Differential expression of glial fibrillary acidic protein (GFAP) in brain metastatic human breast cancer distinguishes astrocytic gene expression from the metastatic gene expression program.

Shahan Mamoor, MS¹

¹shahanmamoor@gmail.com

East Islip, NY USA

Metastasis to the brain is a clinical problem in patients with breast cancer¹⁻³. We mined published microarray data^{4,5} to compare primary and metastatic tumor transcriptomes for the discovery of genes associated with brain metastasis in humans with metastatic breast cancer. We found that the glial fibrillary acidic protein, encoded by GFAP, was among the genes whose expression was most different in the brain metastases of patients with metastatic breast cancer as compared to primary tumors of the breast. GFAP mRNA was present at increased quantities in brain metastatic tissues as compared to primary tumors of the breast. Importantly, expression of GFAP in primary tumors was significantly correlated with patient overall survival, in lymph node positive breast cancer patients but not in lymph node negative breast cancer patients. Expression of two separate astrocyte markers, S100B and ALDH1L1, were not significantly different when comparing brain metastases and primary tumors of the breast. Specific differential expression of an astrocyte marker in brain metastatic breast cancer tissues distinguishes glial gene expression from the brain metastatic gene expression program while demonstrating GFAP as a unique example of overlap between cell-type specific gene expression and transcriptional perturbation in human disease.

Keywords: breast cancer, metastasis, brain metastases, central nervous system metastases, glial fibrillary acidic protein, GFAP, 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 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 brain metastatic breast cancer^{4,5}. We discovered differential and increased expression of the gene encoding the glial fibrillary acidic protein, GFAP, in brain metastatic tissues of patients with metastatic breast cancer.

Methods

We used datasets GSE10893⁴ and GSE52604⁵ 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 $n=11$ primary breast tumors and $n=3$ brain metastases from patients with breast cancer; analysis was performed using platform GPL885. GSE52604 was generated using Agilent-014850 Whole Human Genome Microarray 4x44K G4112F with $n=10$ normal breast tissues and $n=35$ brain metastases from patients with breast cancer; analysis was performed using platform GPL6480; we also compared $n=10$ normal breast tissues and $n=10$ normal brain tissues, as well as $n=10$ normal brain tissues and $n=35$ brain metastases from patients with breast cancer using this dataset. 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 GFAP gene expression was significantly different between primary tumors of the breast and brain 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 GFAP mRNA expression levels with overall survival in $n=452$ lymph node positive breast cancer patients and $n=726$ 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 brain metastasis in human breast cancer in an unbiased fashion and for the discovery of novel therapeutic targets.

GFAP is differentially expressed in the brain metastases of patients with brain metastatic breast cancer.

Through blind, systems-level analysis of published microarray data⁴, we identified the glial fibrillary acidic protein, encoded by GFAP, 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 brain metastases based on significance of difference as compared to primary tumors of the breast in patients with breast cancer, GFAP ranked 2 out of 17418 total transcripts (Chart 1), equating to 99.99% differential expression. Differential expression of GFAP in the brain metastases of patients with metastatic breast cancer was statistically significant (Chart 1; $p=1.70e-07$).

Differential gene expression analyses that utilize primary and metastatic tumor tissues to discover genes associated with metastasis can be challenging due to the relatively heterogeneous cellular composition and constitution of the tumor as compared to normal, untransformed (benign) tissues. Thus, to attempt to validate differential transcriptome-wide differential expression of GFAP in brain metastases in human metastatic breast cancer, we queried a second microarray dataset⁵, here comparing normal breast tissues rather than primary tumors of the breast to brain metastases. Again, we identified GFAP 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 normal breast tissues, GFAP ranked 3970 out of 41093 total transcripts (Chart 2), equating to 90.3% differential expression. Differential expression of GFAP in the brain metastases of patients with metastatic breast cancer was statistically significant (Chart 2; $p=7.56e-05$). Thus, differential expression of GFAP, transcriptome-wide, in the brain metastases of women with metastatic breast cancer was conserved across two independent microarray datasets, both when compared to benign and transformed breast tissues.

Differential expression of GFAP is not due to astrocyte cell contamination, nor is it a reflection of non-specific up-regulation of glial gene expression.

GFAP is a canonical marker of the astrocyte⁸⁻¹⁰. To determine whether differential expression of GFAP could be attributed to contamination of brain metastatic tissues with astrocytes rich in GFAP, or whether up-regulation of GFAP in brain metastatic tissues of patients with breast cancer could be ascribed to a non-specific up-regulation of glial gene expression, we asked whether two separate markers of the astrocyte, S100B and ALDH1L1 were also differentially expressed in brain metastatic tissues when compared to primary tumors of the breast. They were not: S100B ranked 10244 out of 17418 total transcripts, equivalent to 41.2% differential expression ($p=4.02e-01$), while ALDH1L1 ranked 9253 transcripts, equivalent to 46.9% differential expression ($p=3.34e-01$) (Chart 3).

We attempted to further distinguish between cell-type-specific gene expression programs and the brain metastatic gene expression program in human brain metastatic breast cancer by querying a separate gene expression dataset, comparing normal brain tissues to normal breast tissues, as well as comparing brain metastases to normal breast tissues. While astrocyte markers S100B and ALDH1L1 were significantly differentially expressed when comparing normal breast to normal brain, demonstrating organ-specific gene expression (Chart 4: S100B: 432 out of 41093 transcripts (98.9%) ($p=6.43e-14$) and ALDH1L1: 3875 out of 41093 transcripts (90.6%) ($p=7.79e-08$)), when comparing normal breast tissues to brain metastases from humans with brain metastatic breast cancer, expression of S100B and ALDH1L1 was not significantly different (Chart 5: S100B: 32121 out of 41093 transcripts (21.8%) ($p=4.82e-01$), and ALDH1L1: 20931 out of 41093 transcripts (49.1%) ($p=1.00e-01$)). Thus, while GFAP is a marker of the astrocyte, and up-regulated in brain metastatic tissues when compared to the tumors from which they arise, this is a specific phenomena associated with the metastatic gene expression program in human brain metastatic breast cancer, as S100B and ALDH1L1, two separate markers of the astrocyte whose expression distinguish the breast from the brain, are not similarly perturbed in brain metastases.

GFAP is expressed at higher levels in the brain metastases of patients with metastatic breast cancer.

We obtained exact mRNA expression levels for GFAP, in primary tumors of the breast and in brain metastasis of patients with brain metastatic breast cancer to determine direction and statistical significance of change in GFAP expression in brain metastatic tissues. We observed significantly higher expression of GFAP in the brain metastases of patients with breast cancer as compared to primary tumors of the breast: GFAP was expressed at -0.14 ± 0.34 arbitrary units (AU) in primary tumors of the breast, while it was expressed at 1.84 ± 0.49 AU in brain metastatic tissues. The difference in GFAP mRNA levels between primary tumors of the breast and brain metastatic tissues was statistically significant (Figure 1; $p<0.00001$).

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

We performed Kaplan-Meier survival analysis⁸ in $n=452$ lymph node positive breast cancer patients and in $n=726$ lymph node negative breast cancer patients, to evaluate whether GFAP tumor expression was correlated with survival outcomes in humans with breast cancer. We observed a statistically significant correlation between primary tumor expression of GFAP and overall survival (OS) in patients with lymph node positive but not lymph node negative breast cancer (Figure 2). Lymph node positive patients whose primary tumors expressed low levels of GFAP possessed median OS of 55.52 months, while lymph node positive patients whose tumors expressed high levels of GFAP possessed median OS of 40.15 months. This difference in OS based on GFAP tumor expression in patients with lymph node positive breast cancer was statistically significant (Figure 2, Chart 6; logrank p -value: 0.026; hazard ratio: 1.46 (1.04-2.04)). Overall survival was not significantly correlated with GFAP primary tumor expression in patients with lymph node negative breast cancer ((Figure 2, Chart 6; logrank p -value: 0.67; hazard ratio: 0.93 (0.66-1.31)).

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

Discussion

We provided evidence here that the gene encoding the glial fibrillary acidic protein (GFAP) is among those whose expression is most different in the brain metastases of patients with brain metastatic breast cancer when compared to primary tumors of the breast and to normal breast tissues, that GFAP mRNA is present at significantly increased quantities in brain metastatic tissues as compared to primary tumors of the breast, and that primary tumor GFAP expression is significantly correlated with patient survival outcomes in human breast cancer. Though GFAP is an astrocyte marker, its increased expression in brain metastatic tissues in patients with metastatic breast cancer is not a result of astrocytic contamination during tissue sampling, nor is it a reflection of non-specific up-regulation of astrocytic gene expression during metastatic progression from primary breast tumor to brain metastasis, as two other markers of the astrocyte, S100B and ALDH1L1, are not differentially expressed when comparing brain metastasis to primary tumor of the breast. GFAP up-regulation in brain metastasis is a specific transcriptional event associated with brain metastatic breast cancer and the astrocytic gene expression program can be distinguished from the brain metastatic breast cancer gene expression program by studying the differential gene expression patterns of multiple astrocytic markers.

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., Shabalín, 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. 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.
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.
8. Eng, L.F., 1985. Glial fibrillary acidic protein (GFAP): the major protein of glial intermediate filaments in differentiated astrocytes. *Journal of neuroimmunology*, 8, pp.203-214.
9. Brenner, M., Kisseberth, W.C., Su, Y., Besnard, F. and Messing, A., 1994. GFAP promoter directs astrocyte-specific expression in transgenic mice. *Journal of Neuroscience*, 14(3), pp.1030-1037.
10. Lee, Y., Messing, A., Su, M. and Brenner, M., 2008. GFAP promoter elements required for region-specific and astrocyte-specific expression. *Glia*, 56(5), pp.481-493.

Rank: 2
Probe ID: 18119
p-value: 1.70e-07
t: -8.82
B: 7.099786
Gene: GFAP
Gene name: glial fibrillary acidic protein

Chart 1: GFAP is among the genes whose expression is most significantly different, transcriptome-wide, in human brain metastatic breast cancer when comparing brain 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: 3970
probe ID: A_23_P66593
p-value: 7.56e-05
t: -4.3388655
B: 1.222432
Gene: GFAP
Gene name: glial fibrillary acidic protein

Chart 2: GFAP is differentially expressed in brain metastatic breast cancer when comparing brain metastases to normal breast tissues.

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.

S100B

Rank: 10244
probe ID: 7015
p-value: 4.02e-01
t: 8.61e-01
B: -5.773491
Gene: S100B
Gene name: S100 calcium binding protein B

ALDH1L1

Rank: 9253
probe ID: 12139
p-value: 3.34e-01
t: -9.96e-01
B: -5.653382
Gene: ALDH1L1
Gene name: aldehyde dehydrogenase 1 family, member L1

Chart 3: Two separate markers of the astrocyte, are not differentially expressed in human brain metastatic breast cancer when comparing brain 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.

GFAP

Rank: 432

probe ID: A_23_P66593

p-value: 6.43e-14

t: 1.70e+01

B: 22.0362711

Gene: GFAP

Gene name: glial fibrillary acidic protein

S100B

Rank: 1252

probe ID: A_23_P143526

p-value: 4.94e-11

t: 1.21e+01

B: 15.35978145

Gene: S100B

Gene name: S100 calcium binding protein B

ALDH1L1

Rank: 3875

probe ID: A_23_P258887

p-value: 7.79e-08

t: 7.96

B: 7.86958497

Gene: ALDH1L1

Gene name: aldehyde dehydrogenase 1 family, member L1

Chart 4: Astrocyte markers GFAP, S100B, and ALDH1L1 are all differentially expressed when comparing normal brain tissue to normal breast tissue.

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.

S100B

Rank: 32131

probe ID: A_23_P143526

p-value: 4.82e-01

t: 0.7082518

B: -6.474475

Gene: S100B

Gene name: S100 calcium binding protein B

ALDH1L1

Rank: 20931

probe ID: A_24_P941870

p-value: 1.00e-01

t: 1.6776768

B: -5.350086

Gene: ALDH1L1

Gene name: aldehyde dehydrogenase 1 family, member L1

Chart 5: Astrocyte markers S100B and ALDH1L1 are not differentially expressed when comparing brain metastases from humans with metastatic breast cancer to normal breast tissue.

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: GFAP is expressed at significantly higher levels in the brain metastases of patients with metastatic breast cancer when compared to primary tumors of the breast.

The mRNA expression level of GFAP in primary tumors of the breast (left) and in brain metastases of women with metastatic breast cancer (right) is graphically depicted; the result of a statistical test evaluating significance of difference in GFAP expression between primary tumors of the breast and brain metastases, a p -value, is $p < 0.0001$.

Figure 2: Significant correlation between GFAP primary tumor expression and overall survival in patients with breast cancer, in lymph node positive patients but not lymph node negative patients.

Depicted in these Kaplan-Meier plots is the probability of overall survival for $n=452$ total lymph node positive breast cancer patients (left) and $n=726$ lymph node negative breast cancer patients (right), stratified into two groups, based on low or high expression of GFAP 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 GFAP 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.

Low GFAP expression: 55.52 months

High GFAP expression: 40.15 months

Chart 6: High primary tumor expression of GFAP confers inferior median overall survival in lymph node positive breast cancer patients.

The median OS (overall survival) of patients with low primary tumor expression of GFAP and high primary tumor expression of GFAP is listed in this chart, for $n=452$ lymph node positive breast cancer patients.