Multivariate comparison of DNA methylation events in breast cancer stem cells

*Caroline E. Dyar[1] and Austin Y. Shull[2]
[1]Undergraduate Student, Department of Mathematics, Presbyterian College, [2]Assistant Professor, Department of Biology, Presbyterian College
cedyar@presby.edu

Keywords: DNA methylation; epigenetics; cancer stem cell; genomics; breast cancer

Abstract: Breast cancer is the second most common type of cancer diagnosed amongst women, with one in eight women being diagnosed with breast cancer within their lifetime. Due to the prevalence of breast cancer, earlier detection of both the presence and aggressiveness of this disease are important when treating patients. Recent evidence suggests that cancer aggressiveness is associated with the expansion of tumor subpopulation known as cancer stem cells (CSCs), and stem cells are understood to have different DNA methylation patterns within their genome. Thus, the aim of this project was to analyze DNA methylation sites genome wide across 19 breast cancer samples, in order to determine if DNA methylation events correspond directly to increasing aggressiveness among cancer cell lines with enriched CSC populations. Unsupervised PCA and matrix dissimilarity analysis revealed three distinct groups that appear to cluster independently based on relative cancer aggressiveness and CSC enrichment. From this analysis, we determined that the methylation events in the gene body correlated more closely to CSC enrichment when compared to promoter DNA methylation events. To determine which specific methylation events defined the 3 identified clusters, we performed differential methylation using ANOVA and found 1432 differentially methylated promoter probes and 7243 differentially methylated gene body probes. Interestingly, the differentially methylated promoter probes tended to be hypermethylated in the CSC-enriched cluster, while the differentially methylated gene body probes were more hypomethylated. An example of a gene that demonstrated significant hypomethylation throughout the gene body was *MCF2L*. We repeated many of the previous analyses done on the gene body as a whole on solely the methylation events of *MCF2L* and found that *MCF2L* could reconstruct the same clustering patterns demonstrated when using gene body probes indiscriminately. Through these tests, we found that gene body methylation events detail a closer progression with breast cancer aggressiveness compared to the promoter methylation events. Thus, these results could help with earlier prognosis through sequential tracking of gene body methylation events.