Identifying liver metastasis-related hub genes in breast cancer and characterizing *SPARCL1* as a potential prognostic biomarker

4 Mingkuan Chen[1, *], Wenfang Zheng[1, *], Lin Fang[1]

6 [1] Department of Thyroid and Breast Division of General Surgery Shanghai Tenth People's
7 Hospital, Tongji University School of Medicine, Shanghai, China.

9 *These authors contributed equally to this work.

11 Corresponding Author:
12 Lin Fang[1]
13 No. 301 YanChang Middle Road, Jing'an District, Shanghai, 200072, China
14 Email address: fanglin2017@126.com

Abstract

Background: The liver is the third most common metastatic site for advanced breast cancer (BC), and liver metastases predict poor prognoses. However, the characteristic biomarkers of BC liver metastases and the biological role of secreted protein acidic and rich in cysteine-like 1 (*SPARCL1*) in BC remain unclear. The present study aimed to identify potential biomarkers for liver metastasis of BC and to investigate the effect of *SPARCL1* on BC.

Methods: The publicly available GSE124648 dataset was used to identify differentially expressed genes (DEGs) between BC and liver metastases. Gene Ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) enrichment analyses were conducted to annotate these DEGs and understand the biological functions in which they are involved. A protein–protein interaction (PPI) network was constructed to identify metastasis-related hub genes and further validated in a second independent dataset (GSE58708). Clinicopathological correlation of hub gene expression in patients with BC was determined. Gene set enrichment analysis (GSEA) was performed to explore DEG-related signaling pathways. *SPARCL1* expression in BC tissues and cell lines was verified by RT-qPCR. Further *in vitro* experiments were performed to investigate the biological functions of *SPARCL1* in BC cells.

Results: We identified 332 liver metastasis-related DEGs from GSE124648 and 30 hub genes, including *SPARCL1*, from the PPI network. GO and KEGG enrichment analyses of liver-metastasis-related DEGs revealed several enriched terms associated with the extracellular matrix and pathways in cancer. Clinicopathological correlation analysis of *SPARCL1* revealed that its expression in BC was associated with age, TNM stage, estrogen receptor status, progesterone receptor status, histological type, molecular type, and living status of patients. GSEA results suggested that low *SPARCL1* expression in BC was related to the cell cycle, DNA replication, oxidative phosphorylation, and homologous recombination. Lower expression levels of *SPARCL1* were detected in BC tissues compared to adjacent tissues. The *in vitro* experiments showed that *SPARCL1* knockdown significantly increased the proliferation and migration of BC cells, whereas the proliferation and migration were suppressed after elevating the expression of *SPARCL1*.

Conclusion: We identified *SPARCL1* as a tumor suppressor in BC, which shows potential as a target for BC and liver metastasis therapy and diagnosis.

Introduction

Breast cancer (BC) in women has now become the most newly diagnosed malignant tumor worldwide, with both the number of new cases and number of deaths ranking first among all malignant tumors(Sung et al. 2021). Generally, patients with BC have a relatively good prognosis, with approximately 44% of patients with early-stage BC having a nearly 100% 5-year survival rate, whereas once organ metastasis occurs, this survival rate drops sharply to 26% (Miller et al. 2019). The organs most likely to metastasize from BC are the lungs, bone, liver, and brain (Cummings et al. 2014). Metastasis of the liver can cause various fatal complications, including liver failure, intractable ascites, portal vein thrombosis, and malnutrition (Diamond et al. 2009). Thus, occurrence of liver metastases indicates a worse prognosis in patients with BC,

55 with a reported median survival time of approximately 3 years (Zhao et al. 2018a), However, the
56 molecular mechanism of metastasis remains unclear. Identifying potential molecular biomarkers
57 of liver metastasis may provide more accurate information to guide clinical decisions and predict
58 prognosis, as well as provide direction and theoretical support for future research on the
59 mechanism of metastasis. Therefore, identifying DEGs with prognostic significance in BC liver
60 metastases, determining their biological function and evaluating their potential as therapeutic
61 targets remain essential.
62 The Gene Expression Omnibus (GEO) and the Cancer Genome Atlas (TCGA) databases have
63 collected and stored a large amount of tumor sequencing data, which are publicly available. Re-
64 analyzing these sequencing data using bioinformatics methods and mining for differences in
65 genetic information among different samples can help provide a scientific explanation of the
66 occurrence and progression of diseases (Gauthier et al. 2019). Therefore, based on public
67 database and bioinformatics analyses, we screened differentially expressed genes (DEGs)
68 between BC and liver metastasis, and identified hub genes.
69 *SPARCL1*, a member of the secreted protein acidic and rich in cysteine (SPARC) family, is an
70 extracellular matrix (ECM) glycoprotein with a gene located at 4q22. *SPARCL1* has been
71 reported to be widely expressed in normal brain, lung, heart, colon, and muscle tissues(Hu et al.
72 2012a). It is associated with the regulation of cell proliferation and migration through the
73 regulation of embryogenesis, matrix remodeling and tumorigenesis(Liu et al. 2021). Many
74 studies have reported its oncogenic role in the development of various malignancies. However, it
75 has been less studied in BC, especially in BC liver metastasis, highlighting the need to further
76 investigate the role of *SPARCL1* in BC and liver metastasis.
77 In the present study, we first screened the BC liver metastasis-related genes via integrated
78 bioinformatics analysis and identified *SPARCL1* as a key gene. Then the expression of *SPARCL1*
79 and its correlation with the clinical characteristics of BC patients were analyzed. In addition, the
80 biological function of *SPARCL1* in BC was evaluated *in vitro.* In summary, we aimed to identify
81 potential biomarkers of BC liver metastasis and assess the expression pattern and role of
82 *SPARCL1* in BC. These findings are expected to provide a new understanding of BC and liver
83 metastasis, driving future research and providing potential biomarkers for BC diagnosis and
84 therapy.

85 Materials & Methods

86 Discovery and validation datasets

87 In the GEO (http://www.ncbi.nlm.nih.gov/geo/), datasets GSE124648 (Sinn et al. 2019) and
88 GSE58708 (McBryan et al. 2015) were obtained with the following keywords: breast neoplasms,
89 breast cancer, liver metastasis, expression profiling by array, attribute name tissue, and *Homo*
90 *sapiens.* Array data of liver metastases ($N = 16$) and primary tumors (control, $N = 130$) from the
91 GPL96 (HG-U133A; Affymetrix Human Genome U133A Array) platform were selected for
92 differential expression analysis. The array data for GSE58708 consisted of three patients with
93 BC liver metastasis *vs.* three controls as a validation dataset (Table 1).

94 DEG identification

95 R software (version 4.0.3) was employed for bioinformatics analysis. The "*limma*" package was
96 adopted to identify DEGs between BC and liver metastasis (Ritchie et al. 2015), with the
97 following cut-off criteria: |log$_2$FC|>2.0 and adj.*P*.Val<0.05. Similarly, the R package
98 "*TCGAbiolinks*" was used to obtain biological data for BC in the TCGA database (Colaprico et
99 al. 2016), and hub gene expression was verified by analyzing the differentially expressed
100 mRNAs in BC. "*ggplot2*" and "*heatmap*" packages were used for visualizing the DEGs.

101 **Enrichment analysis of DEGs**
102 The "*clusterProfile*" package was employed to conduct gene ontology (GO) and Kyoto
103 Encyclopedia of Genes and Genomes (KEGG) enrichment analyses of the DEGs (Yu et al.
104 2012), and a gene set enrichment analysis (GSEA) method was used for the KEGG enrichment
105 analysis. *p*<0.05 was the criteria set for statistical significance.

106 **Integration of PPI networks and identification of hub genes**
107 The STRING (http://www.string-db.org/) database provides comprehensive information on
108 protein–protein interactions (PPI) (Szklarczyk et al. 2021). A PPI network of DEGs was
109 integrated using the STRING database. Then, the results were visualized and analyzed using
110 Cytoscape (version 3.7.2). The important modules were extracted via the plug-in MCODE, with
111 the cut-off criteria being an MCODE score>5 and nodes>5. The other default parameters were
112 set as Max. Depth=100, K-Core=2, Node score cut-off=0.2, Degree Cut-off=2. The plug-in
113 cytoHubba was employed to identify the top 30 hub genes in the PPI network, with the following
114 cut-off criteria: degree.layout≥25 (MCC algorithm) (Shannon et al. 2003). GEPIA
115 (http://gepia.cancer-pku.cn/) was used to obtain the hub gene expression between BC and normal
116 tissues, with the following default parameters: *p*≤0.01 and |Log$_2$FC|≥1 (Tang et al. 2017). The
117 Kaplan–Meier plotter (https://kmplot.com/analysis/) was used to evaluate the correlation
118 between the expression of the hub genes and patient survival (Lánczky & Győrffy 2021). The
119 median expression was used to divide the patients into high and low expression groups. Then,
120 the Kaplan–Meier curves of overall survival (OS) and recurrence-free survival (RFS) were
121 drawn. Finally, hub genes (*SPARCL1* and *SERPINA1*) with prognostic values were identified.
122 *p*<0.05 using the log-rank test was considered statistically significant.

123 **Validation of hub genes and clinicopathological correlation analysis**
124 A box plot was plotted to verify *SPARCL1* expression in liver metastasis based on the GSE58708
125 dataset. The correlations between *SPARCL1* expression and clinicopathological characteristics of
126 patients with BC were analyzed using TCGA-BRCA clinical data. Then, based on patient
127 survival data and *SPARCL1* levels, the optimal expression threshold was calculated using the
128 "*survminer*" package. The correlation between gene expression and clinicopathological
129 parameters was tested using Pearson's chi-squared test. *p*<0.05 was considered statistically
130 significant.

131 **GSEA**
132 GSEA can be used to evaluate whether a predefined gene set shows statistically significant
133 differences between two groups with different phenotypes. Expression (.gct) and phenotype (.cls)
134 information files for *SPARCL1* were uploaded to the GSEA software (version 4.1.0) to conduct

135 enrichment analysis. The chip platform and gene set database selected were

136 "Human_ENSEMBL_Gene_ID_MSigDB.v7.0" and "c2.cp.kegg.v7.1. symbols.gmt,"

137 respectively. The normalized enrichment score was calculated. Both false discovery rate (FDR q-

138 val) and nominal p-value (NOM p-val) <0.05 were considered statistically significant.

139 **Clinical specimens and cell lines**

140 Tissue samples were obtained from patients with BC who underwent radical mastectomy at

141 Shanghai Tenth People's Hospital (Shanghai, China). A total of 30 pairs of tumor and matched

142 normal adjacent breast tissues were collected and preserved in liquid nitrogen. All clinical

143 samples were obtained with the written informed consent of participants, and the project was

144 approved by the Ethics Committee of Shanghai Tenth People's Hospital (No. 2020-KN174-01).

145 BC cells (MDA-MB-231, BT549, and MCF-7) and mammary epithelial cells (MCF-10A) were

146 purchased from the Cell Bank of Type Culture Collection of Chinese Academy and grown in

147 mammary epithelial basal medium (Cambrex, East Rutherford, NJ, USA) and Gibco Dulbecco's

148 modified Eagle medium (DMEM; Thermo Fisher Scientific, Waltham, MA, USA) supplemented

149 with 10% fetal bovine serum (FBS; Gibco) and 1% penicillin-streptomycin (Enpromise,

150 Shanghai, China). All cells were maintained in a 5% CO_2 incubator at 37° C.

151 **Cell transfection, RNA extraction, and RT-qPCR**

152 Lentiviral vector pLKO.1 and lentiviral packaging kit (Genomeditech, Shanghai, China) was

153 used for construct stable *SPARCL1* knockdown cells. sh-NC and shRNAs targeting *SPARCL1*

154 (*SPARCL1* sh-1 and *SPARCL1* sh-2) were purchased from IBSbio (Shanghai, China) and the

155 sequences were as follows: *SPARCL1* sh-1: 5′- CCCACAATGATAACCAAGAAA-3′,

156 *SPARCL1* sh-2 (sense: 5′- GCAGAGAAATAAAGTCAAGAA-3′. The cells transduced by

157 lentivirus were selected with 2 μg/ml puromycin for 3 days. The pcDNA3.1 vector was used to

158 construct *SPARCL1* overexpression plasmid (IBSbio, Shanghai, China). Plasmids were

159 transfected into BC cells using Invitrogen Lipofectamine 3000 (Thermo Fisher Scientific).

160 A lentiviral short hairpin RNA (shRNA) construct allows for stable and long-term knockdown of

161 the targeted gene. Invitrogen TRIzol reagent (Thermo Fisher Scientific) was used to extract total

162 RNA from cell lines and tissues. HiScript III RT SuperMix for qPCR (Vazyme, Nanjing, China)

163 was used to generate the cDNAs, and Hieff qPCR SYBR Green Master Mix (Yeasen, Shanghai,

164 China) was used to conduct real-time quantitative-polymerase chain reaction (RT-qPCR)

165 following the manufacturer's protocol, with *ACTB* as internal control for normalizing *SPARCL1*

166 expression. We used the following primers to conduct RT-qPCR. *SPARCL1* (forward: 5′-

167 CCAACTGAAGGTACATTGGACAT-3′, reverse: 5′-CTGTGAAGGAACTAACACCAGG-3′)

168 and *ACTB* (forward: 5′-CATGTACGTTGCTATCCAGGC-3′, reverse: 5′-

169 CTCCTTAATGTCACGCACGAT-3′).

170 **MTT, colony formation, wound healing, and Transwell assays**

171 To evaluate the effects of *SPARCL1* on BC cells, post-transfection BC cell lines were cultured in

172 96-well plates at a density of 1,500 cells. Next, 20 μL methylthiazolyldiphenyl-tetrazolium

173 bromide (MTT; Yeasen) was added to each well at 0, 24, 48, 72, and 96 h after inoculation, and

174 the samples were incubated for 4 h at 37 °C in a 5% CO_2 incubator to assess cell viability. Then,

175   150 µL DMSO was added to each well after removing the supernatant. The absorbance was
176   measured using a microplate spectrophotometer (BioTek Instruments, Winooski, VT, USA) at
177   490 nm. Cell proliferation curves were plotted according to the absorbance value.
178   Transfected BC cell lines were prepared as single-cell suspensions, inoculated at a density of 750
179   cells/well until prominent colonies were formed. The cells were fixed with 95% ethanol and
180   stained with 0.1% crystal violet (Yeasen) to detect cell colony formation ability. Representative
181   pictures were recorded, and clone colonies were counted.
182   To detect cell mobility, the cell monolayer was scratched with RNA free tips when the
183   transfected BC cell density reached 90% or more. In the following step, DMEM supplemented
184   with 2% FBS was used as the culture medium. The healing of scratches was observed at 0 and 12
185   h using the same field of view to calculate cell mobility.
186   The transfected BC cells and 500 µL of DMEM containing 10% FBS were added to Transwell's
187   upper and lower chambers (Corning, Corning, NY, USA), respectively. After culturing for 18 h,
188   migrated cells were fixed with 4% paraformaldehyde and stained with 0.1% crystal violet to
189   assess the migratory ability. Representative images were captured using an inverted microscope.

190   **Statistical analysis**
191   Wilcoxon matched-pairs signed-rank test was used to compare the expression of *SPARCL1*
192   between BC and control samples. Unpaired Student's *t*-test was used to compare the expression
193   of *SPARCL1* between MCF-10A and BC cell lines. The results of the MTT assay were analyzed
194   using a two-way analysis of variance. All experiments were repeated three times. The
195   experimental data were analyzed and plotted using Prism v8.3.0 (GraphPad Software, San Diego,
196   CA, USA). $p < 0.05$ was considered statistically significant.

197   # Results

198   **DEG identification**
199   We identified 332 DEGs in the GSE124648 dataset, of which 116 and 216 were upregulated and
200   downregulated genes, respectively (Fig. 1A). A heatmap was used to visualize the top 50 DEGs,
201   as well as the top 50 liver metastasis-related genes in the dataset (Fig. 1B). Among the top 50
202   DEGs, 60% showed significant down expression in liver metastasis compared to BC tissues.
203   Another 40% of highly expressed DEGs showed higher expression compared to BC tissues.
204   These results suggest that downregulated DEGs in BC liver metastasis probably play a more
205   critical role than upregulated DEGs.

206   **GO and KEGG enrichment analysis of DEGs**
207   We further performed functional and pathway enrichment analyses on the 332 DEGs. GO
208   annotation divides gene function into three categories: cellular components (CC), molecular
209   function (MF), and biological process (BP) (Sinn et al. 2019). The top eight enriched GO terms
210   for each category are shown in Figure 2. CC analysis indicated that these DEGs were particularly
211   related to the extracellular matrix, endoplasmic reticulum lumen, collagen-containing
212   extracellular matrix, collagen trimer, and blood microparticles (Fig. 2A). MF analysis showed
213   that the DEGs were mainly involved in extracellular matrix structural constituents,
214   glycosaminoglycan binding, extracellular matrix structural constituents conferring tensile

215 strength, heparin-binding, and collagen binding (Fig. 2B). The BP category was mainly enriched
216 in extracellular structure organization, extracellular matrix organization, wound healing, humoral
217 immune response, and complement activation (Fig. 2C). The KEGG pathway enrichment
218 analysis results further showed that the upregulated DEGs were significantly enriched in
219 neutrophil extracellular trap formation, alcoholic liver disease, and neuroactive ligand−receptor
220 interaction, whereas the downregulated DEGs were significantly enriched in malignancy-related
221 pathways, including pathways in cancer, BC, PI3K−Akt signaling pathway, MAPK signaling
222 pathway, and focal adhesion (Fig. 2D). The above results suggest that these DEGs were mainly
223 enriched in entries associated with extracellular matrix remodeling, and remarkably,
224 downregulated DEGs were significantly enriched in various signaling pathways involved in the
225 regulation of malignant tumor pathogenesis.
226 **PPI network and *SPARCL1* identified as a prognostic-related hub gene**
227 The PPI network contained 332 nodes and 2,972 edges (Fig. S1), and we selected the top three
228 hub modules identified by the MCODE plug-in for display (Fig. 3A−C), where each node
229 represents one DEG and the edges between nodes represent interactions. The top 30 hub genes
230 with a high degree of connectivity were identified from the network (Fig. 3D). The expression of
231 these genes and their modules are listed in Table 2. Among the 30 hub genes, *SERPINA1* and
232 *VCAN* were found to be upregulated in BC, whereas *ALB*, *IGFBP3*, *SPARCL1*, and *FSTL1* were
233 downregulated and no significant difference was discovered in the expression of other hub genes.
234 Survival analyses showed favorable OS and RFS in patients with BC with upregulated
235 *SERPINA1* and *SPARCL1* expression (Fig. 4). However, it was paradoxical that *SERPINA1*
236 showed high expression in both breast cancer and liver metastasis, but predicted a better patient
237 prognosis. Therefore, we selected *SPARCL1* for further evaluation, as the above results also
238 suggested that it was likely to be a diagnostic and prognostic biomarker for BC.
239 **Validation of *SPARCL1* and clinicopathological correlation analysis**
240 To affirm the accuracy of the above results, we externally validated the independent dataset
241 GSE58708. Following differential expression analysis of the GSE58708 dataset, 683 DEGs were
242 obtained, including 410 upregulated and 273 downregulated genes (Fig. S2). *SPARCL1*
243 expression was also significantly downregulated in liver metastasis compared with that in BC
244 tissues ($\log_2$FC=−2.617; Fig. S3A). In the TCGA−BRCA dataset, *SPARCL1* expression in BC
245 tissues was also significantly downregulated compared with that in normal mammary gland
246 tissue (Fig. S3B). The correlation between *SPARCL1* levels and the clinical characteristics of
247 patients with BC was further analyzed and showed that the lower *SPARCL1* expression was
248 significantly related to age, tumor-node-metastasis (TNM) stage, estrogen receptor (ER) status,
249 progesterone receptor (PR) status, histological type, molecular type, and living status, whereas
250 no significant difference in node stage and Her-2 status was observed (Table 3). The above
251 findings suggest that *SPARCL1* expression was downregulated both in BC and liver metastasis,
252 and its low expression correlates with younger age, high TNM stages, and HR negative status of
253 patients, implying that *SPARCL1* may play a tumor suppressive role in BC patients.
254 **GSEA**

255 From the viewpoint of enrichment of gene sets, finding the effects of subtle changes on
256 biological pathways or functions is easier in theory (Subramanian et al. 2005). Therefore, we
257 performed GSEA based on the different expression levels of *SPARCL1* in BC. The results
258 suggested that 50 gene sets were upregulated in the low expression *SPARCL1* phenotype, 47
259 gene sets were significantly enriched at FDR < 25%, and 40 gene sets were significantly
260 enriched at nominal *p*-value < 5%. Figure 5 displays six signaling pathways (DNA replication,
261 cell cycle, oxidative phosphorylation, homologous recombination, spliceosome, and proteasome)
262 that were significantly enriched when *SPARCL1* was downregulated in BC, revealing the
263 potential molecular mechanisms by which *SPARCL1* participates in BC occurrence and
264 progression.
265 **SPARCL1 downregulation in BC tissues and cells**
266 Total RNA was extracted from tissues and cell lines for RT-qPCR to validate *SPARCL1*
267 expression in BC. In comparison with the paired adjacent normal breast tissues, *SPARCL1* was
268 significantly downregulated in BC tissues (*N*=30; Fig. 6A). In comparison with the normal breast
269 epithelial cell line, *SPARCL1* expression was also decreased in the BC cell lines (Fig. 6B). These
270 results are consistent with those from our bioinformatics analysis, suggesting that *SPARCL1* is
271 indeed expressed significantly less in BC compared to in normal breast tissue.
272 **SPARCL1 knockdown-induces proliferation and migration of BC cells *in vitro***
273 To explore the biological function of *SPARCL1*, we used shRNAs to knockdown *SPARCL1* in
274 BC cells. The results suggest that shRNAs could significantly downregulate the level of
275 *SPARCL1* in BC cells compared to sh-NC (Fig. 7A). The results of colony formation and MTT
276 assays suggested that *SPARCL1* inhibition promoted the colony formation and proliferation of
277 BC cells (Fig. 7B and C). Cell migration is an essential step in tumor progression and metastasis.
278 In comparison with the control group (sh-NC), the healing ability of *SPARCL1*-knockdown
279 MDA-MB-231 cells was significantly enhanced (Fig. 7D) and cell migration was significantly
280 increased (Fig. 7E). Collectively, these results suggested that downregulation of *SPARCL1*
281 promoted the colony formation, proliferation, and migration of BC cells.
282 **Overexpression of *SPARCL1* inhibited the proliferation and migration of BC *in vitro*.**
283 To further evaluate the effects of overexpression of *SPARCL1* on the biological function of BC
284 cell lines, *SPARCL1* overexpression plasmid was transfected into BC cells. The results of RT-
285 qPCR showed that the expression of *SPARCL1* was significantly upregulated in the *SPARCL1*
286 OE group compared to the control (vector) group (Fig.8A). MTT assay revealed that the
287 proliferation of BC cells was significantly decreased after the overexpression of
288 *SPARCL1*(Fig.8C), and the results of the colony formation assay showed that overexpression of
289 *SPARCL1* inhibited the colony formation of BC cells (Fig.8B). The results of the wound-healing
290 assay showed that there was a significant decrease in cell migration area after *SPARCL1*
291 overexpression compared to the control group (Fig.8D). The results of Transwell assay showed
292 that the number of cells in the lower chamber was significantly reduced after *SPARCL1*
293 overexpression (Fig.8E). These findings demonstrated that the colony formation, proliferation,
294 and migration of BC cells were inhibited after elevating *SPARCL1* expression.

Discussion

295
296 Bioinformatics analysis of sequencing data can improve our understanding of gene function,
297 including gene expression patterns between different experimental conditions or phenotypes,
298 identification of biological processes related to gene expression, and screening of therapeutic
299 targets and prognostic markers (Hu et al. 2015; Kaifi et al. 2015; Tao et al. 2017). In this study,
300 332 DEGs associated with liver metastasis were identified by differential expression analysis of
301 sequencing data from publicly available datasets. GO enrichment analysis suggested that these
302 DEGs were mostly located in the extracellular matrix and participated in biological processes,
303 including extracellular matrix organization, wound healing, angiogenesis, and humoral immune
304 response. Extracellular matrix organization is closely correlated with the occurrence and
305 progression of cancer; a neatly arranged matrix can promote tumor cell invasion. The tumor
306 suppressor *PTEN* participates in the regulation of matrix remodeling, which is negatively
307 correlated with the arrangement of collagen in human breast tissue (Jones et al. 2019). Tumor
308 angiogenesis can supply nutrients and oxygen essential for tumor growth and metastasis (Weis &
309 Cheresh 2011). Hypoxia-inducible factor-dependent angiogenesis is vital for the invasion,
310 progression, and drug resistance of BC and is closely correlated with its poor prognosis (de Heer
311 et al. 2020). KEGG pathway analysis showed that downregulated DEGs were particularly
312 enriched in malignant tumor-related pathways, such as MAPK signaling, PI3K/AKT signaling,
313 focal adhesion, Wnt signaling, and Hippo signaling. The MAPK signaling pathway is crucial for
314 BC invasion and metastasis, promoting the occurrence and progression of the disease (Cotrim et
315 al. 2013; Jiang et al. 2020; Ke et al. 2021). In BC, more than 70% patients have PI3K signaling
316 pathway alterations and its activation is significantly associated with an HR-negative, basal-like
317 phenotype, high histological grade, and cancer-specific death (López-Knowles et al. 2010). The
318 PI3K/AKT signaling pathway also participated in liver metastasis of various malignant tumors.
319 Indeed, microRNA-582 can promote gastric cancer liver metastasis via the PI3K/AKT/Snail
320 pathway mediated by FOXO-3 (Xie et al. 2020). The c-Met/PI3K/AKT/mTOR axis can activate
321 the liver metastasis-specific cholesterol metabolism pathway in colorectal cancer, providing
322 conditions for tumor cell colonization and growth in the liver (Zhang et al. 2021a). There are also
323 reports on the influence of Wnt, Hippo, and other signaling pathways on the liver metastasis of
324 malignant tumors (Chai et al. 2019; Yuan et al. 2019). Based on the above results, we suggested
325 that these DEGs, especially those which are downregulated, alter the tumor microenvironment
326 primarily through regulating the extracellular matrix, angiogenesis, and humoral immunity, and
327 are involved in the mechanism of BC liver metastasis through multiple signaling pathways,
328 which is moreover a multi-gene co-regulation process rather than a single independent gene or
329 product.
330 Based on the prognostic value of hub DEGs in BC, we identified *SPARCL1*, a member of the
331 secreted protein acidic and rich in cysteine (*SPARC*) family, which is downregulated in both BC
332 and liver metastasis. *SPARCL1* is a newly discovered player in tumors and is mainly associated
333 with physiological processes, such as cell migration, adhesion, and cell proliferation regulation
334 (Gagliardi et al. 2017). *SPARCL1* has been found to be under-expressed in a variety of

335   malignancies and is closely associated with tumorigenesis and metastasis. The expression of

336   *SPARCL1* is downregulated in colorectal cancer and is related to tumor differentiation, stage,

337   distant metastasis, and OS (Hu et al. 2012b; Zhang et al. 2022). *SPARCL1* is also downregulated

338   in prostate cancer, being associated with disease progression, especially in invasive prostate

339   cancer. Moreover, *SPARCL1* can inhibit migration, invasion, and metastasis of prostate cancer

340   (Hurley et al. 2012; Xiang et al. 2013). Similarly, *SPARCL1* was found to be a tumor suppressor

341   in gastric cancer (Li et al. 2012), osteosarcoma (Zhao et al. 2018b), pancreatic cancer (Esposito

342   et al. 2007), lung cancer (Deng et al. 2021), and renal cell carcinoma (Ye et al. 2017). *SPARCL1*

343   expression in colorectal cancer(CRC) liver metastasis is downregulated, with *SPARCL1* being

344   considered the hub gene of liver metastasis and a biomarker with a significant prognostic value

345   (Zhang et al. 2021b). The possible mechanisms underlying the role of *SPARCL1* in liver

346   metastasis include the halt of various activities such as platelet activation driving the acquisition

347   of a metastatic signature, participation in the IGFBP-IGF signaling pathway, and promoting the

348   secretion of granules (endocytosis and exocytosis)(Zhang et al. 2021c). The highly consistent

349   expression pattern of *SPARCL1* in BC and liver metastasis in CRC may suggest an analogous

350   mechanism for *SPARCL1* in BC liver metastasis. Nevertheless, the detail of the mechanism

351   needs to be explored and validated in the future. Furthermore, downregulation of *SPARCL1*

352   expression enhances liver metastasis of malignant gastrointestinal stromal tumor cells (Shen et

353   al. 2018). However, few studies have reported on the expression and function of *SPARCL1* in

354   BC. Consistent with our findings, *SPARCL1* has been reported to be downregulated in BC (Cao

355   et al. 2013). To the best of our knowledge, the expression pattern and role of *SPARCL1* in the

356   liver metastasis of BC have not been reported previously. Our research identified *SPARCL1* as a

357   hub gene for BC liver metastasis and showed low expression in metastasis tissues. In addition,

358   our clinicopathological correlation analysis revealed that the low expression of *SPARCL1* is

359   related to age, TNM stage, ER status, PR status, histological type, molecular type, and survival

360   status of patients with BC, while younger age, higher TNM stage, HR-negative status, and basal-

361   like type were among the unfavorable phenotypes in patients with BC. However, different

362   *SAPRCL1* expression levels were not prognostically significant in individual breast cancer

363   subtypes. We suppose that *SPARCL1* may primarily influence survival through its involvement

364   in determining breast cancer subtypes.

365   GSEA data indicated that pathways such as DNA replication, cell cycle, oxidative

366   phosphorylation, and homologous recombination were significantly enriched when *SPARCL1*

367   was downregulated, revealing potential molecular mechanisms by which *SPARCL1* participates

368   in BC. The most vulnerable cellular process in oncogenic lesions is DNA replication, with

369   oncogene-induced replication stress being a fundamental step and early driver of tumorigenesis

370   (Kotsantis et al. 2018). Cell cycle and DNA replication mechanisms are closely coordinated to

371   ensure correct single-genome replication during cell division, thereby avoiding the occurrence of

372   diseases such as cancer (Dai et al. 2021). Homologous recombination repair is an essential

373   pathway of DNA damage repair, and homologous recombination deficiency is considered an

374 important biomarker and risk factor for BC (den Brok et al. 2017; Shen et al. 2020; Telli et al.
375 2018).

376 The *in vitro* inhibition of BC cell proliferation, migration, and invasion by *SPARCL1* indicates
377 the gene acts as a tumor suppressor in BC. However, the specific mechanisms by which
378 *SPARCL1* inhibits the proliferation and migration of BC cells warrants additional investigations,
379 and the role of *SPARCL1* in BC and liver metastasis needs to be further confirmed. Future
380 experiments are therefore needed to illuminate these critical details. In addition, the
381 transcriptome data in this study were taken from a public database; to obtain more accurate and
382 credible results, the expression and role of hub genes such as *SPARCL1* require further
383 experimental verification using liver metastasis tissues and cell lines, such as tissue microarray
384 technology, animal experiments, etc.

Conclusions

386 In summary, comprehensive screening of DEGs and hub genes revealed that *SPARCL1,* a tumor
387 suppressor in BC, has the potential to become a diagnostic biomarker and therapeutic target of
388 BC and liver metastasis.

Acknowledgements

390 A preprint of this manuscript has previously been published on Research Square
391 (https://www.researchsquare.com/article/rs-2183292/v1)(Mingkuan et al. 2022). We would like
392 to thank Editage (www.editage.cn) for English language editing.

Data Availability

394 The datasets generated during and/or analyzed during the current study are publicly available at
395 Gene Expression Omnibus database (https://www.ncbi.nlm.nih.gov/geo/). The raw
396 measurements are available in the Supplementary File.

References

398 <CA A Cancer J Clinicians - 2021 - Sung - Global Cancer Statistics 2020 GLOBOCAN
399   Estimates of Incidence and Mortality.pdf>.https://doi.org/10.3322/caac.21660
400 Cao F, Wang K, Zhu R, Hu Y-W, Fang W-Z, and Ding H-Z. 2013. Clinicopathological
401   significance of reduced SPARCL1 expression in human breast cancer. *Asian Pacific*
402   *journal of cancer prevention : APJCP* 14:195-200
403 Chai W-X, Sun L-G, Dai F-H, Shao H-S, Zheng N-G, and Cai H-Y. 2019. Inhibition of PRRX2
404   suppressed colon cancer liver metastasis via inactivation of Wnt/β-catenin signaling
405   pathway. *Pathology, Research and Practice*
406   215:152593.https://doi.org/10.1016/j.prp.2019.152593
407 Colaprico A, Silva TC, Olsen C, Garofano L, Cava C, Garolini D, Sabedot TS, Malta TM,
408   Pagnotta SM, Castiglioni I, Ceccarelli M, Bontempi G, and Noushmehr H. 2016.
409   TCGAbiolinks: an R/Bioconductor package for integrative analysis of TCGA data.
410   *Nucleic acids research* 44:e71.https://doi.org/10.1093/nar/gkv1507
411 Cotrim CZ, Fabris V, Doria ML, Lindberg K, Gustafsson JÅ, Amado F, Lanari C, and Helguero
412   LA. 2013. Estrogen receptor beta growth-inhibitory effects are repressed through
413   activation of MAPK and PI3K signalling in mammary epithelial and breast cancer cells.
414   *Oncogene* 32:2390-2402.https://doi.org/10.1038/onc.2012.261
415 Cummings MC, Simpson PT, Reid LE, Jayanthan J, Skerman J, Song S, McCart Reed AE,
416   Kutasovic JR, Morey AL, Marquart L, O'Rourke P, and Lakhani SR. 2014. Metastatic

417   progression of breast cancer: insights from 50 years of autopsies. *The Journal of*
418   *pathology* 232:23-31.https://doi.org/10.1002/path.4288
419   Dai M, Boudreault J, Wang N, Poulet S, Daliah G, Yan G, Moamer A, Burgos SA, Sabri S, Ali S,
420   and Lebrun J-J. 2021. Differential Regulation of Cancer Progression by CDK4/6 Plays a
421   Central Role in DNA Replication and Repair Pathways. *Cancer Res* 81:1332-
422   1346.https://doi.org/10.1158/0008-5472.CAN-20-2121
423   de Heer EC, Jalving M, and Harris AL. 2020. HIFs, angiogenesis, and metabolism: elusive
424   enemies in breast cancer. *The Journal of clinical investigation* 130:5074-
425   5087.https://doi.org/10.1172/JCI137552
426   den Brok WD, Schrader KA, Sun S, Tinker AV, Zhao EY, Aparicio S, and Gelmon KA. 2017.
427   Homologous Recombination Deficiency in Breast Cancer: A Clinical Review. *JCO*
428   *precision oncology* 1.https://doi.org/10.1200/PO.16.00031
429   Deng H, Hang Q, Shen D, Zhang Y, and Chen M. 2021. Low expression of CHRDL1 and
430   SPARCL1 predicts poor prognosis of lung adenocarcinoma based on comprehensive
431   analysis and immunohistochemical validation. *Cancer cell international*
432   21:259.https://doi.org/10.1186/s12935-021-01933-9
433   Diamond JR, Finlayson CA, and Borges VF. 2009. Hepatic complications of breast cancer.
434   *Lancet Oncol* 10:615-621.https://doi.org/10.1016/S1470-2045(09)70029-4
435   Esposito I, Kayed H, Keleg S, Giese T, Sage EH, Schirmacher P, Friess H, and Kleeff J. 2007.
436   Tumor-suppressor function of SPARC-like protein 1/Hevin in pancreatic cancer.
437   *Neoplasia (New York, NY)* 9
438   Gagliardi F, Narayanan A, and Mortini P. 2017. SPARCL1 a novel player in cancer biology.
439   *Critical reviews in oncology/hematology* 109:63-
440   68.https://doi.org/10.1016/j.critrevonc.2016.11.013
441   Gauthier J, Vincent AT, Charette SJ, and Derome N. 2019. A brief history of bioinformatics.
442   *Briefings in bioinformatics* 20:1981-1996.https://doi.org/10.1093/bib/bby063
443   Hu H, Zhang H, Ge W, Liu X, Loera S, Chu P, Chen H, Peng J, Zhou L, Yu S, Yuan Y, Zhang S,
444   Lai L, Yen Y, and Zheng S. 2012a. Secreted protein acidic and rich in cysteines-like 1
445   suppresses aggressiveness and predicts better survival in colorectal cancers. *Clinical*
446   *cancer research : an official journal of the American Association for Cancer Research*
447   18:5438-5448.https://doi.org/10.1158/1078-0432.Ccr-12-0124
448   Hu H, Zhang H, Ge W, Liu X, Loera S, Chu P, Chen H, Peng J, Zhou L, Yu S, Yuan Y, Zhang S,
449   Lai L, Yen Y, and Zheng S. 2012b. Secreted protein acidic and rich in cysteines-like 1
450   suppresses aggressiveness and predicts better survival in colorectal cancers. *Clinical*
451   *cancer research : an official journal of the American Association for Cancer Research*
452   18:5438-5448.https://doi.org/10.1158/1078-0432.CCR-12-0124
453   Hu N, Wang C, Clifford RJ, Yang HH, Su H, Wang L, Wang Y, Xu Y, Tang Z-Z, Ding T, Zhang
454   T, Goldstein AM, Giffen C, Lee MP, and Taylor PR. 2015. Integrative genomics analysis
455   of genes with biallelic loss and its relation to the expression of mRNA and micro-RNA in
456   esophageal squamous cell carcinoma. *BMC genomics*
457   16:732.https://doi.org/10.1186/s12864-015-1919-0
458   Hurley PJ, Marchionni L, Simons BW, Ross AE, Peskoe SB, Miller RM, Erho N, Vergara IA,
459   Ghadessi M, Huang Z, Gurel B, Park BH, Davicioni E, Jenkins RB, Platz EA, Berman
460   DM, and Schaeffer EM. 2012. Secreted protein, acidic and rich in cysteine-like 1
461   (SPARCL1) is down regulated in aggressive prostate cancers and is prognostic for poor
462   clinical outcome. *Proceedings of the National Academy of Sciences of the United States*
463   *of America* 109:14977-14982.https://doi.org/10.1073/pnas.1203525109
464   Jiang W, Wang X, Zhang C, Xue L, and Yang L. 2020. Expression and clinical significance of
465   MAPK and EGFR in triple-negative breast cancer. *Oncol Lett* 19:1842-
466   1848.https://doi.org/10.3892/ol.2020.11274

467 Jones CE, Hammer AM, Cho Y, Sizemore GM, Cukierman E, Yee LD, Ghadiali SN, Ostrowski
468   MC, and Leight JL. 2019. Stromal PTEN Regulates Extracellular Matrix Organization in
469   the Mammary Gland. *Neoplasia (New York, NY)* 21:132-
470   145.https://doi.org/10.1016/j.neo.2018.10.010
471 Kaifi JT, Kunkel M, Das A, Harouaka RA, Dicker DT, Li G, Zhu J, Clawson GA, Yang Z, Reed
472   MF, Gusani NJ, Kimchi ET, Staveley-O'Carroll KF, Zheng S-Y, and El-Deiry WS. 2015.
473   Circulating tumor cell isolation during resection of colorectal cancer lung and liver
474   metastases: a prospective trial with different detection techniques. *Cancer biology &*
475   *therapy* 16:699-708.https://doi.org/10.1080/15384047.2015.1030556
476 Ke J, Han W, Meng F, Guo F, Wang Y, and Wang L. 2021. CTI-2 Inhibits Metastasis and
477   Epithelial-Mesenchymal Transition of Breast Cancer Cells by Modulating MAPK
478   Signaling Pathway. *International Journal of Molecular Sciences*
479   22.https://doi.org/10.3390/ijms222212229
480 Kotsantis P, Petermann E, and Boulton SJ. 2018. Mechanisms of Oncogene-Induced
481   Replication Stress: Jigsaw Falling into Place. *Cancer discovery* 8:537-
482   555.https://doi.org/10.1158/2159-8290.CD-17-1461
483 Lánczky A, and Győrffy B. 2021. Web-Based Survival Analysis Tool Tailored for Medical
484   Research (KMplot): Development and Implementation. *Journal of medical Internet*
485   *research* 23:e27633.https://doi.org/10.2196/27633
486 Li P, Qian J, Yu G, Chen Y, Liu K, Li J, and Wang J. 2012. Down-regulated SPARCL1 is
487   associated with clinical significance in human gastric cancer. *Journal of surgical*
488   *oncology* 105:31-37.https://doi.org/10.1002/jso.22025
489 Liu B, Xiang L, Ji J, Liu W, Chen Y, Xia M, Liu Y, Liu W, Zhu P, Jin Y, Han Y, Lu J, Li X, Zheng
490   M, and Lu Y. 2021. Sparcl1 promotes nonalcoholic steatohepatitis progression in mice
491   through upregulation of CCL2. *The Journal of clinical investigation*
492   131.https://doi.org/10.1172/jci144801
493 López-Knowles E, O'Toole SA, McNeil CM, Millar EKA, Qiu MR, Crea P, Daly RJ, Musgrove
494   EA, and Sutherland RL. 2010. PI3K pathway activation in breast cancer is associated
495   with the basal-like phenotype and cancer-specific mortality. *International journal of*
496   *cancer* 126:1121-1131.https://doi.org/10.1002/ijc.24831
497 McBryan J, Fagan A, McCartan D, Bane FT, Varešlija D, Cocchiglia S, Byrne C, Bolger J,
498   McIlroy M, Hudson L, Tibbitts P, Gaora PÓ, Hill AD, and Young LS. 2015.
499   Transcriptomic Profiling of Sequential Tumors from Breast Cancer Patients Provides a
500   Global View of Metastatic Expression Changes Following Endocrine Therapy. *Clinical*
501   *cancer research : an official journal of the American Association for Cancer Research*
502   21:5371-5379.https://doi.org/10.1158/1078-0432.CCR-14-2155
503 Miller KD, Nogueira L, Mariotto AB, Rowland JH, Yabroff KR, Alfano CM, Jemal A, Kramer JL,
504   and Siegel RL. 2019. Cancer treatment and survivorship statistics, 2019. *CA Cancer J*
505   *Clin* 69:363-385.https://doi.org/10.3322/caac.21565
506 Mingkuan C, Wenfang Z, and Lin F. 2022. Identifying liver metastasis-related hub genes in
507   breast cancer and characterizing SPARCL1 as a potential prognostic biomarker.
508   *PREPRINT (Version 1) available at Research Square [https://doiorg/1021203/rs3rs-*
509   *2183292/v1].*https://doi.org/10.21203/rs.3.rs-2183292/v1
510 Ritchie ME, Phipson B, Wu D, Hu Y, Law CW, Shi W, and Smyth GK. 2015. limma powers
511   differential expression analyses for RNA-sequencing and microarray studies. *Nucleic*
512   *acids research* 43:e47.https://doi.org/10.1093/nar/gkv007
513 Shannon P, Markiel A, Ozier O, Baliga NS, Wang JT, Ramage D, Amin N, Schwikowski B, and
514   Ideker T. 2003. Cytoscape: a software environment for integrated models of
515   biomolecular interaction networks. *Genome research* 13:2498-2504

516   Shen C, Yin Y, Chen H, Wang R, Yin X, Cai Z, Zhang B, Chen Z, and Zhou Z. 2018. Secreted
517   protein acidic and rich in cysteine-like 1 suppresses metastasis in gastric stromal
518   tumors. *BMC Gastroenterology* 18:105.https://doi.org/10.1186/s12876-018-0833-8
519   Shen J, Song R, Chow W-H, and Zhao H. 2020. Homologous recombination repair capacity in
520   peripheral blood lymphocytes and breast cancer risk. *Carcinogenesis* 41:1363-
521   1367.https://doi.org/10.1093/carcin/bgaa081
522   Sinn BV, Fu C, Lau R, Litton J, Tsai T-H, Murthy R, Tam A, Andreopoulou E, Gong Y, Murthy R,
523   Gould R, Zhang Y, King TA, Viale A, Andrade V, Giri D, Salgado R, Laios I, Sotiriou C,
524   Marginean EC, Kwiatkowski DN, Layman RM, Booser D, Hatzis C, Vicente Valero V,
525   and Fraser Symmans W. 2019. SET: a robust 18-gene predictor for sensitivity to
526   endocrine therapy for metastatic breast cancer. *NPJ breast cancer*
527   5:16.https://doi.org/10.1038/s41523-019-0111-0
528   Subramanian A, Tamayo P, Mootha VK, Mukherjee S, Ebert BL, Gillette MA, Paulovich A,
529   Pomeroy SL, Golub TR, Lander ES, and Mesirov JP. 2005. Gene set enrichment
530   analysis: a knowledge-based approach for interpreting genome-wide expression profiles.
531   *Proceedings of the National Academy of Sciences of the United States of America*
532   102:15545-15550
533   Sung H, Ferlay J, Siegel RL, Laversanne M, Soerjomataram I, Jemal A, and Bray F. 2021.
534   Global Cancer Statistics 2020: GLOBOCAN Estimates of Incidence and Mortality
535   Worldwide for 36 Cancers in 185 Countries. *CA Cancer J Clin* 71:209-
536   249.https://doi.org/10.3322/caac.21660
537   Szklarczyk D, Gable AL, Nastou KC, Lyon D, Kirsch R, Pyysalo S, Doncheva NT, Legeay M,
538   Fang T, Bork P, Jensen LJ, and von Mering C. 2021. The STRING database in 2021:
539   customizable protein-protein networks, and functional characterization of user-uploaded
540   gene/measurement sets. *Nucleic acids research* 49:D605-
541   D612.https://doi.org/10.1093/nar/gkaa1074
542   Tang Z, Li C, Kang B, Gao G, Li C, and Zhang Z. 2017. GEPIA: a web server for cancer and
543   normal gene expression profiling and interactive analyses. *Nucleic acids research*
544   45.https://doi.org/10.1093/nar/gkx247
545   Tao Z, Shi A, Li R, Wang Y, Wang X, and Zhao J. 2017. Microarray bioinformatics in cancer- a
546   review. *Journal of BUON : official journal of the Balkan Union of Oncology* 22:838-843
547   Telli ML, Stover DG, Loi S, Aparicio S, Carey LA, Domchek SM, Newman L, Sledge GW, and
548   Winer EP. 2018. Homologous recombination deficiency and host anti-tumor immunity in
549   triple-negative breast cancer. *Breast Cancer Res Treat* 171:21-
550   31.https://doi.org/10.1007/s10549-018-4807-x
551   Weis SM, and Cheresh DA. 2011. Tumor angiogenesis: molecular pathways and therapeutic
552   targets. *Nature medicine* 17:1359-1370.https://doi.org/10.1038/nm.2537
553   Xiang Y, Qiu Q, Jiang M, Jin R, Lehmann BD, Strand DW, Jovanovic B, DeGraff DJ, Zheng Y,
554   Yousif DA, Simmons CQ, Case TC, Yi J, Cates JM, Virostko J, He X, Jin X, Hayward
555   SW, Matusik RJ, George AL, and Yi Y. 2013. SPARCL1 suppresses metastasis in
556   prostate cancer. *Molecular oncology* 7:1019-
557   1030.https://doi.org/10.1016/j.molonc.2013.07.008
558   Xie T, Wu D, Li S, Li X, Wang L, Lu Y, Song Q, Sun X, and Wang X. 2020. microRNA-582
559   Potentiates Liver and Lung Metastasis of Gastric Carcinoma Cells Through the FOXO3-
560   Mediated PI3K/Akt/Snail Pathway. *Cancer management and research* 12:5201-
561   5212.https://doi.org/10.2147/CMAR.S245674
562   Ye H, Wang W-G, Cao J, and Hu X-C. 2017. SPARCL1 suppresses cell migration and invasion
563   in renal cell carcinoma. *Molecular medicine reports* 16:7784-
564   7790.https://doi.org/10.3892/mmr.2017.7535

565   Yu G, Wang L-G, Han Y, and He Q-Y. 2012. clusterProfiler: an R package for comparing
566   biological themes among gene clusters. *Omics : a Journal of Integrative Biology* 16:284-
567   287.https://doi.org/10.1089/omi.2011.0118
568   Yuan L, Zhou M, Wasan HS, Zhang K, Li Z, Guo K, Shen F, Shen M, and Ruan S. 2019. Jiedu
569   Sangen Decoction Inhibits the Invasion and Metastasis of Colorectal Cancer Cells by
570   Regulating EMT through the Hippo Signaling Pathway. *Evidence-based Complementary*
571   *and Alternative Medicine : ECAM* 2019:1431726.https://doi.org/10.1155/2019/1431726
572   Zhang H-P, Wu J, Liu Z-F, Gao J-W, and Li S-Y. 2022. SPARCL1 Is a Novel Prognostic
573   Biomarker and Correlates with Tumor Microenvironment in Colorectal Cancer. *BioMed*
574   *research international* 2022:1398268.https://doi.org/10.1155/2022/1398268
575   Zhang K-L, Zhu W-W, Wang S-H, Gao C, Pan J-J, Du Z-G, Lu L, Jia H-L, Dong Q-Z, Chen J-H,
576   Lu M, and Qin L-X. 2021a. Organ-specific cholesterol metabolic aberration fuels liver
577   metastasis of colorectal cancer. *Theranostics* 11:6560-
578   6572.https://doi.org/10.7150/thno.55609
579   Zhang T, Yuan K, Wang Y, Xu M, Cai S, Chen C, and Ma J. 2021b. Identification of Candidate
580   Biomarkers and Prognostic Analysis in Colorectal Cancer Liver Metastases. *Front Oncol*
581   11:652354.https://doi.org/10.3389/fonc.2021.652354
582   Zhang T, Yuan K, Wang Y, Xu M, Cai S, Chen C, and Ma J. 2021c. Identification of Candidate
583   Biomarkers and Prognostic Analysis in Colorectal Cancer Liver Metastases. *Front Oncol*
584   11:652354.https://doi.org/10.3389/fonc.2021.652354
585   Zhao H-Y, Gong Y, Ye F-G, Ling H, and Hu X. 2018a. Incidence and prognostic factors of
586   patients with synchronous liver metastases upon initial diagnosis of breast cancer: a
587   population-based study. *Cancer management and research* 10:5937-
588   5950.https://doi.org/10.2147/CMAR.S178395
589   Zhao SJ, Jiang YQ, Xu NW, Li Q, Zhang Q, Wang SY, Li J, Wang YH, Zhang YL, Jiang SH,
590   Wang YJ, Huang YJ, Zhang XX, Tian GA, Zhang CC, Lv YY, Dai M, Liu F, Zhang R,
591   Zhou D, and Zhang ZG. 2018b. SPARCL1 suppresses osteosarcoma metastasis and
592   recruits macrophages by activation of canonical WNT/β-catenin signaling through
593   stabilization of the WNT-receptor complex. *Oncogene* 37:1049-
594   1061.https://doi.org/10.1038/onc.2017.403
595

Figure 1

Differentially expressed genes (DEGs) of liver metastasis of breast cancer (BC)
identified from GSE124648.

(A) Volcano plot of the 332 DEGs, where blue and red indicate downregulated and
upregulated DEGs, respectively. The cut-off criteria were adj.P.Val<0.05 and |log$_2$FC|>2.0. (B)
Heatmap of the systematic cluster analysis of the top 50 DEGs.

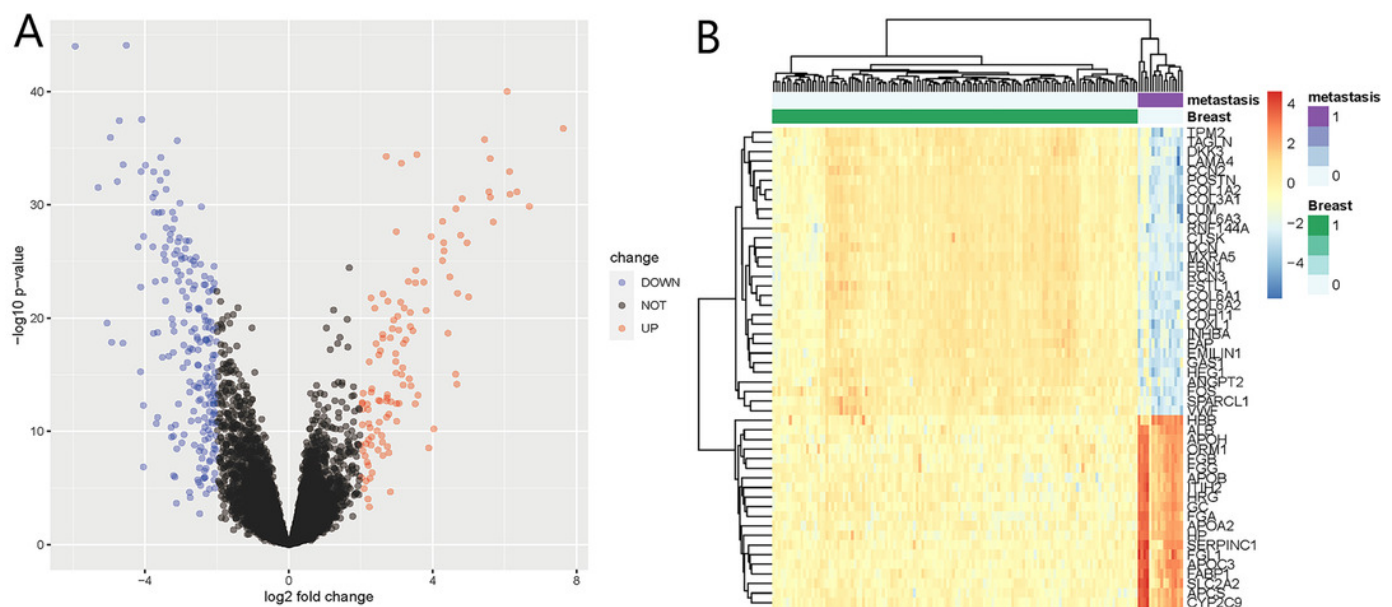



Figure 2

Gene ontology (GO) annotation and Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway enrichment analyses of the DEGs.

(A–C) GO annotation enrichment analyses of the 332 DEGs. CC, cellular component; MF, molecular function; BP, biological process. (D) KEGG pathway enrichment analyses of the DEGs, where blue and red indicate downregulated and upregulated DEGs, respectively.

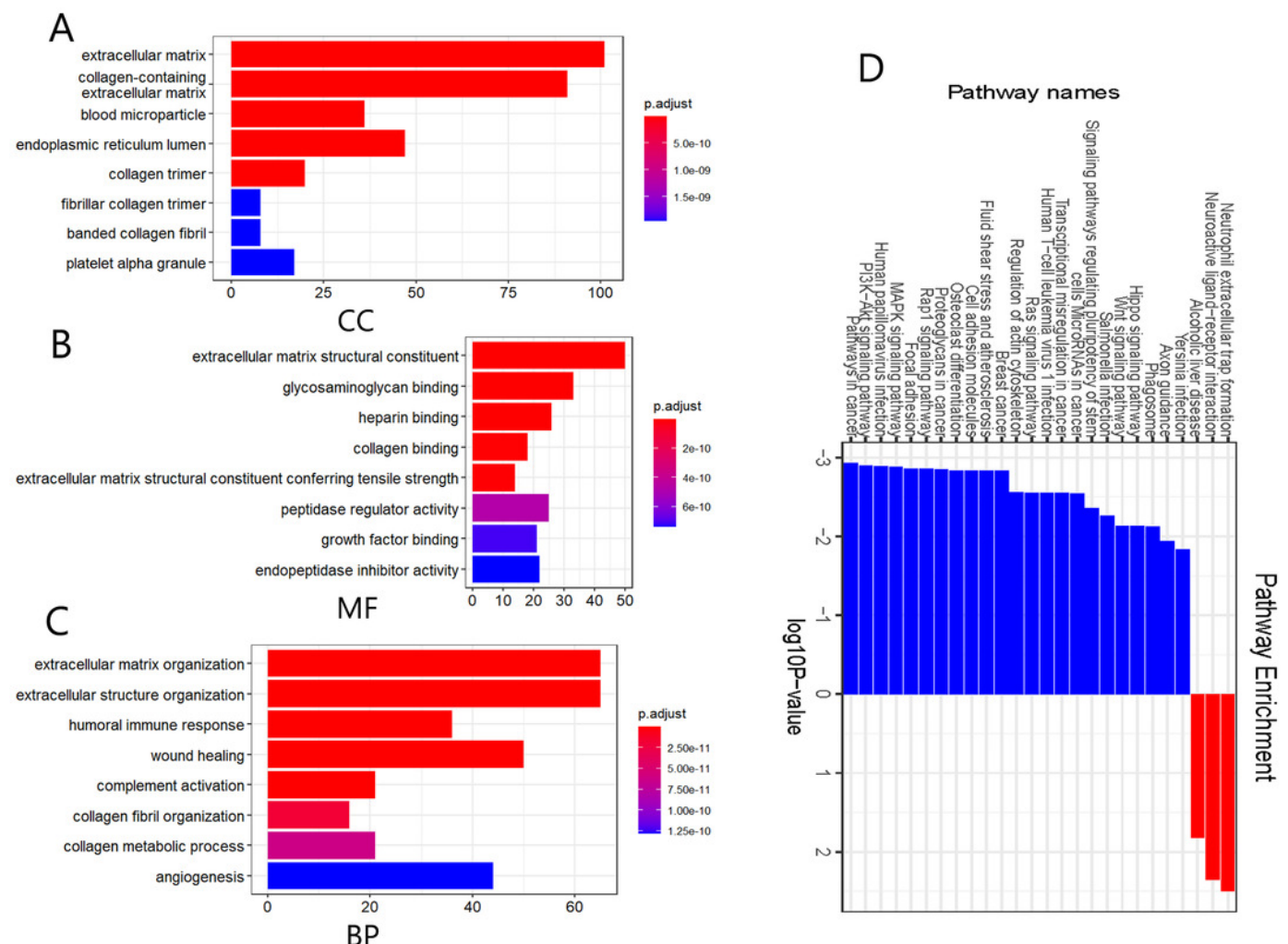



Figure 3

Identification of hub modules and genes from the protein–protein interaction (PPI) network.

(A–C) Top three modules with high scores identified from the PPI network. Blue and red indicate downregulated and upregulated DEGs, respectively. (D) Top 30 hub genes identified from the PPI network. Ordered from high (red) to low (yellow) connectivity degrees.

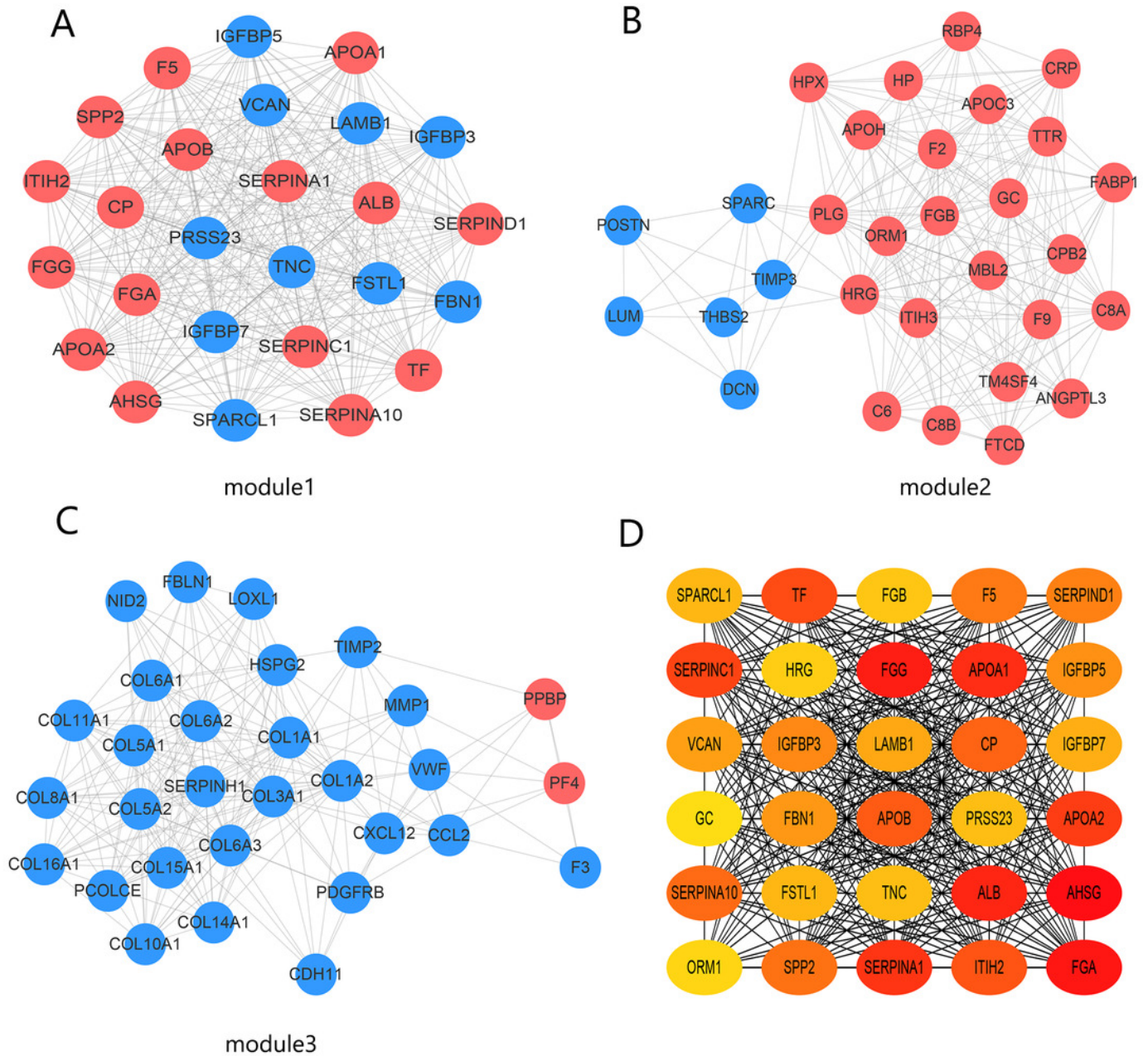



Figure 4

mRNA expression and prognostic value of (A) *SPARCL1* and (B) *SERPINA1*.

Red and black lines represent a patient with high and low gene expression, respectively. OS, overall survival; RFS, recurrence-free survival; HR, hazard ratio (*$p< 0.05$).

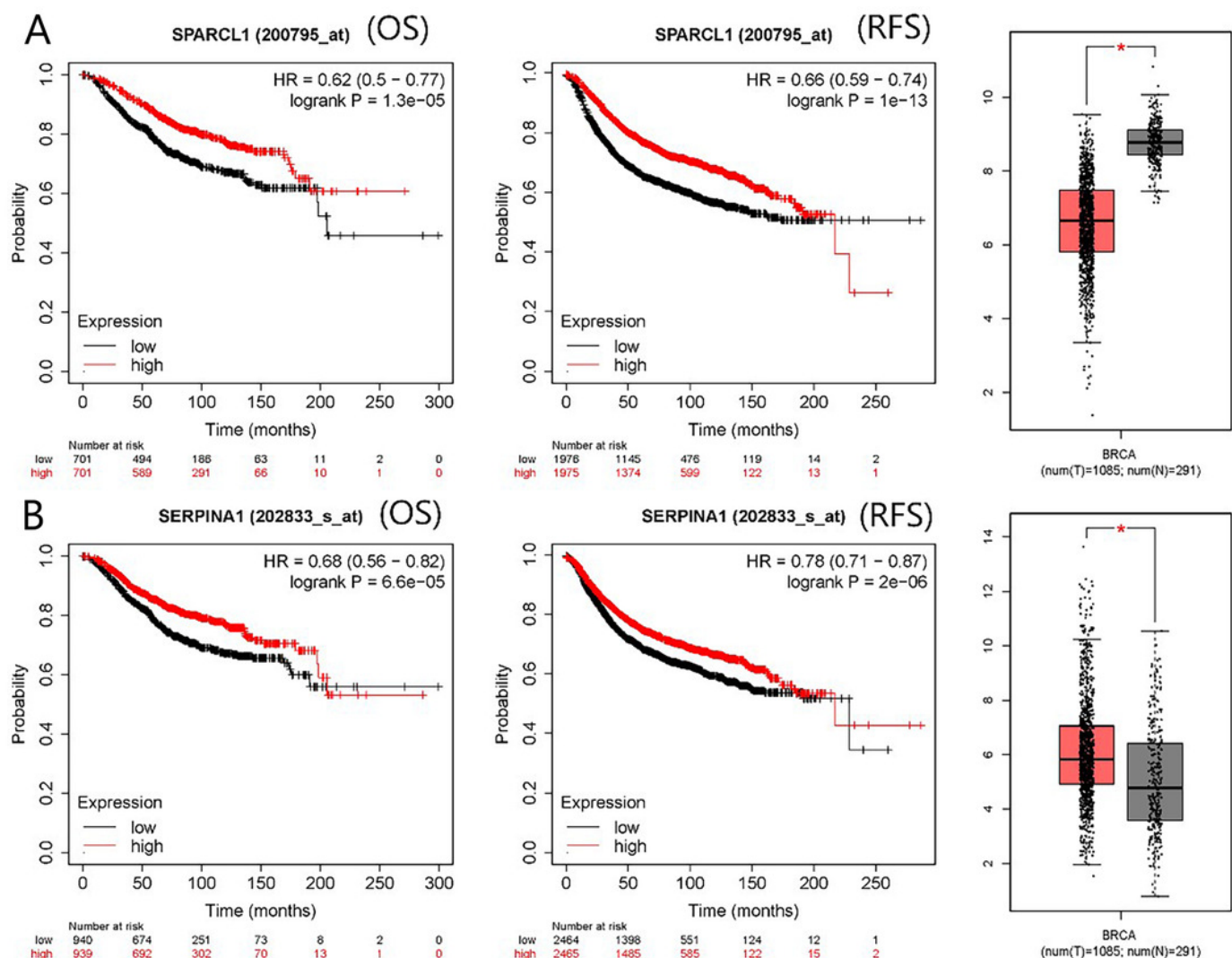



Figure 5

Single-gene gene set enrichment analysis (GSEA) of low *SPARCL1* expression in BC.

NES, normalized enrichment score; FDR, false discovery rate; NOM *p*-val, nominal *p*-value.

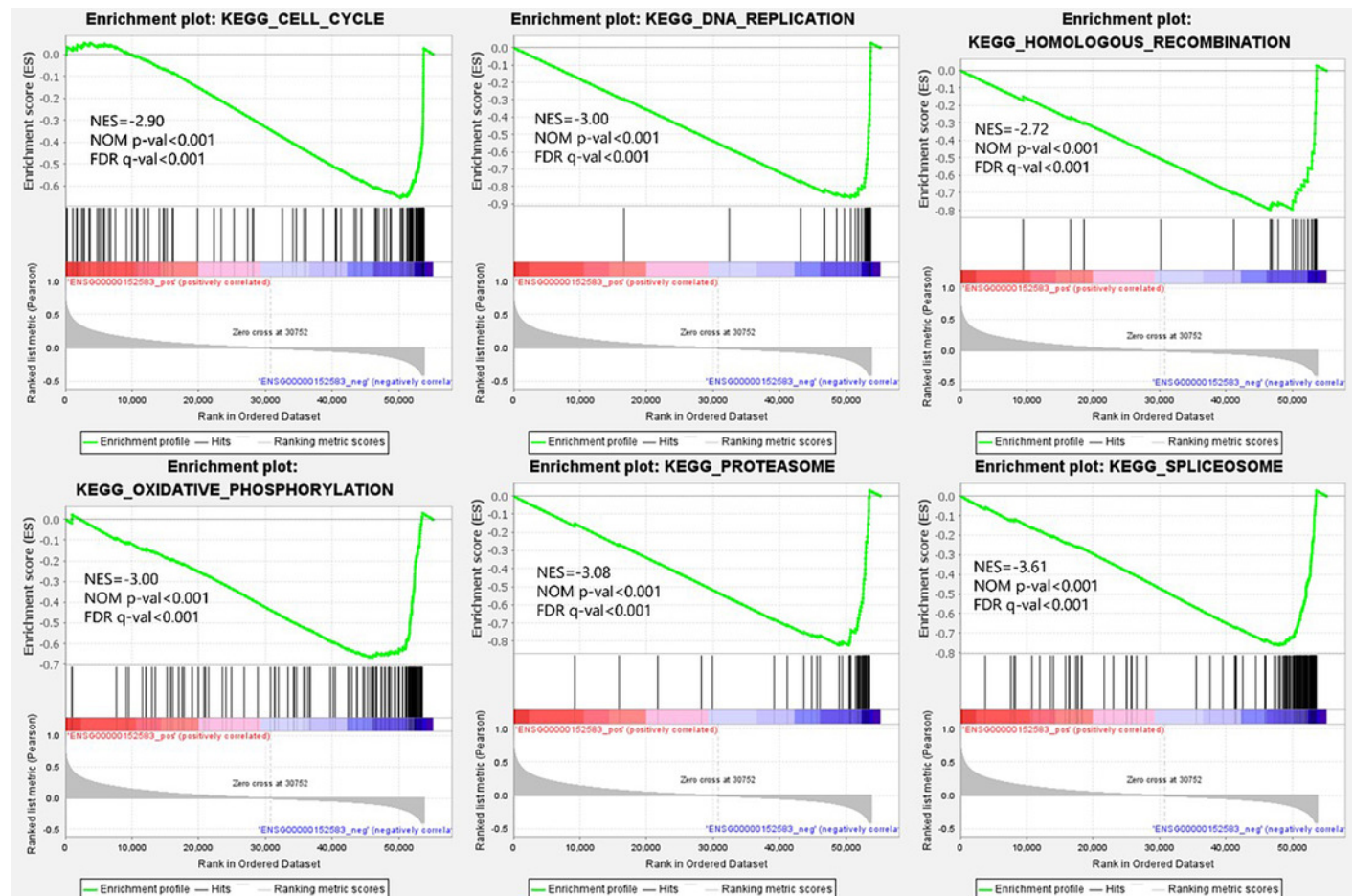



Figure 6

SPARCL1 expression in tissues and cell lines detected by RT-qPCR.

(A) *SPARCL1* expression in BC tissues and matched normal adjacent breast tissues (Wilcoxon matched-pairs signed-rank test). (B) Expression of *SPARCL1* in MCF-10A and breast cancer cell lines (Student's *t*-test). *p<0.05, **p<0.01, ***p<0.001, ****p<0.0001.

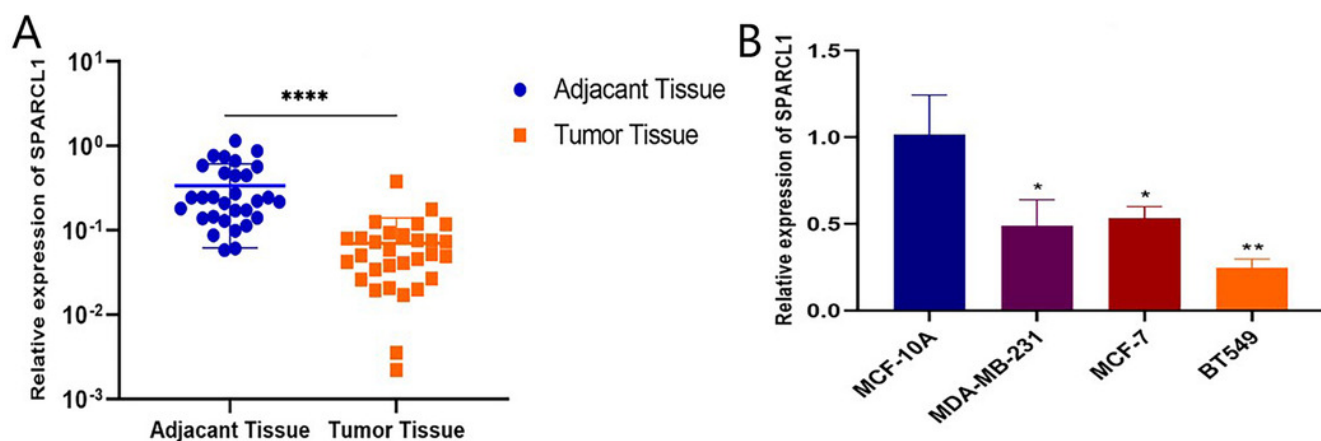



Figure 7

Biological function experiments in BC cells upon *SPARCL1* knockdown.

(A) RT-qPCR assessing the knockdown efficiency of *SPARCL1* shRNA on BC cells. (B) Colony formation assay measuring the proliferation ability of *SPARCL1*-knockdown BC cells. Cell colonies were photographed and their number was counted. (C) MTT assay was performed to detect the proliferation of *SPARCL1*-knockdown BC cells. (D) Wound healing assay was performed to assess the migration of MDA-MB-231 cells. Cells were photographed at 0 and 12 h after scratching, and wound sizes were compared. (E) The migration ability of MDA-MB-231 was evaluated with a Transwell assay, and the number of cells was counted. *$p<0.05$, **$p<0.01$, ***$p<0.001$, ****$p<0.0001$. NC, normal control; sh, short hairpin RNA.

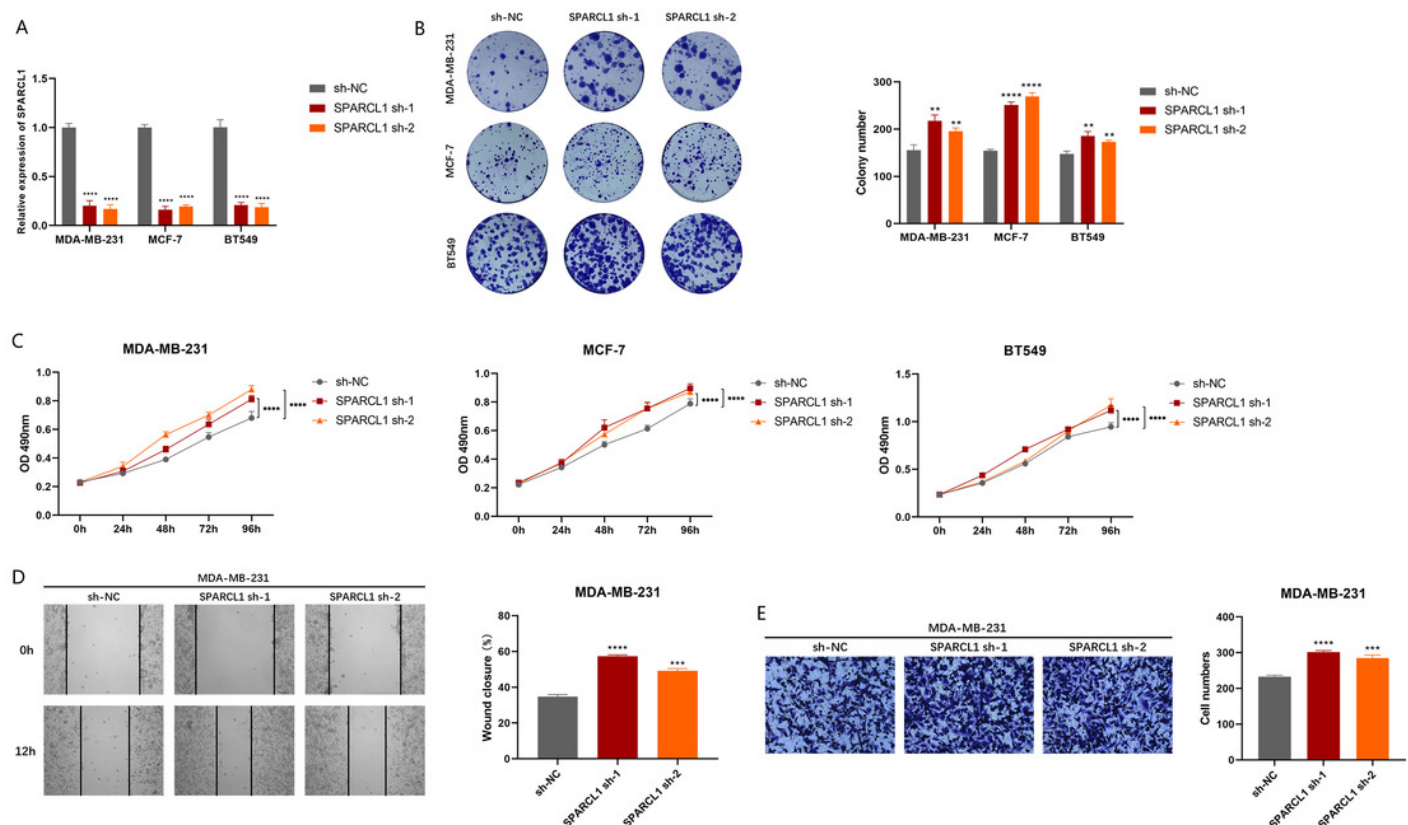
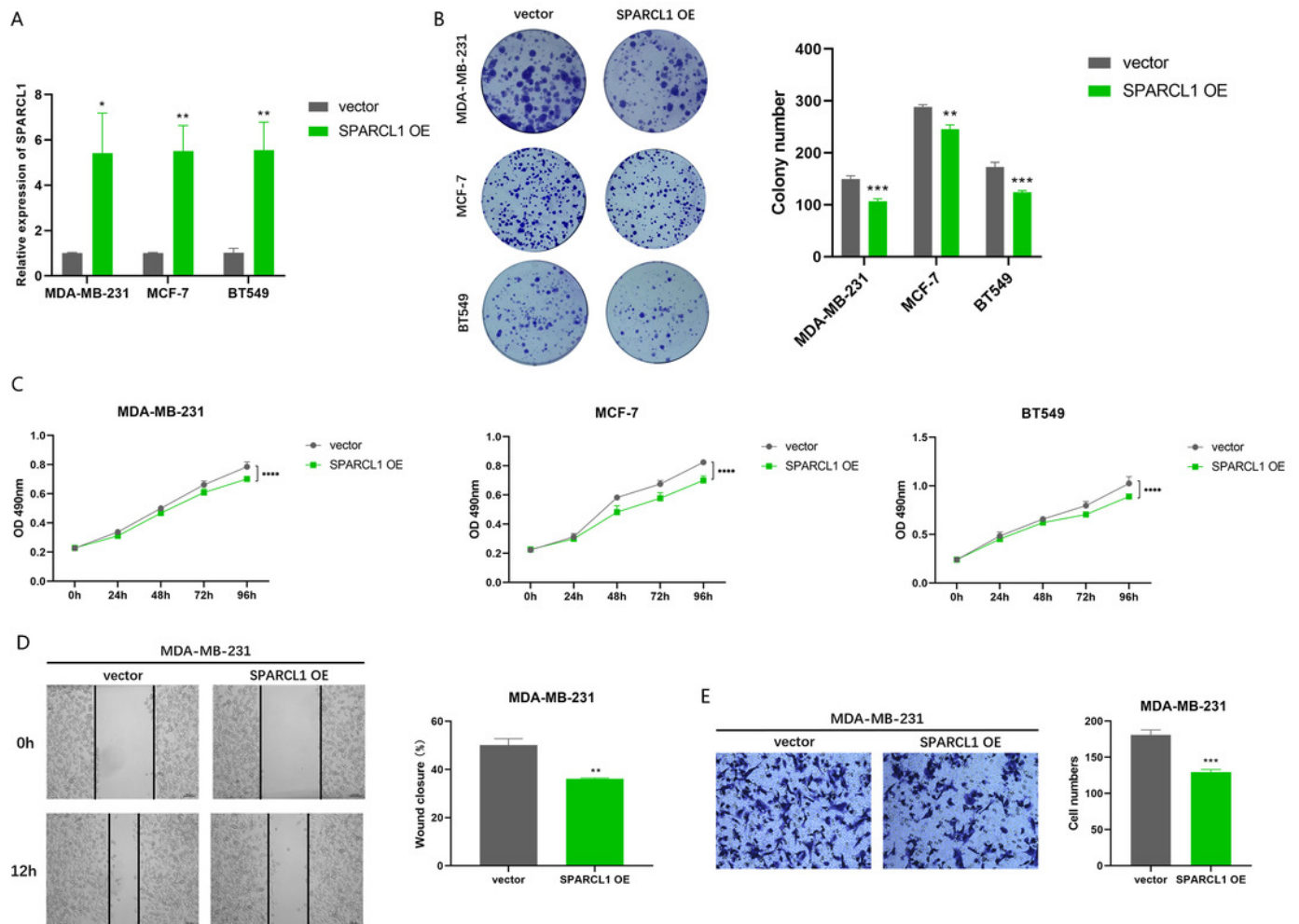



Figure 8

Biological function experiments in BC cells upon *SPARCL1* overexpression.

(A) RT-qPCR assessing the overexpression efficiency of *SPARCL1.* (B) Colony formation assay to detect the effect of overexpression of *SPARCL1* on BC cells colony formation ability. (C) Cell proliferation was assessed in BC cells by MTT assay. (D) Wound healing assay was performed to assess the effect of overexpression of *SPARCL1* on the migration ability of BC cells and the quantitative comparison of the percentage of scratch migration area. (E) Transwell assay was performed to detect the effect of overexpression of *SPARCL1* on the migration ability of BC cells and the quantitative comparison of the number of migrating cells. *$p<0.05$, **$p<0.01$, ***$p<0.001$, ****$p<0.0001$. OE, overexpression.

Discovery and validation of breast cancer (BC) datasets.

1   Table 1: Discovery and validation of breast cancer (BC) datasets.

Datasets	Platform	Experiment type	Number of cases (metastasis/primary)
GSE124648	GPL96	Expression profiling by array	16/130
GSE58708	GPL11154	Expression profiling by high-throughput sequencing	3/3

Top 30 hub genes associated with BC liver metastasis.

1   Table 2: Top 30 hub genes associated with BC liver metastasis.

Gene symbol	log$_2$FC	adj.P.Val	Expression	Module
FGG	6.149436	5.24×10^{-29}	UP	1
APOA1	4.69204	7.81×10^{-21}	UP	1
IGFBP7	−2.9555	5.11×10^{-21}	DOWN	1
APOB	3.954259	1.62×10^{-25}	UP	1
FSTL1	−3.02757	2.78×10^{-28}	DOWN	1
PRSS23	−2.34018	1.94×10^{-10}	DOWN	1
VCAN	−2.54093	5.33×10^{-17}	DOWN	1
TNC	−2.71678	3.74×10^{-10}	DOWN	1
ORM1	6.685187	5.25×10^{-28}	UP	2
IGFBP3	−2.10444	1.15×10^{-15}	DOWN	1
LAMB1	−3.0713	8.26×10^{-22}	DOWN	1
SERPIND1	3.39324	4.63×10^{-13}	UP	1
SPP2	2.236803	4.46×10^{-15}	UP	1
IGFBP5	−2.35175	1.19×10^{-7}	DOWN	1
FGA	4.820105	1.17×10^{-28}	UP	1
GC	6.136118	8.01×10^{-31}	UP	2
FGB	5.56398	3.49×10^{-29}	UP	2
HRG	4.776554	1.28×10^{-25}	UP	2
CP	2.150198	2.51×10^{-5}	UP	1
FBN1	−4.03832	1.62×10^{-25}	DOWN	1
SERPINA10	2.515328	4.67×10^{-10}	UP	1
ALB	7.627917	3.85×10^{-34}	UP	1
APOA2	5.683622	9.91×10^{-27}	UP	1
AHSG	3.816468	2.15×10^{-19}	UP	1
SERPINA1	3.350588	1.20×10^{-13}	UP	1
SPARCL1	−3.75713	8.01×10^{-31}	DOWN	1
TF	4.63786	4.81×10^{-14}	UP	1
F5	2.037343	1.08×10^{-11}	UP	1
SERPINC1	4.646863	8.04×10^{-28}	UP	1
ITIH2	2.98753	6.57×10^{-26}	UP	1

2   Abbreviations: log$_2$FC, log$_2$(Fold Change); adj.P.Val, adjusted P value.

Relationship between clinical characteristics of BC patients from the TCGA database and *SPARCL1* expression level

1   Table 3 Relationship between clinical characteristics of BC patients from the TCGA database
2   and *SPARCL1* expression level

		SPARCL1 expression		
	Total (N = 1,049)	High (N = 432)	Low (N = 617)	*p*-Value
Age (years)				
<55	432 (41.2%)	200 (46.3%)	232 (37.6%)	**0.00592**
≥55	617 (58.8%)	232 (53.7%)	385 (62.4%)	
TNM stage				
I	175 (16.7%)	90 (20.8%)	85 (13.8%)	**0.0027**
II	599 (57.1%)	221 (51.2%)	378 (61.3%)	
III	242 (23.1%)	110 (25.5%)	132 (21.4%)	
IV	20 (1.9%)	5 (1.2%)	15 (2.4%)	
Unknown	13 (1.2%)	6 (1.4%)	7 (1.1%)	
Node stage				
N0–N1	615 (58.6%)	245 (56.7%)	370 (60.0%)	0.139
N2–N3	137 (13.1%)	51 (11.8%)	86 (13.9%)	
Unknown	297 (28.3%)	136 (31.5%)	161 (26.1%)	
ER status				
Negative	170 (16.2%)	31 (7.2%)	139 (22.5%)	**<0.001**
Positive	571 (54.4%)	265 (61.3%)	306 (49.6%)	
Unknown	308 (29.4%)	136 (31.5%)	172 (27.9%)	
PR status				
Negative	238 (22.7%)	58 (13.4%)	180 (29.2%)	**<0.001**
Positive	500 (47.7%)	235 (54.4%)	265 (42.9%)	
Unknown	311 (29.6%)	139 (32.2%)	172 (27.9%)	
HER2 status				
Negative	621 (59.2%)	251 (58.1%)	370 (60.0%)	0.145
Positive	107 (10.2%)	37 (8.6%)	70 (11.3%)	
Unknown	321 (30.6%)	144 (33.3%)	177 (28.7%)	
Histological type				
IDC	752 (71.7%)	260 (60.2%)	492 (79.7%)	**<0.001**
ILC	196 (18.7%)	137 (31.7%)	59 (9.6%)	
Others	100 (9.5%)	35 (8.1%)	65 (10.5%)	
Unknown	1 (0.1%)	0 (0%)	1 (0.2%)	
Molecular type				
Basal-like	94 (9.0%)	11 (2.5%)	83 (13.5%)	**<0.001**
Lum A	219 (20.9%)	135 (31.3%)	84 (13.6%)	
Lum B	120 (11.4%)	21 (4.9%)	99 (16.0%)	
HER2-enriched	52 (5.0%)	14 (3.2%)	38 (6.2%)	
Normal-like	7 (0.7%)	5 (1.2%)	2 (0.3%)	
Unknown	557 (53.1%)	246 (56.9%)	311 (50.4%)	
Living status				
Alive	902 (86.0%)	385 (89.1%)	517 (83.8%)	**0.0185**
Dead	147 (14.0%)	47 (10.9%)	100 (16.2%)	

3   Abbreviations: TCGA, the Cancer Genome Atlas; TNM, tumor-node-metastasis; ER,
4   estrogen receptor; PR, progesterone receptor; HER2, human epidermal growth factor
5   receptor 2; IDC, Invasive ductal carcinoma; ILC, invasive lobular carcinoma; Lum, luminal.