Differential expression of coiled-coil domain containing 178 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 coiled-coil domain containing 178, CCDC178, when comparing primary tumors of the breast to the tissue of origin, the normal breast. CCDC178 was also differentially expressed in the brain metastases of patients with metastatic breast cancer. CCDC178 mRNA was present at significantly lower quantities in tumors of the breast as compared to normal breast tissue. Analysis of human survival data revealed that expression of CCDC178 in primary tumors of the breast was correlated with recurrence-free survival in patients with basal-like subtype cancer, demonstrating a relationship between primary tumor expression of a differentially expressed gene and patient survival outcomes influenced by PAM50 molecular subtype. CCDC178 may be of relevance to initiation, maintenance or progression of cancers of the female breast.

Keywords: breast cancer, CCDC178, coiled-coil domain containing 178, 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. Coiled-coil domain containing 178 emerged as among the most differentially expressed genes in cancer of the female breast.

Methods

We utilized datasets GSE42568² and GSE52604³ for this global differential gene expression analysis of female breast cancer. GSE42568 was generated using Affymetrix Human Genome U133 Plus 2.0 array technology with $n=17$ normal breast tissue biopsies and $n=104$ primary breast tumor biopsies from patients with breast cancer; analysis was performed using platform GPL570; the majority of patients whose tumors were analyzed were age 50 or older. GSE52604 was generated using Agilent-014850 Whole Human Genome Microarray 4x44K G4112F with $n=10$ normal breast tissues and $n=35$ brain metastases from patients with breast cancer; analysis was performed using platform GPL6480. 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 CCDC178 expression was significantly different between primary breast tumors and normal breast tissue using a two-tailed t-test. For Kaplan-Meier survival analysis, we used the Kaplan-Meier plotter online tool⁷ for correlation of CCDC178 mRNA expression levels with recurrence-free survival (RFS) in $n=296$ patients with basal-like subtype cancer, $n=222$ patients with luminal A subtype cancer, $n=200$ patients with luminal B subtype cancer, $n=198$ patients with HER2+ cancer, and $n=27$ patients with normal-like subtype cancer.

Results

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

CCDC178 is differentially expressed in primary tumors of the breast.

Studying the global gene expression profiles of 104 breast cancers from patients aged 31 to 89 revealed that the gene encoding coiled-coil domain containing 178, CCDC178, 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, CCDC178 ranked 38 out of 54675 total transcripts, equating to 99.9% differential expression (Chart 1). Differential expression of CCDC178 in female breast cancer was statistically significant (Chart 1; $p=2.87\text{E-}35$).

Analysis of a second microarray dataset³, here studying global gene expression patterns in the brain metastases of patients with metastatic breast cancer, again revealed significant differential expression of CCDC178 in tumor tissue of patients with human breast cancer (Chart 2). When sorting each of the genes

expressed in the brain metastases of patients with breast cancer based on significance of difference as compared to normal breast tissue, CCDC178 ranked 45 out of 41093 total transcripts, equating to 99.9% differential expression (Chart 2). Differential expression of CCDC178 in the brain metastases of patients with breast cancer was statistically significant (Chart 2; $p=2.41E-14$). These data suggested that differential expression of CCDC178 was not an artifact of a single microarray dataset, nor was it strictly associated with primary tumors of the breast, rather a reflection of the spectrum of transformation in human cancer.

CCDC178 is expressed at significantly lower levels in breast tumors as compared to the breast.

We obtained exact mRNA expression levels for CCDC178 from the breast and from breast tumors to understand the magnitude and direction of CCDC178 expression change. CCDC178 was expressed at lower levels in tumors of the breast as compared to normal breast tissue (Figure 1). Decreased expression of CCDC178 in primary breast tumors was statistically significant (Figure 1: $p<0.0001$). CCDC178 was expressed at 4.95 ± 0.87 arbitrary units (AU) in normal breast tissue, while it was expressed at 2.77 ± 0.38 AU in tumors of the breast. We calculated a mean fold change of 0.56 in CCDC178 mRNA levels when comparing primary tumors of the breast to normal breast tissues.

CCDC178 expression correlates with survival outcomes in basal-like subtype human breast cancer.

We performed Kaplan-Meier survival analysis to study relationships between tumor CCDC178 mRNA expression levels and survival outcomes in patients with breast cancer. We observed a correlation between CCDC178 expression and recurrence-free survival (RFS) in patients with basal-like subtype breast cancer which was statistically significant (Figure 2; log rank p -value: 0.044 for recurrence-free survival, hazard ratio: 0.73 (0.54-0.99) (Fig. 2)). CCDC178 mRNA levels were a positive prognostic indicator in basal-like subtype breast cancer patients. Median RFS was 18 months for basal-like patients with low tumor expression of CCDC178 while median RFS was 31.56 months for basal-like patients with high tumor expression of CCDC178 (Chart 3). CCDC178 primary tumor expression was not correlated with recurrence-free survival in luminal A breast cancer (Figure 2; log rank p -value: 0.5 for RFS, hazard ratio: 0.89 (0.63-1.26) (Fig. 2)), luminal B subtype breast cancer (Figure 2; log rank p -value: 0.82 for RFS, hazard ratio: 0.97 (0.75-1.26) (Fig. 2)), in HER2+ breast cancer (Figure 2; log rank p -value: 0.12 for RFS, hazard ratio: 0.77 (0.56-1.07) (Fig. 2)) or in patients with normal-like breast cancer (Figure 2; log rank p -value: 0.55 for RFS, hazard ratio: 0.74 (0.28-2) (Fig. 2)).

Thus, through comparative transcriptome analysis of primary tumors of the breast and normal breast tissue, we found that differential expression and down-regulation of CCDC178 was among the most significant transcriptional features in primary tumors from patients with breast cancer. CCDC178 expression in primary tumors of the breast was correlated with recurrence-free survival in patients with basal-like subtype disease, with mRNA levels of CCDC178 a positive prognostic indicator for basal-like breast cancer 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 coiled-coil domain containing 178, encoded by CCDC178, is a defining transcriptional feature of human breast cancer: in patients diagnosed after age 50, and in brain metastatic tissues from patients with human breast cancer. CCDC178 was

expressed at significantly lower levels in primary tumors from patients with breast cancer as compared to normal breast tissue. Importantly, in patients with basal-like subtype breast cancers, expression of CCDC178 was correlated with recurrence-free survival. Coiled-coil domain containing 178 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. Clarke, C., Madden, S.F., Doolan, P., Aherne, S.T., Joyce, H., O'driscoll, L., Gallagher, W.M., Hennessy, B.T., Moriarty, M., Crown, J. and Kennedy, S., 2013. Correlating transcriptional networks to breast cancer survival: a large-scale coexpression analysis. *Carcinogenesis*, 34(10), pp.2300-2308.
3. Salhia, B., Kiefer, J., Ross, J.T., Metapally, R., Martinez, R.A., Johnson, K.N., DiPerna, D.M., Paquette, K.M., Jung, S., Nasser, S. and Wallstrom, G., 2014. Integrated genomic and epigenomic analysis of breast cancer brain metastasis. *Plos one*, 9(1), p.e85448.
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: 38
Probe ID: 1558820_a_at
p-value: 2.87E-35
t: 1.77E+01
B: 69.657221
Gene: CCDC178
Gene name: coiled-coil domain containing 178

Chart 1: CCDC178 is differentially expressed in the primary tumors of patients with 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: 45
Probe ID: A_32_P184448
p-value: 2.41E-14
t: 10.8196975
B: 22.432101
Gene: CCDC178
Gene name: coiled-coil domain containing 178

Chart 2: CCDC178 is differentially expressed in the brain metastases of patients with metastatic 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.

Figure 1: CCDC178 is expressed at significantly lower levels in primary breast tumors as compared to normal breast tissue.

The mRNA expression level of CCDC178 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 CCDC178 expression between normal breast tissue and primary tumors of the breast, a *p*-value, listed above.

Figure 2: CCDC178 expression correlates with recurrence-free survival in patients with basal-like subtype breast cancer, but not in patients with luminal A, luminal B, HER2+, or normal-like subtype cancer.

Depicted in this Kaplan-Meier plot is the probability of recurrence-free survival (RFS) for $n=296$ patients with basal-like breast cancer (upper left), $n=198$ patients with HER2+ breast cancer (upper right), $n=222$ patients with luminal A breast cancer (middle left), $n=200$ patients with luminal B breast cancer (middle right), and $n=27$ patients with normal-like breast cancer (below) stratified into two

groups, based on low or high expression of CCDC178 in patient primary tumors. The log rank p-value denoting statistical significance of difference in recurrence-free 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 CCDC178 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.

Basal-like patients

Low CCDC178 expression: 18 months

High CCDC178 expression: 31.56 months

Chart 3: In basal-like subtype breast cancer, median recurrence-free survival is superior in patients with high primary tumor expression of CCDC178.

The recurrence-free survival of $n=296$ basal-like subtype breast cancer patients based on stratification into low or high tumor expression of CCDC178 is listed in this chart.