

•综述•

DOI: 10.13865/j.cnki.cjbmb.2015.02.02

miRNA 靶基因的筛选方法及其相关网络资源

刘卓琦, 万福生, 罗达亚*

(南昌大学基础医学院生物化学与分子生物学教研室, 南昌 330006)

摘要 微小 RNA (microRNA, miRNA) 是一类广泛存在于动植物中, 大小约 22 nt 的单链非编码小分子 RNA。它通过与靶 mRNA 3' 末端非翻译区 (untranslational region, UTR) 结合, 使 mRNA 降解或翻译抑制。大量的研究表明, miRNA 在动植物的生长发育、细胞分化、细胞增殖与凋亡、肿瘤发生等诸多环节发挥着重要的调节作用。目前, miRNA 靶基因的筛选方法主要有生物信息学筛选和实验方法两大类, 且在互联网上拥有大量相关的网络共享资源。本文通过对现有 miRNA 靶基因的筛选方法及其相关网络资源进行整理与比较分析, 以有效指导并辅助 miRNA 领域的研究。

关键词 microRNA; 靶基因; 网络资源

中图分类号 R34

Methods and Their Online Resources for miRNA Targets Screening

LIU Zhuo-Qi, WAN Fu-Sheng, LUO Da-Ya*

(Department of Biochemistry and Molecular Biology, School of Basic Medical Sciences, Nanchang 330006, China)

Abstract microRNAs (miRNAs) are short non-coding RNAs of approximately 22 nt that exist in plants and animals to regulate gene expression through binding to the 3' untranslational region (UTR) of their target mRNA for cleavage or translational repression. miRNAs are involved in various regulatory pathways for development control, cell proliferation, apoptosis and cell differentiation. They are also involved in the occurrence and development of many diseases, including a wide spectrum of cancers. In the present, the screening for miRNA target could be achieved through rational bioinformatic prediction and biological validation experiment, and the results are shared across online databases. We summarized and comparatively analyzed the popular research methods for miRNA target gene screening, to provide an effective guidance and assistance in the field of miRNA research.

Key words microRNA; target gene; online resource

微小 RNA (microRNA, miRNA) 是在多种真核细胞和病毒中发现的一类长度 21 ~ 23 nt 的内源性非编码单链小分子 RNA。它的序列高度保守, 通过在胞质中与 Argonaute 家族蛋白结合形成 miRNA 诱导的沉默复合物 (miRISC) 与靶 mRNA 上 miRNA 识别元件 (miRNA-recognition elements, mREs) 完全或部分互补的结合, 介导转录后水平的基因沉默^[1,2], 包括 mRNA 降解和翻译的抑制。研究表明, miRNA 参与了人类机体多种生物学过程的调控, 包括组织器官发育、细胞凋亡、增殖与分化和免疫调节等, 并且与许多疾病如肿瘤、心血管疾病等的发生、发展及预后密切相关。

miRNA 功能研究的重点在于对其靶基因调控的研究。通过实验结合生物信息学方法, 2005 年

Lewis 等^[3] 预测, miRNA 作用的靶基因超过 5 300 个, 约占整个人类基因总数的 1/3; 2009 年 Friedman 等^[4] 更新了预测成果, 认为 miRNA 在 mRNA 3'UTR

收稿日期: 2014-09-22; 接受日期: 2014-11-20

国家自然科学基金项目 (No. 81160248, No. 81360313); 南昌大学科研基金 (No. JC200835); 南昌大学创新学分子训练项目 (No. 14001814)

* 联系人 Tel: 13576057302; E-mail: luodaya@hotmail.com

Received: September 22, 2014; Accepted: November 20, 2014

Supported by National Natural Science Foundation of China (No. 81160248, No. 81360313); Nanchang University Scientific Research Funds (No. JC200835); Nanchang University Innovation Credits and Research Training Projects (No. 14001814)

* Corresponding author Tel: 13576057302;

E-mail: luodaya@hotmail.com

区的结合位点超过 45 000 个 ,调控了人类 2/3 以上的蛋白质编码基因。如今 ,更多的科学家相信 ,人类几乎所有的编码基因都受到了 miRNA 的调控 ,并且每个 miRNA 可调控数以千计的靶基因。

miRNA 靶基因的筛选方法 ,主要有生物信息学方法和实验方法两大类。近十多年来 ,随着以计算机为工具的生物信息技术的飞速发展 ,在大量的数学基础理论之上 ,创造出了一系列针对 miRNA 靶基因筛选的生物信息类工具 ,为 miRNA 的靶基因研究及功能挖掘提供了良好支撑^[5-6]。与此同时 ,实验方法亦得到了长足的发展 ,最初此类方法仅用于筛选出的个别靶基因的验证 ,而伴随着高通量研究技术在此领域的应用 ,实验方法也可用于靶基因的筛选^[7-9] ,与生物信息学方法相辅相承、相互验证。本文对这两类研究方法及现有的网络共享资源进行了

整理和综述 ,旨在帮助科研工作者更好地共享这些资源 ,提高科研效率。

1 生物信息学方法

生物信息学方法在 miRNA 靶基因筛选领域应用十分广泛 ,设计者们以 miRNA: mRNA 相互作用特点为基础 ,从不同角度 ,结合不同算法设计软件。现有的筛选软件已达几十种 ,并且随着生物信息学的发展 ,不同机构、不同作者采用不同理念设计的筛选软件仍层出不穷 ,在 miRNA 的机制研究中扮演着重要的角色。

根据不同工具的基本原理将生物信息学方法分为 3 大类: 基于序列互补的筛选方法、与统计学原理结合的机器学习类筛选方法及两类方法的联合 ,常见的筛选工具见 Table 1。

Table 1 Online resources for miRNA targets prediction

| Category | Method | Establish | Latest update | Statistical techniques | Advantages | Website |
|-----------------------------|------------------|-----------|---------------|------------------------|--|--|
| Filtering approaches | miRANDA | 2003 | 8/2010 | NA | Not Stringent seed matching; Multiple target sites. | http://www.microRNA.org |
| | TargetScan | 2005 | 6/2012 | NA | Perfect seed complementarity; High interpretability. | http://www.targetscan.org |
| | PITA | 2007 | 2008 | NA | Allowance of a single G-U wobble or mismatch , minimum seed conservation , and flanking setting. | http://genie.weizmann.ac.il/pubs/mir07 |
| | miRWalk | 2011 | 3/2013 | NA | Seed matching; Predicted miRNA binding sites on the complete sequence of all known genes; presenting experimentally validated miRNA interaction. | http://www.umm.uni-heidelberg.de/apps/zmf/mirwalk |
| Machine Learning approaches | PicTar | 2005 | 11/2011 | HMM | Extracted alignment from UCSC dataset; Combinatorial microRNA target predictions. | http://pictar.mdc-berlin.de |
| | miRTarget 2 | 2007 | 4/2012 | SVM | Relied on the Linsley microarray transcriptional dataset; Conjunction with miRDB. | http://mirdb.org |
| | DIANA-microT-CDS | 2009 | 7/2012 | GLM | Based on PAR-CLIP data; miRNA binding location in both 3' UTR and CDS; Multiple target sites. | http://www.microRNA.gr/microT-CDS |
| | TargetSpy | 2009 | 2009 | MultiBoost | Without the usage of seed matching attributes and conservation filtering; Capable of predicting species-specific targets. | http://www.targetspy.org |
| | TargetMiner | 2008 | 5//2012 | SVM | Advanced negative set selection. | http://www.isical.ac.in/~bioinfo_miu/targetminer20.htm |
| Hybrid approaches | MultiMiTar | 2011 | 5/2012 | AMOS-A-SVM | Tackling the limitations of SVM; High prediction performance. | www.isical.ac.in/~bioinfo_miu/multimitar-download.htm |
| | CoMiR | 2012 | 12/2013 | SVM | Including both filtering and machine learning steps. | http://www.benoslab.pitt.edu/comir |

1.1 基于序列互补的筛选方法

基于序列互补的筛选方法(filtering approaches) 是根据 miRNA: mRNA 相互作用的三大特点(种子区配对、序列保守性及 miRNA: mRNA 结合的热稳定性) 来推断可能的靶点。

miRanda^[10] 创建于 2003 年, 是出现较早并且保持持续更新的一款筛选工具。该工具对靶基因的筛选过程分为 3 步: (1) 3'UTR 区与输入的 miRNA 全序列进行配对, 其中对种子区的配对赋值最高, 允许 G: U 错配等非典型配对方式的存在; (2) 计算 miRNA: mRNA 结合的自由能, 并设定阈值; (3) 对 miRNA: mRNA 相互作用的保守性进行最后的筛选。经 3 步筛选后得到的候选靶基因都有一个得分值, 与 miRNA 种子序列匹配度越高的靶基因得分越高, 当 mRNA 上有多个 miRNA 结合位点时也能获得较高的得分。

TargetScan^[11] 出自 David P Bartel 所领导的团队, 更新维护及时, 操作界面简洁。用户可通过输入 miRNA 名称查找其对应的靶基因或者输入基因名称查找对应的 miRNA; 也可根据物种间 miRNA 家族的保守性进行 miRNA 直接选择。筛选步骤依次为“种子区配对——保守性——热动力学”, 该工具强调 miRNA 种子区与 mRNA 3'UTR 区的严格配对。

PITA^[12] 在设计上涵盖了 miRNA: mRNA 结合的基本特点, 同时也考虑了多位点结合等现象。它的与众不同在于以“miRNA 与靶位点的可接近程度”作为主要参数, 除了评估 miRNA: mRNA 结合的热力学稳定性外, 也对 mRNA 的 3'UTR 区的二级结构进行单独预测, 结构越稳定, miRNA 与该区域结合的可能越小。PITA 的操作界面非常友好, 用户可以对种子区最小配对个数、是否允许 G: U 配对、是否允许错配等参数进行设置。此外, 用户可直接输入 UTR 区序列进行筛选, 突破了 3'UTR 区的限制。不足之处在于, PITA 的更新和维护明显不及 miRanda 和 TargetScan。

miRWalk^[13] 数据库构建分成两大模块: 筛选靶点数据库和验证靶点数据库。miRWalk 允许 miRNA 种子区与 mRNA 基因全长进行配对, 包括所有的转录子和线粒体基因。输入项可以是 miRNA 亦可以是基因名称, 用户可在操作界面进行具体的参数设置, 包括配对区域(启动子、5'UTR、CDS、3'UTR 可进行单选和组合)、最小配对种子数、转录子范围、P 值。另外, miRWalk 网站还同时组合其它常用的 8 种筛选工具结果, 给出参考结果。

1.2 机器学习类筛选方法

与统计学原理相结合的计算机学习方法(machine learning approaches) 一直以来都被认为是生物信息学发展的方向, 但是在 miRNA 靶基因筛选领域, 机器学习类方法表现并不突出。miRNA 研究早期, 确定的靶点数据量非常少, 验证靶点数量上达不到计算机学习的要求。随着高通量实验方法的展开和验证靶点数据库的逐步建立, 计算机以 mREs 及非 mREs 数据分别作为正负学习对象, 采用不同的算法, 结合热动力学、miRNA 二级结构和 mREs 位置等信息, 对 miRNA: mRNA 相互作用特点进行分类, 以构建靶点筛选模型。其中常用的计算机分类算法有隐马尔可夫模型(hidden Markov models, HMM)、朴素贝叶斯分类(Naive Bayes, NB) 和支持向量机(support vector machines, SVM) 等方法, 还有一些来自作者的建模和创造。网络上有迹可寻的此类工具非常多, 良莠不齐。

PicTar 创建人 Anders 等^[14] 认为, 筛选工具仅对单个 miRNA 的靶点进行筛选是不准确的。在特定的组织或生物个体发育的某一阶段, 总有小部分(2 ~ 10) miRNA 是共同表达的, 并且 miRNA 与 mRNA 的结合也是多位点的, 不同 miRNA 之间相互影响共同调节 mRNA 的功能。为此, PicTar 以 8 大脊椎动物 3'UTR 区序列为学习对象, 在算法上强调种子区与 3'UTR 区的严格配对, 参考物种间的保守性和热动力学特点, 以 HMM 算法为基础, 计算了多个 miRNAs 同时与 mRNA 结合的最大概率。PicTar 值越大, 表示越可能是其真正靶点。

miRTarget2^[15] 是依据 Linsley 等^[16] 芯片实验数据作为学习对象, 以 SVM 为基础的机器学习类筛选方法, 该方法以 Target Score 值对筛选的靶基因进行评分, 针对 UTR 区上多位点结合情况, 先分别计算 miRNA 与单一位点结合的 P 值再进行组合, 得分越高, 真实靶点的可能性越大。miRTarget2 的筛选结果都保存在 miRDB^[17] 数据库中, 该数据库还包含 miRNA 功能注释部分, 可以对数据进行更进一步挖掘。

DIANA-microT-CDS^[18, 19], 在分析了大量的 PAR-CLIP 实验数据的基础上, 将正确 mREs 和错误 mREs 数据作为计算机学习的对象, 采用了交互验证准确率最高的广义线性模型(generalized linear model, GLM), 对 CDS 和 3'UTR 区域的 mREs 进行综合评分, 得到 miTG 值, 范围介于 0 ~ 1 间, 数值越接近 1, 认为是真正靶点的可能性越大。该工具既可

以筛选单个 miRNA 的靶基因,也可同时筛选多个 miRNA 的靶基因.与单一的 3'UTR 模型相比,灵敏度由 52% 提升到 65%.该工具的结果输出界面同时还有实验验证结果以及来自 miRANDA 和 TargetScan 的筛选结果作为参考.

此外,网络环境较好的机器学习类的筛选工具还有 TargetSpy^[20]、TargetMiner^[21]和 MultiMiTar^[22].

1.3 联合筛选法

将以上两种方法相结合,成为第三类筛选工具,称之为联合筛选工具(Hybrid approaches).

ComiR(combinatorial miRNA targeting)^[23]是两者结合的典范. ComiR 可对一个或一组 miRNA 进行靶点筛选,输入项可以是 miRNA 芯片数据及 miRNA-Seq reads. 该工具组合了 4 种筛选工具,分别为 miRanda、PITA、TargetScan 和 mirSVR. 在每种工具得到一个筛选值后,以 SVM 分类(以 *D. melanogaster* AGO1 IP data set^[24]为学习参照)方法得到 ComiR 值.

2 实验方法

生物学实验方法既可以筛选也可以验证 miRNA 的靶基因.依据结果的可信程度,将这些实验方法分为两类,其中准确性高的实验方法以 miRNA: mRNA 的直接作用为基础,而准确率较低实验方法则是间接检测了 miRNA: mRNA 相互作用.

2.1 准确率高的实验方法

准确率高的实验方法主要有荧光素酶报告基因、RNA 与蛋白质免疫共沉淀技术等.

荧光素酶报告基因方法的基本原理是首先构建荧光素酶表达载体,将希望鉴定的 miRNA 靶基因的 3'UTR 构建到荧光素酶基因的 3'UTR 中;再将构建好的载体转染细胞并改变细胞中相应 miRNA 的表达水平;最后检测荧光素酶的表达情况,以分析转染 3'UTR 中是否含有 miRNA 的靶位点.此方法可以直接鉴定出靶位点,准确率高,是最常用的 miRNA 靶基因的鉴定方法^[25].

RNA 结合蛋白免疫沉淀技术(RNA binding protein immunoprecipitation, RIP),运用针对目标蛋白的抗体把相应的 RNA-蛋白复合物沉淀下来,然后经过分离纯化对结合在复合物上的 RNA 进行分析.2007 年,Beitzinger 等^[26]利用 AGO 蛋白家族既能结合 miRNA 又能结合 mRNA 的特性,分别使用 AGO-1 和 AGO-2 蛋白的单克隆抗体在人类细胞中进行免疫共沉淀,得到与 AGO-1 和 AGO-2 蛋白结合的 mRNA 各 600 余条,然后通过克隆测序对这些

mRNA 进行鉴定.从与 AGO-1 蛋白结合的 mRNA 中随机挑选 6 条进行荧光素酶报告基因检测,发现其中有 5 条都是 miRNA 的靶基因.

随着实验方法的优化,RIP 技术与生物芯片技术及第二代测序技术的结合,衍生出 RIP-Chip、CLIP-seq、PAR-CLIP 等技术,大大提高了 miRNA 靶基因筛选的准确率. CLIP-seq,又称为 HITS-CLIP^[27],即紫外交联免疫沉淀结合高通量测序(crosslinking-immunoprecipitation and high-throughput sequencing).它是将紫外交联免疫共沉淀技术与高通量测序技术相结合的研究系统,可在全基因组水平鉴定与特定 RNA 结合蛋白相互作用的 RNA 分子,适用于对 miRNA 靶基因进行筛选.活细胞内成熟 miRNA 在胞质中与 AGO 家族蛋白结合形成 miRISC,可在紫外照射下与 mRNA 发生耦联,再用 AGO 蛋白的特异性抗体将 RNA-蛋白质复合物沉淀之后,回收其中的 RNA 片段,经添加接头、RT-PCR 等步骤,即可对这些分子进行高通量测序分析. CLIP-Seq 方法能够非常明显地降低 miRNA 结合位点的假阳性并且减小 miRNA 结合位点搜寻空间的范围.

此外,通过对经典 CLIP-seq 的实验流程进行局部调整、修改衍生出许多变异的 CLIP-seq 方法,例如增强了交联强度的 PAR-CLIP^[28],提高准确率的针对 AGO1 蛋白的 CLASH^[29]和能够精确鉴定 RNA 分子中蛋白质结合位点的 iCLIP^[30].

2.2 准确率较低实验方法

准确率较低实验方法通过改变 miRNA 的表达水平,分别检测细胞在 mRNA 水平及蛋白水平的变化,提示 miRNA 作用的靶基因.改变细胞中 miRNA 表达水平通常有两种策略:一个是基于载体或病毒的 miRNA 或 miRNA 反义链序列的过表达;另一个是外源 miRNA 双链或反义抑制剂的转染.在建立了 miRNA 的差异表达之后,在 mRNA 水平检测的方法包括生物芯片、高通量测序以及 RT-PCR 方法;在蛋白水平的检测方法有稳定同位素标记蛋白质谱法(stable isotopic labeling with amino acids in cell culture)^[31]、2D-DIGE 及 Western-Blot 等.

2.3 实验验证靶点数据库

随着大量的 miRNA 靶点被实验方法所筛选和鉴定,这些实验结果被搜集、整理并收录到相关的数据库,供研究者查询.例如,DIAN-TarBase^[32]、MiRecords^[33]、miRWalk、miRNAmap^[34]和 StarBase^[35]等,见 Table 2.

Table 2 Online resources for validated miRNA targets

| Method | Website | Data feature |
|---------------|---|---|
| DIANA-TarBase | http://diana.cslab.ece.ntua.gr/tarbase | Including targets for 21 species. Derived from high throughput experiment. Seamlessly interconnected with other DIANA-lab tools , such as DIANA-microT. |
| miRTarBase | http://mirtarbase.mbc.nctu.edu.tw | Indexing 4270 miRNA gene interactions from 14 species. |
| miRecords | http://mirecords.biolead.org | Hosting 2286 miRNA-gene interactions derived from 9 animal species. Also containing predicted targets calculated with 11 different algorithms. |
| miRNAmap | http://miRNAmap.mbc.nctu.edu.tw | Collecting verified miRNA target genes for 11species. Connected with miRanda , RNAhybrid and TargetScan. |
| StarBase | http://starbase.sysu.edu.cn | A facility for the analysis of sequencing based experiments. Hosting a large number of tools which can assist or extend user experience. |

3 两类方法点评

生物信息学筛选靶基因的 3 类方法中 ,机器学习类的筛选方法是以实验验证的靶点数据库为基础的计算机筛选算法 ,和基于序列互补的筛选方法相比更多地依赖设计者个人理念. 鉴于在参照数据库的选择上缺乏公认的“金标准” ,虽说以“事实(数据)”为基础 ,但相对于其它筛选方法却没有太大的优势. 对机器学习类筛选方法的选择 ,研究者更钟爱更新维护好的筛选工具 ,诸如 PicTar、miRDB 和 DIANA-microT-CDS 等. 联合筛选法是前两种方法的联合 ,并不能从根本上提高筛选的准确率. 因此不论是哪类方法 ,靶基因筛选工具都有两个明显的不足: 一是产生大量的结果 ,但其中大部分结果没有被验证 ,有可能是“假阳性”结果; 二是不同工具筛选结果缺乏一致性. 例如 ,采用 miRanda、TargetScan 和 PicTar 对 hsa-miR-205 的靶基因进行了筛选 ,如图 1 结果显示 ,交集基因只占各自筛选结果的很小一部分. 有学者认为 ,这部分基因最有可能是 hsa-miR-205 的靶基因 ,但也有学者持保留意见 ,认为多种组合求交集的方法只是缩小了后续实验的筛选范围 ,但不能否定其它结果的正确. 这样的做法也许遗漏了很多重要的靶点. 本文更倾向于后者观点 ,因为每种筛选工具的设计理念不同: 基于序列互补的筛选方法的差别集中在 miRNA: mRNA 相互作用位置不同 ,以及每种工具对热动力学、序列保守性等参数的权重不同; 机器学习方法除了各自算法差别外 ,选择的学习数据库也有很大差别. 同时 ,miRNA 有翻译抑制和 mRNA 降解两种功能模式 ,但计算机学习类工具常只针对其中一种模式而设计 ,例如 DIANA-

microT-CDS 采用的是 PAP-CLIP 数据库 ,检测的是翻译水平; miRTarget2 采用的是转录组芯片数据库 ,针对的是 mRNA 水平而设计. 因此 ,取交集只是去除了软件设计上的差异 ,并非提高了准确率.

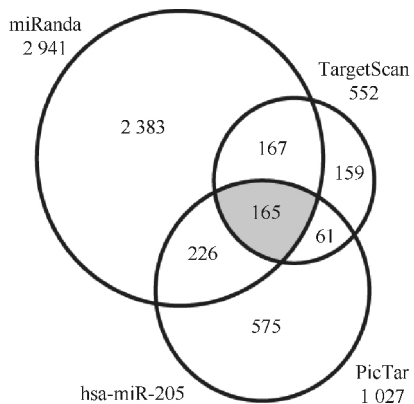


Fig. 1 Computational identification of miR-205 targets
A Venn diagram showing the overlap of predicted targets of hsa-miR-205 between three different online resources

实验方法既可验证靶基因 ,也可筛选靶基因. 除去来自实验室的误差 ,只考虑技术本身 ,在靶基因筛选方面 ,不同的实验方法其准确率亦有高低之分. 准确率高的实验方法检测 miRNA: mRNA 之间的直接作用. 随着实验方法的不断改进 ,以 RIP 为基础的 CLIP-Seq、PAR-CLIP、CLASH 及 iCLIP 方法能够非常明显地降低 miRNA 靶基因的假阳性率. 准确率较低的方法 ,间接检测 miRNA 改变而引起的组织/细胞 mRNA 或蛋白表达的改变 ,此类方法误差主要在于(1) miRISC 介导的转录后水平基因沉默 ,包括 mRNA 的降解及翻译抑制 ,而此类方法常常只检测 mRNA 或蛋白质的单一水平改变; (2) 差异表

达的基因有可能是靶基因的下游基因,即所谓的“间接靶点”; (3) 外源性 miRNA 的浓度不同,对靶基因的作用效率也不同^[36],实验过程中设置的 miRNA 浓度差别对结果也会有影响。

同时,实验方法由于受到来自不同个体、同一个体不同组织细胞或不同发育阶段及环境等多重因素的影响,结果同样缺乏一致性。首先,生物个体的 mRNA 转录存在着时空特异性,并且受到环境等因素的影响,因此,miRNA 在不同组织、不同细胞、生物体的不同生长发育阶段的所对应的“候选靶基因”也可能不同;其次,miRNA 表达也存在时空特异性,并且在组织、细胞中多个 miRNA 常同时表达,因此,生物体中不同的组织细胞及不同的生长发育阶段,miRNA 表达谱亦不同,在与 mRNA 结合的过程中它们之间相互影响,相互竞争,在只对单一 miRNA 进行研究时,不同的实验结果可能大相径庭;再者,根据 Seitz^[37] 提出竞争性内源 RNA 假说,miRNA 的众多靶标(mRNA、假基因转录物、长链非编码 RNA 等转录物)之间存在竞争性抑制,在不同组织细胞中 miRNA 所对应的“靶标”的差异自然会导致靶基因的差异;第四,miRNA 及 mRNA 存在序列多态性。尽管 miRNA 种子区的多态性非常罕见,频率低于 1%,但是 mRNA 3'UTR 区的多态性则非常丰富,常影响其与 miRNA 的结合。Clark 等^[27]认为,即使是那些非常保守的 miRNA,在不同生物甚至个体不同组织之间的行为也相去甚远。本课题组将 miR-205 mimic 分别导入 4 株 miR-205 低表达的乳腺癌细胞株,并以生物芯片检测 mRNA 与对照组之间的表达差异。结果显示,在 4 株乳腺癌细胞株中共同差异表达的 mRNA 不到 1%。

4 小结与展望

综上所述,生物信息学筛选方法与生物学实验筛选方法各有所长,但任何一方都不具有压倒性的优势。可以假设,形形色色的筛选软件所得到的海量结果,大多数可以看作是 miRNA 的“潜在靶基因”,它们之所以得不到实验方法的验证很可能是因为 miRNA 及 mRNA 的时空特异性表达差异造成的。因此,采用动态、联合的观点开发新的生物信息类筛选工具并建设综合性网站,将更有利于 miRNA 的靶基因筛选及功能挖掘。

致谢 特别感谢南昌大学临床医学系 2011 级黄昀、王炜东 2013 级黄明黎、李明辉同学在资料整理过

程中的帮助!

参考文献(References)

- [1] Eulalio A, Huntzinger E, Izaurralde E. Getting to the root of miRNA-mediated gene silencing[J]. *Cell*, 2008, 132(1): 9-14
- [2] Bartel D P. MicroRNAs: target recognition and regulatory functions[J]. *Cell*, 2009, 136(2): 215-233
- [3] Lewis B P, Burge C B, Bartel D P. Conserved seed pairing, often flanked by adenosines, indicates that thousands of human genes are microRNA targets[J]. *Cell*, 2005, 120(1): 15-20
- [4] Friedman R C, Farh K K, Burge C B, et al. Most mammalian mRNAs are conserved targets of microRNAs[J]. *Genome Res*, 2009, 19(1): 92-105
- [5] Bassett A R, Azzam G, Wheatley L, et al. Understanding functional miRNA-target interactions in vivo by site-specific genome engineering[J]. *Nat Commun*, 2014, 5: p. 4640
- [6] Xu W, Wang Z, Liu Y. The characterization of microRNA-mediated gene regulation as impacted by both target site location and seed match type [J]. *PLoS One*, 2014, 9(9): e108260
- [7] Clark P M, Loher P, Quann K, et al. Argonaute CLIP-Seq reveals miRNA targetome diversity across tissue types [J]. *Sci Rep*, 2014, 4: 5947
- [8] Wang T, Chen B, Kim M, et al. A model-based approach to identify binding sites in CLIP-Seq data [J]. *PLoS One*, 2014, 9(4): e93248
- [9] Kaller M, Oeljeklaus S, Warscheid B, et al. Identification of microRNA targets by pulsed SILAC [J]. *Methods Mol Biol*, 2014, 1188: 327-349
- [10] Betel D, Koppal A, Agius P, et al. Comprehensive modeling of microRNA targets predicts functional non-conserved and non-canonical sites[J]. *Genome Biol*, 2010, 11(8): R90
- [11] Garcia D M, Baek D, Shin C, et al. Weak seed-pairing stability and high target-site abundance decrease the proficiency of lsy-6 and other microRNAs[J]. *Nat Struct Mol Biol*, 2011, 18(10): 1139-1146
- [12] Kertesz M, Iovino N, Unnerstall U, et al. The role of site accessibility in microRNA target recognition [J]. *Nat Genet*, 2007, 39(10): 1278-1284
- [13] Dweep H, Sticht C, Pandey P, et al. miRWalk—database: prediction of possible miRNA binding sites by “walking” the genes of three genomes [J]. *J Biomed Inform*, 2011, 44(5): 839-847
- [14] Anders G, Mackowiak S D, Jens M, et al. doRiNA: a database of RNA interactions in post-transcriptional regulation [J]. *Nucleic Acids Res*, 2012, 40(Database issue): D180-186
- [15] Wang X, El Naqa I M. Prediction of both conserved and nonconserved microRNA targets in animals [J]. *Bioinformatics*, 2008, 24(3): 325-332
- [16] Linsley P S, Schelter J, Burchard J, et al. Transcripts targeted by the microRNA-16 family cooperatively regulate cell cycle progression [J]. *Mol Cell Biol*, 2007, 27(6): 2240-2252
- [17] Wang X. miRDB: a microRNA target prediction and functional

- annotation database with a wiki interface [J]. *RNA*, 2008, **14** (6): 1012-1017
- [18] Paraskevopoulou M D, Georgakilas G, Kostoulas N, *et al.* DIANA-microT web server v5.0: service integration into miRNA functional analysis workflows [J]. *Nucleic Acids Res*, 2013, **41** (Web Server issue): W169-173
- [19] Reczko M, Maragkakis M, Alexiou P, *et al.* Functional microRNA targets in protein coding sequences [J]. *Bioinformatics*, 2012, **28** (6): 771-776
- [20] Sturm M, Hackenberg M, Langenberger D, *et al.* TargetSpy: a supervised machine learning approach for microRNA target prediction [J]. *BMC Bioinformatics*, 2010, **11**: 292
- [21] Bandyopadhyay S, Mitra R. TargetMiner: microRNA target prediction with systematic identification of tissue-specific negative examples [J]. *Bioinformatics*, 2009, **25** (20): 2625-2631
- [22] Mitra R, Bandyopadhyay S. MultiMiTar: a novel multi objective optimization based miRNA-target prediction method [J]. *PLoS One*, 2011, **6** (9): e24583
- [23] Coronello C, Benos P V. ComiR: Combinatorial microRNA target prediction tool [J]. *Nucleic Acids Res*, 2013, **41** (Web Server issue): W159-164
- [24] Coronello C, Hartmaier R, Arora A, *et al.* Novel modeling of combinatorial miRNA targeting identifies SNP with potential role in bone density [J]. *PLoS Comput Biol*, 2012, **8** (12): e1002830
- [25] Welch C, Chen Y, Stallings R L. MicroRNA-34a functions as a potential tumor suppressor by inducing apoptosis in neuroblastomacells [J]. *Oncogene*, 2007, **26** (34): 5017-5022
- [26] Beitzinger M, Peters L, Zhu J Y, *et al.* Identification of human microRNA targets from isolated argonaute protein complexes [J]. *RNA Biol*, 2007, **4** (2): 76-84
- [27] Clark P M, Loher P, Quann K, *et al.* Argonaute CLIP-Seq reveals miRNA targetome diversity across tissue types [J]. *Sci Rep*, 2014, **4**: 5947
- [28] Hafner M, Landthaler M, Burger L, *et al.* Transcriptome-wide identification of RNA-binding protein and microRNA target sites by PAR-CLIP [J]. *Cell*, 2010, **141** (1): 129-141
- [29] Helwak A, Kudla G, Dudnakova T, *et al.* Mapping the human miRNA interactome by CLASH reveals frequent noncanonical binding [J]. *Cell*, 2013, **153** (3): 654-665
- [30] Konig J, Zarnack K, Rot G, *et al.* iCLIP reveals the function of hnRNP particles in splicing at individual nucleotide resolution [J]. *Nat Struct Mol Biol*, 2010, **17** (7): 909-915
- [31] Ong S E, Blagoev B, Kratchmarova I, *et al.* Stable isotope labeling by amino acids in cell culture, SILAC, as a simple and accurate approach to expression proteomics [J]. *Mol Cell Proteomics*, 2002, **1** (5): 376-386
- [32] Vergoulis T, Vlachos I S, Alexiou P, *et al.* TarBase 6.0: capturing the exponential growth of miRNA targets with experimental support [J]. *Nucleic Acids Res*, 2012, **40** (Database issue): D222-229
- [33] Xiao F, Zuo Z, Cai G, *et al.* miRecords: an integrated resource for microRNA-target interactions [J]. *Nucleic Acids Res*, 2009, **37** (Database issue): D105-110
- [34] Hsu S D, Chu C H, Tsou A P, *et al.* miRNAmap 2.0: genomic maps of microRNAs in metazoan genomes [J]. *Nucleic Acids Res*, 2008, **36** (Database issue): D165-169
- [35] Li J H, Liu S, Zhou H, *et al.* starBase v2.0: decoding miRNA-ceRNA, miRNA-ncRNA and protein-RNA interaction networks from large-scale CLIP-Seq data [J]. *Nucleic Acids Res*, 2014, **42** (Database issue): D92-97
- [36] Janas M M, Wang B, Harris A S, *et al.* Alternative RISC assembly: binding and repression of microRNA-mRNA duplexes by human Ago proteins [J]. *RNA*, 2012, **18** (11): 2041-2055
- [37] Seitz H. Redefining microRNA targets [J]. *Curr Biol*, 2009, **19** (10): 870-873