# Aberrant *DNMT3B7* Expression Correlates to Age, Race, and Hormone Receptor Status in Breast Cancer Patients

*Christian C.S. Mullin, Jennifer J. Busser, Anna M. Ciampanelli, and Stacey L. Raimondi\**

*Department of Biology, Elmhurst College*

*Students: mullinc@net.elmhurst.edu, jennifer.j.busser@net.elmhurst.edu, ciampanellia@net.elmhurst.edu*
*Mentor: raimondis@elmhurst.edu\**

## ABSTRACT

Invasive breast cancer caused almost 40,000 deaths last year alone. Deciphering a way to better understand aggressive phenotypes of breast cancer could potentially provide a novel approach to increase the efficacy of breast cancer treatments. The aberrant gene *DNA methyltransferase 3B7, DNMT3B7*, has been observed in virtually all cancer types and has been shown to affect multiple facets of breast cancer progression including changes in cell adhesion, cellular proliferation, and anchorage-independent growth. A bioinformatics approach was taken to attempt to determine which clinical parameters are potentially altered by *DNMT3B7* expression in breast cancer patients. Clinical parameters including age at initial diagnosis, menopausal status, race, and the status of three hormone receptors — estrogen, progesterone, and human epidermal growth factor 2 (ER, PR, HER2, respectively) — were collected from The Cancer Genome Atlas (TCGA) and analyzed using T-tests and ANOVA. Our results show increased expression of the aberrant gene, *DNMT3B7,* is correlated with age at initial diagnosis, race, and hormone receptor status (ER negative, PR negative, and HER2 positive) in breast cancer patients, but not with menopausal status. Taken together, these data indicate that *DNMT3B7* expression may be an important marker in tumor progression.

## KEYWORDS

*DNMT3B7*, DNA Methylation; Bioinformatics; Breast Cancer; Tumor Progression; Estrogen Receptor; Progesterone Receptor: Human Epidermal Growth Factor Receptor 2

## INTRODUCTION

Cancer is the term used for a group of different diseases in which abnormal cells divide uncontrollably and may eventually progress to invade other bodily tissues.[1] The National Cancer Institute estimates that 39.6% of all men and women in the United States will develop cancer in their lifetime. Breast cancer is caused by abnormal cell growth in the tissues of the breast, specifically the glands (lobules) and ducts needed for milk production. This type of cancer caused almost 40,000 deaths last year alone, making it the second largest cause of cancer death among women.[2] Breast cancer that spreads to the rest of the breast is classified as invasive or infiltrating. Out of approximately 300,000 cases of breast cancer estimated last year, over 200,000 of those cases were infiltrating.[2] However, once spread to a different and unconnected part of the body, the cancer has metastasized. Metastasis is a more aggressive cancer phenotype because of the difficulty in targeting cancer cells as they move throughout the body. Unfortunately, the five year survival rate for metastatic cancer is only 24%, which is a significant drop from the 99% survival rate for women with cancer localized in only the breast.[2] Therefore, it is imperative that we understand the mechanisms by which breast cancer develops and progresses in order to develop better treatments in the future, especially for patients with invasive or metastatic tumors.

Of particular interest are the genetic factors that control the metastatic process.[3] The abnormal methylation of DNA can cause cells to become cancerous by altering expression of oncogenes and tumor suppressor genes.[4] DNA methyltransferase 3B, or DNMT3B, is one of the three main DNA methyltransferases that catalyze cytosine methylation in eukaryotic cells. It is a *de novo* methylase, which means that it is capable of methylating new sites,[5] and aberrant versions of *DNMT3B* have been found to be highly expressed in cancer cells.[6, 7] *DNMT3B7* is an aberrant transcript formed through the retention of 94bp of intron 10, leading to an early stop codon and a truncated protein. It has been shown to methylate the promoter region of *CDH1/E-cadherin*[6] leading to the loss of E-cadherin expression, which is seen when tumors become more aggressive.[8] Furthermore, previous studies have shown that *DNMT3B7* promotes tumor progression to a more aggressive phenotype in breast cancer cells through changes in cell adhesion, proliferation, and anchorage-independent growth.[8] However, an analysis of the connection between *DNMT3B7* expression and clinical parameters has not been performed to date. Therefore, the purpose of this study was to determine which clinical

It has been shown that as age increases, the probability of getting cancer also increases.[2] The more cycles of replication of DNA, the more opportunities there are for a mutation to occur.[9] Incidence and mortality of breast cancer in women makes a steady climb until the late 70's and then levels off.[2] However, it is more likely for young women to be diagnosed with aggressive tumors that are faster growing, which makes early screening important.[10,11] Premenopausal women are more likely to have a more aggressive form of breast cancer while postmenopausal women have a greater risk of getting breast cancer due to the high endogenous estrogen or testosterone levels.[12–14] Women who started their menstrual cycle early or did not lose their menstrual cycle until later in life also have a greater risk of breast cancer thought to be caused by a longer exposure to reproductive hormones. For every year younger at first menstrual cycle, the risk of breast cancer increases by a factor of 1.05, and for every year older at menopause, the risk of breast cancer increases by a factor of 1.209.[15, 16]

Race is also a factor that is correlated with different five-year survival rates in breast cancer. It has been shown that African Americans have a five year survival rate that is on average about ten percent lower than whites.[2, 17] African American women tend to be diagnosed with a more aggressive disease at a higher rate (39%) compared to non-African American women of any age (16%).[18] Furthermore, premenopausal African American women have a smaller chance (36%) of being diagnosed with less aggressive tumors compared to non-African American women of any age (54%).[18] Therefore, African American women tend to have a higher rate of aggressive breast cancer and a lower rate of the less aggressive cancer type, which may correspond to their altered survival rates.

Breast cancer diagnosis and treatment regimens are based on the expression of three hormone receptors in patients: estrogen receptor (ER), progesterone receptor (PR) and the human epidermal growth factor receptor 2 (HER2). ER and PR are considered tumor suppressor genes, which code for proteins that cause an abnormal cell to stop dividing.[19] With regard to both receptors, it is better to be positive because having a positive status means that these tumor suppressor genes are functioning normally and can respond to targeted therapies.[17] Interestingly, recent studies have shown that ER and PR may be methylated in cancers leading to the loss of expression and more aggressive disease.[19–22] The HER2 gene codes for a protein that, when expressed at normal levels, promotes growth, division and repair of cells. However, if the HER2 gene is amplified, there will be too many HER2 receptors produced, causing cells to grow and divide abnormally. A positive HER2 status is correlated with aggressive breast cancer cells, and this occurs in approximately 20 to 30% of all breast cancers.[23]

Breast cancers are classified based on their expression of these three hormone receptors in one of the following categories, ordered by increasing aggressiveness: Luminal A (ER and/or PR positive, HER2-), Luminal B (ER and/or PR positive, HER2+), HER2 type (ER-, PR-, HER2+), and basal-like/triple negative breast cancer (TNBC; ER-, PR-, HER2-). The first three types can be treated by targeted therapies to the appropriate hormone receptor being expressed, which can lead to higher survival rates. However, in the case of TNBC, targeted therapies do not exist and, therefore, the only treatment option for patients is chemotherapy. Because of this, patients with TNBC typically have a lower survival rate and more aggressive disease.

*DNMT3B7* is correlated with a more aggressive cancer phenotype,[8] so expression of *DNMT3B7* should be greatest in cells that correspond with a clinical risk factor. In order to measure this, we analyzed data that were taken from The Cancer Genome Atlas (TCGA) where patients in clinical settings allowed disclosure of their age at initial diagnosis, menopausal status, race, and status of ER, PR, and HER2 receptors. Based on previous data indicating *DNMT3B7* correlates with tumor progression in breast cancer cells,[8] we hypothesize that the expression of *DNMT3B7* will be higher in patients with a more aggressive phenotype including those who are younger, premenopausal, African American, ER-negative, PR-negative, and HER2-positive.

## METHODS AND PROCEDURES

In order to study gene expression levels in breast cancer patients, we utilized the RNA sequencing (RNAseqV2) data from the breast invasive carcinoma (BRCA) database and all corresponding clinical parameters from The Cancer Genome Atlas (TCGA) data portal.[8] Briefly, data were processed utilizing a custom Python script to provide relative expression levels of the 94bp intron 10 sequence unique to *DNMT3B7. DNMT3B7* expression was measured as reads per kilobase million (RPKM) and analyzed to determine clinical correlations in 971 primary and metastatic breast cancer patient samples. Throughout this study, N values changed due to a lack of available data for one clinical parameter compared to another. The data from the TCGA were grouped into categories consisting of age at initial diagnosis, menopausal status, race, and expression of hormone receptors using TCGA headings when appropriate. *DNMT3B7*

expression was examined and statistical significance was determined by ANOVA or T-test using Sigma Plot. All data were graphed using box plots in which the bottom and top of the box represent the 25th and 75th percentiles of data, respectively. The line within the box indicates the median while the errors bars indicate 10th and 90th percentiles of the data. Finally, the filled dots represent all outliers. Statistical significance is indicated on the graphs, when relevant.

### RESULTS

*Increased DNMT3B7 expression is observed in younger patients*

*DNMT3B7* expression was analyzed with specific age groups at the initial age of diagnosis of breast cancer. The test was done to determine if *DNMT3B7* expression is correlated with a more aggressive phenotype typically observed in younger women.[10, 11] The data obtained from the TCGA datasheet was sorted into age groupings of 25-55 and 56-90+ and the corresponding *DNMT3B7* expression value (measured by RPKM) was analyzed. These age groupings were designated based on the age at which women are more likely to be diagnosed with breast cancer.[1] There is a much higher likelihood of being diagnosed after age 55, so for the purposes of this analysis, we considered women diagnosed between 25 and 55 "young".  A Mann-Whitney Sum test revealed a statistically significant increase in *DNMT3B7* expression in younger patients compared to the older patients (P=0.032; **Figure 1A**). Patients were then grouped in a narrower range of ages — by decade — to determine if changes in *DNMT3B7* expression were observed. An ANOVA test determined that the median values did not display a statistically significant difference (**Figure 1B**).
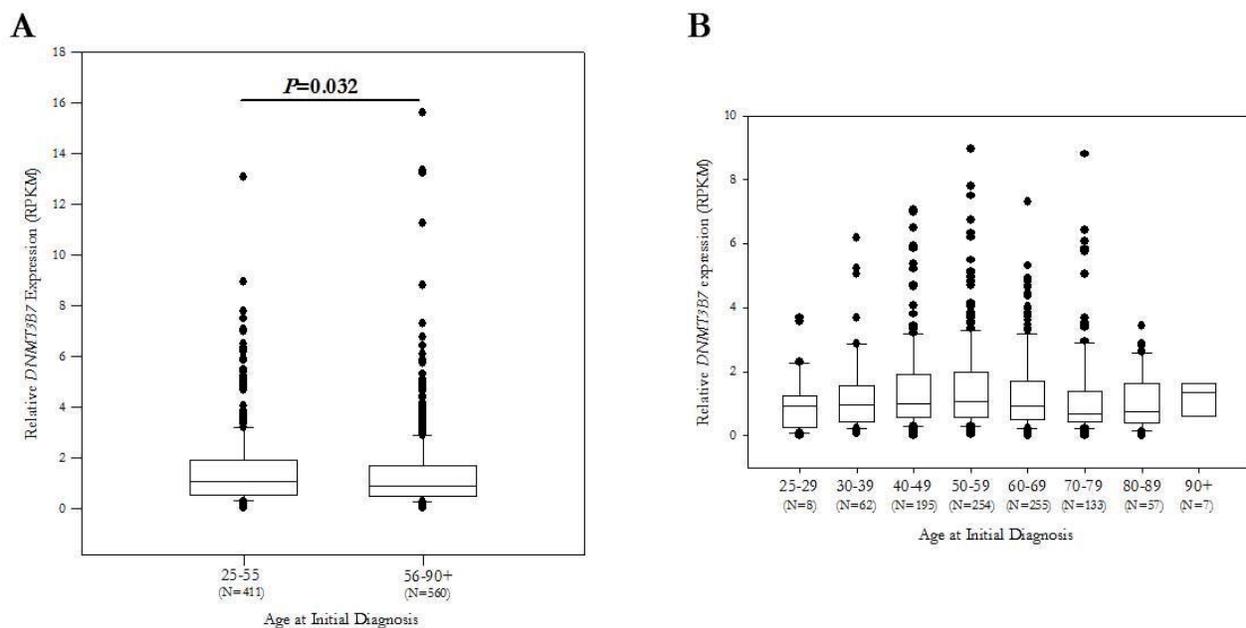


**Figure 1.** *DNMT3B7* expression is higher in younger patients. Patients with available age data were grouped (N for each group shown on graph) and *DNMT3B7* expression was measured as RPKM.  **A)** Patients were grouped as either 25-55 ("young") or 56-90+ ("old").  A Mann-Whitney Sum test indicated a significant difference between groups (*p*=0.032).  **B)** Patients were grouped by decade and an ANOVA on Ranks test was performed showing no significant difference in *DNMT3B7* expression between groups.

*DNMT3B7 expression is not correlated with menopausal status*

Based on the results of our age analysis, above, we wanted to determine if *DNMT3B7* expression was altered based on menopausal status. To that end, all patients were grouped based on their menopausal status — pre, peri, and post — and *DNMT3B7* expression levels were compared. A Kruskal-Wallis ANOVA on Ranks test determined there was no statistically significant difference between the medians of pre-, peri-, and post-menopausal status (**Figure 2**).

*African American women have higher DNMT3B7 expression.*

Because previous studies have shown that African American women typically present with a more aggressive breast tumor,[18] we analyzed *DNMT3B7* expression among races to see if there was a correlation. Patients were grouped by race — Caucasian/White, African American/Black, or Asian — and the corresponding *DNMT3B7* expression was measured. African American breast cancer patients had higher expression of *DNMT3B7* compared to their Caucasian counterparts (*P*=0.048; **Figure 3**).

*DNMT3B7 expression correlates with a more aggressive receptor phenotype.*

*DNMT3B7* expression was analyzed in comparison with ER, PR, and HER2 hormone receptor status. The data were extracted from the TCGA data sheet and separated into positive and negative receptor values in correspondence with *DNMT3B7* expression. Examination of estrogen and progesterone receptor statuses indicated significantly increased expression of *DNMT3B7* in ER- and PR- patients compared to ER+ and PR+ patients, respectively ($P<0.001$; **Figure 4A and 4B**). *DNMT3B7* was also significantly increased in HER2+ patients compared to HER2- patients ($P<0.001$; **Figure 4C**). As previously stated, the loss of ER and PR, as well as amplification of HER2, correspond to a more aggressive phenotype.[17,23]
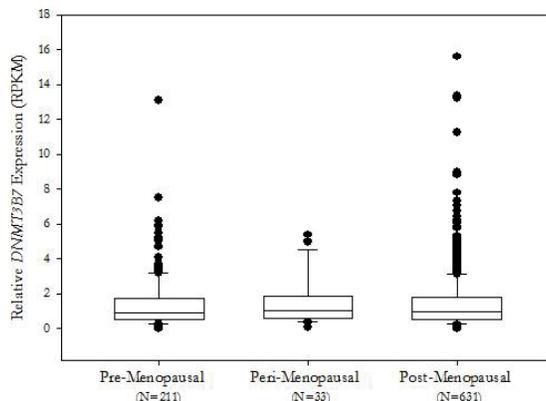


**Figure 2.** Menopausal status does not affect *DNMT3B7* expression. Patients with available data indicating menopausal status (pre-menopause = <6 months since last menstruation; peri-menopause = 6-12 months since last menstruation; post-menopause= >12 months since last menstruation) were grouped based on TCGA headings (N for each group shown on graph) and *DNMT3B7* expression was measured as RPKM. An ANOVA on Ranks test was performed and showed no significant difference between groups.
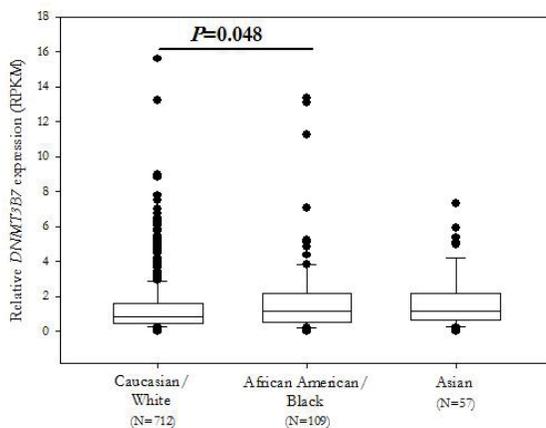


**Figure 3.** *DMNT3B7* expression is increased in African Americans compared to Caucasians. Patients with available race information were grouped based on TCGA headings (N for each group shown on graph) and *DNMT3B7* expression was measured as RPKM. A t-test was performed between each group and significance is indicated on the graph.

*Examination of DNMT3B7 expression in comparison with race and hormone receptor status combined.*

Our previous results showed increased *DNMT3B7* expression in African American women as well as women who were either ER-, PR-, or HER2+. Therefore, we wanted to examine *DNMT3B7* expression when these clinical parameters were combined, as would be found in a realistic clinical setting. All patients for which race and all receptor status data were available were isolated and grouped as shown in **Table 1**. *DNMT3B7* expression was measured as RPKM and means for all groups are indicated. An ANOVA on Ranks with a Tukey's test between all groups showed a difference between African American TNBC women and Caucasian Luminal A ($P<0.001$) and between Asian HER2-type women and Caucasian Luminal A ($P=0.040$). No other groups showed any significant difference. Interestingly, we noticed a difference in the raw number of women diagnosed in each group, as shown in **Table 1**. Specifically, while Caucasian women made up 83% of the total population of all diagnosed patients in our sample with African Americans at 9% and Asians at 8%, African American women made up 20% of the population of women diagnosed with TNBC and Asian

women made up 26% of the HER2-type population. These data indicate that minority women are being diagnosed with the more aggressive forms of breast cancer at a much higher rate than their Caucasian counterparts.
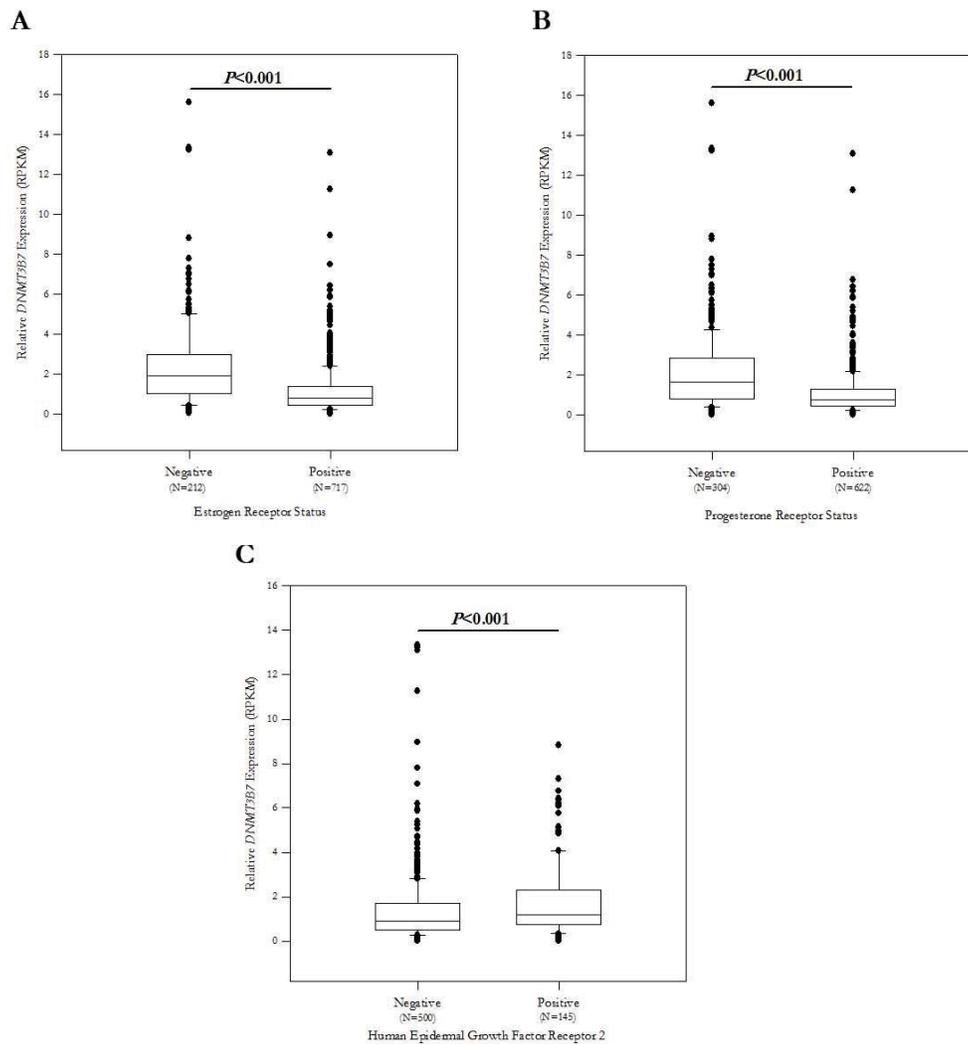


**Figure 4.** *DNMT3B7* expression is increased in ER-, PR-, and HER2+ patients. Patients with available hormone receptor information were grouped based on TCGA headings (N for each group shown on graph) and *DNMT3B7* expression was measured as RPKM. A Mann-Whitney Rank Sum test was performed between each group and significance is indicated on the appropriate graph. *DNMT3B7* expression is increased in **A)** ER- patients, **B)** PR- patients, and **C)** HER2+ patients, which corresponds to a more aggressive phenotype.

| | Luminal A | | | Luminal B | | | HER2-type | | | TNBC | | | Total | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Caucasian | African American | Asian | Caucasian | African American | Asian | Caucasian | African American | Asian | Caucasian | African American | Asian | Caucasian | African American | Asian |
| Mean RPKM | 0.9709 | 2.025 | 1.375 | 1.502 | 1.577 | 1.114 | 2.644 | 2.052 | 2.841* | 2.406 | 2.691** | 1.808 | 1.349 | 1.804 | 1.644 |
| N | 296 | 25 | 21 | 76 | 3 | 5 | 19 | 4 | 8 | 67 | 19 | 8 | 458 | 51 | 42 |
| % of total N for subgroup | 86.55% | 7.31% | 6.14% | 90.48% | 3.57% | 5.95% | 61.29% | 12.90% | 25.81% | 71.28% | 20.21% | 8.51% | 83.12% | 9.26% | 7.62% |

**Table 1.** Examination of a combination of race and hormone receptor status with regard to *DNMT3B7* expression (measured as RPKM) and raw number of patients diagnosed. Statistical significance compared to Caucasian Luminal A is indicated as * (*P*=0.040) or ** (*P*<0.001).

## DISCUSSION

*DNMT3B7* has been shown to promote a more aggressive phenotype in breast cancer cells due to a combination of parameters including changes in cell adhesion, cellular proliferation, and anchorage-independent growth.[8] However, we still do not understand the clinical implications of *DNMT3B7* expression in breast cancer patients. The data described here examine broad clinical correlations between the expression of *DNMT3B7* with respect to age, menopausal status, and race, in addition to ER, PR, and HER2 receptor status. These results allow for a more comprehensive understanding of how *DNMT3B7* promotes tumor progression within breast cancer cells which may lead to better treatment and diagnosis of patients in a clinical setting.

In order to determine if *DNMT3B7* expression was correlated with age at diagnosis, **Figure 1** illustrates the dependence of the aberrant *DNMT3B7* with respect to the age of patients at initial diagnosis. Increased expression of *DNMT3B7* in the younger age bracket when compared to the older age bracket revealed statistically relevant data between the two groupings (P= 0.032; **Figure 1A**). It should be noted that the statistically significant data was seen in **Figure 1A**, which was separated into two age brackets for younger women (25-55 years old) and older women (56-90+). Conversely, when bracketing these same data off into decades, statistically significant data are not shown (**Figure 1B**). We believe the sampling number of the treatment groups was not great enough to provide a proper analysis of whether or not these data were statistically significant. Therefore, we have focused our analysis on **Figure 1A** because of the larger sample size in two groups in order to better understand the relevance between age at initial diagnosis and *DNMT3B7* expression. To support our findings, research has shown that younger women are more apt to develop an advanced stage of breast cancer that is fast growing and lacks estrogen receptors (ER-negative).[10] Furthermore, Hofvind and colleagues demonstrated that younger women develop faster proliferating breast cancers as compared to their older female counterparts and justify the need for earlier screening.[11] This supports the idea that more aggressive phenotypes are seen in younger females than older females and thus supports our claim that higher expression of *DNMT3B7* should be seen in young females.

Further, due to the fact that older females experience a lower ratio of aggressive phenotypes as compared to younger females as supported by **Figure 1A**, it should be expected that older women that have been through menopause would demonstrate decreased expression of *DNMT3B7* when compared to pre-menopausal women. This is not the case as illustrated by **Figure 2**. Menopausal status does not appear to correlate with the aberrant gene *DNMT3B7*, as the data show no statistical significance. These data do not exclude menopausal status as a potential player in the aggressive phenotypes of breast cancer, however it can be said that menopause is not a causative agent with respect to *DNMT3B7* expression in breast cancer patients based on **Figure 2**.

In addition to age and menopausal status, we examined the effect of race on *DNMT3B7* expression in tumorigenic patients. Our results show that aberrant *DNMT3B7* is expressed in Black or African American patients at higher levels than both Asian and White patients (**Figure 3**). This could be due to a multitude of factors. One study suggests that although breast cancer mammography has increased the efficacy of treatment exponentially, it has reduced breast cancer deaths more so in Caucasian or White women as compared to Black or African American women.[25] They claim that this slower decrease in death rate amongst the African American population is due to the accessibility of technological advancements and is linked to the socioeconomic status of the majority of the African American female population.[25] Unfortunately, the information provided in TCGA does not allow us to examine this possibility in our study. It has been documented that Black or African American women show a 12.9% decrease in survival differences when compared to White women in a 5-year study.[25] Additionally, the Black or African American population had a lower survival rate that was unchanged from 1991 to 2005 (P-value < 0.001).[18] Treatment was a factor in the decreased survival rate of African American women, however, only 0.81% of the 12.9% accounted for these data. The treatment plan was thus ruled out as the main contributing factor of the survival difference. Interestingly, the presentation differences at diagnosis were deemed to be of utmost importance in linking race to breast cancer survival, which may correlate with our findings here indicating increased *DNMT3B7* expression.[26] African American women appear to have a higher incidence of TNBC and decreased survival rate when compared to their White or Caucasian counterparts.[26] These results were recapitulated in our own study (**Table 1**) and indicate an area of further study. Our laboratory has previously shown that *DNMT3B7* expression is increased in TNBC patients therefore it is imperative that we understand the mechanism by which this aberrant gene is functioning with regard to this clinical parameter.[8]

To better understand the role of *DNMT3B7* expression in TNBC patients, we compared PR, ER, and HER2 receptor status with respect to the aberrant *DNMT3B7* gene. **Figure 4** elucidates that all three of the receptors (ER, PR, HER2) show statistical significance denoting that the aberrant gene *DNMT3B7* is involved in the regulation of these specific receptors found in breast cancer patients. Breast tumors lacking a functioning estrogen receptor, ER- tumors, are known

to have poor prognoses as well as resistance to endocrine hormone therapies. Previous studies have shown that the ER-phenotype may be caused by abnormal DNA methylation.[27] Specifically, Martínez-Galán and associates have found that the estrogen receptor 1 gene, *ESR1*, is hypermethylated in tumor cells and correlated with the ER- status in patients.[27] This coincides with the function of a *de novo* methylase, like *DNMT3B7,* which can methylate specific sites and is capable of silencing tumor suppressor genes like the endocrine receptors via hypermethylation.[19–21] Taken together, the fact that *DNMT3B7* is highly expressed within ER- patients, and has the ability to methylate new sites, leads to the possibility that *DNMT3B7* could hypermethylate *ESR1* and consequently down-regulate ER in breast cancer patients. Methylation studies of the promoter region of *ESR1* in cells with and without *DNMT3B7* will be performed in the future to investigate the correlation between the *ESR1* pathway and *DNMT3B7*.

Progesterone receptors were thought to always associate with estrogen receptors due to the fact that progesterone receptors are estrogen-induced. It is now understood that PR-negative phenotypes can arise from multiple mechanisms, which include hypermethylation of the PR promoter and loss of heterozygosity of the PR gene.[22] Similar to ER, it is possible that *DNMT3B7* is methylating PR in breast cancer cells, leading to a more aggressive phenotype, and future studies will examine this relationship.

Conversely, HER2-positive phenotypes are also linked to the aberrant *DNMT3B7* gene as shown by **Figure 4C**. It is widely known that the HER family is composed of tyrosine kinases that are involved in intracellular communication cascades that regulate cell survival, cell proliferation, and migration.[28,29] Because *DNMT3B7* has been previously shown to promote cell proliferation,[8] it is possible that *DNMT3B7* and HER2 are operating through the same signaling pathway to promote tumor progression. Furthermore, HER2+ and ER- statuses have been linked together in the HER2 tyrosine kinase pathway.[30] An overexpression of HER2 (HER2+) associates with an ER- phenotype in breast cancer patients, which would aid in linking *DNMT3B7* to the two receptors. Additionally, studies have shown that loss of PR is correlated with a HER2+ phenotype.[31] Therefore, it is possible that *DNMT3B7* is methylating ER and/or PR in breast cancer cells and the loss of these receptors leads to increased HER2 expression. Taken together, this information allows us to believe that *DNMT3B7* is likely to be involved in crosstalk between the receptors leading to aggressive phenotypes.

## CONCLUSIONS
In summation, the aberrant gene, DNMT3B7, has been previously found to be highly expressed and linked to aggressive phenotypes in breast cancer cells.[8] By comparing the given parameters of age, menopausal status, race, and receptor statuses of ER, PR, and HER2, respectively, this study has allowed for a more comprehensive understanding of what was being affected by the aberrant gene DNMT3B7 in a clinical setting. Understanding the big picture of how the aberrant gene interacts with each of these parameters could aid in leading to better and more effective treatment regimens for aggressive breast cancer phenotypes in the future.

## REFERENCES
1. National Cancer Institute. Cancer. *http://www.cancer.gov* (accessed Nov 2014)
2. American Cancer Society, Inc. Breast Cancer Facts and Figures. *http://www.cancer.org* (accessed Nov 2014)
3. Nguyen, D., Massague, J. (2007) Genetic determinants of cancer metastasis, *Nature Rev Genet* 8, 341–352.
4. Jones, P. (1996) DNA methylation errors and cancer, *Cancer Res* 56, 2463–2467.
5. Pawlak M., Jaenisch R. (2011) De novo DNA methylation by DNMT3A and DNMT3B is dispensable for nuclear reprogramming of somatic cells to a pluripotent state, *Genes Dev* 25(1), 1035–1040.
6. Ostler, K.R., Davis, E.M., Payne, S.L., Gosalia, B.B., Exposito-Cespedes, J., Le Beau, M.M., Godley, L.A. (2007) Cancer cells express aberrant DNMT3B transcripts encoding truncated proteins, *Oncogene* 26, 5553–5563.
7. Ostler, K.R., Yang, Q., Looney, T.J., Zhang, L., Vasanthakumar, A., Tian, Y., Kocherginsky, M., Raimondi, S., Demaio, J., Salwen, H., Gu, S., Chlenski, A., Naranjo, A., Gill, A., Peddinti, R., Lahn, B., Cohn, S., Godley, L. (2012) Truncated DNMT3 isoform DNMT3B7 suppresses growth, induces differentiation, and alters DNA methylation in human neuroblastoma, *Cancer Res* 72, 4714–4723.
8. Brambert, P.R., Kelpsch, D.J., Hameed, R., Desai, C., Calafiore, G., Godley, L.A., Raimondi, S.L. (2015) DNMT3B7 expression promotes tumor progression to a more aggressive phenotype in breast cancer cells, *PLoS ONE* 10(1), e0117310.

9.  Venkitaraman, A. (2000) The breast cancer susceptibility gene, BRCA2: At the crossroads between DNA replication and recombination? *Phil Trans: Biol Sciences* 355(1394), 191–198.

10. Freedman R.A., Partridge, A.H. (2013) Management of breast cancer in very young women, *Breast* 22 Suppl 2, S176–S179.

11. Hofvind, S., Vacek, P.M., Skelly, J., Weaver, D.L., Geller, B.M. (2008) Comparing screening mammography for early breast cancer detection in Vermont and Norway, JNCI *J Natl Cancer Inst* 100(15), 1082–1091.

12. Fuhrman, B., Schairer, C., Gail, M., Boyd-Morin, J., Xu, X., Sue, L., Buys, S., Isaacs, C., Keefer, L., Veenstra, T., Berg, C., Hoover, R., Ziegler, R. (2012) Estrogen metabolism and risk of breast cancer in postmenopausal women, *J Natl Cancer Inst* 104(4), 326–339.

13. Key, T. (2011) Endogenous oestrogens and breast cancer risk in premenopausal and postmenopausal women, *Steroids* 76(8), 812–815.

14. Hankinson, S. and Eliassen, A. (2007) Endogenous estrogen, testosterone and progesterone levels in relation to breast cancer risk, *J Steroid Biochem Mol Biol* 106(1-5), 24–30.

15. Kelsey, J., Gammon, M., John, E. (1993) Reproductive factors and breast cancer, Epidemiol Rev 15(1), 36–47.

16. Collaborative Group on Hormonal Factors in Breast Cancer. (2012) Menarche, menopause, and breast cancer risk: individual participant meta-analysis, including 118 964 women with breast cancer from 117 epidemiological studies, *Lancet Oncol* 13(11), 1141–1151.

17. Chevarley, F., White, E. (1997) Recent trends in breast cancer mortality among white and black US women, *Am J Public Health* 87(5), 775–781.

18. Carey, L. A., Perou C.M., Livasy, C.A., Dressler, L.G., Cowan, D., Conway, K., Karaca, G., Troester, M.A., Tse, C.K., Edmiston, S., Deming, S.L.., Geradts, J., Cheang, M.C., Nielson, T.O., Moorman, P.G., Earp, H.S., Millikan, R.C. (2006) Race, breast cancer subtypes, and survival in the Carolina Breast Cancer Study, *JAMA* 295, 2492–2502.

19. Zhu, K., Williams, S. (1998) Methyl-deficient diets, methylated ER genes and breast cancer: A hypothesized association, *Cancer Causes & Control* 9(6), 615–620.

20. Mostoslavsky, R., Bergman, Y. (1997) DNA methylation: regulation of gene expression and role in the immune system, *Biochim Biophys Acta* 1333, 29–50.

21. Issa, J., Zehnbauer, B., Civin, C., Collector, M., Sharkis, S., Davidson, N., Kaufmann, S., Baylin, S. (1996) The estrogen receptor CpG island is methylated in most hematopoietic neoplasms, *Cancer Res* 56, 973–977.

22. Lapidus R.G., Ferguson A.T., Otaviano Y.L., Parl F.F., Smith S.B., Weitzman S.A., Baylin S.B., Isa J.J., Davidson N.E. (1996) Methylation of estrogen and progesterone receptor gene 5' CpG islands correlates with lack of estrogen and progesterone reception gene expression in breast tumors, *Clin Cancer Res* 2, 805–810.

23. Engel, R., Kaklamani, V. (2007) HER2-positive breast cancer, *Drugs* 67(9), 1329–1341.

24. DeSantis, C., Jemal, A., Ward E., Thun, M.J. (2008) Temporal trends in breast cancer mortality by state and race, *Cancer Causes Control* 19, 537–545.

25. Sibler, J.H., Rosenbaum, P.R., Clark, A.S., Giantonio, B.J., Ross, R.N., Teng, Y., Wang, M., Niknam, B.A., Ludwig, J.M., Wang, W., Even-Shoshan, O., Fox, K.R. (2013) Characteristics associated with differences in survival among black and white women with breast cancer, *JAMA* 310(4), 389–397.

26. Veer, L.J., Dai, H., Van de Vijver, M., He, Y., Hart, A., Maro, M., Peterse, H., Van der Kooy, K., Marton, M., Witteveen, A., Schreiber, G., Kerkhoven, R., Roberts, C., Linsley, P., Bernards, R., Friend, S. (2002) Gene expression profiling predicts clinical outcome of breast cancer, *Nature* 415, 530–536.

27. Martínez-Galán, J., Torres-Torres, B., Núñez, M.I., López-Peñalver, J., Del Moral, R., Ruiz De Almodóvar, J.M., Menjón, S., Concha, Á., Chamorro, C., Ríos, S., Delgado, J.R. (2014) *ESR1* gene promoter region methylation in free circulating DNA and its correlation with estrogen receptor protein expression in tumor tissue in breast cancer patients, BMC *Cancer* 14, 59.

28. Yarden, Y., Sliwkowski, M.X. (2001) Untangling the ErbB signaling network, Nat Rev Mol Cell Biol 2(2), 127–137.

29. Fink, M.Y., Chipuk, J.E. (2013) Survival of HER2-positive breast cancer cells: Receptor signaling to apoptotic control centers, *Genes Cancer* 4(5-6), 187–195.

30. Pietras R.J., Arboleda J., Reese D.M., Wongvipat, N., Pegram, M.D., Ramos, L., Gorman, C.M., Parker, M.G., Sliwkowski, M.X., Slamon, D.J. (1995) HER-2 tyrosine kinase pathway targets estrogen receptor and promotes hormone-independent growth in human breast cancer cells, *Oncogene* 10, 2435–2446.

31. Kim H.J., Cui, X., Hilsenbeck S.G., Lee, A.V. (2006) Progesterone receptor loss correlates with human epidermal growth factor receptor 2 overexpression in estrogen receptor-positive breast cancer, *Clin Cancer Res* 12(3pt2), 1013s–1018s.

**ABOUT THE STUDENT AUTHORS**

Christian Mullin graduated from Elmhurst College in May of 2015 with a Bachelor of Science (BS) degree in Biology and a minor in Chemistry. He is currently applying to dental schools.

Jennifer Busser is currently enrolled at Elmhurst College and is working towards graduating with a Bachelor of Science degree in Biology with minors in Chemistry and Medical Humanities in May of 2016. She plans on attending Podiatry School after graduation.

Anna Ciampanelli graduated from Elmhurst College in May 2015 with a Bachelor of Science degree in Biology and a minor in Chemistry. She hopes to attend PA school in the future.

**PRESS SUMMARY**

Invasive breast cancer caused almost 40,000 deaths last year alone. When breast cancer progresses to a more advanced stage, it can be due to a multitude of factors, one of which is epigenetic changes. DNA methylation is an example of an epigenetic change in which a methyl group is added to DNA leading to altered gene expression. Our research focuses on the aberrant DNA methyltransferase, *DNMT3B7,* which is capable of methylating new sites that have not been previously methylated.  Prior research has shown that expression of *DNMT3B7* can turn off important genes that are needed to combat cancerous cells and promotes tumor progression to an aggressive state. However, the effect of *DNMT3B7* expression on clinical parameters was not known. The objective of our research was to analyze data from breast cancer patient samples available in The Cancer Genome Atlas database and attempt to correlate these clinical parameters to *DNMT3B7* expression. Our data demonstrate that increased expression of *DNMT3B7* correlates to age, race, and hormone receptor status. These findings have the potential to lead to better diagnosis and improved treatment regimens in the future.