IRX2 is differentially expressed in the lymph nodes of patients with metastatic 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. The lymph nodes reside between the breast and the brain. 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 the Iroquois homeobox transcription factor IRX2 was among the genes whose expression was most quantitatively different in the lymph node metastases of patients with metastatic breast cancer as compared to primary tumors of the breast. IRX2 mRNA was present at increased quantities in lymph node metastases as compared to primary tumors of the breast. Importantly, expression of IRX2 in primary tumors of the breast was significantly correlated with patient recurrence-free survival. IRX2 differential expression was conserved when comparing normal breast and normal brain, suggesting that IRX2 modulation in lymph node metastatic tissues may perhaps reflect gene expression changes that facilitate metastasis to the brain in humans.

Keywords: breast cancer, metastasis, brain metastases, central nervous system metastases, iroquois homeobox transcription factor, IRX2, 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, comparing to primary tumors and normal breast tissues^{4,5}. We discovered transcriptome-wide differential expression of the gene encoding the homeobox transcription factor IRX2 in lymph node 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 $n=71$ primary breast tumors and $n=7$ lymph node metastases from patients with breast cancer; analysis was performed using platform GPL887. GSE52604 was generated using Affymetrix Human Genome U133A array technology with $n=10$ normal breast tissues and $n=10$ normal brain tissues; analysis was performed using platform GPL6480. 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 IRX2 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 IRX2 mRNA expression levels with recurrence-free survival in $n=1764$ 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.

IRX2 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 the

1 iroquois homeobox 2 transcription factor, encoded by IRX2, as a differentially expressed gene in
2 the lymph node metastatic tissues of humans with breast cancer (Table 1). When sorting each of
3 the genes expressed in lymph node metastases based on significance of difference as compared to
4 primary tumors of the breast in patients with breast cancer, IRX2 ranked 21 out of 18034 total
5 transcripts (Chart 1), equating to 99.9% differential expression. Differential expression of IRX2
in the lymph node metastases of patients with metastatic breast cancer was statistically significant
(Chart 1; $p=5.95e-05$).

6 We queried a second microarray dataset to attempt to validate transcriptome-wide
7 differential expression of IRX2 in the lymph node metastases of patients with metastatic breast
8 cancer. We again observed IRX2 to be among the genes expressed most quantitatively
9 differently in the lymph node metastases of patients with metastatic breast cancer as compared to
10 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, IRX2 ranked
5517 out of 22283 total transcripts (Chart 2), equating to 75.2% differential expression.
Differential expression of IRX2 in the lymph node metastases of patients with metastatic breast
cancer approached the level of statistical significance (Chart 2; $p=3.52e-04$).

11 **IRX2 is expressed at higher levels in the lymph node metastases of patients with metastatic** 12 **breast cancer.**

13 We obtained exact mRNA expression levels for IRX2, in primary tumors of the breast and
14 in lymph node metastases of patients with metastatic breast cancer to determine direction and
15 statistical significance of change in IRX2 expression in lymph node metastatic tissues. We
16 observed significantly higher expression of IRX2 in the lymph node metastases of patients with
17 breast cancer as compared to primary tumors of the breast: IRX2 was expressed at 1.37 ± 1.59
18 arbitrary units (AU) in primary tumors of the breast, while it was expressed at -1.21 ± 1.09 AU in
lymph node metastatic tissues (Figure 1). The difference in IRX2 levels between primary tumors
of the breast and lymph node metastatic tissues was statistically significant (Figure 1;
 $p=0.000081$).

19 **IRX2 expression is significantly correlated with survival outcomes in human breast cancer.**

20 We performed Kaplan-Meier survival analysis⁸ in $n=1764$ breast cancer patients in total, to
21 evaluate whether IRX2 tumor expression was correlated with survival outcomes in breast cancer.
22 We observed a statistically significant correlation between primary tumor expression of IRX2 and
23 recurrence-free survival (RFS) in patients with breast cancer (Figure 2). Patients whose primary
24 tumors expressed low levels of IRX2 possessed median RFS of 34 months, while patients whose
tumors expressed high levels of IRX2 possessed median RFS of 39 months. This difference in
RFS based on IRX2 tumor expression in patients with breast cancer was statistically significant
(Figure 2, Chart 3; logrank p -value: 0.089; hazard ratio: 0.87 (0.75-1.02)).

25 Thus, by mining published microarray data^{4,5} in an unbiased and systematic fashion, we
26 identified IRX2 as among the genes whose expression was most different, transcriptome-wide, in
27 the lymph node metastases of patients with breast cancer, when compared to primary tumors of
28 the breast; we observed significantly increased expression of IRX2 in lymph node metastases as
compared to primary tumors of the breast. Further, we found a significant correlation between
IRX2 expression and patient survival outcomes, as recurrence-free survival was significantly

greater in patients whose primary tumors expressed higher levels of IRX2 as compared to patients whose primary tumors expressed lower levels of IRX2.

These data came to us with interest as we recently described the differential expression of IRX2, IRX3 and IRX5 in diffuse intrinsic pontine glioma⁸, an aggressive and uniformly fatal cancer of the brainstem. We have also recently described the shared differential expression of a network of genes, in both the lymph node and brain metastases of patients with metastatic breast cancer, suggesting that in humans with breast cancer, metastasis to the brain is intrinsically linked to metastasis to the lymph nodes⁹⁻³⁴. This prompted us to ask: is the differential expression of IRX2 that we have found here, in the lymph node metastases of patients with metastatic breast cancer a presage of brain metastasis, or in other words, a reflection of a breast-to-brain transition in metastatic breast cancer?

To answer this, we evaluated global transcription in the normal breast and the normal brain using an independent microarray dataset to determine whether IRX2 or other IRX transcription factors were differentially expressed when comparing breast and the brain. In fact, we found that IRX2, IRX4 and IRX5 were all among the genes whose expression changed most significantly transcriptome-wide when comparing normal breast and normal brain. When sorting each of the genes expressed in the brain based on significance of difference as compared to the breast, IRX2 ranked 742 out of 41093 total transcripts (Chart 3), equating to 98.2% differential expression. Differential expression of IRX2 between the normal breast and normal brain was statistically significant (Chart 3; $p=2.19\text{e-}12$). When sorting each of the genes expressed in the brain based on significance of difference as compared to the breast, IRX4 ranked 228 out of 41093 total transcripts (Chart 4), equating to 99.4% differential expression. Differential expression of IRX4 between the normal breast and normal brain was statistically significant (Chart 4; $p=2.16\text{e-}15$). When sorting each of the genes expressed in the brain based on significance of difference as compared to the breast, IRX5 ranked 214 out of 41093 total transcripts (Chart 5), equating to 99.5% differential expression. Differential expression of IRX5 between the normal breast and normal brain was statistically significant (Chart 5; $p=1.46\text{e-}15$).

Though our computational tool provides exact mRNA expression values for only the 250 most differentially expressed genes, limiting our ability to determine exact mRNA expression levels for IRX2, we were able to determine that both IRX4 and IRX5 were both expressed at significantly higher levels in the breast as compared to the brain. IRX4 was expressed at -3.39 ± 0.65 arbitrary units (AU) in normal breast, while it was expressed at 2.45 ± 0.69 in normal brain (Figure 3). The difference in IRX4 levels between human breast and brain tissues was statistically significant (Figure 3; $p<0.00001$). IRX5 was expressed at -3.89 ± 1.10 arbitrary units (AU) in normal breast, while it was expressed at 3.39 ± 0.43 AU in normal brain (Figure 4). The difference in IRX5 levels between human breast and brain tissues was statistically significant (Figure 4; $p<0.00001$).

Discussion

We provided evidence here that the iroquois homeobox transcription factor IRX2 was 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 IRX2 mRNA is present at significantly increased quantities in lymph node metastatic tissues as compared to primary tumors of the breast, and that primary tumor IRX2 expression is significantly correlated with patient survival outcomes in human breast cancer. Patterns of IRX2 expression among the

1 breast, brain, tumors of the breast, lymph node metastases, and tumors of the brain, with
2 increased IRX2 expression in DIPG tumor when compared to the brain, and with decreased IRX2
3 expression in lymph node metastases when compared to tumors of the breast, but with IRX2
4 among the genes whose expression is most significantly different transcriptome-wide from breast
5 to brain, are indicative of an IRX2 expression gradient with lowest IRX2 expression in the brain
6 and highest IRX2 expression in the breast, suggesting that relative mid-level expression of IRX2
7 in lymph node metastases and tumors of the brain may disrupt homeotic processes influenced by
8 IRX2 gradients. Differential expression of IRX2 when comparing lymph node metastases and
9 primary tumors of the breast, as well as when comparing normal breast and normal brain,
10 together with our previous findings that IRX2 is differentially expressed in diffuse intrinsic
11 pontine glioma, suggests that IRX2 perturbation in lymph node metastases may be a component
12 of a breast-to-brain gene expression program in metastatic breast cancer, further providing one
13 piece of molecular evidence arguing that lymph node metastasis is intrinsically linked to brain
14 metastasis in humans with metastatic breast cancer.
15
16
17
18
19
20
21
22
23
24
25
26
27
28

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. 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. Mamoor, S., 2020. Iroquois transcription factors are perturbed in diffuse intrinsic pontine glioma.
9. Mamoor, S., 2020. The joining chain of multimeric IgA and IgM is differentially expressed in metastatic breast cancer, both in metastases to the brain and to the lymph nodes.
10. Mamoor, S., 2020. Matrix metalloproteinase 14 (MMP14) is differentially expressed in breast cancer metastases.
11. Mamoor, S., 2020. WISP2 is differentially expressed in metastatic breast cancer, both in metastases to the brain and to the lymph nodes.
12. Mamoor, S., 2020. POSTN is differentially expressed in metastatic breast cancer, both in metastases to the brain and to the lymph nodes.

13. Mamoor, S., 2020. CD3E is differentially expressed in metastatic breast cancer, both in metastases to the brain and to the lymph nodes.
14. Mamoor, S., 2020. HSPB6 is differentially expressed in metastatic breast cancer, both in metastases to the brain and to the lymph nodes.
15. Mamoor, S., 2020. SERPINF1 is differentially expressed in metastatic breast cancer, both in metastases to the brain and to the lymph nodes.
16. Mamoor, S., 2020. Runx1T1 is differentially expressed in metastatic breast cancer, both in metastases to the brain and to the lymph nodes.
17. Mamoor, S., 2020. KRT14 is differentially expressed in metastatic breast cancer, both in metastases to the brain and to the lymph nodes.
18. Mamoor, S., 2020. Differential expression of VEGF-A in the brain and lymph node metastases of patients with breast cancer.
19. Mamoor, S., 2020. LYVE1 is differentially expressed in the lymph node and brain metastases of patients with metastatic breast cancer.
20. Mamoor, S., 2020. Rac3 is differentially expressed in the lymph node and brain metastases of patients with metastatic breast cancer.
21. Mamoor, S., 2020. Elastin is differentially expressed in metastatic breast cancer, both in metastases to the brain and to the lymph nodes.
22. Mamoor, S., 2020. Biglycan (BGN) is differentially expressed in metastatic breast cancer, both in metastases to the brain and to the lymph nodes.
23. Mamoor, S., 2020. Versican (VCAN) is differentially expressed in metastatic breast cancer, both in metastases to the brain and to the lymph nodes.
24. Mamoor, S., 2020. Lumican is differentially expressed in metastatic breast cancer, both in metastases to the brain and to the lymph nodes.
25. Mamoor, S., 2020. Dystonin is differentially expressed in metastatic breast cancer, both in metastases to the brain and to the lymph nodes.
26. Mamoor, S., 2020. Decorin is differentially expressed in metastatic breast cancer, both in metastases to the brain and to the lymph nodes.
27. Mamoor, S., 2020. Differential expression of the mutS homolog 6, MSH6, in the brain and lymph node metastases of patients with breast cancer.
28. Mamoor, S., 2020. High mobility group 20B (HMG20B) is a differentially expressed gene in human metastatic breast cancer, in the brain and in the lymph nodes.

- 1 29. Mamoor, S., 2020. Semaphorin 3D is differentially expressed in the lymph node and brain
2 metastases of patients with metastatic breast cancer.
- 3 30. Mamoor, S., 2020. DLK1 is differentially expressed in the lymph node and brain metastases
4 of patients with metastatic breast cancer.
- 5 31. Mamoor, S., 2020. The connective-tissue growth factor is differentially expressed in
6 metastatic breast cancer, both in metastases to the brain and to the lymph nodes.
- 7 32. Mamoor, S., 2020. Traf4 is differentially expressed in the lymph node and brain metastases
8 of patients with metastatic breast cancer.
- 9 33. Mamoor, S., 2020. Glioma tumor suppressor candidate region gene 2 is differentially
10 expressed in the lymph node and brain metastases of patients with metastatic breast cancer.
- 11 34. Mamoor, S., 2020. Vascular endothelial growth factor D, VEGF-D, encoded by FIGF is
12 differentially expressed in metastatic breast cancer, both in metastases to the brain and to the
13 lymph nodes.
- 14
- 15
- 16
- 17
- 18
- 19
- 20
- 21
- 22
- 23
- 24
- 25
- 26
- 27
- 28

Rank: 21
Probe ID: 18107
p-value: 5.95e-05
t: 4.2430237
B: 1.5988
Gene: IRX2
Gene name: iroquois homeobox 2

Chart 1: IRX2 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: 5517
Probe ID: A_23_P156025
p-value: 3.52e-04
t: 3.8537341
B: -0.243284
Gene: IRX2
Gene name: iroquois homeobox 2

Chart 2: IRX2 is differentially expressed in metastatic breast cancer when comparing brain 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: IRX2 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 IRX2 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 IRX2 expression between primary tumors of the breast and lymph node metastases, a p -value, is $p=0.000081$.

Figure 2: Significant correlation between IRX2 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 $n=1764$ total patients stratified into two groups, based on low or high expression of IRX2 in patient primary tumors. 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 IRX2 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 IRX2 expression: 34 months

High IRX2 expression: 39 months

Chart 3: Median recurrence-free survival is inferior in patients with low primary tumor expression of IRX2.

The median RFS (recurrence-free survival) of patients with low primary tumor expression of IRX2 and high primary tumor expression of IRX2 is listed in this chart, for $n=1764$ breast cancer patients.

Rank: 742
Probe ID: A_23_P156025
p-value: 2.19e-12
t: -1.42e+01
B: 18.50422711
Gene: IRX2
Gene name: iroquois homeobox 2

Chart 3: IRX2 is among the genes whose expression changes most significantly transcriptome-wide when comparing normal breast to normal brain.

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: 228
Probe ID: A_23_P110837
p-value: 2.16e-15
t: -2.01e+01
B: 25.38450681
Gene: IRX4
Gene name: iroquois homeobox 4

Chart 4: IRX4 is among the genes whose expression changes most significantly transcriptome-wide when comparing normal breast to normal brain.

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: 214
Probe ID: A_24_P48057
p-value: 1.46e-15
t: -2.05e+01
B: 25.76696772
Gene: IRX5
Gene name: iroquois homeobox 5

Chart 5: IRX5 is among the genes whose expression changes most significantly transcriptome-wide when comparing normal breast to normal brain.

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 3: IRX4 is expressed at significantly higher levels in the breast as compared to the brain.

The mRNA expression level of IRX4 in normal breast (left) and in normal brain (right) is graphically depicted; the result of a statistical test evaluating significance of difference in IRX4 expression between the breast and the brain, a p -value, is $p < 0.00001$.

IRX5

Figure 4: IRX5 is expressed at significantly higher levels in the breast as compared to the brain.

The mRNA expression level of IRX5 in normal breast (left) and in normal brain (right) is graphically depicted; the result of a statistical test evaluating significance of difference in IRX5 expression between the breast and the brain a p -value, is $p < 0.00001$.