Bioinformatics for differential proteomics data between triple-negative breast cancer and luminal A subtype breast cancer

guogang wu
Ansteel Group Hospital https://orcid.org/0000-0001-9803-4425
Lei Zhang
China Medical University
Chunyan Liang
China Medical University
Yao Cheng
China Medical University
Chenguang Lv
Chinese Academy of Medical Sciences & Peking Union Medical College Medical Library
Bo Chen (chenbolaoshi2020@163.com)

Research

Keywords: Breast cancer, Proteomics, DIA, Bioinformatics analysis, IPA

Posted Date: January 15th, 2020

DOI: https://doi.org/10.21203/rs.2.20943/v1

Abstract

Background The heterogeneity between different subtypes of breast cancer has been established. Triple-negative breast cancer (TNBC) is the type of breast cancer with the highest recurrence rate, metastasis rate, and worst prognosis, while luminal A breast cancer is the subtype with the best prognosis. To study the heterogeneity of the two subtypes of breast cancer is important.

Methods We adopted a quantitative proteomics detection method based on data-independent acquisition and screened all differentially proteins between TNBC and luminal A breast cancer. Ingenuity Pathways Analysis was used for analysis.

Results We found that 207 proteins were up-regulated and 326 proteins were down-regulated. Classical pathway analysis showed that Eukaryotic translation initiation factor 2 (EIF2) signaling was significantly activated, in which EIF2α and other molecules were significantly up-regulated. The upstream regulatory analysis predicted 18 strong activators and 63 strong inhibitors, among which v-myc avian myelocytomatosis viral oncogene neuroblastoma derived homolog (MYCN) was the strongest activator and tretinoin was the strongest inhibitor. Based on disease and function analysis, it was found that differentially proteins were enriched in 41 biological processes, with RNA damage and repair and protein synthesis the most significant biological processes. There were 14 diseases or functions that were significantly activated and 35 that were significantly suppressed. Among the diseases or functions, the differentially proteins had the most significant inhibitory effect on osteosarcoma cell death. According to an analysis of the interaction network, 25 molecular interaction networks were found and the network with the highest score mainly affected the diseases and functions of cancer, protein synthesis, and RNA damage and repair.

Conclusion We identified many key proteins with TNBC heterogeneous characteristics and differentially expression, which will provide an abundance of new information for screening TNBC therapeutic specific targets and biomarkers.

Background

Breast cancer is the second major cause of cancer deaths in women after lung cancer. One million patients are estimated to be newly diagnosed with breast cancer each year [1]. Indeed, breast cancer has become one of the major diseases threatening women's health. Triple-negative breast cancer (TNBC) is a heterogeneous phenotype of breast cancer which lacks the expression of estrogen (ER), progesterone (PR), and human epidermal growth factor receptor (HER)2 receptors [2]. TNBC has an aggressive clinical course and poor prognosis [3] with a short 5-year survival and increased 3-year recurrence rates [4]. There is a significantly increased likelihood of distant recurrence in TNBC (33.9%) compared to other subtypes (20.4%) [3]. Considering the malignancy of TNBC and the death rate of metastatic breast cancer, further studies are required to improve the outcome of this subtype of breast cancer.

To discover the heterogeneity between TNBC and luminal A breast cancer, a data-independent acquisition (DIA) proteomics technique was used to compare the difference between TNBC and luminal A breast cancer proteomics. DIA integrates the advantages and characteristics of traditional shotgun and gold standard selective reaction monitoring/multiple reaction monitoring (SRM/MRM) techniques to obtain all fragment information of all ions in the sample without omissions and difference to achieve more accurate protein quantification. DIA proteome technology was also rated as a promising research technology by nature methods in 2015. This method will be more helpful for us to detect and screen all differentially expressed proteins of TNBC versus luminal A subtype.

Bioinformatics is an essential mainstream technology of proteomics research. The Ingenuity Pathways Analysis (IPA) used in this study is the most abundant and annotated biochemical analysis software and database at present. IPA is a proof-of-knowledge based software that helps researchers to model analyze, and understand the complex biological and chemical systems in life science research. IPA had an extensive repository of biological and chemical knowledge offering the researcher access to the most current findings available on genes, drugs, chemicals, protein families, normal cellular and disease processes, and signaling pathways. The software facilitates looking for information on genes and proteins, the impact of genes and proteins on diseases and cellular processes, and the role of genes and proteins in signaling pathways [5]. Such information can improve our understanding of the biological functions of TNBC differentially expressed proteins, contribute to an in-depth study of the heterogeneity of TNBC, and provide information in the search for new tumor targets.

Methods

Breast cancer sample collection

In our study, 52 human breast cancer tissue specimens and 3 para-cancer tissue samples were collected from breast cancer patients who underwent surgical resections in China Medical University. (Shenyang, China). Para-cancer tissues are non-cancerous that are obtained from > 2 cm from the tumor margin. All tissue specimens were retained for pathologic examination with HE staining to determine the pathologic type and other auxiliary tests, such as immunohistochemical and FISH testing. The remaining tissue specimens were washed with cold phosphate buffered saline (GE Healthcare, Beijing, China) to remove the residual blood after surgical resection, then immediately frozen in liquid nitrogen and stored at -80℃ for further analysis. None of the 52 breast cancer patients received pre-operative radiation, chemotherapy, or other treatments. This study was conducted in accordance with the principles expressed in the Helsinki declaration and approved by the Research Ethics Committee of China Medical University. (IRB approval number:AF-SOP-07-1.1-01). Informed consent was obtained from all patients. A summary of clinical and pathologic data for the breast cancer patients is shown in Supplementary Table 1 (red data represent 3 TNBC samples and green data represent 3 luminal A breast cancer samples).

Experimental procedures and strategies

Fifty-five samples were established by Data Dependent Acquisition(DDA) mass spectrometry. Then, based on the constructed database, DIA relative quantitative proteomics analysis was performed on 3 TNBC samples and 3 luminal A breast cancer samples to compare the differences between the two groups and screen out differentially expressed proteins. After that, we continued to perform IPA-based bioinformatics analysis of differentially expressed proteins.

Sample preparation

The tissue was scraped and added with an appropriate amount of SDT lysate, then transferred to a Lysing Matrix A tube. The homogenizer was used for homogenization crushing (24×2, 6.0 m/S, 60s, twice). After ultrasound, the homogenized tissue was placed in a boiling water bath for 10 min, then centrifuged at 14000g for 15 min. The supernatant was filtered through a 0.22-m centrifuge tube and the filtrate was collected. Protein was quantified using the BCA method. The samples were packed and stored at -80℃.

SDS-PAGE electrophoresis

Protein (20 μg) was taken from each sample and added to a 6X sample loading buffer, followed by a boiling water bath for 5 min, separation on 12% SDS-PAGE (constant pressure 250V for 40 min), and staining with Coomassie brilliant blue.

FASP enzyme [6]

A 200-ug protein solution was taken from each sample. DTT was added to a final concentration of 100mM and the solution was placed in a boiling water bath for 5 min, then the samples were cooled to room temperature. Two hundred micron of UA buffer was added to the protein solution and mixed well, then transferred to a 30kD ultrafiltration centrifuge tube and centrifuged at 12,500g for 25 min. The filtrate was discarded (this step was repeated twice). One hundred microliters of IAA buffer in UA was added, oscillated at 600 rpm for 1 min, reacted at room temperature against light for 30 min, and centrifuged at 12,500g for 25 min. One hundred microliters of UA buffer was added and centrifuged at 12,500g for 15 min (this step was repeated twice). One hundred microliters of 0.1M TEAB solution was added and centrifuged at 12,500g for 15 min (this step was repeated twice). Forty microliters of trypsin buffer (4μL g of trypsin in 40 mu L of 0.1M TEAB solution) was oscillated at 600 rpm for 1 min, and placed at 37℃ for 16-18 h. The collection tube was replaced, centrifuged at 12,500g for 15 min, 20μL of 0.1M TEAB solution was added, centrifuged at 12500g for 15 min, and the filtrate was collected. The peptide fragment was desalted with a C 18 cartridge, freeze-dried, and re-dissolved with 40μL of 0.1% formic acid solution. The peptide fragment was quantified.

High PH RP classification

All the peptide compounds were classified using an Agilent 1260 Infinity II HPLC system. Buffer solution A was 10mM $HCOONH_4$ and 5% can (pH 10.0) and buffer solution B was 10mM $HCOONH_4$ and 85% can (pH 10.0).The column was balanced with liquid A and the samples were separated from the column (Waters, XBridge Peptide BEH C18 Column, 130A, 5 m, 4.6 mm X 100 mm) with a flow rate of 1 mL/min. The liquid phase gradient was as follows: using a linear gradient, the column temperature was maintained at 30℃ within 40 min from 5% B to 45% B. Thirty-six components were collected and dried in a vacuum concentrator. After freeze-drying, the samples were re-dissolved with 0.1% formic acid solution in 12 fractions.

DDA mass spectrometry library

Six microliters were taken from each fraction and added to 1 ul of 10× iRT peptide fragment. After mixing, 6 ul of sample was injected, separated by nano-lc, and analyzed by on-line electrospray tandem mass spectrometry. The whole liquid-mass series system was as follows: 1) liquid phase system, Easy nLC system (Thermo Fisher Scientific); 2) mass spectrometry system, q-exactive hf-x (Thermo Fisher Scientific). Buffer solution A was 0.1% formic acid solution and solution B was 0.1% acetonitrile solution (acetonitrile was 80%). The samples were separated by a gradient of non-linear growth in an analytical column (Thermo Fisher Scientific, Acclaim PepMap RSLC 50 um X 15 cm, nano viper, P/N164943) at a flow rate of 300 nL/min for 0-5 min, 1% B for 5-95 min, 1% B to 28% B for 95-110 min, 28% B to 38% B for 110-115 min, 38% B to 100% B for 115-120 min, and 100% B. The electrospray voltage was 2.0kV.

The mass spectrum parameters were set as follows: (1) MS, scan range (m/z)=350–1500; resolution=60,000; AGC target=3e6; maximum injection time=30 ms; included charge states=2-7; filter dynamic exclusion: exclusion duration=30s; and (2) dd-MS2, isolation window=1.6 m/z; resolution=15,000; AGC target=1e5; maximum injection time=45 ms; and NCE=28%.

The Spectronaut Pulsar X (version 12, Biognosys AG) was used to consolidate and analyze the original mass spectrometry data and establish a spectrograph database. The database was Swissprot_human_isoform_201806 (42356 entries), as below: 2018-06, download link: http://www.uniprot.org. Trypsin enzymolysis was set to allow two missing cutting sites. Carbamidomethyl (C), variable modification: Oxidation (M), and acetyl (Protein n-term) n-terminal acetylation. The standard for database construction was 1% precursor false discovery rate(FDR), 1% protein FDR, and 1% peptide FDR.

DIA mass spectrometry

Six samples (TNBC and luminal A breast cancer [3 cases each]) were removed and 6 ul was added to 1 ul of 10× iRT peptide, mixed, separated with nano-lc, and analyzed by on-line electrospray tandem mass

spectrometry. The whole liquid-mass series system was as follows: 1) liquid phase system, Easy nLC system (Thermo Fisher Scientific); and 2) mass spectrometry system, q-exactive hf-x (Thermo Fisher Scientific). Buffer solution A was 0.1% formic acid solution and solution B was 0.1% acetonitrile solution (acetonitrile was 80%). The samples were separated by a gradient of non-linear growth in the analytical column (Thermo Fisher Scientific, Acclaim PepMap RSLC 50um X 15cm,nano viper, P/N164943) at a flow rate of 300 nL/min for 0-5 min, 1% B for 5-95 min, 1% B to 28% B for 95-110 min, 28% B to 38% B for 110-115 min, 38% B to 100% B for 115-120 min, and 100% B. The electrospray voltage was 2.0kV.

The mass spectrum parameters were set as follows: (1) MS, scan range (m/z)=350–1500; resolution=60,000; AGC target=3e6; maximum injection time=50 ms; and (2) DIA, resolution=15,000; AGC target=2e5; maximum injection time=45ms; and NCE=28%.

Bioinformatics Analyses

Differential proteins performed IPA-based classical pathway analysis, upstream regulation analysis, disease and function analysis, regulation effect analysis, and interaction network analysis. A detailed description of these analyses is provided in the supplementary information.

Results

Protein identification

Quantitative results of DIA samples after DDA database construction

A total of 149,700 precursors，113,117 peptides, and 9730 protein groups were detected in the DDA database of 55 samples. DIA analysis was performed on 6 samples (3 TNBC cases and 3 luminal A breast cancer cases) using the constructed database, and the quantitative information obtained is shown in Table 1 and Supplementary Table 2.

Results of sample protein quantification based on DIA

The quantitative results of the 6 breast cancer samples are shown in Supplementary Table 3. The comparison of protein abundance between the two groups indicates that most of the quantitative results of the two groups were close to 1, and the parallelism between the two groups was good（Figure 1）. The volcano map of protein data from the two groups indicates that a significant difference existed between the two groups（Figure 2）. We selected at least one-half of the non-null value data of TNBC versus luminal A breast cancer in the two groups for significant difference analysis, and screened proteins with

expression difference multiples greater than 1.5-fold (up-down) and a P value (t-test) < 0.05 as differentially expressed proteins. The results indicated that 207 proteins were up-regulated and 326 proteins were down-regulated, for a total of 533 proteins (Supplementary Table 3 [red data represent up-regulated proteins and green data represent down-regulated proteins]).

Bioinformatics analysis based on IPA

Classical pathway analysis

Classical pathway analysis based on IPA showed that these differentially expressed proteins were related to 443 signaling pathways. EIF2 signaling was significantly activated [-log(p-value) was 42.6; z-score was 4.439]. Forty-two signaling pathways, including ephrin B signaling, acute phase response signaling, liver X receptor/retinoid X receptor (LXR/RXR) activation, ephrin receptor signaling, G beta gamma signaling, and tec kinase signaling were significantly suppressed (Figure 3 and Supplementary Table 4). We further analyzed the trend change of different protein molecules in the EIF2 signaling pathway and found that EIF2α, EIF4A3, EIF5B, and the poly(A)-binding protein (PABP) were significantly up-regulated. Phosphoinositide 3-kinase(PI3K), Phosphoinositide-dependent protein kinase 1 (PDK1), RAS, and extracellular regulated kinase1/2(ERK1/2) were significantly down-regulated, and the remaining protein molecules were not significantly changed (Figure 4). Previous studies have reported that EIF2 signaling significantly down-regulates EIF2α, Protein kinase R (PKR)-like endoplasmic reticulum kinase(PERK), C/EBP-homologous protein (CHOP), activating transcription factor 4(ATF4), activating transcription factor 3(ATF3), and other molecules, while EIF2ß, GTP, PI3K, PDK1, RAS, ERK1/2, and other molecules are significantly up-regulated (Figure 5), which are inconsistent with the results of the current study. The reason may be that luminal A type breast cancer was selected as the reference in this study, but not the abnormal adjacent tissues.

Upstream regulation analysis

We continued our IPA-based upstream regulatory factor analysis of all differentially expressed proteins and found a total of 505 regulatory factors that might affect the expression of differentially expressed proteins, covering all molecular types, including transcription factors, cytokines, small RNAs, receptors, kinases, chemical molecules, and drugs (Supplementary Table 5). In this study, IPA used the activation z-score algorithm to predict the activation or inhibition of the upstream regulator and reduce the significant prediction caused by random data. Eighteen strong activators, including MYCN, MYC, TCR, lysine-specific histone demethylase 1A (KDM1A), HRAS, mir-122, KRAS, paricalcitol, recombinant Tissue Inhibitors Of Metalloproteinase 3 (TIMP3), and T-box transcription factor(TBX2) were predicted (score ≥ 2). MYCN was predicted to be the strongest activator (score = 6.214), with 48 uniformly activated factors and 8 uniformly suppressed factors (Figure 6). At the same time, 63 strong inhibitors, including tretinoin,

rapamycin-insensitive companion of m TOR (RICTOR), 5-fluorouracil, sirolimus, CD 437, ST1926, microtubule-associated protein tau (MAPT), phenacetin, hexachlorobenzene, and thioacetamide were predicted (score ≤ -2). Tretinoin was predicted to be the strongest inhibitor (score = -5.578), with 37 uniformly suppressed factors.

Disease and functional analysis

In our IPA-based disease and function analysis of differential proteins, we found that there were 41 biological processes with differential protein enrichment. RNA damage and repair, and protein synthesis were the two most significant biological processes (Figure 7). In addition, differential proteins were activated or suppressed in disease and function analysis. Fourteen diseases or functions were significantly activated, including liver lesion (Z - score = 3.465) and morbidity or mortality (3.409). Thirty-five diseases or functions were significantly suppressed, including osteosarcoma cell death (-5.292) and cell movement (-3.335; Figure 8 and Supplementary Table 6). Differential proteins had the most significant relationship with osteosarcoma cell death. We also thought that TNBC may have similar tumor biological characteristics with osteosarcomas, such as a high recurrence rate, high distant metastasis rate, and poor prognosis. We made a network map of the inhibitory effect of up-regulated differential proteins on osteosarcoma cell death. After analysis, we found that 28 kinds of up-regulated differential proteins, including Elongation factor 1-alpha 1(EEF1A1), 60S ribosomal protein L10a(RPL10), 60S ribosomal protein L11(RPL11), 60S ribosomal protein L13a (RPL13A), 60S ribosomal protein L18a(RPL18), 39S ribosomal protein L19(RPL19), 60S ribosomal protein L27a (RPL27A), 60S ribosomal protein L36(RPL3), and 60S Ribosomal Protein L31(RPL31) had an inhibitory effect on osteosarcoma cell death (Figure 9). These findings may provide new insight into the heterogeneity of TNBC.

Regulation effect analysis

In this study, differential proteins were involved in the upstream regulatory network and the downstream function of the possible pathways of action (Supplementary Table 7). We chose the first regulatory network in the analysis of regulatory effects. The network due to the following regulators were able to activate infection by Zaire and inhibit osteosarcoma cell death (Figure 10). Therefore, these regulators should be the focus of attention in future studies involving the heterogeneity of TNBC.

Interaction network analysis

Molecular interaction network analysis based on the IPA showed that of 25 molecular interaction networks (Supplementary Table 8), we chose the highest score of molecular interaction networks to be

displayed (Figure 11). Thirty-three different molecules involved in the network, including 28 up-regulation factors and 5 down-regulation factors, influence the main diseases and the function name for “cancer, protein synthesis, and RNA damage and repair.” We showed that MCM2 and ribosomal 40s subunit is the hub of the interacting network and may play a core role in the interacting network.

Discussion

Compared with luminal A breast cancer, TNBC has a very different prognosis. When the proteomics differences between TNBC and luminal A breast cancer are compared in detail, it will be helpful to better understand the heterogeneity of TNBC, find effective therapeutic targets, and improve the prognosis of TNBC. We used DIA proteomics to compare TNBC with luminal A breast cancer in this study. First, we established 55 samples of DDA in a mass spectrum library. The purpose was to obtain comprehensive information about breast cancer tissue proteins, which not only detected data to be stored permanently and facilitated subsequent retrospective analysis, but will provide a background reference for the quantitative analysis of DIA technical proteins in TNBC and luminal A breast cancer to realize the deep coverage and accurate quantitative analysis of low abundance proteins. After that, we conducted DIA analysis of TNBC and luminal A breast cancer samples. After verifying that the consistency was ideal and the overall identification quantity was in agreement with expectations, we continued to conduct follow-up differential quantitative screening. We showed that 207 proteins were up-regulated and 326 proteins were down-regulated in TNBC in contrast to luminal A breast cancer. This laid the foundation for our follow-up bioinformatics research.

We further analyzed the classic IPA-based pathway of differentially expressed proteins and found that eIF2 signaling was significantly activated, in which eIF molecules, such as EIF2α, eIF4A, and eIF5B, were significantly up-regulated compared with luminal A breast cancer. EIF2α, as the subunit of protein translation initiation factor EIF2, plays an important role in tumor development and development [7]. EIF2αexpression has been reported to increase in various types of solid and hematologic tumors, including lung bronchoalveolar carcinoma [8], Hodgkin's lymphoma [9], gastrointestinal carcinoma [10], and malignant melanoma [11]. EIF2αis mainly distributed in the nucleus and cytoplasm of tumor cells, while the distribution in minimal in the cytoplasm of normal tissues. In addition, EIF2αwas also differently expressed in thyroid cancer in different states, and the expression of EIF2αin invasive thyroid cancer was significantly higher than thyroid papillary cancer [12], which was similar to TNBC and luminal A breast cancer in the current study. EIF2α expression varies in different subtypes of breast cancer or in breast cancer with different degrees of invasion. Even within the same type of tumor,the degree of malignancy may be closely related to EIF2α-dependent protein synthesis. Recent studies have suggested that cancer cells have impaired canonical translation and direct translational machinery to EIF2α-dependent translation when encountering various microenvironmental stresses in tumourigenesis [13] [14]. Chen et al. [15] have shown that mitochondrial stress response (ISR) is the key to survival of tumor cells under stress stimulation, which is closely related to drug resistance of tumor therapy, and EIF2α-dependent directed translation plays an important role in the paclitaxel-mediated ISR process. The up-regulation of EIF2αexpression may suggest that EIF2α-dependent directed translation protein is active

and leads to a higher incidence of chemotherapy resistance in TNBC [15]. Phosphorylation of EIF2αstrongly inhibits the initiation of abnormal protein translation and further inhibits the occurrence and development of tumors [16] [17]. Up-regulated expression of p- EIF2αsuggests a good prognosis for TNBC [18]. We hypothesized that the higher the expression of EIF2α in TNBC compared with luminal A breast cancer may be related to blocking eIF2α phosphorylation; however, the specific mechanism warrants further study.

To find the culprit for the occurrence of differentially expressed proteins, we conducted IPA-based upstream regulatory factor analysis and predicted 18 strong activators and 63 strong inhibitors as candidate targets for further study of TNBC. Among the proteins, MYCN is predicted to be the strongest activator; MYCN is a member of the myc family [19] [20]. The amplification of this gene is related to a variety of tumors, such as neuroblastoma nephroblastoma and other childhood specific tumors, as well as lung cancer, prostate cancer and basal cell carcinoma in adults [21]. Heterogeneity of MYCN has been recently documented to vary from tumor sites during cancer progression and even following treatment [22]. MYCN genetic aberrations are involved in the development of a wide variety of diseases. Many malignancies would benefit from MYCN-targeted therapeutic approaches [21]. Tretinoin is predicted to be the strongest inhibitor; tretinoin is a natural or synthetic vitamin A derivative. Tretinoin can induce malignant tumor cells to mature and differentiate. Clinically, tretinoin can effectively treat acute promyelocytic leukemia, mucosal leukoplakia, and recurrent head and neck squamous cell carcinoma [23]. Many studies have shown that tretinoin can also inhibit the growth of breast cancer cells, promote apoptosis, and partial differentiation [24] [25]. These findings also suggest that MYCN and tretinoin will serve as targets for an in-depth study of TNBC. In addition, the upstream regulatory factor network map indicated the presence of 48 uniformly activated genes in the MYCN gene and 37 uniformly suppressed genes in the tretinoin gene. These common genes also have the potential to serve as biomarkers for TNBC.

We continued to analyze the diseases and functions of differentially expressed proteins based on IPA, and found that the proteins were most significant in the two biological processes of RNA damage and repair and protein synthesis. The proteins could significantly inhibit osteosarcoma cell death. Considering that TNBC may have similar tumor biological characteristics to osteosarcomas, such as a high recurrence rate, high distant metastasis rate, and poor prognosis. We analyzed the regulatory protein and regulatory effect of osteosarcoma cell death and found that EIF4E, MYC, MYCN, RICTOR, and TFAP2A were the main regulatory factors in the regulatory network. EIF4E is a key factor in the formation of the protein translation initiation complex. Over-expression of IF4E is involved in cell growth, cell proliferation, invasion, and apoptosis, and is associated with poor survival [26]. EIF4E is the central aggregation point of multiple signaling pathways. Activation of EIF4E is a key downstream event in phosphorylation and activation of the AKT/mTOR signaling pathway. The RICTOR gene has recently been shown to be amplified in cancer. Over-expression of RICTOR was also positively associated with tumor progression and poor survival in colorectal cancer [27], hepatocellular carcinoma [28], endometrial carcinoma [29], and pituitary adenoma [50]. RICTOR plays a central role in the PI3K/AKT/mTOR pathway in cancer [30]. These are the main nodes in the regulatory network and the main candidates for screening of TNBC

targets. Our follow-up analyzed was based on the interaction of the IPA network and showed 25 interactions. The “cancer, protein synthesis, and RNA damage and repair” network score was the highest. The MCM2 and 40 s ribosomal subunit were the interaction network hub and may play a central role in the network of interactions, so are also prime candidates for our further screening of TNBC targets.

Conclusions

Based on DIA technology and IPA bioinformatics analysis, we identified many key proteins with TNBC heterogeneous characteristics and differential expression, which will provide an abundance of new information for screening TNBC therapeutic specific targets and biomarkers.

Abbreviations

TNBC
triple-negative breast cancer
DIA
data-independent acquisition
IPA
Ingenuity Pathways Analysis
EIF2
eukaryotic translation initiation factor 2

Declarations

Ethics approval and consent to participate

This study was approved by the Research Ethics Committee of China Medical University. (IRB approval number:AF-SOP-07-1.1-01). Informed consent was obtained from all patients.

Consent for publication

Written informed consent for publication was obtained from all participants.

Availability of data and material

All data generated or analysed during this study are included in this published article and its supplementary information files.

Competing interests

The authors declare no competing financial interests.

Funding

This study was supported by the grants from the National Natural Science Foundation of China (No. 81372811) and Science and Technology Agency of Liaoning Province (No. 2013225049).

Authors' contributions

W.G. G. and C. B. designed experiments. Z.L. participated in the conception of the idea. L.C.Y. and C.Y. contributed to the literature search.L.C.G.participated in the conception and coordination. All authors have read and approved the final manuscript for publication.

Acknowledgements

We thank the department of Breast Surgery of The First Hospital of China Medical University for providing human breast cancer samples. We also thank the College of China Medical University and Genechem Co.,Ltd. (Shanghai ,China) for technical assistance in experiments.

References

1. Aysola K , Desai A , Welch C,Xu J , Qin Y , Reddy V , Matthews R , et al. Triple Negative Breast Cancer–An Overview[J]. Hereditary Genet 2013 ;2013 (Suppl 2).
2. Brenton JD , Carey LA , Ahmed AA , Caldas C. Molecular Classification and Molecular Forecasting of Breast Cancer: Ready for Clinical Application? [J]. J. Clin. Oncol 2005,;23:7350–60.
3. Dent R , Trudeau M , Pritchard KI , Hanna WM , Kahn HK , Sawka CA , et al. Triple-negative breast cancer: Clinical features and patterns of recurrence[J]. Clin. Cancer Res 2007;13 (15 pt 1): 4429–34.
4. Lanning NJ , Castle JP , Singh SJ , Leon AN , Tovar EA , Sanghera A , et al. Metabolic profiling of triple-negative breast cancer cells reveals metabolic vulnerabilities[J]. Cancer Metab 2017;5:6.
5. Lv H , Liu L , Zhang Y , Song T , Lu J , Chen X .Ingenuity pathways analysis of urine metabonomics phenotypes toxicity of gentamicin in multiple organs[J].Molecular bioSystems 2010;6(10):2056-67.
6. Wiśniewski JR , Zougman A , Nagaraj N , Mann M.Universal sample preparation method for proteome analysis[J]. Nat Methods 2009;6(5): 359-62.
7. Koromilas AE. Roles of the translation initiation factor eIF2alpha serine 51 phosphorylation in cancer formation and treatment[J]. Biochim Biophys Acta 2015;1849:871–80.
8. Rosenwald I B,Hutzler M J,Wang S,Savas L,Fraire A E.Expression of eukaryotic translation initiation factors 4E and 2alpha is increased frequently in bronchioloalveolar but not in squamous cell carcinomas of the lung[J]Cancer 2001;92(8):2164-71
9. Igor B. Rosenwald,Larisa Koifman,Lou Savas,Jane-Jane Chen,Bruce A. Woda,Marshall E. Kadin. Expression of the translation initiation factors eIF-4E and eIF-2α is frequently increased in neoplastic cells of Hodgkin lymphoma[J]HumPathol 2008;39(6):910-16.

10. Lobo MV, Martin ME, Perez MI, Alonso FJ, Redondo C, Alvarez MI, et al. Levels, phosphorylation status and cellular localization of translational factor eIF2 in gastrointestinal carcinomas[J]. Histochem J 2000;32(3):139–50.
11. Rosenwald IB, Wang S, Savas L, Woda B, Pullman J. Expression of translation initiation factor eIF-2alpha is increased in benign and malignant melanocytic and colonic epithelial neoplasms[J]. Cancer 2003;98(5):1080–1088.
12. Wang S,Lloyd R V,Hutzler M J,Rosenwald I B,Safran M S,Patwardhan N A,et al□Expression of eukaryotic translation initiation factors 4E and 2alpha correlates with the progression of thyroid carcinoma□J□□Thyroid 2002;11(12):1101□07□
13. Pakos-Zebrucka K, Koryga I, Mnich K, Ljujic M, Samali A, Gorman AM. The integrated stress response. [J] EMBO Rep 2016;17(10):1374-1395.
14. Wang M, Kaufman RJ. The impact of the endoplasmic reticulum protein-folding environment on cancer development[J]. Nat Rev Cancer 2014;14(9):581-97.
15. Chen Lin,He Jiang,Zhou Jianhua,Xiao Zhi,Ding Nianhua,Duan Yumei, et al.1EIF2A promotes cell survival during paclitaxel treatment in vitro and in vivo[J]. J Cell Mol Med 2019;29(3):1–12.
16. Holcik Martin. Could the eIF2α-Independent Translation Be the Achilles Heel of Cancer?[J]. Frontiers in oncology 2015;5:264.
17. Boye Erik,Grallert Beáta. eIF2α phosphorylation and the regulation of translation.[J]. Current genetics,2019. https://doi.org/10.1007/s00294-019-01026-1.
18. Guo L , Chi Y , Xue J , Ma L , Shao Z , Wu J.Phosphorylated eIF2α predicts disease-free survival in triple-negative breast cancer patients.[J].Scientific reports 2017;7:44674.
19. Mathsyaraja Haritha□Eisenman Robert N. Parsing Myc Paralogs in Oncogenesis.[J]. Cancer Cell 2016;29(1):1–2.
20. Wilde Blake R,Ayer Donald E. Interactions between Myc and MondoA transcription factors in metabolism and tumourigenesis. Br. J. Cancer 2015;113(11): 1529–33.
21. María Victoria Ruiz-Pérez□Aine Brigette Henley□Marie Arsenian-Henriksson.The MYCN Protein in Health and Disease.[J].Genes 2017;8(4):1-27.
22. Marrano Paula,Irwin Meredith S,Thorner Paul S. Heterogeneity of MYCN amplification in neuroblastoma at diagnosis,treatment, relapse, and metastasis[J].Genes Chromosomes Cancer 2017;, 56(1):28–41.
23. GudasLJ. Retinoids,retinoid-responsive genes,cell differentiation, and cancer[J].Cell Growth Differ 1992; 3(9):655-62.
24. KM Eck, L Yuan, L Duffy, PT Ram, S Ayettey, I Chen, etal.A sequential treatment regimen with melatonin and all-transretinoic acid induces apoptosis inMCF-7 tumour cells[J].Br J Cancer 1998;77(12):2129-37.
25. Van heusden J,Wouters W,Ramaekers F C,Krekels M D,Dillen L,Borgers M, et al.All-trans-retinoic acid metabolites significantly inhibit the proliferation of MCF-7 human breast cancer cells in

vitro[J].BrJCancer 1998;77(1):26-32.

26. Ramalingam Senthilmurugan,Gediya Lalji,Kwegyir-Afful Andrew K,Ramamurthy Vidya P,Purushottamachar Puranik,Mbatia Hannah,etal.First MNKs degrading agents block phosphorylation of eIF4E, induce apoptosis, inhibit cell growth, migration and invasion in triple negative and Her2-overexpressing breast cancer cell lines.[J]. Oncotarget 2014;5(2) :530-43
27. Gulhati P, Bowen KA, Liu J, Stevens PD, Rychahou PG, Chen M,etal. MTORC1 and mTORC2 regulate EMT, motility, and metastasis of colorectal cancer via RhoA and Rac1 signaling pathways[J]. Cancer Res2011;71(9):3246–56.
28. Kaibori M, Shikata N, Sakaguchi T, Ishizaki M, Matsui K, Iida H,et al. Influence of Rictor and raptor expression of mTOR signaling on long-term outcomes of patients with hepatocellular carcinoma[J]. Dig Dis Sci2015;60(4):919–28.
29. Wen SY, Li CH, Zhang YL, Bian YH, Ma L, Ge QL, etal.Rictor is an independent prognostic factor for endometrial carcinoma[J]. Int J Clin Exp Pathol2014;7(5):2068–78.
30. Ahlem Jebali , Nicolas Dumaz.The role of RICTOR downstream of receptor tyrosine kinase in cancers.[J].Molecular cancer 2018;17(1):39.

Tables

Table1 Sample quantitative results

	Precursors	Peptides	Protein Groups
average quantitative	41741	34372	4870
total quantitative	72561	57628	5511
common quantitative	17974	15385	3225

Figures

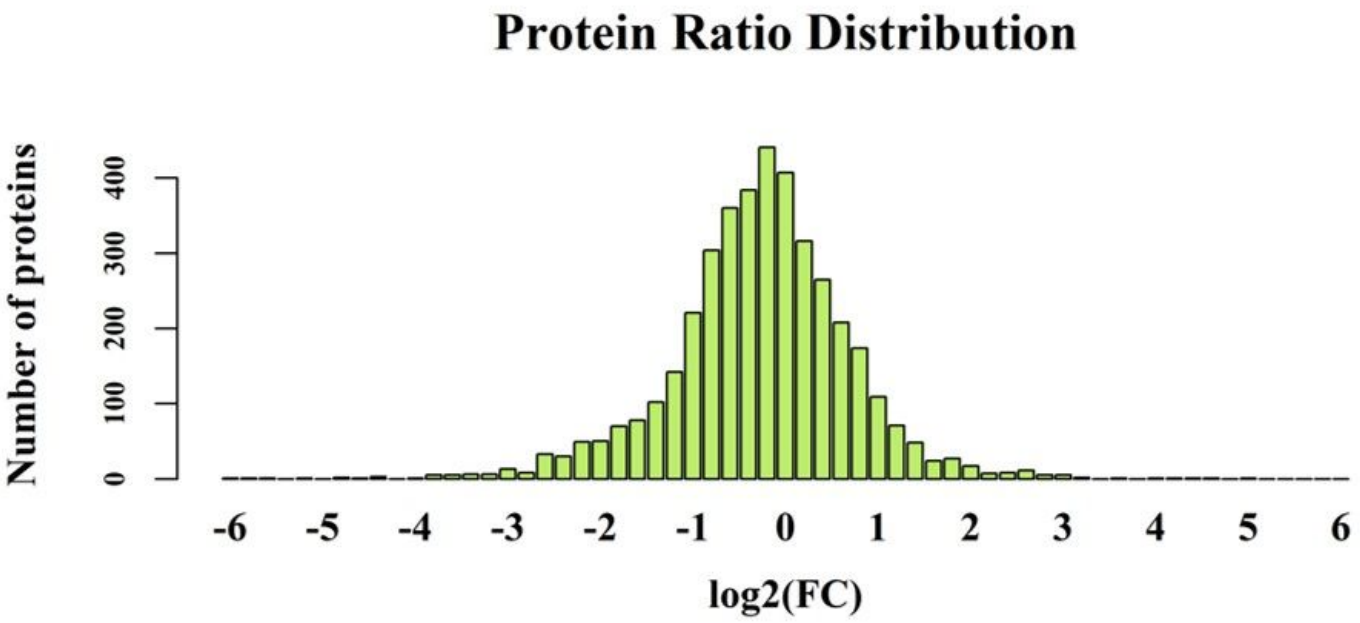

Figure 1. Protein Ratio Distribution

Figure 1

The x-coordinate is the multiple of difference (logarithmic transformation based on 2). The ordinate is the number of identified proteins.

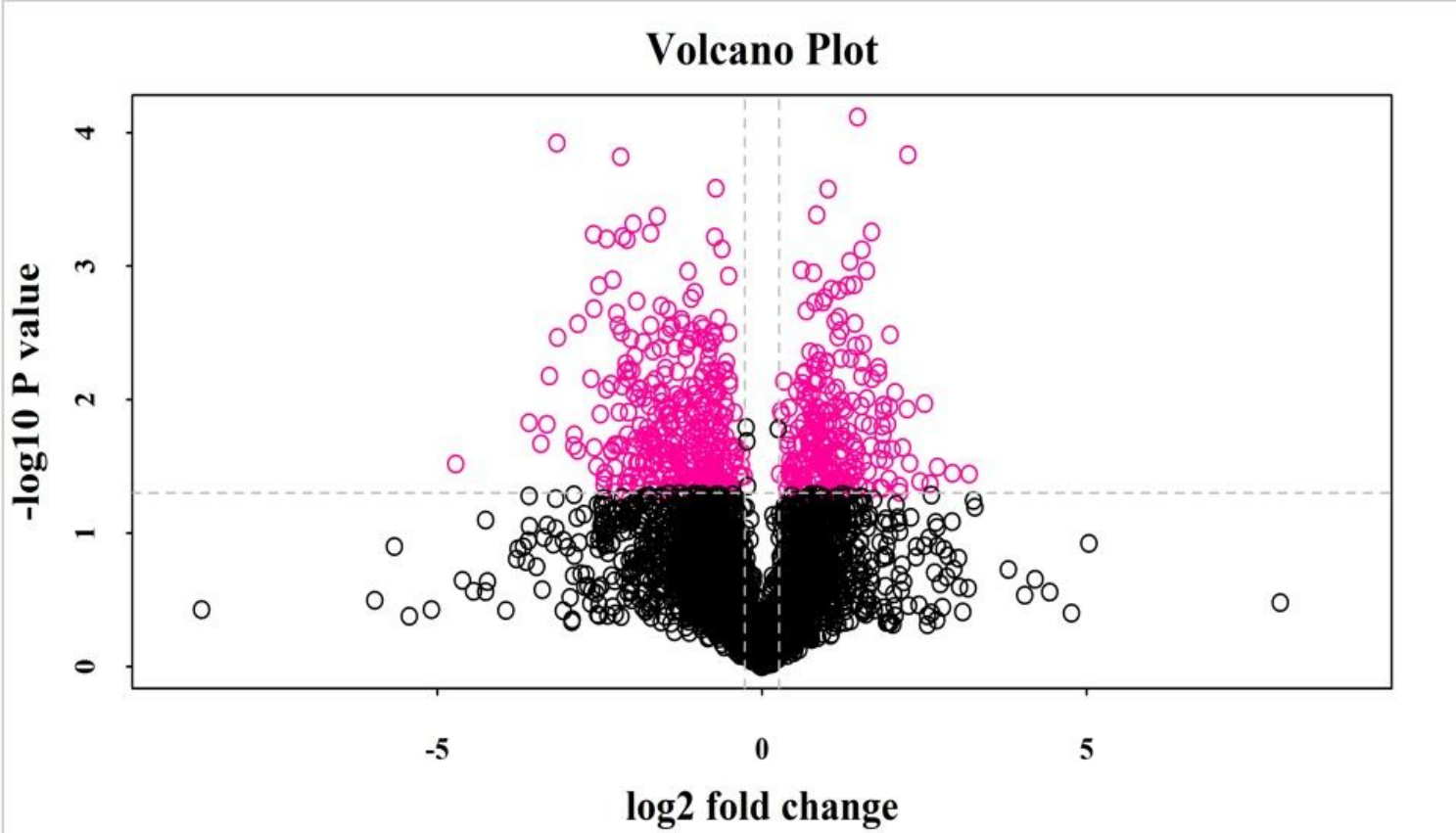

Figure 2. The fold-change and P value of protein expression between the two groups of samples (Volcano Plot).

Figure 2

The fold-change and P value of protein expression between the two groups of samples were used to draw the volcano map to show the significant differences between the two groups of sample data. The x-coordinate is the difference multiple (log base 2) and the y-coordinate is the P value (log base 10).

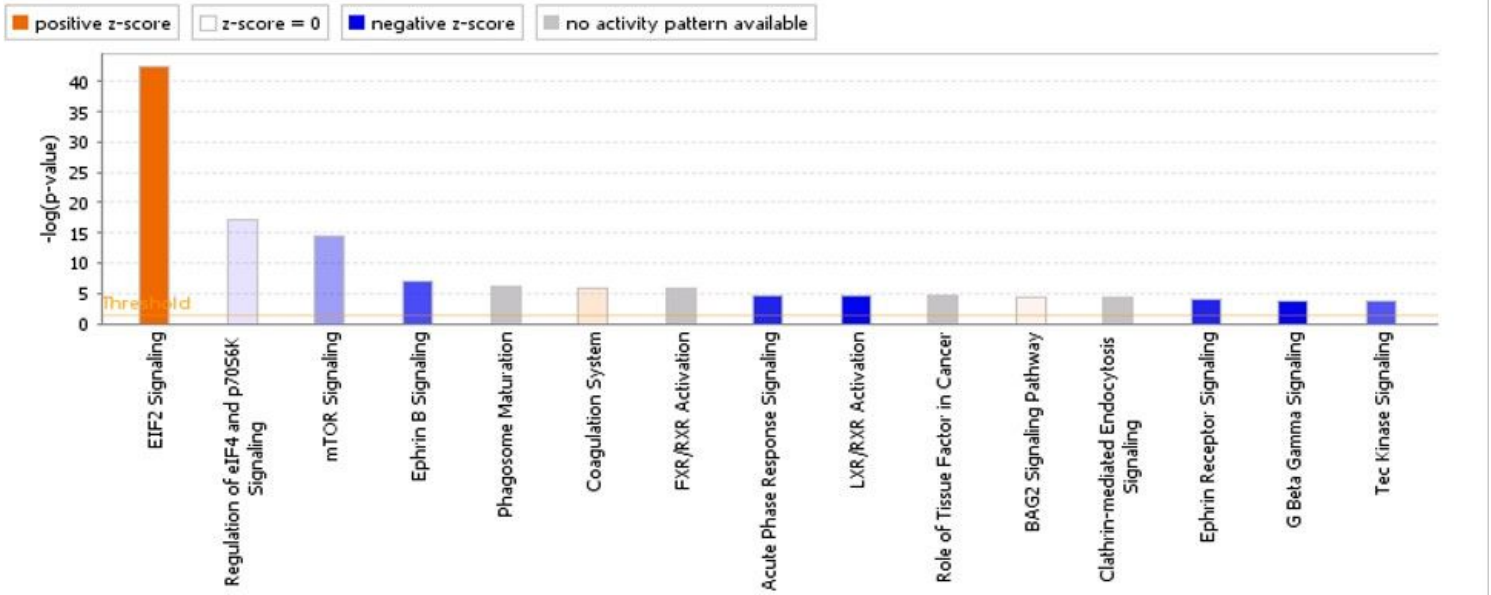

Figure 3. Significant enrichment of differential proteins in some classic pathways.

Figure 3

Significant enrichment of differential proteins in some classic pathways. The x-coordinate is the pathway name, and the y-coordinate is the significance level of enrichment (negative logarithmic transformation with base 10). The orange mark indicates that the pathway is activated (z-score > 0) and the blue mark indicates that the pathway is inhibited (z-score < 0). The depth of orange and blue (or the absolute value of the z-score) represents the degree of activation or inhibition (according to the internal algorithm and criteria of IPA; a z-score ≥ 2 means that the pathway is significantly activated and a z-score ≤ -2 means that the pathway is significantly suppressed). Ratio represents the ratio of the number of differential genes in the signaling pathway to the number of all genes contained in the pathway. EIF2 signaling is significantly activated in this program with a z-score of 4.439.

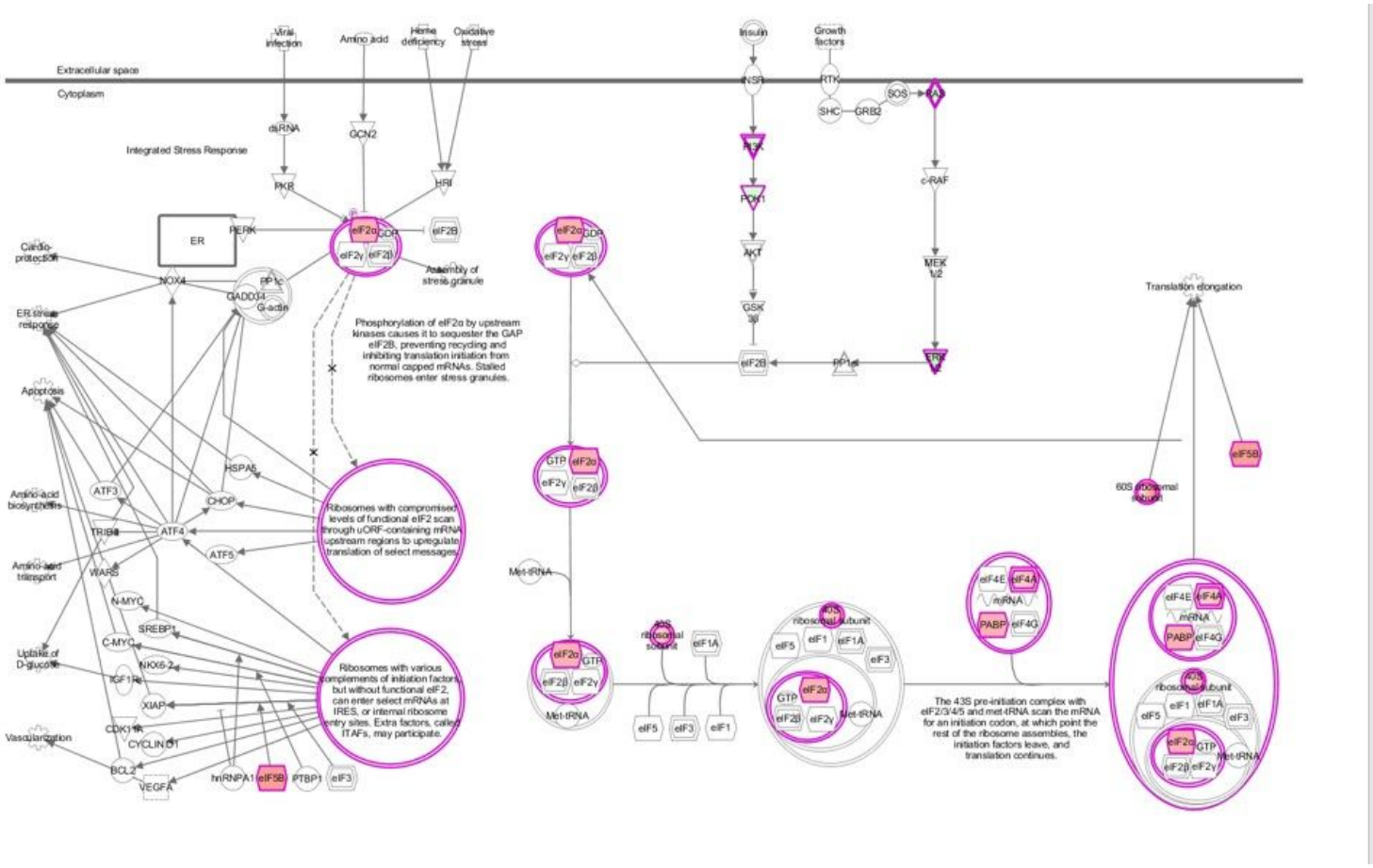

Figure 4

Expression trend of each protein molecule in the EIF2 signaling pathway in the experimental results. Red represents that the protein is significantly up-regulated in the experimental results, and green represents that the protein is significantly down-regulated in the experimental results.

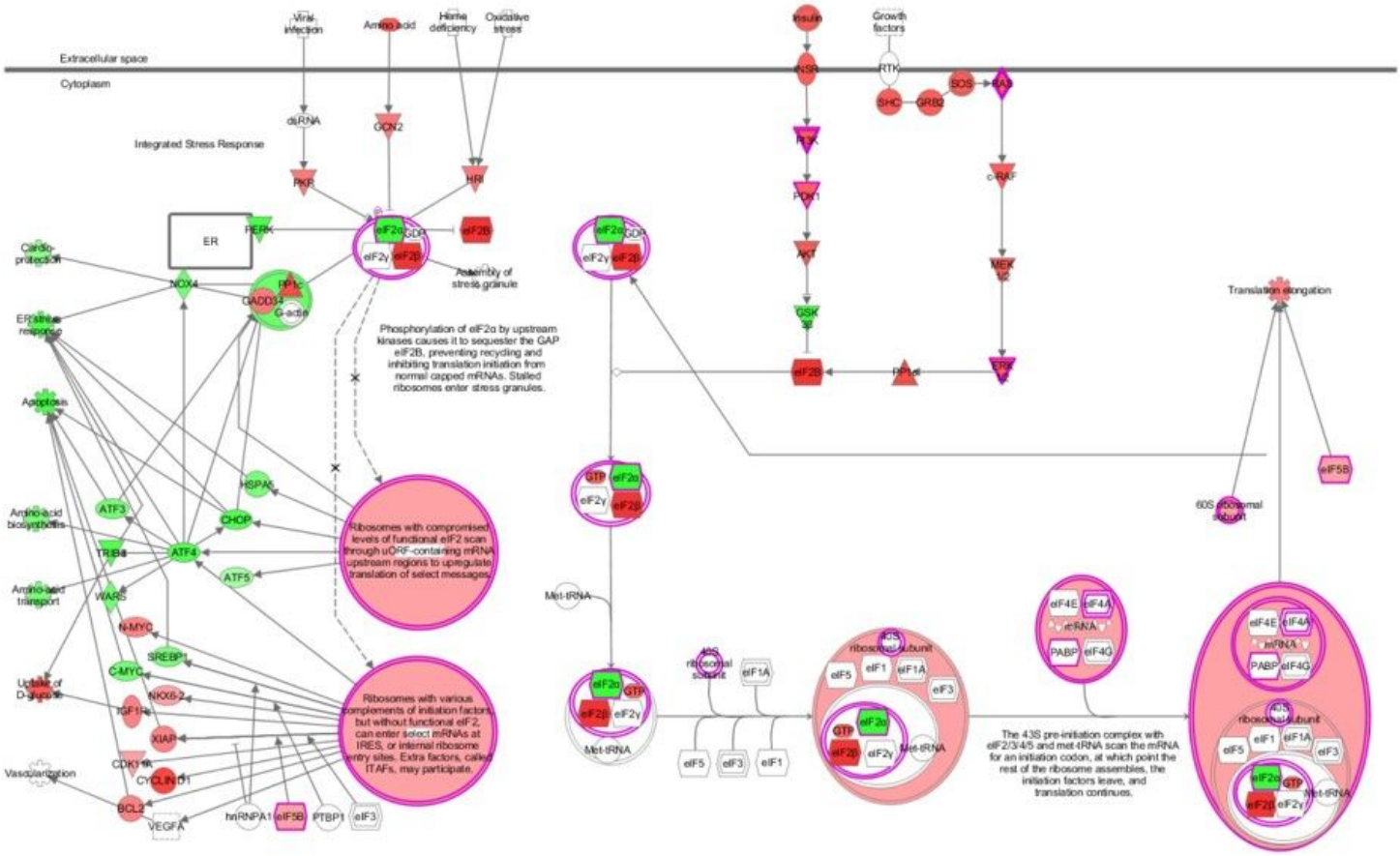

Figure 5

EIF2 signaling shows the trend change in expression of genes involved in this pathway in the activated state in the existing literature, where red represents up-regulated and green represents down-regulated.

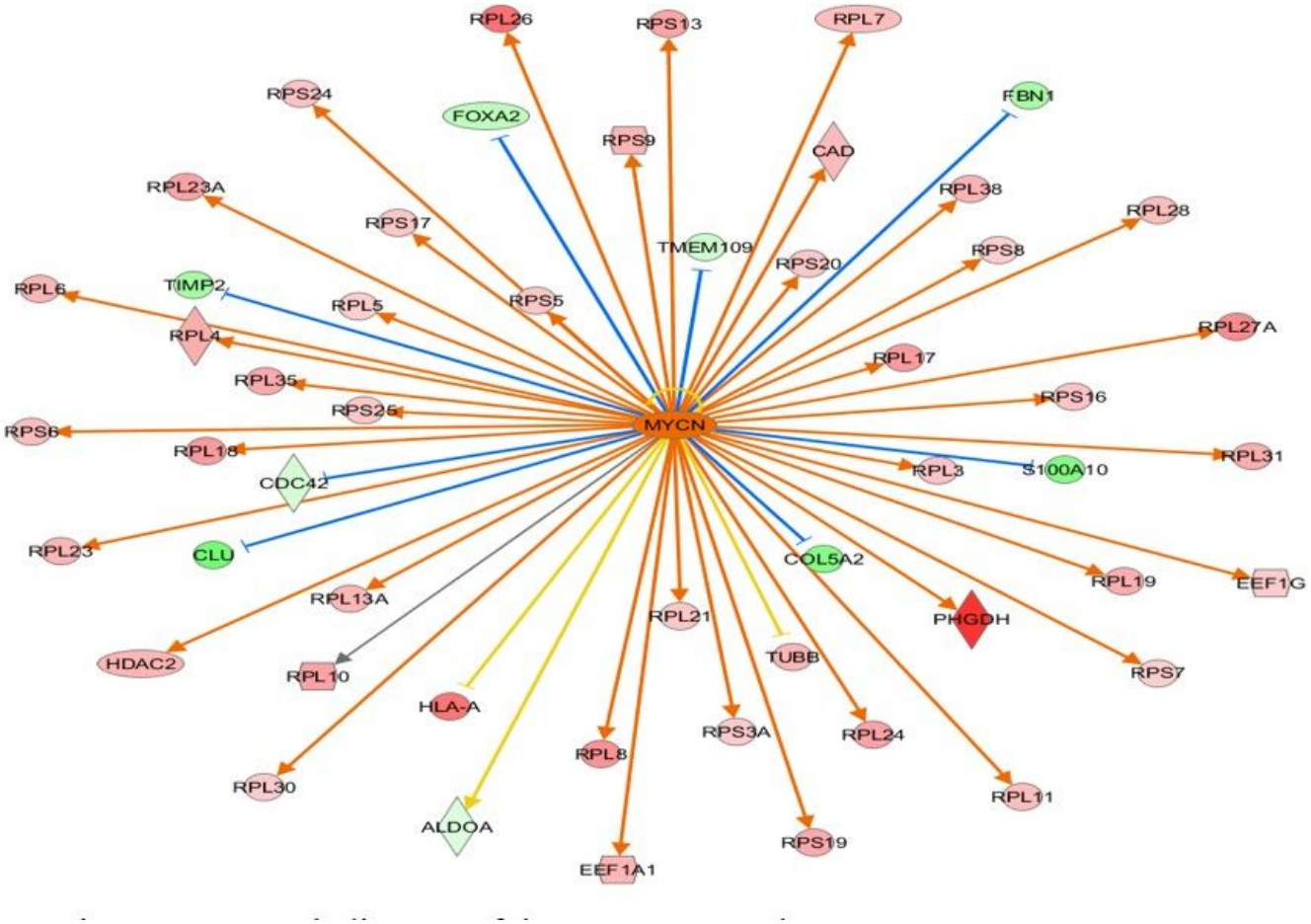

Figure 6

Network diagram of the upstream regulator, MYCN. The upstream regulatory sub-network diagram shows the interactions between upstream regulatory factors and their directly related downstream molecules that co-exist in the dataset. The orange line represents the consistent activation of the upstream regulator and genes, the blue line represents the consistent inhibition of the upstream regulator and genes, the yellow line represents the inconsistent expression trend between the upstream regulator and genes, and the gray line indicates that there is no predictive information related to the expression status in the data set.

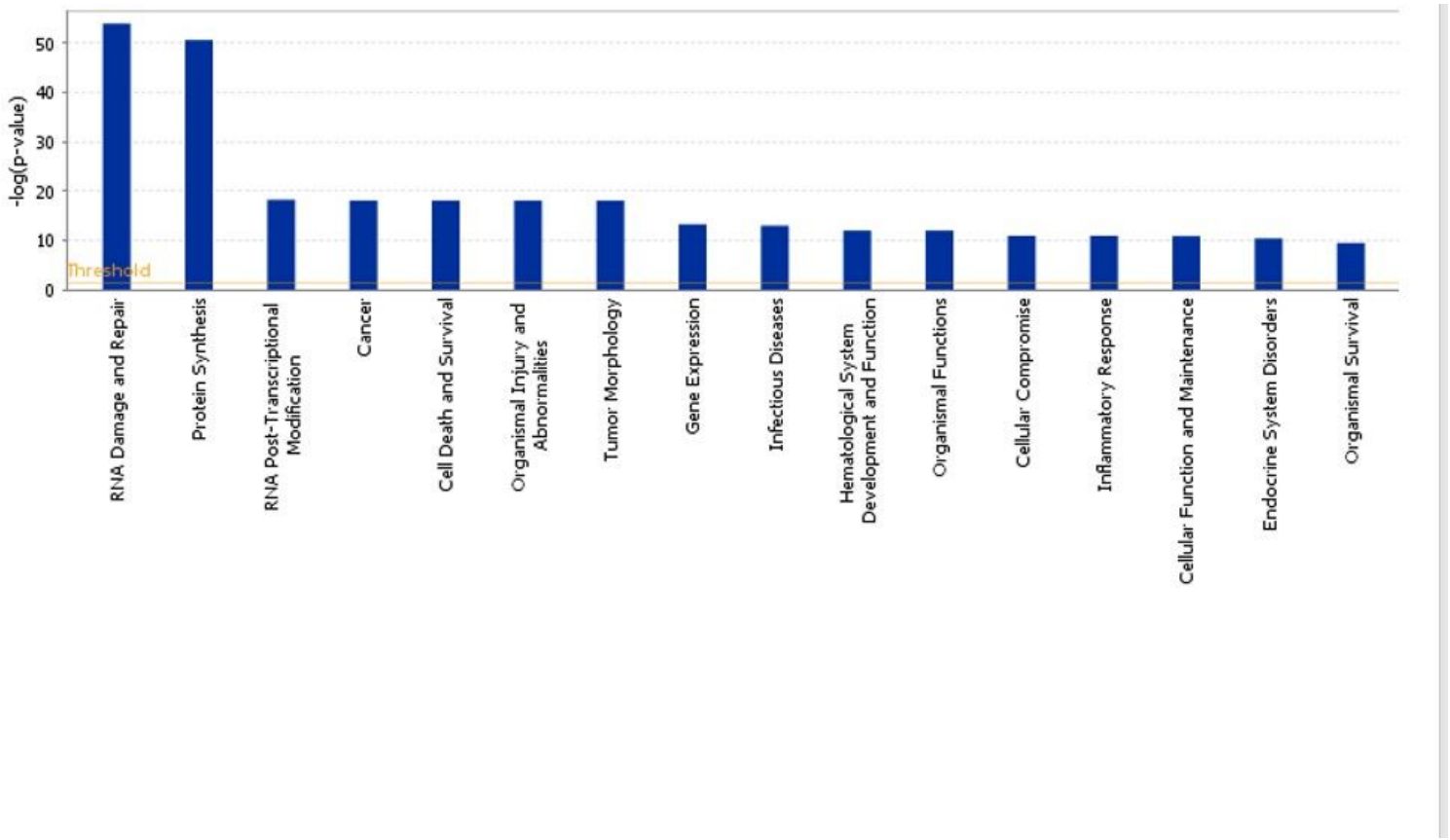

Figure 7

Significant enrichment of differentially expressed proteins in disease and function. The x-coordinate is the pathway name and the y-coordinate is the significance level of enrichment (negative logarithmic transformation with base 10).

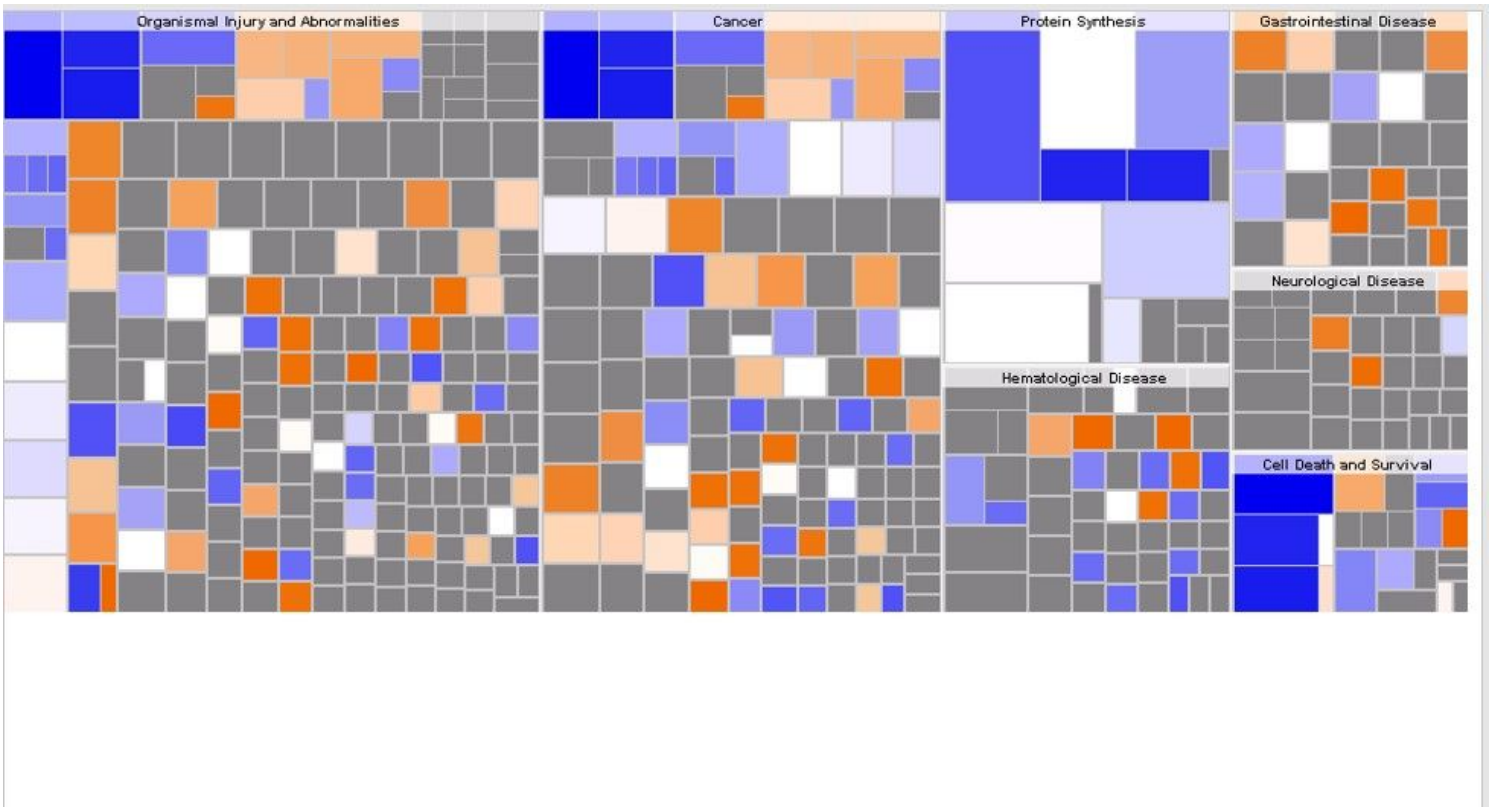

Figure 8

Relationship between disease and functional activation and inhibition of differentially expressed proteins. Orange represents the activation of disease or functional state (z-score > 0), blue represents the inhibition of disease or functional state (z-score < 0), and gray represents the undetermined disease or functional state (z-score could not be calculated). According to the internal algorithm and criteria of IPA, a z-score ≥ 2 means that the disease or function is significantly activated and a z-score ≤ -2 means that the disease or function is significantly suppressed.

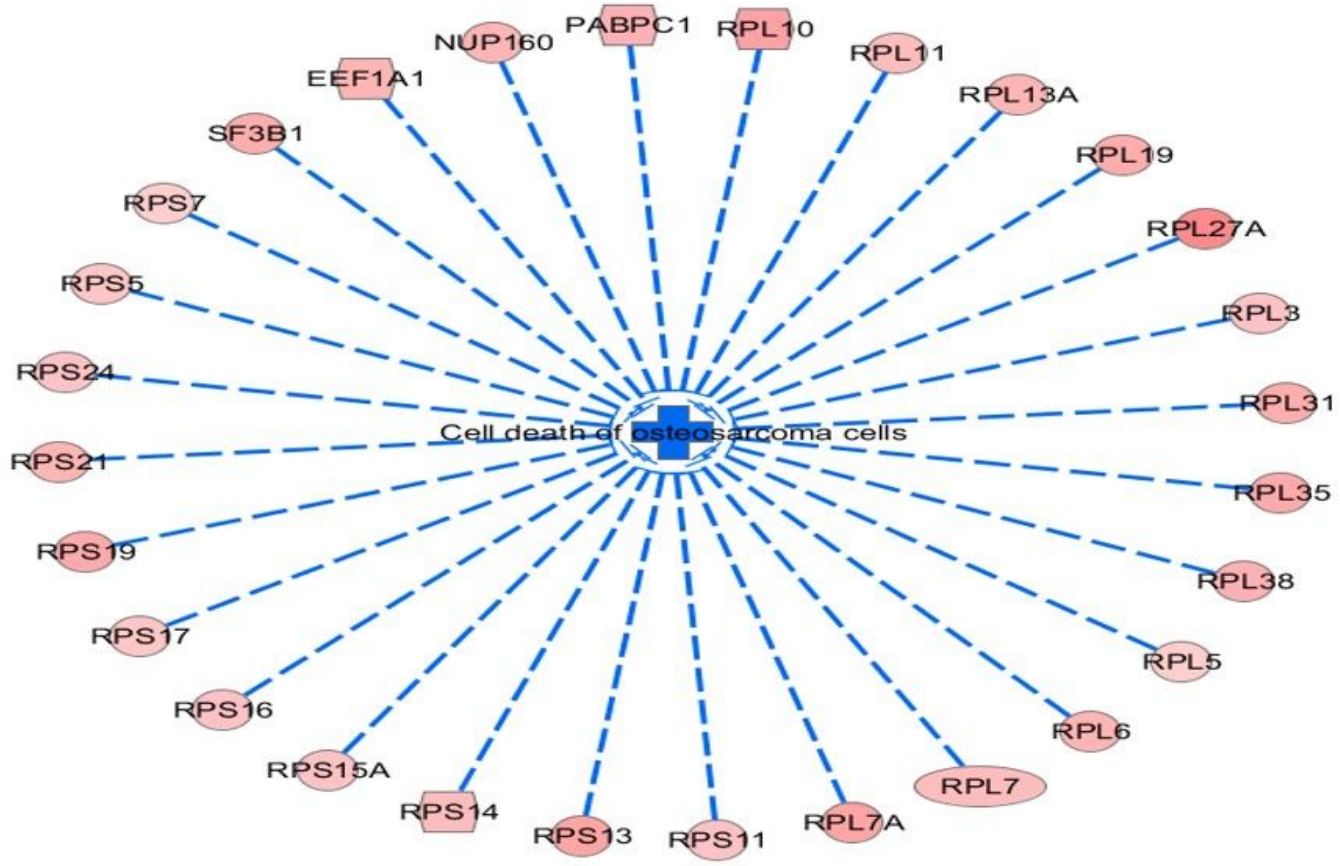

Figure 9

The relationship between the activation and inhibition of differentially expressed proteins and osteosarcoma cell death. In the network are contained in a specified disease or all differences related to the function of genes, and shows them based on the knowledge base of the Ingenuity of possible interactions between relations and the expression of trend changes. The orange line indicates that the change in gene expression level can stimulate the function, the blue line indicates that the change in gene expression level can inhibit the function, the yellow line indicates that the effect of the change in gene expression level on the function is inconsistent with the existing literature reports, and the gray line indicates that the regulatory relationship is unknown.

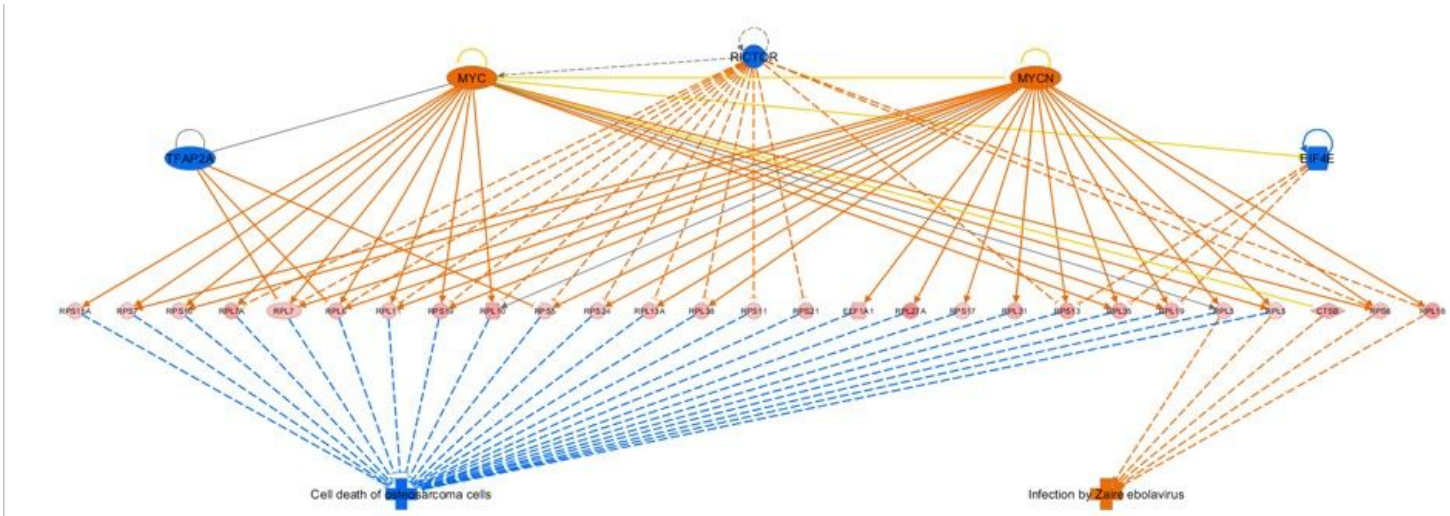

Figure 10

The differentially expressed proteins in the data set were correlated with the interaction between the regulator and osteosarcoma cell death. The orange line indicates that the change in gene expression level can activate the function, the blue line indicates that the change in gene expression level can inhibit the function, the yellow line indicates that the effect of the change in gene expression level on the function is inconsistent with the existing literature reports, and the gray line indicates that the regulatory relationship is unknown.

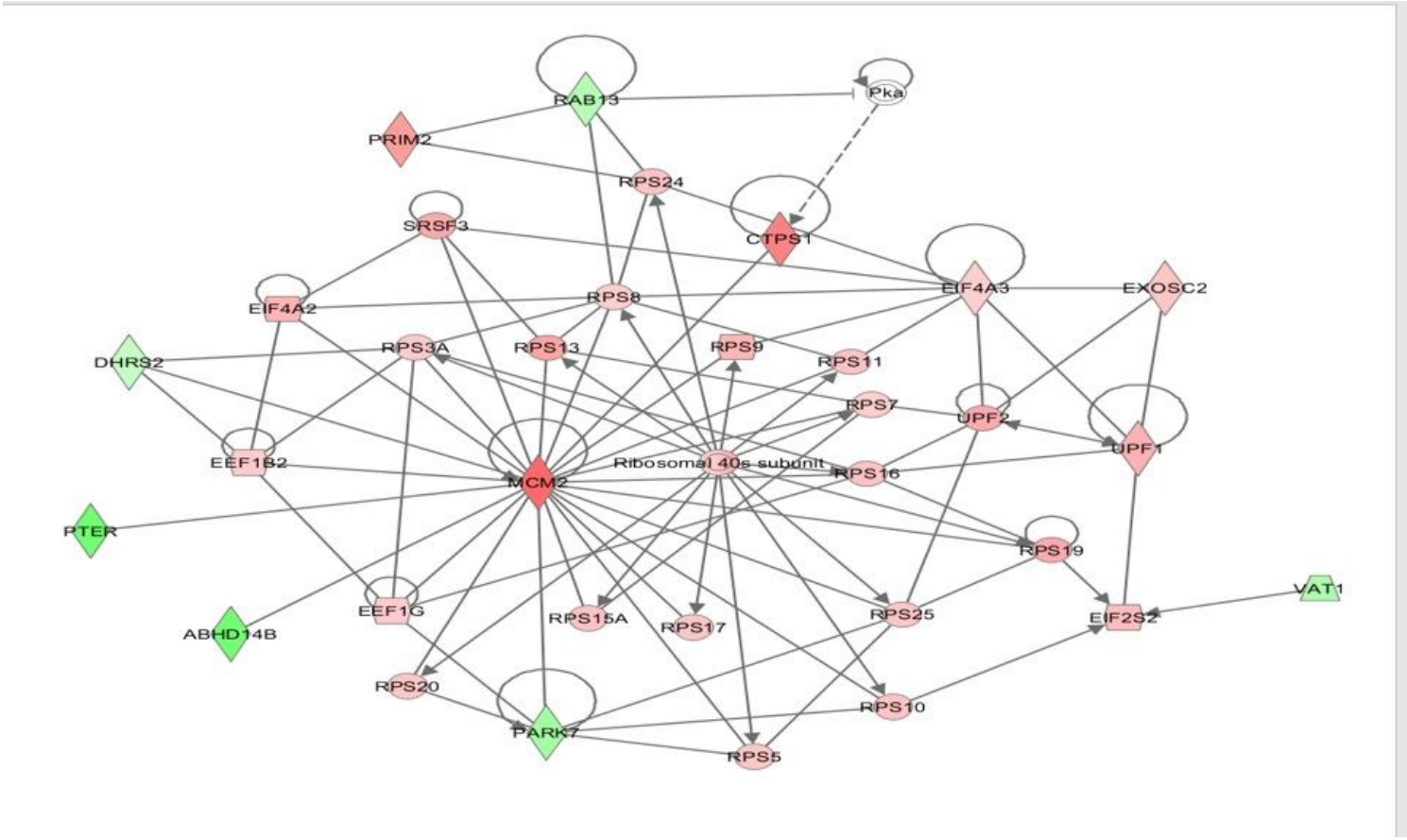

Figure 11

Interaction network of molecules related to cancer, protein synthesis, and RNA damage and repair. The orange line indicates that the change in gene expression level can activate function, the blue line indicates that the change in gene expression level can inhibit function, the yellow line indicates that the effect of the change in gene expression level on function is inconsistent with the existing literature, and the gray line indicates that the regulatory relationship is unknown.

Supplementary Files

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

- supplementarytable7.xlsx
- supplementarytable5.xlsx
- supplementarytable8.xlsx
- supplementarytable6.xlsx
- supplementarytable2..xlsx
- supplementarytable1..xlsx
- supplementarytable3..xlsx
- supplementarytable4.xlsx