Detection of biomarkers related to the diagnosis and prognosis of breast cancer through the combination of gene expression profiling and DNA methylation

Liqiang Qi (jrwangvip@163.com)
Chinese Academy of Medical Sciences and Peking Union Medical College
Bo Sun
National Clinical Research Center for Cancer, Tianjin Medical University Cancer Institute and Hospital
Beibei Yang
National Clinical Research Center for Cancer, Tianjin Medical University Cancer Institute and Hospital
Su Lu
National Clinical Research Center for Cancer, Tianjin Medical University Cancer Institute and Hospital

Article

Keywords: Biomarker, Breast cancer, , DNA methylation, Gene Expression

Posted Date: April 13th, 2022

DOI: https://doi.org/10.21203/rs.3.rs-1516257/v1

Abstract

Purpose

Understanding the DNA changes during the development of breast cancer from the molecular level has certain guidance significance for the diagnosis and treatment of it.

Methods

To detect marker genes of breast cancer, we simultaneously analyzed the gene expression profiling and DNA methylation microarray data of breast cancer downloaded from the Gene Expression Omnibus (GEO) and the cancer genome atlas (TCGA) database. The overlapped genes of differentially expressed genes (DEGs) and differentially methylated genes (DMGs) were screened. Based on these overlapped genes, hypermethylated - low expression "and" hypomethylated - high expression "genes and their methylation sites were screened, which were considered to be the susceptibility genes associated with breast cancer. At the same time, we analyzed the correlation between susceptibility genes expression and clinical indicators, such as Overall Survival (OS), tumor stage and distant metastasis and so on. .

Result

It was found that there were 14 hypermethylated - low expression genes and 6 hypomethylated - high expression genes, among which fifteen genes were associated with patient OS; Seven genes were associated with tumor stage and ten genes were associated with neoplasia, metastasis, recurrence, and regional disease of tumor.

Conclusion

We have identified 20 hypermethylated-low expression and hypomethylated-high expression genes in breast cancer through the integrated analysis of gene expression profiling and DNA methylation. Among them, the expression of ITIH5 and MAMDC2 had association with clinical characteristics, including OS, tumor stage, occurrence, metastasis, recurrence and local regional disease and they were enriched by GO function annotation. They might be regarded as beneficial biomarkers for diagnosis and prognosis of breast cancer.

Introduction

Breast cancer (BC) remains the main cancer-related cause of disease for women, endangering one in 20 globally and as many as one in eight in developed countries[1]. Epigenetic changes are tightly linked to tumorigenesis development and malignant transformation[2]. DNA methylation relies on molecular subtypes to affect gene expression and prognosis of breast cancer patients[3]. Noticeably, targeted tumor suppressor genes are frequently hypermethylated in breast cancer tissues and peripheral blood biospecimens[4]. DNA methylation-based measures of biological age may be important predictors of breast cancer risk[5].

Transcription start site (TSS) promoter of methylation is a promising and sensitive molecular, pan-cancer biomarker that is detectable in tumor tissues and liquid biopsy samples[6]. Methylation of several promoter sequences, including ESR1,APC, HSD17B4, HIC1, RASSF1A, MAD1L1, MAD2L2,MAD2L1, BUB3, BUB1B, BUB1, CDC20, and TTK has been correlated to BC occurrence and development[7–9]. Mechanistically, MAGI2-AS3 (MAGI2 antisense RNA 3) can inhibit the methylation of MAGI2 (membrane-associated guanylate kinase, WW and PDZ domain containing 2) promoter and upregulate the expression level of MAGI2, which is a tumor suppressor gene coexpressed with MAGI2-AS3 in breast cancer, and our findings reveal the role of MAGI2-AS3 in breast cancer and provide potential novel therapeutic targets for metastatic breast cancer intervention[10]. High expression of NEFM (neurofilament medium) correlated with better overall survival (OS) and recurrence-free survival (RFS) in TCGA BRCA and Kaplan-Meier plotter, whereas NEFM DNA methylation with worse OS in TCGA BRCA and NEFM transcriptional expression negatively correlated with DNA methylation[11]. Te hypermethylation in promoters of APC, SFRP1, SFRP2, SFRP5, WIF1, DKK3, ITIH5, and RASSF1A are associated with the development of breast cancer, and studies have found that APC and RASSF1A are common epigenetic biomarkers for early detection of breast cancer[12–15]. It is found that some key hypomethylation sites in enhancer regions and key hypermethylation sites in CpG islands are used to regulate the expression of key genes, such as oncogenes ESR1and ERBB2, and TSGs FBLN2, CEBPA, and FAT4[16]

Alterations in DNA methylation patterns, both at the global genomic level and locispecific, have been successfully explored as molecular biomarkers in cancer management[17]. The investigations identified methylation at Cytoplasmic FMR1 interacting protein 1 (CYFIP1) as a novel epigenetic biomarker candidate for sporadic breast cancer in the Uruguayan population[18]. Novel DNA methylation markers were identified, of which cg12374721 (PRAC2), cg18081940 (TDRD10) and cg04475027 (TMEM132C) show promise as diagnostic and prognostic markers in breast cancer as well as other cancer types[19]. Our findings suggest estrogen receptor-a (ESR1) and paired-like homeodomain transcription factor 2 (PITX2) promoter methylation may be correlated with a worse survival of patients with breast cancer (ESR1: OS, PITX2: OS and metastasis-free survival (MFS)) and the clinical utility of aberrantly methylated ESR1 and PITX2 could be a promising factor for the prognosis of breast cancer[20].

In the study, gene expression and DNA methylation microarrays datasets were incorporated and analyzed through a series of bioinformatics and statistics software. We screened out the overlapped genes of DEGs and DMGs in TCGA and GEO database The clinical significance of these genes was found by analyzing their correlation with clinical characteristics. With this method, we aimed to find novel candidate genes related to the occurrence and progression of breast cancer and they might be regarded as biomarkers to detect occurrence and prognosis of breast cancer.

Materials And Methods

Microarray data

In this study, there were two types of chip datas of BC in TCGA and GEO database: expression spectrum chip and methylation chip data. The gene expression datas included 1102 tumor datas and 113 normal datas, and methylation datas included 123 normal data and 1111 tumor data in TCGA database. The methylation datas selected in GEO database were GSE59901 and GSE141338 using Illumina Human Methylation450 BeadChip. GSE59901 contained 32 samples, 28 breast cancer tissue datas and 4 normal breast tissue datas. GSE141338 contained 41 breast cancer tissue datas and 6 healthy breast tissue datas. Gene differential expression datas selected in GEO database were GSE109169 and GSE29044. GSE109169 included 25 tumor samples and 25 normal samples, using Affymetrix Human Exon 1.0 ST Array. GSE29044 included 67 tumor samples and 36 normal samples, using Affymetrix Human Genome U133 Plus 2.0 Array.

Data Processing And Screening Of Differentially Expressed Genes And Differentially Methylated Sites

The differential expression data of breast cancer in TCGA database were downloaded from the TCGAbiolinks software package of R. For data from different sources, edgeR package was used to standardize the data, and the ComBat method of sva software package was used to conduct batch correction. TCGA database downloaded the methylated B value original file and converted between B value and M value through minifi package.

For microarray data, download the Series Matrix file data and chip annotation file provided by GEO database, and convert probe symbols into gene symbols through chip annotation information. For the case of multiple probes corresponding to the same gene, we select the average expression value of multiple probes as the expression value of the gene. Series Matrix file data and chip annotation file of GEO database. Using the limma package of R to screen differentially expressed genes, the screening threshold is adj.P.Val = 1. For the screening of differential methylation sites, the screening threshold was selected as adj. P. Val = 0.2. The characteristic types of methylation probe screening included N_Shore, Island, S _ Shelf, S _ Shore. Methylation ,differential expression analysis methods and screening criteria of TCGA database were consistent with GEO database.

Analysis Of Gene Expression, Methylation Sites And Clinical Datas

Gene expression and clinical data were analyzed using TPM data converted from Count data. Methylation indicators were analyzed according to the transformation of B value into M value and clinical indicators. Maxstat statistical package was used to calculate the optimal segmentation point for the expression value of each gene. According to the optimal segmentation point, patients were divided into high expression group and low expression group, and patients were divided into high methylation group and low methylation group according to M value of methylation. Survival package and SURvMiner package were used for survival analysis, and GGploT2 and GGPUBR were used for mapping.

Results

DEGs and DMGs analysis in TCGA and GEO database

In TCGA database, there were 5123 differentially expressed genes, 1,952 highly expressed genes and 3,171 low-expressed genes. Meanwhile there were 28,889 differentially methylated loci, 12,957 hypermethylated loci, and 15,932 hypomethylation loci in 7934 genes.

809 differentially expressed genes were obtained in GSE109169 of GEO, including 356 high- expressed genes and 454 low-expressed genes. 1072 differentially expressed genes were obtained in GSE29044, with 465 high-expressed genes and 607 low-expressed genes. In the GSE59901 dataset, 28394 differential methylation sites were obtained, including 19661 hypermethylated sites and 8733 hypomethylated sites. The GSE141338 dataset included 46728 differentially methylated sites, 27028 hypermethylated sites, and 19700 hypomethylated sites.

Table 1 showed the genes and locis with high expression and hypomethylation or low expression and hypermethylation in the 5 filtered database. 23 hypermethylated and low expressed loci distributed in 14 genes were obtained, including APCDD1, IGFBP6, TMEM220, AOX1, ITIH5, LEP, SPRY2, MAMDC2, MYBPC1, TMTC1, EBF1, FAT4, HSPB6, SLIT2. 8 hypomethylated and highly expressed loci distributed in 6 genes were obtained, including MUC1, H2BC5, CTPS1, MAL2, GALNT6, PRC1. Plot Venn was shown in the Fig. 1.

Table 1
Hypermethylated and low expressed or hypomethylated and highly expressed loci in the 5 filtered *database*

	Gene_Symbol	TCGAexpression	GSE20944 expression	GSE109169 expression	TCGA methylation	GSE59901methylation	GSE141338methylation
cg22531371	MUC1	High	high	high	Low	Low	Low
cg22500132	MUC1	High	high	high	Low	Low	Low
cg02386822	MUC1	High	high	high	Low	Low	Low
cg19264571	APCDD1	Low	low	low	high	High	High
cg27409154	IGFBP6	Low	low	low	high	High	High
cg06559575	IGFBP6	Low	low	low	high	High	High
cg24996758	H2BC5	High	high	high	Low	Low	Low
cg24141382	CTPS1	High	high	high	Low	Low	Low
cg02025583	TMEM220	Low	low	low	high	High	High
cg22953017	AOX1	Low	low	low	high	High	High
cg25327343	MAL2	High	high	high	Low	Low	Low
cg21253043	GALNT6	High	high	high	Low	Low	Low
cg10119075	ITIH5	Low	low	low	high	High	High
cg00840332	LEP	Low	low	low	high	High	High
cg06825512	APCDD1	Low	low	low	high	High	High
cg23841186	IGFBP6	Low	low	low	high	High	High
cg01407062	PRC1	High	high	high	Low	Low	Low
cg26814075	LEP	Low	low	low	high	High	High
cg22369786	SPRY2	Low	low	low	high	High	High
cg13870494	MAMDC2	Low	low	low	high	High	High
cg15374435	SPRY2	Low	low	low	high	High	High
cg23891273	MYBPC1	Low	low	low	high	High	High
cg13434308	TMTC1	Low	low	low	high	High	High
cg01135780	EBF1	Low	low	low	high	High	High
cg18607411	SPRY2	Low	low	low	high	High	High
cg17265829	FAT4	Low	low	low	high	High	High
cg02372889	HSPB6	Low	low	low	high	High	High
cg03790250	SLIT2	Low	low	low	high	High	High
cg23901852	FAT4	Low	low	low	high	High	High
cg00251610	EBF1	Low	low	low	high	High	High
cg16126280	EBF1	Low	low	low	high	High	High

Correlation Analysis Between Gene Expression And Clinical Indicators

Correlation between Gene expression and OS

Kaplan-meier curve and Cox univariate test were used to analyse correlation between gene expression and OS. Log-rank test results showed that 15 genes, including TMEM220, MAMDC2, SPRY2, PRC1, APCDD1, MUC1, HSPB6, H2BC5, ITIH5, CTPS1, EBF1, MYBPC1, IGFBP6, LEP and MAL2, were correlated with patient survival. Among them, the expression level of MAL2, CTPS1, PRC1 and MAMDC24 is higher ,and the survival time is shorter; the higher the gene expression level, the longer the survival time of the other genes. The rest is opposite. The Kaplan-meier survival curves were shown in the Fig. 2.

Correlation Between Gene Expression And Tumor Staging

It was showed that HSPB6, PRC1, IGFBP6, MAMDC2, GALNT6, ITIH5 and MYBPC1 were expressed differently in different tumor stages. The scatter diagrams were shown in the Fig. 3.

Correlation Between Gene Expression And New Event

New event contained occurence, metastasis, recurrence and local regional disease of tumor. Genes associated with appearance of new events included H2BC5, CTPS1, TMEM220, MAL2, GALNT6, ITIH5, SPRY2, MAAMDC2, MYBPC1, TMTC1. Correlation between the time of new event to appear and the survival probablity was shown in the Fig. 4. Among them, the higher the expression levels of CTPS1, TMEM220, ITIH5, and TMTC1, the shorter the time for new event to appear.

Kegg And Go Enrichment

20 genes with high expression and hypomethylation or low expression and hypermethylation were not enriched in the KEGG pathway, but they were mostly **enriched by GO function annotation (**Fig. 5**).**

Discussion

Breast cancer seriously affects the quality of life and family harmony, and brings heavy pressure to patients. Explanation of the potential mechanisms of occurrence and progression of breast cancer will have benefit to evaluate diagnosis and prognosis. The occurrence and development of breast cancer is closely related to the expression of many genes and proteins, which is a complex multi-step process[21]. Numerous studies have shown that activation of oncogenes[22], inactivation of tumor suppressor genes[23] and activation of growth factor receptors[24] are of great significance in the breast cancer. Studying the expression and interaction of these genes in breast cancer plays an important role in revealing the mechanism of the occurrence and development of breast cancer[25–27]. Van et al.[28]indicated that the gene-expression profiles they studied was a more formidable predictor of the disease result in young breast cancer patients than standard systems based on clinical and histologic standards.

In the current study, we analyzed the gene expression and DNA methylation profiling data obtained from TCGA and GEO database. There were 20 overlapped genes between *DEGs and DMGs*, among which the down-regulated and hypermethylated genes contained APCDD1, IGFBP6, TMEM220, AOX1, ITIH5, LEP, SPRY2, MAMDC2, MYBPC1, TMTC1, EBF1, FAT4, HSPB6, SLIT2, and the up-regulated and hypomethylated genes contained MUC1, H2BC5, CTPS1, MAL2, GALNT6, PRC1. Because aberrant methylated CpG sites can cause abnormal gene expression, and studies have shown that abnormal expression of certain genes affected the occurrence and development of diseases. So we analyzed the correlation between the expression of hypermethylated-low and hypomethylation-high expression genes and clinical characteristics. 17 genes had correlation with clinical characteristics. There were 15, 7 and 10 genes expression associated with OS, tumor stages and occurrence of new events respectively. Among these genes, only the expression of ITIH5 and MAMDC2 were synchronously correlated with OS, tumor stages and occurrence of new events and they were enriched in extracellular region by GO function annotation. They may develop into biomarkers for diagnosis and prognosis of breast cancer.

Human interalpha trypsin inhibitors (inter-α-trypsin inhibitor, ITI) are a family of plasma protease inhibitor molecules that are synthesized and secreted by the liver in high concentrations (0. 15 – 0. 5 mg /ml) is a protein-glycosaminoglycan-protein (PGP) complex, which exists in plasma[28–29]. Over the past 20 years, the ITI molecular family has been found to play a role in a variety of physiological and pathological processes, including inflammation, tumor formation and metastasis[29]. ITIH5 gene, a new member of ITI family, is a candidate tumor suppressor gene and has been confirmed to be involved in inflammation, tumor formation and metastasis in a number of studies[30]. Himmelfarb et al[31]. studied the expression of ITIH5 mRNA in normal human tissues and breast cancer tissues and found that ITIH5 was mainly expressed in female reproductive system and placental tissues, and the expression of ITIH5 mRNA in normal breast tissues was significantly higher than that in breast cancer tissues. The expression of ITIH5 mRNA was negatively correlated with breast cancer stage, suggesting that ITIH5 gene has a potential role in inhibiting breast cancer development. Heo et al[31] found by real-time fluorescence quantitative RNA in situ hybridization that ITIH5 gene expression was significantly down-regulated in breast cancer and invasive ductal carcinoma of the breast. ITIH5 can be used as a new prognostic marker in invasive non-nodular breast cancer, and down-regulated ITIH5 expression may play a certain role in the development of breast cancer., it has been demonstrated in carcinogenesis that ITIH5 gene expression is down-regulated due to aberrant DNA hypermethylation in breast cancer[32–33]. Promoter methylation-mediated down-regulation of ITIH5 expression is associated with worsening outcomes in breast cancer patients, and thus ITIH5 could be used as a prognostic biomarker[32].

The MAM (meprin/A-5 protein/ receptor protein-tyrosine phosphatase mu) domain is a conserved protein domain found in varied surface proteins[34].. One member of the MAM family, MAMDC2 (MAM domain containing 2), is a putative secretory protein[35]. It is reported that MAMDC2 is correlated with disease-free survival of breast cancer patients by a gene expression analysis[36]. MAMDC2 has a tumour-suppressive function and, as a secretory protein and down-regulated gene, it might be effective as a biomarker for breast cancer diagnosis and prognosis[35].

In conclusion, we have identified 20 hypermethylated-low expression and hypomethylated-high expression genes in breast cancer through the integrated analysis of gene expression profiling and DNA methylation. Among them, the expression of ITIH5 and MAMDC2 had association with many clinical

characteristics, including OS, tumor stages and occurrence of new events and they were enriched in extracellular region. They have the potential to become effective biomarkers for diagnosis and prognosis of breast cancer.

Declarations

Conflict of interests

The authors declare no conflict of interest.

Ethics approval

Not applicable.

Availability of data and material

Data from Gene Expression Omnibus (GEO) are available with accession numbers GSE59901, GSE141338, GSE109169, GSE2904.The Cancer Genome Atlas (TCGA) data of Breast cancer are available from https://portal.gdc.cancer.gov/.

Authors' contributions

L.Q. Q. and B. S. performed the experiments; B.B. Y. analyzed the data. S. L. provided material and methods. L.Q. Q. designed the project and wrote the manuscript; all authors edited and reviewed the final manuscript. Additionally, all authors provided final approval of the version to be submitted.

References

[1] Britt KL, Cuzick J, Phillips KA. Key steps for effective breast cancer prevention. Nat Rev Cancer. 2020 Aug;20(8):417-436.

[2] Zamora A, Alves M, Chollet C, et al. Paclitaxel induces lymphatic endothelial cells autophagy to promote metastasis. Cell Death Dis. 2019;10(12):956

[3] Győrffy B, Bottai G, Fleischer T, Munkácsy G, Budczies J, Paladini L, Børresen-Dale AL, Kristensen VN, Santarpia L. Aberrant DNA methylation impacts gene expression and prognosis in breast cancer subtypes. Int J Cancer. 2016 Jan 1;138(1):87-97.

[4] Vietri MT, D'Elia G, Benincasa G, Ferraro G, Caliendo G, Nicoletti GF, Napoli C. DNA methylation and breast cancer: A way forward (Review). Int J Oncol. 2021 Nov;59(5):98.

[5] Kresovich JK, Xu Z, O'Brien KM, Weinberg CR, Sandler DP, Taylor JA. Methylation-Based Biological Age and Breast Cancer Risk. J Natl Cancer Inst. 2019 Oct 1;111(10):1051-1058.

[6] Manoochehri M, Wu Y, Giese NA, et al. SST gene hypermethylation acts as a pan-cancer marker for pancreatic ductal adenocarcinoma and multiple other tumors: toward its use for blood-based diagnosis. Mol Oncol. 2020;14(6):1252-1267.

[7] Yuan B, Xu Y, Woo J H, et al. Increased expression of mitotic checkpoint genes in breast cancer cells with chromosomal instability. Clinical Cancer Res. 2006, 12(2):405-410.

[8] Widschwendter M, Siegmund KD, Mu¨ller HM, et al. Association of breast cancer DNA methylation profiles with hormone receptor status and response to tamoxifen. Cancer Res. 2004, 64(11):3807.

[9] Fiegl H, Mueller HM, Widschwendter A, et al. P13 DNA methylation in serum of breast cancer patients: an independent prognostic marker. EJC Supplements. 2004;2(1):44-44.

[10] Xu X, Yuan X, Ni J, Guo J, Gao Y, Yin W, Li F, Wei L, Zhang J. MAGI2-AS3 inhibits breast cancer by downregulating DNA methylation of MAGI2. J Cell Physiol. 2021 Feb;236(2):1116-1130.

[11] Li D, Zhao W, Zhang X, Lv H, Li C, Sun L. NEFM DNA methylation correlates with immune infiltration and survival in breast cancer. Clin Epigenetics. 2021 May 17;13(1):112.

[12] Feng W, Orlandi R, Zhao N, Carcangiu ML, Tagliabue E, Xu J, et al. Tumor suppressor genes are frequently methylated in lymph node metastases of breast cancers. BMC Cancer. 2010.

[13] Kloten V, Becker B, Winner K, Schrauder MG, Fasching PA, Anzeneder T, et al. Promoter hypermethylation of the tumor-suppressor genes ITIH5, DKK3, and RASSF1A as novel biomarkers for blood-based breast cancer screening. Breast Cancer Res. 2013;15(1):R4.

[14] Cao X, Tang Q, Holland-Letz T, Gündert M, Cuk K, Schott S, et al. Evaluation of promoter methylation of RASSF1A and ATM in peripheral blood of breast cancer patients and healthy control individuals. Int J Mol Sci. 2018;19(3):900.

[15] Coyle YM, Xie X-J, Lewis CM, Bu D, Milchgrub S, Euhus DM. Role of physical activity in modulating breast cancer risk as defned by APC and RASSF1A promoter hypermethylation in nonmalignant breast tissue. Cancer Epidemiol Biomarkers Prev. 2007;16(2):192–6.

[16] Jin W, Li Q-Z, Zuo Y-C, Cao Y-N, Zhang L-Q, Hou R, et al. Relationship between DNA methylation in key region and the diferential expressions of genes in human breast tumor tissue. DNA Cell Biol. 2019;38(1):49–62.

[17] Berdasco M, Esteller M (2019). Clinical epigenetics: seizing opportunities for translation. Nat Rev Genet. 2019;20:109–27.

[18] Cappetta M, Fernandez L, Brignoni L, Artagaveytia N, Bonilla C, López M, Esteller M, Bertoni B, Berdasco M. Discovery of novel DNA methylation biomarkers for non-invasive sporadic breast cancer detection in the Latino population. Mol Oncol. 2021 Feb;15(2):473-486.

[19] de Almeida BP, Apolónio JD, Binnie A, Castelo-Branco P. Roadmap of DNA methylation in breast cancer identifies novel prognostic biomarkers. BMC Cancer. 2019 Mar 12;19(1):219.

[20] Sheng X, Guo Y, Lu Y. Prognostic role of methylated GSTP1, p16, ESR1 and PITX2 in patients with breast cancer: A systematic meta-analysis under the guideline of PRISMA. Medicine (Baltimore). 2017 Jul;96(28):e7476.

[21] Polyak K. Breast cancer: origins and evolution. J Clin Invest 2007;117:3155-63.

[22] Osborne C, Wilson P, Tripathy D. Oncogenes and tumor suppressor genes in breast cancer: potential diagnostic and therapeutic applications. Oncologist 2004;9:361-77.

[23] Milstein M, Mooser CK, Hu H, et al. RIN1 is a breast tumor suppressor gene. Cancer Res 2007;67:11510-6.

[24] Masuda H, Zhang D, Bartholomeusz C, et al. Role of epidermal growth factor receptor in breast cancer. Breast Cancer Res Treat 2012;136:331-45.

[25] Hedenfalk IA, Ringnér M, Trent JM, et al. Gene expression in inherited breast cancer. Adv Cancer Res 2002;84:1-34.

[26] Van't Veer LJ, Dai H, Van De Vijver MJ, et al. Gene expression profiling predicts clinical outcome of breast cancer. Nature 2002;415:530-6.

[27] Ma XJ, Salunga R, Tuggle JT, et al. Gene expression profiles of human breast cancer progression. Proc Nati Acad Sci USA 2003;100:5974-9.

[28] Van De Vijver MJ, He YD, Van't Veer LJ, et al. A gene-expression signature as a predictor of survival in breast cancer. N Engl J Med 2002;347:1999-2009.

[29] Zhuo L，Kimata K． Structure and function of inter － alpha － trypsin inhibitor heavy chains[J]. Connect Tissue Res，2008，49(5) : 311-320．

[30] Yang Hua，Wang Xinlu，Shi Tiemei. Research and progress of ITIH5 in tumor[J]. MODERN ONCOLOGY, 2018, 26(6):950-954.

[31] Himmelfarb M，Klopocki E，Grube S，et al． ITIH5，a novel member of the inter-α-trypsin inhibitor heavy chain family is downregulated in breast cancer[J]. Cancer Lett，2004，204(1) : 69-77.

[32] Veeck J, Chorovicer M, Naami A et al. The extracellular matrix protein ITIH5 is a novel prognostic marker in invasive node-negative breast cancer and its aberrant expression is caused by promoter hypermethylation. Oncogene 2008;27(6):865-876.

[33] Hamm A, Veeck J, Bektas N et al. Frequent expression loss of Inter-alphatrypsin inhibitor heavy chain (ITIH) genes in multiple human solid tumors: a systematic expression analysis. BMC Cancer 2008;8:25.

[34] Beckmann G, Bork P. An adhesive domain detected in functionally diverse receptors. Trends Biochem Sci. 1993;18(2):40-41.

[35] Lee H, Park BC, Soon Kang J, Cheon Y, Lee S, Jae Maeng P. MAM domain containing 2 is a potential breast cancer biomarker that exhibits tumour-suppressive activity. Cell Prolif. 2020 Sep;53(9):e12883.

[36] Meng L, Xu Y, Xu C, Zhang W. Biomarker discovery to improve prediction of breast cancer survival: using gene expression profiling, meta-analysis, and tissue validation. Onco Targets Ther. 2016;9:6177-6185.

Figures

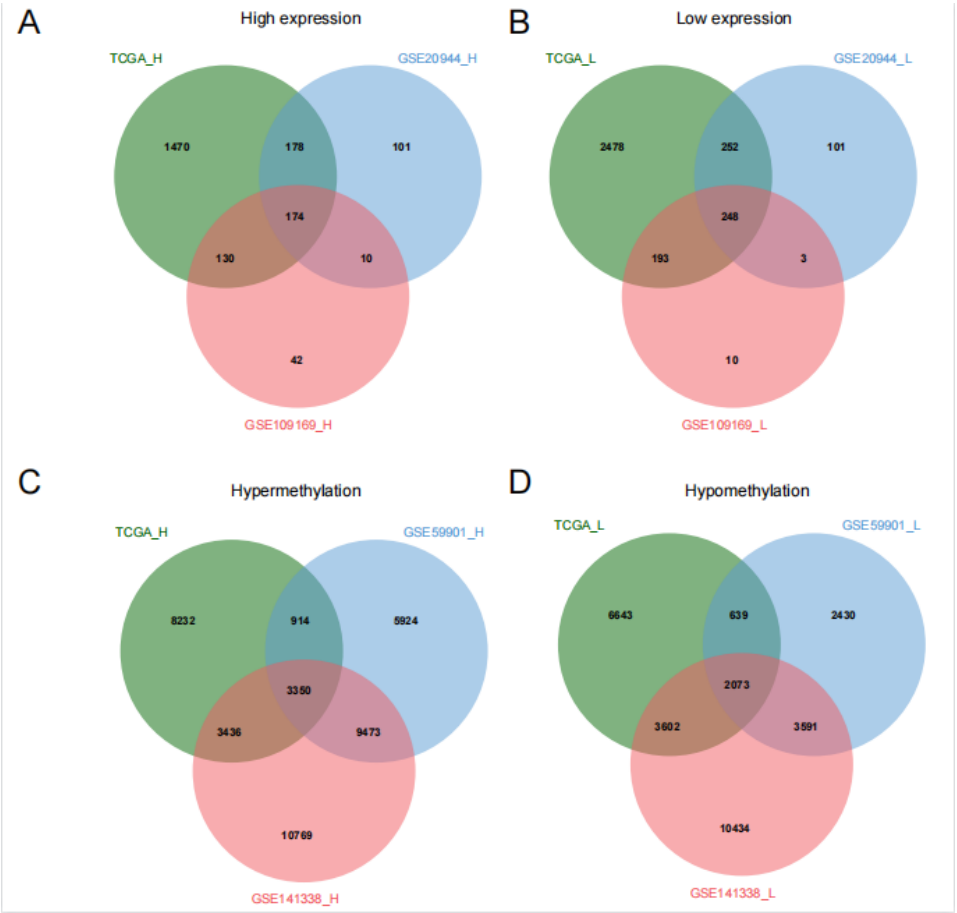

Figure 1

Plot Venn of hypermethylated-low expression and hypomethylated-high expression genes.

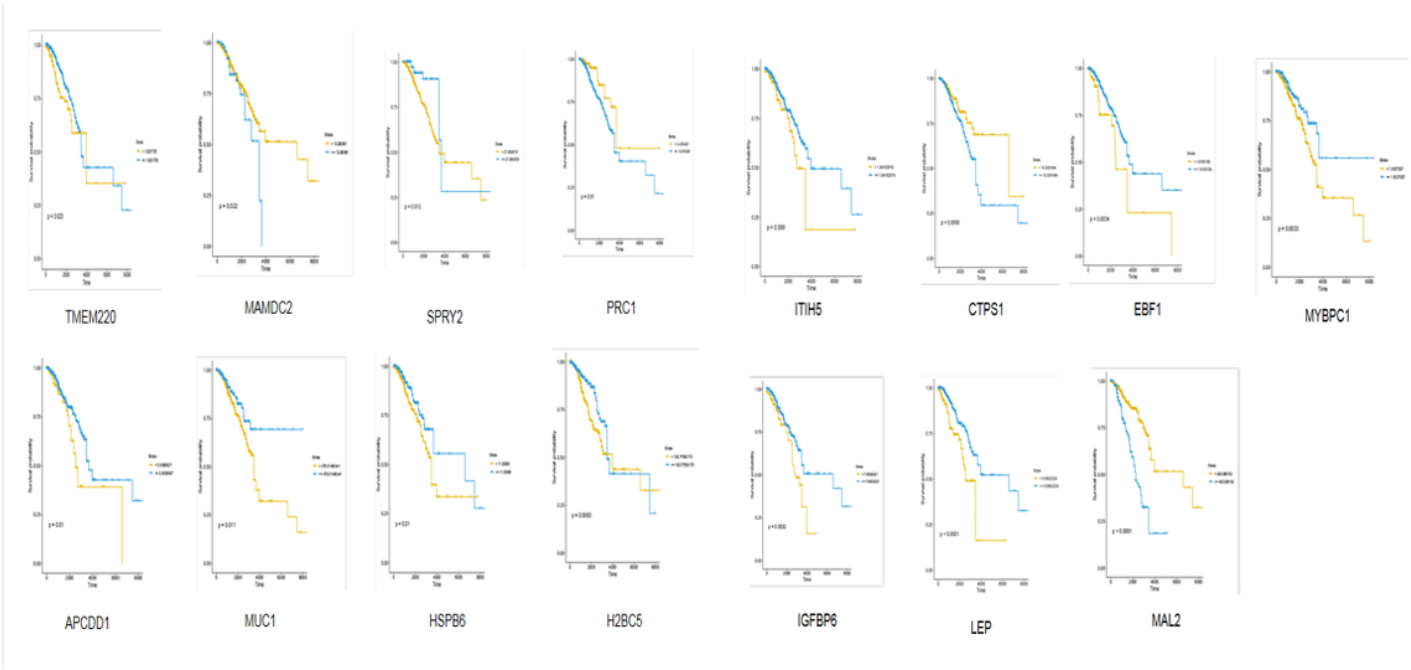

Figure 2

Kaplan-meier survival curves.

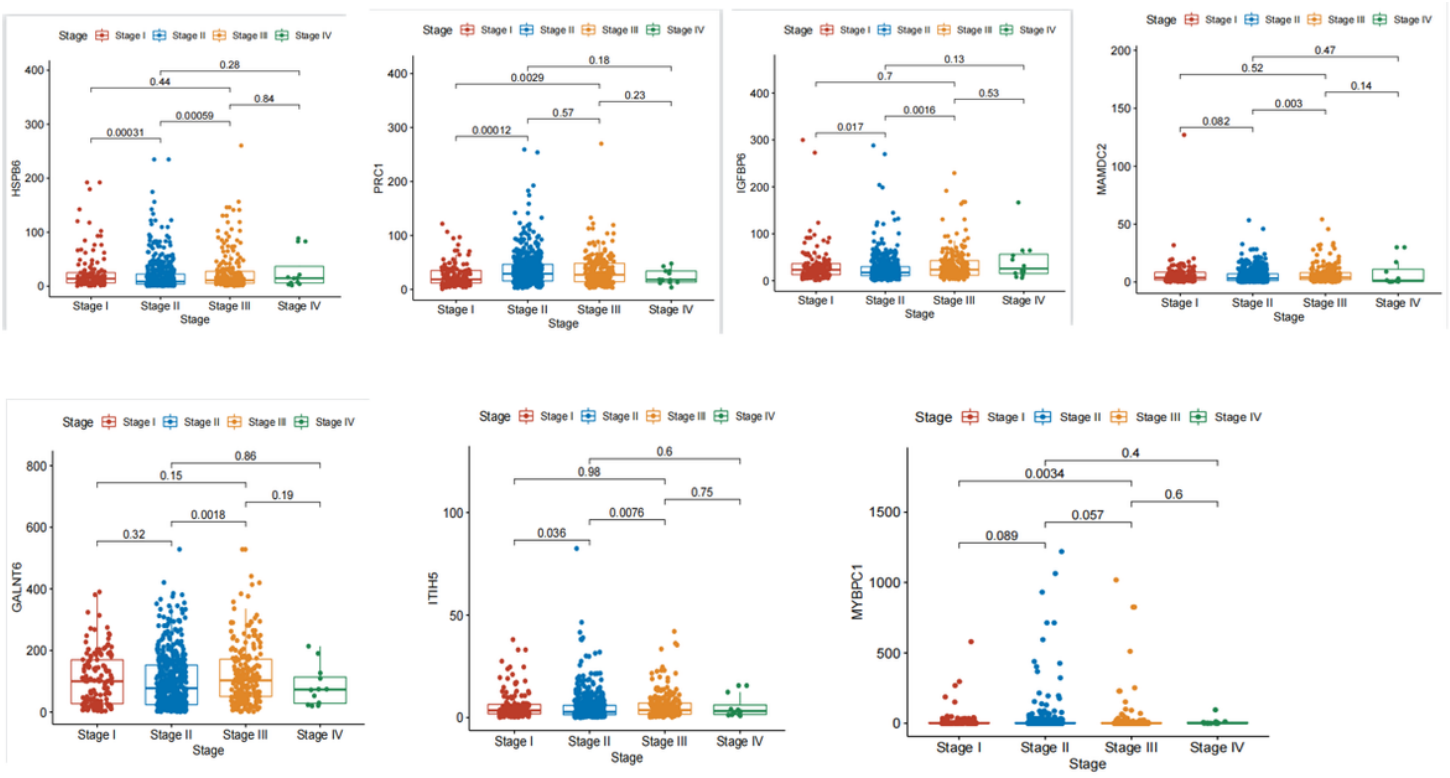

Figure 3

Expression of HSPB6, PRC1, IGFBP6, MAMDC2, GALNT6, ITIH5 and MYBPC1 in different tumor stages.

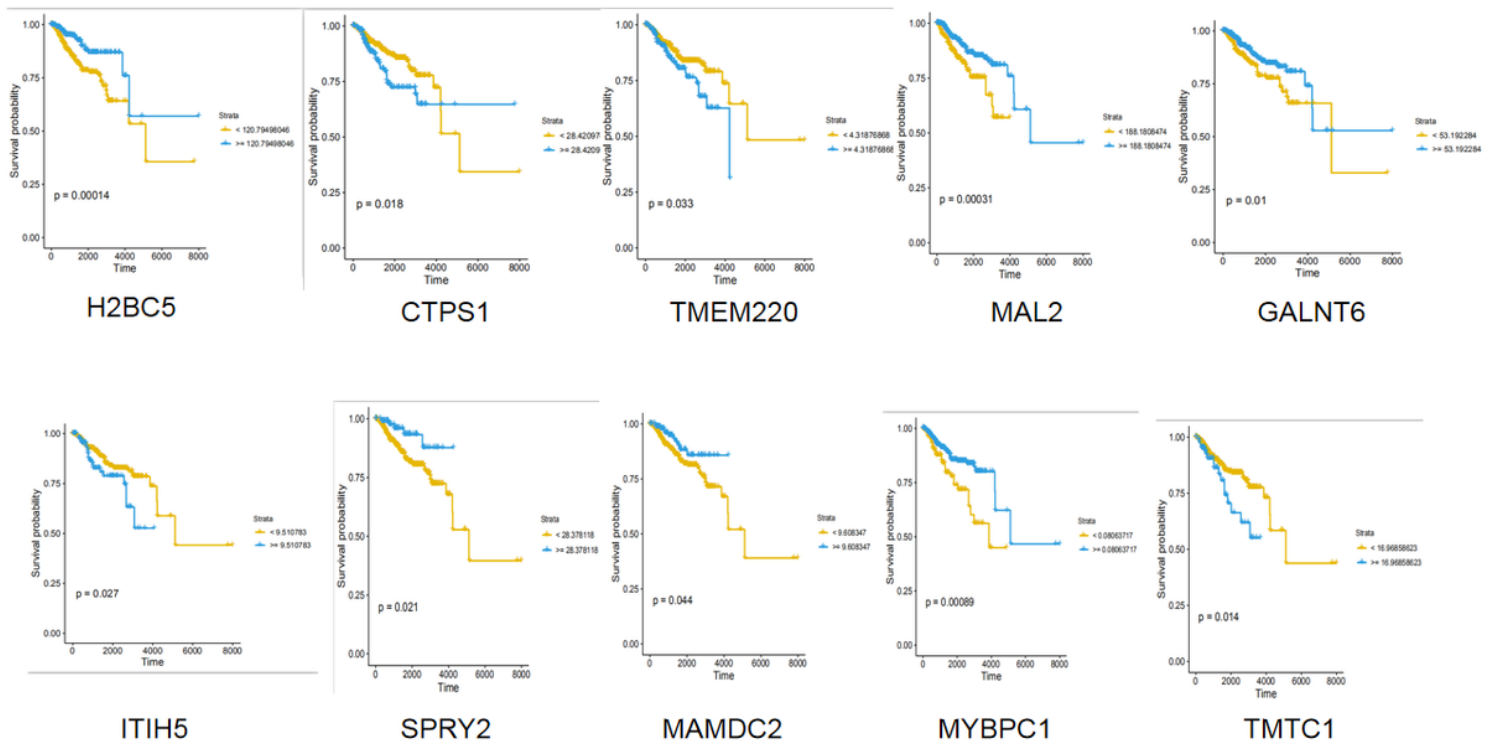

Figure 4

Correlation between the time of new event to appear and the survival probablity.

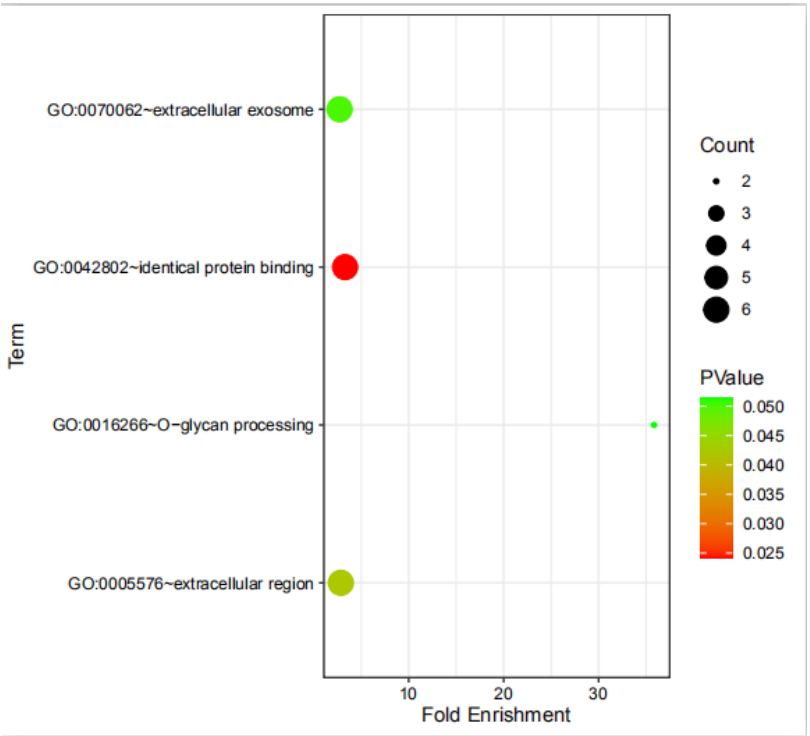

Figure 5

Enrichment analysis of 20 genes with high expression and hypomethylation or low expression and hypermethylation by GO function annotation.

Supplementary Files

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

- table.docx