

Epigenetics and Prognostic Survival Analysis of Muc Gene Family in Breast Cancer Based on Bioinformatics Analysis

Kai Yang

Xuzhou Medical College: Xuzhou Medical University

Xin-Hui Cao

Xuzhou Medical College: Xuzhou Medical University

Ming-Xing Liang

Nanjing Medical University

Di Xu

Nanjing Medical University

Yin-Jiao Fei

Nanjing Medical University

Zhen Liu

Nanjing Medical University

Xiu Chen

Jiangsu Province Hospital and Nanjing Medical University First Affiliated Hospital

jin-hai Tang (✉ jhtang@njmu.edu.cn)

Jiangsu Province Hospital and Nanjing Medical University First Affiliated Hospital

Research

Keywords: muc family, TCGA, breast cancer, prognosis, mutation

Posted Date: May 4th, 2021

DOI: <https://doi.org/10.21203/rs.3.rs-425005/v1>

License: © ⓘ This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Abstract

Background: To explore the epigenetic and biological functions of muc members in breast cancer.

Results: Mutation analysis showed that most of the muc members had missense mutations and the breast cancer patients with muc13 had a poor prognosis. Methylation analysis showed that the methylation expression of muc1, muc4, muc5AC and muc13 was related closely. The differentially expressed muc1, muc13 and muc15 had a poor prognosis and were related to immune infiltrating lymphocytes in breast cancer microenvironment.

Conclusions: The epigenetic basis and biological functions of muc members are closely related to the occurrence and development of breast cancer.

Background

Breast cancer(BC) is the most common cancer in women, accounting for 30% of all cancers[1]. In recent years, more convincing evidence shows that epigenetic modifications such as chromatin remodeling and DNA methylation play key roles in the early stages of breast cancer[2]. Identification of reliable molecular markers for BC prognosis could facilitate individualized treatment.

Mucins(MUC) are a diverse family of high molecular weight glycoproteins widely expressed in epithelial tissues[3]. To date, there are 21 members in the mucin family(<https://www.genenames.org/>). The presence of carbohydrate antigens can often be observed on the surface of tumor cells and is related to the poor prognosis of cancer patients[4]. CA15-3, also known as MUC1, has been identified as a biomarker for breast cancer diagnosis for its elevated serum level[5]. In addition to its protective effect in normal physiology, MUC16 (CA125) contributes to the disease progression and metastasis of certain malignant tumors. It is promising to be regarded as a prognostic target for diagnosis and treatment in cancers, such as gastric cancer, lung cancer, ovarian cancer and cervical cancer[6-10]. Studies have shown that MUC members are widely expressed in gastrointestinal tumors and can serve as potential targets for their prognosis[11-16]. However, the epigenetic and biological functions of muc members in breast cancer have not been systematically studied.

In general, the role of muc members in different cancers has received widespread attention. Here, we use the online public database The Cancer Genome Atlas(TCGA) and online analysis software, combining bioinformatics knowledge, to comprehensively explore the relevant mechanisms of muc members in breast cancer.

Results

Mutation Analysis of Muc Gene Family in Breast Cancer

In breast cancer, a total of 16 muc family members have mutations. Among them, muc16 has the highest number of mutations, followed by muc4. In the analysis of the mutation types, it is found that the muc16, muc4, muc17, and muc5B of the top 5 muc genes had the most common missense mutations, and the mutation of muc2 was mainly the 3'UTR region. The overall analysis showed that non-synonymous mutations per megabase are larger than synonymous mutations (figure1A).

We also analyzed the chromosomal loci of the mutant gene. In the analysis of breast cancer muc mutant genes, we found that chromosomes 1, 3, 4, 6, 7, 11, 12 and 19 have mutations. Among them, the mutation frequency of Chromosomes 11 and 19 were higher than others (shown in red) (figure1B).

For genes with mutations greater than 10, we identified them as highly mutated genes. In breast cancer samples, there were six highly mutated genes in the muc family including muc1, muc12, muc13, muc15, muc16, and muc17. We divided the samples into mutants and wild-types, and analyzed the survival prognosis of these highly mutated genes. The results showed that the muc13 mutation had a poor prognosis which was statistically significant ($p=0.027$). Also, we found that the survival period of mutant genes was significantly shorter than that of genes without mutation (figure2).

Expression Status of MUC Members in Breast Cancer Tissues

Based on the TCGA database, we found that fourteen muc members were expressed in breast cancer tissues. We used the limma package to analyze differentially expressed genes (figure3A) and the pheatmap package for visualization (figure3C). Compared with normal tissues, we discovered that breast cancer tissues with $|\logFC|>1$ and $FDR<0.05$ had significant differences. Among them, muc1, muc5AC, muc5B, muc6, muc13, muc21 were up-regulated, muc7 and muc15 were down-regulated, and others showed no significant difference.

Furthermore, we performed a correlation analysis of muc members expressed in breast cancer, which showed that muc members were highly related to each other (figure3B).

Correlation Analysis of Muc Methylation and Expression in Breast Cancer

Gene promoter region methylation is one of the most common mechanisms that affect gene expression in the development of human cancer. We found that eight muc genes were methylated in breast cancer, of which six genes (muc1, muc5AC, muc5B, muc13, muc15, muc17) were significantly down-regulated ($p<0.05$). At the same time, we conducted a correlation analysis of the methylated muc gene and its expression. The correlation analysis showed that muc1 methylation was negatively correlated with the expression level. Methylation of muc4, muc5AC, and muc13 was positively correlated with expression (all p values <0.05) (figure4).

Prognosis and Clinical Analysis of MUC Members in Breast Cancer

Based on Kaplan-Meier Plot (<http://www.kmplot.com>), an online analysis website dedicated to prognostic analysis of survival, we selected muc members with differential expression in breast cancer for

prognostic survival analysis, and found the high expression of muc1, muc7 and muc13 and the low expression of muc15 represented a poor survival period (figure5).

Also, we made relevant subgroup analysis based on the clinically different types of breast cancer patients. For lymph-positive patients the higher expression of muc15 and muc1 indicated a poor prognosis(Table 1). From the perspective of clinical stage we discovered the higher expression of muc1 gene in patients with clinical stage III represented a poor prognosis(Table 2). We also explored the different breast cancer classification. Among them, the higher expression of muc1 gene in triple-negative breast cancer patients indicated a poor prognosis, and the higher expressions of muc3A, muc7, muc1 in patients with luminalA type indicated poor survival of patients. The expression of muc20, muc13 and muc1 in patients with luminalB type represented a poor prognosis. In addition, the expression of muc3A and muc5AC in patients with her2 overexpression type predicted the poor prognosis (Table 3).

Correlations Between TILs and MUC Members in Breast cancer

In recent years, more and more people believe that tumor immune infiltrating cells play an important role in the occurrence and development of tumors. And they are closely related to the progression and prognosis of cancer. Here we used the Sangerbox tool to analyze statistically differentially expressed muc members and six types of immune infiltrating cells based on the TIMER database. There was a significant negative correlation between muc1 gene expression and immune infiltrating lymphocytes. On the contrary, muc13 and muc15 had a positive correlation with tumor immune infiltrating cells(figure6).

GSEA Multi-channel Enrichment Analysis

In order to explore the potential mechanism of muc members with statistically differential expression in the occurrence and development of breast cancer, we conducted a GSEA multi-pathway enrichment analysis. Here, we focused on the muc1, muc13 and muc15 genes with poor prognosis above. GSEA multi-pathway enrichment analysis indicated that the high expression of muc1 was negatively correlated with the pathways of "cell cycle", "cysteine and methionine metabolism" and "DNA replication", while was positively correlated with the pathways of "peroxisome", "SNARE interactions in vesicular transport" and "sphingolipid metabolism". The high expression of muc13 was positively correlated with the pathways of "antigen processing and presentation", "cell adhesion molecules CAMs", and "cytokine cytokine receptor interaction". The high expression of muc15 was positively correlated to the pathway of "cell adhesion molecules CAMs" and "chemokine signaling pathway", "cytokine cytokine receptor interaction" and "FC gamma r mediated phagocytosis"(figure7).

Discussion

MUC gene family encoded a total of 21 mucins, which were divided into secreted and membrane-related genes[3]. The former type of mucin protein included MUC2, MUC5, MUC6, MUC7, MUC8, MUC9, and MUC19. While MUC1, MUC3, MUC4, MUC11, MUC12, MUC13, MUC15, MUC16, MUC17, MUC20, and MUC21 were bound with cell membranes[17]. Those transmembrane MUC protein contained a

hydrophobic domain[3]. Whether the secreted or membrane-related mucins, they all perform significant role across surface of lumen in human as physical defense. Detailly, different mucins still own special and selective function, which was dependent to highly O-glycosylated amino acids rich in Ser/Thr/Pro[18]. Such sugar modified structure endows plentiful possibilities to serve biological behaviors[19]. Despite of various O-glycosylated modification of mucins, MUC genes share similar characteristics, a repeated sequence of defined nucleotides, which make them a set of gene family on the level of DNA[20, 21].

Increasing clinical evidence have showed that aberrant manifestation of MUC gene family was connected with human cancers. There have been studies found that MUC2 and MUC5AC expression were related with the prognosis for colorectal carcinoma (CRC)[22]. MUC16 has been known as important diagnostic marker of ovarian cancer[23]. And for pancreatic cancer, MUC4 seemed to be hopeful to distinguish adenocarcinoma and chronic pancreatitis[24]. With more and more bioinformation platforms or tools published, we were provided with a bulk of data to trace full expression profiles of MUC gene family in breast cancer and dig out their practical values in clinical. Our study was to demonstrate expression pattern, mutation, methylation and prognosis of MUC gene members, as well as their relationship with tumor immune infiltration and possible biological processes involved in.

According to our results, MUC1 have exhibited difference from other members in breast cancer patients. Consistent with previous studies, MUC1 is overexpressed in breast cancer cells compared to normal breast epithelia, especially in TNBC[25–27]. Nearly 40% breast neoplasm tissues were detected with amplification of MUC1 gene, as well as the upregulation of MUC1 mRNA and protein level[28]. Some studies showed that miR-125b and miR-145 could bind to 3'UTR of MUC1 gene to suppress its expression[29, 30]. There were reports claimed MUC1 was interacted with EGFR on the membrane of cells. Two of these molecules formed complexes through galectin-3 in extracellular region[31]. Then MUC1 cytoplasmic domain would be phosphorylated by EGFR or other tyrosine kinases. In turn, EGFR was activated to enter the nuclear dependent on PI3k promoted by means of its SH2 domain binding with MUC1 cytoplasmic sites[32]. As we mentioned before, MUC1 was negatively relevant with immune response in breast cancer. It has been found that anti-MUC1 antibodies in patients' serum might improve immune reaction of our body by strengthening cytotoxicity, killing those circulating tumor cells carrying MUC1 marker[33]. Detailly, MUC1 could downregulate T cells and NK cells functions directly[34, 35]. Additionally, MUC1 induced PD-L1 through MYC and NF- κ Bp65 pathway was reported to be hopeful method as supplement for immune treatment. Furthermore, increased PD-L1 not only help cancer cells evade immune system, but also contribute to EMT process in TNBC[36, 37].

MUC4, also known as sialomucin complex in rats, keep in low level at nonpregnant mammary gland. When it comes to pregnancy, it will boost up 100 times[38]. Thus, it was presumed MUC4 might be involved in immunity for babies. So, it was strictly regulated in a spatiotemporal pattern. While a higher abundance of MUC4 was also observed in breast cancer metastasized into ascites[39]. Mechanically, MUC4 would combine with ErbB2 following ErbB3 ligand with NRG. Then four of them formed a heteromeric complex to activate PI3K-Akt and MAPK pathways, inhibiting cancer cells apoptosis[40].

During above process, the expression of MUC4 functioned as chaperone for ErbB2, making ErbB2 and ErbB3 hyperphosphated pathologically[41].

In regard to MUC13 and MUC15, even though our data showed they might be correlated with breast cancer from diverse aspect, like mutation, methylation, overall survival and relationship with tumor immune infiltrating cells. However, little was known about how MUC13 and MUC15 affect breast cancer. And we believed it was worth investigating in depth. Briefly, MUC13 was supposed to be involved in gastric cancer[42], colon cancer[43] and ovarian cancer[44]. Aberrant expression was thought to be connected with inhibiting cancer cells apoptosis signal pathway[45]. There were studies revealed MUC15 was upregulated in colorectal tumors [46], papillary thyroid carcinoma[47] and glioma[48]. While it was downregulated in hepatocellular carcinoma due to interaction with EGFR causing PI3K pathway inactive[49].

Conclusions

MRNA Mutation Data of Muc Family Members in The Cancer Genome Atlas (TCGA) Database

TCGA is currently the largest cancer gene information database. We use the GDC tool (<https://portal.gdc.cancer.gov/>) to download breast cancer gene mutation data from the TCGA database, and download 986 breast cancer-related gene mutation samples. We use perl (version 5.28.1) to extract the relevant expression matrix, and use R software (version 4.0.2) to analyze the mutation gene matrix for the mutation frequency, mutation type and chromosome related mutation sites of the MUC gene family. We performed prognostic correlation analysis on muc family genes with mutations greater than ten.

Comparison of the mRNA Expression of the MUC Family in Breast Cancer and Normal Tissues

We downloaded the mRNA expression profile data of 1178 breast cancer-related samples from the TCGA database, including 1066 breast cancer patients and 112 adjacent tissues. We used the edgeR and gplots packages to analyze the differential expression of breast cancer and adjacent samples of muc family members. At the same time, we used the online analysis tool Sangerbox (<http://www.sangerbox.com/tool>) to conduct a correlation analysis of muc family members.

Correlation Between Methylation and mRNA Expression of the MUC Family in Breast Cancer

We downloaded the Illumina Human Methylation 27K breast cancer-related gene methylation data from the TCGA database, and used the limma package to analyze the differential expression of methylated genes. Also, we analyzed the correlation between the muc methylation level and the gene expression level.

Survival Analysis of MUC Members in Breast Cancer

The prognostic survival analysis of muc family members was divided into two aspects. On the one hand, we used online tool analysis Kaplan-Meier Plot (<http://www.kmplot.com>) to analyze the overall prognosis

of muc family members. On the other hand, in order to better study the survival of muc members in breast cancer patients with different factors (including ER, Her2, classification, staging, lymph nodes, etc.), we conducted a single-factor survival prognosis analysis.

Associations Between Tumor Immune Infiltrating Cells (TIICs) and the MUC Family

In order to study the relationship between muc gene family and tumor immune infiltrating cells, we used the Sanger Box tool to perform correlation analysis based on the TIMER platform, including (B cells, T Cells, neutrophils, macrophages, dendritic cells).

Gene Set Enrichment Analysis (GSEA)

In order to evaluate the potential mechanism of the role of MUC family members in breast cancer, we conducted a GSEA (GSEA_4.0.3 version) analysis to determine the potential pathways for the differentially expressed muc gene family of breast cancer in the TCGA database. Annotated gene set file c2.cp.kegg.v7.2.symbols.gmt (from Msig database) was a reference. The number of random combinations of muc gene differential expression was 1000 permutations and the false discovery rate (FDR) was <0.05 to determine the way of significant enrichment.

Methods:

MRNA Mutation Data of Muc Family Members in The Cancer Genome Atlas (TCGA) Database

TCGA is currently the largest cancer gene information database. We use the GDC tool (<https://portal.gdc.cancer.gov/>) to download breast cancer gene mutation data from the TCGA database, and download 986 breast cancer-related gene mutation samples. We use perl (version 5.28.1) to extract the relevant expression matrix, and use R software (version 4.0.2) to analyze the mutation gene matrix for the mutation frequency, mutation type and chromosome related mutation sites of the MUC gene family. We performed prognostic correlation analysis on muc family genes with mutations greater than ten.

Comparison of the mRNA Expression of the MUC Family in Breast Cancer and Normal Tissues

We downloaded the mRNA expression profile data of 1178 breast cancer-related samples from the TCGA database, including 1066 breast cancer patients and 112 adjacent tissues. We used the edgeR and gplots packages to analyze the differential expression of breast cancer and adjacent samples of muc family members. At the same time, we used the online analysis tool Sangerbox (<http://www.sangerbox.com/tool>) to conduct a correlation analysis of muc family members.

Correlation Between Methylation and mRNA Expression of the MUC Family in Breast Cancer

We downloaded the Illumina Human Methylation 27K breast cancer-related gene methylation data from the TCGA database, and used the limma package to analyze the differential expression of methylated

genes. Also, we analyzed the correlation between the muc methylation level and the gene expression level.

Survival Analysis of MUC Members in Breast Cancer

The prognostic survival analysis of muc family members was divided into two aspects. On the one hand, we used online tool analysis Kaplan-Meier Plot (<http://www.kmplot.com>) to analyze the overall prognosis of muc family members. On the other hand, in order to better study the survival of muc members in breast cancer patients with different factors (including ER, Her2, classification, staging, lymph nodes, etc.), we conducted a single-factor survival prognosis analysis.

Associations Between Tumor Immune Infiltrating Cells (TIICs) and the MUC Family

In order to study the relationship between muc gene family and tumor immune infiltrating cells, we used the Sanger Box tool to perform correlation analysis based on the TIMER platform, including (B cells, T Cells, neutrophils, macrophages, dendritic cells).

Gene Set Enrichment Analysis (GSEA)

In order to evaluate the potential mechanism of the role of MUC family members in breast cancer, we conducted a GSEA (GSEA_4.0.3 version) analysis to determine the potential pathways for the differentially expressed muc gene family of breast cancer in the TCGA database. Annotated gene set file `c2.cp.kegg.v7.2.symbols.gmt` (from Msig database) was a reference. The number of random combinations of muc gene differential expression was 1000 permutations and the false discovery rate (FDR) was <0.05 to determine the way of significant enrichment.

Declarations

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Availability of data and materials

Publicly available datasets were analyzed in this study, these can be found in The Cancer Genome Atlas (TCGA) program at <https://portal.gdc.cancer.gov/>.

Competing interests

The authors declare that they have no competing interests.

Funding

This research was supported by the National Key Research and Development Program of China (No. 2016YFC0905900) and National Natural Science Foundation of China (NO.81902987).

Authors' contributions

KY and ML designed the research study and analyzed the data from public database. KY, XC, and ML were involved in data analysis. KY was responsible for writing of manuscript. KY and DX contributed to the revised manuscript. All authors reviewed the manuscript.

Acknowledgements

We would like to express our sincere appreciation to the reviewers for critical comments on this article.

References

1. Siegel RL, Miller KD, Jemal A. Cancer statistics, 2020. *CA Cancer J Clin.* 2020;70(1):7–30.
2. Pasculli B, Barbano R, Parrella P. Epigenetics of breast cancer: Biology and clinical implication in the era of precision medicine. *Semin Cancer Biol.* 2018;51:22–35.
3. Gendler SJ, Spicer AP. Epithelial mucin genes. *Annu Rev Physiol.* 1995;57:607–34.
4. Fu C, et al. Tumor-associated antigens: Tn antigen, sTn antigen, and T antigen. *Hla.* 2016;88(6):275–86.
5. Li X, Xu Y, Zhang L. Serum CA153 as biomarker for cancer and noncancer diseases. *Prog Mol Biol Transl Sci.* 2019;162:265–76.
6. Aithal A, et al. MUC16 as a novel target for cancer therapy. *Expert Opin Ther Targets.* 2018;22(8):675–86.
7. Li X, et al. Association of MUC16 Mutation With Tumor Mutation Load and Outcomes in Patients With Gastric Cancer. *JAMA Oncol.* 2018;4(12):1691–8.
8. Matte I, et al. Ascites from ovarian cancer patients stimulates MUC16 mucin expression and secretion in human peritoneal mesothelial cells through an Akt-dependent pathway. *BMC Cancer.* 2019;19(1):406.
9. Shen H, et al. MUC16 facilitates cervical cancer progression via JAK2/STAT3 phosphorylation-mediated cyclooxygenase-2 expression. *Genes Genomics.* 2020;42(2):127–33.
10. Wang L, et al., *Clinical evaluation and therapeutic monitoring value of serum tumor markers in lung cancer.* *Int J Biol Markers,* 2016. 31(1): p. e80-7.
11. Davison JM, et al. MUC2 expression is an adverse prognostic factor in superficial gastroesophageal adenocarcinomas. *Hum Pathol.* 2014;45(3):540–8.
12. İlhan Ö, et al. Prognostic significance of MUC1, MUC2 and MUC5AC expressions in gastric carcinoma. *Turk J Gastroenterol.* 2010;21(4):345–52.

13. Mikhaleva LM, et al., [*Comparative assessment of the expression of Muc 2, Muc 5AC, and Muc 6 in serrated neoplasms of the colon*]. *Arkh Patol*, 2019. **81**(2): p. 10–17.
14. Shibahara H, et al. Pathobiological implications of mucin (MUC) expression in the outcome of small bowel cancer. *PLoS One*. 2014;9(4):e86111.
15. Terada T. An immunohistochemical study of primary signet-ring cell carcinoma of the stomach and colorectum: II. Expression of MUC1, MUC2, MUC5AC, and MUC6 in normal mucosa and in 42 cases. *Int J Clin Exp Pathol*. 2013;6(4):613–21.
16. Xiao LJ, et al. Clinicopathological and prognostic significance of MUC-2, MUC-4 and MUC-5AC expression in Japanese gastric carcinomas. *Asian Pac J Cancer Prev*. 2012;13(12):6447–53.
17. Andrianifahanana M, Moniaux N, Batra SK. Regulation of mucin expression: mechanistic aspects and implications for cancer and inflammatory diseases. *Biochim Biophys Acta*. 2006;1765(2):189–222.
18. Moniaux N, et al. Structural organization and classification of the human mucin genes. *Front Biosci*. 2001;6:D1192-206.
19. Geisow MJ. The coming of age of glycobiology. *Trends Biotechnol*. 1991;9(7):221–5.
20. Gendler SJ, et al. Cloning of partial cDNA encoding differentiation and tumor-associated mucin glycoproteins expressed by human mammary epithelium. *Proc Natl Acad Sci U S A*. 1987;84(17):6060–4.
21. Swallow DM, et al. The human tumour-associated epithelial mucins are coded by an expressed hypervariable gene locus PUM. *Nature*. 1987;328(6125):82–4.
22. Imai Y, et al. Differential mucin phenotypes and their significance in a variation of colorectal carcinoma. *World J Gastroenterol*. 2013;19(25):3957–68.
23. Yin BW, Lloyd KO. Molecular cloning of the CA125 ovarian cancer antigen: identification as a new mucin, MUC16. *J Biol Chem*. 2001;276(29):27371–5.
24. Ringel J, Löhr M. The MUC gene family: their role in diagnosis and early detection of pancreatic cancer. *Mol Cancer*. 2003;2:9.
25. Kufe DW. MUC1-C oncoprotein as a target in breast cancer: activation of signaling pathways and therapeutic approaches. *Oncogene*. 2013;32(9):1073–81.
26. Rahn JJ, et al. The importance of MUC1 cellular localization in patients with breast carcinoma: an immunohistologic study of 71 patients and review of the literature. *Cancer*. 2001;91(11):1973–82.
27. Ušiaková Z, et al. Circulating tumor cells in patients with breast cancer: monitoring chemotherapy success. *In Vivo*. 2014;28(4):605–14.
28. Bièche I, Lidereau R. A gene dosage effect is responsible for high overexpression of the MUC1 gene observed in human breast tumors. *Cancer Genet Cytogenet*. 1997;98(1):75–80.
29. Rajabi H, et al. MUCIN 1 ONCOPROTEIN EXPRESSION IS SUPPRESSED BY THE miR-125b ONCOMIR. *Genes Cancer*. 2010;1(1):62–8.

30. Sachdeva M, Mo YY. MicroRNA-145 suppresses cell invasion and metastasis by directly targeting mucin 1. *Cancer Res.* 2010;70(1):378–87.
31. Ramasamy S, et al. The MUC1 and galectin-3 oncoproteins function in a microRNA-dependent regulatory loop. *Mol Cell.* 2007;27(6):992–1004.
32. Bitler BG, Goverdhan A, Schroeder JA. MUC1 regulates nuclear localization and function of the epidermal growth factor receptor. *J Cell Sci.* 2010;123(Pt 10):1716–23.
33. von Mensdorff-Pouilly S, et al. Reactivity of natural and induced human antibodies to MUC1 mucin with MUC1 peptides and n-acetylgalactosamine (GalNAc) peptides. *Int J Cancer.* 2000;86(5):702–12.
34. Agrawal B, et al. Expression of MUC1 mucin on activated human T cells: implications for a role of MUC1 in normal immune regulation. *Cancer Res.* 1998;58(18):4079–81.
35. Zhang K, Sikut R, Hansson GC. A MUC1 mucin secreted from a colon carcinoma cell line inhibits target cell lysis by natural killer cells. *Cell Immunol.* 1997;176(2):158–65.
36. Alsuliman A, et al. Bidirectional crosstalk between PD-L1 expression and epithelial to mesenchymal transition: significance in claudin-low breast cancer cells. *Mol Cancer.* 2015;14:149.
37. Rajabi H, et al. MUC1-C oncoprotein activates the ZEB1/miR-200c regulatory loop and epithelial-mesenchymal transition. *Oncogene.* 2014;33(13):1680–9.
38. Price-Schiavi SA, et al. Post-transcriptional regulation of a milk membrane protein, the sialomucin complex (Ascites sialoglycoprotein (ASGP)-1/ASGP-2, rat muc4), by transforming growth factor beta. *J Biol Chem.* 1998;273(52):35228–37.
39. Sherblom AP, Buck RL, Carraway KL. Purification of the major sialoglycoproteins of 13762 MAT-B1 and MAT-C1 rat ascites mammary adenocarcinoma cells by density gradient centrifugation in cesium chloride and guanidine hydrochloride. *J Biol Chem.* 1980;255(2):783–90.
40. Jepson S, et al. Muc4/sialomucin complex, the intramembrane ErbB2 ligand, induces specific phosphorylation of ErbB2 and enhances expression of p27(kip), but does not activate mitogen-activated kinase or protein kinaseB/Akt pathways. *Oncogene.* 2002;21(49):7524–32.
41. Carraway KL, et al. Muc4/sialomucin complex, the intramembrane ErbB2 ligand, in cancer and epithelia: to protect and to survive. *Prog Nucleic Acid Res Mol Biol.* 2002;71:149–85.
42. Shimamura T, et al. Overexpression of MUC13 is associated with intestinal-type gastric cancer. *Cancer Sci.* 2005;96(5):265–73.
43. Walsh MD, et al. The MUC13 cell surface mucin is highly expressed by human colorectal carcinomas. *Hum Pathol.* 2007;38(6):883–92.
44. Chauhan SC, et al. Expression and functions of transmembrane mucin MUC13 in ovarian cancer. *Cancer Res.* 2009;69(3):765–74.
45. Sheng YH, et al. The MUC13 cell-surface mucin protects against intestinal inflammation by inhibiting epithelial cell apoptosis. *Gut.* 2011;60(12):1661–70.
46. Huang J, et al. Overexpression of MUC15 activates extracellular signal-regulated kinase 1/2 and promotes the oncogenic potential of human colon cancer cells. *Carcinogenesis.* 2009;30(8):1452–8.

47. Nam KH, et al. Expression of the membrane mucins MUC4 and MUC15, potential markers of malignancy and prognosis, in papillary thyroid carcinoma. *Thyroid*. 2011;21(7):745–50.
48. Yang WB, Li CY. Correlations of MUC15 overexpression with clinicopathological features and prognosis of glioma. *J Huazhong Univ Sci Technolog Med Sci*. 2014;34(2):254–9.
49. Wang RY, et al. MUC15 inhibits dimerization of EGFR and PI3K-AKT signaling and is associated with aggressive hepatocellular carcinomas in patients. *Gastroenterology*. 2013;145(6):1436-48.e1-12.

Tables

Table 1:Lymph nodes in breast cancer

muc	Lymph	HR(95% CL)	Log-rank P	Case
MUC12	+	1.02(0.79-1.31)	0.88	724
	-	1.26(0.86-1.86)	0.23	496
MUC16	+	0.96(0.79-1.16)	0.66	1133
	-	1.03(0.87-1.22)	0.74	2020
MUC4	+	1.04(0.86-1.27)	0.67	1133
	-	1.04(0.88-1.23)	0.64	2020
MUC3A	+	0.88(0.72-1.07)	0.19	1133
	-	0.91(0.77-1.07)	0.25	2020
MUC20	+	1.13(0.88-1.45)	0.33	724
	-	1.09(0.74-1.6)	0.67	496
MUC5B	+	1.13(0.93-1.37)	0.23	1133
	-	1(0.85-1.18)	0.99	2020
MUC13	+	0.91(0.75-1.11)	0.34	1133
	-	0.98(0.83-1.16)	0.83	2020
MUC5AC	+	1.1(0.9-1.34)	0.35	1133
	-	1.01(0.86-1.2)	0.88	2020
MUC7	+	0.99(0.82-1.21)	0.95	1133
	-	0.85(0.72-1.01)	0.067	2020
MUC15	+	1.4(1.09-1.8)	0.0083 **	724
	-	1.07(0.73-1.57)	0.74	496
MUC6	+	1.09(0.9-1.33)	0.37	1133
	-	0.94(0.79-1.11)	0.47	2020
MUC1	+	0.8(0.66-0.98)	0.027 *	1133
	-	0.92(0.78-1.09)	0.33	2020

Table 2: Stages in breast cancer

muc	Stage	HR(95% CL)	Log-rank P	Case
MUC12	I	0.83(0.29-2.47)	0.74	108
	II	1.2(0.72-1.99)	0.49	227
	III	1.04(0.76-1.41)	0.82	444
MUC16	I	1.04(0.62-1.74)	0.88	345
	II	0.97(0.76-1.23)	0.79	901
	III	1.1(0.89-1.37)	0.38	903
MUC4	I	0.91(0.54-1.54)	0.74	345
	II	0.93(0.73-1.18)	0.56	901
	III	1.11(0.9-1.39)	0.33	903
MUC3A	I	1.28(0.76-2.15)	0.35	345
	II	0.94(0.74-1.2)	0.62	901
	III	1.03(0.83-1.28)	0.81	903
MUC20	I	0.72(0.25-2.06)	0.53	108
	II	1.16(0.7-1.92)	0.57	227
	III	1.28(0.94-1.75)	0.12	444
MUC5B	I	1.5(0.88-2.54)	0.13	345
	II	0.91(0.71-1.15)	0.42	901
	III	1.12(0.9-1.39)	0.32	903
MUC13	I	0.75(0.44-1.26)	0.28	345
	II	0.91(0.71-1.16)	0.43	901
	III	0.92(0.74-1.14)	0.46	903
MUC5AC	I	0.95(0.56-1.59)	0.84	345
	II	0.97(0.77-1.24)	0.82	901
	III	1.1(0.89-1.37)	0.37	903
MUC7	I	0.72(0.42-1.21)	0.21	345
	II	0.92(0.72-1.17)	0.48	901
	III	1.01(0.81-1.25)	0.94	903
MUC15	I	0.52(0.17-1.55)	0.23	108

	II	1.48(0.89-2.46)	0.13	227
	III	1.13(0.83-1.55)	0.43	444
MUC6	I	1.67(0.98-2.83)	0.056	345
	II	0.95(0.74-1.2)	0.65	901
	III	1.19(0.96-1.48)	0.11	903
MUC1	I	0.77(0.46-1.31)	0.34	345
	II	0.98(0.77-1.24)	0.86	901
	III	1.35(1.09-1.68)	0.0069 **	903

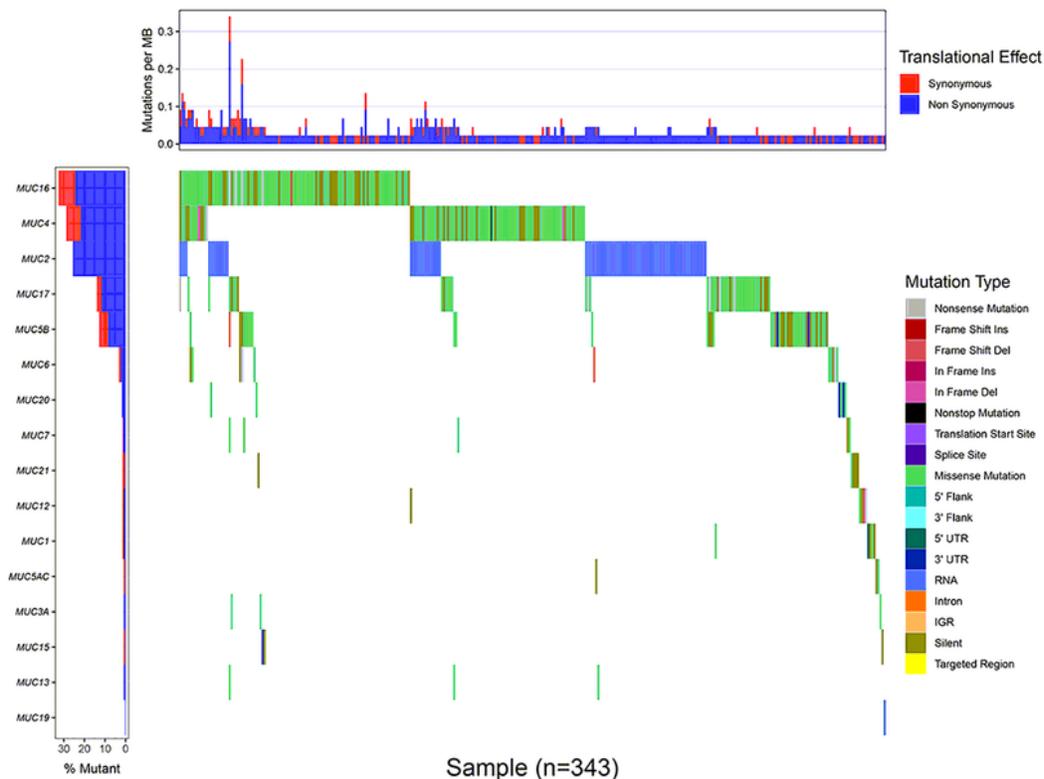
Tables 3:Types in breast cancer

muc	Type	HR(95% CL)	Log-rank P	Case
MUC12	TN	0.77(0.56-1.07)	0.12	360
	LuminalA	0.92(0.72-1.18)	0.53	841
	LuminalB	0.85(0.63-1.16)	0.32	307
	Her2+	0.96(0.61-1.52)	0.87	156
MUC16	TN	1.01(0.78-1.29)	0.96	618
	LuminalA	0.87(0.73-1.03)	0.11	1933
	LuminalB	0.85(0.7-1.03)	0.099	1149
	Her2+	0.93(0.64-1.37)	0.73	251
MUC4	TN	0.88(0.69-1.14)	0.34	618
	LuminalA	0.9(0.76-1.07)	0.23	1933
	LuminalB	0.95(0.78-1.15)	0.6	1149
	Her2+	0.79(0.53-1.16)	0.22	251
MUC3A	TN	0.9(0.7-1.15)	0.39	618
	LuminalA	0.69(0.58-0.81)	1.4E-05 ***	1933
	LuminalB	0.88(0.73-1.06)	0.19	1149
	Her2+	0.62(0.42-0.91)	0.015 *	251
MUC20	TN	1.15(0.83-1.59)	0.4	360
	LuminalA	0.87(0.68-1.11)	0.27	841
	LuminalB	0.72(0.53-0.98)	0.035 *	407
	Her2+	0.72(0.46-1.14)	0.16	156
MUC5B	TN	1.12(0.87-1.44)	0.38	618
	LuminalA	1(0.84-1.18)	0.97	1933
	LuminalB	1.04(0.86-1.26)	0.68	1149
	Her2+	0.84(0.57-1.24)	0.38	251
MUC13	TN	0.78(0.6-1)	0.05	618
	LuminalA	0.9(0.76-1.07)	0.23	1933
	LuminalB	0.7(0.58-0.85)	0.00027 ***	1149
	Her2+	0.74(0.51-1.09)	0.13	251

MUC5AC	TN	1.06(0.83-1.37)	0.63	618
	LuminalA	0.99(0.84-1.17)	0.92	1933
	LuminalB	0.91(0.75-1.1)	0.31	1149
	Her2+	0.57(0.38-0.84)	0.004 **	251
MUC7	TN	0.93(0.72-1.2)	0.57	618
	LuminalA	0.8(0.67-0.95)	0.0092 **	1933
	LuminalB	1.09(0.9-1.32)	0.38	1149
	Her2+	0.8(0.54-1.17)	0.24	251
MUC15	TN	1.06(0.76-1.46)	0.74	360
	LuminalA	0.96(0.75-1.23)	0.76	841
	LuminalB	0.93(0.69-1.27)	0.66	407
	Her2+	0.98(0.62-1.54)	0.92	156
MUC6	TN	1.05(0.82-1.35)	0.7	618
	LuminalA	0.93(0.78-1.1)	0.39	1933
	LuminalB	0.95(0.79-1.16)	0.63	1149
	Her2+	0.93(0.64-1.37)	0.72	251
MUC1	TN	1.47(1.14-1.9)	0.0026 **	618
	LuminalA	0.84(0.71-0.99)	0.042 *	1933
	LuminalB	1.24(1.02-1.51)	0.027 *	1149
	Her2+	1.22(0.83-1.79)	0.3	251

Figures

A



B

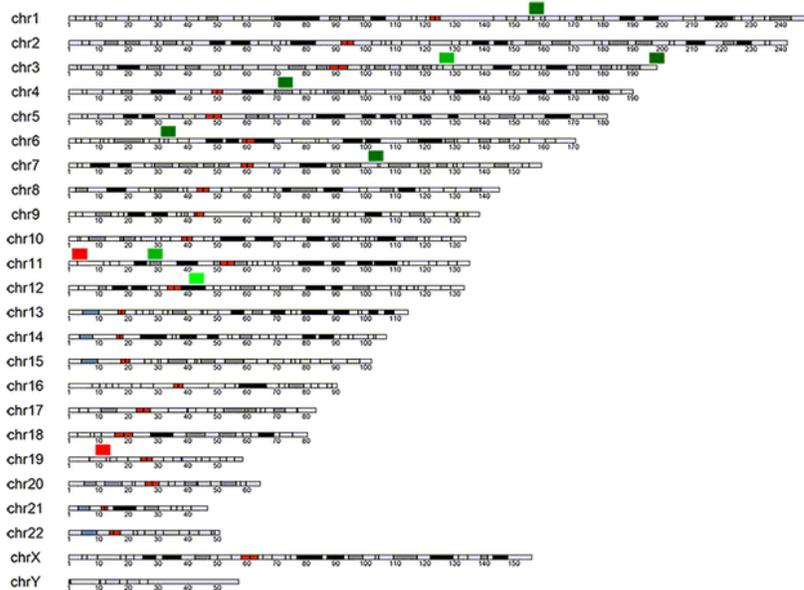


Figure 1

(A)The mutations of MUC family in BC.(B) Frequency of muc gene mutations in BC on human's chromosomes. (The darker the color, the higher the frequency.Red is the highest).

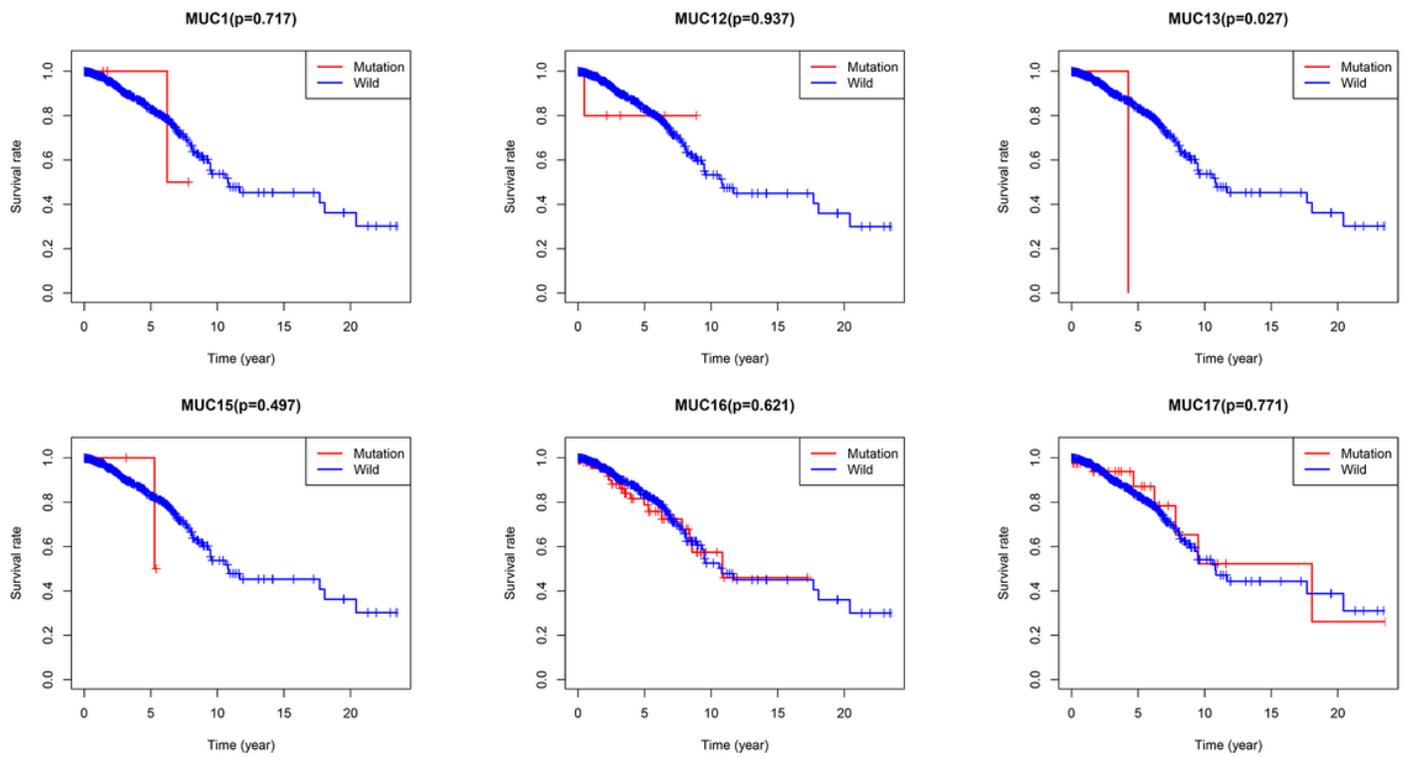


Figure 2

Survival curve of muc gene mutation in breast cancer patients.(Breast cancer patients with muc gene mutations have shorter survival times).

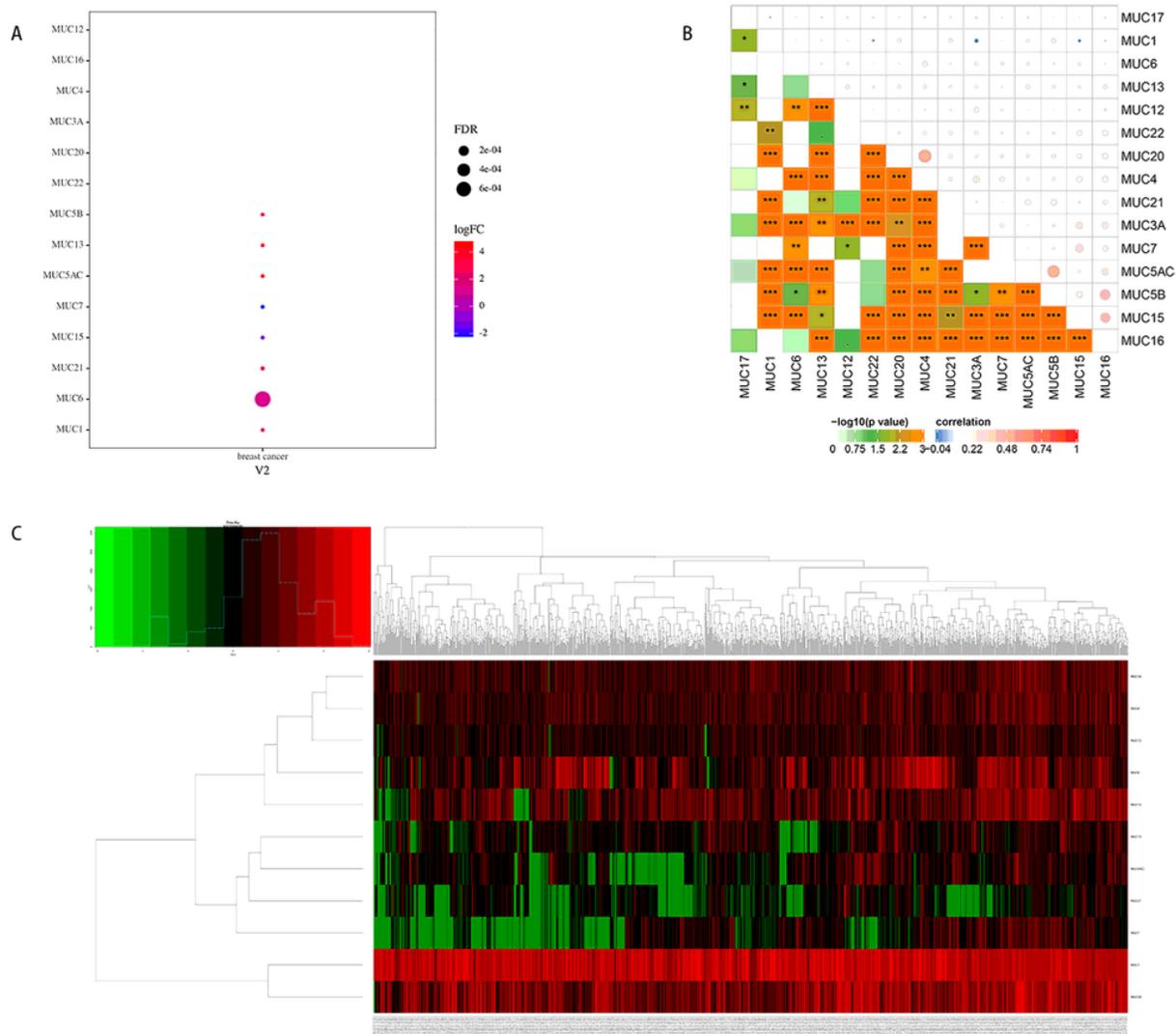


Figure 3

The expression of MUC in BC.(A) Differential expressions of MUC family.(B) Correlation analysis of muc members.(C)The heatmap of differentially expressed MUC.

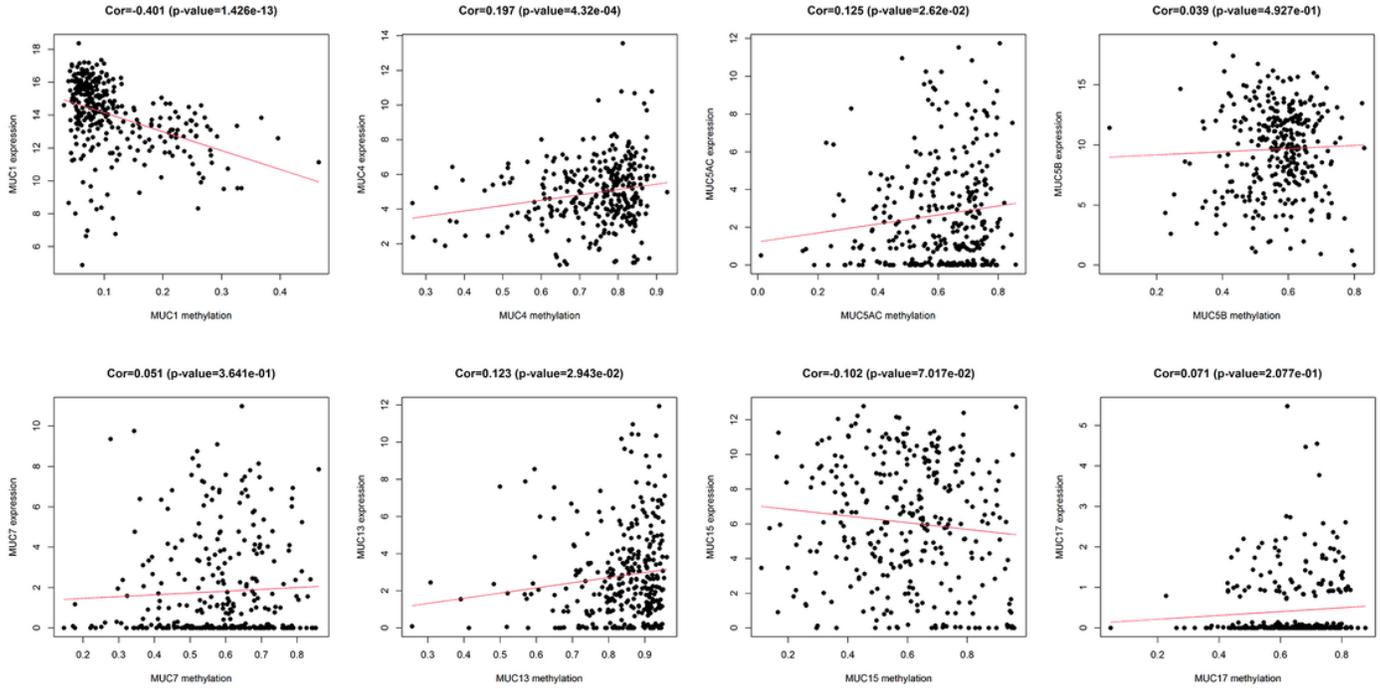


Figure 4

Correlation between methylation and mRNA expression of the MUC family.

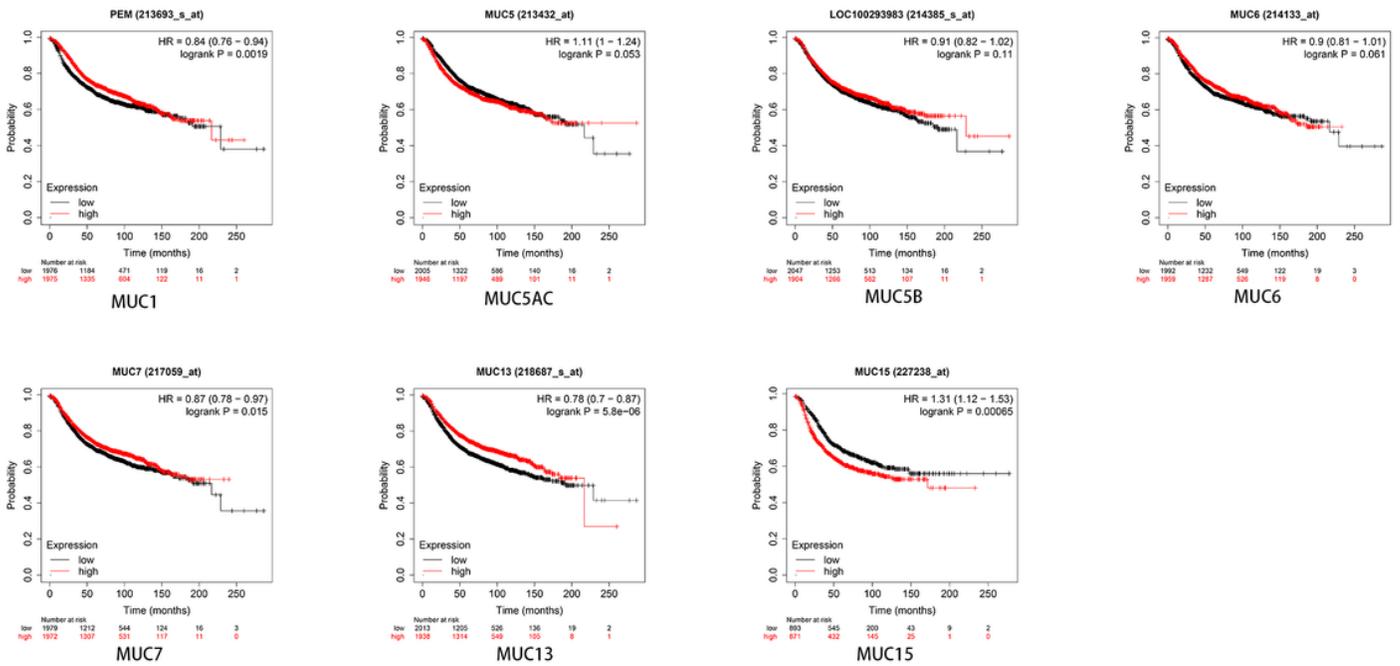


Figure 5

Survival curve between high and low expression of muc gene.

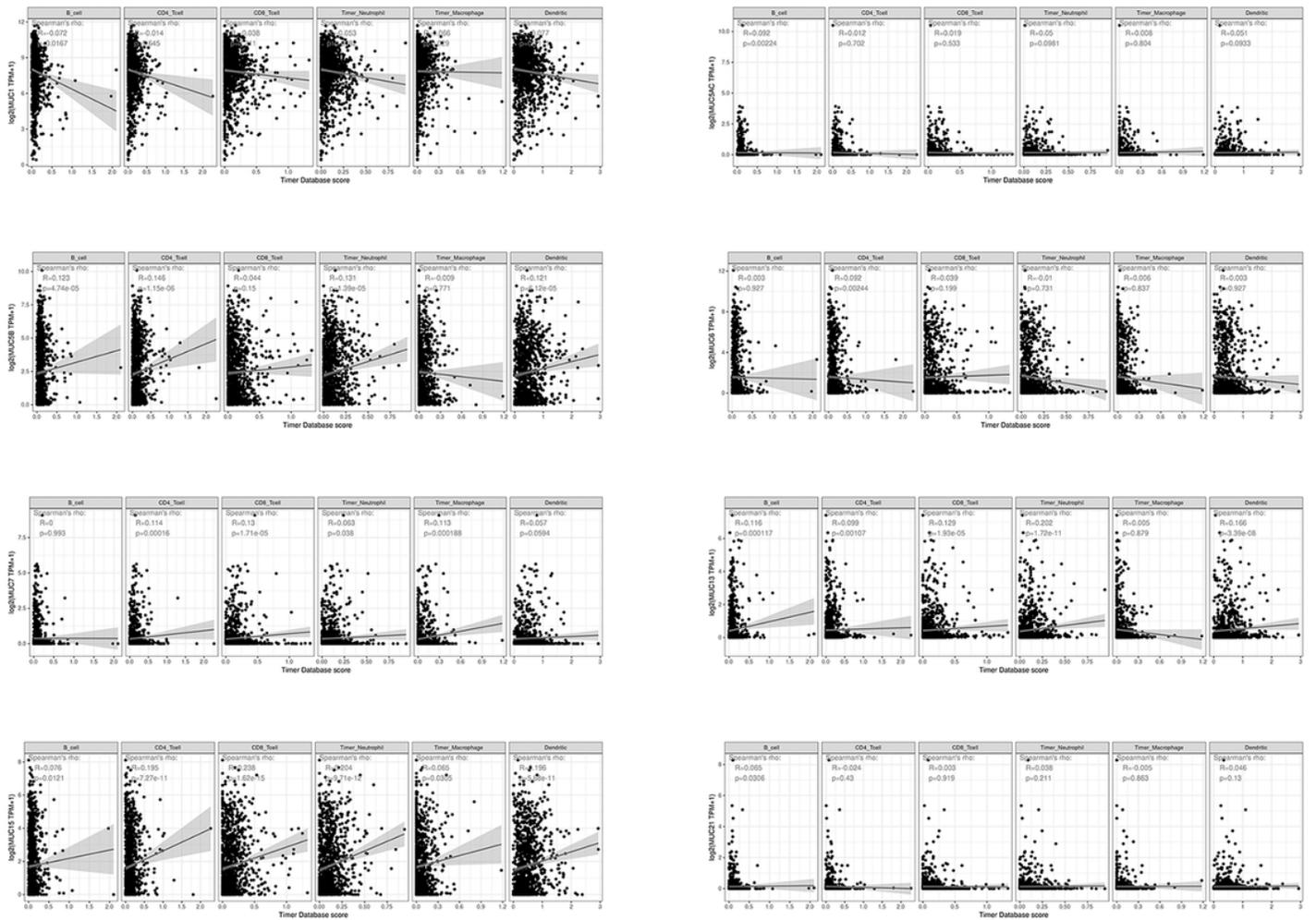


Figure 6

Associations between tumor immune infiltrating cells (TIICs) and the MUC family.

