

# 三阴性乳腺癌miRNA主调控因子的分析鉴定

张 昆, 郭文花, 杨鲁红

山西师范大学生命科学学院, 山西 临汾  
Email: ylh1010309@126.com

收稿日期: 2021年3月24日; 录用日期: 2021年4月8日; 发布日期: 2021年4月25日

## 摘 要

目的: 通过生物信息学方法分析与三阴性乳腺癌发生发展有关的关键miRNA。方法: 从GEO数据库中找到三阴性乳腺癌miRNA数据集GSE121657, 利用GEO2R鉴定差异表达的miRNA, 通过starBase数据库预测出差异常表达的miRNA所对应的靶基因, 并对靶基因进行GO基因功能富集分析以及KEGG信号通路分析, 最后建立编码蛋白互作网络并且从中筛选出关键的miRNA。结果: 筛选出61个差异表达的miRNA, 获得与之相对应的165个靶基因, 这些靶基因主要涉及到细胞的定位、细胞周期、细胞分裂、细胞凋亡等生物学过程, 参与PI3K-Akt、多巴胺能突触、Hippo、MAPK等信号通路的调控, 找到关键性的调控因子hsa-miR-141-3p。结论: 生物信息学分析提示hsa-miR-141-3p可能是三阴性乳腺癌重要的调控因子。

## 关键词

三阴性乳腺癌, miRNA, 生物信息学

# Analysis and Identification of miRNA Master Regulators in Triple Negative Breast Cancer (TNBC)

Kun Zhang, Wenhua Guo, Luhong Yang

School of Life Science, Shanxi Normal University, Linfen Shanxi  
Email: ylh1010309@126.com

Received: Mar. 24<sup>th</sup>, 2021; accepted: Apr. 8<sup>th</sup>, 2021; published: Apr. 25<sup>th</sup>, 2021

## Abstract

**Objective:** The key miRNA in triple negative breast cancer (TNBC) was analyzed by bioinformatics analysis. **Method:** The Gene Expression Omnibus (GEO) dataset (GSE121657) was chosen and ex-

plored to identify differentially expressed miRNAs using GEO2R, and the target genes matching for the miRNAs were predicted by starBase. Enrichment analysis of Gene Ontology (GO) and the Kyoto Encyclopedia of Genes and Genomes (KEGG) was performed by the DAVID online tool. The networks between miRNAs and hub genes were constructed by Cytoscape software. Results: 61 differentially expressed microRNAs were screened and 165 corresponding target genes were obtained. These target genes mainly involved in biological processes such as cell location, cell cycle, cell division, cell apoptosis, etc., and mainly participated in PI3K-Akt, dopaminergic synapse, Hippo, and MAPK signaling pathways. The key regulatory factors hsa-mir-141-3p have been screened by the signal pathway, insulin pathway and MAPK pathway. Conclusion: Bioinformatics analysis suggested miRNA hsa-mir-141-3p may be a good classifier of TNBC and might play key roles in the progression of TNBC.

## Keywords

Triple Negative Breast Cancer (TNBC), miRNA, Bioinformatics

Copyright © 2021 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

## 1. 引言

乳腺癌逐渐成为女性的高发病症，位居于中国女性癌症发病的首位[1]。三阴性乳腺癌(TNBC)是雌激素受体、孕激素受体和原癌基因 Her-2 的检测结果显示为阴性的一种乳腺癌，在乳腺癌的所有分类中占据 10.0%~20.8% [2] [3]。其病原体在患者体内有着高度扩散性和侵袭性，容易在脏器间发生转移，且缺少有用的靶向治疗和内分泌靶点，晚期化疗效果欠佳。由于具有侵袭早、转移快、病死率高的特点，TNBC 是乳腺癌中预后最差的亚型[4] [5]。为了找到有效的治疗策略，降低 TNBC 的复发率、提高预后，许多研究人员做了大量的工作[6] [7]，寻找有效的治疗靶点是 TNBC 有效治疗的关键[8] [9]。

微小 RNA (miRNA)为一段长度约在 19~25 nt 的非编码单链 RNA，靶基因 3'-非编码区(UTR)能与其相结合，因而获得抑制靶基因表达的功能并且能够参与很多种生物过程，比如：细胞的增殖、凋亡和分化等生物过程。microRNA 对于肿瘤发挥着双重作用，着重表现为抑癌和致癌。据探究，miRNA 作为肿瘤抑制因子在 TNBC 中起着诱导肿瘤细胞凋亡和分化的功能；同时，miRNA 调节 TNBC 的进展主要是通过上皮-间充质转化(EMT)参与的，这也许与乳腺癌的诊断以及治疗有关系[10]。导致 TNBC 患者死亡的主要理由是癌细胞的转移，miRNA 有可能通过 EMT 参与调节 TNBC 的侵袭性进展。

生物信息学是生命科学和计算机科学相结合综合性学科。研究证明，整合的生物信息学方法可以帮助研究人员更好地探索疾病潜在的分子生物学机制，为临床提供新的诊断标志物和治疗靶点[11]。本研究以 GEO 数据库中三阴性乳腺癌(TNBC)上皮间充质转换(EMT)模型的 miRNA 芯片数据集 GSE121657 为研究对象，分析三阴性乳腺癌患者差异表达的 miRNA，预测与其相匹配的靶基因，并对靶基因进行生物信息学分析，最终获得调控三阴性乳腺癌的核心 miRNA，以此为三阴性乳腺癌的相关研究供给新的思路。

## 2. 材料与方法

### 2.1. 三阴性乳腺癌(TNBC) miRNA 芯片数据

本研究运用到的芯片数据集 GSE121657 来自于美国国立生物技术信息中心基因公共表达数据库

GEO, 由 Boston University Microarray and Sequencing Resource 上传[12], 为三阴性乳腺癌(TNBC)的上皮-间质转化(EMT)模型的 miRNA 表达数据, 共 6 个样本, 其中 GSM3442850、GSM3442851、GSM3442852 为 MIIshGFP 对照组, GSM3442853、GSM3442854、GSM3442855 为 MIIshSMAD2 处理组。上皮-间质转化使原发癌细胞获得间质特征, 并且在转移部位上皮再生形成继发性肿瘤。在该过程中癌细胞的蛋白转录调节会引起细胞特征性改变, 所以可以通过对 MIIshGFP 对照组与 MIIshSMAD2 处理组两组样本进行差异分析, 并能够提供预测依据和治疗靶点。对该数据集进行质量分析, 结果显示, 数据集中包含的六个样本的中心大致处于同一水平线(图 1), 样本数据总体上差异不大, 对其进行差异分析结果可靠。

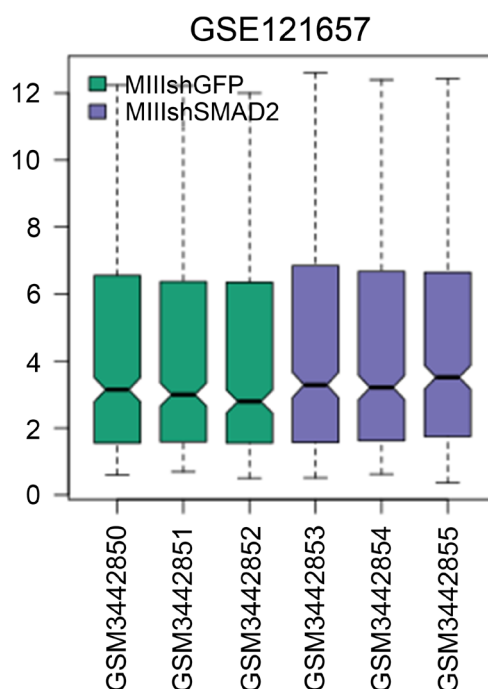


Figure 1. Boxplot graph of GSE121657

图 1. 样本 GSE121657 数据箱线图

## 2.2. 三阴性乳腺癌 miRNA 芯片数据的生物信息学处理

### 2.2.1. 差异表达 miRNA 的获取

运用 GEO 数据库自带的 GEO2R 在线分析工具, 根据  $|\logFC| > 1$ ,  $P\text{-Value} < 0.05$ , 筛查出三阴性乳腺癌患者差异表达的 miRNA。

### 2.2.2. 差异表达 miRNA 的靶基因预测

本研究使用 ENCORI (starBase v3.0) 预测靶基因[13] [14]。starBase 数据库是 miRNA 靶基因预测的综合数据库, 整合了多个数据库如 TargetScan、miRanda 等的信息, 大大降低了预测的假阳性。将之前获得的差异 miRNA 依次输入到 ENCORI 中, 设定 Number of supporting Experiments  $\geq 5$ , Program Number  $\geq 5$  prediction software 进行预测。

### 2.2.3. 差异表达微小 RNA 对应靶基因的相关功能分析、信号通路分析

使用 DAVID 在线工具对预测到的靶基因进行 GO 生物学功能分析和 KEGG 信号通路分析。把上述

的靶基因输入 DAVID, 从 Biological Process、Molecular Function 以及 Cellular Component 三个方面阐明靶基因的生物学功能, 分析其参与调控的信号通路。

#### 2.2.4. 靶基因编码蛋白相互作用(PPI)网络的构建与分析

将靶基因上传至 STRING 在线分析工具, 构建靶基因编码的蛋白质相互作用网络, 并利用 Cytoscape 软件对网络进行处理和筛查, 筛选出较为关键的靶基因, 进一步获得三阴性乳腺癌关键的 miRNA 分子。

### 3. 结果

#### 3.1. 差异表达的 miRNA

经分析共获得 61 个差异 miRNA, 与 MIIIshGFP 对照组相比, 表达上调的 miRNA 有 27 个, 下调的有 34 个, 见表 1 和图 2。

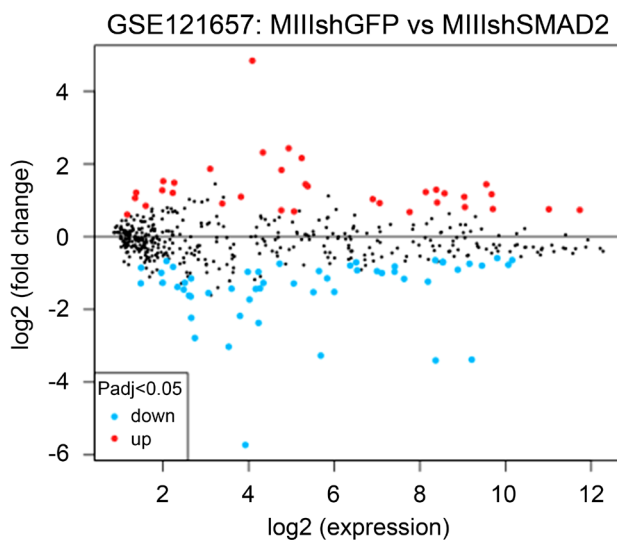
#### 3.2. 差异 miRNA 对应的靶基因

将 61 个差异表达的 miRNA 输入到在线靶基因预测数据库 starBase v3.0 后, 共获得 165 个与之相对应的靶基因。

**Table 1.** Differentially expressed miRNAs

**表 1.** 差异表达的微小 RNA

| Up Regulation (27) | Down Regulation (34) |
|--------------------|----------------------|
| hsa-miR-125b-1-3p  | hsa-miR-4417         |
| hsa-miR-494-3p     | hsa-miR-200c-3p      |
| hsa-miR-3609       | hsa-miR-205-5p       |
| hsa-miR-4433-3p    | hsa-miR-3648         |
| hsa-miR-4492       | hsa-miR-1307-3p      |
| hsa-miR-935        | hsa-miR-342-3p       |
| hsa-miR-224-5p     | hsa-miR-1972         |
| hsa-miR-27b-3p     | hsa-miR-708-5p       |
| hsa-miR-4649-5p    | hsa-miR-92a-1-5p     |
| hsa-let-7a-5p      | hsa-miR-193b-5p      |
| hsa-miR-4463       | hsa-miR-106b-3p      |
| hsa-miR-146a-5p    | hsa-miR-6789-5p      |
| hsa-miR-125b-5p    | hsa-miR-6831-5p      |
| hsa-miR-224-3p     | hsa-miR-500a-5p      |
| hsa-miR-125a-5p    | hsa-miR-503-5p       |
| hsa-let-7c-5p      | hsa-miR-3929         |
| hsa-miR-1271-5p    | hsa-miR-18a-3p       |
| hsa-miR-140-3p     | hsa-miR-141-3p       |
| hsa-miR-100-5p     | hsa-miR-4632-5p      |
| hsa-miR-6780b-5p   | hsa-miR-23a-5p       |
| hsa-miR-99a-5p     | hsa-miR-584-5p       |
| hsa-miR-4485       | hsa-miR-4674         |
| hsa-miR-125a-3p    | hsa-miR-3195         |
| hsa-miR-23b-3p     | hsa-miR-4665-5p      |
| hsa-miR-3135b      | hsa-miR-4669         |
| hsa-miR-6126       | hsa-miR-93-3p        |
| hsa-miR-210-3p     | hsa-miR-4449         |
|                    | hsa-miR-1231         |
|                    | hsa-miR-3180         |
|                    | hsa-miR-378f         |
|                    | hsa-miR-200b-5p      |
|                    | hsa-miR-3187-3p      |
|                    | hsa-miR-30a-5p       |
|                    | hsa-miR-1233-5p      |

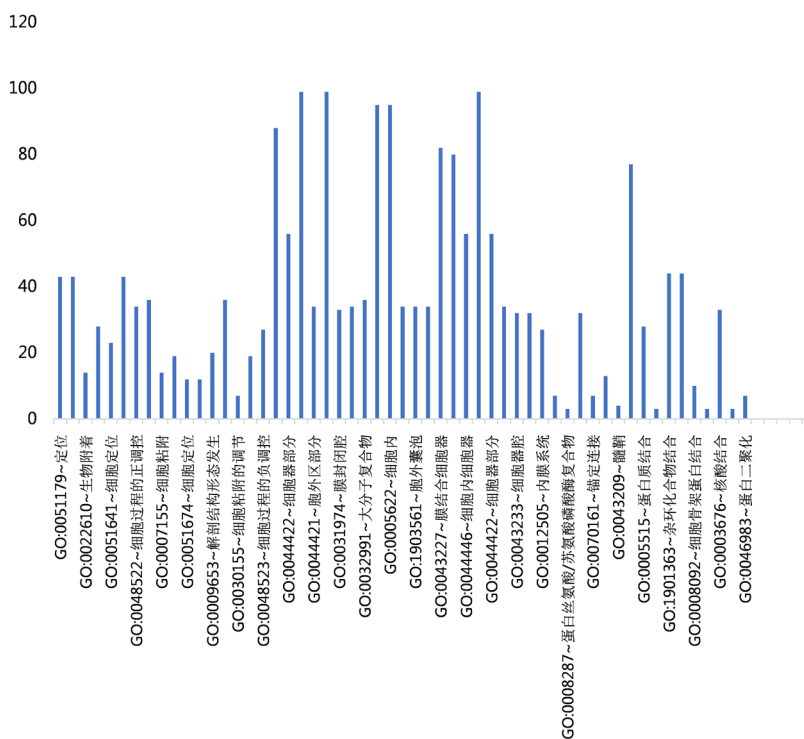


**Figure 2.** Differentially expressed miRNA detection map of GSE121657 (Red represents upregulation, and blue represents downregulation)

**图 2.** GSE121657 差异 miRNA 的表达值图(红色代表上调, 蓝色代表下调)

### 3.3. 基因功能(GO)富集分析

运用 DAVID 对靶基因进行 Gene\_Ontology 功能分析, 结果显示, 165 个靶基因主要参与了细胞的定位、细胞过程的正调控等生物学过程, 参与了细胞器部分、细胞内外组分等细胞成分的构成, 涉及蛋白质及杂环化合物结合等分子功能(如图 3)。



**Figure 3.** GO function enrichment analysis of target genes ( $P < 0.05$ ) histogram

**图 3.** 靶基因的 GO 功能富集分析( $P < 0.05$ )柱状图

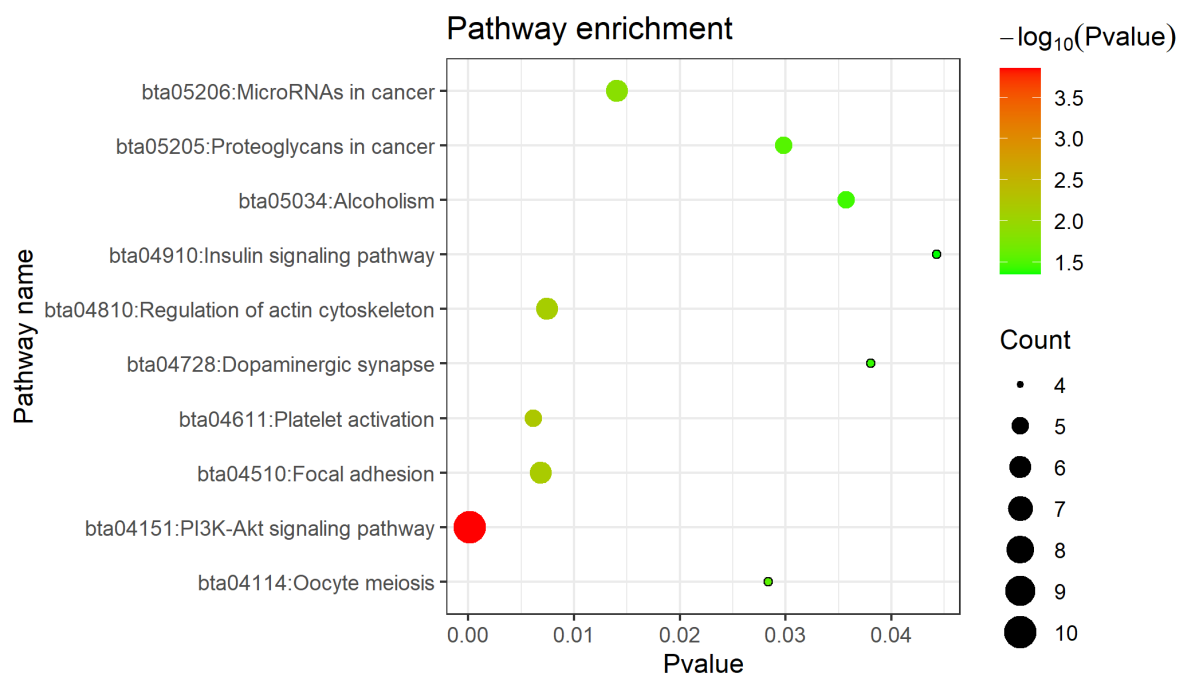
### 3.4. KEGG Pathways 分析

通过 DAVID 在线数据库对 165 个靶点基因进行信号通路富集分析, 以  $P < 0.05$  为筛选标准, 共获得 10 条 KEGG 信号通路, 涉及 PI3K-Akt、血小板活化、黏附斑激酶、激动蛋白细胞骨架的调节、癌组织中的 miRNA 等信号通路(表 2、图 4)。

**Table 2.** Signaling pathways involved in target genes ( $P < 0.05$ )

**表 2.** 靶基因涉及的信号通路( $P < 0.05$ )

| Term                     | Count | P-Value  | Genes  |
|--------------------------|-------|----------|--|
| bta04151: PI3K-Akt 信号通路  | 10    | 1.40E-04 | YWHAG, EIF4EBP1, MCL1, GRB2, CSF1, PPP2CA, NR4A1, LAMC1, GNG5, DDIT4 |
| bta04611: 血小板活化素         | 5     | 0.006164 | TLN1, PPP1CA, FYN, ARHGEF12, PPP1CB                                  |
| bta04510: 粘附斑激酶          | 6     | 0.006858 | TLN1, PPP1CA, GRB2, FYN, LAMC1, PPP1CB                               |
| bta04810: 肌动蛋白细胞骨架的调节    | 6     | 0.007422 | PPP1CA, CFL2, MSN, ARPC5, ARHGEF12, PPP1CB                           |
| bta05206: 癌组织中的 MicroRNA | 6     | 0.01403  | MCL1, GRB2, ZEB2, MARCKS, ZEB1, DDIT4                                |
| bta04114: 卵母细胞减数分裂酶      | 4     | 0.028359 | YWHAG, PPP1CA, PPP2CA, PPP1CB  |
| bta05205: 癌组织中的蛋白多糖      | 5     | 0.029795 | PPP1CA, GRB2, MSN, ARHGEF12, PPP1CB                                  |
| bta05034: 酒精中毒           | 5     | 0.03572  | PPP1CA, GRB2, PPP1CB, PKIA, GNG5                                     |
| bta04728: 多巴胺能突触         | 4     | 0.038072 | PPP1CA, PPP2CA, PPP1CB, GNG5   |
| bta04910: 胰岛素信号通路        | 4     | 0.044296 | PPP1CA, EIF4EBP1, GRB2, PPP1CB                                       |



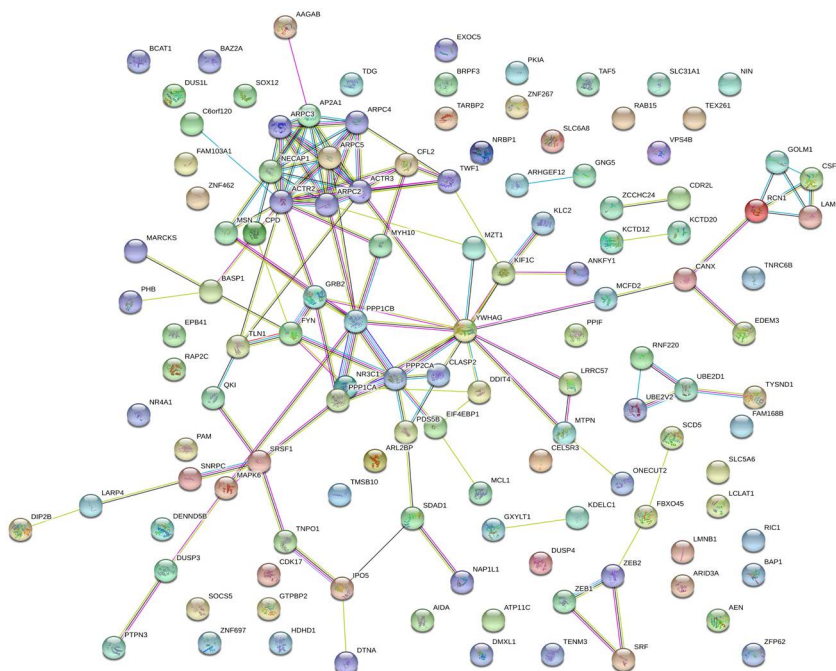
**Figure 4.** KEGG bubble chart of target genes ( $P < 0.05$ )

**图 4.** 靶基因涉及的 KEGG 气泡图( $P < 0.05$ )柱状图

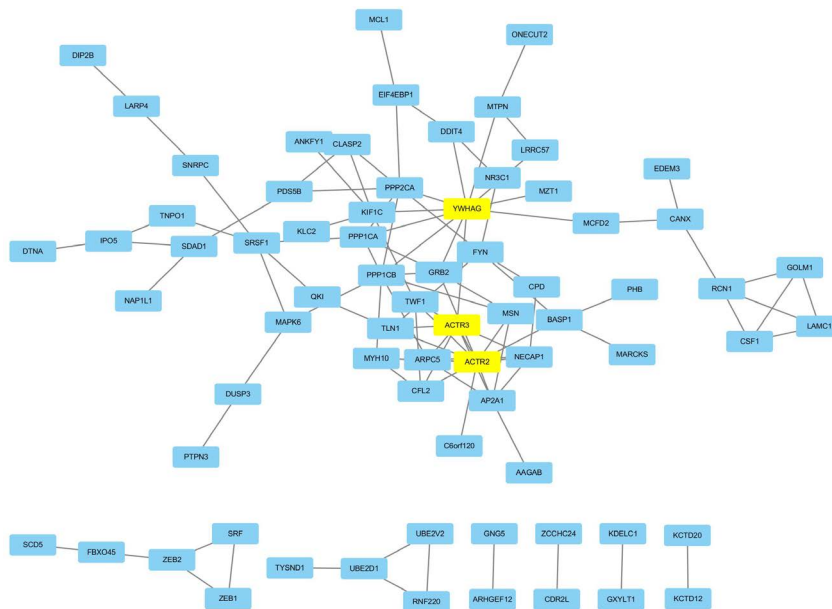
### 3.5. 核心基因及关键 miRNA

将差异表达的 miRNA 所对应的靶基因导入 STRING, 得到一个涵盖 124 个靶基因蛋白节点和 126

条交互关系的蛋白质相互作用网络(图 5), 利用 Cytoscape 对蛋白互作网络进行进一步的分析处理。以 Degree > 10 为筛选条件, 从蛋白互作网络中筛选出 3 个核心靶基因, 分别是 YWHAG、ACTR2、ACTR3 (图 6), 并找到了与这 3 个靶基因相对应的 miRNA, 即 hsa-miR-141-3p, 图 7 显示了该 miRNA 在 TNBC 中的表达情况。

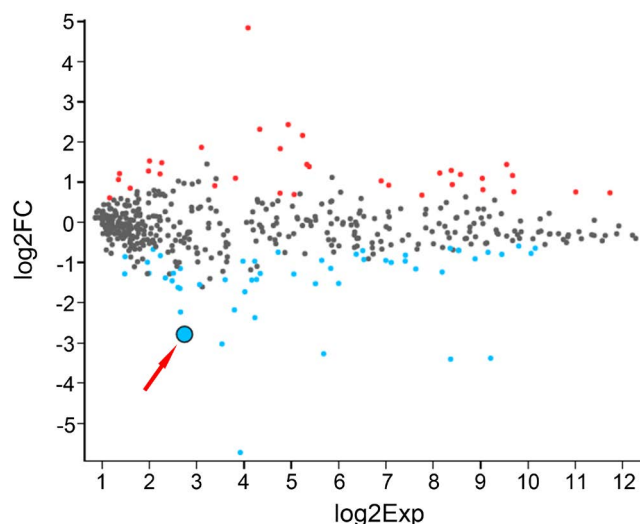


**Figure 5.** Protein interaction network coded by target genes  
**图 5.** 靶基因编码的蛋白互作网络图



**Figure 6.** The protein interaction network including key target genes is shown, and the nodes shown in blue in the figure are the key target genes

**图 6.** 显示出关键靶基因的蛋白互作网络, 图中蓝色显示且形状较大的节点为关键靶基因



**Figure 7.** Location of hsa-miR-141-3p in mean difference plot (plotMD)  
**图 7.** hsa-miR-141-3p 在平均差异图中的位置

#### 4. 讨论

TNBC 是指 ER、PR 以及原癌基因 Her-2 检查结果全都为阴性的乳腺癌，尽管其有许多的生物学特征与 BLBC 相类似，但是 TNBC 和基底细胞样型乳腺癌之间还存在着处在某些免疫表型和基因表达谱上的差异。目前缺乏与之相对应靶向治疗受体，而且对 TNBC 发生的分子机制也没有完整的了解，对于 TNBC 的医治手段有限。为此，对 TNBC 的靶向治疗的探究在如火如荼地开展。

微小 RNA 是一种非编码小 RNA 分子，截止到目前人们已经发现了三千多种微小 RNA，在这之中绝大多数在动植物体内发挥着关键性的调控作用，是最主要的基因表达调控因子之一。预计在人体内将近 2/3 的基因都会受到某一个或一组微小 RNA 的调控[15]。近几年来，有研究表明单个 miRNA 可以同时靶向多个 mRNA，单个 mRNA 可以被不同的 miRNA 协同靶向，miRNA 拥有着高度的保守性，参与多种生物学过程，如细胞增殖、细胞脂肪代谢、糖代谢和胰岛素信号转导等[16]。近些年来，更多的人开始对微小 RNA 在很多相关疾病中的作用增加关注。在本研究中，采取了生物信息学的手段对三阴性乳腺癌的微小 RNA 芯片数据进行了研究，筛选后得到 61 个差异表达微小 RNA 中，预测得到 165 个靶基因，这些靶基因主要涉及细胞的定位、细胞周期、细胞分裂以及凋亡、PI3K-Akt 信号通路等生物学过程。

PI3K-Akt 信号转导通路广泛分散在细胞中，是一条重要的信号传导途径在细胞内。这一通路过度激活后会促进细胞的过度增殖、抑制细胞凋亡、促使细胞分化异常以及参与细胞自噬等导致肿瘤形成和促进肿瘤转移[17] [18]。有学者发现，PI3K-Akt 信号转导通路在 TNBC 的发展中起着十分关键作用[8]，这可能与 TNBC 侵袭性强、预后差的生物学特征和临床特点有紧密关联。所以，找到存在于这一通路中的关键靶点能够为 TNBC 的患者提供新的医疗方法并且拥有相对稳定的医治结果[18]。

通过分析蛋白质互作网络，获得了与 TNBC 相关的核心基因及其对应的 miRNA 分子，即 hsa-miR-141-3p，其在 TNBC 中表达降低(表 1、图 7)。研究表明，hsa-miR-141-3p 可以通过沉默下游靶点来调节多种癌症如前列腺癌、结肠癌、卵巢癌细胞的迁移和侵袭[19] [20] [21] [22]。Wong [12]等人发现，hsa-miR-141-3p 与其他一系列假定的肿瘤抑制 miRNAs 如 hsa-miR-4417、hsa-miR-200c-3p、hsa-miR-210-3p 等在恶性 TNBC 细胞中显著下调，经过对 hsa-miR-4417、miR-200c-3p、miR-210-3p 的 RT-qPCR 验证，认为 hsa-miR-4417 是三阴性乳腺癌的抑癌和预后标志物。而本文生物信息学分析结果显示，hsa-miR-141-3p 有可能才是调控 TNBC 发展的更加重要的 miRNA 分子，需要进一步进行实验验证。



## 参考文献

- [1] 李贺, 郑荣寿, 张思维, 等. 2014 年中国女性乳腺癌发病与死亡分析[J]. 中华肿瘤杂志, 2018, 40(3): 166-171.
- [2] Gelmon, K., Dent, R., Mackey, J.R., *et al.* (2012) Targeting Triple-Negative Breast Cancer: Optimising Therapeutic Outcomes. *Annals of Oncology*, **23**, 2223-2234. <https://doi.org/10.1093/annonc/mds067>
- [3] 黎辉, 韩明其, 陈明光. 三阴性乳腺癌的病理特征及临床预后探讨[J]. 中外医疗, 2020, 39(2): 29-31.
- [4] Liedtke, C., Rody, A. and Gluz, O. (2015) The Prognostic Impact of Age in Different Molecular Subtypes of Breast Cancer. *Breast Cancer Research and Treatment*, **152**, 667-673. <https://doi.org/10.1007/s10549-015-3491-3>
- [5] Cinkaya, A., Akin, M. and Sengul, A. (2016) Evaluation of Treatment Outcomes of Triple-Negative Breast Cancer. *Journal of Cancer Research and Therapeutics*, **12**, 150-154. <https://doi.org/10.4103/0973-1482.154000>
- [6] 曹晓珊, 丛斌斌. 精准医学时代三阴性乳腺癌治疗的研究进展[J]. 中国癌症杂志, 2019, 29(12): 971-976.
- [7] 修萌, 张频. 早期三阴性乳腺癌治疗模式及新辅助治疗研究进展[J/OL]. 肿瘤防治研究, 2020, 47(2): 53-58.
- [8] 李甜, 周钱梅, 张卫红. PI3K/Akt/mTOR 信号通路在三阴性乳腺癌治疗中的研究进展[J]. 中国肿瘤, 2018, 27(1): 40-45.
- [9] Yutaro, G., Aye, T., Hui, O.C.C., *et al.* (2020) Characteristics, Behaviour and Role of Biomarkers in Metastatic Triple-Negative Breast Cancer. *Journal of Clinical Pathology*, **73**. <https://doi.org/10.1136/jclinpath-2019-206078>
- [10] 郑健, 张雪鹏. 微小 RNA 在三阴性乳腺癌中的研究进展[J]. 华北理工大学学报(医学版), 2016, 18(6): 499-504.
- [11] 冷红. 基于生物信息学筛选三阴性乳腺癌差异表达基因[D]: [硕士学位论文]. 泸州: 西南医科大学, 2019.
- [12] Wong, C.K., Gromisch, C., Ozturk, S., *et al.* (2019) MicroRNA-4417 Is a Tumor Suppressor and Prognostic Biomarker for Triple-Negative Breast Cancer. *Cancer Biology & Therapy*, **20**, 1113-1120. <https://doi.org/10.1080/15384047.2019.1595285>
- [13] Zhou, K.R., Liu, S., Cai, L., Bin, L., *et al.* (n.d.) ENCORI: The Encyclopedia of RNA Interactomes.
- [14] Li, J.H., *et al.* (2014) starBase v2.0: Decoding miRNA-ceRNA, miRNA-ncRNA and Protein-RNA Interaction Networks from Large-Scale CLIP-Seq Data. *Nucleic Acids Research*, **42**, D92-D97. <https://doi.org/10.1093/nar/gkt1248>
- [15] Lewis, B.P., Burge, C.B. and Bartel, D.P. (2005) Conserved Seed Pairing, Often Flanked by Adenosines, Indicates That Thousands of Human Genes Are microRNA Targets. *Cell*, **120**, 15. <https://doi.org/10.1016/j.cell.2004.12.035>
- [16] 田亚丽, 古丽海夏·哈勒玛合拜, 张燕, 张奕之. 外泌体微小 RNA 与胰岛素抵抗的研究现状[J]. 中国临床药理学杂志, 2020, 36(8): 1016-1019.
- [17] 武强, 赵全年, 宋卫东. PI3K/Akt/mTOR 信号传导通路在肿瘤的研究进展[J]. 疾病监测与控制, 2013, 7(6): 346-347.
- [18] Massihnia, D., Perez, A., Bazan, V., *et al.* (2016) A Headlight on Liquid Biopsies: A Challenging Tool for Breast Cancer Management. *Tumor Biology*, **37**, 4263-4273. <https://doi.org/10.1007/s13277-016-4856-x>
- [19] Guan, G.Y., Niu, X., Qiao, X., *et al.* (2020) Upregulation of Neural Cell Adhesion Molecule 1 (NCAM1) by hsa-miR-141-3p Suppresses Ameloblastoma Cell Migration. *Medical Science Monitor*, **26**, e923491. <https://doi.org/10.12659/MSM.923491>
- [20] Huang, S., Wa, Q., Pan, J., *et al.* (2017) Downregulation of miR-141-3p Promotes Bone Metastasis via Activating NF-kappaB Signaling in Prostate Cancer. *Journal of Experimental & Clinical Cancer Research*, **36**, 173. <https://doi.org/10.1186/s13046-017-0645-7>
- [21] Liang, Z., Li, X., Liu, S., *et al.* (2019) MiR-141-3p Inhibits Cell Proliferation, Migration and Invasion by Targeting TRAF5 in Colorectal Cancer. *Biochemical and Biophysical Research Communications*, **514**, 699-705. <https://doi.org/10.1016/j.bbrc.2019.05.002>
- [22] Li, J.H., Zhang, Z., Du, M.Z., *et al.* (2018) microRNA-141-3p Fosters the Growth, Invasion, and Tumorigenesis of Cervical Cancer Cells by Targeting FOXA2. *Archives of Biochemistry and Biophysics*, **657**, 23-30. <https://doi.org/10.1016/j.abb.2018.09.008>