Reduced expression of *SFRP1* is associated with poor prognosis and promotes cell proliferation in breast cancer – An integrated bioinformatics approach

Kanagaraj Suganya
Avinashilingam Institute for Home Science and Higher Education for Women
Balraj Sudha
Avinashilingam Institute for Home Science and Higher Education for Women
Arumugam Poornima
Avinashilingam Institute for Home Science and Higher Education for Women
Nachimuthu Senthil Kumar
Mizoram University
Sundaravadivelu Sumathi (sumathi_bc@avinuty.ac.in)
Avinashilingam Institute for Home Science and Higher Education for Women

Research Article

Keywords: Breast cancer, integrative analysis, hub genes, SFRP1, prognostic biomarker

Posted Date: May 27th, 2022

DOI: https://doi.org/10.21203/rs.3.rs-1700477/v1

Abstract

Breast carcinoma is the most frequent form of malignancy in women globally. Exploration of the breast cancer genome tiled the way for the validation of novel cancer biomarkers and to explore various mechanisms involved in the progression of carcinogenesis. The purpose of the research is to find an identification of potential gene linked to breast cancer (BC) progression and prognosis. Three datasets (GSE71053, GSE61724, GSE36295) were downloaded from the Gene Expression Omnibus (GEO) database. An integrated analysis of several gene expression profile datasets was used to find differentially expressed genes (DEGs) in BC and normal breast tissue samples. Protein–protein interaction (PPI) network was used to verify hub genes associated to the pathogenesis and prognosis of BC. The functional enrichment and pathway analysis was performed by FunRich, and cBioPortal. The expression pattern was assessed using COSMIC, GEPIA2 and BC-GenExMiner. The results, revealed that among the hub genes, Secreted Frizzled-related protein 1 (*SFRP1*) was a negative regulator of the Wnt pathway in breast cancer. Loss of SFRP1 may result in abnormal cellular proliferation, migration, and invasion, may trigger cancer cells, leading to progression of the disease, poor prognosis, and therapy resistance. Lastly, the Kaplan–Meier plotter online database demonstrated that expression levels of the *SFRP1* gene were related to lower survival. The findings of this research would provide some directive significance for further investigating the diagnostic and prognostic biomarker to facilitate the molecular targeting therapy of breast cancer; *SFRP1* expression may be effective as a novel prognostic biomarker in early breast cancer.

1. Introduction

Breast cancer (BC) is the most frequent cancer in women and the second leading cause of cancer death. Breast cancer is associated with several risk factors, including long-term fertility, hormonal contraceptive usage, lack of physical activity, and alcohol intake; nonetheless, its etiology and pathophysiology are not fully determined. Several genes and cellular processes are implicated in the genesis and progression of BC (Miller et al., 2019).

The molecular mechanisms underlying the development and progression of BC tumors are unclear. As a result, determining the development of disease and key signaling pathways is vital for developing more efficient diagnostic and treatment strategies. Bioinformatics analysis has been used to promote oncology research in recent years, providing a basis for better disease prevention, early detection, and therapy (Liu et al., 2019). Using bioinformatics tools, we can now screen and identify essential genes comprehensively. Identification of potential genes and pathways linked to BC carcinogenesis and disease prognosis will not only to the discovery of new diagnostic biomarkers and treatment targets but also helps in elucidating the underlying molecular mechanisms (Liu et al., 2020).

Microarrays are especially useful for detecting differentially expressed genes (DEGs) since they can detect gene expression on a global level rapidly. Gene chips are a type of microarray that allows for high-throughput gene expression studies with excellent sensitivity, selectivity, and reliability (Ni et al., 2018).

Microarrays have produced a huge amount of data, and the majority of that data has already been uploaded and preserved in publicly available databases searches. This insight could help researchers understand better the molecular pathways causing BC (Chen et al., 2021).

In this study, we tried to identify novel indicators for prognosis in BC patients as well as the prospective therapeutic targets for this disease. An integrated analysis of DEGs involved in BC will reveal more information about the BC mechanism. These findings could serve as the basis for the development of future BC diagnostic and therapeutic tools.

2. Methods

2.1 Gene expression profile data

GSE71053, GSE61724, and GSE36295 gene expression datasets were screened out based on gene expression omnibus (GEO) datasets, a public repository for data storage containing microarray data (http://www.ncbi.nlm.nih.gov/geo/). The DEGs in BC samples were compared with normal samples using the Limma tool in R language (Ritchie et al., 2015). DEGs were calculated using the following criteria: jlog2FCj ≥ 1 and adjust P value < 0.05.

2.2 Functional enrichment analysis of DEGs

FunRich was used for functional enrichment and interaction network analysis of genes and proteins to elucidate the underlying biological processes and molecular activities of DEGs (Pathan et al., 2017). Meanwhile, P-value < .05 were defined as the cut-off criteria.

2.3 Protein-protein interaction (PPI) analysis

To construct a PPI network, DEG protein products were matched to the search engine for retrieving the interacting genes database (STRING, https://stringdb.org/cgi/input.pl), using a confidence score ≥ 0.9 as the cut-off criterion. The PPI network was visualised using the Cytoscape software (Bao et al., 2020).

2.4 Pathway Analysis

Breast invasive carcinoma datasets (TCGA, PanCancer Atlas) encompassing 1084 samples were chosen from the Cancer Genomics Portal (cBioportal) (http:// www.cbioportal.org) to investigate gene alterations and activities of hub genes in breast cancer. We developed a group using cBioportal to display hub genes in the context of biological interactions derived from public pathway databases (Zhao et al., 2020).

2.5 Analysis of genetic alterations of hub genes

Somatic mutation information from COSMIC (https://cancer.sanger.ac.uk/cosmic) was utilised to examine hub gene alterations in breast cancer (Rothé et al., 2014).

2.6 Analysis of expression level and correlation analysis

The gene expression profiling interactive analysis (GEPIA, http://gepia.cancer-pku.cn/index.html) was used to analyse the hub gene's expression level and correlation. It examines tumour and normal differential expression and was used to show the expression of hub genes in BC and normal tissues. The link between these hub genes was then shown using a boxplot (Pei et al., 2020).

2.7 Assessment of clinicopathological parameters

The Breast Cancer Gene Expression Miner v4.4 (https://bcgenex.centregauducheau.fr/BC-GEM/GEM-Accueil.php?js=1), a DNA microarray and RNA-seq database was used to analyze prognosis based on gene expression. Clinicopathological parameters, such as ER, PR, HER-2 was evaluated (Liu et al., 2020).

2.8 Survival analysis

The hub gene's prognosis values were calculated using the Kaplan–Meier plotter (KM plotter, http://kmplot.com/analysis/) mRNA BC database. According to this software, the relapse-free survival (RFS) and overall survival (OS) information were based on GEO, TCGA, and EGA database. To assess the relationship between gene expression and survival, the hazard ratio (HR) with 95% confidence intervals and log-rank P value were calculated and plotted (Yang et al., 2019).

3. Results

3.1 Identification of differentially expressed genes (DEGs)

Three gene expression profiles were selected and Table 1 shows the comprehensive information about 115 breast cancer and 21 normal tissue samples in the included datasets. A total of 24 DEGs comprising 19 down-regulated and 5 up-regulated genes were retrieved after the integrated analysis of three GEO datasets. Figure 1 shows the volcano plot of the DEGs.

Table 1
Characteristics of three datasets in this study

Expression profile dataset	Platform	Number of Samples	
		Breast cancer	Normal
GSE71053	GPL570[HG-U133_Plus_2] Affymetrix Human Genome U133 Plus 2.0 Array	6	12
GSE61724	GPL6244 [HuGene-1_0-st] Affymetrix Human Gene 1.0 ST Array [transcript (gene) version]	64	4
GSE36295	GPL6244 [HuGene-1_0-st] Affymetrix Human Gene 1.0 ST Array [transcript (gene) version]	45	5

3.2 Functional enrichment analysis of DEGs

FunRich software was used to perform enrichment analysis for up-and down-regulated DEGs after gene integration. Cell division, mitotic nuclear division, kinesins, aurora B signaling, FOXM1 transcription network, signaling by aurora kinases, and M phase signaling were found to be notably abundant in up-regulated DEGs and down-regulated DEGs (Fig. 2).

3.3 PPI network construction and analysis of interrelations between pathways

The STRING database was used to construct a PPI network. A total of 23 nodes and 136 edges were mapped in the PPI network with a local clustering coefficient of 0.739 and a PPI enrichment *P*-value < $1.0e^{-16}$. Figure 3 depicts the information for the PPI network constructed in string and visualized using cytoscape. Gene ontology (GO) analysis showed that the DEGs are involved in the biological process (BPs) and cellular components such as meiotic sister chromatid cohesion and centromeric, actomyosin contractile ring assembly, centrosome separation, mitotic spindle midzome assembly, regulation of mitotic centrosome separation. Moreover, the GO Molecular function (MFs) analysis showed that the DEGs are mainly involved in ATP binding, carbohydrate derivative binding, and anion binding. The DEGs were mainly enriched in pathways such as mitotic prometaphase, M phase, cell cycle, mitotic, and resolution of sister chromatid cohesion.

3.4 Pathway analysis of DEGs generated by cBioPortal

We analyzed the influence of DEGs on biological pathways in the breast cancer dataset containing 1918 samples. Among the DEGs, *SFRP1* involved in the WNT signaling pathway is the down-regulated gene. The Wnt signaling pathway's abnormal activation is linked to the formation of solid tumours including breast cancer. Analysis of breast carcinoma revealed similar frequencies of *SFRP1* loss in breast cancer (101% respectively). Figure 4 depicts the loss of SFRP1 gene activity and its modulation of the Wnt pathway.

3.5 Mutational analysis of *SFRP1*

The *SFRP1* mutations were evaluated in 602 samples from patients with breast cancer. Out of 602 samples, the major types of mutation were found to be a missense substitution (123 samples), synonymous substitution (68 samples), nonsense substitution (10 samples), inframe deletion (10 samples) followed by frameshift deletion (2 samples) (Fig. 5 (a)).

3.6 Validation of down-regulation of *SFRP1* mRNA in breast cancer tissues in TCGA database

Using the GEPIA (Gene Expression Profiling Interactive Analysis) tool, we compared the mRNA expression of *SFRP1* between breast cancer and breast tissues. The results indicated that the expression level of *SFRP1* was down-regulated in breast cancer tissues than in normal tissues and the difference was statistically significant as shown in Fig. 5(b)

3.7 Clinicopathological relevance of *SFRP1* expression in breast cancer patients

Next, the bc-GenExMiner database was used to determine the association between *SFRP1* expression and clinicopathological variables in patients with BC. The results demonstrated that *SFRP1* mRNA expression was negatively associated with estrogen receptor (ER), progesterone receptor (PR), and human epidermal growth factor receptor-2 (HER-2) status as shown in Fig. 6.

3.8 Survival analysis and the prognostic value of *SFRP1*

To investigate the prognostic value of *SFRP1*, the survival analysis was conducted by the K–M plotter platform. Figure 7 shows the K–M survival curves for the *SFRP1* gene (HR = 0.91, P = 0.069). It was found that *the SFRP1* gene was the risky gene for prognosis with HR > 1 and P < .01. Higher expression of *SFRP1* predicts shorter survival times for BC patients.

4. Discussion

Despite advances in treatment, breast cancer remains the most common malignant tumor in women worldwide, with the highest rate of increase in prevalence. The understanding of breast cancer's molecular pathways is critical for its diagnosis, treatment, and prognosis. The use of DNA microarray gene expression profiles to investigate DEGs involved in cancer has yielded useful diagnostic and medical applications (Wang et al., 2019).

In the present study, three gene expression profile datasets (GSE71053, GSE61724, GSE36295) from the GEO database were retrieved and analyzed. The DEGs were identified using the 'limma' R package. The common DEGs were filtered out and 24 hub genes were identified. GO function and pathway enrichment analysis was performed to further analyze the mechanisms of action of these DEGs. These DEGs were associated with the GO BP terms such as cell division, mitotic nuclear division, kinesins, aurora B signaling, FOXM1 transcription network, signaling by aurora kinases, M phase, meiotic sister chromatid cohesion and centromeric, actomyosin contractile ring assembly, centrosome separation, mitotic spindle midzome assembly, regulation of mitotic centrosome separation and response to ATP binding, carbohydrate derivative binding, anion binding as molecular functions terms. Furthermore, the pathways of DEGs were mainly enriched in mitotic prometaphase, M phase, cell cycle, mitotic, and resolution of sister chromatid cohesion.

Of the 24 genes, *SFRP1* (*Secreted Frizzled Related Protein 1)* gene that is closely associated with breast cancer Wnt signaling pathway was identified. One of the most essential mechanisms controlling cell physiologic activities such as division, multiplication, and adhesion is the Wnt/β-catenin signaling pathway (Baharudin et al., 2020). Wnt ligand bind to Frizzled proteins and lipoprotein receptor-related proteins 5 and 6 receptors initiates signaling in normal circumstances. Then, as a transcription cofactor with T-cell factor/lymphoid enhancer factor, β-catenin aggregates and modulates the transcription of genes involved such as c-myc and cyclin D1 (Clemenceau et al., 2020). Abnormal activation of the Wnt/β-

catenin signaling pathway is a common occurrence in malignancy, and also the abnormal methylation state of Wnt antagonists including such Dickkopf proteins, Wnt inhibitory factor1, and SFRPs may contribute to it (Kazi et al., 2019). *SFRP1*, a member of the SFRP family, can inhibit Wnt/-catenin signaling by interfering with Wnt–receptor associations via an N-terminal cysteine-rich domain similar to Frizzled proteins. *SFRP1* is hypermethylated and down-regulated in breast cancer (Wu et al., 2020). *SFRP1* hypermethylation and downregulation are also associated with poor prognosis in breast tumors (Schäfer et al., 2019). Moreover, *SFRP1* is associated with tumor chemotherapy, and some antitumor drugs inhibit cell growth through the re-expression of *SFRP1*. However, new research has revealed that *SFRP1* may also be strongly expressed in carcinomas and enhance tumor development or migration, in breast cancer (van Schie et al., 2020).

Our study demonstrates that *SFRP1* is down-regulated in breast cancer patients which were analyzed in GEPIA and bc-GenExMiner database. Kaplan–Meier analysis showed that patient's low *SFRP1* expression had significantly poorer survival rates. These findings imply that the level of *SFRP1* expression can predict patient prognosis and could be used as a novel therapy target for personalized patient treatment. As a result, *SFRP1* may be linked to breast cancer pathogenesis and could be used as a diagnostic biomarker for the disease.

5. Conclusion

Using an integrated analysis approach of different cohort profile datasets, the current study discovered possible candidate gene *SFRP1* and the pathway involved in BC progression. These findings could contribute to a greater understanding of the biological mechanisms behind BC and the development of a possible biomarker. As a result, more research with higher patient cohorts is needed to validate the findings of this study. To characterize the precise roles of the identified gene, *in vivo* and *in vitro* examination of gene and pathway interaction is required, which might help to confirm gene functions and reveal the mechanisms behind BC.

Abbreviations

BC: Breast cancer; ER: Estrogen receptor; PR: Progesterone receptor, HER-2: Human epidermal growth factor receptor-2; GEO: Gene expression omnibus; DEG: Differentially expressed genes; PPI: Protein–protein interaction; STRING: Search tool for the retrieval of interacting genes/proteins; TCGA: The cancer genome atlas; cBioportal: Cancer genomics portal; COSMIC: Catalogue of somatic mutations in cancer; GEPIA: Gene expression profiling interactive analysis; KM Plotter: Kaplan–meier plotter; RFS: Relapse-free survival; OS: Overall survival; HR: Hazard ratio; EGA: European genome-phenome archive; GO: Gene ontology; BP: Biological process; MF: Molecular function; *SFRP1*: *Secreted Frizzled Related Protein 1*

Declarations

Ethics approval and consent to participate

Not applicable

Consent for publication

Not applicable

Availability of data and material

Datasets analysed in this study was retrieved from NCBI database (URL: https://www.ncbi.nlm.nih.gov/gds).

Competing interests

The authors declare that they have no competing interests.

Funding

Not applicable

Authors' contributions

KS: Processing and writing the original draft of the manuscript. BS: Formal analysis. AP: Methodology. NSK: Review, and editing the manuscript. SS: Conceptualization, supervision, and editing the manuscript. All authors have read and approved the manuscript.

Acknowledgements

The authors are thankful to Avinashilingam Institute for Home Science and Higher Education for Women and Mizoram University for the support throughout the study.

References

1. Miller KD, Nogueira L, Mariotto AB, Rowland JH, Yabroff KR, Alfano CM, Jemal A, Kramer JL, Siegel RL. Cancer treatment and survivorship statistics, 2019. CA: a cancer journal for clinicians. 2019; 69(5):363–85.
2. Liu F, Wu Y, Mi Y, Gu L, Sang M, Geng C. Identification of core genes and potential molecular mechanisms in breast cancer using bioinformatics analysis. Pathology-Research and Practice. 2019; 215(7):152436.
3. Liu S, Liu X, Wu J, Zhou W, Ni M, Meng Z, Jia S, Zhang J, Guo S, Lu S, Li Y. Identification of candidate biomarkers correlated with the pathogenesis and prognosis of breast cancer via integrated bioinformatics analysis. Medicine. 2020; 99(49).

4. Ni M., Liu X., Wu J., Zhang D., Tian J., Wang T. *et al*. 2018 Identification of candidate biomarkers correlated with the pathogenesis and prognosis of non-small cell lung cancer via integrated bioinformatics analysis. Front. Genet. **12**, 469–473.
5. Chen G, Yu M, Cao J, Zhao H, Dai Y, Cong Y, Qiao G. Identification of candidate biomarkers correlated with poor prognosis of breast cancer based on bioinformatics analysis. Bioengineered. 2021; 12(1):5149–61.
6. Ritchie ME, Phipson B, Wu DI, Hu Y, Law CW, Shi W, Smyth GK. limma powers differential expression analyses for RNA-sequencing and microarray studies. Nucleic acids research. 2015; 43(7):e47-.e47
7. Pathan M, Keerthikumar S, Chisanga D, Alessandro R, Ang CS, Askenase P, Batagov AO, Benito-Martin A, Camussi G, Clayton A, Collino F. A novel community driven software for functional enrichment analysis of extracellular vesicles data. Journal of extracellular vesicles. 2017; 6(1):1321455.
8. Bao S, Chen Y, Yang F, Sun C, Yang M, Li W, Huang X, Li J, Wu H, Yin Y. Screening and Identification of Key Biomarkers in Acquired Lapatinib-Resistant Breast Cancer. Frontiers in pharmacology. 2020:1389.
9. Zhao K, Wang M, Wu A. ATP6AP2 is overexpressed in breast cancer and promotes breast cancer progression. Cancer Management and Research. 2020;12:10449.
10. Rothé F, Laes JF, Lambrechts D, Smeets D, Vincent D, Maetens M, Fumagalli D, Michiels S, Drisis S, Moerman C, Detiffe JP. Plasma circulating tumor DNA as an alternative to metastatic biopsies for mutational analysis in breast cancer. Annals of oncology. 2014; 25(10):1959–65.
11. Pei J, Wang Y, Li Y. Identification of key genes controlling breast cancer stem cell characteristics via stemness indices analysis. Journal of translational medicine. 2020; 18(1):1–5.
12. Liu Q, Zhai J, Kong X, Wang X, Wang Z, Fang Y, Wang J. Comprehensive Analysis of the Expressionand Prognosis for TDO2 in Breast Cancer. Molecular Therapy-Oncolytics. 2020; 17:153–68.
13. Yang Y, Li F, Luo X, Jia B, Zhao X, Liu B, Gao R, Yang L, Wei W, He J. Identification of LCN1 as a potential biomarker for breast cancer by bioinformatic analysis. DNA and Cell Biology. 2019; 38(10):1088–99.
14. Wang CC, Li CY, Cai JH, Sheu PC, Tsai JJ, Wu MY, Li CJ, Hou MF. Identification of prognostic candidate genes in breast cancer by integrated bioinformatic analysis. Journal of clinical medicine. 2019; 8(8):1160.
15. Baharudin R, Tieng FY, Lee LH, Ab Mutalib NS. Epigenetics of SFRP1: the dual roles in human cancers. Cancers. 2020; 12(2):445.
16. Clemenceau A, Diorio C, Durocher F. Role of secreted frizzled-related protein 1 in early mammary gland tumorigenesis and its regulation in breast microenvironment. Cells. 2020; 9(1):208.
17. Kazi MM, Vora HH, Gajjar KK, Kobawala TP, Ghosh NR. Frequent methylation of genes encoding wnt pathway antagonists: Secreted frizzled-related protein 1 and dickkopf 3 in invasive breast cancer. Clinical Cancer Investigation Journal. 2019; 8(3):106.

18. Wu ZH, Zhang YJ, Yue JX, Zhou T. Comprehensive Analysis of the Expression and Prognosis for SFRPs in Breast Carcinoma. Cell Transplantation. 2020.
19. Schäfer SA, Hülsewig C, Barth P, von Wahlde MK, Tio J, Kolberg HC, Bernemann C, Blohmer JU, Kiesel L, Kolberg-Liedtke C. Correlation between SFRP1 expression and clinicopathological parameters in patients with triple-negative breast cancer. Future oncology. 2019; 15(16):1921–38.
20. Van Schie EH, Van Amerongen R. Aberrant WNT/CTNNB1 signaling as a therapeutic target in human breast Cancer: weighing the evidence. Frontiers in cell and Developmental Biology. 2020:25.

Figures

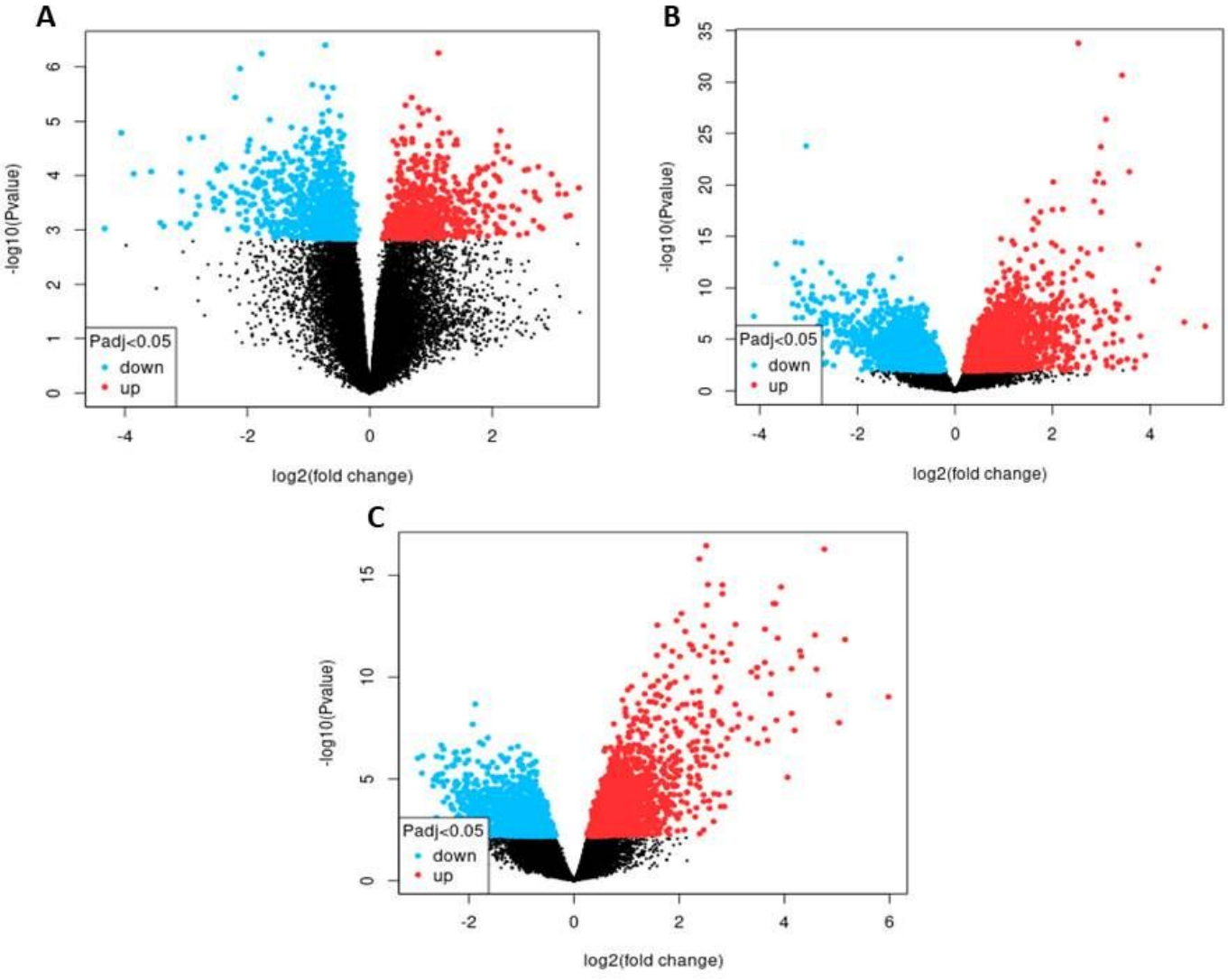

Figure 1

Volcano plot of gene expression profile data in breast cancer samples and normal tissues of DEGs. (a) Volcano plot of GSE71053; (b) Volcano plot of GSE61724; (c) Volcano plot of GSE36295

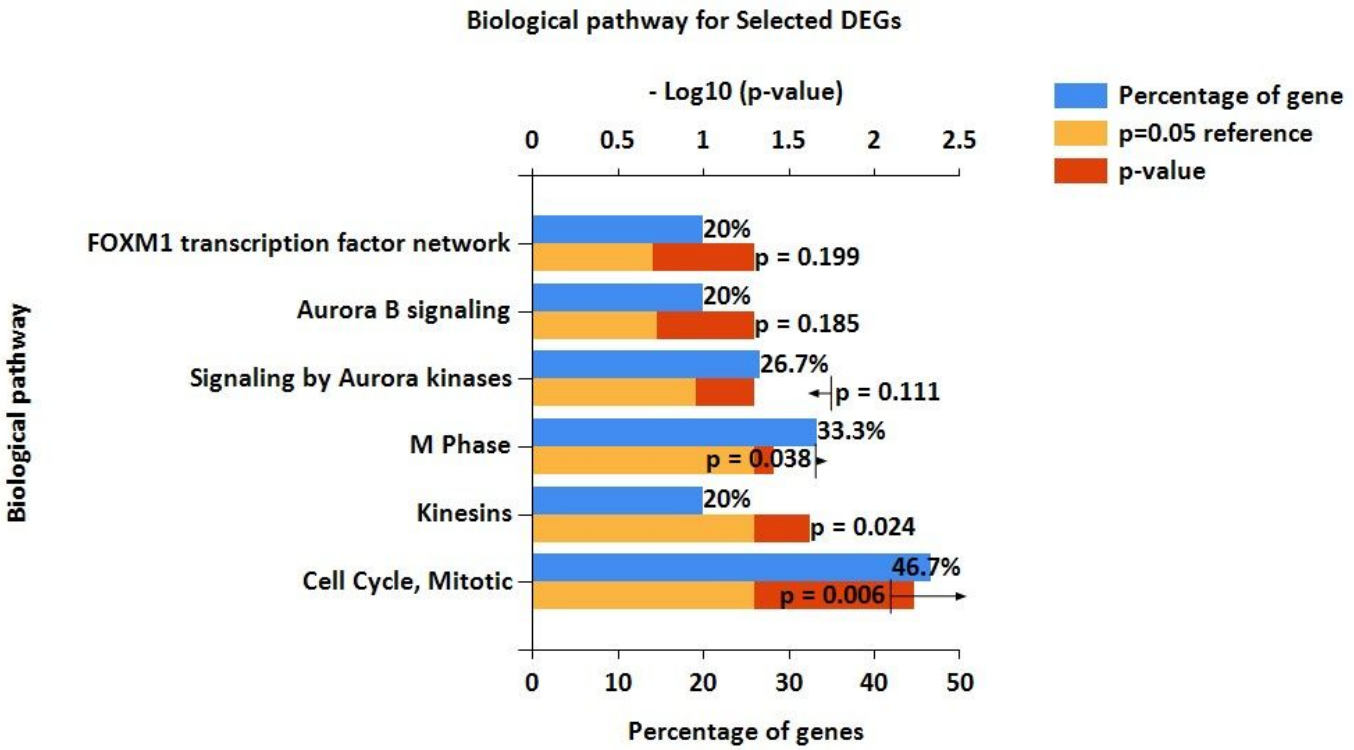

Figure 2

Biological pathway for differentially expressed genes

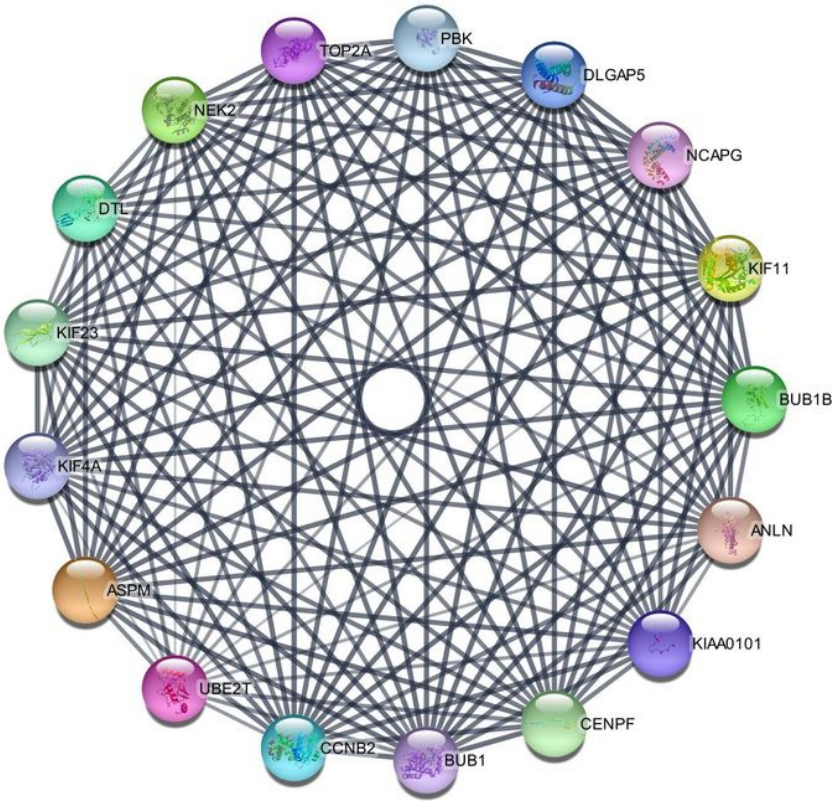

Figure 3

PPI network analysis of DEGs

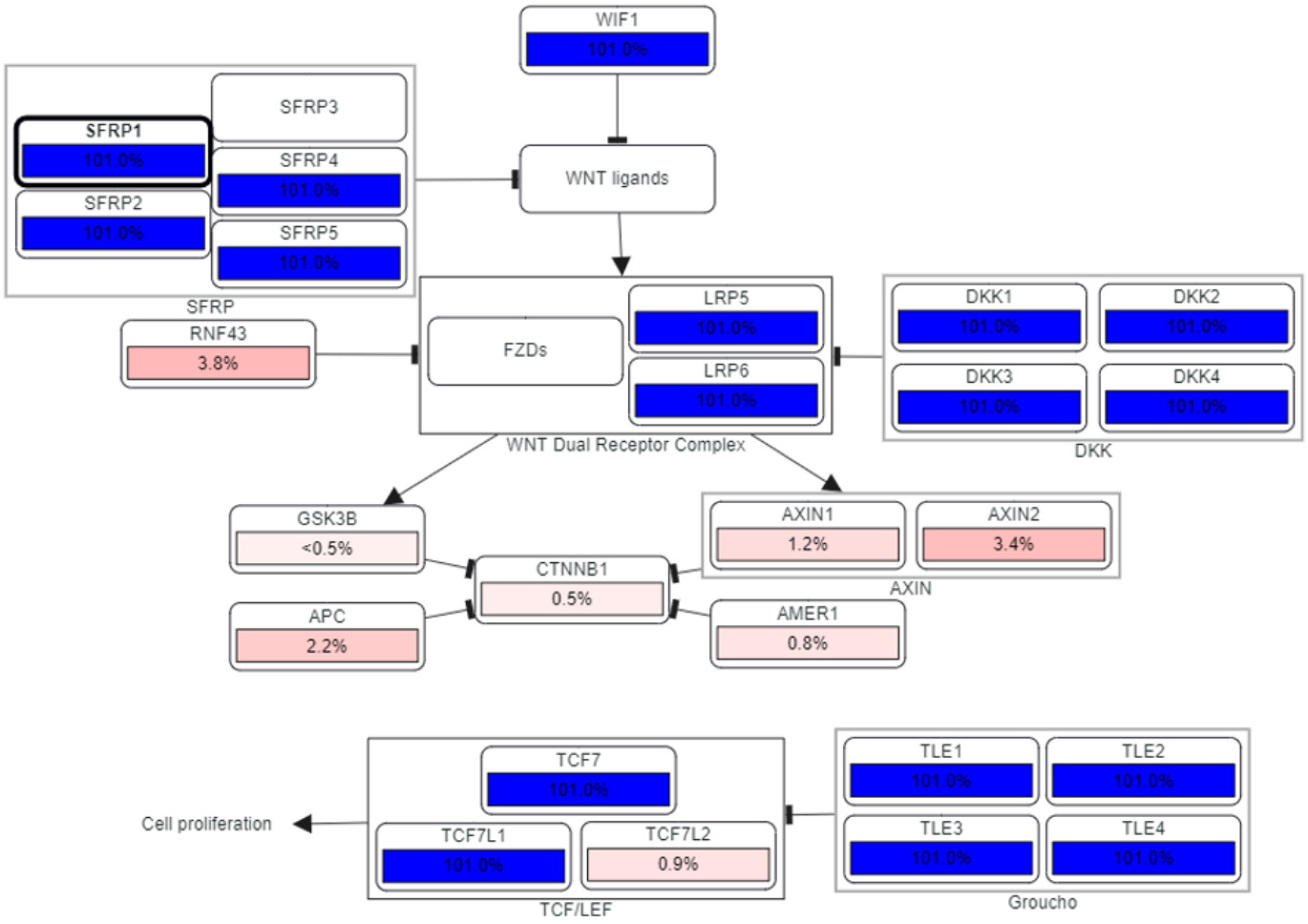

Figure 4

The Wnt pathway analysis and the loss of *SFRP1* lead to cell proliferation in breast cancer

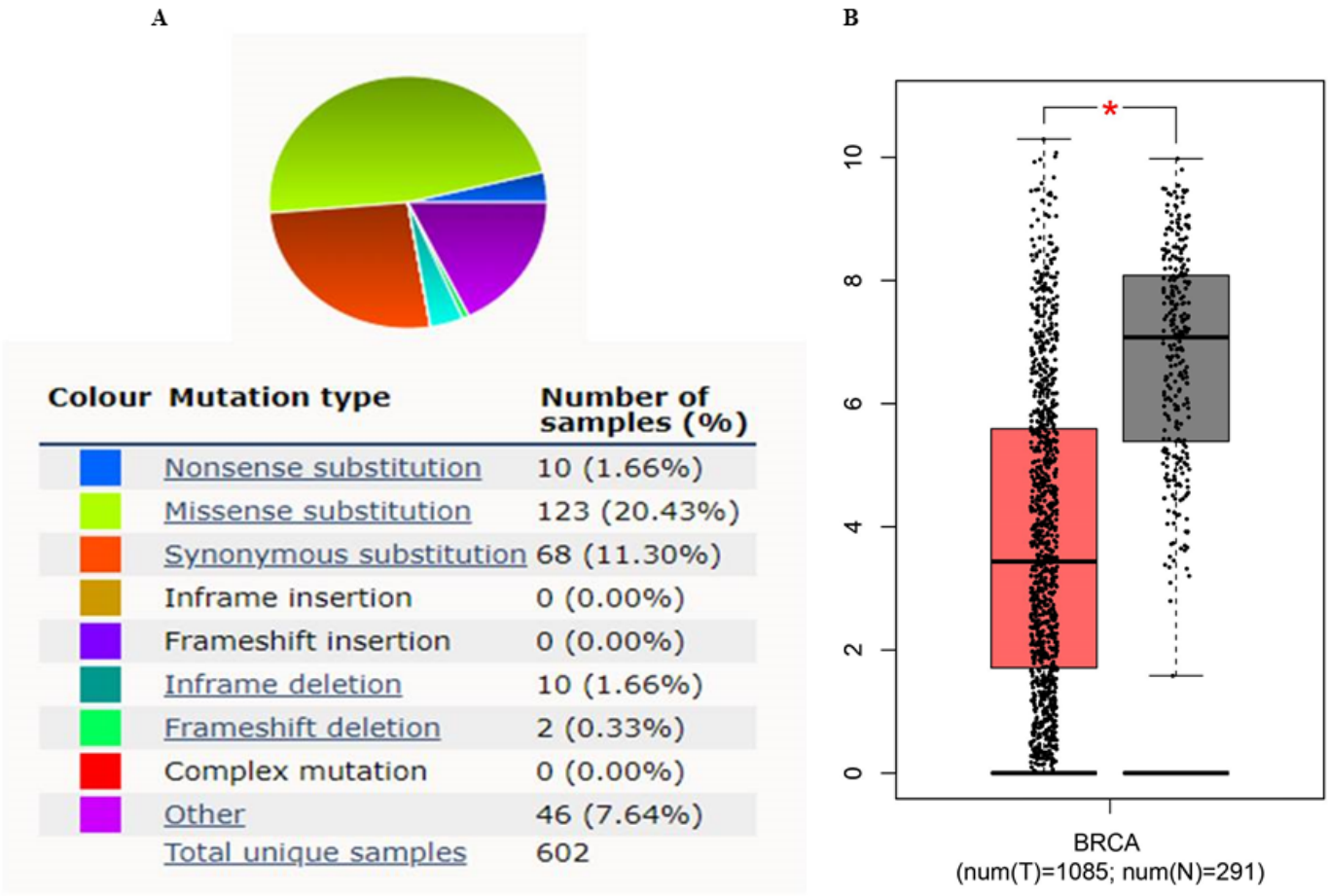

Figure 5

Mutation and expression analysis of *SFRP1*. (a) Breast cancer mutations observed in *SFRP1* using COSMIC database; (b) Expression of *SFRP1* in breast cancer patients. Down-regulated expression of *SFRP1* in breast cancer patients when compared with adjacent normal tissues

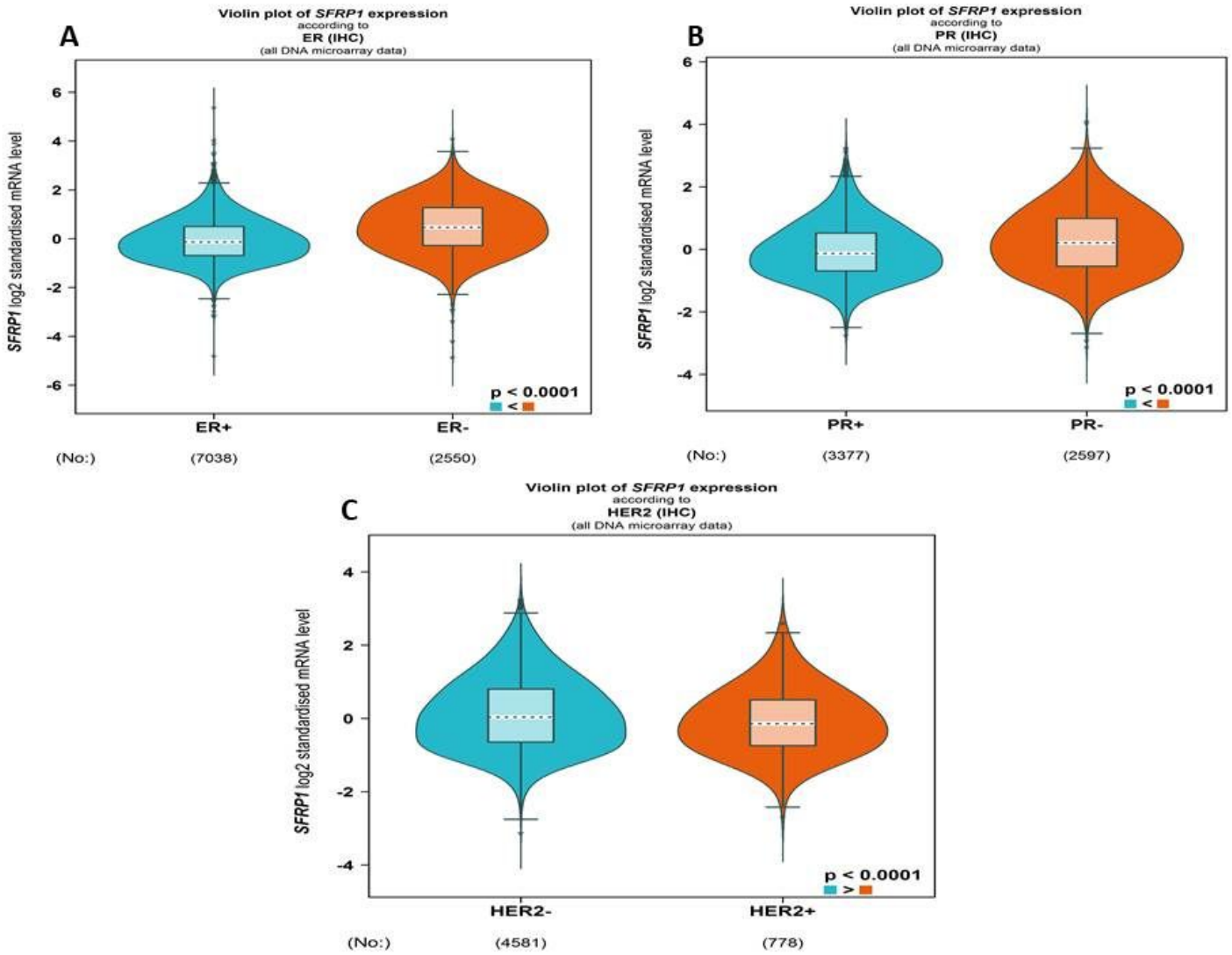

Figure 6

Violin plot depicting *SFRP1* expression among groups of patients expressing receptor status. (a) ER; (b) PR; (c) HER-2 receptors

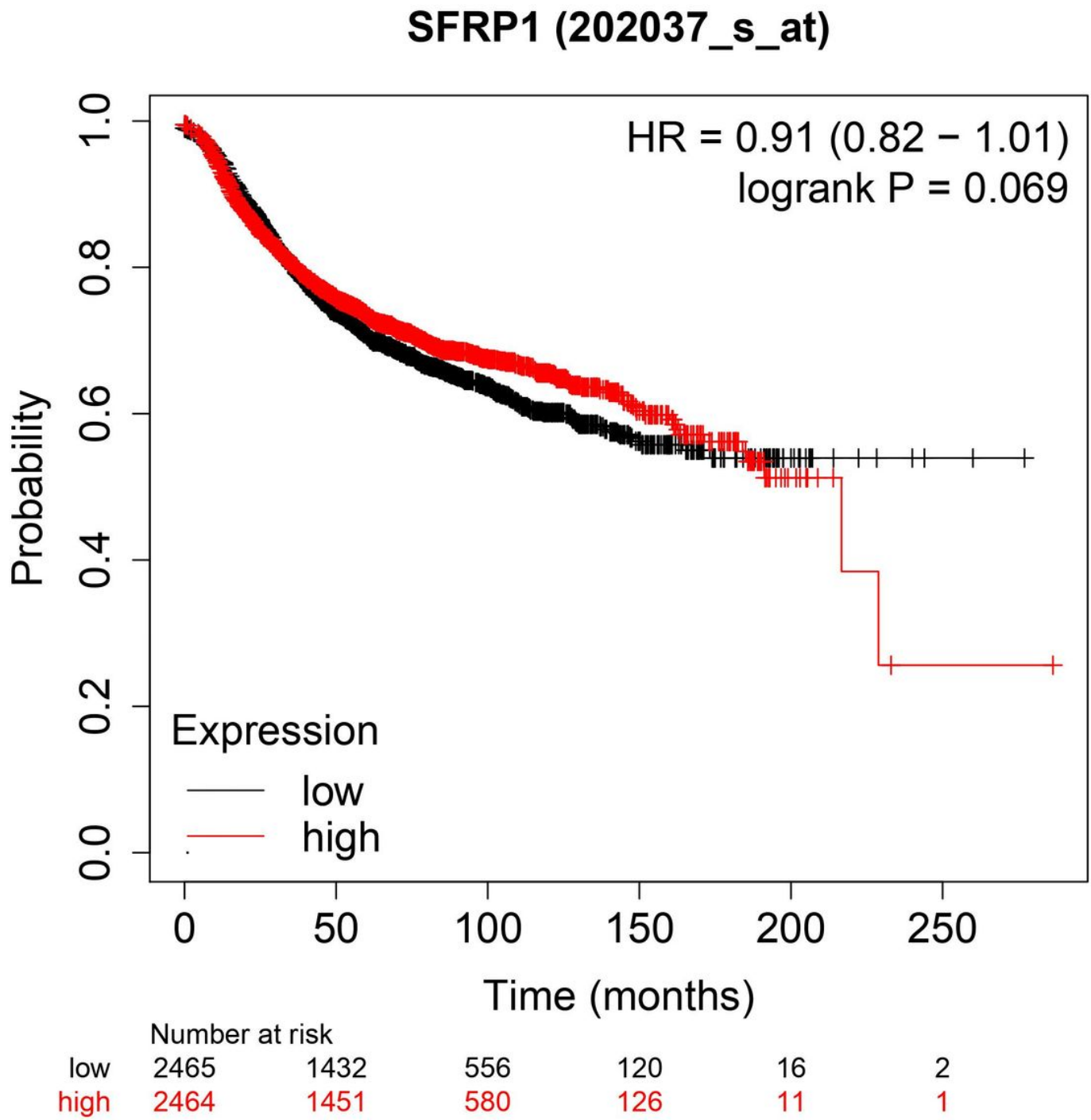

Figure 7

Survival curves of *SFRP1*