**Global Analysis of DNA Methylation by Methyl-CaptureSequencing Reveals Epigenetic change of pancreatic cancer**

Abstract：

**Purpose**:但胰腺癌系统性的全基因层面甲基化研究较少，我们的目标是建立胰腺癌全基因组包括CGI，CGI shore,Orphan CGI,NO CGI promotor等的甲基化图谱，识别胰腺癌中异常甲基化位点。**Method**:我们运用Methylcap-seq方法分别混合10例胰腺癌组织和10例胰腺癌非肿瘤组织建立胰腺癌和胰腺非肿瘤组织DNA甲基化文库，生物信息分析得到胰腺癌中异常甲基化谱式，另外在一组单独胰腺癌临床样品中，运用MSP,BSP，MSRE-qPCR方法对部分选定DMR区域进行甲基化分析和相关基因mRNA，EST表达进行分析。

66807 Hypermethylation DMR and 46815 Hypomethylation DMR were discovered in whole pancreatic cancer genome. 5280 Hypermethylation and 3488 Hypomethylation CpGI region were related gene promoter or gene body. Gene expression of 10 aberrant DMR were related with their methylation status validated by MSRE-qPCR and RT-PCR in 5-aza’ treated pancreatic cancer cell line. 40 DMR were validated in pancreatic cancer and adjacent tissues with MSP and gene ontology analysis showed seven of them were related with regulation of transcription(GO:0006355) which included DLX4,ELAVL2,IRX1,PITX2,SIM2,TBX5,TFAP2C.

**Conclusions**:我们的Methylcap-seq结果显示，在胰腺癌中DMR存在于多种基因区域，可能影响多个已知基因或未知转录本表达，从而引起胰腺细胞多层次的生物学功能异常导致胰腺癌的发生，而这些DMR所涉及的基因或转录本可作为胰腺癌治疗的靶标或诊断的生物学标志物。

Background,

Pancreatic cancer is a highly malignant tumor of the digestive system, the survival rate is very low, and extensive reprogramming of DNA methylation and loss of control is an important feature of cancer. Till now, whole genome wide DNA methylation profile of pancreatic cancer is not come out yet. In present study, we try to establish profiles of DNA methylation, especially in CGI, CGI shore, orphan CGI and non-CGI promoter region so that we could identify aberrant DNA methylation region in pancreatic cancer.

Methods: MBD methylCap/seq was carried out to screen differentially methylated CpG islands in 3 libraries which includes 1 pancreatic cancer mix library, 1 adjacent normal tissue mix library and 1 normal pancreas tissue. MSP and multiplex-BSP validation was performed in independent tumor and adjacent tissues.

**Introduction**

胰腺癌是一种恶性程度极高的消化系统肿瘤，术后5年生存率小于25%,是目前预后最差的实体肿瘤之一。在美国和欧洲每年导致近10万人死于这种癌症，分别占其恶性肿瘤死亡率的第四和第五位[1]，在中国地区胰腺癌的发病率也呈逐年上升的趋势[2]。

生物角度认为遗传主要决定个体间遗传信息稳定传递，在同一个体中不同组织细胞表型的塑造是表观遗传作用的具体展现。先前研究认为一些遗传事件改变，如癌基因的激活和抑癌基因（TSG）的突变和缺失，与肿瘤有密切的关系；由于染色体的改变而影响EGFR[3]、p16INK4A, TP53, MYC, K-RAS2, and AKT2[4]，基因点突变影响K-RAS[5]、BRAF[6]，基因突变影响TP53[7]，都被证明在胰腺癌的发生发展有密切关系。肿瘤作为个体后天出现的一种生物学特殊表型，表观遗传改变同样起着重要作用[8]。DNA甲基化作为表观遗传现象之一，在胰腺癌中已有多项研究，证明在基因层面如p14ARF ，p16INK4a，RASSF1A [9、10] 等由于基因启动子区域发生甲基化失控，导致mRNA的转录异常；在胰腺癌基因组学研究层面，Michael Goggins等运用methylated CpG island amplification结合Agilent’s 244K Human promoter chip-on-chip microarray技术，描述了胰腺癌在全基因范围内的甲基化异常 [11-13]

Michael Goggins等的研究使用甲基化芯片技术平台，其专注点更偏重于基因组基因区域 CGI (CpG island)，其基因组覆盖度低于二代测序技术[14]，对胰腺癌全基因组甲基化谱式描述不够全面，对诸如CpG shore、基因无CGI启动子区、Orphan CGI等在胰腺癌中的甲基化状况改变有所忽略；而现有研究证明一些与肿瘤表型相关或组织特征相关的甲基化改变常发生于上述区域[15,16]。 MethylCap-seq是建立全基因组甲基化谱式新技术，运用甲基化结合蛋白(MBD)特异性富集含甲基化CpG的DNA片段，通过不同盐浓度的洗脱缓冲液,收集不同甲基化密度的DNA片段，对上述富集片段进行2代测序，从而建立全基因组甲基化谱式 [17]。

在本研究中，我们将10例胰腺癌组织和10例胰腺非肿瘤组织分别混合，形成胰腺癌组和胰腺非肿瘤组的研究模型，用本实验室稳定的MethylCap-seq技术[18,19]，建立胰腺癌全基因组甲基化谱式，探寻中国胰腺癌人群中包括与基因相关的CGI，和基因无关的CGI (Orphan CGI), CpG shore、基因无CGI启动子区等高频率甲基化失控的区域，寻找可能受上述甲基化异常区域调控的、参与胰腺癌发生发展的相关基因(fig1)。

**Materials and Methods**

Clinical samples and Cell line

所有临床胰腺癌组织标本来自于上海市交通大学医学院附属仁济医院，共采集18名选择手术治疗的胰腺癌病人，采集日期从2009.5-2011.3，所有患者手术前均未接受化疗或放疗。分别采集胰腺癌患者手术切除肿瘤组织和相距2cm的非肿瘤组织，样本经标签区分后液氮速冻后零下80℃保存。肿瘤组织样本TNM临床分期根据国际抗癌联合会(UICC)的HCC TMN分期标准第6版（见表一）。临床样品收集符合医学伦理规范，所有病人签署临床样品科研知情同意书。DNA preparation from frozen tissues or cell line were made using the conventional proteinase K/organic extraction method as previously described[18]..

**Generation of wide-genome methylation profiling by Methylcap-seq**

取10例胰腺癌肿瘤组织和与其匹配的10例癌旁组织基因组DNA，各取等量基因组DNA分别混合成胰腺癌组(PC)和非肿瘤组(PN).最终PC组和PN组各取基因组DNA 1.2ug。generation of Methylcap-seqwere performed as previously described[18].

**Mapping of Sequence Reads**

We used BWA alignment tools[20].with the default settings to map these 36 bp unpaired reads to the hg19 human genome reference assembly[21]. (UCSC) after removing PCR duplicates with Picard. Samtools[22].and Picard (<http://picard.sourceforge.net>) were used to convert, sort, and index the aligned data.

**Identification and Annotation of DMRs**

Methylation peaks (hypermethylated regions) were identified using MACS [23].and BALM[24].at the same time to increase the detective power of Methylcap-seq while differential methylation region (DMR) was conducted by dual-threshold strategy, in which high-positive threshold (threshold=0.975) and low negative threshold (threshold=0.97). whole genome methylation (methylation of each CpG) were inferred with BALM, which was then used to make pearson correlation analysis among all the samples under R enivironment.The refseq genes (UCSC genes) and corresponding CpG islands (CGIs) were downloaded from Table Browser of the UCSC database[18]. Bed files operation are treated with Bedtools[25]. and some other Perl scripts. All the scripts are available if you sent the email to the authors. The genomic methylation profile generated was uploaded to a public database (Gene Expression Omnibus: ).

Gene ontology categorizations were performed using DAVID Bioinformatics Resources 6.7 (http://david.abcc.ncifcrf.gov/).

**Methylation analysis**

我们使用EpiTect Kit (Qiagen)对1ug胰腺癌组织，胰腺非肿瘤组织和细胞系基因组DNA，按KIT提供的操作程序进行bisulfite-treat。

MSP和BSP引物设计使用在线引物设计软件MethylPrimer (http://www. urogene .org/cgi-bin/methprimer/methprimer.cg)。MSRE-qPCR引物设计使用在线设计软件primer3(http://www.embnet.sk/cgi-bin/primer3\_www.cgi),以上所有本文中使用到的引物序列等信息见(Supplementary tab1).

Jumpstart Taq(Sigma)被应用于BSP (Bisulfite Sequencing PCR)分析，反应体系20ul,反应条件：94℃3 min；94℃20 s，退火20 s，72℃ 20 s，35个循环；72℃5 min。MSP结束后将产物用1.5%的琼脂糖凝胶电泳分析，并将条带割胶回收后，作T-A克隆并送出至少5个克隆进行测序。

Jumpstart Taq(Sigma)被应用于MSP分析,反应体系20ul, MSP反应条件：94℃ 3 min；94℃ 20s，退火20s，72℃ 20 s，35个循环；72℃5 min。MSP结束后将产物用1.5%的琼脂糖凝胶电泳分析，并将条带割胶回收后，作T-A克隆并送测序。

在胰腺癌细胞系经5-AZA处理前后和小样临床肿瘤组织中 ，对Orphan CGI等进行DNA甲基化水平定量分析，我们使用MSRE-qPCR(methylation-sensitive restriction enzyme-based quantitative PCR)方法，具体见前次文章描述[26]。

**Cell culture and** **5-AZA-2-deoxycytidine treatment**

The cell lines used were three Pancreatic Adenocarcinoma cell lines, BxPC-3 (ATCC, CRL -1687),PANC-1 (ATCC,CRL-1469),CFPAC-1 (ATCC,CRL-1918), All the cell lines were cultured in RPMI1640, The media were supplemented with 10% FBS, 100 U/ml penicillin, and 100 U/ml streptomycin. All cell lines were maintained at 37°C in a humidified atmosphere with 5% CO2.

To study whether demethylation could restore expression in BxPC-3,CFPAC-1 and CFPAC-1 cell lines . For CpG demethylation analysis, Exponentially growing cells were seeded at a density of 1.5 times 106 cells/100 mm-diameter dish and allowed to attach overnight. The cells were then treated with freshly prepared 5 mum 5-aza-dC (Sigma) for 3 days.

**RNA isolation and Real-time PCR**

Total RNA was prepared from cultured cells using Trizol reagent according to the manufacture’s instruction (Invitrogen) and then reverse transcribed using oligo(dT) as primer and SuperscriptIITM RNase H-Reverse Transcriptase (Invitrogen). Real-time PCR was carried out with primer pairs for the GAPDH internal control and the primer of EST expression assay. The real-time PCR was carried out as follows: 94℃ for 3min,followed by 40cycles of 94℃ for 10sec,62℃ for 10sec and 72℃ for 15sec. Real-time qPCR was carried out to detectlevels of the corresponding GAPDH,using a SYBR Green PCR Kit(Applied Biosystems) and an ROTOR-GENE 6000 Real-Time PCR System (ROTOR-GENE).

**Statistical analysis**

Statistical calculations were carried out with the SPSS statistical software package (Version 13.0;SPSS,Inc.) 计数资料采用Fisher精确检验进行分析,P<0.05为差异有统计学意义。

**Result**

**胰腺癌组织甲基化组学分析揭示肿瘤中广泛的高甲基化和低甲基化改变**

我们运用Methylcap-seq方法, 分别建立了10例胰腺癌组织和匹配的胰腺非肿瘤组织全基因组甲基化谱式。我们在胰腺癌中得到33,784,358 raw reads，在胰腺非肿瘤组织中得到30,868,151 raw reads。其中胰腺癌中17,267,025 (51.12%) raw reads和胰腺非肿瘤组织16,033,135 (51.94%)raw reads经与人基因组(H19)比对，可被唯一定位。

将上述reads与UCSC定义的28691个CGI进行mapping，结果胰腺癌和胰腺非肿瘤组织分别有3.57%和4.25%的的reads被定位于CGI,，在胰腺癌和胰腺非肿瘤组织中大约有64.31%和64.36%的CGI被reads所覆盖。

胰腺癌和胰腺非肿瘤组织中的差异甲基化区域(DMRs)的获取。我们运用MACS和BALM这2种方法定义甲基化峰(peak)。经比对胰腺癌和胰腺非肿瘤组织所获甲基化峰，我们得到了在胰腺癌中DMRs,包括Hypermethylation的 DMRs 66807个和hypomethylation 的DMRs 46815个,显示在胰腺癌中存在广泛的hypomethylation和局部的hypermethylation (fig2a).进一步观察DMR与基因的关系，我们将Hypermethylation DMR和Hypomethylation DMR分别匹配到相应的TSS， Intragenic和 intergenic(fig2b)。再进一步将其中涉及TSS和Intragenic的DMR在基因与CGI进行关联，我们得到Hypermethylation的DMRs 5280个和hypomethylation DMRs 3488个。这些DMR定位到人基因组的基因结构注释，包括downstream，enhancer，exon, intron, mir, promoter, UTR5(fig 2d). 同时我们发现TSS上下游5K区域中越接近TSS，胰腺癌组织中Hypermethylation的 Peak越多，即甲基化程度越高(fig2c)。

我们的结果与别人的胰腺癌肿瘤甲基化研究比较（不论基因层面还是全基因组层面），下述基因均在本研究胰腺癌组织中找到hypermethylation DMR, 如LHX1, FOXE1, PAX6, BNIP3[27], ALPP, CEBPA[28], CACNA1G[29], CCND2[30]，BAI1,NRN1, PENK, FAM84A,and, ZNF415 [31]。胰腺癌之外的其他肿瘤常常发生甲基化的基因，包括RassF1a, CDKN2A, hHML1,CDH1[32,33]，也出现在本研究的甲基化谱式的DMR之中。

一般认为基因启动子区域CGI甲基化会抑制基因表达，但已有研究表明一些与肿瘤和组织相关的特异性甲基化常发生在基因启动子区域CGI上下游2K的CpG shore中。在本研究中我们观察了该区域中CGI和CpG shore的甲基化状态关系，发现在胰腺癌中有527个基因Promotor区域的CGI和CGI shore（对应2970个DMR）同时发生hypermethylation；有111个基因Promotor区域的CGI（对应189个DMR）发生hypermethylation, 而其上下游CGI shore没有甲基化差异。有1278个基因promotor区域的CGI没有甲基化差异，而在CGI shore含有2320个DMR发生了hypermethylation. 同时我们还观察到，在胰腺癌中有334个基因Promotor区域的CGI和CGI shore（对应1602个DMR），同时发生hypomethylation。有51个基因Promotor区域的CGI（对应88个DMR）发生hypomethylation,而其上下游CGI shore没有甲基化差异。有1256个基因promotor区域的CGI在胰腺癌组织和胰腺非肿瘤组织没有甲基化差异，而在其CGI shore（对应2200个DMR）发生了hypomethylation(fig2e).

进一步我们观察到，在胰腺癌中有多个不含有CGI的基因promotor区域,在其转录起始点上下游5k区域内含有DMR.。其中有615个基因（对应1078个DMR）在胰腺癌中发生了hypermethylation,有384个基因（对应638个DMR）在胰腺癌中发生了hypomethylation.另外我们观察到位于Intergenic 的CGI,我们称为Oaphern CGI,在胰腺癌组织中同样发生了甲基化状态的改变，我们发现在胰腺癌组织中有129个Oaphern CGI（对应133个DMR）发生了Hypermethylation，有95个Oaphern CGI（对应96个DMR）发生了hypomethylation(fig2f)。

**胰腺癌中甲基化异常涉及的GO分析以及信号通路的改变**

通常认为，发生于基因启动子区域的甲基化失控会影响基因表达，所以我们对来自于胰腺癌组织中启动子区域CGI发生异常hypermethylation和hypomethylation的基因进行了GO分析。在biological processes分析中，我们发现hypermethylation genes主要富集在regulation of transcription, DNA- dependent (GO:0006355),regulation of RNA metabolic process(GO:0051252), regulation of transcription(GO:0006355),embryonic morphogenesis(GO:0048598);而hypomethylation genes主要富集在small GTPase mediated signal transduction(GO:0007264),epidermis development(GO:0008544),ectoderm development(GO:0007398), cell aging(GO:0007569).在molecular processes分析中，我们发现hypermethylation genes主要富集在transcription factor activity (GO:0003700), sequence-specific DNA binding(GO:0043565), transcription regulator activity (GO:0030528), DNA binding(GO:0003677)。分别对上述基因进一步进行KEGG Pathway分析，我们发现hypermethylation genes主要富集在以下的KEGG pathway中，Calcium signaling pathway(hsa04020)，Neuroactive ligand-receptor interaction(hsa04080)，Maturity onset diabetes of the young(hsa04950)。而hypomethylation genes主要富集在以下的KEGG pathway中, Melanoma(hsa05218), Bladder cancer(hsa05219)，Pathways in cancer(hsa05200)， Regulation of actin cytoskeleton(hsa04810) (tab2)。

那些启动子区域没有CGI却含有DMR的基因也被进行GO分析，我们发现那些发生Hypermethylation的基因被主要富集在development(GO:0007275), morphogenesis(GO:0009653), transcription, DNA-dependent(GO:0006351), organogenesis(GO:0009887), organ development(GO:0048513)( Supplementary tab3)。Hypomethylation的基因被主要富集在development(GO:0007275), neurogenesis(GO:0007399), morphogenesis(GO:0009653), organogenesis(GO:0009887) organ development(GO:0048513)( Supplementary tab4).分别对上述基因进一步进行KEGG Pathway分析，我们发现那些发生Hypermethylation的基因被主要富集在Neuroactive ligand-receptor interaction(hsa04080),D-Glutamine and D-glutamate metabolism(hsa00471)，而Hypomethylation的基因没有在任何pathway中被富集。

除此之外，还鉴定到一些miRNA在胰腺癌中与发生异常甲基化，其中hypermethylation的有mir-9-3,mir-9-1,mir-124-3,mir-10b,mir-124-2,mir-718,mir-203. hypomethylation的有mir-210,mir-1469,mir-130b,mir-149,mir-1224,mir-564(tab3).

**验证来自于胰腺癌的DMR**

BSP方法被用来验证来自于甲基化组学DMR，对原始建立甲基化组学的胰腺癌组织混合组和胰腺非肿瘤组织组，经亚硫酸盐处理后，使用BSP方法共验证了15个基因的DMR位点(Tab4)，其中11个为hypermethylation，4个为hypomethylation，结果在hypermethylation的11个位点中有7个DMRs与组学信息符合，4个hypomethylation DMRs中有2个与组学信息符合(部分结果见fig3a)，符合率大约60%（9/15）。

我们用甲基转移酶抑制剂5-aza-dc，处理3株胰腺癌细胞株，分别抽提DNA和RNA。用MSRE-qPCR定量分析了包括基因启动子区域CGI和orphan CGI 在内的10个来自于胰腺癌中Hypermethylation DMR (部分信息见fig4a),在3个胰腺癌细胞系中5-aza-dc处理前后的DNA甲基化水平，同时为了确定在胰腺癌中异常甲基化的DMRs和orphan CGI与关联mRNA表达的相关性，我们运用RT-PCR检测了promotor CGI相关基因mRNA以及orphan CGI关联EST表达水平(fig4b)。结果显示，结果有4个orphan CGI和1个promotor CGI经5-aza-dc处理，在3株胰腺癌细胞株中均不同程度的发生了DNA甲基化水平下降，mRNA表达水平上升，提示这些位点的基因（或EST）表达可能受DNA甲基化影响。进一步在methylome group和testing group的胰腺癌组织和胰腺非肿瘤组织中用MSRE-qPCR对上述DMR的进行甲基化状况定量分析，上述4个位点在临床样本中存在甲基化水平差异(fig4c)。

**来自于胰腺癌全基因组甲基化谱式中DMR进行小样本临床胰腺癌组织初步检测**

MSP方法被用来对来自于胰腺癌全基因组甲基化谱式中DMR进行小样本临床胰腺癌组织初步检测，共有8例胰腺癌组织、5例胰腺非肿瘤组织和3例胰腺癌细胞株，进行了40个基因相关DMRs检测。上述DMR来自于基因promotor、mir,intron,exon和CGI shore区域，结果显示有8个基因相关的DMRs在8例胰腺癌组织中至少有1例发生甲基化，而在5例胰腺非肿瘤组织中未出现甲基化；另外有9个基因相关的DMRs在8例胰腺癌组织中至少有2例以上发生甲基化，而在5例胰腺非肿瘤组织中仅有1例出现甲基化。另外来自于TRADD,AGAP2,FAM115A这3个基因启动子区域的DMR在胰腺癌中发生甲基化丢失，其余20个DMRs的在胰腺癌和癌旁之间甲基化没有差异(部分结果fig3b, MSP结果见Supplementary tab2)。

**Discution**

在以前的研究中，混合样品组学被多项研究所使用，被认为是寻找被研究模型表型共性、节约样本、研究微量模型的良好方法[34,35]。在本研究中，我们使用10例胰腺癌组织和10例胰腺癌旁组织，分别混合成胰腺肿瘤组织和非肿瘤组织建立胰腺癌和非肿瘤组织的全基因甲基化谱式，探寻胰腺癌发生后由于表观遗传层面的DNA甲基化发生的改变可能导致的致癌作用。

先前已经有多项研究在多种肿瘤中观察到肿瘤中基因组水平的甲基化变化，如Kim YJ等经甲基化芯片建立前列腺癌细胞全基因组甲基化谱式，经在临床肿瘤样本中筛查，得到EFEMP1受甲基化调控表达，在前列腺癌中有异常甲基化[36]。Walter K等在NSCLC研究中，利用表达谱和全基因组甲基化谱式，在肿瘤细胞和临床肿瘤组织样品中，发现可以利用基因的甲基化表型来确定NSCLC的病理表型[37]。A Formosa1等利用甲基化转移酶抑制构建胰腺癌肿瘤细胞系差异表达谱式，筛选受甲基化调控的mir，研究其在胰腺癌肿瘤发生发展中的功能[38]。

在本研究中我们得到了在胰腺癌组织中与基因和CGI同时关系密切的Hypermethylation DMRs 5280个和hypomethylation DMRs 3488个。我们在对其中位于promotor区域异常甲基化的DMR进行GO分析后发现，异常Hypermethylation的基因主要富集在核酸代谢相关的类别中,这些结果暗示在胰腺癌的发生发展过程中一些与基因转录相关的基因如转录因子等由于本身甲基化现象的发生导致RNA表达降低从而进一步影响其他基因的转录，Helman E 等在NSCLC甲基化研究中证明了这一点[39],Zhao M 研究中也发现有上述被富集基因在多个肿瘤中RNA表达异常，被归类于Tumor suppressor genes (TSGs) [40]。而embryonic morphogenesis (GO:0048598)相关基因的被富集，说明其参与胰腺癌的形态形成过程和后期分化，而hypomethylation genes主要富集在以下，small GTPase mediated signal transduction(GO:0007264), small GTPase家族成员众多，涉及细胞增殖、分化、分裂，迁移，该ontology terms已在多项肿瘤研究中被富集[41]。那些被epidermis development(GO:0008544), ectoderm development(GO:0007398)富集的基因显示其在胰腺癌中的异常去甲基化而导致可能基因表达的上调，导致异常分化，基因组不稳定，细胞不受控制的增殖，该类被富集相关基因表达增加成为肿瘤发生发展过程中已知的一致特征[42]。

我们同时还对来自于胰腺癌组织中DMR(与转录密切相关的promoter)相关基因进行了KEGG pathway 分析，我们发现hypermethylation genes主要富集在Calcium signaling pathway(hsa04020)，Neuroactive ligand-receptor interaction (hsa04080).我们知道Calcium signaling pathway作为重要的信号通路参与调控许多细胞和组织的生理活动，包括肌肉收缩、新陈代谢、分泌以及细胞分裂，已有研究表明也与胰腺癌肿瘤细胞的增殖凋亡密切相关[43]，同时被Neuroactive ligand-receptor interaction所富集的基因主要参与细胞内分泌和外分泌的作用，已在包括meningioma 和胰腺癌的研究中被证明该类基因功能异常[44,45].而在本胰腺癌研究中，hypomethylation genes主要富集在常见的肿瘤异常Pathways 中，Melanoma(hsa05218), Bladder cancer(hsa05219)，Pathways in cancer (hsa05200), 而Regulation of actin cytoskeleton (hsa04810) 相关基因被富集，也印证了先前的研究，该类基因与胰腺癌细胞的侵袭转移密切相关[46,47] (tab2)。

我们对那些启动子区域无CGI却含有DMR的基因所得到的GO分析结果显示，无论是Hypermethylation还是Hypomethylation的基因被一致的主要富集在与器官发育，胚胎形成，转录等生物学功能上。这与Han H等的研究一致[48]，这些区域的甲基化状态被认为是与组织类型和肿瘤发生密切相关，本研究中在999个基因启动子的1716个区域发生甲基化紊乱，可能引起该基因群表达异常从而导致胰腺细胞分化失控，是胰腺癌发生的主要原因之一。

miRNA作为发挥细胞生理功能的重要表观遗传调节因素，其受DNA甲基化调控已有多项肿瘤研究，被证明与肿瘤凋亡、侵袭、转移、复发、耐药等相关。在本研究中我们发现一些miRNA在胰腺癌中发生了异常甲基化，其中hsa-mir-124-3已经被Wang P和Gebauer K证明在胰腺癌中发生hypermethylation，并参与胰腺癌进展、转移、复发 [49,50]. 在本研究中被发现发生异常hypomethylation的mir-130b和mir-210也被Zhao G和Takikawa T 发现在胰腺中过表达且与胰腺癌的增殖侵袭相关[51,52]。还有一些在本研究中被发现异常甲基化的miRNA,在胰腺癌中没有被研究，但已在其他肿瘤中被研究，且受DNA甲基化调控。如本研究中发现miR-9-3,mir-9-1,miR-124,miR-203在胰腺癌组织中hypermethylation, 这些mir在Heller G等的研究中被证明，分别在non-small cell lung cancers,breast cancer,cervical cancer,haematological中发生hypermethylation,导致mir表达下降，从而促进肿瘤的发展，肿瘤细胞的增殖等[53-56]。在本研究中发现的在胰腺癌中发生hypermethylation的mir-10b,在Ma L 等在动物肿瘤模型研究中被认为能抑制肿瘤细胞的转移[57]，而Nakata K 等在胰腺癌的研究中发现其过表达，且可以作为一个临床判定指标[58]。当然还有部分我们发现的miRNA，迄今没有任何研究，这些miRNA在今后的研究中会让我们更全面的了解胰腺癌发生发展的可能机制。

另外我们选取了本研究中在启动子区域甲基化p值最高的40个基因，在另外的7例胰腺癌组织，5例胰腺非肿瘤组织和3例胰腺癌细胞系中用MSP方法进行目的区域甲基化检测，结果有18个基因显示在胰腺非肿瘤组织和胰腺癌组织、胰腺癌细胞系之间有明显的甲基化差异。并对其中甲基化异常基因在5-AZA-2dR处理前后的胰腺癌细胞株中检测基因mRNA表达水平，结果显示基因启动子的甲基化与基因RNA表达水平有相关性。有趣的是，这18个基因中有8个基因：DLX4,ELAVL2,IRX1,PITX2,SIM2,TBX5,TFAP2C,VSX1被GO分析结果：regulation of transcription, DNA-dependen (GO:0006355 )所富集。

上述基因中DLX4属于同源异型盒转录因子DLX家族，DLX4主要在正常妊娠时的绒毛细胞滋养细胞和绒毛外滋养细胞中表达，调节胎盘发育[59]。该基因被研究表明在一些肿瘤中高表达引起肿瘤侵袭转移等[60]，但也有研究观察到该基因在cervical cancer,lung cancer,chronic lymphocytic leukemia,breast cancers中异常甲基化[61-64]，所以DLX4在胰腺癌中的甲基化是否仅是伴随现象还是参与胰腺癌的发生发展有待进一步研究。IRX1属于 Iroquois homeobox protein family，定位于5号染色体短臂5p15.33区。homeobox protein family作为胚胎发育的主控基因在哺乳动物胚胎发育过程中发挥着重要作用，并参与了肿瘤的发生发展[65]。已有研究发现该基因在rheumatoid arthritis, alveolar rhabdomyosarcomas, gastric mucosa, head and neck squamous cell carcinoma中有异常hypermethylation发生[66-69]，但不能确定，该基因是否与Bennett KL的研究一样，IRX1与TGF-β and hypoxia pathways相互作用参与胰腺癌的发生发展[70].SIM2是Down's syndrome的相关基因，其short form (SIM2-s)被发现在多种肿瘤中高表达，被干扰后引起肿瘤细胞增殖抑制和凋亡[71]，而MENG Xian-Fang等研究发现，SIM2基因可能通过抑制神经元细胞中cyclin E-cDK2复合物形成使细胞阻滞于G1期，从而抑制细胞增殖[72]，同时SIM2的甲基化状态也可以作为膀胱癌的判定指标[73]。PITX2是Bicoid homeobox family成员，发挥个体终端发育过程中决定细胞分化方向等遗传调控作用，研究表明PITX2通过调控其下游Cyclin-D1 and C-myc促进肿瘤细胞生长和迁移入侵，参与卵巢癌进展[74]。但同时该基因在breast cancer，non-small cell lung cancer，prostate cancer中发现异常hypermethylation和mRNA表达降低，所以目前我们还不清楚该基因在胰腺癌发生发展中的作用[75-77]。本研究发现hypermethylation的TBX5是一个参与发育过程的保守基因，Yu J等发现其在结肠癌中因hypermethylation而表达下降，在结肠癌细胞株中重新表达TBX5发现其通过BCL2-associated X protein and Granzyme A signaling促进肿瘤细胞凋亡，同时通过上调CDKN2A,MTSS1和下调SNCG,MTA2，抑制肿瘤细胞增殖和转移[78]。TFAP2C是一个结合于特殊DNA序列的转录因子，在脊椎动物胚胎发育过程中发挥着重要的作用，研究表明其在乳腺癌、卵巢癌、前列腺癌等肿瘤组织中高表达，其作为ERα的协同转录因子调控下游的基因，参与肿瘤细胞的增殖，抗凋亡等[79,80]，有趣的是在黑色素瘤中高表达的mir-214通过抑制其靶基因TFAP2C促进黑色素瘤的发展和远处转移，TFAP2C被认为发挥抑癌作用[81]，另外TFAP2C在胃癌和黑色素瘤中存在hypermethylation现象[82,83]，胰腺癌作为一个与性激素通路相关性较小的器官，TFAP2C对胰腺癌发生发展是促进还是抑制有待进一步研究。ELAVL2属于高度保守的神经特定RNA结合蛋白,其通过结合到某些RNA的 3’非翻译区特定序列增加RNA的稳定性来调控RNA的表达[84]，ELAVL2在肿瘤中被研究的较少，Cairns P等发现在SLSC中TFAP2C存在杂合性缺失[85]。以上结果显示在胰腺癌中一批与基因转录相关的转录因子由于异常甲基化导致mRNA表达较低，影响相关基因的转录或转录后调控，从而导致胰腺癌的发生发展。这些基因迄今未见在胰腺癌中进行相关功能和DNA甲基化研究，具体有待进一步研究。

Adrian Bird将CGI分为TSS, intragenic,intergenic,并把后2种CGI定义为orphan CGI[86]，虽然迄今对orphan CGI的功能了解并不清楚，但已有的研究显示这些CGI在基因转录调控，印记基因调控，非编码RNA转录调控，或显示为组织特异性甲基化谱式[87]，本研究中我们也研究了部分orphan CGI在胰腺癌中甲基化状态，发现在肿瘤组织中有hypermethylation，且在肿瘤细胞系中orphan CGI甲基化状态与其附近未注释EST转录水平密切相关，这些受Orphan CGI甲基化调控的EST是否为未知基因或其他元件有待进一步研究。同时我们知道大约仅有6.8%的CpG位于CGI,其余CpG甲基化状态以及生物学功能，迄今研究很少，已有研究表明CGI上下游2K区域的CpG shore中CpG甲基化参与了基因转录调控，构建组织特异性的甲基化谱式，被认为在肿瘤中比promotor和CGI更早发生甲基化改变[19]。以上这些区域在胰腺癌中的甲基化改变情况，在本研究中都被进行了观察，其甲基化状态改变既是胰腺癌基因组甲基化谱式重要组成部分，也提示了这种表观遗传学变化在胰腺癌中可能影响基因或非编码RNA、组织细胞特异分化等从而导致癌症。

综上所述，本研究利用Methycap-seq方法建立了胰腺癌组织和胰腺癌旁组织的全基因组甲基化谱式，展现了在胰腺癌中所发生的全局性DNA甲基化重编程和失控。胰腺癌相比于胰腺非肿瘤组织，在与基因转录密切相关TSS区域CGI、Orphan CGI、CGI shore、无CGI启动子区域均发生了异常的hyper或hypo methylation。以上的区域既为我们提供了深入了解由DNA甲基化调控的基因或非编码RNA参与胰腺癌发生发展的机理，又可能作为胰腺癌早期发现的Biomarker侯选库.

**Captions and legends for tables and figures**

Table 1. Clinicopathologic characteristics of patients with pancreatic cancer

Table 2 Gene ontology enrichment analysis of aberrant methylation in gene promoter

Categorization of significantly(p<0.05)

Table 3.Aberrant methylated miRNA in Pancreatic cancer

\*rpm,reads per million

Table 4. Aberrant methylated DMR in Pancreatic Adenocarcinoma

\*rpm,reads per million

Figure 1. An overview of the experimental approach used to evaluate differential DNA methylation in pancreatic cancer

Figure 2. General characteristics of pancreatic cancer DNA methylation patterns determined by MethylCap-seq..

A, Distribution of DRM in pancreatic cancer genome. B, Distribution of DRM in TSS,Intragenic and Intergenic. C, Distribution of Peaks in TSS region(-5K～+5K).D