



doi:10.3978/j.issn.1005-6947.2017.03.005
http://dx.doi.org/10.3978/j.issn.1005-6947.2017.03.005
Chinese Journal of General Surgery, 2017, 26(3):294-303.

· 基础研究 ·

Affymetrix 基因表达谱芯片技术筛选胰腺癌异常表达基因的研究

易超¹, 依马木买买提江·阿布拉¹, 丁伟¹, 苏雅婷², 晏冬¹, 李海军³

(1. 新疆医科大学第三附属医院 肝胆外科, 新疆 乌鲁木齐 830000; 2. 新疆医科大学第五附属医院 医务部, 新疆 乌鲁木齐 830000; 3. 广东省深圳市罗湖区人民医院 普通外科, 广东 深圳 518001)

摘要

目的: 利用基因表达谱芯片筛选出胰腺癌组织与癌旁正常组织的差异表达基因。

方法: 收集新疆医科大学第三附属医院 2014 年 1 月—2016 年 6 月间手术切除的 10 例胰腺导管细胞癌组织及其相应癌旁正常组织, 用 TRIzol 法提取总 RNA 后采用含 49 395 个基因探针的 Affymetrix 基因表达谱芯片筛选出差异表达基因, 并行 GO 分析和 pathway 分析, 对部分差异表达基因行实时定量 PCR 法验证表达情况。

结果: 样本总 RNA 检测合格, 基因芯片质量评估良好, 检测结果可靠且可重复性高。经芯片降噪处理后共检测 38 079 个基因, 其中存在差异表达的基因共 512 个, 表达上调的基因 419 个, 表达下调的基因 93 个; 共 287 个差异表达基因编码与 3 个 GO 分类(生物学过程、分子功能、细胞组分)相关的蛋白; 共 29 条信号通路存在明显基因差异表达, 涉及 126 个基因。经 PCR 验证, 差异表达基因 CPB1、CELA3B、CPA1、POSTN、PLA2G1B、CTRC 及 SPINK1 的表达情况与基因芯片检测结果吻合。

结论: 胰腺导管细胞癌组织与癌旁正常组织间存在大量差异表达的基因, 这些基因主要与生物学过程、分子功能及细胞组分相关, 且参与了多个细胞信号通路的调控。

关键词

胰腺肿瘤; 基因表达谱; 芯片分析技术; 基因本体; 信号传导
中图分类号: R735.9

Screening of abnormally expressed genes in pancreatic cancer by Affymetrix gene expression chip

YI Chao¹, YIMAMUMAIMAITIJIANG·Abula¹, DING Wei¹, SU Yating², YAN Dong¹, LI Haijun³

(1. Department of Hepatobiliary Surgery, the Third Affiliated Hospital, Xinjiang Medical University, Urumqi, Xinjiang 830000, China;
2. Department of Medical Administration, the Fifth Affiliated Hospital, Xinjiang Medical University, Urumqi, Xinjiang 830000, China;
3. Department of General Surgery, Shenzhen Luohu People's Hospital, Shenzhen, 518001, China)

Abstract

Objective: To screen the differentially expressed genes in pancreatic carcinoma versus paired normal adjacent tissue by gene expression profile chip.

Methods: The surgical specimens of pancreatic ductal cell carcinoma and adjacent normal pancreatic tissue from 10 patients from January 2014 to June 2016 in the Third Affiliated Hospital of Xinjiang Medical University were

基金项目: 国家自然科学基金资助项目(81360328)。

收稿日期: 2016-12-27; 修订日期: 2017-02-19。

作者简介: 易超, 新疆医科大学第三附属医院住院医师, 主要从事肝胆胰腺肿瘤外科临床与基础方面的研究。

通信作者: 李海军, Email: suyating128@163.com

collected, and differentially expressed genes between them were screened by an Affymetrix gene expression profile chip containing 49 395 gene probes after total RNA extraction by TRIzol method, and then, GO and pathway analyses were performed. Finally, some differentially expressed genes were verified by real-time PCR.

Results: The total RNA of samples passed the test, gene chip assessment yielded high-quality information and results of the test were reliable and highly repeatable; 38 079 genes were examined after denoising process of the chips, involving 512 differentially expressed genes that included 419 up-regulated genes and 93 down-regulated genes; there were 287 differentially expressed genes encoding proteins that were related to the three GO ontologies (biological process, molecular function and cellular component), and 29 signaling pathways were associated with significant differential gene expressions, involving 126 genes. The expressions of differentially expressed genes CPB1, CELA3B, CPA1, POSTN, PLA2G1B, CTSC and SPINK1 determined by real-time PCR were consistent with the results of gene chip analysis.

Conclusion: There are a large number of differentially expressed genes between pancreatic ductal cell carcinoma and normal adjacent pancreatic tissue. These genes are mainly associated with biological processes, molecular functions and cellular components, and also involved in the regulation of multiple cellular signaling pathways.

Key words Pancreatic Neoplasms; Gene Expression Profiling; Microchip Analytical Procedures; Gene Ontology; Signal Transduction

CLC number: R735.9

胰腺癌发病率虽较低,但其病死率却位于恶性肿瘤的第4位,且近年来呈现逐年上升的趋势,是恶性程度最高的消化道肿瘤之一^[1-2]。近年来随着医疗技术的发展及对肿瘤发病机制的不断探索,使得各类恶性肿瘤的治疗取得了长足的进展,但因缺乏有效的早期诊断与靶向治疗手段,胰腺癌的预后改善仍不理想,成为严重危害人民生命健康的疾病之一^[3]。因此,探寻在胰腺癌发生发展过程中起关键性作用的调控基因,明确其核心调控环节,阐明胰腺癌恶性生物学特性产生与维持的机制具有极其重要的意义。本研究通过采用Affymetrix基因表达谱芯片检测胰腺癌及其癌旁组织中的基因差异表达情况,以期对胰腺癌早期诊断与靶向治疗基因位点的确定提供实验依据。

1 材料与方法

1.1 试剂与仪器

1.1.1 主要试剂 Agilent RNA 6000 Nano Kit (Agilent公司); GeneChip 3'IVT Express Kit、GeneChip Hybridization Wash and Stain Kit (Affymetrix公司); TRIzol 试剂盒、QIAGEN RNeasy Total RNA Isolation kit、Primer (R&F) (吉凯基因技术有限公司); M-MLV、dNTPs、Rnase Inhibitor (Promega公司); Oligo dT (上海生工生物工

程股份有限公司); Bulge-LoopTM miRNA qPCR Primer Set (广州市锐博生物科技有限公司); SYBR Master Mixture (TAKARA公司)。

1.1.2 主要仪器 Thermo Nanodrop 2000 (Thermo公司); Agilent 2100 (Agilent公司); GeneChip Hybridization Oven 645、GeneChip Fluidics Station 450、GeneChip Scanner 3000 (Affymetrix公司)、Nanodrop 分光光度计 (Thermo公司); 稳压电泳仪 (上海天能科技有限公司); 超细匀浆机 (FLUKO公司); real-time PCR 仪器 (Roche公司)。

1.2 研究样本

收集我院2014年1月—2016年6月间行手术切除的经病理证实的胰腺导管细胞癌组织标本及其对应癌旁正常组织(距离癌组织手术安全边界2 cm以上,且术后病理证实癌旁正常组织内无肿瘤浸润)标本共10组,所有患者均未合并有其他恶性肿瘤性疾病,术前均未接受任何治疗;所有患者均签署我院标本采集与科学研究使用知情同意书后方采集标本用于研究;所有样本提取总RNA经质检合格;标本中男6例,女4例;平均年龄61.38岁;分化程度:高分化5例,中分化2例,低分化3例。

待测样本质检方法与指标:参照TRIzol试剂盒说明提抽上述样本总RNA,并使用QIAGEN

RNeasy Total RNA Isolation Kit纯化总RNA, 利用Nanodrop 2000测定样本总RNA浓度及 A_{260} 与 A_{280} 值以评估样本纯度, 当 $1.7 < A_{260}/A_{280} < 2.2$ 时样品合格; 参照Schroeder等^[4]的方法通过Agilent 2100仪器分析样本总RNA的RIN值与28 S/18 S以检测RNA完整程度, 当 $RIN \geq 7.0$ 且 $28\ S/18\ S > 0.7$ 时样品合格。所有样本满足上述2条为质检合格。

1.3 体外转录 (IVT) 反应

基因芯片检测参照Villegas-Ruiz等^[5]的方法进行: 提取并稀释poly-A RNA (内参照) 后与上述质检合格的样本总RNA混合制备poly-A RNA/样本总RNA混合液, 加入一链合成反应液使用“第一链cDNA合成”程序孵育后合成cDNA, 加入二链合成反应液“第二链cDNA合成”程序, 孵育合成双链cDNA模板, 加入IVT (体外反转录) 反应液“IVT”程序, 反转录获得携带生物素标签的扩增RNA (amplified RNA, aRNA), 参照说明书用GeneChip 3'IVT Express Kit试剂盒纯化aRNA后加入aRNA片段化反应液“片段化 (fragmentation)”程序, 孵育获得片段化产物。芯片杂交: 向芯片中注入130 μ L预杂交液, 45 $^{\circ}$ C杂交炉预杂交10 min, 吸去预杂交液后加入使用“杂交 (hybridization)”程序, 加热的杂交反应液130 μ L, 45 $^{\circ}$ C杂交炉60 r/min杂交16~17 h, 杂

交完成后取出芯片自动洗染, 洗染完成后进行扫描获得数据 (表1), 上述操作均在Affymetrix基因芯片系统平台中完成。

表 1 IVT 反应程序

Table 1 IVT reaction process

程序	参数
第一链 cDNA 合成	42 $^{\circ}$ C, 2 h \rightarrow 4 $^{\circ}$ C保持
第二链 cDNA 合成	16 $^{\circ}$ C, 1 h \rightarrow 65 $^{\circ}$ C, 10 min \rightarrow 4 $^{\circ}$ C保持
IVT	40 $^{\circ}$ C, 16 h \rightarrow 4 $^{\circ}$ C保持
片段化	94 $^{\circ}$ C, 35 min \rightarrow 4 $^{\circ}$ C保持
杂交	98 $^{\circ}$ C, 10 min \rightarrow 45 $^{\circ}$ C, 3 min \rightarrow 45 $^{\circ}$ C保持

1.4 实时定量 PCR 验证

筛选出胰腺癌组织及其癌旁正常组织间差异表达倍数排名前10位的差异表达基因, 依据Primer Bank数据库提供的PCR引物信息合成引物 (表2), 以GAPDH为内参照, 按两步法real-time PCR试剂盒说明检测胰腺癌组织及其相应癌旁正常组织中上述基因的表达情况并绘制统计图, 反应程序为, Hold (预变性): 95 $^{\circ}$ C, 30 s, 1个循环; 两步法PCR: 95 $^{\circ}$ C, 5 s, 60 $^{\circ}$ C, 30 s, 40个循环; Dissociation (溶解): 95 $^{\circ}$ C, 15 s, 60 $^{\circ}$ C, 30 s, 95 $^{\circ}$ C, 15 s, 1个循环。

表 2 实时定量 PCR 基因引物

Table 2 Gene primers for real-time PCR

基因	上游引物	下游引物
CPB1	CCA GCA CGA CCC AGA TTG AC	AAC CCG GCT ATC AAA CTG AGC
CELA3(B)	AGT ACG ACC GTG CTG TGA AG	CAG GGT GTC TCG TTG GGA AG
CPA1	GCT GGT GTT GAG TGT CCT GTT	GGC TAC AGA GAT TCG GAG CAC
CLPS	ATT ATC AAC CTG GAG AAC GGT GA	TAG AGC GTC TTG ACA GAG CAC
CTRB1	CTG TCC AGG ATC GTG AAT GGG	TGC TGA ACT TGG GGT TCT TGA
POSTN	CTC ATA GTC GTA TCA GGG GTC G	ACA CAG TCG TTT TCT GTC CAC
PLA2G1B	TGT GGC AGT TCC GCA AAA TG	CAG CCG TAG TTG TTG TAT TCC A
PNLIPRP2	ACC CTT CAG CGA CCT GTA AAA	GCC AGT GAT TAG TTG GAA GTT GT
CTRC	ATC AGC AAC ACC CGG ACC TA	CAT AGC AGG GGT AGT CCT TGG
SPINK1	TCT ATC TGG TAA CAC TGG AGC TG	ACA CGC ATT CAT TGG GAT AAG T
GAPDH	GGA GCG AGA TCC CTC CAA AAT	GGC TGT TGT CAT ACT TCT CAT GG

1.5 统计学处理

上述芯片检测结果采用R-Project软件进行数据分析并绘图, 采用芯片信号强度分布统计及相对对数信号强度统计进行质量评估; 将所有样本中处于探针组信号强度排序最低20%范围内的探针组作为背景噪音予以滤除, 依照过滤后数据绘制散点图、火山图及层次聚类分析图, 并

采用基于经验贝叶斯分布的线性模型计算差异水平 (P 值); 依据 | 差异倍数 (fold change) | > 3.0 且 $P < 0.05$ 的标准筛选出显著差异表达基因; 采用基因本体 (gene ontology, GO) 注释富集分析对上述显著差异表达基因从参与的生物过程 (biological process, BP)、分子功能 (molecular function, MF) 及细胞组分 (cellular

component, CC) 进行分类, 通过Fisher精确检验评价某基因的富集度显著水平; 采用基于KEGG与BioCarta通路注释富集分析法 (pathway analysis), 以通路为单位, 通过Fisher精确检验分析计算各个通路基因富集度的显著水平, 确定受到显著影响的代谢和信号转导途径。

2 结果

2.1 芯片数据质量评估

芯片信号强度分布曲线见图1, 横坐标代表芯片探针的信号强度区间, 纵坐标代表在不同信号强度区间内芯片探针集数目。由图可见, 不同样本芯片信号强度分布曲线重合度较高, 证实检测数据可靠性良好。

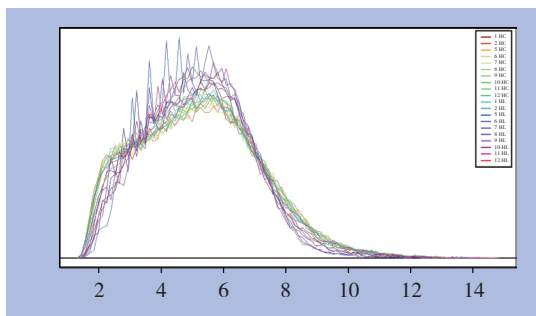


图1 芯片探针信号强度分布曲线图
Figure 1 Signal histogram of the chip probes

芯片相对对数信号强度箱线图见图2, 横坐标代表样本名称, 纵坐标代表相对对数信号强度值, 正中横线表示所有样本芯片的相对对数信号强度平均值, 每个箱外的上下横线表示该样本芯片相对对数信号强度上下90%置信区间, 箱的上下边表示四分位点, 箱正中横线表示中位数。如图所示不同样本芯片相对对数信号强度箱较为接近, 提示该检测数据可重复性良好。

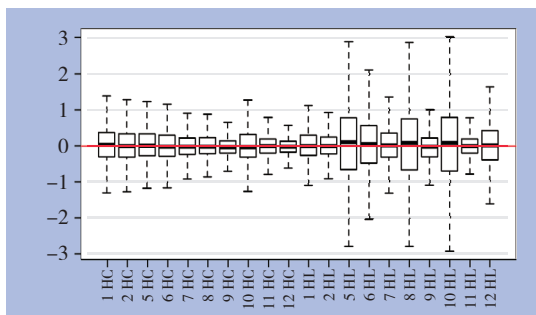


图2 芯片相对对数信号强度箱线图
Figure 2 Box plots of relative signal intensity values

2.2 基因表达谱芯片检测分析结果

本研究中所采用的Affymetrix基因表达谱芯片总计探针组数为49 395组, 经降噪处理后选取其中的38 079组探针检测结果进行基因差异表达分析。依据显著差异性表达基因筛选标准, 筛选出差异表达基因512个, 占检测到的总基因数的1.34%, 其中表达上调的基因419个, 占总有效检测基因数的1.10%, 表达下调的基因93个, 占总有效检测基因数的0.24%。差异表达倍数排名前10位的差异基因详见表3。

表3 差异表达倍数排名前10位的基因

基因符号	基因 ID	差异倍数	变化方向	P
CPB1	1360	33.76	下调	4.34×10^{-5}
CELA3 (B)	23436	28.83	下调	2.26×10^{-5}
CPA1	1357	26.18	下调	5.13×10^{-5}
CLPS	1208	26.06	下调	3.37×10^{-4}
CTRB1	1504	23.28	下调	2.18×10^{-4}
POSTN	10631	19.95	上调	4.68×10^{-6}
PLA2G1B	5319	19.85	下调	4.25×10^{-5}
PNLIPRP2	5408	19.04	下调	1.63×10^{-4}
CTRC	11330	16.95	下调	3.63×10^{-4}
SPINK1	6690	16.10	下调	7.99×10^{-4}

软件分析绘制的散点图 (图3)。此图显示了胰腺癌组织与癌旁正常组织间芯片信号强度在直角坐标系平面上的分布情况, 横坐标表示胰腺癌组织样本, 纵坐标表示对应的癌旁正常组织样本, 图中每个点代表一个探针组在癌组织与癌旁正常组织中的信号表达强度, 平行绿色线代表差异参考线, 两项内区域的点表示无显著表达变化的探针组, 区域外红色的点代表在癌旁组织样本组中相对上调的探针组, 绿色点表示在癌组织样本组中相对表达上调的探针组。

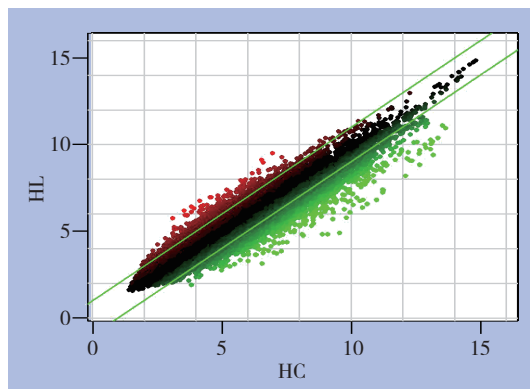


图3 芯片探针信号强度散点分布图
Table 3 Scatter plots of signal intensity values of the chip probes

软件分析绘制的火山图 (volcano plot) (图4), 此图依据两组样本间基因表达差异倍数 (fold change) 和显著性检验的P值两个因素绘制, 横坐标代表差异倍数, 纵坐标代表P值, 数值均经以2位底的对数变换。红色点代表符号筛选标准的差异表达基因, 灰色点为其他无明显差异表达的基因。

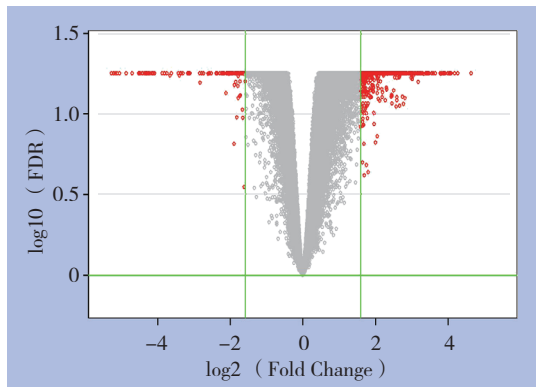


图 4 差异表达基因火山图
Figure 4 Volcano plots of differentially expressed genes

利用显著差异性表达基因筛选标准筛选的差异基因的表达谱对癌组织和癌旁正常组织两组样本进行层次聚类的热图 (图5)。在聚类分析热图中, 每一列代表1个样本, 每一行代表1个差异基因; 上部树状结构是根据差异基因的表达谱, 所有样本的聚集或分类情况; 左侧树状结构表示差异基因的表达模式聚集情况; 红色表示基因的表达程度相对上调, 绿色表示基因的表达程度相对下调, 黑色表示基因的表达没有明显变化, 灰色表示基因的信号强度未检出。

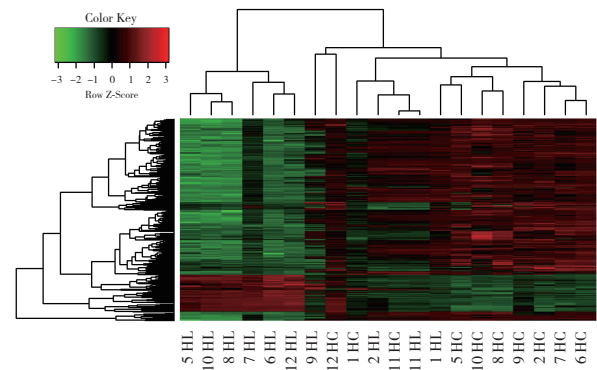


图 5 层次聚类分析图
Figure 5 Hierarchical cluster analyses

2.3 差异表达基因筛生物信息分析

GO分析: 依据Gene Ontology数据库对基因的注释信息, 参与编码与BP、MF、CC这3个GO分类相关蛋白的差异基因共有287个; 如表4 (依据基因富集程度及显著性检验筛选出排名前5的GO分类) 所示, 编码BP相关蛋白的差异基因主要包括编码信号传导功能、炎症反应、蛋白质代谢过程、解剖结构改变及细胞内大分子代谢相关蛋白的基因; 编码MF相关蛋白的差异基因主要包括编码肽酶活性、肽链内切酶活性、丝氨酸肽链内切酶活性、丝氨酸肽酶活性及丝氨酸水解酶活性相关蛋白的基因; 编码CC相关蛋白的差异基因主要包括编码胞外区、细胞膜、质膜、细胞外基质、细胞质相关蛋白的基因。依据基因差异表达倍数及显著性检验结果筛选出上述BP、MF及CC分类中排名前5位的差异表达基因 (表5)。

表 4 GO 分类差异基因富集程度

Table 4 Degrees of enrichment of differentially expressed genes by GO classification

GO 分类	GO 集合名称	基因数	P
BP	SIGNAL_TRANSDUCTION	77	1.28×10^{-28}
	IMMUNE_SYSTEM_PROCESS	31	3.17×10^{-20}
	PROTEIN_METABOLIC_PROCESS	52	1.86×10^{-20}
	ANATOMICAL_STRUCTURE_DEVELOPMENT	45	5.00×10^{-16}
	CELLULAR_MACROMOLECULE_METABOLIC_PROCESS	47	1.33×10^{-15}
MF	PEPTIDASE_ACTIVITY	23	1.35×10^{-18}
	ENDOPEPTIDASE_ACTIVITY	14	2.56×10^{-11}
	SERINE_TYPE_ENDOPEPTIDASE_ACTIVITY	8	9.29×10^{-9}
	SERINE_TYPE_PEPTIDASE_ACTIVITY	8	2.02×10^{-8}
	SERINE_HYDROLASE_ACTIVITY	8	2.42×10^{-8}
CC	EXTRACELLULAR_REGION	50	9.19×10^{-36}
	EXTRACELLULAR_REGION_PART	80	3.50×10^{-25}
	MEMBRANE	63	5.40×10^{-22}
	INTEGRAL_TO_MEMBRANE	19	9.37×10^{-19}
	PLASMA_MEMBRANE	72	1.59×10^{-18}

表5 GO分析中各分类差异倍数排名前5位基因

Table 5 Top five differentially expressed genes for each GO ontology

GO 分类	基因符号	基因 ID	差异倍数	变化方向	P
BP	PRSS2	5645	14.62	下调	7.30×10^{-5}
	RBPJL	11317	14.01	下调	2.54×10^{-5}
	REG1A	5967	13.46	下调	4.36×10^{-5}
	CUZD1	50624	12.63	下调	1.12×10^{-4}
	PDIA2	64714	10.18	下调	6.05×10^{-5}
MF	CPB1	1360	33.76	下调	4.34×10^{-5}
	CPA1	1357	26.18	下调	5.13×10^{-5}
	PLA2G1B	5319	19.85	下调	4.25×10^{-5}
	PNLIPRP2	5408	19.04	下调	1.63×10^{-4}
	CTRC	11330	16.95	下调	3.63×10^{-4}
CC	POSTN	10631	19.95	上调	4.68×10^{-6}
	FN1	2335	15.20	上调	2.49×10^{-5}
	SULF1	23213	11.19	上调	3.31×10^{-6}
	CELA2(B)	51032	10.79	下调	2.58×10^{-4}
	LUM	4060	9.78	上调	4.07×10^{-4}

pathway分析: 依据KEGG与BioCarta数据库提供的信号通路信息, 分析结果显示共29条信号通路存在明显的基因差异表达, 共涉及126个基因, 其中表达上调的基因共124个, 表达下调的基因共2个, 详细结果见表6; 根据通路内差异基因富集程度及差异基因占通路基因相对比率筛选出差异最明显的通路为黏着斑激酶(focal adhesion kinase, FAK)通路, 该通路中存在的差异表达基因详见表7。

表6 pathway分析结果

Table 6 Results of pathway analysis

通路	通路基因	P
KEGG_FOCAL_ADHESION	27	6.30×10^{-22}
KEGG_PATHWAYS_IN_CANCER	22	7.95×10^{-12}
KEGG_REGULATION_OF_ACTIN_CYTOSKELETON	18	1.83×10^{-11}
KEGG_ECM_RECEPTOR_INTERACTION	17	2.05×10^{-17}
KEGG_LEUKOCYTE_TRANSENDOTHELIAL_MIGRATION	14	2.88×10^{-11}
KEGG_CELL_ADHESION_MOLECULES_CAMS	13	1.84×10^{-9}
KEGG_CHEMOKINE_SIGNALING_PATHWAY	12	8.73×10^{-7}
KEGG_COMPLEMENT_AND_COAGULATION_CASCADES	11	1.49×10^{-10}
KEGG_FC_GAMMA_R_MEDIATED_PHAGOCYTOSIS	11	6.26×10^{-9}
KEGG_LEISHMANIA_INFECTIOIN	10	4.15×10^{-9}
KEGG_SMALL_CELL_LUNG_CANCER	10	1.91×10^{-8}
KEGG_LYSOSOME	10	6.22×10^{-7}
KEGG_VIRAL_MYOCARDITIS	8	9.76×10^{-7}
KEGG_ANTIGEN_PROCESSING_AND_PRESENTATION	8	4.44×10^{-6}
BIOCARTA_RHO_PATHWAY	7	3.44×10^{-8}
KEGG_PATHOGENIC_ESCHERICHIA_COLI_INFECTIOIN	7	2.73×10^{-6}
KEGG_RENAL_CELL_CARCINOMA	7	8.70×10^{-6}
KEGG_MELANOMA	7	9.57×10^{-6}
KEGG_CHRONIC_MYELOID_LEUKEMIA	7	1.15×10^{-5}
BIOCARTA_CDC42RAC_PATHWAY	6	9.11×10^{-9}
BIOCARTA_MPR_PATHWAY	6	1.31×10^{-6}
BIOCARTA_SALMONELLA_PATHWAY	5	1.45×10^{-7}
BIOCARTA_ACTINY_PATHWAY	5	1.64×10^{-6}
BIOCARTA_P53HYPOXIA_PATHWAY	5	3.47×10^{-6}
BIOCARTA_ECM_PATHWAY	5	4.35×10^{-6}
BIOCARTA_MCALPAIN_PATHWAY	5	5.39×10^{-6}
KEGG_ASTHMA	5	1.38×10^{-5}
BIOCARTA_CLASSIC_PATHWAY	4	1.07×10^{-5}
BIOCARTA_PLATELETAPP_PATHWAY	4	1.07×10^{-5}

表 7 FAK 通路差异表达基因

基因符号	基因 ID	差异倍数	变化方向	P
FN1	2335	15.20	上调	2.49×10^{-5}
THBS2	7058	10.52	上调	1.89×10^{-5}
COL11A1	1301	9.45	上调	1.79×10^{-5}
COL3A1	1281	8.88	上调	1.95×10^{-5}
LAMC2	3918	7.33	上调	1.73×10^{-5}
COL5A1	1289	7.17	上调	3.42×10^{-5}
COL5A2	1290	7.12	上调	2.34×10^{-5}
COL1A2	1278	7.02	上调	1.45×10^{-5}
COL4A1	1282	6.35	上调	2.53×10^{-6}
COL6A3	1293	4.67	上调	1.10×10^{-4}
RAP1B	5908	4.33	上调	2.80×10^{-5}
COL1A1	1277	4.18	上调	1.43×10^{-4}
MYLK	4638	4.17	上调	1.80×10^{-5}
ITGAV	3685	4.02	上调	2.23×10^{-5}
MYL12A	10627	3.98	上调	1.83×10^{-5}
BIRC3	330	3.85	上调	2.18×10^{-4}
VCL	7414	3.82	上调	2.21×10^{-6}
PIK3R2	5296	3.62	上调	4.51×10^{-5}
COL4A2	1284	3.62	上调	1.60×10^{-6}
CAPN2	824	3.60	上调	1.92×10^{-5}
ITGB1	3688	3.40	上调	4.31×10^{-6}
RAC1	5879	3.30	上调	3.17×10^{-5}
RHOA	387	3.29	上调	7.54×10^{-5}
MET	4233	3.28	上调	5.25×10^{-6}
ITGA1	3672	3.27	上调	9.75×10^{-6}
PIK3CG	5294	3.04	上调	6.18×10^{-5}
ITGB8	3696	3.03	上调	1.71×10^{-5}

2.4 real-time PCR 验证

表 3 所列差异表达基因实时定量 PCR 检测 (图 6), 图中各基因的相对表达量按相对表达量 = $2^{-\Delta CT}$ 公式计算, 其中 ΔCT = 目的基因 CT 值 - 内参基因 CT 值, 由图可见, 除 CLPS、CTRB1 及 PNLIPRP2 基因外, 其余差异基因在胰腺癌组织内的表达水平均明显高于癌旁正常组织。

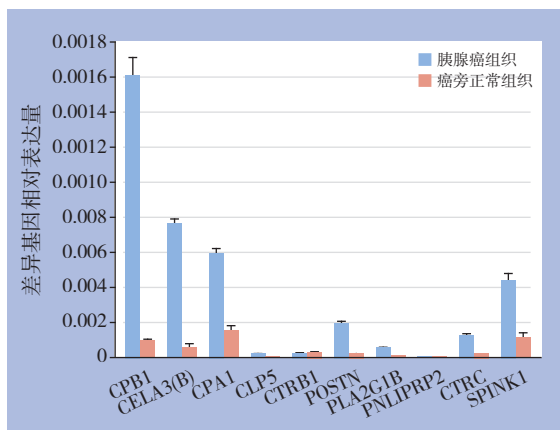


图 6 差异表达基因相对表达量

Figure 6 Relative expression levels of the differentially expressed genes

3 讨论

近年来逐步发展成熟的基因表达谱芯片是目前最常用于该领域研究的工具之一, 它克服了既往针对单一或少数基因分析的缺陷, 最大程度的整合所能采集到的生物信息综合分析疾病发展过程中多基因的表达及其功能^[6-7]。本研究中所采用的 Affymetrix 基因表达谱芯片, 是一种包含目前人类全基因组信息且较之 cDNA 更为准确、灵敏且可信度更高的基因芯片, 是目前应用最为广泛的基因表达谱芯片之一^[8-9]。利用此表达谱芯片, 本研究检测了 10 例胰腺癌组织及其相应癌旁正常组织间基因差异表达的情况, 并通过芯片信号强度及相对对数信号强度分布情况证实了检测分析结果的可靠性与可重复性, 且实时定量 PCR 验证结果也显示, 差异表达倍数排名前 10 位的基因中除 CLPS、CTRB1 及 PNLIPRP2 基因外其余基因的表达情况均与芯片检测结果相一致。依据差异分析结果绘制的散点图与火山图能更为直观的反映差异表达基因在全部检测基因当中的分布情况; 而层次聚类分析则是对样本及芯片检测所得基因定量信息两个维度以相似性为基础的分组与归类: 一方面, 以在不同标本中共同差异表达基因为基础对样本进行双向聚类分析, 结果显示多数癌旁正常组织间与癌组织间均具有相似的表达谱, 而两者之间基因表达谱则存在一定差异, 体现了检测数据结果的合理性; 另一方面, 此结果有助于从基因集合中区分具有不同表达模式的基因子集合, 表达模式相近的基因 (基因簇) 可能具有相似的功能、参与同一生物学途径或在通路内位置接近, 如 EIF4G2 基因簇 (以差异表达倍数最高基因命名) 包含 TXN、ATP2A2、TMEM50A、PTGES3、RHOA、EIF4G2、MOB4、HSPE1-MOB4 及 HSPE1 基因, PSMA4 基因簇则包含 CD55、PON2、HIST1H2BK、CD46、PTTG1IP、SNRPE、ARPC5、DTX3L 及 PSMA4 基因。

基因本体 (gene ontology, GO) 是一个通过不断更新完善的注释词汇数据库对检测基因进行的标准化功能分类体系, 它主要通过参与的生物过程, 分子功能和细胞组分三个主要方面来描述生物体中基因及其产物的功能属性, 通过差异表达基因的 GO 分析可以从上述三个方面对检测到的

异常表达基因集合进行分类,各个分类中基因的比例可在一定程度上反映这些差异表达基因对生物体各种相应生物学功能的影响程度。本研究中检测到与参与生物过程相关的基因中,差异表达最显著的是PRSS2 (protease serine 2) 基因,该基因编码的胰蛋白酶原2是一种丝氨酸蛋白酶,可激活基质金属蛋白酶及间质胶原酶等,直接或间接的水解肿瘤周围的基质蛋白从而促进肿瘤的侵袭转移^[10-12]。胰腺癌研究方面,Cao等^[13]的研究显示PRSS2在胰腺癌组织中的表达明显高于癌旁组织与正常胰腺组织,而本研究的检测结果则显示PRSS2在癌旁正常组织中的表达是胰腺癌组织中的14.62倍,这可能与样本间的个体差异有关,但因本研究所检测的样本量有限,结果还需更进一步的研究证实。分子功能相关基因中,差异表达最显著的是羧肽酶B1 (carboxypeptidase B1, CPB1) 基因,该基因编码的蛋白是一类可选择性水解蛋白或多肽链羧基端赖氨酸和精氨酸的外分泌蛋白酶,目前已有研究证实该酶是胰腺功能紊乱的重要血清学标志物,在多种胰腺疾病的发生发展中起重要作用^[14-15]。在恶性肿瘤研究领域,CPB1的差异表达被证实可能与乳腺癌的发生密切相关^[16],但在胰腺癌方面国内外则鲜有其与CPB1表达相关性的研究报道,本研究结果显示,CPB1在胰腺癌组织中的表达显著下调,癌旁正常组织中CPB1的表达是胰腺癌组织中的33.76倍。编码分子结构相关蛋白的差异表达基因中,POSTN (periostin osteoblast specific factor) 基因在胰腺癌组织中的表达明显上调,是癌旁正常组织中的19.95倍。该基因编码的蛋白最早发现于小鼠成骨细胞系中,是一种由成骨细胞或成骨样细胞分泌的可促进成骨细胞及其前体在骨膜聚集于分化的分泌蛋白,目前已有大量的研究显示,该蛋白在许多组织中均有表达,特别是与恶性肿瘤的发生与转移密切相关。最近的研究^[17]显示,POSTN蛋白的异常表达与胰腺癌的转移潜能及其以后密切相关,其可通过干扰多种信号通路调控作用从而促进肿瘤血管的生成。GO分析可以反映出检测到的差异表达基因在不同生物功能中的富集程度,这对后续差异基因的功能研究有一定的指导意义,然而单纯依据该比例来评价各个分类受影响的显著程度是不准确的,还需要同时考虑各个分

类在总体基因集合(例如:芯片上的全部基因,该物种所有已知的基因等)中的分布情况。

pathway分析较之基因单体分析是更系统、全面地了解细胞的生物学过程、性状或疾病的发生机理、药物作用机制等最直接和必要的途径^[18]。在Pathway分析中,共检测到29条通路存在显著的基因差异表达,其中差异表达基因富集程度最高的是黏着斑激酶(Focal adhesion kinase, FAK)通路。该通路是整蛋白介导的信号转导通路中的关键环节,被证实在多种恶性肿瘤中存在异常表达,其高表达或磷酸化可能与包括胰腺癌在内的多种肿瘤的生长与侵袭转移能力密切相关^[19-21]。最近,Hsieh等^[22]的研究显示,Ovatodiolide(OV)可通过抑制MMP-9与FAK的表达从而阻断NF- κ B和STAT3通路显著抑制胰腺癌细胞的侵袭转移能力,证实FAK通路在胰腺癌的侵袭转移过程中起着关键性的作用。Ning等^[23]在其报道中指出,新发现的泛素水解酶家族成员USP22可能通过调节胰腺癌细胞的上皮间质化从而影响胰腺癌细胞的侵袭与转移,而在此过程中FAK信号通路则起着关键性的作用。FAK信号通路的成员中,本研究检测到共有27个基因存在显著的差异表达,且这些基因均为表达上调,其中以编码纤维连接蛋白1(fibronectin 1, FN1)的基因差异表达倍数最高,为15.20倍。该蛋白是纤维连接蛋白家族的成员之一,是细胞外基质的重要组成部分,在细胞黏附、移、损失修复及维持微血管完整性中起重要作用,此外,它还在多种信号传导通路中扮演着重要的角色,参与细胞的增殖分化及凋亡迁移过程^[24]。目前国内外已有较多研究证实, FN1基因的异常表达及突变与多种恶性肿瘤的增殖分化及侵袭转移密切相关。Wang等^[25]在其基于39例胰腺导管细胞癌与其相应癌旁正常组织基因差异表达数据的分析指出,在蛋白质交互网络(PPI Network)中FN1基因是一个关键的枢纽节点,证实FN1基因在胰腺癌的发生发展中可能发挥着极为重要的作用,这与本研究的检测结果一致。然而,由于国内外相关研究较少,目前尚缺乏进一步的分子生物化学研究证实FN1基因与胰腺癌恶性生物学特性产生与维持过程中的确切关系及其具体的分子机制。

除外上述目前已知的与胰腺癌相关的差异

表达基因外, 本研究还检测到共17个迄今为止尚未见报道在胰腺癌及其癌旁组织中存在明确异常表达的差异基因, 但因检测的样本较少, 研究结果具有一定的局限性, 还需要针对差异表达基因进行进一步的功能验证。本研究通过表达谱芯片检测初步了解了胰腺癌及其癌旁组织中基因表达情况、存在基因表达变化的生物学过程及信号传导通路, 为后续深入分析胰腺癌发生与发展的分子机制奠定了实验基础, 也为早期诊治基因靶点的选择提供了一定的理论依据。

志谢: 感谢新疆医科大学附属肿瘤医院肿瘤研究所全体实验人员及吉凯基因技术有限公司研究人员对本研究顺利开展予以的协助与支持。

参考文献

- [1] 杨尹默. 胰腺癌外科治疗的现状、存在问题与展望[J]. 中国普通外科杂志, 2016, 25(9):1231-1235. doi:10.3978/j.issn.1005-6947.2016.09.001.
Yang YM. Surgical managements of pancreatic cancer:current status and future directions[J]. Chinese Journal of General Surgery, 2016, 25(9):1231-1235. doi:10.3978/j.issn.1005-6947.2016.09.001.
- [2] 黄晓曼. 郑雷教授:胰腺癌的多学科综合诊疗[J]. 中国普通外科杂志, 2016, 25(3):318-320. doi:10.3978/j.issn.1005-6947.2016.03.002.
Huang XM. Professor Lei Zheng: Multidisciplinary management for pancreatic cancer[J]. Chinese Journal of General Surgery, 2016, 25(3):318-320. doi:10.3978/j.issn.1005-6947.2016.03.002.
- [3] Jensen EH, Armstrong L, Lee C, et al. Neoadjuvant interferon-based chemoradiation for borderline resectable and locally advanced pancreas cancer: a Phase II pilot study[J]. HPB (Oxford), 2014, 16(2):131-139. doi: 10.1111/hpb.12086.
- [4] Schroeder A, Mueller O, Stocker S, et al. The RIN: an RNA integrity number for assigning integrity values to RNA measurements[J]. BMC Mol Biol, 2006, 7:3.
- [5] Villegas-Ruiz V, Moreno J, Jacome-Lopez K, et al. Quality control usage in high-density microarrays reveals differential gene expression profiles in ovarian cancer[J]. Asian Pac J Cancer Prev, 2016, 17(5):2519-2525.
- [6] 郑建伟, 杨淑娟, 利小平, 等. 口腔鳞状细胞癌及癌旁正常组织基因表达谱芯片检测[J]. 中国组织工程研究, 2015, 19(27):4365-4370. doi:10.3969/j.issn.2095-4344.2015.27.019.
Zheng JW, Yang SJ, Li XP, et al. DNA chip-based gene expression profiling of oral squamous cell carcinoma tissue and normal paracarcinoma tissue[J]. Journal of Clinical Rehabilitative Tissue Engineering Research, 2015, 19(27):4365-4370. doi:10.3969/j.issn.2095-4344.2015.27.019.
- [7] 马永权, 张雪峰, 蒋会勇. 组织微阵列技术的研究进展[J]. 中国普通外科杂志, 2009, 18(9): 979-983.
Ma YQ, Zhang XF, Jiang HY. Research advances in tissue microarray technique[J]. Chinese Journal of General Surgery, 2009, 18(9): 979-983.
- [8] McCall MN, Murakami PN, Lukk M, et al. Assessing affymetrix GeneChip microarray quality[J]. BMC Bioinformatics, 2011, 12:137. doi: 10.1186/1471-2105-12-137.
- [9] Masuda Y, Yokose S, Sakagami H. Gene Expression Analysis of Cultured Rat-Endothelial Cells after Nd: YAG Laser Irradiation by Affymetrix GeneChip Array[J]. In Vivo, 2017, 31(1):51-54.
- [10] Gou S, Yu J, Wang C, et al. Three female familial cases of solid pseudopapillary tumors with a protease serine 1 gene mutation[J]. Pancreas, 2013, 42(1):168-173. doi: 10.1097/MPA.0b013e3182554276.
- [11] Paliwal S, Bhaskar S, Nageshwar Reddy D, et al. Association Analysis of PRSS1-PRSS2 and CLDN2-MORC4 Variants in Nonalcoholic Chronic Pancreatitis Using Tropical Calcific Pancreatitis as Model[J]. Pancreas, 2016, 45(8):1153-1157. doi: 10.1097/MPA.0000000000000608.
- [12] 赵强, 张志伟, 肖娟, 等. 组织芯片分析PRSS2蛋白在胃癌组织中的表达与意义[J]. 中南医学科学杂志, 2015, 43(3):245-248. doi: 10.15972/j.cnki.43-1509/r.2015.03.002.
Zhao Q, Zhang ZW, Xiao J, et al. Significance and Expression of PRSS2 Proteins Analyzed by Tissue Array in Gastric Cancer Tissue[J]. Medical Science Journal of Central South China, 2015, 43(3):245-248. doi: 10.15972/j.cnki.43-1509/r.2015.03.002.
- [13] Cao J, Xia C, Cui T, et al. Correlations between serum trypsinogen-2 and pancreatic cancer[J]. Hepatogastroenterology, 2015, 62(138):435-440.
- [14] Szabó A, Pilsak C, Bence M, et al. Complex Formation of Human Proelastases with Procarboxypeptidases A1 and A2[J]. J Biol Chem, 2016, 291(34):17706-17716. doi: 10.1074/jbc.M116.743237.
- [15] Nakano E, Geisz A, Masamune A, et al. Variants in pancreatic carboxypeptidase genes CPA2 and CPB1 are not associated with chronic pancreatitis[J]. Am J Physiol Gastrointest Liver Physiol, 2015, 309(8):G688-694. doi: 10.1152/ajpgi.00241.2015.
- [16] Bouchal P, Dvořáková M, Roumeliotis T, et al. Combined Proteomics and Transcriptomics Identifies Carboxypeptidase B1 and Nuclear Factor κ B (NF- κ B) Associated Proteins as Putative Biomarkers of Metastasis in Low Grade Breast Cancer[J]. Mol Cell Proteomics, 2015, 14(7):1814-1830. doi: 10.1074/mcp.

- M114.041335.
- [17] Liu Y, Li F, Gao F, et al. Periostin promotes tumor angiogenesis in pancreatic cancer via Erk/VEGF signaling[J]. *Oncotarget*, 2016, 7(26):40148–40159. doi: 10.18632/oncotarget.9512.
- [18] Kanehisa M, Furumichi M, Tanabe M, et al. KEGG: new perspectives on genomes, pathways, diseases and drugs[J]. *Nucleic Acids Res*, 2016, 45(D1):D353–361. doi: 10.1093/nar/gkw1092.
- [19] 谷小虎, 肖景东, 邢晓静, 等. Biglycan 及 FAK 信号通路对结肠癌细胞侵袭、转移能力的影响及机制[J]. *实用肿瘤学杂志*, 2015, 29(5):444–449. doi:10.11904/j.issn.1002–3070.2015.05.013.
- Gu XH, Xiao JD, Xing XJ, et al. Effects and mechanism of Biglycan and FAK signaling pathway on the invasion and metastasis of colon cancer cells[J]. *Practical Oncology Journal*, 2015, 29(5):444–449. doi:10.11904/j.issn.1002–3070.2015.05.013.
- [20] 赵美娜, 陈公琰. FAK 在恶性肿瘤中的研究进展[J]. *实用肿瘤学杂志*, 2016, 30(2):162–166. doi:10.11904/j.issn.1002–3070.2016.02.015.
- Zhao MN, Chen GY. Research progress on focal adhesion kinase in malignant tumors[J]. *Practical Oncology Journal* 2016, 30(2):162–166. doi:10.11904/j.issn.1002–3070.2016.02.015.
- [21] Yang K, Gao K, Hu G, et al. CTGF enhances resistance to 5-FU-mediated cell apoptosis through FAK/MEK/ERK signal pathway in colorectal cancer[J]. *Onco Targets Ther*, 2016, 9:7285–7295.
- [22] Hsieh YJ, Tseng SP, Kuo YH, et al. Ovatodiolide of *Anisomeles indica* Exerts the Anticancer Potential on Pancreatic Cancer Cell Lines through STAT3 and NF- κ B Regulation[J]. *Evid Based Complement Alternat Med*, 2016, 2016:8680372. doi: 10.1155/2016/8680372.
- [23] Ning Z, Wang A, Liang J, et al. USP22 promotes epithelial-mesenchymal transition via the FAK pathway in pancreatic cancer cells[J]. *Oncol Rep*, 2014, 32(4):1451–1458. doi: 10.3892/or.2014.3354.
- [24] Fagerberg L, Hallström BM, Oksvold P, et al. Analysis of the human tissue-specific expression by genome-wide integration of transcriptomics and antibody-based proteomics[J]. *Mol Cell Proteomics*, 2014, 13(2):397–406. doi: 10.1074/mcp.M113.035600.
- [25] Wang Y, Li Y. Analysis of molecular pathways in pancreatic ductal adenocarcinomas with a bioinformatics approach[J]. *Asian Pac J Cancer Prev*, 2015, 16(6):2561–2567.

(本文编辑 宋涛)

本文引用格式: 易超, 依马木买买提江·阿布拉, 丁伟, 等. Affymetrix 基因表达谱芯片技术筛选胰腺癌异常表达基因的研究[J]. *中国普通外科杂志*, 2017, 26(3):294–303. doi:10.3978/j.issn.1005–6947.2017.03.005

Cite this article as: Yi C, Yimamumaimaitijiang·ABL, Ding W, et al. Screening of abnormally expressed genes in pancreatic cancer by Affymetrix gene expression chip[J]. *Chin J Gen Surg*, 2017, 26(3):294–303. doi:10.3978/j.issn.1005–6947.2017.03.005

本刊 2017 年各期重点内容安排

本刊 2017 年各期重点内容安排如下, 欢迎赐稿。

第 1 期 肝脏肿瘤的临床与基础研究

第 2 期 胆道疾病的诊断与外科手术治疗

第 3 期 胰腺手术及其消化道重建术式

第 4 期 胃肠肿瘤及腹部外科微创治疗

第 5 期 乳腺、甲状腺肿瘤的外科规范化治疗

第 6 期 血管疾病的腔内微创及综合治疗

第 7 期 肝脏外科疾病及肝移植

第 8 期 胆道肿瘤的临床与基础研究

第 9 期 胰腺肿瘤的临床与基础研究

第 10 期 胃肠道肿瘤手术及综合治疗

第 11 期 乳腺、甲状腺疾病的临床与基础研究

第 12 期 血管外科疾病的治疗进展及其他

中国普通外科杂志编辑部