The lipoma HMGIC fusion partner, LHFP, is differentially expressed in primary tumors and lymph node metastases in human breast cancer.

Shahan Mamoor, MS¹

¹shahanmamoor@gmail.com

East Islip, NY USA 11730

Breast cancer is the most frequently diagnosed cancer in women¹. Metastasis is the major cause of death in people diagnosed with cancer². Prognosis for patients with breast cancer worsens as the number of axillary lymph nodes with metastasis increases³. We mined a published microarray dataset⁴ to discover genes associated with metastasis to the lymph nodes, an early event in breast cancer metastasis. We found significant differential expression of the LHFP gene when comparing primary breast tumors to lymph node metastases from women diagnosed with breast cancer. Analysis of a separate microarray dataset⁵ revealed that LHFP was among the genes most differentially expressed when comparing primary tumors of the breast to adjacent normal breast tissue. LHFP expression significantly correlated with median overall survival in patients with breast cancer. LHFP may be of relevance to processes underlying transformation or progression of the primary tumor in human breast cancer, as well as metastasis to the lymph nodes.

Keyword: breast cancer, breast cancer metastasis, lymph node metastasis, LHFP, systems biology of breast cancer, targeted therapeutics in breast cancer.

1 Invasive breast cancer is diagnosed in over quarter of a million women in the United
2 States each year¹. Metastasis is the most common cause of death in humans diagnosed with
3 cancer². The lymphatic system is the primary route by which tumor cells traverse the body⁶.
4 While patients with localized breast cancer are provided a 99% 5-year survival rate, patients with
5 regional breast cancer, cancer that has spread to lymph nodes or nearby structures, are provided
6 an 86% 5-year survival rate^{7,8}. Patients with metastasis to distant sites are provided a 27% 5-
7 year survival rate^{7,8}. We mined published microarray gene expression datasets^{4,5} to discover
8 genes associated with the primary tumor and lymph node metastasis in breast cancer and
9 identified significant differential expression of the lipoma HMGIC fusion partner gene, LHFP.
10 LHFP may be of relevance to breast cancer progression or metastasis to the lymph nodes.
11
12
13
14

15 **Methods**

16
17 We used datasets GSE124648⁴ and GSE109169⁵ for this differential gene expression
18 analysis in conjunction with GEO2R. GSE124648 was generated using Affymetrix Human
19 Genome U133A Array technology with $n=256$ primary breast tumors from patients with breast
20 cancer and $n=67$ lymph node metastases from patients with breast cancer; platform GPL96 was
21 used for this analysis. GSE109169 was generated using Affymetrix Human Exon 1.0 ST Array
22 technology with $n=25$ adjacent control breast tissue and $n=25$ primary breast tumors from
23 patients with breast cancer; platform GPL5175 was used for this analysis. The Benjamini and
24 Hochberg method of p -value adjustment was used for ranking of differential expression but raw
25 p -values were used to assess statistical significance of global differential expression. Log-
26
27
28

transformation of data was auto-detected, and the NCBI generated category of platform annotation was used. A statistical test was performed to evaluate whether LHFP expression was significantly between primary breast tumors and lymph node metastases, or between adjacent control tissue and primary tumors of the breast using a two-tailed, unpaired t-test with Welch's correction. We used PRISM for all statistical analyses of differential gene expression in human breast cancer (Version 8.4.0)(455). For Kaplan-Meier survival analysis, we used the Kaplan-Meier plotter online tool⁹ for correlation of LHFP mRNA expression levels with overall survival in $n=1402$ breast cancer patients.

Results

We mined published microarray data^{4,5} to discover genes associated with progression of breast cancer: between normal breast tissue and primary tumors of the breast, and between primary breast tumors and metastasis of invasive breast cancer to the lymph nodes.

LHFP is differentially expressed in lymph node metastases in invasive breast cancer.

We identified LHFP as among the genes expressed most quantitatively differently, transcriptome-wide, when comparing primary tumors of the breast to lymph node metastases in women diagnosed with breast cancer⁴ (Table 1). When sorting each of the transcripts measured by microarray based on significance of change in expression between lymph node metastases and primary tumors of the breast, LHFP ranked 68 out of 22283 total transcripts (Table 1).

Differential expression of LHFP in lymph node metastases in human breast cancer was statistically significant (Table 1; $p=1.35E-56$).

LHFP is differentially expressed in primary tumors of the breast in invasive breast cancer.

Analysis of a second microarray dataset⁵ revealed that LHFP was also among the genes most differentially expressed in primary tumors of the breast from women diagnosed with invasive breast cancer when compared to adjacent normal breast tissue (Table 2). When sorting each of the transcripts measured by microarray based on significance of change in expression between primary tumors of the breast and adjacent, benign breast tissue, LHFP ranked 110 out of 19076 transcripts measured (Table 2). Differential expression of LHFP in primary tumors of the breast in human breast cancer was statistically significant (Table 2; $p=1.69\text{E-}14$).

LHFP is expressed at significantly lower levels in lymph node metastases when compared to primary tumors of the breast.

We obtained exact mRNA expression levels of the differentially expressed LHFP transcript to directly compare LHFP expression between primary tumors of the breast and lymph node metastases. We also performed a statistical test to evaluate significance of difference in LHFP expression between primary tumors of the breast and lymph node metastases was statistically significant. LHFP was expressed at lower levels in lymph node metastases as compared to primary tumors of the breast; this difference in LHFP mRNA expression was statistically significant (Figure 1; $p<0.0001$). We calculated a fold change of 0.9703 ± 0.3305 in LHFP mRNA expression in lymph node metastases relative to primary breast tumors (Table 1).

LHFP is expressed at significantly lower levels in primary tumors of the breast when compared to adjacent benign breast tissue.

1 We also obtained exact mRNA expression levels of the differentially expressed LHFP
2 transcript to directly compare LHFP expression between primary tumors of the breast and
3 adjacent, benign breast tissue. We also performed a statistical test to evaluate significance of
4 difference in LHFP expression between breast tumors and benign breast tissue in women with
5 invasive breast cancer. LHFP was expressed at lower levels in primary breast tumors; this
6 difference in mRNA expression of LHFP was statistically significant (Figure 2; $p < 0.0001$). We
7 calculated a fold change of 0.85 ± 0.06 in LHFP mRNA expression in primary tumors of the
8 breast relative to adjacent benign breast tissue (Table 2).

13 LHFP expression in primary breast tumors is significantly correlated with median overall
14 survival.

15 We performed Kaplan-Meier survival analysis⁹ using LHFP mRNA expression in primary
16 breast tumors coupled with paired overall survival data from each patient, in 1402 patients in
17 total, to determine whether LHFP tumor expression was correlated with survival outcomes in
18 breast cancer. We found that patients whose tumors expressed lower levels of LHFP possessed
19 significantly shorter overall survival than patients with high tumor expression of LHFP (Figure
20 4). Median overall survival (OS) of patients in the low expression cohort was 37 months, while
21 median OS in patients in the high LHFP expression cohort was 69.33 months (Table 4); this
22 difference in median OS based on LHFP tumor expression in breast cancer was statistically
23 significant (Figure 4; logrank p -value: $1.3e-11$; hazard ratio: 0.68 (0.61-0.76)).
24
25
26
27
28

Discussion

LHFP was identified following analysis of a t(12;13) translocation in lipoma, whose subsequent fusion with the HMGIC gene resulted in production of an HMGIC/LHFP fusion transcript¹¹. LHFP is expressed at relatively high levels in bone and in osteoblasts¹². Bone marrow stromal cells from LHFP-deficient mice manifest higher capacity to differentiate into the osteoblast lineage as evidenced by fibroblast colony formation potential (CFU-U), and increased mineralization as assessed by alizarin red staining *in vitro*, and bone mineral density in these mice is increased as a result of changes in cortical thickness¹². A single nucleotide polymorphism in LHFP, rs9576787, was associated with bone mineral density in the heel¹². Together, investigators concluded that LHFP functions in regulation of osteoblast differentiation and bone mineral density¹².

In gliosarcoma, a study of glial and mesenchymal regions of 13 tumors revealed changes in copy number at 13q13.3-q14.1, a region that encompassed the LHFP gene; these copy number changes were restricted to mesenchymal regions of the tumor¹³. mRNA expression analyses in a larger set of gliosarcomas revealed that LHFP was amplified in 11% (7/64) of gliosarcomas in the mesenchymal portions¹³. A separate study indicated that fusion of LHFP to HMGIC was not an inherent feature of mesenchymal tumors with changes in chromosome structure at 12q13, as HMGIC-LHFP fusion transcripts were absent in pulmonary chondroid hamartomas containing 12q13~q15 and t(12;13)(q14~15;q13) changes¹⁴. Thus, LHFP is implicated in regulation of bone mineral density, in the differentiation potential of osteoblasts, the LHFP locus is fused to the

1 HMGIC gene in human lipoma, producing a fusion transcript, and gliosarcomas can bear copy
2 number changes in the LHFP gene, specifically in mesenchymal portions of tumor.
3

4 By mining independently published gene expression data^{4,5}, we found that the gene
5 encoding the lipoma HMGIC fusion partner, LHFP, was among the genes most differentially
6 expressed in human breast cancer, both in lymph node metastases when compared to primary
7 tumors of the breast as well as in primary tumors of the breast as compared to adjacent benign
8 breast tissue. LHFP tumor expression was significantly correlated with overall patient survival.
9 LHFP appears to be expressed in a gradient-like fashion in invasive breast cancer, and may be of
10 relevance to processes underlying progression of breast cancer from transformation to spread of
11 breast cancer to the lymph nodes.
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28

References

1. DeSantis, C.E., Ma, J., Gaudet, M.M., Newman, L.A., Miller, K.D., Goding Sauer, A., Jemal, A. and Siegel, R.L., 2019. Breast cancer statistics, 2019. *CA: a cancer journal for clinicians*, 69(6), pp.438-451.
2. Mehlen, P. and Puisieux, A., 2006. Metastasis: a question of life or death. *Nature reviews cancer*, 6(6), pp.449-458.
3. Fisher, B., Bauer, M., Wickerham, D.L., Redmond, C.K., Fisher, E.R., Cruz, A.B., Foster, R., Gardner, B., Lerner, H., Margoese, R. and Poisson, R., 1983. Relation of number of positive axillary nodes to the prognosis of patients with primary breast cancer. An NSABP update. *Cancer*, 52(9), pp.1551-1557.
4. 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.
5. Chang, J.W., Kuo, W.H., Lin, C.M., Chen, W.L., Chan, S.H., Chiu, M.F., Chang, I.S., Jiang, S.S., Tsai, F.Y., Chen, C.H. and Huang, P.H., 2018. Wild-type p53 upregulates an early onset breast cancer-associated gene GAS7 to suppress metastasis via GAS7–CYFIP1-mediated signaling pathway. *Oncogene*, 37(30), pp.4137-4150.
6. Rahman, M. and Mohammed, S., 2015. Breast cancer metastasis and the lymphatic system. *Oncology letters*, 10(3), pp.1233-1239.
7. ACS Cancer Facts & Figures 2019. <https://www.cancer.net/cancer-types/breast-cancer-metastatic/statistics>.
8. Survival Rates for Breast Cancer. <https://www.cancer.org/cancer/breast-cancer/understanding-a-breast-cancer-diagnosis/breast-cancer-survival-rates.html>
9. 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.

10. Petit, M.M., Schoenmakers, E.F., Huysmans, C., Geurts, J.M., Mandahl, N. and Van de Ven, W.J., 1999. LHFP, a Novel Translocation Partner Gene of HMGIC in a Lipoma, Is a Member of a New Family of LHFP-like Genes. *Genomics*, 57(3), pp.438-441.
11. Mesner, L.D., Calabrese, G.M., Al-Barghouthi, B., Gatti, D.M., Sundberg, J.P., Churchill, G.A., Godfrey, D.A., Ackert-Bicknell, C.L. and Farber, C.R., 2019. Mouse genome-wide association and systems genetics identifies Lhfp as a regulator of bone mass. *PLoS genetics*, 15(5), p.e1008123.
12. Nagaishi, M., Kim, Y.H., Mittelbronn, M., Giangaspero, F., Paulus, W., Brokinkel, B., Vital, A., Tanaka, Y., Nakazato, Y., Legras-Lachuer, C. and Lachuer, J., 2012. Amplification of the STOML3, FREM2, and LHFP genes is associated with mesenchymal differentiation in gliosarcoma. *The American journal of pathology*, 180(5), pp.1816-1823.
13. Rogalla, P., Lemke, I. and Bullerdiek, J., 2002. Absence of HMGIC-LHFP fusion in pulmonary chondroid hamartomas with aberrations involving chromosomal regions 12q13~q15 and 13q12~q14. *Cancer genetics and cytogenetics*, 133(1), pp.90-93.

Rank	ID	p-value	t	B	FC	Gene	Gene name
68	218656_s_at	1.35E-56	-19.504273	118.15722	0.6564 ± 0.2543	LHFP	lipoma HMGIC fusion partner

Table 1: Differential expression of LHFP in lymph node metastases in breast cancer.

Rank of differential expression, p-value of global differential expression, t, a moderated t-statistic, B, the log-odds of differential expression between the groups compared, fold change of LHFP in lymph node metastases as compared to primary breast tumors, and gene are listed in this chart.

Rank	ID	p-value	t	B	FC	Gene
99	59	8.57E-01	-0.1811	-6.9139	0.9703 ± 0.3305	LHFP

Table 2: Differential expression of LHFP in primary tumors of the breast in breast cancer.

Rank of differential expression, p-value of global differential expression, t, a moderated t-statistic, B, the log-odds of differential expression between the groups compared, fold change of LHFP in primary tumors of the breast as compared to adjacent benign breast tissue, and gene are listed in this chart.

Figure 1: LHFP is expressed at significantly lower levels in lymph node metastases as compared to primary tumors of the breast.

The mRNA expression level of LHFP in primary tumors of the breast from women diagnosed with invasive breast cancer (left) and in lymph node metastases from women diagnosed with breast cancer (right) is graphically depicted with the result of a statistical test evaluating significance of difference in LHFP expression between primary tumors and lymph node metastases, a *p*-value, listed above.

Figure 2: LHFP is expressed at significantly lower levels in primary tumors of the breast as compared to benign adjacent breast tissue.

The mRNA expression level of LHFP in benign, tumor-adjacent breast tissue (left) and in primary tumors of the breast (right) is graphically depicted with the result of a statistical test evaluating significance of difference in LHFP expression between normal breast tissue and primary tumors of the breast, a *p*-value, listed above.

Figure 3: LHFP expression in breast tumors is significantly correlated with overall survival in patients with breast cancer.

Depicted in this Kaplan-Meier plot is the probability of overall survival for $n=1402$ total patients stratified into two groups, based on low or high expression of LHFP 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 LHFP 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 expression cohort (months)	High expression cohort (months)
37	69.33

Table 3: LHFP expression in breast tumors significantly correlates with patient survival.

The median overall survival of $n=1402$ breast cancer patients based on stratification into low or high tumor expression of LHFP is listed in this chart.