Title
Nucleosomes as carriers of epigenetic memory

Author
Schlissel, Gavin

Publication Date
2019

Peer reviewed|Thesis/dissertation

Nucleosomes as carriers of epigenetic memory

by

Gavin S Schlissel

A dissertation submitted in partial satisfaction of the

requirements for the degree of

Doctor of Philosophy

in

Molecular Cell Biology

And the Designated Emphasis

in

Computational and Genomic Biology

in the

Graduate Division

of the

University of California, Berkeley

Committee in charge:

Professor Jasper Rine, Chair
Professor Barbara Meyer
Professor Nick Ingolia
Proffesor Oskar Hallatschek

Summer 2019

Abstract

Nucleosomes as carriers of epigenetic memory

by

Gavin Schwebs Schlissel

Doctor of Philosophy in molecular cell biology

and

Designated emphasis in computational and genomic biology

Professor Jasper D. Rine, Chair

A core assumption in chromatin biology is that nucleosomes store and transmit epigenetic memory of gene regulation through cell division. A requirement of this model is that nucleosomes – modified in place in one generation – continue to occupy the same DNA locus in the subsequent generation. In my graduate work, I asked whether nucleosomes remember their position through cell division. To execute my experiments, I developed a synthetic chromatin modifying enzyme that deposited high affinity covalent chromatin modification at a defined locus. The chromatin-modifying enzyme could be delocalized within minutes by the addition of a small molecule antibiotic, and the synthetic chromatin modification could be tracked for several hours after preventing new label deposition. Using this synthetic enzyme, I labeled nucleosomes at the *GAL10* locus in *S cerevisiae* and tracked their position through DNA replication. I found that nucleosomes remained localized through DNA replication, suggesting that nucleosomes can transmit epigenetic. Furthermore, I found that in the absence of the fork associated nucleosome chaperones Mcm2 and Dpb3, nucleosomes did not remain localized, suggesting that *mcm2 dpb3* mutants cannot transmit epigenetic memory through chromatin. In addition to exploring the fate of nucleosomes during DNA replication, I tracked nucleosomes through transcription and discovered that contrary to published models, nucleosomes are not translated linearly along an open reading frame during transcription.

In the thesis that follows, I discuss the context for our collective intuition about the role of chromatin in transmitting epigenetic memory, and I identify the points at which our collective understanding rests on thin or contradictory data. I then discuss my attempts to develop and apply technology to reach a satisfactory resolution of one of the central unanswered questions in chromatin biology: do nucleosomes remember their position? Lastly I discuss side projects adjacent to chromatin biology that occupied my attention during the slower moments of my core thesis research.

For Paige

Acknowledgments

There are two people who require special acknowledgement, and several more who deserve mention for their varied scientific or personal contributions to my experience in graduate school.

The first person who requires special acknowledgment is Paige. Science is an isolating enterprise, in that very few people understand my daily hope or frustration, or can share in my sense of accomplishment when I've done something I'm proud of. Paige has been the anchor of a parallel identity for me, in which I don't anchor my self-image on my scientific successes or failures. To Paige, I am important and interesting and valuable regardless my progress in lab, and that knowledge expands the radius I feel empowered to explore scientifically or personally. In that way, Paige is the single person who has most encouraged my curiosity, and the single person without whom the course my life would be different.

The second person who requires special acknowledgment is Jasper. Beyond the mechanics of mentorship—weekly lab meetings, organizing lab outings, writing recommendations—Jasper is the nucleus of a scientific culture that is humble without forsaking pride and grand without neglecting detail. Scientifically, he proven to me that new layers of complexity reveal themselves at different scales, and that using new technology to revisit old questions can texture our understanding of fundamental biological processes. He approaches science with a deep sense of optimism ("Let's celebrate that result while it still might be true!") and a offers his personality as a playful counterbalance to tedious topics at hand ("zip Zoop ZOOP!"). Jasper's willingness to inhabit another person's curiosity and give guidance detached from any personal *geist* has attracted a legion of acolytes eager to honor Jasper's generosity by paying it forward to Jasper's more recent trainees; I have already benefited handsomely by this effect, and I anticipate that I will continue to benefit throughout my career.

The last thing I'll say about Jasper revisits the first thing he said to me after I started graduate school. It explains why I joined the lab, and why I am heartbroken to leave it. It is also the kernel of Jasper's outlook that I hope to perpetuate in those around me throughout my career. When I asked Jasper about the tradeoff between working hard in science and having a varied and fulfilling life outside of science, he told me: nothing that you accomplish in science will be worthwhile if takes away from the joy you experience in life, and you'll accomplish dramatically more in science and in life if you can maintain a sense of joy. Under Jasper's mentorship, my scientific and personal lives have been richly vested in joy.

Beyond Paige and Jasper, there are several people who have left deep marks on my approach to science and to life. I'd particularly like to acknowledge Barbara Meyer, who routinely cuts through my delusions or rationalizations and makes me approach scientific or personal questions with a realistic and humble perspective. Barbara also has been an incredible matchmaker for me, and has introduced me to many like minds both in the Berkeley community and during my postdoc search. I'd also like to acknowledge Nick Ingolia, who has been a singular example of the style of science I hope to pursue during my career, by turning new technology to understand old problems. Nick's ability to synthesize the vastness of science and engineering and math and apply it towards understanding biology will continue to set the standard for creative and incisive experimental design throughout my

career. Every time I walked into Nick's office soliciting help on some technical or scientific line of thinking, his advice was just perfect, and inspired by his extremely broad scientific consciousness: I am continually in awe.

In addition to Barbara and Nick, the entire Rine Lab has contributed to the design and execution of my thesis research. Katie Sieverman mentored me during my rotation and taught me the mechanics of yeast culture. Anne Dodson developed the CRASH assay, which was the cornerstone of my first project studying heterochromatin during aging. Also, Anne and I sat in a cluster with Sarah Bissonnette for a year and a half – the two of them provided invaluable personal support as I settled into lab life. Debbie Thurtle-Schmidt stands out in two ways – first, she mentored me over the course of a summer I spent in the Rine lab during college, and set an enduring standard for balance and happiness in an otherwise stressful environment; and second, Debbie was the person who suggested that I use BirA to execute the experiments that became the cornerstone of my thesis. Additionally, Davis Goodnight performed experiments early in the project that served as positive controls throughout the remainder of the technology development process. Before Davis worked with me in his rotation, I felt isolated—like no one I knew really *got* what I was trying to study. By choosing to work with me during his rotation, Davis gave me enormous confidence that the experiments I was designing were interesting and valuable. His specific contributions to my experiments are cited in the text of this thesis.

Lastly I need to thank my parents & siblings, who have impressed on me the singular value of learning, research and teaching. At home, learning was constant, and knowledge spread both down from my parents and horizontally among my siblings. School was always a space that I shared with my siblings, and we could collectively revisit school gossip and share mentors and develop relationships outside of the house. All of those small connections summed to make school feel like a second home for me – a feeling that continues to my university life today.

1 The backdrop of the lab's work

Before he started his lab, Jasper performed a screen as a graduate student for silent information regulator proteins. At the time, it was known that yeast encoded three copies of mating type information at three different loci, and that the mating phenotype of the cell was determined by the allele present at the *MAT* locus, near the centromere on chromosome III (Nasmyth and Tatchell, 1980). The auxiliary copies were encoded on the distal left and right arms of the same chromosome, and were known to be used as the substrate for mating type switching, in which the information at the auxiliary mating type loci *HML* and *HMR* could be copied to the *MAT* locus, thereby changing the mating type of the cell (Kushner et al., 1979; Strathern et al., 1982). Concurrently with Jasper's original *SIR* screen, the yeast field was discovering that the sequence of DNA at *HML* and *HMR* was exactly identical to that at the corresponding allele of *MAT*, and other targeted mutagenesis approaches had identified mutations that affect mating type silencing and map adjacent to – not within – the silent mating type loci (Abraham et al., 1984; Feldman et al., 1984). Thus, around the time that Japer was developing his silencing screen, it seemed likely that transcriptional regulation of the silent mating cassettes involved some kind of transcriptional regulation at-a-distance, rather than simple promoter-dependent gene regulation. Based on this hypothesis, Jasper expected it would be possible to find recessive mutants that contain a loss-of-function variant in the silencing machinery.

To execute the screen, Jasper used a strain that was mutant at *mat*α*1* locus. In this genetic background, loss of silencing from *HML*, which encodes auxiliary *MAT*α information would result in a *MAT*α mating phenotype. If a mutation resulted in loss of silencing at both *HML* and *HMR*, the mutant would express both *MATa* and *MAT***a** information, such that the strain would have a *MAT***a**/α mating phenotype (non-mating). Based on our modern understanding of *HML* regulation, the screen could have discovered two classes of mutations that suppress the *mat*α*1* mutant – mutations that affect silencing *in cis* at *HML*, by mutating regulatory DNA required for silencing, or mutations that affect silencing *in trans*, with the caveat that the mutants must be able to independently effect *HML* and *HMR*[1]. The silencing allele that arose from that suppressor screen exerted its effect *in trans*, and it was capable of regulating the Mat**a** and Matα regulons independently (Rine et al., 1979). That mutant was a null mutation in a gene that Jasper named *SIR1*. Although he likely didn't appreciate it at the time, Jasper was exceptionally lucky to identify *SIR1* – it is anomalous among *SIR* proteins in that it can uncouple silencing at *HML* from *HMR*, and it can result in a semi-stable epigenetic mating type phenotype that enabled his mating-based screen to work.

After the initial screen that identified *SIR1*, Jasper updated the screen design to take advantage of a natural isolate that contained silent copies of *Mat*α at both *HML* and *HMR*, but no silent copies of *MAT***a** information. Working with a version of this strain that was mutated at the *MAT* locus, Jasper was able to screen for mutants that disrupt silencing

[1] Additionally the screen could have identified mutants with aberrant regulation of mating pathway genes, that could mate as *MAT*α without *MAT*α 1 information, but those alleles are less interesting.

without the caveat that the mutants must regulate *HML* and *HMR* independently, and this screen identified *SIR2, SIR3* and *SIR4,* each of which is essential for silencing at *HML* and *HMR* (Rine and Herskowitz, 1987). Beyond identifying genes required for effective mating type silencing, Jasper's work (along with Abrahm's and Feldman's) consolidated the idea that auxiliary mating type cassettes are subject to *silencing* as opposed to non-activation, and that silencing acts over a domain, rather than over individual genes.

Having shown that silencing can act on a locus of indeterminate composition, the field progressed to understand the molecular correlates of silencing, and the biochemical activity of the *SIR* genes. Multiple lines of evidence converged on the idea that histones are important for gene regulation at *SIR*-silenced loci, including the silent mating type cassettes and telomeric domains. Specifically, the multiple labs found that that the N-terminus of H3 and H4 are essential for silencing, and can interact with Sir4 (Hecht et al., 1995; Kayne et al., 1988; Park and Szostak, 2015; Thompson et al., 1994) and the Broach lab discovered stereotypical patterns of histone acetylation that are associated with silent chromatin (Braunstein et al., 1993). These findings spoke to the possibility that *SIR*-silencing acts through chromatin modifications to affect the biochemical properties of chromatin, thereby repressing transcription. The final element of this model fell into place when it was discovered that Sir2 is an enzyme that catalyzes the deacetylation of lysine 16 on histone H4: Sir2 exposes a binding site on the N-terminus of H4, which is occupied by a complex that includes Sir2, Sir3 and Sir4 (Imai et al., 2000; Smith et al., 2002; Stebbins et al., 2002).

In addition to identifying the proteins that play a central role in mating type silencing, the field has moved towards a broadly accepted view of the molecular events that correspond to silencing establishment. Silencing in yeast is initiated by DNA sequence-targeted factors that bind silencers at *HML* and *HMR*. These sequence-targeted factors include the transcription factors Abf1 and Rap1 as well as the Origin Recognition Complex (ORC). The order or co-occurrence of binding among these factors has not been established, but it has been shown that ORC recruits Sir1, which subsequently recruits Sir4, Sir2 and Sir3 (Gartenberg and Smith, 2016). Before Sir2 recruitment, histones at silent mating type loci are acetylated at H4-K16 by Sas2, and Sir2 deacetylation of H4-K16 allows Sir3 to directly bind nucleosomes (Rusché et al., 2002). The Sir3-H4-K16 contact is thought to be the primary structural attachment point between the Sir complex and chromatin, but Sir3 apparently cannot bind H4-K16 at *HML* unless it has been first acetylated by Sas2 and then deacetylated by Sir2 (Armache et al., 2011; Thurtle and Rine, 2014). In addition to Sir3 binding nucleosomes, Sir2 and Sir4 each independently can interact with Sir3, and the three together are thought to form a complex that covers and occludes the silent mating type loci (Moazed et al., 1997; Rusché et al., 2002). Though the Sir2,3,4 complex has been observed in vitro and has been precipitated from cell extracts, it has never been structurally visualized, except small sub-features of the structure like the histone-binding domain of Sir3 or a long helix in the c-terminus of Sir4, and it is impossible to know whether the proteins form a stereotyped complex with defined stoichiometry and geometry, or whether the proteins form a semi-ordered multi-valent condensate (Armache et al., 2011; Johnson et al., 2009).

Once the Sir complex spreads over the auxiliary mating type loci, it is thought to physically occlude the underlying DNA, thereby preventing transcription. Interestingly, the *HML* promoter is accessible to Rap1, and strong promoters like the *GAL1,10* promoter can overcome silencing and allow productive transcription (unpublished observation, Ellie Bondra; Steakley and Rine, 2015; Wang et al., 2015). Thus, it is difficult to establish whether the Sir complex asserts silencing by blocking transcription initiation, transcription elongation or both.

Sidestepping the question of how the Sir complex physically prevents productive transcription, it is clear that transcriptional silencing shows two unique points of correspondence to the cell cycle. First, single-cell silencing assays make it clear that whereas Sir silencing can fail at any time, cells can only establish silencing after passage through S-phase of the cell cycle, even independent of actual DNA synthesis through the silent locus (Kirchmaier and Rine, 2001). This finding supports the idea that the mechanism of silencing establishment is related to the mechanism of DNA replication or mitosis: it is tempting to speculate about the relationship between cohesin, cohesion or condensation and transcriptional silencing and there are some data in support of the idea that cohesin contributes to silencing, but there is no compelling model to connect the process of chromatin condensation to silencing (Wu et al., 2011).

The second point of correspondence between transcriptional silencing and DNA replication circles back to Jasper's foundational silencing screen and Lorraine Plllus's subsequent characterization of the *sir1* mutant phenotype. Lorraine discovered that within a single, genetically identical culture of *MAT***a** *sir1* cells, cells could exhibit either of two phenotypes: they either arrest and shmoo in response to α-factor, or they continue to divide, insensitive of the mating pheromone (Pillus and Rine, 1989). Furthermore, she found that each phenotype "breeds true" through mitosis – that cells after DNA replication show the same phenotype as their progenitors. Thus, *sir1* mutants reveal epigenetic inheritance of the silent state, whereby cells are heritably phenotypically distinct despite having identical DNA sequences. This phenotype speaks to a missing piece in our understanding of DNA replication and chromatin biology: How can a locus *remember* its gene expression pattern during DNA replication?

2 The Piece I wanted to contribute

In choosing a thesis project, I wanted to find a problem that was both germane to yeast mating type gene silencing, and interesting to a broader audience than the yeast mating field. I was particularly inspired by Lorraine Pillus's finding that silencing of *HML* is epigenetically inherited in a *sir1* background, which implied that silencing might be epigenetically inherited in wild-type cells. If mating type silencing could be epigenetically inherited, and mating type silencing involved chromatin modifications, it seemed plausible that chromatin modifications might transmit the epigenetic memory of silencing across generations.

The idea that was beginning to take shape in my mind was a test of a popular framework – the histone code. The histone code model was intuitive by analogy between DNA and chromatin. DNA comprises a linear sequence of distinct elements, and chromatin also comprises a linear array of distinct elements: maybe the linear array of modified nucleosomes on a chromatin fiber stores information as does the sequence of DNA bases, and maybe that information can be replicated and transmitted semi-conservatively as it is in DNA. This model has been advanced by many reviews, and many people take it for granted that the linear array of nucleosomes stores and transmits information on top of the digital information transmitted by DNA replication.

To extend the analogy to DNA: in DNA, we know that the information stored in a purine or pyrimidine base will remember its position relative to the position of other bases, because the DNA bases are covalently attached in a polymer, and barring some chemical interruption of the polymer, the sequence of bases will be stable indefinitely. A DNA base has a defined chemical identity, and that chemical identity is not changed by DNA replication, except in extremely rare circumstances (i.e. spontaneous cytosine deamination). These features of DNA biochemistry underlie the ability of DNA to store and transmit information reliably over hundreds of millions of years of DNA-based life, but neither feature extends to chromatin. Whereas consecutive DNA bases are covalently attached, ensuring that the position of a DNA base relative to its neighbors cannot be changed, nucleosomes in chromatin are not covalently attached to DNA or to one another, meaning the energetic barrier to re-arranging the sequence of nucleosomes on a chromatin fiber is dramatically less than the energetic barrier to re-arranging the sequence of DNA bases in DNA. Similarly, whereas DNA base identity is immutable and can therefore store information indefinitely, countless enzymes have been identified that can chemically modify nucleosomes, suggesting that a nucleosome's biochemical behavior can be re-written easily compared to that of DNA.

I have dwelled on the analogy between information storage in DNA and the possibility of information storage through some kind of linear histone code to illustrate the point that we must not take for granted that the elegance of information storage and transmission in DNA generalizes to information storage and transmission in nucleosomes. Since 1953, it seems that biologists are pre-occupied with searching for parsimonious digital codes – local entropy wells that conform to the Spartan aesthetic that biologists instantly appreciate when we study the chemical structure of the genetic material. In our collective rush to see codes and symmetry and beauty in epigenetic inheritance, the field has taken some foundational questions about chromatin biochemistry for granted, and each is key to our understanding of information transfer through chromatin. Two of the most central questions that remain unanswered are:

1) How long do nucleosomes remember their position relative to DNA?
2) How long do nucleosomes remember their regulatory role?

There have been several attempts at answering the questions above, but each is imperfect. My goal in my PhD research has been to finally answer those questions, and thereby determine to what extent nucleosomes can serve as a carrier of epigenetic information.

2.1 Outstanding ideas about histone dynamics during DNA replication

Because Sir silencing involves local chromatin modifications, and because Sir silencing can be epigenetically inherited in *sir1* mutant cells, it is widely believed that chromatin can store memory of transcriptional silencing and transmit that memory through DNA replication. For this model to be true, chromatin features must remember their location and their regulatory role through DNA replication. Nucleosomes – protein octamers that bind ~146 bp of DNA – are the quantum unit of chromatin, and the statement that "chromatin features must remember their location" could alternately be articulated as: a given nucleosome must remember its position, with the caveat that it is not obvious whether modification of a single nucleosome would have a discrete phenotype[2].

Several studies have tried to understand whether nucleosomes remember their position in the genome after DNA replication, and they have taken many different biochemical or biological approaches. Despite 40 years of effort, the question is unsettled.

The first serious effort to address the question of whether nucleosomes are passaged *in cis* at the replication fork came from biochemical work on DNA replication. After the advent of the SV40 *in vitro* replication assay, multiple labs developed protocols to trace nucleosome occupancy through DNA replication. To achieve this, biochemists purified chromatinized SV40 minichromosomes, and treated the chromosomes with psoralen, which crosslinks DNA at any position where there is no nucleosome bound, then denatured the DNA and visualized it by electron microscopy, counting crosslinked and melted regions to determine how many nucleosomes occupied the DNA. Furthermore, these experiments could be performed in stopped-replication experiments, to capture the position of nucleosomes ahead of and behind the replication fork. This set of experiments was performed under many different sets of conditions: many were performed in histone-depleted extracts, some used alternative translation inhibitors to prevent the synthesis of new histones during the experiment, some were performed with non-replicating carrier DNA present in the reaction and others were performed in the presence of replicating carrier DNA (Gruss et al., 1993). These different approaches routinely arrived at alternative conclusions. Notably, it was shown that one could produce either result – yes or no to the question "are nucleosomes inherited *in cis*" – by performing the reaction in the presence of non-replicating or replicating carrier DNA (Gruss et al., 1993). In short, anyone looking for justification for their opinion on chromatin inheritance could find it in the SV40 chromatin *in vitro* replication literature.

[2] It's plausible, for example, that chromatin-based gene regulation requires multiple nucleosomes or even nucleosomes with associated effector proteins to affect a gene expression state, and thus "chromatin features" should be interpreted broadly.

Later, another style of experiment was tried that combined *in vivo* DNA replication in cancer cell lines with thoughtful *in vitro* analysis. These experiments started by testing the model that the nucleosome core particle (which is a symmetric dimer-of-dimers) splits during DNA replication, semi-conservatively transferring half of each nucleosome to each daughter strand during DNA replication. This model is intuitive by analogy to DNA replication, in which the DNA double helix melts into single strands, one of which ultimately is inherited by each daughter cell. To test the model, Vaughn Jackson synchronized HeLa cells, then switched them into media containing density-labelled amino acids, and allowed DNA replication to proceed. He crosslinked cells, then purified nucleosome particles and analyzed their density. He found that H3/H4 tetramers and H2A/H2b dimers were constituted of entirely dense or entirely light monomers, but that dense H3/H4 tetramers could mix with light H2A/H2B dimers and vice versa (Jackson, 1988). Thus, the nucleosome core tetramer $(H3/H4)_2$ did not dissociate and mix randomly with available H3 or H4 monomers during DNA replication.

Since Jackson's original density labeling experiments, this model has been tested in at least two other experiments. In one particularly clever experiment, Prior *et al* produced recombinant histones *in vitro*, and modified them to contain a proximity-dependent dye at their H3-H3 binding interface (Prior et al., 1980). This dye had the property that if it were near another dye molecule, it would fluoresce green, whereas if it were isolated it would fluoresce blue. They introduced these synthetic histones into *Physarum*, which is a slime mold that is uniquely able to incorporate large amounts of protein from the environment by endocytosis. In this experiment, if the mold incorporates synthetic histones into its chromosomes then replicates its chromosomes and divides, cells will fluoresce green if nucleosomes remain intact during DNA replication, or blue if they split and form heterodimers with natural histones. Their results were exactly consistent with Jackson's, in that they found no evidence for splitting at the nucleosome H3-H3 interface (Prior et al., 1980). Importantly, both Jackson and Prior were working on bulk chromatin. A locus-specific experiment could reveal corner-cases of nucleosome inheritance that break the logic established in these bulk-chromatin experiments.

Whereas Jackson and Prior separately discovered – using very different methods and working in very different organisms – that nucleosome dyad splitting is not a generic feature of DNA replication, at least two other groups discovered that nucleosome dyad splitting can be observed under certain conditions *in vivo*. First, a group working in yeast discovered that Asf1 is involved in activation of the *PHO5* and *PHO8* promoters upon phosphate limitation (Korber et al., 2006). Asf1 is a protein that binds the H3-H3 dyad axis, and thus if Asf1 binds chromatin, it very likely involves disruption of the H3-H3 axis of symmetry. Asf1 is present during transcription and at the replication fork, and plays a role in normal inheritance of transcriptional silencing at *HML*, but it has never been established whether Asf1 at the replication fork acts in disassembling old nucleosomes or whether its role is limited to assembling new nucleosomes from histone H3-H4 dimers (Schulz and Tyler, 2006; Schwabish and Struhl, 2006). The most recent attempt to resolve the point used density labeling and mass spectrometry, and reported that nucleosomes containing the active histone variant H3.3 but not the silent histone variant H3.1 split, allowing mixing between heavy and light-labeled histones (Xu et al., 2010). Furthermore, H3.3 splitting events were rare in the absence of replication, suggesting that H3.3 splitting might be a unique feature of DNA replication (Xu et al., 2010). This phenomenon was observed in HeLa cells and has not been replicated in a more natural cell line, and so should be interpreted cautiously.

So far I have discussed attempts to understand whether an individual nucleosome splits, transmitting half of its information to each strand during DNA replication. A preponderance of the evidence indicates that nucleosomes do not generally split during DNA replication, suggesting that a single nucleosome cannot transmit information to both daughter strands during DNA replication. A related model says that whereas an individual nucleosome cannot split and transmit information to both strands, a pair of nucleosomes will tend to be inherited, such that alternating nucleosomes will be sent to the leading or the lagging strand during DNA replication. The earliest SV40 replication experiments sought to address this question, by counting nucleosome density on psoralen-crosslinked replicating SV40 minichromosomes. Multiple groups found that nucleosome density behind the replication fork was approximately half the density compared to ahead of the fork, however EM images lacked sufficient spatial resolution to make strong claims about alternating strand choice (Cusick et al., 1984; Sugasawa et al., 1992). Instead density labelling experiments by Vaughn Jackson made the first strong attempt to test the model that old nucleosomes are inherited to alternating strands during DNA replication. To test the model, Jackson synchronized cancer cells, then transferred them into medium containing density-labeled amino acids. However unlike previous experiments in which be purified single nucleosomes, in these experiments he purified pairs of nucleosomes (disomes) occupying ~350bp of DNA. He then asked whether the density of pairs of nucleosomes reflected the presence of tracts of all-old or all-new nucleosomes, or whether pairs of nucleosomes reflected random inheritance, in which one would expect one quarter of nucleosome pairs of be all-old, one quarter to be all-new, and one half to be mixed. Jackson found that the density of nucleosome pairs reflected random strand choice at the replication fork, such that on average each nucleosome has a ½ chance of having a neighbor-nucleosome that is either older or younger than itself (Jackson, 1988).

If we accept a strict interpretation of Jackson's data, and suppose that each nucleosome randomly choses its strand at the replication fork, without consideration of the choice of its neighbors, the probability of a tract of *n* old nucleosomes in a row follows a power-law distribution, and rare loci should exist with long tracts of all-old or all-new nucleosomes. No one has ever measured the distribution of tract-sizes for old or new nucleosomes, however Xin Chen's group recently performed super-resolution imaging of chromatin spreads of differentiating Drosophila stem cells, and identified many long tracts of asymmetrically inherited nucleosomes, suggesting that the distribution of tract-lengths likely does not follow a power law distribution (Wooten et al., 2019). Data from recent super-resolution chromatin spreads suggest a mechanism by which cells acquire asymmetric cell fates, however they should be interpreted cautiously because cells were labelled sparsely with fluorescently-labeled histones (a relatively invasive labeling approach), and it is difficult to know whether inheritance of fluorescently labeled histones accurately reflects the inheritance of unlabeled histones. Furthermore, related experiments from *Drosophila* stem cells using similar fusion proteins, if strictly interpreted, require either that each chromosome contains only a single origin of replication, which is not likely based on measured rates of DNA replication and measured rates of cell division, or that multiple replication forks can coordinate their inheritance patterns of ancestral histones, which is difficult to imagine (Tran et al., 2012).

Interpreting the Jackson's and Chen's data together, it seems likely that most nucleosomes choose a strand randomly, but under specific conditions it is possible to bias nucleosome inheritance towards one strand or the other during DNA replication. Two recent papers working in two different organisms with two different biochemical approaches

have identified two different complexes that have a role in distributing old nucleosomes to either of the two daughter strands during DNA Replication. The work – by Anja Groth's lab working in mouse cells and by Zhigao Zhang's lab working in yeast – separately discovered that whereas in wild-type cells, nucleosomes associate evenly with the leading and the lagging strand, in cells that contain mutant *dbp3*, *dpb4* or *mcm2-3A*, old nucleosomes are non-randomly inherited to the leading and lagging strand, respectively, during DNA replication. Based on the data from Groth and Zhang, it is likely that super-resolution microscopy chromosome spreads in *dpb3, dpb4,* or *mcm2-3A* mutants would replicate the Chen lab's work in differentiating *Drosophila* stem cells. It is possible that by regulating the activity of Dpb3, Dpb4, or Mcm2, a cell could exert control over old-nucleosome strand choice, and that cells might do so in a locus-specific manner.

Whereas the work by the Groth and Zhang labs makes a clear argument that nucleosome strand-choice is regulated, the work does not address whether a nucleosome remembers its position along that DNA strand. Whereas Groth and Zhang both relied on tag-swap designs to separately study old and new populations of histones, understanding whether a nucleosome is inherited *in cis* requires some locus-specific labeling of ancestral proteins: bulk labeling is not sufficient. There have been two experiments published recently that sought to determine whether a nucleosome is inherited *in cis*, to understand whether nucleosomes transmit epigenetic memory. The first results were published simultaneously by the Moazed lab and the Allshire lab. Using effectively the same strategy, the two labs built synthetic heterochromatin in *Schizosaccharomyces pombe* by tethering the *pombe* H3-K9 methylase Clr4 to a synthetic TetO sequence in a strain mutant for *epe1,* which encodes the H3-K9me demethylase. In wild-type cells, H3-K9 methylation is established a defined locus, then propagated by Clr4, which can both bind to methylated H3-K9 and methylate H3-K9 on adjacent histones (Ragunathan et al., 2015). To ask whether a chromatin domain is inherited *in cis*, they first established silencing by tethering TetR-Clr4 to the TetO sequence, then released TetR to prevent new Clr4 recruitment. Then, they followed cells through cell division in the presence or the absence of wild-type Clr4. They found that in the absence of new TetR-Clr4 recruitment, Clr4 cells maintain the silent state through cell division, but *clr4* mutant cells lose silencing after cell division (Audergon et al., 2015; Ragunathan et al., 2015). Both experiments depended on the use of an *epe1* mutant, which is broadly deficient in histone demethylation, including on residues that are not targeted by Clr4. Ignoring possible caveats arising from the use of the *epe1* mutant, the published result requires that nucleosomes at the silent locus – which spanned ~45kb – were not completely lost from the locus: if they had been lost, reader-writer based inheritance would fail. However it is impossible to estimate based on their experiments how many nucleosomes must be retained *in cis* to propagate epigenetic silencing: the published data is compatible with the idea that if only a very small number of nucleosomes remembered their location on a 45kb interval, it would be sufficient to re-establish H3-K9-mediated silencing, because Clr4 can cause chromatin domains to expand from a nucleation site by iterative rounds of H3-K9me deposition and spreading.

To understand whether the Moazed and Allshire results reflect efficient nucleosome inheritance *in cis* or some less precise mechanism of position memory, it is necessary track nucleosomes with much spatial resolution than can be done using tethered Clr4. To address this question, the Jane lab purified SV40 circular DNAs with an integrated synthetic Widom 601 nucleosome positioning sequence array. Then, they loaded nucleosomes onto the SV40 chromosomes in a stoichiometry that resulted in a single nucleosome per SV40 molecule, positioned over the Widom sequence. They found that after SV40 DNA replication,

nucleosomes that were deposited at the Wydom sequence remained associated with the SV40 molecule after DNA replication, however each nucleosome's position was randomized along the ~2.5kb SV40 minichromosome (Madamba et al., 2017). Thus in the Jane lab's experiments, nucleosomes are passaged *in cis* during DNA replication, but not in a way that perfectly retains its position in the genome; a nucleosome could move 1kb or more during passage of the replication fork. It is difficult to know if the results from the sparsely chromatinized *in vitro* assembled chromatin reflect what would occur on true chromatin templates: estimates of nucleosome-free regions during chromatin replication suggest that ~1-2 nucleosomes are disrupted around a replication fork, however in the sparsely-chromatinized SV40 the naked DNA around the replication fork reflects ~13 unoccupied nucleosome binding sites (Gruss et al., 1993).

2.2 Outstanding ideas about histone dynamics during transcription

DNA participates in two major chemical reactions in cells: replication and transcription. So far, I have discussed the relationship between chromatin and DNA replication, however RNA transcription poses a related set of questions about the behavior of nucleosomes during DNA transactions.

There are clear similarities between the constraints that affect DNA replication and RNA transcription: both processes involved large protein complexes that move uni-directionally along a DNA fiber, and both processes require access to the nucleobase core of DNA, which requires melting the DNA strand along its axis of symmetry. However, during transcription, unlike during replication in which nucleosomes are inherited along both the old and the new nucleotide polymer, the nascent RNA does not compete with its DNA template for nucleosomes. Furthermore, whereas both DNA replication and RNA transcription rely on topoisomerases to relieve torsional strain along the DNA template, there may be very different requirements for this activity during DNA replication – which occurs once per cell cycle, and involves incredibly long fibers – and RNA transcription, which can occur hundreds of times per cell cycle at a given locus, and produces a comparatively short polymer[3]. In short, there is not a clear expectation that nucleosome behavior at the DNA replication fork should be instructive of nucleosome behavior at the site of RNA transcription.

Biochemical experiments studying transcription through nucleosomes have centered on two questions: are nucleosomes displaced during RNA transcription? And do nucleosomes move relative to DNA as a result of transcription? In one attempt to address this question, Vasily Studitsky loaded 227bp DNA templates with nucleosomes, then initiated transcription along the DNA. Studitsky performed DNase footprinting assays on the DNA-nucleosome complex before and after RNA transcription, and discovered that nucleosomes moved in a retrograde direction relative to RNA transcription along such templates (Studitsky et al., 1994). When they performed analogous experiments with a 262bp

[3] This observation is true in yeast, and is generally true in other organisms. Notably, some extremely long transcripts in animals – including DMD and Titin can significantly longer than the inter-origin distance, and can make transcripts as long as a yeast chromosome (>270kb). Such transcription units might be subject to special topological or torsional constraints to accommodate the exceptionally long RNA polymer.

piece of DNA, they discovered that the result depended on where exactly the nucleosome was loaded on the template: some templates exhibited retrograde transposition, and others showed a less coherent pattern of nucleosome movement during transcription (Studitsky et al., 1994).

If nucleosomes move during transcription, it could reflect two different mechanisms for the interaction between nucleosomes and transcription: nucleosomes could dissociate from DNA then re-associate behind RNA polymerase, or nucleosomes could be partially unwrapped to allow passage of RNA polymerase, without completely dissociating from the DNA. Two groups using very different approaches have established *in vitro* that nucleosomes do not dissociate during RNA transcription. In a first set of experiments, 227bp DNA templates were loaded with nucleosomes, and RNA was produced from the chromatinized DNA in the presence of non-chromatinized DNA templates. In this design, nucleosomes are not incorporated in the naked DNA during transcription, and do not dissociate from the chromatinized template during passage of T7 RNA polymerase (Studitsky et al., 1994). Thus, the T7 RNA polymerase can transcribe through a nucleosomal barrier without causing the nucleosome to be completely removed from the DNA. In a later set of experiments, a ~3kb DNA template was loaded with a single nucleosome, then stretched gently between two beads using optical tweezers. RNA transcription was from the molecule using recombinant *S cerevisiae* RNA polymerase II, and RNA polymerase transcribed until it encountered the nucleosome. After it encountered the nucleosome, it paused until random fluctuations in the nucleosome's position allowed it to enter the nucleosome-bound DNA, but it was able to successfully transcribe through the nucleosome without displacing the nucleosome altogether (Hodges et al., 2009). In the optical trap design, nucleosome loss would be immediately detectable in the increased length of the template molecule stretched between the two beads. Furthermore the optical trap experiments demonstrated that thermal fluctuation in RNA polymerase position could result in occasional backwards slipping of RNA Polymerase along the template, but that backwards slipping is limited by nucleosomes (Hodges et al., 2009). This suggests that a possible role of nucleosomes during transcription is to serve as a ratchet that allows RNA polymerase to move forward as energy is invested in making the RNA chain, but not slide backwards when thermal fluctuations might cause polymerase slipping.

Whereas the data are compelling that RNA polymerase can transcribe DNA without displacing nucleosomes *in vitro*, experiments *in vivo* have shown that nucleosome density is anti-correlated with transcription, such that highly transcribed genes have reduced nucleosome density compared to lowly transcribed genes (Lee et al., 2004).

One experiment attempted to understand whether nucleosomes are moved proccessively by DNA replication, by performing a tag-swap on histone H3, then tracking the position of ancestrally tagged histones. They found that old nucleosomes are enriched at the 5' ends and under-represented at the 3' ends of actively transcribed genes after several generations of growth, and argued that the pattern supports the biochemical observations that nucleosomes move in a retrograde manner during transcription (Radman-Livaja et al., 2011). Furthermore, they found that the 5' – 3' asymmetry in ancestral nucleosomes was limited in a *top1*Δ mutant, which they interpreted to mean that torsional strain induced by transcription opposes the natural retrograde motion of nucleosomes along the DNA fiber. Because tag-swap experiments were performed in cycling cells, these experiments do not effectively isolate the effect of DNA replication from the effect of transcription, and so it is difficult to understand whether arguments that *TOP1* regulates 5' end accumulation of old

histones is a consequence of transcription alone, or whether it is a consequence of the interaction between DNA polymerase and RNA polymerase during S-phase.

Possible interactions between DNA replication forks and RNA polymerase have been speculated to constrain genome organization (Brewer, 1988). For example in the SV40 virus, genes are oriented such that RNA polymerase and DNA polymerase move in the same direction along the chromosome, to avoid head-on collisions (Seidman et al., 1979). In early attempts to understand whether nucleosomes are transmitted to both daughter strands during DNA replication, SV40 replication was monitored in extracts and groups occasionally found contradictory results – either that nucleosomes associated randomly or non-randomly with the replicated DNA strands (Cusick et al., 1984; Riley and Weintraub, 1979). This discrepancy anchored what I consider to be one of the cleverest experiments I've ever seen, in which Michael Seidman, Arnold Levine and Harold Weintraub asked whether the basis for the contradictory result might be occurrence or non-occurrence of transcription in the various extracts in which the SV40 replication was performed. In their paper, Seidman, Levine & Weintraub cultured a cell line that contained an integrated SV40 virus at a single locus. They inhibited new histone synthesis and simultaneously radiolabelled the newly synthesized DNA, and they digested the chromatin with nuclease to reveal where the ancestral nucleosomes were bound, then melted and hybridized the nucleosome-protected DNA to single-stranded DNA that reflected either the Watson or the Crick strand of the SV40 chromosome. They found that the nucleosome protected DNA hybridized to one strand but not the other, suggesting that all old nucleosomes were inherited on one strand during DNA replication. Because transcription in SV40 is oriented in the same direction as DNA replication, they wanted to know whether nucleosome inheritance on the leading strand during DNA replication was a phenomenon of leading strand synthesis, or if it was a consequence of the fact that the leading strand in SV40 synthesis also corresponds to the coding strand for SV40 transcription. To test this notion, they inhibited protein synthesis in chicken cells with cyclohexamide, and labeled new DNA synthesis with ^{3}H-thymidine. They then digested chromatin to obtain nucleosome-protected lagging strand Okazaki fragments, and hybridized the ^{3}H-thymidine, nucleosome-protected DNA to nuclear chicken RNA. Seidman, Levine & Weintraub found that newly synthesized Okazaki fragments did not hybridize with RNA, which they interpreted to mean that nucleosomes were only inherited on the lagging strand if the lagging strand corresponded to the coding strand for RNA synthesis (Seidman et al., 1979). Though the experiment is just about as elegant an experiment as I can imagine by the standards of 1979, their experiment is very contrived in the sense that by inhibiting translation altogether as a way to identify the position of ancestral nucleosomes, they may have inhibited the synthesis of nucleosome segregation factors that oppose the asymmetry they observed. In fact, modern experiments that have labeled new DNA synthesis and new histone synthesis simultaneously (without inhibiting transcription) have not replicated Seidman's observations (Yu et al., 2018).

2.3 The rest of the nucleosome

The $(H3\text{-}H4)_2$ tetramer is referred to as the nucleosome core particle. The remainder of the nucleosome is comprised of two heterodimers of H2A and H2B. H2A and H2B are considered to be peripheral in the nucleosome because they are assembled on DNA after assembly of the H3/H4 tetramer on DNA, and have not been implicated in long-term

memory of transcriptional state (Kaufman and Rando, 2010; Smith and Stillman, 1991). Whereas most covalent modifications that reflect long-term silencing are localized on H3 or H4, H2A and H2B are phosphorylated, ubiquitinated or replaced altogether during DNA damage repair and transcription (Bannister and Kouzarides, 2011). It is widely reported that H2A/H2B dimers are displaced during DNA replication and are slower to associate with daughter strands after passage of the replication fork compared to H3/H4. This model originated in the observation that H3/H4 deposition is coupled to DNA replication in cell extracts, but that H2A/H2B can be incorporated after DNA replication (Smith and Stillman, 1991). However, when SV40 minichromosomes are assembled into chromatin using crosslinked nucleosome octamers, DNA replication can proceed without displacing nucleosomes, suggesting that H2A/H2B dissociation is not a requirement for DNA replication (Vestner et al., 2000). Furthermore, during transcription there is regulated rearrangement of the H2A/H2B dimers, in which the FACT complex (**FA**cilitates **C**hratomatin **T**ranscription) aids RNA polymerase in transcribing through nucleosome barriers (Belotserkovskaya et al., 2003; Hsieh et al., 2013). Surprisingly, H2B accepts extremely large protein tags, and H2B-RFP protein fusions are frequently used to mark chromosomes for live imaging experiments (Bothma et al., 2014). In unpublished research, H2B was found to exchange readily between the DNA-bound and the freely diffusing form, with an average residency time of 30s on DNA (Carl Wu, unpublished observation). Thus there is little expectation that H2A/H2B should store information for any timescale longer than 30s, however recent theoretical work on nuclear search mechanisms predict that H2A/H2B that has been released from DNA might re-bind DNA with little lateral diffusion (Woringer and Darzacq, 2018). Thus, it is possible that H2A/H2B might show some kind of local position memory even if H2A/H2B can exchange between the bound and soluble state.

2.4 Outstanding ideas about how nucleosomes affect gene expression

If nucleosomes remember their position during DNA replication, what information do they transmit from one generation to the next? Since the formulation of the "Histone Code" hypothesis, it has been widely accepted that histone modifications are instructive for gene expression. In fact the only histone post-translational modification that is known to have an autonomous biochemical consequence on DNA-nucleosome transactions independent of any of the chromatin readers is H3-K56-ac, in which H3 is modified at a point of contact between H3 and the DNA, which results in a less stable DNA wrap around the nucleosome (Zhang et al., 2018). But the clear finding from structural work on the nucleosome is that most of the best-studied chromatin modifications, whose presence is correlated with gene expression, are peripheral to the nucleosome and are unlikely to have autonomous effects on gene expression. Instead, most histone marks work by recruiting interacting proteins, and those proteins exert the effect of the histone modifications. Thus, to answer the question: does a histone mark transmit memory of its regulatory state, we must understand the answer to the question: can histone marks autonomously recruit interacting proteins after DNA replication?

To address this question, I consider below some of the major varieties of chromatin modification separately. First, the major mark of silent chromatin in eukaryotes (though

notably not in *S. cerevisiae*), H3-K9-me, enforces long-term gene silencing during development, and typically as a consequence of RNAi (Canzio et al., 2014). Chromatin that is assembled *in vitro* and modified as H3-K9-me is bound by the effector protein HP1, and the resulting complex prevents transcription (Greenstein et al., 2018). In addition to binding HP1, H3-K9-me binds a bifunctional enzyme called Clr4 that can both bind H3-K9-me and deposit methylation on adjacent H3-K9 moieties (Audergon et al., 2015; Ragunathan et al., 2014). Thus, it seems reasonable that if a histone were inherited with H3-K9-me, it would be capable of recruiting proteins to re-establish chromatin state in its vicinity and to effect silencing of the genes in the region, and this property has been observed *in vivo* (Audergon et al., 2015; Ragunathan et al., 2015). Importantly, experiments studying the ability of H3-K9-me to template its own re-establishment were performed in a genetic background deficient in histone demethylation. In wild-type cells, histone demethylation out-competes Clr4 propagation of the H3-K9-me state, and HP1 binding is lost after DNA replication. Therefore, although H3-K9-me can autonomously bind the heterochromatin effector protein HP1, and although it can template its own replication in some genetic backgrounds, it is not autonomously capable of propagating memory of the silent state indefinitely in wild-type cells. Perhaps it's fair to describe this context as a kind of "short term" epigenetic memory that requires maintenance after every round of DNA replication to faithfully transmit memory of silencing.

In *S. cerevisiae* silent mating-type-locus regulation, silencing is enforced not by H3-K9 but instead by the absence of covalent modifications on H4-K16, another histone tail residue that does not directly contact DNA. Like H3-K9me, H4-K16 acts through its binding partner: the Sir complex. However, unlike H3-K9, which effects silencing by directly binding an effector protein in its modified form (H3-K9me), H4-K16 effects silencing by binding the Sir complex in its *un*modified form. In heterochromatin, H4-K16 is subject first to H4-K16 acetylation by Sas2, then H4-K16 deacetylation by Sir2, which might be directly coupled to Sir complex binding of histone tails (Gartenberg and Smith, 2016). Thus the "silent mark" in yeast mating-type regulation is not a histone modification, but rather a *history* of histone modification, in which H4-K16 is first acetylated then subsequently de-acetylated. Based on this observation, it is difficult to argue that H4-K16 can autonomously replicate its chromatin state, or autonomously recruit its effector proteins: if H4-K16 had those abilities, it would autonomously nucleate silencing at many loci throughout the genome. Therefore, even if a silent histone at *HML* or *HMR* remembers its position through DNA replication, it might nonetheless fail to transmit memory of the silent state. Recently, it has been argued that some types of heterochromatin can form an unstructured (and possibly phase separated) aggregate (Larson et al., 2017). If such an aggregate is present at yeast silent mating loci, it is possible that within such an aggregate the high local concentration of Sir complex members could directly bind H4-K16 without the need for acetylation and deacetylation, however there is no evidence of a phase-separated aggregate at yeast silent mating type loci.

In addition to H3-K9me and H4-K16, which are present at silent loci in *S. pombe* and *S. cerevisiae*, respectively, some histone modifications reflect active transcription. One mark that is associated with active transcription is H3-K56ac. Whereas H3-K9me silences transcriptions by binding effector proteins, H3-K56ac is thought to have an intrinsic and autonomous consequence for DNA-nucleosome interactions. Specifically, nucleosomes modified at H3-K56ac have a weaker interaction with the DNA backbone, and might be less stably associated with DNA *in vivo* (Zhang et al., 2018). Although H3-K56ac is associated with transcription, the enzyme that is capable of performing acetylation does so in complex with the nucleosome assembly protein Asf1 (Chih et al., 2005; Zhang et al., 2018). It is

possible that the occurrence of H3-K56ac at actively transcribed genes reflects nucleosome turnover, and deposition-dependent acetylation, rather than some transcription-dependent nucleosome modification. Thus although H3-K56ac has a biophysical, autonomous and intrinsic effect on nucleosome stability, that effect could be irrelevant to the correlated biological activity and could simply reflect transcription rather than instructing it.

3 Tracking histones

To understand whether histones can transmit information, I designed an experiment that could falsify the model that histones remember their position through DNA replication. The experiment required me to track single histones kinetically – it needed to be a kind of pulse chase assay, in which I only labeling a spatially restricted subset of nucleosomes, then following their location in subsequent time points. The first concept I had in designing the experiment was to use a natural histone modifier that is not found in *S cerevisiae*, targeted to a DNA sequence by dCas9. I initially considered building fusion proteins that direct the *S pombe* H3K9 methylase Clr4 to an arbitrary locus using dCas9. This approach could have successfully methylated histones at a defined locus, however there was no obvious way to stop the labeling reaction, without developing a Clr4 inhibitor or making some kind of novel, regulatable allele of either Clr4 or dCas9. On Debbie Thurtle-Schmidt's recommendation, I explored the possibility of using fusion proteins to the *E coli* biotin ligase BirA to direct biotinylation of histones. At the time, the Weissman lab had recently published a ribosome profiling experiment in which they biotinylated ribosomes associated with the endoplasmic reticulum and purified the biotinylated ribosomes to analyze which mRNAs were translated into the lumen of the endoplasmic reticulum separately from mRNAs encoding cytoplasmic proteins. Importantly for our purposes, the Weissman lab demonstrated that biotinylation by BirA could be regulated by modulating the concentration of biotin in the culture medium, suggesting that BirA could be regulated in a way that was impossible for Clr4.

To achieve local biotinylation of histones I planned to fuse a biotin acceptor peptide – the so-called Avi-tag – to every nucleosome in the genome, then use dCas9 fusion proteins to direct the biotin ligase BirA to an arbitrary locus and biotinylate the Avi-tagged histones present at that locus. The dCas9-BirA fusion strategy was a total mess for many reasons, notably that it contained too many moving parts to reasonably allow for iteration on the experimental design. I discuss the many failings of my dCas9-BirA technology development strategy in the subsequent chapter of this thesis.

In the next section, I discuss aspects of technology development that enabled me to track histones *in vivo*, and I describe how I used the technology to answer the question: do nucleosomes remember their position through DNA replication and transcription?

3.1 Labeling heterochromatin domains with Sir4-BirA

After the flameout of my attempt to develop dCas9-BirA fusion proteins, the project received new life when Davis Goodnight joined the lab for a rotation in my third year. Whereas I had previously been attempting to label one nucleosome at an arbitrary locus, Davis's rotation project was to build Sir3-BirA and Sir4-BirA fusion proteins that would biotinylate histones at silent Sir-complex-bound chromatin, including *HML, HMR,* and the telomeres. The advantage to Sir3/Sir4-BirA fusion proteins compared to dCas9-BirA is that we had excellent ChIP-seq data for Sir3 and Sir4, which led us to expect that the proteins should be efficiently localized at their binding sites, and they targeted relatively large regions of the genome compared to the domain targeted by dCas9, which could make pulldown chemistry simpler.

In Davis's rotation, he developed plasmids that encoded Sir3-BirA and Sir4-BirA and found that the fusion proteins complemented *sir3*Δ and *sir4*Δ respectively, suggesting that

the -BirA addition did not affect localization or function of Sir3 and Sir4. Furthermore, Davis showed by ChIP-qPCR that Sir3-BirA biotinylates *HML* more readily in rich media than when cells are grown in the presence of nicotinamide, an inhibitor of heterochromatin function in *S. cerevisiae*. Thus, based on Davis's work it was clear that it would be possible to achieve site-specific biotinylation of nucleosomes *in vivo*.

In addition to demonstrating that BirA fusion proteins could achieve site-specific biotinylation of histones, Davis explored whether biotin concentration could be used to regulate the activity of BirA, as had been previously demonstrated by the Weissman lab. To test the approach, Davis grew strains in limiting biotin, then added biotin to the culture medium for one hour and analyzed biotinylation by streptavidin blot. He found that biotinylation of H3-Avi increased dramatically after addition of biotin to the culture medium, suggesting that biotinylation could be kinetically controlled.

Davis's contribution to the technology development process was twofold – first, mentoring Davis as a rotation student compelled me to think through each experiment more carefully than I had heretofore. As long as I was working alone, I felt like I could cut corners and skip controls to speed up the iteration time; however, when I was advising Davis on how to execute the experiments I was considerably more structured in my thinking, and expected Davis to perform multiple controls at every step. This sense of accountability continued after Davis stopped working on the project, and it made all of the subsequent experiments more rigorous. Second – and this can't be overstated – Davis was the first person to present convincing data that site-specific biotinylation could be achieved *in vivo*, and all of the experiments that followed built on the positive control that Davis developed.

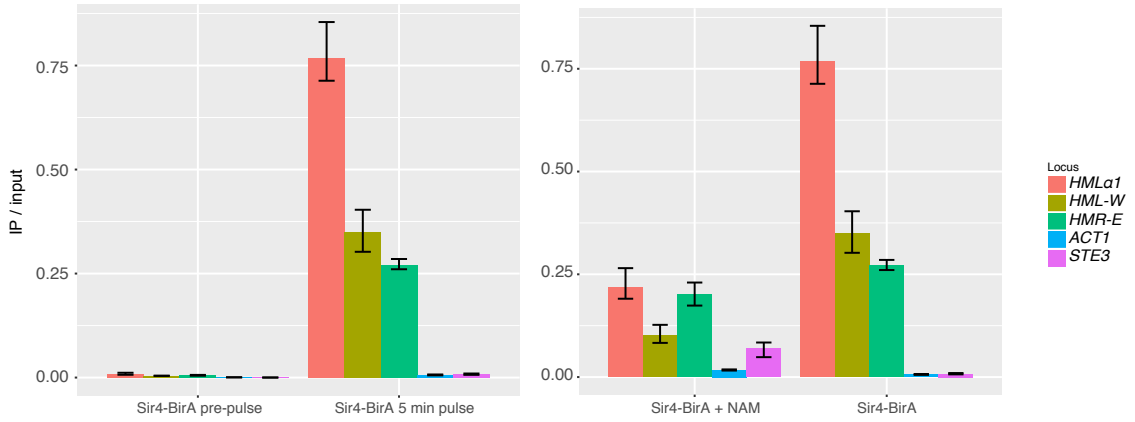

Figure 3-1 – (left) A 5 minute biotin pulse can direct biotinylation of histones at HML and HMR ~100-fold over pre-pulse levels. (right) Biotinylation during a biotin pulse is severely curtailed when Sir4-BirA is delocalized from silent chromatin by pre-growing cells in NAM.

Based on the plasmids Davis developed, I built strains in which Sir3-BirA and Sir4-BirA were encoded in the genome, and I tagged the two H3 genes of yeast, *HHT1* and *HHT2,* at the endogenous loci with the Avi-tag and replicated Davis's findings. I found that biotinylation by Sir4 could be induced ~100-fold by pulsing biotin from 1nM to 200nM for ~5 minutes (Fig 3-1). Furthermore, I found by ChIP-qPCR that Sir4-BirA biotinylated nucleosomes at *HML*, and that biotinylation at *HML* was reduced ~75% in the presence of nicotinamide (Fig 3-1).

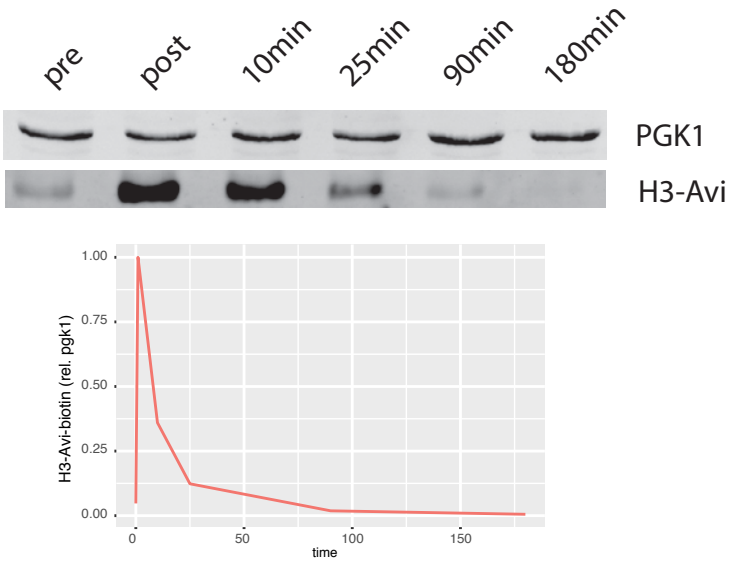

Figure 3-2 - (top) immunoblot for Pgk1 and streptavidin blot for H3-Avi-Biotin showed rapid turnover of biotinylated histones after a transient biotin pulse. Cells were exposed to 20nM biotin for 5 minutes, then washed 3 times in water and resuspended in low biotin medium. (bottom) The same data from the top panel, quantified and normalized to Pgk1 loading control.

Whereas Davis and I separately found that biotinylation can be induced by adding a pulse of free biotin, the pulse-chase experiment required me to stop biotinylation at an appointed time and chase the ancestral, biotinylated nucleosomes. To test whether biotinylation can be restricted to brief window during a biotin pulse, I grew cells with limiting biotin in the medium (~0.3 nM biotin), then spiked in biotin to >20nM for 15 minutes and washed out the biotin by filtration. From the first time I did a biotin-washout

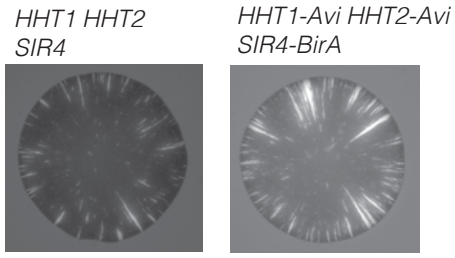

Figure 3-3 – Cells expressing the CRASH assay capture loss of silencing events at *HML* as fluorescent sectors. Representative colonies of wild type or H3-Avi Sir4-BirA strains showed that H3-Avi and Sir4-BirA had a minimal impact on the stability of silencing at *HML*

experiment, it was clear the technique was problematic. In early experiments, I found that a biotin pulse corresponded to an increase in H3-Biotin by immuno-blot, but that H3-Biotin turned over more rapidly than one would expect based on simple dilution of biotinylated histones during cell division (fig 3-2).

In principle the observed loss of signal could reflect that the biotinylated histones are degraded, or that biotin is removed from the Avi tag. Because silencing (as measured by the CRASH assay and by mating) was virtually wild type in strains with biotinylated *HML* and *HMR*, I did not think it was likely that nucleosomes were being degraded, and so I chose to explore the possibility that biotin was being enzymatically removed from the Avi tag after being deposited (fig 3-3). To examine this possibility, I studied the pathways that use biotin in yeast for a clue about what enzymes might be able to break a biotin-lysine bond.

S cerevisiae uses biotin as a cofactor to perform carboxylation reactions, and the two major biotinylated enzymes in yeast are pyruvate carboxylase and acetyl-coA carboxylase. There is an additional protein, Arc1, that is biotinylated, but whose function is unknown. W303 and other common laboratory strains are mutant in the biotin synthesis genes *bio1* and *bio6,* and so are auxotrophic for biotin (Hall and Dietrich, 2007). Many organisms, including humans, encode a biotinidase that can recycle biotin by cleaving biotinylated lysine residues, however a yeast biotinidase has never been discovered. I speculated that if there were a yeast biotinidase, it might be up-regulated in low biotin conditions: yeast pre-grown in low biotin, then exposed to a brief pulse of biotin to biotinylate histones, might show high biotin turnover if a highly expressed biotinidase quickly removes biotin from histones for use in other yeast metabolic processes. Based on homology to known biotinidase enzymes from *Drosophila*, I identified three homologous genes in yeast – *NIT1, NIT2* and *NIT3*– which were reported to encode enzymes that perform reactions similar to the reaction that removes

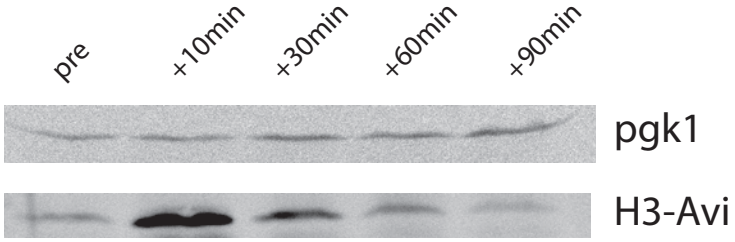

Figure 3-4 Cell expressing *nit1*Δ*0 nit2*Δ*0 nit3*Δ*0* still show rapid H3-Avi-Biotin turnover after a biotin pulse, suggesting that NIT1, NIT2 and NIT3 were not responsible for degrading biotinylated H3-Avi.

biotin from lysine. To determine whether any of the homologs remove biotin from lysine, I made a triple mutant strain with clean deletions *nit1*Δ*0 nit2*Δ*0 nit3*Δ*0*, using Cas9, and tested whether the triple mutant showed reduced turnover of biotinylated H3. Unfortunately, the triple mutant showed the same rapid turnover after a pulse, suggesting that neither *NIT1, NIT2* nor *NIT3* encodes a biotinidase that removes biotin from H3-Avi. (fig 3-4).

Never one to give up on a doomed course of action, I decided to broaden my search for a yeast biotinidase beyond candidate genes. I reasoned that if a biotinidase removes biotin from a target protein, it might be transiently biotinylated, and furthermore that if the biotinidase acts on H3-Avi-Biotin, that the abundance of the biotinylated species should increase in conditions in which H3-Avi is biotinylated. In fact, I observed one protein that

was only biotinylated in conditions in which H3-Avi was also biotinylated and could be detected as a band at ~35 kD on a streptavidin blot (Fig 3-5). To identify the mystery band, I purified proteins that bound to streptavidin and submitted the purified protein for 2-D mass spectrometry. The mass spectrometry experiment confirmed the identity of the major biotinylated proteins – pyruvate carboxylase, acetyl-coA carboxylase, Arc1 and histone H3, but did not reveal an obvious candidate biotinidase with a mass near ~35kD.

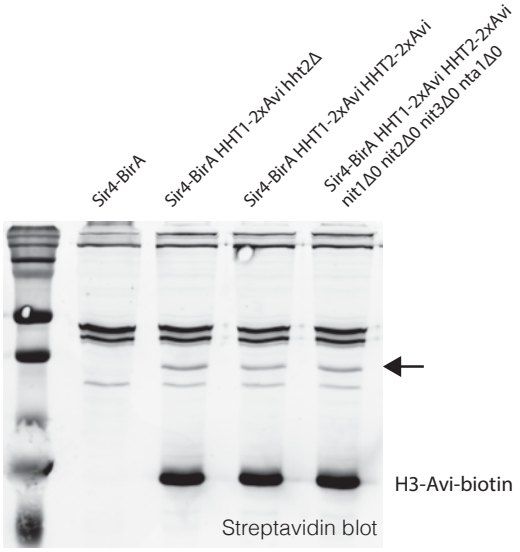

Figure 3-5 – In conditions that biotinylated H3-Avi, a mysterious biotinylated band emerged around ~35kD. I speculated this band is a biotinidase, that is transiently biotinylated as it removes biotin from H3-Avi, but I could not identify the protein. Homology led me to consider NIT1, NIT2, NIT3 and NTA1 as possible biotinidase genes, and I unsuccessfully attempted to identify the band with 2-D mass spectrometry.

Having failed again in my attempt to identify a yeast biotinidase, I devised a final strategy to explore the possibility that biotin is being enzymatically removed from H3-Avi after a biotin pulse. I devised a mutagenic screen, in which I grow cells on CSM + biotin, then replica-plate them onto CSM – biotin + biocytin. Biocytin is a small molecule that is equivalent to biotinylated lysine, and cells that cannot cleave the biotin from the biocytin should fail to grow on the selection. I would screen for failure to grow on biocytin media, and backcross to map the causative mutation.

Although the screen was straightforward, I never ultimately executed it. Concurrently with my efforts to identify a yeast biotinidase, I had been exploring alternative designs for the kinetic labelling strategy that could avoid complication caused by a biotinidase, and those experiments gave me reason to believe that I could achieve my experimental goals without finding the yeast biotinidase.

My search for the yeast biotinidase was motivated by my pilot pulse-chase experiments, in which cells were pre-grown in low biotin medium then briefly exposed to high concentrations of biotin. I suspected that if yeast expresses a biotinidase, it might be transcriptionally regulated by the biotin available in the cell. Thus, if I were to pre-grow cells in high biotin medium, they might not express the biotinidase. Under this model, it should be possible to perform washout experiments—rather than pulse-chase experiments—and achieve kinetic tracking of ancestral histones, while avoiding the possibility of activating a

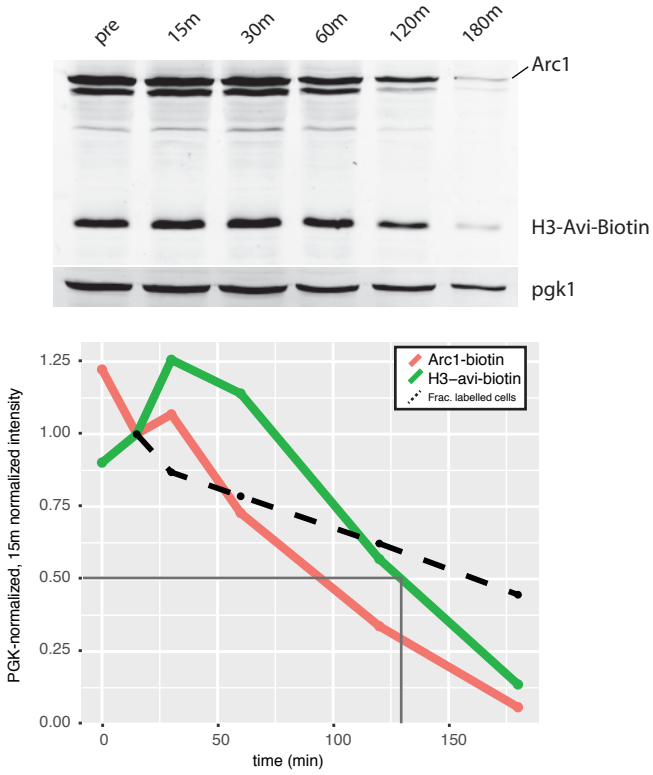

Figure 3-6 – Performing a simple step-down (rather than a biotin pulse) changed the kinetics of biotin removal from H3-Avi, such that H3-Avi-biotin is turned over at the rate predicted by dilution during cell division. This experiment provided the basis for subsequent Sir4-BirA biotin stepdown experiments

low biotin regulon. I tested this idea by pre-growing cells in 10 nM biotin, then washing out the growth medium and recovering cells in 0.33 nM biotin. I monitored the abundance of biotinylated proteins by blotting with streptavidin, and found that in a step-down experiment, dilution of biotinylated proteins – including H3-Avi and the naturally biotinylated Arc1 – was consistent with dilution by growth, without any obvious effect of a potential biotinidase (fig 3-6).

Based on the successful result from the step-down experiment, I was ready to perform the kinetic tracking experiment using Sir4-BirA to label nucleosomes in silent chromatin. To execute the experiment, I grew cells in 10 nM biotin, then arrested them in G1 (fig 3-7). Then I split the culture and held half in an extended arrest, while allowing the other half to cycle. I found that ancestrally labeled nucleosome density of *HML* and *HMR* remained locally enriched through cell division, and that the border between ancestrally biotinylated nucleosomes and adjacent, unmodified nucleosomes was unaffected by DNA replication, suggesting that nucleosomes reoccupy the same DNA sequence after DNA replication that they had occupied before replication (fig 3-7). Interestingly, two biotinylated nucleosomes at *HML* showed much more rapid turnover compared to the locus as a whole, and both nucleosomes were positioned adjacent to the Rap1 binding sites at *HML-E* and at the α1-α2 promoter. (fig 3-8)

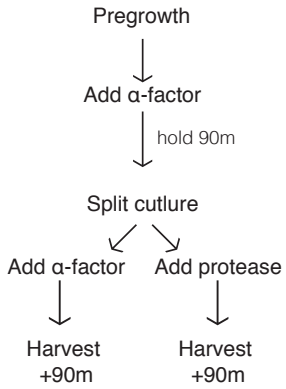

Figure 3-8 – Schematic of step-down design for Sir4-BirA

Overall, the Sir4-BirA labeling experimental result was consistent with the model that nucleosomes are re-incorporated in place after DNA replication, however the interpretation hinged on the question: is nucleosome labeling really *stopped* after I wash out biotin from the growth medium? If there were residual biotinylation activity, it would be isolate and study the population of ancestrally labeled histones. The observation that biotinylated nucleosome density decreased roughly 2-fold during DNA replication was consistent with the idea that nucleosomes remember their position, and that labeling is stopped after the biotin washout; however, it was also consistent with the model that nucleosomes forget their position and that labeling is reduced 50% after the biotin washout. I devised an orthogonal strategy to measure the relative rate of biotinylation by BirA before

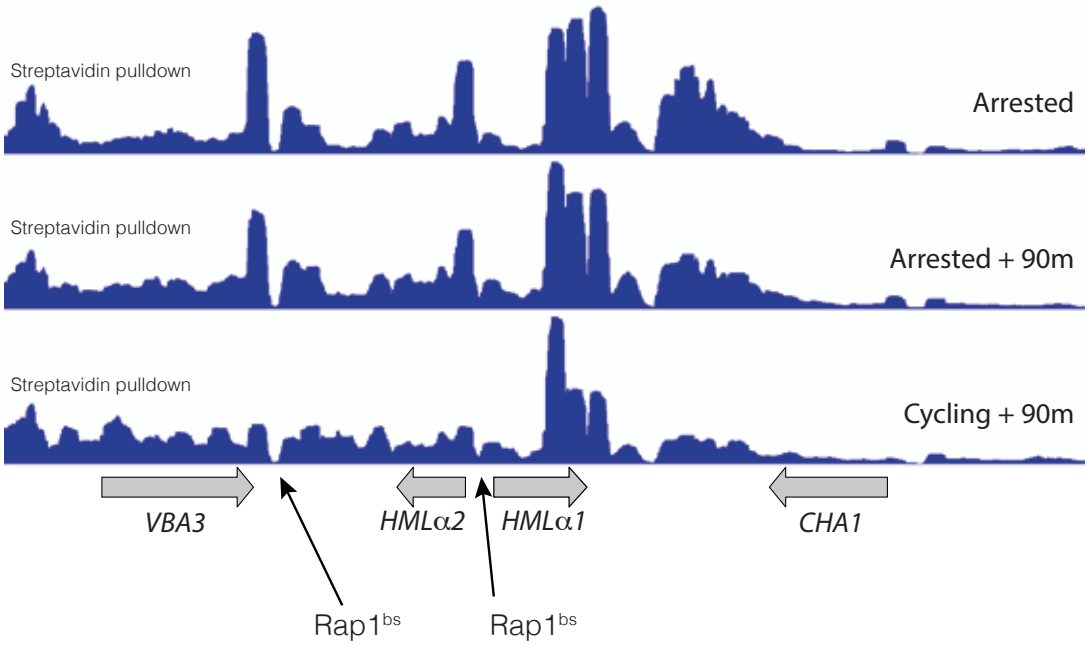

Figure 3-7- Stepdown kinetic labeling of *HML* with Sir4-BirA. Nucleosomes did not move from the silent domain into the adjacent euchromatin during the cell cycle, suggesting that nucleosomes remember their position during DNA replication. Interestingly, two strong H3-Avi-biotin peaks decreased during the cell cycle, and both corresponded to binding sites for Rap1. I suspected – but never proved – that Rap1 competed with nucleosomes for binding at its heterochromatic binding sites.

and after biotin washout, however I never got around to implementing the strategy.

At the same time my Sir4-BirA work was starting to show results, I was starting to get favorable indications from a related set of experiments, in which I labeled a smaller chromatin domain, with even more precise control over local biotinylation activity. These extensions to my Sir4-BirA experiments are the central element of my thesis and is the subject of its next section.

3.2 Labeling euchromatin loci with TetR

Sir4-BirA is expected to bind hundreds of nucleosomes at *HML*, *HMR*, and all of the telomeres. To label a smaller subset of nucleosomes, I needed a more precisely targeted tethering protein than Sir4 to direct BirA to a spatially restricted locus. Furthermore, the ideal tether would be able to bind its target in an arbitrary context, rather than being restricted to binding silent chromatin. Early tech development projects using dCas9 as a tethering protein failed, in part because of my own novice and in part because there were too many moving parts to optimize each element individually. To revisit the goal of labeling an arbitrary locus, I focused the second wave of my tech development efforts on TetR-BirA fusion proteins that would biotinylate histones adjacent to a 19bp TetO sequence. Unlike Sir4-BirA, TetR-BirA could be regulated by TetR localization instead of by biotin concentration. This was advantageous, it that it let me measure the amount of biotinylation that would accumulate in strains that are grown chronically in the "off" state (YPD + doxycycline), without affecting the normal physiology of the yeast. Additionally, strains grow more reliably in YPD than they did in CSM with limiting biotin, and it was considerably easier to time overnight experiments.

In the first version of my TetR-BirA labelling strains, I targeted five operator sites for TetO insertion. Based on Debbie Thurtle-Schmit's MNase-seq data, I targeted linker regions around the *GAL10* promoter and in the *GAL10* coding sequence. I tested the operator insertion strains by streptavidin-seq comparing strains grown in YPD to strains grown in YPD + dox, with the understanding that the ideal strain would label the *GAL10* locus only in YPD and not in YPD + dox. Based on a pilot strepativdin-seq experiment, I found that the strains with the greatest labelling relative to background in YPD compared to YPD + dox corresponded to a TetO insertion in a linker near the middle of the *GAL10* coding sequence (fig 3-9). Unfortunately, the signal over background was not striking: it was about 1.5-fold. Though the signal was repeatable and I believed it was real, it was weak enough that

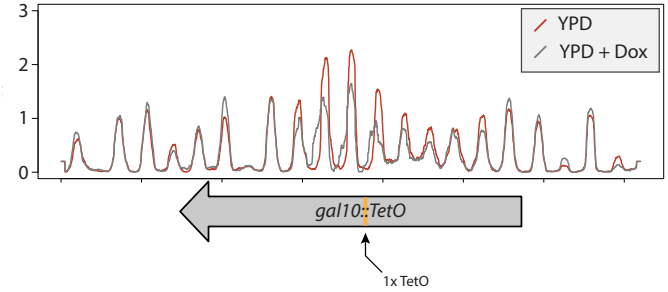

Figure 3-9 – TetR-BirA labels *gal10::TetO* with ~1.5-fold signal:noise in YPD compared to YPD + Doxycycline. Biotinylated histones were detected with streptavidin-coupled Dynabeads C1

I worried that the technology would not produce data with the kind of intuitive clarity that I believed was possible. I decided to focus my attention on improving the quality of the TetR-BirA labeling system, before attempting the complete kinetic labelling-and-tracking experiment.

My attempts to boost the signal:noise ratio for TetR-BirA followed no coherent path. At multiple points I made intuitive hypotheses about how to boost signal or reduce noise, and I tested my hypotheses ideas serially. In retrospect, if I had been focused and asked the incisive questions at the beginning of the technology development process, I could have avoided a lot of poorly construed effort. The incisive question – the one that it took me almost two years to ask – as what fraction of a yeast cell's nucleosomes were biotinylated by the TetR-BirA enzyme. Before I asked that question, I explored many other hypotheses, each of which I thought could explain the modest signal:noise ratio I observed using TetR-BirA.

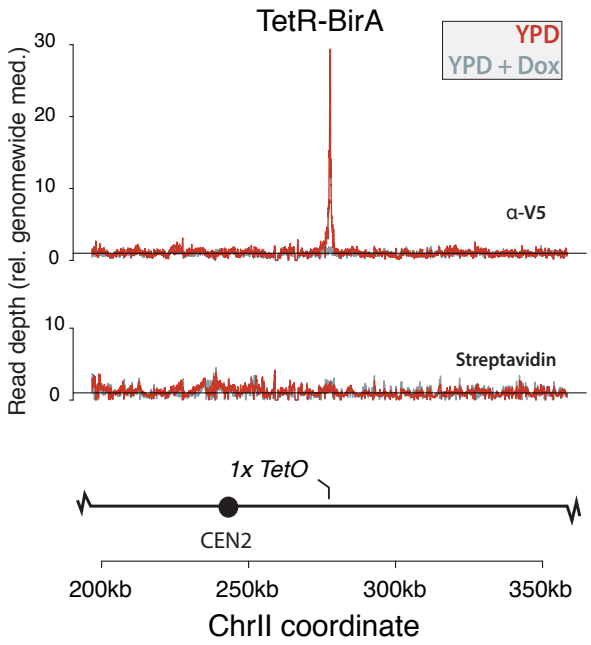

Figure 3-10 – Localization of TetR-BirA (wt) was extremely efficient and was strictly dependent on doxycycline. Although the BirA enzyme was well-localized, I did not detect dramatic local enrichment for biotinylation at the same locus.

My first set of hypotheses were centered on the idea that the signal:noise ratio I observed was low, because the technical background in the assay was too high. I reasoned that technical background could stem either from chromatin binding tube walls nonspecifically, or from chromatin binding the streptavidin-conjugated dynabeads nonspecifically, in spite of the stringent washes. Working under this belief, I tried to develop more effective blocking and washing buffers to block the beads and the tube walls and prevent the technical background. Among the strategies I tested were to use different dynabeads (hydrophilic instead of hydrophobic), to perform the pulldown in low-adhesion tubes, to wash the beads in 5M NaCl, and to perform the pulldown in blocking buffers containing various salts or proteins. In addition to trying a simple BSA blocking buffer, I tried extremely complex blocking buffers that contained either spermidine and salmon sperm DNA (to mimic chromatin electrostatic charges), or *Torulaspora delbrückii* chromatin,

which has the same biochemical properties as *cerevsisiae* chromatin, but which would not be confused for *cerevisiae* if it were to contaminate the final sequencing library. For the blocking buffers that included DNA, Jasper suggested using UV to crosslink DNA, such that it could serve as a blocking buffer, but could not be amplified by PCR during the ultimate library prep, and therefore would not contaminate final sequencing library. Unfortunately none of the blocking buffers changed the signal:noise of the biotinylation reaction.

My second set of hypotheses were based on the idea that the TetO locus was not biotinylated efficiently by the TetR-BirA enzyme. I reasoned that one way the TetR-BirA would fail to biotinylate the TetO sequence were if the enzyme were not expressed highly enough to allow the transcription factor to find its binding site. In my first experiments, I had used the lowly-expressed Sir4 promoter, that resulted in transcription that corresponded to ~1000 TetR-BirA proteins per cell. Based on coarse approximations of the yeast nuclear volume, I estimated that the concentration of TetR-BirA in the nucleus was ~2-5-fold lower than the biochemical k_d of TetR for TetO. Thus, I thought it was possible that the TetR protein never found the TetO sequence in the experimental strains, which could have resulted in low signal at the *gal10::TetO* locus. I replaced the promoter driving TetR-BirA with alternative promoters, designed to drive expression at, or slightly above, the biochemical k_d of TetR for TetO. I found that increasing expression of the TetR-BirA protein increased the biotinylation observable on a streptavidin blot, but didn't dramatically change the signal:noise. Importantly, using the updated TetR-BirA strains I was able to detect binding of TetR-BirA to the *gal10::TetO* sequence, and the binding of TetR-BirA to the *TetO* sequence was dependent on doxycycline (fig 3-10). This suggested that with the updated expression conditions, BirA was efficiently recruited to *TetO*, and any problems detecting biotinylation arise after TetR binding TetO.

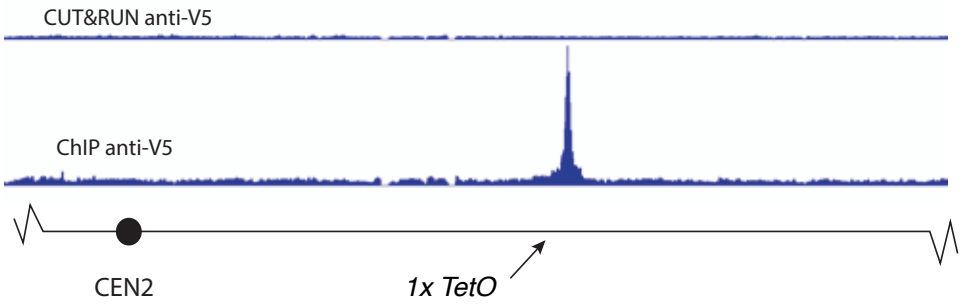

Figure 3-11 – CUT&RUN showed some promise in detecting H3-Avi-biotin, but did not detect TetR-BirA-V5. (Top) Genome browser trace for TetR-BirA-V5 and (bottom) genome browser trace for standard ChIP anti-V5 using agarose beads, for exactly the same strain and condition as above.

In parallel with my experiments testing the effect of TetR-BirA expression level on signal:noise, I was exploring a third hypothesis: that an alternative detection chemistry could detect the label more efficiently, either by being more sensitive to biotinylation or by allowing less technical background. The specific alternative I explored was CUT&RUN, which is essentially a two-step detection chemistry to bind epitopes *in situ* and cut adjacent nucleosome linkers (Skene and Henikoff, 2017). The major advantages of the protocol were that it required less input material, simplifying growth conditions for my experiments, and that it was reported to have very low technical background, allowing for lower sequencing costs and higher signal:noise. Furthermore, the technique relied on an anti-biotin antibody

instead of on streptavidin, which I had heard anecdotally improves signal:noise in pulldown mass spectrometry experiments. In my first CUT&RUN experiments, I was able to detect biotinylation of the *gal10::TetO* locus with better signal:noise than I ever had before – pushing ~10-fold in some experiments, however I was never able to successfully detect the TetR-BirA enzyme at the TetO locus (fig 3-11). For clarity sake, I hoped to perform all experiments using a single set of detection chemistry to monitor TetR-BirA localization in the same samples in which I monitor biotinylated histones, which made CUT&RUN an unattractive option. Furthermore, my attempts to detect biotinylated histones using CUT&RUN yielded variable results, ranging from zero signal to ~10-fold signal to noise, and despite nearly four months of troubleshooting I could not make the protocol work reliably. Thus I returned to pulldown chemistry using streptavidin dynabeads.

At this point, I was almost four years into my experiments attempting to track nucleosomes, and I had run out of tricks. In addition to manipulating blocking buffers, tuning expression levels, and experimenting with CUT&RUN, I tried many additional tricks that were not well justified and didn't solve my technical problem, and so don't merit a longer discussion here. Briefly, they included: LacI-BirA fusions, 2x-Avi tagged histones, ZFN-BirA fusions, dCas9-BirA[4] fusions and low-biotin growth conditions.

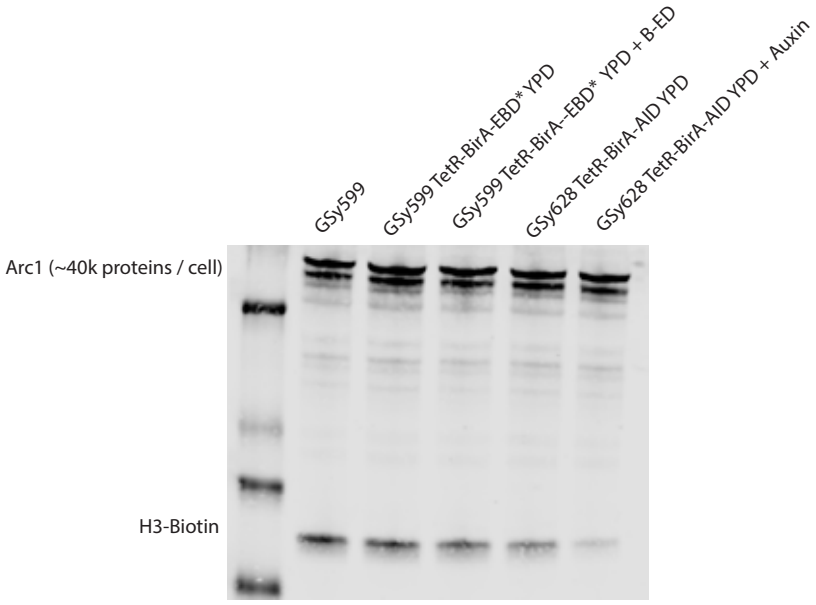

Figure 3-12 – I fused the estrogen receptor estrogen binding domain (EBD) and the auxin degron (AID) to BirA. Of all the strategies I attempted to switchably regulate biotinylation, auxin degradation of BirA showed the most switch-like regulation.

With no other tricks available, I made one more attempt to understand how the biology of BirA – as opposed to the detection chemistry – might influence the signal:noise. Around the same time, Siheng Xiang in the Koshland lab was experimenting using BirA to biotinylate cohesin and had performed band-shift assays to monitor the fraction of cohesion biotinylated by BirA. Band-shift assays were difficult in my case, because the band-shift

[4] My work on dCas9-BirA fusions is presented as an appendix, as a cautionary example of pitfalls in technology development

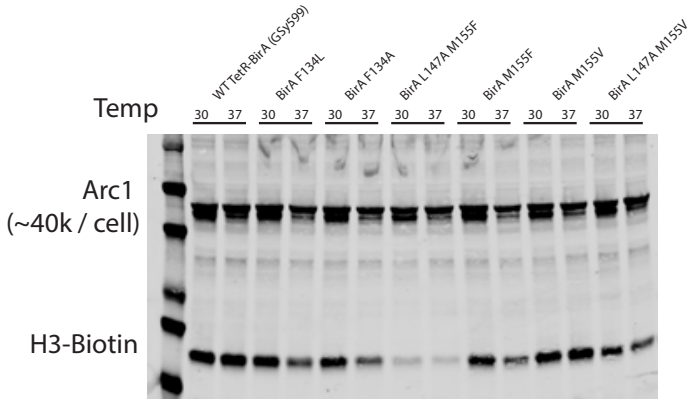

Figure 3-13 – I developed alleles that were predicted to be temperature sensitive for biotinylation. Many were experimentally verified, however none of them showed tight enough switch-like regulation to affect a brief pulse of BirA activity

relied on a ~200kD streptavidin tetramer binding a 18kD biotinylated histone peptide. Instead, I quantified the number of biotinylated H3-Avi molecules per cell by comparing the intensity of the biotinylated H3-Avi band on a streptavidin blot to the intensity of biotinylated bands of known abundance. The quantification was not precise, but based on the relative abundance of biotinylated H3 and biotinylated Arc1 I estimated that ~25% of histones were biotinylated in a given cell.

Based on the quantification, it was clear that my problem was not technical – the BirA enzyme was too active, and broadly biotinylated nucleosomes without specificity for the TetO locus. For the labelling approach to be successful, I would need to balance the rate of catalysis by BirA to the dwell time of TetR on the TetO sequence. Now that the problem was clearly defined, I had several long-shot opportunities to limit background biotinylation by BirA.

To turn down BirA biotinylation activity relative to TetR-TetO binding, I planned multiple approaches. The first set of approaches attempted to use a conditional allele of BirA to prevent biotinylation outside of a brief labeling window. I reasoned that if I can restrict labeling to a discrete time window (~10 minutes), I can accumulate biotin on-target without allowing the enzyme time to biotinylate histones off-target. To implement this approach, I used software to predict temperature sensitive alleles of BirA, and built the corresponding mutations. In addition, I developed strains in which BirA was tagged with the auxin degron, and strains in which BirA was fused to the estrogen binding domain. Based on this strain development, I identified several alleles that were hypomorphic at 37°C, and I demonstrated that the auxin degron limits biotinylation in the presence of auxin (fig 3-12, 3-13).

The auxin degron showed the best switch-like behavior between the induced and the un-induced state, and furthermore because the auxin degradation mechanism is orthogonal to most yeast biology, the auxin degron was a more attractive technology to regulate biotinylation. Based on this reasoning, I performed an additional experiment to test whether auxin could effectively prevent biotinylation by TetR-BirA-AID. For the experiment, I pre-grew cells without auxin, then added auxin and monitored the abundance of biotinylated H3-Avi. Under the model that TetR-BirA-AID is efficiently regulated by auxin, H3-Avi-biotin should decrease 2-fold per cell division and eventually become undetectable during extended growth in auxin. In practice, auxin did decrease biotinylation but did not decrease

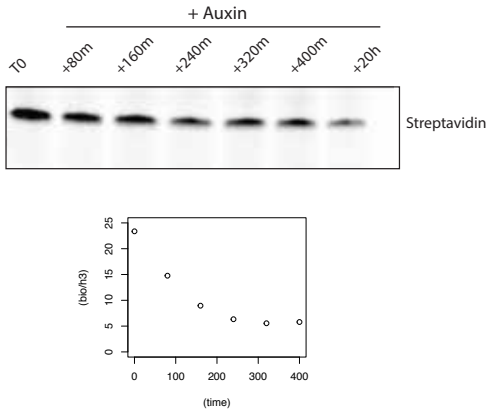

Figure 3-14 – If auxin is an effective regulator of biotinylation, H3-Avi-biotin should decrease 2-fold with every cell cycle until it eventually reaches zero. I found that auxin did effectively regulate biotinylation by BirA, but not tightly enough to drop labeling to undetectable levels. (top) raw data (bottom) quantified H3-Avi-biotin signal, normalized to anti-H3 western (not shown)

it to zero, suggesting that the auxin degron was not an efficient switch, capable of tightly limiting the duration of a biotin label (fig 3-14).

Although none of the switchable BirA alleles showed strong enough switch-like activity to be useful, some of the alleles were clearly hypomorphic, and I reasoned that a hypomorphic allele labeling at steady state might boost signal:noise without the need for a switchable BirA enzyme. Based on this idea, I developed a further set of hypomorphic

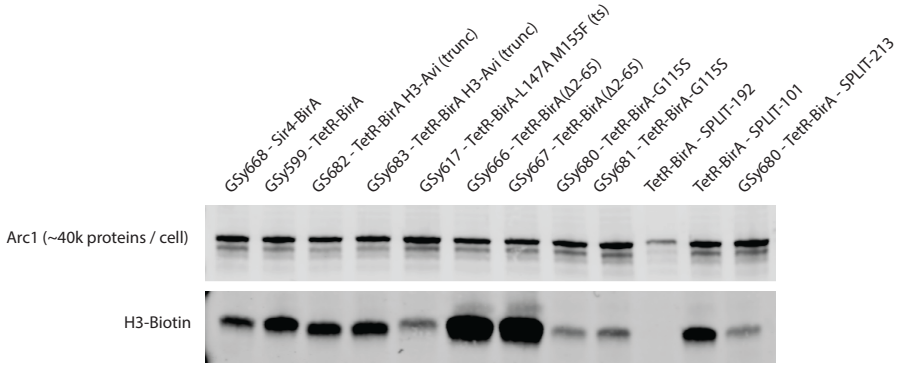

Figure 3-15 – I screened predicted hypomorphic alleles of BirA and split/BirA constructs to identify alleles that resulted in reduced – but nonzero – biotinylation at steady state. The idea hypomorph would act on the rate of catalysis by BirA, rather than stability of TetR-BirA. My previous attempts to develop AID-fusion strains, or TS alleles likely would act through TetR-BirA stability.

strains, including strains in which I truncated the Avi-tag, deleted a domain from BirA, or made point mutants in the BirA enzyme. I also developed three pairs of split-BirA alleles, in which a domain of BirA is removed in each of two TetR-BirA strains, such that the enzyme can only function if both TetR-(s/)BirA alleles are recruited to the same locus at the same time. I tested each of these alleles by performing streptavidin blots, under the assumption that the idea alleles would show reduced – but not completely eliminated – biotinylation of H3-Avi at steady state. Based on this search, I identified two designs that showed dramatically reduced – but nonzero – biotinylation of H3-Avi at steady state (fig 3-15).

Having shown that hypomorphic alleles of BirA decrease bulk H3-Avi biotinylation, I next asked whether strains with reduced biotinylation showed improved signal:noise for biotinylation at the *gal10::TetO* locus. If the hypomorphic strains decreased the rate of biotinylation such that a transient off-target TetR-DNA interaction is too brief to allow biotinylation, whereas an on-target TetR-TetO interaction lasts long enough to allow biotinylation, I would observe increased biotinylation at *gal10::TetO* relative to genome-wide background biotinylation. In fact whereas the wild-type BirA enzyme showed ~1.5-fold signal:noise, hypomorphic TetR-BirA-G115S allele showed ~7-fold signal:noise, and the split TetR-BirA(s/213) allele showed ~35-fold signal:noise (fig 3-16). Furthermore, in the presence of doxycycline, all constructs showed no biotinylation above background at the *gal10::TetO*, suggesting the labelling reaction can be rapidly stopped by addition of doxycycline to the growth medium.

Developing the hypomorphic BirA alleles ended the conceptually difficult phase of the project. The critical insight – which is obvious in retrospect – was that the activity of the BirA labeling enzyme must be balanced with respect to the dwell-time of the DNA tether to achieve site-specific labeling. This constraint likely applies to other experimental tools or natural enzymes with separable DNA-binding and enzymatic activity that would achieve site-specific enzymatic activity.

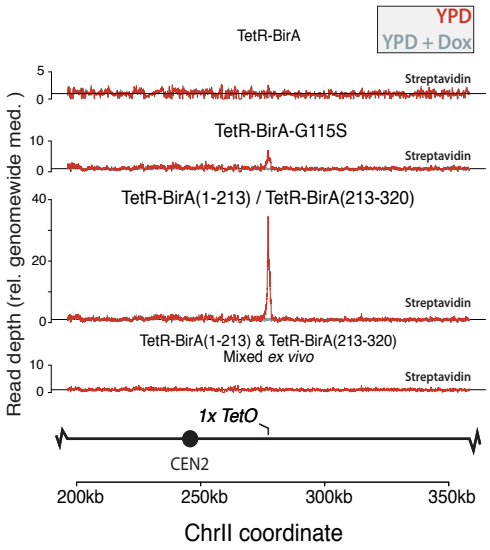

Figure 3-16 – Pulldown & sequencing results for my novel hypomorphic BirA enzymes. TetR-BirA (wt) showed minimal biotinylation at *gal10::TetO* relative to background, whereas TetR-BirA-G115S and TetR-BirA(s/213) showed 7.5-fold and 35-fold signal:noise respectively, When cultures expressing each portion of the TetR-BirA(s/213) allele were mixed *ex vivo*, there was no observed biotinylation suggesting that biotinylation occurred *in vivo* in split BirA strains. I got this data when I was sitting in a coffeeshop waiting for Paige to finish a job interview – I found at that she got the job soon after I got this data. It was a good day.

3.3 Tracking labeled histones through DNA replication in wild type cells

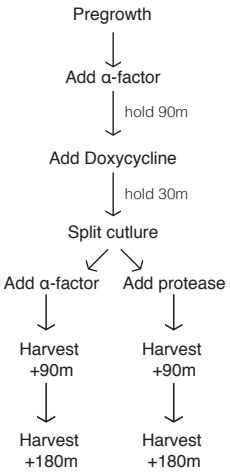

Figure 3-17 – Schematic of TetR-BirA tracking design

With the tool in-hand, I proceeded to execute the experiment that motivated the tool development – tracking nucleosomes through DNA replication. To track nucleosomes, I first labelled histones at the *gal10::TetO* locus, then arrested cells in α-factor. After the cells were arrested, I added doxycycline for 15 minutes to prevent further biotinylation. Next, I split the culture and in one half of the culture I added more α-factor to maintain the cells in a G1-arrest, and in the other half of the culture I added protease to degrade the α-factor, thereby releasing the cells synchronously into S-phase (Fig 3-17). By releasing the arrest using protease instead of by washing out α-factor, I can achieve more rapid, reproducible and synchronous release from arrest compared to washout methods.

To monitor the experiment, I harvested samples during the pre-growth phase at steady state, after α-factor arrest and doxycycline treatment, then again 90m and 180m after splitting the culture. From each sample I harvested, I performed a single chromatin prep that I split and used to precipitate TetR-BirA and H3-Avi-Biotin, to monitor the localization of BirA and the

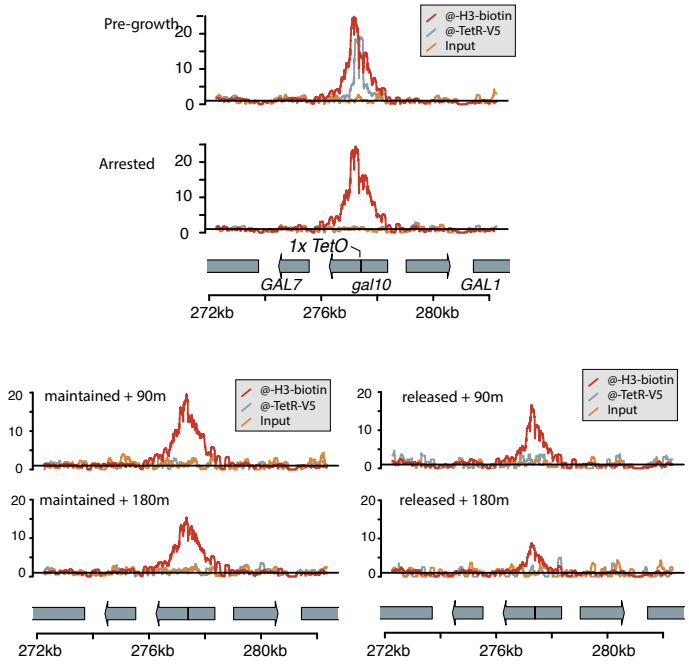

Figure 3-18 – Results for kinetic tracking of H3-Avi-biotin in wild type strains. TetR-BirA was effectively de-localized from *gal10::TetO* by doxycycline in arrested cells, and ancestrally labeled histones remembered their position during two rounds of DNA replication.

localization of the labeled nucleosomes respectively. I found that after DNA replication, labeled histones at the *gal10::TetO* locus were retained at the locus (fig 3-18). Likewise, during an extended G1 arrest, labeled histones were retained at the *gal10::TetO* locus (fig 3-18).

Importantly, the experimental strains showed normal response to α-factor, suggesting that the H3-Avi and TetR-BirA alleles did not affect chromatin stability (fig 3-19).

Figure 3-20 – cells expressing TetR-BirA (s/213) and H3-Avi arrested efficiently in response to mating pheromone, suggesting that the protein labeling machinery does not affect heterochromatin stability or gene regulation

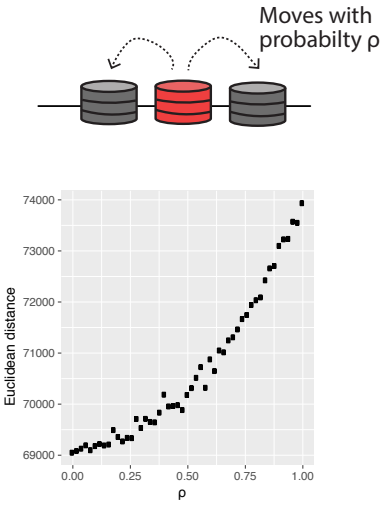

Figure 3-19 – To understand what fraction of nucleosomes move during DNA replication, I simulated nucleosome movement data under a range of assumptions about nucleosome movement. I found that the probability of movement that minimized the difference between the observed data and the simulated data corresponded to no nucleosome movement.

The top-level result that nucleosomes are retained during DNA replication was clear from a qualitative analysis of the streptavidin pulldown sequencing experiment, however I extended the qualitative analysis by performing statistical comparisons between the observed data, and simulated data that could have arisen under a range of assumptions. To perform the analysis, I assumed that every nucleosome observed in the arrested sample was allowed to randomize its position by +/- one position increment, where a position increment was defined as the binomial distribution centered on 165, for 200 Bernoulli trials (fig 3-20). Based on this assumption, I developed hypothetical data for a range of movement probabilities, and I compared each hypothetical data set to the observed data. I computed the similarity between hypothetical and observed data by calculating the Euclidean distance between each data set, using data in which midpoint counts are binned in 10bp bins (although I found the choice of bin size between 5bp and 40bp gave the same results). This analysis revealed that observed data were best described by the model that nucleosomes move +/- 1 position increment with probability 0; stated more succinctly: the data are best explained by the model that nucleosomes do not move during DNA replication (fig 3-20).

3.4 Tracking histones through DNA replication in replisome mutants

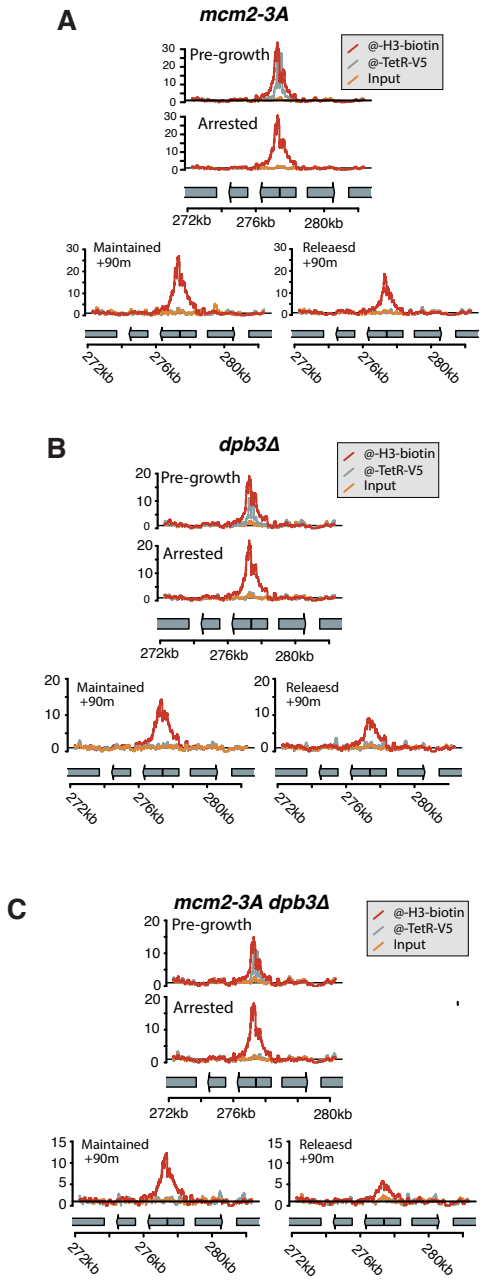

Figure 3-21 – mcm2-3A and *dpb3*Δ strains are deficient in nucleosome chaperone activity at the replication fork, and the mutants show reduced local histone inheritance. The double mutant *mcm2-3A* and *dpb3*Δ show extremely little local histone inheritance.

The histone tracking experiment in wild-type cells was the experiment that motivated the entire technology development process, and the primary goal of my PhD was to obtain a conclusive answer to the question: do nucleosomes remember their position during DNA Replication in wild-type cells. After I answered that question, I turned my attention to related questions that can be easily addressed using the same underlying technology.

Whereas in wild-type cells I found that nucleosomes remembered their position through DNA replication, I was interested in understanding whether it is possible to identify mutants in which the opposite is true: that nucleosomes lose memory of their position during DNA replication. In 2018, two papers were published concurrently that identified roles for Mcm2 and Dpb3 in transferring parental nucleosomes to the daughter strands during DNA Replication (Petryk et al., 2018; Yu et al., 2018). Briefly, based on the work from the Groth and Zhang labs it was clear that in the absence of Mcm2 or Dpb3 chaperone activity parental histones are segregated asymmetrically during DNA replication. It was not clear based on the data whether the old histones that were now asymmetrically inherited reflected 100% of old histones going to a single daughter chromatid, or whether the asymmetric inheritance reflected 50% of histones going to a single daughter chromatid (as in wild-type), and the loss of the remaining 50% to the nucleoplasm. Using the technology I developed to label and track nucleosomes, I could distinguish between the models by labelling nucleosomes at the *gal10::TetO* locus, and tracking their fate in strains expressing *dpb3*Δ or *mcm2-3A*, which show asymmetric nucleosome inheritance as single mutants. Additionally, I could track the fate of nucleosomes in strains expressing both *dpb3*Δ and *mcm2-3A*, which could conceivably show no nucleosome position-memory during DNA replication.

To test whether mutants expressing *dpb3*Δ or *mcm2-3A* transmit nucleosomes without diffusion at the replication fork, I mutated each gene in strains designed to label and track nucleosomes at *gal10::TetO* and repeated kinetic nucleosome tracking experiments as I had performed them for wild type cells. I found that nucleosomes remember their position in *mcm2-3A* or in *dpb3*Δ strains, however the number of parental labeled nucleosomes retained at the *gal10::TetO* locus during DNA replication appeared to be less than wild type cells (fig 3-21). Furthermore the number of parental labeled nucleosomes retained at the *gal10::TetO* locus in double mutant strains was extremely small (fig 3-21).

The results from histone tracking experiments qualitatively supported the model that labeled parental nucleosomes diffuse completely from the replication fork in the absence of the fork-associated nucleosome chaperones Dpb3 or Mcm2. Quantitative analysis of nucleosome position memory over time requires some strategy to normalize the extent of

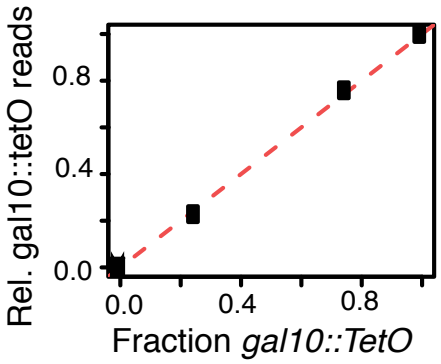

Figure 3-22 – Normalizing reads at *gal10::TetO* relative to background accurately quantifies the extent of biotinylation at *gal10::TetO*. I mixed strains encoding *gal10::TetO* with strains encoding *GAL10* and compared observed biotinylation at the *gal10::TetO* locus relative to the total number of mapped reads to the known mixing fraction. Both strains in the mixture expressed TetR-BirA(s/213), which should result in equivalent background labeling in both mixed strains.

biotinylation at a given locus. I began by naïvely analyzing the reads at the *gal10::TetO* locus per million mapped reads, under the assumption that most reads correspond to background activity by TetR-BirA, and that the background will be proportional to the input cell number in my chromatin prep. A prediction of this model is that if I mix extracts – some of which are biotinylated at *gal10::TetO* and others of which are not biotinylated – the number of biotinylated at *gal10::TetO* relative to background should reflect the proportion of cells competent to label at the *gal10::TetO* locus. Importantly, all cells in the mixing experiment expressed the same *Tet*R-*BirA* alleles, to ensure that background labeling was the same in both mixed populations; they only distinguished themselves in the presence of the *TetO* sequence at *GAL10*.

When I mixed fixed cell populations in defined ratios – either 0% *gal10:TetO,* 25%, 75% or 100%, and counted the number of reads at the *gal10::TetO* locus per million mapped reads the scaled the data to reflect the known quantities 0% and 100%, I found that 25% and 75% *gal10::TetO* in the cell mixture corresponded to 25% and 75% observed biotinylation at the *gal10::TetO* locus, indicating that my naïve approach of measuring reads at *gal10::TetO* per million mapped reads accurately quantified the extent of biotinylation at the *gal10::TetO* locus (fig 3-22). This is a subtle point that is easy to skip over, because it is intuitive to many people who interact with sequencing data; however, it is absolutely critical to the quantitative interpretation of ancestral histone tracking experiments.

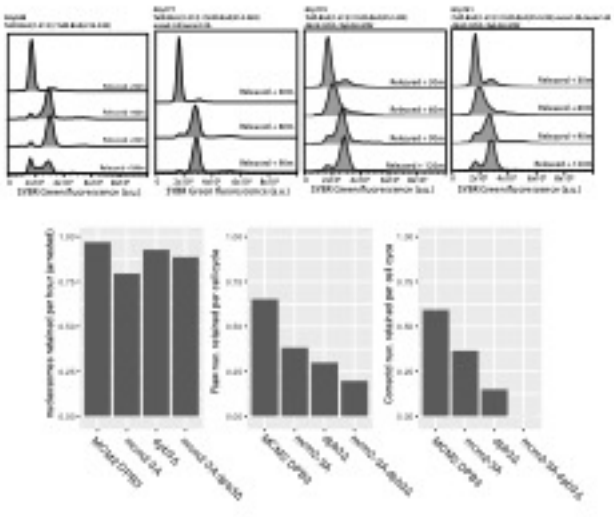

Figure 3-23 – Histone tracking strains each show different efficiency of release from mating pheromone arrest. By correcting for the efficiency of mating pheromone release, I can quantify the fraction of nucleosomes that are retained at the *gal10::TetO* locus during S-phase. (top) Flow cytometry data quantifies the efficiency of release from G1-arrest. (bottom) – (left) Fraction of nucleosomes that are retained per hour of G1-arrest. (center) Raw density of ancestrally labeled nucleosomes after one round of DNA replication. (right) Corrected density of ancestrally labeled nucleosomes, accounting for inefficient release from G1-arrest.

Having shown that the number of reads at *gal10::TetO* per million mapped reads corresponds to the quantitative extend of biotinylation, I compared the extend of biotinylation before and after DNA replication in wild type and replisome mutant strains. I found that in wild type cells after DNA replication, 57% of nucleosomes at the *gal10::TetO* locus were labelled with biotin, compared to 35% and 18% in *mcm2-3A* or *dpb3*Δ strains respectively (fig 3-23). In *mcm2-3A dpb3*Δ double mutant strains after one round of DNA replication, none of the parental labeled nucleosomes were retained at the *gal10::TetO* locus. I

repeated this result once, and obtained similar results: approximately half of nucleosomes at the *gal10::TetO* locus were ancestrally labeled in wild-type cells, an intermediate number were ancestrally labeled in each single mutants, and none were ancestrally labelled in the double mutant. Because of the complexity of the experiment, it was impractical to repeat the experiment enough times to accurately measure the variance in the observed fraction of nucleosomes retained at the *gal10::TetO* locus, but I felt sufficiently comfortable with the quality of the replication data to make the assertion that nucleosomes that are displaced during DNA replication in replisome mutant strains are not locally reincorporated, though it is possible that the *dpb3*Δ and *mcm2-3A* mutants show slightly different extents of nucleosome position memory, which might reflect secondary effects on the replisome that arise from mutating the nucleosome chaperones. Notably, it should be impossible to recover more than 50% ancestrally labeled nucleosomes after one round of replication, but I observed 57% ancestrally labeled nucleosomes at the *gal10::TetO*; thus it is clear that although my estimate of the quantitative extent of nucleosome position memory is imperfect.

Although the genetic relationship between the mutants and the fraction of nucleosomes retained at *gal10::TetO* was consistent and qualitatively evident from the raw data, the quantitative interpretation was limited by the fact that replisome mutant cells did not arrest or release as efficiently as wild type cells (fig 3-23). Thus, to calculate the fraction of nucleosomes retained through DNA replication I needed to estimate the fraction of cells that replicate their DNA. Using SYBR green staining of fixed cells, I was able to quantify the number of cells that replicated their DNA, however flow cytometry is complicated in yeast because cells tend to clump together, and although sonication was effective in relieving clumps it was not perfect. Thus, my ability to appropriate correct for the fraction of cells that replicate their DNA limited my ability to confidently assert the fraction of nucleosomes retained at the *gal10::TetO* locus during DNA replication. To more accurately measure the fraction of nucleosomes retained at the locus in each replisome mutant strain, it would be useful to precisely track which DNA molecules replicated during the experiment by labelling newly replicated DNA with EdU in parallel with the nucleosome labeling and tracking: with EdU incorporated on newly replicated DNA strands, there are multiple biochemical strategies that could separate reads coming from replicated or replicated DNA templates to directly observe the fraction of nucleosomes retained only on the replicated DNA.

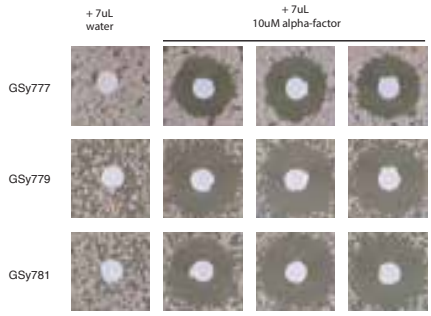

Figure 3-24 – Replisome mutant strains arrested efficiently and durably in response to mating pheromone. Because the halos arrested by α-factor are "clean," i.e. with no small colonies within the pheromone radius, I concluded that silencing is stable, and not transiently lost during prolonged exposure.

There is another observation inherent in my experiments about replisome assembly factors that is not obvious, and about which I did not comment in my publication on the subject. It was astounding that the replisome mutant strains arrest in α-factor at all (fig 3-24). For replisome mutant strains to have arrested in mating pheromone, they must have correctly regulated heterochromatin at *HML* and *HMR*. Furthermore, because the cells show clean halos within the α-factor radius, I concluded that cells that fail to arrest or release as monitored by flow cytometry have some pheromone-independent cell cycle defect. That the replisome single and double mutant strains arrested reasonably well in mating pheromone suggested either that natural co-replication chromatin assembly is not required for maintenance of silencing, or that silencing is re-established faster than it can be lost, even when starting from an unnatural poorly assembled mating type locus. At the time of writing, I did not have sufficient evidence to make a strong claim on the subject, but I favor the hypothesis that nucleosomes do not transmit epigenetic memory of silencing. I am currently performing experiments in *sir1*Δ *dpb3*Δ *mcm2-3A* mutant strains to ask whether—when nucleosome position memory is abolished—silencing at *HML* is nonetheless epigenetically inherited.

3.5 Tracking labeled histones through transcription

The fate of nucleosomes during replication is a central question that has been pursued by dozens of labs over the past 30 years, but the fate of nucleosomes during transcription is comparatively understudied. Based on two sets of experiments performed *in vitro*, it appears that transcription can occur on a nucleosomal template without removing the nucleosome from the DNA, suggesting that melting of the DNA strands to allow passage of RNA polymerase II occurs transiently and locally (Studitsky et al., 1994, 1997). In one set of experiments that probed the fate of nucleosomes during transcription, synthetic chromatin was established using a 227bp linear piece of DNA bound by a single nucleosome assembled *in vitro*. The DNA was than transcribed using a promoter-independent T7 RNA polymerase transcription assay, and the nucleosome was observed to move in a retrograde manner (relative to transcription) by ~10-20bp during transcription by T7 RNA polymerase, however this result was not replicated using slightly longer (262bp) DNA templates (Studitsky et al., 1994). In addition to experiments using mono-nucleosome length templates as a substrate for transcription experiments, more recent experiments have used templates that are ~2kb in length loaded with a single nucleosome, held under tension using an optical trap (Hodges et al., 2009). In the optical trap experiments, a single eukaryotic RNA polymerase II complex was introduced, and the complex was tracked optically as it transcribed the length of the 2kb DNA molecule; additionally, the presence of the nucleosome could be detected by measuring the length of the DNA molecule under moderate tension. The work demonstrated that eukaryotic RNA polymerase II can transcribe DNA without evicting the nucleosome and that after transcription proceeds past a nucleosome the transcribing RNA polymerase II molecule cannot slip or backtrack into the nucleosome footprint as it can on naked DNA (Hodges et al., 2009). Thus, the authors concluded that nucleosomes remain static in their position, and this behavior limits the ability of RNA polymerase II to transiently backtrack during transcription.

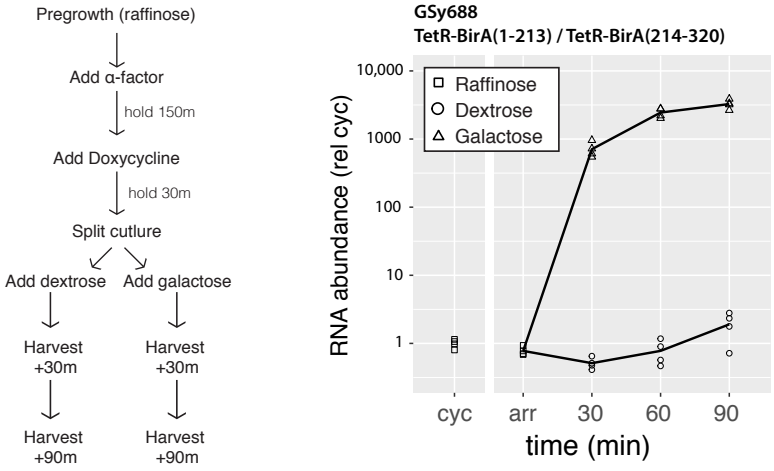

Figure 3-25 – (left) Schematic of nucleosome tracking experiment during transcription. (right) Transcription is induced >1000-fold from *gal10::TetO* in response to galacose.

In contrast to work *in vitro* that found nucleosomes are stably bound to DNA during transcription, one experiment concluded that nucleosomes are translocated in a retrograde manner during transcription *in vivo* (Radman-Livaja et al., 2011). The data that underlie the claim are bulk histone turnover measurements in wild type and topoisomerase mutant yeast: the authors found that ancestral nucleosomes are more prevalent at the 5' end of genes compared to the 3' end of genes, and that effect is minimized in a topoisomerase mutant (Radman-Livaja et al., 2011). To interpret their data, the authors supposed that the retrograde movement of nucleosomes during transcription – observed over ~10-20bp during T7 transcription of 227bp fragments *in vitro* – could be propagated along an open reading frame during transcription *in vivo*. An alternative explanation for their observation would allow that old nucleosomes are concentrated at the 5' end of an ORF because the nucleosomes at the 5' end of an ORF are somehow more stable than those at the 3' end of an ORF in a topoisomerase-dependent manner, without invoking directional movement of nucleosomes during transcription.

Based on the published data, which tracked the entire complement of old histones without locus-specific labeling, it is impossible to distinguish between the processive movement model and the differential stability model, however using the TetR-BirA labeling strategy I developed I could directly test the processive movement model in live cells. To perform the experiment, I labeled nucleosomes at the *gal10::TetO* locus in cells pre-grown in raffinose, to relieve glucose repression of the *gal10::TetO* sequence. Next, I arrested the cells in G1 to prevent DNA replication, and I stopped applying new label. I then split the culture and exposed the cells to either 2% glucose, to repress transcription, or 2% galactose to induce transcription (fig 3-25). By precipitating biotinylated nucleosomes in the repressed or induced transcription state, I could ask directly whether the ancestrally labelled nucleosomes moved relative to the direction of transcription. I found that ancestral nucleosome density at the *gal10::TetO* decreased during transcription, suggesting that some nucleosomes were evicted during transcription; however, I found that the nucleosomes that remained did not move locally along the chromosome (fig 3-26). Thus, my data contradicted the model that nucleosomes move in a retrograde manner relative to transcription *in vivo* but were consistent

with the model that nucleosomes serve as a ratchet, and prevent RNA polymerase II from backtracking by physically constraining its backwards movement.

I didn't attempt to test the model that nucleosomes show different stability at the 5' and 3' ends of an open reading frame, but I built strains that would be useful if someone wanted to test that model in the future. Whereas the *gal10::TetO* locus achieves specific enough labeling to track nucleosomes within an open reading frame, the open reading frame is too short to allow comparative analysis of the 5' end to the 3' end of the ORF in the same strain. Thus, I inserted the *GAL10p* upstream of the *FMP27* gene, which is 7.8kb (*GAL10* is 2.1kb). I found that the *GAL10p* effectively regulated expression of *FMP27*, which would allow me to do similar experiments as I had previously done at the *gal10::TetO* locus by inserting TetO sequences in the *FMP27* ORF, then arresting cells and modulating transcription of *fmp27::TetO*. Because the ORF is considerably longer than *GAL10*, it would be possible to label the 5' end and the 3' end of the ORF, then monitor their relative turnover rates in the same strain and in a topoisomerase mutant, to understand whether differential nucleosome stability might underlie the previously reported asymmetry in ancestral nucleosome occupancy along a transcribed ORF.

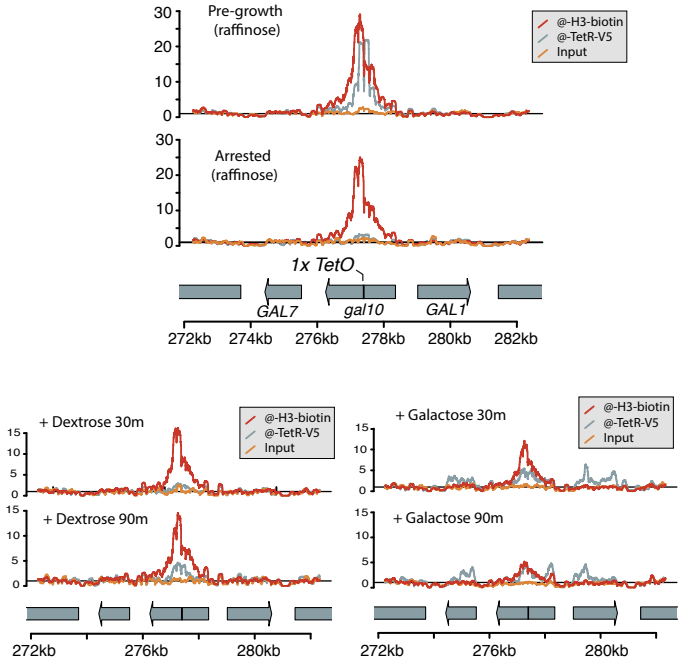

Figure 3-26 – Nucleosome tracking during transcription revealed that nucleosomes did not move linearly along the *gal10::TetO* open reading frame, in contrast with published models.

3.6 Labelling H2B

When I began the project, I was optimistic that I would be able to separately track H3 and H2B: although the nucleosome core tetramer $(H3/H4)_2$ is considered more important for epigenetic inheritance of transcription state, tracking the H2A/H2B dimer would be an interesting point of comparison to understand whether motility of H3 or H2B might reflect their biological roles. Additionally, the H2A/H2B dimer is known to have a specialized role in transcription and DNA double strand break repair, and tracking the fate of ancestrally labeled H2B could reveal details about the relationship between H2B and those central processes (Wyrick and Parra, 2009).

To track the H2A/H2B dimer, I developed strains in which I fused the Avi tag to the n- and c-terminus of H2A and H2B. Whereas H3-Avi fusions grew as well as wild type and arrested efficiently in response to α-factor, H2A and H2B-Avi fusion strains had a growth defect and did not arrest efficiently during exposure to α-factor in raffinose media (Fig 3-27). Additionally, using the same TetR-BirA design that resulted in locus-specific labeling of H3-Avi at *gal10::TetO*, I observed no locus-specific labeling of H2B-Avi.

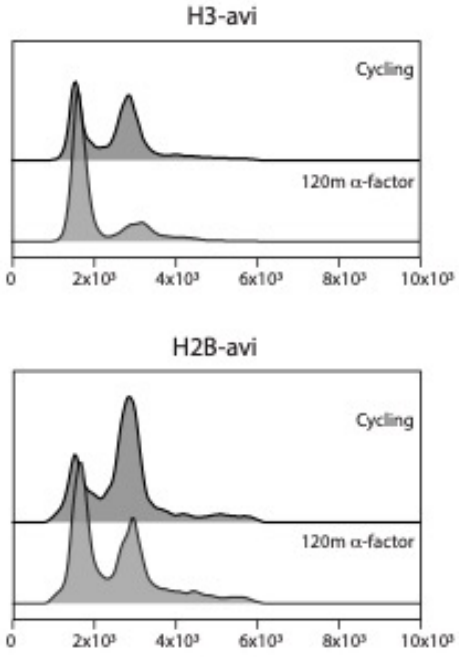

Figure 3-27 – (top) H3-Avi strains grown in raffinose showed efficient cell cycle arrest in response to mating pheromone, suggesting that chromatin biology was unperturbed. (bottom) H2B-Avi strains failed to arrest efficiently in mating pheromone, suggesting that chromatin biology or gene regulation was severely affected by the tagged histones.

A major design criterion for this technology development project was that the chromatin should be unperturbed – it should replicate normal chromatin biology to the extent we can measure. Clearly H2A-Avi and H2B-Avi failed by that criterion, and so I did not include any results from H2B-Avi tracking in my publication on the subject.

At a conference in 2018, I learned of a possible explanation for my failure to track H2B. According to Carl Wu, histones H2B tracked with photo-switchable dyes in live cells only binds chromatin for ~30s before dissociating, suggesting that it should not remember its

location long enough to be detected by the TetR-BirA labeling strategy. Although that observation provides a satisfactory explanation for why I could not track H2A/H2B, it does not explain why the H2A-Avi and H2B-Avi strains grew poorly and failed to arrest.

I have no explanation for why the Avi-tagged histones were apparently deficient, despite the fact that numerous papers rely on H2B-RFP fusions – a fusion I would have expected to be more invasive than an -Avi tag fusion – to monitor nuclear morphology or chromatin compaction. I would caution that I have not read a single paper that uses histone fusion proteins in which the authors have rigorously established that the histone fusion proteins do not interfere with normal chromatin biology, and I am suspicious that H3-GFP or H4-GFP fusion proteins might introduce unappreciated chromatin defects that complicate experiments that rely on the fusions.

3.7 A curious artifact

In the process of developing the histone labeling strategy, I characterized background biotinylation by the BirA fusion proteins. In addition to biotinylating *gal10::TetO*, the TetR-BirA fusion proteins labelled a single additional locus in a doxycycline dependent manner (Fig. 3-28). That locus – *YEL1* – was on the same chromosome as *gal10::TetO*, which led me to wonder whether the labeling occurred as a result of some intra-chromosomal contact, that allowed TetR-BirA to label a locus at a distance. Although *YEL1* labelling occurred with both hypomorphic BirA alleles, the effect was stronger with the TetR-BirA-G115S allele (Fig. 3-28).

If such a mechanism explained the off-target labeling, I expected the intrachromosomal Hi-C contact map to indicate that the two loci show frequent contacts. To test this possibility, I wrote to Jan Skotheim, who had published the best available Hi-C in yeast, to ask for the contact frequency matrix his lab generated in processing their Hi-C data. I queried the contact frequency data, and discovered that there was no meaningful contact between the GAL10 locus and the YEL1 locus in their experiment, although there were meaningful contacts between the GAL10 locus and other loci on ChrII, none of which were biotinylated by my TetR-BirA alleles (Fig. 3-29). Thus, I rejected the hypothesis that chromosome folding caused the off-target labeling I observed at the *YEL1* locus.

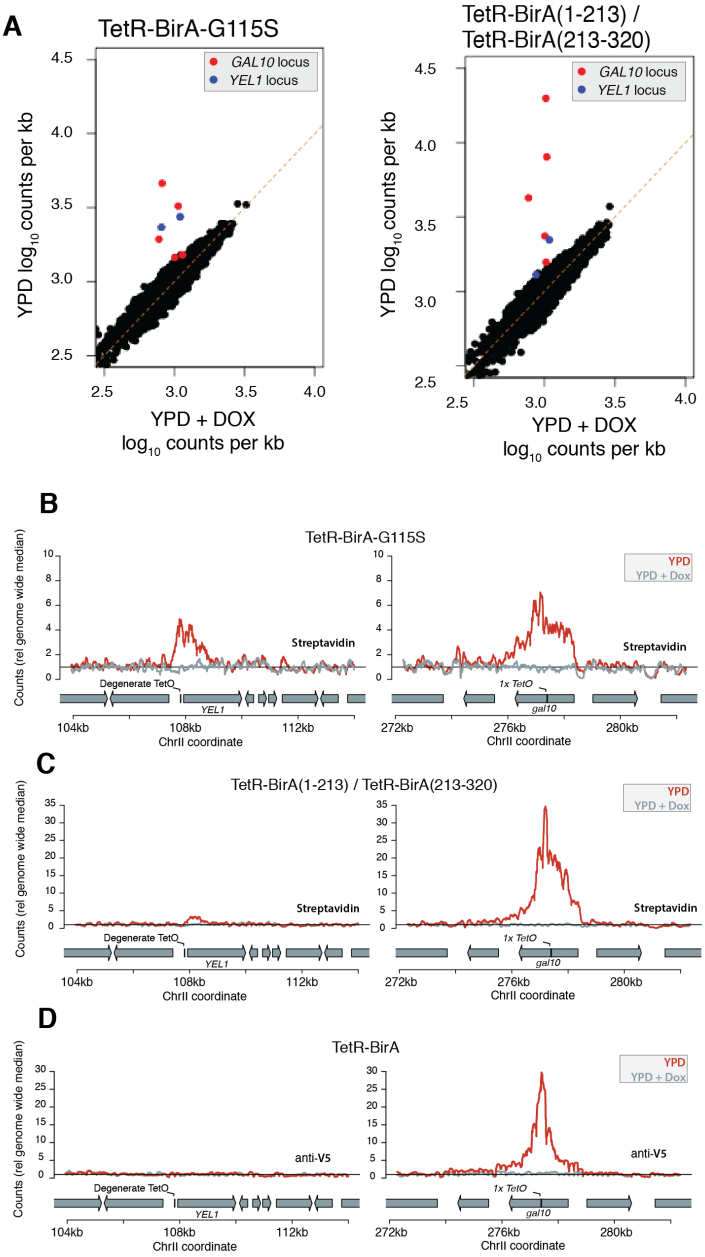

Figure 3-28 – (top) Correlation between labelling in -doxycycline conditions and +doxycycline conditions in 1kb intervals genome wide. Loci that are above the diagonal correspond to loci that are biotinylated only in the absence of doxycycline. Only two loci show this property: the expected *gal10::TetO* target and the unexpected *YEL1* target. (bottom) Local biotinylation at the *YEL1* locus was stronger with TetR-BirA-G115S than with the split TetR-BirA(s/213) design, and no TetR-BirA localization as observed at the *YEL1* locus as measured by V5-ChIP-seq.

The second idea I tested was actually the more obvious possibility: I scanned the *YEL1* locus for *TetO* binding sites that could explain the localized biotinylation, and I found a near consensus *TetO* sequence in the *YEL1* promoter (FIG – lab meeting). Thus, it seemed likely

that the off-target labeling I observed at the *GAL10* locus was really on-target labeling of a cryptic *TetO* sequence that occurred naturally in yeast.

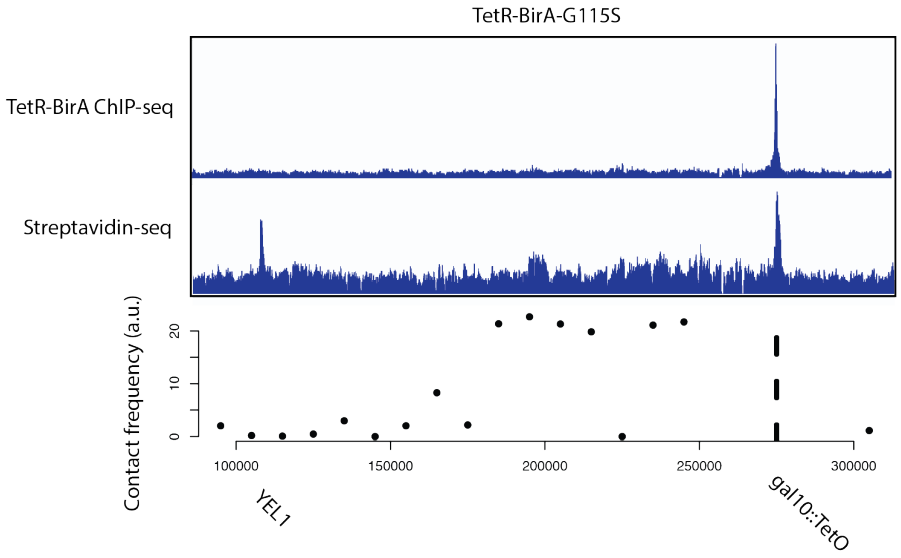

Figure 3-29 – (top) TetR-BirA ChIP-seq identified strong binding at *gal10:TetO* but none at *YEL1*, whereas streptavidin precipitation and sequencing identified biotinylation at both loci. (bottom) The contact frequency between *GAL10* and each intrachromosomal locus was calculated in 10kb bins along chromosome II. The black line corresponds to the *GAL10* locus. *GAL10* has extensive contacts along chromosome II, but no detectable contacts with the *YEL1* locus. Thus, it is unlikely that long-range chromosome contacts explain the observed off-target labeling.

Whereas the *YEL1* locus was biotinylated by TetR-BirA above background, I found no ChIP-seq evidence that the *YEL1* locus was bound by TetR-BirA. Thus, it seemed that the biotinylation of the *YEL1* locus was storing some memory of previous, transient exposure, even in conditions where the transcription factor was not stably bound.

I think this observation might be a clue that could inspire a new understanding of transcription factor biology. Single-molecule live cell imaging in the Tjian and Darzacq labs has revealed that different transcription factors show different dwell times on DNA, and that the dwell time of a transcription factor on DNA is related to its ability to be detected by ChIP-seq, such that short-binding transcription factors are not readily detected by ChIP-seq. Whereas the Tjian and Darzacq labs characterized differences in dwell time between different transcription factors, I think there might be a range of binding dwell times for a given transcription factor, such that the precise dwell time can have a distinct biological activity.

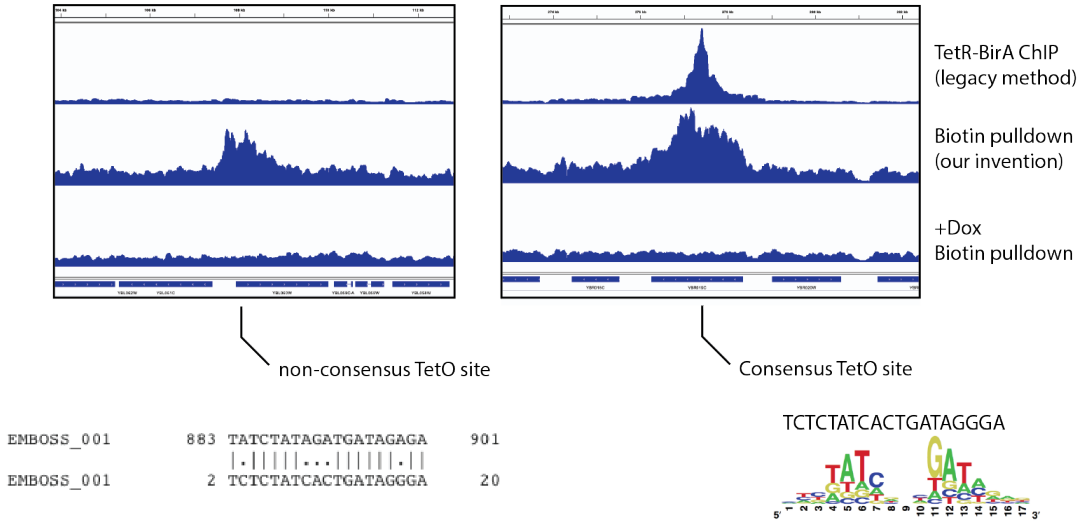

Figure 3-30 – Doxycycline-dependent biotinylation by TetR-BirA-G115S was likely caused by the natural occurrence of a near TetO sequence at the *YEL1* locus.

It is a tautology that transcription factors must have short dwell times on off-target sequences and long dwell times on off-target sequences. It is less clearly defined whether there are binding events that are intermediate between off-target and canonical on-target DNA-binding. Known transcription factor binding sites are generally defined probabilistically, and many variants of each motif occur in the same cell. I propose that those binding sites are not equivalent, and could distinguish themselves in their biological activity on the basis of the duration of the transcription factor binding event they participate in.

Under this model, a possible explanation for why I could not see ChIP-seq signal corresponding to TetR-BirA, but could see biotinylation of the adjacent histones, is that TetR-BirA binds the *YEL1::TetO* too transiently to be captured by ChIP-seq, but durably enough to allow catalytic biotinylation of the adjacent nucleosomes. Furthermore, because off-target *YEL1::TetO* biotinylation was quantitatively less than on-target *gal10::TetO* biotinylation, it is tempting to speculate that the quantitative extent of histone biotinylation is proportional to the duration of a TetR-TetO binding event.

If I was able to accidentally develop a fusion protein that has a distinct biochemical output as a function of transcription factor dwell time, it seems likely that nature would have similarly accidentally developed proteins that show distinct activity on distinct binding sites, perhaps in a way that reflects a hierarchy of dwell times of transcription factors on DNA. Perhaps such a model could explain why some transcription factors seem to induce or repress genes, depending on the precise context of the binding site.

4 The Side Projects

4.1 Replicative lifespan

While doing undergraduate research, I read some of the literature that deals with yeast replicative lifespan. One observation—that old yeast cells fail to mate—was explained as a consequence of age-related changes to mating gene regulation. By luck, I joined the lab at a time when a new tool became available to study silencing at yeast mating type loci, and I used the new tool to re-examine the old model that old yeast cells lose silencing at *HML* and *HMR*. I found that yeast do not lose silencing during aging, and I initiated a collaboration with another lab that showed that age-associated changes in protein folding—not in gene regulation—might underlie age-associated mating phenotypes in yeast.

The story of my research on replicative lifespan in yeast starts shortly after I left the Rine lab for the first time. After spending a summer in the Rine lab, I joined the Kruglyak lab at Princeton to do research for my undergraduate thesis. There was one other undergraduate in the lab—Thúy-Lan Võ Lite—who was working with Joshua Bloom to map genes that affect yeast mitotic aging. She planned to use strains that had been recently developed by Derek Lindstrom and Dan Gottschling to select for long-lived mothers derived from a yeast cross, and map alleles whose inheritance was correlated to replicative lifespan (Lindstrom and Gottschling, 2009). Thúy-Lan and Josh got as far as requesting the reagents from the Gottschling lab, but they had trouble introducing the relevant alleles into our lab strains. Thuy-Lan ultimately tacked to work on other projects, and the project was orphaned at the starting line.

When I started shopping for an undergraduate thesis project, the aging experiment was a top candidate. To do the experiment well seemed ambitious but achievable. In the winter of 2011-2012, I started a deep-dive in the yeast aging literature, and wrote a research proposal for my thesis that justified my experimental approach and projected possible outcomes and interpretations of the data.

One of the contingencies I explored in that proposal was the possibility that the yeast aging process is related to regulation of heterochromatin. This model would connect several lines of evidence—first, old yeast accumulate extrachromosomal rDNA circles that can be detected by agarose gel and DNA blot (Sinclair and Guarente, 1997). Second, alleles of *SIR4*, which participates in heterochromatin regulation and gene silencing, were identified that affect cellular lifespan (Kennedy et al., 1995). Third, old yeast haploid cells failed to mate, and that failure to mate was dependent on the presence of *HML* in *MAT***a** yeast, which suggested that old yeast lose their ability to silence their auxiliary mating type loci (Smeal et al., 1996). Based on these lines of evidence, I expected that in mapping determinants of cell lifespan, I would discover allelic variation in chromatin-associated pathways that contribute to yeast aging.

I started my mapping experiment by building strains of BY4742 and RM11-1a that contained all the genetic bells and whistles we would need to do bulk segregant analysis and lifespan analysis in the same strains—I had to introduce the Cre-EBD, CDC20-loxP, and UBC9-loxP generated by Lindstrom into the bulk segregant analysis strain background generated by Ian Ehrenreich, a former Kruglyak lab postdoc (Ehrenreich et al., 2010;

Lindstrom and Gottschling, 2009). Ehrenreich's strain co-opted the "synthetic genetic array" mating-type-selection circuit developed by Amy Tong in Boone-Andrews lab to select for large numbers of independent haploid *MAT***a** segregants from a cross between two divergent strains BY4742 and RM11-1a (Ehrenreich et al., 2010).

While I was building the strains to perform the mapping experiments, Leonid was contacted by a former colleague to get feedback on a paper he was preparing to submit, which relied on a panel of segregants that the Kruglyak lab had generated for an early QTL mapping project. The paper manually analyzed the lifespan of 88 segregants with known genotypes from an early mapping panel developed by the Kruglyak lab, and discovered that the lifespan difference between BY4742 and RM11-1a is explained by allelic variation at the rDNA locus (Kwan et al., 2013).

At the time, this took a lot of the luster out of my undergraduate thesis project. I became pre-occupied with the idea that the best-case scenario was that I re-discover a known QTL that affects lifespan, and the worse-case scenario was that my results are at odds with the published results, and I run out of time to understand the basis for the disagreement. So, I abandoned the project, in favor of a more conservative project that required less upfront strain construction and was more broadly interesting to the rest of the lab.

The details of my ultimate undergraduate thesis are wholly irrelevant to my graduate thesis work. But the abandoned project—the false start towards mapping genes that control the cellular lifespan—was exceptionally influential in my ultimate graduate school trajectory and would go on to inspire the experiments that led to my first paper.

During my rotation in the Rine lab, I tried to address the question: does meiotic recombination at *HML* or *HMR* result in loss of silencing in *cis*? Eight weeks into my ten-week rotation, I had built strains with markers flanking *HMR* and dissected tetrads for the experimental cross, but I observed vanishingly few crossover events, and vanishingly few loss of silencing events, and the non-conclusion left me unsatisfied. Feeling self-conscious about the status of that experiment, I went looking for a hypothesis that I could test quickly using Anne Dodson's recently developed CRASH assay (Dodson and Rine, 2015). I decided to re-test an established model for yeast aging: that old haploid yeast cells fail to silence *HML*.

The CRASH strain Anne developed captures loss of silencing events with exquisite sensitivity and records a memory of those events by coupling them to a heritable genetic change. Cells that lose silencing—or whose forebears lost silencing—express a green fluorescent protein, whereas cells that have never experienced loss of silencing in their lineage are red (Dodson and Rine, 2015). This design let me determine immediately if the first cell plated on a plate was red or green—if the colony contains any red at all, the founder cell itself must have been red.

Leveraging this property, I dissected pedigrees of the CRASH reporter in the W303 background that Anne used. I was expecting cells to lose silencing 2/3 of the way through their lifespan, meaning that colonies that derive from late daughters are completely green. That is not what I saw. First in three pedigrees, then later in 24 pedigrees in two strain backgrounds, then later in hundreds of pedigrees analyzed by microfluidics, I found no evidence that a cell's age is related to loss of silencing.

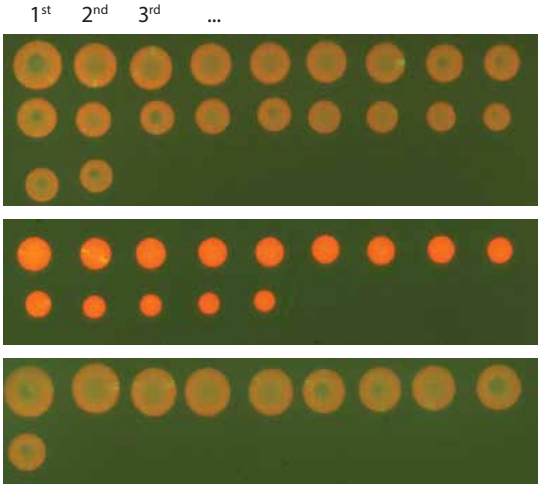

Figure 4-1 – Manual pedigree analysis of a CRASH reporter strain. If a colony is contains any RFP, the founder cell of that colony must have been RFP+, and therefore must never have lost silencing in its pedigree. None of the pedigrees I dissected showed evidence of age-associated loss of silencing.

When I joined the lab and started planning my core thesis project, it became clear that my experiments had a very long time horizon with a relatively high risk profile. Jasper and I agreed that a prudent strategy would be to try to get an early paper based on my rotation work, which would give me a soft landing if my other work were to take too long or fail outright.

Whereas my rotation work was in the W303 strain background, a lot of the previous work on aging—including the work that showed that old cells might lose silencing—was performed in S288C. Therefore, I started by repeating the original pedigree dissection experiment in S288C as well as in W303 diploids, and W303xS288C hybrids. In each case, I used a single copy of the *HMLα2Δ::Cre* transgene and a single copy of the *LoxP-RFP-KanMX-LoxP-GFP* reporter cassette. In 24 pedigrees, I saw no loss of silencing events in any of the strain backgrounds (fig 4-1).

My initial experiments showed that loss of silencing at *HML* was not a common feature of yeast aging, but I wanted to extend the experiment to ask the question: are loss of silencing events more common in old cells compared to young cells? To address this question, I used a microfluidic chip that had been recently published by Myeong Jo in Lidong Qin's lab at Houston Methodist. The device used micro-wells to trap hundreds of yeast mother cells, while supplying them with fresh media and washing their daughters away (Jo et al., 2015). I requested the CAD file to make the microfluidic chip and built the chip to published specifications with help from Naima Azgui and Morgan Delarue in Berkeley's Biomolecular Nanotechnology Center.

In my first experiments working with the microfluidic chip, I attempted to load the microfluidic device with the W303-derived CRASH strain that I had used in manual pedigree analysis experiments. I quickly found that W303 cannot be used in microfluidic experiments because W303 grows in micro-clumps that prevents the media flow from washing daughter cells away in the microfluidic chamber. These micro-clumps had not been previously reported to my knowledge, and at the time I did not know what alleles caused the phenotype. The clumps appear to be a much less severe form of flocculation, where the cells

are well suspended when grown in large liquid cultures, but observation under the microscope reveals many clusters of 2-8 cells. These clumps were not a feature of an S288C-derived CRASH strain, and I transitioned to using the S288C derivative in all of my future microfluidic experiments.

After climbing a steep learning curve—learning how to reliably manufacture microfluidic devices and developing best practices for avoiding dust or bacterial contamination in the microfluidic device—I eventually captured time-lapse microscope images of cells over their entire lifespan. The result obtained using microfluidics was the same as the result obtained by manual pedigree analysis—that loss of silencing is not a feature of yeast aging. However, given the scale of microfluidic experiments, I was also able to observe 13 loss of silencing events that occurred in pedigrees, and ask whether the timing of those events is related to yeast replicative age. I found no evidence that the timing of loss of silencing events is related to cell age (fig 4-2).

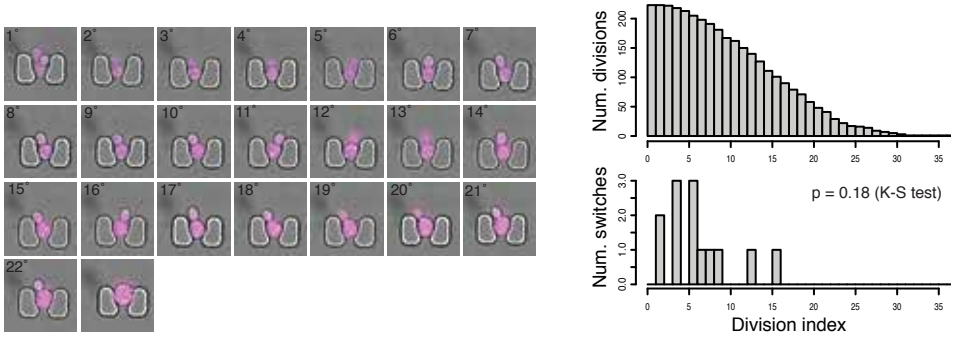

Figure 4-2 – (left) Representative pedigree collected by microfluidic observation of mother cells. (right) histogram of loss of silencing events in microfluidic loss of silencing experiments. I observed no relationship between the number of times a mother cell divided and the occurrence of loss of silencing events.

My data were convincing, but they were at odds with published data and I wanted to explore the reasons why. In the original study that suggested old cells lose silencing, they relied on two forms of evidence. The first was RT-PCR, probing for the spliced *HMR***a***1* message in old *MAT*α haploid cells (Smeal et al., 1996). Smeal found that it is possible to amplify the spliced, processed a1 mRNA in old *MATa* cells, but not in young cells.

I repeated these experiments using modern tools that make the experiment technically easier for me than it was for Smeal in 1996. Like Smeal, I biotinylated the cell wall of a founding population of yeast by incubating them with sulfo-NHS-biotin, which attaches biotin to any solvent-exposed free amine. Unlike Smeal, I had access to the mother enrichment program strain that was developed by Lindstrom in 2009, which kills daughter cells allowing long-term culturing of mother cells (Lindstrom and Gottschling, 2009). In addition, since Smeal had piloted magnetic isolation of biotinylated cells, a number of labs have improved on the method by including a sedimentation step in a Percoll gradient that removes extracellular detritus from the culture (Chen et al., 2003; Lindstrom and Gottschling, 2009; Park et al., 2002). I cultured cells for 48 hours, with one change of media after 24 hours, then purified ~10^6 old *MAT***a** cells, along with matched cells that were grown in the presence of the Sir2 inhibitor nicotinamide (NAM) to compare aging-dependent silencing effects to Sir2-dependent silencing effects. If old yeast showed the same gene expression as NAM-treated yeast, it would support the model that old cells lose silencing. Smeal never did this control in 1996, though he had access to strains that had loss-of-

function mutations in *sir4*, which would have been essentially the same control as the NAM experiment I performed. When I compared *HMR**a1*** expression in young *MAT*α cells to *HMR**a1*** expression in old *MAT*α cells by qRT-PCR, I found that expression of the silent RNA did appear in increase, as was reported by Smeal in 1996. However, the level of *MAT**a1*** mRNA expression never approached the level of expression in cells in which Sir2 activity was inhibited by NAM (fig 4-3). I concluded that the small increase *in HMR**a1*** expression I observed in old cells is likely a consequence of the difficulty associated with dealing with small numbers of cells, and small amounts of RNA.

Another downstream readout of *MATa1* expression is the activity of the mating pathway. If *HMR**a1*** is expressed concurrently with *MAT*α*2*, as would be expected in a cell

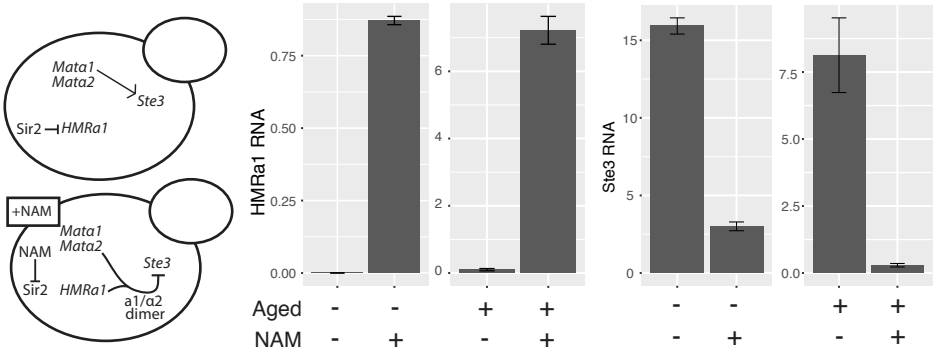

Figure 4-3 – (left) Schematic of silencing-dependent mating gene regulation. (right) qPCR for silent RNAs and downstream targets of regulation by silent genes in young and old cells, normalized to ACT1 expression. I found that aging did not affect silenced RNAs or genes regulated by proteins encoded at silent loci.

that has lost heterochromatin regulation, the cell will adopt a diploid like gene expression program, and therefore repress genes required for mating. Thus, under the model that old cells lose silencing, one would expect to see that the α-factor receptor gene *STE2* is expressed at diploid-like levels. Using the same RNA samples in which I analyzed *HMR**a1*** expression, I asked whether *STE3* expression in old cells reflects a haploid-like or a diploid-like gene expression program, and found that RNA expression in old cells reflects a haploid-like gene expression program, again suggesting that mating type silencing is functional in old cells (fig 4-3).

In addition to regulating the mating pathway, the Sir complex regulates expression of many genes, especially those found at telomeres and those regulated by the mating genes. Rather than test loci one at a time by qRT-PCR, I reanalyzed published RNA sequencing data sets from old cells to ask whether gene expression changes in old cells are related to changes associated with Sir2 mutants. My timing was good to ask this question, in that I could rely on RNA sequencing data generated by Aisha Ellahi shortly before I joined the lab (Ellahi et al., 2015). Using Aisha's data to identify the genes that are regulated by Sir2, I used two aging data sets to ask whether that subset of genes shows interesting regulation during a cell's lifespan.

The first aging data set I analyzed was generated by Georges Janssens in Matthias Heinemann's lab. In their study, Janssens *et al* developed a method to continuously culture yeast mothers immobilized on a magnetic lattice. They sampled yeast mothers during a 72

hour time course and reconstructed a trajectory for gene expression changes during yeast aging (Janssens et al., 2015). In addition to quantifying RNA directly, Janssens *et al* developed a method to understand the contribution of contaminating young or dead cells to their measurements of old cell RNA. This approach let them estimate the maximum likelihood abundance for each RNA species in old yeast, and resulted in eerily smooth kinetic data for each RNA. For my part, I pulled their processed and unprocessed data tables and plotted the kinetic trajectories of each Sir2-regulated mRNA identified by Aisha. From hierarchical clustering (Euclidean distance) of the heatmap of the mRNA kinetic trajectories, I found that there is no coherent kinetic trajectory for genes regulated by Sir2 over a 72-hour aging experiment.

The second data set I analyzed was similar in spirit—it was generated by Payel Sen in Shelley Berger's lab, and it measured RNA abundance in aged yeast populations. Here, however, the RNA sequencing protocol was optimized for short fragments, and the authors documented evidence of H4-K16-dependent aberrant transcription at subtelomeric loci in aging cells (Sen et al., 2015). In re-analyzing Sen's data, I started by asking whether RNA expression in their young cell population matched RNA expression measured in Ellahi's wild-type strain. The RNA expression profiles were very similar (fig 4-4). Next, I asked whether the changes that Ellahi observed when comparing wild-type strains to *sir2*Δ strains were similar to the differences observed when comparing young to old cells. Under the model that Sir2 function is lost as cells age, one would expect the comparison between old cells and young cells to reflect the comparison between *sir2*Δ and wild-type cells. I found that the changes that occur during aging are not similar to the changes that reflect loss of Sir2 function (fig 4-4). This suggested that simple loss of Sir2 function is not sufficient to explain age-associated mRNA phenotypes.

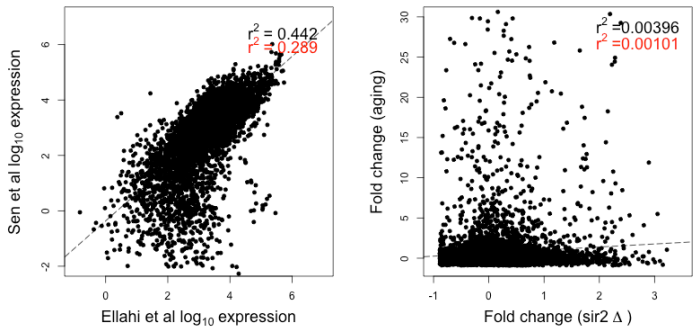

Figure 4-4 – (left) Gene expression is similar in the experimental strains used by Ellahi & Sen in exponentially growing cells. (right) The changes observed during aging by Sen et al were not related to the changes observed in *sir2*Δ strains, suggesting that *sir2*Δ does not phenocopy the aged cell condition.

When it came time to publish the aging experiments, I elected to publish the second analysis. The data I used—published originally by Sen *et al*—came from a lab that had previously argued that Sir2 becomes non-functional as cells age, resulting in loss of transcriptional regulation at telomeres (Dang et al., 2009). I wanted to avoid the criticism that my work reflected something idiosyncratic about the strain I used, so I decided to head it off by using data generated by the very lab that would be likely to be critical of my conclusions (Schlissel et al., 2017).

The story as I've told it so far essentially reflects the status of the project in mid 2015. I had shown that in several strain backgrounds by several methodologies that old yeast

cells do not lose silencing of their heterochromatic auxiliary mating type loci. I had not made any attempt to replicate the original data that gave rise to the model—that old cells fail to mate. In mid 2015, Jasper and I were alerted to the fact that a group in Switzerland—Marek Kryzanowski and Fabrice Caudron in Yves Barral's lab—had started to look at mating in old cells. Kryzanowski and Caudron had shown that old cells do not arrest efficiently in the presence of α-factor, however their daughter cells arrest normally. They determined that this effect was related to aggregation of Whi3, a protein required for cell cycle arrest during mating. Furthermore they demonstrated that the age-associated decrease in mating is actually a decrease in sensitivity to α-factor, a fact that was actually apparent to Smeal in 1996, which he cited as reason to perform mating experiments using abnormally low concentrations of α-factor (Smeal et al., 1996).

We reached out to Yves Barral and decided to finish our experiments in collaboration with one another and submit our results as a single project. Together, we dismantled the model that old cells behave like diploid cells, and replaced it with a new model: that old cells do not show diploid-like mRNA expression, but nonetheless fail to mate as a result of aggregation of a cell cycle regulator (Schlissel et al., 2017). Our paper is a significant step forward in yeast aging research, by redirecting attention away from the erroneous model that metabolic or transcriptional changes render Sir2 incompetent in old cells towards a new model, where age-associated phenotypes can be governed by changes in protein folding.

4.2 dCas9 BirA

My first attempt to track nucleosomes in space and time crashed and burned. Instead of using a well understood protein like Sir4 or the Tet repressor to control the activity of BirA, I attempted to use dCas9. The project failed in a way that should have been predictable in retrospect: at no point during the project did I have a convincing positive control, and I was not careful to establish benchmarks to know if I was making progress towards my goal. Here I recount some of my attempts to engineer a dCas9-BirA fusion protein that would label nucleosomes with spatial and temporal specificity. The experiments inspired the approaches that ultimately became the core of my thesis, and the later work would likely not have been possible without the early failures.

Early experiments in my tethered BirA project explored the possibility of using dCas9 to tether BirA to a defined genomic locus. This approach would have a big advantage compared to Sir4-BirA or TetR-BirA fusions, in that it could be targeted to an arbitrary sequence to explore the behavior of sub-regions of heterochromatin or euchromatin domains. For example, if dCas9-BirA fusions worked well, it would be possible to tether BirA to the *HML-E* silencer, the *HML-I* silencer, and to an arbitrary number of positions between the two, and ask if each nucleosome at *HML* shows consistent behavior in biotin tracking experiments. The same experiment could be performed at a euchromatic locus: by tethering dCas9 to sequences across a gene body, and ask if transcription relates to nucleosome movement as has been suggested by genome-wide nucleosome turnover experiments (Radman-Livaja et al., 2011).

I started by making dCas9 fusions, replacing Mxi1 in a dCas9-Mxi1 plasmid with BirA cloned from XL1-BLUE *E coli*. The dCas9 plasmid was already in a yeast compatible vector and had been previously used effectively in yeast (Gilbert et al., 2013). Separately, I designed a plasmid that would express a guide RNA from the SNR52 promoter, which is a strongly expressed small nuclear RNA promoter. Genes encoding histone H3 in this experiment were deleted in the genome, but I expressed an H3-Avi allele from a HIS plasmid. To test the function of the dCas9 fusion, I grew cells in selective medium to maintain the guide RNA and the dCas9-BirA plasmids and monitored biotinylation of H3-Avi. I built guide RNAs to target each mating type locus, as well as two arbitrary euchromatic loci (*POL4* and *PGK1*). One of my mating-type-specific guide RNAs was designed to target *MATα2*, which was not present in my *Mata* experimental strain.

On reflection, there were huge problems with this experimental design. The biggest problem was that I had no experience with immunoblots, and my blots were consistently uninterpretable. At the time, I was routinely convinced they were *just* good enough blots to interpret, and that my experiments were working well enough to continue the project. Another other major problem was that I had no effective positive control to confirm that I can detect relevant amounts of biotinylation. My plan was to look for signal above background, where background was defined as the biotinylation observed in a strain carrying a *MATα2*-targetted dCas9-BirA without a *MATα2* target available in the cell. In a few experiments, I saw enrichment of 1.5x H3-Avi-biotin over background but had no concept whether 1.5x signal was good or bad. The last major problem that I encountered was that when I tried to do ChIP to identify biotinylated histones, I was not familiar enough with the methodologies to set myself up for success—I did not appreciate differences between types of magnetic beads and I did not understand how important stringent washing is in maximizing the signal to noise in ChIP. I was turning too many knobs in the experimental design simultaneously with no credible strategy to organize the madness.

One of the more off-the-wall ideas that I explored during this period of random-walking through dCas9-BirA was the possibility of using a split BirA allele fused to two different dCas9 proteins to direct biotinylation to a locus determined by the co-occurance of two different sgRNAs. At the time, I thought there was background in the measurement of H3-Avi-Biotin that came from dCas9 sampling the genome randomly, and that this background could be reduced by requiring two dCas9 molecules to simultaneously bind neighboring sequences. This idea was inspired by the FokI approach to genome editing, which similarly depends on the co-binding of two FokI fusion proteins to achieve greater specificity in DNA target selection (Guilinger et al., 2014; Tsai et al., 2014).

To design the split BirA protein, I started from the published crystal structure of the enzyme. Any split should kill the enzymatic activity attributable to either half of the protein, so I found potential split locations (after position 195 and after position 217) that separated two domains that come together to form the active site. The last thing I needed was a method to test enzymatic activity on the condition that the two domains are associated, and for that I turned to a split GFP protein that I was familiar with from my rotation in the Ingolia lab (Blakeley et al., 2012). My plan was to fuse each half of BirA to a domain of GFP, and monitor biotinylation of an Avi-tagged cytoplasmic protein. The ideal technology would have no activity when either BirA domain is independently expressed, no activity when both domains are expressed but not tethered, and full activity when the domains are co-expressed and tethered. The split designs that I executed satisfied the criterion that each domain

shouldn't independently have biotinylation activity, but both designs showed biotinylation in the absence of a tether, suggesting the BirA domains had strong affinity for one another.

There is a clear approach to solve problems like the one I encountered making the split BirA technology. I could have systematically mutated residues at the interaction surface between the split BirA peptides and screened for alleles that are enzymatically active only when the domains are forced together. But for me, the split BirA was designed to solve a problem that I wasn't sure existed, and it seemed like too much effort for a hack that might be difficult to execute, and might not put me any closer to my actual goal: to label nucleosomes with spatial and temporal specificity.

As I approached my first committee meeting, I came to appreciate that I was really flailing without a path to success with the BirA project, so I decided to back off and try a new, more systematic approach. The updated plan was to use Sir3 or Sir4 to tether BirA to silent chromatin, instead of using dCas9. This plan had several advantages. First, because Sir proteins are well-localized in yeast, I had a strong expectation that they would work as effective DNA tethers. Furthermore, because they have previously been used in ChIP with 13x-myc tags at their C-terminus, I had an expectation that the C-terminal BirA fusion would not affect their localization or function (Thurtle and Rine, 2014). Lastly, because the Sir complex binds at least ~100 nucleosomes at silent chromatin and at telomeres, I thought the signal would be 100x stronger than in the case of dCas9, which would make blots for biotinylated H3 more convincing, and by using blots to monitor activity of BirA instead of ChIP would let me iterate much more rapidly on experimental parameters. This was the status of my thinking when Davis Goodnight joined the lab for a rotation. Davis ultimately took on the tech development of Sir3,4-BirA fusions for his rotation project, which gave me a chance to systematically re-think the project with the help of his fresh eyes, and together we developed strong negative controls at every step (though we were still missing positive controls) to convince ourselves that the signal from Sir3,4-BirA was convincingly above background.

After Davis's rotation, I took over the tech development again and the project really started moving, and the results of those experiments were reported in Chapter 1.

4.3 Linear vs circular HML

Plasmids can be linearized by adding sequences that are converted to telomeres in vivo. When *HML*-CRASH is put on a linear plasmid, silencing seems to be come ultra-stable, and adjacent markers on the plasmid are silenced too. It is extremely hard to do experiments without being able to follow plasmids, so most of these experiments are unconvincing. This work makes no strong conclusions, but suggests that silencing is more stable on a linear HML plasmid with telomeres than it is on a circular HML plasmid without telomeres. This work never resolved the relative contribution of telomeres and of DNA topology: either could be responsible for the silencing phenotype.

A half-develop question that became a credible rotation project asked: what is the contribution of DNA topology to silencing? I find it interesting that both *HML* and *HMR* occur near telomeres, and I wondered if the fact that telomeres are less topologically constrained—that they are free to rotate—might contribute to silencing at *HML* or *HMR*.

At the time of this question, I knew of one line of research that involved modifying the topology of *HMR*. In that work, Ann Kirchmaier used the FLP/FRT recombinase system to loop *HMR* out of the genome in cells that were pre-grown in conditions that prevent Sir silencing (Kirchmaier and Rine, 2006). Then she tested the conditions under which silencing was established on the episomal DNA, and determined that silencing could be established, but that it depended on the cells passing through S-phase (Kirchmaier and Rine, 2006). Ann's work built on previous work showing that silencing requires passage through S-phase, but surprisingly replication of the *HMR* episome is not required for silencing establishment (Kirchmaier and Rine, 2006; Lau et al., 2002; Miller and Nasmyth, 1984). Because she was studying the contribution of DNA replication to silencing, Ann's work did not explore whether the episomal DNA shows any special silencing properties compared to chromosomal silent DNA. Since the episomal *HMR* experiments, work in the lab had implied that silent chromatin adopts a higher-order structure during silencing, and in light of the superstructure hypothesis it seemed surprising to me that an episome would be capable of silencing at all.

To reexamine the model that episomal DNA can be silenced, I worked with a rotation student to build a plasmid that contained *HML::CRE* as well as telomere nucleating sequences derived from *Tetrahymena*. The plasmid contained a selectable *HIS3* marker, and between the converging Tetrahymena telomere nucleating sequences was the *URA3* gene. This design allowed us to maintain circular plasmids by selecting for the *HIS3* gene and for *URA3*. At a low rate, the plasmid spontaneously linearizes and linear plasmids can be selected on 5-FOA. We were never convinced of the mechanism by which the plasmid linearizes, but we were told by Vicki Lundblad, who made the original plasmid, and by Martin Kupiec, who works with similar plasmids regularly, that the system just works.

To ask whether DNA topology affects silencing at *HML*, I planned to compare the linear *HML* plasmid to the circular plasmid, and measure silencing with the CRASH assay. To execute the experiment, I transformed the circular *HML-TEL* plasmid, then selected the transformants on –HIS +5-FOA medium to select for plasmids that were linearized, which results in loss of the *URA3* marker. I tested whether the plasmids were in fact linearized by purifying DNA from the cells and digesting the plasmid with BamHI, which cuts the plasmid at one position. After digestion with BamHI, circular plasmids would run as a single band (of approximately 10kb) on an agarose gel, whereas a linearized plasmid would run as two bands approximately 6.5kb and 3.5kb. I monitored the restriction pattern of purified plasmid by DNA-blotting, using biotinylated probes against the *CRE* open reading frame. I detected biotinylated probes using streptavidin conjugated to infrared fluorescent dyes. This experiment was my first and only DNA blot, and it was not an overwhelming success. The ideal experiment would compare a strain with a circular plasmid to a strain with a linear plasmid, analyzing both for their digestion pattern with BamHI. Unfortunately, the sample I loaded to represent the circularized plasmid failed, and so my experiment had no effective circular-plasmid control. Despite the failure of the control experiment, the results from the experimental condition indicated that the plasmid was effectively linearized, as cutting with BamHI resulted in two bands of the appropriate size, based on the plasmid map. Surprisingly, the digested samples contained a band that corresponded to the length of the complete linear *HML-TEL* plasmid. It is possible that this band corresponds to incompletely digested DNA, or it is possible that the plasmid exists either as circular or linear DNA in the same population of cells. I did not repeat the DNA blot with a longer digest to assess those possibilities, but continued under the presumption that the 10kb band resulted from incomplete digestion.

Having shown that the circular plasmid can be transformed, maintained, and linearized in a haploid strain, I asked whether the circular and linear plasmids were stably maintained in mitotic cells. I grew cells to saturation without selection, and spotted a 5-fold dilution series onto solid medium to measure the viability of cells on media designed to select for the plasmid (-His), for the circular plasmid (-His –Ura), or for the linear plasmid (-His + 5-FOA). Strains with linearized plasmids all grew more than strains with circular plasmids on –His +5-FOA media, but one strain (GSy101+ pGS162-L1) grew better than the others by this assay, and that strain appeared to have a general growth advantage on –HIS media compared to all other linearized strains. I excluded GSy101 + pGS162-L1 from further analyses, based on the fact that it was an outlier among the strains with linearized plasmids. I expect that the strain may have picked up a genomic copy of the *HIS3* marker by integration of the plasmid in the genome, but I excluded the strain without formally resolving the disparity in growth phenotypes.

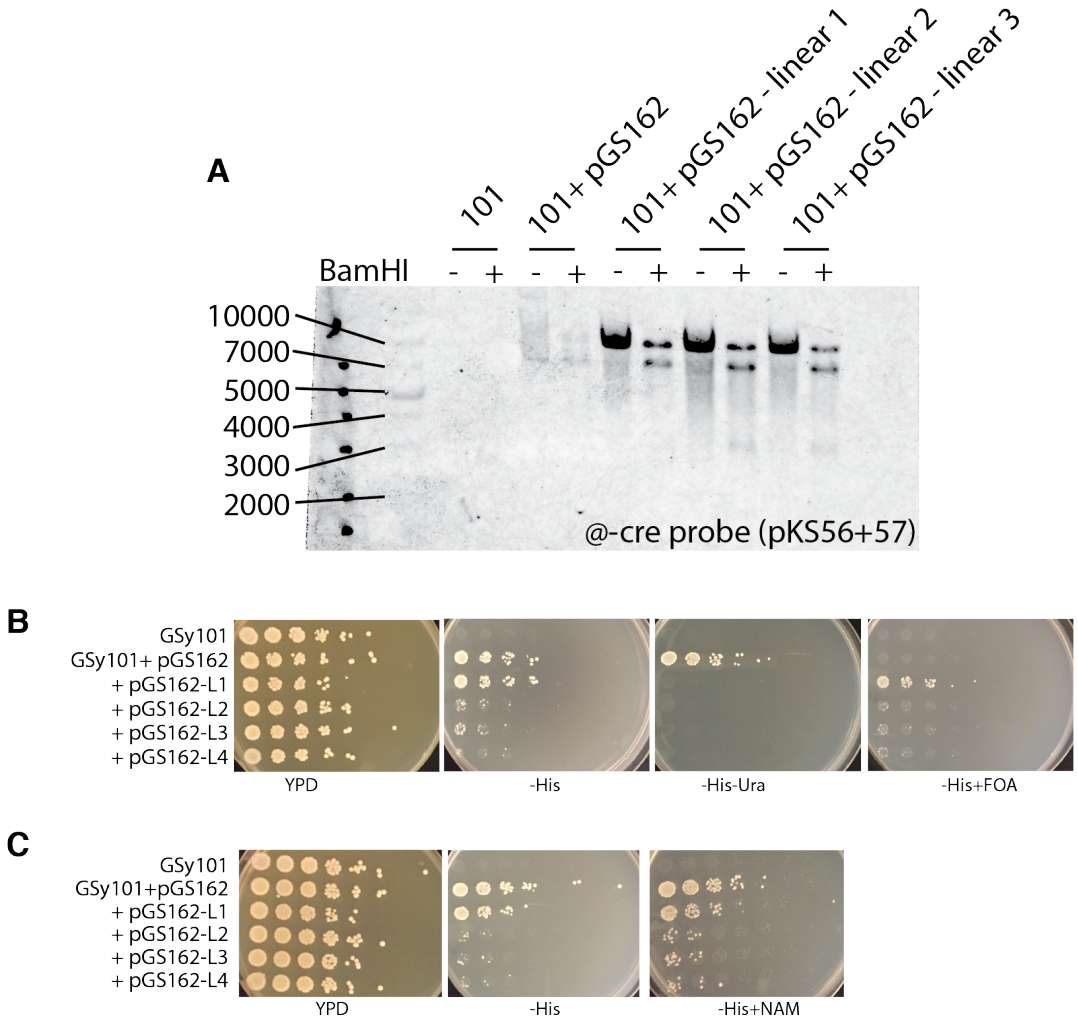

Figure 4-5 Circular and linear plasmids show distinct growth phenotypes. A) +DNA- blot on plasmid DNA prepped from *S. cerevisiae* and cut with BamHI. Linearized plasmids show a banding pattern that is consistent with linear topology. The circular control plasmid did not give enough signal to make an accurate comparison, but linear plasmids showed banding patterns consistent with linear topology. B) Circular plasmids express the *URA3* marker, whereas linear plasmids do not. One candidate (-L2) linearized plasmid grows better than other candidates on –His and on 5-FOA, and was assumed to reflect a suppressor mutation in the strain. C) In linearized plasmids, the *HIS3* marker is silenced by the Sir complex, whereas in circular plasmids the *HIS3* marker is not silenced.

.

The high apparent rate of plasmid loss in strains with the linear *HML*-TEL plasmid could be due to defects in inheriting a linear plasmid, or could alternatively be due to Sir-based silencing blocking expression of the *HIS3* reporter. To address the possibility that the Sir complex silences *HIS3*, thereby making the plasmid appear less stable than it truly is, I repeated the drop dilution experiment using medium supplemented with nicotinamide, which inhibits Sir2 and prevents Sir-based silencing. Strains that contained linear plasmids showed markedly better growth on –HIS +NAM media compared to strains grown on –HIS media, suggesting that Sir silencing on the linearized plasmids affects the expression of the plasmid's *HIS3* marker (fig 4-5). By contrast, NAM had no effect on growth of strains with a circular *HML-TE*L plasmid, suggesting that Sir silencing of the *HIS3* marker is a unique feature of the linear plasmid (fig 4-5).

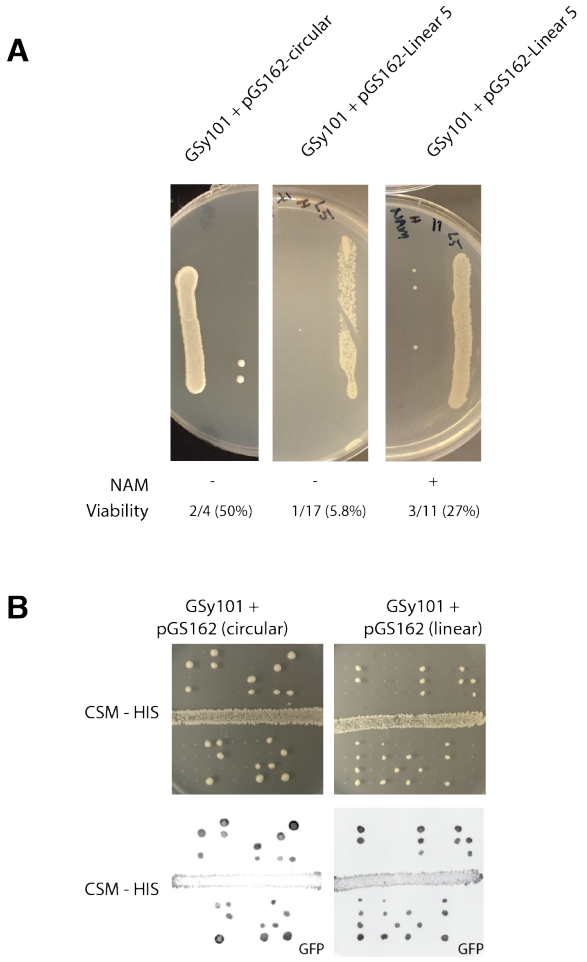

Figure 4-6 A) Zygotes were identified microscopically, and separated by micromanipulation onto media lacking histidine, with or without the addition of nicoctinamide (NAM). The fraction of viable zygotes was scored for each condition and genotype. B) Tetrads were dissected from diploids that contained either the circular (left) or the linear (right) HML-TEL plasmid. Both plasmids showed a high rate of plasmid loss, and the linear plasmid showed non-mendelian inheritance. Neither plasmid showed evidence of silencing, but the result may be deceiving as addressed in the text.

The strain that maintains the linear plasmid is clearly imperfect, in so far as it requires constant selection for the *HIS3* marker to maintain the plasmid. Nonetheless, I thought it would be valuable to back-cross the circular and linear *HML-TEL* plasmids and monitor co-segregation between each plasmid and the silencing phenotype. To execute the experiment, I crossed the derived strains back to a strain with an un-switched LoxP CRASH reporter (RFP+). When picking zygotes, it is common to grow the zygotes on rich medium, then later move them to selective medium to identify the zygotes that retain the plasmid. For a typical plasmid, all zygotes retain the plasmid, and the rate of plasmid loss is said to be on the order of 10^{-3} per generation. When I picked zygotes in this experiment, the circular *HML-TEL* plasmid was maintained like any other. However, when I picked zygotes on rich medium to generate a diploid carrying the linear *HML-TEL* plasmid, no zygotes retained the plasmid. I considered two possible explanations for the effect: 1) that the linear *HML-TEL* plasmid is lost at an elevated rate in diploids compared to the circular *HML-TEL* plasmid, or 2) that the *HIS3* marker used to measure whether the plasmid is retained is not expressed on the linear *HML-TEL* plasmid, because silencing spreads from adjacent *HML* or telomeric silencers.

I addressed both possibilities by picking zygotes onto selective media, instead of onto rich media. To minimize the possibility that the *HML-TEL* plasmid is lost during cell division after mating, I picked zygotes again on –HIS media, to select for the linear *HML-TEL* plasmid from the earliest possible moment. When I did the experiment in this way, I found that among 28 zygotes picked only one formed a colony (3.5% of zygotes). If it were true that the *HIS3* reporter is silenced, I would expect zygotes that do not have Sir silencing activity to be HIS+ at a higher rate than cells with Sir silencing intact. Thus, I picked zygotes on media lacking histidine and supplemented with NAM, and found that three of 22 zygotes formed colonies (14% of zygotes, which is more than one would expect under the null model that NAM does not affect plasmid maintenance; $p=0.02$, binomial test) (fig 4-6). Sir2 activity was sufficient to prevent expression of the *HIS3* marker in some zygotes, but there was likely some non-Sir2 contribution to the zygotes' specific failure to maintain the linearized *HML-TEL* plasmid.

Having generated one diploid strain with a linearized *HML-TEL* plasmid in the absence of NAM, with a red CRASH reporter, I dissected tetrads to ask whether there is a silencing phenotype that co-segregates with the linear plasmid, and I compared the results to a tetrad dissection in which *HML* was maintained on a circular plasmid. The circular plasmid showed a high rate of plasmid loss: many tetrads (8/20, 40%) contained no spores that were His+, but the tetrads that contained His+ colonies each showed two spores with the *HIS3* marker and two without, suggesting that diploid strains maintained zero or one copy of the circular *HML-TEL* plasmid. By contrast the segregation pattern of the linear *HML-TEL* plasmid was non-mendelian. In many tetrads (11/21, 52%), no spores expressed the *HIS3* marker suggesting that the linear *HML-TEL* plasmid is lost at a high rate in diploid cells, as is the circular *HML-TEL* plasmid. Among the tetrads that contain at least one His+ spore, many tetrads showed one, three or four spores that express the His+ marker, suggesting that the plasmid is not segregated faithfully in meiosis.

Among the spores that grew and maintained either the linear or the circular plasmid (which identified by selection on –HIS medium), all spores had an extreme silencing phenotype as measured by the CRASH assay. Among the spores with circular plasmids, the silencing rate was very high but sectors could still be resolved, and had the typical slice-of-pie geometry as one would expect for a colony growing without selective pressure for or against the sectored cells (fig 4-6). However, among the spores with linear plasmids, on –

HIS medium the colonies were virtually completely green, and to the extent that individual sectors could be resolved, they did not show the typical piece-of-pie geometry (Figure 4-6). I interpreted the high occurrence of green cells and the degenerate geometry of the sectors to indicate that the cells, which were explicitly selected for their expression of the *HIS3* marker, were indirectly selected for being green. This result is consistent with drop dilution assays on –HIS and –HIS +NAM, which suggested that expression of the *HIS3* marker is related to Sir complex-mediated silencing, which is active at the adjacent *HML* cassette and at the adjacent telomeres.

Because the plasmid was lost at a high rate, and because I did not have an effective selection for the plasmid that could isolate maintenance of the plasmid from silencing on the plasmid, I could not perform a clean experiment asking whether the topology of the *HML-TEL* plasmid affects silencing of *HML*. On the one hand, the abundance of green cells on medium lacking histidine for cells with a linear *HML-TEL* plasmid points to the interpretation that the linear *HML-TEL* plasmid cannot silence *HML-CRASH* effectively. However I could not exclude the possibility that selection on media lacking histidine selected a subpopulation of cells that cannot silence the linear *HML-TEL* plasmid. To resolve this concern I would have needed to identify a linear plasmid that is maintained with high enough fidelity that I could perform experiments without constantly selecting on –HIS media. I considered the possibility of inserting extra DNA sequence on the linear plasmid to isolate the *HIS3* reporter from *HML* and from the telomeres. My strategy for performing the experiment would have been similar to the strategy used in the synthetic *S. cerevisiae* genome project, with iterative rounds of inserting and replacing markers to grow the chromosome incrementally. Such an approach seemed technically achievable but incredibly dull, and I had more interesting issues to pursue.

As an alternative to the plasmid-based approach, I could have redesigned the experiment to move *HML* and *HMR* away from the telomeres, thus introducing some topological constraint to the motion of the wild-type *HML* or *HMR* loci. If I were to continue with this project, I suspect that approach would be more fruitful. Conveniently, another graduate student is working on a project that would scramble the position of *HML* on chromosome III and if there were a general principle about DNA topology and silencing, I would expect that experiment to reveal it.

Of all the small puzzles that I encountered in this work, one seemed particularly curious: that the linear and circular plasmids showed very different loss rates during mitosis, even in the presence of NAM. When looking at the data, Jasper was reminded of his previous experience in working with *HML* plasmids: according to his recollection, *HML* plasmids that do not contain a centromere are nonetheless very well maintained, and in some cases are more effectively maintained if the plasmid does not include a *CEN* sequence. The explanation that Jasper favored for the phenomenon was that *HML* plasmids form heterochromatin, which can hitchhike on other yeast heterochromatin through cell division, stabilizing the inheritance of the plasmid, and the formation of heterochromatin might inhibit formation of a kinetochore. It is possible that if a chromosome contains both a centromere, such that it aligns at the metaphase plate, and large amounts of heterochromatin which are thought to stick together, that the rate of nondisjunction would be greater than if the plasmid was inherited only through "stickiness" to other heterochromatin, or through normal kinetochore activity as in a normal *CEN* plasmid alone.

4.4 Abf1 acetylation

In a departmental seminar for the department of Nutritional Sciences and Toxicology, Dave Toczyski presented work that sought to identify every regulated acetylation event in the yeast proteome. Part of his approach was to delete each sirtuin and ask which acetylation marks are different in cells that lack sirtuin deacetylases. Those acetylation marks that change are candidates for functional regulation of protein activity by acetylation. In Dave's presentation, a protein name he included as an example of a candidate for a protein that is regulated by acetylation jumped out at me: it was Abf1.

Abf1 was an interesting protein to find on his list for two reasons. First, Dave's list identified proteins that are regulated by Sir2 and its parologs, which are involved in maintaining chromatin structure, my primary interest. Second, Abf1's biological role is a paradox: it is an essential gene and a broad activator of hundreds of yeast genes, but it also plays a role in the establishment of gene silencing at *HML* and *HMR*., I reasoned that Dave's acetylation data might actually reconcile the paradox. It is possible that Abf1 exists in two populations in the cell—one population activates many genes, and a separate population represses genes. My hypothesis was that the populations are distinguished by acetylation of the protein at key residues, and that forcing Abf1 to be acetylated would limit it to behaving as either an activator or a repressor but would disallow it from acting as both.

My approach to testing the hypothesis was straightforward. I would delete the endogenous copy of *ABF1* in a diploid strain that contains the CRASH reporter to measure the stability of silencing, then add back a plasmid that contained either a wild-type copy of *ABF1* or a mutant copy of *ABF1*. The interesting *ABF1* mutants would mimic acetylation or no acetylation of critical lysines as positions 3, 12, 196, 383, 518 and 584. I was able to make all of the mutants concurrently using Gibson Assembly. Then I induced the diploid to sporulate and asked whether haploid segregants that inherit a mutant *ABF1* plasmid showed a different silencing phenotype compared to haploid segregants that inherited a wild-type *ABF1* plasmid, in both cases looking only at the haploid segregants that are deleted for *ABF1* at its chromosomal locus.

In my first experiment, I determined that the acetylation of Abf1 at positions 3, 12, 196, 383, 518 and 584 was not involved in silencing at *HML*, based on the observation that each *ABF1* allele showed similar heterochromatin stability as measured by the CRASH assay (fig 4-8).

Although the CRASH result was convincing evidence that Abf1 acetylation or deacetylation is not required for silencing at *HML*, it did not disprove the idea that Abf1 acetylation might be important for silencing in some contexts. Early work mapping the yeast silencers showed that the Abf1 binding site at the *HMR-E* silencer is only essential for silencing in some genetic contexts. Silencing involves a "two of three" rule, such that having two silencing factors of a possible three at a silencer is sufficient to establish silencing (Brand et al., 1987; McNally and Rine, 1991). Thus, to test whether Abf1 acetylation can be involved in silencing stability, it would make sense to ask whether Abf1 acetylation affects silencing in a context where Abf1 function is known to be strictly required for silencing. To achieve this, I returned to the lab's earlier work, in which Frank McNally built a synthetic silencer that could silence *HMR* effectively, unless the ABF1 binding site was mutated (McNally and Rine, 1991).

To ask whether ABF1 acetylation is involved in silencing in a context where Abf1 is required for silencing, I crossed haploid strains that contained *ABF1* plasmids covering

chromosomal *abf1*Δ alleles to a strain that contains an *HMR-E*-synthetic silencer allele that strictly requires Abf1 binding to silencing *HMRa1* and *HMRa2*. I sporulated the strain and asked whether *Matα* haploid segregants that contain only a mutant *ABF1* allele, and contain the mutant *HMR-E*- synthetic silencer are capable of mating. Because of a quirk of the history, I only performed *HMR-E*-synthetic silencer experiments for K-Q mutants, and not for the K-R mutants (at the time that I did the experiment, I was not confident in the K-R strains, and so I excluded them). I performed quantitative mating assays to measure silencing at *HMR*, and found that *ABF1* K-Q alleles showed similar a similar frequency compared to wild-type, suggesting that Abf1 deacetylation by Sir2 was not involved in the mechanism of

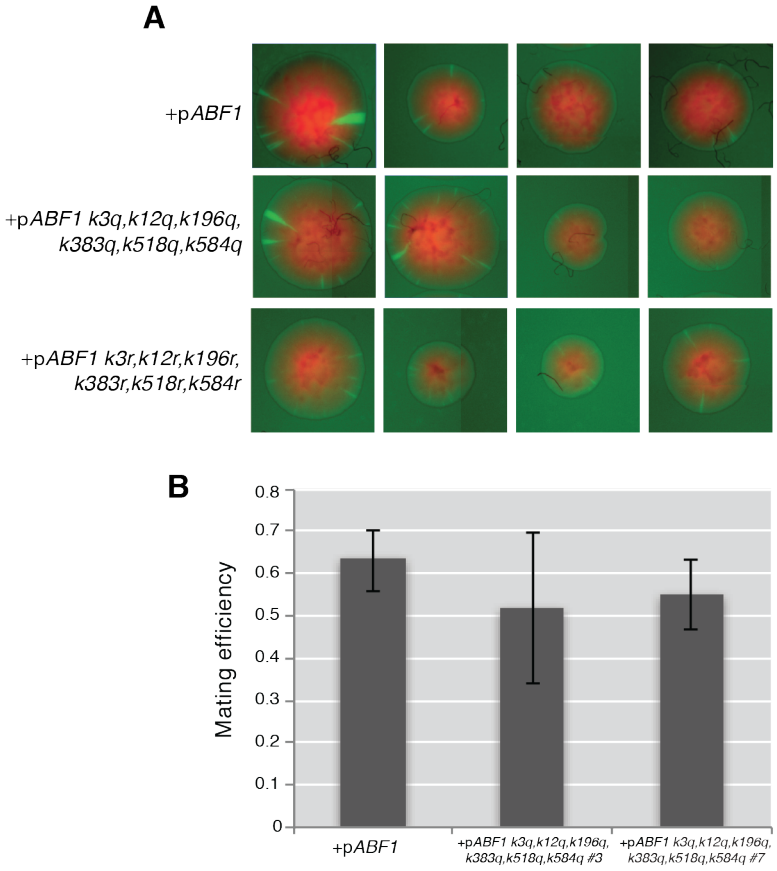

Figure 4-7 A) Strains expressing only mutant *ABF1* show no obvious difference in silencing stability compared to strains expressing wild-type *ABF1*. In all strains, *ABF1* is deleted at its chromosomal locus, and the only copy of *ABF1* is expressed from a centromeric plasmid. Note that the scanned images are fuzzy because this image was captured after replica-plating a tetrad dissection plate B) Mating efficiency of *Mata* strains with an *ABF1*-dependent *HMR-E*-synthetic silencer alleles is plotted for strains expressing wild-type *ABF1* or a mutant *ABF1* that cannot be deacetylated by Sir2. K-Q point mutants in *ABF1* had no impact on silencing of *HMRa* information, even in a context where Abf1 is essential for silencing.

silencing at *HMR*, in a context in which Abf1 was required for silencing (fig 4-7).

When I started thinking about Abf1, I thought Abf1 deacetylation might be a strict requirement for silencing. On paper, it made sense that because Sir2 is proximal to Abf1, is known to deacetylate Abf1, and Abf1 has two functions in the cell that are apparently opposite one another, there must be something functional and interesting happening between Sir2 and Abf1 at the silencers. So when I found that the *ABF1-KQ* and *ABF1-KR*

mutants didn't have a phenotype in the CRASH assay and that *ABF1-KQ* didn't have a mating phenotype in the *HMR-E*-synthetic silencer strain, I decided that it would be better to quickly wrap up the project than to keep hunting until I convince myself there's a subtle phenotype. I set out looking for some reaction that was big and critical for silencing, and I quit after I convinced myself that whatever fruit the project would bear would not be low hanging and pursuing them would distract from my core thesis.

5 Books & media that were influential along the way

- The Eighth Day of Creation – among other things, introduced me to the logic of molecular biology, and introduced me to my (still) favorite experiment I have ever heard of. Pardee & Riley 1969, where they used radioactive decay to degrade DNA in frozen cells.
- The biography of John Boyd – introduced me to von Clausawitz's "water flowing down a hill," and introduced me to some of the rigor of academics, in thinking about Boyd's "patterns of conflict" project, and his love of deep thought. After I read the book, I started dividing my efforts across multiple strategies, following the ones that seem easy and fun and dropping the ones that met resistance, like water flowing down a hill.
- The Black Swan – the book devolves a little into a half-developed discussion of Mandelbrot at the end, but the book includes an incredible passage about "waiting in the antechamber of hope," that handles the friction between people with quarterly psychology (i.e. beep-boop-boppers) and people who live in their minds. It really is a hard thing to communicate what it means to go to work every day expecting and cherishing failure.
- Zen and the art of motorcycle maintenance. It was a terrible and pedantic book that wholly misunderstands induction, and if its worldview were realized science would be equivalent to bean counting. By its counter-example, it convinced me that the kind of science that's fun and beautiful is a lot more like cloud gazing than like bean counting.
- An American Sickness. Boy, what a work of journalism. Just great.
- The Structure of scientific revolutions. The anecdote that I remember from this one is about how our predictions of the planets' positions was worse immediately after the Copernican revolution than it had been before. Science can take a leap forward by taking a step backward. I need to re-read the book and engage with it some more. I read it in a haze during my first year and I'm sure the subtlety evaded me
- In the Garden of Beasts. On the eve of the second world war, Germans repeated as fact that during the first world war, Americans rounded up and executed Germans in public. There was such little communication between German and American civilians that there was no counter-narrative among the general public. They accepted it as fact. We really need to integrate our culture with that of other nations if we want to prevent future conflict. Also if I were ever in a position to appoint ambassadors, I think it would be nice to appoint professors like they did during that era.
- Slaughterhouse five. There is a character in the book who is an American working for the Germans, to explain American psychology. The character does an incredible job of playing with the irony that American poor people believe in the virtue of rich people.
- Just Mercy by Byran Stevenson. Nothing gets me riled up like systematic injustice, and this book is full of reasons to get riled up.
- The Invisible Man by Ralph Ellison. Clever and absurdly hilarious.

6 References

Abraham, J., Nasmyth, K.A., Strathern, J.N., Klar, A.J.S., and Hicks, J.B. (1984). Regulation of mating-type information in yeast. Negative control requiring sequences both 5' and 3' to the regulated region. J. Mol. Biol.

Armache, K.J., Garlick, J.D., Canzio, D., Narlikar, G.J., and Kingston, R.E. (2011). Structural basis of silencing: Sir3 BAH domain in complex with a nucleosome at 3.0 Å resolution. Science (80-.).

Audergon, P.N.C.B., Catania, S., Kagansky, A., Tong, P., Shukla, M., Pidoux, A.L., and Allshire, R.C. (2015). Restricted epigenetic inheritance of H3K9 methylation. Science (80-.).

Bannister, A.J., and Kouzarides, T. (2011). Regulation of chromatin by histone modifications. Cell Res.

Belotserkovskaya, R., Oh, S., Bondarenko, V.A., Orphanides, G., Studitsky, V.M., and Reinberg, D. (2003). FACT facilitates transcription-dependent nucleosome alteration. Science (80-.).

Blakeley, B.D., Chapman, A.M., and McNaughton, B.R. (2012). Split-superpositive GFP reassembly is a fast, efficient, and robust method for detecting protein–protein interactions in vivo. Mol. Biosyst. *8*, 2036.

Bothma, J.P., Garcia, H.G., Esposito, E., Schlissel, G., Gregor, T., and Levine, M. (2014). Dynamic regulation of eve stripe 2 expression reveals transcriptional bursts in living Drosophila embryos. Proc. Natl. Acad. Sci. U. S. A. *111*, 10598–10603.

Brand, A.H., Micklem, G., and Nasmyth, K. (1987). A yeast silencer contains sequences that can promote autonomous plasmid replication and transcriptional activation. Cell.

Braunstein, M., Rose, A.B., Holmes, S.G., David Allis, C., and Broach, J.R. (1993). Transcriptional silencing in yeast is associated with reduced nucleosome acetylation. Genes Dev.

Brewer, B.J. (1988). When polymerases collide: Replication and the transcriptional organization of the E. coli chromosome. Cell.

Canzio, D., Larson, A., and Narlikar, G.J. (2014). Mechanisms of functional promiscuity by HP1 proteins. Trends Cell Biol.

Chen, C., Dewaele, S., Braeckman, B., Desmyter, L., Verstraelen, J., Borgonie, G., Vanfleteren, J., and Contreras, R. (2003). A high-throughput screening system for genes extending life-span. Exp. Gerontol. *38*, 1051–1063.

Chih, L.L., Kaplan, T., Kim, M., Buratowski, S., Schreiber, S.L., Friedman, N., and Rando, O.J. (2005). Single-nucleosome mapping of histone modifications in S. cerevisiae. PLoS Biol.

Cusick, M.E., DePamphilis, M.L., and Wassarman, P.M. (1984). Dispersive segregation of nucleosomes during replication of simian virus 40 chromosomes. J. Mol. Biol.

Dang, W., Steffen, K.K., Perry, R., Dorsey, J.A., Johnson, F.B., Shilatifard, A., Kaeberlein, M., Kennedy, B.K., and Berger, S.L. (2009). Histone H4 lysine 16 acetylation regulates cellular lifespan. Nature *459*, 802–807.

Dodson, A.E., and Rine, J. (2015). Heritable capture of heterochromatin dynamics in Saccharomyces cerevisiae. Elife *4*, e05007.

Ehrenreich, I.M., Torabi, N., Jia, Y., Kent, J., Martis, S., Shapiro, J.A., Gresham, D., Caudy,

A.A., and Kruglyak, L. (2010). Dissection of genetically complex traits with extremely large pools of yeast segregants. Nature *464*, 1039–1042.

Ellahi, A., Thurtle, D.M., and Rine, J. (2015). The Chromatin and Transcriptional Landscape of Native Saccharomyces cerevisiae Telomeres and Subtelomeric Domains. Genetics *200*, 505–521.

Feldman, J.B., Hicks, J.B., and Broach, J.R. (1984). Identification of sites required for repression of a silent mating type locus in yeast. J. Mol. Biol.

Gartenberg, M.R., and Smith, J.S. (2016). The nuts and bolts of transcriptionally silent chromatin in Saccharomyces cerevisiae. Genetics.

Gilbert, L. a, Larson, M.H., Morsut, L., Liu, Z., Gloria, A., Torres, S.E., Stern-ginossar, N., Brandman, O., Whitehead, H., Doudna, J. a, et al. (2013). CRISPR-Mediated Modular RNA-Guided Regualtion of Transcription in Eukaryotes. Cell *154*, 442–451.

Greenstein, R.A., Jones, S.K., Spivey, E.C., Rybarski, J.R., Finkelstein, I.J., and Al-Sady, B. (2018). Noncoding RNA-nucleated heterochromatin spreading is intrinsically labile and requires accessory elements for epigenetic stability. Elife.

Gruss, C., Wu, J., Koller, T., and Sogo, J.M. (1993). Disruption of the nucleosomes at the replication fork. EMBO J. *12*, 4533–4545.

Guilinger, J.P., Thompson, D.B., and Liu, D.R. (2014). Fusion of catalytically inactive Cas9 to FokI nuclease improves the specificity of genome modification. Nat. Biotechnol.

Hall, C., and Dietrich, F.S. (2007). The reacquisition of biotin prototrophy in Saccharomyces cerevisiae involved horizontal gene transfer, gene duplication and gene clustering. Genetics.

Hecht, A., Laroche, T., Strahl-Bolsinger, S., Gasser, S.M., and Grunstein, M. (1995). Histone H3 and H4 N-termini interact with SIR3 and SIR4 proteins: A molecular model for the formation of heterochromatin in yeast. Cell.

Hodges, C., Bintu, L., Lubkowska, L., Kashlev, M., and Bustamante, C. (2009). Nucleosomal fluctuations govern the transcription dynamics of RNA polymerase II. Science (80-.).

Hsieh, F.-K., Kulaeva, O.I., Patel, S.S., Dyer, P.N., Luger, K., Reinberg, D., and Studitsky, V.M. (2013). Histone chaperone FACT action during transcription through chromatin by RNA polymerase II. Proc. Natl. Acad. Sci.

Imai, S., Johnson, F.B., Marciniak, R.A., McVey, M., Park, P.U., and Guarante, L. (2000). Sir2: An NAD-dependent histone deacetylase that connects chromatin silencing, metabolism, and aging. In Cold Spring Harbor Symposia on Quantitative Biology, p.

Jackson, V. (1988). Deposition of newly synthesized histones: hybrid nucleosomes are not tandemly arranged on daughter DNA strands. Biochemistry *27*, 2109–2120.

Janssens, G.E., Meinema, A.C., Gonzalez, J., Wolters, J.C., Schmidt, A., Guryev, V., Bischoff, R., Wit, E.C., Veenhoff, L.M., and Heinemann, M. (2015). Protein biogenesis machinery is a driver of replicative aging in yeast. Elife *4*.

Jo, M.C., Liu, W., Gu, L., Dang, W., and Qin, L. (2015). High-throughput analysis of yeast replicative aging using a microfluidic system. Proc. Natl. Acad. Sci. 201510328.

Johnson, A., Li, G., Sikorski, T.W., Buratowski, S., Woodcock, C.L., and Moazed, D. (2009). Reconstitution of Heterochromatin-Dependent Transcriptional Gene Silencing. Mol. Cell.

Kaufman, P.D., and Rando, O.J. (2010). Chromatin as a potential carrier of heritable information. Curr. Opin. Cell Biol. *22*, 284–290.

Kayne, P.S., Kim, U.J., Han, M., Mullen, J.R., Yoshizaki, F., and Grunstein, M. (1988).

Extremely conserved histone H4 N terminus is dispensable for growth but essential for repressing the silent mating loci in yeast. Cell.

Kennedy, B.K., Austriaco, N.R., Zhang, J., and Guarente, L. (1995). Mutation in the silencing gene SIR4 can delay aging in S. cerevisiae. Cell *80*, 485–496.

Kirchmaier, A.L., and Rine, J. (2001). DNA replication-independent silencing in S. cerevisiae. Science (80-.).

Kirchmaier, A.L., and Rine, J. (2006). Cell cycle requirements in assembling silent chromatin in Saccharomyces cerevisiae. Mol. Cell. Biol. *26*, 852–862.

Korber, P., Barbaric, S., Luckenbach, T., Schmid, A., Schermer, U.J., Blaschke, D., and Hörz, W. (2006). The histone chaperone Asf1 increases the rate of histone eviction at the yeast PHO5 and PHO8 promoters. J. Biol. Chem.

Kushner, P.J., Blair, L.C., and Herskowitz, I. (1979). Control of yeast cell types by mobile genes: A test. Proc. Natl. Acad. Sci.

Kwan, E.X., Foss, E.J., Tsuchiyama, S., Alvino, G.M., Kruglyak, L., Kaeberlein, M., Raghuraman, M.K., Brewer, B.J., Kennedy, B.K., and Bedalov, A. (2013). A Natural Polymorphism in rDNA Replication Origins Links Origin Activation with Calorie Restriction and Lifespan. PLoS Genet. *9*.

Larson, A.G., Elnatan, D., Keenen, M.M., Trnka, M.J., Johnston, J.B., Burlingame, A.L., Agard, D.A., Redding, S., and Narlikar, G.J. (2017). Liquid droplet formation by HP1α suggests a role for phase separation in heterochromatin. Nature.

Lau, A., Blitzblau, H., and Bell, S.P. (2002). Cell-cycle control of the establishment of mating-type silencing in S. cerevisiae. Genes Dev. *16*, 2935–2945.

Lee, C.K., Shibata, Y., Rao, B., Strahl, B.D., and Lieb, J.D. (2004). Evidence for nucleosome depletion at active regulatory regions genome-wide. Nat. Genet.

Lindstrom, D.L., and Gottschling, D.E. (2009). The Mother Enrichment Program: A Genetic System for Facile Replicative Life Span Analysis in Saccharomyces cerevisiae. Genetics *183*, 413–422.

Madamba, E.V., Berthet, E.B., and Francis, N.J. (2017). Inheritance of Histones H3 and H4 during DNA Replication In Vitro. Cell Rep.

McNally, F.J., and Rine, J. (1991). A synthetic silencer mediates SIR-dependent functions in Saccharomyces cerevisiae. Mol. Cell. Biol. *11*, 5648–5659.

Miller, A.M., and Nasmyth, K.A. (1984). Role of DNA replication in the repression of silent mating type loci in yeast. Nature *312*, 247–251.

Moazed, D., Kistler, A., Axelrod, A., Rine, J., and Johnson, A.D. (1997). Silent information regulator protein complexes in Saccharomyces cerevisiae: a SIR2/SIR4 complex and evidence for a regulatory domain in SIR4 that inhibits its interaction with SIR3. Proc. Natl. Acad. Sci.

Nasmyth, K.A., and Tatchell, K. (1980). The structure of transposable yeast mating type loci. Cell.

Park, E.C., and Szostak, J.W. (2015). Point mutations in the yeast histone H4 gene prevent silencing of the silent mating type locus HML. Mol. Cell. Biol.

Park, P.U., McVey, M., and Guarente, L. (2002). Separation of mother and daughter cells. Methods Enzymol. *351*, 468–477.

Petryk, N., Dalby, M., Wenger, A., Stromme, C.B., Strandsby, A., Andersson, R., and Groth, A. (2018). MCM2 promotes symmetric inheritance of modified histones during DNA replication. Science (80-.). *361*, 1389–1392.

Pillus, L., and Rine, J. (1989). Epigenetic inheritance of transcriptional states in S. cerevisiae. Cell.

Prior, C.P., Cantor, C.R., Johnson, E.M., and Allfrey, V.G. (1980). Incorporation of exogenous pyrene-labeled histone into Physarum chromatin: a system for studying changes in nucleosomes assembled in vivo. Cell.

Radman-Livaja, M., Verzijlbergen, K.F., Weiner, A., van Welsem, T., Friedman, N., Rando, O.J., and van Leeuwen, F. (2011). Patterns and Mechanisms of Ancestral Histone Protein Inheritance in Budding Yeast. PLoS Biol. *9*, e1001075.

Ragunathan, K., Jih, G., and Moazed, D. (2014). Epigenetic inheritance uncoupled from sequence-specific recruitment. Science (80-.).

Ragunathan, K., Jih, G., and Moazed, D. (2015). Epigenetic inheritance uncoupled from sequence-specific recruitment. Science (80-.).

Riley, D., and Weintraub, H. (1979). Conservative segregation of parental histones during replication in the presence of cycloheximide. Proc. Natl. Acad. Sci. U. S. A. *76*, 328–332.

Rine, J., and Herskowitz, I. (1987). Four genes responsible for a position effect on expression from HML and HMR in Saccharomyces cerevisiae. Genetics.

Rine, J., Strathern, J.N., Hicks, J.B., and Herskowitz, I. (1979). A suppressor of mating-type locus mutations in Saccharomyces cerevisiae: Evidence for and identification of cryptic mating-type loci. Genetics.

Rusché, L.N., Kirchmaier, A.L., and Rine, J. (2002). Ordered nucleation and spreading of silenced chromatin in Saccharomyces cerevisiae. Mol. Biol. Cell *13*, 2207–2222.

Schlissel, G., Krzyzanowski, M.K., Caudron, F., Barral, Y., and Rine, J. (2017). Aggregation of the Whi3 protein, not loss of heterochromatin, causes sterility in old yeast cells. Science (80-.). *355*.

Schulz, L.L., and Tyler, J.K. (2006). The histone chaperone ASF1 localizes to active DNA replication forks to mediate efficient DNA replication. FASEB J.

Schwabish, M.A., and Struhl, K. (2006). Asf1 Mediates Histone Eviction and Deposition during Elongation by RNA Polymerase II. Mol. Cell.

Seidman, M.M., Levine, A.J., and Weintraub, H. (1979). The asymmetric segregation of parental nucleosomes during chrosome replication. Cell *18*, 439–449.

Sen, P., Dang, W., Donahue, G., Dai, J., Dorsey, J., Cao, X., Liu, W., Cao, K., Perry, R., Lee, J.Y., et al. (2015). H3K36 methylation promotes longevity by enhancing transcriptional fidelity. Genes Dev. *29*, 1362–1376.

Sinclair, D.A., and Guarente, L. (1997). Extrachromosomal rDNA circles--a cause of aging in yeast. Cell *91*, 1033–1042.

Skene, P.J., and Henikoff, S. (2017). An efficient targeted nuclease strategy for high-resolution mapping of DNA binding sites. Elife.

Smeal, T., Claus, J., Kennedy, B., Cole, F., and Guarente, L. (1996). Loss of transcriptional silencing causes sterility in old mother cells of S. cerevisiae. Cell *84*, 633–642.

Smith, S., and Stillman, B. (1991). Stepwise assembly of chromatin during DNA replication in vitro. EMBO J. *10*, 971–980.

Smith, J.S., Brachmann, C.B., Celic, I., Kenna, M.A., Muhammad, S., Starai, V.J., Avalos, J.L., Escalante-Semerena, J.C., Grubmeyer, C., Wolberger, C., et al. (2002). A phylogenetically conserved NAD+-dependent protein deacetylase activity in the Sir2 protein family. Proc. Natl. Acad. Sci.

Steakley, D.L., and Rine, J. (2015). On the Mechanism of Gene Silencing in Saccharomyces cerevisiae . G3: Genes|Genomes|Genetics.

Stebbins, J., Pillus, L., Landry, J., Sternglanz, R., Heller, R.C., Sutton, A., and Tafrov, S.T. (2002). The silencing protein SIR2 and its homologs are NAD-dependent protein

deacetylases. Proc. Natl. Acad. Sci.
Strathern, J.N., Klar, A.J.S., Hicks, J.B., Abraham, J.A., Ivy, J.M., Nasmyth, K.A., and McGill, C. (1982). Homothallic switching of yeast mating type cassettes is initiated by a double-stranded cut in the MAT locus. Cell.
Studitsky, V.M., Clark, D.J., and Felsenfeld, G. (1994). A histone octamer can step around a transcribing polymerase without leaving the template. Cell.
Studitsky, V.M., Kassavetis, G.A., Geiduschek, E.P., and Felsenfeldt, G. (1997). Mechanism of transcription through the nucleosome by eukaryotic RNA polymerase. Science (80-.).
Sugasawa, K., Ishimi, Y., Eki, T., Hurwitz, J., Kikuchi, A., and Hanaoka, F. (1992). Nonconservative segregation of parental nucleosomes during simian virus 40 chromosome replication in vitro. Proc. Natl. Acad. Sci. U. S. A. *89*, 1055–1059.
Thompson, J.S., Ling, X., and Grunstein, M. (1994). Histone H3 amino terminus is required for telomeric and silent mating locus repression in yeast. Nature.
Thurtle, D.M., and Rine, J. (2014). The molecular topography of silenced chromatin in Saccharomyces cerevisiae. Genes Dev. *28*, 245–258.
Tran, V., Lim, C., Xie, J., and Chen, X. (2012). Asymmetric division of Drosophila male germline stem cell shows asymmetric histone distribution. Science (80-.). *338*, 679–682.
Tsai, S.Q., Wyvekens, N., Khayter, C., Foden, J.A., Thapar, V., Reyon, D., Goodwin, M.J., Aryee, M.J., and Joung, J.K. (2014). Dimeric CRISPR RNA-guided FokI nucleases for highly specific genome editing. Nat. Biotechnol. *32*, 569–576.
Vestner, B., Waldmann, T., and Gruss, C. (2000). Histone octamer dissociation is not required for in vitro replication of simian virus 40 minichromosomes. J. Biol. Chem.
Wang, X., Bryant, G., Zhao, A., and Ptashne, M. (2015). Nucleosome avidities and transcriptional silencing in yeast. Curr. Biol.
Wooten, M., Snedeker, J., Nizami, Z., Yang, X., Ranja, R., Urban, E., Kim, J.M., Gall, J., Xiao, J., and Chen, X. (2019). Unidirectional fork movement coupled with strand-specific histone incorporation ensures asymmetric histone inheritance. BioRxiv 242768.
Woringer, M., and Darzacq, X. (2018). Protein motion in the nucleus: from anomalous diffusion to weak interactions. Biochem. Soc. Trans. *46*, 945 LP – 956.
Wu, C.-S., Chen, Y.-F., and Gartenberg, M.R. (2011). Targeted sister chromatid cohesion by Sir2. PLoS Genet. *7*, e1002000.
Wyrick, J.J., and Parra, M.A. (2009). The role of histone H2A and H2B post-translational modifications in transcription: A genomic perspective. Biochim. Biophys. Acta - Gene Regul. Mech.
Xu, M., Long, C., Chen, X., Huang, C., Chen, S., and Zhu, B. (2010). Partitioning of histone H3-H4 tetramers during DNA replicationdependent chromatin assembly. Science (80-.).
Yu, C., Gan, H., Serra-Cardona, A., Zhang, L., Gan, S., Sharma, S., Johansson, E., Chabes, A., Xu, R.M., and Zhang, Z. (2018). A mechanism for preventing asymmetric histone segregation onto replicating DNA strands. Science (80-.).
Zhang, L., Serra-Cardona, A., Zhou, H., Wang, M., Yang, N., Zhang, Z., and Xu, R.M. (2018). Multisite Substrate Recognition in Asf1-Dependent Acetylation of Histone H3 K56 by Rtt109. Cell.