**《Analysis of microRNA-target interactions across diverse cancer types》**

miRNA 是一类通过与目标mRNA结合来调控基因表达的非编码单链RNA分子，肿瘤的产生和发展可能与miRNA的失调有关。虽然可以根据序列的互补性来判断miRNA所作用的mRNA，但细胞类型，条件，和其与辅因子的结合等因素，都会影响实际的作用关系。

为了从肿瘤样本的失调miRNA表达谱中获取（与肿瘤的产生与发展）相关的靶基因和通路，一般采用实验性地扰动细胞系或小鼠中的miRNA的表达的方法，但这种方法忽略了病人产生肿瘤的复杂环境。

因为miRNA会不稳定化或降解目标mRNA，两者的表达理论上呈负相关。可以利用TCGA（The Cancer Genome Atlas ）提供的不同肿瘤中mRNA和miRNA的表达信息，来获得miRNA到mRNA的作用关系的预测。

对分别的癌症类型，这样的分析已颇有成效。而本文旨在利用在不同癌症类型中共有的mRNA与miRNA的联系，找出可能的在不同组织和器官中影响肿瘤发生的共有的miRNA调控过程。

1. **数据来源**

所有的miRNA和mRNA表达数据来自TCGA 数据库，本文使用了9种上皮癌和1种多形性胶质母细胞瘤的数据，每种包括94到671个肿瘤样本，共计3290个样本。

DNA拷贝数数据(aCGH) 和DNA甲基化数据来自firehose。

1. **分析方法与主要结果**

本文设计了一种叫做REC（association recurrence score）的指标，其利用不同类型癌症中的miRNA和mRNA的表达，来推断在各种类型癌症中可能存在的miRNA-target作用关系。

具体地（图1），对每个miRNA-mRNA对，可以建立一个如图的多元线性方程，先消除DNA拷贝数和mRNA附近的启动子甲基化的影响，从而获得miRNA表达向量的系数。

对不同的癌症种类，分别计算这对miRNA-mRNA的相关性并进行排序（图2），综合这些相关性即可获得REC指标。可以看出，REC指标既能筛选出在少数种类癌症中有很高的负相关性的miRNA-mRNA对，也能筛选出在多数种类癌症中都有中等负相关性的miRNA-mRNA对。

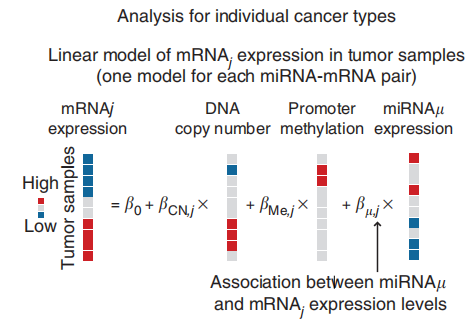


图1：获得一对miRNA与mRNA相关关系的线性模型

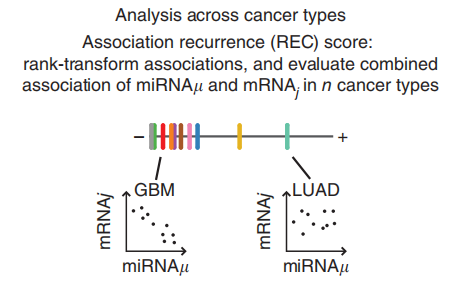


图2：获得一对miRNA与mRNA相关关系的线性模型

为了验证REC指标的有效性，本文对同时在至少5种（共10种）癌症中表达的miRNA-mRNA对计算REC得分（图4），发现有最高负分的10对中有8对被miRanda和TargetScan这种基于序列的方法同时预测为会发生结合或保守结合，其中两对在之前的研究中被发现在癌症环境下的确存在作用关系。

此外从图3可以看出，对所有的癌症类型，观测到的或者预计会发生作用的miRNA-mRNA对都聚集在两者有较高负相关的区间，也就是说对所有10种癌症类型，miRNA与mRNA的负相关越高，越有可能该对miRNA-mRNA存在作用关系。

综上说明一对miRNA-mRNA的REC得分越低，它们就越有可能在所有癌症中存在作用关系，REC指标具有一定的合理性。

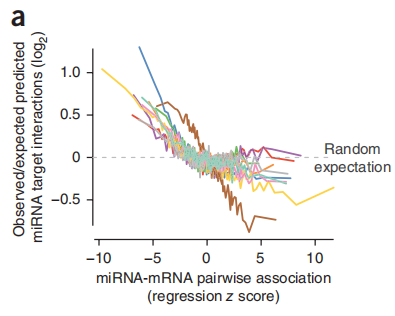


图3：横轴为miRNA-mRNA对的相关性，纵轴为观察到或预计两者会发生作用的对的数量

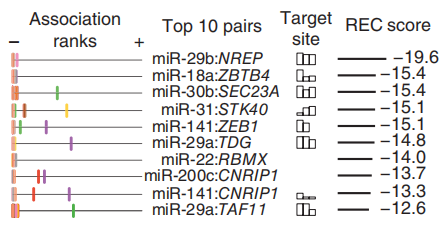


图4：REC得分最低的组合对应的其他target预测方法的结果（Target site）

为了验证REC指标能否真的找出在不同癌症类型中共有的miRNA调控关系。实验选用有很多较低REC得分的miRNA，在不同癌细胞系里对这些miRNA进行抑制或者过表达扰动。

发现（图5）在不同的癌细胞系这些miRNA扰动都会对mRNA产生明显影响，说明在不同癌症类型中共有的miRNA调控关系存在并能被找出，而且泛癌中miRNA与mRNA的相关性的确蕴含了miRNA的调控关系。

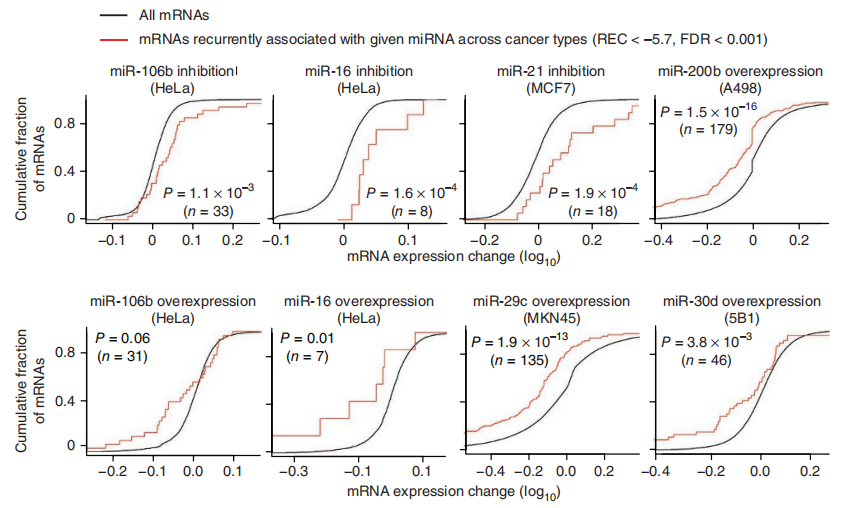


图5

将有较高负REC得分的miRNA-mRNA对，结合之前提到的TargetScan等三种基于序列的方法共同进行筛选（图6），获得143组高可信的miRNA-mRNA对。这些miRNA-mRNA对里，有61对是有实验验证的，甚至有23对被验证为与癌症调控有关。

由143对miRNA-mRNA，40种miRNA和72种mRNA组成的泛癌网络如图7。

因为根据以上方法获得的泛癌网络中，有很高的比例被证实与癌症调控相关，因此剩下被筛选出来的部分可以为miRNA调控癌症的研究提供可能的方向。

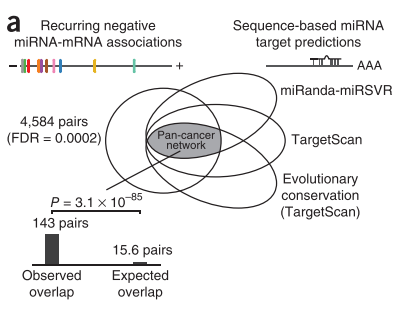


图6

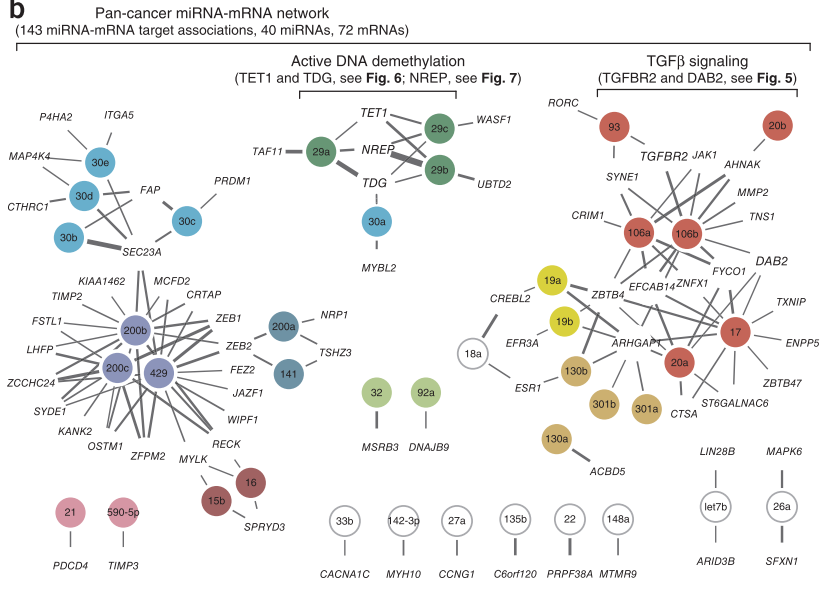


图7：线条越粗说明REC得分越低

此外，为了验证是否有一部分miRNA是由体细胞遗传和表观遗传修饰调控的。本实验（图8）比较了泛癌网络中的miRNA和其它miRNA。发现在不同的癌症类型中，泛癌网络中的miRNA更容易受到DNA拷贝数变化以及启动子DNA甲基化的调控。又因为DNA拷贝数变化以及启动子DNA甲基化与肿瘤的发生过程密切相关，这从某种程度上说明了泛癌网络中的miRNA-mRNA作用关系很可能对不同癌症的肿瘤发生起到一定作用。

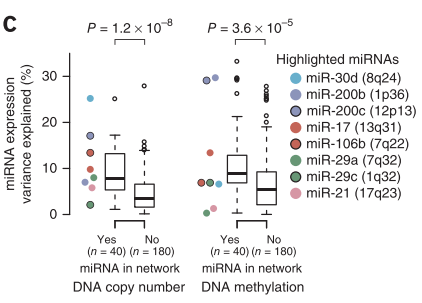


图8

最后，利用得到的泛癌网络进行验证性的研究。包括miR-106 miRNA家族对TGF-β信号通路的调控，miR-29调节激活DNA脱甲基通路，miR-29b和NREP组成双负反馈环等。

1. **总结**

本研究提出病人肿瘤中miRNA与mRNA的相关性可以用于推测存在于多种癌症类型中的miRNA-mRNA作用关系对。并通过比较已有实验和在不同癌细胞系中使用找出的miRNA进行扰动证明了所找出的miRNA-mRNA作用关系对的合理性。

将该统计方法结合基于序列的miRNA靶位点预测方法，对超过3000组肿瘤样本进行处理，获得了包含143对miRNA-mRNA，40种miRNA，72种mRNA的泛癌网络，其中许多的miRNA和miRNA-mRNA作用关系已在癌症领域被广泛研究。而且所得网络中的miRNA更可能被DNA拷贝数变化以及启动子DNA甲基化扰动，与这些作用关系可能与癌症发生有关的假设相符。

本文使用基于10种癌症获得的泛癌网络，做了一些验证实验。但更重要的是，本文提供的构建泛癌网络的方法，能够找出较可能发生的miRNA对mRNA的调控组合，为未来其它对miRNA与癌症产生和发展的关系的研究提供了可能的方向。

参考文献：

[1] Jacobsen A , Silber J , Harinath G , et al. Analysis of microRNA-target interactions across diverse cancer types[J]. Nature Structural & Molecular Biology, 2013, 20(11):1325-U131.