# EPIGENOME-WIDE ASSOCIATION ANALYSIS OF DNA METHYLATION DATA IN EARLY AND LATE-STAGE BREAST CANCER SAMPLES

Authors: Olya Besedina, Mengran Ma, Shannon Nees, Eleanor Zhang

MAY 10, 2019

COLUMBIA UNIVERSITY MEDICAL CENTER - OBGYN

# Epigenome-wide Association Analysis of DNA Methylation Data in Early and Late-stage Breast Cancer Samples

**Authors:** Olya Besedina, Mengran Ma, Shannon Nees, Eleanor Zhang

## Abstract

DNA methylation is crucial in the regulation of gene expression. DNA methylation data may be useful in identifying genes associated with breast cancer progression since abnormal gene expression can lead to malignancy. Data from The Cancer Genome Atlas (TCGA) was analyzed the test for possible associations between breast cancer stage and level of DNA methylation at CpG sites using data from the Illumina 450 platform. This was assessed by analyzing the difference in DNA methylation between early (stage I) and late (stage III) breast cancer samples using an epigenome-wide association study design. Cases included 10 women with advanced stage III Invasive breast cancer and controls included 10 women with early stage I Infiltrating duct carcinoma. *β*-values acquired from the TCGA-BRCA data set were first checked for bimodal distribution, and then transformed using log2(*β)* transformation, such that transformed *β-values* follow normal distribution. A linear regression model was used to test for association between DNA methylation and breast cancer stage. A minimum p value approach was used to select the top 10 significant CpG sites with differences in methylation between stage 1 and stage 3 breast cancer. The significant CpG sites include the following genes: RIC88 (p value = 1.01769e-05), ABCC8 (p value = 1.4192e-05), HMGB2 (p value = 1.434e-05), FAM171A2 (p value = 2.8713e-05), KATNBL1 (p value = 3.4533e-05), G6PC3 (p value = 4.4701e-05), H2FAY (p value = 5.1845e-05), MYO19;PIGW (p value = 5.2051e-05), CASP9 (p value = 7.7232e-05). Details about these genes can be found in table 3. Distribution of all p values from the association test is depicted by Manhattan plot on figure 3. Benjamini Hochberg method was used to adjust raw p-values, however, due to the low sample size (n=20) adjusted p-values were very inflated. A literature search was performed to determine if any of the 13 associated genes had evidence of a role in breast cancer. Five of the thirteen genes genes had previously been reported to be associated with breast cancer progression or response to therapy: *ABCC8, H2AFY, HMGB2, G6PC3* and *CASP9.* Our analysis of DNA methylation was able to detect differences between early and late stages of breast cancer, which were previously reported as clinically important.

## Background

Epigenetic regulation of chromatin structure and, as a result gene expression, is significantly affected by DNA methylation. Alteration of DNA methylation has a great effect on tumorigenesis and tumor-suppression. DNA methylation is a process in which methyl group (CH3) are added to DNA molecule; in humans two out of four DNA bases can undergo methylation - cytosine and adenine and methylation appears to be a default state. CpG islands are sequences of the genome which have content of cytosine and guanine higher than 50%, length higher than 200bp and ratio of observed to expected CpG greater than 0.6. The human genome has approximately 250,000 CpG islands, excluding repeated sequences. About 50% of CpG islands are located in the gene promoter region. It has been shown that methylation of the promoter region is negatively correlated with gene expression, which explains why CpG-dense promoters of actively transcribed genes are never methylated. Methylation can affect transcription by physically preventing binding of transcriptional proteins and by serving as a binding site for proteins, which assist in formation of inactive chromatid (Wikipedia contributors DNA methylation). It is reported that hyper-methylation can cause inactivation of tumor-suppressor genes. Furthermore, DNA methylation is associated with inactivity, but in the case of cancer,

inactivation of tumor suppressor genes means indirect stimulation and allows for proliferation of tumor cells (Celli et al. 2018). Neoplasm is the result of uncontrolled cell division and growth, when tumoral cells begin to invade and destroy surrounding tissues it considered to be malignant. Several studies have demonstrated a role for analysis of DNA methylation data and have shown that DNA methylation assists in the classification of breast cancers and in predicting response to certain types of therapy (Szyf 2012).

**Illumina Infinium HumanMethylation450 BeadChip**

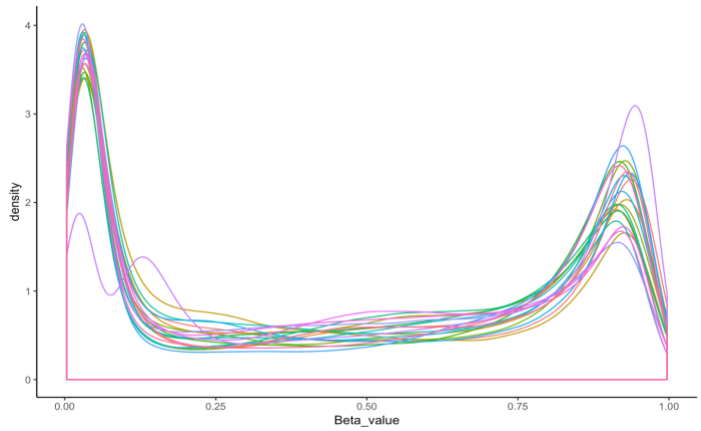
For the samples that we selected from the TCGA study, the Illumina Infinium HumanMethylation450 BeadChip array was utilized which targets about 96% of CpG islands in the human genome. The Illumina 450K BeadChip has two distincts probe types Infinium I and Infinium II. Infinium I targets each CpG site with two 50bp probes which detect methylated (M) and unmethylated (U) intensities, whereas Infinium II uses one probe to separate methylated from unmethylated sites using green and red dyes. $\beta$-value indicates the level of methylation of one CpG site and can be calculated using the following formula: $\beta = M / (M + U + a)$, where a is usually equal to 100. M-value is a logit transformed $M = log2(\beta/(1-\beta))$ (Wang et al. 2018).

**Experimental Design**

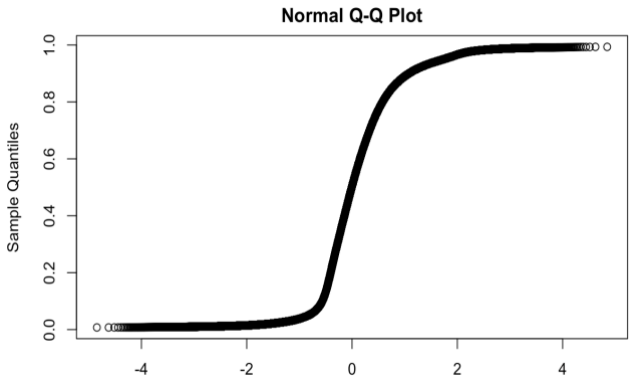
Data were obtained from DNA harmonized methylation data based on the platform of Illumina Bead array 450k (reference human genome hg38) in the TCGA repository. Subjects of uniform gender and ethnicity were selected to control for potential confounders. 10 white non-Hispanic female cases were chosen randomly from the TCGA-BRCA project with advanced stage III A/B infiltrating duct carcinoma, a very common subtype of invasive breast cancer. 10 white non-Hispanic female subjects were selected as the independent control group with the earliest stage (stage 1) of the same subtype of breast cancer. DNA methylation from exclusively breast tissue was included in our study. The goal was to detect those CpG sites where DNA methylation level was significantly different between cases at an advanced stage of breast cancer and controls subjects at early stage in order to determine possible candidate genes that may be involved in breast cancer progression through changes in DNA methylation. For computational efficiency, the total sample size (N = 20) is far from sufficient to capture the association between DNA methylation level and cancer stage progression. Therefore, subjects at either extreme of cancer phenotype, Stage III and Stage I, were chosen as cases and independent controls for the purpose of maximizing effect size and to ultimately boost power in this small study.

**Material and Methods**

Few data preprocessing steps were required for this DNA methylation data since the raw data retrieved from TCGA portal was well prepared and already recalculated into $\beta$ value at each probe site with annotated gene symbols and precise location. To confirm quality control of this dataset, bimodal distributions of $\beta$ values for each subject were plotted in Figure 1. This is consistent with the expected distribution of DNA methylation data because hypermethylation and hypomethylation are generally well separated among all CpG sites (Panel A). These $\beta$ values form beta distributions and therefore deviate from Gaussian distribution as shown in panel B. Hence, further transformation of $\beta$ values (outcome variable) is needed for a linear regression model.




(A) Distribution of $\beta$ values for each subject.




(B) Normal QQ plot of $\beta$ values

**Figure 1**. Quality Control of DNA methylation data of 20 subjects under study

An association test on each CpG site was implemented by applying linear regression and testing the significance of coefficients of the indicator group variable as either case or control. The linear regression model adjusts for age of diagnosis but other factors such as gender and disease subtypes were well controlled in the study design phase so were not included in the regression model. Log 2 transformation was applied to the outcome $\beta$ values in order to improve linear model validity as has been previously described.

Each single coefficient test is supposed to be controlled at 0.05 level. In order to adjust for multiple comparisons, the Benjamini Hochberg (BH) method was applied to adjust raw p values and control the false discovery rate (FDR). Given the very limited sample size, adjusted p values are very likely to be enormously magnified since multiple association tests are performed across over 400k CpG sites on 20 subjects. In this case, BH method was noted to be overly conservative and disallowed any rejection. Therefore, raw p values were ranked from least to greatest. Top 10 most significant raw p values (< 0.00008) were chosen to be candidate sites for further exploration and literature search using the minimum p value method.

Based on the newly selected cut-off, we obtained the top 10 significant CpG sites with corresponding annotated gene symbol and chromosome location and we focused on these 10 specific sites across all individual subjects. Particularly, we examined the genes related to these CpG sites, and explored the published evidence for association between these genes and breast cancer.

## Results

### Clinical Data

We downloaded the clinical data files available for both cases and controls included in our study. Both cases and controls were selected for female sex and white non-Hispanic race/ethnicity in order to match the characteristics between the groups. There were significant differences in the tumor stage between cases and controls as this was the basis for case selection. There were no significant differences between cases and controls in terms of age of diagnosis, follow-up time or death (Table 1 and Figure 2). Because there were no significant differences, these variables were not considered confounders in our regression analysis except for age at diagnosis since there is a significant clinical impact of age on DNA methylation levels. Although there were more deaths in the case group compared to controls, this did not reach statistical significance.

**Table 1.** Clinical Data for 10 cases and 10 controls

| Patient Characteristic | Cases (n=10) | Controls (n=10) | P Value* |
|---|---|---|---|
| **Female sex** | 10 (100%) | 10 (100%) | |
| **White non-Hispanic** | 10 (100%) | 10 (100%) | |
| **Tumor stage**<br>Stage IA<br>Stage IIIA<br>Stage IIIB | <br>0<br>8 (80%)<br>2 (20%) | <br>10 (100%)<br>0<br>0 | **<0.001** |
| **Median age at diagnosis (years)** | 60.5 (IQR 54.2-69.3) | 60.8 (IQR 49.0-64.1) | 0.94 |
| **Median time to last follow-up (years)** | 1.5 (IQR 0.2-6.8) | 2.0 (IQR 1.1-3.8) | 0.72 |
| **Death** | 3 (30%) | 0 | 0.21 |

*P values determined by Fisher's exact test for categorical data and Kruskal-Wallis for continuous data.

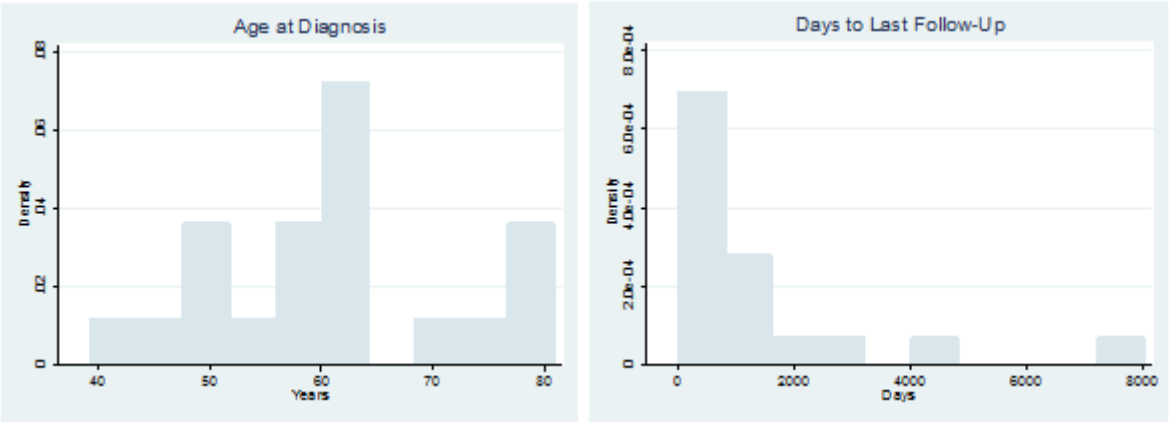



**Figure 2.** Distribution of age at diagnosis and days to last follow-up demonstrating non-normal distribution.

## CpG Sites with Differential Methylation

None of the CpG sites studied in our analysis had significantly different methylation between cases and controls using the Benjamini-Hochberg multiple testing correction. This is likely due to our small sample size and insufficient power to detect small differences. Given this, we decided to explore the 10 CpG sites with the lowest P values to see if there was evidence for their association with breast cancer (Table 2). Unadjusted p values for these 10 sites were all less than $10^{-5}$.

**Table 2.** DNA methylation data for the 10 CpG sites with the minimum p values

| CpG Site | B Value Cases | B Value Controls | logFC | Unadjusted P Value | BH adj P value | Gene Symbol |
|---|---|---|---|---|---|---|
| cg24053061 | 0.04393409 | 0.03065907 | -0.5121730 | 1.017695e-05 | 0.9999825 | RIC8B; RP11-144F15.1 |
| cg03181582 | 0.38848546 | 0.06651329 | -2.5118092 | 1.419112e-05 | 0.9999825 | ABCC8 |
| cg08269316 | 0.09373440 | 0.06474669 | -0.5283102 | 1.433395e-05 | 0.9999825 | HMGB2 |
| cg09014329 | 0.52624834 | 0.28683522 | -0.8843724 | 2.871338e-05 | 0.9999825 | FAM171A2 |
| cg22958262 | 0.15884106 | 0.23940282 | 0.6213344 | 3.453296e-05 | 0.9999825 | KATNBL1 |
| cg21535106 | 0.38463638 | 0.03542464 | -2.8699683 | 3.486304e-05 | 0.9999825 | NA |
| cg18218381 | 0.03965898 | 0.02787680 | -0.5067727 | 4.470060e-05 | 0.9999825 | G6PC3 |
| cg07827630 | 0.28897833 | 0.12365964 | -1.2385538 | 5.184274e-05 | 0.9999825 | CTC-203F4.2; H2AFY |
| cg03705042 | 0.03309506 | 0.02171572 | -0.5703352 | 5.205051e-05 | 0.9999825 | MYO19;PIGW |
| cg24483825 | 0.06900949 | 0.04507329 | -0.6010103 | 7.723168e-05 | 0.9999825 | CASP9 |

The relative methylation level of control (stage I cancer) versus cases (stage III cancer) is measured in log 2 fold change (logFC) in the table. Most of these CpG sites are hypermethylated in the cases compared to control group with the exception of one site associated with *KATNBL1* that demonstrates hypomethylation in the cases. FDR adjusted p value is displayed. It is not surprising to observe that the adjusted p value is inflated by BH adjustment method due to limited sample size. All CpG sites except one unknown gene annotation are in the protein coding region. The distribution of p values across all CpG sites tested with the association test can be better visualized in the Manhattan plot (Figure 3).

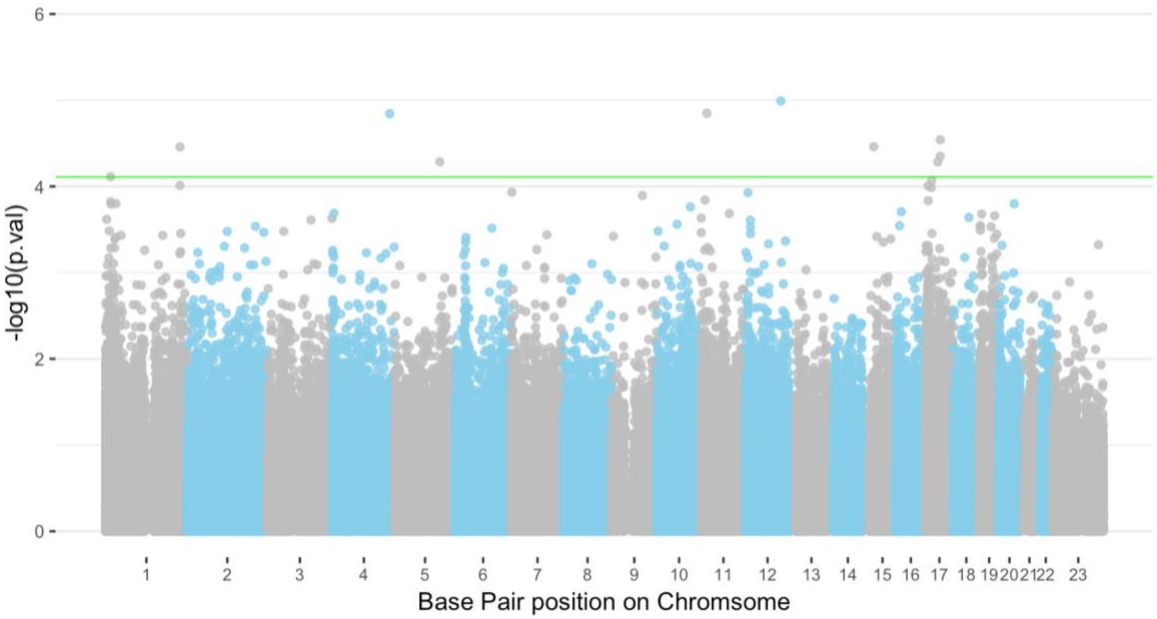



**Figure 3. Manhattan plot.**
P values for all CpG sites across the epigenome. Signals above green line cutoff are the 10 most significant raw p values. *****chromosome X is indexed as 23

Interestingly, several of the genes related to our most significant CpG sites have been implicated in breast cancer in previous studies. Altered expression of these genes has been associated with progression or therapy responsiveness in breast cancer models. Since DNA methylation is an important regulator of gene expression, it is possible that these genes are differentially expressed in our cases compared to controls as a result of their altered methylation. Details on each of the genes that have evidence for association with breast cancer are included below. In addition, if these genes act as tumor suppressors, it would make sense

that there is increased methylation in the more advanced tumor stages which silences expression of those genes.

**Table 3.** CpG Sites with the lowest P values when comparing methylation between cases and controls.

| CpG Site/Position | Chr | Gene | Gene Function | Known Disease Association | Evidence in Breast Cancer |
|---|---|---|---|---|---|
| cg03181582 | chr11 | ABCC8 | Sulfonylurea receptor | Diabetes Mellitus, hyperinsulinemic hypoglycemia | **Yes** |
| cg03705042 | chr17 | MYO19 | Myosin head | None | No |
| cg03705042 | chr17 | PIGW | Acetylation of phosphatidylinositol involved in anchoring of proteins to cell surface | Glycosylphosphatidylinositol biosynthesis defect 11 (Developmental delay and seizures) | No |
| cg07827630 | chr5 | CTC-203F4.2 | Novel Antisense | None | No |
| cg07827630 | chr5 | H2AFY | Histone family | Liebenberg syndrome (upper limb malformation) | **Yes** |
| cg08269316 | chr4 | HMGB2 | High mobility group protein involved in DNA binding and transcription regulation | None | **Yes** |
| cg09014329 | chr17 | FAM171A2 | Unknown | None | No |
| cg18218381 | chr17 | G6PC3 | Glycongenolysis and gluconeogenesis | Autosomal recessive congenital neutropenia (Dursun syndrome) | **Yes** |
| cg21535106 | chr1 | . | | | |
| cg22958262 | chr15 | KATNBL1 | Microtubule-severing ATPase | None | No |
| cg24053061 | chr12 | RIC8B | Unknown | None | No |
| cg24053061 | chr12 | RP11-144F15.1 | Unknown | None | No |
| cg24483825 | chr1 | CASP9 | Apoptosis | None | **Yes** |

ABCC8

ATP-binding cassette, subfamily C, member 8 (*ABCC8*) is a gene that encodes for sulfonylurea receptor 1 protein which is part of the ATP-sensitive potassium channel found in pancreatic beta cells. In general, genes in this class encode cellular transporters. The *ABCC8* receptor controls how much insulin is secreted from the beta cells. There is significant evidence for the association between mutations in *ABCC8* and several

types of diabetes mellitus and well as hyperinsulinemic hypoglycemia and hypoglycemia of infancy (OMIM, Genetics Home Reference). Several studies demonstrate that *ABCC8* may have a role in breast cancer progression or response to therapy. In a study investigating the expression of ATP-binding cassette transporter genes in breast cancer tissue, levels of ABCC8 were found to be associated with tumor grade and the expression of hormone receptors, which are important for cancer treatment and response to chemotherapy (Hlaváč et al. 2013). Dvorak et al studied the expression of multiple ABC genes in cancer tissues including breast cancer and found that *ABCC8* expression was altered in multiple cancer cell lines and associated with worse clinical prognosis in several cases (Dvorak et al. 2017). Another study examined the role of *ACSL4* in drug resistance in breast cancer cell lines and found that in cells that over-expressed *ACSL4*, expression of *ABCC8* was increased along with several other transporters which may be related to drug transport into the cells (Orlando et al. 2019). Given that DNA methylation is an important regulator of gene expression, the difference in methylation of *ABCC8* that we observe between early and late-stage breast cancer tissues may affect the expression of this gene and predispose those with later stage cancer to progression.

H2AFY

H2A Histone Family, member Y (*H2AFY*) is a variant histone that encodes for protein macroH2A1. Histones wrap DNA and play an important role in regulating gene expression and this particular histone has been shown to be involved in X-inactivation. Deletions in this gene have been associated with mild Liebenberg syndrome, a syndrome associated with upper extremity malformations (OMIM, Genetics Home Reference). *H2AFY* is important in the progression of malignant melanoma with normal expression suppressing the development of this disease (Kapoor et al. 2010). A study by Halvorsen et al examined differential DNA methylation in breast cancer patients before and after radiation therapy. When they examined methylation levels before treatment, they found five genes that had methylation levels associated with clinical response to radiation therapy including *H2AFY*. *H2AFY* was also found to have significantly different methylation levels before and after radiation therapy (Halvorsen et al. 2014). Another study demonstrated that overexpression of *H2AFY* in cancer cells led to increased cell proliferation and tumorigenicity and that it interacts with *HER-2*, a gene known the affect the responsiveness of breast cancer tissue to treatment (Li et al. 2012). These data suggest that *H2AFY* plays a role in the response to therapy in breast cancer cells and potentially with progression of breast cancer and other cancers, which would be consistent with our observation that there is differential methylation between early and late stage breast cancers.

HMGB2

High-mobility group box 2 (*HMGB2*) is a member of the high-mobility group box family which encode proteins that have a critical role in regulating DNA expression by binding to DNA (OMIM, Genetics Home Reference). A study examining the role of *HMGB2* in breast cancer showed that it was more highly expressed in the nuclei of breast cancer tumor cells compared to the adjacent normal breast tissue and that higher expression was correlated with increased tumor size and more advance stage. Expression level was correlated with worse prognosis as well (Fu et al. 2018). Another study demonstrated a role for *HMGB2* in determining resistance to endocrine therapy in breast cancer patients (Redmond et al. 2015). Alteration in DNA methylation that affects *HMGB2* could contribute to altered levels of expression and therefore allow progression of tumors. This is consistent with the differential methylation observed in our study.

GPC3

Glucose-6-phosphatase, catalytic, 3 (*G6PC3*) is a ubiquitously expressed catalytic unit for glucose-6=phosphatase which is involved in the last step of the gluconeogenic and glycogenolytic pathways and therefore important for cellular metabolism. Mutations in this gene have been associated with congenital neutropenia (OMIM, Genetics Home Reference). Genes involved in energy metabolism have been implicated

in several cancers. One study examined the commonly mutated tumor suppressor gene *TP53* and cellular metabolism in breast cancer cell lines. They compared gene expression between *TP53* mutated samples and wild-type samples and found that among other alterations, *G6PC3* was downregulated in these cells (Harami-Papp et al. 2016). Thus, *G6PC3* may play a role in altered cellular metabolism observed in tumor cells.

CASP9

Caspase 9, apopotosis-related cysteine protease (*CASP9*) is a member of the caspase family, which includes proteins that regulate cell apoptosis (OMIM, Genetics Home Reference). Changes in this protein disrupt the mitochondrial cell death pathway and mutations in *CASP9* have been implicated in multiple different cancer types (Kesarwani et al. 2011). Brynychova et al showed that *CASP9* was down-regulated in breast cancer tumors and it also affected the expression of hormonal receptors on the breast cancer cells. The ratio of specific isoforms of *CASP9* was associated with survival (Brynychova et al. 2017). One study examined the role of a microRNA miT-182-5p, which was shown to be upregulated in breast cancer cells compared to controls. They demonstrated that inhibition of this microRNA led to upregulation of *CASP9* and therefore induced apoptosis of these cells indicating a potential therapeutic role in breast cancer (Sharifi and Moridnia 2017). Expression of *CASP9* and its role in inducing apoptosis is clearly an important factor in progression and treatment response in multiple tumor types including breast cancer and therefore the difference in methylation that we observe may impact the gene expression in our cases compared to controls.

## Conclusion

Even with a small sample size, we were able to detect differences in DNA methylation between advanced and early stage breast cancer samples. We used a regression-based analysis to test for association between DNA methylation levels and breast cancer stage in early and advanced breast cancer samples selected from the TCGA database in the TCGA-BRCA project. Although none of the sites tested met statistical significance with a Benjamin Hochberg correction for multiple comparison, when we selected the 10 most significant variants using the minimum p value method, we found that 5 of the affected genes have previously been studied in breast cancer samples. For these 5 genes, there was evidence that differential expression affects breast cancer progression or response to therapy. Thus, our comparison of DNA methylation between early and late stage breast cancers found clinically relevant differences consistent with prior evidence. A larger sample size would allow for increased power to detect differences between these two groups.

**References**


Brynychova V, Hlavac V, Ehrlichova M, Vaclavikova R, Nemcova-Furstova V, Pecha V, Trnkova M, Mrhalova M, Kodet R, Vrana D, et al. 2017. Transcript expression and genetic variability analysis of caspases in breast carcinomas suggests CASP9 as the most interesting target. Clin Chem Lab Med 55: 111–122.

Celli F, Cumbo F, Weitschek E. 2018. Classification of Large DNA Methylation Datasets for Identifying Cancer Drivers. Big Data Res 13: 21–28.

Dvorak P, Pesta M, Soucek P. 2017. ABC gene expression profiles have clinical importance and possibly form a new hallmark of cancer. Tumor Biol 39: 101042831769980.

Fu D, Li J, Wei J, Zhang Z, Luo Y, Tan H, Ren C. 2018. HMGB2 is associated with malignancy and regulates Warburg effect by targeting LDHB and FBP1 in breast cancer. Cell Commun Signal 16: 8.

Halvorsen AR, Helland A, Fleischer T, Haug KM, Grenaker Alnaes GI, Nebdal D, Syljuåsen RG, Touleimat N, Busato F, Tost J, et al. 2014. Differential DNA methylation analysis of breast cancer reveals the impact of immune signaling in radiation therapy. Int J cancer 135: 2085–95.

Harami-Papp H, Pongor LS, Munkácsy G, Horváth G, Nagy ÁM, Ambrus A, Hauser P, Szabó A, Tretter L, Győrffy B. 2016. TP53 mutation hits energy metabolism and increases glycolysis in breast cancer. Oncotarget 7: 67183–67195.

Hlaváč V, Brynychová V, Václavíková R, Ehrlichová M, Vrána D, Pecha V, Koževnikovová R, Trnková M, Gatěk J, Kopperová D, et al. 2013. The expression profile of ATP-binding cassette transporter genes in breast carcinoma. Pharmacogenomics 14: 515–529.

Kapoor A, Goldberg MS, Cumberland LK, Ratnakumar K, Segura MF, Emanuel PO, Menendez S, Vardabasso C, LeRoy G, Vidal CI, et al. 2010. The histone variant macroH2A suppresses melanoma progression through regulation of CDK8. Nature 468: 1105–1109.

Kesarwani P, Mandal RK, Maheshwari R, Mittal RD. 2011. Influence of caspases 8 and 9 gene promoter polymorphism on prostate cancer susceptibility and early development of hormone refractory prostate cancer. BJU Int 107: 471–476.

Li X, Kuang J, Shen Y, Majer MM, Nelson CC, Parsawar K, Heichman KA, Kuwada SK. 2012. The atypical histone macroH2A1.2 interacts with HER-2 protein in cancer cells. J Biol Chem 287: 23171–83.

National Library of Medicine (US). Genetics Home Reference [Internet]. Bethesda (MD): The Library; 2010 Apr 30 [cited 2019 May 9]. Available from: https://ghr.nlm.nih.gov/.

Online Mendelian Inheritance in Man, OMIM®. McKusick-Nathans Institute of Genetic Medicine, Johns Hopkins University (Baltimore, MD), {2019 May 9}. World Wide Web URL: https://omim.org/

Orlando UD, Castillo AF, Medrano MAR, Solano AR, Maloberti PM, Podesta EJ. 2019. Acyl-CoA synthetase-4 is implicated in drug resistance in breast cancer cell lines involving the regulation of energy-dependent transporter expression. Biochem Pharmacol 159: 52–63.

Redmond AM, Byrne C, Bane FT, Brown GD, Tibbitts P, O'Brien K, Hill ADK, Carroll JS, Young LS. 2015. Genomic interaction between ER and HMGB2 identifies DDX18 as a novel driver of endocrine resistance in breast cancer cells. Oncogene 34: 3871–3880.

Sharifi M, Moridnia A. 2017. Apoptosis-inducing and antiproliferative effect by inhibition of miR-182-5p through the regulation of CASP9 expression in human breast cancer. Cancer Gene Ther 24: 75–82.

Szyf M. 2012. DNA methylation signatures for breast cancer classification and prognosis. Genome Med 4: 1–12.

Wang Z, Wu X, Wang Y. 2018. A framework for analyzing DNA methylation data from Illumina Infinium HumanMethylation450 BeadChip. BMC Bioinformatics 19: 115.

Wikipedia contributors. "DNA methylation." *Wikipedia, The Free Encyclopedia*. Wikipedia, The Free Encyclopedia, 8 May. 2019. Web.