Correlation between metabolic characteristics and breast cancer subtypes and prognosis

Yuwei Liao  
People's Hospital of Yangjiang

Zhikai Chen  
Hanshan Normal University

Hongfeng Liang  
People's Hospital of Yangjiang

Liye Yang (✉️ yangleeyee@sina.com)  
People's Hospital of Yangjiang

Research Article

Keywords: Breast cancer, abnormal metabolism, PAM50 subtype, prognosis

Posted Date: February 18th, 2022

DOI: https://doi.org/10.21203/rs.3.rs-1364522/v1

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

Background

Breast cancer (BRCA) is a common cancer and the major cause of cancer mortality in females around the world. One important characteristic of BRCA is metabolic reprogramming; however, abnormal metabolism of breast cancer based on large samples has not been reported. In addition, the relationship between metabolic features and the prognosis of BRCA still needs to be explored.

Methods

In this study, we ran dimension reduction for The Cancer Genome Atlas (TCGA) training set from 943 metabolic genes to 100 principle components (PCs) by principle component analysis (PCA). The Pearson correlations were calculated between PCs and clinical information. The least absolute shrinkage and selection operator (LASSO) Cox regression was used to sort out prognostic PCs. The top-loading genes of PCs validated successfully in TCGA validation set were further used for gene enrichment analysis and the abundance of immune cells between high- and low-PC BRCA patients was compared. Finally, the key genes and prognostic PCs in our result were validated by using Joel dataset.

Results

Five metabolic PCs related to subtypes and prognosis of BRCA were found. PC2 and PC4, in which B3GNT5 and RAMP1 were prognostic genes, correlated with basal-like subtypes. PC2 was also correlated with ER- or PR-positive status, indicating the endocrine resistance after therapy. In addition, CCL19 was lowly expressed in Luminal B (LumB) subtype. PAPSS2, enriched in nucleoside bisphosphate biosynthetic process, was a prognostic gene of BRCA and hub metabolic gene exclusively expressed in HER2-positive subtype. PC7, PC13 and PC29 were associated with the prognosis of BRCA and the changes of tumor immune environment. CCR7 chemokine axis, purine metabolism and cytidine catabolic process were the key metabolic pathways according to these prognostic PCs.

Conclusions

Our study revealed that the metabolic PCs were related to subtypes and prognosis of patients. The key pathways and genes of these PCs were potential BRCA biomarkers, which may provide a new direction for the future treatment of BRCA.

Background

Breast cancer (BRCA) is the most common cancer in women around the world [1]. Although the mortality of BRCA has decreased, BRCA is still a leading cause of cancer death in females aged 20–59 years [2–
During malignant transformation of BRCA, metabolic reprogramming was the main characteristic and was different among BRCA subtypes [5]–[8]. Recent studies also revealed that abnormal metabolism was closely associated with dysfunctional immunity in various cancer types [6][9]. However, abnormal metabolism of BRCA in large samples has not been reported yet. In addition, the relationship is also largely unknown between metabolic features and the prognosis of BRCA.

Here, we analyzed 1039 BRCA samples in The Cancer Genome Atlas (TCGA) by principal component analysis (PCA). We identified metabolic PCs that were correlated with histological type, PAM50 subtypes, ER status, PR status, HER2 status and prognosis of BRCA. Moreover, we sorted out the hub genes of these PCs, which could be biomarkers for prognosis and treatment of BRCA.

**Methods**

**Acquisition of information from patients with BRCA**

We obtained TCGA RNA sequencing data (expected count of genes) and relevant clinical information of BRCA patients from the UCSC Xena [10]. To reduce statistical bias in this analysis, BRCA patients without overall survival (OS) or age information were excluded.

The microarray data from Joel’s research [11] was downloaded from GEO (GEO data set GSE10886) for additional validation. Because the follow-up time of patients in the Joel dataset was relatively short (max 118 days), we chose Relapse Free Survival (RFS) instead of OS as a measure of patient’s survival. BRCA patients without RFS or age information were excluded. PAM50 subtypes in Joel dataset were predicted according to the codes provided in their research [11].

**Selection of metabolic-related gene sets**

Four metabolic-related gene sets were downloaded from the “Molecular Signatures Database” module of the GSEA website [12]. The gene sets were as follows: GOBP carbohydrate derivative biosynthetic process, GOBP carbohydrate derivative catabolic process, GOCC lipid droplet, and KEGG fatty acid metabolism (Table S1). A total of 943 metabolic genes in TCGA dataset were used for further analysis.

**Preprocess for expression data of TCGA dataset and Joel dataset**

The TCGA dataset of 943 metabolic genes was randomly divided into training set (75%) and validation set (25%). The expression matrix of the training and validation sets was normalized by counts per million (CPM). Then normalized data was used to run dimension reduction in the training set to 100 PCs by principal component analysis (PCA) using the “prcomp” function in R [13].

For some genes which were not included in Joel dataset but in the 943 metabolic genes, we all replaced them with zero in Joel dataset. Then Joel dataset was scaled by library size normalization and per gene median normalization. The loadings of 100 PCs from training set were used to predict PCs for validation set and Joel dataset.
**Correlation between PCs and subtypes in BRCA**

Finally, the Pearson Correlation Coefficients between PCs and clinical information of BRCA patients were calculated in all datasets. In this process, all categorical variables of clinical information were assigned different levels and converted to the one-hot encoding format. Infiltrating ductal carcinoma, infiltrating lobular carcinoma, and mucinous carcinoma were regarded as level 1 to level 3, respectively. Estrogen receptor (ER)/progesterone receptor (PR)-positive and HER2-negative were considered as level 1 while ER/PR-negative and HER2-positive were level 2. Patients with indeterminate ER/PR status or equivocal HER2 status were filtered out. For classification of the prediction analysis of microarray 50 (PAM50), Luminal A (LumA), Basal, Luminal B (LumB), HER2, and Normal subtypes were regarded as level 1 to level 5, respectively. The levels of these factors above were determined by the results of multivariate Cox regression in the BRCA training set. Then PCs with significant correlation (cor > 0.2 or cor <-0.2) in both training set and validation set were sorted out. Analysis of Variance (ANOVA) was used to determine the significance of the different BRCA subtypes for these PCs.

**Comprehensive prognostic model based on the prognostic metabolic PCs**

The 100 PCs and clinical variables in the training set were firstly normalized by Z-score normalization. LASSO Cox regression was used to eliminate variables unrelated to the survival of BRCA using the R package glmnet [14]. Age, histological type, and 13 prognostic PCs were sorted out to establish the comprehensive risk model by multivariate Cox regression. The training set was utilized to construct a prognostic model while validation set and Joel dataset were applied to validate this established model. The following formula was used to calculate the risk scores:

\[
\text{Risk score} = \text{coef (PC3)} \times \text{PC3} + \text{coef (PC4)} \times \text{PC4} + \ldots + \text{coef (PCn)} \times \text{PCn} + \text{coef (Age)} \times \text{Age} + \text{coef (histological type)} \times \text{histological type},
\]

where coef indicated the coefficient and PCn indicated the PC correlated with survival. According to the median risk scores, patients were divided into low- and high-risk groups. For Joel dataset, “coef (histological type) x histological type” was considered to be zero because of the missing of histological type.

**Kaplan-Meier analyses between two groups**

Patients were divided into two groups according to different criteria, such as risk score of the prognostic model, CPM (Counts Per Million) of gene expression, and values predicted by PCs. R package survival [15] and survminer [16] were used for Kaplan-Meier analyses in this study. The p values were calculated by log-rank test.

**Forest plot**

R package survminer [16] was used to draw the forest plot for the comprehensive 13-PCs Cox proportional risk model.

**Functional analysis**
Generally, genes carried different loadings when they were used for predicting PCs. For each PC, KEGG enrichment analysis and Gene Ontology (GO) enrichment analysis were performed for both positive and negative top 20 loading genes with the R package GDCRNATools [17], KEGG website [18] as a supplementary. \( P<0.01 \) indicated that the pathway was significantly enriched. The top two significant pathways were shown in Venn Diagrams. Only the uppermost pathway was retained if the genes in the two pathways were too similar. If no significant pathway was enriched, top 30 loading genes were involved in the functional analysis.

**Exploration of the prognostic PCs in the immune microenvironment**

The ImmuCellAI [19] was used to estimate cell enrichment score. Wilcoxon rank-sum test was used to analyze the difference of cell enrichment score between the high- and low-PC groups. We also tested whether the gene expression was significantly different between the groups of high- and low-Tr1 cell abundance.

**Comparing the expression of the metabolic genes among the different subtypes of BRCA**

ANOVA was used to test whether the gene expression was significantly different among the different subtypes of BRCA.

**Results**

**Principal component analysis (PCA) of the BRCA expression data of metabolic genes**

The entire workflow was shown in Figure S1. TCGA BRCA expression data (TCGA dataset) was downloaded from UCSC Xena and separated into training set (75%) and validation set (25%). PCA was run in the training set using 943 genes related to carbohydrate metabolism and lipid metabolism. Gene loadings (Table S2) of the top 100 principal components (PCs) in the training set were used to predict PCs for the validation set and Joel dataset. To confirm that PCs could represent the difference among different samples, the first three PCs were used to display the three-dimensional distance. Expectedly, different PAM50 subtypes were separated (Figure 1A-B), indicating the heterogeneity and biological significance of metabolic PCs among BRCA patients.

**Metabolic PCs correlated with clinical information of BRCA patients**

The Pearson Correlation Coefficients were calculated between the clinical information and PCs. Most PCs were associated with histological type, PAM50 subtype, PR status, ER status, and HER2 status in both the training and validation set while no PCs were correlated with tumor stage, age, or tumor metastasis (Table S3).

PC1 and PC12 were correlated with histological type, with p-value less than 2.2e-16 and 4.1e-6, respectively (Figure 2A). The Z-score of infiltrating lobular, infiltrating ductal and mucinous subtype
increased in order in PC1. On the other hand, the Z-scores of infiltrating lobular and mucinous subtypes were similar but higher than infiltrating ductal subtype in PC12.

We divided patients in Joel dataset into PAM50 subtypes according to Joel's method. Then Joel dataset was used for validating the correlation between PCs and PAM50 subtypes. PC1 (p < 2.2e-16), PC2 (p < 2.2-16), and PC4 (p < 2.2-16) were significantly different (Figure 2B-C) among different PAM50 subtypes in both TCGA dataset and Joel dataset. In PC1, the normal subtype was the lowest while LumB subtype was the highest. More importantly, LumB and LumA subtypes could also be distinguished (Figure 2B-C). In PC2, the basal subtype was the largest, followed by HER2 and normal subtypes, and the two luminal subtypes had the least PC2 (Figure 2B-C). In addition, PR-positive and negative patients could be distinguished by PC2 (Figure 2D, p < 2.2e-16), which was capable of separating ER-positive and negative patients (p < 2.2e-16, Figure 2E). In PC4, there was a little difference between TCGA dataset and Joel dataset. The Z-score of normal subtype was not the highest in TCGA dataset (Figure 2B) but was the highest in Joel dataset. The Z-score of basal subtype was the highest in TCGA dataset and the second highest in Joel dataset (Figure 2C). In both two datasets, HER2 was the lowest (Figure 2B-C). Correspondingly, the Z-score in PC4 was lower in HER2-positive patients compared with the counterparts (p = 1.9e-07, Figure 2F). Conversely, HER2-positive patients showed higher Z-scores in PC6 (p < 2.2e-16) and higher PC78 (p = 3.3e-05) (Figure 2F).

**Metabolic PCs correlated with survival of BRCA patients**

Based on the top 100 PCs and clinical information in the training set, the univariate least absolute shrinkage and selection operator (LASSO) model was built. 13 prognostic PCs were sorted out and used to build a multivariate Cox regression model. The Hazard ratios and p values of 13 PCs in multivariate Cox regression were shown in Figure 3A. Based on this prognostic model, prognostic risk scores were calculated for patients in the training and validation sets. Then the patients were divided into high- and low-risk groups according to the median of the prognostic risk scores in each dataset. A significant difference in survival rates was found between high- and low-risk patients in both training (Figure 3B, P < 0.0001) and validation sets (Figure 3C, p = 0.003).

The multivariate Cox regression and forest plot were conducted in the validation set using the PCs in the prognostic model. The p values were significant in PC7, PC13, and PC29, with p =0.0119, 0.0012, and 0.0056, respectively (Figure S2). We further divided the patients into high- and low-PC groups and investigated whether the three PCs were independently correlated with the survival of BRCA patients. The results showed that only PC13 was significant in survival analysis in both training set (p = 3e-04, Figure S3C) and validation set (p = 0.0063, Figure S3D).

**GO and KEGG enrichment for PCs**

To depict the biological characteristics of these PCs, the top 20 positive- and 20 negative-loading genes in these PCs were used for KEGG enrichment and GO enrichment. PC7, PC13, and PC29, which were significant in both training and validation sets by multivariate Cox regression (Figure 3A, Figure S2), were
considered as prognostic PCs and used to run enrichment analysis. All enrichment results were shown in Table S4. For PCs without significantly enriched KEGG pathway (FDR < 0.01), such as PC4, PC6, PC13, PC14, and PC29, top 30 positive- and 30 negative-loading genes were used for enrichment.

The enrichment of positive loading genes of PC1 included the PPAR signaling pathway, regulation of lipolysis in adipocytes, lipid droplet, and lipid catabolic process (Figure 4A, Table S4). On one hand, mucinous subtype and LumB subtype patients had the high activity of these metabolic features with higher Z-scores in PC1. On the other hand, pyrimidine metabolism, carbohydrate derivative catabolic process, fructose, and mannose metabolism, enriched by the negative loading genes, were up-regulated in infiltrating lobular and normal subtypes, both of which had low PC1 values (Figure 4B, Table S4).

According to the enrichment of genes in PC12, chemokine signaling pathway, other types of O-glycan biosynthesis, and cAMP metabolic process were up-regulated in infiltrating lobular and mucinous subtypes while fatty acid metabolism and ferroptosis were up-regulated in infiltrating ductal subtype (Figure 2A, Figure 5C-D, Table S4).

Top-loading genes of PC2 were positively enriched in glycosphingolipid biosynthesis and glycosaminoglycan biosynthesis (Figure 4C). Upregulation of these pathways was a metabolic feature of the basal subtype because of its high PC2 values. Regulation of lipolysis in adipocytes, endocrine resistance pathway, and sphingolipid biosynthetic process were enriched in negative loading genes of PC2 (Figure 4D), indicating that they are metabolically active in subtypes with lower PC2 values, such as ER/PR-positive BRCA patients, LumA, and LumB patients (Figure 2B-E). Glycoprotein metabolic process, glycoprotein biosynthetic process, protein glycosylation, and Golgi lumen were positively or negatively enriched in PC4 (Figure 4 E-F, Table S4), whose Z-scores were higher in basal subtype (Figure 2B-C), but lower in HER2 positive BRCA (Figure 2F).

To investigate the effect of top-loading genes of PCs on the survival of BRCA patients, we ran the Kaplan-Meier survival analysis by comparing patients with high and low expression of each top-loading gene. The p values of survival difference (Table S5) between the two groups were calculated by log-rank test. Genes were marked with colors if the p values were less than 0.05. Accordingly, some genes in these PCs were independently associated with a better prognosis of patients, such as CCL19 in PC4, MT3, HYAL3, UQCC3 in PC1, MMP12, RAMP1 in PC2, FUT7, TYMP, and CCR7 in PC4, while B3GNT5 in PC2, MUCL1 in both PC1 and PC4 were associated with a worse prognosis (Figure 4A-F).

HER2 subtype seemed to be more complicated. PC4, PC6, and PC78 were correlated with HER2-positive BRCA patients (Figure 2F). PC6, correlated with higher expression of lipid droplet pathway, PPAR signaling pathways, and ferroptosis, was also relevant to lower expression of genes enriched in glycoprotein metabolic process and other types of O-glycan biosynthesis (Figure 5A-B, Table S4). PC78 was correlated with higher activity of fatty acid metabolism, nucleoside bisphosphate biosynthetic process, and glycoprotein metabolic process, but with lower activity of endocrine resistance, mannose type O-glycan biosynthesis, glycoprotein biosynthetic process, and glycosaminoglycan metabolic process (Figure 6C-D). Prognostic genes in PC6 and PC78 included MUCL1 and MT3 in glycoprotein metabolic
process, ACSL and ALOX15 in ferroptosis, EHHADH in fatty acid metabolism, PAPSS2 in nucleoside bisphosphate biosynthetic process, FUT9 in both glycoprotein biosynthetic process and mannose type O-glycan biosynthesis (Figure 5A-B, Figure 5C-D). But ADCY8 in PC6 and PC78 was not enriched in any pathway.

PR-positive patients had lower PC2 values (Figure 2D) while ER-positive patients tended to have lower values in both lower PC2 and PC14 (Figure 2E). The glycosphingolipid biosynthesis – lacto and neolacto pathway was enriched in both PC2 and PC14 (Figure 4C, Figure 6B), indicating the important role of this pathway in ER and PR status.

The pathways in PC7, PC13, and PC29 were correlated to the better prognosis of patients. Chemokine signaling pathway, cytidine deaminase activity, and glycoprotein biosynthetic process were enriched in negative loading genes of PC7 (Figure 7B, Table S4). Glycoprotein metabolic process, positive regulation of innate immune response, and fatty acid elongation were enriched by the negative loading genes of PC13 (Figure 7D, Table S4). Aminoglycan catabolic process and protein glycosylation were enriched by the positive loading genes of PC29 (Figure 7F, Table S4).

In contrast, several pathways were correlated with poor survival of patients, such as glycoprotein metabolic process, other types of O-glycan biosynthesis, mucin-type O-glycan biosynthesis and purine metabolism in PC7, retrograde endocannabinoid signaling, membrane lipid metabolic process, glycosaminoglycan biosynthesis- heparan sulfate/heparin, and protein glycosylation in PC13, purine metabolism, other types of O-glycan biosynthesis, and mucopolysaccharide metabolic process in PC29 (Figure 7A, C, E, Table S4). It was noteworthy that purine metabolism was correlated with poor survival of patients in both PC7 and PC29 (Figure 7C, E), indicating that purine metabolism may be an important factor in the survival of BRCA patients.

**Hub genes correlated to the survival of BRCA patients**

To know whether the top-loading genes of subtype-related PCs were correlated with the corresponding subtypes, we selected several prognostic genes for validation. B3GNT5, a positive loading gene in PC2, had high expression in the basal subtype (Figure S4A). RAMP1, a negative loading gene in PC2, was significantly upregulated in LumA and LumB (Figure S4B) while PC2 was lower in LumA and LumB subtypes (Figure 2B-C). CCL19, negative loading gene of PC4, was not lowly expressed in basal subtype but was lowly expressed in LumB subtype (Figure S4C) while PC4 was the highest in basal subtype and LumB subtype was the second higher (Figure 2B-C). The expression of ALOX15 was higher than basal subtype while PC6 was high in HER2 positive patient (Figure S4D). ADCY1, ADCY5, and BCL2, negative loading genes of PC2, were all significantly highly expressed in ER-positive BRCA patients (Figure S4E-G) while PC2 was low in ER positive patients (Figure 2E). PAPSS2 in PC78 positive loading genes was highly expressed in HER2-positive patients (Figure S4I), but EHHADH in PC78 positive loading genes showed no difference between HER2-positive and negative patients (Figure S4H), while PC78 was high in HER2 positive patients (Figure 2F). Therefore, the relationship between PCs and clinical information did not
mean that the top-loading genes of PCs were related to the clinical information and it could only imply the possible relationship.

Next, to sort out the hub prognostic genes that may have therapeutic potential, we examined the top-loading genes in PCs correlated with survival of BRCA patients and enriched in more than one pathway (Figure 7, Table S3). CCR7 and CCL19, PC7’s negative top-loading genes, were correlated with good prognosis and enriched in the CCR7 chemokine axis, chemokine signaling pathway, and glycoprotein biosynthetic process (Figure 7A). Specifically, CCL19 was a negative top-loading gene in two survival-related PCs, PC7 and PC13 (Figure 7B, Figure 7D), and was also downregulated in LumB subtype (Figure S4B). APOBEC3D and APOBEC3G, PC7’s negative loading genes, were enriched in the cytidine catabolic process and correlated with a better prognosis of patients (Figure 7A). The CCR7 chemokine axis was comprised of CCL21 and CCL19 [20]. The CCR7 chemokine axis played an important role in orchestrating the inflammatory and immunological responses [21] as well as cancer progression [20], [22]–[25]. Since PC7’s top negative loading genes that were enriched in the CCR7 chemokine axis were correlated with a worse outcome of BRCA patients, this axis could be a good indicator for the prognosis of BRCA patients. As genes in other pathways might interact with those in the chemokine axis, we sorted out patients with high expression of genes in this axis and investigated whether other top-loading genes in PC7 affected the survival of these patients. As expected, patients had a better prognosis with high expression of CCL21 if other negative loading genes in PC7 were upregulated, such as APOBEC3D, APOBEC3H, or FUT7 (Figure S6 A-C, Table S6). Conversely, they would have a worse prognosis if the positive loading genes in PC7 were overexpressed, such as MUCL1, GPC6, or GALNT15 (Figure S6 D-F, Table S6). Similarly, patients with high expression of CCL19 also had a poor prognosis if GPC6 or GALNT15 was highly expressed (Figure S6 G-H, Table S6).

MUCL1, which was related to the poor survival of patients with BRCA, was the top-loading gene in three prognostic PCs (Figure 7A, D, and E). MMP12 was associated with both glycoprotein metabolic process and positive regulation of innate immune response, and its upregulation indicated a better prognosis. ADCY6 and PAPSS2 in PC13 and PC29 were enriched in purine metabolism and they were correlated with poor survival of BRCA patients (Figure 7C, E). ELOVL2, an important component in PC2, PC7, PC13, PC29, and PC78, seemed to be an indicator for good prognosis in early time but bad prognosis in late time after 3000 to 4000 days (Figure S5A-B).

Cell abundance analysis for patients with high prognosis-related PC values

Patients were divided into high and low groups according to the three validated prognostic PCs. Cell abundance for patients was calculated and compared between the two groups. Patients with high value in PC7 had less abundance in exhausted T (Tex) cells, central memory T (Tcm) cells, gamma delta T cells (Tgd), B cells, induced regulatory T cells (iTreg) cells, CD4 T cells, CD8 T cells, T helper 1 (Th1) cells, natural killer (NK) cells, follicular helper T cells (Tfh) cells and cytotoxic T cells (Tc) (Figure 8A), but showed more abundance in Natural Killer T (NKT) cells, CD8 naive cells, T helper 17 (Th17) cells, dendritic cells (DC), neutrophil cells, monocyte cells and macrophage cells (Figure 8B). These patients
seemed to have less cytotoxic T cell infiltration, consistent with enrichment of negative top-loading genes of PC7 in the CCR7 chemokine axis.

The abundance was higher for Type 1 regulatory T (Tr1) cells and macrophage cells but lower for Mucosal-associated invariant T (MAIT) cells and T helper2 (Th2) cells in patients with high values in PC13 (Figure 8C). Interestingly, cell abundance of Tr1 was also higher in patients with high values in PC29 (Figure 8D), who had worse prognostic outcomes (p=0.068, Figure S7A). The purine metabolism, enriched by the positive top-loading genes in both PC13 and PC29, might play an important role in BRCA patients with high Tr1 cell abundance. Therefore, we compared the expression of purine metabolism-related genes between high and low Tr1 cell abundance groups. The result showed that ADCY7 and PAPSS2 but not ADCY6 were correlated with cell abundance of Tr1 (Figure S7C-D), indicating that metabolic-related PCs and genes were correlated with immunity environment and survival of BRCA patients.

Validation for hub genes and prognostic model by using Joel dataset

Hub genes included in both TCGA dataset and Joel dataset were used for validation. As shown in Figure 9A-E, the Expression of B3GNT5, RAMP1 and BCL2 had significant difference between different PAM50 subtype in Joel dataset (p < 2.2e-16, p = 2e-04, p = 8e-05, respectively). The median expression of PAPSS2 and ALOX15 were high in HER2 subtype although the p values of difference between PAM50 subtypes were not significant (p=0.13 and p=0.098 respectively).

For the reason of short follow-up time, we used RFS instead of OS for validation in Joel dataset. Although the p values were not all significant, we could still get the trends (Figure 9 F-L).

Discussion

Abnormal metabolism was an essential hallmark of breast cancer [6], [26], [27], whose genomic characteristics of different histological types had been documented [28], [29]. But the relationship between abnormal metabolism and survival of BRCA patients was still unclear. The metabolic difference among different histological types or PAM50 subtypes was also scarcely reported in BRCA. In our study, several metabolic PCs correlated with subtypes and prognosis of patients were sorted out.

Of note, although the basal-like subtype of BRCA had a poor prognosis than LumB subtype at 5-years of follow-up, lumB subtype had worse outcome at 10-years[30]. We found that PPAR signaling pathway, regulation of lipolysis in adipocytes, lipid droplet, and lipid catabolic process were highly expressed in LumB subtype. In addition, the expression of CCL19 was lower in LumB subtype compared with other subtypes of BRCA. It had been reported that regulation of lipolysis in adipocytes promoted metabolic remodeling and stimulated the invasiveness of BRCA [31]. PPAR signaling pathway was down-regulated in BRCA compared with normal tissue [32][12]. BRCA patients with low CCL19 expression tended to have a worse prognosis [36]. These findings were consistent with our results.
The basal-like subtype, which accounted for 80% of triple-negative breast cancer [33], was characterized by high expression of proliferation-related genes, keratins typically expressed in skin, moderate expression of HER2-related genes, and low expression of luminal-related genes [33], [34]. The basal subtype had a higher PC2 value, thus, the corresponding cohort had the following metabolic features: lower activity in the regulation of lipolysis in adipocyte, endocrine resistance, and sphingolipid biosynthetic process; higher activity in glycosphingolipid biosynthesis – lacto and neolacto series and glycosaminoglycan biosynthesis – keratan sulfate, glycoprotein metabolic process, and protein glycosylation. On the contrary, these metabolic features were unremarkable in ER-positive, PR-positive, LumA, and LumB patients. Among these pathways, glycosphingolipid biosynthesis – lacto and neolacto series and glycosaminoglycan synthesis were reported to be correlated with LumB and basal-like subtype [35]. We found that B3GNT5 and RAMP1 were potential key prognostic genes for the basal-like subtype. B3GNT5, highly expressed in the basal subtype, was associated with a poor prognosis whereas the lowly expressed gene RAMP1 was contrarily associated.

It was reported that basal-like subtype only contained a small number of hormone receptor-positive BRCA patients while LumA and LumB both accounted for a large number [33], which could explain why PC2 values were high in basal subtype but low in ER/PR, LumA, and LumB subtype in our result. Notably, endocrine resistance was enriched by the PC2 negative loading genes, such as ADCY1, ADCY5, and BCL2, which indicated that endocrine resistance occurred in ER- or PR-positive patients who had received endocrine therapy.

HER2-positive subtype represented approximately 15–20% in BRCA and was characterized by amplification or overexpression of ERBB2/neu [36]. With the advent of trastuzumab, the prognosis of HER2-positive BRCA patients had considerably improved [37]. Here, we found that HER2-positive patients had higher activity in lipid droplet pathway, PPAR signaling pathways, fatty acid metabolism, nucleoside bisphosphate biosynthetic process, gluconeogenesis pathway, and low activity of endocrine resistance, and other types of O-glycan biosynthesis. These unreported metabolic pathways might contribute to the further biological study of HER2-positive BRCA. Notably, PAPSS2 in nucleoside bisphosphate biosynthetic was a prognostic gene and hub metabolic gene exclusively expressed in the HER2-positive subtype. It was correlated with metastasis of BRCA by sulfation pathway [41].

Basal-like subtype was reported to be more susceptible to ferroptosis while HER2-positive and ER-positive BRCA were resistant to ferroptosis [38]. We found that ACSL1 and ALOX15 in the ferroptosis pathway were important metabolic characteristic genes in HER2-positive and infiltrating ductal subtypes, but correlated with poor survival of BRCA patients. ACSL1 was associated with tumor growth by regulation of GM-CSF [38].

Prognostic PCs and genes were sorted out to reveal novel biomarkers or therapeutic targets of BRCA. PC13 and PC29 were correlated with poor survival of BRCA patients and high Tr1 cell abundance. Purine metabolism was enriched in both PC13 and PC29. In this pathway, ADCY6 and PAPSS2 were indicators for a poor prognosis while ADCY7 and PAPSS2 were correlated with cell abundance of Tr1. This result...
was consistent with previous studies. Tr1 was correlated with the impaired immune response in hepatocellular carcinoma [39]. ADCY6 was associated with a worse prognosis involving the DNA methylation-regulated immune processes in luminal-like BRCA [39]. PAPSS2 was correlated with BRCA metastasis in the sulfation pathway [40] while ADCY7 expression was strongly associated with immune cell infiltration [41]. All these results indicated the relationship among purine metabolism pathway, Tr1 abundance, and prognosis of BRCA, which could provide a new therapy direction for BRCA.

Using the genes in the chemokine signaling pathway as biomarkers of BRCA cancer, other PC7's top-loading genes, such as APOBEC3D, APOBEC3H, FUT7, MUCL1, GPC6, and GALNT15, should be comprehensively considered. APOBEC3A was a major cytidine deaminase in breast cancer cells and was possibly a prominent contributor to the APOBEC mutation signature [42], which was promoted by the immune microenvironment [43]. APOBEC3G in the cytidine catabolic process was related to human CD4 + T cells [44]. GPC6 promoted invasive migration of BRCA through NFAT signaling [45]. FUT7 reinforced immunotherapy by reshaping glucose metabolism in triple-negative breast cancer (TNBC) [45]. MUCL1 could promote the proliferation of breast cancer [45]. GALNT15 as a biomarker in BRCA was newly identified by our research.

Recently, it was reported that ELOVL2 was a novel tumor suppressor attenuating tamoxifen resistance in BRCA. Down-regulation of ELOVL2 induced reprogramming of lipid metabolism and contributed to the malignant phenotypes [39]. Consistent with their results, we found that ELOVL2 was a negative top-loading gene in three prognostic PCs and that it was an indicator for good prognosis in early survival time but bad prognosis in long term. This indicated that ELOVL2 was a potentially sensitive indicator.

In summary, our research had identified metabolic PCs which were related to subtypes and prognosis of BRCA. By analyzing these PCs, we discovered metabolic pathways and hub genes that may provide new biomarkers for the guidance of BRCA treatment in the future. The mechanism that impacted on survival and expansion of BRCA should be further clarified. Our research had certain shortcomings and limitations. Firstly, external validation with other clinical datasets would be beneficial. Secondly, the biological significance of PCs had not been fully elucidated. Last but not least, more pathways should be included so that we could reveal the whole picture of the interaction between metabolic PCs and other PCs.

Conclusions

Based on 1039 TCGA samples, our study revealed several metabolic PCs which were related to subtypes and prognosis of BRCA patients. The key pathways and genes within these PCs were described in detail, containing potential biomarkers that were worth studying in the future. These results also provided a new direction for the treatment of different subtypes of BRCA.

Abbreviations
BRCA
Breast cancer
TCGA
The Cancer Genome Atlas
PC
principle component
PCA
principal component analysis
LASSO
least absolute shrinkage and selection operator
LumB
Luminal B
OS
overall survival
RFS
Relapse Free Survival
CPM
counts per million
ER
Estrogen receptor
PR
progesterone receptor
PAM50
prediction analysis of microarray 50
LumA
Luminal A
GO
Gene Ontology
Tex
exhausted T
Tcm
central memory T
Tgd
gamma delta T cells
iTreg
induced regulatory T cells
Th1
T helper 1
NK
natural killer
Declarations

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Availability of data and materials

All datasets used are publicly available and referenced in the methods section.

Competing interests

The authors declare that they have no competing interests.

Funding

High Level Hospital Construction Program of People's Hospital of Yangjiang.

Authors' contributions
Y.L. conducted datasets searches, model discovery and validation analyses. L.Y. and H.L. provided expertise on BRCA. Y.L., Z.C. and L.Y. wrote the paper. All authors read and approved the final manuscript.

Acknowledgements

Not applicable.

References

1. M. Akram, M. Iqbal, M. Daniyal, and A. U. Khan, “Awareness and current knowledge of breast cancer,” Biol. Res., pp. 1–23, 2017, doi: 10.1186/s40659-017-0140-9.

2. C. E. DeSantis et al., “Breast cancer statistics, 2019,” CA. Cancer J. Clin., vol. 69, no. 6, pp. 438–451, 2019, doi: 10.3322/caac.21583.

3. “Breast cancer.” https://www.who.int/news-room/fact-sheets/detail/breast-cancer (accessed Aug. 02, 2021).

4. Y.-S. Sun et al., “Risk Factors and Preventions of Breast Cancer,” Int. J. Biol. Sci., vol. 13, no. 11, p. 1387, 2017, doi: 10.7150/IJBS.21635.

5. D. I. Benjamin, B. F. Cravatt, and D. K. Nomura, “Global profiling strategies for mapping dysregulated metabolic pathways in cancer,” Cell Metab., vol. 16, no. 5, pp. 565–577, 2012, doi: 10.1016/j.cmet.2012.09.013.

6. D. Zhang, X. Xu, and Q. Ye, “Metabolism and immunity in breast cancer,” Front. Med., vol. 15, no. 2, pp. 178–207, 2021, doi: 10.1007/s11684-020-0793-6.

7. Natalya N. Pavlova and Craig B. Thompson, “Emerging metabolic hallmarks of cancer,” Physiol. Behav., vol. 176, no. 1, pp. 139–148, 2018, doi: 10.1016/j.cmet.2015.12.006.

8. C. ES, K. NH, Y. JS, C. SB, K. HS, and Y. Ji, “Breast Cancer Subtypes Underlying EMT-Mediated Catabolic Metabolism,” Cells, vol. 9, no. 9, Sep. 2020, doi: 10.3390/CELLS9092064.

9. S. Cassim and J. Pouyssegur, “Tumor microenvironment: A metabolic player that shapes the immune response,” Int. J. Mol. Sci., vol. 21, no. 1, 2020, doi: 10.3390/ijms21010157.

10. “UCSC Xena.” https://xenabrowser.net/ (accessed Aug. 03, 2021).

11. P. JS et al., “Supervised risk predictor of breast cancer based on intrinsic subtypes,” J. Clin. Oncol., vol. 27, no. 8, pp. 1160–1167, Mar. 2009, doi: 10.1200/JCO.2008.18.1370.

12. G. Sultan, S. Zubair, I. A. Tayubi, H.-U. Dahms, and I. H. Madar, “Towards the early detection of ductal carcinoma (a common type of breast cancer) using biomarkers linked to the PPAR(γ) signaling pathway,” Bioinformation, vol. 15, no. 11, p. 799, Nov. 2019, doi: 10.6026/97320630015799.

13. “R: The R Project for Statistical Computing.” https://www.r-project.org/ (accessed Nov. 26, 2021).

14. N. Simon, J. Friedman, T. Hastie, and R. Tibshirani, “Regularization paths for Cox’s proportional hazards model via coordinate descent,” J. Stat. Softw., vol. 39, no. 5, pp. 1–13, 2011, doi: 10.18637/jss.v039.i05.
15. T. T, “A Package for Survival Analysis in R.” 2020, [Online]. Available: https://cran.r-project.org/package=survival.

16. P. Biecek, Alboukadel Kassambara, Marcin Kosinski, “survminer: Drawing Survival Curves using ‘ggplot2.’” 2021, [Online]. Available: https://cran.r-project.org/package=survminer.

17. R. Li et al., “GDCRNATools: An R/Bioconductor package for integrative analysis of lncRNA, miRNA and mRNA data in GDC,” Bioinformatics, vol. 34, no. 14, pp. 2515–2517, Jul. 2018, doi: 10.1093/bioinformatics/bty124.

18. M. Kanehisa, Y. Sato, and M. Kawashima, “KEGG mapping tools for uncovering hidden features in biological data,” Protein Sci., 2021, doi: 10.1002/PRO.4172.

19. Y. R. Miao et al., “ImmuCellAI: A Unique Method for Comprehensive T-Cell Subsets Abundance Prediction and its Application in Cancer Immunotherapy,” Adv. Sci., 2020, doi: 10.1002/advs.201902880.

20. B. Rizeq and M. I. Malki, “The Role of CCL21/CCR7 Chemokine Axis in Breast Cancer Progression,” Cancers (Basel.), vol. 12, no. 4, Apr. 2020, doi: 10.3390/CANCERS12041036.

21. A. Salem, M. Alotaibi, R. Mroueh, H. A. Basheer, and K. Afarinkia, “CCR7 as a therapeutic target in Cancer,” Biochim. Biophys. acta. Rev. cancer, vol. 1875, no. 1, Jan. 2021, doi: 10.1016/J.BBCAN.2020.188499.

22. B. Rizeq and M. I. Malki, “The Role of CCL21/CCR7 Chemokine Axis in Breast Cancer Progression,” Cancers (Basel.), vol. 12, no. 4, Apr. 2020, doi: 10.3390/CANCERS12041036.

23. M. Al-Jokhadar, A. Al-Mandily, K. Zaid, and E. A. Maalouf, “CCR7 and CXCR4 Expression in Primary Head and Neck Squamous Cell Carcinomas and Nodal Metastases – a Clinical and Immunohistochemical Study,” Asian Pac. J. Cancer Prev., vol. 18, no. 4, pp. 1093–1104, Apr. 2017, doi: 10.22034/APJCP2017.18.4.1093.

24. Y. Ding et al., “Association of CC chemokine receptor 7 with lymph node metastasis of esophageal squamous cell carcinoma,” Clin. Cancer Res., vol. 9, no. 9, pp. 3406–3412, Sep. 2003, Accessed: Nov. 12, 2021. [Online]. Available: https://pubmed.ncbi.nlm.nih.gov/12960129/.

25. Z. J. Shang, K. Liu, and Z. Shao, “Expression of chemokine receptor CCR7 is associated with cervical lymph node metastasis of oral squamous cell carcinoma,” Oral Oncol., vol. 45, no. 6, pp. 480–485, Jun. 2009, doi: 10.1016/J.ORALONCOLOGY.2008.06.005.

26. G. RC and T. A, “Nonessential amino acid metabolism in breast cancer,” Adv. Biol. Regul., vol. 62, pp. 11–17, Sep. 2016, doi: 10.1016/J.JBIOR.2016.01.001.

27. Y.-P. Wang and Q.-Y. Lei, “Perspectives of Reprogramming Breast Cancer Metabolism,” Adv. Exp. Med. Biol., vol. 1026, pp. 217–232, 2017, doi: 10.1007/978-981-10-6020-5_10.

28. S. LEHRER, S. GREEN, F. R. DEMBITZER, P. H. RHEINSTEIN, and K. E. ROSENZWEIG, “Increased RNA Expression of von Willebrand Factor Gene Is Associated With Infiltrating Lobular Breast Cancer and Normal PAM50 Subtype,” Cancer Genomics Proteomics, vol. 16, no. 3, pp. 147–153, May 2019, doi: 10.21873/CGP.20120.
29. R. Barroso-Sousa and O. Metzger-Filho, “Differences between invasive lobular and invasive ductal carcinoma of the breast: Results and therapeutic implications,” Ther. Adv. Med. Oncol., vol. 8, no. 4, pp. 261–266, Jul. 2016, doi: 10.1177/1758834016644156.

30. P. M et al., “Research-based PAM50 signature and long-term breast cancer survival,” Breast Cancer Res. Treat., vol. 179, no. 1, pp. 197–206, Jan. 2020, doi: 10.1007/S10549-019-05446-Y.

31. Y. Y. Wang et al., “Mammary adipocytes stimulate breast cancer invasion through metabolic remodeling of tumor cells,” JCI insight, vol. 2, no. 4, Feb. 2017, doi: 10.1172/JCI.INSIGHT.87489.

32. P. I, M. E, G. I, M. S, K. A, and N. L, “PPARgamma expression in breast cancer: clinical value and correlation with ERbeta,” Histopathology, vol. 46, no. 1, pp. 37–42, Jan. 2005, doi: 10.1111/J.1365-2559.2005.02056.X.

33. A. Prat et al., “Clinical implications of the intrinsic molecular subtypes of breast cancer,” Breast, vol. 24, pp. S26–S35, 2015, doi: 10.1016/j.breast.2015.07.008.

34. et al Curtis, C., “Comprehensive molecular portraits of human breast tumors The Cancer Genome Atlas Network,” Nature, vol. 490, no. 7418, pp. 61–70, 2012, doi: 10.1038/nature11412.Comprehensive.

35. P. IO et al., “Glycan-related gene expression signatures in breast cancer subtypes; relation to survival,” Mol. Oncol., vol. 9, no. 4, pp. 861–876, Apr. 2015, doi: 10.1016/J.MOLONC.2014.12.013.

36. J. Kreutzfeldt, B. Rozeboom, N. Dey, and P. De, “The trastuzumab era: current and upcoming targeted HER2 + breast cancer therapies,” Am. J. Cancer Res., vol. 10, no. 4, p. 1045, 2020, Accessed: Oct. 25, 2021. [Online]. Available: /pmc/articles/PMC7191090/.

37. P. EA et al., “Trastuzumab plus adjuvant chemotherapy for human epidermal growth factor receptor 2-positive breast cancer: planned joint analysis of overall survival from NSABP B-31 and NCCTG N9831,” J. Clin. Oncol., vol. 32, no. 33, pp. 3744–3752, Nov. 2014, doi: 10.1200/JCO.2014.55.5730.

38. R. Ahmad, R. Thomas, F. Al-Rashed, N. Akhter, and F. Al-Mulla, “ACSL1 Regulates TNFα-Induced GM-CSF Production by Breast Cancer MDA-MB-231 Cells,” Biomolecules, vol. 9, no. 10, Oct. 2019, doi: 10.3390/BIOX9100555.

39. H. S et al., “Association of tamoxifen resistance and lipid reprogramming in breast cancer,” BMC Cancer, vol. 18, no. 1, Aug. 2018, doi: 10.1186/S12885-018-4757-Z.

40. Y. Zhang et al., “Enhanced PAPSS2/VCAN sulfation axis is essential for Snail-mediated breast cancer cell migration and metastasis,” Cell Death Differ., vol. 26, no. 3, pp. 565–579, Mar. 2019, doi: 10.1038/S41418-018-0147-Y.

41. Y. Zeng et al., “A Pan-Cancer Analysis of the Prognostic Value and Expression of Adenylate Cyclase 7 (ADCY7) in Human Tumors,” Int. J. Gen. Med., vol. 14, p. 5415, 2021, doi: 10.2147/IJGM.S330680.

42. L. M. Cortez et al., “APOBEC3A is a prominent cytidine deaminase in breast cancer,” PLoS Genet., vol. 15, no. 12, 2019, doi: 10.1371/JOURNAL.PGEN.1008545.

43. N. Roper et al., “APOBEC Mutagenesis and Copy Number Alterations are Drivers of Proteogenomic Tumor Evolution and Heterogeneity in Metastatic Thoracic Tumors,” SSRN Electron. J., Aug. 2018, doi: 10.2139/SSRN.3215351.
44. N. Sakaguchi and K. Maeda, “Germinal Center B-Cell-Associated Nuclear Protein (GANP) Involved in RNA Metabolism for B Cell Maturation,” Adv. Immunol., vol. 131, pp. 135–186, Jan. 2016, doi: 10.1016/BS.AI.2016.02.003.

45. G. K. Yiu, A. Kaunisto, Y. R. Chin, and A. Toker, “NFAT promotes carcinoma invasive migration through glypican-6,” Biochem. J., vol. 440, no. 1, pp. 157–166, Nov. 2011, doi: 10.1042/BJ20110530.

Figures

Figure 1

First three principle components to distinguish PAM50 subtypes.

(A) 3D scatter plot of first three principle components. Samples in different PAM50 subgroups were shown in different colors.

(B) Scatter plot of first three principle components. Samples in different PAM50 subgroups were shown in different colors.
Figure 2

PCs difference among clinical features

(A) Boxplots showed the Z-score difference of PC1 and PC12 among patients with different histological type in TCGA dataset. P value was calculated from ANOVA.

(B) Boxplots showed the Z-score difference of PC1, PC2 and PC4 among patients with different PAM50 subtype in TCGA dataset. P value was calculated from ANOVA.
(C) Boxplots showed the Z-score difference of PC1, PC2 and PC4 among patients with different PAM50 subtype in Joel dataset. P value was calculated from ANOVA.

(D) Boxplots showed the Z-score difference of PC2 among patients with different PR status in TCGA dataset. P value was calculated from ANOVA.

(E) Boxplots showed the Z-score difference of PC2 and P14 among patients with different ER status in TCGA dataset. P value was calculated from ANOVA.

(F) Boxplots showed the Z-score difference of PC4, PC6 and PC78 among patients with different HER2 status in TCGA dataset. P value was calculated from ANOVA.

Figure 3

The risk model for BRCA patients based on PCs

(A) The forest plot of multivariate Cox regression analysis showed basic clinical information, 13 PCs were selected by LASSO-penalized Cox analysis. For calculating Cox model, infiltrating ductal carcinoma, infiltrating lobular carcinoma, and mucinous carcinoma were regarded as 1, 2, and 3.

(B) Kaplan-Meier survival curves of the OS of patients in the high- and low-risk groups in training set.

(C) Kaplan-Meier survival curves of the OS of patients in the high- and low-risk groups in validation set.
**Figure 4**

**GO and KEGG enrichment of top loading genes in prognosis-related PCs**

(A) GO and KEGG enrichment of PC1 positively loading genes. Genes related to bad prognosis and good prognosis were marked with red color and blue color, respectively.
(B) GO and KEGG enrichment of PC1 negatively loading genes. Genes related to bad prognosis and good prognosis were marked with red color and blue color, respectively.

(C) GO and KEGG enrichment of PC2 positively loading genes. Genes related to bad prognosis and good prognosis were marked with red color and blue color, respectively.

(D) GO and KEGG enrichment of PC2 negatively loading genes. Genes related to bad prognosis and good prognosis were marked with red color and blue color, respectively.

(E) GO and KEGG enrichment of PC4 positively loading genes. Genes related to bad prognosis and good prognosis were marked with red color and blue color, respectively.

(F) GO and KEGG enrichment of PC4 negatively loading genes. Genes related to bad prognosis and good prognosis were marked with red color and blue color, respectively.
Figure 5

**GO and KEGG enrichment of top loading genes in prognosis-related PCs**

(A) GO and KEGG enrichment of PC6 positively loading genes. Genes related to bad prognosis and good prognosis were marked with red color and blue color, respectively.

(B) GO and KEGG enrichment of PC6 negatively loading genes. Genes related to bad prognosis and good prognosis were marked with red color and blue color, respectively.
(C) GO and KEGG enrichment of PC12 positively loading genes. Genes related to bad prognosis and good prognosis were marked with red color and blue color, respectively.

(D) GO and KEGG enrichment of PC12 negatively loading genes. Genes related to bad prognosis and good prognosis were marked with red color and blue color, respectively.

Figure 6

GO and KEGG enrichment of top loading genes in prognosis-related PCs
(A) GO and KEGG enrichment of PC14 positively loading genes. Genes related to bad prognosis and good prognosis were marked with red color and blue color, respectively.

(B) GO and KEGG enrichment of PC14 negatively loading genes. Genes related to bad prognosis and good prognosis were marked with red color and blue color, respectively.

(C) GO and KEGG enrichment of PC78 positively loading genes. Genes related to bad prognosis and good prognosis were marked with red color and blue color, respectively.

(D) GO and KEGG enrichment of PC78 negatively loading genes. Genes related to bad prognosis and good prognosis were marked with red color and blue color, respectively.
Figure 7

GO and KEGG enrichment of top loading genes in prognosis-related PCs

(A) GO and KEGG enrichment of PC7 positively loading genes. Genes related to bad prognosis and good prognosis were marked with red color and blue color, respectively.
(B) GO and KEGG enrichment of PC7 negatively loading genes. Genes related to bad prognosis and good prognosis were marked with red color and blue color, respectively.

(C) GO and KEGG enrichment of PC13 positively loading genes. Genes related to bad prognosis and good prognosis were marked with red color and blue color, respectively.

(D) GO and KEGG enrichment of PC13 negatively loading genes. Genes related to bad prognosis and good prognosis were marked with red color and blue color, respectively.

(E) GO and KEGG enrichment of PC29 positively loading genes. Genes related to bad prognosis and good prognosis were marked with red color and blue color, respectively.

(F) GO and KEGG enrichment of PC29 negatively loading genes. Genes related to bad prognosis and good prognosis were marked with red color and blue color, respectively.

Figure 8
Difference of cell abundances between high- and low-PC groups in prognosis-related PCs

(A) Boxplot showed the higher cell abundance in low-PC7 group. P value was calculated from Wilcoxon rank-sum test.

(B) Boxplot showed the higher cell abundance in high-PC7 group. P value was calculated from Wilcoxon rank-sum test.

(C) Boxplot showed the difference of cell abundances between high- and low-PC13 groups. P value was calculated from Wilcoxon rank-sum test.

(D) Boxplot showed the difference of cell abundances between high- and low-PC29 groups. P value was calculated from Wilcoxon rank-sum test.

Figure 9

Validation for different gene expression and prognostic model by using Joel dataset

(A) Boxplot showed the difference of B3GNT5 expression between PAM50 subtypes in Joel dataset. P value was calculated from ANOVA.

(B) Boxplot showed the difference of RAMP1 expression between PAM50 subtypes in Joel dataset. P value was calculated from ANOVA.

(C) Boxplot showed the difference of BCL2 expression between PAM50 subtypes in Joel dataset. P value was calculated from ANOVA.

(D) Boxplot showed the difference of PAPSS2 expression between PAM50 subtypes in Joel dataset. P value was calculated from ANOVA.

(E) Boxplot showed the difference of ALOX15 expression between PAM50 subtypes in Joel dataset. P value was calculated from ANOVA.

(F) Kaplan-Meier survival curves of the RFS of patients in the high- and low-risk score groups in Joel dataset.

(G) Kaplan-Meier survival curves of the RFS of patients in the high- and low-ADCY6 groups in Joel dataset.

(H) Kaplan-Meier survival curves of the RFS of patients in the high- and low-GPC6 groups in Joel dataset.
(I) Kaplan-Meier survival curves of the RFS of patients in the high- and low-ADCY7 groups in Joel dataset.

(J) Kaplan-Meier survival curves of the RFS of patients in the high- and low-PC7 groups in Joel dataset.

(K) Kaplan-Meier survival curves of the RFS of patients in the high- and low-PC13 groups in Joel dataset.

(L) Kaplan-Meier survival curves of the RFS of patients in the high- and low-PC29 groups in Joel dataset.

Supplementary Files

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

- TableS1.xlsx
- TableS2.xlsx
- TableS3.xlsx
- TableS4.xlsx
- TableS5.xlsx
- TableS6.xlsx
- FigureS1.pdf
- FigureS2.pdf
- FigureS3.pdf
- FigureS4.pdf
- FigureS5.pdf
- FigureS6.pdf
- FigureS7.pdf