CA6 is differentially expressed in lymph node metastasis in human breast cancer.

Shahan Mamoor, MS¹

¹shahanmamoor@gmail.com

East Islip, NY USA

Metastasis to the brain is a clinical problem in patients with breast cancer¹⁻³. Between the breast and the brain reside the secondary lymphoid organ, the lymph nodes. We mined published microarray data^{4,5} to compare primary and metastatic tumor transcriptomes for the discovery of genes associated with metastasis to the lymph nodes in humans with metastatic breast cancer. We found that carbonic anhydrase 6, CA6, was among the genes whose expression was most different in the lymph node metastases of patients with metastatic breast cancer as compared to primary tumors of the breast. CA6 mRNA was present at increased quantities in lymph node metastases as compared to primary tumors of the breast. Importantly, expression of CA6 in primary tumors of the breast was correlated with patient post-progression survival, in lymph node negative patients but not in lymph node positive patients. Modulation of CA6 expression may be relevant to the biology by which tumor cells metastasize from the breast to the lymph nodes and the brain in humans with metastatic breast cancer.

Keywords: breast cancer, metastasis, brain metastasis, central nervous system metastasis, lymph node metastasis, carbonic anhydrase 6, CA6, systems biology of breast cancer, targeted therapeutics in breast cancer.

One report described a 34% incidence of central nervous system metastases in patients treated with trastuzumab for breast cancer². More recently, the NEfERT-T clinical trial⁶ which compared administration of either neratinib or trastuzumab in conjunction with paclitaxel demonstrated that in a randomized, controlled setting, in breast cancer patients treated with neratinib, not only was the incidence of central nervous system recurrence significantly lower, the time to central nervous system metastasis was significantly delayed as compared to patients administered trastuzumab⁶. The alarmingly high rate of central nervous system metastasis described, as well as data, both anecdotal² and from a randomized, controlled setting⁶ illustrating that treatment with trastuzumab may be associated with these events demands an enhanced understanding of the transcriptional makeup of brain and lymph node metastatic tissues to support identification of therapeutic targets, whether they are treatment related or not. The lymph nodes reside between the breast and the brain. We performed a global comparative analysis of lymph node metastatic tumor tissues in patients with metastatic breast cancer, compared to primary tumors and normal breast tissues^{4,5}. We discovered transcriptome-wide differential expression of the gene encoding carbonic anhydrase 6, CA6, in lymph node metastatic tissues of patients with metastatic breast cancer.

Methods

We used datasets GSE10893⁴ and GSE124648⁵ for this global differential gene expression analysis of brain metastatic breast cancer in conjunction with GEO2R. GSE10893 was generated using Agilent-011521 Human 1A Microarray G4110A technology with $n=71$ primary breast tumors and $n=7$ lymph node metastases from patients with breast cancer; analysis was performed using platform GPL887. GSE124648 was generated using Affymetrix Human Genome U133A array technology with $n=10$ normal breast tissues and $n=44$ lymph node metastases from patients with breast cancer; analysis was performed using platform GPL96. 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 CA6 gene expression was significantly different between primary tumors of the breast and lymph node metastases in humans with breast cancer using a two-tailed t-test. For Kaplan-Meier survival analysis, we used the Kaplan-Meier plotter online tool⁷ for correlation of CA6 mRNA expression levels with post-progression survival in $n=153$ lymph node positive breast cancer patients and in $n=184$ lymph node negative breast cancer patients.

Results

We performed global comparative transcriptome analysis of metastatic tumor tissues of patients with metastatic breast cancer using published microarray data^{4,5} to describe the transcriptional landscape of metastasis in human breast cancer in an unbiased fashion and for the discovery of novel therapeutic targets.

CA6 is differentially expressed in the lymph node metastases of patients with metastatic breast cancer.

Through blind, systems-level analysis of published microarray data⁴, we identified carbonic anhydrase 6, encoded by CA6, as a differentially expressed gene in the lymph node metastatic tissues of humans with breast cancer (Table 1). When sorting each of the genes expressed in lymph node metastases based on significance of difference as compared to primary tumors of the breast in patients with breast cancer, CA6 ranked 78 out of 18034 total transcripts (Chart 1), equating to 99.6% differential expression. Differential expression of CA6 in the lymph node metastases of patients with metastatic breast cancer was statistically significant (Chart 1; $p=5.23e-04$).

We queried a second microarray dataset⁵ to attempt to validate transcriptome-wide differential expression of CA6 in the lymph node metastases of patients with metastatic breast cancer. We again observed CA6 to be among the genes expressed most quantitatively differently in the lymph node metastases of patients with metastatic breast cancer as compared to primary tumors of the breast. When sorting each of the genes expressed in lymph node metastases based on significance of difference as compared to normal breast tissues, CA6 ranked 4284 out of 22283 total transcripts (Chart 2), equating to 80.8% differential expression. Differential expression of CA6 in the lymph node metastases of patients with metastatic breast cancer approached the level of statistical significance (Chart 2; $p=3.33e-03$).

CA6 is expressed at higher levels in the lymph node metastases of patients with metastatic breast cancer.

We obtained exact mRNA expression levels for CA6, in primary tumors of the breast and in lymph node metastases of patients with metastatic breast cancer to determine direction and statistical significance of change in CA6 expression in lymph node metastatic tissues. We observed significantly increased expression of CA6 in the lymph node metastases of patients with breast cancer as compared to primary tumors of the breast: CA6 was expressed at 0.12 ± 0.32 arbitrary units (AU) in primary tumors of the breast, while it was expressed at 0.59 ± 0.38 AU in lymph node metastatic tissues (Figure 1). The difference in CA6 mRNA levels between primary tumors of the breast and lymph node metastases was statistically significant (Figure 1; $p=0.000459$).

CA6 expression is significantly correlated with survival outcomes in human breast cancer.

We performed Kaplan-Meier survival analysis⁷ in $n=153$ lymph node positive breast cancer patients and $n=184$ lymph node negative breast cancer patients, to evaluate whether CA6 tumor expression was correlated with survival outcomes in breast cancer and to evaluate whether lymph node status influenced correlation of CA6 tumor expression with survival outcomes in human breast cancer. We observed a correlation between primary tumor expression of CA6 and post-progression survival (PPS) in lymph node negative breast cancer, in the upper survival quartile (Figure 2). Lymph node negative patients whose primary tumors expressed low levels of CA6 possessed median PPS of 68.96 months, while lymph node negative patients whose tumors expressed high levels of CA6 possessed median PPS of 34.32 months. This difference in PPS based on CA6 tumor expression in lymph node negative patients with breast cancer was statistically significant (Figure 2, Chart 3; logrank p -value: 0.036 (hazard ratio: 1.63 (1.03-2.58))). CA6 primary tumor expression was not significantly correlated with post-progression survival in lymph node positive patients (Figure 2, Chart 3; logrank p -value: 0.73; hazard ratio: 1.09 (0.66-1.81)).

Thus, by mining published microarray data^{4,5} in an unbiased and systematic fashion, we identified CA6 as among the genes whose expression was most different, transcriptome-wide, in the lymph node metastases of patients with breast cancer, when compared to primary tumors; we observed significantly increased expression of CA6 in lymph node metastases as compared to primary tumors of the breast. Further, we found a significant correlation between CA6 expression and patient survival outcomes, as post-progression survival was significantly greater in patients whose primary tumors expressed low levels of CA6 as compared to patients whose primary tumors expressed high levels of CA6, and this correlation was significant in lymph node negative patients but not in lymph node positive patients.

Discussion

We provided evidence here that carbonic anhydrase 6, CA6, is among the genes whose expression is most different in the lymph node metastases of patients with metastatic breast cancer as compared to primary tumors of the breast, that CA6 mRNA is present at significantly increased quantities in lymph node metastatic tissues as compared to primary tumors of the breast, and that primary tumor CA6 expression is significantly correlated with patient survival outcomes in human breast cancer, in all patients and in lymph node negative patients but not in lymph node positive patients. Evaluation of the effects of genetic depletion of CA6 in mouse models of metastatic breast cancer on metastasis to the central nervous system is merited. Modulation of CA6 expression may be relevant to the processes by which breast cancer cells exit the breast, enter the vasculature and/or lymphatics, reside in the lymph nodes, evade immune clearance, breach the blood-brain barrier and colonize the brain.

References

1. Lin, N.U., Amiri-Kordestani, L., Palmieri, D., Liewehr, D.J. and Steeg, P.S., 2013. CNS metastases in breast cancer: old challenge, new frontiers.
2. Bendell, J.C., Domchek, S.M., Burstein, H.J., Harris, L., Younger, J., Kuter, I., Bunnell, C., Rue, M., Gelman, R. and Winer, E., 2003. Central nervous system metastases in women who receive trastuzumab-based therapy for metastatic breast carcinoma. *Cancer*, 97(12), pp.2972-2977.
3. Tsukada, Y., Fouad, A., Pickren, J.W. and Lane, W.W., 1983. Central nervous system metastasis from breast carcinoma autopsy study. *Cancer*, 52(12), pp.2349-2354.
4. Weigman, V.J., Chao, H.H., Shabalin, A.A., He, X., Parker, J.S., Nordgard, S.H., Grushko, T., Huo, D., Nwachukwu, C., Nobel, A. and Kristensen, V.N., 2012. Basal-like Breast cancer DNA copy number losses identify genes involved in genomic instability, response to therapy, and patient survival. *Breast cancer research and treatment*, 133(3), pp.865-880.
5. Sinn, B.V., Fu, C., Lau, R., Litton, J., Tsai, T.H., Murthy, R., Tam, A., Andreopoulou, E., Gong, Y., Murthy, R. and Gould, R., 2019. SET ER/PR: a robust 18-gene predictor for sensitivity to endocrine therapy for metastatic breast cancer. *NPJ breast cancer*, 5(1), pp.1-8.
6. Awada, A., Colomer, R., Inoue, K., Bondarenko, I., Badwe, R.A., Demetriou, G., Lee, S.C., Mehta, A.O., Kim, S.B., Bachelot, T. and Goswami, C., 2016. Neratinib plus paclitaxel vs trastuzumab plus paclitaxel in previously untreated metastatic ERBB2-positive breast cancer: the NEfERT-T randomized clinical trial. *JAMA oncology*, 2(12), pp.1557-1564.
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: 78
Probe ID: 15868
p-value: 5.23e-04
t: -3.6172825
B: -0.261
Gene: CA6
Gene name: carbonic anhydrase 6

Chart 1: CA6 is differentially expressed in lymph node metastatic breast cancer when comparing lymph node metastases to primary tumors of the breast.

The rank of global differential expression, the probe/transcript ID, the *p*-value with respect to differential expression transcriptome-wide, *t*, a moderated *t*-statistic, *B*, the log-odds of differential expression between the groups compared, the gene and gene name are listed in this chart.

Rank: 4284
Probe ID: 206873_at
p-value: 3.33e-03
t: 3.068338
B: -2.583773
Gene: CA6
Gene name: carbonic anhydrase 6

Chart 2: CA6 is differentially expressed in lymph node metastatic breast cancer when comparing lymph node metastases to normal breast tissues.

The rank of global differential expression, the probe/transcript ID, the *p*-value with respect to differential expression transcriptome-wide, *t*, a moderated *t*-statistic, *B*, the log-odds of differential expression between the groups compared, the gene and gene name are listed in this chart.

Figure 1: CA6 is expressed at significantly higher levels in the lymph node metastases of patients with metastatic breast cancer when compared to primary tumors of the breast.

The mRNA expression level of CA6 in primary tumors of the breast (left) and in lymph node metastases of women with metastatic breast cancer (right) is graphically depicted; the result of a statistical test evaluating significance of difference in CA6 expression between primary tumors of the breast and lymph node metastases, a p -value, is $p=0.000459$.

Figure 2: CA6 primary tumor expression is correlated with post-progression survival in all patients with breast cancer and in lymph node negative patients but not in lymph node positive patients.

Depicted in this Kaplan-Meier plot (**above**) is the probability of post-progression survival in $n=153$ lymph node positive patients (**left**) and $n=184$ lymph node negative patients (**right**) stratified into two groups, based on low or high expression of CA6 in patient primary tumors, in the upper survival quartile. The log rank p -value denoting statistical significance of difference in post-progression 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 CA6 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.

Lymph node negative patients

Low CA6 expression: 68.96 months

High CA6 expression: 34.32 months

Chart 3: In lymph node negative patients, median post-progression survival is inferior in patients with high primary tumor expression of CA6.

The median PPS (post-progression survival) of patients with low primary tumor expression of CA6 and high primary tumor expression of CA6 is listed in this chart, for $n=153$ lymph node positive breast cancer patients and for $n=184$ lymph node negative patients, in the upper survival quartile.