Differential expression of cyclin dependent kinase 1 in cancers of the breast.

Shahan Mamoor, MS¹

¹shahanmamoor@gmail.com

East Islip, NY 11730

Breast cancer affects women at relatively high frequency¹. We mined published microarray datasets^{2,3} to determine in an unbiased fashion and at the systems level genes most differentially expressed in the primary tumors of patients with breast cancer. We report here significant differential expression of the gene encoding cyclin dependent kinase 1, CDK1, when comparing primary tumors of the breast to the tissue of origin, the normal breast. CDK1 was also differentially expressed in the tumor cells of patients with triple negative breast cancer. CDK1 mRNA was present at significantly higher quantities in tumors of the breast as compared to normal breast tissue. Relative to the two currently therapeutically targeted cyclin dependent kinases CDK4 and CDK6, CDK1 was more differentially expressed and present at higher quantity than CDK4 and CDK6, both in the tumors of patients with early-onset breast cancer and in the tumors of patients with triple negative breast cancer. Analysis of human survival data revealed that expression of CDK1 in primary tumors of the breast was correlated with overall survival in patients with stage III cancer, demonstrating an intricate relationship between correlation of primary tumor expression with overall survival based on stage of disease (progression). CDK1 may be of relevance to initiation, maintenance or progression of cancers of the female breast.

Keywords: breast cancer, CDK1, cyclin dependent kinase 1, systems biology of breast cancer, targeted therapeutics in breast cancer.

Invasive breast cancer is diagnosed in over a quarter of a million women in the United States each year¹ and in 2018, breast cancer was the leading cause of cancer death in women worldwide⁴. While patients with localized breast cancer are provided a 99% 5-year survival rate, patients with regional breast cancer, cancer that has spread to lymph nodes or nearby structures, are provided an 86% 5-year survival rate^{5,6}. Patients with metastasis to distant sites, like the brain, are provided a 27% 5-year survival rate^{5,6}. Understanding how primary tumors are most transcriptionally different from the tissue from which they originate, the breast, can facilitate development of novel diagnostic and therapeutics to promote early detection and enhanced treatment, and contribute to efforts to prevent progression to metastatic stages. We mined published microarray data^{2,3} to understand at the transcriptome level and in an unbiased fashion genes most differentially expressed in primary tumors of the breast as compared to normal breast tissue. cyclin dependent kinase 1 emerged as among the most differentially expressed genes in cancer of the female breast.

Methods

We utilized datasets GSE109169² and GSE38959³ for this differential gene expression analysis of female breast cancer. GSE109169 was generated using Affymetrix Human Exon 1.0 ST Array technology with $n=25$ normal breast tissue and $n=25$ tumors of the breast; analysis was performed using platform GPL5175. The tissues from this first dataset are paired tissues (25 tumors matching 25 breast tissues from 25 patients). GSE38959 was generated using Agilent-014850 Whole Human Genome Microarray 4x44K G4112F technology with $n=13$ samples of normal mammary gland ductal cells and $n=30$ samples of tumor cells from patients with triple negative breast cancer; analysis was performed using platform GPL4133. The Benjamini and Hochberg method of p -value adjustment was used for ranking of differential expression but raw p -values were used to assess statistical significance of global differential expression. Log-transformation of data was auto-detected, and the NCBI generated category of platform annotation was used. A statistical test was performed to evaluate whether CDK1 expression was significantly different between primary breast tumors and breast tissue using a two-tailed t-test. For Kaplan-Meier survival analysis, we used the Kaplan-Meier plotter online tool⁷ for correlation of CDK1 mRNA expression levels with overall survival (OS) in $n=404$ patients with basal subtype cancer, $n=794$ patients with luminal A subtype cancer, $n=515$ patients with luminal B subtype cancer, and $n=166$ patients with HER2+ cancer.

Results

We performed discovery of genes associated with breast cancer in females by mining two independently published microarray datasets^{2,3}.

CDK1 is differentially expressed in primary tumors of the breast.

Comparison of 25 normal breast tissues to 25 tumors of the breast² from patients with early-onset breast cancer revealed that the gene encoding cyclin dependent kinase 1, encoded by the CDK1 gene, was among the genes most differentially expressed in tumors of the breast in human breast cancer (Chart 1). When sorting each of the genes expressed in tumors of the breast based on significance of difference as compared to normal breast tissue, CDK1 ranked 212 out of 19076 total transcripts, equating to 98.9% differential expression (Chart 1). Differential expression of CDK1 in female breast cancer was statistically significant (Chart 1; $p=4.84E-13$).

Analysis of a second microarray dataset³, here in tumor cells of patients with triple negative breast cancer as compared to normal mammary gland ductal cells, again revealed significant differential expression of CDK1 in human breast cancer (Chart 2). When sorting each of the genes expressed in the

tumor cells of patients with triple negative breast cancer based on significance of difference as compared to normal mammary gland ductal cells, CDK1 ranked 94 out of 45015 total transcripts, equating to 99.8% differential expression (Chart 2). Differential expression of CDK1 in the tumor cells of patients with triple negative breast cancer was statistically significant (Chart 2; $p=8.12E-13$). These data suggested that differential expression of CDK1 was not an artifact of a single microarray dataset, nor was it associated with one type of breast cancer, rather a general feature of cancers of the breast.

CDK1 is expressed at significantly higher levels in breast tumors as compared to the breast.

We obtained exact mRNA expression levels for CDK1 from the breast and from breast tumors² to understand the magnitude and direction of CDK1 expression change. CDK1 was expressed at higher levels in tumors of the breast as compared to normal breast tissue (Figure 1). Increased expression of CDK1 in primary breast tumors was statistically significant (Figure 1: $p<0.0001$). CDK1 was expressed at 4.03 ± 0.48 arbitrary units (AU) in normal breast tissue, while it was expressed at 6.46 ± 1.21 AU in tumors of the breast. We calculated a mean fold change of 1.60 in CDK1 mRNA levels when comparing primary tumors of the breast to normal breast tissues.

We also obtained exact mRNA expression levels for CDK1 from normal mammary ductal epithelial cells and from the sorted tumor cells of patients with triple negative breast cancer breast tumors³ to understand the direction of CDK1 expression change in triple negative breast cancer. CDK1 was also expressed at significantly higher levels in triple negative breast tumor cells as compared to normal cells of the breast (Figure 2).

CDK1 expression is higher than both CDK4 and CDK6 expression in human breast cancer.

As treatments for metastatic human breast cancer now include inhibitors of the cyclin dependent kinases CDK4 and CDK6, we asked if CDK4 and CDK6 were also differentially expressed, both in early-onset breast cancer and in triple negative breast cancer, and if so, how this compared to relative to differential expression of CDK1 that we found here. In early onset breast cancer, while differential expression of CDK1 was ranked 212 out of 19076 transcripts or 98.9% differential expression, differential expression of CDK4 was ranked 1869 out of 19076 transcripts, or 90.2% differential expression, and differential expression of CDK6 was ranked 5203 out of 19076 transcripts, or 72.7% differential expression (Chart 1 and 1A). In triple negative breast cancer breast cancer, while differential expression of CDK1 was ranked 94 out of 45015 transcripts or 99.8% differential expression, differential expression of CDK4 was ranked 1668 out of 45015 transcripts, or 96.3% differential expression, and differential expression of CDK6 was ranked 6145 out of 45015 transcripts, or 86.3% differential expression (Chart 2 and 2A). Thus, CDK1 was more differentially expressed than CDK4 and CDK6 in each case, and accordingly, in the tumors of both patients with early-onset breast cancer and triple negative breast cancer, transcript for CDK1 was present at higher levels than was transcript for CDK4 and CDK6.

CDK1 expression correlates with survival outcomes in stage III human breast cancer.

We performed Kaplan-Meier survival analysis in an attempt to correlate CDK1 mRNA expression levels with survival outcomes in patients with breast cancer, evaluating interactions between overall survival, disease progression (cancer stage) and tumor expression of CDK1. We observed a statistically significant correlation between CDK1 expression and overall survival (OS) in patients with stage III breast cancer (Figure 3; log rank p -value: 0.038 for overall survival, hazard ratio: 1.85 (1.02-3.36) (Fig. 3)). CDK1 mRNA levels were a negative prognostic indicator in stage III breast cancer patients. Median OS was 73.57 months for stage III patients with low tumor expression of CDK1 as compared to stage III patients with high tumor expression of CDK1, who possessed median OS of 34.93 months (Chart 3).

CDK1 primary tumor expression was not correlated with overall survival in stage I (Figure 3; log rank p -value: 0.54 for OS, hazard ratio: 1.37 (0.5-3.78) (Fig. 3), stage II (Figure 2; log rank p -value: 0.55 for OS, hazard ratio: 1.17 (0.71-1.92) (Fig. 3) or stage IV breast cancer (Figure 2; log rank p -value: 0.91 for OS, hazard ratio: 0.94 (0.29-3.05) (Fig. 3)).

Thus, through comparative transcriptome analysis of primary tumors of the breast and normal breast tissue, we found that differential expression and up-regulation of CDK1 was among the most significant transcriptional features in primary tumors from patients with breast cancer. CDK1 expression in primary tumors of the breast was correlated with overall survival in patients with stage III breast cancers, with CDK1 primary tumor mRNA levels a negative prognostic indicator for stage III patients.

Discussion

Invasive breast cancer is a medical problem with a 27% 5-year survival rate for women whose disease has spread to distant sites^{5,6}. To facilitate understanding of the basic transcriptional differences between primary tumors of the breast and the tissues from which these tumors originate, normal breast tissues, we performed comparative transcriptome analysis using two independently published microarray datasets^{2,3}, providing evidence here that differential expression of cyclin dependent kinase 1, encoded by

CDK1, is a defining transcriptional feature of human breast cancer: in early onset-breast cancer, and in triple negative breast cancer. CDK1 was expressed at significantly higher levels in primary tumors from patients with breast cancer, as compared to normal breast. In patients with stage III breast cancer, primary tumor expression of CDK1 was correlated with overall survival. Cyclin dependent kinase 1, and the molecular processes to which it pertains to, may be relevant to the initiation or progression of human breast cancer.

References

1. DeSantis, C.E., Ma, J., Goding Sauer, A., Newman, L.A. and Jemal, A., 2017. Breast cancer statistics, 2017, racial disparity in mortality by state. *CA: a cancer journal for clinicians*, 67(6), pp.439-448.
2. Chang, J.W., Kuo, W.H., Lin, C.M., Chen, W.L., Chan, S.H., Chiu, M.F., Chang, I.S., Jiang, S.S., Tsai, F.Y., Chen, C.H. and Huang, P.H., 2018. Wild-type p53 upregulates an early onset breast cancer-associated gene GAS7 to suppress metastasis via GAS7–CYFIP1- mediated signaling pathway. *Oncogene*, 37(30), pp.4137-4150.
3. Komatsu, M., Yoshimaru, T., Matsuo, T., Kiyotani, K., Miyoshi, Y., Tanahashi, T., Rokutan, K., Yamaguchi, R., Saito, A., Imoto, S. and Miyano, S., 2013. Molecular features of triple negative breast cancer cells by genome-wide gene expression profiling analysis. *International journal of oncology*, 42(2), pp.478-506.
4. Bray, F., Ferlay, J., Soerjomataram, I., Siegel, R.L., Torre, L.A. and Jemal, A., 2018. Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA: a cancer journal for clinicians*, 68(6), pp.394-424.
5. ACS Cancer Facts & Figures 2019.
<https://www.cancer.net/cancer-types/breast-cancer/metastatic/statistics>.
6. Survival Rates for Breast Cancer. <https://www.cancer.org/cancer/breast-cancer/understanding-a-breast-cancer-diagnosis/breast-cancer-survival-rates.html>
7. Györfy, B., Lanczky, A., Eklund, A.C., Denkert, C., Budczies, J., Li, Q. and Szallasi, Z., 2010. An online survival analysis tool to rapidly assess the effect of 22,277 genes on breast cancer prognosis using microarray data of 1,809 patients. *Breast cancer research and treatment*, 123(3), pp.725-731.

Rank: 212
Probe ID: 3248289
p-value: 4.84E-13
t: -9.61
B: 19.5087
Gene: CDK1
Gene name: cyclin dependent kinase 1

Chart 1A: CDK1 is differentially expressed in the primary tumors of patients with early-onset human breast cancer.

Rank of differential expression, probe ID, p-value of global differential expression, t, a moderated t-statistic, B, the log-odds of differential expression between the groups compared, gene and gene name are listed in this chart.

Rank: 1869
Probe ID: 3458783
p-value: 4.64E-07
t: -5.77
B: 5.945375
Gene: CDK4
Gene name: cyclin dependent kinase 4

Rank: 5203
Probe ID: 3061319
p-value: 7.33E-04
t: 3.59
B: -1.173433
Gene: CDK6
Gene name: cyclin dependent kinase 6

Chart 1B: CDK4 and CDK6 are also differentially expressed in the primary tumors of patients with early-onset human breast cancer, but to a lesser extent than CDK1.

Rank of differential expression, probe ID, p-value of global differential expression, t, a moderated t-statistic, B, the log-odds of differential expression between the groups compared, gene and gene name are listed in this chart.

Rank: 94
Probe ID: 43582
p-value: 8.12E-13
t: -9.8461714
B: 19.054753
Gene: CDK1
Gene name: cyclin dependent kinase 1

Chart 2: CDK1 is differentially expressed in the tumor cells of patients with triple-negative breast cancer.

Rank of differential expression, probe ID, p-value of global differential expression, t, a moderated t-statistic, B, the log-odds of differential expression between the groups compared, gene and gene name are listed in this chart.

Rank: 1668
Probe ID: 28096
p-value: 1.89E-06
t: -5.4680844
B: 4.765718
Gene: CDK4
Gene name: cyclin dependent kinase 4

Rank: 6145
Probe ID: 28438
p-value: 1.13E-03
t: -3.4775168
B: -1.365421
Gene: CDK6
Gene name: cyclin dependent kinase 6

Chart 2B: CDK4 and CDK6 are also differentially expressed in the primary tumor cells of patients with triple negative breast cancer, but to a lesser extent than CDK1.

Rank of differential expression, probe ID, p-value of global differential expression, t, a moderated t-statistic, B, the log-odds of differential expression between the groups compared, gene and gene name are listed in this chart.

Figure 1: CDK1 is expressed at significantly higher levels in primary breast tumors as compared to normal breast tissue.

The mRNA expression level of CDK1 in normal breast tissue (left) and in primary tumors of the breast (right) is graphically depicted with the result of a statistical test evaluating significance of difference in CDK1 expression between normal breast tissue and primary tumors of the breast, a p -value, is $p<0.00001$.

Figure 2: CDK1 is expressed at significantly higher levels in the tumor cells of patients with triple negative breast cancer as compared to normal mammary ductal (breast) epithelial cells.

The mRNA expression level of CDK1 in normal cells of the breast (left) and in isolated tumor cells of patients with triple negative breast cancer (right) is graphically depicted with the result of a statistical test evaluating significance of difference in CDK1 expression between normal cells of the breast and primary triple negative breast tumor cells of the breast, a *p*-value, listed above.

Figure 3: CDK1 expression correlates with overall survival in patients with stage III breast cancer, but not in patients with stage I, II, or IV cancer.

Depicted in this Kaplan-Meier plot is the probability of overall survival (OS) for patients with stage I breast cancer (upper left), patients with stage II breast cancer (upper right), patients with stage III breast cancer (lower left), and patients with stage IV breast cancer (lower right), stratified into two groups, based on low or high expression of CDK1 in patient primary tumors, in the upper survival tertile. The log rank p-value denoting statistical significance of difference in overall survival when comparing the two groups, as well as hazard ratio for this comparison is listed above. Listed below is the number of patients at risk (number of patients alive) per interval, after stratification based on CDK1 expression; in the first interval, number at risk is number of patients alive; in each subsequent interval, number at risk is the number at risk less those who have expired or are censored.

1 Low CDK1 expression: 73.57 months

2 High CDK1 expression: 34.93 months

3 **Chart 3: In stage III breast cancer, median overall survival is superior in patients with low primary**
4 **tumor expression of CDK1.**

5 The overall survival of stage III breast cancer patients based on stratification into low or high tumor
6 expression of CDK1, in the upper survival tertile, is listed in this chart.
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28