Expression of LINC00968 correlates with overall survival in normal-like breast cancer.

Shahan Mamoor, MS¹

¹shahanmamoor@gmail.com

East Islip, NY 11730

Breast cancer affects women at relatively high frequency¹. By mining published microarray data^{2,3}, we have previously found significant differential expression of the gene encoding long intergenic non-protein coding RNA 968, LINC00968, when comparing primary tumors of the breast to the tissue of origin, the normal breast and in the brain metastases of patients with metastatic breast cancer⁴. Here we demonstrate through analysis of human survival data that expression of LINC00968 in primary tumors of the breast is correlated with overall survival in patients with normal-like subtype cancer, demonstrating a relationship between primary tumor expression of a differentially expressed non-coding RNA and patient survival outcomes influenced by PAM50 molecular subtype. LINC00968 may be of relevance to initiation, maintenance or progression of cancers of the female breast.

Keywords: breast cancer, LINC00968, long intergenic non-protein coding RNA 968, 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 worldwide⁵. 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^{6,7}. Patients with metastasis to distant sites, like the brain, are provided a 27% 5-year survival rate^{6,7}. 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, and report here correlation between expression of LINC00968 in primary tumors of the breast and overall patient survival in human breast cancer.

Methods

For Kaplan-Meier survival analysis, we used the Kaplan-Meier plotter online tool⁸ for correlation of LINC00968 mRNA expression levels with overall survival (OS) 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. We had utilized datasets GSE42568² and GSE52604³ for our previous global differential gene expression analysis of female breast cancer⁴.

Results

We performed discovery of genes associated with breast cancer in females by mining two independently published microarray datasets²⁻⁴; here we integrated these findings with comprehensive analysis of patient survival outcomes based on PAM50 molecular subtype⁸.

LINC00968 is differentially expressed in primary tumors of the breast.

We previously found, through global examination of gene expression profiles of 104 breast cancers from patients aged 31 to 89, that the gene encoding long intergenic non-protein coding RNA 968, LINC00968, was among the genes most differentially expressed in tumors of the breast in human breast cancer⁴. When sorting each of the genes expressed in tumors of the breast based on significance of difference as compared to normal breast tissue, LINC00968 had ranked 24 out of 54675 total transcripts, equating to 99.96% differential expression⁴. Differential expression of LINC00968 in female breast cancer was statistically significant⁴.

Analysis of a second microarray dataset^{3,4}, here studying global gene expression patterns in the brain metastases of patients with metastatic breast cancer, again revealed significant differential expression of LINC00968 in tumor tissue of patients with human breast cancer⁴. 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, LINC00968 had ranked 220 out of 41093 total transcripts, equating to 99.5% differential expression⁴. Differential expression of LINC00968 in the brain metastases of patients with breast cancer was statistically significant⁴. These data had suggested that differential expression of LINC00968 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 breast cancer⁴.

LINC00968 expression correlates with survival outcomes in normal-like subtype human breast cancer.

We performed Kaplan-Meier survival analysis to study relationships between tumor LINC00968 mRNA expression levels and survival outcomes in patients with breast cancer. We observed a correlation between LINC00968 expression and overall survival (OS) in patients with normal-like subtype breast cancer which was statistically significant (Figure 1; log rank p -value: 0.05 for overall survival, hazard ratio: 4.24 (0.88-20.45) (Fig. 1)). LINC00968 mRNA levels were a negative prognostic indicator in normal-like subtype breast cancer patients. OS was, on average, markedly greater for normal-like patients with low tumor expression of LINC00968 as compared to normal-like patients with high tumor expression of LINC00968 (Chart 1). LINC00968 primary tumor expression was not correlated with overall survival in basal-like breast cancer (Figure 1; log rank p -value: 0.73 for OS, hazard ratio: 0.92 (0.58-1.47) (Fig. 1)), in luminal A subtype breast cancer (Figure 1; log rank p -value: 0.66 for OS, hazard ratio: 1.17 (0.58-2.38) (Fig. 1)), in luminal B breast cancer (Figure 1; log rank p -value: 0.99 for OS, hazard ratio: 1 (0.57-1.75) (Fig. 1)) or in patients with HER2+ breast cancer (Figure 1; log rank p -value: 0.75 for OS, hazard ratio: 0.92 (0.53-1.57) (Fig. 1)).

Thus, we found here that primary tumor expression of a non-coding RNA that we had previously demonstrated as among the most differentially expressed in human breast cancer was correlated with overall survival in patients with normal-like subtype disease, with mRNA levels of LINC00968 a negative prognostic indicator for normal-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^{6,7}. 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 had previously performed comparative transcriptome analysis using two independently published microarray datasets²⁻⁴, providing evidence that differential expression of long intergenic non-protein coding RNA 968, encoded by LINC00968, 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 cancer⁴.

Importantly, we demonstrated here through examination of survival outcomes in each PAM50 molecular subtype that expression of LINC00968 was correlated with overall survival in patients with normal-like subtype breast cancers. Long intergenic non-protein coding RNA 968 and the molecular processes to which it pertains to may be relevant to the initiation or progression of human breast cancer, particularly in patients with normal-like 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. Mamoor, S., 2020. LINC00968 is differentially expressed in brain metastatic breast cancer.
5. 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.
6. ACS Cancer Facts & Figures 2019.
<https://www.cancer.net/cancer-types/breast-cancermetastatic/statistics>.
7. Survival Rates for Breast Cancer. <https://www.cancer.org/cancer/breast-cancer/understanding-a-breast-cancer-diagnosis/breast-cancer-survival-rates.html>
8. 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.

Figure 1: LINC00968 expression correlates with overall survival in patients with normal-like subtype breast cancer, but not in patients with basal, luminal A, luminal B, or HER2+ subtype cancer.

Depicted in this Kaplan-Meier plot is the probability of overall survival (OS) 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 LINC00968 in patient primary tumors. 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 LINC00968 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.