

基于TCGA数据库结肠癌*RBL1*基因 甲基化水平的诊断与预后分析

潘冉冉 季慧慧 黄天怡 胡豪畅 段世伟*

(宁波大学医学院医学遗传中心, 宁波 315211)

摘要 视网膜母细胞瘤样蛋白1(RBL1/p105)是视网膜母细胞瘤蛋白质家族的成员之一, *RBL1*通常被认为是抑癌基因, 在细胞增殖、凋亡、分化中发挥重要作用。该文探讨*RBL1*基因在结肠癌诊断和预后中的作用和可能的机制。利用癌症基因组图谱(TCGA)数据库中的结肠癌病例资料和芯片数据, 筛选*RBL1*基因差异甲基化位点, 并通过Cox回归模型研究甲基化位点与结肠癌患者预后之间的关系。该研究发现, 待研究的癌组织*RBL1*基因启动子甲基化水平高于癌旁组织(0.120 ± 0.012 vs 0.113 ± 0.008 , $P=0.000\ 04$)。随后, 研究者采用受试者工作特征(ROC)曲线来评价*RBL1*基因启动子甲基化水平的诊断价值, 用曲线下面积(AUC)作为诊断价值的评判标准。研究发现, 启动子区甲基化的AUC达到0.732, 对应的灵敏度为80.5%、特异度为55.3%。Cox多因素分析结果发现, cg04086771高甲基化是结肠癌患者预后的独立危险因素($P=0.002$)。该研究通过对TCGA数据库的挖掘, 发现*RBL1*基因启动子区甲基化均值可能与结肠癌发病风险有关, 同时*RBL1*基因的甲基化位点的甲基化水平对结肠癌的预后有影响。

关键词 结肠癌; TCGA; *RBL1*基因; 甲基化; 生物标志物; 预后

Analysis of Potential Diagnostic and Prognostic Value of the Methylation Levels of *RBL1*: A Study Based on TCGA Database

Pan Ranran, Ji Huihui, Huang Tianyi, Hu Haochang, Duan Shiwei*

(Medical Genetics Center, School of Medicine, Ningbo University, Ningbo 315211, China)

Abstract Retinoblastoma-like protein 1 (RBL1/p105) is a member of the retinoblastoma protein family. *RBL1* is generally considered to be a tumor suppressor gene and plays an important role in cell proliferation, apoptosis and differentiation. This article explores the role of *RBL1* promoter methylation in diagnosis and prognosis of colon cancer and its potential molecular mechanism. Using the methylation and expression microarray data from The Cancer Genome Atlas (TCGA), we screened for differential methylation sites of the *RBL1* in colon cancer, and researched the role of *RBL1* methylation sites in the prognosis of colon cancer patients using Cox regression model. This study found that the promoter methylation level of *RBL1* in tumor tissues was higher than that in adjacent no-tumor tissues (0.120 ± 0.012 vs 0.113 ± 0.008 , $P=0.000\ 04$). Subsequently, we used the receiver operating characteristic (ROC) curve to evaluate the diagnostic value of the *RBL1* promoter methylation, using the area under the curve (AUC) as the criterion for the diagnostic value. We found that the promoter methylation of *RBL1* could reach an AUC of 0.732 with a sensitivity of 80.5% and a specificity of 55.3%. Cox multiple regression

收稿日期: 2018-07-18 接受日期: 2018-12-14

宁波大学王宽诚基金资助的课题

*通讯作者。Tel: 0574-87609950, E-mail: duanshiwei@nbu.edu.cn

Received: July 18, 2018 Accepted: December 14, 2018

This work was supported by the K.C. Wong Magna Fund in Ningbo University

*Corresponding author. Tel: +86-574-87609950, E-mail: duanshiwei@nbu.edu.cn

网络出版时间: 2019-02-21 15:30:11

URL: <http://kns.cnki.net/kcms/detail/31.2035.Q.20190221.1529.034.html>

found that cg04086771 hypermethylation was an independent risk factor for prognosis in colon cancer patients ($P=0.002$). This study found that the average methylation status of *RBL1* promoter region may associate with the risk of colon cancer. Meanwhile, we found that the methylation level of *RBL1* had an important impact in the prognosis of colon cancer.

Keywords colon cancer; TCGA; *RBL1* gene; DNA methylation; biomarker; prognosis

结肠癌是人类高发的恶性肿瘤之一, 历年的发病率和病死率呈上升趋势^[1]。目前结肠癌的分子致癌机制尚未完全阐明, 但有研究表明, 结肠癌的发生和发展会伴有遗传、表观遗传和环境因素的失调^[2]。表观遗传学作为遗传和环境因素之间的桥梁, 在癌症的发生和发展中起重要作用。DNA甲基化作为表观遗传学重要组成部分, 是结肠癌发生和发展过程中的基本分子机制之一^[3-6], 其主要表现为肿瘤抑制基因功能失调^[7]以及致癌基因的激活^[8]。

*RBL1/P130*是抑癌基因*Rb*基因家族成员之一, 作为关键的肿瘤抑制因子, 在细胞增殖、凋亡、分化中发挥重要作用^[9-10]。细胞周期失调是肿瘤发生、发展的重要特征之一^[11]。RB蛋白可以通过与转录调节因子E2F家族成员的相互作用来抑制G₁-S期转变^[12-13]。它的失活和移位被认为是多种类型肿瘤的基础, 包括肺癌^[14-15]、前列腺癌^[11]、结直肠癌^[16]等。DNA甲基化作为最重要的表观遗传修饰之一^[17], 主要发生在启动子区, 导致肿瘤抑制基因的沉默^[18-19], 在癌症的发生、发展中起重要作用^[20-21]。

粪便隐血试验(FOBT)和结肠镜检查是目前广泛应用于检查结肠癌的两种方法, 前者结果不稳定、易产生假阳性^[22], 后者可能会导致严重并发症, 如肠道消化道出血、肠穿孔等, 且患者依从性较差^[23]。同时, 大量研究表明, 基因甲基化水平与癌症的发病风险相关^[24-26]。

因此, 本研究通过对TCGA(The Cancer Genome Atlas)数据库的数据进行挖掘, 研究*RBL1*基因cg位点甲基化水平在结肠癌中的诊断价值, 探索*RBL1*基因甲基化水平与结肠癌预后的关联和机制, 为今后结肠癌诊断、预后相关生物标志物的研究提供科学依据。

1 材料与方法

1.1 TCGA数据库

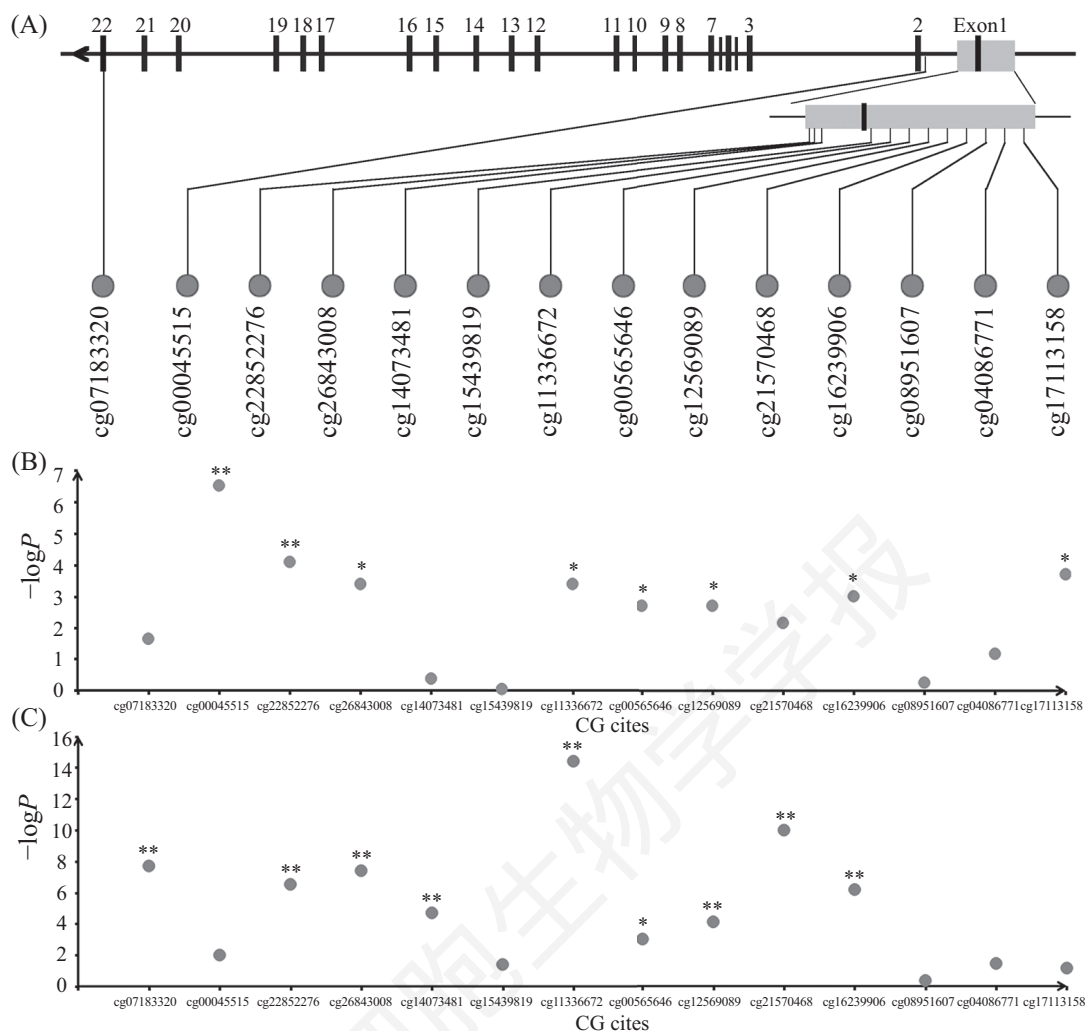
本研究采用从UCSC Xena网站(<https://xenabrowser.net/>)下载的数据集, 该网站以TCGA数据为基础提供较

为完善的多项信息检索服务。入选样本采用统一的标准: (1)病理确诊为腺癌的结肠癌患者; (2)有完整的预后随访资料以及甲基化数据; (3)具有临床资料。最终确定了297名结肠癌患者。

下载数据集包括: (1)结肠癌的临床预后数据; (2)基于美国Illumina公司Human Methylation 450K BeadChip(HM450K)甲基化芯片数据检测的全基因组甲基化的数据; (3)Illumina HiSeq_RNA-Seq Version 2转录组测序数据。从UCSC Xena网站可获得*RBL1*基因上14个不同cg位点的甲基化数据, 每个cg位点的基因坐标基于UCSC Genome Browser GRCh37(<https://genome-cancer.soe.ucsc.edu/>), 详见图1。因此, 本文主要探索*RBL1*基因上共计14个甲基化位点(cg位点)与结肠癌诊断、预后之间的关联。

基因启动子被定义为从转录起始位点(transcription start site, TSS)开始的上下游的一段DNA序列^[27], 能活化RNA聚合酶, 使之与模板DNA准确地结合并具有转录起始的特异性, 是基因转录调控的核心区域^[28]。我们选取TSS上游2 000 bp以及下游500 bp的区域作为*RBL1*基因的启动子区域(-2 000 bp~+499 bp)。14个cg位点中有12个位于*RBL1*基因启动子区域, 详见图1。相较于基因体区域, 基因的启动子区域与基因的调控和特异性表达有着十分密切的关系^[27]。Danino等^[29]发现, 核心启动子在调控多细胞动物基因表达中起着关键作用。考虑到单个cg位点的甲基化水平难以代表基因启动子的甲基化水平, 我们用位于启动子区域的12个cg位点的甲基化的平均值代表*RBL1*基因启动子的甲基化水平, 用于综合考虑*RBL1*基因启动子的甲基化水平。

Human Methylation 450K BeadChip(HM450K)甲基化芯片数据中每个cg位点对应的甲基化水平可用 β 值来反映, 其中 $\beta=U/(M+U+1)$, U为未甲基化探针的信号值, M为甲基化探针的信号强度^[30]。DNA甲基化 β 值是0和1之间的连续变量, 表示甲基化珠子类型的强度与组合的基因座强度的比率。



A: 基因组图谱; B: 病例组与对照组各cg位点甲基化水平差异; C: 各CG位点甲基化水平与mRNA表达水平相关性分析。** $P < -\log(0.05/N)$, N: CG位点数量, **** $P < 0.0001$ 。

A: genome mapping; B: the difference of methylation levels of CG sites between the case and the control; C: correlation analysis between methylation level of CG sites and mRNA expression level. ** $P < -\log(0.05/N)$, N: number of CG sites, **** $P < 0.0001$ 。

图1 RBL1基因各cg位点分析

Fig.1 Annotation of cg sites in RBL1 gene

1.2 统计分析

采用One-sample kolmogrov-smirnov test对计量资料进行正态性检验, 服从正态分布的计量资料表示为均数±标准差, 不服从正态分布的计量资料表示为中位数(四分位数间距)[Median(P_{25} , P_{75})]。服从正态分布的计量资料病例-对照组间比较采用两独立样本t检验, 否则采用Mann-Whitney U非参数检验。受试者工作特征(ROC)曲线用于评价RBL1基因甲基化水平的诊断价值。二分类或多分类变量进行单因素生存分析时, 用Kaplan-Meier法绘制生存曲线, 并计算术后生存率; 用Log-rank检验进行两组或多组生存曲线的比较。连续变量(如甲基化水平)进行单因素生存分析时,

采用Cox比例风险模型。Cox模型还可用于多因素分析, 在分析预后不良相关的因素时, 可先进行单因素分析, 单因素分析有统计学意义的因素以及根据专业判断有必要纳入的因素(甲基化水平)再进行多因素分析。统计检验均为双侧检验, $P < 0.05$ 时差异有统计学意义。

2 结果

2.1 纳入分析的人群基本特征

本研究共计纳入335例样本, 其中结肠癌组织共计297例, 作为对照的癌旁组织有38例。297例结肠癌患者平均年龄为 65.00 ± 13.31 岁, 其中男性有158例, 女性136例, 3例性别资料缺失; 180例患者没有

淋巴结转移, 79例患者发生淋巴结转移, 38例淋巴结转移资料缺失; 黄种人11例, 白种人206例, 黑种人59人, 21例种族资料缺失; I期(IA和IB) 44例, II期(II、IIA和IIIB) 114例, III期(III、IIIA、IIIB和IIIC) 86例, IV期(IV和IVA) 41例, 12例TNM分期资料缺失。随访队列中最长随访时间为4 502天, 有69例死亡, 217例存活, 11例失访, 中位生存期为2 746天。

2.2 与结肠癌相关的cg位点

CpGs的基因组坐标基于UCSC网站数据(<https://genome-cancer.soe.ucsc.edu/>)(图1A)。我们用位于基因启动子区域的12个cg位点的甲基化水平的均值代表*RBL1*基因启动子甲基化水平, 结果发现, 癌组织*RBL1*基因启动子甲基化水平高于癌旁组织(0.120 ± 0.012 vs 0.113 ± 0.008 , $P=4E-6$, 表1)。采用两独立样本*t*检验/Mann-Whitney U非参数检验筛选具有差异的甲基化位点, 发现cg00045515、cg00565646、cg07183320、cg11336672、cg12569089、cg16239906、cg17113158、cg22852276、cg26843008的甲基化水平在结肠癌组织与癌旁组织之间的差异具有统计学意义(图1B)。

2.3 各cg位点甲基化水平与mRNA表达水平的相关性分析

用Spearman/Pearson检验*RBL1*基因各位点的甲基化水平和*RBL1*基因表达水平相关性(图1C)。研究发现, cg00045515($r=0.148$, $P=0.010$)、cg22852276($r=0.293$, $P=3E-7$)、cg26843008($r=0.312$,

$P=4E-8$)、cg00565646($r=0.187$, $P=0.001$)、cg12569089($r=0.227$, $P=8E-5$)、cg21570468($r=0.362$, $P=1E-10$)、cg16239906($r=0.282$, $P=7E-7$)、cg04086771($r=0.121$, $P=0.036$)的甲基化水平与基因表达水平呈正相关, cg07183320($r=-0.320$, $P=2E-8$)、cg14073481($r=-0.247$, $P=2E-5$)、cg15439819($r=-0.119$, $P=0.040$)、cg11336672($r=-0.433$, $P=4E-15$)的甲基化水平与基因表达水平呈负相关(图1C)。

2.4 各cg位点甲基化水平的诊断与预后分析

我们采用受试者工作特征(ROC)曲线来评价*RBL1*基因启动子甲基化水平的诊断价值, 用曲线下面积(AUC)作为诊断价值的评判标准。我们发现, 启动子区甲基化的AUC达到0.732, 甲基化水平(PMR)为0.092时取到ROC最佳点, 对应的灵敏度为80.5%、特异度为55.3%($P=3E-6$, 图2)。

Kaplan-Meier生存分析结果显示: 男、女患者的中位生存时间差异存在统计学意义($\chi^2=3.893$, $P=0.048$, 图3A); 无淋巴结转移患者的中位生存时间与淋巴结已转移患者的中位生存时间差异存在统计学意义($\chi^2=8.119$, $P=0.004$, 图3B); I期患者、II期患者、III期、IV期患者的中位生存时间存在明显递减趋势, 差异具有统计学意义($\chi^2=32.729$, $P=4E-7$, 图3C); 黄种人、白种人或黑种人的中位生存时间未见明显差异($\chi^2=0.053$, $P=0.974$, 图3D)。经多变量Cox比例风险模型校正后, 发现cg04086771高甲基化是结肠

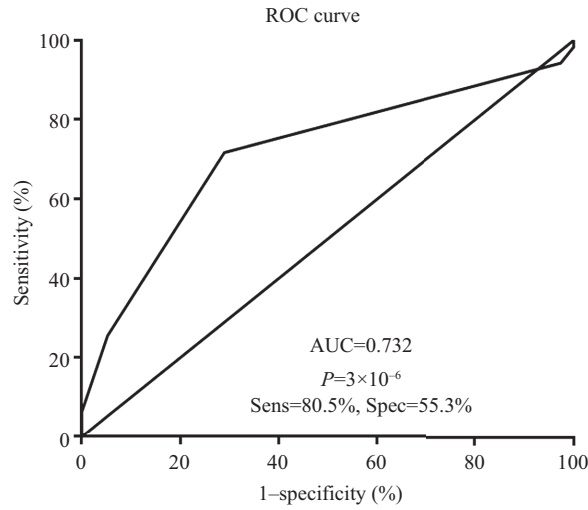
表1 病例组与对照组的差异分析

Table 1 Variation analysis between case and control

变量 Variables	N(病例/对照) Number (case/control)	病例组 Case	对照组 Control	P值 P value
Promoter	297/38	0.100 ± 0.010	0.093 ± 0.007	4×10⁻⁶
cg07183320 ^a	297/38	0.938 (0.925, 0.947)	0.943 (0.933, 0.951)	0.023
cg00045515 ^a	297/38	0.553 ± 0.170	0.457 ± 0.087	3×10⁻⁷
cg22852276 ^b	297/38	0.043 ± 0.009	0.038 ± 0.007	8×10⁻⁵
cg26843008 ^b	297/38	0.057(0.051, 0.064)	0.052(0.047, 0.056)	4×10⁻⁴
cg14073481 ^b	297/38	0.018 ± 0.004	0.019 ± 0.003	0.420
cg15439819 ^b	297/38	0.013 ± 0.003	0.013 ± 0.002	0.901
cg11336672 ^b	297/38	0.020 ± 0.003	0.021 ± 0.002	4×10⁻⁴
cg00565646 ^b	297/38	0.048 (0.041, 0.058)	0.043 (0.039, 0.046)	0.002
cg12569089 ^b	297/38	0.070 (0.057, 0.087)	0.056 (0.052, 0.067)	0.002
cg21570468 ^b	297/38	0.049 ± 0.011	0.044 ± 0.010	0.007
cg16239906 ^b	297/38	0.051 ± 0.011	0.044 ± 0.011	0.001
cg08951607 ^b	297/38	0.040 (0.031, 0.048)	0.041 (0.031, 0.052)	0.552
cg04086771 ^b	297/38	0.057 (0.048, 0.071)	0.052 (0.049, 0.060)	0.067
cg17113158 ^b	297/38	0.714 ± 0.086	0.678 ± 0.048	2×10⁻⁴

粗体代表 $P<0.05$ 。a表示cg位点位于基因体区域; b表示cg位点位于基因启动子区域。

Bold represents $P<0.05$; a represents that cg sites are located in the gene body region; b represents that cg sites are located in the gene promoter region.



ROC曲线显示甲基化水平为0.092时取到ROC最佳点, 对应的灵敏度为80.5%, 特异度为55.3%。

ROC curve showed that the optimal point was obtained when the methylation level was 0.092, with the corresponding sensitivity of 80.5% and specificity of 55.3%.

图2 *RBL1*基因甲基化水平在结肠癌中的诊断价值

Fig.2 Diagnostic value of the methylation of *RBL1* gene in the colon cancer

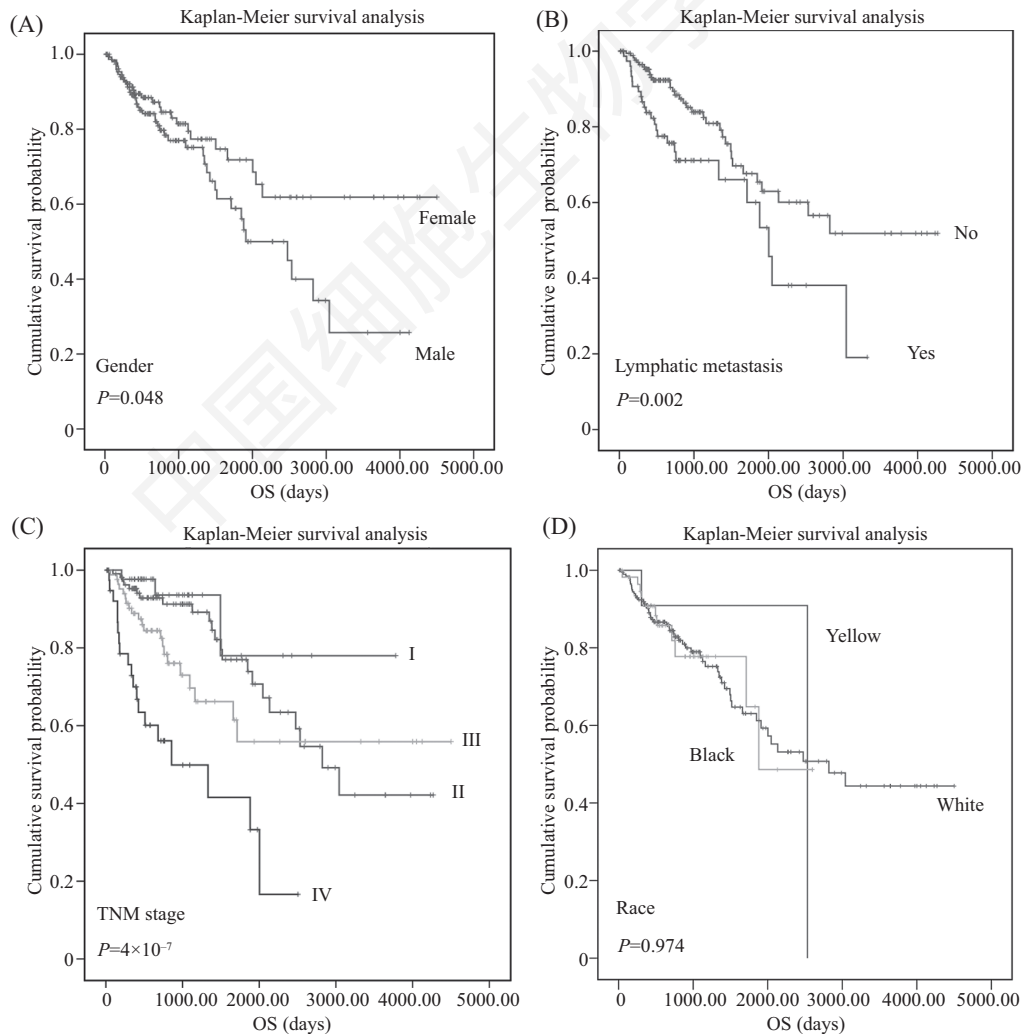


图3 297例结肠癌患者累计生存概率曲线

Fig.3 Cumulative survival probability curve of 297 patients with colon cancer

表2 影响结肠癌预后的COX多因素分析

Table 2 Multivariate COX analysis of prognostic factors in colon cancer

CpG	总体生存率				无病生存率			
	Overall survival (OS)				Disease-free survival (DFS)			
	<i>P</i> [*]	β	<i>SE</i>	<i>P</i> [#]	<i>P</i> [*]	β	<i>SE</i>	<i>P</i> [#]
cg07183320	0.176	10.387	8.831	0.239	0.075	-5.432	6.445	0.399
cg00045515	0.363	-0.459	0.766	0.550	0.134	0.284	0.913	0.756
cg22852276	0.622	27.462	19.884	0.167	0.572	7.405	22.130	0.738
cg26843008	0.256	-18.465	18.186	0.310	0.125	3.737	3.112	0.230
cg14073481	0.318	-52.660	72.009	0.465	0.696	-29.150	72.359	0.687
cg15439819	0.173	42.046	119.530	0.725	0.220	151.442	102.538	0.140
cg11336672	0.410	-77.990	80.586	0.333	0.786	-64.819	78.815	0.411
cg00565646	0.518	6.283	11.551	0.586	0.749	-15.034	11.693	0.199
cg12569089	0.991	4.827	6.039	0.424	0.417	1.615	6.334	0.799
cg21570468	0.621	-23.514	16.812	0.162	0.200	-33.478	18.281	0.067
cg16239906	0.072	-31.384	17.741	0.077	0.280	-10.457	17.279	0.545
cg08951607	0.747	11.566	15.280	0.799	0.803	0.941	14.384	0.978
cg04086771	0.069	12.176	3.989	0.002	0.799	7.376	6.059	0.223
cg17113158	0.731	0.928	2.006	0.644	0.682	0.911	1.995	0.648

粗体代表 $P < 0.05$ 。总体生存率是指从随机化开始至因任何原因引起死亡的时间; 无病生存率是指从随机化开始至疾病复发或由于疾病进展导致患者死亡的时间; *: 连续变量(如甲基化水平)进行单因素生存分析时, 采用Cox比例风险模型; #: 单因素分析有统计学意义的因素以及根据专业判断有必要纳入的因素(甲基化水平)再进行多因素分析, 采用Cox比例风险模型。

Bold represents $P < 0.05$. Overall survival refers to the time from the beginning of randomization to death due to any cause; Disease-free survival rate refers to the time from the beginning of randomization to the recurrence of disease or the death of patients due to disease progression; *: Cox proportional risk model is used for univariate survival analysis of continuous univariate variables (such as methylation level); #: Cox proportional risk model was used to do multivariate analysis, which includes univariate analysis of statistically significant factors and factors that are necessary to be included according to professional judgment (methylation level).

癌独立危险因素($P=0.002$, 表2)。此外, 我们在远期转移结肠癌患者(IV期)中探索*RBL1*基因异常甲基化对其预后的影响, 发现cg26843008($P=0.028$, 表3)、cg04086771($P=0.022$, 表3)是影响结肠癌远期转移患者预后的独立因素。

3 讨论

本研究通过对TCGA数据库进行挖掘, 初步发现, *RBL1*基因14个cg位点里面有9个位点甲基化水平在病例对照组之间存在差异。选择位于*RBL1*基因启动子区的12个cg位点的均值来代表其启动子区的总体甲基化水平, 发现癌组织*RBL1*基因启动子区甲基化水平高于癌旁组织。在此基础上分析甲基化水平在结肠癌中的诊断价值, 发现*RBL1*基因启动子甲基化水平可用于结肠癌筛检, 具有较高的灵敏度(80.5%)。最后发现, 位于*RBL1*基因启动子区的cg04086771与结肠癌预后存在显著关联。

结肠癌的早期诊断和治疗仍是亟待解决的难题, 特别是结肠癌很大程度上可以在早期被治疗。有研究表明, 定期结肠镜检查来切除结肠癌前病变

组织是预防结肠癌的有力工具^[31]。结肠镜检查技术由于成本太高, 具有侵入性和劳动密集性, 显然无法作为人群范围的快速筛查工具。因此, 在肿瘤早期发现快速、低成本的无创性诊断和预后生物标志物是目前的研究热点之一。抑癌基因的DNA高甲基化已被描述为结肠癌主要的表观遗传改变之一^[32], 在多种样本中可以被检测到, 例如肿瘤组织、血液、血浆和粪便, 并且这些标记物中的一些已经在临床环境中被使用。我们的研究发现, *RBL1*基因启动子区在结肠癌组织中高甲基化, 具有较高的灵敏度(80.5%), 表明其作为结肠癌生物标志物的潜在应用价值, 并为进一步研究提供依据。

结肠癌是常见的恶性肿瘤之一, 严重影响患者的生命健康。晚期结肠癌患者的中位生存时间通常低于两年^[33], 我们在远期转移结肠癌患者中进行生存分析发现, 中位生存时间为858天(图3C)。因此, 寻找新的治疗靶点对靶向治疗药物开发、患者预后判断具有重要意义。有研究发现, 在前列腺癌细胞株中*RBL1*基因显著高甲基化^[11]。Mileo等^[34]在肝癌细胞中发现, *RBL1*基因呈现高甲基化, 高甲基化导

表3 RBL1基因甲基化在远期转移(IV和IVA期)结肠癌患者中的COX多因素分析

Table 3 Multivariate COX analysis of RBL1 gene in colon cancer patients with long-term metastatic (IV and IVA stage)

CpG	总体生存率				无病生存率			
	Overall survival (OS)				Disease-free survival (DFS)			
	P*	β	SE	P#	P*	β	SE	P#
cg07183320	0.164	-20.365	316.673	0.949	0.277	817.100	642.416	0.203
cg00045515	0.497	23.590	23.049	0.306	0.418	-31.642	66.509	0.634
cg22852276	0.286	159.005	333.058	0.633	0.545	-1 954.972	1 951.181	0.316
cg26843008	0.421	-1 065.171	483.771	0.028	0.453	1 739.411	1 907.131	0.362
cg14073481	0.078	305.632	860.078	0.722	0.177	1 435.801	1 554.108	0.356
cg15439819	0.475	-434.128	513.898	0.398	0.657	-1 029.481	2 792.491	0.712
cg11336672	-	-	-	-	-	-	-	-
cg00565646	0.619	-50.025	182.973	0.785	0.263	-1 646.241	1 605.524	0.305
cg12569089	0.061	164.666	91.828	0.073	0.186	1 119.019	1 001.089	0.264
cg21570468	0.242	-514.439	382.272	0.178	0.738	-2 113.169	2 081.822	0.310
cg16239906	0.890	86.974	376.146	0.817	0.497	-3 929.032	3 722.043	0.291
cg08951607	0.162	430.964	321.622	0.180	0.928	-2 294.187	2 614.888	0.380
cg04086771	0.204	466.809	203.223	0.022	0.996	99.801	448.692	0.824
cg17113158	0.628	-35.490	28.185	0.208	0.550	29.230	83.253	0.726

粗体代表 $P < 0.05$ 。总体生存率是指从随机化开始至因任何原因引起死亡的时间; 无病生存率是指从随机化开始至疾病复发或由于疾病进展导致患者死亡的时间; *: 连续变量(如甲基化水平)进行单因素生存分析时, 采用Cox比例风险模型; #: 单因素分析有统计学意义的因素以及根据专业判断有必要纳入的因素(甲基化水平)再进行多因素分析, 采用Cox比例风险模型。

Bold represents $P < 0.05$. Overall survival refers to the time from the beginning of randomization to death due to any cause; Disease-free survival rate refers to the time from the beginning of randomization to the recurrence of disease or the death of patients due to disease progression; *: Cox proportional risk model is used for univariate survival analysis of continuous univariate variables (such as methylation level); #: Cox proportional risk model was used to do multivariate analysis, which includes univariate analysis of statistically significant factors and factors that are necessary to be included according to professional judgment (methylation level).

致RBL1基因表达沉默, 从而促进肝癌发生发展。本次研究我们发现, 癌组织RBL1基因启动子区甲基化水平高于癌旁组织。同时, 通过分析RBL1基因甲基化水平对结肠癌预后的影响我们发现, cg04086771高甲基化是结肠癌的独立危险因素, 这可能是由于作为抑癌基因的RBL1基因高甲基化导致其表达水平下降, 进而起到促进癌症进展的作用引起的。因此, 我们进一步猜想, RBL1基因可能可以作为结肠癌的治疗靶点, 此发现有待进一步实验验证。

值得注意的是, 这是第一项通过TCGA数据库研究RBL1基因在结肠癌患者的癌组织和癌旁组织中的甲基化状态的研究。然而因为常规外周血标志物(CEA、CA199、CA125等)资料的缺乏, RBL1基因异常甲基化的诊断价值需要进一步研究。癌症患者的血浆循环DNA(cf-dna)可能起源于循环肿瘤细胞, 这可以反映微转移和癌细胞侵袭的潜力^[35-36]。在血浆cf-DNA中检查RBL1基因甲基化可以避免组织取样偏差, 因此应在此方面进行下一步验证^[37]。此外, 本研究采用的TCGA数据库是全基因组甲基化水平的高通量检测平台, 具有探针密度高、结果可靠等优点, 但其样本对中国人群的代表性难以评价。意

识到研究的局限性和亮点使我们认为本研究提出的结论仅仅只是一个起点, 在中国人群中的验证试验是我们接下来的研究重点。

总之, 本研究采用现有的TCGA数据库数据, 对RBL1基因与诊断、预后有关的cg位点进行了初步挖掘, 发现一些与结肠癌相关的甲基化位点, 为今后结肠癌诊断、预后研究提供一些方向。

参考文献 (References)

- 1 Chen W, Zheng R, Baade PD, Zhang S, Zeng H, Bray F, *et al.* Cancer statistics in China, 2015. *CA Cancer J Clin* 2016; 66(2): 115-32.
- 2 Carethers JM, Jung BH. Genetics and genetic biomarkers in sporadic colorectal cancer. *Gastroenterology* 2015; 149(5): 1177-90 e3.
- 3 Okugawa Y, Grady WM, Goel A. Epigenetic alterations in colorectal cancer: emerging biomarkers. *Gastroenterology* 2015; 149(5): 1204-25 e12.
- 4 Brenner H, Kloor M, Pox CP. Colorectal cancer. *Lancet* 2014; 383(9927): 1490-502.
- 5 Zhou D, Tang W, Su G, Cai M, An HX, Zhang Y. PCDH18 is frequently inactivated by promoter methylation in colorectal cancer. *Sci Rep* 2017; 7(1): 2819.
- 6 Rezvani N, Alibakhshi R, Vaisi-Raygani A, Bashiri H, Saidijam M. Detection of SPG20 gene promoter-methylated DNA, as a novel

- epigenetic biomarker, in plasma for colorectal cancer diagnosis using the MethyLight method. *Oncol Lett* 2017; 13(5): 3277-84.
- 7 Li M, Gao F, Xia Y, Tang Y, Zhao W, Jin C, *et al.* Filtrating colorectal cancer associated genes by integrated analyses of global DNA methylation and hydroxymethylation in cancer and normal tissue. *Sci Rep* 2016; 6: 31826.
- 8 Li L, Li C, Mao H, Du Z, Chan WY, Murray P, *et al.* Epigenetic inactivation of the CpG demethylase TET1 as a DNA methylation feedback loop in human cancers. *Sci Rep* 2016; 6: 26591.
- 9 Baldi A, Boccia V, Claudio PP, De Luca A, Giordano A. Genomic structure of the human retinoblastoma-related Rb2/p130 gene. *Proc Natl Acad Sci USA* 1996; 93(10): 4629-32.
- 10 Classon M, Harlow E. The retinoblastoma tumour suppressor in development and cancer. *Nat Rev Cancer* 2002; 2(12): 910-7.
- 11 Mishra DK, Chen Z, Wu Y, Sarkissyan M, Koeffler HP, Vadgama JV. Global methylation pattern of genes in androgen-sensitive and androgen-independent prostate cancer cells. *Mol Cancer Ther* 2010; 9(1): 33-45.
- 12 Buchkovich K, Duffy LA, Harlow E. The retinoblastoma protein is phosphorylated during specific phases of the cell cycle. *Cell* 1989; 58(6): 1097-105.
- 13 Sun A, Bagella L, Tutton S, Romano G, Giordano A. From G0 to S phase: a view of the roles played by the retinoblastoma (Rb) family members in the Rb-E2F pathway. *J Cell Biochem* 2007; 102(6): 1400-4.
- 14 Barh D, Jain N, Tiwari S, Field JK, Padin-Iruegas E, Ruibal A, *et al.* A novel in silico reverse-transcriptomics-based identification and blood-based validation of a panel of sub-type specific biomarkers in lung cancer. *BMC Genomics* 2013; 14 Suppl 6: S5.
- 15 Modi S, Kubo A, Oie H, Coxon AB, Rehmatulla A, Kaye FJ. Protein expression of the RB-related gene family and SV40 large T antigen in mesothelioma and lung cancer. *Oncogene* 2000; 19(40): 4632-9.
- 16 Eyvani H, Moghaddaskho F, Kabuli M, Zekri A, Momeny M, Tavakkoly-Bazzaz J, *et al.* Arsenic trioxide induces cell cycle arrest and alters DNA methylation patterns of cell cycle regulatory genes in colorectal cancer cells. *Life Sci* 2016; 167: 67-77.
- 17 Tokarz P, Blasiak J. Role of DNA methylation in colorectal cancer. *Postepy Biochem* 2013; 59(3): 267-79.
- 18 Pfeifer GP, Rauch TA. DNA methylation patterns in lung carcinomas. *Semin Cancer Biol* 2009; 19(3): 181-7.
- 19 Naumov VA, Generozov EV, Zaharjevskaya NB, Matushkina DS, Larin AK, Chernyshov SV, *et al.* Genome-scale analysis of DNA methylation in colorectal cancer using Infinium HumanMethylation450 BeadChips. *Epigenetics* 2013; 8(9): 921-34.
- 20 Sabir M, Baig RM, Ali K, Mahjabeen I, Saeed M, Kayani MA. Retinoblastoma (RB1) pocket domain mutations and promoter hyper-methylation in head and neck cancer. *Cell Oncol (Dordr)* 2014; 37(3): 203-13.
- 21 Jovanovic J, Ronneberg JA, Tost J, Kristensen V. The epigenetics of breast cancer. *Mol Oncol* 2010; 4(3): 242-54.
- 22 Li Y, Song L, Gong Y, He B. Detection of colorectal cancer by DNA methylation biomarker SEPT9: past, present and future. *Biomark Med* 2014; 8(5): 755-69.
- 23 Committee ASoP, Fisher DA, Maple JT, Ben-Menachem T, Cash BD, Decker GA, *et al.* Complications of colonoscopy. *Gastrointest Endosc* 2011; 74(4): 745-52.
- 24 Hrasovec S, Hauptman N, Glavac D, Jelenc F, Ravnik-Glavac M. TMEM25 is a candidate biomarker methylated and down-regulated in colorectal cancer. *Dis Markers* 2013; 34(2): 93-104.
- 25 Warren JD, Xiong W, Bunker AM, Vaughn CP, Furtado LV, Roberts WL, *et al.* Septin 9 methylated DNA is a sensitive and specific blood test for colorectal cancer. *BMC Med* 2011; 9: 133.
- 26 Takane K, Midorikawa Y, Yagi K, Sakai A, Aburatani H, Takayama T, *et al.* Aberrant promoter methylation of PPP1R3C and EFHD1 in plasma of colorectal cancer patients. *Cancer Med* 2014; 3(5): 1235-45.
- 27 Kadonaga JT. Perspectives on the RNA polymerase II core promoter. *Wiley Interdiscip Rev Dev Biol* 2012; 1(1): 40-51.
- 28 Juven-Gershon T, Kadonaga JT. Regulation of gene expression via the core promoter and the basal transcriptional machinery. *Dev Biol* 2010; 339(2): 225-9.
- 29 Danino YM, Even D, Ideses D, Juven-Gershon T. The core promoter: At the heart of gene expression. *Biochim Biophys Acta* 2015; 1849(8): 1116-31.
- 30 Croes L, Beyens M, Fransens E, Ibrahim J, Vanden Berghe W, Suls A, *et al.* Large-scale analysis of DNFA5 methylation reveals its potential as biomarker for breast cancer. *Clin Epigenetics* 2018; 10: 51.
- 31 Brenner H, Stock C, Hoffmeister M. Effect of screening sigmoidoscopy and screening colonoscopy on colorectal cancer incidence and mortality: systematic review and meta-analysis of randomised controlled trials and observational studies. *BMJ* 2014; 348: g2467.
- 32 Kulis M, Esteller M. DNA methylation and cancer. *Adv Genet* 2010; 70: 27-56.
- 33 Lee YC, Lee YL, Chuang JP, Lee JC. Differences in survival between colon and rectal cancer from SEER data. *PLoS One* 2013; 8(11): e78709.
- 34 Mileo AM, Mattarocci S, Matarrese P, Anticoli S, Abbruzzese C, Catone S, *et al.* Hepatitis C virus core protein modulates pRb2/p130 expression in human hepatocellular carcinoma cell lines through promoter methylation. *J Exp Clin Cancer Res* 2015; 34: 140.
- 35 Couraud S, Vaca-Paniagua F, Villar S, Oliver J, Schuster T, Blanche H, *et al.* Noninvasive diagnosis of actionable mutations by deep sequencing of circulating free DNA in lung cancer from never-smokers: a proof-of-concept study from BioCAST/IFCT-1002. *Clin Cancer Res* 2014; 20(17): 4613-24.
- 36 Jung K, Fleischhacker M, Rabien A. Cell-free DNA in the blood as a solid tumor biomarker—a critical appraisal of the literature. *Clin Chim Acta* 2010; 411(21/22): 1611-24.
- 37 Ling ZQ, Lü P, Lu XX, Yu JL, Han J, Ying LS, *et al.* Circulating methylated XAF1 DNA indicates poor prognosis for gastric cancer. *PLoS One* 2013; 8(6): e67195.