Research Square

Preprints are preliminary reports that have not undergone peer review. They should not be considered conclusive, used to inform clinical practice, or referenced by the media as validated information.

Prognostic Significance of SQSTM1 in Breast Cancer: A Comprehensive Analysis

Yang Liu
Department of Pharmacy, the First Affiliated Hospital of Xi'an Jiaotong University, Xi'an, Shaanxi,China

Qian Du
Department of Pharmacy, the First Affiliated Hospital of Xi'an Jiaotong University, Xi'an, Shaanxi 710061, PR China

Dan Sun
Department of Pharmacy, the First Affiliated Hospital of Xi'an Jiaotong University, Xi'an, Shaanxi 710061, PR China

Ruiying Han
Department of Pharmacy, the First Affiliated Hospital of Xi'an Jiaotong University, Xi'an, Shaanxi 710061, PR China

Mengmeng Teng
Department of Pharmacy, the First Affiliated Hospital of Xi'an Jiaotong University, Xi'an, Shaanxi 710061, PR China

Ying Zhang
Department of Pharmacy, the First Affiliated Hospital of Xi'an Jiaotong University, Xi'an, Shaanxi 710061, PR China

Yuzhu Dong
Department of Pharmacy, the First Affiliated Hospital of Xi'an Jiaotong University, Xi'an, Shaanxi 710061, PR China

Siying Chen
Department of Pharmacy, the First Affiliated Hospital of Xi'an Jiaotong University, Xi'an, Shaanxi 710061, PR China

Haisheng You
Department of Pharmacy, the First Affiliated Hospital of Xi'an Jiaotong University, Xi'an, Shaanxi 710061, PR China

Yalin Dong (dongyalin@mail.xjtu.edu.cn)
Department of Pharmacy, the First Affiliated Hospital of Xi'an Jiaotong University, Xi'an, Shaanxi 710061, PR China https://orcid.org/0000-0001-5619-0788

Research

Keywords: SQSTM1, breast cancer, prognosis, expression profiling, bioinformatics

Posted Date: June 30th, 2020

DOI: https://doi.org/10.21203/rs.3.rs-37254/v1

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

Abstract

Background: *SQSTM1* (Sequestosome 1, p62) is degraded by activated autophagy and involved in the progression of in various types of cancers. However, the prognostic role and underlying regulation mechanism of *SQSTM1* in the progression and development of breast cancer remain unclear.

Methods: In this study, 1336 samples with available mRNA data from Molecular Taxonomy of Breast Cancer International Consortium (METABRIC) database and 27 formalin fixation and paraffin embedding (FFPE) tissue samples from the First Affiliated Hospital of Xi'an Jiaotong University were collected to evaluate *SQSTM1* expression in mRNA and protein levels. Kaplan–Meier and Cox regression were used for revealing prognostic value in three independent breast cancer independent datasets. Tumor Immune Estimation Resource (TIMER) database and Gene Set Variation Analysis (GSVA) was used to explore the relationship of *SQSTM1* mRNA expression and immune infiltration level in breast cancer. Dysregulation mechanisms of *SQSTM1* were also explored including copy number variation (CNV), somatic mutation, epigenetic alterations and other transcription and post-transcription level using multiple datasets. Finally, Gene Set Enrichment Analysis (GSEA) was constructed to elucidate functional regulating performance of *SQSTM1* in breast cancer.

Results: The results showed that mRNA and protein level of *SQSTM1* were significantly elevated in breast cancer and receiver operating characteristic (ROC) curve showed that p62 may act as diagnostic biomarker. Lower expression of *SQSTM1* predicted better outcome through multiple datasets. It was also found that *SQSTM1* correlated with immune infiltrates in breast cancer. Moreover, CNV and methylation of *SQSTM1* DNA was correlated with *SQSTM1* dysregulation and act as prognostic factors for breast cancer patients. Yet, somatic mutation status of *SQSTM1* didn't show any prognostic relevance. We also identified diverse transcription factors that directly bound to *SQSTM1* DNA and the miRNAs which may regulate *SQSTM1* mRNA. Finally, functional enrichment analysis revealed that *SQSTM1* is related to cell signal transduction, oxidative stress and autophagy in breast cancer.

Conclusion: Our findings revealed that overexpression of *SQSTM1* significantly to poor survival and immune infiltrations in breast cancer. In addition, *SQSTM1* plays a key role in the progression of breast cancer and might be a promising biomarker for the diagnosis and personalized treatment of breast cancer patients.

Introduction

Breast cancer is the one of the most frequently diagnosed malignant cancers and the leading cause of cancer-related deaths among women worldwide (1). The Global Cancer Statistics 2018 reported that 2,088,849 new cases and 626,679 deaths of breast cancer occurred globally in 2018 (2).The combination of surgery, chemotherapy and other strategies have made remarkable progress during the past few years. However, the clinical outcome of breast cancer patients still remains poor due to lack of reliable tumor biomarkers and personalized therapies. It is widely known that there is high heterogeneity in breast

cancer with different clinical, histological, and prognostic characteristics (3). Therefore, exploring effective diagnostic and therapeutic biomarkers to help stratify patients and optimize appropriate therapy strategies is significantly urgent.

SQSTM1, also called p62, has been reported as an adaptor protein involved in autophagy and played as a central hub in various signal pathway and regulate multiple effectors, such as NF-kappaB and mTOR (4, 5). Dysregulation of *SQSTM1* is considered to be involved in various types of cancers, including hepatocellular carcinoma, lung adenocarcinoma, breast cancer, colon cancer(6–10). In human liver tissues, high expression of *SQSTM1* correlates with rapid recurrence of resectable hepatocellular carcinoma which demonstrated its oncogenic role (11). According to a meta-analysis, high expression of *SQSTM1* was associated with poor overall survival in lung cancer and might be useful to predict prognosis of lung cancer (12). On the other hand, previous study indicated that *SQSTM1* enhanced breast cancer stem-like properties and promoted breast cancer metastasis promoter by binding vimentin (13, 14). However, other researchers also revealed that *SQSTM1* expression didn't show significant difference between breast cancer tissues and healthy adjacent tissues (15). These findings suggest the *SQSTM1* hold the promise as a novel biological and therapeutic marker and also needed further studies to elucidate its detailed role in human cancers.

The aim of our study is to shed light on the impact of *SQSTM1* in the development and prognostic outcome of breast cancer. We evaluated *SQSTM1* expression in mRNA and protein levels. It was found that *SQSTM1* was upregulated in breast cancer tissues and indicated a poor prognosis in patients with breast cancer. Based on this, we investigated the mechanism of the dysregulation of *SQSTM1* in breast cancer by identifying diverse regulation levels. Furthermore, GSVA and GSEA analysis were constructed to elucidate functional regulating performance of *SQSTM1* in breast cancer. Our finding indicated that *SQSTM1* is a novel therapeutic biomarker and may be useful for improvement of breast cancer treatment (Fig. 1).

Materials And Methods

Patients and samples selection

This study was approved by the Human Ethics Committee of the First Affiliated Hospital of Xi'an Jiaotong University. Between 2011 and 2017, paraffin-embedded breast cancer tissues (n = 27) and adjacent non-tumorous tissues (n = 27) from 27 patients underwent breast cancer surgery at Department of Breast Surgery was collected in our study.

SQSTM1 mRNA expression analysis

SQSTM1 mRNA expression in multiple The Cancer Genome Atlas (TCGA) tumors and adjacent normal tissues was explored by "DiffExp" module of TIMER (https://cistrome.shinyapps.io/timer/). The mRNA expression (microarray) differences between breast cancer tissues and normal tissues were selected from

the Gene Expression Omnibus (GEO) database (GSE54002, GSE42568) using GEOquery. *SQSTM1* mRNA expression, *SQSTM1* CNV data, *SQSTM1* mutation profile as well as complete clinical data of breast cancer patients were downloaded from METABRIC using cBioPortal (https://www.cbioportal.org/) and TCGA using UCSC Xena (http://xena.ucsc.edu/) via R (version 3.6.3).

Immunohistochemical staining

SQSTM1 expression in protein level (p62) was estimated by immunohistochemical (IHC). All tissues were formalin-fixed paraffin-embedded and cut in 3-μm sections. All the experiment procedures were performed based on manufacturer's protocol. The rabbit anti-SQSTM1 antibody was provided by Abcam (ab121146). The extent of positively stained cells was graded as: 0 (positive cells 0–5% of the cells), 1 (6–25% of the cells), 2 (26–50% of the cells), 3 (51–75% of the cells), 4 (76–100% of the cells). The staining intensity score was classified by four grades: negative, 0; weak, 1; medium, 2; and strong, 3. Expression levels of p62 were determined by final staining scores, which were calculated by multiplying the positive cells scores and intensity and ranging from 0 to 9.

Immune Infiltration of SQSTM1 in breast cancer

The correlation of *SQSTM1* expression with the tumor-infiltrating levels (B cells, CD4 + T cells, CD8 + T cells, neutrophils, macrophages and dendritic cells) in BRCA (Breast Invasive Carcinoma) was evaluated by the module of TIMER. Based on the previous study, we conducted the gene set variance analysis (GSVA) to estimate the correlation between *SQSTM1* expression and immune cell abundance in breast cancer tissue samples using GSVA package in R software (16).

The prognostic value of SQSTM1 mRNA in public databases

To investigate the prognostic role of *SQSTM1* in breast cancer, Kaplan-Meier (Log-rank tests) analysis was conducted to determine the prognostic significance using METABRIC and GEO breast cancer cohorts (GSE1456, GSE9195).

Transcription factors and miRNA identification

GCBI (https://www.gcbi.com.cn) is a web which integrates diverse genetic, clinical and bioinformatic data (17). In this study, we used GCBI to identify transcription factors which interacts with SQSTM1. The Cistrome (http://cistrome.org/db/) is a platform which includes ChIP-seq, DNase-seq and ATAC-seq data from multiple public databases of human and mouse (18). In our study, we used Cistrome's Chip-seq data to confirm the transcription factors directly bound to *SQSTM1* DNA. The miRDB (http://www.mirdb.org/),

DIANA tool (www.microrna.gr), TargetScan (http://www.targetscan.org/) are online resources for miRNAs and targets predictions. In this study, they were used to predict potential microRNAs targeting on SQSTM1 mRNA.

Functional enrichment analysis

The limma package in R was performed for TCGA dataset differential expression analysis. The cutoff for log (Fold change) FC in our study was 0.477 and a *P* value < 0.05 was statistically significant. The co-expressed genes related to *SQSTM1* in breast cancer was retrieved from the Coexpedia (http://www.coexpedia.org/). The Gene-set enrichment analysis for Gene ontology (GO), the Broad Molecular Signatures Database (MSigDB) and Kyoto Encyclopedia of Genes and Genomes pathway (KEGG) was performed by clusterProfiler package and the GSEA program (version 3.0).

Statistical analysis

The statistical analysis in this study was performed by SPSS (version 19.0) software. The association between expression of *SQSTM1* and clinical parameters were analyzed by Chi-square test. Multivariate logistic regression was used to find independent influence factors for *SQSTM1* mRNA expression. Screening for prognostic factors for breast cancer patients was performed by univariate and multivariate Cox regression and visualized by Review Manager (version 5.3). Overall survival of breast cancer was examined with Kaplan–Meier curve and compared by log-rank test using GraphPad prism (version 7.0) software. A value of $P < 0.05$ was considered to be statistically significant.

Results

SQSTM1 is overexpressed in patients with breast cancer at RNA and protein levels

We first explored *SQSTM1* mRNA expression distribution in various tumors using TIMER dataset. It was shown that *SQSTM1* mRNA was significantly higher in most common tumor tissues compared with normal tissues. It is noteworthy that *SQSTM1* acted as an oncogene in different subtypes (Luminal, Her2, basal) of breast cancer (Fig. 2A). Furthermore, same results were validated in 2 independent GEO cohorts (GSE54002, GSE42568) when compared *SQSTM1* mRNA expression between breast cancer tissues and normal tissues (Fig. 2B).

SQSTM1 expression in protein level was examined by IHC staining. As shown in Fig. 3, p62 was mainly expressed in the cytoplasm of breast cancer cells and significantly overexpressed in breast cancer tissues compared with adjacent non-tumorous tissues ($P < 0.001$). The diagnostic performance of p62 for distinguishing breast cancer from non-breast cancer was assessed by ROC analysis with AUC of 0.846 (95% CI = 0.760–0.933, $P < 0.001$). Our results indicated that p62 can be used as a diagnostic biomarker

(Fig. 3). All the above data demonstrated the oncogene role of *SQSTM1* in breast cancer and can be utilized as a predictive tool for achieving precision medicine.

Correlation of SQSTM1 mRNA expression with clinicopathologic characteristics

In order to explore the clinical significance of *SQSTM1* in breast cancer, we explored RNA-seq data from METABRIC database with *SQSTM1* mRNA expression (n = 1336) and detailed clinical information. Using Chi-square test, we assessed the correlation between *SQSTM1* mRNA expression and clinical-pathologic characteristics. As were shown in Table 1, the expression level of *SQSTM1* mRNA was significantly associated with ER status ($P = 0.018$), hormone therapy ($P = 0.006$).Multivariate logistic regression indicated that ER status (OR = 1.762, $P < 0.05$) and PR status (OR = 0.743, $P < 0.05$) were independent influence factors of *SQSTM1* mRNA expression in breast cancer patients (Table 2).

Table 1
Relationship between *SQSTM1* mRNA expression and clinicopathological characteristics

Variables	Number	*SQSTM1*		*P* value
		Low	High	
Age				0.064
< 50	306	239	67	
≥ 50	1030	750	280	
ER				0.018
Positive	1024	742	282	
Negative	312	247	65	
PR				0.348
Positive	695	522	173	
Negative	641	467	174	
HER2				0.329
Positive	165	117	48	
Negative	1171	872	299	
Chemotherapy				0.252
YES	299	229	70	
NO	1037	760	277	
Hormone therapy				0.006
YES	818	584	234	
NO	518	405	113	
Radio therapy				0.869
YES	900	665	235	
NO	436	324	112	
Neoplasm Histologic Grade				0.304
1 + 2	640	482	158	
3	696	507	189	

Abbreviation: ER: estrogen receptor; PR: progesterone receptor; HER-2: Human epidermal growth factor receptor-2

Variables	Number	*SQSTM1*		*P* value
Nottingham prognostic index				0.629
≤ 5.4	1192	880	312	
> 5.4	144	109	35	
Primary Tumor Laterality				0.858
Left	687	510	177	
Right	649	479	170	
Tumor stage				0.766
0,1,2	1218	903	315	
3, 4	118	86	32	

Abbreviation: ER: estrogen receptor; PR: progesterone receptor; HER-2: Human epidermal growth factor receptor-2

Table 2
Correlation of *SQSTM1* mRNA expression with clinicopathological characteristics in by multivariate logistic regression analysis.

variables	*P* value	OR (95%CI)
Age (≥ 50 vs < 50)	0.346	1.177 (0.839, 1.651)
ER status (Positive vs Negative)	0.007	1.762 (1.164, 2.667)
PR status (Positive vs Negative)	0.043	0.743 (0.557, 0.990)
HER-2 status (Positive vs Negative)	0.173	1.308 (0.889, 1.926)
Chemotherapy (YES vs NO)	0.879	0.972 (0.672, 1.405)
Hormone therapy (YES vs NO)	0.161	1.234 (0.919, 1.656)
Radio therapy (YES vs NO)	0.839	1.029 (0.783, 1.352)
Neoplasm Histologic Grade (3 vs 1 + 2)	0.089	1.280 (0.693, 1.702)
Nottingham prognostic index (> 5.4 vs ≤ 5.4)	0.247	0.758 (0.474, 1.211)
Primary Tumor Laterality (Right vs Left)	0.701	1.050 (0.819, 1.345)
Tumor stage (3 + 4 vs 0 + 1 + 2)	0.604	1.133 (0.708, 1.812)

Abbreviation: ER: estrogen receptor; PR: progesterone receptor; HER-2: human epidermal growth factor receptor-2; OR: Odds ratio

Table 3
Significant neighboring genes associated with SQSTM1 in breast cancer

Rank	Gene	ENSEMBLE ID	Description	Score
1	LAMTOR2	ENSG00000116586	Late endosomal/lysosomal adaptor, MAPK and MTOR activator 2	3.668
2	PIR	ENSG00000087842	pirin	2.132
3	GULP1	ENSG00000144366	GULP, engulfment adaptor PTB domain containing 1	2.073
4	TXNRD1	ENSG00000198431	thioredoxin reductase 1	1.747
5	FAM129A	ENSMUSG00000026483	family with sequence similarity 129 member A	1.714
6	ARL4C	ENSG00000188042	ADP ribosylation factor like GTPase 4C	1.633
7	SEL1L3	ENSG00000091490	sel-1 suppressor of lin-12-like 3 (C. elegans)	1.606
8	SLPI	ENSG00000124107	secretory leukocyte peptidase inhibitor	1.460
9	KIZ	ENSG00000088970	kizuna centrosomal protein	1.429
10	SUPT4H1	ENSG00000213246	SPT4 homolog, DSIF elongation factor subunit	1.394
11	RTCA	ENSG00000137996	RNA 3'-terminal phosphate cyclase	1.376
12	PPP1R15A	ENSG00000087074	protein phosphatase 1 regulatory subunit 15A	1.364
13	DNAJB9	ENSG00000128590	DnaJ heat shock protein family (Hsp40) member B9	1.329
14	CASP4	ENSG00000196954	caspase 4	1.263
15	HHLA3	ENSG00000197568	HERV-H LTR-associating 3	1.254
16	IDH1	ENSG00000138413	isocitrate dehydrogenase 1 (NADP+)	1.253
17	PTEN	ENSG00000171862	phosphatase and tensin homolog	1.224
18	HMOX1	ENSG00000100292	heme oxygenase 1	1.213
19	GFPT1	ENSG00000198380	glutamine–fructose-6-phosphate transaminase 1	1.062

Screening for prognostic factors for breast cancer patients

Using univariate and multivariate cox regression analysis, we identified age at diagnosis, ER status, PR status, HER2 status, chemotherapy, hormone therapy, neoplasm histologic grade, Nottingham prognostic index, tumor stage, *SQSTM1* mRNA expression influenced the OS of breast cancer patients. Moreover, age at diagnosis, HER2 status, radio therapy, Nottingham prognostic index, tumor stage, *SQSTM1* mRNA were independent prognostic factors for breast cancer patients (Fig. 4).

Prognostic value of SQSTM1 in breast cancer

In our METABRIC cohort, by plotting Kaplan-Meier curve, we found that breast cancer patients with higher *SQSTM1* mRNA expression (median survival time = 130.7 months) tended to have a worse overall survival (OS) than patients with lower *SQSTM1* mRNA expression (median survival time = 172.9 months, $P < 0.001$) (Fig. 5A). By using GEO database, we validated the prognostic role of *SQSTM1* in breast cancer patients. Lower *SQSTM1* expression indicated favorable prognosis (OS, DFS and RFS) in breast cancer from 2 independent cohorts (GSE1456, GSE9195) (Fig. 5B **and C**). All these results indicated that high *SQSTM1* mRNA expression may be a poor prognostic biomarker of breast cancer.

Association between *SQSTM1* expression and immune infiltration level in breast cancer

Tumor microenvironment has been demonstrated to serve as a “complex network” of different tumor cells, extracellular matrix components, chemotactic factor and other types of cells which forms the basis for tumor cancer cell proliferation and metastasis (19). Here, we analyzed the correlation between *SQSTM1* expression and immune infiltration levels in breast cancer. As were shown in Fig. 6A, *SQSTM1* expression was inversely associated with infiltrating levels in breast cancer. It is noteworthy that *SQSTM1* expression has positive correlation with tumor purity in breast cancer.

Next, we compared *SQSTM1* expression in breast cancer patients with different GSVA score (lowest 25% versus highest 25%) of multiple immune cells through TCGA dataset. The results indicated that higher *SQSTM1* expression was significantly correlated with higher infiltration of memory B cell, activated CD4 + T cell and neutrophil. However, it was shown different result in activated CD8 + T cells (Fig. 6B). These results further demonstrated that *SQSTM1* may serve as an immune modulatory role in breast cancer and large-scale projects are still urgently needed in the near future.

The mechanism of *SQSTM1* expression dysregulation in patients with breast cancer

Numerous studies have indicated that CNV unbalanced gene expression by disrupting the structure of gene coding regions. Next, we evaluated the copy number alterations of *SQSTM1* using a cohort of 1904 breast cancer patients from METABRIC database (Shallow Deletion, n = 192; Diploid, n = 1460; Amplification, n = 31; Gain, n = 221). We found that 56.8% (252/444) patients in the altered group harboring *SQSTM1* amplification/gain. This result indicated that the amplification/gain of gene copy

numbers was likely to be one of the main mechanisms of over-expression of *SQSTM1* in breast cancer patients. Consistently, breast cancer patients with *SQSTM1* amplification/gain exhibited higher *SQSTM1* mRNA expression compared with shallow deletion and diploid (no alteration) group. By drawing the Kaplan–Meier survival curve, the results revealed that patients with *SQSTM1* amplification/gain significantly associated with worse overall survival compared with other groups (Fig. 7).

In order to explore the potential clinical significance of *SQSTM1* mutation, we first evaluated its mutation profile in the METABRIC database. The results showed that there was no *SQSTM1* mutation in the selected patients. Next, patients obtained from TCGA database with mutation profiles were validated. Compared with the high-frequency altered genes such as *PIK3CA*, *AKT1* and *PTEN*, *SQSTM1* mutation frequency is rare and has no predictive value on the prognosis of breast cancer patients ($P = 0.338$) (Fig. 8).

In addition to point mutations and CNV, epigenetic changes (especially DNA methylation) also play an important role in regulating specific genes expression and the development of breast cancer. We then investigated characteristics of the *SQSTM1* promoter methylation in breast cancer. First, the heat map of the *SQSTM1* methylation value used different probes were drawn from TCGA dataset. The Kaplan-Meier survival analysis showed that patients with lower methylation of *SQSTM1* experienced longer overall survival and disease specific survival significantly, which further suggested that the high expression of *SQSTM1* plays a critical prognostic role in breast cancer (Fig. 9). All of the above data showed that upregulation of *SQSTM1* expression involved in the development and progression of breast cancer.

Regulation of SQSTM1 in other transcription and post-transcription level

Next, we investigated what transcription factors might regulate *SQSTM1* in the upstream in breast cancer by GCBI platform. It can predict the transcription factors through the Transfac database from 2000 bp upstream and 500 bp downstream of the start site based on the transcript of each gene (Ensembl database). First, we identified the transcription factors which have the highest grade among the predicted genes (Supplementary Fig. 1). Based on this analysis, we then used Chip-seq data of Cistrome and confirmed that CTCF, ERG, EP300, E2F1, FOXA1 can directly bind to *SQSTM1* DNA in breast cancer (**Supplementary Table 1**).

In addition, using miRDB, DIANA tools, Targetscan databases, we explored what miRNAs were involved in the post-transcription regulation of SQSTM1. Notably, we set strict screening criteria for the databases (miRDB: Score > 70, Targetscan: context + + score <-0.4. context + + score percentile > 98, DIANA tools: miTG score > 0.8). Finally, 6 common miRNAs (miR-106b-5p, miR-20a-5p, miR-106a-5p, miR-93-5p, miR-17-5p, miR-20b-5p) were identified in three datasets (**Supplementary Fig. 2**).

SQSTM1 is related to cell signal transduction, oxidative stress and autophagy

To clarify the biological molecular mechanism of *SQSTM1* in breast cancer, we first performed differential gene expression analysis based on LIMMA package in samples with high expression of *SQSTM1* (N = 552) and low expression of *SQSTM1* (N = 552) from TCGA database. Our analysis found that a total of 387 genes were significantly up-regulated and 561 genes were significantly down-regulated (Fig. 10A). In addition, *SQSTM1* was observed to be associated with various signal transduction pathways according to KEGG analysis, such as JAK/STAT and PI3K/Akt, which was consistent with previous reports (Fig. 10B). Next, we used the MSigDB Hallmark gene set (Fig. 10C) for GSEA. The results showed that compared with high expression levels of *SQSTM1*, low levels of *SQSTM1* were significantly related to oxidative phosphorylation, peroxisome, DNA repair and reactive oxygen species pathway.

Coexpedia is a distinct co-expression database which offers biomedical hypotheses through medical subject headings. In our study, the co-expression genes in breast cancer associated with *SQSTM1* were explored from Coexpedia database in order to clarify the underlying regulation network and mechanism of breast cancer. Through exploring GSE12237, GSE7848 and GSE14018, a total of 19 genes, such as LAMTOR2, PIR, and GULP1 were identified (**Supplementary Fig. 3**) (**Table 4**).

The GO analysis based on *SQSTM1* and its related genes were then constructed. The top 20 Go terms enrichment of the gene lists was showed in **Supplementary Fig. 4**. The most significantly enriched GO terms of BP, CC and MF for *SQSTM1* and co-expressed genes were negative regulation of endoplasmic reticulum unfolded protein response (GO: 1900102; $P = 7.12E-05$), ionotropic glutamate receptor binding (GO:0035255; $P = 0.0005$), DSIF complex (GO:0032044; $P = 0.0017$) and amphisome (GO:0044753, $P = 0.001693$), respectively. Altogether, these data indicated that *SQSTM1* is related to cell signal transduction, oxidative stress and autophagy thus plays a key role in the progression of breast cancer.

Discussion

Breast cancer remains a global health concern as a type of aggressive tumor. Over the past few years, a great number of studies have demonstrated the molecular characteristics of breast cancer with genetic and clinical heterogeneity which restrict the accuracy of typical morphological and pathological classification. However, newly molecular targeted drugs by identifying and discovering diagnostic and prognostic biomarkers have offered new fields of breast cancer treatment. Therefore, discovering new therapeutic target involved in the progression of tumor to improve the prognosis of breast cancer is urgent nowadays.

In the current study, we explored the clinical significance of *SQSTM1* based on RNA expression data from METABRIC/GEO databases and protein expression data from our hospital cohort. We found that *SQSTM1* mRNA and protein level were significantly higher in breast cancer tissues than adjacent non-tumorous tissues. It was also showed that *SQSTM1* was a high-risk factor and could be an independent prognostic

factor in patients with breast cancer using univariate and multivariate Cox analyses. Besides, we also found high *SQSTM1* mRNA expression predicted poor outcome in breast cancer patients through multiple databases. By plotting ROC curve, we observed the AUC value for p62 was 0.846 which was a potential predictor of breast cancer. All these data suggest that *SQSTM1* might be a therapeutic target for breast cancer.

Numerous studies have demonstrated the role of autophagy activity on tumor cells involving in modulating functions of T cells such as CD8 + cytotoxic T cells and regulatory T cells (20, 21). Although it is widely believed that breast cancer was a relatively non-immunogenic cancer and showed poor response to immunotherapy, immuno-oncology focused on tumor-infiltrating lymphocytes have showed remarkable progress in treatment of breast cancer recently, especially for those with hormone receptor negative subtypes (22). Since the role of *SQSTM1* in immunity of breast cancer remains unclear, in this study, we found that *SQSTM1* mRNA expression was correlated with diverse immune infiltration levels significantly. In addition, we also found that higher *SQSTM1* expression was significantly correlated with higher infiltration of memory B cell, activated CD4 + T cell and neutrophil. These data indicated the underlying role of *SQSTM1* in the breast cancer microenvironment

Then, we tried to investigate the mechanisms of *SQSTM1* dysregulation. Through examining its CNV, DNA methylation and somatic mutation status in patients with breast cancer, it was found that copy number amplification/ gain of *SQSTM1* could be the key driver mechanism for its overexpression. In addition, we also observed *SQSTM1* CNV and methylation status were significantly associated with survival of breast cancer patients. In the future, detection of *SQSTM1* copy number amplification, methylation, and overexpression status may provide new guidelines of evaluation and adjustment of breast cancer treatment strategies.

Predicting the transcription factors regulating *SQSTM1* made it possible to get a better understanding of the gene expression patterns and regulation mechanisms in breast cancer. In our study, we identified CTCF, ERG, EP300, E2F1, FOXA1 can directly bind to *SQSTM1* DNA in breast cancer and may help complement the regulatory network thus develop novel effective targeted therapeutic strategies for patients.

miRNAs are a class of endogenous non-coding RNAs that regulate the expression of different transcription factors at the post-transcriptional level. Since 2002, George and other scientists first reported that miRNAs are dysregulated in tumors, more and more scholars devoted to studying the role of miRNAs in tumorigenesis and development (23). Current findings showed that miRNAs were dysregulated in a variety of malignant tumors, and the regulation of tumorigenesis and affected various activities, including tumor cell proliferation, invasion and metastasis, drug resistance, angiogenesis and immune escape. Previous study demonstrated that miR-17/20/93/106 targeted *SQSTM1* and promoted hematopoietic cell expansion (24) thus implicated *SQSTM1* expression was regulated by miRNAs in different malignancies. In our study, we identified 6 common miRNAs (miR-106b-5p. miR-20a-5p, miR-

106a-5p, miR-93-5p, miR-17-5p, and miR-20b-5p) by different databases and more convincing evidences in the future which may lead to novel therapeutic strategies for breast cancer.

A large number of *in vivo* and *in vitro* studies have reported that *SQSTM1* can promote tumor development and malignant phenotypes such as tumor growth, invasion, migration and apoptosis inhibition via multiple signal transduction pathways (13, 25, 26). Previous studies have confirmed the double-edged sword effect of autophagy in regulating tumors which highlighted the importance of *SQSTM1* expression pattern. It has recently been reported that the oxidation of *SQSTM1* promoted its oligomerization via disulfide-linked conjugates then activated autophagy which facilized cell homeostasis and survival under oxidative stress from aging or cancer (27). In addition, an increase of *SQSTM1* in autophagy-deficient cells directly bonded to and inhibited nuclear RNF168, an E3 ligase essential for histone H2A ubiquitination and DNA damage responses (28). Combining previous results and our findings, it can be concluded that autophagy defects with *SQSTM1* accumulation can impair the DNA repair activity of cells thus leading to tumorigenesis.

Then, a total of 19 co-expressed genes of *SQSTM1* in breast cancer were explored by Coexpedia. The highest score gene was LAMTOR2 which is a convergence point for RAF/MEK/ERK and PI3K/AKT/mTOR pathways (29). Numerous studies have verified the significance of two signaling pathway in the progression of breast cancer. In addition, Lin and his colleagues have reported that *LAMTOR2* interacted with *SQSTM1* and was required for recruiting *TAX1BP1* to autophagosomes (30). However, to the best of our knowledge, there is no study about the associated between *LAMTOR2* and *SQSTM1* in breast cancer and the underlying regulation network requires further investigations.

Furthermore, the potential biological processes mainly involved in regulation of endoplasmic reticulum unfolded protein response has also been discussed. Unfolded protein response (UPR) is a protective cellular response activated by endoplasmic reticulum stress (31). *SQSTM1* was known to protect cells against tunicamycin (TM)-mediated oxidative damage through Nrf2 activation (32, 33). This finding further demonstrated the importance of *SQSTM1* in regulating oxidative stress of tumor progression.

Accumulating evidence demonstrated that *SQSTM1* dysregulation involved in multiple tumor progression. In the current study, we found that *SQSTM1* acted as an oncogene in breast cancer. Importantly, the overexpression of *SQSTM1* is related to poor prognosis. We also explored the dysregulation mechanism of *SQSTM1* and found CNV and methylation might be the potential targets for patients. In the upstream of *SQSTM1*, several transcription factors and miRNA have also been identified. The miRNA mimic/inhibitor might be a promising target for new cancer therapy in the future. The precise mechanism of *SQSTM1* needed further *in vivo* and *in vitro* experiments to elucidate its biological function. These results have provided such an exciting future, in which focusing on *SQSTM1* profiling might on one day help tailor therapy strategies and achieve better management of breast cancer.

Conclusion

This study indicated that *SQSTM1* is a promising diagnostic and prognostic target in breast cancer patients by exploring multiple cohorts. Overexpression of *SQSTM1* was correlated with tumor progression, poor survival, immune infiltrations in breast cancer. Multiple mechanisms involved in transcription and post-transcription levels are responsible for the dysregulation of *SQSTM1*. In addition, elevated *SQSTM1* was also associated with cell signal transduction, oxidative stress and autophagy in breast cancer. Our study provides new insights into the biological and clinical characteristics of *SQSTM1* in breast cancer. Further large-scale multicentre clinical trials and studies are urgently needed to reveal complete gene expression profiles and provide therapeutic regimens for breast cancer patients.

Abbreviations

SQSTM1: Sequestosome 1; METABRIC: Molecular Taxonomy of Breast Cancer International Consortium; FFPE: formalin fixation and paraffin embedding; METABRIC: Molecular Taxonomy of Breast Cancer International Consortium; FFPE: formalin fixation and paraffin embedding; TIMER: Tumor Immune Estimation Resource; GSVA: Gene Set Variation Analysis; CNV: copy number variation; GSEA: Gene Set Enrichment Analysis; ROC: receiver operating characteristic; GEO: Gene Expression Omnibus; ER: estrogen receptor; PR: progesterone receptor; HER-2: Human epidermal growth factor receptor-2; OS: overall survival; DFS: disease-free survival; RFS: relapse Free Survival; OR: odds ratio; HR: hazard ratio; CI: confidence interval; UPR: Unfolded protein response.

Declarations

Acknowledgement

None

Authors’ contributions

Yang Liu retrieved the relevant data and drafted the manuscript. Qian Du and Dan Sun designed the structure of this article. Ruiying Han and Mengmeng Teng critically revised the manuscript. Ying Zhang and Yuzhu Dong analyzed the data. Siying Chen and Haisheng You reviewed the final manuscript prior to submission. All authors read and approved the final manuscript.

Funding

The work is supported by National Natural Science Foundation of China (No., 81502616, 81672954), the Science and Technology Development of Shaanxi Province, China (No. 2020SF-210) and the Science and Technology Plan Project of Xi’an [No. 2019114613YX001SF044 (15)]

Availability of data and materials

The data used to support the findings of this study are included within the article.

Ethics approval and consent to participate

The study protocol was in accordance with ethical standards of 1964 Helsinki Declaration and approved by the Human Ethics Committee of the First Affiliated Hospital of Xi'an Jiaotong University. Informed consent was obtained from all individual participants included in the study.

Consent for publication

None

Competing interests

The authors declare that they have no conflict of interest.

References

1. Waks AG, Winer EP. Breast Cancer Treatment: A Review. Jama. 2019;321(3):288-300.
2. Bray F, Ferlay J, Soerjomataram I, Siegel RL, Torre LA, Jemal A. Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA: a cancer journal for clinicians. 2018;68(6):394-424.
3. Yeo SK, Guan JL. Breast Cancer: Multiple Subtypes within a Tumor? Trends in cancer. 2017;3(11):753-60.
4. Duran A, Linares JF, Galvez AS, Wikenheiser K, Flores JM, Diaz-Meco MT, et al. The signaling adaptor p62 is an important NF-kappaB mediator in tumorigenesis. Cancer cell. 2008;13(4):343-54.
5. Duran A, Amanchy R, Linares JF, Joshi J, Abu-Baker S, Porollo A, et al. p62 is a key regulator of nutrient sensing in the mTORC1 pathway. Molecular cell. 2011;44(1):134-46.
6. Moscat J, Karin M, Diaz-Meco MT. p62 in Cancer: Signaling Adaptor Beyond Autophagy. Cell. 2016;167(3):606-9.
7. Moscat J, Diaz-Meco MT. p62 at the crossroads of autophagy, apoptosis, and cancer. Cell. 2009;137(6):1001-4.
8. Inoue D, Suzuki T, Mitsuishi Y, Miki Y, Suzuki S, Sugawara S, et al. Accumulation of p62/SQSTM1 is associated with poor prognosis in patients with lung adenocarcinoma. Cancer science. 2012;103(4):760-6.

9. Luo RZ, Yuan ZY, Li M, Xi SY, Fu J, He J. Accumulation of p62 is associated with poor prognosis in patients with triple-negative breast cancer. OncoTargets and therapy. 2013;6:883-8.
10. Park JM, Huang S, Wu TT, Foster NR, Sinicrope FA. Prognostic impact of Beclin 1, p62/sequestosome 1 and LC3 protein expression in colon carcinomas from patients receiving 5-fluorouracil as adjuvant chemotherapy. Cancer biology & therapy. 2013;14(2):100-7.
11. Umemura A, He F, Taniguchi K, Nakagawa H, Yamachika S, Font-Burgada J, et al. p62, Upregulated during Preneoplasia, Induces Hepatocellular Carcinogenesis by Maintaining Survival of Stressed HCC-Initiating Cells. Cancer cell. 2016;29(6):935-48.
12. Wang BJ, Tang YD, Yu BY, Gui D, Xu H. Expression of autophagy-related factor p62 for lung cancer diagnosis and prognosis: A systematic review and meta-analysis. Mathematical biosciences and engineering : MBE. 2019;16(6):6805-21.
13. Li SS, Xu LZ, Zhou W, Yao S, Wang CL, Xia JL, et al. p62/SQSTM1 interacts with vimentin to enhance breast cancer metastasis. Carcinogenesis. 2017;38(11):1092-103.
14. Xu LZ, Li SS, Zhou W, Kang ZJ, Zhang QX, Kamran M, et al. p62/SQSTM1 enhances breast cancer stem-like properties by stabilizing MYC mRNA. Oncogene. 2017;36(3):304-17.
15. Claude-Taupin A, Fonderflick L, Gauthier T, Mansi L, Pallandre JR, Borg C, et al. ATG9A Is Overexpressed in Triple Negative Breast Cancer and Its In Vitro Extinction Leads to the Inhibition of Pro-Cancer Phenotypes. Cells. 2018;7(12).
16. Charoentong P, Finotello F, Angelova M, Mayer C, Efremova M, Rieder D, et al. Pan-cancer Immunogenomic Analyses Reveal Genotype-Immunophenotype Relationships and Predictors of Response to Checkpoint Blockade. Cell reports. 2017;18(1):248-62.
17. Huang H, Zheng J, Shen N, Wang G, Zhou G, Fang Y, et al. Identification of pathways and genes associated with synovitis in osteoarthritis using bioinformatics analyses. Scientific reports. 2018;8(1):10050.
18. Zheng R, Wan C, Mei S, Qin Q, Wu Q, Sun H, et al. Cistrome Data Browser: expanded datasets and new tools for gene regulatory analysis. Nucleic acids research. 2019;47(D1):D729-D35.
19. Junttila MR, de Sauvage FJ. Influence of tumour micro-environment heterogeneity on therapeutic response. Nature. 2013;501(7467):346-54.
20. Zhong Z, Sanchez-Lopez E, Karin M. Autophagy, Inflammation, and Immunity: A Troika Governing Cancer and Its Treatment. Cell. 2016;166(2):288-98.
21. Ladoire S, Enot D, Senovilla L, Chaix M, Zitvogel L, Kroemer G. Positive impact of autophagy in human breast cancer cells on local immunosurveillance. Oncoimmunology. 2016;5(6):e1174801.
22. Burugu S, Asleh-Aburaya K, Nielsen TO. Immune infiltrates in the breast cancer microenvironment: detection, characterization and clinical implication. Breast cancer. 2017;24(1):3-15.
23. Calin GA, Dumitru CD, Shimizu M, Bichi R, Zupo S, Noch E, et al. Frequent deletions and down-regulation of micro- RNA genes miR15 and miR16 at 13q14 in chronic lymphocytic leukemia. Proceedings of the National Academy of Sciences of the United States of America. 2002;99(24):15524-9.

24. Meenhuis A, van Veelen PA, de Looper H, van Boxtel N, van den Berge IJ, Sun SM, et al. MiR-17/20/93/106 promote hematopoietic cell expansion by targeting sequestosome 1-regulated pathways in mice. Blood. 2011;118(4):916-25.
25. Zhang J, Yang S, Xu B, Wang T, Zheng Y, Liu F, et al. p62 functions as an oncogene in colorectal cancer through inhibiting apoptosis and promoting cell proliferation by interacting with the vitamin D receptor. Cell proliferation. 2019;52(3):e12585.
26. Jiang X, Huang Y, Liang X, Jiang F, He Y, Li T, et al. Metastatic prostate cancer-associated P62 inhibits autophagy flux and promotes epithelial to mesenchymal transition by sustaining the level of HDAC6. The Prostate. 2018;78(6):426-34.
27. Carroll B, Otten EG, Manni D, Stefanatos R, Menzies FM, Smith GR, et al. Oxidation of SQSTM1/p62 mediates the link between redox state and protein homeostasis. Nature communications. 2018;9(1):256.
28. Wang Y, Zhang N, Zhang L, Li R, Fu W, Ma K, et al. Autophagy Regulates Chromatin Ubiquitination in DNA Damage Response through Elimination of SQSTM1/p62. Molecular cell. 2016;63(1):34-48.
29. De Araujo ME, Erhart G, Buck K, Muller-Holzner E, Hubalek M, Fiegl H, et al. Polymorphisms in the gene regions of the adaptor complex LAMTOR2/LAMTOR3 and their association with breast cancer risk. PloS one. 2013;8(1):e53768.
30. Lin CY, Nozawa T, Minowa-Nozawa A, Toh H, Aikawa C, Nakagawa I. LAMTOR2/LAMTOR1 complex is required for TAX1BP1-mediated xenophagy. Cellular microbiology. 2019;21(4):e12981.
31. Park JS, Oh SY, Lee DH, Lee YS, Sung SH, Ji HW, et al. p62/SQSTM1 is required for the protection against endoplasmic reticulum stress-induced apoptotic cell death. Free radical research. 2016;50(12):1408-21.
32. Mozzini C, Cominacini L, Garbin U, Fratta Pasini AM. Endoplasmic Reticulum Stress, NRF2 Signalling and Cardiovascular Diseases in a Nutshell. Current atherosclerosis reports. 2017;19(8):33.
33. Cominacini L, Mozzini C, Garbin U, Pasini A, Stranieri C, Solani E, et al. Endoplasmic reticulum stress and Nrf2 signaling in cardiovascular diseases. Free radical biology & medicine. 2015;88(Pt B):233-42.

Figures

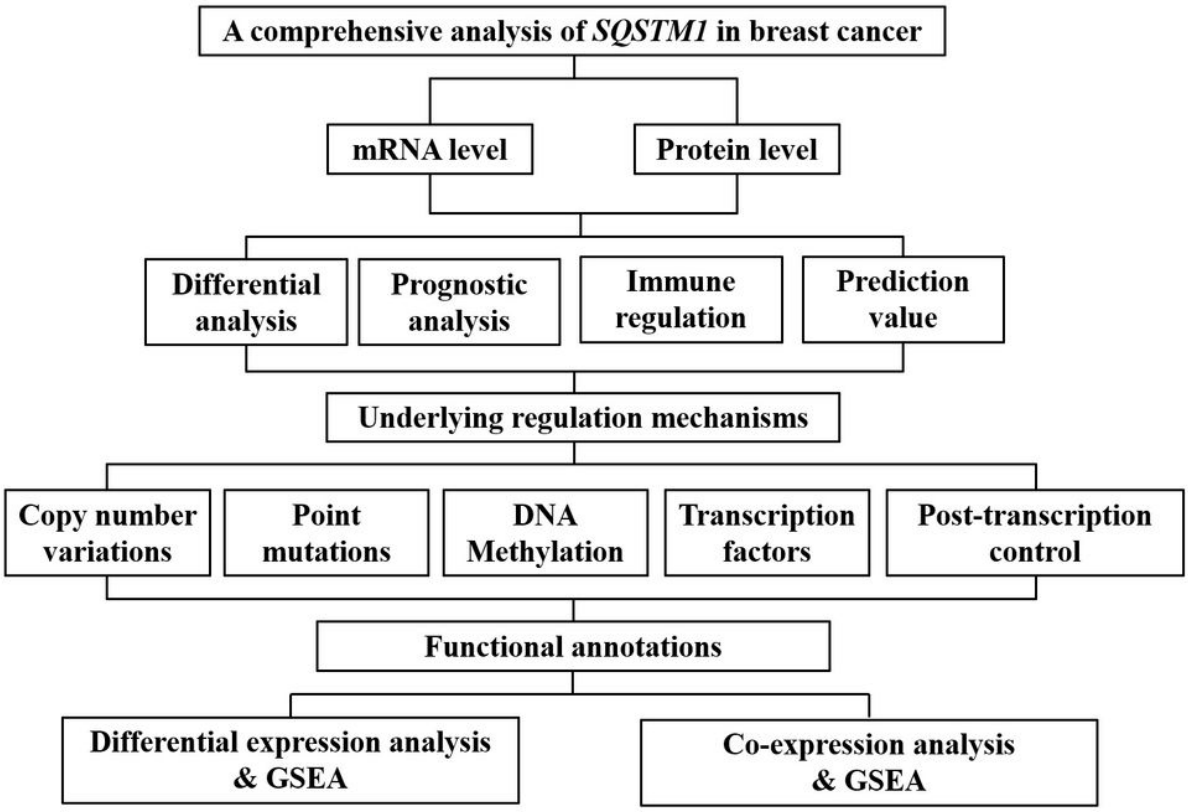

Figure 1

Workflow of this study.

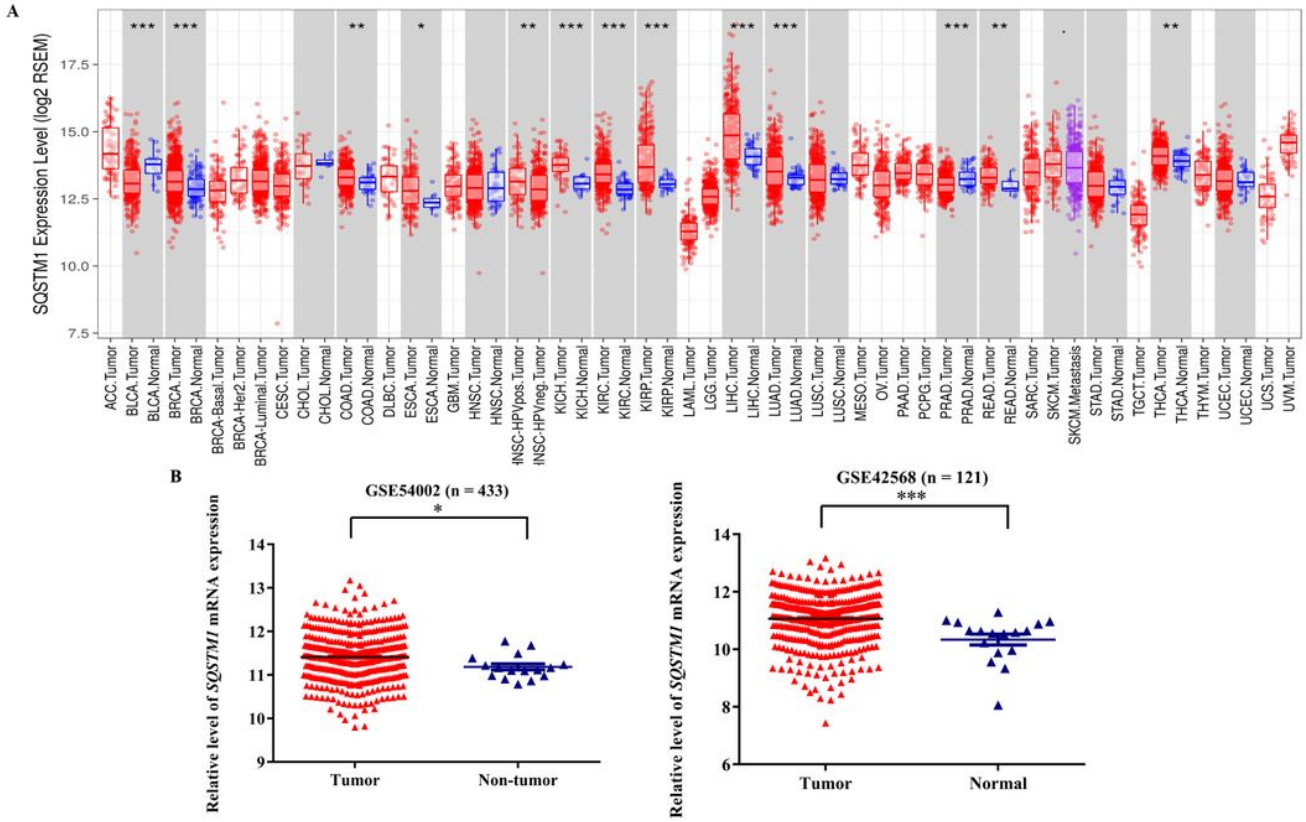

Figure 2

The expression profiling analysis of SQSTM1. (A)The pan-cancer expression profiling of SQSTM1 mRNA in human cancers. The below row refers to the standard abbreviation of tumor in TCGA. The color refers to the tumor (red) or normal (blue). (B) SQSTM1 mRNA expression in breast tumor tissues and normal tissues. P-value significant codes: *** $P \leq 0.001$, **$P < 0.01$, * $P < 0.05$.

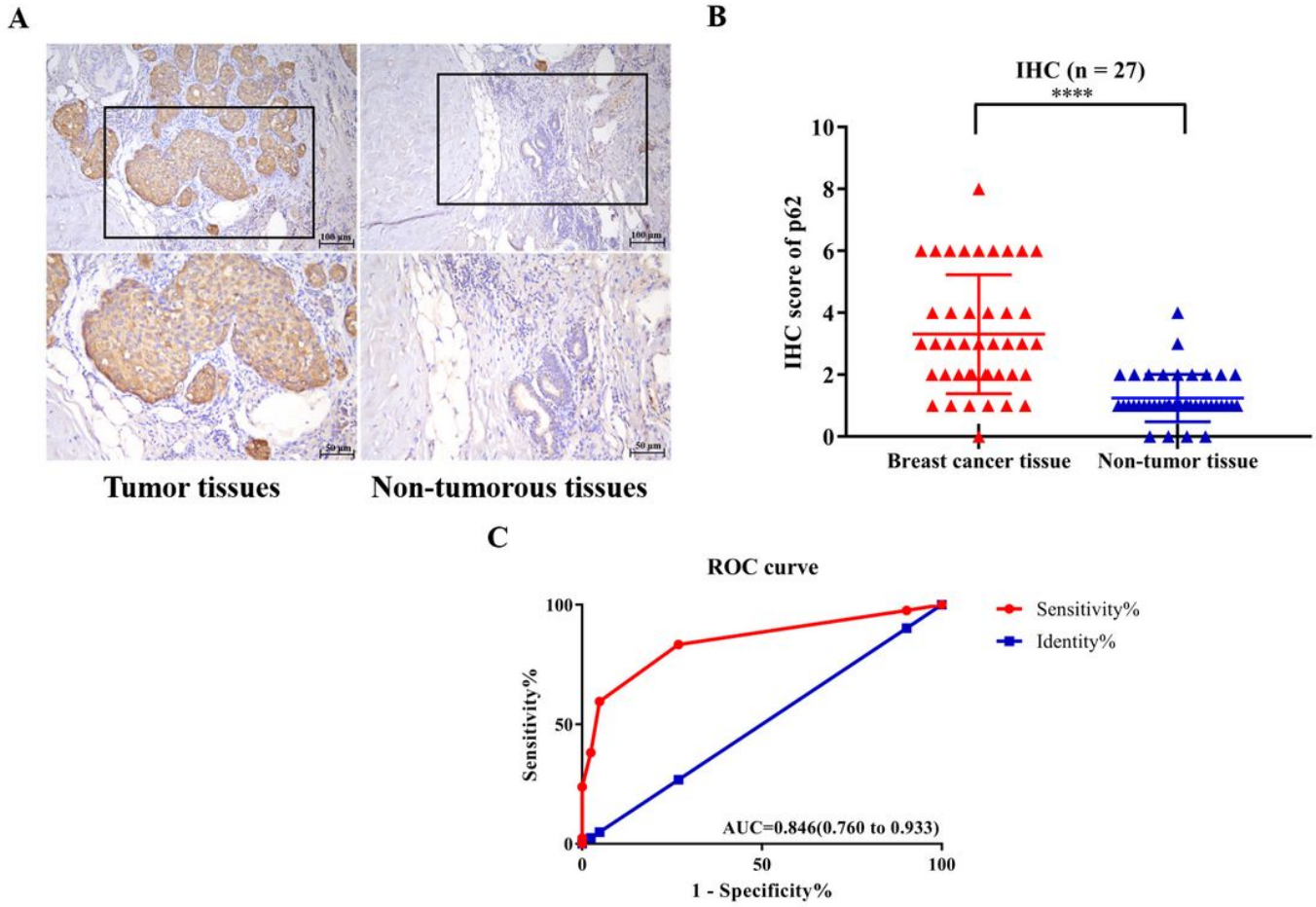

Figure 3

p62 was overexpressed in breast cancer tissues. (A) p62 expression in representative tissue samples. (B) A scatter gram about p62 expression in breast cancer tissues and adjacent non-tumorous tissues. (C) The clinical significance of p62 evaluated by ROC.

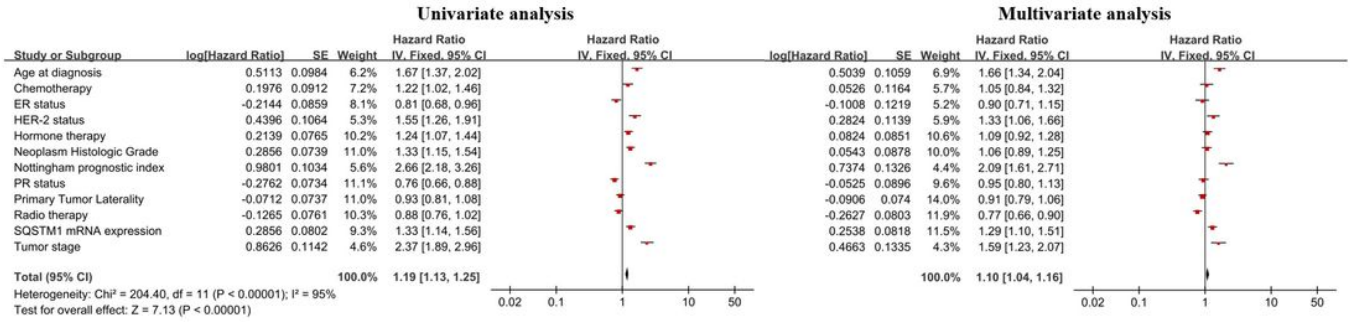

Figure 4

Forest plot shows the result of SQSTM1 in breast cancer survival by univariate Cox regression and multivariate Cox regression analysis.

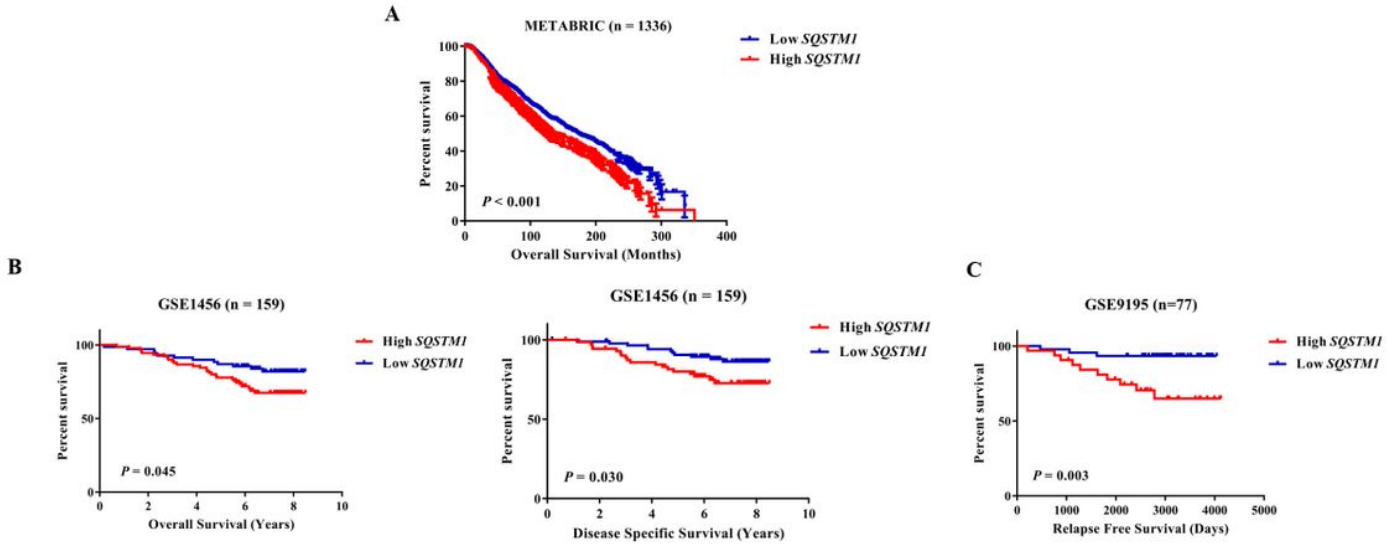

Figure 5

Kaplan-Meier survival curves comparing the high and low SQSTM1 expression in breast cancer. (A) SQSTM1 expression and OS in METABRIC dataset. (B) SQSTM1 expression and OS, DSS, RFS from GEO database.

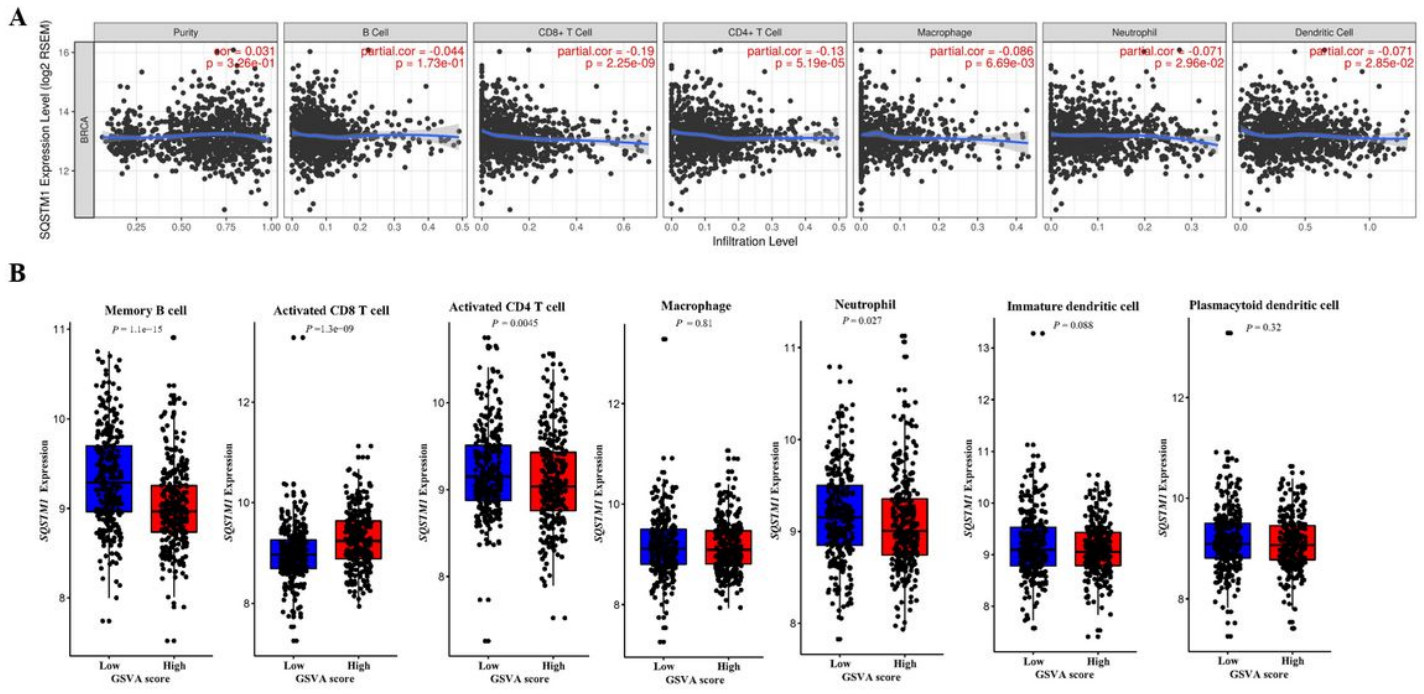

Figure 6

The role of SQSTM1 in immunity regulation of breast cancer. (A) Correlation of SQSTM1 expression with immune infiltration level in the TIMER database (B) SQSTM1 expression in breast cancer patients with lowest 25% versus patients with highest 25% GSVA score of different immune cells.

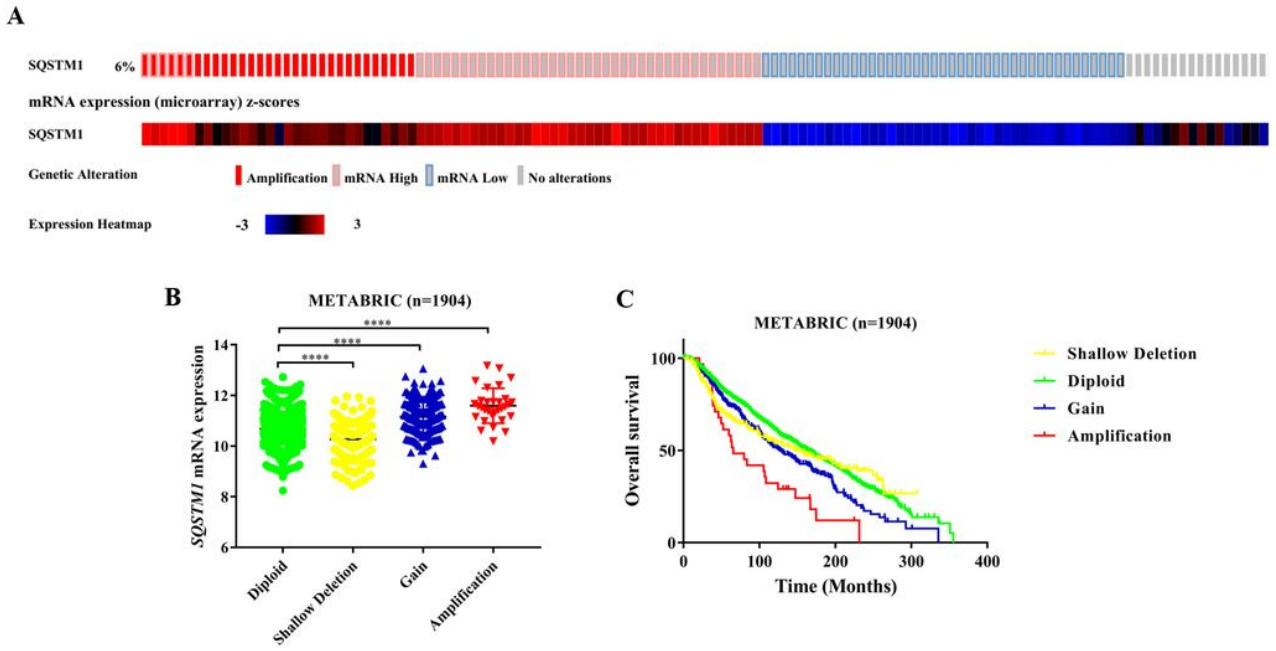

Figure 7

SQSTM1 CNV status is significantly corelated with SQSTM1 mRNA expression. (A) cBioPortal OncoPrint plot showing the distribution of SQSTM1 CNV. (B) Dot plot showing the correlation between SQSTM1 copy number values and mRNA expression values. (C) Kaplan-Meier survival curve showing SQSTM1 CNV status and OS from METABRIC database.

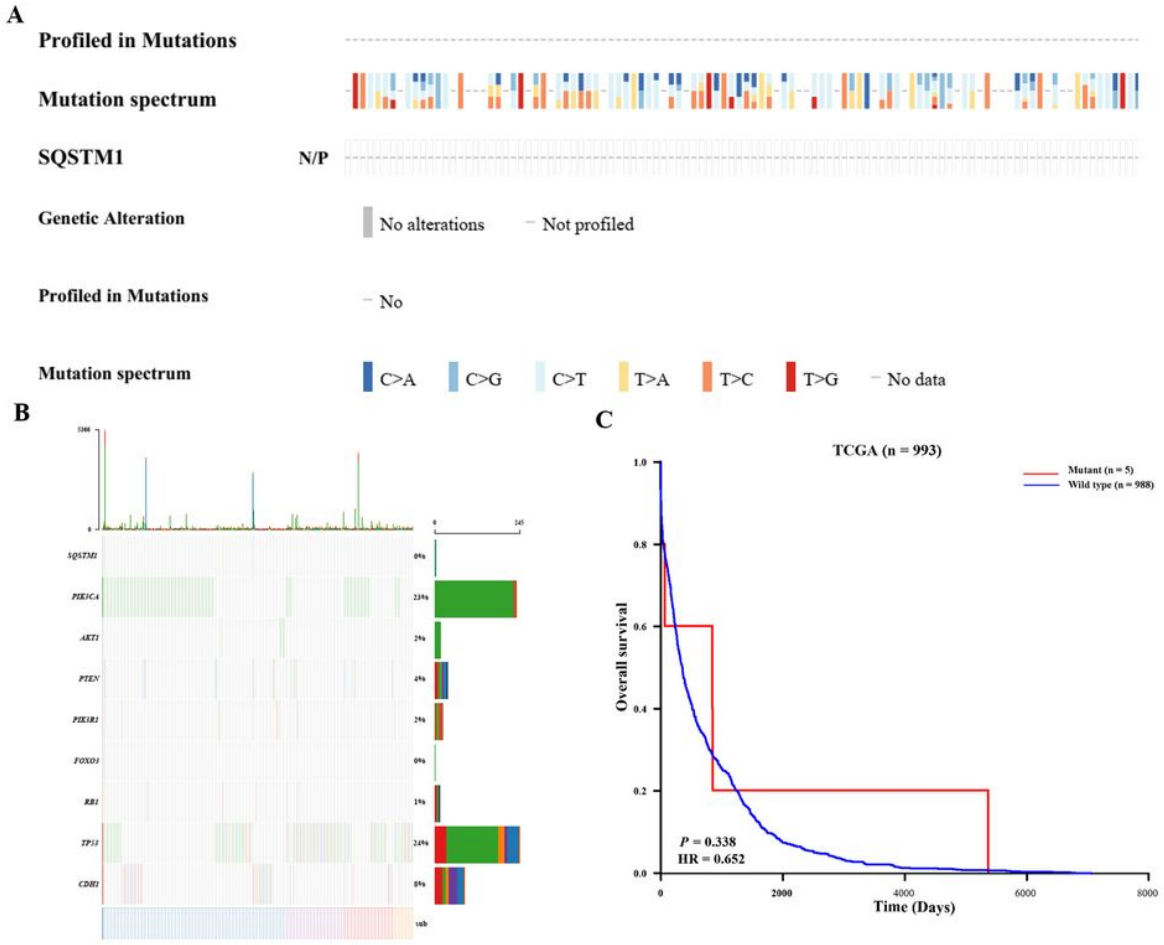

Figure 8

Genomic alterations of SQSTM1 in breast cancer. (A) cBioPortal OncoPrint plot showing the distribution of SQSTM1 mutations form METABRIC. (B) OncoPrint of SQSTM1 and PI3K, AKT1, PTEN, PIK3R1, FOXO3, RB1, TP53 and CDH1 in breast cancer from TCGA. (C) Kaplan-Meier survival curve showing SQSTM1 mutation status and OS from TCGA.

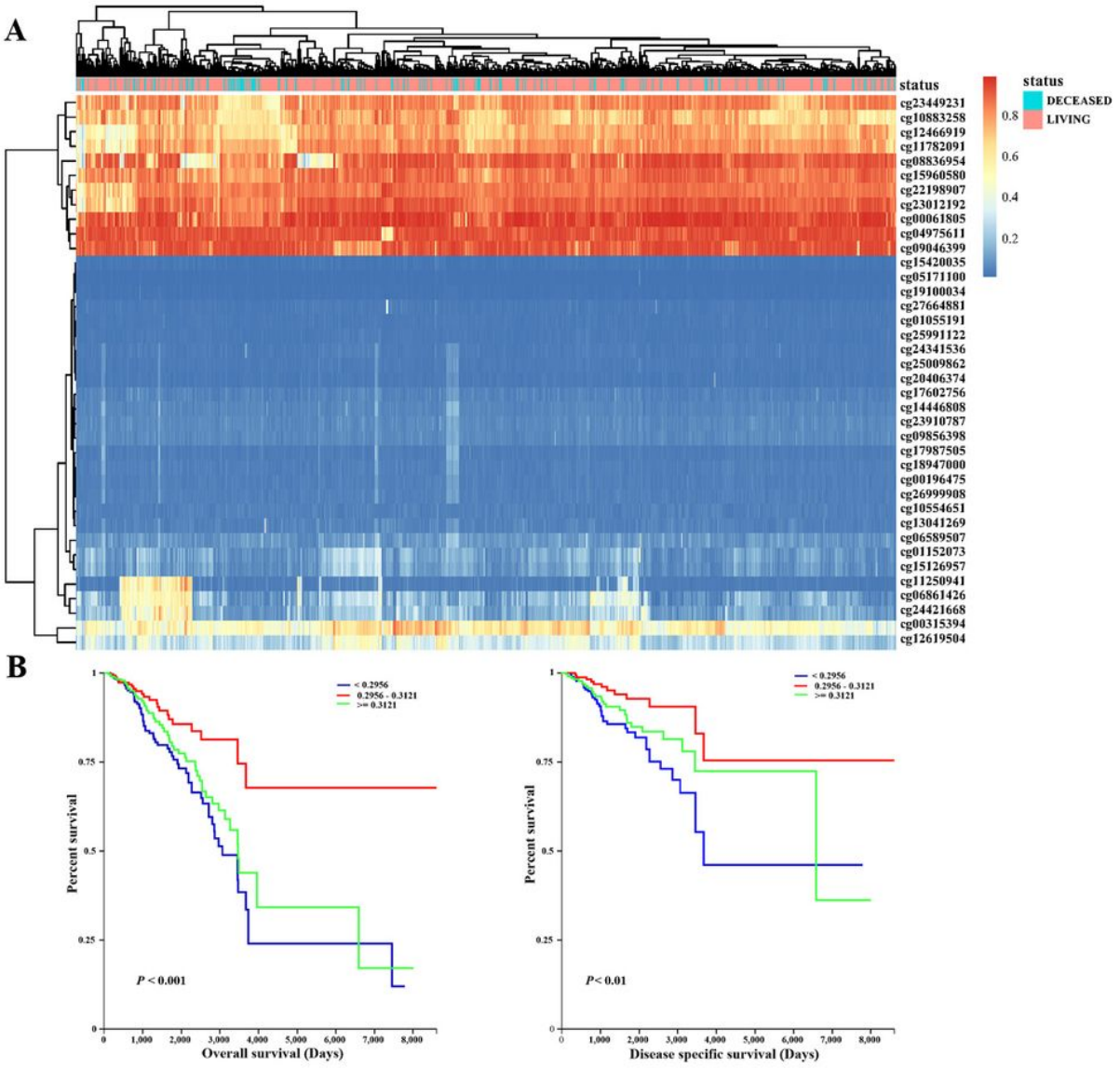

Figure 9

The Heatmap and Kaplan-Meier curves of SQSTM1 methylation. (A) The heatmap shows SQSTM1 methylation profile in the TCGA database, determined by UCSC Xena. (B) Kaplan-Meier plot shows high SQSTM1 methylation is favorable prognostic factor for breast cancer patients.

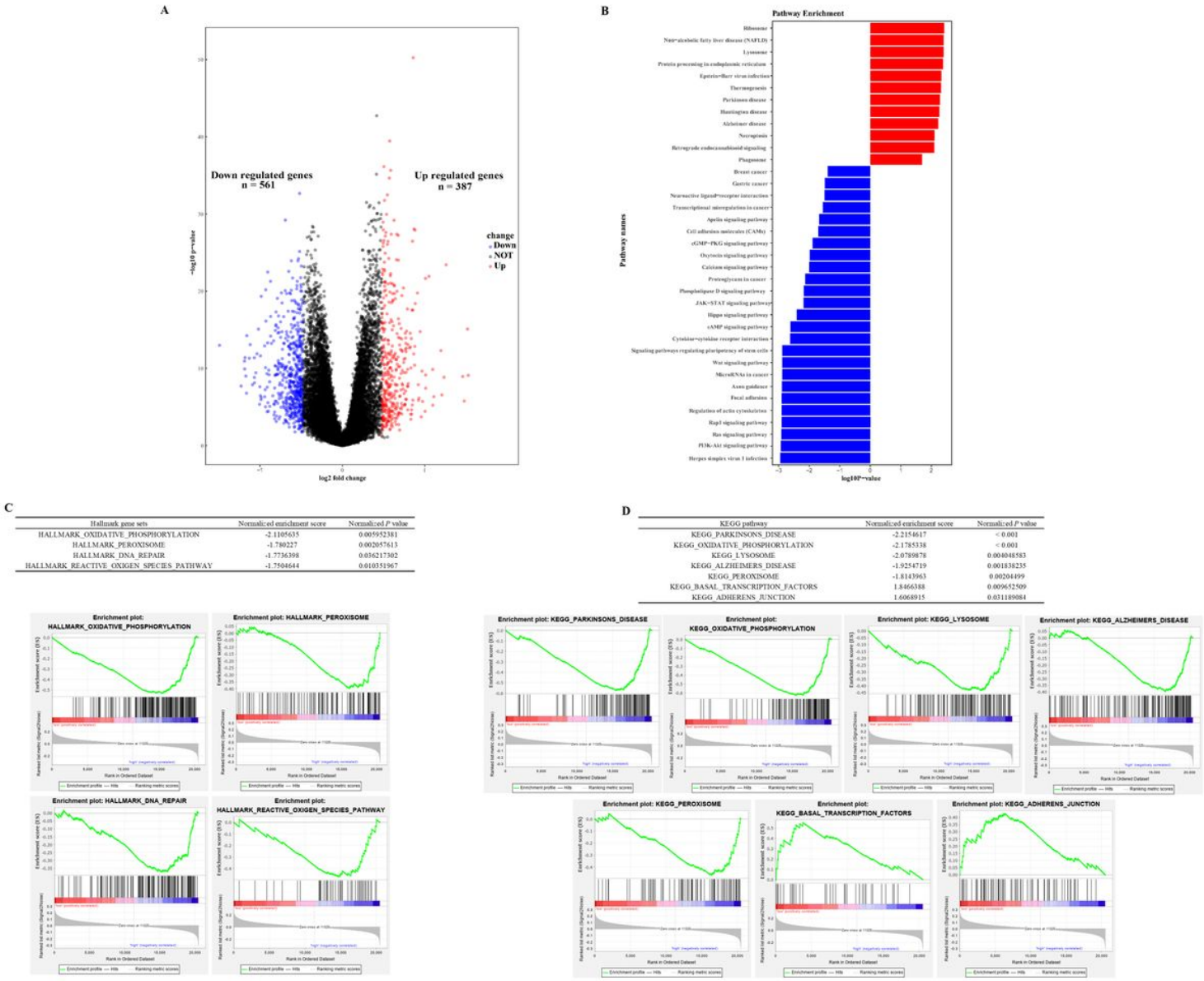

Figure 10

Gene set enrichment analysis between SQSTM1 high- and low- expression samples. (A) Volcano plot of differentially expressed genes between breast cancer samples harboring SQSTM1 high- and low-expression. The x-axis specifies the log2 fold-changes (FC) and the y-axis specifies the negative logarithm to the base 10 of P-values. Red and blue dots represent genes expressed at significantly higher or lower levels, respectively. (B) GSEA KEGG enrichment terms for up-regulated (top) and down-regulated (bottom) genes, respectively. (C) GSEA comparing SQSTM1 high- and low-expression by using hallmark gene sets. GSEA table (top) showing the top four hallmark gene sets from MSigDB. (D) GSEA comparing SQSTM1 high- and low-expression using KEGG pathway gene sets. GSEA result table (top) showing top seven KEGG pathways gene sets from MSigD.

Supplementary Files

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

- Supplementaryfiles.docx