CA2 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 following treatment: dead or alive. We observed significant transcriptome-wide differential expression of carbonic anhydrase II, encoded by CA2 when comparing the primary tumors of triple negative breast cancer patients dead or alive. CA2 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, CA2, carbonic anhydrase II.

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) (2, 3). Though not completely identical, triple negative breast cancer shares significant overlap with the basal or basal-like molecular subtype of human breast cancer (3). 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 (2, 3). 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 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 CA2 expression was significantly different between the primary tumors of TNBC patients that were dead or alive using a two-tailed t-test.

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.

CA2 is differentially expressed in triple negative breast cancer and negatively correlates with survival.

When comparing primary tumors of 87 live patients with triple negative breast cancer with that of 123 dead TNBC patients we identified carbonic anhydrase II, *CA2*, 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, CA2 ranked 86 out of 53617 total transcripts, equivalent to 99.8% differential expression (Chart 1). CA2 differential expression in the tumors of patients based on survival outcomes was statistically significant (Figure 1; $p=0.00213876$).

CA2 is expressed at significantly lower levels in the breast tumors of TNBC patients when comparing based on survival.

We obtained exact mRNA expression levels for CA2 from breast tumors of TNBC patients dead and alive to understand the magnitude and direction of CA2 expression change based on survival outcomes. CA2 was expressed at lower levels in tumors of TNBC patients alive as compared to that of TNBC patients dead (Figure 1). Decreased expression of CA2 in primary breast tumors was statistically significant (Figure 1: $p=0.0054$). CA2 was expressed at 0.26 ± 1.25 arbitrary units (AU) in the primary

tumors of TNBC patients dead, while it was expressed at -0.17 ± 0.79 AU in the primary tumors of TNBC patients alive.

Thus, we found through unbiased comparative transcriptome analysis that CA2 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, and that CA2 was expressed at significantly lower levels in the primary tumors of TNBC patients alive as compared to TNBC patients dead.

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. CA2 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. Blind systems-level analyses, like this, can help shed light on the fundamental molecular characteristics of a disease whose racial propensities demand further explanation.

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. 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.
3. 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.

Rank: 86
Probe ID: 17070456
p-value: 0.00213876
t: 3.0702561
B: -3.2
Gene: CA2
Gene name: carbonic anhydrase II

Chart 1: CA2 is differentially expressed at the transcriptome-level 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: CA2 is expressed at significantly lower levels in the primary tumors of TNBC patients alive as compared to the primary tumors of TNBC patients dead.

The mRNA expression level of CA2 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 CA2 expression between the primary tumors of TNBC patients dead and alive, a *p*-value, listed here: $p=0.0054$.