

## 一种检测 microRNA 表达的 微阵列芯片的研制及应用\*

骆明勇<sup>1,2)</sup> 田志刚<sup>6)</sup> 徐智<sup>6)</sup> 张亮<sup>1,3)\*\*</sup> 王应雄<sup>2)\*\*</sup> 程京<sup>1,3,4,5)</sup>

<sup>1)</sup>生物芯片北京国家工程研究中心, 北京 102206; <sup>2)</sup>重庆医科大学公共卫生学院, 重庆 400016; <sup>3)</sup>博奥生物有限公司, 北京 102206;  
<sup>4)</sup>清华大学医学系统生物学研究中心, 北京 100084; <sup>5)</sup>清华大学生物科学与技术系, 北京 100084; <sup>6)</sup>北京大学第三医院, 北京 100083)

**摘要** MicroRNA (miRNA) 是一组内源性的非编码 RNA, 主要功能是采用降解靶 mRNA 和抑制靶 mRNA 的翻译两种作用方式调控基因的表达. 研究表明 miRNA 与动植物的发育、细胞生长分化和凋亡、脂肪代谢以及人类重大疾病密切相关. 检测特定状态下的组织或细胞的 miRNA 表达谱对深入研究 miRNA 的生物学功能具有重要的意义. 研制出了一种检测 miRNA 表达的微阵列芯片, 实验结果显示该芯片具有良好的重复性、灵敏度和特异性. 利用这个芯片检测了人脑、肝脏和心脏等组织以及 HeLa、HepG2 和 HL60 等细胞株的 miRNA 表达谱, 聚类分析结果显示, miRNA 表达模式具有组织特异性, 不同个体的同一组织的 miRNA 表达谱可以聚类在一起. 最后, 用 RT-PCR (reverse transcription-PCR) 技术检测了不同组织和细胞株中 miR-122a、miR-9 和 miR-124a 的表达, 结果显示它们分别在肝脏或脑组织中特异高表达, 和芯片结果一致.

**关键词** microRNA, 微阵列, 聚类分析, RT-PCR

**学科分类号** Q75

MicroRNA (miRNA) 是一组生物体基因组编码的内源的非编码小 RNA (non-coding small RNA), 成熟的 miRNA 长度 19~25 nt, 它们广泛存在于病毒、植物、到高等哺乳动物中. 在哺乳动物细胞中, 大多数 miRNA 是由基因间 DNA 序列编码, 也有部分 miRNA 是由位于基因内含子的 DNA 序列编码. 编码 miRNA 的基因组序列在 RNA 聚合酶 II 作用下产生原始 miRNA 转录本 (primary transcripts, pri-miRNA). pri-miRNA 在细胞核内被一种称为 “Drosha” 的 RNA 内切酶 III 处理, 产生 60~70 nt 的 miRNA 前体 (pre-miRNA), miRNA 前体从细胞核转运到细胞质中, 然后在细胞质被称为 Dicer 的第二种 RNA 内切酶 III 加工剪切为长度 19~25 nt 的成熟 miRNA. miRNA 主要采用降解靶 mRNA 和抑制靶 mRNA 的翻译两种作用方式调控基因的表达. 近年的研究发现, miRNA 在动植物的发育、细胞生长分化和凋亡、脂肪代谢等生命活动过程中发挥重要的调控作用<sup>[1~9]</sup>. 例如在肿瘤发生机制方面, miRNA 可能是潜在的癌基因或者抑癌基因, 并且由于 miRNA 是作为基因表达和蛋白质翻译过程中

的调节分子, 在肿瘤的发生过程中起到一个调控的枢纽作用, 已经有专家预测, 把 miRNA 作为肿瘤生物治疗的靶分子将比编码基因作为靶分子更加有效<sup>[10]</sup>. 目前人基因组 DNA 编码的 miRNA 已经确定的有 300 多个, 预计人基因组编码的 miRNA 超过 1 000 个. 为了深入研究这些 miRNA 的功能, 首先需要我们能够准确、快捷地检测 miRNA 在不同组织、不同发育阶段以及各种生理状态下的表达情况.

目前研究 miRNA 表达的方法有克隆测序、RNA 印迹以及 RNA 酶保护分析 (RNase protection assay, RPA). 近来通过设计茎环结构逆转录引物进行实时定量 PCR (real-time quantitative PCR) 检测 miRNA 表达的方法也已经建立起来<sup>[11]</sup>. 但以上的方法耗时耗力, 检测的通量不高. 在 mRNA 水平高通

\*国家科技攻关项目资助(2005BA711A05-04).

\*\* 通讯联系人.

张亮. Tel: 010-80726787, E-mail: lzhang@capitalbio.com

王应雄. Tel: 023-68485001, E-mail: wyx61221@yahoo.com.cn

收稿日期: 2006-06-04, 接受日期: 2006-08-26

量检测基因表达的生物芯片技术在 20 世纪 90 年代发展以来,发展到今天已经非常成熟,并在基因表达检测方面得到了广泛应用.2004 年以来,利用生物芯片技术来检测 miRNA 在不同组织、不同发育时相、不同疾病状态下表达的报道日渐增多.已有的 miRNA 芯片技术的具体路线差异很大,例如用来制备检测 miRNA 芯片的基片包括膜、玻璃片、微珠体等,检测 miRNA 来源的材料有的直接利用总 RNA,有的需要先把低分子质量 RNA 分离出来.探针的设计有的是正义链序列,有的是反义链序列.芯片上探针的长度有的为 20 nt 左右,有的为 40 nt 左右,也有的为 60 nt 左右.在检测技术上,有同位素检测、化学发光检测以及荧光检测.在标记方法上,有的采用随机引物反转录标记,有的采用 RNA 接头后通过 PCR 方法掺入标记,也有的在 RNA 末端直接进行标记.这些 miRNA 芯片的技术路线差别较大,但它们都在 miRNA 检测中发挥了重要作用<sup>[12~20]</sup>.

由于 miRNA 在生命活动中的重要调节功能,特别是 miRNA 与肿瘤发生的密切关系,国际上一些研究机构正在积极寻找能作为肿瘤生物治疗靶标的 miRNA 并取得知识产权的保护,例如美国俄亥俄州立大学 Croce 实验室,就申请了专利以保护 miR-15 和 miR-16 在慢性淋巴白血病和前列腺癌中诊断和治疗的潜在应用(美国专利申请号 20040152112).国外对 miRNA 的研究进展较快,芯片技术是大规模、高通量检测 miRNA 表达情况的理想技术,而目前国内还没有对 miRNA 芯片技术研究的相关报道.本研究旨在研制一种稳定、准确而操作简洁的 miRNA 检测微阵列芯片,为深入研究 miRNA 的功能提供高效的技术方法.

## 1 材料与方法

### 1.1 材料

HeLa、HepG2 和 HL60 细胞株为生物芯片北京国家工程研究中心冻存;人肝脏、脑和心肌组织总 RNA 购自美国 Ambion 公司(Cat#. 7960、7962 和 7966),部分组织样本来自北京大学第三医院;Trizol 试剂和 M-MLV 逆转录酶购自美国 Invitrogen 公司;T4 RNA 连接酶和低分子质量 DNA marker 购自美国 NEB 公司;5'-(p)CU-Cy3-3' RNA 在美国 Dharmacon 公司合成;miRNA Probe Construction kit (Cat#. 1550) 购自美国 Ambion 公司;SmartArray™ 48 芯片点样仪、LuxScan™ 10K-A 激

光共聚焦芯片扫描仪、芯片杂交盒、BioMixer™ II 芯片杂交清洗仪和 DNA 点样液都为博奥生物公司产品;LifterSlip™ 杂交盖片购自美国 Erie 公司;数据分析用软件 SpotData™ 软件为博奥生物公司产品,Significance Analysis of Microarrays 分析软件和 Cluster 3.0 软件下载自 Stanford University.

### 1.2 实验方法

**1.2.1 细胞培养.** HeLa 细胞用 DMEM、HepG2 细胞和 HL60 细胞用 RPMI 1640 培养基,所有培养基都含有 10% 的 FBS、50 mg/L 链霉素和 50 U/ml 青霉素,细胞在 37℃、CO<sub>2</sub> 浓度为 5% 的培养箱中培养.

**1.2.2 芯片探针的设计及芯片的点制.**

成熟的 miRNA 序列来自 miRNA Registry Release 7.0 (<http://microrna.sanger.ac.uk>, 2005 年 9 月),包括 313 条人、261 条小鼠和 196 条大鼠的 miRNA 序列.由于三者的 miRNA 保守性较强,具有很多共同的 miRNA 序列,去冗余后得到 387 条 miRNA 序列,另外我们还收集了 Xie 等<sup>[21]</sup>2005 年发表在 Nature 上预测的 miRNA 序列 122 条,最终在芯片上一共设计了 509 条成熟 miRNA 序列对应的探针.由于 miRNA 长度在 19~25 nt 左右,在设计探针时没有太多的选择,因此我们把所有的 miRNA 成熟序列都设计在探针中.考虑到探针较短的情况下,影响到探针在芯片表面的空间伸展,我们在探针上加上了与 miRNA 无关的序列使探针长度扩展到 40 nt,并在 5' 端进行氨基修饰.同时设计了 tRNA 和 U6 RNA 的探针作为内源阳性对照(内标).另外合成了与所有 RNA 没有同源性的 8 条 DNA 序列,然后用 Ambion 公司的 miRNA Probe Construction kit 体外转录生成 8 条长度为 30 nt 左右的 RNA.这 8 条 RNA 作为外源阳性对照(外标),以不同的浓度掺入到待检测的 RNA 样品中.

所有探针都在德国 MWG 公司合成,然后以 40 μmol/L 的浓度溶于 DNA 点样液中.探针经 SmartArray™ 48 芯片点样仪点制在经过醛基修饰的玻璃片上,每个探针重复 3 个点.点好的芯片通过探针上的氨基和芯片表面的醛基发生的希夫碱反应进行固定化处理,然后室温保存备用.

**1.2.3 靶 RNA 的制备及芯片杂交.**

细胞和组织总 RNA 抽提采用 Trizol 试剂,然后测定总 RNA 的 A<sub>260</sub>:A<sub>280</sub> 比值检测其纯度,其比值大约在 1.8~2.0,并用琼脂糖电泳检测其完整性.然后用聚乙二醇 6000(PEG 6000)分离出低分子质量

RNA (low molecular weight RNA, LMW RNA), 分离后的 RNA 片段用变性聚丙烯酰胺凝胶电泳 (PAGE) 鉴定其分离效果.

分离后的 LMW RNA 采用 T4 RNA 连接酶按照 Thomson 等<sup>[17]</sup>的方法进行标记. 简要的过程如下: 20  $\mu\text{g}$  总 RNA 分离出的 LMW RNA 用 500 ng CU RNA 和 2 单位 T4 RNA 连接酶进行标记, 然后用 0.3 mol/L 醋酸钠和 2.5 倍体积的无水乙醇进行沉淀, 最后溶于 15  $\mu\text{l}$  杂交液中 (3 $\times$ SSC、0.2% SDS 和 15% 甲酰胺).

杂交采用 LifterSlip™ 杂交盖片, 杂交时将杂交盒置于辅助杂交仪上, 辅助杂交仪可以促进盖片下的杂交液的流动. 在 42 $^{\circ}\text{C}$  水浴杂交过夜或 16 h 以上, 然后在 42 $^{\circ}\text{C}$  用洗液 I (2 $\times$ SSC、0.2% SDS)、洗液 II (0.2 $\times$ SSC) 分别洗 4 min, 然后离心甩干芯片.

#### 1.2.4 图像获得和数据处理方法.

芯片扫描采用 LuxScan™ 10K-A 激光共聚焦芯片扫描仪, 扫描参数设置以视觉获得 U6 的信号强度一致为准. 获得芯片的 TIFF 图后, 用 SpotData™ 软件提取数据.

在芯片数据处理时, 先筛选除去弱信号点. 由于一个探针点是由多个像素的信号构成的, 一个探针点中必须有超过 30% 的像素信号高于此探针本底背景信号值 $\pm 2\text{SD}$ , 此探针的信号才被视为有效

信号, 进入后续的数据分析. 然后用每张芯片上所有有效探针点信号的平均值进行归一化, 以便芯片之间的比较. 对于每一个样品, 我们进行 2 次重复的芯片杂交, 一个探针在每张芯片上重复了 3 个点, 这样每个 miRNA 就得到 6 个信号值. 为了确定各张芯片之间的 miRNA 表达差异的显著性, 我们用 SAM (significance analysis of microarrays) 分析软件中的 two class unpaired 进行芯片数据分析. miRNA 表达的聚类分析采用每张芯片平均值归一化后的数据, 用 Cluster 3.0 软件进行分析<sup>[21]</sup>.

**1.2.5 miRNA 表达的 RT-PCR (reverse transcription-PCR) 的验证.** miRNA 的 RT-PCR 方法参照 Chen 等<sup>[14]</sup>的方法. 简要过程如下: 先针对要检测的 miRNA 设计一条带茎环结构的逆转录引物以及相应的一对 PCR 引物 (序列见表 1). 然后进行逆转录, 10  $\mu\text{l}$  逆转录体系用 M-MLV 逆转录酶 100 单位、50 nmol/L 茎环结构的逆转录引物、总 RNA 50 ng. 最后用相应的 PCR 引物进行 PCR, PCR 条件为先 95 $^{\circ}\text{C}$  变性 10 min, 然后进入 PCR 循环, 95 $^{\circ}\text{C}$  15 s, 60 $^{\circ}\text{C}$  1 min, 进行 30 个循环. 同时采用 U6 RNA 作为内标<sup>[23]</sup>, 进行归一化. 我们得到的加长的 miRNA PCR 产物长度为 59 bp, U6 的 PCR 产物长度为 94 bp.

**Table 1 Primers used to perform RT-PCR of miRNAs**

Primer		Sequence (5'→3')
hsa-miR-122a	RT primer	GTCGTATCCAGTGCAGGGTCCGAGGTATTTCGCACTGGATACGACACAAAC
	PCR primer	Forward: GCCGTGGAGTGTGACAATGGT Reverse: GTGCAGGGTCCGAGGT
hsa-miR-9	RT primer	GTCGTATCCAGTGCAGGGTCCGAGGTATTTCGCACTGGATACGACTCATAAC
	PCR primer	Forward: CCCGTCTTTGGTTATCTAGCTGT Reverse: GTGCAGGGTCCGAGGT
hsa-miR-124a	RT primer	GTCGTATCCAGTGCAGGGTCCGAGGTATTTCGCACTGGATACGACTGGCAT
	PCR primer	Forward: CGTTAAGGCACGCGGTGA Reverse: GTGCAGGGTCCGAGGT
U6	PCR primer	Forward: CTCGCTTCGGCAGCAC Reverse: AACGCTTCACGAATTTGCGT

## 2 结 果

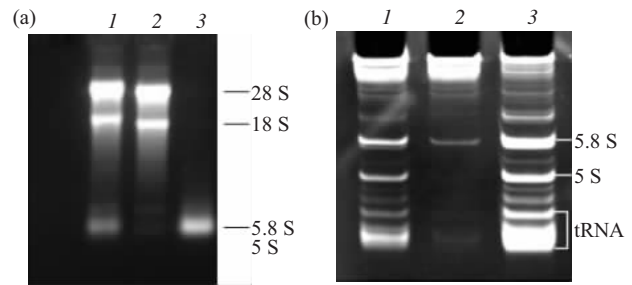
### 2.1 低分子质量 RNA 的分离和标记

采用 PEG 6000 分离低分子质量 RNA (LMW RNA), 分离前后的 RNA 分别进行了琼脂糖电泳和

PAGE (图 1), 可以发现用 PEG 6000 离心分离后 LMW RNA 得到了很好的富集. 同时, 为了检测分离过程对 miRNA 的分离富集效果, 我们把总 RNA 分离出来的两部分 RNA (低分子质量 RNA 和高分子质量 RNA) 分别进行标记和芯片杂交 (图 2a,

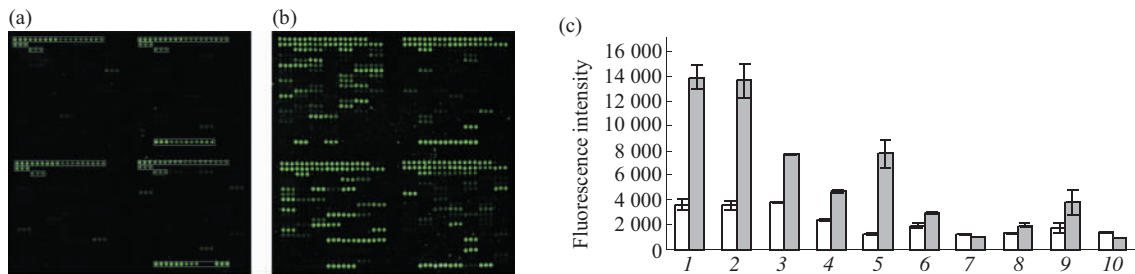


2b). 杂交结果显示, 用 LMW RNA 得到的芯片杂交信号信噪比较好, 而从 HMW RNA 的杂交结果来看, 只有几个探针出了较弱的信号, 我们把这几个探针的信号和 LMW RNA 杂交中相对应的探针的信号进行了比较分析, HMW RNA 中的信号远远低于 LMW RNA 中的信号(图 2c). 通过生物信息比对分析, 可以发现, 在人 mRNA 中有很多和 miRNA 有极高的序列相似性, 因此我们推测 HMW RNA 在芯片上产生的杂交信号是由于 HMW RNA 中众多的 mRNA 与 miRNA 探针发生交叉杂交的结果. 综合以上结果, 我们认为使用 PEG 6000 能有效地对 miRNA 进行分离富集.



**Fig. 1 Electrophoresis analysis of RNAs isolated by PEG6000 centrifugation method**

Two fractions, low molecular weight RNA (LMW RNA) and high molecular weight RNA (HMW RNA), were fractionated from total RNA after PEG6000 centrifugation method. 0.5  $\mu$ g of each fraction was resolved on a 1.2% denaturing agarose gel (a) or 15% denaturing polyacrylamide gel (b) (1: Total RNA; 2: HMW RNA; 3: LMW RNA).

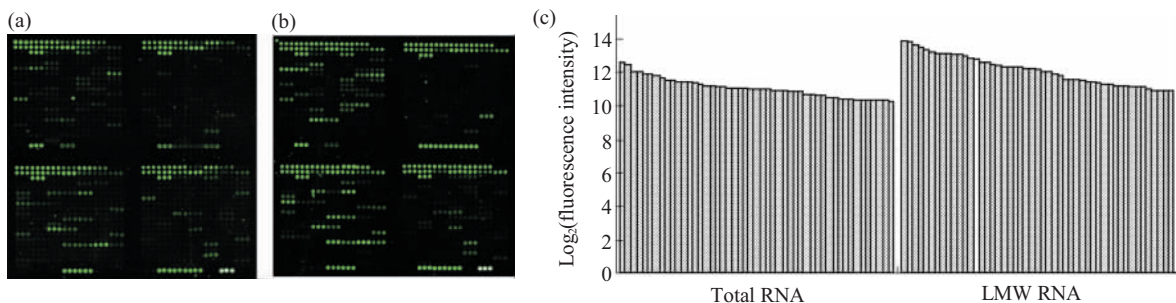


**Fig. 2 Validation of the LMW RNA isolating method**

Total RNA was fractionated into two fractions with PEG: high molecular weight RNA (HMW RNA) and low molecular RNA (LMW RNA). To validate the isolating method, HMW RNA (a) and LMW RNA (b) were labeled and hybridized to the miRNA microarray. Several probes have weak signal in HMW RNA hybridization picture (note: the probes marked with rectangles are the positive controls instead of the miRNA probes), their fluorescence intensity in HMW RNA and LMW RNA hybridization was compared (c). □: HMW RNA; ■: LMW RNA. 1: hsa-miR-23a; 2: hsa-miR-23b; 3: hsa-miR-155; 4: hsa-miR-188; 5: hsa-miR-15b; 6: hsa-miR-320; 7: hsa-miR-520e; 8: hsa-miR-134; 9: hsa-miR-125b; 10: hsa-miR-365.

我们比较了用总 RNA 和分离富集后的 LMW RNA 进行标记杂交的差别. 分别用 10  $\mu$ g HeLa 细胞总 RNA(图 3a)和用从 10  $\mu$ g HeLa 细胞总 RNA 分离富集后的 LMW RNA (图 3b) 进行标记杂交. 从结果可以看出, 用 LMW RNA 进行标记杂交得到

的信号强度大大高于直接用总 RNA 进行的标记杂交(图 3c), 并且信噪比也大大提高, 所以分离富集 LMW RNA 这一步在我们的 miRNA 芯片实验中是必需的, 我们后面的实验都用 LMW RNA 进行标记杂交.



**Fig. 3 Comparison of the performance of the miRNA microarray with total RNA and LMW RNA**

Total RNA (10  $\mu$ g) from HeLa cells (a) and LMW RNA (b) isolated from the same amount of total RNA were labeled and hybridized to the miRNA microarray, respectively. The signal intensity of hybridization with LMW RNA was much stronger than that with total RNA. The distribution of fluorescence intensity was also shown in (c). The results demonstrated that the step of isolating LMW RNA was necessary in our miRNA microarray experiments.

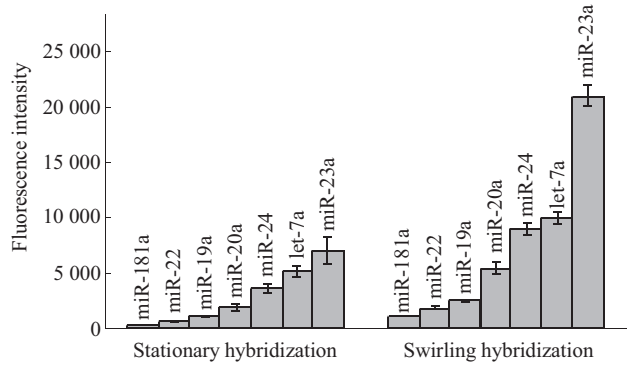
### 2.2 芯片杂交方式的优化

一般芯片的杂交方式是恒温下静置过夜杂交. 在静置的状态下, 盖片下的杂交液是通过自由扩散使探针和靶分子之间相互碰撞实现杂交, 杂交效率较低. 针对这样的问题, 我们改进了芯片的杂交方式, 使用了一种三维旋转摇动的杂交仪 BioMixer™ II, 这样可以促进盖片下的杂交液的流动, 大大提高了芯片的杂交效率. 我们选取了几个 HeLa 细胞中表达丰度不同的 miRNA, 对它们在芯片静置杂交和旋转流动杂交后的实验结果进行了比较(图 4), 可以看出流动杂交的方式大大地提高了芯片的信号强度, 从而提高了芯片检测的灵敏度.

### 2.3 miRNA 芯片的重复性

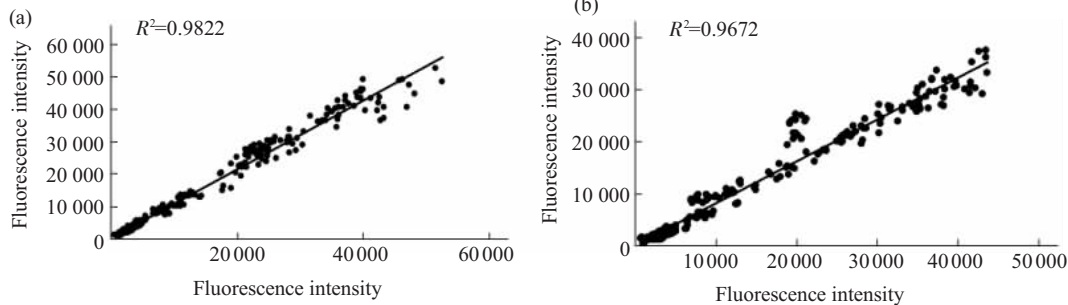
为了检测芯片的重复性, 我们用 HeLa 细胞和脑组织总 RNA 进行了重复分离、标记和杂交试验. 在芯片数据处理时, 先筛选除去弱信号点, 然后作散点图, 作线性回归后计算相关系数  $R^2$ . 实验结果显示, 任意 2 次重复试验的相关系数  $R^2=0.9753\pm$

0.0081, 细胞和脑组织样品各 2 次重复杂交信号散点图如图 5, 可以看出芯片的重复性很好.



**Fig. 4 Optimizing the hybridization method of the miRNA microarray**

The signal intensity of several miRNAs in stationary hybridization and swirling hybridization was compared. The signal intensity of swirling hybridization was much stronger than that of stationary hybridization. By using this method, the performance of the microRNA microarray was greatly improved.



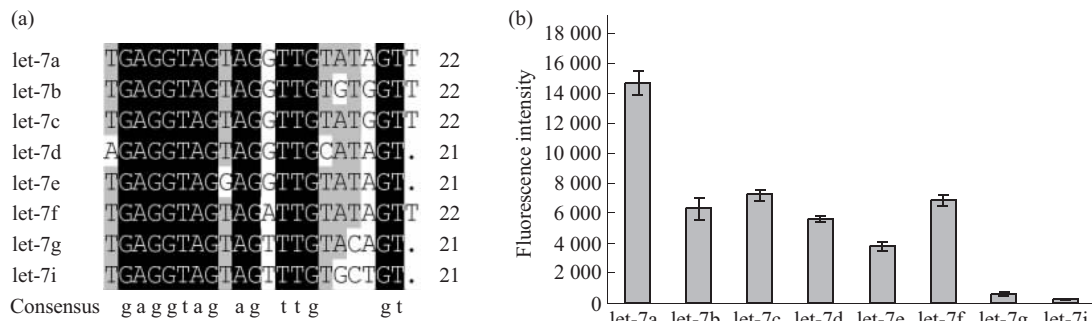
**Fig. 5 Reproducibility of the miRNA microarray**

(a) Duplicate hybridizations of 20  $\mu$ g HeLa total RNA. (b) Duplicate hybridizations of 20  $\mu$ g brain total RNA.

### 2.4 芯片的特异性和灵敏度

由于许多 miRNA 序列具有很高的同源性, 用芯片检测 miRNA 时, 检测的特异性问题是一个难

点. 为了观察芯片检测的特异性, 我们选择了 Let-7 家族 miRNA 进行实验, Let-7 家族有 8 个成员, 序列具有高度的同源性(图 6a). 我们用 Ambion 公司



**Fig. 6 Specificity of the miRNA microarray**

The microarray was probed with 1.4 nmol/L Let-7a transcript and detection was performed as described in **Materials and methods**. Sequences of Let-7 family miRNAs are seen in (a). Fluorescence intensity of different Let-7 are plotted in (b).

的 miRNA Probe Construction kit 通过体外转录制备了 Let-7a 对应的 miRNA, 然后用 1.4 nmol/L 的浓度与芯片进行杂交, 结果如图 6b. 序列分析显示, Let-7a 和 Let-7c、Let-7f 之间都只相差一个碱基, 杂交结果显示, Let-7c 和 Let-7f 这 2 个探针与 Let-7a 的交叉反应信号, 在所有除 Let-7a 自身之外的 Let-7 家族中信号强度最高, 但序列完全匹配 Let-7a 探针的信号强度仍然是 Let-7c、Let-7f 这 2 个探针信号强度的 2.1 倍以上, 以上结果说明芯片可以区分到靠近探针中间位置一个碱基的差异.

我们用制备不同浓度的 Let-7a 对芯片的灵敏度进行了检测, 在低至 70 pmol/L 时, 芯片可以检测出荧光信号, 检测限达 1 fmol, 从 70 pmol/L 到 3500 pmol/L 芯片具有很好的线性(图 7a). 由于 miRNA 在细胞或组织总 RNA 中所占的比例非常小, 为了检测我们芯片完成一次分析需要的最低总

RNA 的量, 我们用 5、10、20 和 40  $\mu\text{g}$  HeLa 细胞总 RNA 分别进行 LMW RNA 分离、标记和杂交, 结果的信号分布如图 7b. 我们可以看出, 对于高表达的 miRNA, 5  $\mu\text{g}$  总 RNA 就可以得到不错的信号. 但对于低表达的 miRNA, 5  $\mu\text{g}$  总 RNA 杂交得到的信号较弱或没信号, 而用 20  $\mu\text{g}$  总 RNA 杂交时整体信号都不错. 这和 Castoldi 等<sup>[24]</sup>用丹麦 Exiqon 公司的 miRNA 芯片得到的结论类似, 即在 RNA 量少的情况下, 对于高表达的 miRNA 检测还是有效的, 但会遗漏一些低表达 miRNA 的信息. 所以我认为, 在 RNA 来源量比较充足的情况下, 用 20  $\mu\text{g}$  进行杂交比较合适, 除非特别指明之处, 本实验都用 20  $\mu\text{g}$  总 RNA 进行杂交(由于组织中 miRNA 的表达量比细胞高, 组织总 RNA 的用量可以比细胞略少).

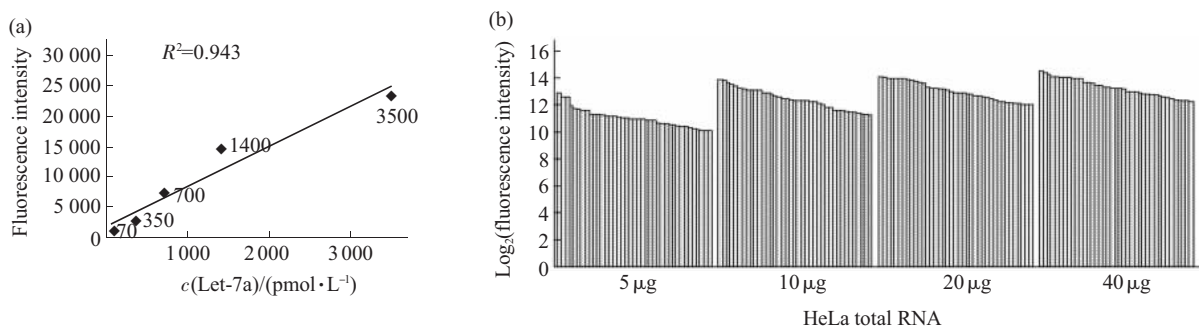


Fig. 7 Sensitivity of the miRNA microarray

(a) The microarrays were hybridized with different concentrations of Let-7a transcript and detection was performed as described in **Materials and methods**. Plots of fluorescence intensity  $\log_2$  with calculated linear regression line was shown. (b) Different quantities of HeLa total RNA were hybridized to the miRNA microarrays. The distribution of fluorescence intensity was shown.

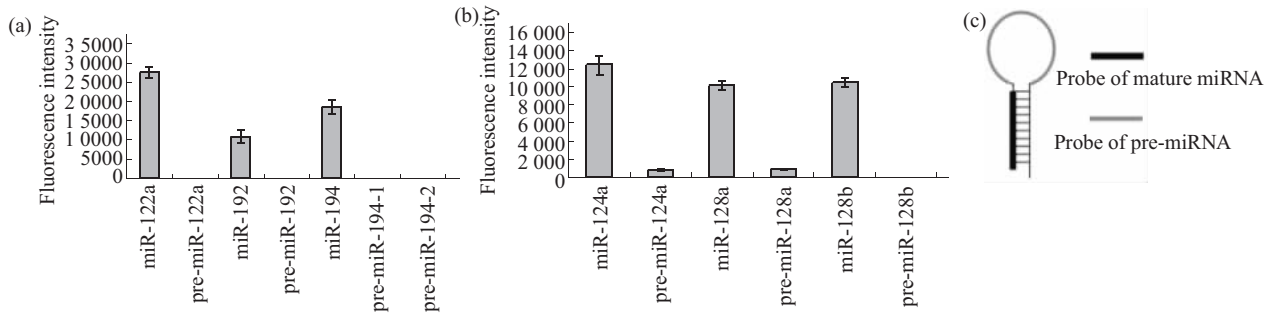
## 2.5 芯片探针对 miRNA 成熟体和 miRNA 前体的区分

我们设计的探针针对的是 miRNA 的成熟体序列, 但是 miRNA 的成熟体序列是被包含在 miRNA 前体序列中的. 为了确定探针检测的是成熟体还是前体, 我们针对肝脏高表达的 miR-122a、miR-192 和 miR-194 以及脑高表达的 miR-124a 和 miR-128a/b 等 miRNA 又设计了只针对它们前体中特有序列的探针, 探针设计的位置见图 8c. 然后用肝脏组织和脑组织 RNA 进行杂交试验, 结果如图 8a 和 8b, 成熟体的探针信号都很强, 而针对前体的几个探针没信号或信号较弱. 这说明芯片检测的

信号主要来自于 miRNA 成熟体, 而 miRNA 前体基本不被我们的芯片检测出来.

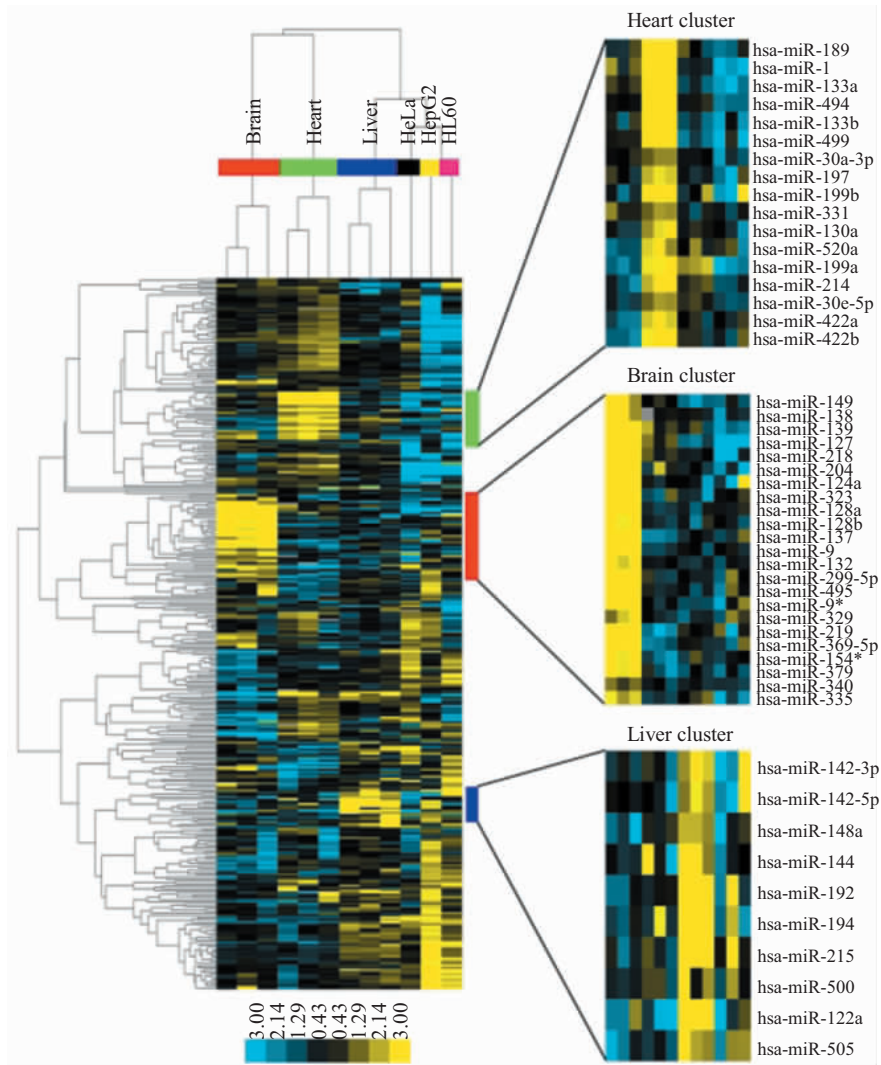
## 2.6 不同的细胞和组织的 miRNA 表达谱及聚类分析

我们利用这个检测 miRNA 表达的微阵列芯片对人不同细胞株和组织的 miRNA 表达谱进行了分析, 结果显示, 不同的细胞或组织都有独特的 miRNA 表达谱, 各种细胞或组织都有高表达或特异表达的 miRNA. 我们用非监督等级聚类 (unsupervised hierarchical clustering) 可以将不同个体的同一组织聚类到一起, 而将不同组织区分开 (图 9).



**Fig. 8 Validating the signal derived from mature miRNAs on the miRNA microarray**

We designed mature miRNA and pre-miRNA probes of miR-122a, 192, 194, 124a, 128a/b which express highly in liver or brain. Liver RNA (a) and brain RNA (b) was hybridized to the miRNA microarray respectively. Fluorescence signal of pre-miRNAs was absent or weak. The locations of probes of mature miRNA and pre-miRNA were shown in (c).



**Fig. 9 Clustering analysis of miRNAs profiles of different human tissues and cell lines**

The raw data which was normalized with mean of per-array was  $\log_2$  transformed and median centered by arrays and genes using CLUSTER 3.0 software, then further analyzed with hierarchical clustering with average linkage. Yellow denotes high expression and blue denotes low expression relative to the median.



### 2.7 组织特异表达的 miRNA 的 RT-PCR 验证

我们采用 RT-PCR 的方法检测了 miR-122a、miR-9 和 miR-124a 在细胞株、肝脏和脑中的表达, 采用内标 U6 进行归一化. PCR 的结果见图 10,

miR-122a 特异地在肝组织中高表达, miR-9 和 miR-124a 特异地在脑组织中高表达, 与芯片结果一致.

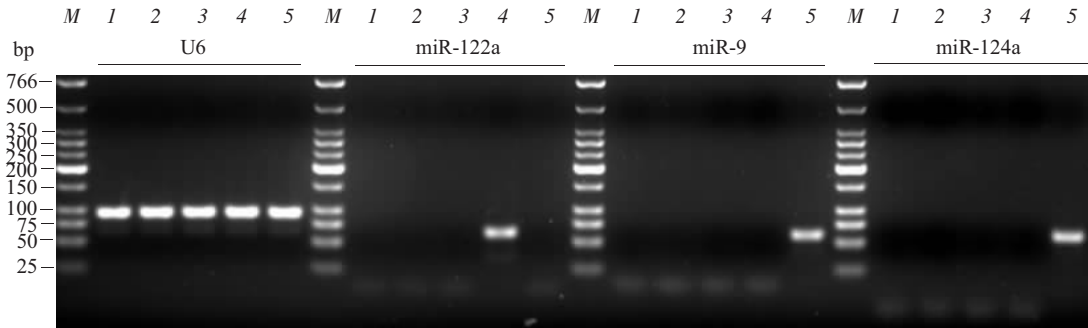


Fig. 10 RT-PCR of miRNAs

We detected miR-122a, miR-9 and miR-124a in different cell lines and tissues with RT-PCR. U6 RNA is the internal standard. The length of miRNAs PCR products is 59 bp and that of U6 is 94 bp. 1: HeLa. 2: HepG2. 3: HL60. 4: Liver. 5: Brain. M: Low molecular weight DNA ladder.

## 3 讨 论

通过本研究, 我们成功地研制出了一种用于检测 miRNA 表达的微阵列芯片. 与国外文献报道的同类芯片比较<sup>[13,17~20]</sup>, 我们的芯片涵盖了 miRNA Registry Release 7.0 数据库中所有的人、小鼠和大鼠的 miRNA, 另外我们还加进来了 122 条预测的 miRNA, 所以芯片结果可以提供更多的 miRNA 的表达信息. 并且随着 miRBase 数据库的更新, 我们的芯片探针也会随着更新. 由于生物芯片技术是一门多学科交叉的技术, 其他相关学科的发展必将带动它的发展. 在本研究中我们就采用自己设计的能在 X、Y 和 Z 轴上进行三维旋转的杂交仪优化了 miRNA 芯片的杂交方式, 把芯片在水浴中静置杂交改为三维旋转的流动杂交方式, 改善了杂交液的流动性, 大大提高了芯片的杂交效率. 这个方法也可以应用到其他中高密度芯片的杂交中.

成熟的 miRNA 由于短小, 并且没有 mRNA 的 poly(A)尾巴, 不能用 Oligo d(T)逆转录来进行特异性标记, 所以 miRNA 的标记比较困难. 现在采用的标记方法有: 在 miRNA 两端加接头, 然后进行 PCR 扩增和标记<sup>[14,18]</sup>; 用标记的随机引物逆转录进行标记<sup>[15]</sup>, 或者用随机引物逆转录后用末端转移酶 (TdT)进行标记<sup>[19]</sup>; 另外有一类方法就是直接标记 RNA, 有采用 Poly(A) RNA 聚合酶进行 3' 端标记<sup>[16]</sup>, 也有采用 T4 RNA 连接酶为 miRNA 端连接上一段荧光修饰的 RNA 进行标记<sup>[17]</sup>. 直接标记

RNA 方法的优点是中间酶学反应步骤少, 没有引进扩增步骤, 避免了这些步骤可能造成的偏差. 本文中我们采用了 T4 RNA 连接酶在 miRNA 的 3' 端连接上一段 CU 的二核苷酸 RNA 分子, CU 二核苷酸的 5' 端进行磷酸化修饰, 3' 端进行荧光 Cy3 修饰. 这种标记方法具有很高的特异性, 只能对 3' 端的 OH 进行标记, 操作简便、快捷, 并且中间的纯化步骤少而简单, 降低了中间纯化步骤可能造成的 miRNA 丢失.

是否从总 RNA 中分离出 LMW RNA 来进行 miRNA 芯片检测, 在文献中报道不一, 这也和采用的标记方法相关. Liu 等<sup>[19]</sup>采用总 RNA 直接进行标记杂交; Thomson 等<sup>[17]</sup>的研究报道也认为 LMW RNA 的分离并不十分必要. 我们的实验结果显示, 用分离后的 LMW RNA 标记杂交的信噪比大大优于用总 RNA 直接标记杂交. 分析其可能的原因有: 总 RNA 中其他丰度高的 RNA 在连接步骤中可能影响 miRNA 与 CU 二核苷酸连接效率; 由于 miRNA 在序列上和多个高分子质量的 mRNA 具有很高的相似性, 所以芯片上的探针容易与总 RNA 中的 mRNA 发生交叉杂交反应导致非特异性杂交, 降低了芯片杂交信号的信噪比. 因此我们认为把 LMW RNA 分子从总 RNA 分离出来再进行标记有利于提高 miRNA 检测的灵敏度和特异性.

我们在用 Let-7 家族进行的特异性实验中可以发现, 在用浓度 1 400 pmol/L 的 Let-7a 进行杂交



时, Let-7a 的探针信号强度都在 Let-7 家族其他探针杂交信号的 2.1 倍以上, 能够区分到中间一个碱基的差异. 但对于高丰度表达的 miRNA, 与它同源性高的探针的非特异信号即使没有完全匹配的探针信号强, 也可能通过芯片数据分析中筛除弱信号的阈值值而进入到后续的分析. 但序列高度同源的 miRNA 往往调控的是相同的靶 mRNA, 已有文献报道针对芯片检出序列高相似性的 miRNA, 把它们作为功能相似的一个 miRNA 簇看待<sup>[25]</sup>. 所以这种非特异反应不会影响我们的芯片结果对 miRNA 功能的分析.

我们用 PEG 6000 分离的低分子质量 RNA 与分别针对 miRNA 前体和成熟体的探针进行杂交, 结果显示, 针对前体的探针信号较弱或无信号, 说明芯片的信号主要来自 miRNA 成熟体. Barad 等<sup>[26]</sup>也在芯片上设计了前体探针和成熟体探针来进行试验, 最后的结果也显示芯片探针的信号都基本来自 miRNA 成熟体. 分析其原因可能是: miRNA 前体的发夹结构不利于和探针杂交, 另外推测 miRNA 前体在细胞中存在的时间短暂而不易被检测到.

miRNA 的组织特异表达的模式已在其他研究中报道, 我们用我们的芯片分析了 HeLa、HepG2 和 HL60 等细胞以及肝、脑和心脏等组织的 miRNA 表达谱. 实验结果和其他报道一致<sup>[13,17~20]</sup>: 肝特异表达的有 miR-122a/b、miR-148 和 miR-194 等; 脑特异表达的有 miR-9、miR-124a 等; 心脏特异表达的有 miR-1 等. 我们对各种细胞和组织的 miRNA 表达谱进行了聚类分析, 结果显示, 来源于不同个体的同一组织可以聚类到一起, 这样可以利用 miRNA 表达谱把不同来源的组织区分开. miRNA 这种表达的空间特异性与其在基因组上位置相关, 位置相近的有可能在转录时共转录而具有相同的表达模式, 也可能与其在特定部位的特定生理功能密切相关, 这需要对其功能进行深入研究才能够得到理解. RNA 印迹和 RT-PCR 技术虽然是低通量的检测技术, 但芯片检测的结果仍需要它们进行验证. 我们采用 RT-PCR 检测了 miR-122a、miR-9 和 miR-1 在各细胞和组织中的表达, 结果显示 RT-PCR 和芯片的结果完全一致, 有力地验证了芯片的结果. 检测成熟 miRNA 的 PCR 方法是最近才发展起来的方法<sup>[11]</sup>, 它是验证 miRNA 芯片数据或检测单个 miRNA 表达的高灵敏度的方法, 另外也可以采用 SYBR Green I 荧光染料或 TaqMan 探针等方法进行实时荧光定量检测.

综上所述, 我们研制出了一种灵敏度高, 特异性较好的检测 miRNA 表达的微阵列芯片, 它是检测 miRNA 表达的强有力的工具. 为高通量研究不同生理状态下的细胞或组织的 miRNA 表达情况提供了技术手段, 将大大促进 miRNA 的生物学功能研究. 如果将 miRNA 检测芯片与基因表达谱芯片技术相结合<sup>[27]</sup>, 有利于分析阐明 miRNA 作用的靶基因以及 miRNA 在肿瘤等疾病和各种生命活动过程中发挥作用的分子机制.

### 参考文献

- 1 Bartel D P. MicroRNAs: genomics, biogenesis, mechanism, and function. *Cell*, 2004, **116** (2): 281~297
- 2 Chen C Z, Li L, Lodish H F, *et al.* MicroRNAs modulate hematopoietic lineage differentiation. *Science*, 2004, **303** (5654): 83~86
- 3 Calin G A, Dumitru C D, Shimizu M, *et al.* Frequent deletions and down-regulation of micro-RNA genes miR15 and miR16 at 13q14 in chronic lymphocytic leukemia. *Proc Natl Acad Sci USA*, 2002, **99** (24): 15524~15529
- 4 Yang M, Li Y, Padgett R W. MicroRNAs: Small regulators with a big impact. *Cytokine Growth Factor Rev*, 2005, **16** (4~5): 387~393
- 5 Kasashima K, Nakamura Y, Kozu T. Altered expression profiles of microRNAs during TPA-induced differentiation of HL-60 cells. *Biochem Biophys Res Commun*, 2004, **322** (2): 403~410
- 6 Wienholds E, Plasterk R H. MicroRNA function in animal development. *FEBS Lett*, 2005, **579** (26): 5911~5922
- 7 Miska E A. How microRNAs control cell division, differentiation and death. *Curr Opin Genet Dev*, 2005, **15** (5): 563~568
- 8 Xu P, Guo M, Hay B A. MicroRNAs and the regulation of cell death. *Trends Genet*, 2004, **20** (12): 617~624
- 9 Volinia S, Calin G A, Liu C G, *et al.* A microRNA expression signature of human solid tumors defines cancer gene targets. *Proc Natl Acad Sci U S A*, 2006, **103** (7): 2257~2261
- 10 Chen C Z. MicroRNAs as oncogenes and tumor suppressors. *N Engl J Med*, 2005, **353** (17): 1768~1771
- 11 Chen C, Ridzon D A, Broomer A J, *et al.* Real-time quantification of microRNAs by stem-loop RT-PCR. *Nucleic Acids Res*, 2005, **33** (20): e179.
- 12 Krichevsky A M, King K S, Donahue C P, *et al.* A microRNA array reveals extensive regulation of microRNAs during brain development. *RNA*, 2003, **9** (10): 1274~1281
- 13 Liu C G, Calin G A, Meloon B, *et al.* An oligonucleotide microchip for genome-wide microRNA profiling in human and mouse tissues. *Proc Natl Acad Sci USA*, 2004, **101** (26): 9740~9744
- 14 Miska E A, Alvarez-Saavedra E, Townsend M, *et al.* Microarray analysis of microRNA expression in the developing mammalian brain. *Genome Biol*, 2004, **5** (9): R68
- 15 Babak T, Zhang W, Morris Q, *et al.* Probing microRNAs with

- microarrays: tissue specificity and functional inference. *RNA*, 2004, **10** (11): 1813~1819
- 16 Shingara J, Keiger K, Shelton J, *et al.* An optimized isolation and labeling platform for accurate microRNA expression profiling. *RNA*, 2005, **11** (9): 1461~1470
- 17 Thomson J M, Parker J, Perou C, *et al.* A custom microarray platform for analysis of microRNA gene expression. *Nat Methods*, 2004, **1** (1): 47~53
- 18 Baskerville S, Bartel D P. Microarray profiling of microRNAs reveals frequent coexpression with neighboring miRNAs and host genes. *RNA*, 2005, **11** (3): 241~247
- 19 Sun Y, Koo S, White N, *et al.* Development of a micro-array to detect human and mouse microRNAs and characterization of expression in human organs. *Nucleic Acids Res*, 2004, **32** (22): 2486~2494
- 20 Lu J, Getz G, Miska E A, *et al.* MicroRNA expression profiles classify human cancers. *Nature*, 2005, **435** (7043): 834~838
- 21 Xie X, Lu J, Kulbokas E J, *et al.* Systematic discovery of regulatory motifs in human promoters and 3' UTRs by comparison of several mammals. *Nature*, 2005, **434** (7031): 338~345
- 22 Eisen M B, Spellman P T, Brown P O, *et al.* Cluster analysis and display of genome-wide expression patterns. *Proc Natl Acad Sci USA*, 1998, **95** (25): 14863~14868
- 23 Jiang J, Lee E J, Gusev Y, *et al.* Real-time expression profiling of microRNA precursors in human cancer cell lines. *Nucleic Acids Res*, 2005, **33** (17): 5394~5403
- 24 Castoldi M, Schmidt S, Benes V, *et al.* A sensitive array for microRNA expression profiling (miChip) based on locked nucleic acids (LNA). *RNA*, 2006, **12** (5): 913~920
- 25 Chen J F, Mandel E M, Thomson J M, *et al.* The role of microRNA-1 and microRNA-133 in skeletal muscle proliferation and differentiation. *Nat Genet*, 2006, **38** (2): 228~233
- 26 Barad O, Meiri E, Avniel A, *et al.* MicroRNA expression detected by oligonucleotide microarrays: system establishment and expression profiling in human tissues. *Genome Res*, 2004, **14** (12): 2486~2494
- 27 Lim L P, Lau N C, Garrett-Engle P, *et al.* Microarray analysis shows that some microRNAs downregulate large numbers of target mRNAs. *Nature*, 2005, **433** (7027): 769~773

## Construction and Application of a Microarray for Profiling MicroRNA Expression\*

LUO Ming-Yong<sup>1,2</sup>, TIAN Zhi-Gang<sup>6</sup>, XU Zhi<sup>6</sup>,  
ZHANG Liang<sup>1,3</sup>\*\*, WANG Ying-Xiong<sup>2</sup>\*\*, CHENG Jing<sup>1,3,4,5</sup>

<sup>1</sup>National Engineering Research Center for Beijing Biochip Technology, Beijing 102206, China;

<sup>2</sup>School of Public Health, Chongqing University of Medical Science, Chongqing 400016, China;

<sup>3</sup>CapitalBio Corporation, Beijing 102206, China;

<sup>4</sup>Medical Systems Biology Research Center, Tsinghua University, Beijing 100084, China;

<sup>5</sup>Department of Biological Sciences and Biotechnology, Tsinghua University, Beijing 100084, China;

<sup>6</sup>Peking University Third Hospital, Beijing 100083, China)

**Abstract** MicroRNAs (miRNAs) are a class of endogenous non-coding RNA, which regulate target gene expression via mRNA degradation and translational repression. MicroRNA has been linked to the development of plants and animals, to cell growth and apoptosis, to fat metabolism and other significant physiological processes. Misregulation of microRNA function might contribute to human diseases. Profiling microRNA expression will facilitate the study of biological functions of microRNAs. To detect microRNA expression, a high-density oligonucleotide microarray was constructed, which contains oligonucleotides corresponding to 313 human microRNAs, 261 mouse microRNAs, 196 rat microRNAs and 122 predicted microRNAs. Firstly, high molecular weight RNA was removed from total RNA by polyethylene glycol (PEG) precipitation, and low molecular weight RNA (LMW RNA) was obtained. Then LMW RNA was directly labeled based on a published method that uses T4 RNA ligase to couple LMW RNA to a fluorescence modified dinucleotide. By using a special spotting solution in combination with a swirling hybridization method, the performance of the microRNA microarray was greatly improved. The results of the microRNA microarray experiments indicated its good reproducibility, sensitivity and specificity. It was also showed that the signal of the microRNA microarray was derived from mature microRNAs, but not from their precursors. Using the microarray, microRNA expressions in human brain, heart, liver and in HeLa, HepG2, HL60 cells were profiled. The results revealed a good overall concordance with previously published data. Clustering analysis showed the tissue specificity pattern of microRNA expression, and the same

tissue from different individuals was clustered together. Then the expression of miR-122a, miR-9 and miR-124a in human tissues and cell lines with RT-PCR (reverse transcription-PCR) were validated. The results which showed the liver-specific expression of miR-122a and the brain-specific expression of miR-9 and miR-124a were consistent with the microarray results.

**Key words** microRNA, microarray, clustering, RT-PCR

---

\*This work was supported by a grant from The National Tackle Key Projects (2005BA711A05-04).

\*\*Corresponding author .

ZHANG Liang. Tel: 86-10-80726787, E-mail: lzhang@capitalbio.com

WANG Ying-Xiong. Tel: 86-23-68485001, E-mail: wyx61221@yahoo.com.cn

Received: June 4, 2006 Accepted: August 26, 2006