

cDNA 芯片研发及其在肥胖患者脂肪基因差异表达研究中的应用*

赵 黄¹⁾ 骆天红¹⁾** 王侃侃²⁾ 张 济²⁾ 郑培丞²⁾ 赵春军²⁾
刘优萍¹⁾ 郑 升¹⁾ 顾燕云¹⁾ 李 果¹⁾ 罗 敏¹⁾

(¹⁾ 上海第二医科大学附属瑞金医院, 上海市内分泌代谢病临床研究中心, 上海市内分泌研究所,
医学基因组学国家重点实验室, 上海 200025;

²⁾ 上海第二医科大学附属瑞金医院, 上海市血液病研究所, 医学基因组学国家重点实验室, 上海 200025)

摘要 为了加快基因功能的研究, 利用已有的来源于不同组织的 cDNA 克隆, 并通过交换和购买补充了低丰度和染色体覆盖不完全的部分 cDNA, 研制开发出具有相当代表性、覆盖较完全的高密度 cDNA 表达型基因芯片, 每张芯片上含有 384 个质控 DNA 和 12 630 个 cDNA 探针, 其中包括 12 508 个 Unigene 和 122 个表达序列标签 (EST)。利用这些芯片, 对肥胖患者及正常人内脏脂肪组织基因表达谱进行了初步研究, 并发现在肥胖患者内脏脂肪组织差异表达的基因, 其中上调的有与凋亡相关的基因、与免疫有关的基因以及与能量代谢有关的基因等, 而下调的主要是与脂肪酸及胆固醇合成有关的基因, 对这些基因进一步的功能研究将为阐明肥胖发生机制奠定基础。

关键词 cDNA 芯片, 肥胖, 脂肪组织, 差异表达

学科分类号 R58

基因芯片是近年来逐渐发展起来的高通量地研究基因差异表达的一种新的技术手段。与以往研究基因表达的方法相比, 表达型基因芯片可经过一次杂交试验, 获知样品中各种不同 RNA 的表达情况, 具有大规模、高通量、快速简便等特点。为了加快基因功能的研究, 我们利用上海血液学研究所和国家人类基因组南方研究中心的来源于人造血干/祖细胞^[1]、下丘脑-垂体-肾上腺^[2]、肝脏和肝癌^[3]等不同组织的各种全长 cDNA, 或部分测序的 cDNA 克隆共 10 000 多个, 并针对低丰度和染色体覆盖不完全的部分, 适当补充了 Resgene 公司的部分 cDNA 克隆, 经过优化研制开发出具有相当代表性、覆盖较完全的 cDNA 表达芯片。利用这些芯片, 对不同体重指数患者的内脏脂肪组织基因表达谱进行研究, 并初步探明了肥胖患者内脏脂肪组织差异表达的基因。

1 材料和方法

1.1 材料及样本

1.1.1 材料: 芯片点样仪 (Generation III Array Spotter, Amersham Pharmacia Biotech 公司)、芯片扫描仪 (GenePix™ 4000B, Axon Instruments)、扫描分析软件 (GenePix™ Pro 4.1)、自动测序仪 (Perkin Elmer 377)、玻璃芯片 (TURNER Designs,

Full Moon BioSystems FMB)、芯片质控对照 (Lucidea Microarray ScoreCard v1.1 Amersham Pharmacia Biotech 公司)、Spike mix (Cy3 或 Cy5 标记 PRK 1161 Amersham Biosciences 公司)、dNTP set 100 mmol/L (Gibco BRL Life Technologies Gibco BRL)、逆转录酶 Superscript II (Gibco BRL)、RNA 酶抑制剂 (华美生物工程公司)、Cy Dyes-monofunctional (Amersham Biosciences 公司)、Cot I (Invitrogen 公司), 实时定量 PCR 仪 (iCycle, BioRad 公司)。

1.1.2 cDNA 克隆: 来源于人造血干/祖细胞、下丘脑-垂体-肾上腺、肝脏和肝癌等不同组织的各种全长 cDNA 或部分测序的 cDNA 克隆共 10 000 多个, 并针对低丰度和染色体覆盖不完全的部分, 适当补充了 Resgene 公司的部分 cDNA 克隆。

1.1.3 脂肪组织样本: 选择 2 名正常人、2 名单纯肥胖患者网膜脂肪组织, 其中正常人体重指数小于 23, 无糖尿病、高血压及脂代谢紊乱等, 单纯肥胖患者体重指数大于 27, 无糖尿病、高血压及

* 国家重点基础研究发展规划项目 (973) (2002BA711A05) 和国家自然科学基金资助项目 (30100084, 30100085)。

** 通讯联系人。

Tel: 021-64370045-611119, E-mail: luotianhong68@yahoo.com.cn

收稿日期: 2004-06-16, 接受日期: 2004-08-30

脂代谢紊乱等。所有标本按 Trizol 说明书操作抽提总 RNA 备用。另将来自于 5 名正常人、4 名单纯肥胖患者、4 名单纯 2 型糖尿病患者及 2 名肥胖伴 2 型糖尿病患者的内脏脂肪组织, 亦按上述方法抽提总 RNA, 并将这些标本的 RNA 等量混合后, 做成参照样品 RNA 池。

1.2 方法

1.2.1 cDNA 模板制备: 将带有目的 cDNA 片段的菌液在 5% MR-2001 抗性高浓度培养基中孵育, 菌液经 PCR 鉴定是否带有目的基因片段, 若 PCR 产物电泳时见到一条清晰的条带, 则继续通过测序鉴定克隆, 经双向测序后得到长度超过 500 bp 的序列, 运用 Blast 程序将测序结果分别和 Unigene 库、GenBank 库以及 dbEST 库比较再次鉴定序列, 同源性 > 90%, 且配对碱基数 > 150 bp 的序列归类备选。

1.2.2 点样: 在 Full Moon BioSystems (FMB) 的多功能表达芯片玻璃基质 (Multi-Functional Microarray Slide) 上, 分 2 个点样区域, 每个区域又分为 12 个亚区 (矩阵), 每个亚区点样密度为 34 列 × 16 行, 点样总数为 13 056。点样完成后在点样仪中放置约 20 ~ 30 min 以使探针自然干燥, 经 254 nm 紫外光照射交联, 使点样的探针 cDNA 固定于玻片。

1.2.3 样品标记: 通过直接标记法, 在逆转录过程中, 分别在两管样品上标记 Cy3 和 Cy5 荧光染料。标记反应条件如下: 总 RNA 50 μg, 0.5 g/L 的 oligo dT 2 μl, Spike mixCy3-label* 5 μl (或 Spike mixCy5-label 5 μl), 用无 RNA 酶的水补足体积至 21.5 μl, 70℃ 保温 5 min, 然后分别加入 5 × 缓冲液 8 μl, 0.1 mol/L DTT 4 μl, 2 mmol/L dATP, dCTP, dGTP, 1 mmol/L dTTP 2 μl, 逆转录酶 Superscript II (200 U/μl) 2 μl, RNA 酶抑制剂 (20 ~ 40 U/μl) 1 μl, Cy3-dUTP (或 Cy5-dUTP) 1.5 μl, 42℃ 保温 90 min, 然后 70℃ 保温 10 min, 加入 2.5 mmol/L NaOH 4 μl, 37℃ 保温 15 min, 加入 2 mmol/L HEPES 20 μl, 反应结束后合并两管样品, 经 Millipore 纯化柱纯化, 最后用 20 μl 去离子水洗脱。

* Spike mix 为 Lucidea™ Microarray ScoreCard™ 提供的与质控 DNA 探针相对应的质控 mRNA, 含有线性范围对照 (DR, dynamic range) 和比率对照 (RC, ratio control) 样品。其混合物的浓度及比例见表 1。

Table 1 Control mRNA spike mix components

Sample	Cy3: Cy5 ratio	ρ (spike mix) / ($\mu\text{g} \cdot \text{L}^{-1}$)	
		Cy3	Cy5
1DR	1:1	6 600	6 600
2DR	1:1	2 000	2 000
3DR	1:1	200	200
4DR	1:1	66	66
5DR	1:1	20	20
6DR	1:1	6.6	6.6
1RC	1:3	200	600
2RC	3:1	600	200
3RC	1:10	200	2 000
4RC	10:1	2 000	200

1.2.4 杂交: 制备杂交液 20 × SSC 6 μl, Cot I 10 μl, Poly A 2 μl, 1 mmol/L HEPES 1 μl, 10% SDS 1 μl, 样品 20 μl, 总体积 40 μl, 13 400 g 离心 3 min, 95℃ 变性 2 min, 65℃ 保温 10 min。将杂交液加于芯片一侧, 加盖玻片, 放入盛有少量 3 × SSC 的密封盒内, 65℃ 杂交过夜。

1.2.5 图像采集: 采用 GenePix™ 4000B Scanner 扫描仪对芯片进行扫描, 并采用 GenePix™ Pro4.1 分析软件自动进行杂交点搜寻、排列, 最后人工校对排列杂交点, 并作标记。

1.2.6 数据标化: 采用整张芯片上所有点的信号作标化, 满足下列条件的数据作进一步分析。a. 中位数比值 (ratio of medians)、比值中位数 (medians of ratio)、回归比率 (regression ratio) 三者之间的差异不大于 10%; b. 信噪比 (SNR) ≥ 2。

1.2.7 芯片质量控制: 运用 Lucidea Microarray ScoreCard v1.1 软件对各个质控点进行信号整合、标化, 并评估芯片的检测范围、重复性、灵敏度、特异性等参数。在芯片左侧点样区, 12 个点样矩阵第一行前 32 个点为 Lucidea™ Microarray ScoreCard™ 提供的质控点, 点样克隆如表 2 所示。

1.2.8 正常人及肥胖患者脂肪组织基因表达谱分析: 将 2 名正常人及 2 名肥胖患者脂肪组织的 RNA 以上述方法进行反转录, 并标记 Cy5 荧光染料。另取等量参照样品 RNA 以同样方法进行反转录, 并标记 Cy3 荧光染料, 再按前述方法, 将 2 名正常人 cDNA 分别与参照样品 cDNA 共同孵育, 并与芯片杂交, 然后通过图像采集、数据标化后得到 2 名正常人脂肪组织基因表达谱。同样, 将 2 名肥

Table 2 Quality control DNA provided by Lucidea™ Microarray ScoreCard™		
Number	Control probe	Gene Name
1	1PC	Total human DNA/complex target
2	1NC	<i>Arabidopsis thaliana</i> protein G1p
3	1DR	Yeast Intergenic Region from ChromXI
4	2DR	Yeast Intergenic Region from ChromXIV
5	3DR	Yeast Intergenic Region from ChromVII
6	4DR	Yeast Intergenic Region from ChromVIII
7	5DR	Yeast Intergenic Region from ChromXII
8	6DR	Yeast Intergenic Region from ChromXIII
9	1RC	Yeast Intergenic Region from ChromXIII
10	2RC	Yeast Intergenic Region from ChromXIII
11	3RC	Yeast Intergenic Region from ChromXIII
12	4RC	Yeast Intergenic Region from ChromXIII
13	2NC	PolydA oligonucleotide
14	1HG	Actin gamma-1 HS. 14376
15	2HG	Glyceraldehyde-3-phosphate dehydrogenase Hs. 169476
16	3HG	Ubiquinol-cytochrome c reductase core protein II Hs. 173554
17	4HG	Casein kinase II beta polypeptide Hs. 165843
18	Reserved	Empty
19	Reserved	Empty
20	Reserved	Empty
21	Reserved	Empty
22	5HG	Est highly similar to NY-REN-37 antigen Hs. 173684
23	3NC	Empty
24	6HG	Human hydroxymethyl glutaryl-CoA lyase Hs. 831
25	7HG	Neuroblastoma RAS viral oncogene homolog Hs. 69855
26	8HG	Eukaryotic translation initiation factor 4A, isoform 1 Hs. 129673
27	9HG	Ubiquinol-cytochrome c reductase core protein II Hs. 173554
28	10HG	Enoyl Coenzyme A hydrolase, short chain 1 Hs. 76394
29	11HG	Actin gamma-1 HS. 14376
30	4NC	<i>Bacillus subtilis</i> diaminopimelate decarboxylase lysA
31	5NC	<i>Bacillus subtilis</i> phenylalanine biosynthesis associated protein pheB
32	2PC	Total human DNA/complex target

PC: Positive control; NC: Negative control; DR: Dynamic range; RC: Ratio control; HG: Housekeeping gene.

胖患者 cDNA 分别与参照样品 cDNA 共同孵育并与芯片杂交后, 得到 2 名肥胖患者脂肪组织基因表达谱, 再将 2 名正常人及 2 名肥胖患者的基因表达谱各基因信号分别取均值后, 得到正常人及肥胖患者两组的基因表达谱, 将两组每个对应点数据相除后, 即可得到两组人群基因差异表达谱, 其中大于 2 或小于 0.5 的点表示在肥胖患者内脏脂肪组织中上调 200% 或下调 50% 以上的基因.

1.2.9 实时定量 PCR 验证芯片分析结果: 对芯片分析找到的差异表达基因, 进一步用 BioRad 的 iCycle 定量 PCR 仪进行实时定量 PCR 验证. 首先将用于芯片试验的 2 名正常人和 2 名肥胖患者脂肪组织的 RNA 逆转录成 cDNA. 反应过程如下: 10 × TaqMan RT 缓冲液 10 μl, 25 mmol/L MgCl₂ 22 μl, 2.5 mmol/L dNTPs 20 μl, 50 μmol/L 随机引物 5 μl, 20 U/L RNA 酶抑制剂 2 μl, 50 U/μl 逆转录酶 2.5 μl, 2 μg RNA, 无 Rnase 水 38.5 μl, 混匀后, 25℃ 10 min, 48℃ 30 min, 95℃ 5 min. 得到 cDNA 后, 再进行实时定量 PCR, 反应如下: PCR Master mix 12.5 μl, 1 μmol/L 正向引物 2.5 μl, 1 μmol/L 反向引物 2.5 μl, Bio-Rad calibration dye 0.25 μl, cDNA 2 μl, H₂O 5.25 μl, 总体积 25 μl (每个反应用相应标本的 RNA 做无扩增对照), 实时定量 PCR 结果分析参数设置如下: a. 荧光本底信号设定为前 10 个循环的荧光均值; b. 设定 PCR 第 3 ~ 15 个循环的荧光信号标准差的 10 倍为阈值; c. 每个反应管内的荧光信号到达设定的域值时所经历的循环数为 C_t 值, 以管家基因 GAPDH 来校正每个样品的 C_t 值; d. 通过校正后的 C_t 值比较每个基因在正常人和肥胖患者中的表达差异. 为保证定量 PCR 结果的可靠性, 我们用一个正常人标本逆转录得到的 cDNA 进行连续 5 个梯度稀释 (1:1, 1:8, 1:8², 1:8³, 1:8⁴), 并同时定量 PCR 反应, 扩增 GAPDH 基因, 以检验定量 PCR 系统的可靠性和准确性. 上游引物为 5' TTCGACAGTCAGCCGCATC TTCTT3', 下游引物为 5' CAGGCGCCCAATAC GACCAAATC 3'.

2 结 果

2.1 芯片制备

2.1.1 点样分布: 采用 Amersham Pharmacia Biotech 公司的 Generation III Array Spotter 点样仪, 结合 ASC-XT (Array Spotter Control-Extended

Technology) 软件, 保证芯片制作能采用高密度点样模式, 在 54.25 mm × 18.36 mm 的玻片点样范围中分成两个分别为 8.93 mm × 54.25 mm 的点样区域 (Spot Sets), 每个区域由 12 根点样针分别点出

12 个小区 (矩阵), 每个小区共有 cDNA 探针 34 × 16 = 544 个, 整张芯片共可点样 13 056 个。图 1 中左边 12 个矩阵第一行前 32 个探针为质控探针。

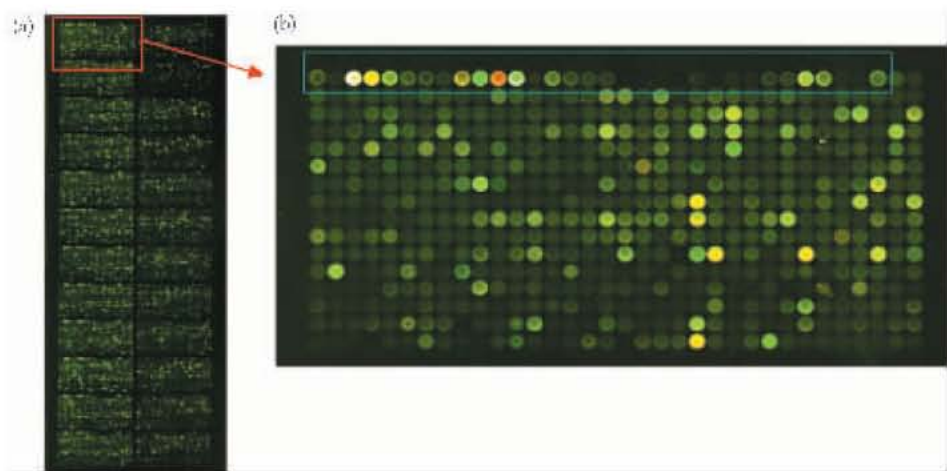


Fig. 1 Microarray image generated by GenePix™Pro scanner

(a) Image of the whole chip. (b) Enlarged image of the first block, spots framed in blue were quality controls provided by Lucidea™ Microarray ScoreCard™. All the spots had regular shape with smooth edge, neighboring spots were clearly distinguished without overlapping. Background was low and homogenous.

2.1.2 点样探针: 共 12 630 个 cDNA 克隆, 其中已经有染色体定位信息的共 9 522 个克隆, 分布于整个基因组的 22 条常染色体和 X, Y 两条性染色体上, 基本覆盖了整个基因组。基因分类参考 Gene Ontology (GO) 分类方法, 将芯片上的探针分为分子功能 (molecular function)、生物过程 (biology process) 和细胞成分 (cellular component)

3 大类, 其中分子功能又分为细胞分裂 (cell division)、细胞信号传导 (cell signalling/ cell communication)、细胞结构/运动 (cell structure/ mobility)、细胞器/防御 (cell organism/defense)、基因/蛋白质表达 (gene/protein expression)、代谢 (metabolism)、假基因 (pseudogene) 以及功能未知的基因 (unclassified) 等 8 类 (表 3)。

Table 3 Classification and proportion of probes spotted on the miroarray

Category I	Category II	Number	Proportion/%
Molecular function	Cell Division	595	4.711
	Cell signaling/cell communication	1 504	11.908
	Cell structure/motility	339	2.684
	Cell/organism defense	487	3.856
	Gene/protein expression	1 081	8.559
	Metabolism	1 485	11.758
	Pseudogene	30	0.238
	Unclassified	6 914	54.743
Biology process		80	0.633
Cellular component		115	0.911
Total		12 630	100

2.1.3 基因芯片的质量控制: a. mRNA 浓度与杂交信号的相关性. 在 5 μ l spike mixture 中分别加入 33 pg、100 pg、330 pg、1 000 pg、10 000 pg 及 33 000 pg 质控 mRNA, 当 mRNA 在 33 ~ 10 000 pg

的范围内, 信号强度与 mRNA 的量呈良好的线性关系 ($r_{\text{Cy3}} = 0.988$, $P < 0.001$, $r_{\text{Cy5}} = 0.991$, $P < 0.001$), 而当 mRNA 的量达到 33 000 pg 时, 杂交信号的强度趋于饱和 (图2). b. 灵敏度和特异

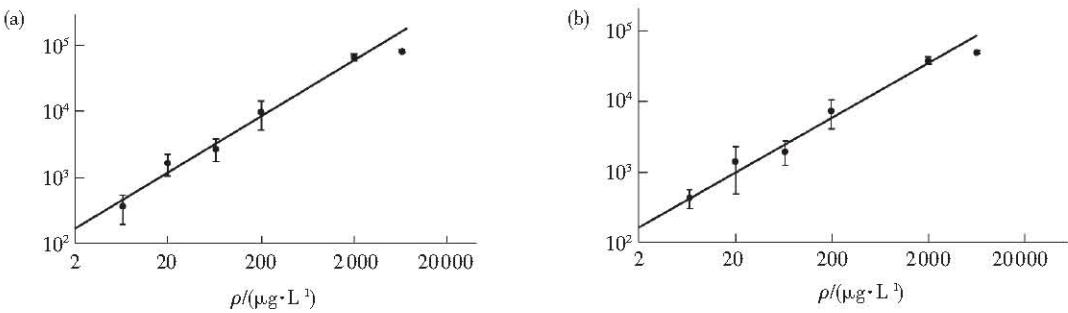


Fig. 2 Cy3 and Cy5 signal intensity plotted by concentration of mRNA
(a) Cy3; (b) Cy5.

性. 上述动力范围试验表明, 质控 mRNA 为 6.6 $\mu\text{g} \cdot \text{L}^{-1}$ spike mixture (约相当于每个细胞 5.67 拷贝) 时, 杂交信号仍能被准确检测 (信噪比 > 2), 显示该芯片具有较高的灵敏度. 同时, 阳性对照及 11 个管家基因均出现明显的杂交信号,

而阴性对照均无杂交信号出现, 表明该芯片具有很好的特异性 (图1). c. 重复性. 我们对 Lucidea Microarray ScoreCard 提供的同一个管家基因 (Actin, HG1 和 HG11) 在所有矩阵中的信号强度进行比较, 结果显示两种荧光标记信号的变异系数都小于默认阈值 20%, 表明该芯片具有较好的重复性. d. 差异表达率. 我们利用 Lucidea Microarray ScoreCard 提供的比率对照 (RC), 研究芯片检测基因差异表达的准确率. 结果显示, 10 个已知比率的探针其杂交信号经标化后检测到的差异表达率与预计值相比, 误差小于预计值的 50%, 表明该芯片能准确地测定基因的差异表达 (表4).

Table 4 The observed and normalized value for the spiked control analyzed by Lucidea Microarray ScoreCard

Control (Cy3/Cy5)	Expected	Observed	Normalized
1RC (1:3)	0.33	0.804 326	0.450 928
2RC (3:1)	3	6.279 557	3.255 861
3RC (1:10)	0.1	0.251 654	0.163 594
4RC (10:1)	10	13.386 51	7.082 63
1DR (1:1)	1	1.454 747	0.974 317
2DR (1:1)	1	1.520 987	0.999 106
3DR (1:1)	1	2.017 651	1.109 689
4DR (1:1)	1	2.137 618	0.977 616
5DR (1:1)	1	2.431 959	0.946 931
6DR (1:1)	1	3.119 992	0.899 486

2.2 肥胖患者内脏脂肪组织基因差异表达谱研究
2.2.1 芯片分析结果: 利用上述自制 cDNA 芯片, 我们比较了肥胖患者与正常人 (临床资料见表5) 内脏脂肪组织基因表达谱的差异, 结果正常人脂肪组织中共检测到 5 504 个基因 (两张芯片信噪比均 > 2), 肥胖患者脂肪组织中检测到 6 015 个基因 (两张芯片信噪比均 > 2) (图3, 4), 肥胖患者中

Table 5 Clinical data of normal subjects and obese patients

Group	Sex	Age	Height/cm	Weight/kg	BMI/ ($\text{kg} \cdot \text{m}^{-2}$)	FBG/ ($\text{mmol} \cdot \text{L}^{-1}$)
normal subject	female	42	151	52	22.8	4.86
normal subject	female	44	162	58	22.1	4.63
obese patient	female	51	150	62	27.56	5.51
obese patient	female	46	165	76	27.9	5.8

上调的基因共 119 个，其中上调 300% 以上的有 22 个基因，下调的基因共 46 个，其中下调 67% 以上的有 5 个（表 6）。

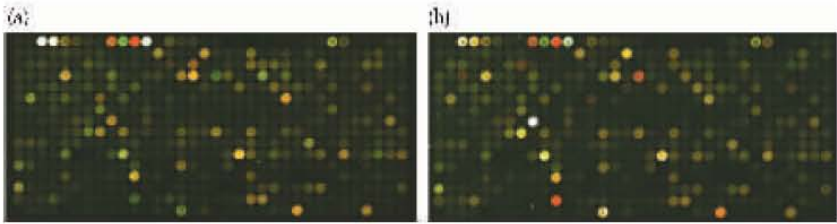


Fig. 3 Hybridization image of the same block of microarray from normal subject and obese patient
(a) normal subject; (b) obese patient.

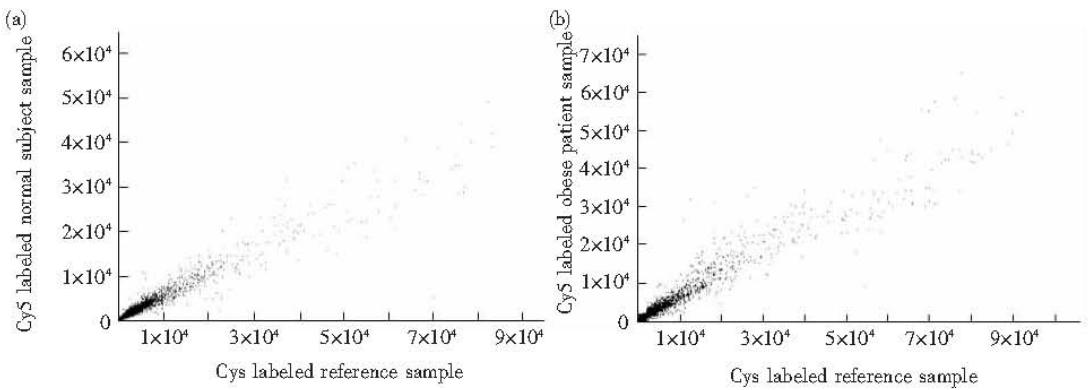


Fig. 4 Scattered plot graph of Cy3/Cy5 hybridizing signal
(a) normal subject; (b) obese patient.

Table 6 Microarray results

Gene name	Obese/ Normal	Category	Gene name	Obese/ Normal	Category
S100A11	2.9	Cell division	Coagulation factor Ⅲ	2.3	Cell structure/motility
BIRC2	2.3	Cell division	PRG1	2.2	Cell structure/motility
CAV2	2.2	Cell division	CCT2	2.1	Cell structure/motility
MYC	2.2	Cell division	MACF1	2.0	Cell structure/motility
PDCD6IP	2.2	Cell division	PA26	3.8	Cell/organism defense
NCKAP1	2.1	Cell division	LY117 (LST1)	3.6	Cell/organism defense
TACC1	2.1	Cell division	FCGR3B (FCGR3A)	2.9	Cell/organism defense
EMP1	2.1	Cell division	HSPA1A	2.4	Cell/organism defense
HBA2	5.2	Cell signaling/communication	HLA-DMB	2.4	Cell/organism defense
CTGF	3.2	Cell signaling/communication	LILRB3	2.3	Cell/organism defense
DTR	3.0	Cell signaling/communication	HLA-C	2.2	Cell/organism defense
ANXA1	2.8	Cell signaling/communication	SAA1	2.2	Cell/organism defense
TIEG	2.8	Cell signaling/communication	HLA-A	2.1	Cell/organism defense
RGSZ	2.6	Cell signaling/communication	SAA2	2.1	Cell/organism defense
CLIC4	2.6	Cell signaling/communication	FCG	2.1	Cell/organism defense
FTH1	2.5	Cell signaling/communication	MICA	2.1	Cell/organism defense

Continued					
Gene name	Obese/ Normal	Category	Gene name	Obese/ Normal	Category
PRKAR1A	2.5	Cell signaling/communication	HSPA8	2.0	Cell/organism defense
P23	2.5	Cell signaling/communication	KLF4	8.9	Gene/protein expression
Nidogen enactin	2.5	Cell signaling/communication	SCAMP5	5.5	Gene/protein expression
ATP5F1	2.3	Cell signaling/communication	NFE2L2	3.1	Gene/protein expression
HINT1	2.2	Cell signaling/communication	LRRFIP1	2.9	Gene/protein expression
G3BP	2.1	Cell signaling/communication	FHL1	2.6	Gene/protein expression
IKBKAP	2.1	Cell signaling/communication	ZFP36L1	2.5	Gene/protein expression
ATP5J	2.1	Cell signaling/communication	SSB	2.4	Gene/protein expression
NUP214	2.1	Cell signaling/communication	TFDP2	2.4	Gene/protein expression
FSTL1	2.0	Cell signaling/communication	SFRS7	2.3	Gene/protein expression
CRYAB	3.9	Cell structure/motility	RNASE1	2.1	Gene/protein expression
B2M	2.9	Cell structure/motility	EGR1	2.1	Gene/protein expression
CALD1	2.7	Cell structure/motility	BTF3	2.1	Gene/protein expression
SPON1	2.4	Cell structure/motility	GC20	2.0	Gene/protein expression
ARPC2	2.3	Cell structure/motility	NFE2L1	2.0	Gene/protein expression
SCO1	15.2	Metabolism	NCOA4	2.1	Unclassified
ADAMTS1	5.2	Metabolism	clone	2.1	Unclassified
PDK4	3.4	Metabolism	FLJ33282	2.1	Unclassified
ACP5	3.1	Metabolism	VIK	2.1	Unclassified
FTL	2.9	Metabolism	RAB6A	2.0	Unclassified
SAT	2.9	Metabolism	CCNB1	0.50	Cell Division
RISC	2.7	Metabolism	UTS2	0.45	Cell signaling/communication
PDK4	2.7	Metabolism	RANGAP1	0.46	Cell signaling/communication
SDCBP	2.4	Metabolism	TNFSF11	0.46	Cell signaling/communication
ASNS	2.3	Metabolism	VAPA	0.47	Cell signaling/communication
CTSB	2.3	Metabolism	ATP1B1	0.49	Cell signaling/communication
ARF4	2.2	Metabolism	C4A	0.38	Cell/organism defense
EST	2.2	Metabolism	RIP60	0.41	Cell/organism defense
TST	2.1	Metabolism	CML66	0.44	Cell/organism defense
UGP2	2.1	Metabolism	NOP56 (NOL5A)	0.47	Gene/protein expression
ASAH1	2.1	Metabolism	SRP46	0.47	Gene/protein expression
HSD11B1	2.0	Metabolism	SCD	0.35	Metabolism
ETR101	4.7	Unclassified	FASN	0.38	Metabolism
cDNA: FLJ23460 fis	4.7	Unclassified	SUCLG2	0.46	Metabolism
cDNA	4.1	Unclassified	PDHB	0.47	Metabolism
IER3	3.7	Unclassified	HMGCS1	0.45	Metabolism
KIAA1649 protein	3.4	Unclassified	HUXP1015 B08	0.27	Unclassified
BBX	3.4	Unclassified	AV707795	0.28	Unclassified
Formetallothionein	3.1	Unclassified	LOC55884	0.30	Unclassified
HAN11	3.0	Unclassified	CA2	0.35	Unclassified

Continued

Gene name	Obese/ Normal	Category	Gene name	Obese/ Normal	Category
HSPC072	3.0	Unclassified	FBP17	0.37	Unclassified
EST	2.9	Unclassified	FLJ23109	0.38	Unclassified
KIAA0220	2.8	Unclassified	EST	0.39	Unclassified
C11orf10	2.7	Unclassified	CloneMGC: 16362	0.39	Unclassified
H3F3A	2.7	Unclassified	MGC13096	0.41	Unclassified
AV704909 ADB	2.6	Unclassified	PRG4	0.42	Unclassified
ESTs,	2.6	Unclassified	Similar to hemopexin	0.43	Unclassified
USP24	2.6	Unclassified	FLJ20432	0.43	Unclassified
FLJ30532	2.5	Unclassified	N/A	0.43	Unclassified
cDNA clone	2.5	Unclassified	DC6	0.43	Unclassified
ITGAV	2.5	Unclassified	BAC RP11-426K3	0.43	Unclassified
NXP2	2.5	Unclassified	SLC27A4	0.44	Unclassified
CRI1	2.4	Unclassified	KIDINS220	0.44	Unclassified
Membrane-spanning4	2.4	Unclassified	SYNCOILIN	0.44	Unclassified
SUI1	2.3	Unclassified	ESTs	0.45	Unclassified
EST	2.3	Unclassified	RERE	0.46	Unclassified
DKFZP586A0522	2.3	Unclassified	KIAA1128	0.46	Unclassified
WDR18	2.3	Unclassified	MGC2488	0.46	Unclassified
ESTs	2.3	Unclassified	GCSH	0.47	Unclassified
HNRPH3	2.2	Unclassified	ZNF397	0.47	Unclassified
EHD4	2.2	Unclassified	KIF21A	0.48	Unclassified
MT1B	2.2	Unclassified	LOC56270	0.49	Unclassified
SMURF1	2.2	Unclassified	FLJ14308	0.49	Unclassified
HARS2	2.2	Unclassified	BAC RP11-466M21	0.49	Unclassified
ESTs	2.2	Unclassified	EST390446	0.49	Unclassified
ESTs	2.1	Unclassified	PP591	0.50	Unclassified
PSMA3	2.1	Unclassified			

(> 2 up-regulated, <0.5 down-regulated)

2.2.2 实时定量 PCR 验证芯片结果: a. 实时定量 PCR 系统可靠性的验证. 我们对来自同一样本连续 5 个梯度稀释的 cDNA 进行实时定量 PCR 反应 (GAPDH), 每个稀释度做复管, 并计算其 *Ct* 值与 cDNA 起始浓度的相关性, 结果表明, *Ct* 值与起始浓度具有很好的相关性 ($r = 0.989$), 表明该实时定量 PCR 系统能很好地反映 cDNA 的起始浓度 (图 5). b. 差异表达基因的验证. 为了验证芯片的结果, 我们进一步对 6 个上调的基因及 1 个下调的基因进行实时定量 PCR 检测, 我们利用芯片实验 2 个正常人标本和 2 个肥胖患者标本, 同时进行

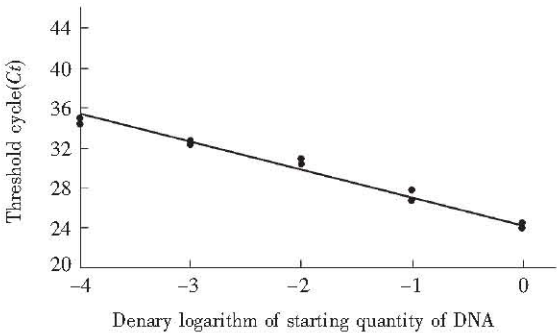


Fig. 5 Correlation of *Ct* value and starting quantity of DNA

实时定量 PCR 反应, 结果发现实时定量 PCR 的结果与芯片的结果具有较好的相关性, 即芯片中上调

的基因在定量 PCR 中也上调, 芯片中下调的基因在定量 PCR 中也下调 (图 6, 7).

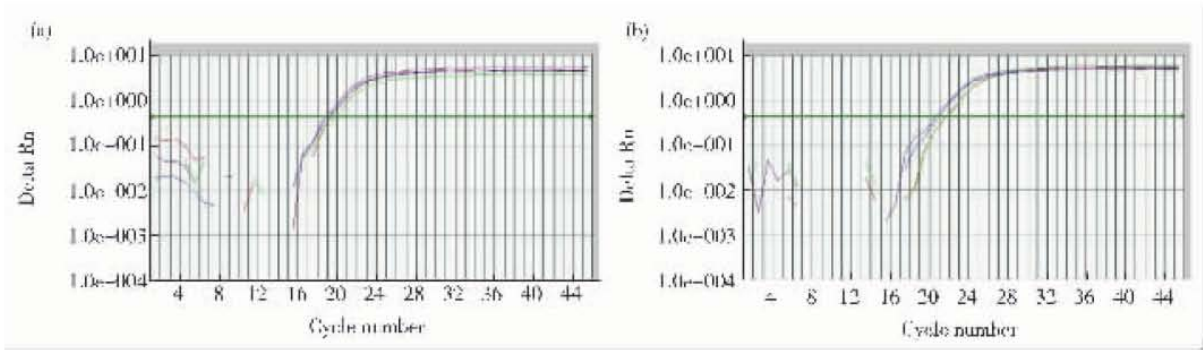


Fig. 6 Dynamic curves of Real-time PCR of SCD gene from normal subjects and diabetic patients
(a) Normal subjects. (b) Obese patients.

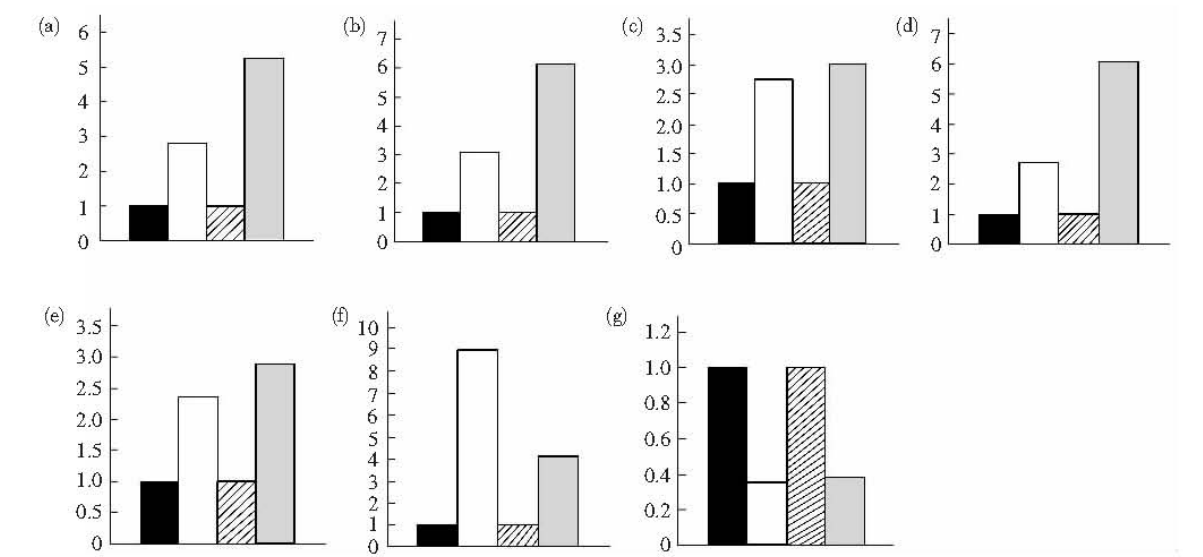


Fig. 7 Comparison of microarray and real-time PCR results, which showed good correlation between them
(a) ANX1; (b) ACP5; (c) TIEG; (d) PDK4; (e) ATP5F1; (f) KLF4; (g) SCD. Genes that have been shown to be up-regulated in microarray were also up-regulated in real-time PCR, while genes which were down-regulated in microarray were also down-regulated in real-time PCR.
■: Microarray, normal subjects; □: Microarray, obese patients; ▨: Real-time PCR, normal subjects; ▩: Real-time PCR, obese patients.

3 讨 论

表达型基因芯片是一种高通量的基因表达分析技术. 由于其能同时对大量基因进行高效快速分析, 得到较为全面的基因表达谱信息, 因此这一技术已被越来越广泛地应用于不同状态下组织基因表达谱的研究^[4].
制备本次实验芯片的 cDNA 克隆, 来源于上海血液学研究所和国家人类基因组南方研究中心建立的人造血干/祖细胞、下丘脑-垂体-肾上腺、肝脏

和肝癌等 cDNA 文库, 并从中筛选出来源于不同组织的各种全长 cDNA 或部分测序的 cDNA 近万个, 这些 cDNA 序列信息都运用 Blast 软件与 GenBank 数据库进行核对, 并通过 PCR 扩增、纯化及双向测序, 确认探针序列后才最后被采用, 因而保证了芯片上点样探针信息的可靠性. 同时为了保证所构建芯片在人类基因组的代表性和覆盖率, 针对低丰度表达和覆盖不完全的部分, 通过交换和购买, 适当补充了部分 cDNA 克隆. 我们对所点样的探针按 Gene Ontology (GO, <http://www.geneontology.org/>)

标准分成3大类,对于芯片上涉及多种功能或功能不明确的克隆则参考了其他多个大型数据库(如NCBI、EMI、EMBL、Gnecard等)的资料来选择可能的最主要的功能类别。

本次实验的芯片介质采用了美国Turner Designs有限公司基因芯片分部Full Moon BioSystems(FMB)的多功能基因表达芯片玻璃基质(multi-functional microarray slide)。这种新型玻璃基质采用了最新表面镀层技术,在玻璃表面镀上一层非常均匀的高分子聚合物,自身荧光背景极低,可使探测浓度范围达到4至5个数量级,特别适合自行设计制造高精度的生物芯片。点样仪采用了Molecular Dynamics的Generation III Array Spotter,结合ASC-XT(array spotter control-extended technology)软件的高密度点样模式,每块玻片最多能点样38 400个探针。本次实验制备的人cDNA高密度芯片目前共有探针12 630个,不同的质控对照384个。

本次实验中样品标记采用Cy3-dUTP及Cy5-dUTP双色荧光作标记,这两种荧光标记物在逆转录过程中具有较高的掺入率,且它们的激发光和发射光波长差别较大,因而可通过光栅将两者很好地区分开来,保证了在一次实验中同时检测受试细胞和对照细胞的RNA量。

杂交反应的质量和效率直接关系到芯片结果的准确性^[5]。影响杂交反应的因素有多种,如寡核苷酸探针密度、支持介质与杂交序列间的间隔序列长度、杂交序列长度、GC含量、探针浓度、核酸二级结构等。我们在实验过程中通过反复比较,优化了杂交条件,即保证了较高的信号强度,又得到了较高的信噪比,从而获得了可靠的实验数据。

为了进一步保证芯片的质量,我们在芯片的12个矩阵中加入了Lucidea™ Microarray ScoreCard™质控点。Lucidea™ ScoreCard是试验中用来确认数据质量的提高和标化工具,包含对照试剂和分析软件,可以应用于从点样到数据分析的整个芯片工作流程。我们在实验中利用Microarray ScoreCard软件对点样芯片进行了鉴定,结果显示,该芯片能检测从33 pg到10 000 pg的特异序列,并呈现良好的线性关系,表明该芯片有很好的动态检测范围及极高的检测灵敏度。重复试验显示,同一个管家基因(Actin, HG1和HG11)的两种荧光信号变异系数都小于默认阈值20%,且标化后的质控点(RC及DR)Cy3/Cy5的比值与预期值相比,差异均小于

预计值的0.5倍,表明该芯片具有很好的重复性和准确性,另外,阳性对照均出现明显的杂交信号,而阴性对照均无杂交信号出现,表明芯片具有很好的特异性。上述结果表明,我们实验室自己开发研制的高通量基因芯片具有代表性较强、覆盖较全面、重复性较好、可信度较高的优点,可以用于相关的基因表达谱研究。

在此基础上,利用这些自制的芯片比较了肥胖患者与正常人内脏脂肪组织基因表达谱的差异,结果发现,肥胖患者中上调的基因共119个,其中上调300%以上的有22个基因,下调的基因共46个,其中5个下调67%以上。在上调的基因中,我们发现多个与机体免疫调节有关的基因。如膜联蛋白1(annexin 1),这种蛋白质属于膜联蛋白超家族成员,这类蛋白质具有多个(通常为4个)重复功能域,可以和钙及磷脂结合。膜联蛋白1对于细胞生长、分化、细胞信号传递、花生四烯酸的释放等都具有重要的作用。而它的合成、磷酸化及胞内分布又受到了糖皮质激素的调控,除此以外,有证据表明,膜联蛋白1还介导了糖皮质激素对ACTH及其他激素的抑制作用,目前认为,膜联蛋白1是糖皮质激素的第二信使,在机体的免疫及抗炎过程中起着重要的作用^[6]。另外还有多个人类白细胞相关抗原,如HLA-DMB、HLA-C等均表现为上调,这一结果表明,肥胖患者脂肪组织的免疫功能趋于活跃。最近的研究表明,脂肪组织在免疫调节方面也扮演了重要的角色,脂肪组织能分泌多种和免疫调节有关的细胞因子,如IL-6、IL-8、TNF α 等。有研究报道肥胖患者体重减轻后,血清中IL-6和TNF α 水平降低,但肥胖患者免疫功能活跃究竟有何生理及病理意义仍有待进一步研究^[7]。另外我们发现多个与细胞凋亡有关的基因上调,如早期生长反应1基因(early growth response 1, EGR1),这是一种G0/G1转换调节基因,也是肿瘤抑制基因,EGR1可通过与TGF β 1及纤维连接蛋白(fibronectin, FN1)的启动子结合而调控这两个基因表达^[8]。EGR1依赖的TGF β 1基因表达可抑制肿瘤细胞生长。很多人类肿瘤细胞中极少或不表达EGR1,因而EGR1具有抑制生长和转化并诱导凋亡的作用。另一个上调的与细胞凋亡有关的基因为TGF β 诱导的早期相基因(TGF β inducible early gene, TIEG),这是一种Kruppel样锌指结构转录因子,它的表达受TGF β 1的调控,而TGF β 家族具有抗增殖及诱导凋亡的作用,TIEG的过度表达

会诱导细胞凋亡^[9]。

除此以外,我们还发现几个与能量及脂代谢有关的基因在肥胖患者中上调,如丙酮酸脱氢酶激酶(pyruvate dehydrogenase kinase 4, PDK4),该酶在调节葡萄糖氧化过程中起开关作用,PDK4通过使丙酮酸脱氢酶磷酸化而抑制其活性,阻断丙酮酸进入三羧酸循环,从而抑制葡萄糖氧化作用。例如,地松鼠(*Spermophilus tridecemlineatus*)冬眠时,其白色脂肪中PDK4基因的表达增加1500%,蛋白量增加800%,因而冬眠时脂肪酸代替碳水化合物成为能量的主要来源^[10]。我们的研究表明,肥胖患者的脂肪细胞中葡萄糖的氧化代谢受到抑制,脂肪酸代替碳水化合物成为能量的主要来源。

值得注意的是,在为数不多的下调基因中,就有3个基因分别和胆固醇及脂肪酸的合成有关,它们分别是3-羟基-3-甲基戊二酰辅酶A合成酶(HMG-CoA synthase)、脂肪酸合成酶及硬脂酰辅酶A脱氢酶(stearoyl-CoA desaturase),这3个酶分别是胆固醇及脂肪酸合成途径中的关键酶,这表明在肥胖患者中有生脂基因表达的下调,这一结果和国外学者在动物实验中得到的结果一致,如Nadler等^[11]利用Affymetrix寡核苷酸芯片筛查消瘦和肥胖小鼠脂肪组织基因表达谱的差异时发现,在肥胖小鼠脂肪组织中,与脂代谢相关的一些基因表达显著下降,这些基因包括ATP柠檬酸裂解酶、甘油3磷酸脱氢酶以及硬脂酰辅酶A脱氢酶等。提示肥胖小鼠脂肪细胞中这些基因的表达水平和脂肪细胞分化过程正好相反,类似于未成熟的前脂肪细胞。营养过度导致的肥胖病人虽然脂肪组织增加,但有功能的成熟脂肪细胞数目却减少,致使脂肪细胞更倾向于胰岛素抵抗,如在遗传性肥胖小鼠模型中,肝脏内不但甘油三酯水平升高,而且许多生脂基因表达水平也升高^[12],如SREBP-1、脂肪酸合成酶、ATP-柠檬酸裂解酶和苹果酶等,而这些基因在脂肪细胞中表达降低,迫使肝脏在脂质代谢中替代了脂肪细胞的作用。

为了进一步验证芯片的结果,我们对部分上调和下调的基因用荧光实时定量PCR方法(Real-Time PCR)进行了验证,我们挑选了6个在肥胖患者中上调的基因和1个下调的基因进行了验证,结果发现所有7个基因均得到了证实,即6个在芯

片实验中上调的基因在实时定量PCR中也上调,1个在芯片实验中下调的基因在实时定量PCR中也下调,这一结果表明,我们的芯片实验结果非常可靠。

综上所述,我们利用已有的cDNA文库建立了具有代表性较强、覆盖全面、重复性好、可信度高的cDNA表达芯片,并利用这种芯片,发现了肥胖患者内脏脂肪组织差异表达的基因,这些基因的发现将为进一步研究肥胖发病机制奠定基础。

参 考 文 献

- 1 Mao M, Fu G, Wu J S, *et al.* Identification of genes expressed in human CD34 (+) hematopoietic stem/progenitor cells by expressed sequence tags and efficient full-length cDNA cloning. *Proc Natl Acad Sci USA*, 1998, **95** (14): 8175 ~ 8180
- 2 Hu R M, Han Z G, Song H D, *et al.* Gene expression profiling in the human hypothalamus-pituitary-adrenal axis and full-length cDNA cloning. *Proc Natl Acad Sci USA*, 2000, **97** (17): 9543 ~ 9548
- 3 Xu X R, Huang J, Xu Z G, *et al.* Insight into hepatocellular carcinogenesis at transcriptome level by comparing gene expression profiles of hepatocellular carcinoma with those of corresponding noncancerous liver. *Proc Natl Acad Sci USA*, 2001, **98** (26): 15089 ~ 15094
- 4 Duggan D J, Bittner M, Chen Y D, *et al.* Expression profiling using cDNA microarrays. *Nat Genet*, 1999, **21** (1 Suppl): 10 ~ 14
- 5 David M. Gene expression studies using microarrays: principles, problems, and prospects. *Adv Physiol Educ*, 2002, **26**: 256 ~ 270
- 6 Roviezzo F, Getting S J, Paul-Clark M J S, *et al.* The annexin-1 knockout mouse: what it tell us about the inflammatory response. *J Physiol Pharmacol*, 2002, **53** (4 Pt 1): 541 ~ 553
- 7 Samuelsson L, Gottsater A, Lindgarde F. Decreasing levels of tumour necrosis factor alpha and interleukin 6 during lowering of body mass index with orlistat or placebo in obese subjects with cardiovascular risk factors. *Diabetes Obes Metab*, 2003, **5** (3): 195 ~ 201
- 8 Liu C, Yao J, de Belle I, *et al.* The transcription factor EGR-1 suppresses transformation of human fibrosarcoma HT1080 cells by coordinated induction of transforming growth factor-beta-1, fibronectin, and plasminogen activator inhibitor-1. *J Biol Chem*, 1999, **274** (7): 4400 ~ 4411
- 9 Tachibana I, Imoto M, Adjei P N, *et al.* Overexpression of the TGF-beta-regulated zinc finger encoding gene, TIEG, induces apoptosis in pancreatic epithelial cells. *J Clin Invest*, 1997, **99** (10): 2365 ~ 2374
- 10 Buck M J, Squire T L, Andrews M T. Coordinate expression of the PDK4 gene: a means of regulating fuel selection in a hibernating mammal. *Physiol Genomics*, 2002, **8** (1): 5 ~ 13
- 11 Nadler S, Stoehr J P, Schueler K L, *et al.* The expression of adipogenic genes is decreased in obesity and diabetes mellitus. *Proc Natl Acad Sci USA*, 2000, **97** (21): 11371 ~ 11376
- 12 Shimomura I, Bashmakov Y, Horton J D. Increased levels of nuclear SREBP-1c associated with fatty livers in two mouse models of diabetes mellitus. *J Biol Chem*, 1999, **274** (42): 30028 ~ 30032

Development of High-density cDNA Microarray and Its Application in The Study of Differential Gene Expression Profile of Adipose Tissue From Obese Patients *

ZHAO Yu¹⁾, LUO Tian-Hong¹⁾ **, WANG Kan-Kan²⁾, ZHANG Ji²⁾, ZHENG Pei-Zheng²⁾, ZHAO Chun-Jun²⁾,
LIU You-Ping¹⁾, ZHENG Sheng¹⁾, GU Yan-Yun¹⁾, LI Guo¹⁾, LUO Min¹⁾

(¹⁾ Ruijin Hospital Attached to Shanghai Second Medical University, Shanghai Endocrine and Metabolic Diseases Clinical Research Center,
Shanghai Institute of Endocrinology, State Key Laboratory of Medical Genomics, Shanghai 200025, China;

(²⁾ Ruijin Hospital Attached to Shanghai Second Medical University, Shanghai Institute of Hematology,
State Key Laboratory of Medical Genomics, Shanghai 200025, China)

Abstract To accelerate study of gene function, a home-made high-density cDNA microarray was developed, using the cDNA clones from different types of tissues. In order to cover as far as possible the whole genome, cDNA clones that essentially represented genes with low abundance were introduced from Resgene Company. Totally 384 quality control DNA and 12 630 cDNA fragments, which includes 12 508 Unigenes and 122 ESTs, were spotted on the chip, after optimization. The home-made cDNA microarray was evaluated with Microarray ScoreCard and proved to be representative and reliable. These microarrays were then used to study the differential gene expression profiles of visceral adipose tissues between obese patients and normal subjects. The results showed that genes involved in immune-regulation, apoptosis and metabolism of energy were up-regulated in obese patients, while genes involved in lipid synthesis were down-regulated in obese patients. Further exploration of these genes will greatly improve the understanding of the pathogenesis of obesity.

Key words cDNA microarray, obesity, adipose tissue, differential expression

* This work was supported by grants from The Special Funds for Major State Basic Research of China (2002BA711A05) and The National Natural Sciences Foundation of China (30100084, 30100085).

** Corresponding author. Tel: 86-21-64370045, E-mail: luotianhong68@yahoo.com.cn

Received: June 16, 2004 Accepted: August 30, 2004