APOBEC3D expression associates with survival in triple negative breast cancer.

Shahan Mamoor, MS¹

¹shahanmamoor@gmail.com

East Islip, NY USA

We mined published microarray data (1) to understand the most significant gene expression differences in the tumors of triple negative breast cancer patients based on survival at time of analysis: dead or alive. The apolipoprotein B mRNA editing enzyme catalytic subunit 3D *APOBEC3D* emerged as among the most significant differences, transcriptome-wide, when comparing the primary tumors of triple negative breast cancer patients with disparate survival outcomes. Importantly, APOBEC3D expression was significantly correlated with overall survival in basal subtype breast cancer, a molecular subtype sharing significant overlap with triple negative breast cancer. APOBEC3D, a lentiviral restriction factor (2, 3), may be of relevance as a biomarker or as a molecule of interest in understanding the etiology or progression of triple negative breast cancer.

Keywords: triple negative breast cancer, TNBC, systems biology of breast cancer, targeted therapeutics in breast cancer, APOBEC3D, apolipoprotein B mRNA editing enzyme catalytic subunit 3D.

Triple negative breast cancer is defined by lack of expression of the receptors for the hormones progesterone and estrogen (PR negative and ER negative), as well as for the human epidermal growth factor receptor 2 (HER2 negative) (4, 5). Though not completely identical, triple negative breast cancer shares significant overlap with the basal or basal-like molecular subtype of human breast cancer (5). TNBC is more frequently diagnosed in women of African descent (black women) for reasons not understood, and their disease course is similarly more aggressive than in white women with triple negative breast cancer (4, 5). To understand the transcriptional biology of triple negative breast cancer in an unbiased fashion and at the systems level, we mined published microarray data paired with patient survival data (1), performing comparative transcriptome analysis based on survival at time of data collection: dead or alive. We present here a finding from the preliminary results of this analysis.

Methods

We used dataset GSE142102 (1) for this differential gene expression analysis in conjunction with GEO2R. GSE142102 was generated using Affymetrix Human Gene 2.1 ST Array technology; in this analysis, we used data from $n=87$ tumors from TNBC patients that were dead at time of analysis and $n=123$ tumors from TNBC patients that were alive at time of analysis. 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 APOBEC3D expression was significantly different between the primary tumors of TNBC patients that were dead or alive using a two-tailed t -test. For Kaplan-Meier survival analysis, we used the Kaplan-Meier plotter tool (6) for correlation of APOBEC3D mRNA expression levels with overall survival (OS) in $n=417$ patients with basal subtype breast cancer and with recurrence-free survival in $n=278$ patients with basal subtype breast cancer.

Results

We performed comparative transcriptome analysis using published microarray data (1) to understand in an unbiased fashion the most significant gene expression differences in the primary tumors of patients with triple negative breast cancer based on survival outcomes.

APOBEC3D is differentially expressed in triple negative breast cancer and positively correlates with survival.

When comparing primary tumors of 87 patients with triple negative breast cancer, alive at time of data collection and 123 TNBC patients, dead at time of data collection, we identified the apolipoprotein B mRNA editing enzyme catalytic subunit 3D, *APOBEC3D*, as among the most significant transcriptional differences based on survival in human triple negative breast cancer. When ranking each of the transcripts whose expression was measured by microarray based on significance of difference in expression between the tumors of patients dead and alive, APOBEC3D ranked 94 out of 53617 total transcripts, equivalent to 99.8% differential expression (Chart 1). APOBEC3D differential expression in the tumors of patients based on survival outcomes was statistically significant (Figure 1; $p=0.00236086$).

APOBEC3D is expressed at significantly higher levels in the breast tumors of TNBC patients when comparing based on survival.

We obtained exact mRNA expression levels for APOBEC3D from breast tumors of TNBC patients dead and alive to understand the magnitude and direction of APOBEC3D expression change based on survival outcomes. APOBEC3D was expressed at higher levels in tumors of TNBC patients alive as compared to that of TNBC patients dead (Figure 1). Increased expression of APOBEC3D in primary breast tumors was statistically significant (Figure 1: $p=0.0019$). APOBEC3D was expressed at -0.21 ± 0.89 arbitrary units (AU) in the primary tumors of TNBC patients alive, while it was expressed at 0.21 ± 1.05 AU in the primary tumors of TNBC patients dead.

APOBEC3D expression correlates with overall survival in basal subtype human breast cancer.

We referenced a bioinformatics tool containing Kaplan-Meier survival data of human breast cancer patients to understand more thoroughly whether a relationship existed between APOBEC3D primary tumor expression and patient survival in human breast cancer. We found that APOBEC3D expression was significantly correlated with overall survival in patients with basal or basal-like human breast cancer, a breast cancer molecular subtype which shares significant overlap with triple negative breast cancer (Figure 2; log rank p -value: 0.037 for overall survival, hazard ratio: 1.64 (1.03-2.62) (Fig. 2)). APOBEC3D mRNA levels were a negative prognostic indicator in basal subtype breast cancer patients. Median OS was 77.4 months for basal subtype patients with low tumor expression of APOBEC3D as compared to 40.77 months for basal subtype patients with high tumor expression of APOBEC3D. APOBEC3D expression was also significantly correlated with recurrence-free survival in patients with basal subtype human breast cancer (Figure 2; log rank p -value: 0.0029 for overall survival, hazard ratio: 0.63 (0.46-0.85) (Fig. 2)). With regard to recurrence-free survival (RFS), and in contrast to overall survival, APOBEC3D mRNA levels were a positive prognostic indicator in basal subtype breast cancer patients. Median RFS was 16 months for basal subtype patients with low tumor expression of APOBEC3D as compared to 27 months for basal subtype patients with high tumor expression of APOBEC3D.

Thus, we found through unbiased comparative transcriptome analysis that APOBEC3D was among the genes whose expression was most significantly different in the primary tumors of triple negative breast cancer patients alive when compared to that of TNBC patients dead, that APOBEC3D was expressed at significantly higher levels in the primary tumors of TNBC patients alive as compared to TNBC patients dead, and that APOBEC3D expression was correlated with overall survival and recurrence-free survival in the basal or basal-like molecular subtype, but in a contrary manner.

Discussion

It is not understood why women of African descent (black women) are diagnosed with triple negative breast cancer more frequently than white women, and why this disease is more aggressive in black women. APOBEC3D may be of relevance as a biomarker in identifying patients who are more likely to survive, perhaps as part of a panel of genes used for prognostic stratification, and the molecule itself may be important to the biology of triple negative breast cancer etiology or more likely its progression. APOBEC3D is a lentiviral restriction factor (2, 3). We previously demonstrated that APOBEC3D is among the genes most differentially expressed in the lymph node metastases of patients with metastatic breast cancer, with significantly increased expression in metastasis to the lymph nodes as compared to primary tumors of the breast (7). While transcriptome-wide differential expression revealed

APOBEC3D expression levels as a positive prognostic indicator within the one cohort of patients with TNBC whose tumors were analyzed here, Kaplan-Meier survival analysis encompassing a larger set of patients with basal-like breast cancer demonstrated the existence of two patient populations based on APOBEC3D primary tumor expression: an APOBEC3D high-expressing population, with significantly inferior overall survival as compared to an APOBEC3D low-expressing population. We suggest that this indicates the existence of two patient populations in the basal-like or triple negative breast cancer population: one of which involves a lentiviral or retroviral infection (e.g., HIV-2, HTLV family, or SIV-like) as a causative agent or contributory factor, that tumors from this population are APOBEC3D high-expressing (and thus associated with inferior overall survival), but that within this population, higher expression of APOBEC3D (and putative antiviral restriction properties associated with this elevated expression) results in or is associated with superior survival outcomes.

References

1. Purrington, K.S., Knight III, J., Dyson, G., Ali-Fehmi, R., Schwartz, A.G., Boerner, J.L. and Bandyopadhyay, S., 2020. CLCA2 expression is associated with survival among African American women with triple negative breast cancer. *PloS one*, 15(4), p.e0231712.
2. Hultquist, J.F., Lengyel, J.A., Refsland, E.W., LaRue, R.S., Lackey, L., Brown, W.L. and Harris, R.S., 2011. Human and rhesus APOBEC3D, APOBEC3F, APOBEC3G, and APOBEC3H demonstrate a conserved capacity to restrict Vif-deficient HIV-1. *Journal of virology*, 85(21), p.11220
3. Sato, K., Takeuchi, J.S., Misawa, N., Izumi, T., Kobayashi, T., Kimura, Y., Iwami, S., Takaori-Kondo, A., Hu, W.S., Aihara, K. and Ito, M., 2014. APOBEC3D and APOBEC3F potently promote HIV-1 diversification and evolution in humanized mouse model. *PLoS Pathog*, 10(10), p.e1004453.
4. Foulkes, W.D., Smith, I.E. and Reis-Filho, J.S., 2010. Triple-negative breast cancer. *New England journal of medicine*, 363(20), pp.1938-1948.
5. Dietze, E.C., Sistrunk, C., Miranda-Carboni, G., O'regan, R. and Seewaldt, V.L., 2015. Triple-negative breast cancer in African-American women: disparities versus biology. *Nature Reviews Cancer*, 15(4), pp.248-254.
6. 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.
7. Mamoor, S., 2021. APOBEC3D is differentially expressed in the lymph nodes of patients with metastatic breast cancer.

Rank: 94
Probe ID: 16930173
p-value: 0.00236086
t: -3.0406273
B: -3.23
Gene: APOBEC3D
Gene name: apolipoprotein B mRNA editing enzyme catalytic subunit 3D

Chart 1: APOBEC3D is among the genes most differentially expressed transcriptome-wide when comparing the primary tumors of TNBC patients dead and alive.

Rank of differential expression, probe ID, *p*-value of global differential expression, *t*, a moderated *t*-statistic, *B*, the log-odds of differential expression between the groups compared, gene and gene name are listed in this chart.

Figure 1: APOBEC3D is expressed at significantly higher levels in the primary tumors of TNBC patients alive as compared to the primary tumors of TNBC patients dead.

The mRNA expression level of APOBEC3D in the primary tumors of TNBC patients dead (left) and in primary tumors of TNBC patients alive (right) is graphically depicted with the result of a statistical test evaluating significance of difference in APOBEC3D expression between the primary tumors of TNBC patients dead and alive, a p -value, is $p=0.0019$.

Figure 2: APOBEC3D expression correlates with overall survival in basal subtype human breast cancer.

Depicted in this Kaplan-Meier plot is the probability of overall survival (OS) for $n=417$ breast cancer patients of basal subtype, based on low or high expression of APOBEC3D 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 APOBEC3D 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 APOBEC3D expression: 77.4 months
High APOBEC3D expression: 40.77 months

Chart 2: Median overall survival is superior in basal subtype breast cancer patients with low primary tumor expression of APOBEC3D.

The median OS (overall survival) of $n=417$ basal subtype breast cancer patients, with low primary tumor expression of APOBEC3D and high primary tumor expression of APOBEC3D is listed in this chart.

Figure 3: APOBEC3D expression correlates with recurrence-free survival in basal subtype human breast cancer.

Depicted in this Kaplan-Meier plot is the probability of recurrence-free survival (RFS) for $n=278$ breast cancer patients of basal subtype, based on low or high expression of APOBEC3D 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 APOBEC3D 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 APOBEC3D expression: 16 months

High APOBEC3D expression: 27 months

Chart 3: Median recurrence-free survival is superior in basal subtype breast cancer patients with high primary tumor expression of APOBEC3D.

The median RFS (recurrence-free survival) of $n=278$ basal subtype breast cancer patients, with low primary tumor expression of APOBEC3D and high primary tumor expression of APOBEC3D is listed in this chart.