Differential expression of KIAA0101 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 KIAA0101 when comparing primary tumors of the breast to the tissue of origin, the normal breast. KIAA0101 was also differentially expressed in the tumor cells of patients with triple negative breast cancer. KIAA0101 mRNA was present at significantly higher quantities in tumors of the breast as compared to normal breast tissue. Analysis of human survival data revealed that expression of KIAA0101 in primary tumors of the breast was correlated with distant metastasis-free survival in patients with HER2+ and luminal A subtype cancer, but in a contrary manner, demonstrating a complex relationship between correlation of primary tumor expression with overall survival based on molecular subtype. KIAA0101 may be of relevance to initiation, maintenance or progression of cancers of the female breast.

Keywords: breast cancer, KIAA0101, 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. KIAA0101 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 KIAA0101 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 KIAA0101 mRNA expression levels with distant metastasis-free 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}.

KIAA0101 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 KIAA0101 was among those 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, KIAA0101 ranked 229 out of 19076 total transcripts, equating to 98.8% differential expression (Chart 1). Differential expression of KIAA0101 in female breast cancer was statistically significant (Chart 1; $p=8.28E-13$).

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 KIAA0101 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, KIAA0101 ranked 1338 out of 45015 total transcripts,

equating to 97.0% differential expression (Chart 2). Differential expression of KIAA0101 in the tumor cells of patients with triple negative breast cancer was statistically significant (Chart 2; $p=6.38E-07$). These data suggested that differential expression of KIAA0101 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.

KIAA0101 is expressed at significantly higher levels in breast tumors as compared to the breast.

We obtained exact mRNA expression levels for KIAA0101 from the breast and from breast tumors² to understand the magnitude and direction of KIAA0101 expression change. KIAA0101 was expressed at higher levels in tumors of the breast as compared to normal breast tissue (Figure 1). Increased expression of KIAA0101 in primary breast tumors was statistically significant (Figure 1: $p<0.0001$). KIAA0101 was expressed at 4.52 ± 0.81 arbitrary units (AU) in normal breast tissue, while it was expressed at 7.05 ± 1.11 AU in tumors of the breast. We calculated a mean fold change of 1.56 in KIAA0101 mRNA levels when comparing primary tumors of the breast to normal breast tissues.

KIAA0101 expression correlates with survival outcomes in HER2+ and luminal A human breast cancer.

We performed Kaplan-Meier survival analysis in an attempt to correlate KIAA0101 mRNA expression levels with survival outcomes in patients with breast cancer, evaluating interactions between molecular subtype and distant metastasis-free survival. We observed a statistically significant correlation between KIAA0101 expression and distant metastasis-free survival (DMFS) in patients with luminal A breast cancer (Figure 2; log rank p -value: $8.3e-05$ for distant metastasis-free survival, hazard ratio: 1.7 (1.3-2.23) (Fig. 2)). KIAA0101 mRNA levels were a negative prognostic indicator in luminal A subtype breast cancer patients. Median DMFS was 149 months for luminal A subtype patients with low tumor expression of KIAA0101 while median DMFS was 88.64 months for luminal A subtype patients with high tumor expression of KIAA0101 (Chart 3). We also observed a correlation between KIAA0101 expression and distant metastasis-free survival (DMFS) in patients with HER2+ breast cancer that trended towards statistical significance (Figure 2; log rank p -value: 0.085 for distant metastasis-free survival, hazard ratio: 0.65 (0.39-1.07) (Fig. 2)). KIAA0101 mRNA levels were a positive prognostic indicator in HER2+ breast cancer patients. Median DMFS was 24 months for HER2+ patients with low tumor expression of KIAA0101 while median DMFS was 38.31 months for HER2+ patients with high tumor expression of KIAA0101 (Chart 3). KIAA0101 primary tumor expression was not correlated with distant metastasis-free survival in basal subtype (Figure 2; log rank p -value: 0.52 for DMFS, hazard ratio: 0.9 (0.66-1.23) (Fig. 2)), or in luminal B subtype breast cancer (Figure 2; log rank p -value: 0.34 for DMFS, hazard ratio: 1.15 (0.87-1.52) (Fig. 2)).

Thus, through comparative transcriptome analysis of primary tumors of the breast and normal breast tissue, we found that differential expression and up-regulation of KIAA0101 was among the most significant transcriptional features in primary tumors from patients with breast cancer. KIAA0101 expression in primary tumors of the breast was correlated with distant metastasis-free survival in patients with luminal A and HER2+ cancers, with low expression of KIAA0101 associating with superior survival in luminal A breast cancer but with high expression of KIAA0101 associating with superior survival in HER2+ breast cancer.

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

1 between primary tumors of the breast and the tissues from which these tumors originate, normal breast
2 tissues, we performed comparative transcriptome analysis using two independently published microarray
3 datasets^{2,3}, providing evidence here that differential expression of the KIAA0101 gene is a defining
4 transcriptional feature of human breast cancer: in early onset-breast cancer, and in triple negative breast
5 cancer. KIAA0101 was expressed at significantly higher levels in primary tumors from patients with
6 breast cancer as compared to normal tissues of the breast. In patients with luminal A or HER2_subtype
7 breast cancers, expression of KIAA0101 was correlated with distant metastasis-free survival. KIAA0101
8 and the molecular processes to which it pertains to may be relevant to the initiation or progression of
9 human breast cancer.
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28

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: 229
Probe ID: 3629103
p-value: 8.28E-13
t: -9.46
B: 18.978864
Gene: KIAA0101
Gene name: KIAA0101

Chart 1: KIAA0101 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: 1338
Probe ID: 14018
p-value: 6.38E-07
t: -5.7873333
B: 5.821028
Gene: KIAA0101
Gene name: KIAA0101

Chart 2: KIAA0101 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.

Figure 1: KIAA0101 is expressed at significantly higher levels in primary breast tumors as compared to normal breast tissue.

The mRNA expression level of KIAA0101 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 KIAA0101 expression between normal breast tissue and primary tumors of the breast, a p -value, is $p<0.0001$.

Figure 2: KIAA0101 expression correlates with distant metastasis-free survival in patients with HER2+ or luminal A subtype breast cancer, but not in patients with basal or luminal B subtype cancer.

Depicted in this Kaplan-Meier plot is the probability of distant metastasis-free survival (DMFS) for $n=571$ patients with basal subtype breast cancer (upper left), $n=178$ patients with HER2+ breast cancer (upper right), $n=1260$ patients with luminal A breast cancer (lower left), and $n=756$ patients with luminal B breast cancer (lower right), stratified into two groups, based on low or high expression of KIAA0101 in patient primary tumors. The log rank p-value denoting statistical significance of difference in distant metastasis-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 KIAA0101 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.

Luminal A

Low KIAA0101 expression: 149 months

High KIAA0101 expression: 88.64 months

HER2+

Low KIAA0101 expression: 24 months

High KIAA0101 expression: 38.31 months

Chart 3: In luminal A subtype breast cancer, median distant metastasis-free survival is superior in patients with low primary tumor expression of KIAA0101, but in HER2+ patients, median distant-metastasis free survival is superior in patients with high primary tumor expression of KIAA0101.

The distant metastasis-free survival of $n=1260$ luminal A subtype breast cancer patients and $n=178$ HER2+ breast cancer patients based on stratification into low or high tumor expression of KIAA0101 is listed in this chart.