Proceedings of The National Conference
On Undergraduate Research (NCUR) 2018
University of Central Oklahoma
Edmond, Oklahoma
April 5-7, 2018

Profiling DNA methylation of DNMTs in normal lung tissues from African Americans and European Americans with lung cancer

Amy E. Boles
Department of Biology
Lafayette College
Easton, Pennsylvania

Faculty Advisor: Dr. Khadijah A. Mitchell

Abstract

Background:African Americans (AA) have higher incidence and mortality rates for lung cancer compared with European Americans (EA), despite smoking less. DNA methyltransferases (DNMT) are proteins that add methyl groups to DNA, causing gene silencing. Overexpression of DNMTs, and increased DNA methylation is associated with multiple cancers. The biological basis of the disparity can be explored by comparing DNA methylation of *DNMTs* between normal and lung tumor tissues by race. Hypotheses: H1: *DNMTs* are differentially methylated and expressed in normal lung tissues from AA compared with EA. H2: *DNMTs* are differentially methylated and expressed in normal lung tissues compared with lung tumors. Methodology: A total of 61 paired normal and lung tumor tissues from AA (n=5) and EA (n=56) were extracted using TCGA data for DNA methylation and mRNA-Seq data. Data were imported to Partek Genomics Suite 7.0 and differential methylation analyses (1-way ANOVA and paired t-tests) were performed. GraphPad Prism 7.03 was used to perform differential gene expression analysis. Cellular pathway association analysis was performed using STRING (an on-line bioinformatic tool). Results: *DNMT* genes are differentially methylated in normal lung tissues from AA compared with EA, with *DNMT1* showing a decrease from EA to AA and then *DNMT3A* and *DNMT3B* showed an increase. However, *DNMT* genes show no racial differences in gene expression. When considering tissue-specific differences, *DNMT* genes are differentially methylated in normal lung tissues compared with lung tumors, with *DNMT1* and *DNMT3B* showing decreased DNA methylation, while *DNMT3A* had increased DNA methylation. As seen in literature, *DNMT1* and *DNMT3B* cooperate to drive DNA methylation in the human genome validating the overexpression of *DNMT1* and *DNMT3A* genes from normal to tumor tissues. More genes are impacted by DNA methylation in the High DNMT1 expression group. STRING analysis showed that low and high DNMT1 expressors had different cellular pathways impacted by DNA methylation. Conclusions: While racial differences in DNA methylation from normal tissues can't explain the increased lung cancer risk in African Americans, DNA methylation profiles and associated pathways are found to vary according to tissue-type. Discussion: Considering differences in levels of DNMT expression could act as a form of precision medicine through the grouping of patients based on high and low expression.Future Directions: Further works will include obtaining additional samples for racial studies to gain more power for the AA cohort. For the tissue specific studies, other works will look into correlating the *DNMT* gene expression with the associated DNMT protein expression status through primary samples and cell lines. Other studies will look into identifying molecular subgroups based on the *DNMT1* and *DNMT3B* expression status with histology, smoking status and survival rate. Finally, this information will be validated in an additional cohort.

Keywords: DNA, DMNT, Lung Tissue

1. Introduction

Lung cancer has the highest incidence and mortality rates in men and the third highest incidence and then second highest mortality rate in women across the world [1]. However, when you break that incidence up by race, African Americans (AA) have the highest incidence rate compared to any minority group [2]. Frequently, smoking is attributed as the leading cause of lung cancer. However, when breaking up smoking rates, AA actually smoke less than Americans of European descent (EA), and this difference is more pronounced when broken up by race and gender [3]. So, smoking alone cannot explain this disparity. In addition to smoking, other known risk factors for lung cancer include environmental exposures, such as exposure to radon gas, secondhand smoke, exposure to asbestos and air pollution, and a positive family history, suggesting there could be biological determinants involved [4]. Shared biology amongst AA and EA may help to explain the racial disparity.

There are 6 key biological hallmarks of a cancer cell compared with a normal cell. Two specific ones include proliferative signaling and evading growth suppressors. These lead to activation of oncogenes, which cause cancer, and inhibition of tumor suppressors, which stop cancer. Each of these hallmarks lead to increase cancer phenotype [5]. Some individuals have an increased risk of cancer based on biological predispositions that can be found in their normal tissues. This increased risk can vary by population.

One way these hallmarks are exhibited are through DNA methylation. DNA methylation is the addition or removal of chemical tags to the gene in order to silence or express them, respectively [6]. Increased DNA methylation, also known as hypermethylation, in tumor suppressor genes is often associated with various cancers [6]. DNA Methyltransferases, or *DNMTs*, are the proteins that drive and carry out DNA methylation in the body [7]. These genes have been shown to be hypermethylated in other cancer types with a known racial disparity, such as breast and colon cancers [8, 9]. However, to date, no research has been done on the *DNMT* family in lung cancer health disparities. So we hypothesized that *DNMTs* are differentially methylated & expressed in normal lung tissues from AA compared with EA to address the methylation and expression status of *DNMTs* in normal lung tissues from AA and EA. Later in the experiment, it was additionally hypothesized that *DNMT* genes are differentially methylated & expressed in normal lung tissues to address tissue- specificity and DNMT regulation in normal lung tissues and lung tumors, independent of race.

2. Materials and Methods

We used primary normal lung tissues from the The Cancer Genome Atlas Cohort, or TCGA. TCGA is a publicly available dataset with over 33 cancers, including lung cancer. For the first hypothesis, we downloaded Illumina 450K Methylation Array data from TCGA for 5 AA and 56 EA lung cancer patients. Using Partek Genome Suite 7.0, a differential methylation analysis was performed and a 1-way ANOVA was done to determine statistical significance of the analysis. For the same cohorts, mRNA-sequencing data was downloaded from TCGA in order to perform a differential expression analysis on Partek Genome Suite 7.0. Statistical significance was determined with a two-tailed t-test with Welch correction. For the second hypothesis, the same procedures were followed for differential methylation and expression analyses as well as statistical significance, however, the two cohorts were 61 normal tissues and 61 lung tumors, independent of race. Additionally, DNMT expression data was input into the STRING analysis website in order to determine the functional protein association.

3. Results

Our differential methylation analysis revealed racial differences in DNA methylation for all 3 *DNMTs*. Interestingly, we saw *DNMT1*, the maintenance methyltransferase had decreased DNA methylation in AA, while the *de novo* methyltransferases, *DNMT3A* and *DNMT3B*, had increased DNA methylation in this population.

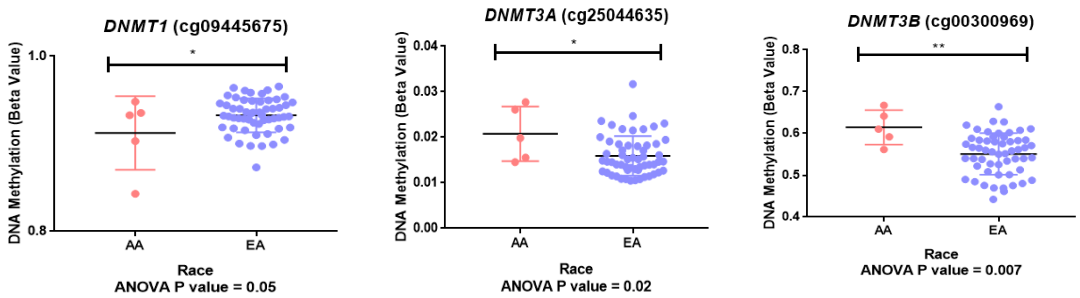


Figure 1. Racial difference are indicated in differential methylation analysis race comparison for the *DNMT* gene family

However, these racial differences in DNA methylation did not translate to differences in gene expression.

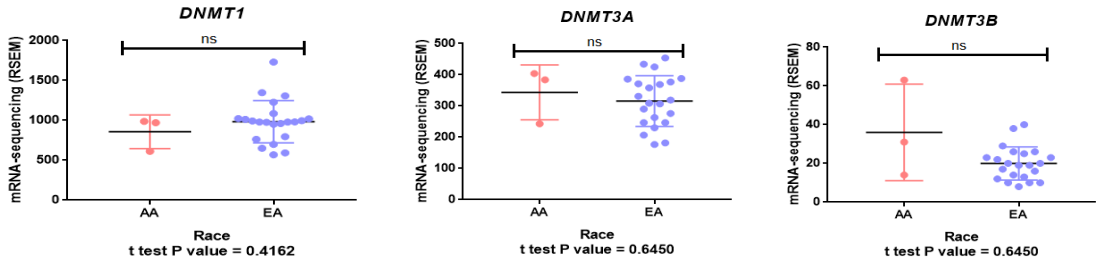


Figure 2. No racial differences are seen in gene expression for the *DNMT* gene family

When visiting the second hypothesis, the differential methylation analysis showed that *DNMT*s genes are differentially methylated in normal lung tissues compared with lung tumors. *DNMT1* and *DNMT3B* had decreased DNA methylation, while *DNMT3A* had increased DNA methylation.

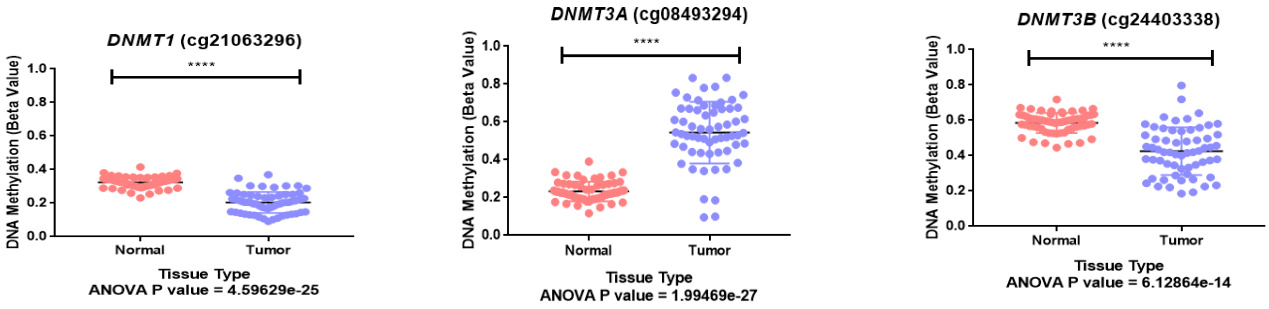


Figure 3. DNA methylation differential analyses indicate there are tissue specific differences in the *DNMT* gene family

The differential expression analysis showed that all *DNMT* genes are overexpressed from normal tissue to tumor tissue.

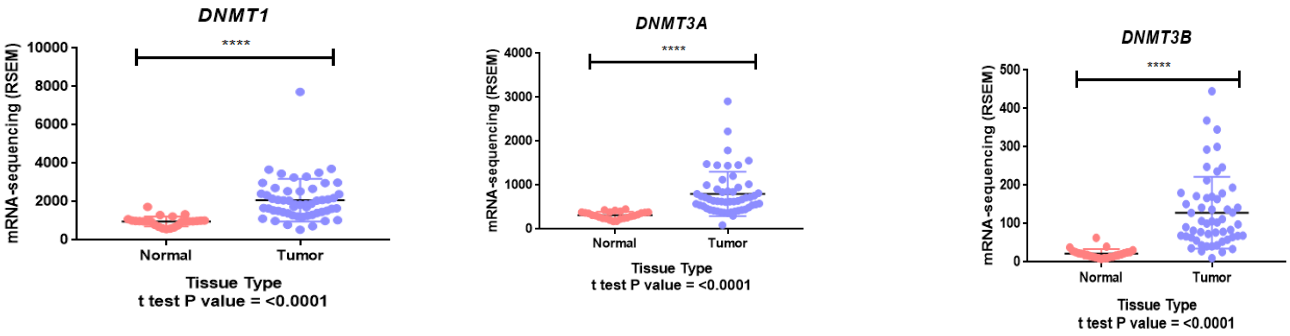
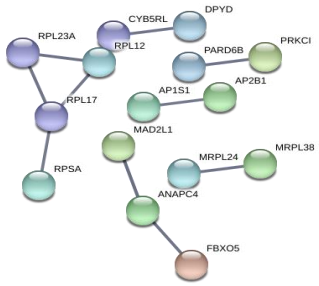
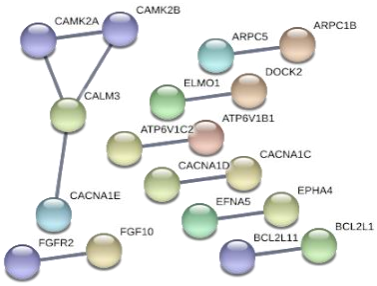


Figure 4. All genes of the DNMT gene family are overexpressed from normal lung tissues to lung tumor tissues in differential expression analyses

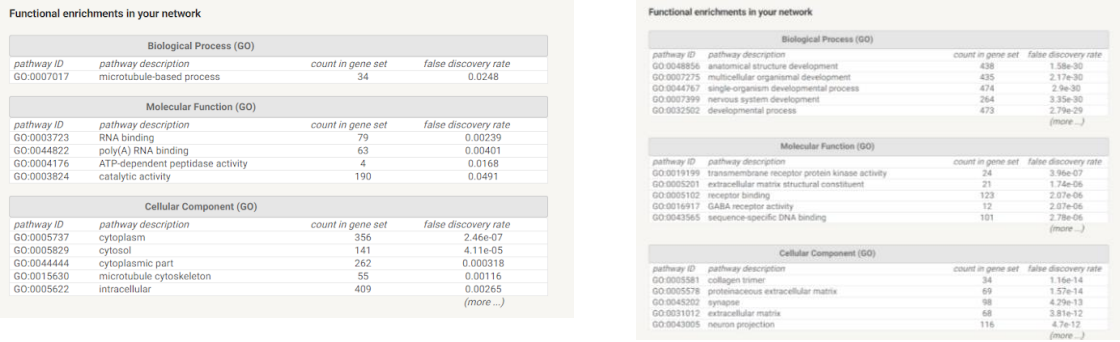
DNMT1 gene expression showed the highest level of gene expression. In an effort to identify new molecular subgroups, we categorized patients as either low or high *DNMT1* expressors. Through STRING analysis, a online tool that looks at functional protein associations, we found that both low and high *DNMT1* expressors use DNA methylation to regulate different cellular pathways.



Molecular subgroup: Low DNMT1 expression in lung tumors
High methylation in normal (*n* = 666, P value = 0.001)
Most likely oncogenes



Molecular subgroup: High DNMT1 expression in lung tumors
High methylation in tumor (*n* = 4175, P value = 0.001)
Most likely tumor suppressor genes



Functional enrichments in your network

Biological Process (GO)

pathway ID	pathway description	count in gene set	false discovery rate
GO:0007017	microtubule-based process	34	0.0248

Molecular Function (GO)

pathway ID	pathway description	count in gene set	false discovery rate
GO:0003723	RNA binding	79	0.00239
GO:0044822	poly(A) RNA binding	63	0.00401
GO:0004176	ATP-dependent peptidase activity	4	0.0168
GO:0003824	catalytic activity	190	0.0491

Cellular Component (GO)

pathway ID	pathway description	count in gene set	false discovery rate
GO:0005737	cytoplasm	356	2.46e-07
GO:0005829	cytosol	141	4.11e-05
GO:0044444	cytoplasmic part	262	0.000318
GO:0015630	microtubule cytoskeleton	55	0.00116
GO:0005622	intracellular	409	0.00265
			(more ...)

Functional enrichments in your network

Biological Process (GO)

pathway ID	pathway description	count in gene set	false discovery rate
GO:0048856	anatomical structure development	438	1.58e-30
GO:0007275	multicellular organismal development	435	2.17e-30
GO:0044767	single-organism developmental process	474	2.9e-30
GO:0007399	nervous system development	264	3.35e-30
GO:0032502	developmental process	473	2.79e-29
			(more ...)

Molecular Function (GO)

pathway ID	pathway description	count in gene set	false discovery rate
GO:0019199	transmembrane receptor protein kinase activity	24	3.96e-07
GO:0005201	extracellular matrix structural constituent	21	1.74e-06
GO:0005102	receptor binding	123	2.07e-06
GO:0016917	GABA receptor activity	12	2.07e-06
GO:0043565	sequence-specific DNA binding	101	2.78e-06
			(more ...)

Cellular Component (GO)

pathway ID	pathway description	count in gene set	false discovery rate
GO:0005581	collagen trimer	34	1.16e-14
GO:0005578	proteinaceous extracellular matrix	69	1.57e-14
GO:0045202	synapse	98	4.29e-13
GO:0031012	extracellular matrix	68	3.81e-12
GO:0043005	neuron projection	116	4.7e-12
			(more ...)

Figure 5. STRING analysis determined high and low *DNMT1* expression patient groups' DNA methylation patterns impact different cellular pathways

4. Discussion

The race specific analysis result is most likely due to a lack of power. The initial cohort was 5 AA, however, as we moved to the differential expression analysis, we had fewer normal lung tissue samples from AA with mRNA-sequencing data compared with DNA methylation data. When considering the tissue specific analysis, this expression result validates what was expected for *DNMT1* and *DNMT3B*. Additionally, this corresponding DNA methylation and gene expression data for *DNMT1* and *DNMT3B* validated published data from Dr. Bert Vogelstein's lab at Johns Hopkins School of Medicine stating that *DNMT1* and *DNMT3B* work together to methylate the human genome [10]. We found more genes were impacted by DNA methylation in the high *DNMT1* expressing patients. This finding correlates with the known function of *DNMT1*, which is to methylate DNA. Knowledge that there is a difference in the cellular pathways impacted by different expressors indicates which genes become silenced in lung cancer and imply a difference in tumor biology.

5. Conclusion

While *DNMT* genes are differentially methylated in normal lung tissues from AA to EA, *DNMT* genes show no racial differences in gene expression, indicating that racial differences cannot explain this disparity. However, tissue specific analyses showed that there is differential methylation and associated differential expression from normal tissue to tumor tissue, independent of race. Considering differences in levels of *DNMT* expression could act as a form of precision medicine through the grouping of patients based on high and low expression.

6. Future Directions

For the race-specific analysis, we are hoping to work with collaborators to increase normal lung tissue samples from AA and repeat analysis with additional power. For the tissue-specific analysis, we are hoping to determine if *DNMT* protein expression status correlates to *DNMT1* gene expression using TCGA protein data, and normal lung and lung cancer cell line data. Additionally, we're hoping to identify new molecular subgroups in order to correlate *DNMT1* and *DNMT3B* expression status with histology, smoking status, and survival data. Finally, these findings will be validated in an additional cohort.

7. References

1. American Cancer Society. “Cancer Facts and Figures 2017.” American Cancer Society , pp. 1–76, 2017.
2. Center for Disease Control. “United States Cancer Statistics.” CDC, 2018.
3. Ahmed Jamal et al. “Current Cigarette Smoking Among Adults — United States, 2016.” Center for Disease Control Morbidity and Mortality Weekly Report (MMWR), 2018.
4. Bristol-Myers Squibb. “Lung Cancer by the Numbers.” Bristol-Myers Squibb, 2017.
5. Flavahan WA et al. “Epigenetic plasticity and the hallmarks of cancer.” *Science*. 357(6348), 2017.
6. Lam, Kevin et al. “DNA methylation markers in colorectal cancer state-of-the-art.” *Atlas of Science*. 2017.
7. Zhang, Wu and Jie Xu. “DNA methyltransferases and their roles in tumorigenesis.” *Biomarker Research*. 5(2017):1, 2017.
8. Huidobro, C et al. “A DNA methylation signature associated with aberrant promoter DNA hypermethylation of DNMT3B in human colorectal cancer.” *European Journal of Cancer*. 48(14):2270-2281, 2012.
9. Naghitorabi, M et al. “Quantitative evaluation of DNMT3B promoter methylation in breast cancer patients using differential high resolution melting analysis.” *Research in Pharmaceutical Sciences*. 8(3): 167–175, 2013.
10. Rhee, Ina et al. “DNMT1 and DNMT3b cooperate to silence genes in human cancer cells.” *Nature*. 416: 552-556, 2002.