The immunoglobulin light chain kappa locus is differentially expressed in the primary tumors of patients with breast cancer and this is an intrinsic transcriptional feature of the primary tumor.

Shahan Mamoor, MS¹

¹shahanmamoor@gmail.com

East Islip, NY USA

We recently described differential expression and down-regulation of the immunoglobulin light chain kappa constant locus (IGKC) in brain metastases of patients with breast cancer¹. Immunoglobulin gene expression is generally thought of as restricted to cells of lymphocytic lineage². Expression of IGKC has previously been reported in primary tumors of the breast in patients with breast cancer, but ascribed to the presence of tumor-infiltrating lymphocytes³⁻⁵. We analyzed tumor gene expression at the transcriptome level using independent microarray datasets^{6,7} here to determine in an unbiased fashion whether IGKC expression was differential in the primary tumors of patients with breast cancer and whether this could be ascribed to the presence of lymphocytes in tumor tissues studied. We provide evidence here that IGKC is differentially expressed in the primary tumors of patients with breast cancer and that this is a genetic event intrinsic to the primary tumor, not the result of the presence of lymphocytes.

Keyword: breast cancer, systems biology of breast cancer, IGK, IGKC, immunoglobulin gene expression, immunoglobulin gene expression in human cancer, targeted therapeutics in breast cancer.

1 It has been suggested that the source of immunoglobulin gene expression observed in primary
2 tumors of the breast was attributable to tumor-infiltrating lymphocytes⁴, but we recently discovered that
3 among the genes whose expression was most different between primary tumors of the breast and
4 metastases to the brain was an immunoglobulin gene segment¹, with substantial down-regulation of IGKC
5 expression in metastases to the brain as compared to primary tumors of the breast, suggesting IGK
6 expression was an inherent feature of human breast cancer. We mined published microarray datasets^{6,7} to
7 determine in an unbiased fashion and at the transcriptome level whether IGKC expression observed in
8 human breast cancer was tumor-intrinsic or a byproduct of lymphocyte contamination/infiltration.

9 **Methods**

10 We used datasets GSE124648⁶ and GSE42568⁷ for this comparative transcriptome analysis of
11 differentially expressed genes in human breast cancer. GSE42568 was generated using Affymetrix
12 Human Genome U133 Plus 2.0 Array technology with $n=17$ normal breast tissues and $n=104$ primary
13 tumors of the breast; analysis was performed using platform GPL570. GSE124648 was generated using
14 Affymetrix Human Genome U133A Array technology with $n=10$ normal breast tissue and $n=126$ tumors
15 of the breast and; analysis was performed using platform GPL96. The Benjamini and Hochberg method
16 of p -value adjustment was used for ranking of differential expression but raw p -values were used to assess
17 statistical significance of global differential expression. Log-transformation of data was auto-detected,
18 and the NCBI generated category of platform annotation was used.

19 **Results**

20 We performed comparative transcriptome analysis of primary tumors of the breast and normal
21 breast tissue using published microarray data^{6,7} to investigate immunoglobulin gene expression in human
22 breast cancer and ascertain the source of this immunological gene expression.

23 **The gene encoding the immunoglobulin light chain kappa locus is differentially expressed in the 24 primary tumors of patients with breast cancer.**

25 We identified the gene encoding the immunoglobulin light chain kappa constant region, IGKC, as
26 among the genes whose expression was most quantitatively different at the messenger RNA (mRNA)
27 level, transcriptome-wide, when comparing primary tumors of the breast to normal breast tissues in
28 humans (Table 1). When sorting each of the genes expressed in primary tumors of the breast based on
significance of change in expression as compared to normal breast tissue, IGKC ranked 2088 out of
22283 total transcripts (Table 1), equating to 90.6% differential expression. Differential expression of
IGKC in primary tumors of the breast was statistically significant (Table 1; $p=4.52E-05$).

We queried a second microarray dataset to ascertain whether differential expression of IGKC at
the transcriptome level could be validated using an independent source of data. Here we found that IGK,
the immunoglobulin light chain kappa complex, was detected as among the genes whose expression was
most quantitatively different transcriptome-wide when comparing primary tumors of the breast to normal
breast tissues in humans (Table 2). When sorting each of the genes expressed in primary tumors of the
breast based on significance of change in expression as compared to normal breast tissue, IGK ranked
9384 out of 54675 total transcripts (Table 2), equating to 82.8% differential expression. Differential
expression of IGK in primary tumors of the breast was statistically significant (Table 2; $p=2.26E-04$).

Differential expression of IGKC and IGK are inherent genetic features of the primary tumor, as CD4 and CD19 are not differentially expressed when comparing primary tumor tissues of the breast to normal breast tissues

Broadly speaking, transformation and metastasis follow a linear progression, from normal breast tissue to primary tumor and then to metastasis. Since we found that IGKC was among the genes most differentially expressed in metastatic breast cancer between brain metastatic tissues and the tissues from which metastases arise, primary tumors of the breast, we then assessed whether expression of IGKC that we observed in primary tumors was significantly different from that of the tissue from which it arose, the normal breast. Indeed, we found that this was the case. However, for multiple reasons, expression of immunoglobulins is thought to be restricted to hematopoietic cells of the lymphocytic lineage, and more specifically, to B-lymphocytes. We then asked, could we attribute expression of IGKC to tumor-infiltrating B-lymphocytes associated with anti-humoral response since we studied bona fide whole patient tumor tissues from humans with breast cancer. In fact, we could not: CD19, a marker of the B-cell, was detected by microarray analysis in both transcriptome-wide datasets we analyzed, but expression of CD19 was not significantly different when comparing primary tumors of the breast and normal breast tissues (Table 3 and Table 4). When sorting each of the genes expressed in primary tumors of the breast based on significance of difference as compared to normal breast tissues, CD19 ranked 13533 out of 22283 total transcripts (Table 3), equating to 39.3% differential expression, and in a second dataset, CD19 ranked 49559 out of 54675 total transcripts, equating to 9.4% differential expression (Table 4). Expression of CD19 in primary tumors of the breast was not significantly different, across two independent microarray datasets (Table 3: $p=2.33E-01$ and Table 4: $p=8.13E-01$). CD20, another B-cell marker, was not detected by either microarray dataset we utilized.

Though immunoglobulins are expressed by B-lymphocytes and not T-lymphocytes, immunoglobulin rearrangement events at the genetic level are most similar to that of expression and rearrangement of the T-cell receptor (TCR) locus and major histocompatibility (MHC) locus, found in T-lymphocytes. T-lymphocytes can broadly be classified into two major lineages: the cytotoxic CD8+ T-lymphocyte, and the CD4+ helper T-lymphocyte, the cell type most responsible for providing assistance and instruction to B-lymphocytes in the production of immunoglobulins. While we did not observe detectable expression of CD8 in either microarray dataset we analyzed, we did observe expression of the T-lymphocyte marker CD4, but its expression was not significantly different between primary tumors of the breast and normal breast tissues (Table 5 and Table 6). When sorting each of the genes expressed in primary tumors of the breast based on significance of difference as compared to normal breast tissues, CD4 ranked 14779 out of 22283 total transcripts (Table 5), equating to 33.7% differential expression, and in a second dataset, CD4 ranked 31496 out of 54675 total transcripts (Table 6), equating to 42.4% differential expression. Expression of CD4 in primary tumors of the breast was not significantly different, across two independent microarray datasets (Table 5: $p=3.08E-01$ and Table 6: $p=2.5E-01$).

Thus, we discovered transcriptome-level differential expression of IGK and IGKC in human breast cancer, and conclude that the differential expression of immunoglobulin light chain is a characteristic transcriptional feature of primary tumors of the breast, and not due to the presence of infiltrating lymphocytes. IGK differential expression in breast cancer is a genetic event intrinsic to the primary tumor.

Discussion

We provided evidence in this letter that the immunoglobulin light chain kappa gene segment is differentially expressed in the primary tumors of humans with breast cancer. We recently discovered that

1 among the genes whose expression, or lack thereof, most significantly defines the brain metastases of
2 patients with metastatic breast cancer was IGKC, the immunoglobulin light chain kappa constant gene
3 segment¹. We are now investigating whether other immunoglobulin gene segments and genes associated
4 with immunoglobulin gene expression in B-lymphocytes are differentially expressed (transcriptionally
5 activated) in primary tumors in human breast cancer, and plan to evaluate through assessment of the
6 tumor transcriptome at the systems-level what agents - infectious or chemically-based (e.g., olfactory
7 receptor ligand) - trigger expression of an immunoglobulin gene in the breast.
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28

References

1. Mamoor, S., 2020. IGKC, the immunoglobulin kappa constant gene segment, is differentially expressed in brain metastatic breast cancer.
2. Chen, Z., Qiu, X. and Gu, J., 2009. Immunoglobulin expression in non-lymphoid lineage and neoplastic cells. *The American journal of pathology*, 174(4), pp.1139-1148.
3. Schmidt, M., Hellwig, B., Hammad, S., Othman, A., Lohr, M., Chen, Z., Boehm, D., Gebhard, S., Petry, I., Lebrecht, A. and Cadenas, C., 2012. A comprehensive analysis of human gene expression profiles identifies stromal immunoglobulin κ C as a compatible prognostic marker in human solid tumors. *Clinical Cancer Research*, 18(9), pp.2695-2703.
4. Pandey, J.P., Kistner-Griffin, E., Black, L., Namboodiri, A.M., Iwasaki, M., Kasuga, Y., Hamada, G.S. and Tsugane, S., 2014. IGKC and Fc γ R genotypes and humoral immunity to HER2 in breast cancer. *Immunobiology*, 219(2), pp.113-117.
5. Chen, Z., Gerhold-Ay, A., Gebhard, S., Boehm, D., Solbach, C., Lebrecht, A., Battista, M., Sicking, I., Cotarelo, C., Cadenas, C. and Marchan, R., 2012. Immunoglobulin kappa C predicts overall survival in node-negative breast cancer. *PLoS One*, 7(9), p.e44741.
6. 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.
7. Clarke, C., Madden, S.F., Doolan, P., Aherne, S.T., Joyce, H., O'driscoll, L., Gallagher, W.M., Hennessy, B.T., Moriarty, M., Crown, J. and Kennedy, S., 2013. Correlating transcriptional networks to breast cancer survival: a large-scale coexpression analysis. *Carcinogenesis*, 34(10), pp.2300-2308.
8. Komatsu, M., Yoshimaru, T., Matsuo, T., Kiyotani, K., Miyoshi, Y., Tanahashi, T., Rokutan, K., Yamaguchi, R., Saito, A., Imoto, S. and Miyano, S., 2013. Molecular features of triple negative breast cancer cells by genome-wide gene expression profiling analysis. *International journal of oncology*, 42(2), pp.478-506.

Rank	ID	p-value	t	B	Gene	Gene name
2088	214669_x_at	4.52E-05	-4.216145	1.59108	IGKC	immunoglobulin kappa constant

Table 1: IGKC is differentially expressed in primary tumors of the breast in human breast cancer.

The rank of differential expression, probe ID, *p*-value with respect to differential expression, *t*, a moderated *t*-statistic, *B*, the log-odds of differential expression between the two groups compared, gene and gene name are listed in this chart.

Rank	ID	p-value	t	B	Gene	Gene name
9384	215906_at	2.26E-04	-3.80	-0.3542899	Igk	immunoglobulin kappa chain complex

Table 2: IGK is differentially expressed in primary tumors in human breast cancer.

The rank of differential expression, probe ID, *p*-value with respect to differential expression, *t*, a moderated *t*-statistic, *B*, the log-odds of differential expression between the two groups compared, gene and gene name are listed in this chart.

Rank	ID	p-value	t	B	Gene	Gene name
13533	206398_s_at	2.33E-01	1.1981693	-5.97476	CD19	CD19 molecule

Table 3: CD19 is not differentially expressed in primary tumors in human breast cancer.

The rank of differential expression, probe ID, *p*-value with respect to differential expression, *t*, a moderated *t*-statistic, *B*, the log-odds of differential expression between the two groups compared, gene and gene name are listed in this chart.

Rank	ID	p-value	t	B	Gene	Gene name
49559	206398_s_at	8.13E-01	2.37E-01	-7.1714324	CD19	CD19 molecule

Table 4: CD19 is not differentially expressed in primary tumors in human breast cancer.

The rank of differential expression, probe ID, *p*-value with respect to differential expression, *t*, a moderated *t*-statistic, *B*, the log-odds of differential expression between the two groups compared, gene and gene name are listed in this chart.

Rank	ID	p-value	t	B	Gene	Gene name
14779	203547_at	3.08E-01	-1.0242372	-6.1647	CD4	CD4 molecule

Table 5: CD4 is not differentially expressed in primary tumors in human breast cancer.

The rank of differential expression, probe ID, *p*-value with respect to differential expression, *t*, a moderated *t*-statistic, *B*, the log-odds of differential expression between the two groups compared, gene and gene name are listed in this chart.

Rank	ID	p-value	t	B	Gene	Gene name
31496	203547_at	2.5E-01	-1.15	-6.5345511	CD4	CD4 molecule

Table 6: CD4 is not differentially expressed in primary tumors in human breast cancer.

The rank of differential expression, probe ID, p -value with respect to differential expression, t , a moderated t -statistic, B , the log-odds of differential expression between the two groups compared, gene and gene name are listed in this chart.