Characterization of Purmorphamine's Effects on Chromatin Structure

April 7, 2017

Haley Higginbotham

Characterization of Purmorphamine's Effects on Chromatin Structure

Abstract

The Hedgehog signaling pathway is a cellular pathway that plays a key role in cell differentiation and proliferation. This project intended to study how purmorphamine, a molecule which activates the Hedgehog signaling pathway, affects chromatin structure. Chromatin is composed of DNA and protein complexes called nucleosomes. Chromatin structure is therefore directly related to the characteristics of nucleosomes.

In order to characterize purmorphamine's effects on chromatin structure, nucleosome characteristics, namely position and micrococcal nuclease (MNase) sensitivity, were mapped and compared for purmorphamine-treated and untreated (control) sample tissues. Towards that end, the samples were subject to digestion via varied concentrations of the enzyme MNase. MNase digestion selectively cuts the DNA between nucleosomes (internucleosomal DNA), and is thus commonly used in the industry to isolate the fragments of DNA wrapped around individual nucleosomes (mononucleosomally protected DNA). Sequencing the fragments released by different MNase concentrations permits their location within the genome to be determined, enabling an analysis of individual nucleosome's "sensitivity" to MNase digestion and position relative to the rest of the genome. A comparison can then be made between these nucleosome characteristics in treated and untreated cells.

Nucleosome sensitivity and distribution are characteristics of chromatin structure which can indicate the accessibility of certain segments of DNA, and by extension the regulation of gene expression. As a result, it was predicted that purmorphamine-treated cells would display

nucleosome sensitivity and distribution changes that created a chromatin structure more favorable for transcription of purmorphamine-activated genes. Such structural changes could theoretically include increased nucleosome sensitivity near the transcription start sites (TSS) of genes activated in the Hedgehog pathway or decreased nucleosome occupancy at those locations.

Arrangements were made to have the samples sequenced at Florida State University, however, an unforeseen lack of human and financial resources forced the university to withdraw from this commitment late in the experimental process. Instead, the Dennis Lab provided surrogate data from samples sequenced in a similar, previous experiment. The samples in that experiment were processed using the same procedures as this experiment and so the data they provide is used to illustrate what data from the intended samples might look like and what sort of conclusions can be drawn from this type of data.

Introduction

Before delving into the study itself, it is necessary to establish a foundation of background information regarding chromatin structure, transcription, the relationship between nucleosome characteristics and gene expression, and the Hedgehog signaling pathway.

Chromatin Structure

DNA in the nucleus is organized into chromosomes, which are composed of chromatin. Chromatin fibers are composed of nucleosomes and the DNA which links these nucleosomes (Figure 1). A nucleosome consists of 146 base pairs (bp) of DNA wrapped approximately 1.5 times around a histone octamer, which is a protein complex formed from eight subunits called histones (Luger, Mäder, R. Richmond, Sargent, & T. Richmond, 1997). It may be helpful to envision chromatin as a string of Christmas lights. In this analogy, the bulbs represent

nucleosomes, the wire represents the strand of DNA, and the tangled mass formed by the unwound lights could be compared to the chromatin fiber.

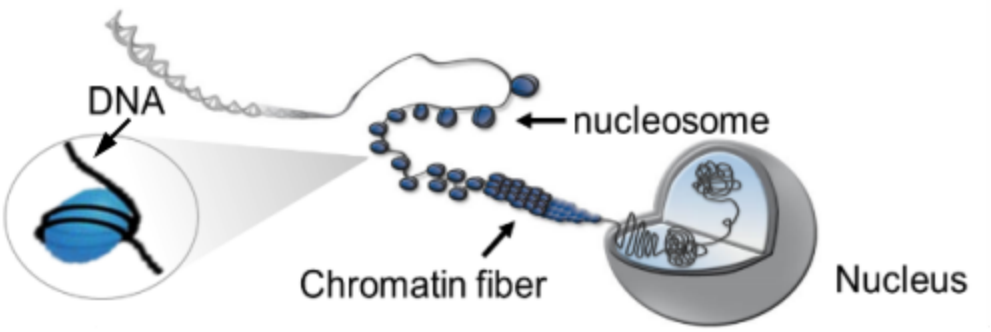

Figure 1. Chromatin Structure

Transcription

Segments of DNA called genes contain the biochemical instructions needed for cells to produce particular proteins. In order for eukaryotic cells to start "reading" these instructions, they must first find the "bookmark." Proteins called transcription factors serve as "bookmarks" by binding to regions of DNA called promoters, which occur right before the start of genes. The binding of a specific class of transcription factors, general transcription factors, enables the binding of RNA polymerase, forming what is known as the transcription initiation complex (TIC). RNA polymerase, the cell's scribe, then "writes" a copy of the gene for the cell to "read." This copy is composed of RNA, a molecule structurally very similar to DNA, hence the name RNA polymerase (Nature Education, n.d.-a) (Figure 2). The process of producing the RNA transcripts of genes is called transcription, and is the first step in protein production. Regulation of transcription is the most common way in which gene expression is controlled (Nature Education, n.d.-b).

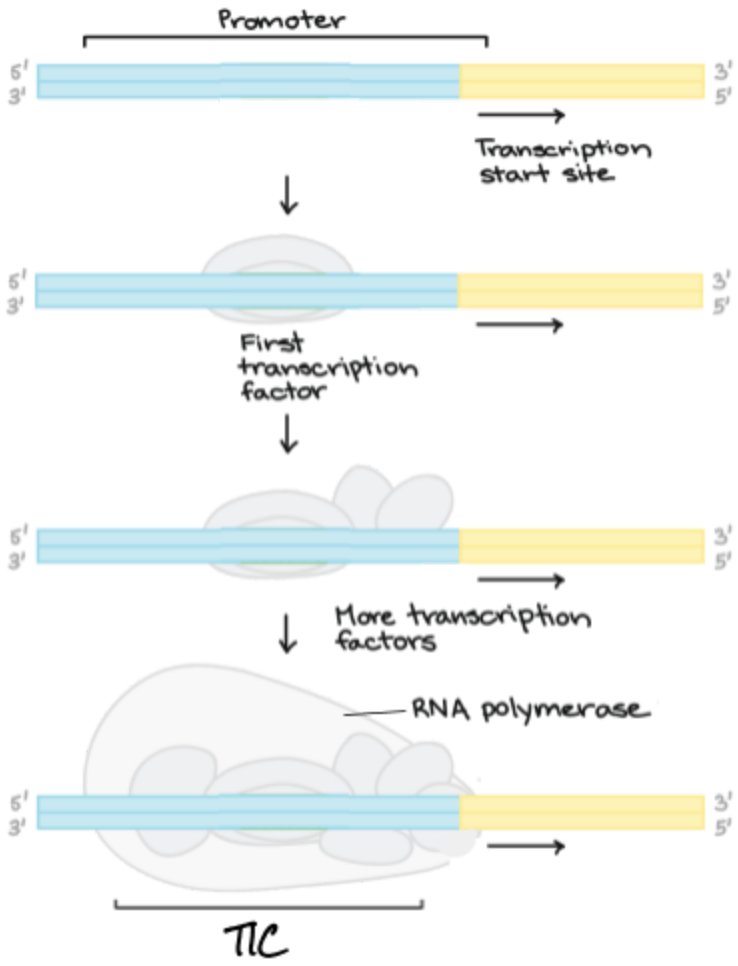

Figure 2. Transcription. Adapted from *Khan Academy*, n.d., Retrieved from https://www.khanacademy.org/science/biology/gene-expression-central-dogma/transcription-of-dna-into-rna/a/stages-of-transcription.

Nucleosome Characteristics and Gene Expression

Epigenetics is the study of factors other than DNA sequence that affect gene expression. Past research in epigenetics has identified chromatin structure as one such factor, as correlations have been found between nucleosome characteristics, such as position and MNase sensitivity, and the level of transcription of genes.

Past research has uncovered evidence of how typical nucleosome positions differ in active and inactive genes. As noted by a study conducted through the University of North Carolina at Chapel Hill, "during rapid…[cellular] growth [and replication], the level of nucleosome occupancy is inversely proportional to the transcriptional initiation rate at the promoter." Nucleosomes have been found to be more prevalent in the promoter region of inactive genes and, conversely, less prevalent in the promoter of active genes (Lee, Shibata, Rao, Strahl, & Lieb, 2004). These findings indicate that nucleosome occupancy in the promoter region of a gene may inhibit transcription of the gene.

The mechanism by which nucleosome position exerts this effect has not been definitively proven, but one well-supported theory postulates that occupancy in the promoter region can restrict access to the promoter for transcription factors. DNA is negatively charged while the histones that form nucleosomes are positively charged, resulting in a tight bond between them. If part of a gene's promoter region is wrapped around a nucleosome, the general transcription factors may not be able to bind to that DNA to promote RNA polymerase binding and initiate transcription, and thus the gene may not be able to be expressed, or "turned on" (Mellor, 2005; Phillips & Shaw, 2008). Nucleosomes therefore may exert their influence on transcription by regulating access to certain segments of DNA.

In order to discuss individual nucleosomes surrounding a TSS, a number-based nomenclature system has been devised in which nucleosome number increases as distance from the TSS increases, with upstream (before the TSS) elements denoted as negative and downstream (after the TSS) elements as positive. For example, the first nucleosome upstream of the TSS is designated as -1 while the first nucleosome downstream the TSS is designated as +1.

Upstream nucleosomes located in the promoter region of a gene, especially the -1 nucleosome, can regulate promoter access. As a result of this key role, the -1 nucleosome in particular has been observed to undergo extensive modification during transcription in order to permit promoter access. These changes can be a simple repositioning or even complete eviction. A study by researchers of the National Heart, Blood, and Lung Institute, for instance, provides evidence of this role as their results showed that, when compared with the occupancy of the relatively stable -2 nucleosome, -1 nucleosome occupancy was 40% lower in expressed genes but only 16% lower in unexpressed genes (Schones et al., 2008). Accordingly, it is predicted that the results of the present study will show lower levels of expression for genes with higher levels of nucleosome occupancy, especially in the -1 position.

Another nucleosome characteristic explored in this study is MNase sensitivity. Heavy (high concentration) and light (low concentration) MNase digests produce different populations of mononucleosomally protected fragments, thereby indicating differential nucleosome sensitivity to MNase digestion. MNase sensitivity has been proposed as a measure of DNA accessibility, and by extension the probability of transcription of individual genes. In support of this view, a study by affiliates of Massachusetts General Hospital, Harvard Medical School, and Brown University found that the regions of DNA with the highest MNase sensitivity were those associated with active transcription. Additionally, this study confirmed that the correlation between MNase sensitivity and transcriptional activity was consistent between different types of organisms (Mieczkowski et al., 2016).

Further evidence of the utility of measuring MNase sensitivity was provided by a study from Florida State University which found that MNase sensitive nucleosomes tended to be

located at either end of a gene, with MNase resistant nucleosomes in the middle. This nonrandom distribution of MNase sensitivity appears to indicate that such regions play a role in gene regulation. This study went further, however, and established that the MNase sensitivity of nucleosomes near TSSs is positively correlated with the expression level of genes (Vera et al., 2014). It is therefore predicted that the results of the present study will reveal higher levels of nucleosome sensitivity near the TSS of highly expressed genes, which in this case would be those activated by the Hedgehog signaling pathway.

The Hedgehog Signaling Pathway

The Hedgehog (Hh) signaling pathway is a chain of cellular events that play a critical role in embryonic development, differentiation, and cellular regulation and repair, and the malfunction of which has been connected to certain types of cancers (Corcoran, 2006).

In general, the process of cellular signaling is initiated by the binding of a signal molecule, called a ligand, to a protein called a receptor. This binding triggers a cascade of cellular events, culminating in a cellular response, such as gene transcription (Mason, Losos, Singer, Raven, & Johnson, 2011a).

Two transmembrane (extending across the cellular membrane) proteins that play a critical role in the Hh signaling pathway are called Patched (PTCH) and Smoothened (SMO). PTCH is a receptor for Hh ligand, and in the absence of Hh binding PTCH inhibits SMO. When Hh ligand binds to PTCH the inhibition of SMO is lifted, initiating a series of cellular events, the specifics of which have yet to be elucidated. This cellular relay culminates in the binding of certain transcription factors, classified as Gli transcription factors (Kiecker, Graham, & Logan, 2016) (Figure 3). Gli transcription factor binding enables the expression of specific target genes, most

notably genes involved in cellular growth, as mentioned previously.

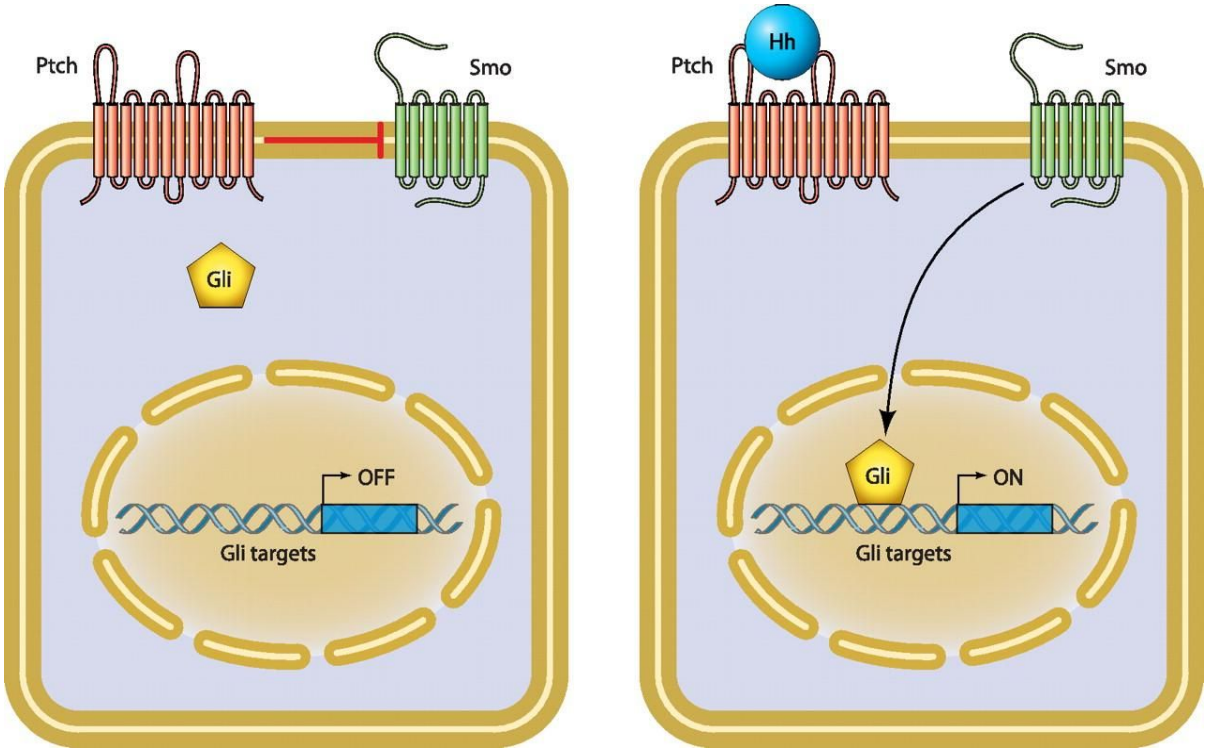

Figure 3. The Hedgehog Signaling Pathway. Reprinted from "Hedgehog Signaling and Maintenance of Homeostasis in the Intestinal Epithelium," by N. V. Büller, S. L. Rosekrans, J. Westerlund, and G. R. van den Brink, 2012, *American Physiological Society, 27*(3), p. 150.

A study by researchers at Stanford played a key role in uncovering the role of PTCH in the Hh pathway. They discovered that inactivation of the gene coding for PTCH in mice resulted in the activation of two Hh target genes. These results indicated that PTCH serves to repress Hh signaling, and by extension the genes activated by the Hh pathway. Without PTCH to repress SMO, the Hh pathway is continually active rather than activated at specific timepoints in development as normal. This aberrant activation can have deadly consequences. Mice with two copies of the mutated PTCH gene "died during embryogenesis and were found to have open and

overgrown neural tubes." Mice that had only one mutated copy of the gene and one functional copy grew to be abnormally large and some developed "hindlimb defects or cerebellar medulloblastomas [, malignant brain tumors]" (Goodrich, Milenković, Higgins, & Scott, 1997). These results imply that malfunction of Hh signaling results in severe birth defects and tumor proliferation.

Since inappropriate activation of the Hedgehog pathway has been identified as playing a role in promoting tumor growth in many types of cancers, treatments are being developed based on the inhibition of this pathway. While two drugs that inactivate the pathway by inhibiting SMO, Vismodegib and Erismodegib, have been approved by the US Food and Drug Administration for the treatment of basal cell carcinoma, a type of skin cancer, in humans, they have been found to have limited efficacy in some cases. Most drugs for Hh inhibition are still in the early stages of testing to determine their efficacy and safety (Rimkus, Carpenter, Qasem, Chan, & Lo, 2016).

Purmorphamine is a molecule developed by researchers to activate the Hedgehog signaling pathway by directly binding to and activating SMO (STEMCELL Technologies, n.d.). While the mechanisms by which this powerful molecule interacts with the Hedgehog pathway are understood, its effects on chromatin structure (i.e. nucleosome characteristics — since nucleosomes make up chromatin fibers) have not previously been examined. This study therefore intends to reveal the latent changes on the potential of a cell to express certain genes, as wrought by purmorphamine through modification of nucleosome characteristics.

Rationale

Past studies in epigenetics have suggested that changes in nucleosome distribution and

sensitivity to the enzyme MNase can confer altered biochemical potential to the cell, as the accessibility of individual genes/binding sites are increased or decreased (Vera et al., 2014). However, the effect of purmorphamine-mediated activation of the Hedgehog signaling pathway on this chromatin structure has never been studied, despite the profound influence of this pathway on human development. Studying how the sensitivity and distribution of nucleosomes are altered by activation of this pathway may provide a greater understanding of how cellular potential is affected by exposure to this powerful molecule, of the Hedgehog signaling pathway, and of epigenetic mechanisms as a whole. Therefore, this study uses an experimental approach in order to garner new insight into the nucleosomal changes wrought by exposure to purmorphamine, and the consequent activation of the influential Hedgehog pathway.

Methods

In order to investigate purmorphamine's effects on chromatin structure, treated and untreated cells were first harvested and crosslinked with formaldehyde. Then two different concentrations of MNase were used to cleave the DNA in between nucleosomes. The mononucleosomally protected DNA was then isolated. The isolated fragments were then to be sequenced and aligned to a reference genome (the set of all of an organism's DNA). Alignment to the human genome allows nucleosome sensitivities and positions to be mapped and analyzed. The original samples were not sequenced however. Instead, surrogate data was obtained from previously sequenced samples processed with the same procedure. The analysis of this surrogate data serves to show the kind of information that could theoretically be revealed when the original samples are sequenced.

Harvesting and Crosslinking Cells

This study utilizes GM12878 cells. The GM12878 cell line is derived from human B-lymphocytes, which are cells that play a key role in immune response. This cell line was selected because GM12878 cells are widely available and relatively easy to maintain, and the Hedgehog pathway has been shown to have a profound influence on B-lymphocytes. For example, researchers at Sapporo Medical University School of Medicine in Japan found that Hh signaling increases the production of B-lymphocytes (Kobune et al., 2008). Additionally, researchers from the Genomics Institute of the Novartis Research Foundation found that blocking the Hedgehog pathway inhibited growth or reduced the mass of tumors in mice with lymphoma, which is cancer that originates in lymphocytes (Dierks et al., 2007).

The GM12878 cells were treated with 1μM purmorphamine for 24 hours. Then 2 x 10^7 cells from both the purmorphamine and control groups were harvested and crosslinked in 1% formaldehyde solution. Crosslinking produces chemical bonds that secure the nucleosomes, and other proteins such as transcription factors, in place so that they do not shift along the DNA or relocate post-treatment. The efficacy of crosslinking with formaldehyde has been confirmed by research conducted at Yale University. This study crosslinked DNA with formaldehyde to identify the binding sites of the transcription factor Zeste. The resultant data was compared to an established method of crosslinking, exposure to UV light, and was found to be accurate by this standard (Toth & Biggin, 2000).

After the addition of formaldehyde, the samples were incubated at room temperature for 10 minutes, after which the crosslinking was stopped with 125 mM glycine. The cells were then pelleted by centrifugation for 10 minutes at 3000 rpm and re-suspended in 2 ml of nucleus isolation buffer, which contains Triton™ x 100 — a detergent that serves to break up the cellular

membrane while preserving the nuclear membrane.

MNase Digestion and Isolation of Mononucleosomally Protected DNA

In order to study purmorphamine's effects on nucleosomes' MNase sensitivity, 300 µl (~ 3 x 10^6 cells) from each treatment group were subjected to digestion by 300 units/mL MNase (the heavy digest) and incubated for 5 minutes at 37ºC. Then, an additional 300 µl from each treatment group were digested with 10 units/mL MNase (the light digest) and incubated for 5 minutes at 37ºC. The MNase reactions were stopped with 10 mM ethylenediaminetetraacetic acid (EDTA). EDTA binds to the calcium ions in the solution making them unavailable to MNase, which requires their presence in order to cleave DNA.

Proteins were then denatured — the term for when a protein unfolds, losing its characteristic shape — and the crosslinks were removed by the addition of a 1% sodium dodecyl sulfate solution and overnight incubation at 60ºC. A phenol-chloroform extraction was then performed on the de-crosslinked samples in order to isolate the DNA. The phenol-chloroform extraction is performed by adding phenol and chloroform, centrifuging the samples to separate the sample into layers (phases), and then removing the aqueous layer — which contains the DNA — to a new tube. This procedure is repeated until no middle layer is visible after centrifugation, indicating that the sample contains pure DNA. Next, ethanol is added, the samples are set on ice for 30 minutes, and then they are placed in the centrifuge for 10 minutes at 30 rpm so that the DNA forms a pellet at the bottom of the tube (Figure 4).

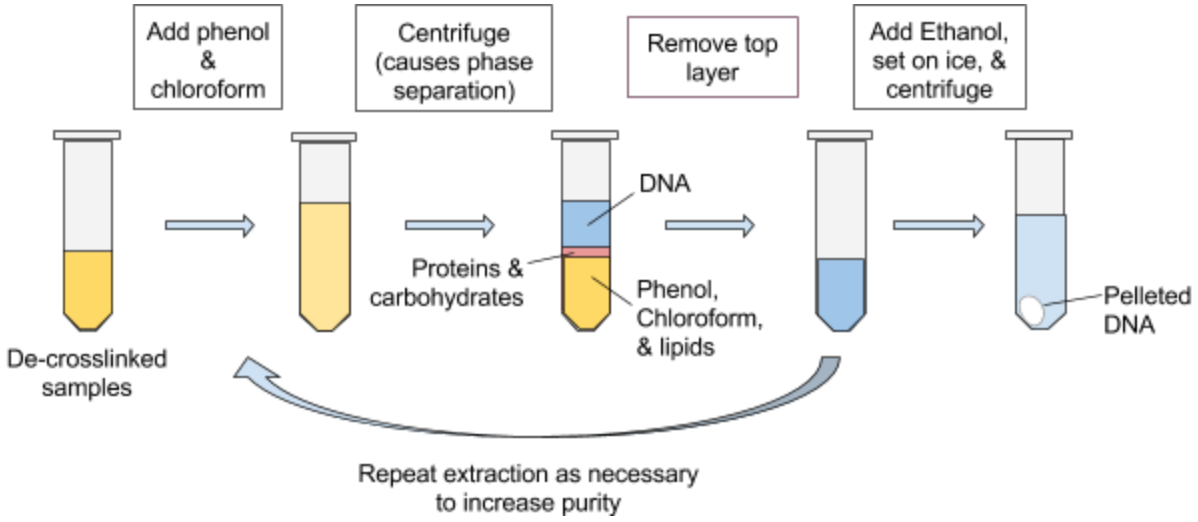

Figure 4: DNA Isolation

The leftover solution was vaporized via 7 minutes in a SpeedVac™ (a centrifugal evaporator), and the DNA was re-suspended in 100µl of Tris-EDTA, a buffer commonly used to store DNA. The populations of mononucleosomally protected DNA fragments produced by each digestion were then isolated using a PippinHT (a device used to separate DNA fragments based on size). For this experiment the PippinHT was set up to isolate DNA fragments between 90 and 250 bp, because nucleosomes cover roughly 150 bp of DNA. The mononucleosomally protected fragments were isolated so that they could later be aligned to the reference genome to map nucleosome characteristics.

Preparation for Sequencing

Before the sequence of the DNA is found, the fragments must be replicated repeatedly to increase the concentration of DNA in the samples. A greater concentration is desirable in that it will produce a stronger, clearer signal in the sequencer. Replication, however, requires that the fragments undergo a few modifications first. The samples were prepared for sequencing on a

Biomek® 4000 using the NEBNext® Ultra™ DNA Library Prep Kit for Illumina® (NEB #E7370S/L). The Biomek® 4000 is machine that automates the process, saving time and providing greater consistency in sample preparation. The Library Prep Kit adds DNA fragments with a known sequence (called adaptors), to both ends of each fragment already in the sample. Extra adaptors were then removed through the use of AMPure® XP beads. AMPure® XP beads can be attracted by a magnet and have a carboxyl coating that binds to DNA when polyethylene glycol (PEG) and salt are added to the solution. The amount of salt and PEG determines what size of DNA fragments bind to the beads (Gonzalez, 2016). A magnet is then used to attract the beads, which are bound to the DNA fragments of the desired size, while unwanted elements are washed away.

The samples are then heated up so that the double-stranded DNA fragments — which now have adaptors at each end — split into individual strands. The individual strands will be replicated by DNA polymerase, which is similar to RNA polymerase except the copy it produces is composed of DNA. However, DNA polymerase cannot begin copying the DNA unless at least a small fragment of DNA is already attached to the template strand (the DNA to be replicated) (Mason, Losos, Singer, Raven, & Johnson, 2011b). Therefore, fragments called primers, which have a sequence complementary to that the adaptors, are added to the solution. Primers bind to the adaptors on each DNA fragment so that DNA polymerase can then replicate the fragments in the samples. Additionally, the primers contain a short sequence called a barcode, which serves as an identifier for the DNA in a sample, so that multiple samples can be sequenced at once and still be easily distinguished from each other. The process of splitting the double stranded DNA into single strands, binding primers, and allowing DNA polymerase to replicate the fragments is

called a polymerase chain reaction (PCR) (Figure 5). PCR was performed for 8 cycles and then the samples were cleaned up using AMPure® XP beads to remove extra elements, such as primers. Each PCR cycle doubles the number of DNA fragments in solution and so the number of fragments increases exponentially.

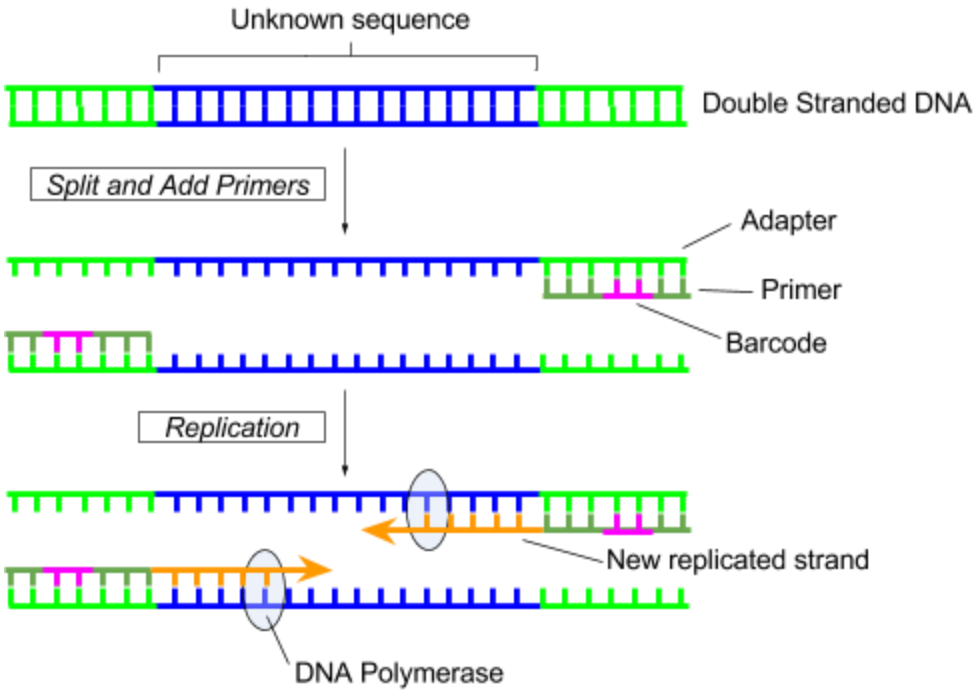

Figure 5. Polymerase Chain Reaction

A Change of Plans

The isolated and replicated mononucleosomally-protected DNA fragments were then to be sequenced on an Illumina® HiSeq™ and aligned to the human genome. Arrangements were made to have the samples sequenced at Florida State University, however, an unforeseen lack of human and financial resources forced the university to withdraw from this commitment late in the experimental process. Instead, the Dennis Lab provided surrogate data from samples

sequenced in a similar, previous experiment. That experiment used the same procedures as this one to compare the epigenetic differences between cancerous and noncancerous breast epithelial tissue (the outer layer of skin). The conclusions drawn from this data will not be identical to those that would have been reached through data from the original samples. However, analysis of the surrogate data is intended to illustrate the kind of information that could be provided by data from the original samples.

Additionally, the RNA present in the original samples before MNase digestion was to be sequenced to provide data about the expression level of genes. Transcription produces RNA copies of genes and so sequencing RNA can indicate which genes are being expressed, a useful measure to correlate with any observations of nucleosome characteristics. The RNA expression data was to be provided by an experiment conducted in parallel to this one by an outside source, due to time constraints and the level of expertise required for RNA processing. However, because the original samples were not sequenced, the decision was made to halt progress on this experiment as well. Instead surrogate RNA expression data was obtained from the similar, previous experiment which also provided the data on nucleosome characteristics.

Data Visualization

In each tissue, the different concentrations of MNase produced different populations of DNA fragments, which were then sequenced and aligned to the human genome. The results from the two digest conditions can be combined to produce a line along the genome, the height of which indicates the number of fragments produced at that point in the genome. Because the DNA fragments were protected from MNase digestion by individual nucleosomes, peaks in the line indicate nucleosome position.

Fragment alignment can also reveal nucleosome sensitivity, however, as the line produced by the heavy digest can be subtracted from that of the light digest to suggest the sensitivity of nucleosomes in individual locations. Fragments released primarily under heavy digest conditions are known as "resistant" to MNase digestion. "Sensitive" nucleosomes, on the other hand, are easily released by MNase, and so these mononucleosomally protected fragments are found primarily in the light digest — these fragments become <150 bp in a heavy digest (Figure 6).

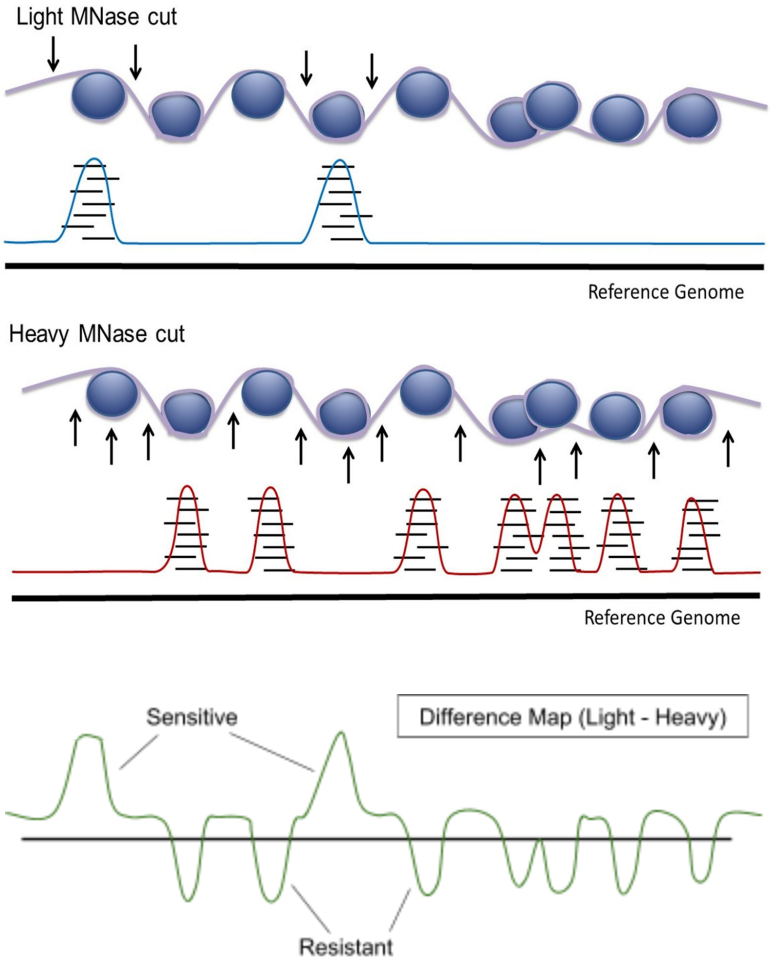

Figure 6. Aligning Fragments to the Reference Genome Reveals Nucleosome Position and Sensitivity.

Rather than mapping nucleosome characteristics at every point in the human genome, this study focuses on nucleosomes near TSSs, as these nucleosomes have the most influence on transcription. A heatmap, a graphic that displays numerical data as color, can be utilized to visualize nucleosome information for all TSS regions at once. The height of the aforementioned nucleosome position or sensitivity line is indicated by color on the heatmap, and the heatmaps organize the TSS regions into clusters with similar results. The heatmaps produced for this study show nucleosome position and sensitivity for a window of ±1 kb (kilobases) from each TSS. The TSS regions of the genome are stacked vertically such that each TSS falls along a vertical line through the middle of the heatmap. A window of 2 kb should be sufficient for this study as the core promoter region of a gene only lays around 30 bp upstream of the TSS and other important promoter elements (promoter-proximal elements) are found "within 100 to 200 bp of the [TSS]" One kb greatly exceeds this 100 to 200 bp estimate, enabling some data to be obtained about regions located much farther from the TSS. Regions called enhancers and silencers can also promote or inhibit transcription of particular genes, but unlike promoters, these regions can be located far from the TSS, "sometimes 50 kb or more" (Griffiths, et al., 2000). A window of ±1 kb could enable collection of data about some of these regions, but a thorough analysis of enhancer and silencers is not feasible for this study due to the vast number of bases that would have to be sequenced and the nonuniform distribution of such regions. Additionally, data about nucleosome characteristics farther than 1 kb from a TSS may not be very useful. Nucleosomes beyond 1 kb from the TSS show increasingly random distributions, potentially indicating a loss of the functional significance of nucleosome position at such distances from TSSs (Jiang & Pugh, 2009).

Results

Figure 7 contains two nucleosome sensitivity heatmaps produced from the surrogate data, the first of which displays nucleosome sensitivities for the noncancerous tissue (MCF10A). The second heatmap displays nucleosome sensitivities for the cancerous tissue (MCF10CA1a). Each heatmap is accompanied by a separate heatmap indicating the level of expression for each gene. Red indicates higher levels of expression. A boxplot also accompanies the heatmaps and indicates the expression levels for each cluster of genes in the heatmaps (labeled 1-7).

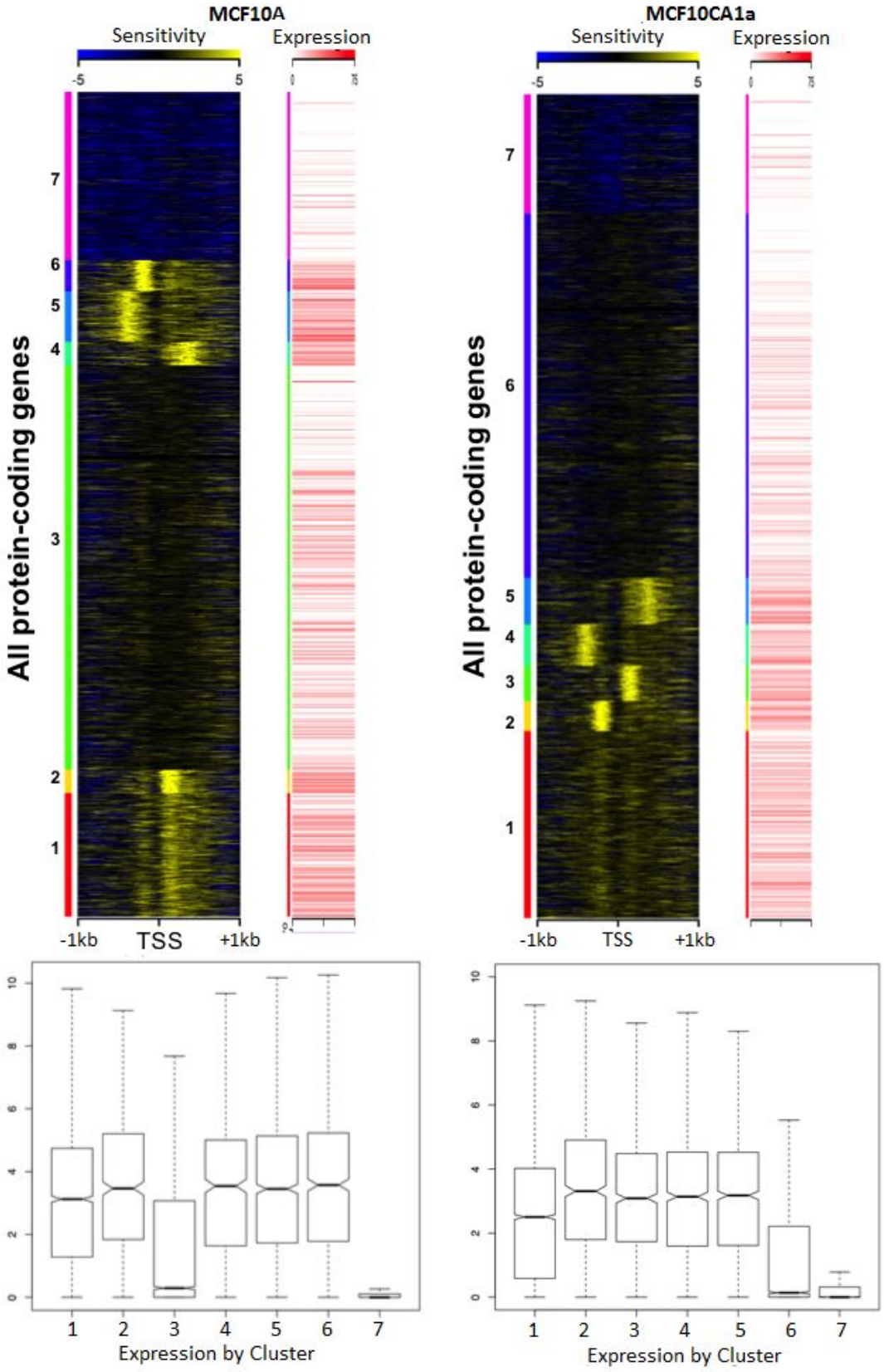
MCF10A
Sensitivity
Expression
-5
5
All protein-coding genes
-1kb
TSS
+1kb
MCF10CA1a
Sensitivity
Expression
-5
5
All protein-coding genes
-1kb
TSS
+1kb
Expression by Cluster
Expression by Cluster

Figure 7. Nucleosome Sensitivity Heatmaps. Yellow indicates nucleosome sensitivity whereas blue indicates resistivity.

As indicated by the boxplot under the first heatmap, the most expressed genes in the noncancerous tissue were grouped into cluster 6, though by an extremely narrow margin. The least expressed genes in the noncancerous tissue were in cluster 7, with very few RNA transcripts detected. In the cancerous tissue, the most expressed genes were grouped into cluster 2, while the least expressed genes were in cluster 7. The relative expression levels indicated by the boxplots are supported by the red and white expression heatmaps, which show less expression (white) in cluster 7 for both tissues, and more expression (red) in clusters 6 and 2 for the noncancerous and cancerous tissues respectively.

For both tissues, the cluster with the greatest expression level (6 for noncancerous and 2 for cancerous tissue) has a strongly positioned, sensitive -1 nucleosome – as indicated by the prominent yellow bar just left of the TSS. The clusters with least expressed genes (7 for both tissues) also have the greatest level of nucleosome resistivity to MNase (indicated in blue).

Figure 8 contains a heatmap of the overall nucleosome positions for the noncancerous tissue. The accompanying boxplot and expression heatmap reveals some correlations between nucleosome occupancy and expression levels. The genes with the highest level of expression, clusters 3-7, also had strongly positioned nucleosomes – indicated by prominent yellow bars. Of these clusters, those with strongly positioned nucleosomes upstream of the TSS, clusters 4 and 6, had slightly higher levels of expression. The genes with the lowest expression levels, cluster 1, had the fewest nucleosomes.

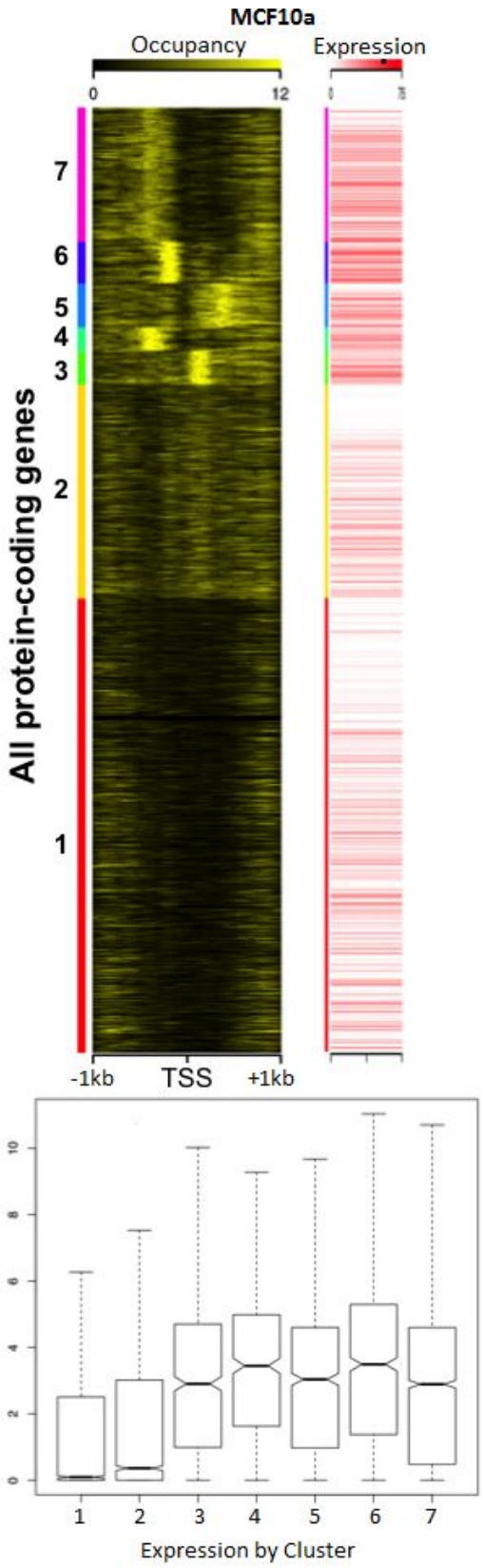
MCF10a
Occupancy
Expression
0
12
All protein-coding genes
7
6
5
4
3
2
1
-1kb
TSS
+1kb
Expression by Cluster

Figure 8. Nucleosome Occupancy. Yellow indicates nucleosome occupancy.

The tissues were then compared using Venny, a venn diagram program, and GOrilla, a tool for comparing the functions of groups of genes. GOrilla results indicated that the most expressed clusters of genes in both tissues were associated with the same functions. The genes in the least expressed clusters from Figure 7 were then compared for both tissues. The results revealed a difference between the functions of the genes in these clusters for the two tissues; the least expressed genes in the cancerous tissue, unlike those in the noncancerous tissue, shared a functional affiliation with cytokine activity.

Discussion

The expression and nucleosome position data provided in Figure 8 illustrate an association between nucleosome positioning and gene expression. Clusters with greater nucleosome occupancy, especially upstream of the TSS, showed the greatest expression levels. These findings run counter to the predicted results and conflict with those of previous studies which have found decreased nucleosome occupancy in the promoter region of actively-expressed genes.

The expression and sensitivity data provided in Figure 7 illustrate an association between nucleosome sensitivity and gene expression. Clusters with greater MNase sensitivity, especially at the -1 nucleosome, showed the greatest expression levels, while clusters with resistant nucleosomes showed the lowest expression levels. These results agree with the current understanding of MNase sensitivity as an indicator of DNA accessibility.

Modifications to chromatin structure confer altered cellular potential for gene expression,

and so it was predicted that the epigenetic differences in the cancerous tissue, as compared to the noncancerous tissue, would correspond to enhanced survival of the cancerous cells. For instance, increased nucleosome sensitivity near the transcription start site (TSSs) of genes that promote tumor growth or limit immune response, would suggest greater ease of nucleosome repositioning/modification, thus facilitating access to those gene segments, transcription, and ultimately cancer cell survival and proliferation. Therefore, it would be expected that the cancerous tissue might show increased nucleosomal sensitivity near the TSSs of such genes, thereby increasing the cell's potential to express those genes. Conversely, genes which promote immune response or cause cell death are expected to have greater resistance to MNase digestion as such conditions would decrease the cell's potential to express those genes, thereby reducing the threat to cancer cell survival. These predictions were confirmed by an analysis of correlations between the level of expression and the function of genes in the cancerous tissue.

The most expressed genes in the two tissues were associated with the same functions. The cancerous tissue however, showed decreased expression of cytokine-related genes relative to the noncancerous tissue. Cytokines are a group of proteins involved in cellular communication and interaction during immune response. Types of cytokines include chemokines, interferons, interleukins, lymphokines, tumor necrosis factor, and arguably some growth factors (Sino Biological, n.d.). The relative repression of cytokine activity related genes in the cancerous tissue is explained well by the function of those genes.

For example, one group of cytokine-related genes repressed in the cancerous tissue was interferons. Alpha interferons, are a group classified as Type 1 interferons (M. Hunt & R. Hunt, 2014). Type 1 interferons have been found to induce apoptosis, cell death, in tumors and inhibit

angiogenesis, the formation of new blood vessels, thereby cutting off nourishment to tumors (Hervas-Stubbs et al., 2011). Interferons thus can retard the proliferation of tumors, leading them to be used somewhat successfully as a cancer therapy. The repression of cytokine-related genes in the cancerous tissue is a stunning demonstration of cancer's ability to alter cellular potential for its own benefit. The repression of the cytokine-related genes observed in this study, serves to promote the survival and proliferation of cancerous cells by hindering the proper function of immune defense mechanisms.

Other cytokines repressed in the cancerous tissue included tumor necrosis factor superfamily members 10 and 8 (TNFSF10 & TNFSF8), both of which have been identified as promoting apoptosis. As noted by a study from the University of North Carolina School of Medicine "[e]pigenetic silencing of these genes (directly or indirectly) would convey a survival advantage to the cancer cells through elimination of pro-apoptotic signaling. TNFSF10...and TNFRSF8 have been shown to be subject to [epigenetic]...silencing in cancer cells" (Hunter, Edson, & Coleman, 2014). A mechanism commonly associated with epigenetic silencing is the methylation of nucleosomes in promoters. Methylation is the addition of the chemical group, methyl, and is theorized to promote a more condensed chromatin state, decreasing access to promoter regions. The results of this study showed that TNFSF10 and TNFSF8 were less expressed in the cancerous tissue, and displayed increased nucleosome resistivity. These findings therefore confirm that nucleosome resistivity can provide a useful measure of epigenetic silencing and may correspond to nucleosome methylation.

While the methods used in this experiment were chosen deliberately to optimize results, one limit to the analysis conducted in this study is GC content bias. DNA contains four different

kinds of bases; adenine (A), guanine (G), cytosine (C), and thymine (T). As noted in a study from researchers of the Berlin-based Max Planck Institute for Molecular Genetics, MNase tends to preferentially cleave regions of DNA with higher GC content, thus producing relatively higher fragment counts during sequencing (Chung et al., 2010). To account for this bias, some studies opt to normalize fragment counts based on relative GC content. However, due to limits in time and processing ability, this study does not normalize based on GC content.

Future Direction and Conclusion

The analysis of cancerous and noncancerous breast tissue supported the efficacy of MNase as a measure of DNA accessibility and provided evidence against the claim that nucleosome occupancy in a promoter obstructs transcription. Furthermore, results from this study demonstrated the decreased expression of cytokine-related genes in cancerous breast tissue and correlated this decreased expression with nucleosome resistivity near the TSS of the associated genes. These results provide evidence to support the idea that cancer is marked by nucleosomal changes that alter gene expression to inhibit immune function.

An analysis of results from the original samples could similarly provide evidence indicating correlations between changes in nucleosome characteristics and gene regulation in response to activation of the Hedgehog pathway. Knowledge of the implications of Hedgehog activation could provide valuable information about the pathway's role in tumor growth and cellular development, ultimately aiding in the development of novel therapeutic and diagnostic procedures. Future investigation could sequence the purmorphamine-treated and control samples as originally intended. Additionally, future studies could explore the mechanisms by which specific changes in nucleosome sensitivity and position occur.

Acknowledgements

Special thanks to the Young Scholars Program for giving me this amazing research opportunity, to Florida State University for the generous contribution of their lab facilities and resources, to Dr. Jonathan Dennis for his mentorship and instruction in wet-lab techniques, and to Lauren Cole for her invaluable guidance, support, and surrogate data.

References

Chung H-R, Dunkel I, Heise F, Linke C, Krobitsch S, Ehrenhofer-Murray AE, et al. (2010) The effect of micrococcal nuclease digestion on nucleosome positioning data. PLoS ONE 5(12): e15754. doi:10.1371/journal.pone.0015754

Corcoran RB, Scott MP. Oxysterols stimulate Sonic Hedgehog signal transduction and proliferation of medulloblastoma cells. *Proceedings of the National Academy of Sciences of the United States of America*. 2006;103(22):8408-8413. doi:10.1073/pnas.0602852103.

Dierks, C., Grbic, J., Zirlik, K., Beigi, R., Englund, N. P., Guo, G., . . . Warmuth, M. (2007). Essential role of stromally induced Hedgehog signaling in B-cell malignancies. *Nature Medicine, 13*(8), 944-951. doi:10.1038/nm1614

Gonzalez, O. G. (2016). *Mononucleosomal fragment retrieval and what it entails about chromatin structure and its relation with genome response* (Unpublished master's thesis). Florida State University. Retrieved from diginole.lib.fsu.edu/islandora/object/fsu%3A309378

Goodrich, L. V., Milenković, L., Higgins, K. M., & Scott, M. P. (1997). Altered neural cell fates and medulloblastoma in mouse patched mutants [Abstract]. *Science, 277*(5329), 1109-1113. doi:10.1126/science.277.5329.1109

Griffiths, Miller, Suzuki, et al. (2000). Transcription: An overview of gene regulation in eukaryotes. In *An Introduction to Genetic Analysis* (7th ed.). New York, NY: W. H. Freeman.

Hervas-Stubbs, S., Perez-Gracia, J. L., Rouzaut, A., Sanmamed, M. F., Bon, A. L., & Melero, I.

(2011). Direct effects of Type I interferons on cells of the immune system. *Clinical Cancer Research, 17*(9), 2619-2627. doi:10.1158/1078-0432.ccr-10-1114

Hunter, D., Edson, L., & Coleman, W. (2014). Loss of tumor necrosis factor superfamily genes in breast cancer cell lines (1047.8). *The FASEB Journal, 28*(1).

Hunt, Margaret, and Richard Hunt. "Interferon." *Microbiology and Immunology On-line*. University of South Carolina School of Medicine, 8 Jan. 2014. Web. 06 Apr. 2017.

Jiang, C., & Pugh, B. F. (2009). Nucleosome positioning and gene regulation: advances through genomics. *Nature Reviews. Genetics*, *10*(3), 161–172. http://doi.org/10.1038/nrg2522

Kiecker, C., Graham, A., & Logan, M. (2016). Differential cellular responses to Hedgehog signaling in vertebrates—What is the role of competence? *Journal of Developmental Biology*, *4*(4), 36. doi:10.3390/jdb4040036

Kobune, M., Kato, J., Kawano, Y., Sasaki, K., Uchida, H., Takada, K., . . . Niitsu, Y. (2008). Adenoviral vector-mediated transfer of the Indian Hedgehog gene modulates lymphomyelopoiesis in vivo. *Stem Cells, 26*(2), 534-542. doi:10.1634/stemcells.2007-0741

Lee, C., Shibata, Y., Rao, B., Strahl, B. D., & Lieb, J. D. (2004). Evidence for nucleosome depletion at active regulatory regions genome-wide [Abstract]. *Nature Genetics, 36*(8), 900-905. doi:10.1038/ng1400

Luger, K., Mäder, A., Richmond, R., Sargent, D., & Richmond, T. (1997). Crystal structure of the nucleosome core particle at 2.8 Å resolution [Abstract]. *Nature, 389*(6648), 251-260. doi:10.1038/38444

Mason, K. A., Losos, J. B., Singer, S. R., Raven, P. H., & Johnson, G. B. (2011a). Cell

communication. In *Biology* (9th ed., pp. 168-183). New York, NY: McGraw-Hill.

Mason, K. A., Losos, J. B., Singer, S. R., Raven, P. H., & Johnson, G. B. (2011b). DNA: The genetic material. In *Biology* (9th ed., pp. 256-275). New York, NY: McGraw-Hill.

Mellor, J. (2005). The dynamics of chromatin remodeling at promoters. *Molecular Cell, 19*(2), 147-157. doi:10.1016/j.molcel.2005.06.023

Mieczkowski, J., Cook, A., Bowman, S. K., Mueller, B., Alver, B. H., Kundu, S., . . . Tolstorukov, M. Y. (2016). MNase titration reveals differences between nucleosome occupancy and chromatin accessibility. *Nature Communications, 7*, 11485. doi:10.1038/ncomms11485

Nature Education. (n.d.-a). Transcription / DNA transcription. Retrieved from http://www.nature.com/scitable/definition/transcription-dna-transcription-87

Nature Education. (n.d.-b). Transcription factor / transcription factors. Retrieved March 28, 2017, from http://www.nature.com/scitable/definition/general-transcription-factor-transcription-factor-167

Phillips, T. & Shaw, K. (2008) Chromatin remodeling in eukaryotes. Nature Education 1(1):209

Rimkus, T. K., Carpenter, R. L., Qasem, S., Chan, M., & Lo, H.-W. (2016). Targeting the Sonic Hedgehog signaling pathway: Review of Smoothened and GLI inhibitors. *Cancers*, *8*(2), 22. http://doi.org/10.3390/cancers8020022

Schones, D. E., Cui, K., Cuddapah, S., Roh, T., Barski, A., Wang, Z., . . . Zhao, K. (2008). Dynamic regulation of nucleosome positioning in the human genome. *Cell, 132*(5), 887-898. doi:10.1016/j.cell.2008.02.022

Sino Biological. (n.d.). Types of cytokines. Retrieved from http://www.sinobiological.com/Types-of-cytokines.html

STEMCELL Technologies. (n.d.). Purmorphamine. Retrieved from https://www.stemcell.com/purmorphamine.html#section-product-documents

Toth, J., & Biggin, M. D. (2000). The specificity of protein–DNA crosslinking by formaldehyde: *in vitro* and in *Drosophila* embryos. *Nucleic Acids Research*, *28*(2), e4.

Vera, D. L., Madzima, T. F., Labonne, J. D., Alam, M. P., Hoffman, G. G., Girimurugan, S. B., . . . Bass, H. W. (2014, October 31). Differential nuclease sensitivity profiling of chromatin reveals biochemical footprints coupled to gene expression and functional DNA elements in maize. *The Plant Cell, 26*(10), 3883-3893. doi:10.1105/tpc.114.130609