Differential expression of a pseudogene of TPTE2 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 transmembrane phosphoinositide 3-phosphatase and tensin homolog 2 pseudogene 2, TPTE2P2, when comparing primary tumors of the breast to the tissue of origin, the normal breast. TPTE2P2 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 TPTE2P2 in primary tumors of the breast was correlated with overall survival in patients with stage III breast cancer, demonstrating a relationship between primary tumor expression of a differentially expressed gene and patient survival outcomes influenced by cancer progression. TPTE2P2 may be of relevance to initiation, maintenance or progression of cancers of the female breast.

Keywords: breast cancer, TPTE2P2, transmembrane phosphoinositide 3-phosphatase and tensin homolog 2 pseudogene 2, 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. Transmembrane phosphoinositide 3-phosphatase and tensin homolog 2 pseudogene 2 emerged as among the most differentially expressed genes in cancer of the female breast.

Methods

We utilized datasets GSE109169² and GSE72653³ for this differential gene expression analysis of female breast cancer in conjunction with GEO2R. 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). GSE72653 was generated using Agilent-014850 Whole Human Genome Microarray 4x44K G4112F with $n=8$ normal epithelium and $n=8$ tumors of the breast from patients with breast cancer; the tissues from this second dataset are paired tissues (8 tumors matching 8 normal breast epithelial tissues from 8 patients); 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 TPTE2P2 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 TPTE2P2 mRNA expression levels with overall survival (OS) based on stage of disease in $n=1090$ breast cancer patients.

Results

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

TPTE2P2 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 transmembrane phosphoinositide 3-phosphatase and tensin homolog 2 pseudogene 2, encoded by the TPTE2P2 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, TPTE2P2 ranked 118 out of 19076 total transcripts, equating to 99.4% differential expression (Chart 1). Differential expression of TPTE2P2 in female breast cancer was statistically significant (Chart 1; $p=2.54E-14$).

To attempt to validate differential transcriptome-wide differential expression of TPTE2P2 in human breast cancer, we queried a second microarray dataset³, here comparing normal breast tissues and primary tumors of the breast. Again, we identified TPTE2P2 as a differentially expressed gene in the tumor tissues

of patients with breast cancer (Chart 2). When sorting each of the genes expressed in primary tumors of the breast based on significance of difference as compared to normal breast tissues, TPTE2P2 ranked 11593 out of 41093 total transcripts (Chart 2), equating to 71.8% differential expression. Differential expression of TPTE2P2 in the primary tumors of patients with breast cancer trended towards statistical significance (Chart 2; $p=7.34E-02$). These data indicated that differential expression of TPTE2P2 was not an artifact of a single microarray dataset, and perhaps a general feature of cancers of the breast.

TPTE2P2 is expressed at significantly lower levels in breast tumors as compared to the breast.

We obtained exact mRNA expression levels for TPTE2P2 from the breast and from breast tumors² to understand the magnitude and direction of TPTE2P2 expression change. TPTE2P2 was expressed at lower levels in tumors of the breast as compared to normal breast tissue (Figure 1). Decreased expression of TPTE2P2 in primary breast tumors was statistically significant (Figure 1; $p<0.0001$). TPTE2P2 was expressed at 7.15 ± 0.22 arbitrary units (AU) in normal breast tissue, while it was expressed at 6.52 ± 0.20 AU in tumors of the breast. We calculated a mean fold change of 0.91 in TPTE2P2 mRNA levels when comparing primary tumors of the breast to normal breast tissues.

TPTE2P2 expression correlates with survival outcomes in stage III human breast cancer.

We performed Kaplan-Meier survival analysis in an attempt to correlate TPTE2P2 mRNA expression levels with survival outcomes in patients with breast cancer, evaluating correlations between stage of disease and overall survival. We observed a correlation between TPTE2P2 expression and overall survival (OS) in patients with stage III breast cancer which trended towards statistical significance, in the lower survival quartile (Figure 2; log rank p -value: 0.081 for overall survival, hazard ratio: 0.58 (0.31-1.08) (Fig. 2)). TPTE2P2 mRNA levels were a positive prognostic indicator in stage III breast cancer. Median OS was 29.3 months for stage III patients with low tumor expression of TPTE2P2 while median OS was 62.8 months for stage III patients with high tumor expression of TPTE2P2 (Chart 3). TPTE2P2 primary tumor expression was not significantly correlated with overall survival in stage I breast cancer (Figure 2; log rank p -value: 0.99 for OS, hazard ratio: 0.99 (0.31-3.13) (Fig. 2)), stage II breast cancer (Figure 2; log rank p -value: 0.54 for OS, hazard ratio: 0.85 (0.51-1.43) (Fig. 2)), stage IV breast (Figure 2; log rank p -value: 0.28 for OS, hazard ratio: 2.26 (0.49-10.35) (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 TPTE2P2 was among the most significant transcriptional features in primary tumors from patients with breast cancer. TPTE2P2 expression in primary tumors of the breast was correlated with overall survival in patients with stage III breast cancer, with TPTE2P2 primary tumor mRNA levels a positive 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 transmembrane phosphoinositide 3-phosphatase and tensin homolog 2 pseudogene 2, encoded by TPTE2P2, is a defining transcriptional feature of human breast cancer. TPTE2P2 was expressed at significantly lower 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 TPTE2P2 was correlated with overall survival. Transmembrane

phosphoinositide 3-phosphatase and tensin homolog 2 pseudogene 2 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. Abdalla, M., Tran-Thanh, D., Moreno, J., Iakovlev, V., Nair, R., Kanwar, N., Abdalla, M., Lee, J.P., Kwan, J.Y.Y., Cawthorn, T.R. and Warren, K., 2017. Mapping genomic and transcriptomic alterations spatially in epithelial cells adjacent to human breast carcinoma. *Nature communications*, 8(1), pp.1-11.
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. Nagy, Á., Munkácsy, G. and Györffy, B., 2021. Pancancer survival analysis of cancer hallmark genes. *Scientific reports*, 11(1), pp.1-10.

Rank: 118
Probe ID: 3514879
p-value: 2.54E-14
t: 1.05E+01
B: 22.414128
Gene: TPTE2P2
Gene name: transmembrane phosphoinositide 3-phosphatase and tensin homolog 2 pseudogene 2

Chart 1: TPTE2P2 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: 11593
Probe ID: A_24_P145268
p-value: 7.34E-02
t: 1.9089038
B: -4.363239
Gene: TPTE2P2
Gene name: transmembrane phosphoinositide 3-phosphatase and tensin homolog 2 pseudogene 2

Chart 2: TPTE2P2 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.

Figure 1: TPTE2P2 is expressed at significantly lower levels in primary breast tumors as compared to normal breast tissue.

The mRNA expression level of TPTE2P2 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 TPTE2P2 expression between normal breast tissue and primary tumors of the breast, a p -value, is $p < 0.0001$.

Figure 2: TPTE2P2 expression correlates with overall survival in patients with stage III breast cancer, but not in patients with stage I, stage II, or stage IV breast 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 TPTE2P2 in patient primary tumors, in the lower survival quartile. 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 TPTE2P2 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.

Stage III

Low TPTE2P2 expression: 29.3 months

High TPTE2P2 expression: 62.8 months

Chart 3: In stage III breast cancer, median overall survival is superior in patients with high primary tumor expression of TPTE2P2.

The overall survival of stage III breast cancer patients based on stratification into low or high tumor expression of TPTE2P2 is listed in this chart, in the lower survival quartile.