

# 中国女性外周血 *EMR3* 基因甲基化水平与乳腺癌的相关性

周夏婕<sup>1</sup>, 雷水芳<sup>1</sup>, 黎立喜<sup>2</sup>, 徐天<sup>3</sup>, 顾万建<sup>3</sup>, 马飞<sup>2</sup>, 杨蓉西<sup>1</sup>

<sup>1</sup>南京医科大学公共卫生学院流行病学与卫生统计学系, 江苏 南京 211166; <sup>2</sup>国家癌症中心//国家肿瘤临床医学研究中心//中国医学科学院北京协和医学院肿瘤医院肿瘤内科, 北京 100021; <sup>3</sup>江苏省中医院检验科, 江苏 南京 210029

**摘要:**目的 探究中国人外周血中 *C19orf57*、*MAP9*、*EMR3*、*NEK6* 和 *PCOLCE2* 基因的甲基化水平与乳腺癌的相关性。方法 通过病例对照研究收集了 258 例早期乳腺癌患者和 272 名健康女性的外周血, 并采用基质辅助激光解吸电离飞行时间质谱技术 (MALDI-TOF-MS) 定量检测基因 CpG 位点的甲基化水平。使用 logistics 回归模型校正协变量分析 DNA 甲基化水平与乳腺癌的相关性。Spearman 秩相关用于分析基因甲基化水平与年龄的关系。使用非参数检验分析不同临床特征组别的甲基化水平差异。结果 与欧美人群的乳腺癌 *EMR3* 高甲基化相反, *EMR3* 基因的低甲基化与中国女性的乳腺癌相关, 但该相关性仅存在于 ≥50 岁的女性中 (每减少 10% 的甲基化, *EMR3*\_CpG\_1: OR=1.40; *EMR3*\_CpG\_2: OR=2.31; *EMR3*\_CpG\_3: OR=2.76,  $P < 0.05$ )。 *EMR3* 的甲基化水平与肿瘤分期、大小、淋巴转移、ER、PR、HER2、Ki67 等没有或仅有弱相关。 *C19orf57* 甲基化与乳腺癌无相关性。结论 外周血 *EMR3* 基因的低甲基化与中国女性乳腺癌之间存在相关性, 尤其是年龄较大或绝经后女性。  
**关键词:** 外周血; *EMR3*; DNA 甲基化; 乳腺癌; 基质辅助激光解吸电离飞行时间质谱技术

## Peripheral blood *EMR3* gene methylation level is correlated with breast cancer in Chinese women

ZHOU Xiajie<sup>1</sup>, LEI Shuifang<sup>1</sup>, LI Lixi<sup>2</sup>, XU Tian<sup>3</sup>, GU Wanjian<sup>3</sup>, MA Fei<sup>2</sup>, YANG Rongxi<sup>1</sup>

<sup>1</sup>Department of Epidemiology and Biostatistics, School of Public Health, Nanjing Medical University, Nanjing 211166, China; <sup>2</sup>Department of Medical Oncology, National Cancer Center/National Clinical Research Center for Cancer/Cancer Hospital, Chinese Academy of Medical Sciences & Peking Union Medical College, Beijing 100021, China; <sup>3</sup>Department of Clinical Laboratory, Jiangsu Provincial Hospital of Chinese Medicine, Nanjing 210029, China

**Abstract: Objective** To explore the association of methylation levels of *C19orf57*, *MAP9*, *EMR3*, *NEK6* and *PCOLCE2* genes in peripheral blood with breast cancer (BC) in Chinese women. **Methods** We collected peripheral blood samples from 258 early-stage BC patients and 272 healthy women. Agena matrix-assisted laser desorption ionization time-of-flight mass spectrometry (MALDI-TOF-MS) was utilized to quantitatively measure the methylation levels of CpG sites in the genes. The association between DNA methylation and BC was analyzed using a logistic regression model adjusted for covariants. Spearman's correlation analysis was performed to analyze the association between the gene methylation levels and age. The methylation levels of the genes in the BC patients with different clinical characteristics were investigated using non-parametric tests. **Results** In stead of *EMR3* gene hypermethylation as found in BC patients as found in the Caucasian population, *EMR3* gene hypomethylation was found to correlate with BC in Chinese women, but this correlation was significant only in women beyond the age of 50 years (for every 10% reduction of the methylation level, *EMR3*\_CpG\_1: OR=1.40; *EMR3*\_CpG\_2: OR=2.31; *EMR3*\_CpG\_3: OR=2.76,  $P < 0.05$ ). *EMR3* methylation was not or was only weakly correlated with tumor stage, size, lymphatic metastasis, ER, PR, HER2, or Ki67. Our data did not show a correlation between *C19orf57* methylation and BC. **Conclusion** Peripheral blood *EMR3* gene hypomethylation is associated with BC in Chinese women, especially in those at an old age and in postmenopausal women.

**Keywords:** peripheral blood; *EMR3*; DNA methylation; breast cancer; MALDI-TOF

收稿日期: 2021-06-06

基金项目: 国家自然科学基金(81803303, 81874122); 江苏省自然科学基金(BK20180674); 中国医学科学院医学与健康科技创新工程-重大协同创新项目(2017-12M-3-004); 江苏省基金(WSW-009)

Supported by National Natural Science Foundation of China (81803303, 81874122).

作者简介: 周夏婕, 在读硕士研究生, E-mail: xiajiezhou413@163.com; 雷水芳, 在读硕士研究生, E-mail: LeiSF2019@njmu.edu.cn。周夏婕、雷水芳共同为第一作者

通信作者: 杨蓉西, 教授, 博士生导师, E-mail: rongxiyang@njmu.edu.cn; 马飞, 教授, 主任医师, 博士生导师, E-mail: drmafei@126.com

2020年最新数据显示, 乳腺癌已超过肺癌成为全球发病率最高的癌症, 也是导致癌症患者死亡的第5大原因。在女性群体中, 乳腺癌约占所有癌症新发病例的1/4, 在440万女性癌症患者死亡病例中占比15.5%<sup>[1]</sup>。乳腺癌的早诊断早治疗是决定其预后和生存的关键因素<sup>[2]</sup>。目前用于乳腺癌早诊的技术有钼靶X射线成像技术、超声诊断技术、核磁共振成像技术等<sup>[3]</sup>。钼靶X射线成像技术对于40岁以下或乳房密度较高的女性效果不佳<sup>[4]</sup>。而我国多数女性乳腺腺体致密并且发病人群年龄相对较低<sup>[5, 6]</sup>, 因此该技术并不完全适用。超声诊断

技术则受影像质量以及人为主观判断等因素影响较大<sup>[7]</sup>。核磁共振成像(MRI)在乳腺癌研究中被广泛使用,但因其成本高、耐受性有限等,不适用于全人群筛查<sup>[8,9]</sup>。因此,寻找一种高效客观且适用范围广的乳腺癌早诊技术显得尤为重要。

表观遗传学包括DNA甲基化、染色质重塑、组蛋白修饰和非编码RNA等,它在不改变DNA序列的前提下产生可遗传变异<sup>[10]</sup>。DNA甲基化作为一种重要的表观遗传机制,通常指在甲基转移酶(DNMT)的作用下,将甲基基团转移到胞嘧啶的C5位置,形成5-甲基胞嘧啶<sup>[11]</sup>。DNA甲基化的改变与癌症的发生发展相关<sup>[12,13]</sup>。对于组织样本来讲,抑癌基因的启动子区高甲基化往往导致转录抑制和基因表达降低<sup>[14,15]</sup>,而全基因组DNA低甲基化则可能通过增加染色体的不稳定性来促进肿瘤形成<sup>[16]</sup>。在癌症早期,特异的DNA甲基化即有可能发生改变,如果可以通过特定技术检测到,则可能是理想的癌症早诊生物标志物<sup>[17,18]</sup>。

Xu等<sup>[19]</sup>使用Illumina 27K芯片在美国Sister Study前瞻性队列的血液样本中发现了250个跟乳腺癌相关的甲基化位点(FDR<0.05)。其中的5个甲基化位点(cg03430067, cg03616357, cg07072643, cg08287471, cg19709625)在他们随后的基于2776例外周血样本的Illumina 450K芯片数据中再次得到了验证(FDR<0.01),并且他们在欧洲EPIC-Italy前瞻性队列数据库中也验证了外周血中这5个甲基化位点与乳腺癌之间的相关性<sup>[20]</sup>。由于DNA甲基化会受不同遗传背景及生活习惯、环境等其他因素所影响<sup>[21,22]</sup>,因此有价值且有必要对这5个甲基化位点与乳腺癌的相关性在其他种族中进行验证。Guan等<sup>[23]</sup>也强调,在研究甲基化与疾病的相关性时,对研究结果进行多轮独立验证是有必要的。经文献检索,目前尚未找到这5个甲基化位点相关基因(*C19orf57*、*MAP9*、*EMR3*、*NEK6*、*PCOLCE2*)的甲基化水平与乳腺癌相关的独立研究。本研究通过病例对照设计来探讨这些基因的甲基化水平与中国女性乳腺癌之间的相关性,从而评估这些基因甲基化作为中国女性乳腺癌早期检测生物标志物的可能性。

## 1 资料和方法

### 1.1 研究对象

从中国医学科学院肿瘤医院收集女性乳腺癌患者258例(中位年龄45.5岁)病例纳入标准:手术后病理诊断为乳腺癌、有完整的临床数据记录、尚未开始放疗或化疗、无其他癌症病史且无来自其他器官的癌症转移。病例排除标准:手术后病理诊断为乳腺良性肿瘤和乳腺腺病等。收集乳腺癌患者临床资料包括:初诊年龄、手术病理诊断结果、乳腺癌亚型分类、乳腺癌分期、雌激素

受体情况(ER)、孕酮受体情况(PR)、人表皮生长因子受体-2情况(HER2)、Ki67水平等。临床资料显示,我们的患者多数处于早期,其中I期43.8%,II期36.9%,III期18.9%,IV期0.4%。

272名女性对照来自江苏省中医院健康体检中心(中位年龄45.0岁)。对照纳入标准:自述健康、无肿瘤史、无自身免疫病史、血液各项指标正常。采集研究对象外周血全血,置于含乙二胺四乙酸(EDTA)的管中,防止血液凝固,并于-80℃保存。本研究的伦理经南京医科大学和中国医学科学院伦理委员会讨论通过。

### 1.2 确定甲基化位点相关基因

我们将5个CpG位点分别输入UCSC网站中的“human GRCh37/hg19”数据库中,根据CpG位点位于某基因上或距离最近,确定其为相关基因。5个CpG位点与相关基因的各自关系如下:cg03430067位于*C19orf57*基因的启动子区;cg03616357位于*MAP9*基因的第一外显子区;cg07072643位于*EMR3*基因的第一外显子区;cg08287471位于*NEK6*基因的第一内含子区;cg19709625位于*PCOLCE2*基因的第一外显子区。

### 1.3 全基因组DNA提取及亚硫酸氢盐处理

全基因组DNA按照南京腾辰公司的DNA提取试剂盒(TANTICA, China)说明书进行提取。使用EZ-96 DNA Methylation Gold Kit(Zymo Research, USA)试剂盒按照说明书对上述提取好的DNA进行亚硫酸氢盐处理。

### 1.4 基因甲基化水平测定

使用Yang等<sup>[24,25]</sup>描述的基质辅助激光解吸电离飞行时间质谱(MALDI-TOF-MS, Agena Bioscience, USA)对基因甲基化水平进行定量检测。使用特异引物(GenScript, China)对亚硫酸氢盐处理后的DNA进行目标序列扩增。具体各引物的序列如表1所示。反应条件如下:95℃预变性4min;接下来45次循环,包括95℃下变性20s,56℃下退火30s,72℃下延伸2min;最后,在72℃下延伸5min。聚合酶链式反应(PCR)产物按照Agena EpiTYPER Assay的标准程序处理,然后用Nanodispenser点样至硅基靶板上。数据由MassARRAY系统采集,使用EpiTYPER v1.2软件进行数据分析。该方法检测到的大多数质量峰只包含一个CpG位点,少数质量峰包含两个CpG位点。例如:C19orf57\_CpG\_7和C19orf57\_CpG\_8经过EpiTYPER处理后位于同一片段。因此,质量峰代表的是C19orf57\_CpG\_7和C19orf57\_CpG\_8两位点的平均甲基化水平(在表1中记为C19orf57\_CpG\_7,8)。实验操作中,对病例和对照样本采取平行处理。

从Xu等<sup>[20]</sup>研究的5个CpG位点对应的5个基因中选取了4个与肿瘤、免疫相关的基因(*C19orf57*、*MAP9*、

*EMR3*、*NEK6*)。每个基因设计两对引物,引物及扩增片段区域内的CpG位点均未与已知的单核苷酸多态性(SNP)重合。引物扩增序列尽可能包含目的CpG位点,但由于目的位点在CpG岛或扩增序列包含目的位点的引物无法设计的原因,*C19orf57*和*NEK6*的引物扩增序列均未包含目的CpG位点(cg03430067、cg08287471),而尝试了附近的区域。在预实验中,对8对引物(表1)进行了PCR扩增,随后进行质谱甲基化水平检测。质谱检测结果显示其中6个扩增片段的甲基化值低于质谱稳定检测区间(小于0.1)。仅有2个扩增片段(对应基因:*C19orf57*和*EMR3*)满足要求,可进行进一步的实验。*C19orf57*基因上的片段位于cg03430067下游165 bp的位置,长度为217 bp。使用飞行时间质谱定量检测其

9个CpG位点的甲基化水平。*EMR3*基因上的片段长度为269 bp,其上5个CpG位点(其中*EMR3\_CpG\_1*为目的CpG位点:cg07072643)的甲基化水平得到检测。对于*C19orf57*和*EMR3*两个基因,我们先进行了第一轮验证(病例:88;对照:94)随后进行数据分析。发现*C19orf57*基因的甲基化水平与乳腺癌不相关,而*EMR3*基因的低甲基化与乳腺癌相关。因此,我们仅对*EMR3*基因进行了独立的第二轮验证(病例:170;对照:178;第二轮验证的样本跟第一轮验证的样本无重合)。随后对*EMR3*基因的数据分析包括年龄相关、年龄分层以及临床数据分析均为两次验证的样本量相加(病例:258;对照:272)的结果。

表1 基因引物序列

Tab.1 Sequence of primers for amplification of the genes

Gene	Primer name	Left primer	Right primer
<i>C19orf57</i>	C19orf57_244	TTTTGAAGATGTATAAGAGGAAAAGGA	AACCAAAAACCAATAAACCCACTTA
	C19orf57_217	GTAGTGTAGTGTGGGGTTTGTAT	TTAAAAACCCTCTATCCTTATCAAAA
<i>MAP9</i>	MAP9_290	GAGTTGTTTATTTTGGTTTGGGTTT	TCCTATCTCCTAAACAACAACACAT
	MAP9_271	TTTTTTAAAGATTTAGGAAGGTGATG	ACCCTAAAAATCAAAAACCAAAACAT
<i>EMR3</i>	EMR3_269	TTTGGAAGAAGTAATGGTTTTTGTA	CCCAAAAATAATCAACAATAACTAAACA
	EMR3_291	GAAAAGTAGAATTATGAGTTTGATGG	CCAATTATTAAACCAATAATCCCT
<i>NEK6</i>	NEK_297	TGTGTGTATTTTAAAGTGATTGTGTGT	CTCTATCAAACCAACCAAACTCTA
	NEK_183	GTTTTTTTTGGGTTTGGATTTTT	AAAATTAATACTACAAAACCTCATCAT

### 1.5 统计学分析

所有的统计分析都在SPSS 25.0版中进行。使用校正年龄和批间差的logistic回归模型来比较病例组和对照组之间各位点的DNA甲基化水平差异。使用Spearman秩相关分析甲基化与年龄的相关性。使用Kruskal-Wallis检验或Mann-Whitney U检验分析同一临床特征不同分组后的甲基化水平是否有差异。所有统计分析均采用双侧检验, $P < 0.05$ 被认为差异具有统计学意义。

## 2 结果

### 2.1 *C19orf57*甲基化水平与乳腺癌的相关性

研究对象共包括88例早期乳腺癌患者以及94例对照的外周血。使用logistics回归模型校正年龄后,分析结果显示*C19orf57*基因上9个位点的甲基化水平在病例组与对照组间无统计学差异,提示外周血*C19orf57*的甲基化水平与中国女性的乳腺癌不相关(表2)。因

此,并未对*C19orf57*基因进行扩大样本的研究。

### 2.2 *EMR3*甲基化水平与乳腺癌的相关性

研究对象与上述*C19orf57*基因保持一致,为88例乳腺癌患者以及94例对照。Logistics回归模型校正年龄后,结果显示*EMR3\_CpG\_2*、*EMR3\_CpG\_3*、*EMR3\_CpG\_4*三个位点的低甲基化与乳腺癌相关(每减少10%的甲基化;*EMR3\_CpG\_2*:OR=1.96,95% CI: 1.10~3.50, $P=0.022$ ;*EMR3\_CpG\_3*:OR=3.10,95% CI: 1.51~6.34, $P=0.002$ ;*EMR3\_CpG\_4*:OR=1.51,95% CI: 1.05~2.16, $P=0.026$ ,表2),其他的*EMR3*甲基化位点未观察到与乳腺癌的相关关系。然而,在我们将样本量扩大至合计258例乳腺癌患者以及272例对照后,logistics回归模型校正年龄与批间差的结果却显示5个位点的甲基化程度均与乳腺癌不相关(表3)。

### 2.3 *EMR3*甲基化与年龄和激素水平的相关性

在乳腺癌病例中,*EMR3\_CpG\_2*和*EMR3\_CpG\_4*两个位点的甲基化水平与年龄呈负相关(*EMR3\_CpG\_2*:

表2 散发性乳腺癌患者与对照 *C19orf57* 和 *EMR3* 基因甲基化水平差异Tab.2 Comparison of methylation levels of *C19orf57* and *EMR3* genes between 88 sporadic BC patients and 94 control subjects

CpG sites	BC cases (n=88)	Controls (n=94)	OR (95% CI)*	P
	median (IQR)	median (IQR)	per-10% methylation	
C19orf57_CpG_1	0.33 (0.26-0.40)	0.35 (0.29-0.42)	1.13 (0.88-1.47)	0.339
C19orf57_CpG_2	0.31 (0.22-0.40)	0.31 (0.25-0.36)	0.95 (0.73-1.24)	0.721
C19orf57_CpG_3	0.45 (0.35-0.53)	0.48 (0.42-0.52)	1.37 (1.00-1.88)	0.052
C19orf57_CpG_4	0.87 (0.75-0.98)	0.86 (0.78-0.95)	1.17 (0.93-1.48)	0.190
C19orf57_CpG_5	0.22 (0.17-0.30)	0.25 (0.20-0.31)	1.16 (0.87-1.55)	0.312
C19orf57_CpG_6	0.39 (0.29-0.50)	0.40 (0.33-0.45)	1.02 (0.78-1.32)	0.903
C19orf57_CpG_7,8	0.42 (0.36-0.49)	0.44 (0.39-0.49)	1.43 (0.98-2.07)	0.061
C19orf57_CpG_10	0.86 (0.80-0.91)	0.85 (0.80-0.88)	0.98 (0.66-1.45)	0.907
EMR3_CpG_1	0.52 (0.42-0.65)	0.50 (0.44-0.59)	0.84 (0.65-1.08)	0.172
EMR3_CpG_2	0.00 (0.00-0.02)	0.00 (0.00-0.09)	1.96 (1.10-3.50)	<b>0.022</b>
EMR3_CpG_3	0.06 (0.03-0.08)	0.08 (0.05-0.13)	3.10 (1.51-6.34)	<b>0.002</b>
EMR3_CpG_4	0.16 (0.11-0.24)	0.20 (0.15-0.27)	1.51 (1.05-2.16)	<b>0.026</b>

\*Logistic regression, adjusted for age. The bold values indicate  $P < 0.05$ .表3 散发性乳腺癌患者与对照 *EMR3* 基因甲基化水平差异Tab.3 Comparison of methylation levels of *EMR3* gene between 258 sporadic BC cases and 272 control subjects

CpG sites	BC cases (n=258)	Controls (n=272)	OR (95% CI) *	P
	median (IQR)	median (IQR)	per-10% methylation	
EMR3_CpG_1	0.51 (0.44-0.60)	0.51 (0.44-0.60)	0.97 (0.84-1.12)	0.675
EMR3_CpG_2	0.02 (0.00-0.08)	0.03 (0.00-0.10)	1.08 (0.84-1.37)	0.560
EMR3_CpG_3	0.08 (0.05-0.13)	0.09 (0.05-0.13)	1.08 (0.84-1.38)	0.576
EMR3_CpG_4	0.19 (0.14-0.26)	0.19 (0.14-0.26)	0.92 (0.76-1.10)	0.350
EMR3_CpG_5	0.09 (0.07-0.13)	0.09 (0.07-0.12)	0.75 (0.56-1.00)	0.051

\*Logistic regression, adjusted for age and different batches for the measurements.

Spearman  $\rho = -0.189$ ,  $P = 0.002$ ; *EMR3\_CpG\_4*: Spearman  $\rho = -0.128$ ,  $P = 0.040$ , 表4)。而在对照组中, *EMR3\_CpG\_1*、*EMR3\_CpG\_3* 和 *EMR3\_CpG\_4* 三个位点的甲基化水平与年龄呈正相关(*EMR3\_CpG\_1*: Spearman  $\rho = 0.243$ ,  $P = 5.0E-05$ ; *EMR3\_CpG\_3*: Spearman  $\rho = 0.201$ ,  $P = 0.001$ ; *EMR3\_CpG\_4*: Spearman  $\rho = 0.153$ ,  $P = 0.012$ , 表4)。

为进一步探究 *EMR3* 甲基化水平与乳腺癌的相关性是否受年龄影响, 我们以 50 岁为界对研究人群进行年龄分层:  $< 50$  岁组和  $\geq 50$  岁组。前者包括乳腺癌患者 194 例, 健康女性 167 名; 后者包括乳腺癌患者 64 例, 健康女性 105 名。在  $< 50$  岁组别中, 5 个位点均未显示出 *EMR3* 甲基化与乳腺癌的相关性。而在  $\geq 50$  岁的人群中, *EMR3\_CpG\_1*、*EMR3\_CpG\_2* 和 *EMR3\_CpG\_3* 的低甲基化与乳腺癌相关, 且相关性具有统计学意义(每减少 10% 的甲基化; *EMR3\_CpG\_1*: OR=1.40, 95% CI: 1.03~1.89,  $P = 0.032$ ; *EMR3\_CpG\_2*: OR=2.31, 95% CI: 1.17~4.55,  $P = 0.016$ ; *EMR3\_CpG\_3*: OR=2.76, 95% CI:

1.37~5.56,  $P = 0.005$ , 表5)。其他 2 个 *EMR3* 位点的甲基化水平尚未在  $\geq 50$  岁人群中观察到与乳腺癌的相关性(表5)。

#### 2.4 *EMR3* 甲基化与乳腺癌临床特征的关系

Stage I, Stage II 与 Stage III & IV 三组的乳腺癌患者在 *EMR3\_CpG\_2* 位点的甲基化水平存在显著的递减 ( $P = 0.022$ , 表6)。但是其他的 4 个 *EMR3* CpG 位点并未显示出这一趋势。另外, 我们的结果显示 *EMR3* 的甲基化水平与肿瘤大小、淋巴转移、ER、PR、HER2、Ki67 没有相关性(表6)。

### 3 讨论

Xu 等<sup>[20]</sup> 报道了血液中 5 个甲基化位点 (cg03430067, cg03616357, cg07072643, cg08287471, cg19709625) 的 DNA 甲基化水平与乳腺癌相关。本项目旨在中国人群中研究外周血中这 5 个位点相关基因 (*C19orf57*、*MAP9*、*EMR3*、*NEK6*、*PCOLCE2*) 的甲基化水平与乳腺癌的关联。试图说明以下几个方面: (1) 不

表4 *EMR3*甲基化与年龄相关性  
Tab.4 Correlation between *EMR3* methylation and age

CpG sites	BC cases (n=258)		Controls (n=272)	
	Spearman rho	P	Spearman rho	P
EMR3_CpG_1	-0.042	0.500	0.243	<b>5.0E-05</b>
EMR3_CpG_2	-0.189	<b>0.002</b>	0.059	0.329
EMR3_CpG_3	-0.120	0.054	0.201	<b>0.001</b>
EMR3_CpG_4	-0.128	<b>0.040</b>	0.153	<b>0.012</b>
EMR3_CpG_5	0.076	0.226	0.109	0.071

\*Spearman rank correlation. The bold values indicate  $P < 0.05$ .

表5 258例散发性乳腺癌患者和272名健康女性*EMR3*甲基化年龄分层分析  
Tab.5 Methylation difference of *EMR3* between 258 sporadic BC cases and 272 controls stratified by age

CpG sites	EMR3_CpG_1	EMR3_CpG_2	EMR3_CpG_3	EMR3_CpG_4	EMR3_CpG_5
Participants < 50 years old					
BC cases (n=194) median (IQR)	0.51 (0.45-0.60)	0.03 (0.00-0.09)	0.09 (0.05-0.15)	0.20 (0.15-0.26)	0.09 (0.07-0.13)
Controls (n=167) median (IQR)	0.50 (0.42-0.57)	0.03 (0.00-0.10)	0.08 (0.05-0.13)	0.18 (0.14-0.25)	0.08 (0.07-0.12)
OR (95% CI)* per-10% methylation	0.87 (0.71-1.07)	0.89 (0.64-1.23)	0.86 (0.60-1.24)	0.82 (0.63-1.05)	0.96 (0.76-1.21)
P	0.185	0.471	0.416	0.119	0.721
Participants ≥ 50 years old					
BC cases (n=64) median (IQR)	0.50 (0.42-0.64)	0.00 (0.00-0.03)	0.07 (0.03-0.11)	0.17 (0.11-0.24)	0.09 (0.07-0.16)
Controls (n=105) median (IQR)	0.54 (0.48-0.64)	0.03 (0.00-0.10)	0.10 (0.06-0.16)	0.21 (0.14-0.27)	0.10 (0.07-0.15)
OR (95% CI)* per-10% methylation	1.40 (1.03-1.89)	2.31 (1.17-4.55)	2.76 (1.37-5.56)	1.31 (0.87-1.97)	0.94 (0.57-1.53)
P	<b>0.032</b>	<b>0.016</b>	<b>0.005</b>	0.197	0.788

\*Logistic regression, adjusted for age and different batches for the measurements. The bold values indicate  $P < 0.05$ . IQR: Interquartile range.

同种族的甲基化水平与疾病之间的关联可能不同；(2) Xu等<sup>[20]</sup>采用Illumina芯片技术筛选发现了乳腺癌相关的甲基化位点,但并未对这些位点进行逐一验证。我们通过飞行时间质谱(MALDI-TOF)技术对其报道的甲基化位点以及周边的位点都进行了逐一的验证,从而提升了数据的可靠性和可重复性；(3)说明候选基因研究对于我国乳腺癌早期检测潜在生物标志物的发现具有重要意义。同时提示,用于检测的分子标志物不建议直接套用国外人群的研究结果,而需在本国人群中先进行验证。本研究通过文献检索和因为甲基化值低于质谱稳定检测区间排除了3个基因(MAP9、NEK6、PCOLCE2,详见研究材料和方法)。在小样本研究的88例散发性乳腺癌患者和94名健康对照中,发现C19orf57甲基化水平与乳腺癌不相关。在后续合计258例散发性乳腺癌患者和272名健康女性的研究中则显示,EMR3基因的低甲基化与乳腺癌的相关性受年龄影响,该相关性仅存在于≥50岁的人群中。在Xu等<sup>[19,20]</sup>的研究中,910位参与者的Illumina 27K芯片数据显示EMR3(cg07072643)在乳腺癌患者中的甲基化水平高于健康女性,其后的Illumina 450K芯片数据并未展示该位点的甲基化水平具体数值。但我们的研究却发现EMR3的甲基化水平在大于等于50岁乳腺癌患者中是

低于健康女性的。我们的结果和Xu等<sup>[20]</sup>的报道有明显的差异,可能源于高加索人群和中国人之间的种族差异<sup>[21]</sup>、也可能源于不同的生活方式<sup>[26]</sup>以及环境<sup>[22]</sup>等因素对甲基化水平的影响。

EMR3基因全称adhesion G protein-coupled receptor E3(又名ADGRE3),编码一种表皮生长因子,属于粘附性G蛋白耦合受体,是G蛋白耦合受体大家族中的成员,该家族还包括CD97, EMR1, EMR2和EMR4。该基因在中性粒细胞、单核细胞和巨噬细胞中表达水平最高<sup>[27]</sup>。有研究发现相对于癌旁组织来说,EMR3在肝细胞癌组织中处于低甲基化状态<sup>[28]</sup>。TCGA数据库中肝癌的数据分析提示,EMR3的表达水平与免疫细胞渗透有关<sup>[29]</sup>。EMR3的基因敲除实验证明,该基因在胶质母细胞瘤中的作用与细胞侵袭有关<sup>[30]</sup>。Singh等<sup>[31]</sup>通过全基因组甲基化测量发现食道癌组织中CD97的启动子区处于低甲基化状态且该基因呈高表达。Ward等<sup>[32]</sup>报道组织中CD97的表达上调与甲状腺癌的发展以及侵袭性有关。类似的研究表明CD97是宫颈癌的独立预后指标,与其不良预后以及侵袭性有关<sup>[33]</sup>。目前我们尚未找到EMR3基因的甲基化与乳腺癌相互作用的文章。我们推测与CD97同一家族的EMR3甲基化水平改变可能同样通过增加侵袭性与肿瘤的发展相

表6 散发性乳腺癌患者临床特征与EMR3甲基化的关联分析

Tab.6 Association between EMR3 methylation and clinical characteristics of sporadic BC cases

Clinical characteristics (n)	Group (n)	Median of age	Median of methylation levels				
			EMR3_CpG_1	EMR3_CpG_2	EMR3_CpG_3	EMR3_CpG_4	EMR3_CpG_5
Tumor stage (249)							
	Stage I (109)	46.0	0.50	0.02	0.07	0.19	0.08
	Stage II (92)	45.0	0.51	0.02	0.08	0.21	0.09
	Stage III & IV (48)	47.0	0.56	0.00	0.07	0.19	0.09
	<i>P</i>	0.773	0.088	<b>0.022</b>	0.602	0.800	0.500
Tumor size (249)							
	T0&1 (153)	46.0	0.50	0.01	0.08	0.20	0.09
	T2&3&4 (96)	45.0	0.53	0.02	0.08	0.19	0.09
	<i>P</i>	0.386	0.589	0.346	0.729	0.400	0.930
Lymph node involvement (249)							
	N0 (139)	45.0	0.50	0.02	0.08	0.19	0.08
	N1&2&3 (110)	46.0	0.53	0.01	0.08	0.20	0.09
	<i>P</i>	0.908	0.234	0.403	0.514	0.754	0.338
Ki67 (251)							
	Ki67≤20% (97)	46.0	0.51	0.02	0.07	0.19	0.08
	Ki67>20% (154)	45.0	0.51	0.02	0.08	0.20	0.09
	<i>P</i>	0.920	0.539	0.791	0.312	0.984	0.227
ER (254)							
	ER negative (52)	46.0	0.48	0.02	0.08	0.20	0.08
	ER positive (202)	45.0	0.51	0.02	0.08	0.19	0.09
	<i>P</i>	0.332	0.247	0.568	0.503	0.652	0.399
PR (254)							
	PR negative (65)	46.0	0.49	0.02	0.08	0.20	0.08
	PR positive (189)	45.0	0.51	0.01	0.08	0.19	0.09
	<i>P</i>	0.474	0.276	0.494	0.404	0.799	0.261
HER2 (254)							
	HER2 negative (197)	45.0	0.51	0.01	0.08	0.20	0.09
	HER2 positive (57)	46.0	0.51	0.02	0.07	0.19	0.09
	<i>P</i>	0.609	0.758	0.499	0.925	0.772	0.801
Triple-negative (254)							
	Triple-negative (33)	45.0	0.48	0.02	0.08	0.21	0.09
	Non-triple-negative (221)	46.0	0.51	0.02	0.08	0.19	0.09
	<i>P</i>	0.812	0.621	0.802	0.335	0.233	0.492

关。可惜我们没有新鲜外周血样本以及RNA样本,也没有找到相关数据库能提供乳腺癌患者外周血中EMR3基因甲基化与表达的关系。而且本研究的结果受限于样本量较小与回顾性病例对照的研究设计。因此未来在前瞻性队列中收集乳腺癌各项影响因素的数据,并进行相关研究是十分有必要的。

本研究发现EMR3基因甲基化水平与乳腺癌的相关性仅在50岁以上女性中存在(表5)。亚洲妇女平均绝经年龄约为49周岁<sup>[34]</sup>,围绝经期通常是指绝经前3~4年<sup>[35]</sup>,也就是45~49周岁。Key等<sup>[36]</sup>发现中国女性的雌

激素水平随年龄的增加而降低。遗憾的是,我们并没有雌激素水平数据。因此暂定50作为年龄分层分析的界值。年龄是甲基化改变的重要因素<sup>[37]</sup>同时,雌激素水平也跟女性年龄密切相关,尤其是青春期以及绝经期的激素水平变化<sup>[38]</sup>。有研究表明DNA甲基化的改变与青春期过渡和生殖激素水平的变化亦极具相关性<sup>[39-41]</sup>。并且有关激素与乳腺癌的研究也指出,跟乳腺相关的激素(如雌二醇、睾酮、催乳素等)与乳腺癌患病风险呈正相关关系<sup>[42, 43]</sup>。根据以上证据,我们推测50岁以上女性群体(绝大多数处于绝经期)EMR3甲基化与乳腺癌的相

关性可能与年龄及其自身雌激素水平有关。这与 Yin 等<sup>[44]</sup>发现的 *S100P* 和 *HYAL2* 基因甲基化水平与乳腺癌的相关关系受激素影响结果类似。然而 *EMR3* 基因甲基化与乳腺癌相关的具体机制还有待后续的功能实验研究。

在既往研究中, Xu 等<sup>[19, 20]</sup>并未比较不同临床特征乳腺癌患者的 *EMR3* 甲基化水平差异。我们的研究还针对中国乳腺癌患者不同临床特征的 *EMR3* 甲基化水平进行了分组比较, 结果显示不同分期的乳腺癌患者 *EMR3* 甲基化水平在个别位点上存在差异。同时, *EMR3* 的甲基化水平与肿瘤大小、淋巴转移、ER、PR、HER2、Ki67 也没有相关性。由此可见, 外周血 DNA 甲基化的改变是发生在肿瘤早期的, 并且跟肿瘤的亚型关系不大。如果能够找到高效特异的外周血甲基化信号, 则有可能是潜在的肿瘤早筛早诊标志物。

综上所述, 本次研究在中国女性人群中发现外周血 *EMR3* 基因的低甲基化与早期乳腺癌之间存在相关性, 尤其存在于年龄较大或者绝经以后的女性中。我们的工作为进一步探索 *EMR3* 基因甲基化与乳腺癌的作用机制提供了线索, 同时也再次提示基因甲基化与疾病之间的相关性需在不同人种中进行独立验证。

#### 参考文献:

- [1] 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): 209-49.
- [2] Getachew S, Tesfaw A, Kaba M, et al. Perceived barriers to early diagnosis of breast Cancer in south and southwestern Ethiopia: a qualitative study[J]. BMC Women's Heal, 2020, 20(1): 1-8.
- [3] Wang LL. Early diagnosis of breast cancer[J]. Sensors, 2017, 17(7): 1572.
- [4] Medical Advisory Secretariat. Cancer screening with digital mammography for women at average risk for breast cancer, magnetic resonance imaging (MRI) for women at high risk: an evidence-based analysis[J]. Ont Health Technol Assess Ser, 2010, 10: 1-55.
- [5] Bae JM, Kim EH. Breast density and risk of breast cancer in Asian women: a meta-analysis of observational studies[J]. J Prev Med Public Health, 2016, 49(6): 367-75.
- [6] Kwong A, Cheung P, Chan S, et al. Breast cancer in Chinese women younger than age 40: are they different from their older counterparts [J]? World J Surg, 2008, 32(12): 2554-61.
- [7] Jalalian A, Mashohor S, Mahmud R, et al. Foundation and methodologies in computer-aided diagnosis systems for breast cancer detection[J]. Excli J, 2017, 16: 113-37.
- [8] Kuhl CK. Abbreviated magnetic resonance imaging (MRI) for breast cancer screening: rationale, concept, and transfer to clinical practice[J]. Annu Rev Med, 2019, 70: 501-19.
- [9] Valdora F, Houssami N, Rossi F, et al. Rapid review: radiomics and breast cancer[J]. Breast Cancer Res Treat, 2018, 169(2): 217-29.
- [10] Goldberg AD, Allis CD, Bernstein E. Epigenetics: a landscape takes shape[J]. Cell, 2007, 128(4): 635-8.
- [11] Moore LD, Le T, Fan GP. DNA methylation and its basic function [J]. Neuropsychopharmacology, 2013, 38(1): 23-38.
- [12] Cheung HH, Lee TL, Rennert OM, et al. DNA methylation of cancer genome[J]. Birth Defect Res C, 2009, 87(4): 335-50.
- [13] Robertson KD. DNA methylation and human disease[J]. Nat Rev Genet, 2005, 6(8): 597-610.
- [14] Kim SY, Han YK, Song JM, et al. Aberrantly hypermethylated tumor suppressor genes were identified in oral squamous cell carcinoma (OSCC)[J]. Clin Epigenetics, 2019, 11(1): 1-12.
- [15] Min HL, Kim J, Kim WH, et al. Epigenetic silencing of the putative tumor suppressor gene GLDC (Glycine dehydrogenase) in gastric carcinoma[J]. Anticancer Res, 2016, 36(1): 179-87.
- [16] Gaudet F. Induction of tumors in mice by genomic hypomethylation [J]. Science, 2003, 300(5618): 489-92.
- [17] Laird PW. The power and the promise of DNA methylation markers [J]. Nat Rev Cancer, 2003, 3(4): 253-66.
- [18] Severi G, Southey MC, English DR, et al. Epigenome-wide methylation in DNA from peripheral blood as a marker of risk for breast cancer[J]. Breast Cancer Res Treat, 2014, 148(3): 665-73.
- [19] Xu ZL, Bolick SCE, DeRoo LA, et al. Epigenome-wide association study of breast cancer using prospectively collected sister study samples[J]. JNCI: J Natl Cancer Inst, 2013, 105(10): 694-700.
- [20] Xu Z, Sandler DP, Taylor JA. Blood DNA methylation and breast cancer: a prospective case-cohort analysis in the sister study[J]. J Natl Cancer Inst, 2020, 112(1): 87-94.
- [21] BCD Lynn, C Bodelon, RM Pfeiffer, et al. Differences in genome-wide DNA methylation levels in breast milk by race and lactation duration[J]. Cancer Prev Res, 2019, 12(11): 781-90.
- [22] Martin EM, Fry RC. Environmental influences on the epigenome: exposure-associated DNA methylation in human populations [J]. Annu Rev Public Health, 2018, 39: 309-33.
- [23] Guan Z, Yu HX, Cuk K, et al. Whole-blood DNA methylation markers in early detection of breast cancer: a systematic literature review [J]. Cancer Epidemiol Biomarkers Prev, 2019, 28(3): 496-505.
- [24] Yang R, Pfützte K, Zucknick M, et al. DNA methylation array analyses identified breast cancer-associated *HYAL2* methylation in peripheral blood[J]. Int J cancer, 2015, 136: 1845-55.
- [25] Yang RX, Stöcker S, Schott S, et al. The association between breast cancer and *S100P* methylation in peripheral blood by multicenter case-control studies[J]. Carcinogenesis, 2017, 38(3): 312-20.
- [26] Sapienza C, Issa JP. Diet, nutrition, and cancer epigenetics[J]. Annu Rev Nutr, 2016, 36: 665-81.
- [27] Matmati M, Pouwels W, van Bruggen R, et al. The human EGF-TM7 receptor *EMR3* is a marker for mature granulocytes [J]. J Leukoc Biol, 2007, 81(2): 440-8.
- [28] Archer KJ, Mas VR, Maluf DG, et al. High-throughput assessment of CpG site methylation for distinguishing between HCV-cirrhosis and HCV-associated hepatocellular carcinoma [J]. Mol Genet Genom, 2010, 283(4): 341-9.
- [29] Xiang SX, Li J, Shen J, et al. Identification of prognostic genes in the tumor microenvironment of hepatocellular carcinoma[J]. Front Immunol, 2021, 12: 653836.

- [30] Kane AJ, Sughrue ME, Rutkowski MJ, et al. EMR-3: a potential mediator of invasive phenotypic variation in glioblastoma and novel therapeutic target[J]. *Neuroreport*, 2010, 21: 1018-22.
- [31] Singh V, Singh LC, Vasudevan M, et al. Esophageal cancer epigenomics and integrative analysis of genome-wide methylation and expression in high risk northeast Indian population[J]. *OMICS: A J Integr Biol*, 2015, 19(11): 688-99.
- [32] Ward Y, Lake R, Martin PL, et al. CD97 amplifies LPA receptor signaling and promotes thyroid cancer progression in a mouse model[J]. *Oncogene*, 2013, 32(22): 2726-38.
- [33] He Y, Wang W, Xu L, et al. Immunohistochemical Expression and Prognostic Significance of CD97 and its Ligand DAF in Human Cervical Squamous Cell Carcinoma[J]. *Int J Gynecol Pathol Off J Int Soc Gynecol Pathol*, 2015, 34: 473-9.
- [34] Schoenaker DA, Jackson CA, Rowlands JV, et al. Socioeconomic position, lifestyle factors and age at natural menopause: a systematic review and meta-analyses of studies across six continents [J]. *Int J Epidemiol*, 2014, 43(5): 1542-62.
- [35] Dunneram Y, Greenwood DC, Cade JE. Diet, menopause and the risk of ovarian, endometrial and breast cancer[J]. *Proc Nutr Soc*, 2019, 78(3): 438-48.
- [36] Key T, Chen J, Wang D, et al. Sex hormones in women in rural China and in Britain[J]. *Br J Cancer*, 1990, 62(4): 631-6.
- [37] Unnikrishnan A, Freeman WM, Jackson J, et al. The role of DNA methylation in epigenetics of aging[J]. *Pharmacol Ther*, 2019, 195: 172-85.
- [38] Park SJ, Goldsmith LT, Weiss G. Age-related changes in the regulation of luteinizing hormone secretion by estrogen in women [J]. *Exp Biol Med Maywood N J*, 2002, 227(7): 455-64.
- [39] Huen KR, Harley K, Kogut K, et al. DNA methylation of LINE-1 and Alu repetitive elements in relation to sex hormones and pubertal timing in Mexican-American children [J]. *Pediatr Res*, 2016, 79(6): 855-62.
- [40] Almstrup K, Lindhardt Johansen M, Busch AS, et al. Pubertal development in healthy children is mirrored by DNA methylation patterns in peripheral blood[J]. *Sci Rep*, 2016, 6(1): 1-12.
- [41] Yuan XL, Li ZH, Ye SP, et al. Genome-wide DNA methylation analysis of pituitaries during the initiation of puberty in gilts [J]. *PLoS One*, 2019, 14(3): e0212630.
- [42] Key T, Appleby P, Barnes I, et al. Endogenous sex hormones and breast cancer in postmenopausal women: reanalysis of nine prospective studies[J]. *J Natl Cancer Inst*, 2002, 94: 606-16.
- [43] Gompel A. Hormones and breast cancer [J]. *Hum Reprod Updat*, 2004, 10(4): 281-93.
- [44] Yin QM, Yang XQ, Li LX, et al. The association between breast cancer and blood-based methylation of S100P and HYAL2 in the Chinese population[J]. *Front Genet*, 2020, 11: 977.

(编辑:吴锦雅)