Generation of Reporter Constructs to Characterize the Role of Histone H3 and H4 Residues in cell-to-cell Variability in Gene Expression

A Major Qualifying Project Report: Submitted to the Faculty of the

WORCESTER POLYTECHNIC INSTITUTE

in partial fulfillment of the requirements for the Degree of Bachelor of Science By

Jared G. Curtin Date:

Approved by:

Oliver Rando, Ph. D Dept. of Biochemistry and Molec. Pharmacology UMass Medical School Major Advisor

Destin Heilman, Ph. D Dept. of Chemistry and Biochemistry WPI Project Advisor

Abstract:

Histones are strongly basic proteins responsible for the packaging and ordering of DNA inside eukaryotic cells. They are highly modifiable multifunctional proteins. 147 DNA base pairs are wound around an octamer of core histone proteins (H3, H4, H2A, H2B). Post-translational modifications allow enormous potential for functional responses. Little is known about the effects of histone mutations on cell-to-cell variability of gene expression. To characterize these mutations, yeast strains with fluorescent reporter constructs were generated. Yellow fluorescent protein (Venus) was fused to a gene of interest, and red fluorescent protein (Cherry) was fused to a gene with no ties to histone modification in gene expression. Expression of the fluorescent reporter constructs in a comprehensive library of histone H3, and H4 mutants will be used to characterize cell-to-cell variability. Through quantitative single-cell measurements of protein abundance using fluorescence microscopy speculations into the roles of histone modifications on gene expression can be made.

Acknowledgements:

I'd like to thank Professor Oliver Rando at UMass Medical School for giving me the change to work in his lab. I would also like to Sarah Swygert who worked with me during my time in the Rando lab.

Table of Content:

2	Abstract
3	Acknowledgements
5-17	Introduction
18-21	Materials and Methods
22-27	Results
28-30	Discussion
37-38	References
5	Figure 1
7	Figure 2
10	Figure 3
15	Figure 4
16	Figure 5
17	Figure 6
31	Figure 7
32	Figure 8
33	Figure 9
34	Figure 10
45	Figure 11
46	Figure 12

Introduction:

The human genome is an enormous molecule consisting of 3.2 billion-bp that together have a total contour length of 1 meter. Inside the cell DNA is condensed into a form 40,000 times shorter than when unpackaged. The DNA molecule is a two-stranded polymer of deoxynucleotides (one coding for the genetic inheritance and its compliment stand) held together by phosphodiester bonds between monobasic phosphate groups and the deoxyribose sugars. Hydrogen bonds hold the base pairs together (A to T and G to C), and orient them into layers. The stacking interactions of the nucleic acids hold the two DNA stands together to form a double helix, as seen in figure 1.



Figure 1: Diagram of the Structure of DNA. Diagram shows the pairing of bases (colored rungs on the ladder) and the overall helical structure of DNA. The deoxyribose sugars are bonded sideways through phosphate groups, and the opposite stand is bonded through base pairing. [22] Two grooves adjacent to the nucleotides run in between the matching helical strands of the DNA backbone. The two strands of the double helix are not directly opposite each other causing the grooves to be unevenly spaced. The major groove is 22 Å wide and the other, the minor groove, is 12 Å wide [21].

DNA can adopt different conformations. The most biologically abundant form is B-DNA. The molecules are flexible and can adopt many different degrees of curvature effectively altering the topology of the DNA. Altering the number of twists around the helical axis add strains to the DNA causing it to writhe at a superhelical axis. The contortions force the DNA to wrap around itself in response to the added torsion. Extra helical twist hold the nucleotides to more tightly together and leads to positive supercoiling, while removing twist leads to negative supercoiling. Most DNA has slightly negative supercoiling to facilitate easier access to the genome in order for vital biological processes to occur. Topoisomerases are enzymes that can alter the topological state of DNA to control transcription and translation, or alleviate strain on the molecule. In eukaryotes additional method for DNA compression are necessary.

The next order of packaging is the nucleosome. DNA is the is wound around the nucleosome in what is commonly referred to as the "beads on a string" structure; packages of approximately 147 DNA base pairs wound around an octamer of histone proteins. The core histone proteins, H3, H4, H2A, and H2B are the fundamental units of DNA compaction inside the cell. The nucleosome beads are connected with linker DNA that is approximately 20 to 60 base pairs long. Linker histones: H1 and H5 also play a role in nucleosome compaction by interacting with the Linker DNA to form higher ordered structures. Histones are some of the most highly conserved proteins in eukaryotes, and have vital roles in nucleosome construction as well as higher order chromatin structure and gene regulation. Histones are highly modifiable multifunctional proteins. They can be made to have a range of different functions inside the cell through

post-translational modifications. Different modifications will give rise to differences in cell-to-cell variability in gene expression.

Histones are strongly alkaline proteins. The interactions of DNA with histone proteins can be characterized in five distinct ways. First H2B, H3, and H4, positively charged helix-dipoles from alpha helices interact with negatively charged phosphate groups on DNA. Second hydrogen bonds form between the amide group on the main chain of histone proteins and the DNA backbone. Third the nonpolar deoxyribose sugars of the DNA interact with the histone proteins (Serine and Threonine). Fourth, there are salt links and hydrogen bonds between side chains of basic amino acids (especially lysine and arginine) and phosphate oxygens on DNA. Lastly non-specific minor groove insertions of the H3 and H2B N-terminal tails into two minor grooves each on the DNA molecule. Each interaction determines how histone proteins will package and order the DNA into the structural unit of the nucleosome.





DNA nonspecifically binds to the nucleosome although in large sequences there are preferences that govern nucleosome positioning. This is though to be due to watermediated interactions between the histone proteins and the DNA. Adenine and Thymine regions of the DNA are more favorably compressed into the inner minor grooves. When bound to the nucleosome DNA is rotated in 1.67 negative superhelical turns to maximize compression. DNA is distorted within the nucleosome core. Twist defects and nonuniform bends are created through the distribution and strength of DNA-histone interactions. The overall twist of nucleosomal DNA varies from a value of 9.4 to 10.9 bp per turn. In vivo, DNA positioning can be found in the same location, which suggests that there are histone-loading complexes. Nucleosomes are mobile along the DNA molecule, ATP-dependent chromatin remodeling enzymes have been identified that are able to translationally relocate nucleosomes onto adjacent sequences, disrupting the histone-DNA contacts, and generating negative superhelical stress on chromatin and DNA.

Chromatosomes are a higher order of DNA compaction. They are a combination of the nucleosome and linker histones. Under non-physiological conditions, the nucleosome along with 20 to 60 base pairs of linker DNA can form the 10nm "beads on a string structure". With linker histones, especially H1, which contacts the DNA entering and exiting the nucleosome, the " beads on a string" structure becomes a 30nm dihelical structure. These groups of nucleoprotein complexes are called chromatin.

In its compact form, chromatin is referred to as a 30nm fiber. Inside the cell it is found in two forms as heterochromatin (compact) and euchromatin (extended). Currently the accepted model for the structure of chromatin is the nucleosomes lying perpendicular to the axis of the fiber with the linker histones arranged internally. The stability of the

30nm fiber relies heavily on the regular positioning of nucleosomes on DNA. Linker DNA is also critical to the fiber's stability because it resists bending and rotation. The nucleosome requires separations that permit rotation and folding into required orientations without excessive strain on the DNA. Recently using electron microscopy, it was proven that different lengths of Linker DNA give rise to different topologies of the 30nm fiber [14]

During the life cycle of the cell chromatin compaction varies greatly. In interphase, most of the chromatin is loosely condensed in the form euchromatin. It is the phase the cell spends the majority of its time in, replicating DNA for cellular division and expressing its genome. Euchromatin in interphase is organized into large loops found mostly in its 30nm form, containing approximately 50 to 100 Kb of DNA. Approximately 10% of the euchromatin, is in a more decondensed state. It unfolds into the 10nm "beads on a sting" structure when traversed by RNA polymerase engaged in transcription. Gene expression in eukaryotes is closely linked to chromatin structure[1].

The structure of chromatin and its layout of the genome inside the nucleus are not random. Evidence supports certain genes within the nucleus associate together. "In chromosomes with aggregated gene clusters, gene deserts preferentially align with the nuclear periphery, providing evidence for chromosomal region architecture by specific associations with functional nuclear domains" (Shopland, Lynch). This allows distant genes to physically interact and mediate expression and other chromosomal functions. The layout of the genome however is not well characterized.



Figure 3: Chromatin Structure- illustrates the structure of the 30nm fiber, how the condensation of the "beads on a string " structure of nucleosomes becomes chromatin. [5]

DNA is packaged into chromatin to fit inside the nucleus of the cell. Chromatin plays key roles in many cell processes like the transcrition, DNA replication, and DNA repair. Modifications to the structure of chromatin affect DNA packaging. Specific enzymes complexes can be recruited to manipulate the DNA, and effect gene regulation and expression. Modifications to chromatin are due to chemical changes in the histone proteins. By making covalent modifications to the histones, certain regions of the genome can be turned on or off. These modifications include methylation, acetylation, and phosphorylation; as well as many others.

Histones can be modified at many sites. There are over 60 different residues where modifications have been observed. However this is a large understatement to the number of actual modifications. Amino acids like lysine and arginine have shown up to three different forms of methylation. This allows enormous potentials for functional responses [4]These responses either activate or silence transcription by the establishment of global chromatin environments or the orchestration of DNA based biological tasks. Recently bivalent modifications (possessing both activating and repressing modifications) have been detected. It changes how we look at modifications, and adds a new element of complexity. Repressing modification usually take precedence over activating ones. The manifestation of modifications depends on the signaling conditions from the cell.

Histone modifications are still an area of active study. Lysine acetylation has been studied longest. Lysine is positively charged and interacts with the negatively charged DNA, binding it tightly. Acetylation removes the positive charge of lysine and has the largest effect on chromatin unraveling compared to other modifications. This allows transcription mechanisms to come in and interact with the DNA. The formation of the 30nm fiber and higher order structures is prevented by acetylation of H4K16. Acetyltransferases usually modify multiple lysine residues. Actively transcribed genes are almost always associated with high level of acetylation. Typically acetylation is observed at the N-terminal tail of histones, although the Rtt109 enzyme acetylates the H3K56 core domain and is thought to be associated with genome stability and DNA replication. Acetylation occurs at active sites on the histone, and orchestrates DNA based process inside the cell. Less is known about lysine methylation. Lysine methyltransferases have a greater specificity than acetyltransferases, and usually only methylate one lysine residue on a single histone. An emerging view is that methylation recruits protein complexes to aid in activation translation. It is also believed to establish global chromatin environments that can silence the chromatin.

Lysine methylation can have a range of different functions inside the cell. Methylated H3K4, and H3K36, are involved in activation of transcription. Both H3K4me and H3K36me are important in transcriptional elongation; H3K36me also recruits an

EAF3 protein that brings Rpd35 deacetylase to prevent transcription at inappropriate start sites. Methylated H3K4 and H3K36 are typically accompanied with acetylation Conversely methylated H3K9, and H3K27 are involved in repression. H3K9me silences active DNA, repression involves the recruitment of methylating enzymes and HP1. H3K9me, and HP1 recruitment are always repressive. Bivalent modifications of H3K27 and H3K4 have opposite functions, causing low levels of gene expression. Methylation H3K20 is involved in gene repression. The Lysine residue can be single, double, or triple methylated, but it looses its abilities as a repressor when triply methylated. It is believed that methylation also plays a large role in development because more is viewed in embryonic cells than differentiated ones.

Little is known about the role of phosphorylation in gene expression. Phosphorylated H3S10 has been demonstrated in the activation of NFKB- regulated genes. Modifications on histones allow accessibility to and recognition of the damaged DNA. Phosphorylation of the histone occurs for many Kb around the site of damage and in one of the earliest signs seen. It recruits complexes to the phosphorylation sites to repair breaks in the DNA. Two important phosphorylation events occur in the replication cycle of mammalian cells. The first is during mitosis Aurora B kinase phosphorylates H3S10, and simultaneously displaces HP1 from H3K9me (normally responsible for chromatin compaction). The second is phosphorylation of H3T3 by the Haspin kinase. Metaphase chromosomal alignment would not be possible without this process.

Chromatin modifications are very important to our understating of gene regulation and function. The development of a global chromatin environment, helps to divide the genome into functional groups; the accessible active DNA-protein complex

(euchromatin) that can be used in transcription, and the inaccessible inactive DNAprotein complex (heterochromatin) that is not used in transcription. The second group consists of silenced transcriptional DNA that has silenced through methods like histone methylation, and structural DNA. It plays a major role in specific jobs like DNA replication, and gene expression. Euchromatin is a large portion of the genome, is usually accompanied by high levels of acetylation and methylation that allow the nucleosomes to unravel and participate it transcription. Heterochromatin is accompanied by low levels of acetylation, methylation, and phosphorylation. Due to its tight packing, heterochromatin prevents degradation by proteins and serves to protect the integrity of the chromosomes during mitosis.

DNA is responsible for passing on genetic material and phenotypic traits over time. Histone modifications can also affect the phenotype and gene expression of the cell, giving it different characteristics than a cell with an identical set of DNA. These characteristics differ greatly from genetic inheritance. They can be gained or lost many multiples of times faster than genetic mutation.

Many fundamental biological pathways are influenced by histone modifications. There are over a hundred different possible covalent modifications that appear on a histone. Modifications tend to occur simultaneously and lead us to contemplate the roles of these modifications in the control of transcriptional noise when looking at gene expression in bulk cultures. Despite all the research in this field, the effects of these modifications are still widely unknown.

Typically studies on chromatin are carried out on bulk cultures. In an effort understand the processes behind chromatin modifications and their effects on

transcription levels are studied and quantized in large cell cultures with characterized modifications and compared to those without modifications. This illustrates differences in levels of gene expression and induction time but withholds much of the information about the inner workings of the cell. Specifically, the variation between cells in a population; this information is lost in bulk cultures. A more in depth approach must be taken to truly understand the processes of the cell at a chemical level. In an attempt to better understand the effects of histone variations, a closer look has to be taken at individual cells. In fact, chromatin packing has been implicated as a major player in cell-to-cell variation in gene expression of yeast.



Figure 4: Fluorophore Structure- (A) Venus is a yellow fluorescent protein, consisting of a beta barrel with an alpha helix running down the middle containing the chromophore. The protein molecule is 239 peptides in length. Venus has an excitation wavelength at 515nm and an emission wavelength at 528nm. Venus will be used to tag a gene of interest. [14](B) Cherry is a red fluorescent Protein consisting of a beta barrel with an alpha helix running down the middle containing the chromophore. The protein molecule is 236 peptides in length. Cherry has an excitation wavelength at 587nm and an emission wavelength at 610nm. Cherry is used as the tag for a control gene. [18]

Fluorescence is generally used as a non-destructive means for tracking biological molecules. Fluorescence is the emission of light in the visible spectrum by a molecule or atom that has absorbed a different wavelength of light. George Gabriel Strokes coined the term in 1852 from observations of mineral fluorite. In most cases, higher energy light is absorbed by a molecule or atom. The energetically excited structure emits a photon of a larger wavelength as it relaxes back to its ground state. Few cellular components are naturally fluorescent (also known as intrinsically fluorescent). Many cellular components

can be marked extrinsically with a fluorophore (a fluorescent molecule like a protein, or small molecule). Typically the structure consists of a beta barrel structure with an alpha helix running through the center, which contains the chromophore (the molecule responsible for it's color). Fluorescence is rare in many molecules. Fluorescence spectroscopy, unlike other forms of spectroscopy, is extremely sensitive and measurements can be detected down to a single molecule.

Yeast containing the BY4734 genetic background will be used as the model organism to study cellular behavior. Fluorescent reporter constructs will be generated to tag a housekeeping gene, and specific genes of interest inside yeast. The housekeeping gene, selected to have little variability in gene expression, will be tagged with the Cherry fluorescent protein. The genes of interest will be tagged with the Venus fluorescent protein. The constructed reporter constructs will be mated against the H3 and H4 histone mutant library generated by the Boeke lab as seen in figure 5 and 6 (Dai. et al).



Figure 5: Histone H3 and H4 mutant library the generated fluorescent reporter constructs will be mated against. "The library consists of an alanine scan with other systematic reside swaps and systematic tail deletions, totaling 486 mutants."[2] The white boxes show the original 135-peptide sequence for histone H3, and the original 102-peptide sequence for histone H4. Under in color are the amino acids used as substitutions.

Through the fluormetric analysis of the each reporter construct, the cell-to-cell variability of gene expression can be quantified and the roles of a variety of histone mutant can be characterized. Looking at the ratio of YFP to RFP we will determine the levels of transcription for our tagged proteins inside the cell versus the developmental stage of the yeast. From the data gathered about fluoresce reporters speculations can be made into the roles of histone modification and their effects on gene expression and non-genetic inheritance.





Materials and Methods:

I. Plasmid Construction

The fluorescent tag, Cherry, contained in the plasmid pKT355 was obtained from Kurt Thorn. LB-Amp liquid media was inoculated and allowed to grow overnight at 37°C. The plasmid was extracted in abundance using the Mega Plasmid Preparation Kit from Boston BioProducts. A double restriction digest with the restriction endonucleases: PmeI and MluI were preformed over night at 35°C on the pKT355 plasmid. The digest was then analyzed by gel electrophoresis. The band containing the Cherry protein (~3500 base pairs) was extracted using the Geneclean III Kit from MP Biomedicals. The concentration of the extracted doubly digested pKT355 plasmid was determined using gel quantification.

Next pGB006 (contained plasmid with selectable marker Trp1) made by Gwendolyn Bennett was grown up overnight at 37°C in LB-Amp cultures. The cultures were mini prepped via the Sigma Gene Elute Kit. Trp1 was amplified out of pGB006 using Platinum Pfx DNA Polymerase (Invitrogen) and the primers: TrpPmeEco3' and TrpMluBgl5' (cycling parameters: initial denaturation at 94°C for 3min., denaturation at 94°C for 15sec., annealing at 56°C for 30 sec., amplification at 68°C for 3 min repeated for 29 cycles, and a hold at 4°C). Trp1 PCR product was run out on a gel and gene cleaned. The purified product was then doubly digested with PmeI and MluI at 37°C for 2.5hrs. The reaction was cleaned using the MinElute Reaction Clean Up Kit by Qiagen.

Ligations of pKT355 and Trp1 were preformed (1:1, 3:1, 6:1) overnight at 16°C, and then heat inactivated at 65°C for 20min.

II. E.coli Transformation

50uL of competent E. coli cells (DH5α) for each ligation were thawed in separate tubes on ice. 20uL of ligations were added, and the transformations were incubated on ice for 30min. Next they were heat shocked for 1.5min at 42°C. The transformations were placed back on ice and 250uL SOC was added. Then they were allowed to recover an hour at 37°C. The transformations were plated on LB-Amp (ampicillin is selectable marker) plates and incubated at 37°C overnight.

Five colonies from pKT355 3:1 plate were used to inoculate 5mL LB-AMP cultures. The cultures were mini prepped using the Sigma Genelute Mini Prep Kit. Conformational restriction digestions of the transformed plasmid using SalI, ClaI, BgIII, and SspI at 37°C for 2.5 hours were done as seen in figure 8, and the plasmid was sent out for sequencing. Sequencing came back positive and the plasmid will now be called pSGS002 (contains Cherry and Trp1).

III. Yeast Construction

LB-Amp liquid media was inoculated with pSGS002 and allowed to grow overnight at 37°C. The plasmid was isolated using the Mega Plasmid Preparation Kit from Boston BioProducts. Specific 60 base pair primers were made to tag each gene of interest (TEF1 and YLR027C). A PCR was set up (cycling parameters: initial denaturation at 94°C for 3min., denaturation at 94°C for 30 sec., annealing at 58°C for 30 sec., amplification at 68°C for 2 min 10 sec. repeated for 5 cycles, then denaturation 94°C

for 30 sec., annealing at 68°C for 30 sec., amplification at 68°C for 2 min. 10 sec. for 25 cycles, a final amplification at 68°C for 10 min and a hold at 4°C) using Platinum Pfx DNA Polymerase (Invitrogen) for TEF1, and YLR027C. The products were run on a gel and the bands were cut out and extracted via GenecleanIII (Obiogene).

Yeast constructs will be made following the same protocol for pGB001(venues tagged with Leu2) with 60 base pair primers for: GLK1, HSP104, AGA1, FIG1, MET3 and HXK1.

IV. Yeast Transformation-

Yeast Transformations were carried out following the High-Efficiency Yeast Transformation protocol from the Rando Lab. On the first day, a single colony of yeast strain, BY4734, was grown in a 3mL culture overnight. The second day, 50mL of culture were inoculated in the evening to obtain an appropriate OD the next day. The target OD was approximately the (number of transformation) X (0.1) for a 50mL culture. On the third day, cultures at the correct OD were poured into 50mL conical and pelleted at 4000rpm for 2min. The pellet was washed in 25mL water and pelleted again. Next the pellet was resuspended by vortex in 1mL of 100mM LiOAc and transferred to a 1.5mL tube. Then pelleted with a short 10sec spin. The liquid was poured off and the pellet was resuspended up to (250/(0.833X0.6))X(OD) in 100mM LiOAc by vortex. The resuspension was then incubated at 30°C for 20min. Meanwhile, denature Salmon Sperm DNA (10mg/mL) by boiling at 95°C for 10min. 50µL aliquots were made of cells in LiOAc, 10µL of Salmon Sperm DNA (10mg/mL), and 1µg of transforming DNA was added for each transformation. Next the 50µL aliquots were vortexed and incubate at 30°C for 20min. Meanwhile 3mL of 100mM LiOAc/40% PEG was made and , combined

with 2.4mL 50%PEG, 0.3mL of 1M LiOAc, and 0.3mL of water. Next 300µL of 100mM LiOAc/40%PEG was added to the transformations. They were vortexed and incubated at 30°C for 20min. After they were incubated 40µL of DMSO was added to each one. The transformation were vortexed again, and heat shocked at 42°C for 20min. Next the cells were pelleted with a 10sec spin. The pelleted cells were washed with 500µL water and pelleted again. Finally the cells were resuspended in 100µL of water and spread on selective plates (-Leu for pGB001 (Venus), and –Trp for pSGS002 (Cherry)).

Yeast strain, BY4734, was transformed with pSGS002 for TEF1 was determined by conformational PCR and gel electrophoresis of PCR products. The new strain of yeast with Cherry and Trp1 will now be called SGSY001. This processes was repeated for BY4734 transformed with pSGS002 for YLR027C. The new strain with Venus and Leu2 will now be called SGSY002.

This strategy will be continued for the proteins labeled with Venus. Transformations will occur on SGSY001 and SGSY002 instead of BY4734 to generate yeast strains with a YFP-fused protein of interest and a RFP-fused control protein.

Results:

I. Plasmid Construction

In order to facilitate detection of gene activation, plasmid constructs were built to serve as a template for generation of homologous recombination constructs. The first plasmid was built to contain the AMP bacterial resistance marker, as well as the Trp1 gene for yeast selection and the cherry fluorescence gene (reporter). This plasmid was assembled using a two-way ligation scheme via the PmeI and MluI restriction enzymes on the plasmid pKT355 (containing the AMP bacterial resistance gene, cherry fluorescence gene, and Kan gene for yeast selection) obtained by Kurt Thorn. For our purposes the plasmid was modified to include the yeast Trp1 gene in place of Kan. Plasmids containing useful Trp1 genes (pGB006) were obtained from Gwendolyn Bennett.

PmeI (3' end) and MluI (5' end) were selected due to their single digestion sites on pKT355 outside of Kan that would not affect either the Amp resistance gene or the cherry fluorescence gene. After the double restriction digestion was preformed, the digested plasmid was run out on a 1% agarose gel. Two bands were seen, one larger band at approximately 3.5Kb with the cherry fluorescence gene and one smaller band at approximately 500bp with the Kan gene. The larger band was gel purified using GeneCleanIII (Obiogene). Next Trp1 was amplified from pGB006 using PCR methods.

Specific primers containing a PmeI restriction site on the 3' stand and a MluI restriction site on the 5' strand were used in conjunction with Platinum Pfx DNA Polymerase (Invitrogen) were used to amplify the Trp1 gene out of pGB006 (with an

extra 300bp upstream of Trp1 to include the Trp1 promoter). The PCR product was run out on a 1% agarose gel, and a band at approximately 1Kb with the amplified Trp1 gene was observed. The band was purified using GeneCleanIII (Obiogene), and doubly digested with PmeI and MluI. The purified fragments for our Trp1 and pKT355 digestions were ligated using T4 DNA Ligase (NEB).

The full ligation mixtures as shown in figure 7A containing the AMP bacterial resistance gene, cherry fluorescence gene and yeast Trp1 selection gene were used in transforming competent E. coli cells (DH5 α).

II. E. coli Transformation

E. coli cells transformed with our constructed plasmid were grown on ampicillin plates for selection purposes. The DH5α stain used in this experiment did not naturally contain an AMP resistance gene, and therefore would die on ampicillin plates. Normal colonies from our transformed plates were chosen at random and grown up in LB-AMP broth. The plasmids were purified from our selected colonies using Genelute Mini Prep Kit (Sigma). We ran them out 1% agarose gel, bands at approximately 4.5Kb were observed. Six diagnostic digestions (A, B, C, D, E, and F) using Sall, ClaI, BglII, and SspI, ran out on 1% agarose gel with bands at approximately 2.4Kb, 1.9Kb, 3.3Kb, 1Kb 3.8Kb, 0.5Kb, 3.4Kb, 1Kb, 2.4Kb, 1.9Kb, 2.9Kb, and1.5Kb respectively, confirmed our transformed E. coli indeed carried the our constructed plasmid to be named pSGS002 as seen in figure 8. Yeast transformation could now be carried out using homologous recombinant PCR products from pGB001 and pSGS002 using specific primers that anneal at the forward and reverse primer sites as seen in figure 7.

III. Yeast Construction

In order to generate fluorescently tagged genes, specialized primers were designed to target specific genes of interest inside yeast genome and allow homologous recombination with the constructed plasmids. The first consideration was to attach the gene for a fluorescent protein to a gene of interest, and the second was to prevent frame shifts that would mutate the fusion protein upon translation causing it to be disadvantageous in later stages of this experiment. The primers were designed to be 20bps specific to the newly created plasmids (upstream of the fluorescent protein and downstream of the yeast selectable marker), and 40bps specific to a homologous region of each gene as seen in figure 9A.

The plasmid pSGS002 was used to tag the housekeeping genes, TEF1 (Translation Elongation Factor) and YLR027C (Aspartate AminoTransferase), which have shown little variability in gene expression with histone modifications [10]. These will be used as control for cell size. The PCR products amplified from pSGS002 with the specialized primers will homologously recombine with their counter part inside the yeast genome. Figure 9B illustrates how this process will occur, linking TEF1 and YLR027C to the cherry fluorescent gene as seen in figure 9C and 9D. The TEF1-cherry or YLR027C-cherrry fusion proteins will give us a base line to normalize our data as we characterize gene expression.

The PCR products obtained from amplification of pSGS002 with TEF1 and YLR027C primers using Platinum Pfx DNA Polymerase (Invitrogen) were run out on 1% agarose gel. Bands at approximately 2Kb were observed, and confirmed that correct

amplification had occurred. They were then extract using GeneCleanIII (Obiogene) and used to transform yeast from the BY4734 strain to create the cherry tagged fluorescent constructs SGSY001 (cherry-TEF1) and SGSY002 (cherry-SLR027C).

The plasmid pGB001, created by Gwendolyn Bennett while working at the Rando lab, was created to tag genes that exhibit high cell-to-cell variation in gene expression [10]. For the initial study, six genes were chosen: HXK1 (Hexokinase isoenzyme) HSP104 (Heat Shock Protein) GLK1 (Glucokinase) FIG1 (Factor-Induced Gene) AGA1 (a cell-specific sexual Agglutination) MET3 (Methionine requiring). The PCR products amplified out of pGB001 using the specialized primers will homologously recombine with their counter parts in a similar fashion to the PCR products from pSGS002 as seen in figure 9A and 9B. The Venus tagged genes will give precise data about the varying level of gene expression between histone mutants.

PCR of pGB001 with HXK1, HSP104, and GLK1 primers using Platinum Pfx DNA Polymerase (Invitrogen) was preformed. The products were run on a 1% agarose gel and examined. Bands at approximately 2Kb were observed, verifying successful amplification. The bands were extracted using GenecleanIII (Obiogene), and used to doubly transform SGSY001 and SGS002.

IV. Yeast Transformation

To determine the success of the yeast transformations with pSGS002 PCR products were grown on Trp- plates. The BY4734 strain has Ura, His, Met, Leu, and Trp knockouts. Media was made to have all the essential amino acid for growth of the

SGSY001 and SGSY002. The transformed colonies that grew contained the Trp1 gene crucial to growth on the minimal media plates. Normal colonies were selected at random and grown in YPD without tryptophan. Next mini preps were preformed to extract the genomic DNA from the transformed yeast. PCR primers were designed to analyze the effectiveness of the transformations.

Three reactions (A, B, and C) were run, as seen in figure 8A, on wild type BY4734 and the transformations, to support their validity. The results were run out on 1% agarose gel and evaluated. Reaction A amplified outside the homologous region. A band was seen in wild type BY4734 at approximately 500bps in both controls for the TEF1 and YLR027C transformations, as seen in figure 10B and 10C. Reaction B began amplification outside the homologous region and terminated inside the cherry fluorescent gene. No bands were observed for wild type BY4734, but for the TEF1 and YLR027C transformations, bands were seen at approximately 800bps as seen in figure 10B and 10C. Finally reaction C began amplifying inside the Trp1 gene and concluded outside the homologous region. Again, it can be shown through figure 10B and 10C, that no bands were detected for wild type BY4734, and witnessed at 500bps respectively in both transformed strains.

The data obtained, confirmed through PCR methods the successfulness of both amplifications of pSGS002 with the TEF1 and YLR027C primers. The attainment of positive results for the two yeast transformations preformed from pSGS002 grants the generation of the cherry-TEF1 and cherry-YLR027C fluorescent reporter constructs. The new strains are named SGSY001, and SGSY002 respectively.

Transformations of SGSY001 and SGSY002 proceeded with the PCR products from pGB001 with HXK1, HSP104, and GLK1 primers to create doubly transformed reporter constructs. Similar to the cherry tagged constructs, the newly transformed yeast were grown on minimal media plates. They lacked tryptophan, and also leucine. Effective transformations contained the Trp1 gene from pSGS002 and the Leu2 gene from pGB001. Colonies grew for all three transformations of SGSY001, and for the GLK1 in SGSY002. Further confirmation through PCR is needed to positively prove the transformation.

Discussion:

Histones are strongly basic proteins responsible for the packaging and ordering of DNA inside eukaryotic cells. They are highly modifiable multifunctional proteins. 147 DNA base pairs are wound around an octamer of core histone proteins (H3, H4, H2A, H2B). Post-translational modifications allow enormous potential for functional responses inside the cell. Little is known about the effects of histone mutations on cell-to-cell variability in gene expression. To characterize these mutations, in particular, the ones occurring on histones H3 and H4, yeast strains with fluorescent reporter constructs were generated.

To discern the effects of histone mutants we wanted to two distinct fluorescent proteins. For our purposes, we chose, the yellow fluorescent protein, Venus, and the red fluorescent protein, Cherry. They had different excitation and emission wavelengths that would not interfere with quantification of fluorescence for each individual protein, and relatively high or moderate brightness . As seen in figure 11, Cherry has an excitation wavelength at 587nm and an emission wavelength at 610nm. Venus has an excitation wavelength at 515nm and an emission wavelength at 528nm.Our goal was to create two fusion proteins, one made up of Cherry linked to a control gene, and the other made up of Venus linked to a gene of interest.

The Cherry fusion protein needed to have stable expression throughout the cell cycle, to give a base line for protein abundance that could be used to normalize the data we obtain in regards to the varying levels of gene expression between cells. TEF1 and YLR027C were selected as intensity controls for cell size. They are low noise proteins due to their exhibition of little variability in gene expression. Conversely, the Venus

fusion protein needed to exhibit high cell-to-cell variation. The genes: HXK1, GLK1, HSP104, FIG1, AGA1, and MET3 were chosen. AGA1 and MET3 were know to be environmentally-inducible noisy genes, while others had shown expression regulated by histone tails, H4 point mutations [3], and tail deletion mutants [15].

The first experimental step taken in the construction the Cherry and Venus fluorescent reporter constructs was the development of plasmids in E.coli. Gwendolyn Bennett did early work in the production of a pGB001. Contained within the plasmid was Venus and Leu2. Next we began the development of pSGS002, a plasmid similar to pGB001, but instead of Venus and Leu2 it contained Cherry and Trp1.

The completed plasmids, pGB001 and pSGS002, were then used to generate PCR products that homologously recombined within the yeast genome. Specialized primers were used to amplify the fluorescent gene and selectable marker that target specific genes of interest inside the cell. The primers were designed to fuse the fluorescent gene to the gene of interest, and allow yeast selection using specific media.

We were able to generate Cherry tagged TEF1 and YLR027C reporter constructs. Further work is needed to confirm and continue the generation double transformed yeast strains containing Venus and Cherry reporter constructs. In the future we will mate the fluorescent reporter constructs against a comprehensive library of histone H3, and H4 mutants (figure 5), and characterize the cell-to-cell variability in gene expression.

Through quantitative single-cell measurements of protein abundance using fluorescence microscopy speculations into the roles of histone modifications on gene expression can be made. Venus abundance levels will be monitored to identify variability

between cells. The data gathered will then be normalized to Cherry abundance levels to factor out any incongruency due to cell size.

Using the H3 and H4 histone mutant library in conjunction with our generated fluorescent reporter constructs will give functional data about the effects of histone mutations on gene expression. Through Fluorescence microscopy we will determine the cell-to-cell variability in protein abundance and characterize the roles of H3 and H4 histone mutations compared to wild type strains.

The ultimate goal is to analyze strains for epigenetic regulation and expression across cell lineages via microfluidic devices. The yeast will be continuously supplied with fresh media flowing over them to eliminate environmental factors as a cause for variability in expression. Individual yeast will be trapped in long channels 1.5 microns wide forcing the yeast to grow in lines with the daughter cell adjacent to the mother as seen in figure 12. Simple inspection of fluorescence images can reveal the switching rates of epigenetically regulated genes in the system. Correlations between mother and daughter gene expression will be measured. This will enable the time behavior of epigenetically regulated genes to be measured over many generations, the roles of various histone residues in tuning epigenetic switching rates can be characterized.



Figure 7: Generated plasmids to be used in construction of fluorescent tagging constructs. Both plasmids contain the AMP bacterial resistance gene, and forward and reverse primer sites to be used in PCR amplification on homologous recombination constructs for yeast transformation. (A) pSGS002 contains Cherry and a Trp1 selectable marker. This plasmid will be used to tag a housekeeping gene as a control to measure the progress of the cell cycle. (B) pGB001 contains Venus and a Leu2 selectable marker. This plasmid will be used to tag a gene with expression tied to histone modifications.



Figure 8: Diagnostic digestion to verify size and identity of

pSGS002. Six reactions (A, B, C, D, E, F) were run at 37 C for 2.5 hours to confirm the transformation (A) Restriction double digest with SalI and ClaI had expected bands at 2.4Kb and 1.9Kb. (B) Restriction double digest with SalI and BglII had expected bands at of 3.3Kb and 1Kb. (C) Restriction double digest with SalI and SspI had expected bands at 3.8Kb and 0.5Kb. (D) Restriction double digest with ClaI and BglII had expected bands at 3.4Kb and 1Kb. (E) Restriction double digest with ClaI and SglII had expected bands at 2.4Kb and 1.9Kb (F) Restriction double digest with SspI and BglII had expected bands at 2.9Kb and 1.5Kb.



Figure 9: Reporter constructs amplified out of pSGS002 designed to tag TEF1 and YLR027C in BY4734(Ura, His, Met, Leu, and Trp knockouts) (A) PCR

amplification out of the plasmid. In green are the 40 bps specific homologous sequences for our gene of interest attached to our 20bps specific PCR primers. Being amplified out of pSGS002 are the cherry fluorescent gene, in red, used to tag a housekeeping gene, and the yeast Trp1 gene, in yellow, used as a selectable marker (B) Homologous recombination with our gene of interest. The PCR product will knock out the original sequence and replace it with one containing cherry, the fluorescent reporter, and Trp1, the selectable marker. The cherry fluorescent gene is linked to the housekeeping gene of interest, and will create a fusion protein when translation occurs. (C) SGSY001: TEF1 tagged with cherry fluorescent gene and yeast Trp1gene selectable marker. (D) SGSY002: YLR027C tagged with cherry fluorescent gene and yeast Trp1 gene selectable marker.



Figure 10: Confirmation through PCR of SGSY001 and SGSY002. (A)

Diagram of the positioning of the primers used to prove generation of Cherry tagged gene of interest. The arrows are sites were primers for reactions A, B, and C annealed. In the diagram the tagged gene is shown on the right in purple, the cherry fluorescent gene is shown in red , and the Trp1gene (used as a selectable marker) is shown in yellow. (B) Confirmation through PCR of the generation of Cherry-tagged TEF1 strain (SGSY001). In tagged strain: reaction A should be 2297bps (not observed), reaction B should be 820bps, and reaction C should be 403bps. In wild type, only reaction A should be observed at roughly 500bps (C) Confirmation through PCR of the generation of Cherry-tagged YLR027C strain (SGSY002). In tagged strain: reaction A should be 2424bps (not observed), reaction B should be 855bps, and reaction C should be 601bps. In wild type , only reaction A should be 601bps. In wild type , only reaction A should be 900bps

Excitation and Emission Spectra of Fluorescent Proteins								
Excitation Spectral Profiles			Emission Spectral Profiles					
100 uotiduos 60 40 0 0		Cerulean Venus EBFP -			mCl	sabira ange herry plum 40 malized 60 lized 20 on 0		
300 400 500 600 700 400 500 600 700 800								
Protein	Excitation Maximum	Emission Maximum	Molar Extinction	Quantum	in vivo	Relative Brightness		
GFP (wt)	(nm) 395/475	(nm) 509	Coefficient 21,000	0.77	Monomer*	(% of EGFP) 48		
Blue Fluorescent Pro	teins							
EBFP Sapphire	383	445 511	29,000	0.31	Monomer*	27		
T-Sapphire	399	511	44,000	0.60	Monomer*	79		
Cyan Fluorescent Pro	439	476	32,500	0.40	Monomer*	39		
mCFP	433	475	32,500	0.40	Monomer	39		
Cerulean	433	475	43,000	0.62	Monomer*	79		
CyPet AmCyan1	435	477	35,000	0.51	Monomer* Tetramer	53		
Midori-Ishi Cyan	472	495	27,300	0.90	Dimer	73		
mTFP1 (Teal)	462	492	64,000	0.85	Monomer	162		
EGFP	484	507	56,000	0.60	Monomer*	100		
AcGFP	480	505	50,000	0.55	Monomer*	82		
TurboGFP	482	502	70,000	0.53	Monomer*	110		
Azami Green	487	509	57,500	0.68	Monomer*	116		
ZsGreen	493	505	43,000	0.91	Tetramer	117		
Yellow Fluorescent P	roteins							
EYFP	514	527	83,400	0.61	Monomer*	151		
Venus	515	528	92,200	0.57	Monomer*	156		
mCitrine	516	529	77,000	0.76	Monomer	174		
PhiYEP	517	530	104,000	0.77	Monomer*	238		
ZsYellow1	529	539	20,200	0.42	Tetramer	25		
mBanana	540	553	6,000	0.70	Monomer	13		
Kusabira Orange	548	559	51,600	0.60	Monomer	92		
mOrange	540	550	71.000	0.00	Manager	146		
dTomato	548	562	69.000	0.69	Dimer	146		
tdTomato (Tandem)	554	581	138,000	0.69	Monomer	283		
DsRed	558	583	75,000	0.79	Tetramer	176		
DsRed2	563	582	43,800	0.55	Tetramer	72		
DsRed-Express (T1)	555	584	38,000	0.51	Tetramer	58		
mTangerine	568	585	38,000	0.30	Monomer	34		
mStrawberry	574	596	90,000	0.29	Monomer	78		
AsRed2	576	592	56,200	0.05	Tetramer	8		
mRFP1	584	607	50,000	0.25	Monomer	37		
mCherry	584	610	44,000	0.20	Dimer	26		
HcRed1	588	618	20,000	0.015	Dimer	1		
mRaspberry	598	625	86,000	0.15	Monomer	38		
HcRed-Tandem	590	637	160,000	0.04	Monomer	19		
mPlum AO143	590	649	41,000	0.10	Monomer	12		
AG143	595	655	90,000	0.04	retramer	Neak Dimer		

Figure 11: Excitation and Emissions Spectra for Fluorescent Proteins. Cherry fluorescent protein has an excitation wavelength at 587nm and an emission wavelength at 610nm. Venus fluorescent protein has an excitation wavelength at 515nm and an emission wavelength at 528nm. (A) Excitation and Emission spectra. (B) Characteristic table of fluorescent proteins. [11]



Figure 12: Microfluidic device to be used in the analysis of epigenetic regulation and expression switching rates. The channels where the yeast would be grown are 1.5 microns wide, not quite wide enough for two yeast. The cells are forced to grown in a line, gene expression on a cell-to-cell basis can be easily quantified using fluorescent reporter constructs over many generations.

References:

- 1. Cooper, Geoffrey, and Hausman, Robert. "The Cell, a Molecular Approach". Sinauer Associates Inc, 2007. Print.
- Dai, J. Hyland, E. Yuan, D. Huang, H. Bader, J. Boeke, J. (2008). "Probing Nucleosome Function: A Highly Versatile Library of Synthetic Histone H3 and H4 Mutants". Cell 134: 1066-1078
- 3. Dion, M.F., et al., "Genomic characterization reveals a simple histone H4 acetylation code." Proc Natl Acad Sci USA, 2005
- Kouzarider, T. (2007). "Chromatin Modifications and Their Function" Cell 128: 693-705
- 5. Klug, William, Michael Cummings, Charlotte Spencer, and Michael Palladino. Essentials of Genetics. Benjamin-Cummings Pub Co, 2009. Print.
- Laurence D, J, R. (1966) "Interactions of Calf-Thymus Histone Fractions in Aqueous Solution with 8-Anilinonaphthalene-1 -sulphonic Acid" Biochem. J. (1966) 99, 419
- Lindsay S. Shopland, Christopher R. Lynch, Kevin A. Peterson, Kathleen Thornton, Nick Kepper, Johann von Hase, Stefan Stein, Sarah Vincent, Kelly R. Molloy, Gregor Kreth, Christoph Cremer, Carol J. Bult, and Timothy P. O'Brien. (2006) "Folding and organization of a contiguous chromosome region according to the gene distribution pattern in primary genomic sequence". J Cell Biol. 2006 July 3; 174(1): 27–38.
- 8. Lomvardas, S. Thanos, D. (2002). "Modifying Gene Expression Programs by Altering Core Promoter Chromatin Architecture". Cell 110: 261-271
- 9. Luger K, Mader AW, Richmond RK, Sargent DF, Richmond TJCrystal structure of the nucleosome core particle at 2.8 A resolutionNature v389, p.251-260
- 10. Newman, J.R., et al., "Single-cell proteomic analysis of S. cerevisiae reveals the architecture of biological noise". Nature, 2006. 441(7095):p.840-6
- 11. Piston, David, W. Claxton, Nathan S. Olenych, Scott G. Davidson, Michael W. (2004). "The Fluorescent Protein Color Palette" Olympus FluoView Resource Center: The Fluorescent Protein Color Palette http://www.olympusfluoview.com/applications/fpcolorpalette.html>
- 12. Radman-Livaja, M. Rando, O. "Nucleosome positioning: how is it established, and why does it matter "?
- 13. Rekas, A., Alattia, J.R., Nagai, T., Miyawaki, A., Ikura, M. (2002)"Crystal structure of Venus, a yellow fluorescent protein with improved maturation and reduced environmental sensitivity". J.Biol.Chem. 277: 50573-50578
- Routh, Andrew, Sandin, Sara, Rhodes, Daniela (2008). "Nucleosome repeat length and linker histone stoichiometry determine chromatin fiber structure". PNAS vol. 105 no. 26 8872–8877
- 15. Sabet, N., et al., "Global and specific transcriptional repression by the histone H3 amino terminus in yeast." Proc Natl Acad Sci USE, 2003. 100(7):p.4084-9
- 16. Sarah R. Edwards, Thomas J. Wandless. "Dicistronic regulation of fluorescent

proteins in the budding yeast Saccharomyces cerevisiae". Yeast, Volume 27 Issue 4, Pages 229 – 236 Published Online: 17 Dec 2009 Copyright © 2010 John Wiley & Sons, Ltd.

- Schild-Poulter, S. Sassone-Corsi, P. Granger-Schnarr, M. Schnarr, M (1996).
 "Nucleosome assembly on the human c-fos promoter interferes with transcription factor binding". Nucleic Acid Research Vol. 24, No. 23
- Shu, X., Shaner, N.C., Yarbrough, C.A., Tsien, R.Y., Remington, S.J.(2006)" Novel chromophores and buried charges control color in mFruits". Biochemistry 45: 9639-9647
- 19. Stokes, G. G. (1852). "On the Change of Refrangibility of Light". Philosophical Transactions of the Royal Society of London 142: 463–562
- 20. Voet, Donald, Judith G. Voet, and Charlotte W. Pratt. "Fundamentals of Biochemistry: Life at the Molecular Level." Hoboken, NJ: Wiley, 2008. Print.
- Wing R, Drew H, Takano T, Broka C, Tanaka S, Itakura K, Dickerson R (1980). "Crystal structure analysis of a complete turn of B-DNA". Nature 287 (5784): 755–8.
- 22. Yan, Yiu. "DNA, RNA and Chromosomes." Musings (2009): n. pag. Web. 26 Apr 2010.