1.00
CT	96(20.43)			108(16.19)	1.33(0.97~1.81)	0.074	
TT	2(0.43)			4(0.60)	0.67(0.12~3.73)	0.652	
显性模型					1.30(0.96~1.77)	0.092	
隐性模型					0.64(0.12~3.54)	0.610	
相加模型					1.25(0.94~1.68)	0.130	
rs11202916	鳞癌			AA	55(35.48)	230(34.64)	1.00
		AG	76(49.03)	315(47.44)	1.16(0.77~1.74)	0.484	
		GG	24(15.48)	119(17.92)	0.88(0.51~1.54)	0.663	
		显性模型			1.08(0.73~1.58)	0.703	
		隐性模型			0.81(0.49~1.34)	0.421	
		相加模型			0.98(0.75~1.27)	0.862	
		腺癌	腺癌	AA	179(38.25)	230(34.64)	1.00
	AG			227(48.50)	315(47.44)	0.94(0.72~1.22)	0.629
	GG			62(13.25)	119(17.92)	0.63(0.44~0.92)	0.016
	显性模型					0.85(0.66~1.09)	0.205
	隐性模型					0.66(0.47~0.92)	0.015
	相加模型					0.83(0.69~0.98)	0.031

注：^a括号外数据为人数，括号内数据为构成比(%)；^b调整因素为年龄、性别和吸烟状态

表达可抑制癌细胞的增殖、侵袭和迁移,加速细胞凋亡,这些信息提示 *ACTA2* 可以发挥抑癌基因的作用,这与 GEPIA 数据库的结果一致。此外,高表达的 *ACTA2* 还可通过调节 c-MET 和 FAK 的表达影响肺癌细胞转移,提示 *ACTA2* 可能作为转移性肺癌的预后生物标志物和治疗靶点^[13]。本研究发现,rs11202916 的不同基因型可能会影响 *ACTA2* 和 miRNA-1306 之间的表达相关性,从而在肺癌的发生发展中起到一定的调控作用。目前,已有部分研究关注了 miRNA-1306 在不同癌症中发挥的作用,并进行了相应的生物学功能探索。He 等^[14]的研究发现 miRNA-1306-3p 可通过靶向结合 *FBXL5* 基因,促进上皮间质转化(EMT)过程从而调控肺癌的发生发展。Li 等^[15]的研究显示,miRNA-1306 可以调控 *SIX1* 基因的表达,即 miRNA-1306 表达水平的下调可以促进 *SIX1* 基因的表达,而 *SIX1* 基因与多种肿瘤发病的关联已经被证实。还有研究发现,miRNA-1306 可影响 TGF- β /SMAD 信号通路活性^[16],而 TGF- β 的过度表达与肺癌的不良预后有关^[17]。然而,miRNA-1306 在肺癌发生中的生物学功能尚未明确,本研究的结果可能为后续更加深入的生物学机制探讨提供基础数据。

ADRA1A 是肾上腺素受体 1A 蛋白,属于 $\alpha 1$ -肾上腺受体家族^[18]。它在生物体内发挥的功能主要有调控血压、收缩血管以及收缩平滑肌等^[19]。有研究发现,*ADRA1A* 基因在多种癌症的发生发展过程中发挥了重要的作用。例如,Chen 等^[20]的研究指出,当 *ADRA1A* 基因的启动子区域出现高甲基化时,*ADRA1A* 基因的表达随之下降,从而促进了肝癌的发生发展。有研究指出,*ADRA1A* 基因高表达能够明显降低肝癌患者的死亡率^[21],提示 *ADRA1A* 基因可能在肝癌的发生发展中发挥抑癌基因的作用。同时,GEPIA 数据库提示 *ADRA1A* 基因在 LUSC 和 LUAD 癌组织中表达水平均显著低于癌旁组织,提示 *ADRA1A* 基因在肺癌的发生发展中也发挥着相似的抑癌基因作用。本研究已预测出,rs6998591 的不同基因型可能会影响 *ADRA1A* 基因和 miRNA-1247 之间的表达相关性。目前,已有多篇研究指出 miRNA-1247 在多种癌症发生发展中起到抑制癌症发生的作用^[22-24]。有研究显示,与健康对照组相比,miRNA-1247-5p 在非小细胞肺癌患者血浆中显著高表达,提示其可以作为早期诊断非小细胞肺癌的生物标志物^[25]。此外,Chu 等^[22]的研究

表明,高表达的 miRNA-1247 可以靶向结合 Wnt3 并抑制细胞增殖,进而影响肝癌的发生发展过程。此外,还有学者发现 Wnt3 可以通过调控 EMT 过程来促进肺腺癌的侵袭与转移,进而影响肺腺癌患者的预后^[26]。后续需要进一步明确 *ADRA1A* 基因与 miRNA-1247 在肺癌发生发展中的调控作用。

本研究主要有两个优点:基于 regQTL 调控理论,在目前已报道的肺癌易感区域中筛选发挥 regQTL 功能的 SNPs,为发现更多的肺癌易感区域内致病位点提供基础数据;本研究利用两阶段的病例-对照研究设计开展人群易感性研究(以中国大样本肺癌 GWAS 数据为初筛辅以其他人群数据进行验证),增加了研究结果的准确性和可靠性。

本研究存在局限性。首先,本研究是以医院为基础的病例-对照研究,可能导致入院率偏倚,后续还需在以社区人群为基础的更大样本中评价研究结果的稳定性;其次,人群中获得的关联研究结果还需后续生物功能学实验来进一步评价其潜在的机制。再次,尽管 rs6998591 在相加模型下与肺癌发病风险具有统计学意义,但是由于验证阶段样本量的限制,病例组仅有 3 个 rs6998591 突变纯合型 TT 携带者(占 0.48%),对照组也仅有 4 个 rs6998591 突变纯合型 TT 携带者(占 0.60%),所以,该位点相加模型具有统计学意义的效应可能主要是由突变杂合型 CT 所体现(病例组 CT 型占 21.44%,对照组 CT 型占 16.19%),后续还需要更大的样本量来明确在共显性模型下 rs6998591 突变纯合型 TT 与肺癌发病风险之间的关联。最后,regQTL-SNPs 数据库是基于 TCGA 数据分析得到,不同种族人群的遗传变异差异可能会对候选 SNPs 的选择产生影响。尽管千人基因组数据库提示在全世界范围内不同种族人群中,上述 2 个阳性 regQTL-SNPs 突变方向与中国汉族人群一致率均较高,但后续还是需要其他种族人群中深入评价上述 2 个阳性 regQTL-SNPs 与肺癌易感性的关联。

综上所述,本研究基于最新的 regQTL 调控理论,通过两阶段病例-对照研究设计,发现肺癌 GWAS 已报道区域中 2 个发挥 regQTL 功能的潜在致病位点(*ACTA2* 基因上的 rs11202916 和 *ADRA1A* 基因上的 rs6998591)与肺癌易感性显著相关。研究结果对识别和评价肺癌的遗传易感因素,阐明肺癌的发生机制具有重要指导价值。

利益冲突 所有作者均声明不存在利益冲突

参 考 文 献

- [1] 陈万青, 李贺, 孙可欣, 等. 2014 年中国恶性肿瘤发病和死亡分析[J]. 中华肿瘤杂志, 2018, 40(1):5-13. DOI:10.3760/cma.j.issn.0253-3766.2018.01.002.
Chen WQ, Li H, Sun KX, et al. Report of cancer incidence and mortality in China, 2014[J]. Chin J Oncol, 2018, 40(1):5-13. DOI:10.3760/cma.j.issn.0253-3766.2018.01.002.
- [2] Sung H, Ferlay J, Siegel RL, et al. Global cancer statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries[J]. CA A Cancer J Clin, 2021, 71(3): 209-249. DOI: 10.3322/CAAC.21660.
- [3] Yang DW, Liu Y, Bai CX, et al. Epidemiology of lung cancer and lung cancer screening programs in China and the United States[J]. Cancer Lett, 2020, 468: 82-87. DOI: 10.1016/j.canlet.2019.10.009.
- [4] Mucci LA, Hjelmberg JB, Harris JR, et al. Familial risk and heritability of cancer among twins in Nordic countries[J]. JAMA, 2016, 315(1): 68-76. DOI: 10.1001/jama. 2015.17703.
- [5] Dai JC, Lv J, Zhu M, et al. Identification of risk loci and a polygenic risk score for lung cancer: a large-scale prospective cohort study in Chinese populations[J]. Lancet Respir Med, 2019, 7(10):881-891. DOI:10.1016/S2213-2600(19)30144-4.
- [6] 沈洪兵, 靳光付. 全基因组关联研究与复杂疾病风险预测的现状与展望[J]. 中华流行病学杂志, 2011, 32(7):643-649. DOI:10.3760/cma.j.issn.0254-6450.2011.07.002.
Shen HB, Jin GF. Genome-wide association study (GWAS) and risk prediction of complex disease: advances and prospects[J]. Chin J Epidemiol, 2011, 32(7):643-649. DOI: 10.3760/cma.j.issn.0254-6450.2011.07.002.
- [7] 宋庆峰, 张红星, 马亦龙, 等. 复杂疾病的遗传易感基因区域的精细定位[J]. 遗传(北京), 2014, 36(1):2-10. DOI: 10.3724/SPJ.1005.2014.0011.
Song QF, Zhang HX, Ma YL, et al. Fine mapping of complex disease susceptibility loci[J]. Hereditas (Beijing), 2014, 36(1):2-10. DOI:10.3724/SPJ.1005.2014.0011.
- [8] Wilk G, Braun R. RegQTLs: Single nucleotide polymorphisms that modulate microRNA regulation of gene expression in tumors[J]. PLoS Genet, 2018, 14(12): e1007837. DOI:10.1371/journal.pgen.1007837.
- [9] Hu ZB, Wu C, Shi YY, et al. A genome-wide association study identifies two new lung cancer susceptibility loci at 13q12.12 and 22q12.2 in Han Chinese[J]. Nat Genet, 2011, 43(8):792-796. DOI:10.1038/ng.875.
- [10] Yu YH, Mao LP, Lu X, et al. Functional variant in 3'UTR of FAM13A is potentially associated with susceptibility and survival of lung squamous carcinoma[J]. DNA Cell Biol, 2019, 38(11):1269-1277. DOI:10.1089/dna.2019.4892.
- [11] Yuan SM. α -Smooth muscle actin and *acta2* gene expressions in vasculopathies[J]. Braz J Cardiovasc Surg, 2015, 30(6): 644-649. DOI:10.5935/1678-9741.20150081.
- [12] Luo LL, Wang M, Li XP, et al. A novel mechanism by which *ACTA2-AS1* promotes cervical cancer progression: acting as a ceRNA of miR-143-3p to regulate SMAD3 expression [J]. Cancer Cell Int, 2020, 20:372. DOI:10.1186/s12935-020-01471-w.
- [13] Lee HW, Park YM, Lee SJ, et al. Alpha-smooth muscle actin (*ACTA2*) is required for metastatic potential of human lung adenocarcinoma[J]. Clin Cancer Res, 2013, 19(21): 5879-5889. DOI:10.1158/1078-0432.CCR-13-1181.
- [14] He ZJ, Li W, Chen H, et al. miR-1306-3p targets FBXL5 to promote metastasis of hepatocellular carcinoma through suppressing snail degradation[J]. Biochem Biophys Res Commun, 2018, 504(4): 820-826. DOI: 10.1016/j.bbrc. 2018.09.059.
- [15] Li L, Liang YC, Kang L, et al. Transcriptional regulation of the warburg effect in cancer by SIX1[J]. Cancer Cell, 2018, 33(3):368-385.e7. DOI:10.1016/j.ccell.2018.01.010.
- [16] Yang L, Du X, Liu L, et al. miR-1306 mediates the feedback regulation of the TGF- β /SMAD signaling pathway in granulosa cells[J]. Cells, 2019, 8(4):298. DOI: 10.1186/s12885-019-5917-5.
- [17] Li J, Shen C, Wang X, et al. Prognostic value of TGF- β in lung cancer: systematic review and Meta-analysis[J]. BMC Cancer, 2019, 19(1):691. DOI:10.1186/s12885-019-5917-5.
- [18] Perez DM. Structure-function of alpha1-adrenergic receptors[J]. Biochem Pharmacol, 2007, 73(8): 1051-1062. DOI:10.1016/j.bcp.2006.09.010.
- [19] Docherty JR. Subtypes of functional α 1-adrenoceptor[J]. Cell Mol Life Sci, 2010, 67(3): 405-417. DOI: 10.1007/s00018-009-0174-4.
- [20] Chen GQ, Fan XX, Li YR, et al. Promoter aberrant methylation status of *ADRA1A* is associated with hepatocellular carcinoma[J]. Epigenetics, 2020, 15(6/7): 684-701. DOI:10.1080/15592294.2019.1709267.
- [21] He JF, Zhao HC, Deng DF, et al. Screening of significant biomarkers related with prognosis of liver cancer by lncRNA-associated ceRNAs analysis[J]. J Cell Physiol, 2020, 235(3):2464-2477. DOI:10.1002/jcp.29151.
- [22] Chu YK, Fan WN, Guo WW, et al. miR-1247-5p functions as a tumor suppressor in human hepatocellular carcinoma by targeting Wnt3[J]. Oncol Rep, 2017, 38(1): 343-351. DOI:10.3892/or.2017.5702.
- [23] Taddei ML, Cavallini L, Ramazzotti M, et al. Stromal-induced downregulation of miR-1247 promotes prostate cancer malignancy[J]. J Cell Physiol, 2019, 234(6):8274-8285. DOI:10.1002/jcp.27679.
- [24] Shi S, Lu Y, Qin Y, et al. miR-1247 is correlated with prognosis of pancreatic cancer and inhibits cell proliferation by targeting neuropilins[J]. Curr Mol Med, 2014, 14(3): 316-327. DOI: 10.2174/1566524014666140228120014.
- [25] Dong XH, Chang MH, Song XG, et al. Plasma miR-1247-5p, miR-301b-3p and miR-105-5p as potential biomarkers for early diagnosis of non-small cell lung cancer[J]. Thorac Cancer, 2021, 12(4): 539-548. DOI: 10.1111/1759-7714.13800.
- [26] Xu JM, Lv W, Hu YJ, et al. Wnt3a expression is associated with epithelial-mesenchymal transition and impacts prognosis of lung adenocarcinoma patients[J]. J Cancer, 2017, 8(13):2523-2531. DOI:10.7150/jca.18560.