CD226 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 lymph node metastasis in humans with metastatic breast cancer. We found that cluster of differentiation 226, encoded by CD226, 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. Analysis of a separate microarray dataset revealed that CD226 was also differentially expressed in brain metastatic tissues⁵. CD226 mRNA was present at increased quantities in lymph node metastatic tissues as compared to primary tumors of the breast. Expression of CD226 in primary tumors was correlated with patient overall survival in lymph node negative patients but not in lymph node positive patients. Modulation of CD226 expression may be relevant to the biology by which tumor cells metastasize from the breast to the lymph nodes in humans with metastatic breast cancer.

Keywords: breast cancer, metastasis, lymph node metastases, central nervous system metastases, cluster of differentiation 226, CD226, 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 lymph node metastatic tissues to support identification of therapeutic targets, whether they are treatment related or not. We performed a global comparative analysis of primary and metastatic tumors in patients with lymph node metastatic breast cancer^{4,5}. We discovered differential and increased expression of the gene encoding cluster of differentiation 226, CD226, in lymph node metastatic tissues of patients with metastatic breast cancer.

Methods

We used datasets GSE10893⁴ and GSE125989⁵ for this global differential gene expression analysis of lymph node metastatic breast cancer in conjunction with GEO2R. GSE10893 was generated using Agilent-011521 Human 1A Microarray G4110A technology with $n=11$ primary breast tumors and $n=3$ lymph node metastases from patients with breast cancer; analysis was performed using platform GPL885. GSE125989 was generated using Affymetrix Human Genome U133A 2.0 array technology with $n=16$ primary breast tumors and $n=16$ brain metastases from patients with breast cancer; the samples utilized by investigators for this microarray dataset were paired, matched tumor and metastasis samples; analysis was performed using platform GPL571. 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 CD226 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 CD226 mRNA expression levels with overall survival in $n=1309$ lymph node negative patients and $n=889$ lymph node positive 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 lymph node metastasis in human breast cancer in an unbiased fashion and for the discovery of novel therapeutic targets.

CD226 is differentially expressed in the lymph node metastases of patients with lymph node metastatic breast cancer.

Through blind, systems-level analysis of published microarray data⁴, we identified cluster of differentiation 226, encoded by CD226, as a differentially expressed gene in the breast 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, CD226 ranked 207 out of 18034 total transcripts (Chart 1), equating to 98.9% differential expression. Differential expression of CD226 in the lymph node metastases of patients with metastatic breast cancer was statistically significant (Chart 1; $p=2.53e-03$).

To attempt to validate differential transcriptome-wide differential expression of CD226 in human metastatic breast cancer, we queried a second microarray dataset⁵; here, we identified CD226 as a differentially expressed gene in the brain metastatic tissues of patients with breast cancer (Chart 2). When sorting each of the genes expressed in brain metastases based on significance of difference as compared to primary tumors of the breast, CD226 ranked 2613 out of 22277 total transcripts (Chart 2), equating to 88.3% differential expression. Differential expression of CD226 in the brain metastases of patients with metastatic breast cancer trended towards statistical significance (Chart 2; $p=8.14e-02$). Thus, transcriptome-wide differential expression of CD226 in the metastatic tissues of women with metastatic breast cancer was conserved across two independent microarray datasets, in the lymph nodes and in the brain, indicating that this may likely be a *bona fide* biological phenomena associated with metastatic breast cancer in humans.

CD226 is expressed at higher levels in the lymph node metastases of patients with metastatic breast cancer.

We obtained exact mRNA expression levels for CD226, in primary tumors of the breast and in lymph node metastasis of patients with lymph node metastatic breast cancer to determine direction and statistical significance of change in CD226 expression in lymph node metastatic tissues. We observed significantly higher expression of CD226 in the lymph node metastases of patients with breast cancer as compared to primary tumors of the breast: CD226 was expressed at 0.27 ± 0.28 units (AU) in primary tumors of the breast, while it was expressed at 0.63 ± 0.30 AU in lymph node metastatic tissues. The difference in CD226 mRNA levels between primary tumors of the breast and lymph node metastatic tissues was statistically significant (Figure 1; $p=0.002058$).

CD226 expression is significantly correlated with survival outcomes in human breast cancer.

We performed Kaplan-Meier survival analysis⁷ in $n=889$ lymph node positive patients and $n=1309$ lymph node negative patients, to evaluate whether CD226 tumor expression was correlated with survival outcomes in breast cancer. We observed a correlation between primary tumor expression of CD226 and overall survival (OS) in lymph node negative patients with breast cancer (Figure 2). Lymph node negative patients whose primary tumors expressed low levels of CD226 possessed median OS of 120 months, as compared to 197.62 months for lymph node negative patients whose tumors expressed high levels of CD226. This difference in OS based on CD226 tumor expression in lymph node negative patients with breast cancer was statistically significant (Figure 2, Chart 3; logrank p -value: 0.055; hazard ratio: 0.71 (0.5-1.01)). Primary tumor expression of CD226 was not correlated with RFS in lymph node positive breast cancer patients (Figure 2, Chart 3; logrank p -value: 0.87; hazard ratio: 1.03 (0.74-1.43)).

Thus, by mining published microarray data^{4,5} in an unbiased and systematic fashion, we identified cluster of differentiation 226, encoded by CD226, 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 of the breast; we observed significantly increased expression of CD226 in lymph node metastases as compared to primary tumors of the breast. Further, we observed a correlation between CD226 expression and patient survival outcomes in lymph node negative patients but not in lymph node positive patients.

Discussion

We provided evidence here that cluster of differentiation 226, encoded by CD226, is among the genes whose expression is most different in the lymph node metastases of patients with lymph node metastatic breast cancer, that CD226 mRNA is present at significantly increased quantities in lymph node metastatic tissues as compared to primary tumors of the breast, and that primary tumor CD226 expression is correlated with patient survival outcomes in lymph node negative human breast cancer. Evaluation of the effects of genetic depletion of CD226 in mouse models of metastatic breast cancer on metastasis to the lymph nodes and to the central nervous system is merited. Modulation of CD226 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. Iwamoto, T., Niikura, N., Ogiya, R., Yasojima, H., Watanabe, K.I., Kanbayashi, C., Tsuneizumi, M., Matsui, A., Fujisawa, T., Iwasa, T. and Shien, T., 2019. Distinct gene expression profiles between primary breast cancers and brain metastases from pair-matched samples. *Scientific reports*, 9(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: 207
Probe ID: 20783
p-value: 2.53e-03
t: -3.1190631
B: -1.5942
Gene: CD226
Gene name: cluster of differentiation 226 homolog

Chart 1: CD226 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: 2613
probe ID: 207315_at
p-value: 8.14e-02
t: 1.7908173
B: -4.5189
Gene: CD226
Gene name: cluster of differentiation 226

Chart 2: CD226 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, 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: CD226 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 CD226 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 CD226 expression between primary tumors of the breast and lymph node metastases is $p=0.002058$.

Figure 2: CD226 primary tumor expression correlates with overall survival in patients with lymph node negative breast cancer but not in lymph node positive patients.

Depicted in these Kaplan-Meier plots is the probability of overall survival in $n=889$ patients with lymph node positive breast cancer (**left**) and $n=1309$ patients with lymph node negative breast cancer (**right**), stratified into two groups, based on low or high expression of CD226 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 each graph is the number of patients at risk (number of patients alive) per interval, after stratification based on CD226 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 CD226 expression: 120 months

High CD226 expression: 197.62 months

Chart 3: Median overall survival is inferior in lymph node negative breast cancer patients with low primary tumor expression of CD226.

The median OS (overall survival) of patients with low primary tumor expression of CD226 and high primary tumor expression of CD226 is listed in this chart, for $n=889$ lymph node negative breast cancer patients.