

·信息研究·

## 基于生物芯片检测结直肠癌相关基因的甲基化

陈 伟, 王 磊, 叶俊文, 侯煜杰, 谭淑云, 汪普宁, 黄美近\*, 汪建平  
(中山大学 附属第六医院//附属胃肠肛门医院 结直肠外科, 广东 广州 510655)

**摘 要:**【目的】应用微矩阵基因芯片技术筛选结直肠癌甲基化基因。【方法】分别提取 6 对结直肠癌组织和癌旁正常组织 DNA, 超声破碎及甲基化 DNA 富集后与甲基化芯片(NimbleGen human DNA Methylation 3 × 720K CpG)进行杂交, 杂交信号经扫描等处理后采用生物信息学对检测结果进行分析。并对筛选出来的基因进行验证。【结果】一共发现有 2 296 个高甲基化基因, 其中有 293 个基因无文献报道。生物信息学分析显示, 所有的甲基化基因随机分布在 24 对染色体上。聚类分析结果显示, 某些在细胞分裂和肿瘤发生发展中起重要作用的生理过程中存在大量甲基化差异基因, 如 DNA 修复, 细胞周期和侵袭等。初步验证发现, Opmcl 基因在结直肠癌组织及癌旁正常组织中存在甲基化差异。【结论】在结直肠癌中运用 DNA 甲基化芯片进行甲基化基因筛查是一种有效的方法。

**关键词:**甲基化; 结直肠癌; 芯片; Opmcl

**中图分类号:**R73      **文献标志码:**A      **文章编号:**1672-3554(2014)06-0920-05

### Screening for Epigenetically Masked Genes in Colorectal Cancer by Microarray

CHEN Wei, WANG Lei, YE Jun-wen, HOU Yu-jie, TAN Shu-yun, WANG Pu-ning, HUANG Mei-jin\*,  
WANG Jian-ping

(Colorectal Surgery Department, The Sixth Affiliated Hospital, Sun Yat-sen University, Guangzhou 510655, China)

**Abstract:** 【Objective】 We used the DNA methylation microarrays to investigate the hypermethylated genes in colorectal cancer and adjacent normal mucosa. 【Method】 DNA was extracted from six pairs of CRC and adjacent normal mucosa and sonicated into fragments ranging in size from 300-1000 bp, followed by methylated DNA Immunoprecipitation analysis and then sent to NimbleGen for Microarray hybridization according to their standard protocol. The hybridized array is then scanned and the resulting image analyzed. The biologic functions were analyzed by bioinformatics methods. In addition, preliminary validation studies were done in six pairs of samples by MSP (methylation-specific PCR). 【Results】 Compared with the normal tissues, there are 2 296 genes were hypermethylated, in which 293 hypermethylated genes were unreported. All these genes were randomly distributed on all the chromosomes. According to gene ontology analysis, the genes involved in some physiological processes which play important roles in the cell division and the development of tumor are methylated, such as DNA repair, cell cycle and invasion et c. Preliminary validation, the Opmcl in thirty top-ranking genes was shown hypermethylated in six pairs of colorectal cancer (CRC) and adjacent normal mucosa. 【Conclusions】 High density DNA methylation microarrays is an effective method for screening aberrantly methylated genes in CRC.

**Key words:** methylation; colorectal cancer; microarray; Opmcl

[J SUN Yat-sen Univ(Med Sci), 2014, 35(6):920-924]

甲基化是基因表观遗传学改变最常见的类型,在某些情况下,它是肿瘤抑制基因失活的唯一机制。基因高甲基化可以影响很多种细胞通路,如

DNA 修复(hMLH1<sup>[1]</sup>和 MGMT<sup>[2]</sup>), 细胞周期(p16 基因<sup>[3]</sup>和 p14<sup>[4]</sup>)和细胞凋亡(DAPK 的<sup>[5]</sup>和 APAF-1<sup>[6]</sup>)。此外,基因低甲基化也已被确定为多种恶性

收稿日期:2014-06-16

基金项目:国家自然科学基金(81172040);中山大学“985 项目”;广东转化医学公共平台项目(4202037)

作者简介:陈伟,博士,住院医师,研究方向:胃肠肿瘤; \* 通信作者:黄美近,博士,副主任医师, E-mail: 13924073322@139.com

肿瘤相关,如 LINE-1<sup>[7]</sup>。基因甲基化可以被用作肿瘤的早期诊断、判断预后及化疗疗效预测<sup>[8-10]</sup>。本研究采用全基因组甲基化芯片的方法来测定结直肠癌特征性的甲基化谱改变,为进一步研究结直肠癌发生发展的分子机制有着积极的意义。

## 1 材料和方法

### 1.1 患者样本

6对结直肠癌组织及癌旁正常黏膜(距离癌组织有7~8 cm)来自于中山大学附属第六医院术后标本。标准如下:①为了排除X染色体上基因的甲基化,我们选择了6名男性结直肠癌患者(年龄范围为52~63岁, S=59)。②根据最新的TNM标准, T3N0M0 6例,其中高分化腺癌2例,中分化腺癌3例,低分化腺癌1例。③患者未接受术前化疗和/或放疗治疗。组织离体后30 min内置于液氮中,随后保存于-80℃,并经病理检查证实。

### 1.2 组织DNA的提取和超声破碎

取结直肠癌组织及相应的癌周正常组织各约50 mg,用基因组DNA提取试剂盒提取DNA。将抽提得到的基因组DNA用Nanodrop分光光度仪测定浓度,样本的 $D_{260/280}$ 值在1.7~2.0,将DNA定量到200 ng/ $\mu$ L,取200  $\mu$ L,超声破碎到300~1000 bp片段大小。

### 1.3 甲基化DNA的富集

将超声好的DNA样品取500 ng约20  $\mu$ L加入干净离心管中,再加入内参DNA-M及U各2  $\mu$ L。轻微振动以悬浮MBD2b蛋白-sepharose-4B,在4℃下亲和层析仪上结合2 h后用洗涤缓冲液洗涤,500  $\times g$ 离心30 s后,收集离心后的洗脱缓冲液,为富集得到的甲基化DNA。

### 1.4 富集样品的WGA扩增及荧光杂交

采用全基因组扩增试剂盒WGA4(Sigma)进行甲基化DNA富集并扩增。分别将甲基化DNA和其他试剂混合,制备成杂交液,将杂交液加入到芯片(NimbleGen human DNA Methylation 3  $\times$  720K CpG)中,杂交16 h后取出芯片,加入洗脱液和染色液各组分,在全自动洗脱染色工作站中完成洗脱和染色的步骤。该芯片同时也含有阳性、阴性以及非CpG对照区域,实现更好的质控。

### 1.5 芯片扫描及数据分析

芯片用GenePix 4000B Scanner双通道激光扫

描仪进行扫描,采用图像分析软件(NimbleScan v2.3 software)对芯片像进行分析,把图像转化为数字信号,计算每个基因点的表达差异Ratio, Ratio=cy5/cy3,筛选出Ratio大于2(高甲基化基因)或小于0.5(低甲基化基因)的基因点。

### 1.6 甲基化PCR

在6对标本中对筛选出来的基因进行初步验证。采用TIANamp Genomic DNA Kit提取组织标本DNA, EZDNZ甲基化试剂盒对DNA进行甲基化修饰。OPCML体系:2 $\times$  Taq PCR Master Mix 12.5  $\mu$ L; OpCml-m1及m2各1  $\mu$ L; 甲基化处理后的DNA 2.5  $\mu$ L; ddH<sub>2</sub>O 8  $\mu$ L; OpCml基因PCR反应条件:94℃ 10 min 预变性, 62℃下退火45 s, 72℃延伸10 min;以扩增产物经1%琼脂糖凝胶电泳,运用Gel Doc 1000凝胶图像成像仪(美国Bio-Rad公司)。引物序列如下,OpCml-m1: CGTTTAGTTTTT CGTGCGTTC; OpCml-m2: CGAAAACGCGCAACC GACG; 片段大小:129 bp。

## 2 结果

### 2.1 甲基化基因在染色体上的分布

一共有2 296高甲基化和2 348低甲基化基因,甲基化差异基因随机分布在每个染色体上(表1),1号染色体包含了大部分的高甲基化基因(232个基因),随后是19号染色体(149个基因),11号染色体(144个基因),2号染色体(141个基因),3号染色体(127个基因)。此外,1号染色体包含了大部分的低甲基化基因(215个基因),其次是2号染色体(155个基因),11号染色体(147个基因),6号染色体(140个基因)。表2为排名前30的高甲基化基因和低甲基化基因。

### 2.2 甲基化基因的聚类分析

聚类分析显示差异甲基化基因分布在各种分子通路中(表3、4)。根据富集指数,低甲基化基因可参与某些分子通路,如核糖核酸酶P活性、DNA修复及细胞周期等。高甲基化分布在磷酸酶活性、转录激活活性及T细胞受体信号转导通路等。

### 2.3 基因甲基化状态检测

经过初步筛选以后,我们选择排名前29位的甲基化基因进行层次聚类分析(图1)。进一步验证发现,OPCML基因在结直肠癌组织及癌旁正常黏膜存在甲基化差异(图2)。

表 1 甲基化基因在染色体上的分布

Table 1 The number of methylated genes in each chromosome

| chromosome | Total | Hypomethylation | Hypermethylation |
|------------|-------|-----------------|------------------|
| chr1       | 444   | 212             | 232              |
| chr2       | 296   | 155             | 141              |
| chr3       | 251   | 124             | 127              |
| chr4       | 195   | 99              | 96               |
| chr5       | 245   | 134             | 111              |
| chr6       | 252   | 140             | 112              |
| chr7       | 217   | 109             | 108              |
| chr8       | 170   | 80              | 90               |
| chr9       | 188   | 88              | 100              |
| chr10      | 184   | 94              | 90               |
| chr11      | 291   | 147             | 144              |
| chr12      | 222   | 110             | 112              |
| chr13      | 79    | 44              | 35               |
| chr14      | 154   | 74              | 80               |
| chr15      | 157   | 74              | 83               |
| chr16      | 196   | 107             | 89               |
| chr17      | 60    | 27              | 33               |
| chr18      | 80    | 46              | 34               |
| chr19      | 286   | 137             | 149              |
| chr20      | 148   | 65              | 83               |
| chr21      | 60    | 27              | 33               |
| chr22      | 106   | 62              | 44               |
| chrX       | 153   | 87              | 66               |
| chrY       | 19    | 10              | 9                |
| Total      | 4 644 | 2 348           | 2 296            |

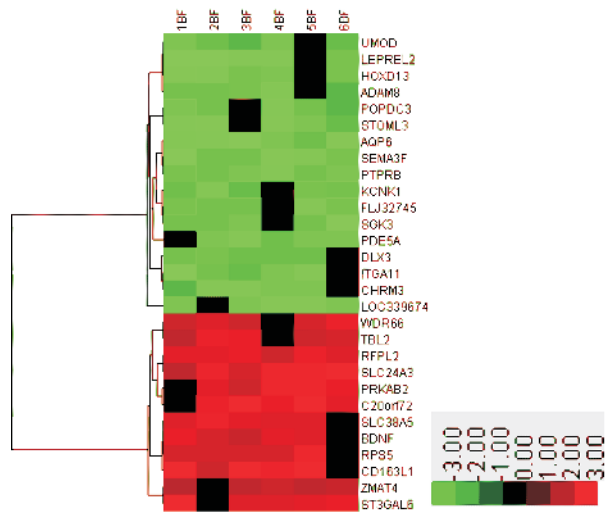


图 1 排名前 29 位的甲基化基因层次聚类图

Fig.1 Hierarchical clustering of 29 top-ranking genes

The color in each well represents relative methylation level of each gene (vertical axis) in each paired sample (horizontal axis); more intense color reflects wider differences between hyper- and hypo-methylation. Red, hypermethylation; green, hypomethylation; black, no differential; 1BF, The first pair of CRC and adjacent normal mucosa

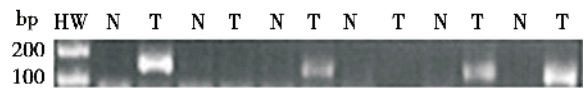


图 2 Opm1 在 6 对结直肠癌组织及癌旁组织中的甲基化分析

Fig.2 Methylation specific polymerase chain reaction (MSP) analysis results in the six paired tissues

T: tumor tissue; N: normal mucosa

### 3 讨论

结直肠癌的发生是一个多步骤、多环节的过程。其分子机制涉及癌基因、抑癌基因、转移相关基因、生长因子及其受体、粘附分子及细胞外基

质、肿瘤血管、机体免疫等多个环节。上述任何环节都可用来研究预测指标，干预任何环节均有可能预防或延缓结直肠癌的发生和发展<sup>[1]</sup>。

本研究通过建立具有组织特异性的结直肠癌

表 2 排名前 30 位的高甲基化基因和低甲基化基因

Table 2 Thirty top-ranking hypermethylated and hypomethylated genes

| Hypermethylation   | Hypomethylation  |
|--|--|
| RFPL2, SLC24A3, WDR66, ZMAT4, SLC38A5, TBL2, PRKAB2, BDNF, RPS5, ST3GAL6, CD163L1, C20orf72, KIAA0774, CNN1, ELF3, PLEC1, SLC8A1, DEFB109, HISPPD2A, NASP, CD40, TTMB, FARSLA, ZNF100, NGFR, CYP11B2, KCNAB1, RPS6KA1, SELV, OPCML | SEMA3F, AQP6, PTPRB, POPDC3, UMOD, LEPREL2, ADAM8, DLX3, ITGA11, LOC339674, HOXD13, PDE5A, KCNK1, STOML3, FLJ32745, CHRM3, SGK3, BAZ1B, CRTAC1, KIAA0652, ZCCHC9, LOC338809, BNC2, ZBTB7C, MSLN, RGS3, C20orf43, PDE5A, C5orf3, MOSPD1 |

表 3 低甲基化基因聚类分析

Table 3 The ontologic categories of hypomethylated genes

|                              | Term                                       | Count                     | %    | <i>P</i> |
|------------------------------|--|---------------------------|------|----------|
| GOTERM_BP_FAT                | cell cycle                                 | 78                        | 4.42 | 0.370    |
|                              | cell-cell signaling                        | 70                        | 3.96 | 0.059    |
|                              | DNA repair                                 | 31                        | 1.76 | 0.310    |
|                              | positive regulation of signal transduction | 21                        | 1.19 | 0.960    |
|                              | negative regulation of T cell activation   | 4                         | 0.23 | 0.780    |
|                              | cell cycle checkpoint                      | 8                         | 0.45 | 0.780    |
|                              | regulation of neurological system process  | 26                        | 1.47 | 0.005    |
|                              | water transport                            | 7                         | 0.40 | 0.003    |
|                              | transmission of nerve impulse              | 47                        | 2.66 | 0.016    |
|                              | ammonium transport                         | 4                         | 0.23 | 0.014    |
|                              | response to organic substance              | 89                        | 5.04 | 0.009    |
|                              | GOTERM_MF_FAT                              | identical protein binding | 84   | 4.76     |
| ribonuclease P activity      |  | 6                         | 0.34 | 0.001    |
| adenylate cyclase activity   |  | 5                         | 0.28 | 0.022    |
| receptor antagonist activity |  | 4                         | 0.23 | 0.081    |
| KEGG_PATHWAY                 | Dilated cardiomyopathy                     | 18                        | 1.02 | 0.016    |
|                              | ErbB signaling pathway                     | 13                        | 0.74 | 0.210    |
|                              | mTOR signaling pathway                     | 9                         | 0.51 | 0.180    |
|                              | alpha-Linolenic acid metabolism            | 4                         | 0.23 | 0.300    |
|                              | Retinol metabolism                         | 8                         | 0.45 | 0.350    |
|                              | Non-small cell lung cancer                 | 7                         | 0.40 | 0.520    |

表 4 高甲基化基因的聚类分析

Table 4 The ontologic categories of hypermethylated genes

|                             | Term                                       | Count | %     | <i>P</i> |
|-----------------------------|--|-------|-------|----------|
| GOTERM_BP_FAT               | cell cycle                                 | 75    | 4.25  | 0.600    |
|                             | positive regulation of signal transduction | 27    | 1.53  | 0.770    |
|                             | cell-cell signaling                        | 82    | 4.64  | 0.002    |
|                             | DNA repair                                 | 25    | 1.98  | 0.820    |
|                             | negative regulation of T cell activation   | 4     | 0.23  | 0.800    |
|                             | cell cycle checkpoint                      | 11    | 0.62  | 0.420    |
|                             | response to radiation                      | 36    | 2.04  | 0.001    |
|                             | cell morphogenesis                         | 54    | 3.06  | 0.002    |
|                             | actin filament-based process               | 37    | 2.10  | 0.008    |
|                             | protein amino acid dephosphorylation       | 21    | 1.19  | 0.038    |
| GOTERM_MF_FAT               | amino acid binding                         | 13    | 0.74  | 0.013    |
|                             | Purine NTP-dependent helicase activity     | 16    | 0.91  | 0.057    |
|                             | transcription activator activity           | 58    | 3.28  | 0.005    |
|                             | phosphatase activity                       | 33    | 1.87  | 0.073    |
| KEGG_PATHWAY                | T cell receptor signaling pathway          | 23    | 1.30  | 0.003    |
|                             | ErbB signaling pathway                     | 14    | 0.79  | 0.150    |
|                             | Cell adhesion molecules (CAMs)             | 17    | 0.96  | 0.360    |
|                             | Histidine metabolism                       | 6     | 0.34  | 0.200    |
|                             | Nucleotide excision repair                 | 8     | 0.45  | 0.200    |
|                             | Alzheimer's disease                        | 20    | 1.13  | 0.410    |
| integral to plasma membrane | 144  | 8.15  | 0.002 |          |

基因表达谱, 筛选并利用生物信息学手段对差异甲基化基因进行分析, 为结直肠癌的基因治疗提供可能靶点。初步分析发现 2 296 个高甲基化基因 ( $\log_2\text{ratio} > 1$ ) 和 2 348 个低甲基化基因 ( $\log_2\text{ratio} < -1$ ), 其中 293 个高甲基化和 313 个低甲基化基因未见文献报道。聚类分析发现差异甲基化基因涵盖了许多不同的功能群落, 表明在结直肠癌的发生及发展中, 有多种类型的基因参与调控, 而非单一因素所致。

本研究结果发现, 在甲基化差异有显著性的基因中, 许多都与肿瘤的发生、发展过程有密切关系。例如, ATM 是一种重要的细胞周期检查点激酶, 其 ATM 启动子甲基化在乳腺癌<sup>[12]</sup>和滤泡性淋巴瘤<sup>[13]</sup>有报道, 但在结直肠癌中尚未有报道。HRK 在细胞凋亡通路上扮演重要角色, 其甲基化状态在结直肠癌中和胃肠中均有报道。此外, NBN 启动子甲基化是一种很有前途的生物标志物, 其在 DNA 损伤和染色体的完整性维护起着至关重要的作用<sup>[14]</sup>。Opcml 在卵巢癌中研究较多, 认为是卵巢癌中新的候选肿瘤抑制基因。Opcml 在卵巢癌中主要以等位基因缺失和 CpG 岛甲基化的方式失活<sup>[15]</sup>。在本研究中, 我们首次发现 Opcml 在结直肠癌组织中存在甲基化, 提示此基因可能在结直肠癌的发生中也有重要的意义, 但其具体机制及临床意义仍需进一步研究。

本研究利用甲基化芯片高灵敏度、高通量的特点, 初步筛出了一系列可能与结直肠癌有关的基因。下一步我们将结合生物信息学挑选出可能有意义的基因, 并对其差异表达水平做进一步的鉴定, 以期发现结直肠癌的相关基因, 为结直肠癌的诊断和治疗提供新的思路和线索。

#### 参考文献:

- [1] Bhat AA, Wani HA, Beigh MA, et al. Epigenetic promoter methylation of hmlh1 gene in human gut malignancies: A comparative study [J]. J Invest Biochemistry, 2013, 2(2): 101-108.
- [2] Xiong HL, Liu XQ, Sun AH, et al. Aberrant DNA Methylation of P16, MGMT, hMLH1 and hMSH2 Genes in Combination with the MTHFR C677T Genetic Polymorphism in Gastric Cancer [J]. Asian Pacific J Cancer Prev, 2013, 14(5): 3139-3142.
- [3] Hong L, Ahuja N. DNA Methylation Biomarkers of Stool and Blood for Early Detection of Colon Cancer[J]. Gen Test Mole Biomarkers, 2013, 17(5): 401-406.
- [4] Chaar I, Amara S, Elamine OE, et al. Biological significance of promoter hypermethylation of p14/ARF gene: Relationships to p53 mutational status in Tunisian population with colorectal carcinoma [J]. Tumor Biol, 2013, 11(5): 1-11.
- [5] Esposito A, Bardelli A, Criscitiello C, et al. Monitoring tumor-derived cell-free DNA in patients with solid tumors: Clinical perspectives and research opportunities [J]. Cancer Treat Rev, 2013, 11(5):211-216.
- [6] Fernald K, Kurokawa M. Evading apoptosis in cancer [J]. Trends in Cell Biol, 2013, 23(12): 620-633.
- [7] Irahara N, Nosho K, Baba Y, et al. Precision of pyrosequencing assay to measure LINE-1 methylation in colon cancer, normal colonic mucosa, and peripheral blood cells[J]. J Mole Diagn, 2010, 12(2): 177-183.
- [8] Kalimutho M, Di Cecilia S, Blanco GDV, et al. Epigenetically silenced miR-34b/c as a novel faecal-based screening marker for colorectal cancer[J]. BJC, 2011, 104(11): 1770-1778.
- [9] Özata D M, Xie H, Lindfors U, et al. miR-185 and miR-133b deregulation is associated with overall survival and metastasis in colorectal cancer [J]. Inter J Oncol, 2011, 39(9): 311-318.
- [10] Rossi L, Bonmassar E, Faraoni I. Modification of miR gene expression pattern in human colon cancer cells following exposure to 5-fluorouracil in vitro [J]. Pharmac Res, 2007, 56(3): 248-253.
- [11] Fidler I J. Cancer biology: invasion and metastasis[J]. Clinical Oncology, 2000, 2(9): 29-53.
- [12] Vo QN, Kim WJ, Cvitanovic L, et al. The ATM gene is a target for epigenetic silencing in locally advanced breast cancer [J].Oncogene, 2004, 23 (58):9432-9437.
- [13] Chim CS, Wong KY, Loong F, et al. Absence of ATM hypermethylation in mantle cell and follicular lymphoma [J]. Leukemia, 2005, 19(5): 880-882.
- [14] Pardini B, Naccarati A, Polakova V, et al. NBN 657del5 heterozygous mutations and colorectal cancer risk in the Czech Republic[J]. Mutat Res, 2009, 666 (1-2): 64-67.
- [15] McKie AB, Vaughan S, Zanini E, et al. The OPCML tumor suppressor functions as a cell surface repressor-adaptor, negatively regulating receptor tyrosine kinases in epithelial ovarian cancer[J]. Cancer Discov, 2012, 2 (2): 156-171.