

Identification of Key Genes and Pathways in Breast Cancer Based on Bioinformatic Analysis

Chi Pan

Jiangsu Taizhou People's Hospital

Qingtao Ni (✉ nqt1990@126.com)

Jiangsu Taizhou People's Hospital <https://orcid.org/0000-0003-0263-5993>

Primary research

Keywords: breastcancer, keygene, miRNA, pathways, bioinformatics

Posted Date: July 1st, 2020

DOI: <https://doi.org/10.21203/rs.3.rs-38610/v1>

License:  This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

**Identification of key genes and pathways in breast cancer based on
bioinformatic analysis**

Chi Pan , QingTao Ni^{2*}

¹Department of General Surgery, Jiangsu Taizhou People's Hospital, Taizhou, China

²Department of Oncology, Jiangsu Taizhou People's Hospital, Taizhou, China

*Correspondence to:

Qingtao Ni, MD

Department of Oncology

Jiangsu Taizhou People's Hospital

Hailing South Road 399

225300 Taizhou, China

Phone: +86 18852656280

Email: nqt1990@126.com

ABSTRACT

Breast cancer(BC) is the most frequent cancer type in women. However, the pathogenesis of BC is still not well understood. Thus, we aim to explore key genes and pathways in relation to BC. We used the Gene Expression Omnibus (GEO) database to identify the differently expression of genes in the carcinogenesis and progression of BC. Then, bioinformatics analysis was performed to determine the key genes targets and pathways associated with BC. The gene expression profile of the hub genes in human tumor was displayed using GEPIA. Finally, the expression of hub genes, correlation between genes and miRNA were screened via the miRDB, MirTarBase and DIANA Tools. We screened 159 downregulated genes and 55 upregulated genes in BC patients among the 4 datasets. The enriched functions of the DEGs involved in cell proliferation, positive regulation of Akt signaling and extracellular exosome, PPAR signaling pathway and AMPK signaling pathway. 3 hub genes were screened out by construction of PPI network. MELK were found to be upregulated, and CLU and NTRK2 were downregulated. Further verification showed that MELK displayed higher levels in almost all tumors. We found correlation between these hub genes and the miRNAs. All in all, three key genes closely related to the incidence of BC were identified, and the results could provide new potential molecular targets for the diagnosis and treatment of BC. In particular, MELK is regulated by multiple miRNAs and participate in the development of BC.

Keywords: breast cancer, key gene, miRNA, pathways, bioinformatics

Running title: key genes in breast cancer

Introduction

Breast cancer (BC) is the most common cancer type and leading cause of cancer deaths in women. Regarding estimates in the United States, around 279,100 new diagnoses of BC are expected in women for 2020, and around 42,690 women are likely to die from this disease(Siegel et al., 2020). The estimated data has increased than in past year(Siegel et al., 2020). The prognosis of BC patients has improved with the improvement of early diagnosis, surgical techniques, neoadjuvant chemotherapy and molecular target gene therapy. A variety of studies has demonstrated that genetic backgrounds are involved in the development of BC. Up to 10% of BC patients are caused by a mutation of genes(Childers et al., 2017). In addition to BRCA1/2, the expression of protein kinase R (PKR) and phosphatidylinositol 4-kinase 2-alpha (PI4K2A) gene play an important role in the prognosis of BC(Pataer et al., 2020). Inhibiting the activity of fatty acid transport protein 1 (FATP1) can inhibit the ability of BC cells to obtain fatty acids and lead to the decrease of cell viability(Mendes et al., 2019). A vitro experiments show that combining the highly specific HER3 (human epidermal growth factor Receptor3) with siRNA can effectively inhibit the proliferation of HER3 positive BC cells(Nachreiner et al., 2019).However, BC is still the main cause of death in female cancer patients because of the lack of effective methods for early diagnosis and effective treatment strategies in the late stage. Therefore, it is crucial to figure out more potential molecular targets to provide theoretical basis for molecular targeted treatment of BC.

Due to the development of microarray technology and sequencing, bioinformatic analysis has been widely used to seek the differentially expressed genes (DEGs) and related signaling pathways. In order to obtain DEGs in BC, we selected four microarray datasets from Gene Expression Omnibus (GEO). Subsequently, GO, KEGG pathway enrichment and PPI network analysis were performed to help us understand the molecular mechanisms underlying carcinogenesis and found therapeutic targets. The workflow of the detailed analysis is summarized in Fig. 1.

Methods

Analysis of Gene microarray. In order to examine the potentially important genes in BC, four gene expression profile of BC datasets [GSE65194 , GSE15852, GSE42568 and GSE10810] were selected by retrieving Gene Expression Omnibus (GEO database) (<http://www.ncbi.nlm.nih.gov/geo>). The GSE65194 dataset included 152 BC samples and 10 healthy samples. GSE15852 included 43 BC samples and 43 healthy samples. GSE42568 included 104 BC samples and 16 healthy samples. GSE10810 included 31 BC samples and 27 control healthy samples.

Identification of DEGs. We performed analysis of DEGs from BC samples compared to noncancerous samples using GEO2R (<http://www.ncbi.nlm.nih.gov/geo/geo2r>). Adjusted P value < 0.01 and \log_2FC (fold change) >1 were considered as statistical significance.

GO and KEGG pathway analysis of DEGs. To determine the functional categories of genes, the database (DAVID Gene Functional Classification Tool, <http://david.abcc.ncifcrf.gov/>) was used. Gene Ontology (GO) enrichment analysis was performed using DAVID online database. P<0.05 was considered as statistical significance. Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway analysis of DEGs was retrieved from KEGG web server

(<http://www.genome.jp/kegg/>).

PPI network analysis. Search Tool for the Retrieval of Interacting Genes (STRING database, V10.5; <http://string.gdb.org/>) is an online database. We performed PPI network of DEGs using STRING database, interactions with a combined score ≥ 0.4 were considered as statistical significance. Cytoscape is a software platform for integrated models interaction networks. PPI networks were drawn using Cytoscape and the most significant module in the PPI networks was identified to be significant with a Molecular Complex Detection (MCODE) score >5 , degree cut-off=2, node score cut-off=0.2, Max depth=100 and k-score=2. Genes with most significant module are considered as hub genes. Then, the GO enrichment and KEGG pathway analysis for these hub genes were performed using DAVID.

Analysis of the hub genes. The cellular component and biological processes analysis of hub genes were performed using Biological Networks Gene Oncology tool (BiNGO), a Cytoscape v3.0.3 plugin. Hierarchical clustering of hub genes was made by using UCSC (University of California, Santa Cruz) Cancer Genomics Browser (<http://genome-cancer.ucsc.edu>). The overall survival (OS) and disease free survival curves of hub genes were plotted using the cBioPortal. Furthermore, to verify the expression level of the hub genes in different kinds of tumors, We used the GEPIA (gene expression profiling interactive analysis) database (<http://gepia.cancer-pku.cn/>).

miRNA-Gene-Pathway Network Construction. In order to obtain potential miRNA targeting, we applied online prediction tools utilizing miRDB (<http://mirdb.org/>), MirTarBase (<http://mirtarbase.mbc.nctu.edu.tw>), and DIANA Tools (<http://diana.imis.athena-innovation.gr/DianaTools/>). Subsequently, miRNA-gene-pathway network was constructed based on Venn Diagrams (<http://bioinformatics.psb.ugent.be/webtools/Venn/>).

Results

Identification of DEGs in BC. After standardization of the microarray results, DEGs (1,672 in GSE65194, 332 in GSE42568, 4,014 in GSE15852 and 6,018 in GSE10810) were identified. The overlap among the 4 datasets contained 214 genes as shown in the Venn diagram (Fig. 2A), consisting of 159 downregulated genes and 55 upregulated genes between BC samples and normal samples.

GO and KEGG analysis. GO analysis results showed that changes of DEGs were significantly enriched in cell proliferation, positive regulation of protein kinase B (Akt) signaling and extracellular exosome (Table 1). KEGG pathway analysis revealed that the downregulated DEGs were mainly enriched in PPAR (peroxisome proliferators-activated receptor) signaling pathway and AMPK (Adenosine 5'-monophosphate-activated protein kinase) signaling pathway.

PPI network analysis. We performed the PPI network analysis for DEGs (Fig. 2B). The tool Cytoscape was used to obtain the most significant module (Fig. 2C). According to the results we found the functional analyses of genes involved in this module were mainly enriched in exosome (Table 2).

Hub gene analysis. The full name and function for these hub genes are presented in Table 3.

The biological process analysis of the hub genes is shown in Fig. 3A. Hierarchical clustering showed that the hub genes could basically differentiate the BC samples from the normal samples (Fig. 3B). Then, the OS analysis of the hub genes was performed using Kaplan-Meier curve. BC patients with MELK (maternal embryonic leucine zipper kinase) alteration showed better overall survival. BC patients with CLU (Clusterin) and NTRK2 (neurotrophic tyrosine kinase receptor 2) alteration showed worse OS (Fig. 4A). Nonetheless, BC patients with CLU and NTRK2 alteration showed worse disease-free survival (Fig. 4B). However, this observation was not statistically significant (Fig. 4). The gene expression profile of MELK, CLU and NTRK2 in human tumor was displayed using GEPIA. We found that MELK displayed higher levels in almost all tumors except LAML(Acute Myeloid Leukemia) as compared with the matched normal tissues (Fig. 5A B). CLU displayed higher levels in GBM(Glioblastoma multiforme), LGG(Brain Lower Grade Glioma) and THCA(Thyroid carcinoma) as compared with the matched normal tissues. Similarly, NTRK2 displayed higher levels in LGG and THCA as compared with the matched normal tissues.

Identification of novel miRNAs. Common targets of three online databases such as miRDB, MirTarBase and DIANA Tools were filtered out by getting intersection elements, as demonstrated in Venn Diagram (Fig. 6). These three databases contain computationally predicted miRNA-target interactions. Only one miRNA (hsa-miR-22-3p) appeared to have a significant nominal p-value in three databases for NTRK2 (Fig. 6A). For the CLU gene, two miRNAs (hsa-miR-6817-3p, hsa-miR-7110-3p) were curated from three databases (Fig. 6B). For MELK, miRTarBase indicates 41 distinct target genes and DIANA Tools indicates 30, with 16 targets in common between both approaches (hsa-miR-4645-3p, hsa-miR-497-3p, hsa-miR-485-3p, hsa-miR-548x-3p, hsa-miR-205-3p, hsa-miR-4803, hsa-miR-466, hsa-miR-548aj-3p, hsa-miR-889-3p, hsa-miR-4311, hsa-miR-520d-5p, hsa-miR-505-3p, hsa-miR-212-3p, hsa-miR-4719, hsa-miR-524-5p, hsa-miR-132-3p) (Fig. 6C).

Discussion

BC is the most common malignant tumor in female (Siegel et al., 2020). BC exist familial clustering individual and race differences, which suggest that gene mutation plays an important role in carcinogenesis of BC. Therefore, targeted therapy may have great implications for the treatment of BC. However, how to discover the potential gene mutation is still a big challenge for targeted therapy(Gu et al., 2016).Microarray technology has rapidly evolved during the past decade, which help us to explore the genetic alterations in BC.

In this present study, a total of 214 DEGs were identified from the 4 datasets in GEO database, including 159 downregulated genes and 55 upregulated genes. GO and KEGG analysis were performed to explore interactions among the DEGs. GO enrichment analysis revealed that changes in the most significant modules were mainly enriched in cell proliferation, positive regulation of Akt signaling and extracellular exosome, while changes in KEGG were mainly enriched in PPAR signaling pathway and AMPK signaling pathway. As we all know that cell proliferation plays an important role in the carcinogenesis of tumors. Previous research has indicated regulation of the Akt signaling pathway may inhibit the growth and aggressiveness of BC cells(Wang et al., 2018). In our previous research, we found that exosomes may be

selectively enriched with miRNA in BC(Ni et al., 2018). Recent study has shown that genes involved in PPAR signaling pathway can be as a potential biomarker for the early diagnosis of BC(Sultan et al., 2019). The reduction in the expression levels of PPAR potentially prevent the tumorigenesis of BC(Pirouzpanah et al., 2019). AMPK activation-dependent mechanisms may participate in the regulation of growth of BC cells(Pham et al., 2020). Taken together, all these observations are consistent with our results.

We screened 3 DEGs (MELK, CLU and NTRK2) as hub genes. MELK is a member of the snf1/AMPK family of protein serine/threonine kinases and highly associated with accelerated proliferation of cancer stem cells (CSC) in various organs(Ganguly et al., 2014). It has been considered to be a potential therapeutic target in cancer since 2005(McDonald and Graves, 2020). Previous research has showed that MELK has the ability of inhibition of different cell cycle occurred in BC cells(Li et al., 2018). The expression level of CLU were appreciably higher in cancer stem cells(Samson et al., 2019). CLU was proven associated with tumor size, lymph node and the clinical stage in triple-negative breast cancer(TNBC)(Zhang et al., 2012). NTRK2 is the gene that encodes for the neurotrophin receptor Tyrosine Kinase receptor B(Pattwell et al., 2020). It may exacerbate symptoms in BC patients receiving chemotherapy. We assessed the expression of MELK, CLU and NTRK2 in relation to overall and disease-free survival. Gene alteration in MELK showed increased in overall and disease free survival. However, in the present study, those observations were not statistically significant. Another bioinformatic analysis found MELK were related to worse OS in BC(Young et al., 2017). CLU might be a predictive factor for recurrence in <T2 stage BC(Yom et al., 2009). Recent study showed that NTRK2 is related to the prognosis of invasive BC(Gao et al., 2019). We speculate that the reason may be that this database focused on differential gene expression and biological function, but not on outcomes. In addition, inadequate length of follow-up in BC patients with longer OS time may be one of the reasons. Oncomine analysis showed that higher levels of MELK, CLU and NTRK2 were associated with tumor grade, indicating vital roles of MELK, CLU and NTRK2 in the carcinogenesis or progression of BC. As further proof of the importance of MELK, CLU and NTRK2, we verified these genes in GEPIA. MELK is an important gene in the carcinogenesis and development of almost all tumors. CLU and NTRK2 in LGG and THCA were abnormal expression, similarly. These results further suggest the importance of these genes in carcinogenesis in tumor.

Further exploration of the pathway of these genes is required. Therefore, we screened in miRDB, MirTarBase and DIANA Tools to find relevant miRNA-gene interactions, these being visualized as networks. Using two databases, we identified 16 miRNA that are associated with MELK, such as miR-205. MiR-205 inhibit the proliferation and promote the apoptosis of BC cells, but its mechanism is not clear(Du YE et al., 2017).Other study also showed that mir-205 can induce PI3K / Akt signaling pathway in BC by acting on HER3(Orang et al., 2019). In three databases, miRDB, MirTarBase and DIANA Tools, we only found miR-22 has relationship with NTRK2. Previous work shows that miR-22 already be considered as a tumor suppressor or promoter in different cancers(Orang et al., 2019). Furthermore, multiple findings suggesting miR-22 have been implicated in non-small-cell lung cancer (NSCLC) through gene regulation(Ding et al., 2020; He et al., 2020). However, the role of miR-22 in BC remains unclear and need further exploration. For the CLU gene, miR-6817 and miR-7110 were

curated from these three databases. MiR-6817 has not yet received experimental attention. At present, studies focusing on miR-7110 are still few in number, only found that miR-7110 was a novel prognostic, diagnostic and therapeutic target for Type I diabetes mellitus(Priyanka et al., 2018). These findings present crucial targets for BC therapy and have significant implications. However, further studies are needed to elucidate the biological function and pathway of these genes in BC.

Conclusions

All in all, the present study was designed to identify DEGs that may be involved in the carcinogenesis or progression of BC. Three key genes closely related to the incidence of BC were identified, and the results could provide new potential molecular targets for the diagnosis and treatment of BC. In particular, MELK is regulated by multiple miRNAs and participate in the development of BC. Therefore, it would be interesting to elucidate the role of MELK and miRNA pathway in BC in further researches.

Abbreviations

ACC	Adrenocortical carcinoma
Akt	protein kinase B
AMPK	Adenosine 5'-monophosphate-activated protein kinase
BC	Breast cancer
BiNGO	Biological Networks Gene Oncology tool
BLCA	Bladder Urothelial Carcinoma
BRCA	Breast invasive carcinoma
CESC	Cervical squamous cell carcinoma and endocervical adenocarcinoma
CHOL	Cholangiocarcinoma
CLU	Clusterin
COAD	Colon adenocarcinoma
CSC	cancer stem cells
DEGs	differentially expressed genes
DLBC	Lymphoid Neoplasm Diffuse Large B-cell Lymphoma
ESCA	Esophageal carcinoma
FATP1	fatty acid transport protein 1
GBM	Glioblastoma multiforme
GEO	Gene Expression Omnibus
GEPIA	gene expression profiling interactive analysis
GO	Gene Ontology
HER3	human epidermal growth factor Receptor3
HNSC	Head and Neck squamous cell carcinoma

KEGG	Kyoto Encyclopedia of Genes and Genomes
KICH	Kidney Chromophobe
KIRC	Kidney renal clear cell carcinoma
KIRP	Kidney renal papillary cell carcinoma
LAML	Acute Myeloid Leukemia
LGG	Brain Lower Grade Glioma
LIHC	Liver hepatocellular carcinoma
LUAD	Lung adenocarcinoma
LUSC	Lung squamous cell carcinoma
MCODE	Molecular Complex Detection
MELK	maternal embryonic leucine zipper kinase
NSCLC	non-small-cell lung cancer
NTRK2	neurotrophic tyrosine kinase receptor 2
OS	overall survival
OV	Ovarian serous cystadenocarcinoma
PAAD	Pancreatic adenocarcinoma
PCPG	Pheochromocytoma and Paraganglioma
PI4K2A	phosphatidylinositol 4-kinase 2-alpha
PKR	expression of protein kinase R
PPAR	peroxisome proliferators-activated receptor
PRAD	Prostate adenocarcinoma
READ	Rectum adenocarcinoma
SARC	Sarcoma
SKCM	Skin Cutaneous Melanoma
STAD	Stomach adenocarcinoma
TGCT	Testicular Germ Cell Tumors
THCA	Thyroid carcinoma
THYM	Thymoma
TNBC	triple-negative breast cancer
UCEC	Uterine Corpus Endometrial Carcinoma
UCS	Uterine Carcinosarcoma
UCSC	University of California, Santa Cruz

Ethics Statement

Not applicable.

Consent to Publish

Not applicable.

Availability of Data and Material

The data that support the findings of this study are available from the corresponding author upon reasonable request.

Competing interests

The authors declare that the research was conducted in the absence of any commercial or financial relationships

Funding

Not Applicable

Authors' Contributions

Q.N. and C.P. contributed equally to this work.

Acknowledgments

Not applicable.

Disclosure

The author reports no conflicts of interest in this work.

References:

- Childers CP, Childers KK, Maggard-Gibbons M, Macinko J. 2017. National Estimates of Genetic Testing in Women With a History of Breast or Ovarian Cancer. *J CLIN ONCOL* 35(34):3800-3806.
- Deng JL, Xu YH, Wang G. 2019. Identification of Potential Crucial Genes and Key Pathways in Breast Cancer Using Bioinformatic Analysis. *FRONT GENET* 10:695.
- Ding S, Huang H, Xu Y, Shen L, Zhong C, Zheng S. 2020. Neuroepithelial Cell Transforming Gene 1 Acts as an Oncogene and Is Mediated by miR-22 in Human Non-Small-Cell Lung Cancer. *BIOMED RES INT* 2020:1648419.
- Du YE, Tu G, Yang G, Li G, Yang D, Lang L, Xi L, Sun K, Chen Y, Shu K, Liao H, Liu M, Hou Y. 2017. MiR-205/YAP1 in Activated Fibroblasts of Breast Tumor Promotes VEGF-independent Angiogenesis through STAT3 Signaling. *THERANOSTICS* 7(16):3972-3988.
- Ganguly R, Hong CS, Smith LG, Kornblum HI, Nakano I. 2014. Maternal embryonic leucine zipper kinase: key kinase for stem cell phenotype in glioma and other cancers. *MOL CANCER THER* 13(6):1393-1398.
- Gao C, Li H, Zhuang J, Zhang H, Wang K, Yang J, Liu C, Liu L, Zhou C, Sun C. 2019. The construction and analysis of ceRNA networks in invasive breast cancer: a study based on The Cancer Genome Atlas. *CANCER MANAG RES* 11:1-11.
- Gu G, Dustin D, Fuqua SA. 2016. Targeted therapy for breast cancer and molecular mechanisms of resistance to treatment. *CURR OPIN PHARMACOL* 31:97-103.
- Guo J, Zhang J, Yang T, Zhang W, Liu M. 2020. MiR-22 suppresses the growth and metastasis of bladder cancer cells by targeting E2F3. *Int J Clin Exp Pathol* 13(3):587-596.
- He W, Zhang Y, Xia S. 2020. LncRNA NNT-AS1 promotes non-small cell lung cancer progression through regulating miR-22-3p/YAP1 axis. *THORAC CANCER* 11(3):549-560.
- Li G, Yang M, Zuo L, Wang MX. 2018. MELK as a potential target to control cell proliferation in triple-negative breast cancer MDA-MB-231 cells. *ONCOL LETT* 15(6):9934-9940.
- McDonald IM, Graves LM. 2020. Enigmatic MELK: The controversy surrounding its complex role in cancer. *J BIOL CHEM*.
- Mendes C, Lopes-Coelho F, Ramos C, Martins F, Santos I, Rodrigues A, Silva F, Andre S, Serpa J. 2019. Unraveling FATP1, regulated by ER-beta, as a targeted breast cancer innovative therapy. *Sci Rep* 9(1):14107.
- Nachreiner I, Hussain AF, Wullner U, Machuy N, Meyer TF, Fischer R, Gattenlohner S, Meinholt-Heerlein I, Barth S, Tur MK. 2019. Elimination of HER3-expressing breast cancer cells using aptamer-siRNA chimeras. *EXP THER MED* 18(4):2401-2412.
- Ni Q, Stevic I, Pan C, Muller V, Oliveira-Ferrer L, Pantel K, Schwarzenbach H. 2018. Different signatures of miR-16, miR-30b and miR-93 in exosomes from breast cancer and DCIS patients. *Sci Rep* 8(1):12974.
- Orang AV, Petersen J, McKinnon RA, Michael MZ. 2019. Micromanaging aerobic respiration and glycolysis in cancer cells. *MOL METAB* 23:98-126.
- Pataer A, Ozpolat B, Shao R, Cashman NR, Plotkin SS, Samuel CE, Lin SH, Kabil NN, Wang J, Majidi M, Fang B, Roth JA, Vaporciyan AA, Wistuba II, Hung MC, Swisher SG. 2020. Therapeutic targeting of the PI4K2A/PKR lysosome network is critical for misfolded protein clearance and survival in cancer

- cells. *ONCOGENE* 39(4):801-813.
- Pattwell SS, Konnick EQ, Liu YJ, Yoda RA, Sekhar LN, Cimino PJ. 2020. Neurotrophic Receptor Tyrosine Kinase 2 (NTRK2) Alterations in Low-Grade Gliomas: Report of a Novel Gene Fusion Partner in a Pilocytic Astrocytoma and Review of the Literature. *Case Rep Pathol* 2020:5903863.
- Pham DV, Raut PK, Pandit M, Chang JH, Katila N, Choi DY, Jeong JH, Park PH. 2020. Globular Adiponectin Inhibits Breast Cancer Cell Growth through Modulation of Inflammasome Activation: Critical Role of Sestrin2 and AMPK Signaling. *Cancers (Basel)* 12(3).
- Pirouzpanah S, Asemani S, Shayanfar A, Baradaran B, Montazeri V. 2019. The effects of Berberis vulgaris consumption on plasma levels of IGF-1, IGFBPs, PPAR-gamma and the expression of angiogenic genes in women with benign breast disease: a randomized controlled clinical trial. *BMC Complement Altern Med* 19(1):324.
- Priyanka P, Panagal M, Sivakumar P, Gopinath V, Senthilkumar S. 2018. Identification, expression, and methylation of miR-7110 and its involvement in type 1 diabetes mellitus. *CANCER GENE THER*:229-234.
- Samson AAS, Park S, Kim S, Min D, Jeon NL, Song JM. 2019. Liposomal co-delivery-based quantitative evaluation of chemosensitivity enhancement in breast cancer stem cells by knockdown of GRP78/CLU. *J LIPOSOME RES* 29(1):44-52.
- Siegel RL, Miller KD, Jemal A. 2019. Cancer statistics, 2019. *CA Cancer J Clin* 69(1):7-34.
- Siegel RL, Miller KD, Jemal A. 2020. Cancer statistics, 2020. *CA Cancer J Clin* 70(1):7-30.
- Sultan G, Zubair S, Tayubi IA, Dahms HU, Madar IH. 2019. Towards the early detection of ductal carcinoma (a common type of breast cancer) using biomarkers linked to the PPAR(gamma) signaling pathway. *Bioinformation* 15(11):799-805.
- Wang L, Zhou S, Xie J, Gao H, Wang F, Zhou J, Wang Y, Wang H. 2017. BRCA1 germline mutations dominate familial breast cancer patients in Henan China. *J THORAC DIS* 9(12):5295-5299.
- Wang X, Xiong W, Tang Y. 2018. Tunicamycin suppresses breast cancer cell growth and metastasis via regulation of the protein kinase B/nuclear factor-kappaB signaling pathway. *ONCOL LETT* 15(4):4137-4142.
- Yom CK, Woo HY, Min SY, Kang SY, Kim HS. 2009. Clusterin overexpression and relapse-free survival in breast cancer. *ANTICANCER RES* 29(10):3909-3912.
- Young EE, Kelly DL, Shim I, Baumbauer KM, Starkweather A, Lyon DE. 2017. Variations in COMT and NTRK2 Influence Symptom Burden in Women Undergoing Breast Cancer Treatment. *BIOL RES NURS* 19(3):318-328.
- Zhang D, Zhang D, Sun B, Sun B, Zhao X, Zhao X, Cui Y, Cui Y, Xu S, Xu S, Dong X, Dong X, Zhao J, Zhao J, Meng J, Meng J, Jia X, Jia X, Chi J, Chi J. 2012. Secreted CLU is associated with the initiation of triple-negative breast cancer. *CANCER BIOL THER* 13(5):321-329.

Figure 1 Workflow of the analysis process

Figure 2. Venn diagram and module analysis of PPI network. (A) DEGs were selected in 4 datasets, with an overlap of 214 genes. (B) The PPI network of DEGs was constructed using

Cytoscape. (C) The most significant module was obtained from PPI network with 20 nodes and 80 edges. Red represents upregulated genes and blue represents downregulated genes.

Figure 3. Hub genes analysis. (A)The biological process analysis of hub genes was constructed using BiNGO. The color depth of nodes refers to the corrected P-value of ontologies. The size of nodes refers to the numbers of genes that are involved in the ontologies. (B)The heat map of hub genes between BC samples and normal samples. Red colour represents upregulation of genes and green colour represents downregulation of genes.

Figure 4. (A) Overall survival and (B) disease-free survival analyses of hub genes were performed using cBioPortal online platform.

Figure 5. (A)Interactive Bodymap. The median expression of tumor and normal samples in bodymap. (B) The gene expression profile across all tumor samples and paired normal tissues (Bar plot). The height of bar represents the median expression of certain tumor type or normal tissue.

Figure 6. Venn diagram showing the intersection of gene targets for each hub gene from miRDB, MirTarBase and DIANA Tools. (A) The 3 datasets showed an overlap of 1 miRNA for NTRK2. (B) The 3 datasets showed an overlap of 2 miRNA for CLU. (C) The 2 datasets showed an overlap of 16 miRNA for MELK.

Figures

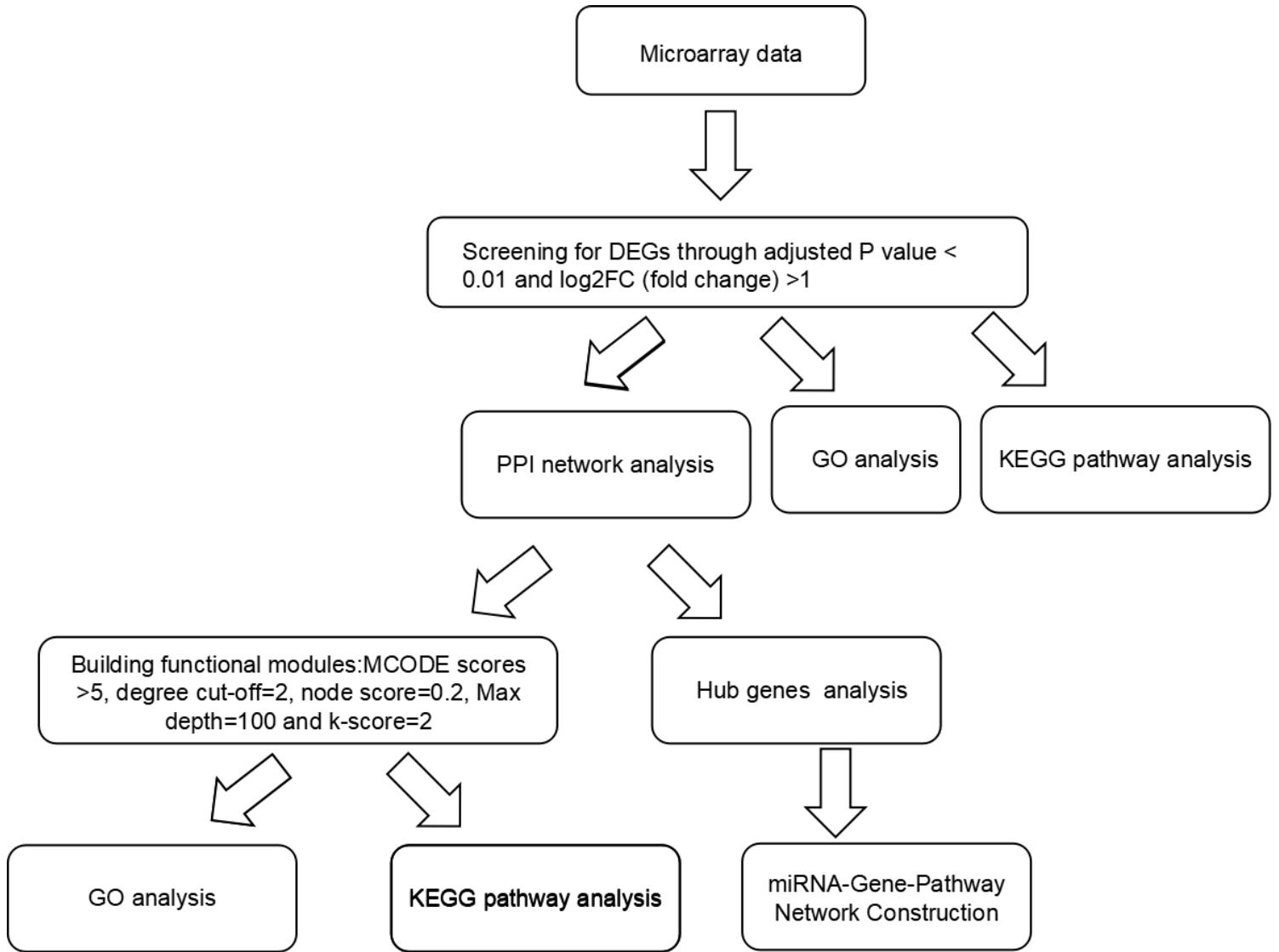


Figure 1

Workflow of the analysis process

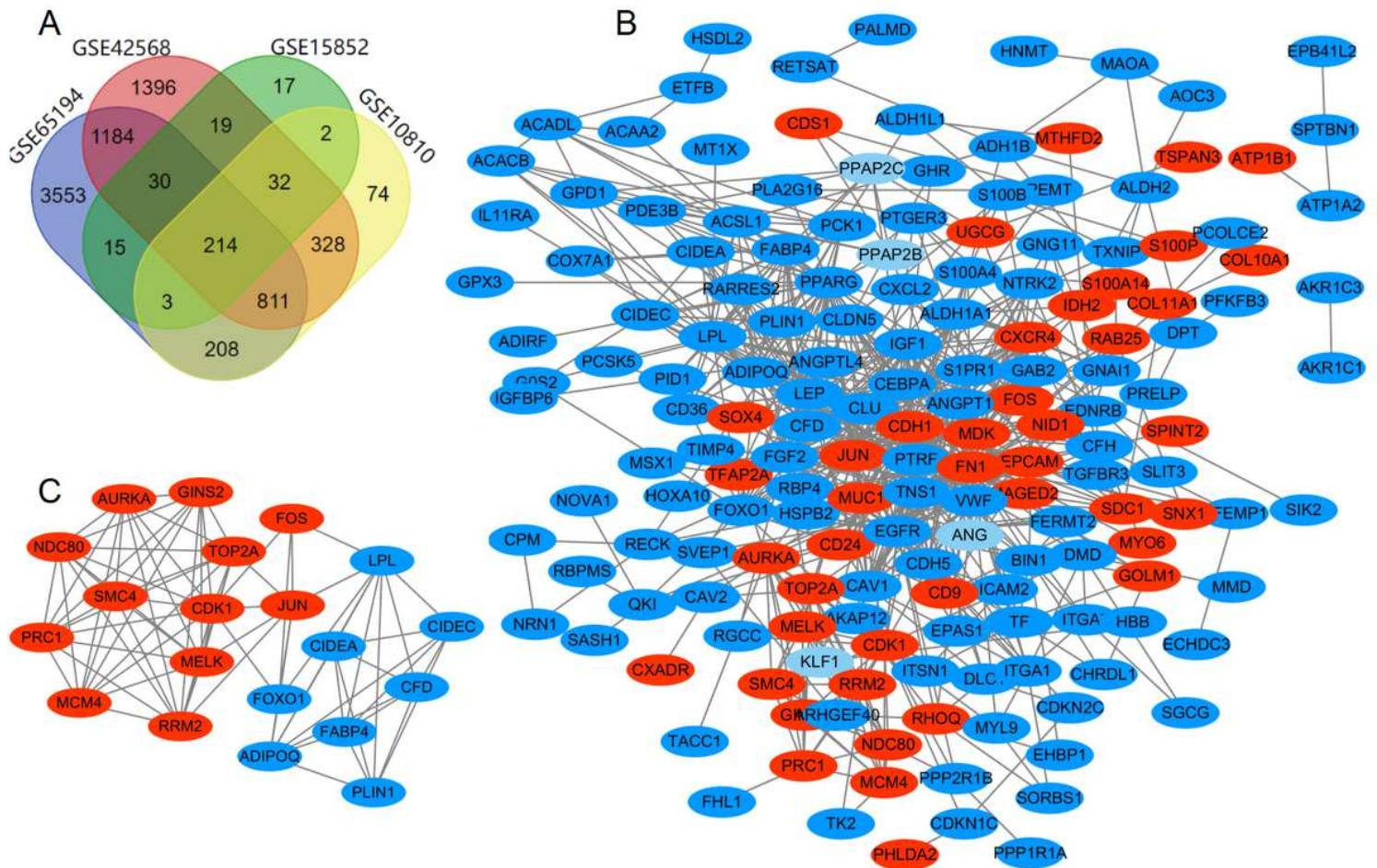
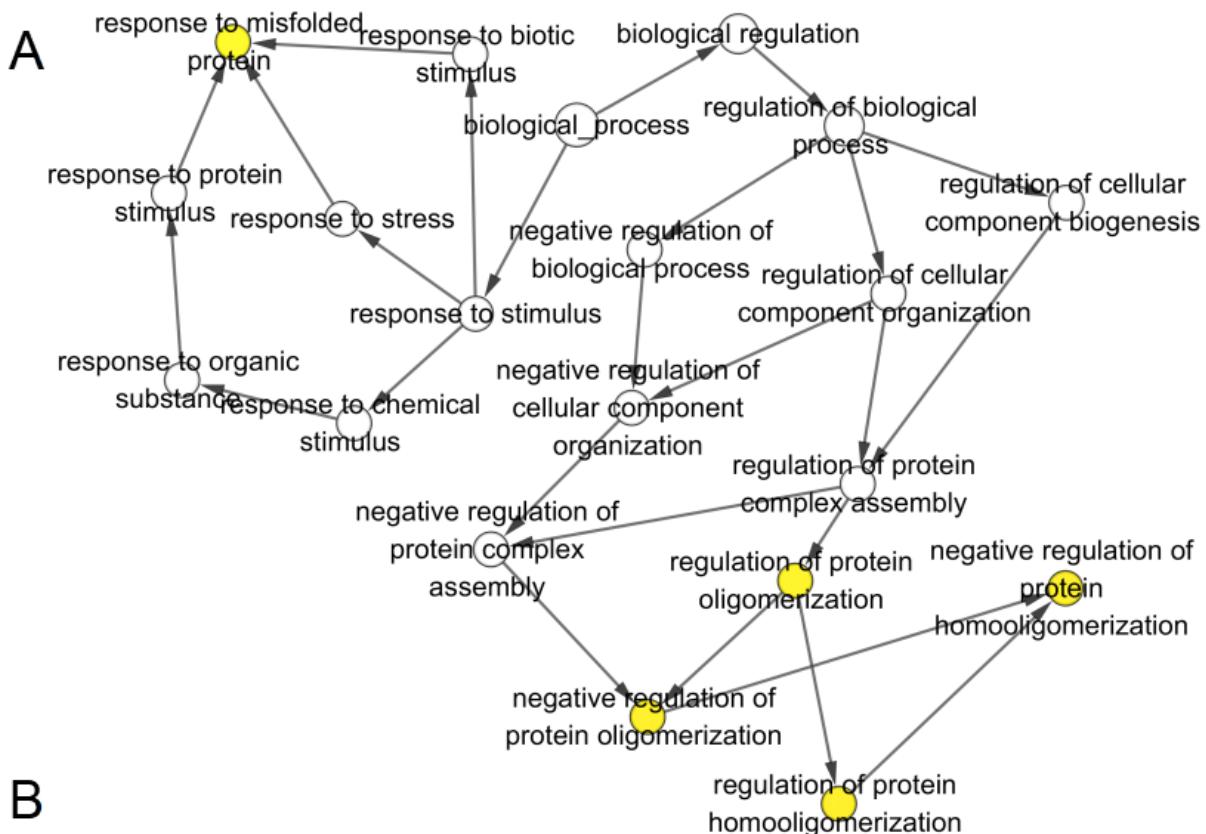


Figure 2

Venn diagram and module analysis of PPI network. (A) DEGs were selected in 4 datasets, with an overlap of 214 genes. (B) The PPI network of DEGs was constructed using Cytoscape. (C) The most significant module was obtained from PPI network with 20 nodes and 80 edges. Red represents upregulated genes and blue represents downregulated genes.



B

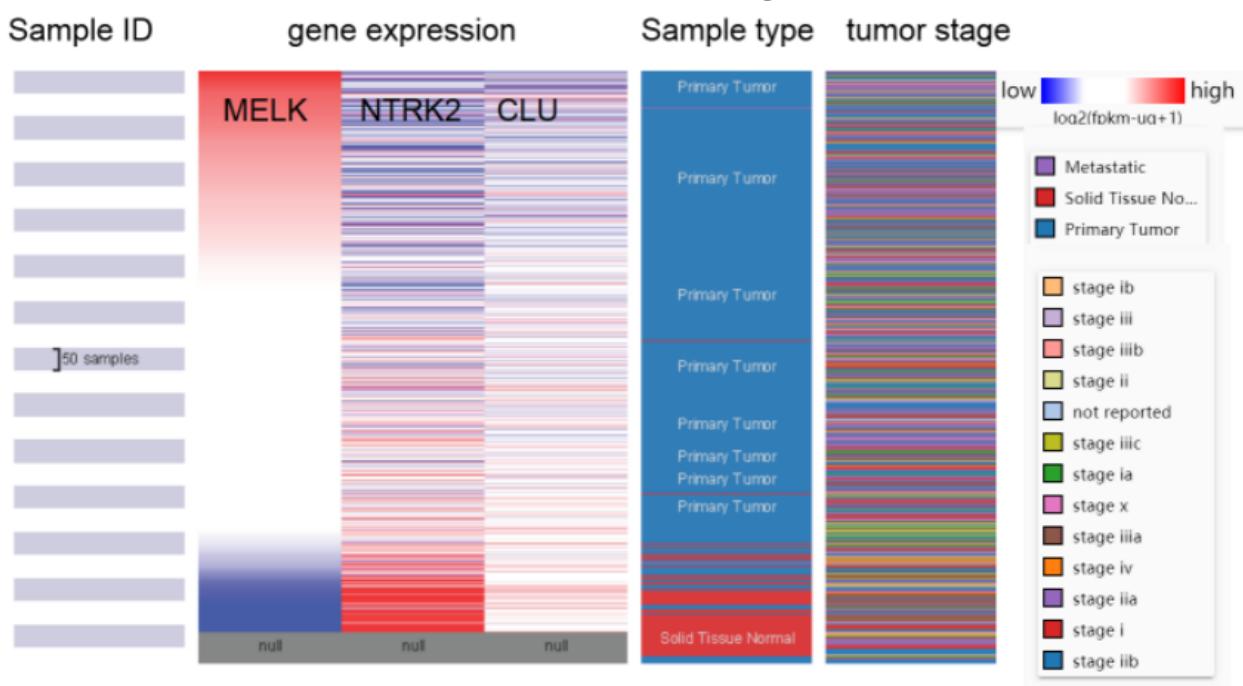


Figure 3

Hub genes analysis. (A)The biological process analysis of hub genes was constructed using BiNGO. The color depth of nodes refers to the corrected P-value of ontologies. The size of nodes refers to the numbers of genes that are involved in the ontologies. (B)The heat map of hub genes between BC samples and normal samples. Red colour represents upregulation of genes and green colour represents downregulation of genes.

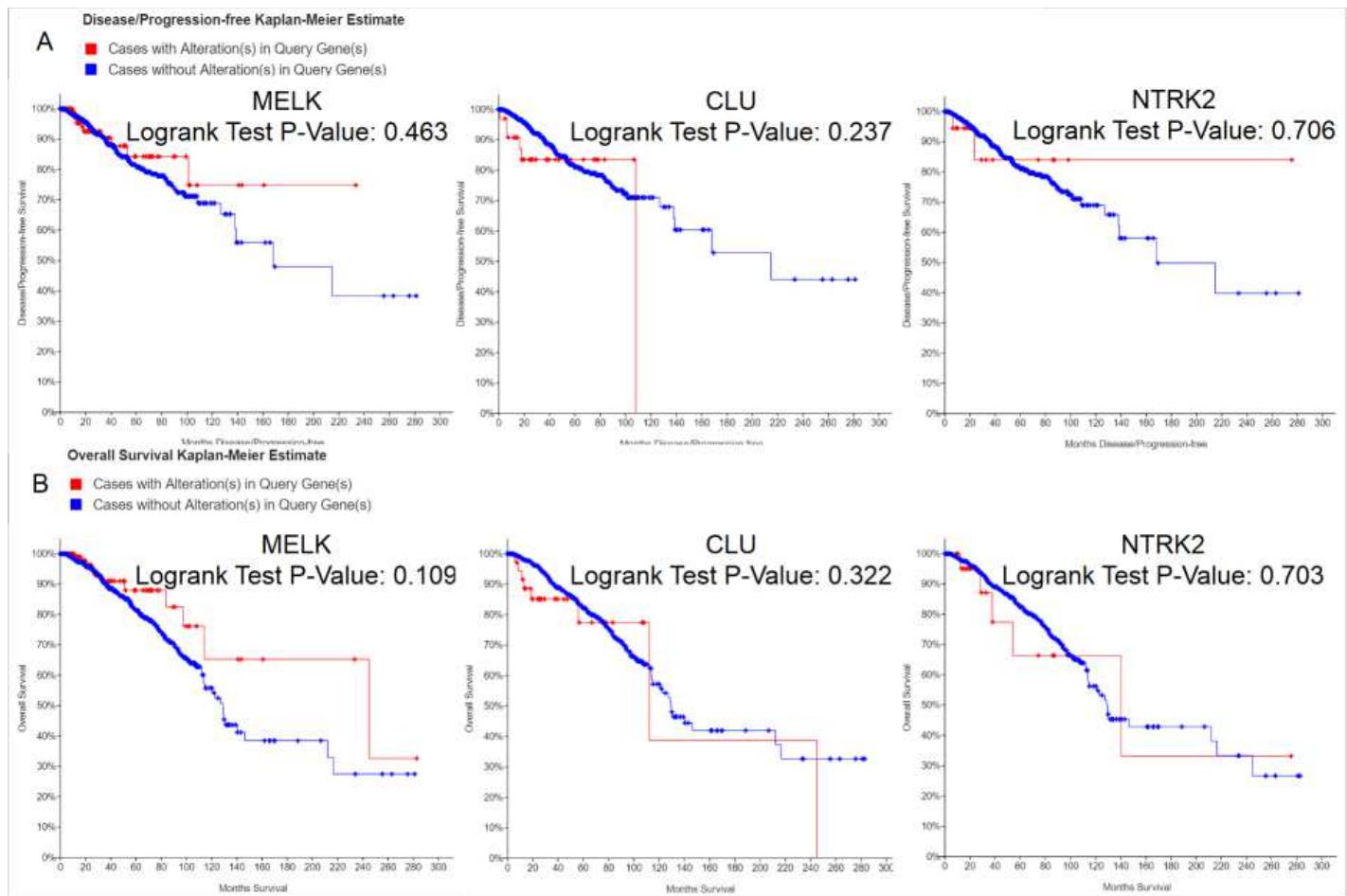


Figure 4

(A) Overall survival and (B) disease-free survival analyses of hub genes were performed using cBioPortal online platform.

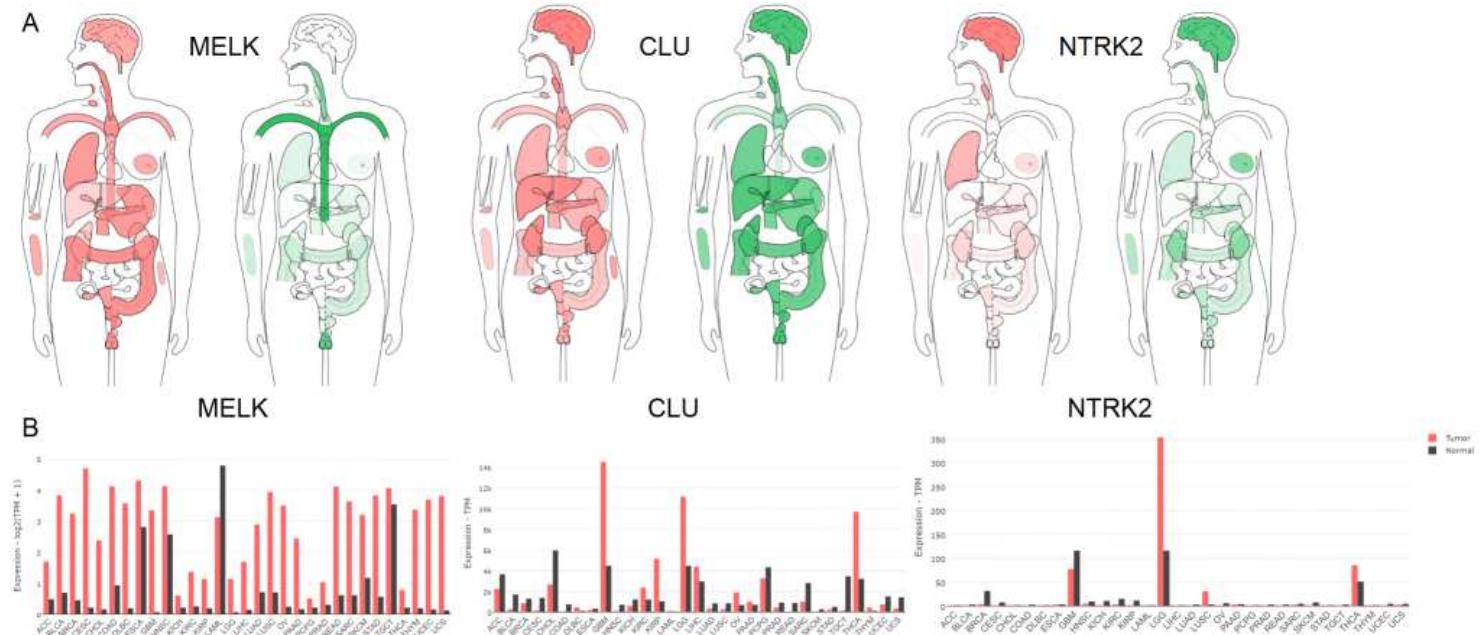


Figure 5

(A) Interactive Bodymap. The median expression of tumor and normal samples in bodymap. (B) The gene expression profile across all tumor samples and paired normal tissues (Bar plot). The height of bar represents the median expression of certain tumor type or normal tissue.

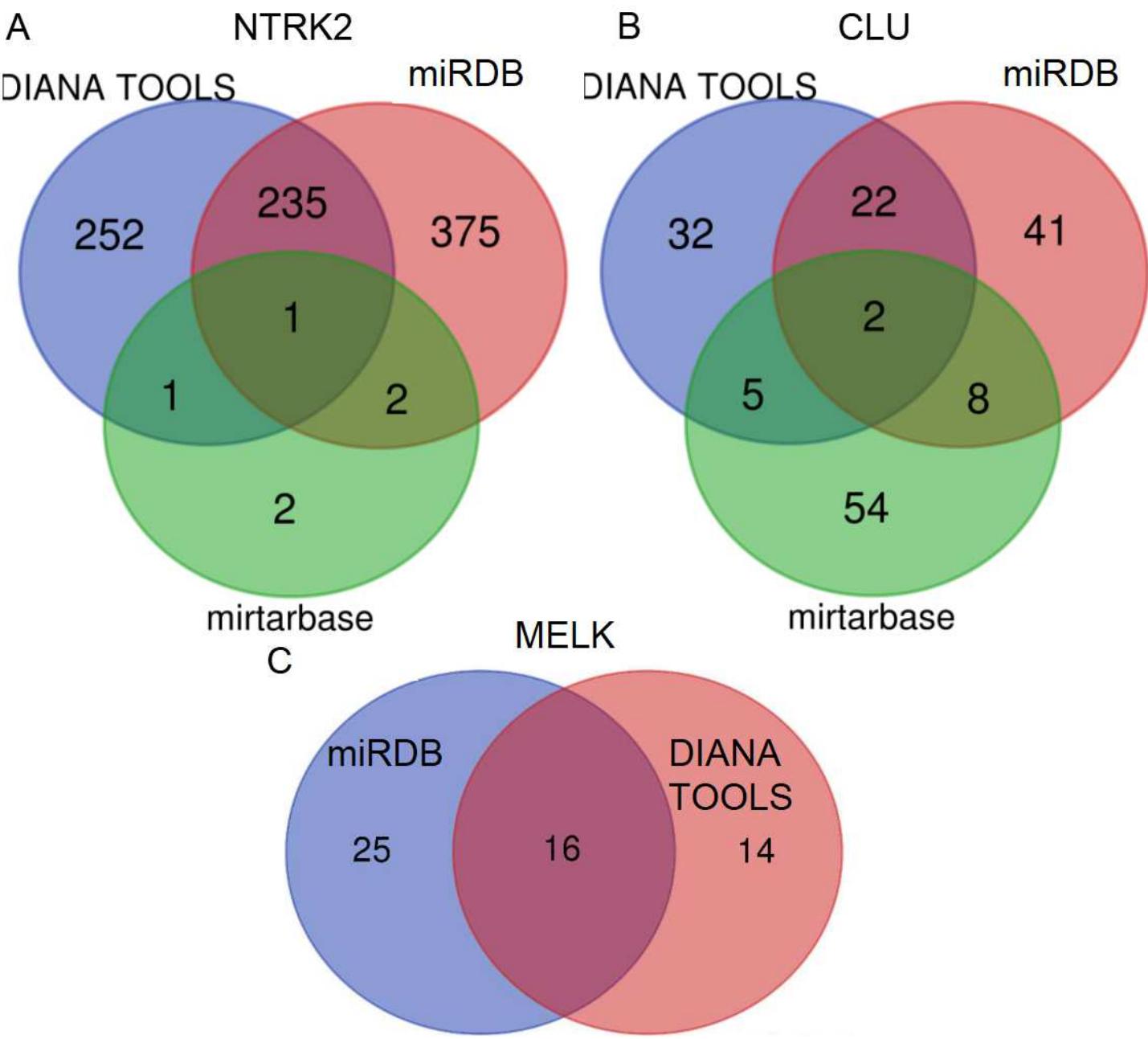


Figure 6

Venn diagram showing the intersection of gene targets for each hub gene from miRDB, MirTarBase and DIANA Tools. (A) The 3 datasets showed an overlap of 1 miRNA for NTRK2. (B) The 3 datasets showed an overlap of 2 miRNA for CLU. (C) The 2 datasets showed an overlap of 16 miRNA for MELK.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- Table1.docx
- Table2.docx
- Table3.docx