Biomedicine & Pharmacotherapy 107 (2018) 117–128

Contents lists available at ScienceDirect

Biomedicine & Pharmacotherapy

journal homepage: www.elsevier.com/locate/biopha

Bioinformatics-based interaction analysis of miR-92a-3p and key genes in tamoxifen-resistant breast cancer cells

Jinjing Cun[a], Qifeng Yang[a,b,*]

[a] Department of Breast Surgery, Qilu Hospital, Shandong University, Jinan, 250012, PR China
[b] Pathology Tissue Bank, Qilu Hospital, Shandong University, Jinan, 250012, PR China

ARTICLE INFO

Keywords:
MiR-92a-3p
MYC
Breast cancer
Tamoxifen resistance
Bioinformatics analysis
Gene Expression Omnibus database

ABSTRACT

The abnormal expression of miR-92a-3p was detected in multiple cancers. However, the biological role and underlying mechanism of miR-92a-3p in tamoxifen-resistant cells are still unknown. The main objective of our study was to find potential miR-92a-3p regulating pathways involved in tamoxifen resistance and to construct their regulatory network using bioinformatics. Four gene expression profiles were retrieved from GEO database and the GEO2R tool was used for analysis. GSE41922 and GSE42072 were applied to investigate aberrant miR-92a-3p expression in breast cancer serum and tissue. We found that miR-92a-3p expression was higher in breast cancer serum or tissue than in healthy volunteer serum or adjacent normal tissue, and high expression of miR-92a-3p could predict poor prognosis of breast cancer patients. In our qRT-PCR validation, we found that miR-92a-3p was upregulated in tamoxifen-resistant cells. MiR-92a-3p might play a role in tamoxifen resistance. In order to find the relationship between miR-92a-3p and some key genes and their potential molecular mechanisms in tamoxifen-resistant cells. The microarray data GSE26459 and GSE28267 were analyzed to determine the differentially expressed genes (DEGs) or miRNAs (DEMs). Furthermore, the related long non-coding RNAs (lncRNAs) were screened with starBase v2.0. Finally,microRNA.org,miRDB, targetminer and targetscan were applied to predict the targets of miR-92a-3p. Through analysis, we find that miR-92a-3p may be used as a potential biomarker for early detection of cancer and monitoring the efficacy of endocrine therapy.

1. Introduction

Breast cancer is the most common malignancy and is the leading cause of cancer-related death among females worldwide. About 50%–70% of breast cancers are thought to be estrogen receptor (ER) positive. Tamoxifen, a selective estrogen receptor modulator, is the first-line hormone therapy for these patients, can reduce recurrence and mortality rates of breast cancer. However, approximately 30% of patients treated with tamoxifen eventually develop resistance and make the treatment invalid. The presence of resistance has limited the therapeutic application of tamoxifen and ultimately affected the prognosis of breast cancer patients. Therefore, it is crucial to study the molecular mechanisms of tamoxifen resistance and assess more effective treatment strategies, to achieve a better survival rate of patients with breast cancer.

MicroRNAs (miRNAs) regulate gene expression mainly by matching the sequence of target genes, either completely or partially, resulting in protein translation inhibition or mRNA degradation. A single miRNA can regulate multiple target genes, while different miRNAs are able to regulate the same target gene to coordinate and control the regulation of protein suppression. MiRNAs often seem to be located in regions of genomic instability and are found to be closely related to various types of cancer-related cell processes, including proliferation, cell-cycle control, apoptosis, metastasis, angiogenesis and drug resistance, which indicates that miRNAs serve an important role in the pathogenesis of human cancer and therefore are good choices for cancer diagnosis and treatment.

The miR-17–92 cluster is composed of six miRNAs (miR-17, miR-18a, miR-19a, miR-20a, miR-19b, miR-92a-3p) and is located on chr13q31.3 within the third intron of the C13orf25/MIR17HG

Abbreviations: ER, estrogen receptor; miRNAs, microRNAs; QRT-PCR, quantitative real-time polymerase chain reaction; Ct, threshold cycle; DEGs, differentially expressed genes; DEMs, differentially expressed miRNAs; PPI, protein-protein interaction; GEO, Gene Expression Omnibus; FC, fold change; STRING, The Search Tool for the Retrieval of Interacting Genes; OS, overall survival; RFS, relapse-free survival

* Corresponding author at: Department of Breast Surgery, Qilu Hospital, Shandong University, Wenhua Xi Road No. 107, Jinan, 250012 Shandong Province, PR China.

E-mail address: qifengy_sdu@163.com (Q. Yang).

https://doi.org/10.1016/j.biopha.2018.07.158
Received 29 May 2018; Received in revised form 19 July 2018; Accepted 31 July 2018

(Chromosome 13 open reading frame 25) gene [1]. Because of its oncogenic activity in several types of cancer,the cluster is considered as an oncomiR. miR-92a-3p is a key component of miR-17–92 cluster. Up to now, several studies have investigated the clinical impact of miR-92a-3p in several types of cancers and have demonstrated the important role of miR-92a-3p in carcinogenesis. However, the exact roles of the miR-92a-3p in tamoxifen resistance have not yet been studied, systematic analysis of miR-92a-3p is required.

Data from high-throughput platforms, including microarrays and next-generation sequences, provide an unprecedented choice for understanding cancer progression and finding functional networks in cancer. They can provide amounts of useful information if analyzed using bioinformatics methods.But they are generally expensive. Therefore, Microarray data integrated from public databases, such as the GEO database, is popular among researchers. Bioinformatics analysis plays a key role in large-scale gene expression profiling, which lays a foundation for biological interpretation. With the availability and accumulation of more and more biomedical data, the computational methods will be more and more powerful for the future precision medicine strategy.

In this study, a series of resources and techniques of bioinformatics and computational biology were used to acquire the key genes in tamoxifen resistance,to explore potential molecular mechanisms and to construct regulatory networks between miR-92a-3p and mRNAs/lncRNAs. Through these comprehensive bioinformatic analyses, miR-92a-3p may be regarded as an important therapeutic target and that it exerts its functions in tamoxifen-resistant cells by targeting different potential genes. However, Larger samples and further investigation are needed to verify this conclusion.

2. Materials and methods

2.1. Cell culture

MCF-7 cells were obtained from ATCC (American Type Culture Collection, Manassas, VA, USA), and were maintained in DMEM medium (Macgene, Beijing, China) supplemented with 10% fetal bovine serum, 100 U/ml penicillin and 100 μg/ml streptomycin. M7R-tam cells were originally derived from MCF-7 cells by selection in increasing concentrations of tamoxifen (Sigma, St. Louis, MO, USA) for more than 1 year. Cell lines were cultured at 37 °C in a humidified atmosphere containing 5% carbon dioxide.

2.2. Microarray data

We searched the GEO database (https://www.ncbi.nlm.nih.gov/geo/) [2], a public functional genomics data repository for high-throughput gene expression datasets, using the following keywords: “breast cancer miRNA” or “breast cancer tamoxifen resistance” (study keyword), “Homo sapiens” (organism). After a systematic review, four gene expression profiles (GSE41922, GSE42072, GSE26459, GSE28267) were downloaded for further analysis. GSE41922 was the miRNA expression data with 22 serum from healthy volunteer and 32 pre-operative serum from breast cancer patients [3]. The dataset GSE42072, based on the platform of GOL16249, consisted of 7 pairs of breast cancer tissue and adjacent normal tissue [3]. GSE26459 was the mRNA expression data including 3 pairs of cell lines that were either sensitive or resistant to tamoxifen [4]. The dataset GSE28267, based on the platform of GPL7723, included MCF-7 cells and tamoxifen-resistant cells (LY2) [5].

2.3. Data processing

GEO2R (http://www.ncbi.nlm.nih.gov/geo/geo2r/) is an interactive web tool for comparing two or more groups of samples in a GEO dataset in order to identify DEGs or DEMs under the same experimental conditions. Morpheus (https://software.broadinstitute.org/morpheus/) is a matrix visualization and analysis platform designed to support visual data exploration. Heat map and boxplot of DEGs or DEMs can be generated using the online tool.

2.4. PPI network construction

The STRING database (http://string-db.org/) is a biological database of known and predicted protein–protein interactions (PPI), which aims to provide a critical assessment of the interaction between proteins, including direct (physical) and indirect (functional) associations [6]. Cytoscape is a popular open-source software that is used to visually explore the interaction networks of biomolecules [7]. We firstly predicted the genes that were associated with MYC by means of STRING, a combined score > 0.4 was set as the cut-off criterion. Then putative genes were intersected with the DEGs of GSE26459. Overlapping genes were used to build a PPI network. Finally, the PPI network was visualized using the Cytoscape software 3.6.1.

2.5. Venn diagram

Venny 2.1.0 is freely accessible at http://bioinfogp.cnb.csic.es/tools/venny/index.html. It can summarize the overlap between two to four lists, enabling researchers to quickly observe similarities and differences between the data sets they are analyzing. Each circle represents a data set, and the overlap between the circles corresponds to the overlap between the data sets.We used Venny 2.1.0 to find the intersection of the predicted genes associated with MYC and the DEGs of GSE26459, to identify the genes that affected the expression of MYC in tamoxifen-resistant cells. At the same time, a Venn diagram was conducted to examine the overlap of predicted target genes lists from four online prediction databases, to increase the accuracy of the forecast. We finally attained the potential targets of miR-92a-3p by gaining intersection of 144 predictive genes from online prediction databases and the DEGs of GSE26459, to find the target genes that indeed mediated the roles of miR-92a-3p in tamoxifen-resistant cells.

2.6. Quantitative reverse transcription-polymerase chain reaction

Total RNA from cultured cells was extracted using TRIzol reagent (Invitrogen) according to the manufacturer's protocol. 500 ng of total RNA was reverse transcribed. Subsequent real-time PCR (RT-PCR) was performed on Applied Biosystems 7900HT instrument. Results were normalized to U6 (small nuclear RNA) for miRNA or GAPDH for mRNA. All reactions were carried out in triplicate and relative gene expressions were analyzed by the $2^{-\Delta\Delta Ct}$ method.

2.7. MiRNA target prediction

The target genes of miR-92a-3p were predicted with miRDB (http://mirdb.org/miRDB/) [8], TargetMiner (https://www.isical.ac.in/~bioinfo_miu/targetminer20.htm) [9], TargetScan (http://www.targetscan.org/vert_71/) and microrna.org(http://34.236.212.39/microrna/home.do),they) [10]. They were commonly used to predict the targets of miRNAs. Only the target genes identified by all of the above prediction programs were considered to be the putative targets of miR-92a-3p. Subsequently, a miRNA–mRNA regulatory network depicting interactions between miR-92a-3p and their potential targets was visualized using Cytoscape 3.6.1 software.

2.8. Survival analysis

The Kaplan-Meier Plotter Database (KMPD) is an online survival analysis tool. It can be used to screen miRNA/mRNA related to the overall survival (OS) or relapse-free survival (RFS) of breast cancer patients. The web-tool allows selection of patients by receptors status,