Oncology and Translational Medicine
DOI 10.1007/s10330-020-0458-8

April 2021, Vol. 7, No. 2, P59–P64

ORIGINAL ARTICLE

Differentially expressed genes analysis and target genes prediction of miR-22 in breast cancer*

Tao Fan[1], Chaoqi Wang[2] (Co-first author), Kun Zhang[3], Hong Yang[1], Juan Zhang[1], Wanyan Wu[1], Yingjie Song[4] (✉)

[1] Department of Oncology, The People's Hospital of China Three Gorges University, The First People's Hospital of Yichang, Yichang 443000, China
[2] Department of Urinary Surgery, Affiliated Hospital of Inner Mongolia University for the Nationalities, Tongliao 028007, China
[3] Department of Orthopedics, The People's Hospital of China Three Gorges University, The First People's Hospital of Yichang, Yichang 443000, China
[4] Department of General Surgery, The People's Hospital of China Three Gorges University, The First People's Hospital of Yichang, Yichang 443000, China

Abstract

Objective miR-22 is highly active in breast cancer, especially in the luminal B and HER2 subtypes. However, the detailed potential of the use of target genes for miR-22 in breast cancer are still unclear. In this study, we aimed to discover potential genes and the miRNA-DEGs network of miR-22 in breast cancer using bioinformatics approaches.

Methods Analysis of microarray data GSE17508 (including 3 miR-22 knockout samples and 3 controls) obtained from the Gene Expression Omnibus (GEO) database was performed. Differentially expressed genes (DEGs) between the miR-22 knockout samples and the three control samples were detected using GEO2R. The gene ontology (GO) functional enrichment analysis and protein-protein interaction (PPI) network of DEGs were performed using the online tool Metascape and STRING database, separately. The miR-22 and DEG networks were obtained from the miRNet database. Cytoscape software was used to construct and analyze a merged miRNA-DEG network. The online tools database, mirDIP 4.1, was used to predict miR-22 target genes.

Results Certain DEGs and miRNAs may be potential targets for predicting and treating miR-22 expressed breast cancer.

Conclusion We constructed a prognostic model of rectal adenocarcinomas based on four immune-related lncRNAs by analyzing the data based on TCGA database, with high prediction accuracy. We also identified two biomarkers with poor prognosis (PXN-AS1 and AL158152.2) and one biomarker with good prognosis (LINC01871).

Received: 5 October 2020
Revised: 25 October 2020
Accepted: 15 November 2020

Key words: bioinformatics; breast cancer; MCF7 cells; MiR-22

MicroRNAs (miRNAs) are a class of noncoding single-stranded RNA molecules with a length of about 22 nucleotides encoded by endogenous genes, which coordinate multiple gene expression programs through gene regulation [1–2]. There are over 1,700 identified miRNAs in the human genome that are associated with a wide variety of human cancers, such as breast, lung, and colon cancer [2–3]. Research suggests that miRNAs may be related to the pathogenesis of cancer, tumor growth, and metastasis, and play the role as oncogenes or tumor suppressor genes [4]. Therefore, the identification of miRNA targets is considered to be key in improving our understanding of the regulatory effects of miRNAs. miR-22 has been identified as a regulator of lipid and folic acid metabolism in breast cancer cells through the systematic integration of the molecular spectrum [5]. Song *et al*. found that miR-22 regulated breast cancer stemness and metastasis via TET (ten eleven translocation) family

✉ Correspondence to: Yingjie Song. Email: 1601340054@qq.com
*Supported by the Joint Fund Project of Hubei Provincial Health Committee (No. WJ2019H510) and the Natural Science Foundation of Inner Mongolia Autonomous Region (No. 2015MS0877), China.

dependent chromatin remodeling [6]. Experiments and clinical studies show that miR-22 promoted epithelial mesenchymal transition and tumor invasion and metastasis [7]. Hence, exploring miR-22 target genes is important for targeted therapy of breast cancer.

In this study, the differentially expressed genes (DEGs) from microarray data in the GEO database were identified between knockout or duplex miR-22 samples in human breast cancer MCF-7 cell lines. Online tools including mirDIP 4.1 software were used to predict miR-22 target genes. We aimed to explore certain differential genes, and miRNAs may be potential targets for predicting and treating miR-22 expressed breast cancer.

Materials and methods

Microarray data filtering eligible data set

We searched and downloaded microarray data GSE17508 (including three miR-22 knockout samples and three controls) from the GEO database (www.ncbi.nlm.nih.gov/geo/). The datasets used human breast cancer MCF-7 cell lines. Microarray data were used to analyze the eligible data set and identify the gene expression patterns.

DEGs' screening

DEGs were independently screened using the GEO2R online tool in the GEO database. In our study, DEGs between miR-22 knockout samples and controls were screened and selected by the cut-off point of adj. *P* value < 0.05 and |log FC|> 0.5.

Functional enrichment and protein-protein interaction (PPI) analysis

The functions of DEGs were further analyzed using the online tool Metascape (http://metascape.org/gp/index.html#/main/step1). The terms with *P*-value < 0.01, minimum count of 3, and enrichment factor > 1.5 were collected and grouped into clusters based on their membership similarities. Furthermore, we obtained the PPI of differential genes using the Search Tool in the Retrieval of Interacting Genes database (STRING, http://string-db.org/), and one big pairing picture was generated. In this study, DEGs with a confidence score of > 0.4 were selected to construct the PPI network.

Established miRNA-DEG network

The DEGs and has-miR-22 were uploaded to the miRNet 2.0 database (www.mirnet.ca/faces/home.xhtml) to acquire a list of miRNA-DEG pairs, and one big pairing picture was generated. The picture was merged with the PPI picture, and a complete and huge miRNAs-DEG network was generated using the Cytoscape 3.6.1. software [8]. Then, the Molecular Complex Detection (MCODE) plugin was used to module clustering analysis to detect the potential functional modules in the network. In the MCODE process, the cut-off value of degree was set to 2, and the cut-off value of node score was set to 0.2.

Predicted target genes

To obtain potential target genes, hsa-miR-22 and the DEGs were uploaded to the online tools database mirDIP 4.1 (http://ophid.utoronto.ca/mirDIP/index_confirm.jsp). Researchers can predict target genes of miRNA with the help of this online tools database. We only selected the gene symbol whose confidence class is considered very high.

Results

DEG screening between miR-22 knockout samples and controls

An assessment of data normalization and cross-comparability (Table 1), and then DEGs analysis was carried out.

We applied an online tool (http://www.heatmapper.ca/expression/) analysis to discern the differential expression of genes. A total of 40 DEGs were identified in the profiles, including 38 upregulated and 2 downregulated DEGs in breast cancer MCF7 cells (Fig. 1).

Gene ontology and functional enrichment analysis

Pathway and process enrichment analysis was performed using GO Biological Processes, GO Cellular Components, GO Molecular Functions, KEGG Functional Sets, and KEGG Pathway Ontology sources.

Heatmap-selected GO showed a defense response to virus (GO: 0051607) and interferon signaling (R-HSA-913531) being the top two pathways (Fig. 2a). The heatmap-selected GO parent showed a multi-organism process (GO: 0051704), immune system process (GO: 0002376), and signaling (GO: 0023052) as the top three pathways (Fig. 2b).

Construction of miRNA-DEG network analysis

Taking the selected 40 DEGs and has-miR-22 into account, we identified 393 PPI pairs using the STRING database as well as a large network of 284 genes and 751 miRNAs using miRNet. Two big pairing pictures mentioned above were merged using Cytoscape software to obtain a complete miRNA-DEG network in Cytoscape. The MCODE plugin in Cytoscape was used to perform the module clustering of the miRNA-DEG network mentioned above, and then the key functional modules of the network were evaluated.

Two modules were identified and showed certain differential genes (Fig. 3 and 4), such as PARP14 (Poly-

Table 1 Data normalization and cross-comparability

ID	miR-22 knockout			miR-22 duplex		
	GSM436499	GSM436500	GSM436501	GSM436502	GSM436503	GSM436504
1405_i_at	0.671169	2.09665	1.6046	191.673	83.8896	143.034
1555895_at	61.5801	42.5199	44.7025	5.12442	5.76263	5.50906
201641_at	299.389	255.064	267.642	1764.3	1416.32	1435.47
201649_at	189.132	169.375	173.936	2753.45	1971.27	2119.3
202411_at	231.34	254.555	170.62	5884.85	4094.31	4084.7
202446_s_at	746.161	585.523	633.731	4041.83	3528.97	3308.84
202863_at	147.238	136.453	132.913	809.03	896.412	745.35
202869_at	204.487	181.35	139.1	2422.99	1932.6	1909.51
203595_s_at	147.951	142.789	138.859	932.521	756.007	840.578
203964_at	205.023	135.525	185.118	2314.57	1679.31	1899.64
204972_at	26.7859	19.5057	9.41348	1287.85	814.213	1048.49
205660_at	68.5446	40.9388	46.0824	1289.54	794.51	1163.13
208012_x_at	280.53	273.367	246.022	1606.66	1515.08	1492.88
209417_s_at	82.8806	86.9507	91.2375	1103.68	771.531	831.984
209969_s_at	222.731	254.367	227.426	1324.41	1401.89	1295.05
213293_s_at	0.930746	2.59175	1.11313	149.973	79.6659	105.26
214329_x_at	21.8184	19.769	21.3362	229.147	157.749	167.125
218400_at	302.368	241.323	240.435	2008.15	1504.66	1623.47
218543_s_at	111.989	95.0014	115.551	1077.07	792.782	864.665
218943_s_at	112.554	104.418	127.537	1023.37	822.671	978.666
218986_s_at	26.0235	35.4281	33.9382	576.956	616.786	542.616
219209_at	69.2003	90.0639	64.2059	637.783	646.895	599.096
219211_at	76.3393	65.6202	45.6518	1113.93	893.935	944.552
219863_at	25.5529	25.9988	30.3203	435.618	270.39	445.39
223220_s_at	117.359	100.177	95.087	1108.72	949.663	830.072
224701_at	69.3057	45.6731	66.2717	747.816	526.984	605.155
226702_at	32.3924	43.4558	35.3575	1516.01	1143.79	1376.78
226757_at	28.6421	16.0251	25.8946	1015.75	864.11	1089.9
228152_s_at	29.6723	24.0253	19.3472	429.784	281.131	393.66
228230_at	37.0041	35.0767	46.5385	292.262	209.932	258.172
228531_at	28.14	25.3732	22.6905	413.528	307.086	378.416
228617_at	21.1541	14.1225	20.0689	1040.53	693.619	801.349
229450_at	23.0721	30.9242	40.6104	3772.66	2927.48	3140.0
234987_at	63.3734	40.6403	54.9785	494.176	445.555	481.173
235684_s_at	16.245	15.971	19.0761	3.1126	3.62215	3.08859
236156_at	1.32972	0.587631	1.35526	25.3288	24.3493	33.6445
38241_at	1.0171	0.639231	0.931997	28.4628	16.7508	19.8981
53720_at	51.1619	48.4386	40.9648	367.055	295.637	312.053

adenosine diphosphate-ribose polymerase 14), SAMHD1 (Sterile alpha motif and histidine/aspartic acid domain-containing protein 1), CCL5 (C-C chemokine ligand 5), and TNFSF10 (Tumor Necrosis Factor superfamily 10).

The four DEGs, PARP14, SAMHD1, CCL5, and TNFSF10, in these modules also occurred in the GO terms enriched above, which were associated with multi-organism immune processes and signaling pathways. PARP14 and SAMHD1 targeted miRNAs, such as miR-21-3p, miR-138-5p, miR-130a-3p, miR-155-5p, miR-452-5p, and miR-124-3p. We also observed a key hub, miR-146a-5p, that interacted with CCL5 and TNFSF10.

Predicted target genes

We identified that upregulated OAS1 and downregulated DEGs may be potential target genes of miR-22-3p. These were predicted using the online tools database mirDIP 4.1, with the confidence class selected as very high (Table 2).

Discussion

Breast cancer is the most common malignant tumor in women worldwide, and metastasis is the main cause of death. Increasing evidence indicates that miR-22 is

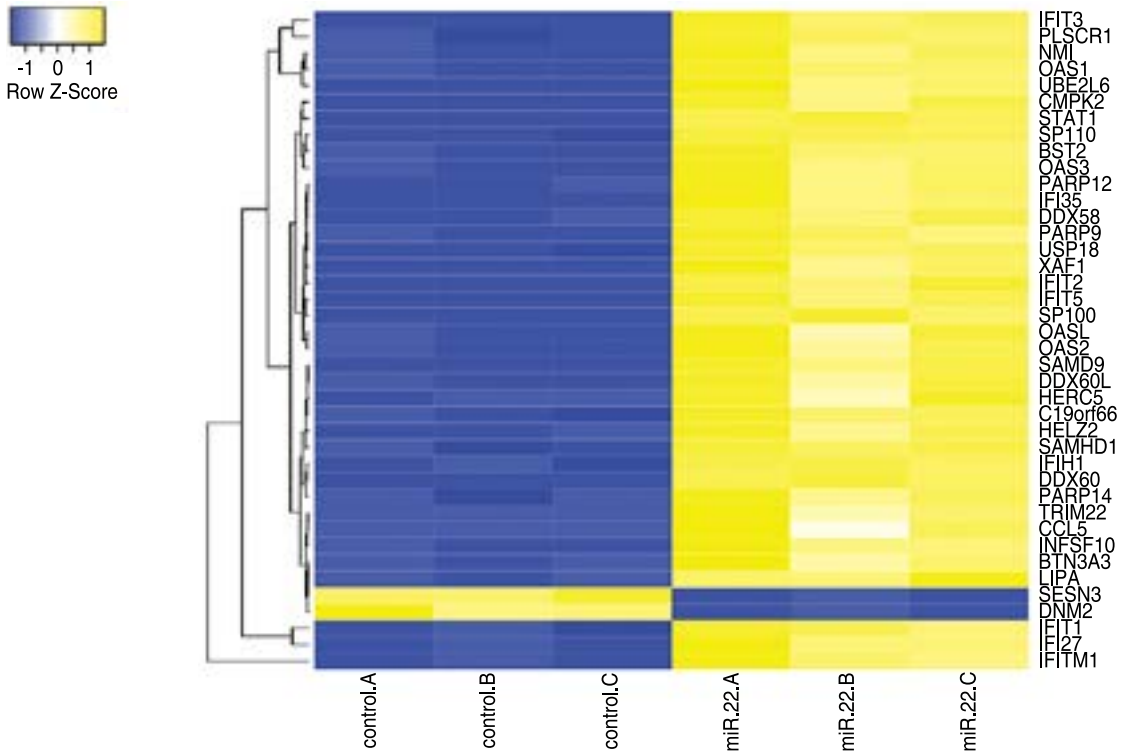

Fig. 1 DEGs expression profiling in breast cancer MCF7 cells.

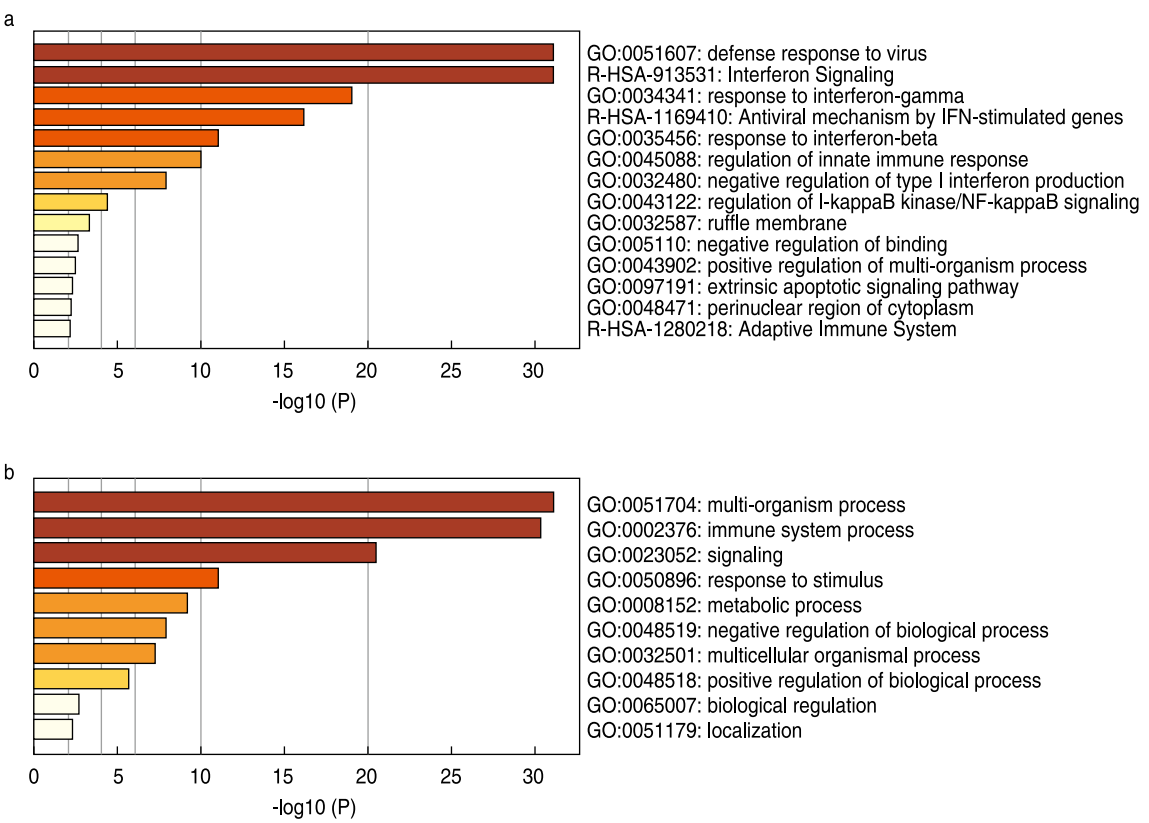

Fig. 2 Distribution of gene ontology terms for the DEGs. (a) heatmap-selected GO; (b) heatmap-selected GO parent

Table 2 The predicted target genes of hsa-miR-22-3p

Gene	Uniprot	MicroRNA	Rank	Source	Confidence score	Confidence class
OAS1	P00973	hsa-miR-22-3p	0.2208986	microrna.org	0.1054261	Very High
SESN3	P58005	hsa-miR-22-3p	0.090834	miRcode	0.1209991	Very High

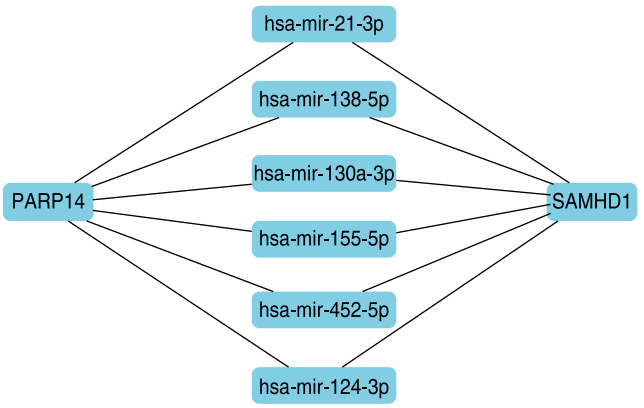

Fig. 3 An outstanding hub of DEGs, PARP14 and SAMHD1, interacting with their miRNAs.

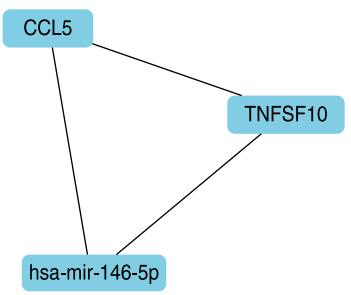

Fig. 4 Another outstanding hub of DEGs, CCL5 and TNFSF10, interacting with miR-146a-5p.

upregulated in breast tumors compared to its marginal non-tumor counterparts, which are associated with poor overall survival [9–10]. However, the detailed potential target genes of miR-22 in breast cancer are still unclear. Thus, it is very important to study the molecular mechanism of breast cancer metastasis and progression to develop a therapeutic strategy for breast cancer patients.

As the miRNA-DEG network analysis showed, expression of miR-22 was closely related to the PARP14, SAMHD1, CCL5, and TNFSF10 genes in breast cancer MCF7 cells. PARP14 uses nicotinamide adenine dinucleotide (NAD+) as a metabolic substrate to modify the target protein by single ADP ribosylation, which is involved in cellular reactions and signaling pathways in the immune system. SAMHD1 is considered an intrinsic viral limiting factor that inhibits the process of viral infection, including retrovirus replication, packaging, and transmission. Several studies have demonstrated that PARP14 and SAMHD1 are associated with the development of multiple types of cancer, such as lung, colon, breast, myeloma, and pancreatic cancer among others [11–12]. In the present study, through MCODE analysis, we found that PARP14 and SAMHD1 interact with some miRNAs, such as miR-21-3p, miR-138-5p, miR-130a-3p, miR-155-5p, miR-452-5p, and miR-124-3p. We found that miR-21-3p may be the core of the module because the survival analysis of miRNAs shows that the overexpression of miR-21 can significantly reduce the overall survival rate of breast cancer patients [13].

CCL5 is secreted by breast cancer cells and rarely expressed in epithelial cells of normal ducts or benign breast masses [14]. Increasing studies showed that the CCL5/CCR5 axis is involved in tumor growth, migration and angiogenesis in different tumor types, including breast cancer. In addition, CCL5/CCR5 is closely related to the recruitment of tumor-associated immune-suppressive cells and promoting the construction of the tumor microenvironment [14]. We found that CCL5 may also be the core of the module, through MCODE analysis, CCL5 interacts with TNFSF10 and miR-146a-5p. Moreover, in order to predict the target gene of miR-22 in DEGs, through the use of the online tools database mirDIP 4.1, we found that upregulated OAS1 and downregulated SESN3 may be potential target genes of miR-22-3p.

Conclusion

In summary, certain DEGs and miRNAs may be potential targets and biomarkers for predicting and treating miR-22 expressed breast cancer. However, further studies such as disease models and PCR experiments are necessary to verify these findings.

Conflicts of interest

The authors indicated no potential conflicts of interest.

References

1. Kim VN. MicroRNA biogenesis: coordinated cropping and dicing. Nat Rev Mol Cell Biol, 2005, 6: 376–385.
2. Denli AM, Tops BB, Plasterk RH, *et al*. Processing of primary microRNAs by the microprocessor complex. Nature, 2004, 432: 231–235.
3. He K, Li WX, Guan DG, *et al*. Regulatory network reconstruction of five essential microRNAs for survival analysis in breast cancer

by integrating miRNA and Mrna expression datasets. Funct Integr Genomics, 2019, 19: 645–658.
4. Meltzer PS. Cancer genomics: small RNAs with big impacts. Nature 2005, 435: 745–746.
5. Koufaris C, Valbuena GN, Pomyen Y, *et al*. Systematic integration of molecular profiles identifies miR-22 as a regulator of lipid and folate metabolism in breast cancer cells. Oncogene, 2016, 35: 2766–2776.
6. Song SJ, Ito K, Ala U, *et al*. The oncogenic microRNA miR-22 targets the TET2 tumor suppressor to promote hematopoietic stem cell self-renewal and transformation. Cell Stem Cell, 2013; 13: 87–101.
7. Farazi TA, Ten HJJ, Brown M, *et al*. Identification of distinct miRNA target regulation between breast cancer molecular subtypes using AGO2-PAR-CLIP and patient datasets. Genome Biol, 2014, 15: R9.
8. Saito R, Smoot ME, Ono K, *et al*. A travel guide to cytoscape plugins. J Nat Methods, 2012, 9: 1069–1076.
9. Damavandi Z, Torkashvand S, Vasei M, *et al*. Aberrant expression of breast development-related MicroRNAs, miR-22, miR-132, and miR-212, in breast tumor tissues. J Breast Cancer, 2016, 19: 148–155.
10. Pandey AK, Zhang Y , Zhang S, *et al*. TIP60-miR-22 axis as a prognostic marker of breast cancer progression. Oncotarget, 2015, 6: 41290–41306.
11. Qin W, Wu HJ, Cao LQ, *et al*. Research progress on PARP14 as a drug target. Front Pharmacol, 2019, 10: 172.
12. Zhang Z, Zheng L, Yu Y, *et al*. Involvement of SAMHD1 In dNTP homeostasis and the maIntenance of genomic Integrity and oncotherapy (Review). Int J Oncol, 2020, 56: 879-888.
13. He K, Li WX, Guan DG, *et al*. Regulatory network reconstruction of five essential microRNAs for survival analysis in breast cancer by integrating miRNA and mRNA expression datasets. Funct Integr Genomics, 2019, 19: 645–658.
14. Aldinucci D, Borghese C, Casagrande N. The CCL5/CCR5 axis in cancer progression. Cancers (Basel), 2020; 12: 1–30.

DOI 10.1007/s10330-020-0458-8
Cite this article as: Fan T, Wang CQ, Zhang K, *et al*. Differentially expressed genes analysis and target genes prediction of miR-22 in breast cancer. Oncol Transl Med, 2021, 7: 59–64.