RESEARCH PAPER

OPEN ACCESS Check for updates

Identification of significant genes and therapeutic agents for breast cancer by integrated genomics

Xiao Sun, Zhenzhen Luo, Liuyun Gong D, Xinyue Tan, Jie Chen, Xin Liang, and Mengjiao Cai

Department of Oncology, The First Affiliated Hospital, Xi'an Jiaotong University, Xi'an, Shanxi P.R. China

ABSTRACT

Breast cancer is the most commonly diagnosed malignancy in women; thus, more cancer prevention research is urgently needed. The aim of this study was to predict potential therapeutic agents for breast cancer and determine their molecular mechanisms using integrated bioinformatics. Summary data from a large genome-wide association study of breast cancer was derived from the UK Biobank. The gene expression profile of breast cancer was from the Oncomine database. We performed a network-wide association study and gene set enrichment analysis to identify the significant genes in breast cancer. Then, we performed Gene Ontology analysis using the STRING database and conducted Kyoto Encyclopedia of Genes and Genomes pathway analysis using Cytoscape software. We verified our results using the Gene Expression Profile Interactive Analysis, PROgeneV2, and Human Protein Atlas databases. Connectivity map analysis was used to identify small-molecule compounds that are potential therapeutic agents for breast cancer. We identified 10 significant genes in breast cancer based on the gene expression profile and genome-wide association study. A total of 65 small-molecule compounds were found to be potential therapeutic agents for breast cancer.



1. Introduction

Breast cancer is a frequently diagnosed cancer in women with a family history [1]. Breast cancer is a heterogeneous disease with different molecular subtypes and biological behaviors. Gene microarray technology and immunohistochemical techniques have classified breast cancers into different types [2]. The estrogen receptor (ER) is the most important prognostic and predictive immunohistochemical marker in breast cancer. ER-negative tumors tend to be of higher histological grade, are more sensitive to chemotherapy, and are more likely to metastasize to visceral organs [3,4]. Breast cancer does not

ARTICLE HISTORY

Received 24 March 2021 Revised 12 May 2021 Accepted 13 May 2021



CONTACT Mengjiao Cai 🛛 caimengjiao@xjtu.edu.cn 🖃 Department of Oncology, The First Affiliated Hospital, Xi'an Jiaotong University, Yan Ta Road 227, Xi'an, Shaanxi P.R. China 710061

Supplemental data for this article can be accessed here.

^{© 2021} The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group.

This is an Open Access article distributed under the terms of the Creative Commons Attribution License (http://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

have a poor prognosis, and there is no lack of therapeutic targets. ER positive tumors represent about 70% of all breast cancers and there are a lot of therapeutic targets, as well as for HER2 positive breast cancer (about 20% of all BC). The only subtype lacking for target therapies is the triple negative subtype [5,6]. There is an urgent need to find available drugs and clarify their molecular mechanisms in breast cancer treatment.

Most previous studies have focused on identifying novel prognostic markers and drug targets for breast cancer [7-9]. Sulaiman et al. [10] reported that a synthetic azaspirane targets the Janus kinase/signal transducer and activation of transcription 3 pathway in breast cancer. Huang et al. [11] demonstrated that the Gah-PLC δ 1 signaling axis drives metastatic progression in breast cancer. However, due to toxicity, cost, the chemical effects of novel prognostic markers and drug targets for breast cancer that need further research [12], not all previous findings contribute to breast cancer treatment; breast cancer still lacks therapeutic targets and with poorer prognosis. And there is still an urgent need to identify additional therapeutic and prognostic targets in breast cancer [13].

Genome-wide association studies (GWAS) are widely used to characterize the genetic mechanisms that underlie complex diseases. Integrative analyses of GWAS data are rapidly becoming a standard approach to explore the genetic basis of disease susceptibility [14]. Network-wide association studies (NetWAS) can identify relevant disease-gene associations by integrating tissuespecific networks and GWAS results [15,16]. Prior studies have shown that the networkassociated analysis of GWAS data is highly efficient when used to identify novel causal genes of complex diseases [17,18].

In this study, to better understand the molecular mechanisms and find therapeutic agents for breast cancer, we identified novel candidate therapeutic agents for breast cancer treatment by integrating genomic data with drug database analysis. In total, 65 small-molecule compounds were identified, including trichostatin A, LY-294,002, econazole, prestwick-1082, and vorinostat. Our study demonstrates the usefulness of this approach for evaluating the relationship among genes, diseases, and drugs. These findings will pave the way for the discovery of potential therapeutic targets for breast cancer.

2. Methods

2.1 Summary of GWAS datasets in breast cancer

The UK Biobank is a large, population-based prospective UK study, which was established to identify genetic and nongenetic determinants of various diseases. It comprises approximately 500,000 individuals with extensively detailed phenotypes. Their genotypes were determined using an array that included 847,441 genetic polymorphisms, enabling the identification of novel genetic variants in a uniformly genotyped and phenotyped cohort of unprecedented size [19]. Using data from the UK Biobank, samples from the participants were genotyped on the UK Biobank Axiom array and UK BiLEVE custom array. Genotype imputation was conducted with IMPUTE software against the UK10K haplotype panel and the 1000 Genomes Project phase 3 panel. GWAS analysis was performed by SNPTEST using a logistic regression model. genome-wide Α geneassociation study was performed using the MAGMA gene analysis tool, and multiple genes and genetic variants were identified. The Icelandic GWAS dataset from the deCODE Genetics genealogical database was based on whole-genome sequencing using Illumina technology. Finally, meta-analysis of small nucleotide polymorphisms (SNPs) in the UK Biobank and deCODE sample was performed using the METAL analysis tool [20].

The atlas of genetic associations in the UK Biobank (GeneATLAS, http://geneatlas.roslin.ed. ac.uk) helps researchers effectively analyze UK Biobank results without high computational costs. It also allows users to query genome-wide association results for 9,113,133 genetic variants and download GWAS summary statistics for more than 30 million imputed genetic variants (>23 billion phenotype-genotype pairs) [21]. We downloaded large-scale GWAS breast cancer summary data from the atlas of genetic associations. Detailed descriptions of sample characteristics, experimental designs, statistical analyses, and quality control can be found in previous studies.

2.2 Gene expression datasets

Oncomine (https://www.oncomine.org) is a cancer microarray database and web-based data mining platform for facilitating discovery. In this study, differentially expressed genes (DEGs) in breast cancer were identified by comparing cancer samples to respective normal samples using the Oncomine database. The heatmap of significant DEGs in breast cancer was driven from the Oncomine.

2.3 Identification of significant genes in breast cancer

NetWAS (https://hb.flatironinstitute.org/netwas/) integrates tissue-specific networks and significant GWAS association results, and identifies relevant disease-gene associations based on genomics. Briefly, SNP-level association statistics were converted into gene-level statistics (gene-based *P* values), which then were integrated with tissuespecific networks to predict the causal genes [18]. Greene *et al.* [13] demonstrated that NetWAS is more accurate than GWAS alone. In this study, we identified the most relevant genes in breast cancer using NetWAS.

2.4 Kyoto Encyclopedia of Genes and Genomes pathway and Gene Ontology analyses

Cytoscape is one of the most successful network biology analysis and visualization tools. It exposes more than 270 core functions and 34 applications as REST-callable functions with standardized JSON interfaces supported by Swagger documentation [22]. CluePedia, a plug-in in Cytoscape, can search for certain Kyoto Encyclopedia of Genes and Genomes (KEGG) signaling pathways of certain genes by calculating linear and nonlinear statistical dependencies from experimental data [23]. KEGG signaling pathways were identified by CluePedia. Search Tool for the Retrieval of Interacting Genes (STRING) (https://string-db. org/cgi/input.pl) is an online tool that for Gene ontology (GO) analysis in gene sets [24,25]. GO is a commonly used bioinformatics tool that provides comprehensive information on the gene function of individual genomic products based

on defined features consisting of three domains: biological process (BP), cellular component (CC), and molecular function (MF) [26]. We conducted GO analysis using the STRING database.

2.5 Analysis of the correlation between significant genes and breast cancer

Gene Expression Profiling Interactive Analysis (GEPIA, http://gepia.cancer-pku.cn) is a web server for analyzing RNA-sequencing expression data of 9,736 tumors and 8,587 normal samples from The Cancer Genome Atlas and Genotype-Tissue Expression projects, using a standard processing pipeline [27]. The Human Protein Atlas (HPA, www.proteinatlas.org) is an immunohistochemistry-based map of protein expression profiles in normal tissues, cancer tissues, and cell lines, and provides a resource for pathology-based biomedical research, including protein biomarker discovery [28–30]. Correlations between significant genes and breast cancers were analyzed with GEPIA and HPA.

2.6 Analysis of the correlation between significant gene expression and overall survival

PROGgeneV2 (http://www.compbio.iupui.edu/ proggene), a tool that can be used to predict the prognostic implication of genes in cancers, is written in PHP5 with a MySQL database backend, which stores gene expression data, covariates data, and metadata for cataloged studies in the form of relational database tables. Survival analysis in PROGgeneV2 is done using the backend R script; users can input multiple genes and use combined analysis to create survival plots for difgenes of interest [31]. We ferent used PROGgeneV2 to analyze the relationship between overall survival and genes that were overexpressed and underexpressed in breast cancer.

2.7 Drug prediction analysis

CMap (https://portals.broadinstitute.org/cmap/) is a collection of genome-wide transcriptional expression data from cultured human cells treated with bioactive small molecules and simple patternmatching algorithms that together enable the discovery of functional connections among drugs, genes, and diseases through the transitory feature of common gene expression changes [32–34]. We used CMap to identify small-molecule compounds as potential therapeutic agents to target the significant genes in breast cancer.

3. Results

3.1 Identification of significant DEGs in breast cancer

To identify the significant DEGs in breast cancer, we retrieved GWAS summary data (C50-C50) of breast cancer from the UK Biobank, and microarray expression profiles of breast cancer from the Oncomine database. C50-C50 contained 10,478 malignant neoplasm of breast cases and 235,016 controls for the analyses, and the data were consolidated and normalized (Figure 1(a,b)).

From NetWAS of GWAS summary data, we converted SNP-level association statistics into gene-level statistics (gene-based P values) and identified the 127 most relevant genes in breast cancer (Table 1). A total of 1019 overexpressed genes (Supplementary Figure 1) and 1019 under-expressed genes (Supplementary Figure 2) were identified by Oncomine. The top 20 DEGs in breast cancer compared to the normal controls are shown in a heatmap (Figure 2).

After overlapping the 127 most relevant genes with the 2038 DEGs in breast cancer, we identified 10 significant genes (*CLDN7*, *MLLT10*, *RBM33*, *SH3RF1*, *SSBP4*, *UBE2Z*, *BMPER*, *FGF7*, *MSRB3*,



Figure 1. A) Q-Q Plot of C50-C50 (β = 1.05, λ mean = 1.07, λ median = 1.06) Containing 9,113,133 imputed variants that passed quality control (QC), with a P different than 0; b) Manhattan PLOT (IMPUTED) CONTAINING all (QC and non-QC) 30,798,054 imputed variants; c) Venn diagram of significant genes (CLDN7,MLLT10,RBM33,SH3RF1,SSBP4, UBE2Z,BMPER,FGF7,MSRB3,TNRC6B) in breast cancer (BC).

 Table 1. Significant genes in breast cancer identified using NetWAS.

Table 1. (Continued).

DVL2 I 0.23598 SHR I 0.038987 SNR B I 0.16157 PIDD1 I 0.038947 ADSL I 0.16157 PIDD1 I 0.038947 ACADVL I 0.160621 TFAP2D I 0.037515 ACADVL I 0.149804 PARD66 I 0.037515 ACADVL I 0.149804 PATD1 I 0.034815 GCI3 I 0.149804 PATD2 I 0.034535 THEMINIAL I 0.14358 DLX2 I 0.033405 TICAS I 0.116397 NULRICG3 I 0.031321 TICAS I 0.116397 NULRICG3 I 0.032435 TICAS I 0.016700 ELOF1 I 0.026497 ATGBB I 0.106700 ELOF1 I 0.026497 SATGBC I 0.010717 DLL4 I 0.027807 SAT	Gene symbol	Training label	NetWAS Score	Gene symbol	Training label	NetWAS Score
SNF8 1 0.174439 NPLCca 1 0.03947 ADSL 1 0.16157 PIDD1 1 0.03847 ORAAE1 1 0.160651 THAP2D 1 0.03847 ACADVL 1 0.149640 PARD6R 1 0.03486 ACADVL 1 0.149615 PTEJ 1 0.03485 TIMMA 1 0.149616 PNDC1 1 0.03485 GCS 1 0.148613 POTEJ 1 0.03385 TICALS 1 0.128671 NELNGC3 1 0.03385 TICALS 1 0.013451 RASSF3 1 0.03124 DIDO1 1 0.16950 ELOF1 1 0.02469 DIDO1 1 0.107490 SPIB42 1 0.02469 GRO3R 1 0.024697 1 0.02345 SCADE1 1 0.098124 FZD7 1 0.024697 GRO3R 1	DVL2	1	0.235936	SH3RF1	1	0.039898
ADSL 1 0.16137 PDDT 1 0.033907 ACADVL 1 0.160621 TFAP2D 1 0.033907 ACADVL 1 0.140601 TFAP2D 1 0.033907 ACADVL 1 0.140601 TFAP2D 1 0.03481 CR3A 1 0.140505 PNC1 1 0.034835 TMEMINIA 1 0.140605 PNC1 1 0.033803 SPDYE3 1 0.120807 NEUROC3 1 0.031241 PHF23 1 0.101637 RAS573 1 0.031241 PHF23 1 0.106706 ELOF1 1 0.029807 DDO1 1 0.106404 SPTB2 1 0.029807 SABBP 1 0.106404 SPTB2 1 0.026407 CAND 1 0.03645 FRP2 1 0.026407 SABBP 1 0.016451 KLC4 1 0.023515 SACADYA	SNF8	1	0.174439	NPI OC4	1	0.039647
012AE1 1 0.160621 MIET 1 0.033715 012A5 1 0.149644 PARD6B 1 0.03494 GIG3 1 0.149613 POTEL 1 0.03494 GIG3 1 0.140566 FNOC1 1 0.033865 GIG3 1 0.128071 NEUROG3 1 0.033865 TRAM 1 0.128071 NEUROG3 1 0.033865 SPUYE3 1 0.128071 NEUROG3 1 0.03144 PIF23 1 0.116377 RASS73 1 0.03247 DID01 1 0.106509 WH33 1 0.02649 SPTBN2 1 0.02649 SPTBN2 1 0.02649 SA230BP 1 0.010717 DL4 1 0.03357 SA230BP 1 0.010717 DL4 1 0.03357 SA230BP 1 0.010717 DL4 1 0.03357 SA222A17 <td>ADSL</td> <td>1</td> <td>0.16157</td> <td></td> <td>1</td> <td>0.032047</td>	ADSL	1	0.16157		1	0.032047
ACADVL 1 0.14904 TAADD 1 0.034994 TRIMA 1 0.14904 PADO66 1 0.034994 TRIMA 1 0.140556 FNDC1 1 0.034835 TMEMIGIA 1 0.140956 FNDC1 1 0.033803 SPDYE3 1 0.128971 NEUBOG3 1 0.031343 SPDYE3 1 0.016357 RASSF3 1 0.031344 PH723 1 0.016056 REG 1 0.029867 DDO1 1 0.160596 WIT3 1 0.027907 SYDSB 1 0.101494 SPBN2 1 0.026649 DDO1 1 0.04941 FUC21 1 0.023902 GRIC2IP 1 0.038176 PLCE1 1 0.023602 SYDSBP 1 0.039059 ACIT4 0.023157 SCAMP2 1 0.023612 GRIC2IP 1 0.039059 REG 1 <t< td=""><td>OR2AE1</td><td>1</td><td>0.160822</td><td>MIEE1</td><td>1</td><td>0.038407</td></t<>	OR2AE1	1	0.160822	MIEE1	1	0.038407
OR2AS 1 0.148013 PAROGE 1 0.0348 GC3 1 0.148013 POTE 1 0.0348 GC4 1 0.143088 DV2 1 0.03385 TRGA5 1 0.120586 FNC1 1 0.03137 SPDVE3 1 0.120586 SPEG 1 0.03137 LRRD1 1 0.116377 RAST33 1 0.03146 DID01 1 0.105079 WNT3 1 0.027907 KRT1 1 0.106476 EUC1 1 0.026479 SAP30BP 1 0.100171 DL4 1 0.023357 GRD2P 1 0.004211 KLRC4 1 0.023357 SC2SAT7 1 0.039066 PSTK 1 0.023357 TMEM164A 1 0.039365 NEUROPA 1 0.023453 SCAMP2 1 0.039066 NEUROPA 1 0.018473 SCAMP2	ACADVL	1	0.160651	TFAP2D	1	0.030407
TRIM4 1 0.148013 POTD 1 0.0348 GC3 1 0.146056 PHDC1 1 0.034935 TMEMIGIA 1 0.13087 NEUPOG3 1 0.033835 SPDYE3 1 0.128071 NEUPOG3 1 0.03134 SPDYE3 1 0.11637 RASFIS 1 0.02986 AFG96 1 0.011637 RASFIS 1 0.02986 AFG96 1 0.106706 BLOF1 0.02849 DIDO1 1 0.100740 SPTBN2 1 0.028649 CCND1 1 0.100471 RLKA 1 0.023602 SK25A17 1 0.093676 PSTK 1 0.023151 TKEM184A 1 0.093676 PSTK 1 0.023151 SCXMP2 1 0.093676 PSTK 1 0.023151 SCXMP2 1 0.093757 PSCQ2 1 0.016477 SCAMP2 <td< td=""><td>OR2A5</td><td>1</td><td>0.149804</td><td>PARDER</td><td>1</td><td>0.037313</td></td<>	OR2A5	1	0.149804	PARDER	1	0.037313
GC3 1 0.04398 TMEM161A 1 0.014308 TMEM161A 1 0.014308 TMEM161A 1 0.014308 SPVFG3 1 0.013080 SPVFG3 1 0.013087 RSST3 1 0.013137 RSST3 1 0.013137 RSST3 1 0.013137 RSST3 1 0.013137 RSST3 1 0.013147 RSST3 1 0.013147 RSST3 1 0.013147 RSST3 1 0.02346 EUC1 1 0.023476 SPT8QP 1 0.010449 SPT8Q2 1 0.023467 SP308P 1 0.010449 SPT8Q2 1 0.023467 SP308P 1 0.010449 SPT8Q2 1 0.023467 SP308P 1 0.010449 SPT8Q2 1 0.023467 SP308P 1 0.010449 SCQNP 1 0.004911 ELG TMEM16AA 1 0.093976 RST3 1 0.02357 TMEM16AA 1 0.093976 RST2 1 0.023577 TMEM16AA 1 0.093976 RST2 1 0.023577 TMEM16AA 1 0.093976 SCQNP 1 0.003976 SCQNP 1 0.003977 SCQNP 1 0.003978 SCQNP 1 0.003978 SCQ	TRIM4	1	0.148613	POTEI	1	0.034554
TMEMIGIA 1 0.13088 DX2 1 0.03389 SPDVE3 1 0.128071 NEUPOG3 1 0.031303 SPDVE3 1 0.116527 RASJF3 1 0.03134 PHF23 1 0.011657 RASJF3 1 0.02987 AFG98 1 0.011690 WNT3 1 0.02987 DIDO1 1 0.106706 ELOF1 1 0.02849 DIDO1 1 0.10071 D.024649 SPSN2 1 0.02649 SKP30EP 1 0.009634 FLCE1 1 0.02315 SKP30EP 1 0.093676 PSTN2 1 0.02315 SKZ5A17 1 0.093676 PSTK 1 0.02315 SCXMP2 1 0.093676 NEUFCDA 1 0.02315 SCXMP1 1 0.093676 NEUFCDA 1 0.01637 SCXMP2 1 0.043734 SCGD3 1 0.01637	GJC3	1	0.146056	FNDC1	1	0.034535
ITGAS 1 0.120580 SPUPE3 1 0.031537 LRRD1 1 0.110586 SPEC 1 0.031537 LRRD1 1 0.116357 RASSF3 1 0.031537 LRRD1 1 0.116367 RLSSF3 1 0.03957 ATG9B 1 0.106706 ELOF1 1 0.039574 SAP30DP 1 0.103454 ICFIP2 1 0.024607 SAP30DP 1 0.100449 SPTNN2 1 0.024607 SAP30DP 1 0.009812 FLD7 1 0.024607 SAP30DP 1 0.098376 PyTK 1 0.023315 SCAMP2 1 0.093005 ASICA 1 0.023457 SCAMP2 1 0.093076 PyTK 1 0.014457 SCAMP2 1 0.093056 ASICA 1 0.014637 LPP1 1 0.08524 SCASD3 1 0.014632 <t< td=""><td>TMEM161A</td><td>1</td><td>0.143088</td><td></td><td>1</td><td>0.03455</td></t<>	TMEM161A	1	0.143088		1	0.03455
SPDVE3 1 0.10366 Dec 0.031337 IRRD1 1 0.11637 RASEF2 1 0.03124 PHF33 1 0.11637 RASEF2 1 0.03124 ATG96 1 0.106706 WNT3 1 0.023976 DID01 1 0.106706 WNT3 1 0.024967 SPD0EP 1 0.100449 SFBVU2 1 0.02467 CCND1 1 0.098912 DLL4 1 0.024867 SEX25A17 1 0.098912 PLC1 1 0.022315 SCAM72 1 0.093056 PSTK 1 0.022315 SCAM72 1 0.093055 NEURO24 1 0.014473 LP11 1 0.08324 RSCSD3 1 0.014473 SCAM72 1 0.024654 SCCSD3 1 0.0148473 PEG10 1 0.08324 SCCSD3 1 0.014682 LP11 1 <td>ITGA5</td> <td>1</td> <td>0.128071</td> <td>NEUROG3</td> <td>1</td> <td>0.033203</td>	ITGA5	1	0.128071	NEUROG3	1	0.033203
LRRD1 1 0.116337 PLS3F2 1 0.01374 PHF33 1 0.106706 ELOT1 1 0.039676 ATG98 1 0.106706 ELOT1 1 0.039676 CND01 1 0.10649 WNT3 1 0.026449 KRT1 1 0.10649 SPTBN2 1 0.02649 CND1 1 0.10949 SPTBN2 1 0.02649 GRID2IP 1 0.096814 PLCE1 1 0.0223357 TMEM184A 1 0.093056 SSIL4 1 0.0223357 TMEM184A 1 0.093056 SSIL4 1 0.02243 SCAMP2 1 0.093056 NEUROD4 1 0.018473 SCAMP2 1 0.083947 SLCSD3 1 0.018473 SCAMP2 1 0.083989 NMKS 1 0.016417 SCAMP2 1 0.083989 NMKS 1 0.016417	SPDYE3	1	0.120586	SPEG	1	0.033203
PHF23 1 0.11639 COV 1 0.02976 AlFG9B 1 0.106706 ELOF1 1 0.029767 DIDO1 1 0.106706 ELOF1 1 0.027907 KRIT1 1 0.106449 SPTBN2 1 0.02649 CCND1 1 0.100471 SPTSN2 1 0.026407 GRD2P 1 0.099812 FLD 1 0.024107 GRD2P 1 0.099812 FLD 1 0.024107 SLCSAT7 1 0.098306 PSTK 1 0.022115 SCZD4 1 0.939306 PSTK 1 0.022115 SCZD4 1 0.939306 NURIO14 1 0.021315 SCMP1 1 0.939306 SCSD3 1 0.0148473 SCAMP2 1 0.048324 SCSD3 1 0.014843 IPP1 1 0.048324 SCSD3 1 0.014852 IPM2	LRRD1	1	0.116357	RASSES	1	0.03124
ATG98 1 0.106706 ELOF1 1 0.022149 KRT1 1 0.103645 IGFBP2 1 0.02649 KRT1 1 0.103445 IGFBP2 1 0.02649 SAP308P 1 0.100171 DIL4 1 0.024667 GR102P 1 0.098334 PLCE1 1 0.023692 SLCSA177 1 0.099166 PSTK 1 0.022163 TMEM184A 1 0.093066 PSTK 1 0.0221315 SCCAMP2 1 0.093066 NEUROD4 1 0.0221315 SCCMP1 1 0.083944 RSC03 1 0.016897 SCAMP2 1 0.083984 ADC73 1 0.016532	PHF23	1	0.111639	CBX7	1	0.03121
DIDO1 1 0.105009 WNT3 1 0.027607 KRT1 1 0.106449 SPTBN2 1 0.02649 CCND1 1 0.100149 SPTBN2 1 0.02647 GRID2IP 1 0.099812 FZD7 1 0.022667 MAFK 1 0.099812 FZD7 1 0.022357 SCAMP2 1 0.093666 PSTK 1 0.022315 SCAMP2 1 0.093005 NEUROD4 1 0.022131 CCWPW1 1 0.083244 SCASD3 1 0.016478 SCAMP2 1 0.085924 SCASD3 1 0.016478 LRP1 1 0.086398 ADCG2 1 0.016478 SWL17 1 0.085924 SCASD3 1 0.016478 NUC17 1 0.085924 SCASD3 1 0.016479 MUC17 1 0.065686 FAM71E2 1 0.016522 C	ATG9B	1	0.106706	FLOF1	1	0.029070
KRIT 1 0.103645 IGFBP2 1 0.022449 SAP308P 1 0.100171 DLL4 1 0.02607 GRID2IP 1 0.099812 FZD7 1 0.0224607 GRID2IP 1 0.099812 FZD7 1 0.0223692 SLCSA17 1 0.099306 ASIC4 1 0.022315 SCXD41 1 0.093096 ASIC4 1 0.022743 TKMEM184A 1 0.093096 ASIC4 1 0.022743 SCXDP2 1 0.093096 ASIC4 1 0.022743 SCMP2 1 0.093096 ASIC4 1 0.022743 SCMP2 1 0.085924 SIC2503 1 0.016473 FEG10 1 0.082857 ASICG2 1 0.0164825 MUC12 1 0.080398 MIMIN5 1 0.015532 MUC12 1 0.062894 SSIP4 1 0.014533	DID01	1	0.105009	WNT3	1	0.020110
CCND1 1 0.100449 SPTIN2 1 0.02607 SAP30BP 1 0.100171 DLL 1 0.024507 GRID2IP 1 0.096812 FZD7 1 0.024507 MARK 1 0.096812 FZD7 1 0.0224165 MARK 1 0.099612 FLCE 1 0.022357 TMEMIBAA 1 0.093005 NEUROD4 1 0.022135 SCAMP2 1 0.093005 NEUROD4 1 0.0118987 LRP1 1 0.085924 SC3503 1 0.016428 SCMP33 1 0.085924 SC3503 1 0.016534 FNR2 1 0.085937 ABCG2 1 0.016534 MUC12 1 0.065968 SMM/S 1 0.015532 MUC12 1 0.06586 FAM71E2 1 0.014252 CSK 1 0.06584 SEP44 1 0.014525 CA	KRIT1	1	0.103645	IGERP2	1	0.026449
SAP308P 1 0.100171 DL4 1 0.024367 MAFK 1 0.098814 PLCE1 1 0.02367 SLC25A177 1 0.098814 PLCE1 1 0.023367 SLC25A177 1 0.093076 PSTK 1 0.022315 SLC25A17 1 0.093076 PSTK 1 0.022743 SLC25A17 1 0.093076 NEROD4 1 0.022743 SLC25A17 1 0.093076 NEROD4 1 0.022743 SCXUPVI 1 0.083574 SLC35D3 1 0.0168473 PEG10 1 0.088398 SLC35D3 1 0.0168473 MUC17 1 0.080398 SMIN5 1 0.016354 MUC12 1 0.066909 Clarof60 1 0.015322 MUC12 1 0.066794 Clarof60 1 0.014373 MUC12 1 0.066784 SSP4 1 0.014372 FGFR2 1 0.062734 KTK17 1 0.006984	CCND1	1	0.100449	SPTRN2	1	0.02607
GRID2IP 1 0.099812 FZD7 1 0.024165 SLC25A17 1 0.094211 KLRC4 1 0.023052 SLC25A17 1 0.093676 PSTK 1 0.023157 TMEM184A 1 0.093676 PSTK 1 0.022165 SCAMP2 1 0.093005 NEUROD4 1 0.018473 SCAMP2 1 0.093005 NEUROD4 1 0.018473 LRP1 1 0.088344 RSP03 1 0.016478 SCMPW1 1 0.084988 ADCY3 1 0.016478 FNP2 1 0.080699 DNAH9 1 0.016354 TNRC6B 1 0.080699 DNAH9 1 0.015522 C7orf61 1 0.066384 SPIN12 1 0.01477 AVL 1 0.066384 GPIN12 1 0.01452 C7orf61 1 0.066374 SUSP14 1 0.01452	SAP30BP	1	0.100171	DI14	1	0.024507
MARK 1 0.096834 PLCE1 1 0.023602 SLC25A17 1 0.093676 PSTK 1 0.023357 TMEM184A 1 0.093076 PSTK 1 0.022713 TSC22D4 1 0.093096 ASIC4 1 0.022713 ZCWPW1 1 0.092518 ZDHHC24 1 0.018937 ZCWPW1 1 0.085924 SLC35D3 1 0.016417 FNF2 1 0.088938 ADCY3 1 0.016398 MUC17 1 0.080398 SMMMS 1 0.015522 C7orf61 1 0.066666 FAM71E2 1 0.01477 MUC12 1 0.066784 SSBP4 1 0.014352 CSK 1 0.066788 USHBP1 1 0.014373 PLIMA 1 0.066784 SSBP4 1 0.01313 CSK 1 0.067349 KCNK17 1 0.003467 <t< td=""><td>GRID2IP</td><td>1</td><td>0.099812</td><td>FZD7</td><td>1</td><td>0.024165</td></t<>	GRID2IP	1	0.099812	FZD7	1	0.024165
SLC25A17 1 0.094211 KLRC4 1 0.023357 TMEMTIBAA 1 0.093066 PSTK 1 0.023215 SCAMP2 1 0.093066 ASIC4 1 0.022743 SCAMP2 1 0.093076 NEURODA4 1 0.021315 ZCMPW1 1 0.083344 RSP03 1 0.016897 LRP1 1 0.083344 RSP03 1 0.016897 PIR10 1 0.083344 RSC320 1 0.016834 RBM33 1 0.084988 ADCY3 1 0.016354 MUC17 1 0.080609 DNAH9 1 0.016354 MUC12 1 0.070727 CTDNEP1 1 0.015522 C7or6f1 1 0.066949 C12orf80 1 0.014757 AVIL 1 0.066378 CLDN7 1 0.014757 CSK 1 0.06538 ZHF18 1 0.014219 AZGP1 1 0.06538 ZHF18 1 0.009944 <t< td=""><td>MAFK</td><td>1</td><td>0.096834</td><td>PLCF1</td><td>1</td><td>0.023692</td></t<>	MAFK	1	0.096834	PLCF1	1	0.023692
TMEMI8AA 1 0.093676 PSTK 1 0.023213 TSC22D4 1 0.093005 NEUROD4 1 0.022373 SCAMP2 1 0.093005 NEUROD4 1 0.022313 ZCWPVI 1 0.093005 NEUROD4 1 0.023135 ZCWPVI 1 0.093905 NEUROD4 1 0.018473 ZCWPVI 1 0.085924 SLC35D3 1 0.016428 RBM33 1 0.08699 DNAH9 1 0.016354 TINRC6B 1 0.080398 SMIMS 1 0.015322 Cordef1 1 0.06099 C120rf80 1 0.014252 Cordef1 1 0.062894 SSBP4 1 0.014552 Cordef1 1 0.062179 DLG4 1 0.014553 OR12D2 1 0.062179 DLG4 1 0.014553 VNIL 1 0.062179 DLG4 1 0.01453 OR12D2 1 0.065178 UNF18 1 0.009367	SLC25A17	1	0.094211	KI RC4	1	0.023357
TSC2D4 1 0.093096 ASIC4 1 0.022749 SCAMP2 1 0.093005 NEURDD4 1 0.021315 ZCWPVI 1 0.092518 ZDHHC24 1 0.018987 LRP1 1 0.088344 RSP03 1 0.016328 RBM33 1 0.084988 ADCY3 1 0.016324 FNIP2 1 0.08009 DNAH9 1 0.016354 TNRC68 1 0.080398 SMIMS 1 0.015522 VILC12 1 0.07277 CTONEP1 1 0.014825 CSK 1 0.06686 FAM71E2 1 0.01477 AVIL 1 0.06686 FAM71E2 1 0.01477 AVIL 1 0.066381 CLDN7 1 0.014572 GFGR2 1 0.06588 ZNF18 1 0.014572 AZGP1 1 0.06588 ZNF18 1 0.01336 PULM4 1 0.065788 ZNF18 1 0.0039367 ATPSVOA4 <td>TMEM184A</td> <td>1</td> <td>0.093676</td> <td>PSTK</td> <td>1</td> <td>0.023215</td>	TMEM184A	1	0.093676	PSTK	1	0.023215
SCAMP2 1 0.093005 NEUROD4 1 0.003315 ZCWPWIT 1 0.0920518 ZOHHC24 1 0.018973 LRP1 1 0.08344 RSPO3 1 0.018973 PEG10 1 0.083924 SLC35D3 1 0.016473 FNIP2 1 0.082357 ABCG2 1 0.016354 TNRC68 1 0.080398 SMIMS 1 0.015352 MUC12 1 0.06090 DNAH9 1 0.015352 Corlof1 1 0.060949 C120r80 1 0.014825 CV12 1 0.067048 SRPH12 1 0.014825 CSK 1 0.067179 DLC4 1 0.013153 PILRB 1 0.061788 USHBP1 1 0.01333 PDLM4 1 0.06583 ZNF18 1 0.019335 MMD 1 0.05583 PHE20 1 0.003945 MM	TSC22D4	1	0.093096	ASIC4	1	0.022743
ZCWPW1 1 0.092518 ZDHHC24 1 0.01897 LRP1 1 0.083344 RSP03 1 0.018473 PG10 1 0.085924 SLC35D3 1 0.018478 RBM33 1 0.084998 ADCY3 1 0.016334 TNRC6B 1 0.080609 DNAH9 1 0.01532 MUC17 1 0.080398 SMIMS 1 0.01532 MUC12 1 0.070727 CTINEP1 1 0.014522 C70r61 1 0.062894 SSP4 1 0.01477 AVIL 1 0.062179 DL64 1 0.01313 PLIB8 1 0.060581 ZNF18 1 0.010934 MND 1 0.057349 KCLN77 1 0.012121 AZ6P1 1 0.057349 KCLN77 1 0.012219 AZ5P1 1 0.057349 KCKK17 1 0.009342 CARS <td>SCAMP2</td> <td>1</td> <td>0.093005</td> <td>NFUROD4</td> <td>1</td> <td>0.021315</td>	SCAMP2	1	0.093005	NFUROD4	1	0.021315
LRP1 1 0.088344 RSP03 1 0.018473 PEG10 1 0.088324 SLG3D3 1 0.018278 PEG10 1 0.085924 SLG3D3 1 0.016278 FNP2 1 0.084998 ADCY3 1 0.016417 FNP2 1 0.082357 AECG2 1 0.01634 MUC17 1 0.080398 SMMS 1 0.01532 MUC12 1 0.070727 CTDNEP1 1 0.015522 C7arf61 1 0.060499 C12orf80 1 0.014572 FGFR2 1 0.066484 GPRIN2 1 0.014772 FGFR2 1 0.062894 SSBP4 1 0.014572 FGFR2 1 0.05685 ZNF18 1 0.013133 PDLM4 1 0.060558 ZNF18 1 0.010693 WNT2 1 0.05583 PHE20 1 0.00984 MMD 1 0.0557349 KCNK17 1 0.00984 MMD 1 0.0557349 KCNK17 1 0.00984 FMMD 1 0.0557349 KCNK17 1 0.00984 FMMD 1 0.0557349 KCNK17 1 0.00984 FMMD 1 0.0557349 KCNK17 1 0.00984 FMMB 1 0.0557349 KCNK17 1 0.00984 FMMD 1 0.0557349 KCNK17 1 0.00984 FMMD 1 0.0557349 KCNK17 1 0.00984 FMMA8 1 0.05566 PNLP 1 0.0009367 ATP6V0A4 1 0.055734 PHE20 1 0.009367 ATP6V0A4 1 0.055734 ZBTB2 1 0.008477 FMFV17L2 1 0.053732 RBM43 1 0.007988 FMM48 1 0.055666 PNLP 1 0.000578 FMPER 1 0.055618 CASC10 1 0.000734 FMH42 1 0.055618 CASC10 1 0.0007176 FMLDA2 1 0.048975 MLLT10 1 0.00578 FMF41 1 0.048975 MLLT10 1 0.00578 FMF41 1 0.048975 MLLT10 1 0.00578 FMF41 1 0.048749 SFAG4 1 0.00578 FMF45 1 0.044874 SFAG4 1 0.00578 FMF45 1 0.044874 SFAG4 1 0.00578 FMF45 1 0.044874 SFAG4 1 0.00578 FMF45 1 0.044975 MLLT10 0.00578 FMF45 1 0.044975 MLLT10 1 0.00578 FMF45 1 0.044975 MLLT10 1 0.00578 FMF45 1 0.0	ZCWPW1	1	0.092518	ZDHHC24	1	0.018987
PEG10 1 0.085924 SLC35D3 1 0.016228 RBM33 1 0.084988 ADCY3 1 0.016354 RBM33 1 0.082357 ABCG2 1 0.016354 TNRC6B 1 0.080398 SMIM5 1 0.015532 MUC17 1 0.060394 Cloarf80 1 0.015532 MUC12 1 0.070727 CTDNEP1 1 0.014825 CSK 1 0.065686 FAM71E2 1 0.014572 AVIL 1 0.062894 SSBP4 1 0.014573 AVIL 1 0.062894 SSBP4 1 0.014533 OR12D2 1 0.062894 SSBP4 1 0.014533 PLIM4 1 0.062894 SSBP4 1 0.014533 PVIT2D2 1 0.06583 ZNF18 1 0.01693 MNT2 1 0.05513 PHF20 1 0.009367 ATF	LRP1	1	0.088344	RSPO3	1	0.018473
RBM33 1 0.084988 ADCY3 1 0.016417 FNIP2 1 0.0802357 ABCG2 1 0.016334 TNRC6B 1 0.080509 DNAH9 1 0.016334 MUC17 1 0.070727 CTONEP1 1 0.015522 C7orf61 1 0.065686 FAM71E2 1 0.014727 AVIL 1 0.065686 FAM71E2 1 0.014572 Corf61 1 0.062494 SSBP4 1 0.01477 AVIL 1 0.062484 SSP4 1 0.01477 GR12D2 1 0.062179 DLG4 1 0.01316 PLIAB 1 0.060538 ZNF18 1 0.012219 MD1 1 0.057349 KCNK17 1 0.009367 ATP6V0A4 1 0.057349 KCNK17 1 0.009367 ATP6V0A4 1 0.057352 RBM43 1 0.007948 <	PEG10	1	0.085924	SI C35D3	1	0.016828
FNIP2 1 0.082357 ABCG2 1 0.016354 TNRC6B 1 0.080609 DNAH9 1 0.016354 MUC17 1 0.080398 SMIM5 1 0.015522 Cordf61 1 0.070277 CTDNFP1 1 0.014572 Cordf61 1 0.066866 FAM71E2 1 0.014777 AVIL 1 0.062894 SSBP4 1 0.014572 FGR2 1 0.062894 SSBP4 1 0.014493 OR12D2 1 0.062179 DLG4 1 0.014533 PLIM4 1 0.060588 ZNF18 1 0.012219 AZCP1 1 0.057349 KCNK17 1 0.00944 MMD 1 0.057349 KCNK17 1 0.009442 CAPS 1 0.057349 KCNK17 1 0.009479 MPVT2 1 0.057359 PHE20 1 0.007724 PM	RBM33	1	0.084988	ADCY3	1	0.016417
TNRC6B 1 0.080609 DNAH9 1 0.016198 MUC17 1 0.080398 SMIM5 1 0.015532 MUC12 1 0.070727 CTDNEP1 1 0.01552 C7orf61 1 0.065686 FAM71E2 1 0.014825 CSK 1 0.064848 GPRIN2 1 0.014873 OR12D2 1 0.062179 DL64 1 0.01316 PILRB 1 0.060558 ZNF18 1 0.01219 AZCP1 1 0.060558 ZNF18 1 0.01693 WNT2 1 0.053739 KCNK17 1 0.009442 CARS 1 0.0537349 KCNK17 1 0.009195 BMPER 1 0.0537349 RCNK17 1 0.009442 CARS 1 0.053732 RBM43 1 0.007944 PV17L2 1 0.053766 PNLP 1 0.007244 P4HA	FNIP2	1	0.082357	ABCG2	1	0.016354
MUC17 1 0.080398 SMMS 1 0.015332 MUC12 1 0.070727 CTDNEP1 1 0.015332 Cordf01 1 0.069049 C12orf80 1 0.014825 CSK 1 0.064848 GPIN2 1 0.014825 CSK 1 0.062894 SSBP4 1 0.014825 GRIZD2 1 0.062179 DLG4 1 0.013133 PULM4 1 0.060831 CLDN7 1 0.012219 AZCP1 1 0.06584 ZNF18 1 0.009347 MMD 1 0.057349 KCNK17 1 0.009367 ATP6V0A4 1 0.055112 TNF1 1 0.009367 ATP6V0A4 1 0.053732 RBM43 1 0.009367 ATP6V0A4 1 0.053732 RBM43 1 0.007948 RBM48 1 0.05366 PNLIP 1 0.007716 P	TNRC6B	1	0.080609	DNAH9	1	0.016198
MUC12 1 0.070727 CTDNEP1 1 0.015522 C7orf61 1 0.069049 C12orf80 1 0.014825 CSK 1 0.06568 FAM71E2 1 0.014777 AVIL 1 0.064848 GPRIN2 1 0.014772 FGR2 1 0.062179 DLG4 1 0.013133 PLIB8 1 0.061788 USHBP1 1 0.013133 PDLIM4 1 0.060588 ZNF18 1 0.01219 MMD 1 0.057349 KCNK17 1 0.009367 ATP6V0A4 1 0.055112 TNP1 1 0.009367 MPUT2 1 0.055372 RBM43 1 0.007349 MPV17L2 1 0.055372 RBM43 1 0.007349 MPV17L2 1 0.055372 RBM43 1 0.007746 PHLA2 1 0.055381 CASC10 1 0.007375 <	MUC17	1	0.080398	SMIM5	1	0.015532
C7orf61 1 0.069049 C12orf80 1 0.014825 CSK 1 0.065486 FAM71E2 1 0.014777 AVIL 1 0.064848 GPRN2 1 0.014572 FGFR2 1 0.062894 SSBP4 1 0.01316 OR12D2 1 0.062179 DLG4 1 0.01313 PDLIM4 1 0.060831 CLDN7 1 0.012219 AZGP1 1 0.05588 ZNF18 1 0.00984 MMD 1 0.05512 TNL6 1 0.009155 BMPER 1 0.055432 PHE20 1 0.009442 CARS 1 0.055132 TNP1 1 0.009155 BMPER 1 0.055132 RBM43 1 0.007244 PH1A2 1 0.055163 CASC10 1 0.00776 PH4A2 1 0.055163 CASC10 1 0.005735 ISYNA1	MUC12	1	0.070727	CTDNEP1	1	0.015522
CSK 1 0.065686 FAM71E2 1 0.014777 AVIL 1 0.064848 GPRIN2 1 0.014572 AVIL 1 0.062844 SSB44 1 0.014573 GR12D2 1 0.062179 DLG4 1 0.01316 PILBR 1 0.06558 USHBP1 1 0.012219 AZGP1 1 0.06558 ZNF18 1 0.00984 MMD 1 0.055427 TTLL6 1 0.009442 CARS 1 0.055112 TNP1 1 0.009956 MMPER 1 0.055372 RBM43 1 0.007746 PV17L2 1 0.053696 PNLIP 1 0.007746 PHLDA2 1 0.053696 PNLIP 1 0.007746 PHLDA2 1 0.05613 ELL 1 0.005735 ISYNA1 1 0.048975 MLIP10 1 0.005735 USP251	C7orf61	1	0.069049	C12orf80	1	0.014825
AVIL 1 0.064848 GPRIN2 1 0.014572 FGFR2 1 0.062894 SSBP4 1 0.014453 OR12D2 1 0.062179 DLG4 1 0.013133 PLIR8 1 0.060831 CLDN7 1 0.012219 AZGP1 1 0.060831 CLDN7 1 0.00984 MMD 1 0.05583 ZNF18 1 0.00984 MMD 1 0.05513 TTLL6 1 0.00987 ATF6V0A4 1 0.055112 TNP1 1 0.009798 BMPER 1 0.055372 RBM43 1 0.007744 PV17L2 1 0.055366 PNLIP 1 0.007744 P4HA2 1 0.055081 CASC10 1 0.007744 P4HA2 1 0.0505181 ELL 1 0.005735 NAP1L4 1 0.048975 MLT10 1 0.005735 SVNA1	CSK	1	0.065686	FAM71E2	1	0.014777
FGFR2 1 0.062894 SSBP4 1 0.014453 OR12D2 1 0.062179 DLG4 1 0.01313 PILRB 1 0.060831 CLDN7 1 0.01219 AZGP1 1 0.060831 CLDN7 1 0.012219 AZGP1 1 0.060558 ZNF18 1 0.009842 MMD 1 0.05583 PHF20 1 0.009442 CARS 1 0.05583 PHF20 1 0.009472 MMD 1 0.05583 PHF20 1 0.009477 ATF6V0A4 1 0.05572 ZBB2 1 0.008477 MPV17L2 1 0.053732 RBM43 1 0.00776 PHLA2 1 0.050561 CASC10 1 0.007176 PHLDA2 1 0.050563 ELL 1 0.005583 UBE2Z 1 0.048027 DF311 1 0.005735 ISYNA1 1 0.048027 DF311 1 0.005333 MEFCE <td< td=""><td>AVIL</td><td>1</td><td>0.064848</td><td>GPRIN2</td><td>1</td><td>0.014572</td></td<>	AVIL	1	0.064848	GPRIN2	1	0.014572
OR12D2 1 0.062179 DLG4 1 0.01316 PILRB 1 0.061788 USHBP1 1 0.013133 PDLIM4 1 0.060578 CLDN7 1 0.01219 AZGP1 1 0.060558 ZNF18 1 0.00984 MMD 1 0.05543 PHE20 1 0.009944 CARS 1 0.05512 TNP1 1 0.009947 ATP6V0A4 1 0.05512 TNP1 1 0.009442 CARS 1 0.055437 PHE20 1 0.009477 MPV17L2 1 0.053732 RBM43 1 0.007948 MAB4 1 0.050581 CASC10 1 0.007176 PHLDA2 1 0.050581 CASC10 1 0.005181 NAP1L4 1 0.049917 SKIDA1 1 0.005735 ISYNA1 1 0.048024 SPAG4 1 0.005735 ISYNA1 </td <td>FGFR2</td> <td>1</td> <td>0.062894</td> <td>SSBP4</td> <td>1</td> <td>0.014453</td>	FGFR2	1	0.062894	SSBP4	1	0.014453
PILRB 1 0.061788 USHBP1 1 0.013133 PDLIM4 1 0.060831 CLDN7 1 0.012219 AZGP1 1 0.060558 ZNF18 1 0.010693 WNT2 1 0.057349 KCNK17 1 0.009442 CARS 1 0.057349 KCNK17 1 0.009367 ATP6V0A4 1 0.055112 TNP1 1 0.009798 BMPER 1 0.053732 RBM43 1 0.007748 RBM48 1 0.053666 PNLIP 1 0.007746 P4HA2 1 0.050581 CASC10 1 0.007746 PHLDA2 1 0.0506163 ELL 1 0.005735 ISWNA1 1 0.049917 SKIDA1 1 0.005735 ISWNA1 1 0.048267 ODF3L1 1 0.005735 ISWNA1 1 0.048267 ODF3L1 1 0.005735 <	OR12D2	1	0.062179	DLG4	1	0.01316
PDLIM4 1 0.060831 CLDN7 1 0.012219 AZGP1 1 0.060558 ZNF18 1 0.010693 MMD 1 0.057349 KCNK17 1 0.009442 CARS 1 0.05583 PHF20 1 0.009472 CARS 1 0.055112 TNP1 1 0.008477 MPV17L2 1 0.053732 RBM43 1 0.007948 BMPER 1 0.050581 CASC10 1 0.00776 PHLDA2 1 0.05066 PLIP 1 0.00776 PHLDA2 1 0.050581 CASC10 1 0.00776 PHLDA2 1 0.050581 CASC10 1 0.00776 PHLDA2 1 0.049975 MLIP1 1 0.005735 ISYNA1 1 0.048287 ODF311 1 0.00735 ISYNA1 1 0.047283 TAR8 1 0.00333 MEPCE <td>PILRB</td> <td>1</td> <td>0.061788</td> <td>USHBP1</td> <td>1</td> <td>0.013133</td>	PILRB	1	0.061788	USHBP1	1	0.013133
AZGP1 1 0.060558 ZNF18 1 0.010693 WNT2 1 0.058427 TLL6 1 0.00944 MMD 1 0.057349 KCNK17 1 0.009420 CARS 1 0.05583 PHF20 1 0.009367 ATP6V0A4 1 0.055112 TNP1 1 0.008477 MPV17L2 1 0.053732 RBM43 1 0.007244 P4HA2 1 0.050581 CASC10 1 0.007244 P4HA2 1 0.050581 CASC10 1 0.00588 NAP1L4 1 0.049917 SKIDA1 1 0.00588 NR2F6 1 0.048975 MLL110 1 0.00533 UBE2Z 1 0.048975 MLL110 1 0.00533 UBE2Z 1 0.048975 MLL110 1 0.00533 UYP31A1 1 0.047283 TAAR8 1 0.00333 MEPCE 1 0.047283 TAAR8 1 0.003333 MEPCE	PDLIM4	1	0.060831	CLDN7	1	0.012219
WNT2 1 0.058427 TTLL6 1 0.00984 MMD 1 0.05739 KCNK17 1 0.009442 CARS 1 0.05583 PHF20 1 0.009367 ATP6V0A4 1 0.055112 TNP1 1 0.008477 MPVT7L2 1 0.053732 RBM43 1 0.007998 RBM48 1 0.050581 CASC10 1 0.007176 PHLDA2 1 0.05061 ELL 1 0.00598 NAP1L4 1 0.049917 SKIDA1 1 0.00583 UBE2Z 1 0.048287 OPS11 1 0.005735 ISYNA1 1 0.047283 TAAR8 1 0.005733 CYP51A1 1 0.047283 TAAR8 1 0.003733 GIP 1 0.047158 UBALD2 1 0.002579 GIP 1 0.047158 UBALD2 1 0.002363 CYP251 <td>AZGP1</td> <td>1</td> <td>0.060558</td> <td>ZNF18</td> <td>1</td> <td>0.010693</td>	AZGP1	1	0.060558	ZNF18	1	0.010693
MMD 1 0.057349 KCNK17 1 0.009442 CARS 1 0.05583 PHF20 1 0.009367 ATP6V0A4 1 0.055112 TNP1 1 0.009367 MPER 1 0.054279 ZBTB2 1 0.008477 MPV17L2 1 0.053732 RBM43 1 0.007244 P4HA2 1 0.050163 ELL 1 0.007244 PHLDA2 1 0.050163 ELL 1 0.007581 NAP1L4 1 0.049917 SKIDA1 1 0.00583 NAP1L4 1 0.048975 MLLT10 1 0.00533 ISYNA1 1 0.048024 SPAG4 1 0.00533 ISYNA1 1 0.047283 TAAR8 1 0.03333 MEPCE 1 0.047384 LMAN1L 1 0.02579 GIP 1 0.045344 LYPD5 1 0.002363 CCS	WNT2	1	0.058427	TTLL6	1	0.00984
CARS 1 0.05583 PHF20 1 0.009367 ATP6V0A4 1 0.055112 TNP1 1 0.009195 BMPER 1 0.054279 ZBTB2 1 0.008477 MPV17L2 1 0.053366 PNLIP 1 0.007244 P4HA2 1 0.050681 CASC10 1 0.007176 PHLDA2 1 0.050163 ELL 1 0.00581 NAP1L4 1 0.049917 SKIDA1 1 0.005335 UBE2Z 1 0.048975 MLLT10 1 0.005331 UBE2Z 1 0.048275 ODF3L1 1 0.005333 ISYNA1 1 0.047283 TAAR8 1 0.003731 CYP51A1 1 0.047283 TAAR8 1 0.002579 GIP 1 0.046384 LVPD5 1 0.002363 CCS 1 0.045444 LVPD5 1 0.002363 CCS<	MMD	1	0.057349	KCNK17	1	0.009442
ATP6V0A4 1 0.055112 TNP1 1 0.009195 BMPER 1 0.054279 ZBTB2 1 0.008477 MPV17L2 1 0.053732 RBM43 1 0.007988 RBM48 1 0.053696 PNLIP 1 0.007244 P4HA2 1 0.050163 ELL 1 0.007176 PHLDA2 1 0.050163 ELL 1 0.00598 NAP1L4 1 0.049917 SKIDA1 1 0.00583 UBEZZ 1 0.04827 ODF3L1 1 0.00533 CYP51A1 1 0.047283 TAAR8 1 0.003733 MEPCE 1 0.047158 UBALD2 1 0.00333 MEPCE 1 0.047158 UBALD2 1 0.002579 GIP 1 0.047158 UBALD2 1 0.002579 GIP 1 0.047158 USALD2 1 0.002363 CCS <td>CARS</td> <td>1</td> <td>0.05583</td> <td>PHF20</td> <td>1</td> <td>0.009367</td>	CARS	1	0.05583	PHF20	1	0.009367
BMPER 1 0.054279 ZBTB2 1 0.008477 MPV17L2 1 0.053732 RBM43 1 0.007998 RBM48 1 0.053696 PNLIP 1 0.007244 P4HA2 1 0.050581 CASC10 1 0.007176 PHLDA2 1 0.050163 ELL 1 0.00598 NAP1L4 1 0.049917 SKIDA1 1 0.00583 UBE2Z 1 0.048287 ODF3L1 1 0.00533 CYP51A1 1 0.047263 TAAR8 1 0.003733 CYP51A1 1 0.047283 TAAR8 1 0.00333 MEPCE 1 0.047158 UBALD2 1 0.003033 GIP 1 0.047158 UBALD2 1 0.002579 GIP 1 0.047174 MSRB3 1 0.002563 CCS 1 0.041477 MSRB3 1 0.001837 KCNN4 </td <td>ATP6V0A4</td> <td>1</td> <td>0.055112</td> <td>TNP1</td> <td>1</td> <td>0.009195</td>	ATP6V0A4	1	0.055112	TNP1	1	0.009195
MPV17L2 1 0.053732 RBM43 1 0.007998 RBM48 1 0.053696 PNLIP 1 0.007244 P4HA2 1 0.050581 CASC10 1 0.007176 PHLDA2 1 0.050163 ELL 1 0.00598 NAP1L4 1 0.049917 SKIDA1 1 0.00598 NR2F6 1 0.048975 MLLT10 1 0.00533 UBE2Z 1 0.048024 SPAG4 1 0.00503 CYP51A1 1 0.047283 TAAR8 1 0.003731 CYP2S1 1 0.047158 UBALD2 1 0.002579 GIP 1 0.047344 LYPD5 1 0.002363 CCS 1 0.041477 MSRB3 1 0.001425 NYAP1 1 0.041477 MSRB3 1 0.001425 NYAP1 1 0.040597 BAZ2A 1 0.000831 RBM39 </td <td>BMPER</td> <td>1</td> <td>0.054279</td> <td>ZBTB2</td> <td>1</td> <td>0.008477</td>	BMPER	1	0.054279	ZBTB2	1	0.008477
RBM48 1 0.053696 PNLIP 1 0.007244 P4HA2 1 0.050581 CASC10 1 0.007176 PHLDA2 1 0.050163 ELL 1 0.006518 NAP1L4 1 0.049917 SKIDA1 1 0.00598 NR2F6 1 0.04827 ODF3L1 1 0.005735 UBE2Z 1 0.048287 ODF3L1 1 0.00503 CYP51A1 1 0.047283 TAAR8 1 0.003711 CYP2S1 1 0.046038 LMAN1L 1 0.002579 GIP 1 0.0443749 FGF7 1 0.002363 CCS 1 0.0443749 FGF7 1 0.001425 NYAP1 1 0.041465 SLC2A5 1 0.001425 NYAP1 1 0.041658 ZNF420 1 0.000831 RBM39 1 0.040597 BAZ2A 1 0.000786 BBS1<	MPV17L2	1	0.053732	RBM43	1	0.007998
P4HA2 1 0.050581 CASC10 1 0.007176 PHLDA2 1 0.050163 ELL 1 0.006518 NAP1L4 1 0.049917 SKIDA1 1 0.00598 NR2F6 1 0.048975 MLLT10 1 0.00583 UBE2Z 1 0.048287 ODF3L1 1 0.005735 ISYNA1 1 0.047283 TAAR8 1 0.005731 CYP51A1 1 0.047158 UBALD2 1 0.00333 MEPCE 1 0.047158 UBALD2 1 0.002579 GIP 1 0.045434 LYPD5 1 0.002363 CCS 1 0.043749 FGF7 1 0.001422 GDF15 1 0.041477 MSRB3 1 0.001425 NYAP1 1 0.041088 ZNF420 1 0.000842 SLC2A4 1 0.040597 BAZ2A 1 0.000786 BM39<	RBM48	1	0.053696	PNLIP	1	0.007244
PHLDA2 1 0.050163 ELL 1 0.006518 NAP1L4 1 0.049917 SKIDA1 1 0.00598 NR2F6 1 0.048975 MLLT10 1 0.00583 UBE2Z 1 0.048287 ODF3L1 1 0.005735 ISYNA1 1 0.047283 TAAR8 1 0.003711 CYP51A1 1 0.047158 UBALD2 1 0.00333 MEPCE 1 0.046038 LMAN1L 1 0.002363 GIP 1 0.045434 LYPD5 1 0.002363 CCS 1 0.043749 FGF7 1 0.001942 GDF15 1 0.041477 MSRB3 1 0.001837 KCNN4 1 0.041088 ZNF420 1 0.000842 SLC2A4 1 0.040597 BAZ2A 1 0.000786 RBM39 1 0.040581 RNF175 1 0.000378	P4HA2	1	0.050581	CASC10	1	0.007176
NAP1L4 1 0.049917 SKIDA1 1 0.00598 NR2F6 1 0.048975 MLLT10 1 0.00583 UBE2Z 1 0.048287 ODF3L1 1 0.005735 ISYNA1 1 0.048024 SPAG4 1 0.00503 CYP51A1 1 0.047283 TAAR8 1 0.003711 CYP2S1 1 0.046038 LMAN1L 1 0.002579 GIP 1 0.045434 LYPD5 1 0.002363 CCS 1 0.041477 MSRB3 1 0.001425 GDF15 1 0.041465 SLC22A5 1 0.001425 NYAP1 1 0.040597 BAZ2A 1 0.000831 RBM39 1 0.040581 RNF175 1 0.000786 BBS1 1 0.000378 R 0.000378	PHLDA2	1	0.050163	ELL	1	0.006518
NR2F6 1 0.048975 MLLT10 1 0.00583 UBE2Z 1 0.048287 ODF3L1 1 0.005735 ISYNA1 1 0.048024 SPAG4 1 0.00503 CYP51A1 1 0.047283 TAAR8 1 0.003711 CYP2S1 1 0.047158 UBALD2 1 0.003033 MEPCE 1 0.046038 LMAN1L 1 0.002579 GIP 1 0.043749 FGF7 1 0.001942 GDF15 1 0.041477 MSRB3 1 0.001837 KCNN4 1 0.041465 SLC22A5 1 0.001425 NYAP1 1 0.040597 BAZ2A 1 0.000831 RBM39 1 0.040581 RNF175 1 0.000786 BBS1 1 0.000378 1 0.000378	NAP1L4	1	0.049917	SKIDA1	1	0.00598
UBE2Z 1 0.048287 ODF3L1 1 0.005735 ISYNA1 1 0.048024 SPAG4 1 0.00503 CYP51A1 1 0.047283 TAAR8 1 0.003711 CYP2S1 1 0.047158 UBALD2 1 0.003033 MEPCE 1 0.046038 LMAN1L 1 0.002579 GIP 1 0.043749 FGF7 1 0.001942 GDF15 1 0.041477 MSRB3 1 0.001425 NYAP1 1 0.041088 ZNF420 1 0.000842 SLC2A4 1 0.040597 BAZ2A 1 0.000831 RBM39 1 0.040581 RNF175 1 0.000786 BBS1 1 0.000378 1 0.000378	NR2F6	1	0.048975	MLLT10	1	0.00583
ISYNA1 1 0.048024 SPAG4 1 0.00503 CYP51A1 1 0.047283 TAAR8 1 0.003711 CYP2S1 1 0.047158 UBALD2 1 0.003033 MEPCE 1 0.046038 LMAN1L 1 0.002579 GIP 1 0.043749 FGF7 1 0.001942 GDF15 1 0.041477 MSRB3 1 0.001837 KCNN4 1 0.041465 SLC22A5 1 0.001425 NYAP1 1 0.040597 BAZ2A 1 0.000831 RBM39 1 0.040581 RNF175 1 0.000786 BBS1 1 0.000378 SL 0.000378	UBE2Z	1	0.048287	ODF3L1	1	0.005735
CYP51A1 1 0.047283 TAAR8 1 0.003711 CYP2S1 1 0.047158 UBALD2 1 0.003033 MEPCE 1 0.046038 LMAN1L 1 0.002579 GIP 1 0.045434 LYPD5 1 0.002363 CCS 1 0.043749 FGF7 1 0.001942 GDF15 1 0.041477 MSRB3 1 0.001837 KCNN4 1 0.041465 SLC22A5 1 0.001425 NYAP1 1 0.040597 BAZ2A 1 0.000842 SLC2A4 1 0.040597 BAZ2A 1 0.000786 RBM39 1 0.040581 RNF175 1 0.000378	ISYNA1	1	0.048024	SPAG4	1	0.00503
CYP2S1 1 0.047158 UBALD2 1 0.003033 MEPCE 1 0.046038 LMAN1L 1 0.002579 GIP 1 0.045434 LYPD5 1 0.002363 CCS 1 0.043749 FGF7 1 0.001942 GDF15 1 0.041477 MSRB3 1 0.001425 NYAP1 1 0.041088 ZNF420 1 0.000842 SLC2A4 1 0.040597 BAZ2A 1 0.000831 RBM39 1 0.040581 RNF175 1 0.000366 BBS1 1 0.000378 BBS1 1 0.000378	CYP51A1	1	0.047283	TAAR8	1	0.003711
MEPCE 1 0.046038 LMAN1L 1 0.002579 GIP 1 0.045434 LYPD5 1 0.002363 CCS 1 0.043749 FGF7 1 0.001942 GDF15 1 0.041477 MSRB3 1 0.001837 KCNN4 1 0.041465 SLC22A5 1 0.001425 NYAP1 1 0.041088 ZNF420 1 0.000842 SLC2A4 1 0.040597 BAZ2A 1 0.000786 RBM39 1 0.040581 RNF175 1 0.000378	CYP2S1	1	0.047158	UBALD2	1	0.003033
GIP 1 0.045434 LYPD5 1 0.002363 CCS 1 0.043749 FGF7 1 0.001942 GDF15 1 0.041477 MSRB3 1 0.001837 KCNN4 1 0.041465 SLC22A5 1 0.001425 NYAP1 1 0.041088 ZNF420 1 0.000842 SLC2A4 1 0.040597 BAZ2A 1 0.000786 RBM39 1 0.040581 RNF175 1 0.000378	MEPCE	1	0.046038	LMAN1L	1	0.002579
CCS 1 0.043749 FGF7 1 0.001942 GDF15 1 0.041477 MSRB3 1 0.001837 KCNN4 1 0.041465 SLC22A5 1 0.001425 NYAP1 1 0.041088 ZNF420 1 0.000842 SLC2A4 1 0.040597 BAZ2A 1 0.000786 RBM39 1 0.040581 RNF175 1 0.000378 (Continued)	GIP	1	0.045434	LYPD5	1	0.002363
GDF15 1 0.041477 MSRB3 1 0.001837 KCNN4 1 0.041465 SLC22A5 1 0.001425 NYAP1 1 0.041088 ZNF420 1 0.000842 SLC2A4 1 0.040597 BAZ2A 1 0.000786 RBM39 1 0.040581 RNF175 1 0.000378 (Continued) BBS1 1 0.000378	CCS	1	0.043749	FGF7	1	0.001942
KCNN4 1 0.041465 SLC22A5 1 0.001425 NYAP1 1 0.041088 ZNF420 1 0.000842 SLC2A4 1 0.040597 BAZ2A 1 0.000831 RBM39 1 0.040581 RNF175 1 0.000786 Continued) BBS1 1 0.000378	GDF15	1	0.041477	MSRB3	1	0.001837
NYAP1 1 0.041088 ZNF420 1 0.000842 SLC2A4 1 0.040597 BAZ2A 1 0.000831 RBM39 1 0.040581 RNF175 1 0.000786 Continued) BBS1 1 0.000378	KCNN4	1	0.041465	SLC22A5	1	0.001425
SLC2A4 1 0.040597 BAZ2A 1 0.000831 RBM39 1 0.040581 RNF175 1 0.000786 (Continued) BBS1 1 0.000378	NYAP1	1	0.041088	ZNF420	1	0.000842
RBM39 1 0.040581 RNF175 1 0.000786 (Continued) BBS1 1 0.000378	SLC2A4	1	0.040597	BAZ2A	1	0.000831
(Continued) <u>BBS1</u> 1 0.000378	RBM39	1	0.040581	RNF175	1	0.000786
			(Continued)	BBS1	1	0.000378



Figure 2. Heatmap of significant different expression genes in BC. a)Top 20 Over-expressed genes; b) Top 20 Under-expressed genes.

and *TNRC6B*) in breast cancer (Figure 1c). Among them, *CLDN7*, *MLLT10*, *RBM33*, *SH3RF1*, *SSBP4*, and *UBE2Z* were overexpressed; and *BMPER*, *FGF7*, *MSRB3*, and *TNRC6B* were underexpressed.

3.2 GO and KEGG enrichment analyses of significant DEGs in breast cancer

To explore the roles of the significant DEGs in breast cancer, we played GO and KEGG

enrichment analyses. BP analysis revealed that the significant genes in breast cancer were mainly enriched in the Wnt signaling pathway, calcium-modulating pathway, protein repair, gene silencing by microRNA (miRNA), mRNA cleavage involved in gene silencing by miRNA, and positive regulation of epithelial cell proliferation involved in lung morphogenesis (Table 2). MF analysis showed that significant genes were enriched in functions related to oxidoreductase activity, acting on a sulfur group of donors and disulfide as acceptor, and phosphoinositide 3-kinase (PI3K) and PIK3CA activities (Table 2). CC analysis showed that significant genes were enriched in P-bodies. KEGG analysis revealed that significant genes in breast cancer were mainly involved in pathways in cancer, breast cancer, gastric cancer, melanoma, the PI3K/Akt signaling pathway, mitogen-activated protein kinase (MAPK) signaling pathway, Ras signaling pathway, tight junctions, and ubiquitin-mediated proteolysis (Figure 3).

3.3 Correlation between significant DEGs and breast cancer

To verify the significant DEGs of breast cancer, we further explore the DEGs. Consistent with the identification of significant genes, protein profiling in breast cancer samples from the HPA using immunohistochemistry showed that the gene expression of *CLDN7*, *RBM33*, *SH3RF1*, and *UBE2Z* was significantly enriched in breast cancer, whereas there was no significant enrichment of *FGF7* and *TNRC6B* (Figure 4).

The significant DEGs (*CLDN7*, *BMPER*, *FGF7*, *MSRB3*) in breast cancer samples compared to normal samples also showed coincident results of significant gene identification (Figure 5).

3.4 Correlation between overall survival and significant DEGs in breast cancer

In the analysis of the correlation between overall survival and significant DEG expression (*CLDN7*, *MLLT10*, *RBM33*, *SH3RF1*, *SSBP4*, and *UBE2Z*) in breast cancer, we found a shorter survival time based on GSE5881 (Figure 6a) and GSE42568 (Figure 6b) (P< 0.05). The significantly

underexpressed genes (BMPER, FGF7, MSRB3, and TNRC6B) in breast cancer were correlated with a longer survival time based on GSE42568 (Figure 6c) and GSE37751 (Figure 6d) (P> 0.05), this correlation cannot be demonstrated. The overall survival analysis combined significantly underexpressed genes (BMPER, MSRB3, and based TNRC6B) in breast cancer on GSE1456_U133B (Figure 6e) and GSE3494_U133B (Figure 6f); combined FGF7 and TNRC6B based on GSE3494 U133A (Figure 6g); and combined single FGF7 based on GSE9893 (Figure 6h) were also correlated with longer survival in breast cancer (P < 0.05).

3.5 Drug prediction analysis

To identify potential small-molecule compounds with therapeutic effects on breast cancer, drug prediction analysis was performed using CMap. A total of 65 drugs were predicted, and the 10 most significant were trichostatin A, LY-294,002, econazole, Prestwick-1082, vorinostat, lomefloxacin, clorsulon, amantadine, thiostrepton, and orciprenaline (Table 3).

4. Discussion

Breast cancer is the most commonly diagnosed malignancy in women worldwide and is the main cause of cancer-related death in women [35–37]. Although there are a lot of effective therapeutic agents for breast cancer, breast cancer remains a major health problem and is a top biomedical research priority [38–40], as there is an urgent need for effective breast cancer treatments.

In this study, we identified 10 significant genes (CLDN7, MLLT10, RBM33, SH3RF1, SSBP4, UBE2Z, BMPER, FGF7, MSRB3, and TNRC6B) in breast cancer using combined GWAS data and profiling of DEGs. Protein profiling in breast cancer samples from the HPA using immunohistochemistry and analysis of significant DEGs in breast cancer samples compared to normal samples from GEPIA further verified the results. Significantly overexpressed (CLDN7, genes MLLT10, RBM33, SH3RF1, SSBP4, and UBE2Z) were correlated with shorter survival, whereas underexpressed genes (BMPER, FGF7, MSRB3,

Table	2.	Gene	ontology	(GO)	enrichment	result	of	significant
genes	in	breast	cancer.					

	Biological processes	
		false
		discovery
term ID	term description	rate
GO:0007223	Wnt signaling pathway, calcium modulating pathway	0.0024
GO:0030091	protein repair	0.0024
GO:0035195	gene silencing by miRNA	0.0024
GO:0035279	mRNA cleavage involved in gene silencing by miRNA	0.0024
GO:0060501	positive regulation of epithelial cell proliferation involved in lung morphogenesis	0.0024
GO:0035278	miRNA mediated inhibition of translation	0.003
GO:0060213	positive regulation of nuclear-transcribed mRNA poly(A) tail shortening	0.003
GO:0010463	mesenchymal cell proliferation	0.0039
GO:0060445	branching involved in salivary gland morphogenesis	0.0039
GO:0031069	hair follicle morphogenesis	0.0078
GO:2,000,026	regulation of multicellular organismal development	0.0082
GO:0051173	positive regulation of nitrogen compound metabolic process	0.0168
GO:0036092	phosphatidylinositol-3-phosphate biosynthetic process	0.019
GO:0010604	positive regulation of macromolecule metabolic process	0.0194
GO:0031325	positive regulation of cellular metabolic process	0.0194
GO:0048522	positive regulation of cellular process	0.0194
GO:0051254	positive regulation of RNA metabolic process	0.0194
GO:0060688	regulation of morphogenesis of a branching structure	0.0194
GO:0007267	cell-cell signaling	0.0199
GO:1,903,313	positive regulation of mRNA metabolic process	0.0242
Molecular funct	tions	
GO:0016671	oxidoreductase activity, acting on a sulfur group of donors, disulfide as acceptor	0.0083
GO:0016303	1-phosphatidylinositol-3-kinase activity	0.0406
GO:0046934	phosphatidylinositol-4,5-bisphosphate 3-kinase activity	0.0406
Cellular compo	nents	
GO:0000932	P-body	0.0029

and *TNRC6B*) were correlated with longer survival in breast cancer.

Consistent with our findings, previous studies have shown that some of these genes play important roles in the development of breast cancer. For example, Bernardi *et al.* [41] showed that *CLDN7* is associated with a shorter time to recurrence, suggesting its contribution to the aggressiveness of breast cancer. In a GWAS, Guo *et al.* [42] identified common genetic loci for breast cancer risk including SSBP4. Whole transcriptome

analysis by Bauer et al. [43] demonstrated that BMPER plays a possible therapeutic role in breast cancer. Fu et al. [44] demonstrated that acetylation, expression and recruitment of FGF7 promoters induce cancer growth and progression. Zhu et al. [45] found that targeting FGF7 can exert oncogenic functions in breast cancer. A previous study showed that the ZEB1-MSRB3 axis is related breast cancer genome stability to [46]. Interestingly, other DEGs in breast cancer identified in this study, including MLLT10, RBM33, SH3RF1, UBE2Z, and TNRC6B, have not been proven in previous studies. We believe that these are potentially novel key genes in breast cancer.

BP analysis in GO annotation indicated that the 10 significant genes are mainly enriched in the Wnt signaling pathway, which plays an important role in the occurrence and development of many cancers. Inhibiting this pathway can suppress breast cancer growth and metastasis [47–49]. MF analysis of GO suggested that the DEGs were most significantly enriched in functions related to oxidoreductase activity. The redox reaction is accompanied by tumor development. CC analysis of GO annotation showed that the 10 DEGs were enriched in P-bodies. A previous study suggested that P-body disassembly correlates with breast cancer progression [50].

KEGG analysis of the 10 DEGs showed their enrichment in breast cancer, gastric cancer, melanoma, the PI3K/Akt signaling pathway, MAPK signaling pathway, Ras signaling pathway, tight junctions, and ubiquitin-mediated proteolysis. Some of these pathways contribute to the development of breast cancer. For example, the PI3K pathway is found in many types of cancer and plays an important role in breast cancer cell proliferation [51]. Ras signaling is a key determinant of poor survival in breast cancer patients [52]. Abnormal MAPK signaling plays a core role in the regulation of growth and survival, and the development of drug resistance in triplenegative breast cancer [53].

The aim of this work was to identify significant genes and potential therapeutic agents for breast cancer based on genomics. We found 65 potentially small-molecule compounds to reverse significant genes in breast cancer. The 10 most significant drugs were trichostatin A, LY-



KEGG pathway of significant genes in BC

Figure 3 KEGG pathway analysis of significant genes (CLDN7, MLLT10, RBM33, SH3RF1, SSBP4, UBE2Z, BMPER, FGF7, MSRB3 and TNRC6B) in BC.

294,002, econazole, Prestwick-1082, vorinostat, lomefloxacin, clorsulon, amantadine, thiostrepton, and orciprenaline. Consistent, with our study, it has been reported that trichostatin A, a histone deacetylase inhibitor, has therapeutic potential in breast cancer [54]. Jiang et al. [55] showed that trichostatin A sensitizes ER-negative breast cancer cells to tamoxifen. LY294002, a specific inhibitor of the PI3K pathway, can decrease the rate of cell growth and increase the therapeutic sensitivity in MCF7 cells expressing wild-type p53, which may be useful for the treatment of breast cancer [56]. Econazole is a novel PI3K/AKT signaling pathway inhibitor, which can be used to overcome adriamycin resistance and improve chemotherapy sensitivity in breast cancer [57]. A preclinical study showed that vorinostat can prevent the formation of brain metastases in breast cancer [58]. Yang et al. [59] suggested that thiostrepton is a promising agent for triple-negative breast cancer. Kwok et al. [60] showed that thiostrepton selectively targets breast cancer cells through inhibition of Forkhead box M1 expression. However, some of the predicted drugs, such as Prestwick-1082, lomefloxacin, clorsulon, amantadine, and orciprenaline, have not been shown to directly play a role in breast cancer. Thus, future studies are needed to confirm our findings.

Compared to previous studies [61–63], we conducted an analysis combining genomic data with drug database analysis to identify novel candidate therapeutic agents for breast cancer treatment. Our study demonstrates the usefulness of this approach for evaluating the relationship among genes, diseases, and drugs. These findings will pave the way for the discovery of potential therapeutic targets for breast cancer.

5. Conclusion:

Combined analyses of network-wide association studies, gene expression profiles, and drug databases are helpful for identifying potential therapeutic agents for diseases. This method is a new paradigm that can guide future research directions.





FGF7





Figure 4. The immunohistochemistry of significant genes(CLDN7, RBM33, SH3RF1, UBE2Z, FGF7 and TNRC6B) in BC.



Figure 5. The different expression of significant genes genes (CLDN7, BMPER, FGF7 and MSRB3) in breast cancer samples to normal samples, the red box mean in breast cancer samples and the black box mean in normal samples. *:p < 0.05.

Acknowledgements

The authors acknowledge the National Natural Science Foundation of China 81972861, Innovation Capacity Support Plan of Shaanxi Province 2018TD-002. We thank LetPub (www.letpub.com) for its linguistic assistance during the preparation of this manuscript.



Figure 6. Overall survival analysis combined multiple genes expression. a)Combined significant Over-expressed genes(CLDN7, MLLT10, RBM33, SH3RF1, SSBP4 and UBE2Z) in BC based on GSE58812; b)Combined significant Over-expressed genes(CLDN7, MLLT10, RBM33, SH3RF1, SSBP4 and UBE2Z) in BC based on GSE42568; c)Combined significant under-expressed genes(BMPER, FGF7, MSRB3 and TNRC6B) in BC based on GSE42568; d)Combined BMPER,FGF7,MSRB3 and TNRC6B in BC based on GSE1456_U133B; f)Combined BMPER,MSRB3 and TNRC6B in BC based on GSE1456_U133B; f)Combined BMPER,MSRB3 and TNRC6B in BC based on GSE3494_U133B; g)Combined FGF7 and TNRC6B in BC based on GSE3494_U133A; h)Combined FGF7 in BC based on GSE9893.

Table 3. The most significant drugs provided by cmap to reverse core genes of breast cancers.

		anrichment	- j	co o cifi citu
	Illedii	ennchment	h	specificity
trichostatin A	-0.44/	-0.422	0	0.3069
LY-294,002	-0.389	-0.402	0	0.1534
econazole	0.716	0.91	0.00006	0.0154
Prestwick-1082	0.749	0.949	0.00012	0
vorinostat	-0.508	-0.568	0.00038	0.2832
lomefloxacin	-0.60/	-0./35	0.0007	0
clorsulon	0./14	0.85	0.000/2	0
amantadine	0.659	0.838	0.00107	0.0063
thiostrepton	-0.738	-0.831	0.00147	0.037
orciprenaline	0.589	0.801	0.00298	0.0058
thiamphenicol	0.428	0.725	0.00352	0.1173
khellin	0.444	0.713	0.00459	0.0181
thiethylperazine	0.576	0.767	0.00567	0.011
felbinac	0.476	0.763	0.00599	0.0468
Chicago Sky Blue 6B	0.472	0.762	0.00605	0.0061
vinburnine	0.41	0.756	0.0068	0.0351
scriptaid	-0.688	-0.849	0.00683	0.0833
Prestwick-1103	0.338	0.747	0.00786	0.0435
naringenin	0.444	0.745	0.00796	0.0585
adiphenine	0.385	0.683	0.00811	0.2819
terazosin	0.561	0.74	0.00875	0.0112
oxolamine	0.339	0.739	0.00897	0.0385
monobenzone	-0.557	-0.736	0.00961	0.0203
chenodeoxycholic acid	0.326	0.731	0.01034	0.0495
rifabutin	-0.623	-0.823	0.01104	0.125
levonorgestrel	-0.38	-0.61	0.01158	0.0874
cinnarizine	0.413	0.72	0.01233	0.0146
oxybuprocaine	-0.563	-0.715	0.01339	0.0225
metformin	0.335	0.475	0.01352	0.0311
memantine	0.374	0.7	0.01677	0
acetohexamide	0.292	0.698	0.01719	0.0059
proxyphylline	-0.515	-0.694	0.01846	0.0792
R-atenolol	0.273	0.688	0.01969	0
quinostatin	-0.756	-0.9	0.02	0.1832
vinblastine	0.41	0.781	0.02129	0.0774
colecalciferol	-0.407	-0.684	0.02158	0.0217
dexibuprofen	-0.501	-0.681	0.02248	0.0173
BCB000040	0.616	0.679	0.02308	0.0057
levopropoxyphene	0.28	0.676	0.02401	0.0063
sulconazole	-0.396	-0.676	0.02455	0.0855
nadolol	0.550	0.673	0.02133	0.0055
hycanthone	0.510	0.67	0.02514	0.1814
CP-863 187	0.618	0.67	0.02012	0.0704
karakoline	0.010	0.557	0.02707	0.0767
otamiyan	0.517	0.557	0.02034	0.0207
methylergometrine	-0.407	-0.662	0.0292	0.0208
homochlorcyclizing	-0.407	-0.002	0.02952	0.0337
fluoromotholono	0.321	0.00	0.03009	0.1007
nuorometrioione	-0.401	-0.050	0.03223	0.057
producoi	-0.382	-0.549	0.03200	0.0001
flosoinido	-0.406	-0.591	0.03308	0.2864
	0.35	0.547	0.03329	0.0058
tetraetnylenepentamine	-0.384	-0.542	0.03655	0.0229
Josamycin	0.469	0.588	0.03679	0.1235
torasemide	-0.418	-0.645	0.03/96	0.0545
nadide	0.511	0.637	0.04225	0.0643
TOSTOSAI	0.459	0.637	0.04265	0.0142
rescinnamine	-0.592	-0.723	0.04363	0.0248
diethylcarbamazine	0.396	0.634	0.04416	0.0252
cyclopentolate	-0.499	-0.632	0.0444	0.05
harmol	0.275	0.634	0.04444	0.0853
metixene	0.288	0.632	0.04534	0.1593
idoxuridine	0.474	0.572	0.0458	0.0838
LM-1685	-0.571	-0.714	0.04795	0.0238
iohexol	0.253	0.628	0.04808	0.1657
fenbendazole	0.497	0.626	0.0489	0.11

Funding

This work was supported by the National Science Foundation of China [81972861]; The scientific development funding of the first affiliated hospital of Xi'an Jiaotong University (2020QN-07); The fundamental research funds for the central universities (xzy012020044).

Availability of data and materials

All materials are available by the corresponding author.

Declarations

Ethical statement

Our study did not require an ethical board approval because it did not contain human or animal trials.

Consent for publication

Not applicable.

Disclosure Statement

The authors declare no competing interests.

Highlights

- (1) Combined analyses of network-wide association studies, gene expression profiles, and drug databases.
- (2) A useful approach for evaluating the relationship among genes, diseases, and drugs.

These findings will pave the way for the discovery of potential therapeutic targets for breast cancer.

ORCID

Liuyun Gong (http://orcid.org/0000-0002-6795-8235

References

- Nyblade L, Stockton M, Travasso S, et al. A qualitative exploration of cervical and breast cancer stigma in Karnataka, India. BMC Womens Health. 2017;17(1):58.
- [2] Cancer Genome Atlas Network. Comprehensive molecular portraits of human breast tumours. Nature.2012;490(7418):61–70.

- [3] Hess KR, Pusztai L, Buzdar AU, et al. Estrogen receptors and distinct patterns of breast cancer relapse. Breast Cancer Res Treat. 2003;78(1):105–18.
- [4] Andre F, Job B, Dessen P, et al. Molecular characterization of breast cancer with high-resolution oligonucleotide comparative genomic hybridization array. Clin Cancer Res. 2009;15(2):441–51.
- [5] Cocco S, Piezzo M, Calabrese A, et al. Biomarkers in triple-negative breast cancer: state-of-the-art and future perspectives. Int J Mol Sci. 2020;21(13):13.
- [6] Piezzo M, Cocco S, Caputo R, et al. Targeting cell cycle in breast cancer: CDK4/6 INHIBITORS. Int J Mol Sci. 2020;21(18):18.
- [7] Fu CH, Lin R-J, Yu J, et al. A novel oncogenic role of inositol phosphatase SHIP2 in ER-negative breast cancer stem cells: involvement of JNK/vimentin activation. Stem Cells. 2014;32(8):2048–60.
- [8] Sirinian C, Papanastasiou AD, Schizas M, et al. RANKc attenuates aggressive properties of ER-negative breast cancer by inhibiting NF-kappaB activation and EGFR signaling. Oncogene. 2018;37(37):5101–5114.
- [9] Gao S, Ge A, Xu S, et al. PSAT1 is regulated by ATF4 and enhances cell proliferation via the GSK3beta/beta-catenin/cyclin D1 signaling pathway in ER-negative breast cancer. J Exp Clin Cancer Res. 2017;36(1):179.
- [10] Nb S, Cd, M, S, B, et al. An azaspirane derivative suppresses growth and induces apoptosis of ER-positive and ER-negative breast cancer cells through the modulation of JAK2/STAT3 signaling pathway. Int J Oncol. 2016;49(3):1221–9.
- [11] Sp H, Py, L, Cj, K, et al. The Gah-PLCô1 signaling axis drives metastatic progression in triple-negative breast cancer. J Hematol Oncol. 2017;10(1):114.
- [12] Cj H, Yl H, Yf H, et al. Molecular mechanisms of anticancer effects of phytoestrogens in breast cancer. Breast cancer research and treatment. 2018;19(3):323–332.
- [13] Espinal AC, Buas MF, Wang D, et al. FOXA1 hypermethylation: link between parity and ER-negative breast cancer in African American women? Breast Cancer Res Treat. 2017;166(2):559–568.
- [14] Hirschhorn JN, Daly MJ. Genome-wide association studies for common diseases and complex traits. Nat Rev Genet. 2005;6(2):95–108.
- [15] Greene CS, Krishnan A, Wong AK, et al. Understanding multicellular function and disease with human tissue-specific networks. Nat Genet. 2015;47(6):569–76.
- [16] Wang J, Zheng J, Wang Z, et al. Inferring gene-disease association by an integrative analysis of eQTL genome-wide association study and protein-protein interaction data. Hum Hered. 2018;83(3):117–129.
- [17] Song A, Yan J, Kim S, et al. Network-based analysis of genetic variants associated with hippocampal volume in Alzheimer's disease: a study of ADNI cohorts. BioData Min. 2016;9(1):3.

- [18] Ma C, Gu C, Huo Y, et al. The integrated landscape of causal genes and pathways in schizophrenia. Transl Psychiatry. 2018;8(1):67.
- [19] Sudlow C, Gallacher J, Allen N, et al. UK biobank: an open access resource for identifying the causes of a wide range of complex diseases of middle and old age. PLoS Med. 2015;12(3):e1001779.
- [20] Hammerschlag AR, Stringer S. Genome-wide association analysis of insomnia complaints identifies risk genes and genetic overlap with psychiatric and metabolic traits. Nat Genet. 2017;49(11):1584–1592.
- [21] Canela-Xandri O, Rawlik K, Tenesa A. An atlas of genetic associations in UK Biobank. Nat Genet. 2018;50(11):1593–1599.
- [22] Otasek D, Morris JH, Bouças J, et al. Cytoscape Automation: empowering workflow-based network analysis. Genome Biol. 2019;20(1):185.
- [23] Bindea G, Galon J, Mlecnik B. CluePedia Cytoscape plugin: pathway insights using integrated experimental and in silico data. Bioinformatics. 2013;29(5):661–3.
- [24] Szklarczyk D, Franceschini A, Wyder S, et al. STRING v10: protein-protein interaction networks, integrated over the tree of life. Nucleic Acids Res. 2015;43(D1): D447-52. Database issue.
- [25] Szklarczyk D. Morris J H, Cook H, et al. The STRING database in 2017: quality-controlled protein-protein association networks. *made broadly accessible*. Nucleic Acids Res. 2017;45(D1):D362–d368.
- [26] Y, Z, Y Z, Y F, et al. Identification of biomarkers, pathways and potential therapeutic agents for white adipocyte insulin resistance using bioinformatics analysis. Adipocyte. 2019;8(1):318–329.
- [27] Z, T, C L, B K, et al. GEPIA: a web server for cancer and normal gene expression profiling and interactive analyses. Nucleic Acids Res. 2017;45(null):W98–W102.
- [28] Ponten F, Jirstrom K, Uhlen M. The human protein atlas-a tool for pathology. J Pathol. 2008;216 (4):387-93.
- [29] Uhlen M, Zhang C, Lee S, et al. A pathology atlas of the human cancer transcriptome. Science. 2017;357 (6352):6352.
- [30] Thul PJ, Åkesson L, Wiking M, et al. A subcellular map of the human proteome. Science. 2017;356 (6340):6340.
- [31] Goswami CP, Nakshatri H. PROGgeneV2: enhancements on the existing database. BMC Cancer. 2014;14 (1):970.
- [32] Lamb J, et al. The connectivity map: using gene-expression signatures to connect small molecules, genes, and disease. Science. 2006;313(5795):1929–35.
- [33] Musa A, Ghoraie LS, Zhang S-D, et al. A review of connectivity map and computational approaches in pharmacogenomics. Brief Bioinform. 2018;19 (3):506-523.
- [34] Raghavan R, Hyter S, Pathak HB, et al. Drug discovery using clinical outcome-based connectivity mapping:

application to ovarian cancer. BMC Genomics. 2016;17(1):811.

- [35] North CM, Christiani DC. Women and lung cancer: what is new? Semin Thorac Cardiovasc Surg. 2013;25 (2):87–94.
- [36] Woolston C. Breast cancer. Nature. 2015;527(7578): S101.
- [37] Tsikouras P, Zervoudis S, Manav B, et al. Cervical cancer: screening, diagnosis and staging. J Buon. 2016;21(2):320–5.
- [38] Anastasiadi Z, Lianos GD, Ignatiadou E, et al. Breast cancer in young women: an overview. Updates Surg. 2017;69(3):313–317.
- [39] Musselwhite LW, Oliveira CM, Kwaramba T, et al. Racial/ethnic disparities in cervical cancer screening and outcomes. Acta Cytol. 2016;60(6):518–526.
- [40] Kossai M, Leary A, Scoazec J-Y, et al. Ovarian Cancer:
 a heterogeneous disease. Pathobiology. 2018;85 (1-2):41-49.
- [41] Bernardi M, Logullo A, Pasini F, et al. Prognostic significance of CD24 and claudin-7 immunoexpression in ductal invasive breast cancer. Oncol Rep. 2012;27 (1):28–38.
- [42] Guo X, Lin W, Bao J, et al. A Comprehensive cis-eQTL analysis revealed target genes in breast cancer susceptibility loci identified in genome-wide association studies. Am J Hum Genet. 2018;102(5):890–903.
- [43] Bauer D, Mazzio E, Soliman KFA. Whole transcriptomic analysis of apigenin on TNFα immuno-activated MDA-MB-231 breast cancer cells. Cancer Genomics Proteomics. 2019;16(6):421–431.
- [44] Fu R, Han C, Ni T, et al. A ZEB1/p53 signaling axis in stromal fibroblasts promotes mammary epithelial tumours. Nat Commun. 2019;10(1):3210.
- [45] Zhu Y, Yang L, Chong Q-Y, et al. Long noncoding RNA Linc00460 promotes breast cancer progression by regulating the miR-489-5p/FGF7/AKT axis. Cancer Management and Research. 2019;11(p):5983–6001.
- [46] Morel AP, Ginestier C, Pommier RM, et al. A stemness-related ZEB1-MSRB3 axis governs cellular pliancy and breast cancer genome stability. Nat Med. 2017;23(5):568–578.
- [47] Lin Q, He Y, Wang X, et al. Targeting pyruvate carboxylase by a small molecule suppresses breast cancer progression. Adv Sci (Weinh). 2020;7(9):1903483.
- [48] Khawaled S, Nigita G, Distefano R, et al. Pleiotropic tumor suppressor functions of WWOX antagonize metastasis. Signal Transduction and Targeted Therapy. 2020;5(1):43.
- [49] Eyre R, Alférez DG, Santiago-Gómez A, et al. Microenvironmental IL1β promotes breast cancer metastatic colonisation in the bone via activation of Wnt signalling. Nature Communications. 2019;10 (1):5016.
- [50] Cougot N, Daguenet É, Baguet A, et al. Overexpression of MLN51 triggers P-body disassembly and formation

of a new type of RNA granules. J Cell Sci. 2014;127 (21):4692-701. .

- [51] A, K, V LN, N N, et al. Benzo[b]furan derivatives induces apoptosis by targeting the PI3K/Akt/mTOR signaling pathway in human breast cancer cells. chemistry SRNJB. 2016;66(undefined):124–131.
- [52] KI W, JR A, JC L, et al. Ras signaling is a key determinant for metastatic dissemination and poor survival of luminal breast cancer patients. Cancer Res. 2015;75 (22):4960–72.
- [53] W, C, J L, Q H, et al. AMP-activated protein kinase: a potential therapeutic target for triple-negative breast cancer. Breast Cancer Res. 2019;21(1):29.
- [54] Alao J, Stavropoulou A, Lam E, et al. Histone deacetylase inhibitor, trichostatin A induces ubiquitin-dependent cyclin D1 degradation in MCF-7 breast cancer cells. Mol Cancer. 2006;5(p):8.
- [55] Jang E, Lim S, Lee E, et al. The histone deacetylase inhibitor trichostatin A sensitizes estrogen receptor alpha-negative breast cancer cells to tamoxifen. Oncogene. 2004;23(9):1724–36.
- [56] Zheng L, Ren J, Li H, et al. Downregulation of wild-type p53 protein by HER-2/neu mediated PI3K pathway activation in human breast cancer cells: its effect on cell proliferation and implication for therapy. Cell Res. 2004;14(6):497–506.
- [57] Dong C, Chen Y, Ma J, et al. Econazole nitrate reversed the resistance of breast cancer cells to Adriamycin through inhibiting the PI3K/AKT signaling pathway. Am J Cancer Res. 2020;10(1):263–274.

- [58] Palmieri D, Lockman P, Thomas F, et al. Vorinostat inhibits brain metastatic colonization in a model of triple-negative breast cancer and induces DNA double-strand breaks. Clinical cancer res. 2009;15 (19):6148–57.
- [59] Yang N, Zhou T, Lei X, et al. Inhibition of sonic hedgehog signaling pathway by thiazole antibiotic thiostrepton attenuates the CD44+/CD24-stem-like population and sphere-forming capacity in triple-negative breast cancer. Cell Physiol Biochem. 2016;38(3):1157–70.
- [60] Kwok J, Myatt S, Marson C, et al. Thiostrepton selectively targets breast cancer cells through inhibition of forkhead box M1 expression. Mol Cancer Ther. 2008;7 (7):2022–32.
- [61] Xu X, Zhang J, Zhang Z, et al. Systems pharmacology in combination with proteomics reveals underlying mechanisms of Xihuang pill against triple-negative breast cancer. Bioengineered. 2020;11 (1):1170–1188.
- [62] Ye Q, Han X, Wu Z. Bioinformatics analysis to screen key prognostic genes in the breast cancer tumor microenvironment. Bioengineered. 2020;11 (1):1280–1300.
- [63] Yan S, Xu J, Liu B, et al. Long non-coding RNA BCAR4 aggravated proliferation and migration in esophageal squamous cell carcinoma by negatively regulating p53/p21 signaling pathway. Bioengineered. 2021;12(1):682-696.