Research Article

Expression and Gene Regulation Network of Metabolic Enzyme Phosphoglycerate Mutase Enzyme 1 in Breast Cancer Based on Data Mining

Yongxuan Wang,¹ Xifeng Xiong,² Xing Hua,¹ and Wei Liu³

¹Department of Pathology, Guangzhou Red Cross Hospital, Jinan University, Guangzhou, Guangdong 510220, China
²Guangzhou Institute of Traumatic Surgery, Guangzhou Red Cross Hospital, Jinan University, Guangzhou, Guangdong 510220, China
³Department of Breast Surgery, Guangzhou Red Cross Hospital, Jinan University, Guangzhou, Guangdong 510220, China

Correspondence should be addressed to Xing Hua; huaxingtime@126.com and Wei Liu; lwei7@mail2.sysu.edu.cn

Received 19 October 2020; Revised 9 January 2021; Accepted 19 January 2021; Published 10 February 2021

Academic Editor: Fu-Ming Tsai

Copyright © 2021 Yongxuan Wang et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

The metabolic enzyme phosphoglycerate mutase enzyme 1 (PGAM1) is a key enzyme in the glycolysis pathway, and glycolysis is closely related to cancer progression, suggesting that PGAM1 may have important functions in breast cancer. We used sequencing data from the Oncomine database and UALCAN database to analyze the expression of PGAM1 and its influence on the clinicopathological characteristics of breast cancer. LinkedOmics was used to identify genes related to PGAM1 expression, kinases, miRNAs, and transcription factors that were significantly related to PGAM1 through GSEA. cBioPortal was used to identify the alternation frequency and form of PGAM1 in breast cancer. The expression level of PGAM1 in breast cancer was significantly higher than that in normal tissues. Moreover, the expression level of PGAM1 is closely related to the molecular subtype and TP53 mutation status. The expression level of PGAM1 in HER2-positive and triple-negative tumors was significantly higher than that of luminal type. The expression level of PGAM1 in TP53-mutant tumors was higher than that in non-TP53-mutant tumors. In addition, the overall survival of patients with high PGAM1 expression was significantly worse than that of patients with low expression ($P = 0.0077$). Through GSEA analysis, we found multiple kinases, miRNAs, and transcription factors significantly related to PFKFB4. cBioPortal analysis showed that the mutation rate of PGAM1 in breast cancer was relatively low (4%), and the main form of mutation was high mRNA expression. This study suggests that PGAM1 is a potential diagnostic and prognostic marker in breast cancer. Through data mining, we revealed the potential regulatory network information of PGAM1, laying a foundation for further research on the role of PGAM1 in breast cancer.

1. Introduction

Breast cancer is the most common malignant tumor in women in the world and the leading cause of cancer-related deaths in women [1]. Although the survival of early breast cancer has been significantly improved, there are still some patients who subsequently relapse and metastasize. The development of various targeted drugs has prolonged the survival time of patients and has made breakthrough progress in the treatment of advanced breast cancer. However, patients with advanced breast cancer inevitably develop primary or continued resistance to targeted drugs. The pathogenesis of breast cancer is extremely complex, involving processes such as cell cycle regulation and signal transduction, reflecting the function and interaction of multiple genes in multiple steps. Therefore, identifying more molecular markers for breast cancer is expected to develop more new molecular targeted therapeutic drugs.

Phosphoglycerate mutase enzyme 1 (PGAM1) is a vital glycolytic protein that catalyzes the reversible reaction of 3-phosphoglycerate (3-PG) and 2-phosphoglyceride (2-PG) [2, 3]. The regulatory role of PGAM1 in aggressive tumors...
TCGA Breast statistics
Over-expression gene rank: 3923 (in top 20%)
Reporter: A_23_P138620

1. Breast (61)
   P value: 4.25E-7
t-test: 5.186
Fold change: 1.365

Curtis Breast statistics
Over-expression gene rank: 661 (in top 4%)
Reporter: ILMN_2112417

1. Breast (144)
   P value: 2.03E-5
t-test: 5.767
Fold change: 1.594

Figure 1: Continued.
Ma Breast 4 statistics

Over-expression gene rank: 3122 (in top 17\%)

\[ \text{Reporter: } \text{g45035752}_3\text{p}_\text{a}_\text{at} \]

\[ P \text{ value: 0.016} \]

\[ t\text{-test: 2.416} \]

\[ \text{Fold change: 1.421} \]

Farmer Breast statistics

Over-expression gene rank: 661 (in top 4\%)

\[ \text{Reporter: } \text{200886}_s\_\text{at} \]

\[ P \text{ value: 0.045} \]

\[ t\text{-test: 1.765} \]

\[ \text{Fold change: 1.106} \]

1. Breast (14)
2. Invasive ductal breast carcinoma (9)

(c)

1. Breast-like subtype of invasive breast carcinoma (16)
2. Luminal-like subtype of invasive breast carcinoma (27)

(d)

**Figure 1**: Continued.
has received increasing attention in recent years. To date, many studies have shown that PGAM1 is highly expressed in various tumors such as pancreatic cancer [3], oral squamous cell carcinoma [4], and hepatocellular carcinoma [5]. These studies indicate that PGAM1 is a novel oncogene that may have a regulatory role in breast cancer. However, the expression of PGAM1 in breast cancer and its effect on the prognosis are currently unclear. Therefore, we studied PGAM1 expression and mutations in breast cancer patient data in The Cancer Genome Atlas (TCGA) and various public databases. Using multidimensional analysis, we analyzed PGAM1-related genomic changes and functional networks in breast cancer. Therefore, our results may reveal new molecular targets for breast cancer diagnosis and treatment.

2. Materials and Methods

2.1. Oncomine Analysis. Oncomine (http://www.oncomine.org) has 715 gene expression datasets and data from 86,733 cancer tissues and normal tissues [6]. It is currently a widely used bioinformatics data analysis platform. Oncomine integrated RNA-seq and DNA-seq data from GEO database, TCGA database, and published literature. After logging in Oncomine, we can see a search box and filter on the left side of the webpage. The filter catalog is divided into several parts, including “Primary Filters”, “Sample Filters”, “Dataset Filters”, and “Concept Filters”. Primary Filters can be used to select analysis types, datasets, data sources, cancer types, etc. For example, in order to analyze the expression of PGAM1 in breast cancer, selecting “Breast cancer”, “Cancer vs. Normal”, and “Clinical Specimen” in the data filter is to know the expression level of PGAM1 in breast cancer tissues and normal tissues. We analyzed the mRNA expression of PGAM1 in breast cancer in multiple cohorts in the Oncomine 4.5 database, including TCGA breast cancer, Curtis breast cancer, Ma Breast 4, and Farmer Breast datasets. The difference in the expression of PGAM1 in breast cancer tissue and the corresponding normal tissue was evaluated, and the difference related to \( P < 0.05 \) was considered significant. Fold change was used for differential expression analysis.

2.2. ENCORI Analysis. ENCORI database (http://starbase.sysu.edu.cn/index.php) is an open platform for studying miRNA-ncRNA, miRNA-mRNA, ncRNA-RNA, RNA-RNA, RBP-ncRNA, and RBP- in CLIP-seq, degradome-seq, and RNA-RNA interaction group data. Combining gene expression data of 32 cancers derived from 10882 RNA-seq and 10546 miRNA-seq data, ENCORI allows researchers to perform pan-cancer analysis of RNA-RNA and RBP-RNA interactions. ENCORI also provides a platform for survival and differential expression analysis of miRNA, IncRNA, pseudogenes, and mRNA. In this study, ENCORI database was used to further verify the expression level of PGAM1 in breast cancer compared with normal tissues. At the same time, the ENCORI database was used to evaluate the effect
Expression of PGAM1 in BRCA based on sample types

(a) Expression of PGAM1 in BRCA based on breast cancer subclasses

(b) Expression of PGAM1 in BRCA based on nodal metastasis status

(c) Figure 2: Continued.
Expression of PGAM1 in BRCA based on TP53 mutation status

(d)

Expression of PGAM1 in BRCA based on patient’s race

(e)

Expression of PGAM1 in BRCA based on nodal metastasis status

(f)

Figure 2: PGAM1 transcription levels in a subgroup of breast cancer patients, stratified according to molecular subtype, stage, TP53 mutation status, and other criteria (UALCAN). (a) Box plot shows the relative expression of PGAM1 in normal tissues and breast cancer tissues. (b) Box plot shows the relative expression of PGAM1 in normal tissues or breast cancer tissues of different molecular subtypes. (c) Box plot shows the relative expression of PGAM1 in normal individuals or breast cancer patients with different N stages. (d) Box plot shows the relative expression of PGAM1 in normal tissues or breast cancer tissues with different TP53 mutation states. (e) Box plot shows the relative expression of PGAM1 in normal tissues or breast cancer tissues of Caucasian, African American, or Asian ethnicity. (f) Box plot shows the relative expression of PGAM1 in normal tissues or stage 1, 2, 3, or 4 breast cancer tissues. Data are presented as the mean ± SE. *P < 0.05, **P < 0.01, ***P < 0.001.
Figure 3: Continued.
2.3. UALCAN Analysis. UALCAN (http://ualcan.path.uab.edu) includes RNA-seq data of 31 cancer types in the TCGA database and has corresponding clinical pathological characteristics [7]. It is mainly based on the relevant cancer data in the TCGA database for biomarker identification, expression difference analysis, survival analysis, etc. The analysis platform can be used to analyze the relationship between a single gene and cancer stage, tumor grade, age, or other clinicopathological characteristics. This study used the UALCAN platform to analyze the association between PGAM1 and the clinicopathological characteristics of breast cancer, including molecular subtype, stage, and TP53 mutation status. For example, in the “Analysis” module, we enter “PGAM1” in the input box, select the cancer type as “breast cancer”, and finally click the “Explore” button. For the input genes, relevant analyzable items will be jumped out, as well as information links of genes in other databases.

2.4. LinkedOmics Analysis. The LinkedOmics database (http://www.linkedomics.org/login.php) is a web-based platform that can be used to analyze 32 TCGA cancer-related cubes [8]. It usually includes five steps, namely, “Select Cancer Cohort”, “Select Search Dataset”, “Select Search Dataset Attribute”, “Select Target Dataset”, and “Select Statistical Method”. During the analysis, we selected “TCGA_BRCA”, “HiSeq RNA”, “PGAM1”, “RNAseq”, and “Pearson Correlation test”, respectively. In this study, the LinkFinder module was used to analyze the differentially expressed genes related to PGAM1 in the breast cancer cohort of the TCGA database, and the genes
Antigen processing and presentation
Chromosome segregation
Granulocyte activation
Mitotic cell cycle phase transition
Ribonucleotide metabolic process
Mitochondrial transport
Pattern specification process
Multicellular organismal signaling
Muscle organ development
Heart morphogenesis

Normalized enrichment score

Mitochondrial inner membrane
Condensed chromosome
Spindle
Endoplasmic reticulum-Golgi intermediate
Tertiary granule
Mitochondrial matrix
Myelin sheath
Chromosomal region
Vesicle lumen
Coated vesicle

Electron transfer activity
Nuclease activity
Cofactor binding
GTPase activity
Transferase activity, transferring glycosyl groups
Cell adhesion molecule binding
Catalytic activity, acting on RNA
Magnesium ion binding
DNA-binding transcription activator activity, RNA polymerase II-specific
Insulin receptor binding

Normalized enrichment score

Figure 4: Continued.
related to the significant expression of PGAM1 in breast cancer were analyzed according to the Pearson correlation coefficient. The results were displayed in volcano maps and heat maps. We used LinkFinder module gene set enrichment analysis (GSEA) to perform GO (cell component (CC), biological process (BP), and molecular function (MF)) analysis of differentially related genes, KEGG signal pathway analysis, kinase target enrichment, miRNA target enrichment, and transcription factor target enrichment.

2.5. GeneMANIA Analysis. GeneMANIA (http://www.genemania.org) was used to construct a protein-protein interaction (PPI) network, generate hypotheses about the function of PGAM1 gene, and analyze protein networks that may interact with PGAM1 protein. The PPI network constructed by GeneMANIA is based on many publicly available large-scale biological datasets to find related genes. These protein-protein interactions include physical interactions, coexpression, predicted potential effects, cross-linking between signaling pathways, and colocalized expression in cells. But GeneMANIA does not provide quantitative values between related proteins. These proteins include possible direct or indirect interactions with PGAM1 protein, similar to coexpression, colocalization, or direct binding.

Figure 4: GO annotation and KEGG pathway significantly enriched by PGAM1 in breast cancer. GSEA was used to analyze the GO annotation and KEGG pathway of PGAM1 coexpressed genes in breast cancer. (a) Biological process. (b) Cell component. (c) Molecular function. (d) KEGG pathway analysis.

Table 1: PGAM1-related kinases, miRNAs, and transcription factor-target networks in breast cancer (LinkedOmics).

| Enriched category       | Gene set                  | Leading-edge-num | FDR     |
|-------------------------|---------------------------|------------------|---------|
| Kinase target           | Kinase_CDK1               | 32               | 0.0081983 |
|                         | Kinase_PLK1               | 16               | 0.024595 |
|                         | Kinase_CHEK1              | 14               | 0.025620 |
|                         | Kinase_AURKB              | 12               | 0.028182 |
|                         | Kinase_BRAF               | 5                | 0.030060 |
|                         | ACAACTT, miR-382          | 8                | 0.0084138 |
|                         | TATCTGG, miR-488          | 5                | 0.17389 |
| miRNA target            | ACAACCT, miR-453          | 3                | 0.20824 |
|                         | ACTGAAA, miR-30A-3P, miR-30E-3P | 20   | 0.26267 |
|                         | GCGCCTT, miR-525, miR-524 | 4                | 0.26281 |
|                         | VSFACI_01                | 19               | 0.072263 |
|                         | V5RSRFC4_Q2              | 34               | 0.085812 |
| Transcription factor target | YCATTTAA_UNKNOWN       | 59               | 0.10194 |
|                         | V$MEF2_03                | 38               | 0.10457 |
|                         | CTAWWWATA_V5RSRFC4_Q2    | 42               | 0.11103 |
GeneMANIA report
Created on: 16 September 2020 11:00:10
Last database update: 13 March 2017 00:00:00
Application version: 3.6.0

Physical interactions
Co-expression
Predicted
Co-localization

Pathway
Genetic interactions
Shared protein domains

Functions
N/A

Figure 5: Continued.
Figure 5: Continued.

**Search parameters**

| Organism     | Homo sapiens (human) |
|--------------|----------------------|
| Genes        | PGAM1                |
| Network weighting | Automatically selected weighting method |
| Networks A   |                      |
| Bahr-Bowler-2013 , Bailey-Hieter-2015 , Bandyopadhyay-Ideker-2010 , Bantscheff-Drewes-2011 , Barr-Knapp-2009 , Barrios-Rodiles-Wrana-2005 , Behrends-Harper-2010 , Behzadnia-Lührmann-2007 , Bennett-Harper-2010 , Benzinger-Herneking-2005 , Berggård-James-2006 , Bett-Hay-2013 , Bhatnagar-Attie-2014 , Bild-Nevins-2006 B , BIOGRID-SMALL-SCALE-STUDIES , BIOGRID-SMALL-SCALE-STUDIES , Blandin-Richard-2013 , Blomen-Brummelkamp-2015 , Blomen-Brummelkamp-2015 , Bogachek-Weigel-2014 , Boldrick-Rehman-2002 , Bonacci-Soubeyran-2014 , Bouwmeester-Superti-Furga-2004 , Brajnovic-Drewes-2004 , Brehme-Superti-Furga-2009 , Bruderer-Hay-2011 , Burington-Shaughnessy-2008 , Butland-Hayden-2014 , Byron-Humphries-2012 |
| C            |                      |
| Cai-Conaway-2007 , Camargo-Brandon-2007 , Campos-Reinberg-2015 , Cao-Chinnaiyan-2014 , Carmon-Liu-2014 , CELL_MAP , Chen-Brown-2002 , Chen-Ge-2013 , Chen-Huang-2014 , Chen-Zhang-2013 , Christianson-Kopito-2011 , Cloutier-Coulombe-2013 , Colland-Gauthier-2004 , Corominas-Iakovcheva-2014 , Couzens-Gingras-2013 , Cox-Rizzino-2013 , Coyaud-Raught-2015 |
| D            |                      |
| Danielsen-Nielsen-2011 , Dart-Wells-2015 , de Hoog-Mann-2004 , Diner-Cristea-2015 , Dobbin-Giordano-2005 , Drissi-Boisvert-2015 , Dyer-Sobral-2010 |
| E            |                      |
| Emanuele-Elledge-2011 , Emdal-Olsen-2015 , Ewing-Figeys-2007 |
| F            |                      |
| Fenner-Prehn-2010 , Floyd-Pagliarini-2016 , Foerster-Ritter-2013 , Fogerons-Lange-2013 , Foster-Marshall-2013 , Freibaum-Taylor-2010 |
| G            |                      |
| Gabriel-Baumgrass-2016 , Galligan-Howley-2015 , Gao-Reinberg-2012 , Gautier-Hall-2009 , Giannone-Liu-2010 , Glatter-Gstaiger-2009 , Gloeckner-Ueffing-2007 , |
Figure 5: Continued.
Figure 5: Continued.
### Genes

| Gene       | Description                                                                 | Rank |
|------------|------------------------------------------------------------------------------|------|
| PGAM1      | phosphoglycerate mutase 1 [Source:HGNC Symbol;Acc:HGNC:8888]                 | N/A  |
| PGAM2      | phosphoglycerate mutase 2 [Source:HGNC Symbol;Acc:HGNC:8889]                 | 1    |
| GAPDH      | glyceraldehyde-3-phosphate dehydrogenase [Source:HGNC Symbol;Acc:HGNC:4141]  | 2    |
| GPI        | glucose-6-phosphate isomerase [Source:HGNC Symbol;Acc:HGNC:4458]            | 3    |
| PGAM4      | phosphoglycerate mutase family member 4 [Source:HGNC Symbol;Acc:HGNC:21731]  | 4    |
| ALDOA      | aldolase, fructose-bisphosphate A [Source:HGNC Symbol;Acc:HGNC:414]          | 5    |
| TPI1       | triosephosphate isomerase 1 [Source:HGNC Symbol;Acc:HGNC:12009]             | 6    |
| AP2M1      | adaptor related protein complex 2 mu 1 subunit [Source:HGNC Symbol;Acc:HGNC:564] | 7    |
| BPGM       | bisphosphoglycerate mutase [Source:HGNC Symbol;Acc:HGNC:1093]               | 8    |
| TYMS       | thymidylate synthetase [Source:HGNC Symbol;Acc:HGNC:12441]                  | 9    |
| TIGAR      | TP53 induced glycolysis regulatory phosphatase [Source:HGNC Symbol;Acc:HGNC:1185] | 10   |
| CHUK       | conserved helix-loop-helix ubiquitous kinase [Source:HGNC Symbol;Acc:HGNC:1974] | 11   |
| PGK1       | phosphoglycerate kinase 1 [Source:HGNC Symbol;Acc:HGNC:8896]                | 12   |
| PRCP       | prolylcarboxypeptidase [Source:HGNC Symbol;Acc:HGNC:9344]                   | 13   |
| PGAM5      | PGAM family member 5, mitochondrial serine/threonine protein phosphatase [Source:HGNC Symbol;Acc:HGNC:28763] | 14   |
| NDUFB8     | NADH:ubiquinone oxidoreductase subunit B8 [Source:HGNC Symbol;Acc:HGNC:7703] | 15   |
| PFKP       | phosphofructokinase, platelet [Source:HGNC Symbol;Acc:HGNC:8878]            | 16   |
| HSD17B10   | hydroxysteroid 17-beta dehydrogenase 10 [Source:HGNC Symbol;Acc:HGNC:4800]   | 17   |
| MDM2       | MDM2 proto-oncogene [Source:HGNC Symbol;Acc:HGNC:6973]                      | 18   |
| PKM        | pyruvate kinase, muscle [Source:HGNC Symbol;Acc:HGNC:9021]                  | 19   |
| LDHA       | lactate dehydrogenase A [Source:HGNC Symbol;Acc:HGNC:6535]                  | 20   |

*Figure 5: Continued.*
## Networks

### Physical Interactions

| Reference                                      | Percentage |
|------------------------------------------------|------------|
| Matsumoto-Nakayama-2005                        | 67.64%     |
| Singh-Moore-2012                               | 5.54%      |
| Petschnigg-Stagljar-2014                       | 4.12%      |
| Whisenant-Salomon-2015                         | 3.90%      |
| Lu-Zhang-2013                                  | 2.70%      |
| Phillips-Corn-2013                             | 2.43%      |
| Wagner-Choudhary-2011                          | 1.88%      |
| Behzadnia-Lührmann-2007                        | 1.86%      |
| Leung-Jones-2014                               | 1.55%      |
| Maréchal-Zou-2014                              | 1.43%      |
| Berggård-James-2006                            | 1.41%      |

Physical Interactions with 311 interactions from BioGRID

Physical Interactions with 301 interactions from iRefIndex

Physical Interactions with 237 interactions from BioGRID

Physical Interactions with 281 interactions from iRefIndex

Physical Interactions with 134 interactions from BioGRID

Physical Interactions with 1,158 interactions from iRefIndex

Physical Interactions with 112 interactions from iRefIndex

Physical Interactions with 190 interactions from iRefIndex

Physical Interactions with 976 interactions from iRefIndex

---

**Figure 5:** Continued.
| Physical Interactions | 67.64% |
|-----------------------|--------|
| Berggård-James-2006   | 1.28%  |
| Neganova-Lako-2011    | 1.21%  |
| McCracken-Blencowe-2005 | 1.16%  |
| Mak-Moffat-2010       | 1.03%  |
| Murakawa-Landthaler-2015 | 1.00%  |
| Liéres-Lamond-2010    | 0.99%  |
| Nicholson-Hupp-2014   | 0.93%  |
| Jones-MacBeath-2006   | 0.87%  |
| McFarland-Nussbaum-2008 | 0.83%  |
| Guarani-Harper-2014   | 0.83%  |
| Hill-Livingston-2014  | 0.83%  |

**Figure 5: Continued.**
| Publication                          | Percentage |
|------------------------------------|------------|
| Bett-Hay-2013                       | 67.64%     |
| The P-body component USF2/PAN2 is a novel regulator of HIF1α mRNA stability. Bett et al (2013). *Biochim Biophys Acta* |
| Physical Interactions with 319 interactions from iRefIndex | 0.79%      |
| Stees-Gevaert-2014                  | 0.73%      |
| A COFIRADIC protocol to study protein ubiquitination. Stees et al (2014). *J Proteomics* |
| Physical Interactions with 1,337 interactions from iRefIndex | 0.67%      |
| Danielsen-Nielsen-2011              | 0.65%      |
| Mass spectrometric analysis of lysine ubiquitylation reveals promiscuity at site level. Danielsen et al (2011). *Mol Cell Proteomics* |
| Physical Interactions with 2,479 interactions from iRefIndex | 0.65%      |
| Jeronimo-Coulombe-2007              | 0.65%      |
| Systematic analysis of the protein interaction network for the human transcription machinery reveals the identity of the 7SK capping enzyme. Jeronimo et al (2007). *Mol Cell* |
| Physical Interactions with 699 interactions from BioGRID | 0.62%      |
| Xu-Ye-2012                          | 0.61%      |
| SCITA recognizes a canonical ubiquitin-like domain in the Bag6-Ubi4A-Trcl35 complex to promote endoplasmic reticulum-associated degradation. Xu et al (2012). *Cell Rep* |
| Physical Interactions with 235 interactions from iRefIndex | 0.60%      |
| Gloeckner-Ueffing-2007              | 0.60%      |
| A novel tandem affinity purification strategy for the efficient isolation and characterization of native protein complexes. Gloeckner et al (2007). *Proteomics* |
| Physical Interactions with 100 interactions from BioGRID | 0.59%      |
| Fogeran-Lange-2013                  | 0.59%      |
| LGALS8BP regulates centrosome biogenesis and centrosome hypertrophy in cancer cells. Fogeran et al (2013). *Nat Commun* |
| Physical Interactions with 578 interactions from iRefIndex | 0.59%      |
| Freibaum-Taylor-2010                 | 0.59%      |
| Global analysis of TDP-43 interacting proteins reveals strong association with RNA splicing and translation machinery. Freibaum et al (2010). *J Proteome Res* |
| Physical Interactions with 216 interactions from iRefIndex | 0.59%      |
| Zhou-Conrads-2004                    | 0.59%      |
| An investigation into the human serum "interactome". Zhou et al (2004). *Electrophoresis* |
| Physical Interactions with 158 interactions from iRefIndex | 0.55%      |
| Brehme-Superti-Furga-2009            | 0.55%      |
| Charting the molecular network of the drug target Bcr-Abl. Brehme et al (2009). *Proc Natl Acad Sci U S A* |
| Physical Interactions with 578 interactions from iRefIndex | 0.55%      |
| Yu-Chow-2013                         | 0.55%      |
| VCP phosphorylation-dependent interaction partners prevent apoptosis in Helicobacter pylori-infected gastric epithelial cells. Yu et al (2013). *PLoS One* |
| Physical Interactions with 272 interactions from iRefIndex | 0.55%      |
| Hegele-Stelzl-2012 B                 | 0.55%      |
| Dynamic protein-protein interaction wiring of the human spliceosome. Hegele et al (2012). *Mol Cell* |
| Physical Interactions with 500 interactions from BioGRID | 0.55%      |

*Figure 5: Continued.*
| Physical Interactions | 67.64% |
|-----------------------|--------|
| **Weinmann-Meister-2009** | 0.54% |
| Importin 8 is a gene silencing factor that targets argonaute proteins to distinct mRNAs. Weinmann et al (2009). *Cell* |
| Physical Interactions with 96 interactions from BioGRID |

| **Narayan-Bennett-2012** | 0.53% |
|--------------------------|--------|
| Short-chain 3-hydroxacyl-coenzyme A dehydrogenase associates with a protein super-complex integrating multiple metabolic pathways. Narayan et al (2012). *PLoS One* |
| Physical Interactions with 110 interactions from BioGRID |

| **Udeshi-Carr-2012** | 0.52% |
|----------------------|--------|
| Methods for quantification of in vivo changes in protein ubiquitination following proteasome and deubiquitase inhibition. Udeshi et al (2012). *Mol Cell Proteomics* |
| Physical Interactions with 554 interactions from iRefIndex |

| **Kristensen-Foster-2012** | 0.51% |
|-----------------------------|--------|
| A high-throughput approach for measuring temporal changes in the interactome. Kristensen et al (2012). *Nat Methods* |
| Physical Interactions with 7,115 interactions from BioGRID |

| **Agrawal-Sedivy-2010** | 0.50% |
|--------------------------|--------|
| Proteomic profiling of Myc-associated proteins. Agrawal et al (2010). *Cell Cycle* |
| Physical Interactions with 104 interactions from iRefIndex |

| **Varjosalo-Superti-Furga-2013** | 0.49% |
|-----------------------------------|--------|
| Interlabatory reproducibility of large-scale human protein-complex analysis by standardized AP-MS. Varjosalo et al (2013). *Nat Methods* |
| Physical Interactions with 483 interactions from BioGRID |

| **Rowbotham-Mermoud-2011** | 0.46% |
|-----------------------------|--------|
| Maintenance of silent chromatin through replication requires SWI/SNF-like chromatin remodeler SMARCAD1. Rowbotham et al (2011). *Mol Cell* |
| Physical Interactions with 114 interactions from iRefIndex |

| **Arroyo-Aloy-2014** | 0.44% |
|----------------------|--------|
| Charting the molecular links between driver and susceptibility genes in colorectal cancer. Arroyo et al (2014). *Biochim Biophys Acta Genom Med* |
| Physical Interactions with 598 interactions from iRefIndex |

| **Barr-Knapp-2009** | 0.43% |
|---------------------|--------|
| Large-scale structural analysis of the classical human protein tyrosine phosphatase. Barr et al (2009). *Cell* |
| Physical Interactions with 164 interactions from iRefIndex |

| **Emdal-Olsen-2015** | 0.41% |
|----------------------|--------|
| Temporal proteomics of NGF-TrkA signaling identifies an inhibitory role for the E3 ligase Cbl-b in neuroblastoma cell differentiation. Emdal et al (2015). *Sci Signal* |
| Physical Interactions with 1,919 interactions from BioGRID |

| **Cox-Rizzino-2013** | 0.40% |
|----------------------|--------|
| The SOX8-interactome in brain cancer cells identifies the requirement of MSH2 and USP9X for the growth of brain tumor cells. Cox et al (2013). *PLoS One* |
| Physical Interactions with 280 interactions from iRefIndex |

| **Giannone-Liu-2010** | 0.39% |
|-----------------------|--------|
| Figure 5: Continued. |
| Physical Interactions | 67.64% |
|-----------------------|--------|
| **Giannone-Liu-2010** |        |
| The protein network surrounding the human telomere repeat binding factors TRF1, TRF2, and POT1. Giannone et al (2010). *PLoS One* |
| Physical Interactions with 279 interactions from iRefIndex |
| **Havugimana-Emili-2012** | 0.39% |
| A census of human soluble protein complexes. Havugimana et al (2012). *Cell* |
| Physical Interactions with 13,716 interactions from BioGRID |
| **IREF-DIP** | 0.39% |
| Physical Interactions with 4,470 interactions from iRefIndex |
| **Li-Dorf-2011 A** | 0.39% |
| Mapping a dynamic innate immune protein interaction network regulating type I interferon production. Li et al (2011). *Immunity* |
| Physical Interactions with 400 interactions from BioGRID |
| **Persaud-Rotin-2009** | 0.38% |
| Comparison of substrate specificity of the ubiquitin ligases Nedd4 and Nedd4-2 using proteome arrays. Persaud et al (2009). *Mol Syst Biol* |
| Physical Interactions with 239 interactions from iRefIndex |
| **Feorster-Ritter-2013** | 0.37% |
| Characterization of the EGFR interaction reveals associated protein complex networks and intracellular receptor dynamics. Feorster et al (2013). *Proteomics* |
| Physical Interactions with 159 interactions from iRefIndex |
| **Zhao-Krug-2005** | 0.36% |
| Human ISG15 conjugation targets both IFN-induced and constitutively expressed proteins functioning in diverse cellular pathways. Zhao et al (2005). *Proc Natl Acad Sci U S A* |
| Physical Interactions with 150 interactions from iRefIndex |
| **Wan-Emili-2015** | 0.36% |
| Panorama of ancient metazoan macromolecular complexes. Wan et al (2015). *Nature* |
| Physical Interactions with 16,682 interactions from BioGRID |
| **Fenner-Prehn-2010** | 0.36% |
| Expanding the substantial interactions of NEMO using protein microarrays. Fenner et al (2010). *PLoS One* |
| Physical Interactions with 133 interactions from iRefIndex |
| **Roux-Burke-2012** | 0.35% |
| A promiscuous biotin ligase fusion protein identifies proximal and interacting proteins in mammalian cells. Roux et al (2012). *J Cell Biol* |
| Physical Interactions with 115 interactions from iRefIndex |
| **Loch-Strickler-2012** | 0.33% |
| A microarray of ubiquitylated proteins for profiling deubiquitylase activity reveals the critical roles of both chain and substrate. Loch et al (2012). *Biochem Biophys Acta* |
| Physical Interactions with 145 interactions from iRefIndex |
| **IREF-MATRIXDB** | 0.32% |
| Physical Interactions with 249 interactions from iRefIndex |

*Figure 5: Continued.*
Physical Interactions

Diner-Cristea-2015

Interactions of the Antiviral Factor Interferon Gamma-Inducible Protein 16 (IFIT1) Mediate Immune Signaling and Herpes Simplex Virus-1 Immunosuppression. Diner et al. (2015). Mol Cell Proteomics

Hein-Mann-2015

A human interactome in three quantitative dimensions organized by stoichiometries and abundances. Hein et al. (2015). Cell

Lipp-Guthrie-2015

SR protein kinases promote splicing of nonconsensus introns. Lipp et al. (2015). Nat Struct Mol Biol

Li-Chen-2015

Proteomic analyses reveal distinct chromatin-associated and soluble transcription factor complexes. Li et al. (2015). Mol Syst Biol

Wang-Xu-2015

Interaction of amyotrophic lateral sclerosis/frontotemporal lobar degeneration-associated fused-in-sarcoma with proteins involved in metabolic and protein degradation pathways. Wang et al (2015). Neurobiol Aging

Roy-Pardo-2014

hnRNP A1 couples nuclear export and translation of specific mRNAs downstream of FGFR-2/Smad3 signalling. Roy et al (2014). Nucleic Acids Res

Kotlyar-Jurisica-2015

In silico prediction of physical protein interactions and characterization of interactome orphans. Kotlyar et al (2015). Nat Methods

Barrios-Rodiles-Wrana-2005

High-throughput mapping of a dynamic signaling network in mammalian cells. Barrios-Rodiles et al (2005). Science

BIOGGRID-SMALL-SCALE-STUDIES

Physical Interactions with 58,871 interactions from BioGRID

Sowa-Harper-2009

Defining the human double ubiquitinating enzyme interaction landscape. Sowa et al (2009). Cell

Physical Interactions with 1,569 interactions from BioGRID

Ramachandran-LaBaer-2004

Self-assembling protein microarrays. Ramachandran et al (2004). Science

Physical Interactions with 112 interactions from iRefIndex

Oshikawa-Nakayama-2012

Proteome-wide identification of ubiquitylated sites by conjugation of engineered lysine-null ubiquitin. Oshikawa et al (2012). J Proteome Res

Physical Interactions with 116 interactions from iRefIndex

Bonacci-Soubeyran-2014

Physical Interactions with 116 interactions from iRefIndex

Figure 5: Continued.
| Physical Interactions                  | 67.64% |
|---------------------------------------|--------|
| **Bonacci-Soubeyran-2014**            |        |
| Identification of new mechanisms of cellular response to chemotherapy by tracking changes in post-translational modifications by ubiquitin and ubiquitin-like proteins. Bonacci et al (2014). *J Proteome Res* |        |
| Physical Interactions with 937 interactions from iRefIndex |        |
| **Kim-Gygi-2011**                     | 0.26%  |
| Systematic and quantitative assessment of the ubiquitin-modified proteome. Kim et al (2011). *Mol Cell* |        |
| Physical Interactions with 1,345 interactions from iRefIndex |        |
| **Taipale-Lindquist-2012**            | 0.26%  |
| Quantitative analysis of Hsp90-client interactions reveals principles of substrate recognition. Taipale et al (2012). *Cell* |        |
| Physical Interactions with 716 interactions from iRefIndex |        |
| **Wilker-Yaffe-2007**                 | 0.25%  |
| 14-3-3-sigma controls mitotic translation to facilitate cytokinesis. Wilker et al (2007). *Nature* |        |
| Physical Interactions with 110 interactions from iRefIndex |        |
| **Jin-Pawson-2004**                   | 0.24%  |
| Proteomic, functional, and domain-based analysis of in vivo 14-3-3 binding proteins involved in cytoskeletal regulation and cellular organization. Jin et al (2004). *Curr Biol* |        |
| Physical Interactions with 236 interactions from iRefIndex |        |
| **Oliviero-Cagnety-2015**             | 0.24%  |
| The variant Polycomb Repressor Complex 1 component PCGF1 interacts with a pluripotency sub-network that includes DPPA4, a regulator of embryogenesis. Oliviero et al (2015). *Sci Rep* |        |
| Physical Interactions with 675 interactions from BioGRID |        |
| **IREF-BIND**                         | 0.24%  |
| Physical Interactions with 3,659 interactions from iRefIndex |        |
| **Hutchins-Peters-2010**              | 0.24%  |
| Systematic analysis of human protein complexes identifies chromosome segregation proteins. Hutchins et al (2010). *Science* |        |
| Physical Interactions with 1,783 interactions from BioGRID |        |
| **Tong-Morain-2014**                  | 0.23%  |
| Proteomic analysis of the epidermal growth factor receptor (EGFR) interactome and post-translational modifications associated with receptor endocytosis in response to EGFR and stress. Tong et al (2014). *Mol Cell Proteomics* |        |
| Physical Interactions with 271 interactions from iRefIndex |        |
| **IREF-MPPI**                         | 0.23%  |
| Physical Interactions with 382 interactions from iRefIndex |        |
| **Li-Haura-2013**                     | 0.21%  |
| Perturbation of the mutated EGFR interactome identifies vulnerabilities and resistance mechanisms. Li et al (2013). *Mol Syst Biol* |        |
| Physical Interactions with 433 interactions from BioGRID |        |
| **Woods-Monteiro-2012**               | 0.21%  |
| Clarifying the landscape of tandem HRCT-domain-mediated protein interactions. Woods et al (2012). *Sci Signal* |        |
| Physical Interactions with 919 interactions from iRefIndex |        |
| **Xie-Cong-2013**                     | 0.19%  |

**Figure 5: Continued.**
| Physical Interactions | 67.64% |
|-----------------------|--------|
| **Xie-Cong-2013**     |        |
| Deniquitinase FAM/USP9X interacts with the E3 ubiquitin ligase SMURF1 protein and protects it from ligase activity-dependent self-degradation. Xie et al (2013). *J Biol Chem* |        |
| Physical Interactions with 168 interactions from iRefIndex |        |
| **Malovannaya-Qin-2010** | 0.19% |
| Streamlined analysis schema for high-throughput identification of endogenous protein complexes. Malovannaya et al (2010). *Proc Natl Acad Sci U S A* |        |
| Physical Interactions with 224 interactions from iRefIndex |        |
| **Chen-Zhang-2013** | 0.18% |
| Quantitative study of the interaction of PKC involved in the EGF-induced tumor cell chemotaxis. Chen et al (2013). *J Proteome Res* |        |
| Physical Interactions with 189 interactions from iRefIndex |        |
| **Koch-Hermeking-2007** | 0.17% |
| Large-scale identification of c-MYC-associated proteins using a combined TAP/MassPIT approach. Koch et al (2007). *Cell Cycle* |        |
| Physical Interactions with 175 interactions from iRefIndex |        |
| **Xiao-Lefkowitz-2007** | 0.17% |
| Functional specialization of beta-arrestin interactions revealed by proteomic analysis. Xiao et al (2007). *Proc Natl Acad Sci U S A* |        |
| Physical Interactions with 402 interactions from iRefIndex |        |
| **Napolitano-Meroni-2011** | 0.17% |
| Functional interactions between ubiquitin E2 enzymes and TRIM proteins. Napolitano et al (2011). *Biochem J* |        |
| Physical Interactions with 81 interactions from BioGRID |        |
| **IREF-PUBMED** | 0.16% |
| Physical Interactions with 571 interactions from iRefIndex |        |
| **Varjosalo-Getaiger-2013** | 0.16% |
| The protein interaction landscape of the human CMGC kinase group. Varjosalo et al (2013). *Cell Rep* |        |
| Physical Interactions with 936 interactions from iRefIndex |        |
| **So-Colwill-2015** | 0.15% |
| Interative analysis of kinase networks in TRAIL-induced apoptosis provides a source of potential targets for combination therapy. So et al (2015). *Sci Signal* |        |
| Physical Interactions with 647 interactions from BioGRID |        |
| **Scholz-Taylor-2016** | 0.14% |
| F11 regulates cellular metabolism through hydroxylation of the Deniquitinase OTUB1. Schotla et al (2016). *PLoS Biol* |        |
| Physical Interactions with 134 interactions from BioGRID |        |
| **Huttlin-Gygj-2015** | 0.14% |
| The BioPlex Network: A Systematic Exploration of the Human Interactome. Huttlin et al (2015). *Cell* |        |
| Physical Interactions with 23,389 interactions from BioGRID |        |
| **Ingham-Pawson-2005** | 0.14% |
| WW domains provide a platform for the assembly of multiprotein networks. Ingham et al (2005). *Mol Cell Biol* |        |
| Physical Interactions with 289 interactions from iRefIndex |        |

**Figure 5:** Continued.
| Physical Interactions | 67.64% |
|-----------------------|--------|
| Tsai-Cristea-2012     | 0.13%  |
| Functional proteomics establishes the interaction of BRDT with chromatin remodeling complexes and expands its role in regulation of RNA polymerase I transcription. Tsai et al (2012). *Mol Cell Proteomics* |
| Physical Interactions with 655 interactions from iRefIndex |
| Yatim-Benkirane-2012  | 0.12%  |
| NOTCH1 nuclear interactome reveals key regulators of its transcriptional activity and oncogenic function. Yatim et al (2012). *Mol Cell* |
| Physical Interactions with 131 interactions from iRefIndex |
| Thompson-Luchansky-2014| 0.12%  |
| Quantitative Lys- Gly-Gly (dGly) proteomics coupled with inducible RNAi reveals ubiquitin-mediated proteolysis of DNA damage-inducible transcript 4 (DDIT4) by the E3 ligase HUWE1. Thompson et al (2014). *J Biol Chem* |
| Physical Interactions with 555 interactions from iRefIndex |
| Floyd-Paglharini-2016 | 0.12%  |
| Mitochondrial protein interaction mapping identifies regulators of respiratory chain function. Floyd et al (2016). *Mol Cell* |
| Physical Interactions with 1,508 interactions from BioGRID |
| Bandyopadhyay-Idedker-2010 | 0.12%  |
| A human MAP kinase interactome. Bandyopadhyay et al (2010). *Nat Methods* |
| Physical Interactions with 611 interactions from iRefIndex |
| van Wijk-Timmers-2009 | 0.11%  |
| A comprehensive framework of E3-RING E3 interactions of the human ubiquitin-proteasome system. van Wijk et al (2009). *Mol Biol Cell* |
| Physical Interactions with 301 interactions from iRefIndex |
| IREF-INTACT            | 0.11%  |
| Physical Interactions with 56,287 interactions from iRefIndex |
| Lee-Songyang-2011      | 0.11%  |
| Genome-wide YFP fluorescence complementation screen identifies new regulators for telomere signaling in human cells. Lee et al (2011). *Mol Cell Proteomics* |
| Physical Interactions with 694 interactions from iRefIndex |
| Grossmann-Stelzl-2015  | 0.11%  |
| Phospho-tyrosine dependent protein-protein interaction network. Grossmann et al (2015). *Mol Syst Biol* |
| Physical Interactions with 622 interactions from BioGRID |
| Liu-Wang-2012          | 0.11%  |
| Proteomic identification of common RCP ubiquitin ligase FbxO8-interacting glycoproteins in three kinds of cells. Liu et al (2012). *J Proteome Res* |
| Physical Interactions with 586 interactions from iRefIndex |
| Povlsen-Choudhary-2012 | 0.10%  |
| Systems-wide analysis of ubiquitylation dynamics reveals a key role for PAR1 ubiquitylation in DNA-damage bypass. Povlsen et al (2012). *Nat Cell Biol* |
| Physical Interactions with 562 interactions from iRefIndex |
| Oštáhl-Ovádi-2011      | 0.10%  |
| Interactions of pathological hallmark proteins: tubulin polymerization promoting protein/τ25, beta-amyloid, and alpha-synuclein. Oštáhl et al (2011). *J Biol Chem* |

(Figure 5: Continued.)
| Physical Interactions | p-value |
|-----------------------|---------|
| Oláh-Ovádi-2011       | 67.64%  |
| Hayes-Urbé-2012       | 0.10%   |
| Joshi-Cristea-2013    | 0.10%   |
| Roy-Parent-2013       | 0.10%   |
| Ewing-Figueys-2007    | 0.10%   |
| Zanon-Pichler-2013    | 0.09%   |
| Brajenovic-Drewes-2004| 0.09%   |
| Greco-Cristea-2011    | 0.09%   |
| Wu-Li-2007            | 0.09%   |
| Behrends-Harper-2010  | 0.09%   |
| Nakayasu-Adkins-2013  | 0.09%   |
| Ouyang-Gill-2009      | 0.08%   |

**Figure 5**: Continued.
| Physical Interactions | 67.64% |
|-----------------------|--------|
| **Ouyang-Gill-2009**  |        |
| Physical Interactions with 105 interactions from BioGRID |
| **Perez-Hernandez-Yáñez-Mé-2013** | 0.08% |
| The intracellular interactions of tetraspanin-enriched interdomains reveals their function as sorting machines toward exosomes. Perez-Hernandez et al (2013). J Biol Chem |
| Physical Interactions with 450 interactions from iReffindex |
| **Reyniers-Taymans-2014** | 0.08% |
| Differential protein-protein interactions of LRRK1 and LRRK2 indicate roles in distinct cellular signaling pathways. Reyniers et al (2014). J Neurochem |
| Physical Interactions with 102 interactions from iReffindex |
| **Bruderer-Hay-2011** | 0.08% |
| Purification and identification of endogenous polySUMO conjugates. Bruderer et al (2011). EMBO Rep |
| Physical Interactions with 106 interactions from iReffindex |
| **Couzens-Gingras-2013** | 0.08% |
| Protein interaction network of the mammalian Hippo pathway reveals mechanisms of kinase-phosphatase interactions. Couzens et al (2013). Sci Signal |
| Physical Interactions with 364 interactions from BioGRID |
| **Bennett-Harper-2010** | 0.08% |
| Dynamics of cullin-RING ubiquitin ligase network revealed by systematic quantitative proteomics. Bennett et al (2010). Cell |
| Physical Interactions with 4,367 interactions from BioGRID |
| **Christianson-Kopito-2011** | 0.07% |
| Defining human ERAD networks through an integrative mapping strategy. Christianson et al (2011). Nat Cell Biol |
| Physical Interactions with 260 interactions from iReffindex |
| **Albers-Koepli-2005** | 0.07% |
| Automated yeast two-hybrid screening for nuclear receptor-interacting proteins. Albers et al (2005). Mol Cell Proteomics |
| Physical Interactions with 238 interactions from iReffindex |
| **Thalapilly-Dusetti-2008** | 0.07% |
| Identification of multi-SUMO domain-containing protein interactome in pancreatic cancer: a yeast two-hybrid approach. Thalapilly et al (2008). Proteomics |
| Physical Interactions with 104 interactions from iReffindex |
| **Gupta-Pelletier-2015** | 0.07% |
| A Dynamic Protein Interaction Landscape of the Human Centrosome-Cilium Interface. Gupta et al (2015). Cell |
| Physical Interactions with 307 interactions from BioGRID |
| **Vinayagam-Wanker-2011** | 0.06% |
| A directed protein interaction network for investigating intracellular signal transduction. Vinayagam et al (2011). Sci Signal |
| Physical Interactions with 2,576 interactions from BioGRID |
| **Li-Dorf-2014** | 0.06% |
| TRIM69 regulates microRNA activity by ubiquitination of TNRC6. Li et al (2014). J Exp Med Acta Sci U S A |
| Physical Interactions with 470 interactions from iReffindex |
| **Katsogiannou-Rocchi-2014** | 0.06% |

(q)

**Figure 5: Continued.**
| Physical Interactions | 67.64% |
|-----------------------|-------|
| **Katsogiannou-Rocchi-2014** |  |
| The functional landscape of Hsp27 reveals new cellular processes such as DNA repair and alternative splicing and proposes novel anticancer targets. Katsogiannou et al (2014). *Mol Cell Proteomics* |  |
| Physical Interactions with 217 interactions from iRefIndex |  |
| **Arbuckle-Grant-2010** | 0.05% |
| The SH3 domain of postsynaptic density 95 mediates inflammatory pain through phosphatidylinositol-3-kinase recruitment. Arbuckle et al (2010). *EMBO Rep* |  |
| Physical Interactions with 268 interactions from iRefIndex |  |
| **Bantscheff-Drewes-2011** | 0.05% |
| Chemospectroscopic profiling of HDAC inhibitors reveals selective targeting of HDAC complexes. Bantscheff et al (2011). *Nat Biotechnol* |  |
| Physical Interactions with 103 interactions from BioGRID |  |
| **Low-Heck-2014** | 0.04% |
| A systems-wide screen identifies substrates of the SCF TrCP ubiquitin ligase. Low et al (2014). *Sci Signal* |  |
| Physical Interactions with 221 interactions from BioGRID |  |
| **Lau-Romai-2012** | 0.04% |
| PKC promotes oncogenic functions of ATF2 in the nucleus while blocking its apoptotic function at mitochondria. Lau et al (2012). *Cell* |  |
| Physical Interactions with 134 interactions from iRefIndex |  |
| **Kiri-Görlich-2015** | 0.04% |
| A deep proteomic perspective on CRMP1-mediated nuclear export and nucleocytoplasmic partitioning. Kiri et al (2015). *Elife* |  |
| Physical Interactions with 1,036 interactions from BioGRID |  |
| **Woodsmith-Sanderson-2012** | 0.04% |
| Systematic analysis of dimeric E3-RING interactions reveals increased combinatorial complexity in human ubiquitination networks. Woodsmith et al (2012). *Mol Cell Proteomics* |  |
| Physical Interactions with 212 interactions from iRefIndex |  |
| **Zhang-Zou-2011** | 0.03% |
| A bead-based approach for large-scale identification of in vitro kinase substrates. Zhang et al (2011). *Proteomics* |  |
| Physical Interactions with 162 interactions from iRefIndex |  |
| **Leng-Wang-2014** | 0.03% |
| A proteomics strategy for the identification of FAT10-modified sites by mass spectrometry. Leng et al (2014). *J Proteome Res* |  |
| Physical Interactions with 144 interactions from iRefIndex |  |
| **Bogachek-Weigel-2014** | 0.02% |
| Sumoylation pathway is required to maintain the basal breast cancer subtype. Bogachek et al (2014). *Cancer Cell* |  |
| Physical Interactions with 134 interactions from iRefIndex |  |
| **Hanson-Clayton-2014** | 0.02% |
| Identifying biological pathways that underlie primordial short stature using network analysis. Hanson et al (2014). *J Mol Endocrinol* |  |
| Physical Interactions with 1,887 interactions from iRefIndex |  |
| **Markson-Sanderson-2009** | 0.02% |
| Analysis of the human E2 ubiquitin-conjugating enzyme protein interaction network. Markson et al (2009). *Genome Res* |  |

(Figure 5: Continued.)
| Physical Interactions | 67.64% |
|-----------------------|--------|
| Markson-Sanderson-2009 | Physical Interactions with 799 interactions from iRefIndex |
| Golebiowski-Hay-2009 | 0.02% |
| System-wide changes to SUMO modifications in response to heat shock. Golebiowski et al (2009). *Sci Signal* |
| Physical Interactions with 351 interactions from iRefIndex |
| Yang-Chen-2010 | 0.02% |
| Proteomic dissection of cell type-specific H3AX-interacting protein complex associated with hepatocellular carcinomas. Yang et al (2010). *J Proteome Res* |
| Physical Interactions with 190 interactions from BioGRID |
| Ravasi-Hayashizaki-2010 | 0.02% |
| An atlas of combinatorial transcriptional regulation in mouse and man. Ravasi et al (2010). *Cell* |
| Physical Interactions with 635 interactions from iRefIndex |
| Arroyo-Aloy-2015 | 0.02% |
| Systematic identification of molecular links between core and candidate genes in breast cancer. Arroyo et al (2015). *J Mol Biol* |
| Physical Interactions with 600 interactions from iRefIndex |
| IREF-HPRD | 0.02% |
| Physical Interactions with 34,206 interactions from iRefIndex |
| Pichlmair-Superti-Furga-2012 | 0.02% |
| Viral immune modulators perturb the human molecular network by common and unique strategies. Pichlmair et al (2012). *Nature* |
| Physical Interactions with 14 interactions from BioGRID |
| Virok-Fülöp-2011 | 0.02% |
| Protein array based interactions analysis of amyloid-β indicates an inhibition of protein translation. Virok et al (2011). *J Proteome Res* |
| Physical Interactions with 299 interactions from BioGRID |
| Vermeulen-Mann-2010 | 0.01% |
| Quantitative interaction proteomics and genome-wide profiling of epigenetic histone marks and their readers. Vermeulen et al (2016). *Cell* |
| Physical Interactions with 131 interactions from iRefIndex |
| Grant-2010 | 0.01% |
| Identification of SUMOylated proteins in neuroblastoma cells after treatment with hydrogen peroxide or ascorbate. Grant (2010). *BMB Rep* |
| Physical Interactions with 114 interactions from iRefIndex |
| Meck-Piwnica-Worms-2004 | 0.01% |
| Comprehensive proteomic analysis of interphase and mitotic 14-3-3-binding proteins. Meck et al (2004). *J Biol Chem* |
| Physical Interactions with 359 interactions from iRefIndex |
| Soler-López-Aloy-2011 | 0.01% |
| Interactome mapping suggests new mechanistic details underlying Alzheimer’s disease. Soler-López et al (2011). *Genome Res* |
| Physical Interactions with 312 interactions from iRefIndex |
| Yu-Vidal-2011 | 0.01% |
| Next-generation sequencing to generate interactome datasets. Yu et al (2011). *Nat Methods* |

(a) Figure 5: Continued.
| Physical Interactions | 67.64% |
|-----------------------|--------|
| Yu-Vidal-2011         |        |
| Physical Interactions with 1,108 interactions from BioGRID |
| Weimann-Stelzl-2013 A | 0.00%  |
| A Y2H-seq approach defines the human protein methyltransferase interactions. Weimann et al (2013). Nat Meth |
| Physical Interactions with 111 interactions from BioGRID |
| Wong-O'Bryan-2012     | 0.00%  |
| Intersectin (ITSN) family of scaffolds function as molecular hubs in protein interaction networks. Wong et al (2012). PloS One |
| Physical Interactions with 111 interactions from IRefIndex |
| Kim-Major-2015        | 0.00%  |
| Substrate trapping proteomics reveals targets of the TrCP2,FBXW11 ubiquitin ligase. Kim et al (2015). Mol Cell Biol |
| Physical Interactions with 111 interactions from IRefIndex |
| Co-expression         | 13.50% |
| Ramaswamy-Golub-2001  | 1.02%  |
| Multichannel cancer diagnosis using tumor gene expression signatures. Ramaswamy et al (2001). Proc Natl Acad Sci U S A |
| Co-expression with 275,113 interactions from supplementary material |
| Wang-Maris-2006       | 0.95%  |
| Integrative genomics identifies distinct molecular classes of neuroblastoma and shows that multiple genes are targeted by regional alterations in DNA copy number. Wang et al (2006). Cancer Res |
| Co-expression with 264,533 interactions from GEO |
| Mallon-McKay-2013     | 0.86%  |
| StemCellDB: the human pluripotent stem cell database at the National Institutes of Health. Mallon et al (2013). Stem Cell Res |
| Co-expression with 565,865 interactions from GEO |
| Bild-Nevins-2006 B    | 0.85%  |
| Oncogenic pathway signatures in human cancers as a guide to targeted therapies. Bild et al (2006). Nature |
| Co-expression with 280,083 interactions from GEO |
| Burington-Shaughnessy-2008 | 0.80% | |
| Tumor cell gene expression changes following short-term in vivo exposure to single agent chemotherapeutics are related to survival in multiple myeloma. Burington et al (2008). Clin Cancer Res |
| Co-expression with 290,538 interactions from GEO |
| Dobkin-Giordano-2005  | 0.78%  |
| Interlaboratory comparability study of cancer gene expression analysis using oligonucleotide microarrays. Dobkin et al (2005). Clin Cancer Res |
| Co-expression with 444,931 interactions from GEO |
| Bahr-Bowler-2013      | 0.70%  |
| Peripheral blood mononuclear cell gene expression in chronic obstructive pulmonary disease. Bahr et al (2013). Am J Respir Cell Mol Biol |
| Co-expression with 374,949 interactions from GEO |
| Alizadeh-Staudt-2000  | 0.69%  |
| Distinct types of diffuse large B-cell lymphomas identified by gene expression profiling. Alizadeh et al (2000). Nature |
| Co-expression with 90,336 interactions from supplementary material |

(Figure 5: Continued.)
| Study                  | Co-expression | Gene Expression Details                                                                 |
|-----------------------|---------------|-----------------------------------------------------------------------------------------|
| Innocenti-Brown-2011  | 0.69%         | Identification, replication, and functional fine-mapping of expression quantitative trait loci in primary human liver tissue. Innocenti et al (2011). PLoS Genet Co-expression with 693,785 interactions from GEO |
| Rieger-Chu-2004       | 0.66%         | Toxicity from radiation therapy associated with abnormal transcriptional responses to DNA damage. Rieger et al (2004). Proc Natl Acad Sci U S A Co-expression with 259,974 interactions from GEO |
| Noble-Diehl-2008      | 0.65%         | Regional variation in gene expression in the healthy colon is dysregulated in ulcerative colitis. Noble et al (2008). Gut Co-expression with 691,539 interactions from GEO |
| Roth-Zlotnik-2006     | 0.61%         | Gene expression analysis reveals molecular relationships among 28 regions of the human CNS. Roth et al (2006). Neuron Co-expression with 699,062 interactions from GEO |
| Boldrick-Rehalten-2002| 0.60%         | Stereotyped and specific gene expression programs in human innate immune responses to bacteria. Boldrick et al (2002). Proc Natl Acad Sci U S A Co-expression with 111,737 interactions from supplementary material |
| Perou-Botstein-2000   | 0.59%         | Molecular portraits of human breast tumours. Perou et al (2000). Nature Co-expression with 185,268 interactions from supplementary material |
| Smirnov-Cheung-2009   | 0.58%         | Genetic analysis of radiation-induced changes in human gene expression. Smirnov et al (2009). Nature Co-expression with 461,800 interactions from GEO |
| Wang-Cheung-2015      | 0.57%         | Genetic variation in insulin-induced kinase signaling. Wang et al (2015). Mol Syst Biol Co-expression with 411,047 interactions from GEO |
| Chen-Brown-2002       | 0.51%         | Gene expression patterns in human liver cancers. Chen et al (2002). Mol Biol Cell Co-expression with 282,241 interactions from supplementary material |
| Perou-Botstein-1999   | 0.50%         | Distinctive gene expression patterns in human mammary epithelial cells and breast cancers. Perou et al (1999). Proc Natl Acad Sci U S A Co-expression with 65,689 interactions from supplementary material |
| Wu-Garvey-2007        | 0.48%         | The effect of insulin on expression of genes and biochemical pathways in human skeletal muscle. Wu et al (2007). Endocrine Co-expression with 267,199 interactions from GEO |
| Rosenwald-Staudt-2001  | 0.40%         | Relation of gene expression phenotype to immunoglobulin mutation genotype in B cell chronic lymphocytic leukemia. Rosenwald et al (2001). J Exp Med Co-expression with 114,494 interactions from supplementary material |

(u)

**Figure 5: Continued.**
| Predicted                                      | Accuracy |
|-----------------------------------------------|----------|
| I2D-vonMering-Bork-2002-High-Yeast2Human       | 6.35%    |
| Comparative assessment of large-scale data sets of protein-protein interactions. von Mering et al (2002). Nature |          |
| Predicted with 1,196 interactions from I2D      |          |
| I2D-vonMering-Bork-2002-Medium-Yeast2Human     | 0.89%    |
| Comparative assessment of large-scale data sets of protein-protein interactions. von Mering et al (2002). Nature |          |
| Predicted with 3,009 interactions from I2D      |          |
| I2D-BioGRID-Yeast2Human                       | 0.59%    |
| BioGRID: a general repository for interaction datasets. Stark et al (2006). Nucleic Acids Res |          |
| Predicted with 13,434 interactions from I2D     |          |
| I2D-Chen-Pawson-2009-PiwiScreen-Mouse2Human    | 0.50%    |
| Mouse Piwi interactome identifies binding mechanism of Tdrk Tudor domain to arginine methylated Piwi. Chen et al (2009). Proc Natl Acad Sci U S A | | Predicted with 31 interactions from I2D | |
| I2D-INNATEDB-Mouse2Human                      | 0.46%    |
| InnateDB: facilitating systems-level analyses of the mammalian innate immune response. Lynn et al (2008). Mol Syst Biol | |
| Predicted with 1,451 interactions from I2D      |          |
| I2D-Tarassov-PCA-Yeast2Human                   | 0.43%    |
| An in vivo map of the yeast protein interactome. Tarassov et al (2008). Science | |
| Predicted with 440 interactions from I2D        |          |
| I2D-vonMering-Bork-2002-Low-Yeast2Human        | 0.38%    |
| Comparative assessment of large-scale data sets of protein-protein interactions. von Mering et al (2002). Nature | |
| Predicted with 16,063 interactions from I2D     |          |
| Wu-Stein-2010                                  | 0.34%    |
| A human functional protein interaction network and its application to cancer data analysis. Wu et al (2010). Genome Biol | |
| Predicted with 87,829 interactions from supernmental material | |
| I2D-IntAct-Mouse2Human                        | 0.31%    |
| The IntAct molecular interaction database in 2010. Aranda et al (2010). Nucleic Acids Res | |
| Predicted with 3,427 interactions from I2D      |          |
| Stuart-Kim-2003                                | 0.27%    |
| A gene-coexpression network for global discovery of conserved genetic modules. Stuart et al (2003). Science | |
| Predicted with 24,872 interactions from supernmental material | |
| I2D-Yu-Vidal-2008-GoldStd-Yeast2Human          | 0.22%    |
| High-quality binary protein interaction map of the yeast interactome network. Yu et al (2008). Science | |
| Predicted with 386 interactions from I2D        |          |
| I2D-Krogan-Greenblatt-2006-Core-Yeast2Human    | 0.19%    |
| Global landscape of protein complexes in the yeast Saccharomyces cerevisiae. Krogan et al (2006). Nature | |
| Predicted with 1,823 interactions from I2D      |          |
| I2D-BIND-Rat2Human                             | 0.19%    |
| BIND—a data specification for storing and describing biomolecular interactions, molecular complexes and pathways. Bader et al | |

**Figure 5: Continued.**
| Predicted                                           |       |
|----------------------------------------------------|-------|
| I2D-BIND-Rat2Human                                 | 6.35% |
| (2000). Bioinformatics                             |       |
| Predicted with 548 interactions from I2D           |       |
| I2D-BioGRID-Worm2Human                             | 0.18% |
| BioGRID: a general repository for interaction datasets. Stark et al (2006). Nucleic Acids Res |       |
| Predicted with 932 interactions from I2D           |       |
| I2D-MINT-Rat2Human                                 | 0.17% |
| MINT: a Molecular INTeraction database. Zanconi et al (2002). FEBS Lett |       |
| Predicted with 572 interactions from I2D           |       |
| I2D-MINT-Mouse2Human                               | 0.15% |
| MINT: a Molecular INTeraction database. Zanconi et al (2002). FEBS Lett |       |
| Predicted with 971 interactions from I2D           |       |
| I2D-BIND-Mouse2Human                               | 0.14% |
| BIND—a data specification for storing and describing biomolecular interactions, molecular complexes and pathways. Bader et al (2000). Bioinformatics |       |
| Predicted with 1,186 interactions from I2D         |       |
| I2D-BioGRID-Mouse2Human                            | 0.12% |
| BioGRID: a general repository for interaction datasets. Stark et al (2006). Nucleic Acids Res |       |
| Predicted with 286 interactions from I2D           |       |
| I2D-BIND-Yeast2Human                               | 0.11% |
| BIND—a data specification for storing and describing biomolecular interactions, molecular complexes and pathways. Bader et al (2000). Bioinformatics |       |
| Predicted with 1,541 interactions from I2D         |       |
| I2D-Krogan-Greenblatt-2006-NonCore-Yeast2Human     | 0.09% |
| Global landscape of protein complexes in the yeast Saccharomyces cerevisiae. Krogan et al (2006). Nature |       |
| Predicted with 1,786 interactions from I2D         |       |
| I2D-IntAct-Fly2Human                               | 0.09% |
| The IntAct molecular interaction database in 2010. Aranda et al (2010). Nucleic Acids Res |       |
| Predicted with 3,912 interactions from I2D         |       |
| I2D-Formstecher-Daviet-2005-Embryo-Fly2Human       | 0.07% |
| Protein interaction mapping: a Drosophila case study. Formstecher et al (2005). Genome Res |       |
| Predicted with 491 interactions from I2D           |       |
| I2D-MGI-Mouse2Human                                | 0.07% |
| Ontological visualization of protein-protein interactions. Drakakis et al (2005). BMC Bioinformatics |       |
| Predicted with 726 interactions from I2D           |       |
| I2D-IntAct-Rat2Human                               | 0.01% |
| The IntAct molecular interaction database in 2010. Aranda et al (2010). Nucleic Acids Res |       |
| Predicted with 1,052 interactions from I2D         |       |
| I2D-MINT-Worm2Human                                | 0.00% |

(w)

Figure 5: Continued.
Predicted 6.35%

1D-MINT-Worm2Human
MINT: a Molecular INTeraction database. Zanconi et al (2002). *FEBS Lett*
Predicted with 1,178 interactions from 1D

Co-localization 6.17%

Zhang-Shang-2006 2.58%
The catalytic subunit of the proteasome is engaged in the entire process of estrogen receptor-regulated transcription. Zhang et al (2006). *EMBO J*
Co-localization with 53 interactions from BioGRID

Schadt-Shoemaker-2004 1.65%
A comprehensive transcript index of the human genome generated using microarrays and computational approaches. Schadt et al (2004). *Genome Biol*
Co-localization with 60,126 interactions from GEO

Johnson-Shoemaker-2003 1.11%
Genome-wide survey of human alternative pre-mRNA splicing with exon junction microarrays. Johnson et al (2003). *Science*
Co-localization with 436,332 interactions from GEO

Chen-Huang-2014 0.82%
Using an in situ proximity ligation assay to systematically profile endogenous protein-protein interactions in a pathway network. Chen et al (2014). *J Proteome Res*
Co-localization with 559 interactions from BioGRID

Pathway 4.35%

Wu-Stein-2010 1.54%
A human functional protein interaction network and its application to cancer data analysis. Wu et al (2010). *Genome Biol*
Pathway with 78,010 interactions from supplementary material

REACTOME 1.33%
Pathway with 24,913 interactions from Pathway Commons

NCI_NATURE 0.58%
Pathway with 10,122 interactions from Pathway Commons

CELL_MAP 0.46%
Pathway with 598 interactions from Pathway Commons

IMID 0.41%
Pathway with 1,073 interactions from Pathway Commons

HUMANCYC 0.03%
Pathway with 680 interactions from Pathway Commons

Genetic Interactions 1.40%

Vizeacoumar-Moffat-2013 0.41%
A negative genetic interaction map in isogenic cancer cell lines reveals cancer cell vulnerabilities. Vizeacoumar et al (2013). *Mol Syst Biol*
Genetic Interactions with 201 interactions from BioGRID

BIOGGRID-SMALL-SCALE-STUDIES 0.38%

(x)

**Figure 5: Continued.**
2.6. cBioPortal Analysis. cBioPortal is an open-access network analysis database that can be used for comprehensive exploration of genomics data of a variety of cancers [9]. The cBioPortal website currently stores DNA copy number data (assuming discrete values for each gene, such as “deep deletion” or “amplification” and log2 levels), mRNA and microRNA expression data, nonsynonymous mutations, protein levels and phosphoprotein levels (RPPA) data, DNA methylation data, and limited clinical data. cBioPortal greatly reduces the access barriers between complex genomic data and cancer researchers. It facilitates quick, intuitive, and high-quality access to molecular profiles and clinical prognostic correlations of large-scale cancer genome projects. In this study, we used cBioPortal to analyze the frequency and type of PGAM1 gene alternation in 817 breast cancer specimens in TCGA data.

2.7. Cancer Cell Line Encyclopedia (CCLE) Analysis. Cancer Cell Line Encyclopedia database (https://portals.broadinstitute.org/ccle) is an online database jointly developed by Broad Institute and Novartis Research Foundation. At present, the database has collected and visualized the genetic information of more than 1100 cell lines, including copy number and mRNA expression (RNAseq).

3. Results

3.1. PGAM1 Expression in Breast Cancer. We initially evaluated PGAM1 mRNA levels in multiple breast cancer studies from TCGA and Gene Expression Omnibus (GEO). The analysis flowchart was shown in Supplementary Figure 1. The data in TCGA database showed that the mRNA expression of PGMA1 in breast cancer tissues was significantly higher than that in normal tissues (Figure 1(a), P < 0.001). The Curtis Breast (Figure 1(b), P < 0.001) and Ma Breast 4 (Figure 1(c), P = 0.016) databases further validated that the expression level of PGAM1 in breast cancer tissue was significantly higher than that in breast tissue. Therefore, PGAM1 expression can be used as a potential diagnostic indicator for breast cancer. In addition, we also observed that the expression of PGAM1 in the basal-like subtype was significantly higher than that of the luminal-like subtype (Figure 1(d), P = 0.045). In another database, we also found that the expression level of PGAM1 in breast cancer was significantly higher than that in normal tissues (Figure 1(e), P < 0.001). In addition, the overall survival of patients with high PGAM1 expression was significantly worse than that of patients with low expression (Figure 1(f), P = 0.0077). Finally, we evaluated...
the expression level of PGAM1 in multiple breast cancer cell lines through the CCLE database (Supplementary Figure 2).

We then analyzed the relationship between the expression level of PGAM1 and the clinicopathological characteristics of patients, and found that PGAM1 was significantly related to molecular subtypes and TP53 mutation status (Figure 2). The expression of PGAM1 in HER2-positive and triple-negative tumors was significantly higher than that in luminal breast cancer (Figure 2(b), *P* < 0.001). The expression level of PGAM1 in TP53-mutant breast tumors was significantly higher than that in non-TP53-mutant tumors (Figure 2(d), *P* < 0.001). However, there was no significant difference in the expression level of PGAM1 between tumor tissues of different N and TNM stages.

3.2. Analysis of GO and KEGG Pathways of Breast Cancer Coexpressed Genes Related to PGAM1. The functional module of LinkedOmics was used to analyze the mRNA sequencing data of 526 breast cancer patients in TCGA. As shown in the volcano map (Figure 3(a)), 1925 genes (dark red dots) showed a significant positive correlation with PGAM1, while 1427 genes (dark green dots) showed a significant negative correlation (false discovery rate (FDR) < 0.01). The heat maps, respectively, show 50 important gene sets that are positively and negatively correlated with PGAM1 (Figures 3(b) and 3(c)). Among them, ACTR1A, PGD, RRM2, NUP93, and SLC25A5 are the top 5 genes with the most significant positive correlation, while RBM16, ZMAT1, SF1, FLJ21062, and DMTF1 are the top 5 genes with the most significant negative correlation. Similarly, we analyzed the mRNA levels of ACTR1A, PGD, RRM2, NUP93, and SLC25A5 through the CCLE database (Supplementary Figure 3). In addition, the mRNA expression levels of RBM16, ZMAT1, SF1, and DMTF1 were also analyzed by the CCLE database (Supplementary Figure 4). These results indicate that PGAM1 may have an extensive gene regulatory network.

Biological process analysis showed that the biological functions of differentially expressed genes positively related to PGAM1 focused on antigen processing and presentation, chromosome separation, granulocyte activation, mitotic cell cycle phase transition, ribonucleotide metabolism, and mitochondrial transport (Figure 4(a)). Cell component analysis showed that differentially expressed genes related to PGAM1 played a structural role in the inner mitochondrial membrane, the intrinsic components of the organelle membrane, and condensed chromosomes (Figure 4(b)). Molecular function analysis showed that the biological functions of differentially expressed genes positively related to PGAM1 were concentrated in electron transfer activity, nucleotidyl transferase activity, nuclease activity, etc. (Figure 4(c)). KEGG analysis showed that differentially expressed genes positively related to PGAM1 were significantly enriched in phagocytosis, glutathione metabolism, Toll-like receptor signal transduction pathways, lysosomes, and human cytomegalovirus infection pathways (Figure 4(d)).

3.3. PGAM1 Networks of Kinase, miRNA, or Transcription Factor Targets in Breast Cancer. In order to further explore the regulatory network related to PGAM1 in breast cancer, we found a number of kinases, miRNAs, and transcription factors statistically related to PGAM1 through GSEA analysis. The five most significant kinases related to PGAM1 are cyclin-dependent kinase 1 (CDK1), polo-like kinase 1 (PLK1), checkpoint kinase 1 (CHEK1), aurora kinase B (AURKB), and B-Raf protooncogene, serine/threonine kinase (BRAF). The most significant miRNAs associated with PGAM1 are miR-382, miR-488, miR-453, miR-30A-3P, miR-30E-3P, miR-525, and miR-524. The most significant transcription factors associated with PGAM1 are FAC1, RSRFC4, and MEF2 (Table 1). The protein interaction network constructed by GeneMANIA revealed that PGAM1 mainly interacts with glycolysis-related enzymes, such as PGK1, LDHA, and PKM (Figure 5).

3.4. Frequency and Type of PGAM1 Gene Alteration in Breast Cancer. Then, based on the sequencing data from breast cancer patients in the TCGA database, we used cBioPortal to determine the type and frequency of PGAM1 changes in breast cancer. Of the 817 (4.2%) breast cancer patients, 34 had PGAM1 gene alternation (Figure 6). These changes include mRNA upregulation in 22 cases (16%), downregulation in 11 cases (10%), and 1 case (5%) of deep deletion. Therefore, upregulation of mRNA is the most common type of PGAM1 alteration in breast cancer.

4. Discussion

Breast cancer is the most common malignant tumor in women worldwide. Although the prognosis of breast cancer has been significantly improved in recent years, most breast cancers still inevitably recur and metastasize even after comprehensive treatment. Currently, there is an urgent clinical need to identify more molecular markers for accurate diagnosis and prognosis prediction of breast cancer. As we all know, even when there is sufficient oxygen, cancer cells still metabolize glucose for energy through glycolysis. Tumor-targeted glycolysis is considered a potential therapeutic strategy for tumor targeting. PGAM1 is a key enzyme for glycolysis, and a large number of studies suggest that PGAM1 is closely related to the malignant progression of tumors [10–13]. However, whether PGAM1 has a regulatory role in breast cancer remains unclear. In order to gain a deeper understanding of the diagnostic and prognostic value of PGAM1 in breast cancer, we conducted a bioinformatics analysis of the public sequencing database on the effect of PGAM1 on the clinicopathological characteristics and prognosis of breast cancer, and discussed the correlation of PGAM1 in breast cancer. The regulatory network laid the foundation for further elucidating the mechanism of PGAM1 in breast cancer.

We first investigated the correlation between the expression level of PGAM1 and the clinicopathological characteristics of breast cancer patients. Analysis of transcript sequencing data from 137 clinical samples in the TCGA database found that the level of PGAM1 mRNA in breast cancer was significantly higher than that in normal breast tissue. The Curtis Breast and Ma Breast 4 databases further validated that the expression level of PGAM1 in breast cancer tissue
was significantly higher than that in breast tissue. These results strongly suggest that PGAM1 can be used as a diagnostic biomarker for breast cancer. Subsequently, we also found that PGAM1 was closely related to aggressive clinicopathological features. For example, we found that the expression level of PGAM1 in HER2-positive and triple-negative subtypes was higher than that of luminal subtypes. Interestingly, we also found that the expression of PGAM1 in TP53-mutant tumors was significantly higher than that in non-TP53-mutant breast cancers. As a tumor suppressor gene, TP53 is the most common mutant gene in breast cancer [14]. TP53 mutation can lead to cell cycle arrest, apoptosis, metabolism, DNA repair, and cell senescence of breast cancer cells [15, 16]. In addition, TP53 mutations are also associated with drug resistance and poor prognosis of breast cancer [17, 18]. Therefore, PGAM1 is likely to exert its cancer-promoting effect through TP53-dependent signaling pathways, but the specific mechanism is worthy of further study. Finally, survival analysis further confirmed that the expression of PGAM1 was significantly related to the overall survival (OS) of patients. This shows that PGAM1 can not only be used as a biomarker for breast cancer diagnosis but also as a biomarker for predicting the prognosis of breast cancer.

We used GSEA to perform enrichment analysis on PGAM1 and obtained multiple kinases, miRNAs, and transcription factors that are significantly associated with it. Our results indicate that the functional network of PGAM1 is mainly involved in metabolism, biological regulation, response to stimuli, and cell proliferation. We found that PGAM1 in breast cancer was related to kinase networks including CDK1, PLK1, and CHEK1. These kinases could participate in the regulation of mitosis, DNA damage, and intracellular signal transduction. CDK1 is an essential cyclin-dependent kinase, which regulates the cell cycle by interacting with specific cell cycle regulator cyclins. CDK1 expression is elevated in a variety of cancers, often leading to unrestricted proliferation of malignant tumor cells [19–21]. PLK1 is a serine threonine kinase that plays a vital role in centrosome maturation, mitotic chromosome separation, and mitosis, and is closely related to the occurrence and development of malignant tumors [22–24]. As a kinase, CHEK1 also plays an important role in tumor progression by regulating the cell cycle [25, 26].

Abnormalities in these pathways are closely related to tumor progression. In addition, we also identified several miRNAs related to PGAM1. These small RNAs could participate in the posttranscriptional regulation of gene expression and then affect tumor progression. A large number of miRNAs have been reported to be related to tumor proliferation, apoptosis, cell cycle, invasion, metastasis, drug resistance, and angiogenesis. In fact, miR-382, miR-488, and miR-453 can be used as diagnostic and prognostic markers for malignant tumors.

Considering the significant correlation between PGAM1 and the prognosis of patients, PGAM1 may play an important regulatory role in the malignant progression of breast cancer. We then explored the statistically relevant genes in the expression of PGAM1. Correlation analysis suggests that PGAM1 is positively correlated with the expression of some oncogenes and is a key regulator that affects cancer proliferation, invasion, metastasis, and patient survival. These results further suggest that PGAM1 can promote the malignant progression of breast cancer by interacting with other oncogenes.

Our research inevitably has some limitations. First of all, our results rely on online databases, but the database provides less information about patient treatment options, so we cannot explore whether treatment has an impact on gene expression. Secondly, although our results indicate that PGAM1 may be a potential diagnostic and prognostic marker for breast cancer. We have not conducted in vivo and in vitro studies to verify whether PGAM1 is a true oncogene. The results of this study still need to be verified by a large number of experiments. Thirdly, the specific role and molecular mechanism of PGAM1 in breast cancer have not been thoroughly explored in this study, and more molecular biology studies are needed to further explore the molecular details of PGAM1. Fourthly, GeneMANIA does not provide quantitative values between related proteins. Finally, the impact of PGAM1 on survival was mainly based on univariate analysis, without adjusting for confounding factors such as age.

In conclusion, this study provides preliminary evidence for PGAM1 as a potential marker for the prognosis of breast cancer. At the same time, our results indicate that PGAM1 is significantly associated with several tumor-associated kinases (such as CDK1 and PLK1), miRNA (such as miRNA-382), and transcription factors (such as RSRFC4) in breast cancer tissues. However, the analysis results of this study based on data mining still need to be verified by more studies, including functional tests and molecular mechanisms, which will help to further clarify the regulatory role of PGAM1 in breast cancer.

Data Availability
All data generated or analyzed in this study are included in this article.

Conflicts of Interest
All authors of this study stated that they had no competing interests.

Authors’ Contributions
YW, WL, XX, and XH are responsible for the research design. YW, WL, and XX are responsible for collecting, analyzing, and interpreting the data. YW, WL, and XX are the main contributors to the writing of manuscripts. The final draft was read and approved by all authors. Yongxuan Wang and Xifeng Xiong contributed equally to this work.

Acknowledgments
This study was supported by the National Natural Science Foundation of China (81902802, XX), the Medical Science...
and Technology Research Foundation of Guangdong (A2018063, XX), the research grants of Traditional Chinese Medicine Bureau of Guangdong Province (20191260, XX), the Medical and Health Science and Technology Project of Guangzhou (20191A011016, WL), and the research grants of Guangzhou Municipal Health and Family Planning Commission (20181A010017, XX and 20201A011020, XX).

Supplementary Materials

Supplementary 1. Figure 1. the analysis flowchart.

Supplementary 2. Figure 2. the mRNA expression of PGAM1 in multiple breast cancer cell lines according to Cancer Cell Line Encyclopedia (CCLE) analysis.

Supplementary 3. Figure 3. the mRNA expression of ACTR1A, PGD, RRM2, NUP93, and SLC25A5 in multiple breast cancer cell lines according to Cancer Cell Line Encyclopedia (CCLE) analysis.

Supplementary 4. Figure 4. the mRNA expression of RBM16, ZMAT1, SF1, and DMTF1 in multiple breast cancer cell lines according to Cancer Cell Line Encyclopedia (CCLE) analysis.

References

[1] R. L. Siegel, K. D. Miller, and A. Jemal, “Cancer statistics, 2019,” CA: A Cancer Journal for Clinicians, vol. 69, no. 1, pp. 7–34, 2018.

[2] X. Liu, X. Tan, P. Liu, Y. Wu, S. Qian, and X. Zhang, “Phosphoglycerate mutase 1 (PGAM1) promotes pancreatic ductal adenocarcinoma (PDAC) metastasis by acting as a novel downstream target of the PI3K/Akt/mTOR pathway,” Oncology Research, vol. 26, no. 7, pp. 1123–1131, 2018.

[3] X. Liu, Y. Weng, P. Liu et al., “Identification of PGAM1 as a putative therapeutic target for pancreatic ductal adenocarcinoma metastasis using quantitative proteomics,” Oncotargets and Therapy, vol. Volume 11, pp. 3345–3357, 2018.

[4] D. Zhang, H. Wu, X. Zhang et al., “Phosphoglycerate mutase 1 predicts the poor prognosis of oral squamous cell carcinoma and is associated with cell migration,” Journal of Cancer, vol. 8, no. 11, pp. 1943–1951, 2017.

[5] F. Ren, H. Wu, Y. Lei et al., “Quantitative proteomics identification of phosphoglycerate mutase 1 as a novel therapeutic target in hepatocellular carcinoma,” Molecular Cancer, vol. 9, no. 1, p. 81, 2010.

[6] D. R. Rhodes, S. Kalyana-Sundaram, V. Mahavisno et al., “Oncomine 3.0: genes, pathways, and networks in a collection of 18,000 cancer gene expression profiles,” Neoplasia, vol. 9, no. 2, pp. 166–180, 2007.

[7] D. S. Chandrashekhar, B. Bashel, S. Balasubramanya et al., “UALCAN: a portal for facilitating tumor subgroup gene expression and survival analyses,” Neoplasia, vol. 19, no. 8, pp. 649–658, 2017.

[8] S. V. Vasaikar, P. Straub, J. Wang, and B. Zhang, “LinkedO-mics: analyzing multi-omics data within and across 32 cancer types,” Nucleic Acids Research, vol. 46, no. D1, pp. D956–D963, 2018.

[9] J. Gao, B. A. Aksoy, U. Dogrusoz et al., “Integrative analysis of complex cancer genomics and clinical profiles using the cBioPortal,” Science Signaling, vol. 6, no. 269, pp. p11, 2013.

[10] F. Li, H. Yang, T. Kong et al., “PGAM1, regulated by miR-3614-5p, functions as an oncogene by activating transforming growth factor-β (TGF-β) signaling in the progression of non-small cell lung carcinoma,” Cell Death & Disease, vol. 11, no. 8, p. 710, 2020.

[11] S. Ohba, T. A. Johannessen, K. Chatla, X. Yang, R. O. Pieper, and J. Mukherjee, “Phosphoglycerate mutase 1 activates DNA damage repair via regulation of WIP1 activity,” Cell Reports, vol. 31, no. 2, p. 107518, 2020.

[12] K. Huang, Q. Liang, Y. Zhou et al., “A novel allosteric inhibitor of phosphoglycerate mutase 1 suppresses growth and metastasis of non-small-cell lung cancer,” Cell Metabolism, vol. 30, no. 6, pp. 1107–1119.e8, 2019.

[13] Q. Sun, S. Li, Y. Wang et al., “Phosphoglyceric acid mutase-1 contributes to oncogenic mTOR-mediated tumor growth and confers non-small cell lung cancer patients with poor prognosis,” Cell Death and Differentiation, vol. 25, no. 6, pp. 1160–1173, 2018.

[14] Cancer Genome Atlas Network, “Comprehensive molecular portraits of human breast tumours,” Nature, vol. 490, no. 7418, pp. 61–70, 2012.

[15] A. Shahbandi, H. D. Nguyen, and J. G. Jackson, “TP53 mutations and outcomes in breast cancer: reading beyond the headlines,” Trends Cancer, vol. 6, no. 2, pp. 98–110, 2020.

[16] E. R. Kastenhuber and S. W. Lowe, “Putting p53 in context,” Cell, vol. 170, no. 6, pp. 1062–1078, 2017.

[17] A. di Gennaro, V. Damiano, G. Brisotto et al., “A p53/miR-30a/ZEB2 axis controls triple negative breast cancer aggressiveness,” Cell Death and Differentiation, vol. 25, no. 12, pp. 2165–2180, 2018.

[18] S. Lin, L. Yu, X. Song et al., “Intrinsic adriamycin resistance in p53-mutated breast cancer is related to the miR-30c/FANC-F/REV1-mediated DNA damage response,” Cell Death & Disease, vol. 10, no. 9, p. 666, 2019.

[19] P. Zhang, H. Kawakami, W. Liu et al., “Targeting CDK1 and MEK/ERK overcomes apoptotic resistance in BRAF-mutant human colorectal cancer,” Molecular Cancer Research, vol. 16, no. 3, pp. 378–389, 2018.

[20] J. Y. Qian, J. Gao, X. Sun et al., “KIAA1429 acts as an oncogenic factor in breast cancer by regulating CDK1 in an N6-methyladenosine-independent manner,” Oncogene, vol. 38, no. 33, pp. 6123–6141, 2019.

[21] R. Zhang, H. Shi, F. Ren et al., “The aberrant upstream pathway regulations of CDK1 protein were implicated in the proliferation and apoptosis of ovarian cancer cells,” Journal of Ovarian Research, vol. 10, no. 1, p. 60, 2017.

[22] E. Montaudon, J. Nikitorowicz-Buniak, L. Sourd et al., “PLK1 inhibition exhibits strong anti-tumoral activity in CCND1-driven breast cancer metastases with acquired palbociclib resistance,” Nature Communications, vol. 11, no. 1, p. 4053, 2020.

[23] A. Parrilla, M. Barber, B. Majem et al., “Aurora borealis (bora), which promotes PLK1 activation by Aurora A, has an oncogenic role in ovarian cancer,” Cancers, vol. 12, no. 4, p. 886, 2020.

[24] R. Affatato, L. Carrassa, R. Chilì, M. Lupi, V. Restelli, and G. Damia, “Identification of PLK1 as a new therapeutic target in mucinous ovarian carcinoma,” Cancers, vol. 12, no. 3, p. 672, 2020.

[25] X. Yang, Y. Pan, Z. Qiu et al., “RNF126 as a biomarker of a poor prognosis in invasive breast cancer and CHEK1 inhibitor
efficacy in breast cancer cells,” Clinical Cancer Research, vol. 24, no. 7, pp. 1629–1643, 2018.

[26] T. M. Abdel-Fatah, F. K. Middleton, A. Arora et al., “Untangling the ATR-CHEK1 network for prognostication, prediction and therapeutic target validation in breast cancer,” Molecular Oncology, vol. 9, no. 3, pp. 569–585, 2015.