21-hydroxylase, encoded by CYP21A2, is a differentially expressed gene in human metastatic breast cancer, in the brain and in the lymph nodes.

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 21-hydroxylase, also known as cytochrome P450, family 21, subfamily A, polypeptide 2, encoded by CYP21A2, was among the genes whose expression was most different in the brain and lymph node metastases of patients with metastatic breast cancer. CYP21A2 mRNA was present at increased quantities in brain metastatic tissues as compared to primary tumors of the breast. Importantly, expression of CYP21A2 in primary tumors was significantly correlated with patient recurrence-free survival in patients with breast cancer. Modulation of CYP21A2 expression may be relevant to the biology by which tumor cells metastasize from the breast to the brain while evading immune clearance in the lymph nodes in humans with metastatic breast cancer. These data are one piece of evidence suggesting a common ancestor or tumor clone for brain and lymph node metastases that originate from the primary tumor, alluding to patterns in developmental origin and migratory pathways through the lymph node in human brain metastatic breast cancer.

Keywords: breast cancer, metastasis, brain metastases, central nervous system metastases, lymph node metastases, 21-hydroxylase, CYP21A2, cytochrome P450, family 21, subfamily A, polypeptide 2, 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 21-hydroxylase, CYP21A2, in brain 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 $n=11$ primary breast tumors and $n=3$ brain metastases from patients with breast cancer; analysis was performed using platform GPL885. 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 CYP21A2 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 CYP21A2 mRNA expression levels in primary tumors of the breast with recurrence-free survival in $n=3951$ breast cancer patients.

Results

We performed global comparative transcriptome analysis of metastatic and primary tumor tissues of patients with metastatic breast, as well as normal tissues of the breast 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.

CYP21A2 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 21-hydroxylase, also known as cytochrome P450, family 21, subfamily A, polypeptide 2, encoded by CYP21A2, as a differentially expressed gene in the brain 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, CYP21A2 ranked 159 out of 17418 total transcripts (Chart 1), equating to 99.1% differential expression. Differential expression of CYP21A2 in the brain metastases of patients with metastatic breast cancer was statistically significant (Chart 1; $p=3.66E-04$).

CYP21A2 is differentially expressed in the lymph node metastases of patients with breast cancer.

To determine whether differential expression of CYP21A2 in human breast cancer could be validated in a second microarray dataset, and whether this was specific to brain metastases or could be observed in other anatomical locations, we queried a second microarray dataset⁵, here comparing normal breast tissues to lymph node metastases. Again, we identified a probe recognizing CYP21A2 and a CYP21A2 pseudogene as differentially expressed in the lymph node metastatic tissues of patients with breast cancer (Chart 2). When sorting each of the genes expressed in lymph node metastases based on significance of difference as compared to normal breast tissues, CYP21A2 ranked 7007 out of 22283 total transcripts (Chart 2), equating to 68.6% differential expression. Differential expression of CYP21A2/CYP21A2 pseudogene in the lymph node metastases of patients with metastatic breast cancer approached the level of statistical significance (Chart 2; $p=2.72E-02$). Thus, differential expression of CYP21A2, transcriptome-wide, in metastatic tissues of women with metastatic breast cancer was conserved across two independent microarray datasets, in the brain and in the lymph nodes.

CYP21A2 is expressed at higher levels in the brain metastases of patients with metastatic breast cancer.

We obtained exact mRNA expression levels for CYP21A2, 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 CYP21A2 expression in brain metastatic tissues. CYP21A2 was expressed at higher levels in the brain metastases of patients with breast cancer as compared to primary tumors of the breast, and this difference was statistically significant (Figure 1; $p=0.000516$).

CYP21A2 expression is significantly correlated with survival outcomes in human breast cancer.

We performed Kaplan-Meier survival analysis⁸ in 3951 breast cancer patients in total, to evaluate whether CYP21A2 tumor expression was correlated with survival outcomes in breast cancer. We observed a statistically significant correlation between primary tumor expression of CYP21A2 and recurrence-free survival (RFS) in patients with breast cancer (Figure 2). Patients whose primary tumors expressed low levels of CYP21A2 possessed median RFS of 38 months, while patients whose tumors expressed high levels of CYP21A2 possessed median RFS of 61 months. This difference in RFS based on CYP21A2 tumor expression in patients with breast cancer was statistically significant (Figure 2, Chart 3; logrank p -value: $4.2e-6$; hazard ratio: 0.77 (0.69-0.86)).

Thus, by mining published microarray data^{4,5} in an unbiased and systematic fashion, we identified 21-hydroxylase, encoded by CYP21A2, as among the genes whose expression was most different, transcriptome-wide, in the brain and lymph node metastases of patients with breast cancer, when compared to primary tumors of the breast and to normal breast tissues, respectively; we observed significantly increased expression of CYP21A2 in brain metastases as compared to primary tumors of the breast. Further, we found a significant correlation between CYP21A2 expression and patient survival outcomes, as recurrence-free survival was significantly greater in patients whose primary tumors expressed higher levels of CYP21A2 as compared to patients whose primary tumors expressed lower levels of CYP21A2.

Discussion

We provided evidence here that CYP21A2 is among the genes whose expression is most different in the brain and lymph node metastases of patients with metastatic breast cancer, that CYP21A2 mRNA is present at significantly increased quantities in brain metastatic tissues as compared to primary tumors of the breast, and that primary tumor CYP21A2 expression is significantly correlated with patient survival outcomes in human breast cancer. Evaluation of the effects of genetic depletion of CYP21A2 in mouse models of metastatic breast cancer on metastasis to the central nervous system is merited. Modulation of CYP21A2 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. These data are one piece of evidence that allude to a common cell or cell(s) of origin for brain and lymph node metastases in human metastatic breast cancer.

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: 159
Probe ID: 17830
p-value: 3.66E-04
t: -4.51
B: 0.333346
Gene: CYP21A2
Gene name: cytochrome P450, family 21, subfamily A, polypeptide 2

Chart 1: CYP21A2 is differentially expressed in 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: 7007

probe ID: 214622_at

p-value: 2.72E-02

t: -2.268475

B: -4.486496

Gene: CYP21A1P///CYP21A2

Gene name: cytochrome P450 family 21 subfamily A member 1, pseudogene///cytochrome P450 family 21 subfamily A member 2

Chart 2: CYP21A2 is differentially expressed in the lymph node metastases of patients with breast cancer when comparing lymph node 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.

Figure 1: CYP21A2 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 CYP21A2 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 CYP21A2 expression between primary tumors and brain metastases is $p=0.000516$.

Figure 2: Significant correlation between CYP21A2 primary tumor expression and recurrence-free survival in patients with breast cancer.

Depicted in this Kaplan-Meier plot is the probability of recurrence-free survival for $n=3951$ total breast cancer patients stratified into two groups, based on low or high expression of CYP21A2 in primary tumors of patients. The log rank p -value denoting statistical significance of difference in recurrence-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 CYP21A2 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 CYP21A2 expression: 38 months

High CYP21A2 expression: 61 months

Chart 3: Median recurrence-free survival is inferior in patients with low primary tumor expression of CYP21A2.

The median RFS (recurrence-free survival) of patients with low primary tumor expression of CYP21A2 and high primary tumor expression of CYP21A2 is listed in this chart, for $n=3951$ breast cancer patients.