UC Berkeley
UC Berkeley Electronic Theses and Dissertations

Title

The characterization of LUTI mRNAs in budding yeast meiosis

Permalink

https://escholarship.org/uc/item/8pm09254

Author

Tresenrider, Amy Moreda

Publication Date

2019

Peer reviewed|Thesis/dissertation

The characterization of LUTI mRNAs in budding yeast meiosis

by

Amy M Tresenrider

A dissertation submitted in partial satisfaction of the

requirements for the degree of

Doctor of Philosophy

in

Molecular and Cell Biology

in the

Graduate Division

of the

University of California, Berkeley

Committee in charge:

Assistant Professor Elçin Ünal, Chair
Professor Jasper Rine
Assistant Professor Nicholas Ingolia
Professor N. Louise Glass

Summer 2019

The characterization of LUTI mRNAs in budding yeast meiosis

Abstract

The characterization of LUTI mRNAs in budding yeast meiosis

by

Amy M Tresenrider

Doctor of Philosophy in Molecular and Cell Biology

University of California, Berkeley

Assistant Professor Elçin Ünal, Chair

Differentiation programs such as meiosis depend on extensive gene regulation to mediate cellular morphogenesis. The molecular events associated with gene regulation in budding yeast meiosis are particularly well-documented at both the transcriptional and the translational levels. Through the integrated study of these two steps in gene regulation, a novel mode of gene repression was identified whereby an mRNA does not produce the protein coded within it, but rather plays a purely regulatory role (Chapter 2, Chapter 3).

We first discovered this by investigating the regulation of the kinetochore gene *NDC80* (Chapter 2). The program of meiosis requires transient removal, and thus inactivation, of the outer kinetochore, the complex that connects microtubules to chromosomes. We found that in budding yeast, this occurs by reducing the abundance of a limiting subunit, Ndc80. Central to this mechanism is the developmentally controlled transcription of an alternate *NDC80* mRNA isoform (Long Undecoded Transcript Isoform, LUTI), which itself cannot produce protein due to regulatory upstream ORFs (uORFS) in its extended 5'-leader. Instead, transcription of the LUTI represses the canonical *NDC80* mRNA expression in *cis* through the deposition of Set1-dependent histone H3K4 dimethylation, Set2-dependent H3K36 trimethylation, and increased nucleosome occupancy to establish a repressive chromatin state in the downstream canonical *NDC80* promoter and thereby inhibit Ndc80 protein synthesis.

We further investigated whether other genes may be regulated in a similar fashion (Chapter 3). We identified 74 genes that have 5'-extended leaders in meiotic prophase, all but two of which have at least one ATG uORF, and the vast majority of which are translationally repressed. While the translational down-regulation of these extended transcripts is near universal, the transcriptional repression of the canonical isoform is more variable. The features important for transcriptional repression at the *NDC80* locus are associated with down regulation of the canonical transcript in the newly identified LUTIs. More specifically, H3K36 trimethylation and a decrease in strong positioning of the +1 nucleosome correlate with a

greater decrease of the canonical transcript upon LUTI induction. Higher LUTI abundance also correlates with greater repression.

Together we have delved deeply into the regulation of how a novel regulatory mRNA affects the expression of a single gene, and we provide evidence that the same mechanics of gene repression occur at additional loci throughout the genome during meiotic prophase.

To Kyle and Mochi

Contents

List of Figures

List of Tables

Acknowledgments

First and foremost, I need to thank my advisor Elçin Ünal. A careful, motivated, and inspiring scientist, she has helped to shape me into the scientist that I am today. Her energy cultivates an open and engaging atmosphere of support and inquiry. It has allowed me to grow leaps and bounds from where I started. Her respect and unwavering investment in my growth built up my scientific confidence year over year as I gained more and more independence. I could not imagine getting through graduate school without Elçin as my advisor. The lens through which I understand, appreciate, and participate in science is forever changed for the better because of her mentorship.

Second, I need to thank my collaborator, classmate, bay-mate, and close friend Jingxun Chen. Her experimental insights and careful scientific thought have pushed me to think deeper and harder about my own work. I've become a better scientist having her as my peer. Thanks for putting up with me, laughing with me, and for helping me to take care of my lab plants. I hope to continue crossing paths with her throughout our careers.

As part of the joint Brar-Ünal lab, I also need to thank my second mentor Gloria Brar. Gloria was always happy to chat with me whenever I had questions about uORFs and/or ribosome profiling. Her insights during my lab meetings and one-on-one meetings consistently led me down a new and better path.

I also need to thank the rest of the lab both past and present. I especially want to thank Eric Sawyer and Leon Chan for their careful readings of my work over the years and their technical expertise. I want to thank the final two parts of the Brar-Ünal lab tetrad: Kelsey van Dalfsen and Ze Cheng. To Vicki Jorgensen, I thank her for her never-ending energy and northern expertise. It was an immense honor to watch her grow from an undergrad with no laboratory expertise to an independent technician capable of driving her own experiments and projects. I want to thank all participants of our lab's Genomics Club for fostering an environment in which to exchange our ideas, problems, and solutions to genome-wide analyses in a friendly and constructive manner. I also want to thank all those who will continue to investigate LUTIs even after I've left the lab, with a special recognition of my former rotation students: Anthony Harris, Kate Morse, and Amanda Su. Lastly, I need to thank our wonderful lab manager Christiane Brune for insuring the lab ran smoothly and that we had whatever was needed for us to be able to perform our experiments.

I thank my thesis committee members Jasper Rine, Nick Ingolia, and Louise Glass for their insightful comments and input. I thank my collaborators at the Francis Crick Institute: Minghao Chia and Folkert van Werven. It was a pleasure working with you while we fleshed out the *NDC80* story. I additionally thank you for your expertise with Transcript leader sequencing.

I must also thank all of the friends I have made throughout graduate school. Madeleine Jensen and Rebecca Lamothe, I could not have asked for better roommates during the first two years of graduate school. Rebecca, I will always cherish our weekly boba dates.

To my family, I thank them for their continued support of my ambitions. To my parents, your trust and confidence in my abilities helped me through the best and worst times of graduate school. To my sisters, never stop being your goofy selves. I love you all.

Finally, to my husband Kyle, words cannot express how thankful I am for your constant love and support. I can however thank you for sharing your coding expertise with me over the last couple of years as I learned to program. Each and every day you make me a better person. I love you and Mochi so much.

Chapter 1

Introduction

The following chapter contains published material from a publication that I am the first author on (Tresenrider & Ünal 2018). The article is distributed under the terms of the Creative Commons Attribution License (CC BY 4.0), which permits unrestricted use and redistribution provided that the original author and source are credited.

1.1 The many levels of gene regulation

The time and location of gene expression affects how and when organisms differentiate their cells into distinct lineages. Regulation of this process occurs at numerous levels including but not limited to the chromatin context, transcription factor expression, RNA stability, translational efficiency, protein stability, and protein activity. While each of these have an effect on the final level and localization of a protein output, transcription factors are considered to be the dominant drivers of gene regulation throughout development. Some of the most famous examples include the Hox genes and the Yamanaka factors. Hox gene expression patterns in the developing embryo dictate where along the anterior-posterior axis individual body parts will be in the adult organism (Krumlauf 1994). The Yamanaka factors Oct4, Sox2, KLF4, c-Myc are essential for maintaining stem cell pluripotency (Takahashi & Yamanaka 2006). Their expression is sufficient to reprogram differentiated cells into an undifferentiated state.

Studies of these and other transcription factor families have provided invaluable insights into how gene activation can drive developmental programs. However, although important, transcriptional regulation is not the end-all be-all to discovering the inner molecular workings of a cell. With advancing genomics techniques, it has become possible to measure multiple levels of gene expression with matched biological samples. Through this, recent studies have indicated that levels of translation and protein abundance don't always correlate well with transcript abundance (Ingolia *et al.* 2009; Brar *et al.* 2012; Cheng *et al.* 2018; Floor & Doudna 2016). One explanation for this phenomenon will be explored in-depth throughout

this dissertation, using budding yeast meiosis as a model system. As an introduction to those findings, below is a review of the major steps in gene regulation, an introduction to meiosis in budding yeast, and an overview of how investigating gene regulation at multiple levels and over time can lead to the discovery of new biology not possible when studying a single mode in isolation.

Packaging of DNA into chromatin

DNA is the genetic material of the cell. Even in budding yeast, one of the smallest eukaryotic genomes, there are $\sim$ 12 million base pairs (bps) of DNA. Amazingly, the long polymer is packaged into the nucleus of a cell without becoming impossibly tangled. It achieves such a feat in part through the structuring of DNA strands around proteins. The smallest unit of such organization is the nucleosome. Conserved across all eukaryotic cells, each nucleosome is composed of a protein core with 147 bp of DNA wrapped around it (Cutter & Hayes 2015). Most nucleosome cores contain two copies of each of the four major histone proteins: H2A, H2B, H3, and H4. Together they assemble as a highly basic octamer, ideal for interacting with negatively charged DNA. Alone, the organization of DNA in this way compacts the DNA, but it also serves as a platform for further condensation throughout the cell cycle (Hansen *et al.* 2018; Prieto & Maeshima 2019).

In addition to organizing DNA, nucleosomes serve to regulate gene expression. Their regulatory abilities come from 1) their positioning and occupancy on the DNA and 2) the post-translational modifications that they bear. It is well established that although nucleosomes are arrayed like beads on a string across the entirety of a genome, their spacing is variable. In addition to the 147 bps of DNA that wrap around each nucleosome, linker DNA exists to bridge the gap between two successive nucleosomes. Constitutively silenced genomic regions, such as areas of heterochromatin, have more regularly spaced and tightly packed nucleosomes (Henikoff 2000). Euchromatic regions, or regions with active transcription, have more irregularly spaced and dynamic nucleosomes, especially during changes in gene expression.

Generally, regions of active transcription have a pattern of decreased nucleosome occupancy across gene promoters. These nucleosome depleted regions (NDRs) are flanked by well-positioned nucleosomes on either side (a +1 nucleosome directly downstream, a -1 nucleosome directly upstream) (Lee *et al.* 2004; Lee *et al.* 2007; Yuan *et al.* 2005). The presence of such exposed regions of DNA facilitates the binding of transcription factors and general transcriptional machinery which would otherwise be occluded from sites with high nucleosome occupancy (Field *et al.* 2008; Tirosh & Barkai 2008; Cairns 2009; Struhl & Segal 2013). Positioning of nucleosomes around the promoters in such a manner relies on features inherent to the DNA sequence as well as the chromatin-associated proteins (Cairns 2009; Struhl & Segal 2013).

Common sequence features found at promoters are poly(dA:dT) tracts. Budding yeast in particular contain large strings of poly(dA:dT) in gene promoters (Field *et al.* 2008). These tracts are stiff and inflexible making them unfavorable for wrapping around a nucleosome (Nelson *et al.* 1987; Suter *et al.* 2000). This alone can be enough to leave some yeast promoters depleted of nucleosomes when assaying *in vitro* reconstituted chromatin (Zhang *et al.* 2009; Kaplan *et al.* 2009). In cells, those same promoters have larger than average NDRs and well-positioned +1 nucleosomes, which are attributes associated with constitutively high levels of gene expression. Promoters with smaller NDRs and poorly positioned +1 nucleosomes frequently don't have poly(dA:dT) tracts (Field *et al.* 2008). They are enriched in conditionally expressed genes. Instead of relying on unchanging features inherent to the DNA sequence, these conditional promoters must activate only when a gene's expression is needed. They therefore would not want to rely solely on the unchanging sequence of DNA. Rather, promoters of conditionally expressed genes may have transcription factor binding sites covered either entirely or partially by nucleosomes (Field *et al.* 2008; Tirosh & Barkai 2008). The sites may only become accessible upon nucleosome remodeling.

Even most genes that are not conditionally expressed require more than sequence context to fully recapitulate their NDRs and the *in vivo* positioning of +1 and -1 nucleosomes (Zhang *et al.* 2009; Kaplan *et al.* 2009). Only a complex interplay between DNA sequence, chromatin remodelers, and the act of transcription itself results in the proper positing of nucleosomes. Chromatin remodelers are a family of proteins that have the ability to reposition nucleosomes by sliding, ejecting, or changing the nucleosome composition (Cairns 2009; Narlikar *et al.* 2013). The first identified chromatin remodeler Snf2 was discovered in a yeast screen for genes that inhibited expression of the *SUC2* gene, which is required for growth on sucrose (Neigeborn & Carlson 1984). Upon further characterization, cells lacking Snf2 or cells harboring mutants that disrupted the enzyme's ATPase activity exhibited altered nucleosome positioning (Laurent *et al.* 1993; Hirschhorn *et al.* 1992). Sequence comparison revealed that Snf2 belonged to a larger group of highly conserved Snf2-like ATP-dependent DNA binding proteins (Laurent *et al.* 1992). Much effort has been made to understand the function of this family of proteins and to explore the unique activities of many different remodelers. Using genome-wide techniques, the patterns of remodeler binding with respect to genes and how they affect nucleosome position has come to light (Yen *et al.* 2012; Gkikopoulos *et al.* 2011). For example, Chd1 and Isw1 do not appear to affect position of the +1 nucleosome but they are required for proper spacing of nucleosomes over gene bodies (Gkikopoulos *et al.* 2011). To the contrary, Isw2 localizes almost exclusively to the +1 nucleosome which moves slightly more 3' in *isw1Δ* cells (Yen *et al.* 2012). Other remodelers such as *INO80* or *SNF2* help to maintain NDRs. When they are deleted, +1 nucleosomes on average move to a more 5' position which therefore shrinks the NDR (Yen *et al.* 2012). More work remains to uncover how such specificity is achieved, and what the full effects are on gene expression but it's clear that remodeler-mediated nucleosome movements are important since mutations in chromatin remodelers, especially SWI/SNF components, are frequently observed in cancer (Kadoch *et al.* 2013; Garraway & Lander 2013).

A growing body of work has begun to appreciate how the binding of transcription factors can act as nucleosome remodeling events. Previously, transcription factor binding was correlated with DNA accessibility: nucleosome-bound DNA was thought to prevent binding while nucleosome-free DNA could easily be bound by non-histone proteins such as transcription factors. This mantra is true for many proteins but not all. Some transcription factors can robustly interact with DNA that is wrapped around histones. These transcription factors, referred to as "pioneer" factors, are commonly drivers of cell differentiation. Presumably because in order initiate activation of a new cell state, the transcription factor may have to induce transcription from repressed promoters with nucleosome occupancy at transcription factor binding sites. The first characterized pioneer factor was FoxA, a driver of liver cell differentiation (Cirillo *et al.* 2002). It was shown that *in vitro* FoxA can bind to nucleosomal DNA at the albumin promoter and induce opening of compacted nucleosomal DNA (Cirillo *et al.* 1998; Cirillo *et al.* 2002). Since then the list of pioneer factors has grown to include other transcription factors involved in cell fate determination including a subset of the world-renowned "Yamanaka factors" (Takahashi & Yamanaka 2006; Soufi *et al.* 2015). Oct3/4, Sox2, and Klf4 are all able to bind nucleosomes (Soufi *et al.* 2015). This makes sense because together these factors are sufficient to reactivate regions of DNA that had been silenced in differentiated fibroblasts. While *in vitro* studies indicated that FoxA can open chromatin structure in the absence of ATP-dependent remodelers, it remains unclear whether most other pioneer transcription factors have remodeling activity on their own or whether they help to recruit remodelers.

Studies in yeast have recently systematically identified transcription factors that can associate with nucleosome bound DNA to alter nucleosome positioning, similar to pioneer factors (Yan *et al.* 2018). Of these characterized nucleosome displacing factors (NDFs), those with the highest nucleosome displacing activity also have the highest DNA binding selectivity and expression levels. Some weaker NDFs can compensate if multiple binding sites are found near each other. Interestingly, weaker NDFs, such as Rfx1 or Cbf1, appear to only bind DNA during replication before stable histones assemble over binding sites (Yan *et al.* 2018). Stronger NDFs do not show such cell-cycle dependence, demonstrating the multiple pathways by which NDFs alter nucleosome positioning. Continued study of this class of transcription factors will help us to understand how DNA sequence, chromatin remodelers, and transcription factors all work together to make DNA accessible for transcription.

The histone code

In addition to their positioning being crucial for proper gene expression, the histone proteins in nucleosomes can be post-translationally modified. All of the four core histones have a globular domain (Cutter & Hayes 2015). These domains assemble with each other to form a puck-like shape around which the DNA wraps (Cutter & Hayes 2015). Protruding from the core are flexible histone "tails." These tails contain numerous residues that can be covalently modified. Modifications of histones most commonly include methylation

and acetylation, but also ubiquitylation, phosphorylation, ADP-ribosylation, sumoylation, and deimindation to name a few (Kouzarides 2007). Together, the writing, reading, and functional outcomes associated with these modifications was coined the histone code in 2000 (Strahl & Allis 2000). Almost two decades later, much of the histone code has been filled in, but new modifications are still being uncovered, and our understanding of even some of the earliest identified chromatin modifications continues to deepen (Kebede *et al.* 2015).

Of the two most studied marks, acetylation was first to be linked to transcription. For decades it had been known that histones get acetylated, and that acetylated histones are preferentially associated with regions undergoing active transcription (Struhl 1998). It wasn't until Brownell & Allis (1995) identified and purified the first histone acetyl transferase (HATs) that the true link between transcription and acetylation began to unfold. After their discovery in *Tetrahymena* a slew of other previously identified transcriptionally associated proteins were classified as HATs (Brownell *et al.* 1996; Mizzen *et al.* 1996; Bannister & Kouzarides 1996; Ogryzko *et al.* 1996). Not long after the discovery of the first HATs, the first histone deacetylases (HDACs) were found, in humans (Taunton *et al.* 1996), and then in yeast (Rundlett *et al.* 1996), thus laying the groundwork for studying how the balance between deacetylation and acetylation of histones interplays to regulate gene expression.

Second to acetylation, an explosion in the study of histone methylation quickly followed. Originally identified as proteins associated with heterochromatin, the first confirmed histone methyltransferases (HMTs) were homologs of the *Drosophila Su(var)3-9* (Rea *et al.* 2000). Rea *et al.* (2000) demonstrated that mouse Suv39h2, and human SUV39H1 had HMTase activity specific for H3 at lysine 9 (H3K9). Since then, H3K9me3 has been linked to the formation of heterochromatin in organisms as divergent as fission yeast and humans (Nakayama *et al.* 2001; Lachner *et al.* 2002). It is interesting to note that the chromatin modification alone does not have a strong affect on gene expression. Rather, the modification serves as a base to recruit other components of a pathway. In the case of H3K9me3, the modification helps to recruit heterochromatin protein 1 (HP1) to establish and maintain a repressed heterochromatic state (Bannister *et al.* 2001; Lachner *et al.* 2002).

Methylation is not only found in regions of heterochoromatin though. Most pertinent to this work are the co-transcriptional methylation events found at H3 lysine 4 (H3K4) and H3 lysine 36 (H3K36). The discussion below will focus on the genes responsible for laying these modifications in budding yeast, but it is important to note that the methylation of these lysine residues regulates gene expression in higher eukaryotes as well. In budding yeast, all H3K4 methylation is catalyzed by a single HMT: Set1. Set1, a member of the Set1/COMPASS complex, was identified as the first H3K4 HMT in any organism by Roguev *et al.* (2001). Set1 had previously been associated with ensuring proper telomere length, silencing at the mating type loci, and progression through sporulation (Nislow *et al.* 1997). It has also been associated with the regulation of meiotic double strand break (DSB) formation (Borde *et al.* 2009; Sommermeyer *et al.* 2013; Acquaviva *et al.* 2013). Its role in laying mono-,

di-, and tri- methylation is in part why it has so many pleiotropic effects in the cell (Dehé & Géli 2006). However, focusing only on regions of active transcription, Set1 is recruited by phosphorylated Ser5 (S5-P) in the C-terminal domain (CTD) of RNA polymerase II (RNA pol II) where at and just downstream of promoters it deposits H3K4me3 (Ng *et al.* 2003; Krogan *et al.* 2003). It is still unknown exactly what role this modification plays or how it functions, but it co-localizes with high levels of acetylation over gene promoters and is highly conserved across evolution (Eissenberg & Shilatifard 2010).

The second Set1-dependent chromatin modification, H3K4me2, is found just downstream of promoters and near the 5' end of actively transcribed genes (Ng *et al.* 2003; Krogan *et al.* 2003; Pokholok *et al.* 2005). Dimethylation of H3K4 recruits the histone deacetylase complex Set3C (Pijnappel *et al.* 2001; Kim & Buratowski 2009). The namesake member of the complex Set3 is an intriguing protein containing both a SET domain and a PHD domain (Pijnappel *et al.* 2001). PHD domains most commonly bind H3K4 methylated histones, while SET domains are traditionally associated with methyltransferase activity (Rea *et al.* 2000; Shi *et al.* 2006). It is through the PHD domain that Set3 recruits the Set3C complex to chromatin, but intriguingly, to date, no methyltransferase activity has been associated with Set3C. Instead, the presence of Set3C is most associated with temporal gene repression at sites of overlapping transcription upon shifts in carbon source (Pijnappel *et al.* 2001; Kim *et al.* 2012; Kim *et al.* 2017).

While the Set1-H3K4me2-Set3C pathway can function independently of other chromatin modifications to temporally repress gene expression, it can also work in conjunction with a second co-transcriptional epigenetic pathway (van Werven *et al.* 2012). The second pathway, that of Set2-H3K36me3-Rpd3S is made up of similar classes of proteins and complexes. Set2 is a HMT with specificity for K36 of histone H3 (Strahl *et al.* 2002). A pattern of increased H3K36me3 across the coding domains of transcribed genes relies on the recruitment of Set2 to phosphorylated Ser2 in the CTD of RNA pol II where it associates with actively elongating polymerase (Li *et al.* 2003; Pokholok *et al.* 2005; Kizer *et al.* 2005). H3K36me3 then helps to recruit Rpd3S, an HDAC, to deacetylate gene bodies (Carrozza *et al.* 2005; Keogh *et al.* 2005). Originally identified as a method to repress cryptic initiation, this pathway is now also associated with gene repression at sites of overlapping transcription from both intergenic sense transcripts and antisense transcripts (Venkatesh *et al.* 2016; Carrozza *et al.* 2005; Kim *et al.* 2017; van Werven *et al.* 2012). Active investigation into the biological role of such non-coding transcription and transcripts continues. The biological function of non-coding transcripts with regard to meiotic entry in budding yeast meiosis will be discussed later in greater depth.

The biological function of uORFs

After an mRNA transcript is produced, it is trafficked out of the nucleus to be translated into protein by the ribosome. An mRNA's first contact with the ribosome occurs through

the transcript's 5' m7G cap. The cap is initially bound by the heterotrimeric eIF4F complex which recruits the 43S preinitiation complex. Upon cap binding, the preinitiation complex begins to scan in a 5' -> 3' direction along the mRNA searching for an initiation site, most commonly an AUG. Once found, the initiation factors are released, the large subunit (60S) of the ribosome is recruited to the mRNA, and translation can begin. The ribosome will continue translating the mRNA until it reaches one of three stop codons (UAG, UAA, UGA). These foundational aspects of translation have been well-established for decades and are found in countless textbooks, but it is critical that we do not forget that there is still much to learn about translation and its regulation.

With the advent of methods such as ribosome profiling, which allows the tracking of both where and how much translation is occurring across the genome, we now know that significant translation occurs outside of traditionally expected regions (Ingolia *et al.* 2009; Ingolia *et al.* 2014; Brar *et al.* 2012). For example, the signal to begin translation occurs not uncommonly at near-cognate start sites, sites in which the codon matches two of the three bases in AUG (Ingolia *et al.* 2011; Brar *et al.* 2012). One of the most abundant types of translation outside of annotated coding domain sequences (CDSs) is the translation of upstream open reading frames or uORFs in what was termed the 5'UTR, but will be referred to as the 5'-leader in the remainder of this work (Ingolia *et al.* 2009; Ingolia *et al.* 2011; Brar *et al.* 2012). This uORF translation is pervasive in budding yeast meiosis, but even in larger eukaryotic genomes such as human and drosophila about 50 % of genes have evidence of uORF translation (Calvo *et al.* 2009; Chew *et al.* 2016; Johnstone *et al.* 2016).

This has lead to increased intrigue into what role uORFs may play in a cell. Initially, there was much interest as to what possible functions the short peptides may play. To date, the consensus in the field is that in the vast majority of cases, the small peptides produced from uORFs are not functional; rather, the act of their translation plays a regulatory role by decreasing translation at downstream AUGs. The theory is that translation of a uORF engages the ribosome in a translation cycle. Upon reaching the uORF's stop codon, the ribosome may stay on the RNA, or it may fall off. Even if it does remain, it may not be able to reassemble all the machinery required to initiate another round of translation. Thus, it may never initiate translation at a downstream ORF.

The best know example of uORF-mediated translational repression comes from the yeast gene *GCN4* whose transcript is deferentially translated during amino acid starvation compared to growth in rich media (Mueller & Hinnebusch 1986). *GCN4* codes for a transcription factor that turns on genes required for amino acid synthesis. Under nutrient-rich conditions, amino acids are freely available to yeast cells, so they have no need to synthesize them. Therefore, it is advantageous to repress *GCN4*. These demands are reversed once amino acids become sparse, increasing the demand for *GCN4* expression. Such regulation in this case relies on 4 uORFs in the 5'-leader of *GCN4*. The first and fourth have been investigated in the greatest depth because even if the 2nd and 3rd are mutated, the pattern of

GCN4 expression is essentially identical to wild-type (Mueller & Hinnebusch 1986). The first uORF and the fourth uORFs are robustly translated in nutrient rich conditions. It is thought that scanning occurs after the uORF1 and that translational reinitiation is possible at the fourth uORF; however, reinitiation cannot occur at the gene's ORF. Once amino acids are no longer available externally, the first uORF is still translated, and scanning still occurs after uORF1 translation, but due to low levels of the ternary complex, which contains the charged Met-tRNA required for initiating another round of translation, the ribosome is not prepared to reinitiate translation as quickly (Mueller & Hinnebusch 1986; Hinnebusch 1997). uORF4 is not translated. By the time the *GCN4* ORF is reached though, the ribosome has all elements required to reinitiate translation (Hinnebusch 1997), leading to increased levels of Gcn4 protein.

Instances of known uORF-mediated repression are not limited to budding yeast. In mammalian cells, ATF4 regulation follows a very similar pattern to that observed for *GCN4* (Harding *et al.* 2000). It was known that upon stresses such as the unfolded protein response or amino acid starvation, a general shutdown of translation occurs. Harding *et al.* (2000) observed just the opposite when investigating the transcription factor ATF4: its translation increased under stress. They showed that this was due to the presence of 2 uORFs in mouse or 3 uORFs in humans. In both cases, the last uORF overlapped but was out of frame with the coding domain such that translation of it led to severe repression of ATF4 CDS translation. They proposed that, similar to *GCN4* in yeast, the decrease in translation initiation under stressful conditions allowed the ribosome to scan past the uORFs and initiate further downstream at the CDS (Vattem & Wek 2004). Two more examples of attenuated uORF-mediated repression upon stress come from TBF1 in *Arabidopsis* and *zip-2* in *C. elegans* (Dunbar *et al.* 2012; Pajerowska-Mukhtar *et al.* 2012). In both cases, the increase in translation of the protein-coding ORF occurred after infection, providing even more biological events during which uORF represssion plays a highly functional biological purpose. All of the above genes just so happen to be transcription factors. At first surprising, it is not unreasonable that cells would want to regulate transcription factors at multiple points during gene expression since they have the power to initiate extensive changes in cell state. Cells may want to couple such drastic changes in genes expression to varying inputs in order to only initiate new programs of gene expression when truly necessary. Still, the fact that all four examples, from organisms as disparate as fungi, metazoans, and plants regulate the same type of proteins is intriguing.

With examples of uORFs in varied organisms, and the presence of uORF translation in the 5'-leaders of so many genes, a search for the defining features of uORF-mediated repression commenced. Reports repeatedly observed that the presence of a translated ATG uORF was associated with translational repression, and that increasing numbers of uORFs in a leader correlated with greater repression (Calvo *et al.* 2009; Chew *et al.* 2016; Johnstone *et al.* 2016). Consistently, greater distances between the termination of the uORF and the start of the CDS were associated with poor repression, possibly because it allowed more

time for the recruitment of a newly charged ternary complex. Ultimately though, we still do not fully understand the rules of uORF-mediated repression. It remains impossible to know with certainty whether a transcript will be unaffected, partially repressed, or translationally shutdown in a uORF-dependent manner just by observing the sequence. This problem is most likely difficult to address because each uORF is unique in it is repressive ability and on top of that most uORF-containing genes have multiple uORFs. This hurdle has been recognized, and attempts have been made to analyze genes with just a single uORF in order to extract out repressive trends (Chew *et al.* 2016). However, this model would also miss cooperative or opposing interactions between uORFs which may occur frequently as is the case with *GCN4*. It will be interesting to follow this problem going forward to see if indeed it will be possible to better understand the uORF code.

1.2   Regulation of the meiotic differentiation program

With the knowledge that changes in gene expression help to determine cell fate during differentiation, these processes are common subjects when studying gene regulation. One such program, gametogenesis (also know as meiotic differentiation), is the focus of the work described in the following chapters. Gametogenesis is the process by which sex cells, or gametes, form, and it is essential for the generation of progeny by sexual reproduction. Conserved across vast walks of life, sexual reproduction is a driving force through which genetic diversity can be increased. It allows the mixing of the genetic material of individuals to produce new combinations of alleles, some of which may benefit an organism's or species' chance of survival. Not limited to metazoans or plants or even to multi-cellular organisms, this highly conserved developmental program occurs in various single-celled organisms including budding and fission yeasts as well as cilliates like *Tetrahymena* and more recently discovered the closest living ancestor of animals, choanoflagellates (Woznica *et al.* 2017).

One key aspect of gametogenesis is the reductional division of genetic material through the meiotic program. Meiosis initiates with the replication of DNA and generation of programmed DNA double strand breaks. After completion of S-phase, cells enter an extended prophase, during which they condense their chromosomes and line up homologous chromosomes next to each other. Upon repair of the DNA breaks, crossovers form, whereby the DNA from one homologous chromosome is repaired by exchanging DNA strands with the other homologous chromosome. This event in which chromosomes mix together is called recombination. It along with the independent assortment of chromosomes accounts for the majority of genetic mixing that is the hallmark of sexual reproduction. After recombination is complete, the chromosomes are divided in two consecutive nuclear divisions called meiosis I and meiosis II: in meiosis I homologous chromosomes separate and in meiosis II sister chromatids are pulled apart. The resulting cell(s) thus have half the genetic material of the progenitor cell, and the chromosomes of the progeny are recombined mixtures of the two homologous chromosomes found in the progenitor.

In addition to being a driver of genetic diversity and evolution, the study of meiosis is valuable for better understanding human health. Chromosomal abnormalities that arise due to errors in meiosis occur in at least 10% of human pregnancies, accounting for 35% of miscarriages and 4% of still births (Hassold & Hunt 2001; Nagaoka *et al.* 2012). The most common birth defect caused by the mis-segregation of autosomes during meiosis is Down Syndrome, one of the most prevalent conditions responsible for impaired development in humans (Hassold & Hunt 2001; Nagaoka *et al.* 2012; Herbert *et al.* 2015). Incidence of Down Syndrome and other aneuploidies are currently on the rise as women give birth later and later in life (Herbert *et al.* 2015). The reasons for which maternal age is highly correlated with an increased likelihood of aneuploidy are still under active investigation (Hassold & Hunt 2001; Nagaoka *et al.* 2012; Herbert *et al.* 2015).

With how essential meiosis is to the generation of new life, it is not surprising that there is interest in understanding the regulatory events underpinning the process. In mammals, it is known that retinoic acid plays a role in signaling cells to enter meiosis in both male and female gametes through its role in the induction of STRA8 expression (Anderson *et al.* 2008). STRA8 expression is known to be necessary for oogenensis and spermatogenesis in mice (Anderson *et al.* 2008). It has no homolgy to any other known protein, but a Gal4 DNA binding domain fusion with STRA8 demonstrated some ability to stimulate transcription (Tedesco *et al.* 2009). In a more native context, it was recently determined to be a *bone fide* transcription factor (Kojima *et al.* 2019). Other meiosis-specific genes are known in mammals, but there are no documented meiosis-specific transcription factors. Further studies in mice and humans are essential to further our grasp of meiotic regulation but such studies can be difficult. Only a small number of cells will undergo the process at one time. Compounded upon that, in humans and mice, female meiosis starts during embryo development but does not finish until after fertilization making meiosis in a single cell take months to years, for mice, and decades, for humans.

Due to such difficulties, alternative model organisms are utilized to garner insights into this conserved developmental process. Budding yeast is one such model of choice. It is especially amenable to the study of meiosis because large numbers of cells can be made to go through the process synchronously and to completion within 24 hours. This is in large part due to decades of research with regard to the regulation of meiotic entry in budding yeast which hinges on expression of the meiosis-specific transcription factor Ime1. Ime1, was originally discovered in a cDNA screen for genes that when expressed can induce sporulation independent of any mating type control (Kassir *et al.* 1988). In order to act as the gatekeeper of meiosis, it has one of the longest "promoters" in budding yeast at over 2 kb. Through its promoter, it can integrate nutritional and genetic signals in order to assess if a cell should enter the meiotic program. The four conditions that must be met in order for meiosis to progress are: 1) the cells must be MATa/MATα diploid, 2) the cells must be able to generate ATP through cellular respiration (this is not required for mitotic growth in budding yeast), 3) the cells must be starved of nitrogen, and 4) they must not have access to a fermentable

carbon source (van Werven & Amon 2011). Further work demonstrated that the nutrient control of Ime1 expression relies on the downregulation of both PKA and TORC1 pathways for glucose and nitrogen, respectively (Weidberg *et al.* 2016).

By studying the ability of cells to induce Ime1 expression in diploid cells specifically, an unexpected regulatory mechanism was uncovered which in part explains why the gene's promoter is so long. Encompassing almost the entirety of the *IME1* promoter is a long non-coding (lnc) lncRNA *IRT1* (van Werven *et al.* 2012). Expressed from the same strand as *IME1*, these two transcripts are mutually exclusive. If *IRT1* is on, it shuts down expression of *IME1* through an increase in H3K4me2 and H3K36me3 and change in nucleosome position (van Werven *et al.* 2012). Cleverly, cells have enlisted a transcription factor to regulate the repressive *IRT1* transcript depending on mating type of the cell. This gene *RME1* is expressed in haploid cells, but in diploid cells, it is repressed by the a1/α2 complex, which is only able to form in cells with both MATa and MATα loci. After this discovery, the presence of a second lncRNA upstream of *IRT1* was found. The second transcript *IRT2* is expressed on the same strand as *IRT1* and *IME1*, but it is only found in diploid cells upon starvation (Moretto *et al.* 2018). Instead of preventing *IME1* expression, *IRT2* helps to induce *IME1* by repressing *IRT1* (Moretto *et al.* 2018). Working in a feed-forward loop, Ime1 plays a role in robustly inducing *IRT2*. This is not the only site in which ncRNAs play a role in regulating meiotic entry. An antisense RNA is transcribed across the meiotic RNA methyltransferase gene *IME4* in haploid cells (Hongay *et al.* 2006). Using the same a1/α2 complex used to shut off *RME1* in diploid cells, the antisense transcript is repressed in diploid cells which then allows for transcription of *IME4* to proceed (Hongay *et al.* 2006). It is apparent that ncRNAs are central to proper entry into the meiotic program in budding yeast and that changes to the chromatin landscape, be it chromatin modifications and or nucleosome remodeling at promoters of sites with overlapping transcription are key. This idea will be further explored in the chapters to come.

On top of having a non-canonical promoter, Ime1's function as a transcription factor relies on a non-canonical interaction with a secondary transcriptional regulator, Ume6. Early meiotic genes that turn on in response to Ime1 expression frequently have a URS1 site in their promoters (5'-TCGGCGGCT-3'), but there is no evidence that Ime1 can directly bind to that sequence (Williams *et al.* 2002; Strich *et al.* 1994). In fact, Ime1 has no known DNA binding activity (Smith *et al.* 1993). Rather the mitotic repressor Ume6 binds to those URS1 sites (Williams *et al.* 2002; Strich *et al.* 1994). In the absence of Ime1, Ume6 recruits the HDAC Rpd3 through Sin3 causing strong transcriptional repression to sites where it is bound (Kadosh & Struhl 1997). Upon Ime1 expression, the function of Ume6 switches. Ime1 is recruited to and interacts with Ume6 in such a manner to both prevent the Sin3/Rpd3 dependent repression and provide an activating domain to induce transcription (Washburn & Esposito 2001; Rubin-Bejerano *et al.* 1996; Bowdish *et al.* 1995). Being that Ume6 is bound to many meiosis-specific genes, this serves as an efficient way to couple strong repression of a set of genes in one condition to strong activation of that same set of genes in a secondary

condition.

Towards the end of meiotic prophase, as the cells are prepared to enter the meiotic divisions, Ime1/Ume6 induce the expression of a second meiosis-specific transcription factor Ndt80 (Chu & Herskowitz 1998). Ndt80 is responsible for activating the cascade of genes required for the meiotic divisions as well as gamete formation by interacting with mid-sporulation elements (MSEs) in the promoters of target genes (Chu & Herskowitz 1998). Similar to the URS1 binding motif, a MSE is necessary for both the mitotic repression and the meiotic gene expression at a subset of Ndt80 target genes (Pierce *et al.* 1998; Xie *et al.* 1999). At those genes, the DNA binding protein, Sum1, is recruited to an extended MSE motif where, along with its binding partners Rmf1 and Hst1, it represses gene expression through the HDAC activity of Hst1 (Xie *et al.* 1999; McCord *et al.* 2003). Unlike the relationship between Ime1 and Ume6, Sum1 is not converted to a co-activator upon interaction with Ndt80. Rather, Ndt80 and Sum1 compete with each other for binding to MSE sites (Pierce *et al.* 2003; Wang *et al.* 2005). Additionally, programmed removal of Sum1 from the DNA occurs as cells exit meiotic prophase and enter the meiotic divisions opening up the sites for Ndt80-dependent regulation (Winter 2012). What is clear from both the Ime1/Ume6 and Ntd80/Sum1 gene expression circuits is that meiotic genes are tightly regulated both inside and outside of the meiotic program, requiring mechanisms to timely induce strong expression when the genes are needed, as demonstrated by the use of Ime1 and Ndt80 transcriptional activators, but also repress them in mitosis as is performed by Ume6 and Sum1.

These few transcription factors cannot possibly account for all patterns of gene expression observed in meiosis. Practically every yeast gene is expressed at some point during meiosis and almost 2/3 of them change in their expression by at least 10-fold (Brar *et al.* 2012). In addition, during meiosis, transcription frequently begins either upstream of canonical TSSs or even within ORFs (Brar *et al.* 2012). The extent to which known vs unknown transcription factors and transcriptional mechanisms mediate these patterns of gene expression is an active area of research in the lab. With the advent of ribosome profiling and genome-wide quantitative mass spectrometry, it has been possible to begin unraveling how transcription, translation, and protein abundance interact to determine the final gene expression output in a single developmental program (Brar *et al.* 2012; Cheng *et al.* 2018). To our knowledge this has made meiosis in budding yeast the single developmental process in any organism to have all three measurements on the genome-wide scale sampled to a quantitative depth. It positions budding yeast meiosis as an ideal system for the further study of gene regulatory mechanisms.

A holistic approach to studying gene expression

By considering the current knowledge of the field with regard to both transcriptional and translational gene regulation in the context of meiosis, a biologically relevant developmental program, discoveries of novel insights into biology are possible even in the most well-studied

of all eukaryotes. Traditionally, the fields outlined above were studied independently. Transcription, translation, and chromatin biology are all massive and formidable fields in their own right that could be extensively expanded upon. It can make branching out to a new step of the gene regulatory process overwhelming. We will argue in the following chapters though that it is essential. This stems from the observation that protein levels are frequently not correlated with mRNA levels (Ingolia *et al.* 2009; Brar *et al.* 2012). By investigating this phenomenon at the locus of the kinetochore gene *NDC80*, we uncovered a 5'-extended mRNA that is expressed in meiosis. This mRNA rather than producing protein plays a purely regulatory role as it is translationally repressed by uORFs. Further, its transcription represses downstream transcription initiation at the more proximal canonical promoter through a mechanism that relies on Set2 and Set3, similar to what was described at the Ime1 locus (van Werven *et al.* 2012). By regulating Ndc80 in this way, the toggling between transcript isoforms can affect the function of the entire kinetochore complex in a manner that depends on progression through meiosis. This is described and expanded upon in Chapter 2. In Chapter 3, the same regulatory mechanism is investigated but on a genome-wide scale. A pipeline for the discovery of 5'-extended mRNAs was developed and the identified transcripts are assessed to what extent they are both translationally and transcriptionally repressed. This work will conclude in Chapter 4 with an outlook on the future of similar modes of gene regulation and a perspective on how important it is to integrate studies of transcriptional, translational, and post-translational gene regulation to fully understand the inner working of a cell.

Chapter 2

Kinetochore Inactivation by a Repressive mRNA

The following chapter contains published material from two publications that I am an author on (Chen *et al.* 2017; Chia *et al.* 2017). I am a co-first author on Chen *et al.* (2017) and second author on Chia *et al.* (2017). Both articles are distributed under the terms of the Creative Commons Attribution License (CC BY 4.0), which permits unrestricted use and redistribution provided that the original author and source are credited.

2.1   Introduction

Cellular differentiation programs depend on temporally controlled waves of gene activation and inactivation. These waves in turn drive the morphogenetic events that ultimately transform one cell type into another. Differentiation models ranging from *Bacillus subtilis* sporulation to mouse embryogenesis have elucidated how transcription factor handoffs temporally activate the expression of gene clusters (Errington 2003; Zernicka-Goetz *et al.* 2009). In comparison, much less is understood about how gene repression is coordinated with the transcription factor-driven waves of gene expression and how this inactivation is mechanistically achieved.

One critical morphogenetic event that relies on inactivation is the loss of kinetochore function during meiotic prophase. The kinetochore is a protein complex that binds to centromeric DNA and serves as the attachment site for spindle microtubules to mediate chromosome segregation (Musacchio & Desai 2017) (Figure 2.1A). In multiple systems, it has been shown that kinetochores do not bind to microtubules in meiotic prophase (Asakawa *et al.* 2005; Meyer *et al.* 2015; Kim *et al.* 2013; Miller *et al.* 2012; Sun *et al.* 2011). Furthermore, this temporal inactivation is achieved through removal of the outer kinetochore, the site where microtubule attachments occur (Asakawa *et al.* 2005; Kim *et al.* 2013; Meyer *et al.* 2015; Miller *et al.* 2012; Sun *et al.* 2011) (Figure 2.1B). In the presence of a spindle,
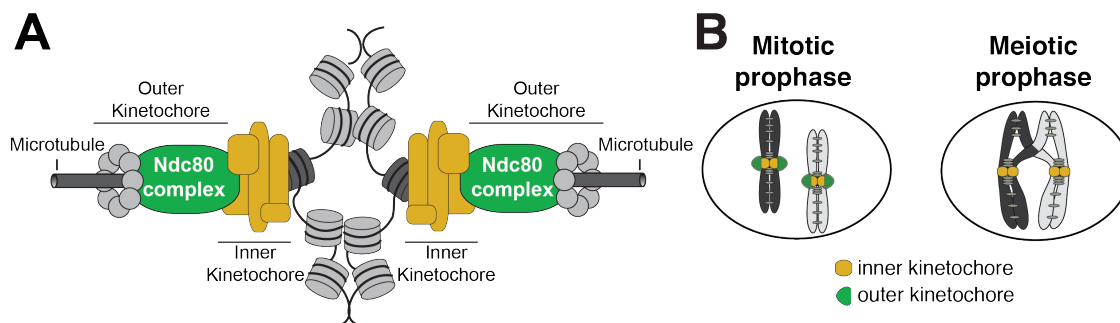
Figure 2.1: Schematics of kinetochore structure and dynamic behavior. A) kinetochores assembled on the centromere and attached to microtubules. B) During mitosis, the outer kinetochores are fully assembled, while in meiotic prophase, the outer kinetochores disassemble.

cells that fail to disassemble the outer kinetochore undergo catastrophic missegregation of meiotic chromosomes, underlying the essential nature of kinetochore downregulation during meiotic prophase (Miller *et al.* 2012). Importantly, the kinetochore is reactivated when the outer kinetochore reassembles upon transition from prophase to the meiotic divisions. How the initial removal and subsequent reassembly of the outer kinetochore is coordinated with the meiotic gene expression program is unknown.

Budding yeast provides a powerful model to address how the dynamic regulation of kinetochore function is integrated into the meiotic gene expression program. Entry into meiosis marks a clear cell-fate transition defined by the induction of Ime1, a master transcription factor. Ime1 activates the expression of genes involved in DNA replication and meiotic recombination (Kassir *et al.* 1988; van Werven & Amon 2011). Successful completion of recombination, in turn, induces a second transcription factor, Ndt80, which activates the expression of genes involved in meiotic divisions and gamete development (Chu & Herskowitz 1998; Xu *et al.* 1995). Thus, the landmark morphogenetic events in budding yeast meiosis are coordinated by the relay between these two transcription factors. Furthermore, a high-resolution map of the gene expression waves that drive meiosis has been generated for budding yeast (Brar *et al.* 2012). Importantly, analysis of this dataset revealed that, of the 38 genes that encode kinetochore subunits, *NDC80* displays the most regulated expression pattern between meiotic prophase and the subsequent division phases (Miller *et al.* 2012).

Ndc80 is the namesake member of an evolutionarily conserved complex that forms the microtubule-binding interface of the outer kinetochore (Tooley & Stukenberg 2011) (Figure 2.1A). Numerous lines of evidence indicate that the tight regulation of *NDC80* is essential for the timely function of kinetochores during meiosis. First, the decline of Ndc80 protein in meiotic prophase correlates with the dissociation of the outer kinetochore from the chromosomes (Kim *et al.* 2013; Meyer *et al.* 2015; Miller *et al.* 2012). Second, even though the other outer kinetochore subunits are expressed in meiotic prophase, they do not localize

to the kinetochores (Meyer *et al.* 2015). Third, the subsequent increase in Ndc80 protein coincides with outer kinetochore reassembly (Meyer *et al.* 2015; Miller *et al.* 2012). Finally, in the presence of a spindle, misexpression of *NDC80* in prophase disrupts proper meiotic chromosome segregation (Miller *et al.* 2012). Together, these results indicate that *NDC80* regulation is necessary for the proper timing of kinetochore function in meiosis and highlight the importance of controlling Ndc80 protein levels during meiotic differentiation.

Here we uncovered how the timely function of kinetochores is achieved through the regulation of Ndc80 protein synthesis during budding yeast meiosis. This mechanism is based on the use of two *NDC80* mRNA isoforms, which have opposite functions and display distinct patterns of expression. In addition to the canonical protein-translating *NDC80* mRNA, we found that meiotic cells also expressed a 5'-extended *NDC80* isoform. Despite carrying the entire *NDC80* open reading frame (ORF), this alternate isoform cannot produce Ndc80 protein due to the presence of regulatory upstream ORFs (uORFs) in its extended 5'-leader. Rather, its transcription plays a repressive role to inhibit transcription of the canonical *NDC80* mRNA in *cis* by establishing a repressive chromatin state, dependent on Set2 and Set3, to thereby inhibit Ndc80 protein synthesis. Furthermore, we found that the expression of the 5'-extended isoform was activated by the meiotic initiator transcription factor Ime1. Upon exit from prophase, the mid-meiotic transcription factor Ndt80 activated the expression of the canonical *NDC80* mRNA isoform. Taken together, this study uncovers how *NDC80* gene repression is achieved and how inactivation and subsequent reactivation of the kinetochore is coordinated with the transcription factor-driven waves of meiotic gene expression.

2.2   Results

Two distinct NDC80 transcript isoforms exist in meiosis

To dissect the molecular mechanism for the strict temporal regulation of the *NDC80* gene in meiosis, we first took advantage of the high-resolution mRNA-seq and ribosome profiling dataset generated for budding yeast meiosis (Brar *et al.* 2012). Analysis of this dataset revealed the presence of meiosis-specific RNA-seq reads that extend to 500 base pairs (bp) upstream of the *NDC80* ORF (Figure 2.2A). These reads appeared after meiotic entry and persisted until the end of meiosis, but were absent during vegetative growth (Figure 2.3A, vegetative) or starvation (Figure 2.3A, MATa/MATa).

To monitor the different RNA molecules generated from the *NDC80* locus, we performed northern blotting. In the absence of meiotic progression, when cells were subject to nutrient poor conditions, we detected only a single *NDC80* transcript throughout the starvation regime (no $CuSO_4$, Figure 2.4A). However, in cells undergoing synchronous meiosis, two distinct *NDC80* transcript isoforms became evident: a longer, meiosis-specific isoform, and a shorter isoform that was also present under non-meiotic conditions (Figure 2.2B and Fig-
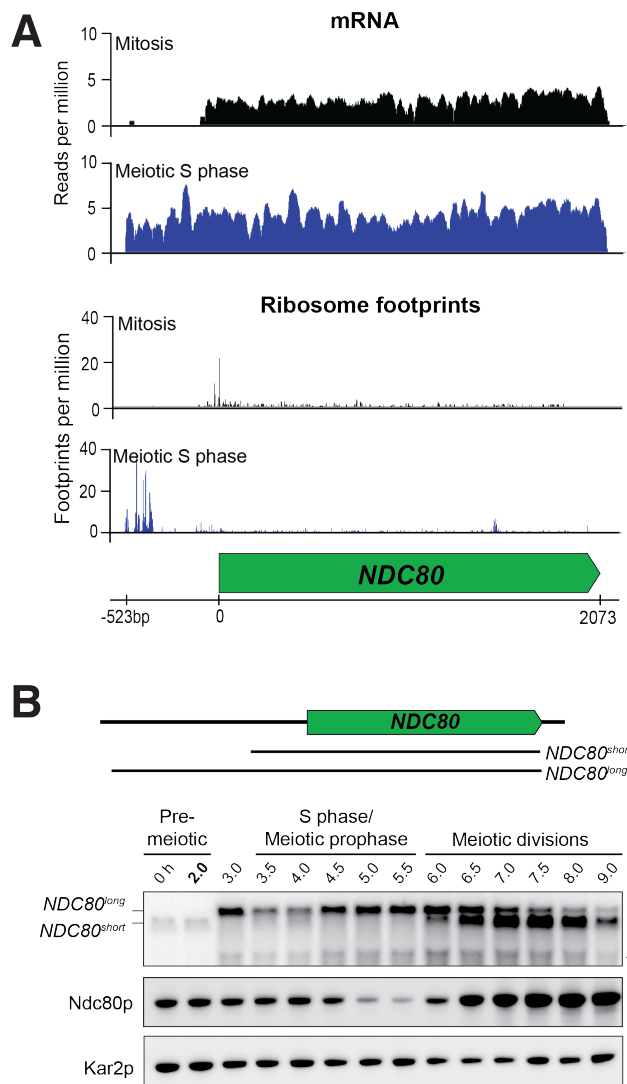
Figure 2.2: Two distinct *NDC80* transcripts are expressed during meiosis. A) Ribosome profiling and mRNA-seq reads over the *NDC80* locus during vegetative growth (top track) or meiotic S phase (bottom track). Data are derived from Brar *et al.* (2012). B) *NDC80* mRNA isoforms and Ndc80 levels in meiosis. *NDC80*^{long} and *NDC80*^{short} levels were determined by northern blot, and Ndc80 level was determined by anti-V5 immunoblot at the indicated time points. To induce meiotic entry, *IME1* and *IME4* expression was induced in the strain UB1337 by addition of $CuSO_4$ 2 hours after cells were transferred to SPO. *SCR1*, loading control for northern blot. Kar2, loading control for immunoblot. One of the two repeated experiments is shown. * indicates a smaller RNA product, which likely represents a truncated form of *NDC80*^{long}.
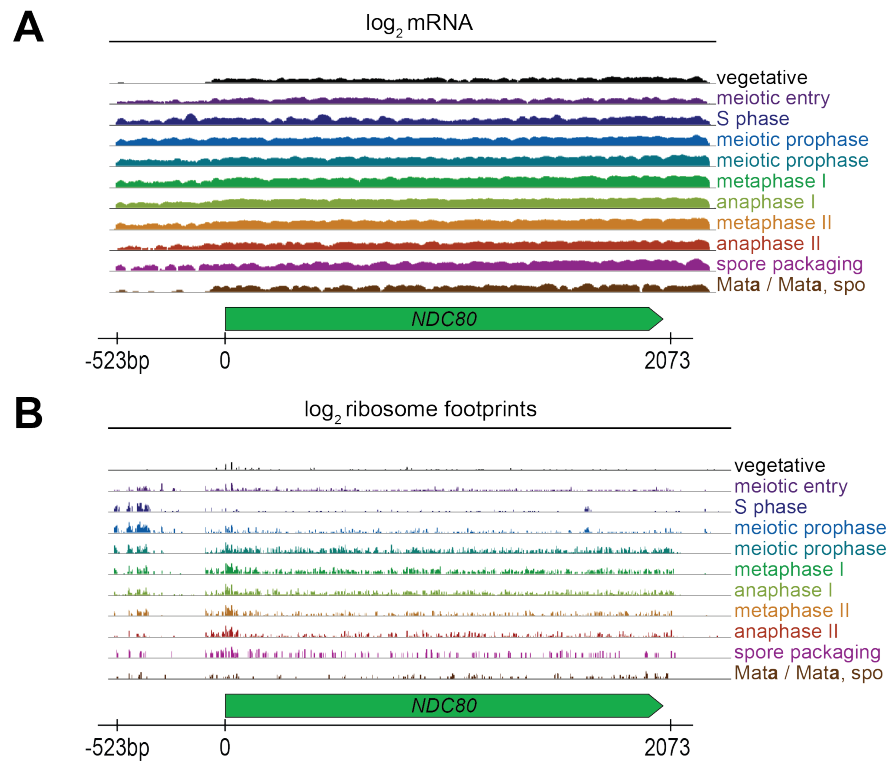
Figure 2.3: A) mRNA-seq and B) Ribosome profiling reads over the *NDC80* locus, during vegetative growth, starvation (MATa/MATa), and throughout meiosis. Data are derived from Brar *et al.* (2012).

ure 2.4A). The longer isoform appeared after meiotic entry, persisted throughout meiotic prophase and gradually disappeared during the meiotic divisions. The shorter isoform was present in vegetative cells prior to meiotic entry, but was weakly expressed during S phase and meiotic prophase. Its abundance dramatically increased during the meiotic divisions (Figure 2.2B and Figure 2.4). Interestingly, the Ndc80 protein levels were noticeably higher during the meiotic stages when the shorter transcript was the predominant isoform, but lower when the longer transcript was predominant (Figure 2.2B). Altogether, this reveals two interesting trends: 1) In meiosis, the expression of the long and short *NDC80* isoforms are anti-correlated. 2) Ndc80 protein levels positively correlate with the presence of the short isoform and negatively correlate with the long isoform (Figure 2.2B).

The long *NDC80* isoform is unable to produce Ndc80 protein due to translation of its upstream ORFs

The negative correlation between the longer *NDC80* isoform and Ndc80 protein levels suggested that this longer isoform was unable to support the synthesis of Ndc80 protein. In
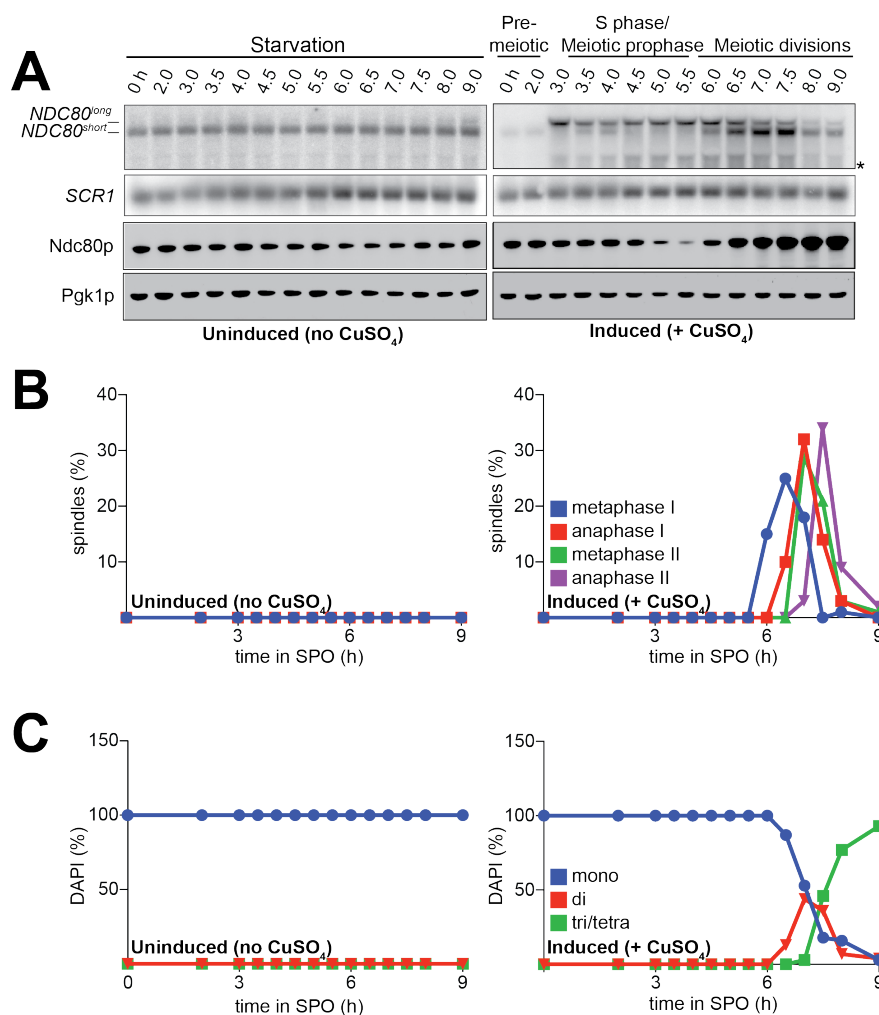
Figure 2.4: The abundance of *NDC80* mRNA isoforms and Ndc80 protein level in starvation versus in meiosis. A) *NDC80^{long}* and *NDC80^{short}* were detected by northern blot, and Ndc80, by anti-V5 immunoblot. *SCR1*, loading control for northern blot. Pgk1, loading control for immunoblot. Cells harboring the *pCUP-IME1 pCUP-IME4* system (UB1337) were transferred into SPO at 0 hour. After 2 hours, the culture was split into two: In one half, *IME1* and *IME4* expression was never induced, and thus cells stayed in starvation. In the other half, *IME1* and *IME4* expression was induced by addition of $CuSO_4$. * indicates a smaller RNA product, which likely represents a truncated form of *NDC80^{long}*. B) Percentage of cells with metaphase I, anaphase I, metaphase II, or anaphase II spindles at each time point of the experiment in A. C) Percentage of cells with 1 (mononucleates), 2 (binucleates), or 3/4 (triad/tetranucleates) nuclei at the end of meiosis for the experiment in A, determined by counting cells stained with DAPI. In all analyses, 100 cells were counted per time point, per condition. Results from one of the two repeated experiments are shown.

addition to the *NDC80* ORF, the longer isoform contains nine uORFs, each with an AUG start codon. The first six of these uORFs, those closest to the 5' end of the mRNA, have ribosome profiling signatures consistent with them being translated in meiosis (Figure 2.5A). Upstream start codons in transcript leaders can capture scanning ribosomes to alternate reading frames, thereby restricting ribosome access to the main ORF (Arribere & Gilbert 2013; Calvo *et al.* 2009; Johnstone *et al.* 2016).

We mutated the start codon of the first six uORFs (*Δ6AUG*) to test whether translation of the uORFs within the longer *NDC80* isoform represses translation of Ndc80 protein from this mRNA. In the *Δ6AUG* strain, the negative correlation between the long isoform and Ndc80 protein level persisted (Figure 2.5B), potentially because translation of the remaining three uORFs could still repress translation of the ORF. Indeed, when all nine AUGs were mutated, Ndc80 protein became highly abundant during meiotic prophase, even though the long isoform remained the predominant *NDC80* transcript in these cells (Figure 2.5B). These results demonstrate that although the longer isoform of *NDC80* contains the entire ORF, the presence of the uORFs in its 5'-leader prevents Ndc80 translation from this mRNA.

Next, we tested whether the repressive role of the uORFs resulted from the act of translation or the peptides encoded by these uORFs. We modified the long isoform, such that it still contained all the upstream AUG start codons, but each start codon was followed by a single amino acid and then immediately by a stop codon (mini uORF). Thus, this construct retained the translation ability of the uORFs but rendered them incapable of producing a peptide chain. We found that Ndc80 levels were still reduced during meiotic prophase in the mini uORF strain (Figure 2.5B). Therefore, translation of the uORFs represses translation of the *NDC80* ORF from the long *NDC80* isoform, rendering this isoform unable to synthesize Ndc80 protein.

Our analyses so far demonstrate that the two *NDC80* mRNA isoforms differ with regards to their size and ORF coding capacity. The shorter isoform is capable of translating *NDC80* ORF. In contrast, although the longer isoform contains the entire ORF, it does not support Ndc80 synthesis. The coding information is not decoded from this isoform because uORF translation prevents ribosomes from accessing the actual ORF. To signify the unique features of each *NDC80* transcript isoform, we named the short mRNA *NDC80^{PROX}*, and the longer mRNA *NDC80^{LUTI}* for long un-decoded transcript isoform.

NDC80^{LUTI} expression is necessary to downregulate *NDC80^{PROX}*

Given that *NDC80^{LUTI}* does not appear to produce Ndc80 protein, we set out to understand why meiotic cells express this mRNA isoform. Based on the observation that the expression levels of these two isoforms are anti-correlated, we posited that the transcription of *NDC80^{LUTI}* represses *NDC80^{PROX}*. To test this hypothesis, we first eliminated *NDC80^{LUTI}* production by deleting its promoter along with different portions of
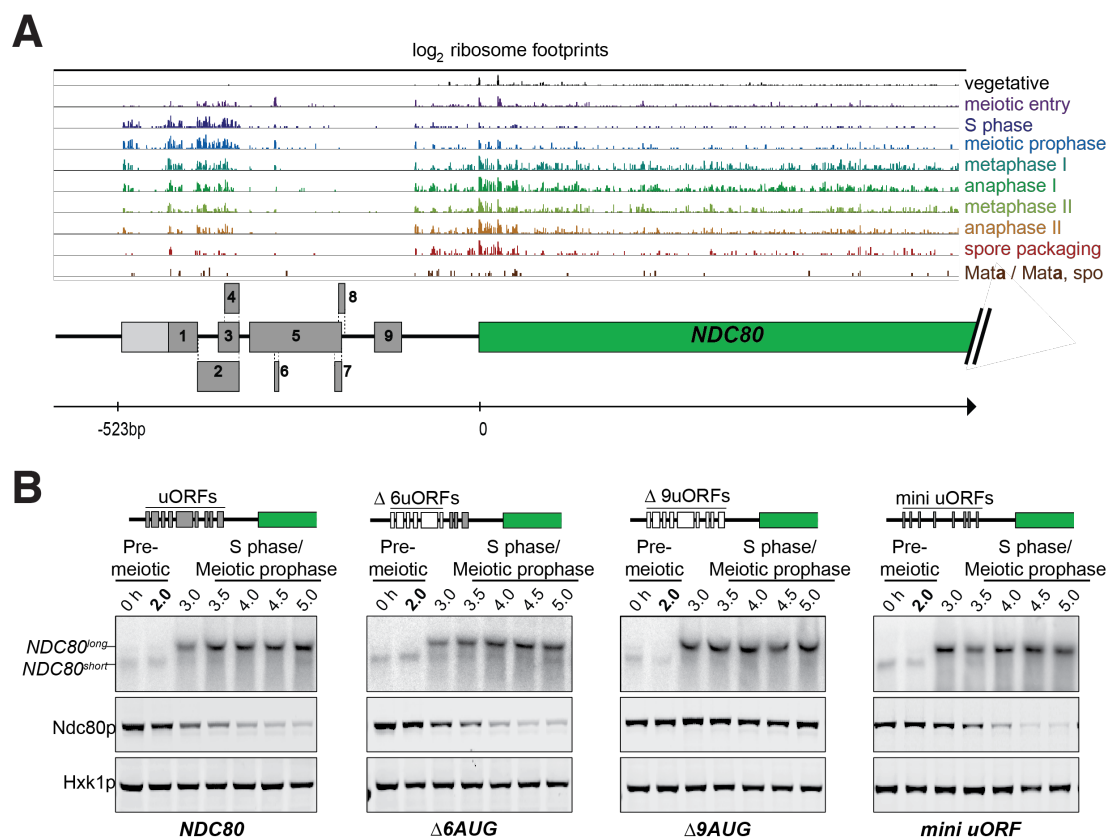
Figure 2.5: The longer NDC80 mRNA isoform is unable to synthesize Ndc80 protein due to translation of its AUG uORFs. A) Ribosome profiling reads over the *NDC80* locus throughout meiosis Brar *et al.* (2012). The dark grey boxes indicate the locations of the nine AUG uORFs. The lighter grey box indicates the location of a uORF with a near cognate start site (AUU) and ribosome footprints characteristic of translated regions. The positions of these uORFs in relation to the *NDC80* ORF are drawn proportional to the axis, with the start of *NDC80* ORF marked as 0, and the approximate 5' end of $NDC80^{LUTI}$ mRNA marked as -523 bp. B) $NDC80^{PROX}$, $NDC80^{LUTI}$, and Ndc80 abundance during synchronous meiosis (as described in Figure 2.2B) in wild type (UB6190), $\Delta 6AUG$ (UB6181), $\Delta 9AUG$ (UB6183), and mini uORF (UB9243) strains. In the $\Delta 6AUG$ and $\Delta 9AUG$ strains, the first 6 or 9 uORF AUGs in the 5'-leader of $NDC80^{LUTI}$ were converted to AUCs, respectively. The mini uORF construct contained all 9 uORF start sites in the $NDC80^{LUTI}$ leader; however, the third codon of each of the 9 uORFs was mutated to a stop codon. One of the two repeated experiments is shown.

```
-790  CAATTCTATATGGATATGCAGTATAAAAACTTGTATAAAACAGTGCAGTATCAACACACA  -731

-730  CCACTTCTTGTACACATAAATATGTACGCTACATCGAGAAATGTTCATTAATCTCTTGTT  -671

-670  GGAAAATATGACAAAGAAAAAAACTCGTTAAAGTAACATTTACCCGGATATCTGTTCAG   -611

                          URS1
-610  CCAACATTATAAAAAAGATGGGCGCTTAATTTCCACTCTTACAGATATCACTCCTTCTAA  -551
         Δ-600 to -300:
         Δ-600 to -479:

-550  TTAACTTACATCAGTAAGAAAGCAAAAAGAGCAACAATTCGCCACAAGAAGGTCTCTGTA  -491
                        NDC80^LUTI TSS

-490  GGGTCAATAGTAGCAGGGCTCTTGTCAGGGAAAGACTGTGAATGCAAAGGAAAGACGGTT  -431
                                               C   T A

-430  GGAGGTCCGAAAAGCATTACTGATGTTAGGATTGAAGAGAAATTCACCGTACTATGTGCC  -371
                         C    T                           C   T

-370  AGATGGTTATAACTGAACATTGAAATAATATCGTGATTATGGGCTGGTTGAGAGCCCCGT  -311
        GA  C   TG                           C   A

-310  TAAGTATTGTTAAAGATGTAAGCAGTCTATTGCGATTAGCGGTTAAAATACCATTTCAAG  -251
                         C

                                                          MSE
-250  GGTTACCCATACTGAAAGTGCCAAAAGAAAAAAAGATAACGTATGGAATGATATAGTCAC  -191
                                               C    C        A A

-190  AAAACTAATATTTCCTAAAATAAACAAACAACACCAAAAATGGTAGAAATAATTGGAATA  -131
                                               C   T G

-130  CATTCACAGGAGAGGTAGAATCGTCCCTGTAAAAAAAAATTGAAACGGAACAAGAAAAATA  -71
                      NDC80^PROX                NDC80^PROX
                         TSS1                      TSS2
-70   ACTGACGTTCCTTTCTTTAAACGTTAGGGTGTACTATAATATATTCCATAAGAGATACCC  -11

-10   GTTTCTATAA  -1
```

Figure 2.6: Annotated upstream intergenic region of the NDC80 locus and engineered mutations used in this study. Transcription start sites (TSS) for $NDC80^{LUTI}$ and $NDC80^{PROX}$, as estimated from a published RNA-seq dataset (Brar *et al.* 2012), are highlighted in grey. The $\Delta NDC80^{LUTI}$ constructs used in this study are marked: blue for the -600 to -300 deletion (Figure 2.7A) and orange for the -600 to -479 deletion (Figure 2.7B). Red letter Cs represent the G to C mutations engineered to abolish uORF translation in the $\Delta 6AUG$ and $\Delta 9AUG$ strains (Figure 2.5B). Purple letters represent the nucleotides mutated in the mini uORF strain. The site at which the terminator sequence was inserted is indicated by a stop sign.
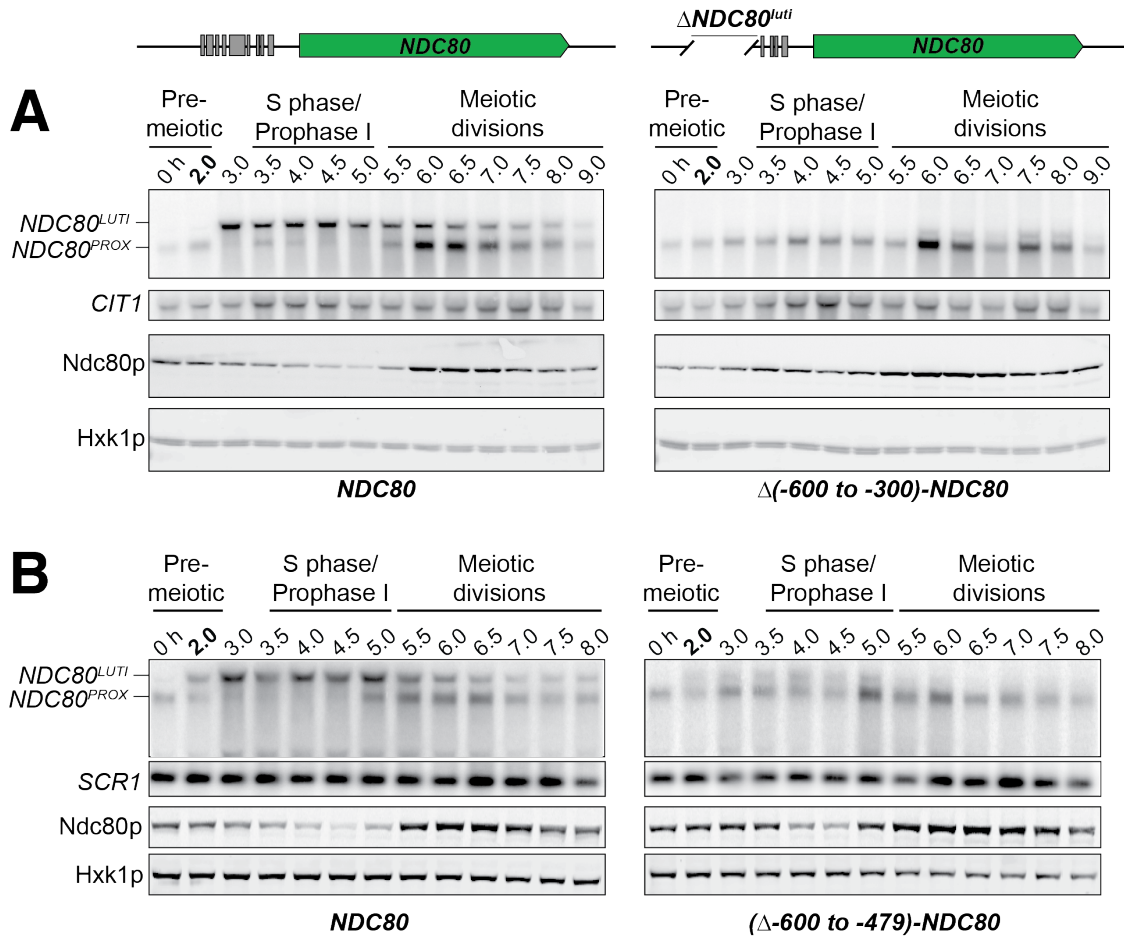
Figure 2.7: $NDC80^{LUTI}$ is necessary to downregulate $NDC80^{PROX}$. $NDC80^{PROX}$, $NDC80^{LUTI}$, and Ndc80 abundance during synchronous meiosis (as described in Figure 2.2B) in wild type cells (FW1902) and in $\Delta NDC80^{LUTI}$ cells, in which A) 300-600 bps (FW1871) or B) 479-600 bps (UB6079) upstream of the Ndc80 translation start site were deleted. Ndc80 level was determined by anti-V5 immunoblot. Hxk1, loading control for immunoblot. One of the two repeated experiments is shown. A) $CIT1$, loading control for northern blot. B) $SCR1$, loading control for northern blot.
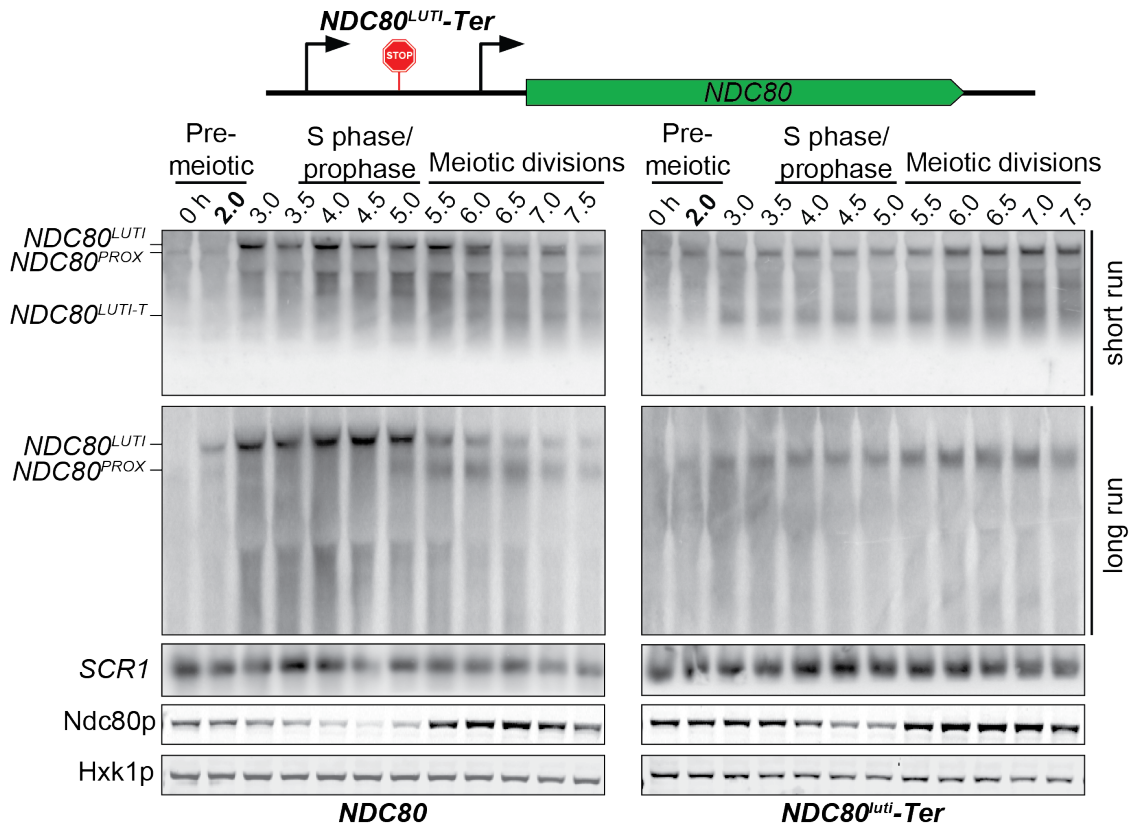
Figure 2.8: Premature termination of $NDC80^{LUTI}$ prevents $NDC80^{PROX}$ downregulation. $NDC80^{PROX}$, $NDC80^{LUTI}$, and Ndc80 abundance during synchronous meiosis in wild type cells (UB6190) and in $NDC80^{LUTI}$-Ter cells (UB6077), which harbor a terminator sequence inserted after the second uORF of $NDC80^{LUTI}$. Ndc80 protein level was determined by anti-V5 immunoblot. $SCR1$, loading control for northern blot. Hxk1, loading control for immunoblot. Top (short run): the gel was run for 1.5 hours. Middle: (long run) the gel was run for 3 hours. Note that the $NDC80^{LUTI}$ and $NDC80^{PROX}$ isoforms could be sufficiently resolved only in the long run conditions, while the truncated $NDC80^{LUTI}$ transcript due to early termination ($NDC80^{LUTI}$ term) could only be detected in the short run conditions. One of the two repeated experiments is shown.

the $NDC80^{LUTI}$ transcript ($\Delta NDC80^{LUTI}$, Figure 2.6). As shown by northern blotting, $NDC80^{PROX}$ was detected during meiotic prophase in two different $\Delta NDC80^{LUTI}$ mutant strains (Figure 2.7A-B). Accordingly, Ndc80 protein levels increased throughout meiotic prophase (Figure 2.7A-B).

Additionally, we inserted a termination sequence 220 bp downstream of the $NDC80^{LUTI}$ transcription start site ($NDC80^{LUTI}$-Ter). We observed that, upon early termination of $NDC80^{LUTI}$, $NDC80^{PROX}$ mRNA and Ndc80 protein persisted in meiotic prophase (Figure 2.8). This observation suggests that continuous transcription through the $NDC80^{PROX}$ promoter is necessary for $NDC80^{PROX}$ repression. It also indicates that the repression of $NDC80^{PROX}$ is not due to competition between the $NDC80^{PROX}$ promoter and the $NDC80^{LUTI}$ promoter for RNA polymerase and the general transcription machinery. Altogether, we conclude that expression of the $NDC80^{LUTI}$ mRNA is required to repress the $NDC80^{PROX}$ transcript and reduce Ndc80 protein levels during meiotic prophase.

$NDC80^{LUTI}$ represses transcription in *cis*

By what mechanism does $NDC80^{LUTI}$ reduce the steady-state level of $NDC80^{PROX}$? We posited that $NDC80^{LUTI}$ acts in *cis* based on other instances of overlapping transcription in budding yeast (Martens *et al.* 2004; van Werven & Amon 2011; Bird *et al.* 2006). To test this, we engineered strains to have one wild type $NDC80^{LUTI}$ allele and another allele in which the promoter of $NDC80^{LUTI}$ has been deleted ($\Delta NDC80^{LUTI}$). In order to monitor Ndc80 protein levels, we inserted a 3V5 epitope as a C-terminal fusion to $NDC80$ in either the wild type or the $\Delta NDC80^{LUTI}$ allele. If $NDC80^{LUTI}$ functions in *trans*, then Ndc80-3V5 should be downregulated to the same extent in both strains. Instead, we found that Ndc80-3V5 was downregulated only when $NDC80^{LUTI}$ was generated on the same chromosome, directly upstream of $NDC80$-3V5 (Figure 2.9B). This result demonstrates that $NDC80^{LUTI}$-mediated repression occurs in *cis*, since $NDC80^{LUTI}$ cannot reduce Ndc80 protein expression from a copy of $NDC80$ on another chromosome (Figure 2.9C).

Transcription of $NDC80^{LUTI}$ correlates with reduced binding of TFIIB and repressive chromatin in the $NDC80^{PROX}$ promoter

The mechanism by which $NDC80^{LUTI}$represses the downstream $NDC80^{PROX}$ promoter might be related to a transcriptional interference mechanism during which intergenic transcription or transcription over promoter regions establishes a repressive chromatin state and prevents transcription factors from binding (Martens *et al.* 2004; Hainer & Martens 2011; van Werven *et al.* 2012). Similar to transcriptional interference, $NDC80^{LUTI}$-mediated repression of $NDC80^{PROX}$ is exclusively *cis*-dominant (Martens *et al.* 2004; van Werven *et al.* 2012; Chen *et al.* 2017). To further investigate whether the mechanism of $NDC80^{LUTI}$-mediated gene repression also shares other features of transcriptional interference, we tested whether $NDC80^{LUTI}$ transcription alters the association of transcription factors with the
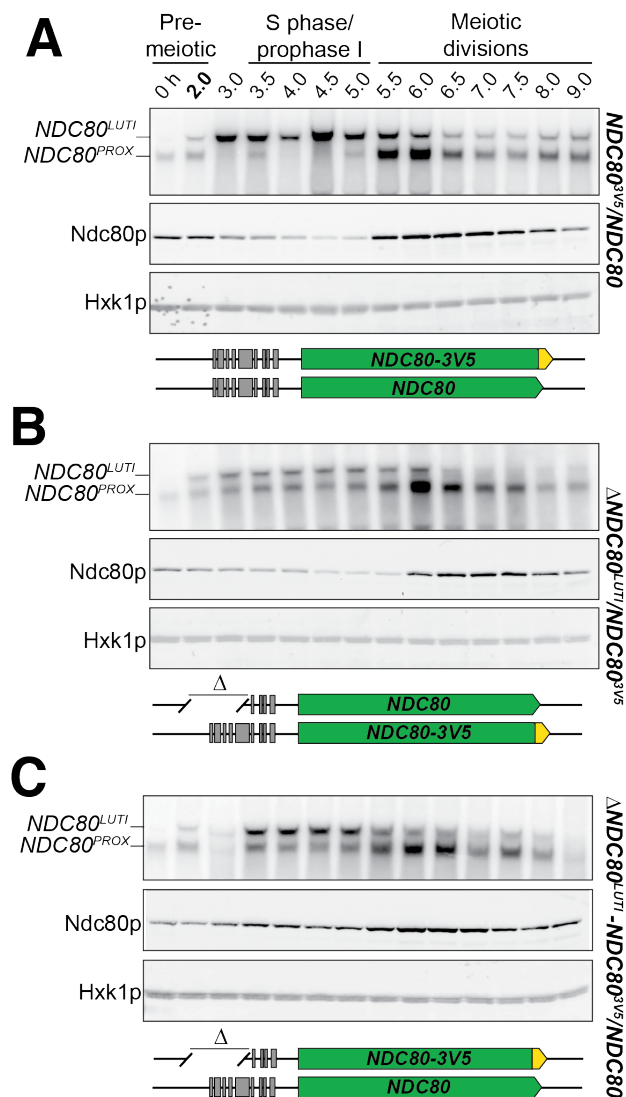
Figure 2.9: $NDC80^{LUTI}$ represses $NDC80^{PROX}$ expression in *cis*. Meiosis was induced and samples were collected and processed as in (Figure 2.2B). Ndc80 level was determined by anti-V5 immunoblot. Hxk1, loading control. Three yeast strains were used in this experiment: A) a strain (FW1900) with one *NDC80-3V5* allele and one wild type *NDC80* allele, B) a strain (FW1899) with one *NDC80-3V5* allele and one $\Delta NDC80^{LUTI}$ allele, in which 300-600 bps upstream of the Ndc80 translation start site were deleted, and C) a strain (FW1923) with one $\Delta NDC80^{LUTI}$-*NDC80-3V5* allele, which has the aforementioned 300-600 bps deletion, and one wild type *NDC80* allele. One of the two repeated experiments is shown.
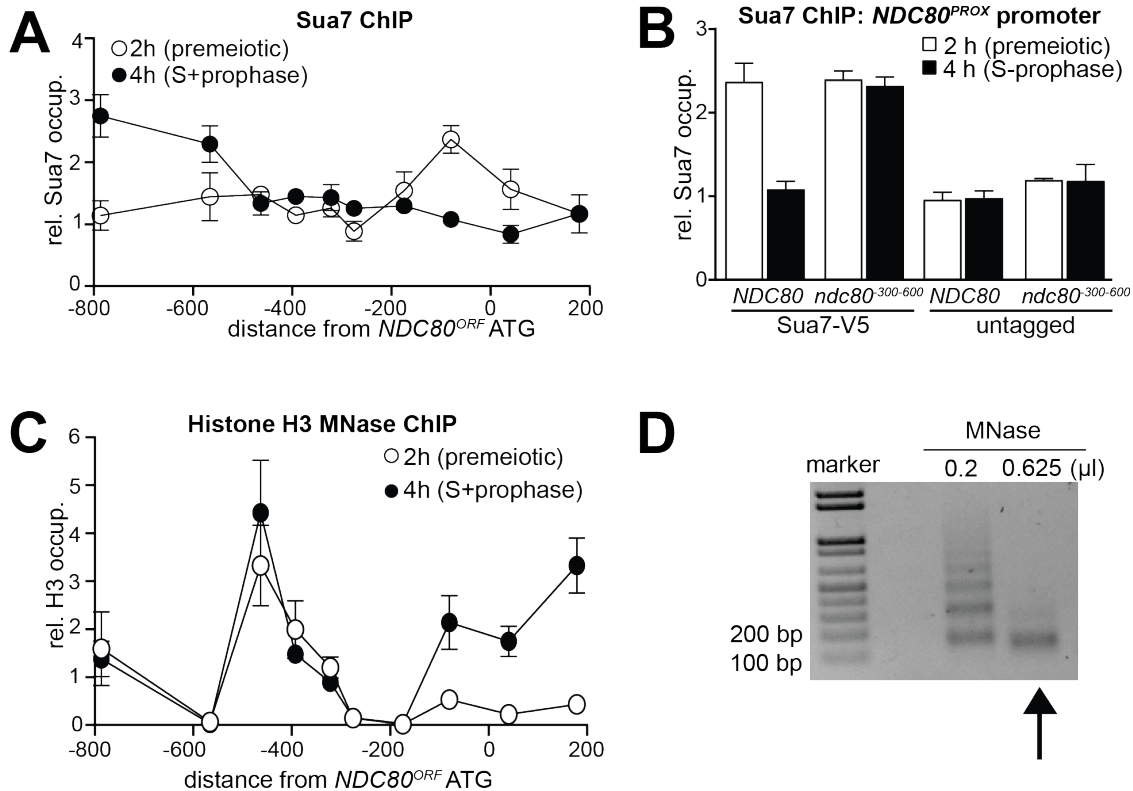
Figure 2.10: A repressive chromatin landscape over $NDC80^{PROX}$ promoter. A) $NDC80^{LUTI}$ transcription correlates with reduced TFIIB (Sua7) binding at the $NDC80^{PROX}$ promoter. Cells harboring Sua7 tagged with three copies of V5 (Sua7-V5) were induced to undergo meiosis synchronously as described in 2.2B (FW2957). Samples for chromatin immunoprecipitation were taken at two hours (2h, premeiotic) and four hours after transfer to sporulation medium (SPO) (4h, S+prophase). DNA fragments were quantified by qPCR using primer pairs spanning the $NDC80$ locus, and were normalized over the HMR locus. The midpoint position of each primer pair is indicated in the x-axis. The mean normalized signal from three independent experiments plus the standard error of the mean is displayed. B) $NDC80^{LUTI}$ transcription is required to inhibit Sua7 binding in the $NDC80^{PROX}$ promoter during meiotic prophase. Similar to A except that a mutant strain harboring a deletion upstream in the $NDC80$ promoter region (Δ*(-600 to -300)-NDC80*, FW5530) and untagged strains (FW1902 and FW1868) were included. A primer pair directed against a $NDC80^{PROX}$ core promoter was used. C) Chromatin structure at the $NDC80$ locus was determined by ChIP of histone H3 on micrococcal nuclease (MNase) treated extracts (FW1902). Samples were taken at the same times as in A. Mononucleosome extracts were used for ChIP assays with histone H3 antibodies. The recovered DNA fragments were quantified by qPCR using the primer pairs in A relative to a no MNase input. The signals from each primer pair were normalized over a primer pair directed against the *PHO5* core promoter. The midpoint position of each primer pair is indicated in the x-axis. The mean signal from three independent experiments plus the standard error of the mean is displayed. D) Example of extract with mononucleosomes prepared from cells (FW1902, 2h (premeiotic)). The arrow indicates the extract that was used for subsequent ChIP analysis.

$NDC80^{PROX}$ promoter. The binding of the basal transcription factor Sua7 (TFIIB), which is homologous to human TFIIB, changed during meiosis across the *NDC80* locus (Figure 2.10A). Before entry into meiosis, Sua7 was bound to the core promoter of $NDC80^{PROX}$. However, after *IME1* and *IME4* induction (four hours in SPO) when $NDC80^{LUTI}$ transcription occurred and cells underwent meiotic S phase, Sua7 binding to the $NDC80^{PROX}$ core promoter (around -100bp from AUG) was reduced while binding to the $NDC80^{LUTI}$ promoter (around -600bp from AUG) increased (Figure 2.10B). It is worth noting, that the signal for Sua7 binding also showed a peak at -800 bp, which may be due to fluctuation in expression of the adjacent *PAN6* gene in the divergent direction. Next, we examined Sua7 binding at the $NDC80^{PROX}$ promoter in a mutant that does not transcribe $NDC80^{LUTI}$ (*Δ(-600 to -300)-NDC80*) (Chen *et al.* 2017). In the *Δ(-600 to -300)-NDC80* mutant, no change in Sua7 binding around the $NDC80^{PROX}$ promoter was observed after induction of *IME1* and *IME4* (Figure 2.10B). This result shows that $NDC80^{LUTI}$ prevents TFIIB recruitment at the $NDC80^{PROX}$ promoter during early meiosis.

The reduction in TFIIB recruitment to the $NDC80^{PROX}$ promoter could be due to the establishment of a repressive chromatin state. For example, transcription of an intergenic ncRNA across the *SER3* promoter directs nucleosome assembly in the promoter, which is essential for *SER3* repression in budding yeast (Hainer & Martens 2011). Therefore, we examined how the chromatin structure in the $NDC80^{PROX}$ promoter is modified by $NDC80^{LUTI}$ transcription. To identify where the nucleosomes stably associate with the *NDC80* locus, we performed chromatin immunoprecipitation (ChIP) of histone H3 on micrococcal nuclease (MNase) treated chromatin extracts (see material and methods for details) (Figure 2.10C and Figure 2.10D). In premeiotic cells (2h) we detected a relatively low signal around the core promoter of $NDC80^{PROX}$, which is indicative of a nucleosome free region (NFR) and consistent with active $NDC80^{PROX}$ transcription. During meiotic prophase (4h), when $NDC80^{LUTI}$ was transcribed, the signal around the core promoter increased, indicating that nucleosome occupancy was increased. These findings are consistent with the notion that transcription of $NDC80^{LUTI}$ inhibits TFIIB recruitment and establishes a repressive chromatin state at the $NDC80^{PROX}$ promoter.

$NDC80^{PROX}$ transcription promotes Set1-dependent histone H3 lysine 4 dimethylation and Set2-dependent lysine 36 trimethylation in the $NDC80^{PROX}$ promoter

Co-transcriptional recruitment of chromatin modifying enzymes regulates the chromatin state of genes in the wake of elongating RNA polymerase II. For example, repressive chromatin marks, such as histone 3 lysine 4 dimethylation (H3K4me2) and lysine 36 trimethylation (H3K36me3) are deposited co-transcriptionally within gene bodies by the Set1 and Set2 methyltransferases (Kim & Buratowski 2009; Hampsey & Reinberg 2003). The histone deacetylase complexes Set3C and Rpd3S recognize H3K4me2 and H3K36me3, respectively,
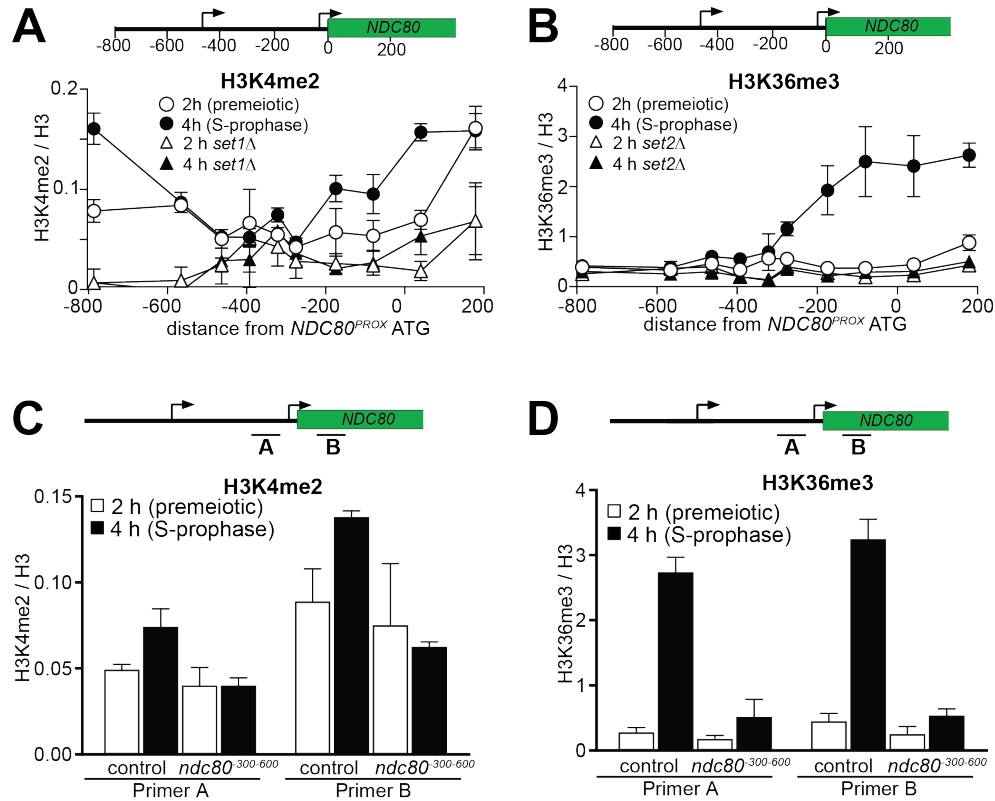
Figure 2.11: Transcription of $NDC80^{LUTI}$ promotes H3K4me2 and H3K36me3 in the promoter and 5' region of $NDC80^{PROX}$. A) Wild-type (FW1902) and $set1\Delta$ (FW3033) cells were induced to undergo meiosis synchronously. Samples for chromatin immunoprecipitation were taken at two hours (2h (premeiotic), no $NDC80^{LUTI}$ transcription) and four hours in sporulation medium (SPO) (4h (S+prophase), $NDC80^{LUTI}$ transcription). The recovered DNA fragments were quantified by qPCR using ten different primer pairs scanning the $NDC80$ locus. The midpoint position of each primer pair is indicated in the x-axis. The mean enrichment from three independent experiments plus the standard error of the mean for each primer pair is displayed. The H3K4me2 signal was normalized over histone H3. B) Similar to A, except that histone H3 lysine 36 trimethylation (H3K36me3) abundance was determined by ChIP. Wild-type (FW1902) and $set2\Delta$ (FW1472) cells harboring the $pCUP-IME1/pCUP-IME4$ alleles were used for the analysis. C) Similar to A except that the ChIP for H3K4me2 was performed in control cells (FW1902) and cells harboring a deletion upstream in the $NDC80$ promoter region ($\Delta(-600\ to\ -300)-NDC80$, FW1868). For the analyses we used primer pairs directed against the $NDC80^{PROX}$ promoter (A), and the 5' region of the $NDC80$ gene (B). The mean fold enrichment from three independent experiments plus the standard error of the mean for each primer pair are displayed. The signals were normalized to the levels of H3. D) Similar to C except that H3K36me3 levels were determined by ChIP.
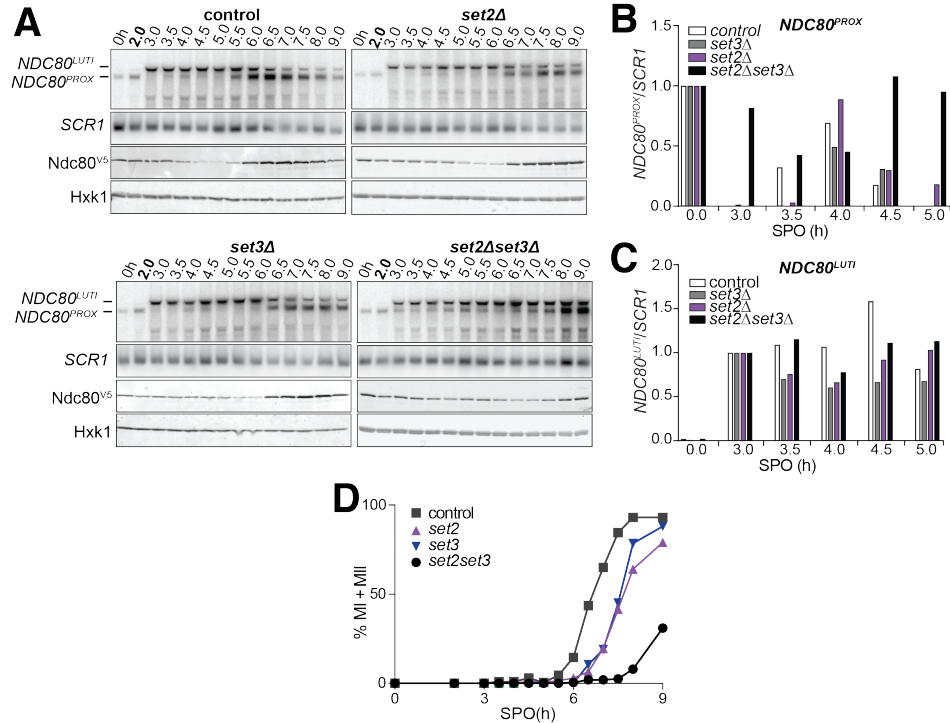
Figure 2.12: Set2 and Set3 mediate $NDC80^{LUTI}$ induced gene repression of $NDC80^{PROX}$. A) . Control (FW1902), $set2\Delta$ (FW2929), $set3\Delta$ (FW2928) and $set2\Delta set3\Delta$ (FW1922) cells harboring $pCUP\text{-}IME1/pCUP\text{-}IME4$ and $NDC80\text{-}V5$ were grown in rich medium, transferred to presporulation medium, and then shifted to SPO medium. After 2 hours, *IME1* and *IME4* expression were induced, and samples for northern and western blot analyses were taken at the indicated time points. Northern blot membranes were prepared and hybridized with a probe that detects both $NDC80^{LUTI}$ and $NDC80^{PROX}$ transcripts. As a loading control, membranes were also hybridized with *SCR1*. Ndc80 protein was detected with anti-V5 antibodies and Hxk1 levels were determined with anti-hexokinase antibodies. B-C) Quantification of (B) $NDC80^{PROX}$ and (C) $NDC80^{LUTI}$ levels in the experiment described in A. Signals are normalized to *SCR1*. To control for variation in overall signal between different northern blots, the $NDC80^{PROX}$ signal at the 0 hour time point was set to one, and the $NDC80^{LUTI}$ signal at the 3 hour time point was set to one. D) $set2\Delta set3\Delta$ mutants undergo meiosis with delayed kinetics. Kinetics of meiotic divisions (MI+MII) in control (FW1902), $set2\Delta$(FW2929), $set3\Delta$ (FW2928) and $set2\Delta set3\Delta$ (FW1922) cells harboring p$CUP\text{-}IME1/pCUP\text{-}IME4$ and $NDC80\text{-}V5$. Cells were prepared for meiosis as in A. After 2 hours, *IME1* and *IME4* expression were induced, and samples were taken at the indicated time points, fixed, and stained with DAPI. The percentage of cells with one, two or more DAPI masses was determined for at least 200 cells per time point.

and repress cryptic transcription from chromatin carrying these modifications (Carrozza *et al.* 2005; Keogh *et al.* 2005; Kim & Buratowski 2009; Govind *et al.* 2010). Set1/Set3C and Set2/Rpd3S have also been implicated in transcription-coupled repression of gene promoters (Kim *et al.* 2012; van Werven *et al.* 2012; Ard & Allshire 2016; Houseley *et al.* 2008). To investigate whether $NDC80^{PROX}$ mediated repression of $NDC80^{PROX}$ also requires Set1/Set3C and Set2/Rpd3S, we measured the distribution of H3K4me2 and H3K36me3 marks at the *NDC80* locus (Figure 2.11A and 2.11B). We observed almost no enrichment in the $NDC80^{PROX}$ promoter of either marks in premeiotic cells (labelled 2h in Figure 2.11A and 2.11B), but H3K36me3 and to a lesser extent, H3K4me2 increased at the $NDC80^{PROX}$ promoter in meiotic prophase cells (labeled 4h in Figure 2.11A and 2.11B). As expected, the enrichment of H3K4me2 and H3K36me3 depended on Set1 and Set2, respectively (Figure 2.11A and 2.11B, *set1Δ* and *set2Δ*). In *Δ(-600 to -300)-NDC80* cells that do not express $NDC80^{PROX}$, the deposition of the H3K36me3 and H3K4me2 marks in the $NDC80^{PROX}$ promoter was reduced (Figure 2.11C and 2.11D). Thus $NDC80^{PROX}$ transcription promotes the deposition of repressive H3K4me2 and H3K36me3 marks within the $NDC80^{PROX}$ promoter.

Because H3K36me3 and H3K4me2 marks localize to the $NDC80^{PROX}$ promoter when $NDC80^{PROX}$ is transcribed, we examined whether Set1/Set3C and Set2/Rpd3S contribute to $NDC80^{PROX}$ repression. Since Set1 also plays an important role in meiotic recombination, we deleted *SET3* to test how the Set1/Set3C pathway regulates the *NDC80* locus (Sommermeyer *et al.* 2013; Borde *et al.* 2009; Acquaviva *et al.* 2013). In the *set2Δset3Δ* double mutant, but not the single mutants, both $NDC80^{PROX}$ and $NDC80^{LUTI}$ transcripts were detected throughout multiple time points in early meiosis, and the steady-state level of Ndc80 protein remained high (Figure 2.12A-C, compare the time points from two to five hours between control and mutant cells). The *set2Δset3Δ* double mutant cells entered and underwent meiosis with delayed kinetics (Figure 2.12). Thus, it is possible that a population of cells never entered meiosis and continued to express the mitotic $NDC80^{PROX}$ mRNA isoform. We improved the kinetics of meiosis by adopting a different meiotic synchronization protocol (Figure 2.13A). Instead of growing cells in pre-sporulation medium, we shifted them directly to sporulation medium after they reached saturation in nutrient rich conditions. We then induced *IME1* and *IME4*. This synchronization procedure reduced the delay in meiotic divisions (compare Figure 2.13B and 2.12D). In addition, meiotic S phase was completed in more than 75 percent of cells after 6 hours, indicating that the majority of cells had entered meiosis (Figure 2.13A). Importantly, $NDC80^{LUTI}$ mediated repression was still compromised in *set2Δset3Δ* double mutant cells despite improved synchrony of meiosis (Figure 2.13C-E, compare the time points from three-to five hours for the control with three-to six hours for the mutant cells). Further analyses of selective time-points (3.5 and 4.5 hours) confirmed that there were significant differences in $NDC80^{PROX}$ levels between the control and the *set2Δset3Δ* double mutant, but not the single mutants (Figure 2.13F-G).

Previous work showed that the *set2Δ* mutant exhibits increased nucleosome dynamics leading to de-repression of cryptic promoters (Venkatesh *et al.* 2012). In addition, the
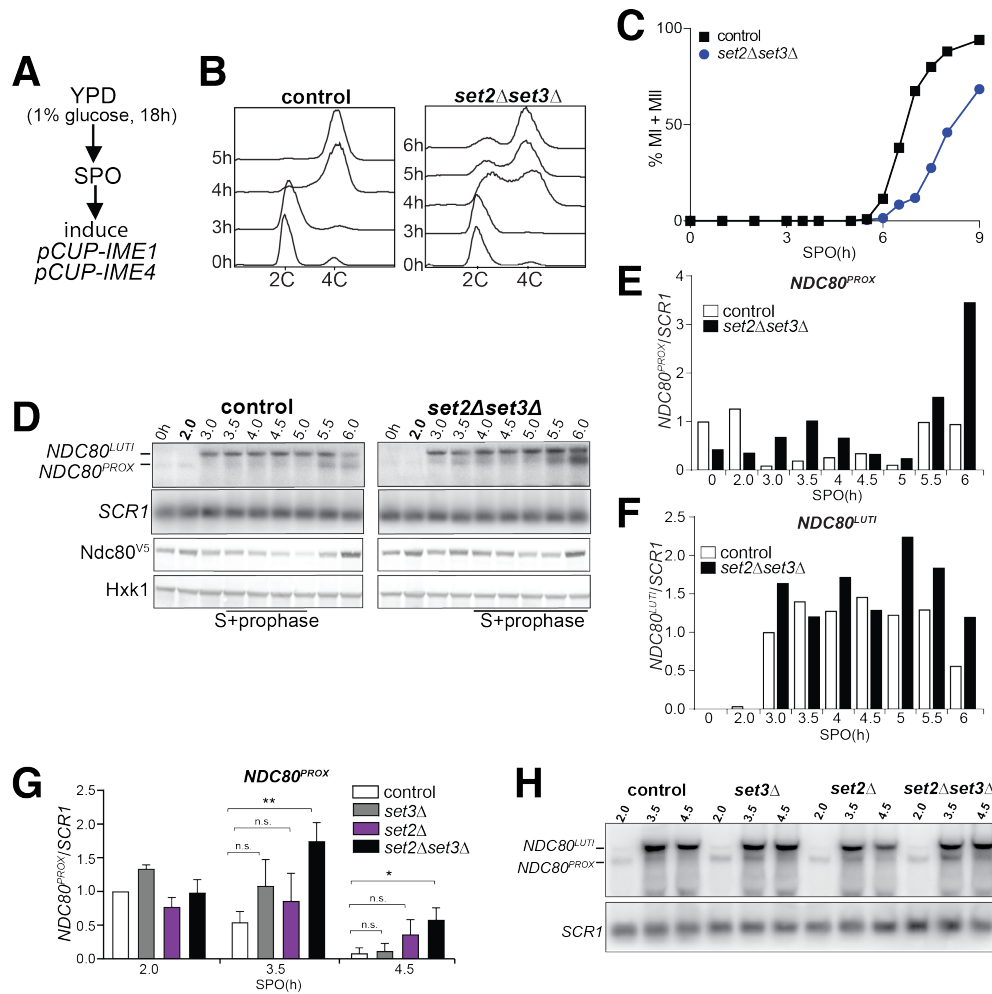
Figure 2.13: Meiotic synchronization with a direct YPD to SPO protocol. A) Scheme of protocol. B-H) Meiosis was induced as in A. B) Flow cytometry of DNA content from meiosis induced with control (FW1902) and *set2Δset3Δ* (FW1922) strains. C) Kinetics of meiotic divisions (MI+MII) in control (FW1902) and *set2Δset3Δ* (FW1922) cells. DAPI masses were counted for at least 200 cells per time point. D) RNA was detected with a probe specific to *NDC80*. As a loading control, membranes were hybridized with a probe for *SCR1*. Ndc80 protein was detected with anti-V5 antibodies and Hxk1 levels were determined with anti-hexokinase antibodies. E-F) Quantification of (E) $NDC80^{PROX}$ and (F) $NDC80^{LUTI}$ from the northern blot in D. Signals are normalized to *SCR1*. To control for variation in overall signal between different northern blots, the $NDC80^{PROX}$ signal at 0 hour was set to one, and the $NDC80^{LUTI}$ signal at 3 hour was set to one. G-H) Selective time points were taken for northern blot analysis (H) and quantification (G) from strains in Figure 2.12. As a loading control, the northern membranes were hybridized with *SCR1*. The $NDC80^{PROX}$ levels were quantified (G) and data from three independent experiments plus the standard error of the mean (SEM) is displayed. One-tailed, unpaired t-tests were conducted. * denotes p-value < 0.05. ** denotes p-value < 0.01. n.s. denotes not significant.
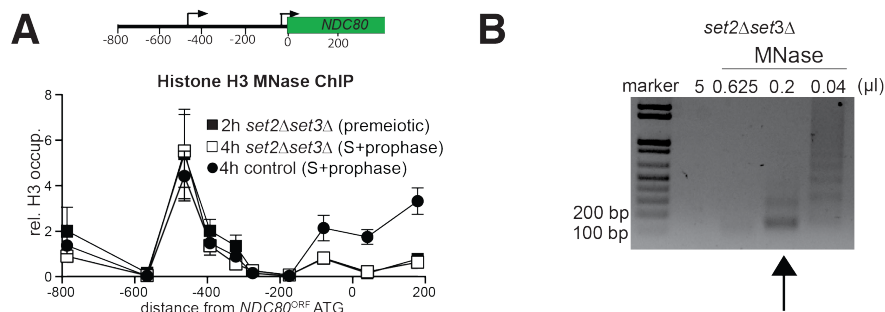
Figure 2.14: $NDC80^{LUTI}$ transcription requires Set2 and Set3 to establish a repressive chromatin state at the promoter of $NDC80^{PROX}$. A) Chromatin structure at the $NDC80$ locus was determined by ChIP of histone H3 on micrococcal nuclease (MNase) treated extracts in control (FW1902) and $set2\Delta set3\Delta$ (FW1922) cells. Samples were taken prior to $IME1/IME4$ induction at 2 hours in SPO (2h, premeiotic) and after induction at 4 hours in SPO (4h, S + prophase), fixed with formaldehyde, and mononucleosome fragments were isolated. The recovered DNA fragments were quantified by qPCR using ten different primer pairs directed against the $NDC80$ locus relative to a no MNase input. The signals from each primer pair were then normalized over a primer pair directed against the $PHO5$ core promoter. The midpoint position of each primer pair is indicated in the x-axis. The mean signal from three independent experiments plus the standard error of the mean for each primer pair is displayed.

$set3\Delta$ mutant displays reduced histone H3 density in the 5' region of transcribed genes (Kim & Buratowski 2009). Set2 and Set3 are also required for transcription coupled chromatin changes in the $IME1$ promoter by the long noncoding RNA $IRT1$ (van Werven *et al.* 2012). These findings prompted us to examine whether Set2 and Set3 are necessary for $NDC80^{LUTI}$ mediated nucleosome assembly in the $NDC80^{PROX}$ promoter. We found that even though $NDC80^{LUTI}$ was efficiently transcribed in $set2\Delta set3\Delta$ cells during early meiosis (Figure 2.12A-B and 2.13C-D), repressive chromatin was not established at the $NDC80^{PROX}$ promoter (Figure 2.14A-B).

The $NDC80^{LUTI}$ leader is sufficient to downregulate gene expression at the $NUF2$ locus

Since $NDC80^{LUTI}$ is necessary to repress $NDC80^{PROX}$ during meiosis, we next investigated whether the $NDC80^{LUTI}$ leader is sufficient to regulate other genes in meiosis. We replaced the promoter and 5'-leader of $NUF2$, the gene encoding the binding partner of Ndc80, with the promoter and 5'-leader region of $NDC80^{LUTI}$ ($NDC80^{LUTI}$-$NUF2$). In wild type cells, a single $NUF2$ mRNA species was expressed in meiotic prophase, a stage when $NUF2$ mRNA levels and Nuf2 protein levels were stable (Figure 2.15A-C). In contrast, $NDC80^{LUTI}$-$NUF2$ cells expressed a longer mRNA ($NUF2^{LUTI}$) in meiotic prophase (Figure 2.15A), and the abundance of $NUF2^{PROX}$ transcripts was reduced by 60% compared to

that in the pre-meiotic stage (Figure 2.15B), a reduction level similar to that of the Nuf2 protein (Figure 2.15C). This result demonstrates that the promoter and 5'-leader sequence of $NDC80^{LUTI}$ is sufficient to downregulate another protein in meiotic prophase.

Master meiotic transcription factors Ime1 and Ndt80 regulate $NDC80^{LUTI}$ and $NDC80^{PROX}$ expression, respectively

Since the timely expression of $NDC80^{LUTI}$ and $NDC80^{PROX}$ is crucial to establish the temporal pattern of Ndc80 protein levels in meiosis, we next investigated which transcription factors directly control $NDC80^{LUTI}$ and $NDC80^{PROX}$ expression. In *S. cerevisiae*, meiotic gene expression is orchestrated by two master transcription factors: Ime1 and Ndt80 (Chu & Herskowitz 1998; Kassir *et al.* 1988; Xu *et al.* 1995). Diploid MATa/MATα cells initiate meiosis by expressing *IME1* in response to nutrient deprivation (van Werven & Amon 2011). Interestingly, *IME1* expression correlated with the time of $NDC80^{LUTI}$ expression, suggesting that Ime1 might regulate $NDC80^{LUTI}$ transcription. Indeed, deletion of *IME1* abolished $NDC80^{LUTI}$ production and resulted in persistent levels of $NDC80^{PROX}$ transcript and Ndc80 protein (Figure 2.16A-B).

Ime1 does not directly bind to DNA, but functions as a co-activator for Ume6 (Washburn & Esposito 2001). In the absence of Ime1, Ume6 represses early meiotic genes in mitosis by binding to a consensus site called the upstream repressive sequence (URS1) in the promoters of these genes. Upon meiotic entry and subsequent interaction with Ime1, the Ume6-Ime1 complex activates the transcription of these early meiotic genes (Park *et al.* 1992; Bowdish *et al.* 1995). Given the close relationship between Ime1 and Ume6, we inspected the 5' intergenic region of *NDC80* and identified a consensus site for Ume6 583 bp upstream of the Ndc80 translation start site (Figure 2.17 and 2.6), within the $NDC80^{LUTI}$ promoter. ChIP analysis revealed that Ume6 binding was enriched over the predicted URS1 site in mitosis and early meiosis (Figure 2.18A), whereas Ume6 binding was undetectable within the $NDC80^{PROX}$ promoter (Figure 2.18B). Deletion of the URS1 site (*ndc80-urs1Δ*) completely abolished Ume6 binding to the $NDC80^{LUTI}$ promoter (Figure 2.18C), but did not affect another Ime1-Ume6 target gene *IME2* (Figure 2.18D). Consistent with the role of Ume6 as a transcriptional repressor in mitosis, deletion of the URS1 site resulted in leaky expression of $NDC80^{LUTI}$ during vegetative growth (Figure 2.18E) and reduced expression of $NDC80^{PROX}$ (Figure 2.18E). Abolishing Ume6 binding eliminated strong induction of $NDC80^{LUTI}$ in meiosis (Figure 2.18E), causing moderately increased levels of $NDC80^{PROX}$ transcript by northern blot and Ndc80 protein in meiotic prophase (Figure 2.18E). We conclude that similar to early meiotic genes, Ime1 and Ume6 directly regulate the transcription of $NDC80^{LUTI}$.

The second key meiotic transcription factor, Ndt80, is required for meiotic chromosome segregation and spore formation (Chu & Herskowitz 1998; Xu *et al.* 1995). Expression of *NDT80* occurs shortly before the reappearance of $NDC80^{PROX}$ transcript. Within the budding yeast lineage, an Ndt80 consensus site, called the mid-sporulation element (MSE), was
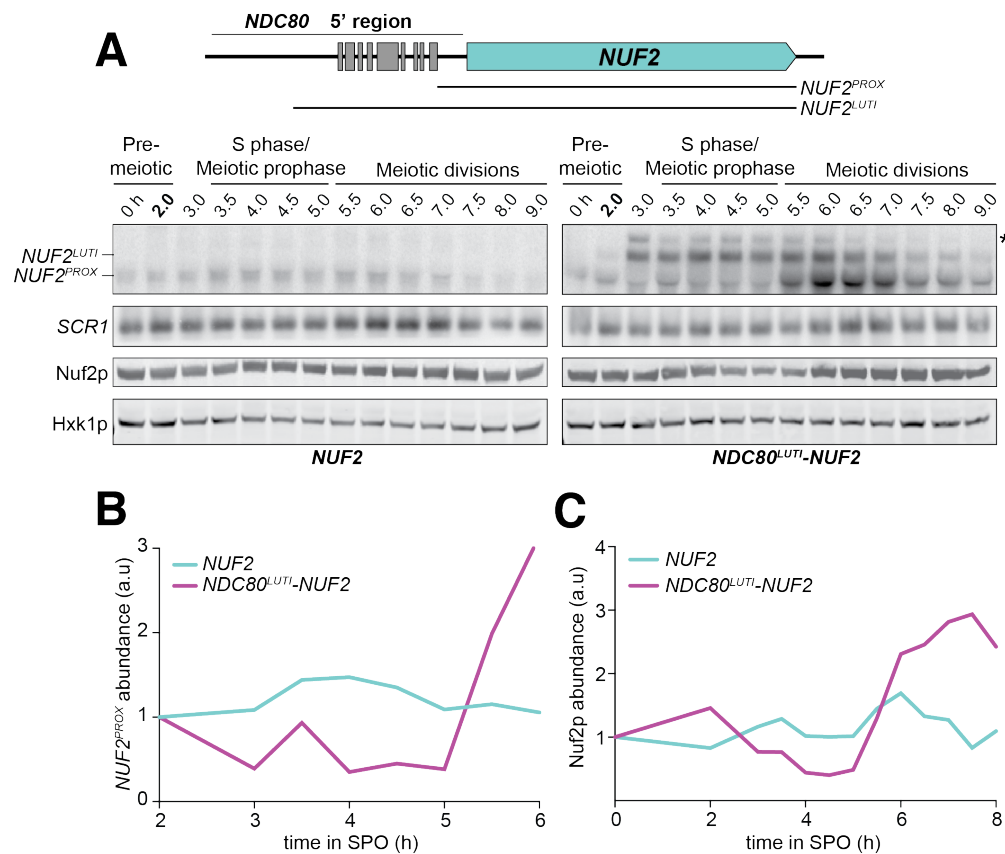
Figure 2.15: The $NDC80^{LUTI}$ leader is sufficient to downregulate $NUF2$ expression. A) A LUTI mRNA is produced by the $NDC80^{LUTI}$-$NUF2$ fusion construct ($NUF2^{LUTI}$) in meiosis. To generate the $NDC80^{LUTI}$-$NUF2$ construct, the promoter and leader sequence of $NDC80^{LUTI}$ (1000 bps directly upstream of the $NDC80$ ORF start site) was placed immediately upstream of the $NUF2$ coding region. $NUF2^{LUTI}$ and $NUF2^{PROX}$ expression was detected by northern blot, and Nuf2 was detected by anti-V5 immunoblot. $SCR1$, loading control for northern blot. Hxk1, loading control for immunoblot. Samples were taken when the wild type (UB5103) and $NDC80^{LUTI}$-$NUF2$ (UB5101) cells were undergoing synchronous meiosis. * indicates a band of unknown origin. One of the two repeated experiments is shown. B) $NUF2$ signal was first normalized to $SCR1$. This normalized value was set to 1 for the time point collected 2 hours after entry into SPO (t2) and all the subsequent time points were then calibrated relative to t2. One of the two repeated experiments is shown. C) Quantification of Nuf2 protein abundance from the experiment shown in (A). For each time point, Nuf2 signal was first normalized to Hxk1. This normalized value was set to 1 for the 0 hour time point (t0), and all the subsequent time points were calibrated relative to t0.
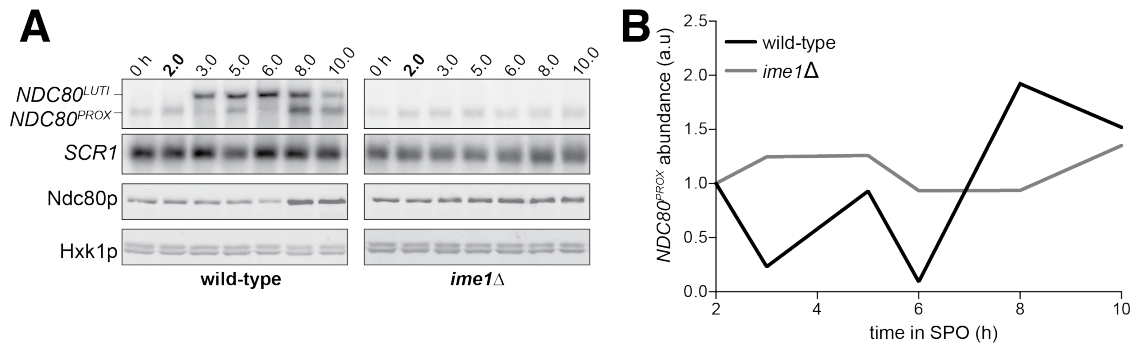
Figure 2.16: The meiosis-specific transcription factor Ime1 is necessary for $NDC80^{LUTI}$ expression. A) $NDC80^{PROX}$, $NDC80^{LUTI}$, and Ndc80 abundance during meiosis in $pCUP\text{-}IME1$ $pCUP\text{-}IME4$ (FW1902) and $pCUP\text{-}IME4$ $ime1\Delta$ (FW3058) cells. Expression from the $pCUP$ promoter was induced 2 hours after cells were transferred to SPO. One of the two repeated experiments is shown. B) $NDC80^{PROX}$ signal was first normalized to $SCR1$. The normalized value for the 2 hour time point (immediately prior to $IME1$ and $IME4$ induction) was set to 1, and all the subsequent time points were calibrated relative to this time point.
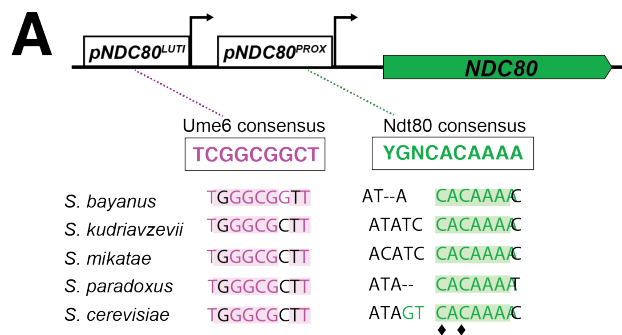
Figure 2.17: Meiotic regulators have conserved binding motifs in the $NDC80^{LUTI}$ and $NDC80^{PROX}$ promoters. Putative Ume6 (URS1) and Ndt80 (MSE) binding sites are present in the intergenic region upstream of *NDC80*. Colored bases match the consensus binding sequences. Highlighted areas indicate the conserved regions across all five Saccharomyces species by Clustal analysis (RRID:SCR_001591). The black diamonds indicate the two sites mutated from C to A in the ndc80-mse strain.
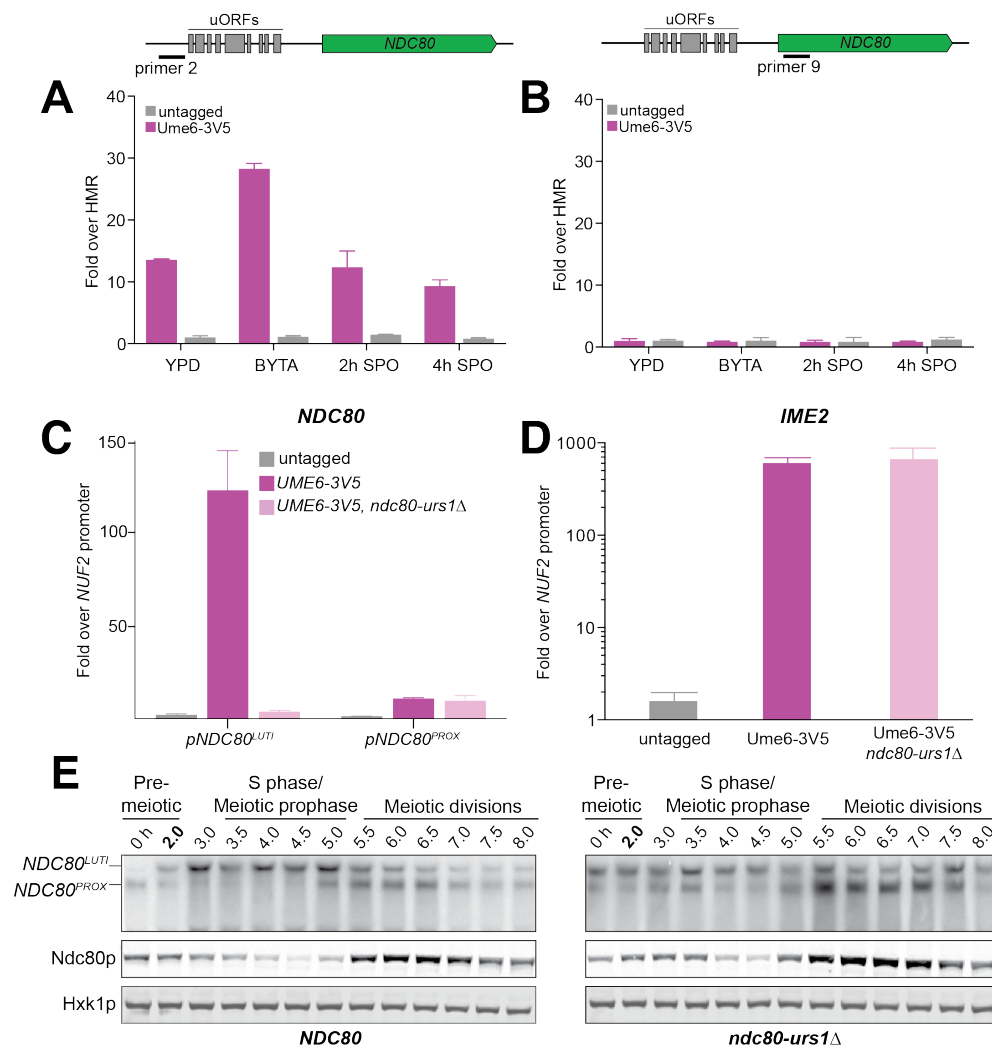
Figure 2.18: Ume6-3V5 chromatin immunoprecipitation of untagged (FW1511) and *UME6-3V5* (FW1208) strains from samples collected during exponential growth in YPD, during stationary phase in BYTA, and after transfer to SPO. The recovered DNA fragments from the Ume6-3V5 ChIP were quantified by qPCR using a primer set specific for the A) $NDC80^{LUTI}$ promoter and the B) *NDC80* coding region. Enrichment was normalized to *HMR*. The mean fold enrichment over *HMR* from three independent experiments, as well as the standard error of the mean, is displayed. C) Ume6-3V5 chromatin immunoprecipitation in untagged (UB2531), *UME6-3V5* (UB3301), and *UME6-3V5 ndc80-urs1Δ*(UB6760) strains. Cells were harvested from BYTA. The DNA fragments recovered from the Ume6-3V5 ChIP were quantified by qPCR using primer pairs specific for the $NDC80^{LUTI}$ promoter and the $NDC80^{PROX}$ promoter. Enrichment was normalized to the signal from the *NUF2* promoter. The mean fold enrichment three independent experiments, as well as the standard error of the mean, is displayed. D) Same as in C, but with primers specific to *IME2*. E) $NDC80^{PROX}$, $NDC80^{LUTI}$, and Ndc80 levels during synchronous meiosis (as described in Figure 2.2B) in wild type cells (UB6190) and *ndc80-urs1Δ*cells (UB6075).
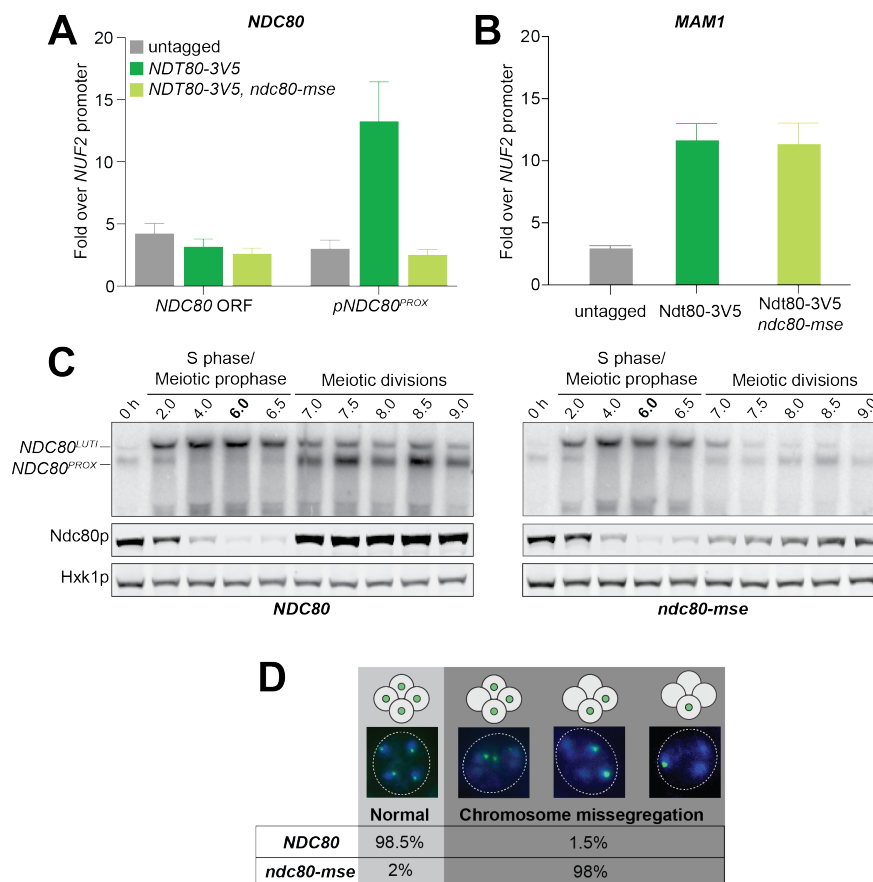
Figure 2.19: Ndt80 induces $NDC80^{PROX}$ before the meiotic divisions. A) Ndt80-3V5 chromatin immunoprecipitation in untagged (UB7997), *NDT80-3V5* (UB7999), and *NDT80-3V5 ndc80-mse* strains (UB7496). After 5 hours in SPO, *NDT80* expression was induced with β-estradiol. One hour after Ndt80 induction, cells were fixed with formaldehyde and chromatin extracts were prepared. The recovered DNA fragments were quantified by qPCR using two primer pairs: one specific for the $NDC80^{PROX}$ promoter ($pNDC80^{PROX}$) and one specific to the *NDC80* coding region (*NDC80* ORF). Enrichment at these loci was normalized to the signal from the *NUF2* promoter. The mean fold enrichment over the *NUF2* promoter from three independent experiments, as well as the standard error of the mean, is displayed. B) Same as A, but with primers for the promoter of *MAM1*. C) $NDC80^{PROX}$, $NDC80^{LUTI}$, and Ndc80 level during meiosis in wild type (UB4074) and *ndc80-mse* (UB3392) strains. Cells were transferred to SPO at 0 hour and released from pachytene arrest at 6 hours by addition of β-estradiol. D) Chromosome segregation accuracy in wild type (UB5876) and *ndc80-mse* (UB5437) strains was determined by counting homozygous CENV-GFP dots in tetranucleates. Samples were taken 7.5 hours after transfer to SPO. The fraction of tetranucleates that displayed normal segregation (one GFP dot in each nucleus), or missegregation (multiple or zero GFP dots in any of the four nuclei) was quantified. The average fraction of normal segregation or missegregation from two independent experiments is shown. Over 100 cells were counted per strain, per experiment.
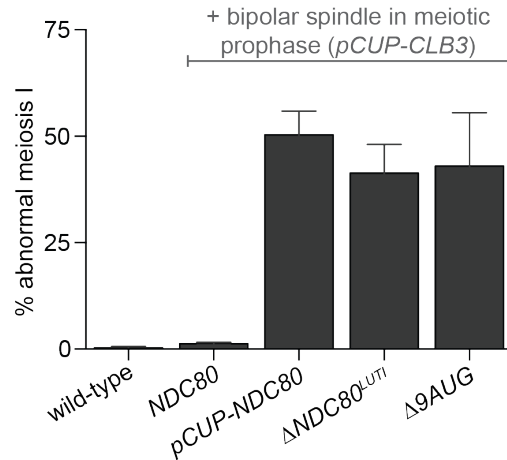
Figure 2.20: Temporal regulation of Ndc80 level by $NDC80^{LUTI}$ and $NDC80^{PROX}$ in meiosis is required for proper meiotic chromosome segregation. Sister chromatid segregation in wild type (UB2942), *pCUP-CLB3* (UB877), *pCUP-CLB3 pCUP-NDC80* (UB880), *pCUP-CLB3 ΔNDC80^{LUTI}* (UB2940), and *pCUP-CLB3 Δ9AUG* (UB2936) cells. Cells were induced to sporulate by transferring to SPO, and 6 hours later, expression of the cyclin Clb3 was induced by addition of $CuSO_4$. Immediately after induction, cells were released from pachytene by addition of β-estradiol. Samples were taken 1 hour 45 minutes after the release. Premature segregation of sister chromatids in meiosis I (abnormal meiosis I) was detected as two separated GFP dots in binucleates, one in each nucleus. The average fraction of binucleates that displayed sister segregation in meiosis I from three independent experiments, as well as the standard error of the mean, was graphed. 100 cells were counted per strain, per experiment.

identified at 184 bp upstream of the Ndc80 translation start site (Figure 2.17 and Figure 2.6), within the $NDC80^{PROX}$ promoter. One hour after Ndt80 expression was induced in the *pGAL-NDT80* GAL4-ER system, Ndt80 binding was enriched over the predicted MSE by ChIP analysis; moreover, mutations in the MSE (*ndc80-mse*) led to a complete loss of Ndt80 enrichment (Figure 2.19A), but did not affect another Ndt80 target gene *MAM1* (Figure 2.19B). Furthermore, the defect in Ndt80 binding to the $NDC80^{PROX}$ promoter reduced both $NDC80^{PROX}$ transcript and Ndc80 protein levels during the meiotic divisions (Figure 2.19C). This resulted in the inability of *ndc80-mse* cells to properly segregate their chromsomes (Figure 2.19D). These results demonstrate that Ndt80 directly induces $NDC80^{PROX}$ expression after meiotic prophase, and this timely induction of $NDC80^{PROX}$ elevates the levels of Ndc80 protein to those necessary for the meiotic divisions.
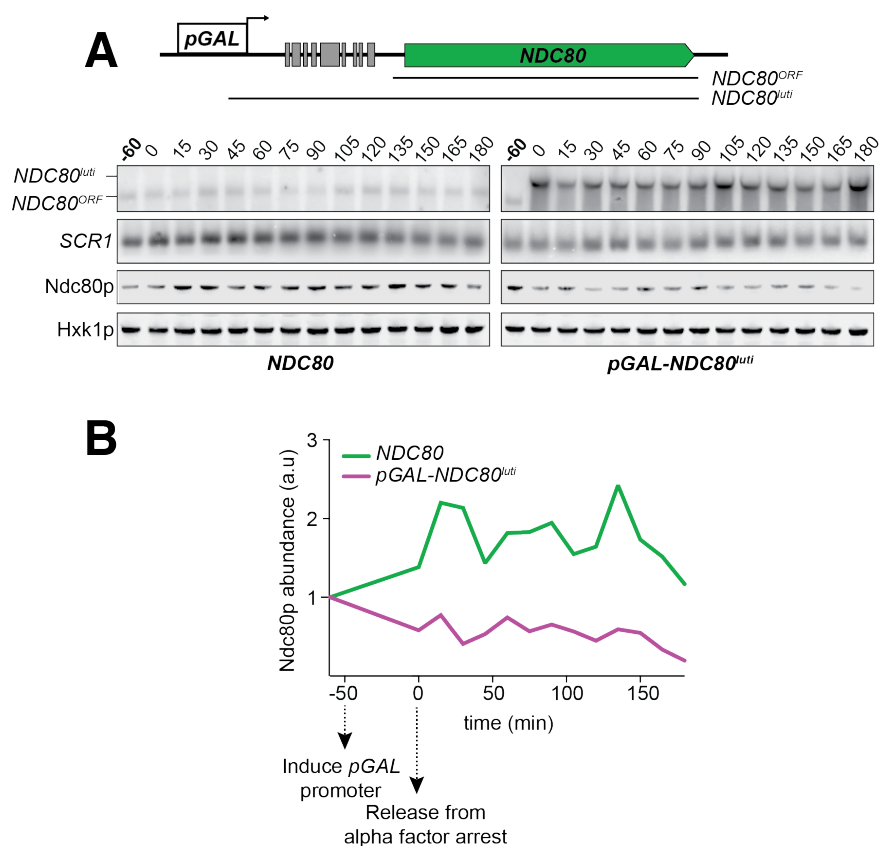
Figure 2.21: $NDC80^{LUTI}$ is sufficient to downregulate $NDC80^{PROX}$ in mitosis. A) $NDC80^{PROX}$, $NDC80^{LUTI}$, and Ndc80 levels when $NDC80^{LUTI}$ is expressed in synchronous mitosis. MATa wild type control (UB2389) and $pGAL$-$NDC80^{LUTI}$ (UB2388) cells, both harboring the Gal4-ER fusion protein, were arrested in G1 with -factor. $pGAL$ expression was induced 2 hours later by addition of β-estradiol (-60 min). One hour after the β-estradiol addition (0 min), cells were released from G1 arrest. One of the two repeated experiments is shown. B) Quantification of Ndc80 abundance from the experiment shown in A. For each time point, Ndc80 signal was first normalized to Hxk1. This normalized value was set to 1 for the first time point at -60 minutes (t-60, the time of β-estradiol addition) and all the subsequent time points were then calibrated relative to t-60.
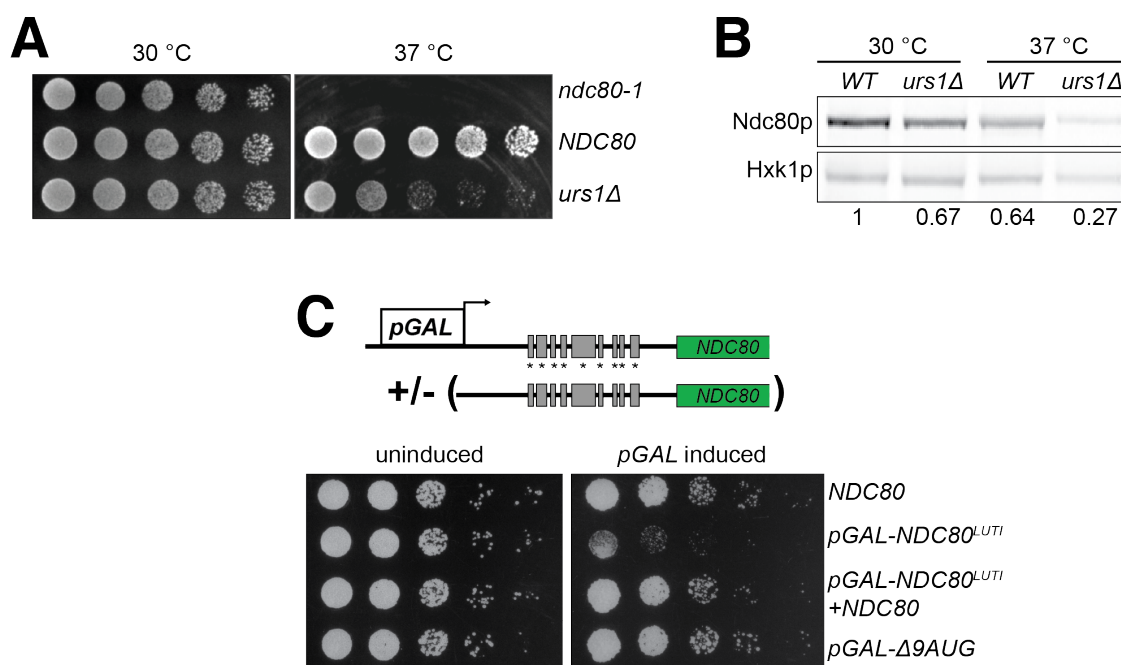
Figure 2.22: Misexpression of $NDC80^{LUTI}$ outside of meiosis causes severe growth defects. A) Growth phenotype of *ndc80-urs1Δ* cells at 30 °C and 37 °C. Temperature-sensitive *ndc80-1* (UB494), wild type (UB3262), and *urs1Δ* (UB4212) cells were serially diluted and grown on nutrient rich medium (YPD) plates at 30 °C or 37 °C for 2 days. (B) Ndc80 level in wild type (UB3262) and *urs1Δ* (UB4212) cells grown at 30 °C or 37 °C. For each condition, equal OD_{600} of cells were taken, and Ndc80 was visualized by anti-V5 immunoblot. Hxk1, loading control. WT, wild type. The number under each lane is the ratio of the relative Ndc80 levels (normalized to Hxk1 levels) compared with that of wild type at 30 °C. The results of one representative repeat from two independent experiments are shown. C) Growth phenotype of haploid control (UB1240), *pGAL-NDC80^{LUTI}* (UB1217), *pGAL-NDC80^{LUTI}* with a second copy of *NDC80* at the *LEU2* locus (UB8001), and *pGAL-Δ9AUG* (UB1323). Cells were serially diluted and grown on YEP-raffinose/galactose (YEP-RG) plates (uninduced) or YEP-RG plates supplemented with β-estradiol (*pGAL* induced) at 30 °C for 2 days.
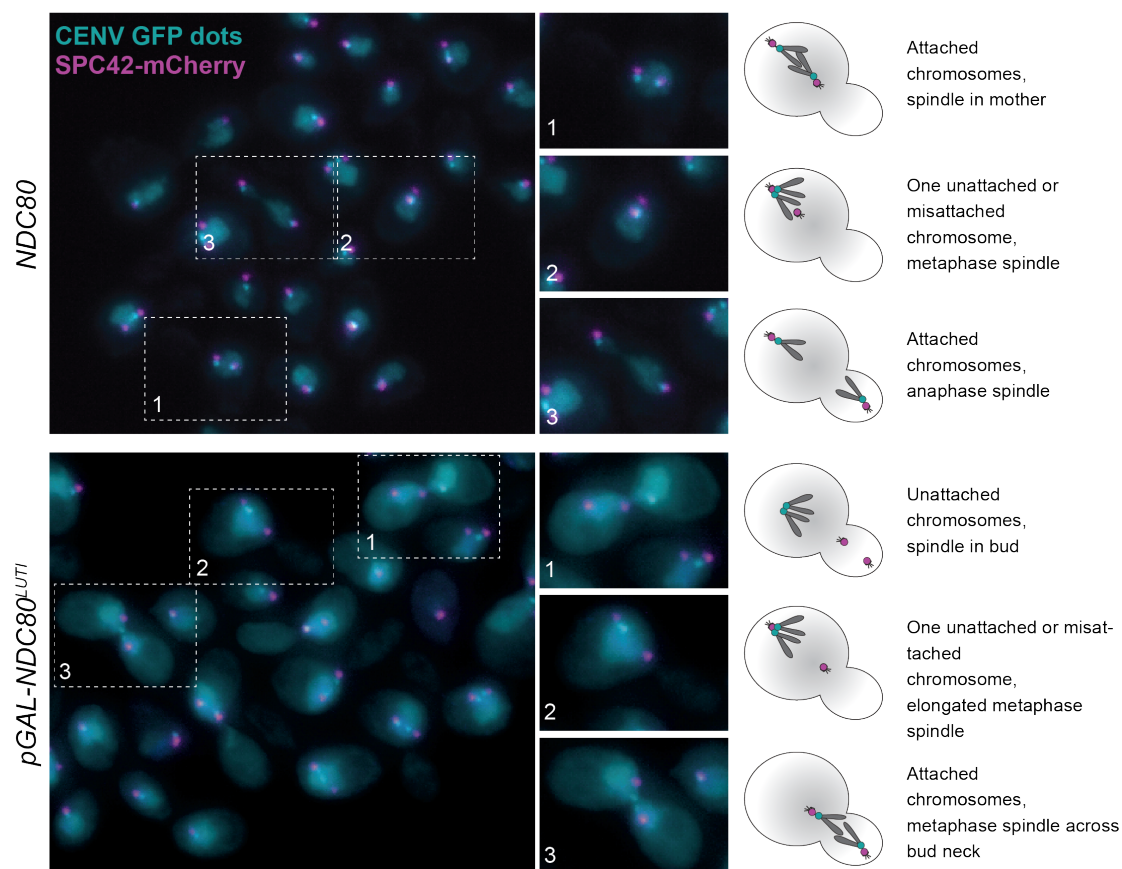
Figure 2.23: Abnormal kinetochore-chromosome attachment upon $NDC80^{LUTI}$ misexpression. Phenotypic characterization of cells expressing $pGAL$-$NDC80^{LUTI}$. Both the control (UB8682) and $pGAL$-$NDC80^{LUTI}$ (UB8684) cells harbor homozygous CENV-GFP dots and Spc42-mCherry (spindle pole body marker). The strains were grown overnight in YEP-RG, and samples were collected at 0 hour and 6 hours after $pGAL$ induction by β-estradiol. Representative images of wild type cells and the cells expressing $NDC80^{LUTI}$ after 6 hours of $pGAL$ induction. Enlarged images of the boxed regions are shown in the middle. To the right are schematics of the microtubule-kinetochore attachment status in each class of phenotype observed.
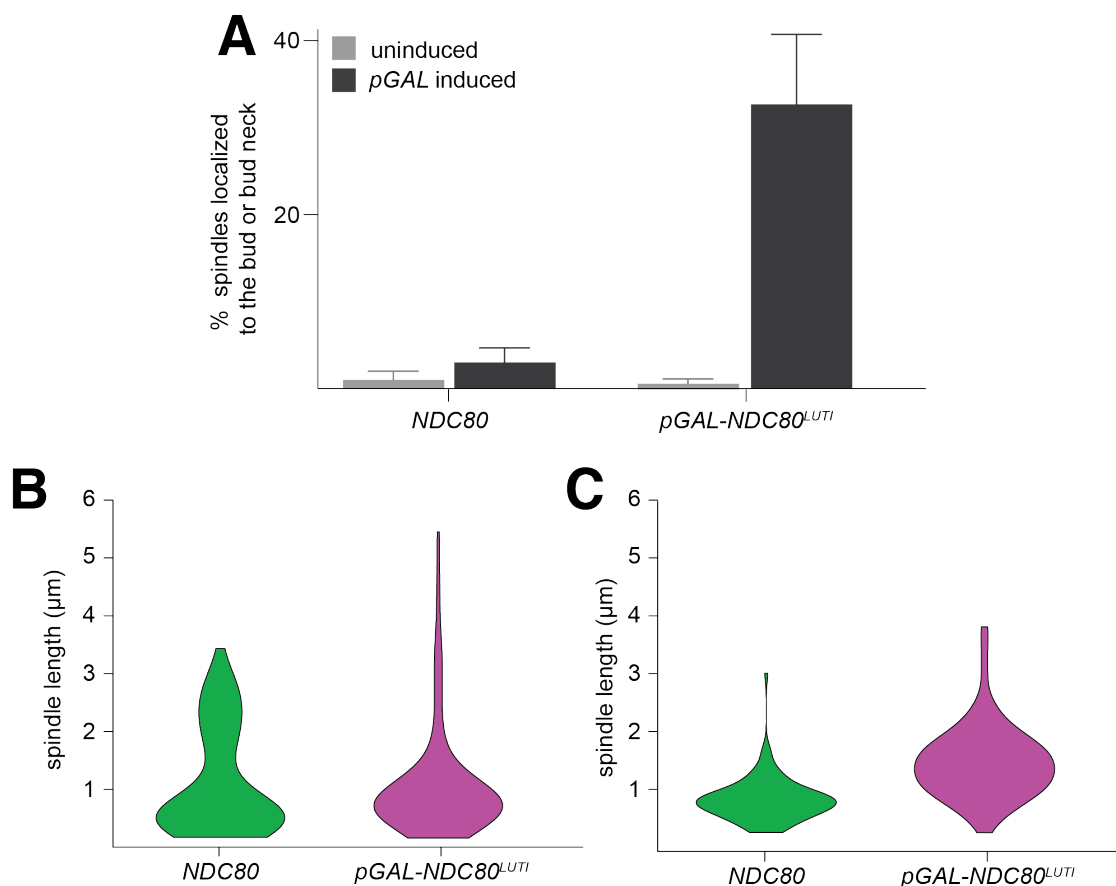
Figure 2.24: Quantification of kinetochore disfunction upon $NDC80^{LUTI}$ misexpression. Quantification of the spindle localization data shown in 2.23. A) Among the cells with separated spindle poles, the percentage of cells that had a spindle shorter than 2 μm and were abnormally localized (i.e. across the bud neck or entirely within the bud) is displayed. 100 cells were counted per strain, for each condition. The average percentage and the standard deviation from three independent experiments are shown. B) Quantification of spindle length in control (UB8682) and $pGAL\text{-}NDC80^{LUTI}$ (UB8684) cells harboring Spc42-mCherry (spindle pole body marker). Cells were grown overnight in YEP-raffinose/galactose (YEP-RG). $pGAL$ was induced with β-estradiol at time 0 hour. Spindle lengths were measured in all the cells with separated spindle poles. A representative replicate out of three independent experiments was graphed as a violin plot. 100 cells were analyzed per strain, per experimental replicate. C) Quantification of the spindle length in cells with at least one chromosome V not attached to a spindle pole body. Cells with either both CENV-GFP dots associated with a single spindle pole or both CENV-GFP dots completely dissociated from either spindle pole. This allows analysis of populations of cells that are either in S-phase/early mitosis (after SPB duplication, but before chromosome alignment) or are unable to properly attach their chromosomes. A representative replicate out of three independent experiments was graphed as a violin plot. 100 cells were analyzed per strain, per replicate.

Temporal regulation of $NDC80^{LUTI}$ and $NDC80^{PROX}$ expression is essential for the proper timing of kinetochore function

Since Ndc80 appears to be the limiting subunit of the kinetochore, we posited that the regulated expression of $NDC80^{LUTI}$ and $NDC80^{PROX}$ serves to inactivate and reactivate kinetochores, respectively, through modulating Ndc80 protein levels. In budding yeast, kinetochores are inactive in meiotic prophase (Miller *et al.* 2012), but they can be activated upon Ndc80 overexpression (Miller *et al.* 2012). We asked whether functional kinetochores could also be generated in meiotic prophase if cells failed to express $NDC80^{LUTI}$ ($\Delta NDC80^{LUTI}$) or expressed a version of $NDC80^{LUTI}$ that could translate Ndc80 protein ($\Delta 9AUG$). Both conditions caused an increase in Ndc80 levels in meiotic prophase (Figure 2.5B and 2.7A-B). Further, we observed that over 50% of the $\Delta NDC80^{LUTI}$ or $\Delta 9AUG$ cells in a sensitized strain background displayed abnormal chromosome segregation in meiosis I (Figure 2.20A), suggesting premature kinetochore activity in meiotic prophase. The extent of this phenotype was indistinguishable from that when Ndc80 was overexpressed in meiotic prophase (pCUP-NDC80) (Figure 2.20A). Therefore, repression of $NDC80^{PROX}$ by $NDC80^{LUTI}$ transcription is crucial to inhibit untimely kinetochore function during meiotic prophase.

Unlike $NDC80^{PROX}$ transcript, $NDC80^{LUTI}$ is absent in vegetative growth due to repression by Ume6 (Figure 2.18E). We hypothesized that $NDC80^{LUTI}$ is repressed during the mitotic cell cycle because its expression could inactivate kinetochore function. Indeed, when the Ume6 repressor-binding site within the $NDC80^{LUTI}$ promoter was deleted ($urs1\Delta$), these cells grew similar to wild type cells at 30 °C, but they had a severe growth defect at 37 °C due to reduced Ndc80 levels (Figure 2.22A and 2.22B). Thus, the repression of $NDC80^{LUTI}$ by Ume6 is critical for the fitness of mitotically dividing cells.

When $NDC80^{LUTI}$ was strongly induced in vegetative growth using the inducible *GAL1-10* promoter, these cells had a severe growth defect and reduced Ndc80 (Figure 2.22C, Figure 2.21). In wild type cells synchronously progressing through the mitotic cell cycle, a single mRNA isoform, $NDC80^{PROX}$, was present at all stages (Figure 2.21A, left panel). In contrast, the $NDC80^{PROX}$ transcript became undetectable in $pGAL-NDC80^{LUTI}$ cells one-hour after $NDC80^{LUTI}$ induction (Figure 2.21A, right panel). Four hours after induction, Ndc80 protein levels were reduced to 20% of the initial level, while in wild type cells it was increased to 116% (Figure 2.21B). Based on these data, we conclude that $NDC80^{LUTI}$ expression is sufficient to repress $NDC80^{PROX}$ outside of meiosis. The reduction in $NDC80^{PROX}$ expression, in turn, causes reduced synthesis of Ndc80 protein, thus essentially turning off the *NDC80* gene. This defect was rescued by a second copy of *NDC80* at an ectopic locus, consistent with the notion that $NDC80^{LUTI}$-mediated repression of $NDC80^{PROX}$ occurs in *cis* (Figure 2.22C and Figure 2.9). Induction of the uORF-free $NDC80^{LUTI}$ ($\Delta 9AUG$) caused no appreciable growth defect (Figure 2.22C), consistent with the observation that the $\Delta 9AUG$ cells could express Ndc80 protein (Figure 2.5).
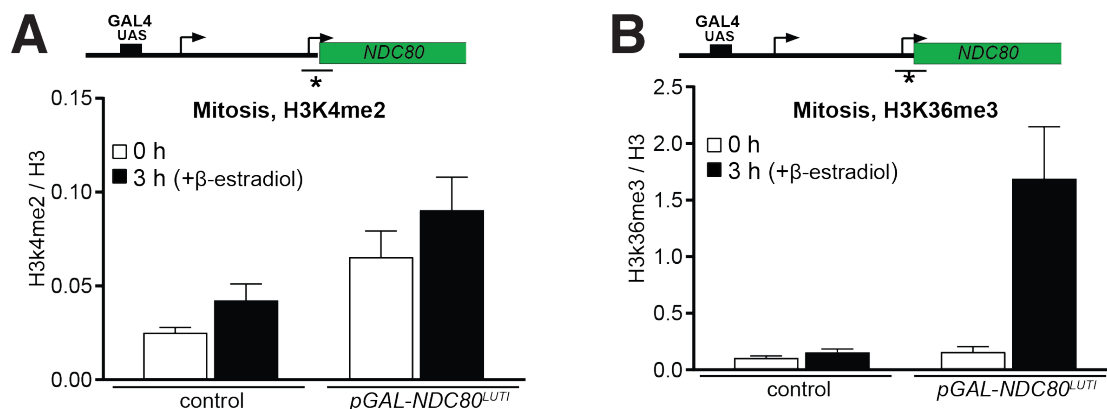
Figure 2.25: H3K4me2 and H3K36me3 during LUTI induction in mitosis. A) Transcription of $NDC80^{LUTI}$ during mitotic growth leads to increased H3K4me2 levels in the $NDC80^{PROX}$ promoter. Control cells (UB91) and cells harboring $NDC80^{LUTI}$ driven by the *GAL1-10* promoter ($pGAL\text{-}NDC80^{LUTI}$) (UB3338) were grown to exponential phase in rich medium with raffinose and galactose (YP-RG). Subsequently, cells were treated with β-estradiol to activate the *GAL1-10* promoter. Samples were taken at 0 and 3 hours after induction of $NDC80^{LUTI}$ for ChIP. H3K4me2 and histone H3 levels were determined using a primer pair directed against the $NDC80^{PROX}$core promoter. The mean enrichment from three independent experiments plus the standard error of the mean are displayed. B) Similar to A except that H3K36me3 levels were determined by ChIP.

The inducible nature of the *GAL1-10* promoter allowed us to directly test whether the growth defect associated with the mitotic $NDC80^{LUTI}$ expression arose from defects in kinetochore function. We performed fluorescence microscopy to track spindle length (Spc42-mCherry) and chromosome segregation (CENV-GFP dots). Cells expressing $NDC80^{LUTI}$ displayed a range of kinetochore-microtubule attachment defects (Figure 2.23, bottom panel). In cells with separated spindle pole bodies, 30% of the cells expressing $NDC80^{LUTI}$ had metaphase spindles (≤ 2 μm) improperly localized to either the bud or the bud neck, whereas only 3% of the wild type cells displayed this phenotype (Figure 2.24A). Furthermore, in cells expressing $NDC80^{LUTI}$, an abnormal distribution of spindle length was observed, characteristic of a metaphase arrest (Figure 2.24B). Spindle elongation was also observed prior to chromosome capture, suggesting improper kinetochore function (Figure 2.24C). Collectively, these analyses revealed that the strict temporal regulation of $NDC80^{LUTI}$ and $NDC80^{PROX}$ transcription in both mitosis and meiosis is essential to ensure the proper timing of kinetochore function and high fidelity chromosome segregation.

$NDC80^{LUTI}$-dependent downregulation $NDC80^{PROX}$ requires Set2 and Set3 in mitosis

To investigate if the mitotic and meiotic methods of LUTI-based gene repression are conserved at the *NDC80* locus, we examined whether deposition of the H3K36me3 and H3K4me2
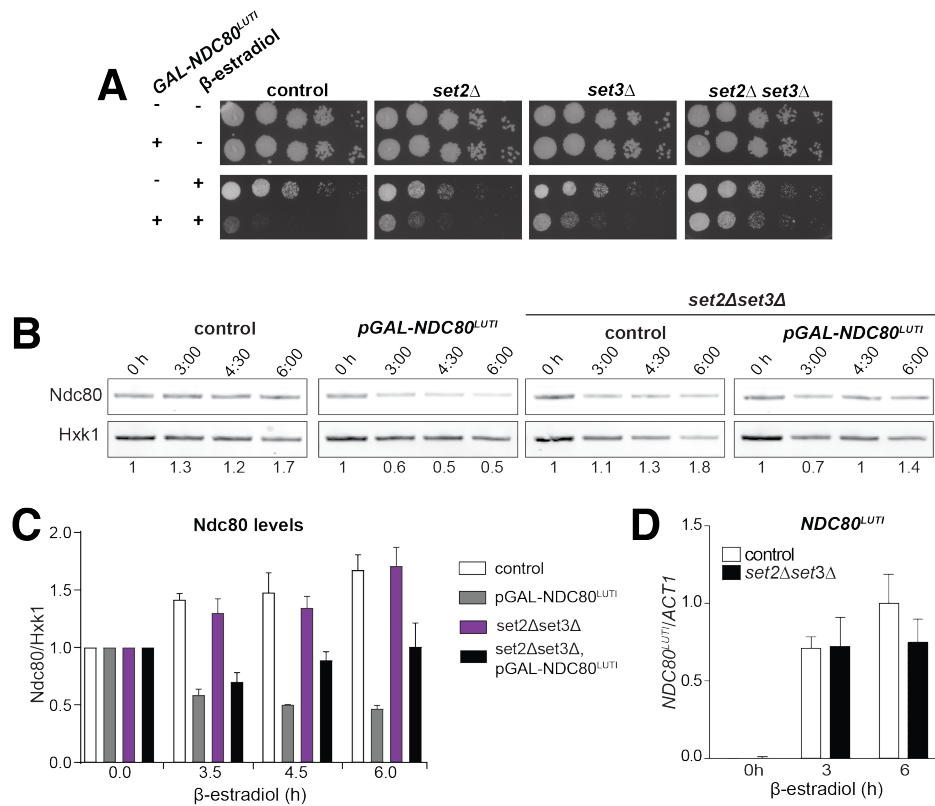
Figure 2.26: A) The lethality of ectopic $NDC80^{LUTI}$ expression is rescued by $set2\Delta set3\Delta$. Spot assays of control cells, which harbor a wild-type $NDC80$ locus, with *SET2 SET3* (UB1252), $set2\Delta$ (UB3545), $set3\Delta$ (UB3547), or $set2\Delta set3\Delta$ (UB3549); as well as cells expressing $NDC80^{LUTI}$ from the heterologous *GAL* promoter ($pGAL$-$NDC80^{LUTI}$) with *SET2 SET3* (UB1218), $set2\Delta$ (UB1236), $set3\Delta$ (UB1237), and $set2\Delta set3\Delta$ (UB1235). These cells also expressed Gal4 fused to estrogen receptor (Gal4-ER) to activate the *GAL1-10* promoter. Cells were grown overnight on YP-glycerol plates and spotted on YP + raffinose + galactose (YP-RG) plates in the absence or presence of β-estradiol. B) Immunoblot of Ndc80 for *SET2 SET3* cells harboring an *NDC80-V5* allele (UB1240) or harboring a $pGAL$-$NDC80^{LUTI}$-*V5* allele (UB1217), as well as $set2\Delta set3\Delta$ cells with *NDC80-V5* (UB8110) or $pGAL$-$NDC80^{LUTI}$-*V5* (UB8114). Cells were grown in YP-RG and treated with β-estradiol. Ndc80 protein levels were determined with anti-V5 antibodies. Hxk1 levels were detected with anti-hexokinase antibodies. C) Quantification of B. Ndc80 signals were first normalized to Hxk1 and were then normalized with respect to the 0 hour time point. The mean from three independent experiments plus the standard error of the mean are displayed. D) Levels of $NDC80^{LUTI}$ are not affected in the $set2\Delta set3\Delta$ double mutant. Control (UB1217) and $set2\Delta set3\Delta$ (UB8114) cells harboring $NDC80^{LUTI}$ driven by the *GAL1-10* promoter were grown in YP-RG medium and treated with β-estradiol during exponential growth. RNA was extracted, reverse transcribed, and the $NDC80^{LUTI}$ mRNA levels were determined by qPCR. Signals were normalized to *ACT1*. The mean from three independent experiments plus the standard error of the mean are displayed.

was observed upon $NDC80^{LUTI}$ induction in mitosis. Using the $pGAL\text{-}NDC80^{LUTI}$ construct, $NDC80^{LUTI}$ was expressed in rich nutrient conditions in cells harboring the Gal4-ER chimeric transcription factor, which responds to β-estradiol. We observed a moderate increase of H3K4me2 in cells that harbored $NDC80^{LUTI}$ compared to control cells, which was independent of $NDC80^{LUTI}$ transcription (Figure 2.25A). One explanation is that the *GAL1-10* promoter is leaky and can increase H3K4me2 levels without induction with β-estradiol. In contrast to H3K4me2, H3K36me3 levels were strongly enriched in the $NDC80^{PROX}$ promoter when $NDC80^{LUTI}$ was induced (Figure 2.25B). Control cells harboring the wild-type $NDC80^{LUTI}$ promoter did not show increased H3K36me3 levels. We conclude that the deposition of H3K36me3, but not H3K4me2, is independent of the identity of the promoter that directs $NDC80^{LUTI}$ transcription. Taken together, deposition of repressive chromatin marks in the $NDC80^{PROX}$ promoter requires $NDC80^{LUTI}$ transcription.

We also tested whether Set2 and Set3 are required for repression of $NDC80^{PROX}$ when $NDC80^{LUTI}$ is expressed ectopically during vegetative growth. Although $NDC80^{LUTI}$ transcription was lethal in wild-type control cells, there was a partial growth rescue in either $set2\Delta$ or $set3\Delta$ cells. Importantly, growth was restored almost completely in $set2\Delta set3\Delta$ (Figure 2.26A), due to de-repression of Ndc80 protein expression despite elevated transcription of the $NDC80^{LUTI}$ mRNA (Figure 2.26B-D). Altogether, these results show that both Set2 and Set3 are necessary for efficient repression of the $NDC80^{PROX}$ promoter by $NDC80^{LUTI}$ transcription during mitosis, and that the LUTI mRNA-mediated mode of transcription repression outlined above is not restricted to meiosis.

Gene repression by $NDC80^{LUTI}$ transcription is tunable

Work from *Escherichia coli* showed that gene regulation by transcriptional interference is not binary with an on or off state, but can be utilized to fine-tune gene expression levels (Bordoy *et al.* 2016; Hao *et al.* 2016). The work prompted us to investigate whether transcriptional interference by $NDC80^{LUTI}$ could also be tunable, thus enabling incremental changes in $NDC80^{PROX}$ expression levels. To scale the level of $NDC80^{LUTI}$ expression, we used a tightly controlled, inducible system. The system utilizes a heterologous, chimeric transcriptional activator (LexA-ER-AD) whose activity is induced in a concentration-dependent manner by β-estradiol (Ottoz *et al.* 2014). Varying the number of LexA-binding sites (lexO) in the $NDC80^{LUTI}$ promoter and titrating the concentration of β-estradiol, enabled scalable transcriptional induction of $NDC80^{LUTI}$ where the growth defect caused by $NDC80^{LUTI}$ expression in mitosis was more severe with elevated concentrations of β-estradiol and higher number of lexO sites in the $NDC80^{LUTI}$ promoter (Figure 2.27A-B). The higher the LUTI transcription, the greater the inhibition of $NDC80^{PROX}$ expression. Thus, modulating $NDC80^{LUTI}$ transcription levels allows scalable transcriptional repression of $NDC80^{PROX}$ in a population of cells.

During transcription nucleosomes are disassembled and reassembled by histone chaper-
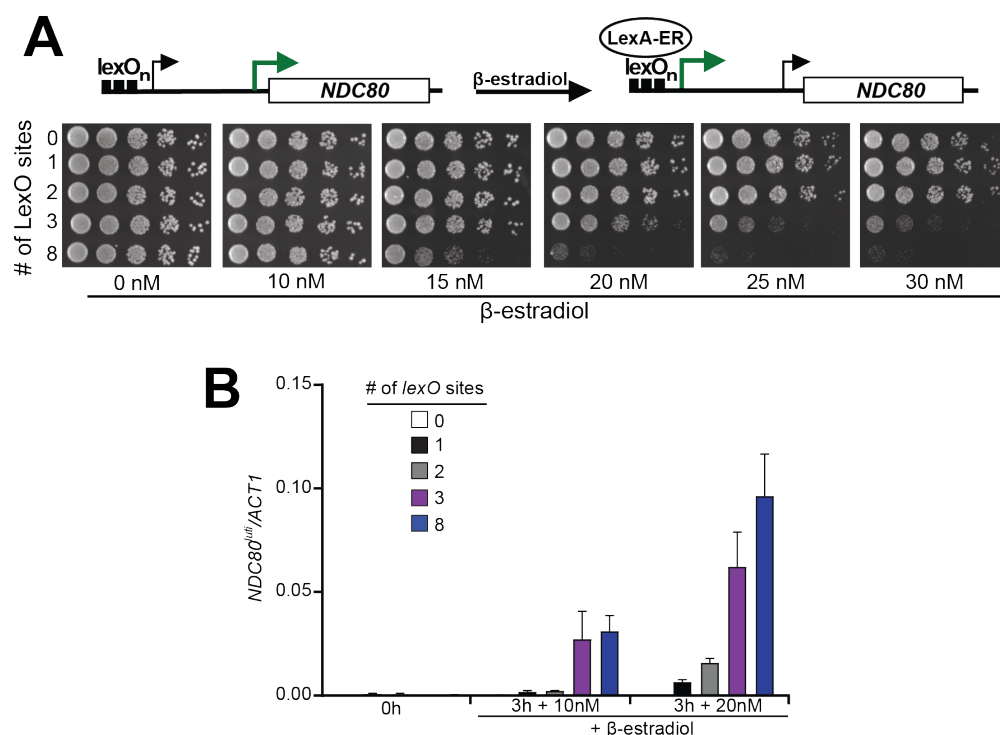
Figure 2.27: Gene repression by $NDC80^{LUTI}$ is tunable. A) Adjustable expression of $NDC80^{LUTI}$ using the LexA-lexO system. Spot assay of cells harboring 0, 1, 2, 3, or 8 lexO binding sites in the $NDC80^{LUTI}$ promoter (UB8374, UB8358, UB8362, UB8366, UB8370) in the presence of different concentrations of β-estradiol. These cells also expressed LexA fused to an activation domain (AD) and the human estrogen receptor (ER) (LexA-ER-AD). Cells were spotted on YPD plates in the absence or presence of different concentrations of β-estradiol. B) $NDC80^{LUTI}$ levels in the presence of variable number of lexO sites in the $NDC80^{LUTI}$ promoter. Cells harboring 0, 1, 2, 3, or 8 lexO and LexA-ER-AD (UB8374, UB8358, UB8362, UB8366, and UB8370) were grown in YPD overnight. Subsequently, cells were diluted and exponentially growing cells were treated with 10 or 20 nM β-estradiol for 3 hours. RNA was extracted, reverse transcribed, and $NDC80^{LUTI}$ mRNA levels were determined by quantitative PCR. Signals were normalized to *ACT1*. The mean from three independent experiments plus the standard error of the mean are displayed.
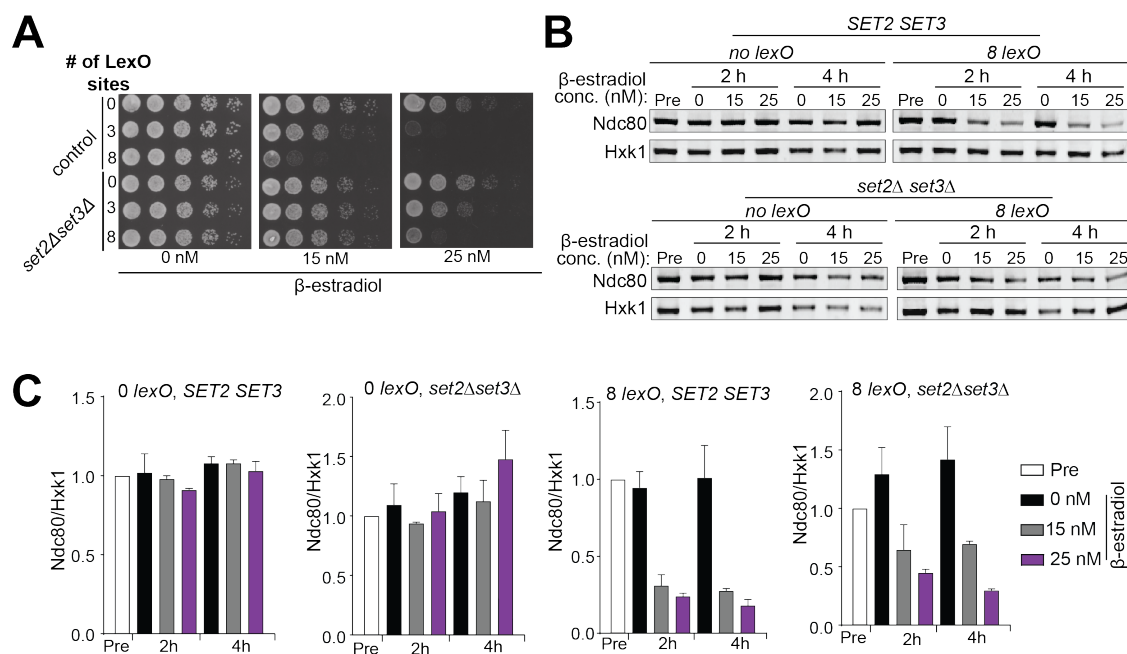
Figure 2.28: Massive over-expression of $NDC80^{LUTI}$ leads to $SET2$ and $SET3$ independent repression A) Adjustable expression of $NDC80^{LUTI}$ using the LexA-lexO system in $set2\Delta set3\Delta$ cells. Spot assay of cells harboring 0, 3, or 8 lexO binding sites at the $NDC80^{LUTI}$ promoter in $SET2$ $SET3$ (UB8374, UB8366, UB8370) or $set2\Delta set3\Delta$cells (UB8691, UB8686 and UB8693). These cells also expressed LexA fused to an activation domain (AD) and the human estrogen receptor (ER) (LexA-ER-AD). Cells were spotted on YPD plates in the absence or presence of different concentrations of β-estradiol. B) Ndc80 protein level in $SET2$ $SET3$ cells harboring none (UB12945) or 8 lexO sites (UB12949), or $set2\Delta set3\Delta$cells with none (UB12947) or 8 lexO sites (UB12951). All four strains carry LexA-ER-AD. Ndc80 protein was detected by anti-V5 immunoblot. Hxk1 levels were used as a loading control. Pre denotes pre-induction. Exponentially growing cells were treated with ethanol, 15 nM, or 20 nM β-estradiol. Samples were taken at 2 hours or 4 hours after β-estradiol induction. Ndc80 level was normalized to Hxk1 level, and the number under each lane shows the Ndc80/Hxk1 ratio normalized to that in the pre-induction condition. C) Quantification of B. The mean of two independent experiments plus standard error of the mean are displayed.

ones that associate with RNA polymerase (Venkatesh & Workman 2015). Therefore, higher levels of $NDC80^{LUTI}$ transcription could lead to an increased rate of nucleosome deposition in the $NDC80^{PROX}$ promoter and thus scalable $NDC80^{PROX}$ repression. If so, then sufficiently high levels of $NDC80^{LUTI}$ transcription should be sufficient for repressing $NDC80^{PROX}$ without requiring Set1/Set3C and Set2/Rpd3S to maintain repressive chromatin. Cells with both pathways compromised (*set2Δset3Δ*) and harboring three or eight lexO sites did not show a growth defect when exposed to intermediate levels of β-estradiol (15 nM), whereas control cells did (Figure 2.28A). This result was expected because in the *set2Δset3Δ* mutant background $NDC80^{LUTI}$-mediated repression is impaired (also see Figure 2.26). Surprisingly at higher concentrations of β-estradiol (25 nM), *set2Δset3Δ* mutant cells harboring three lexO sites exhibited a moderate growth defect while cells with eight lexO sites exhibited a severe growth defect. We also measured the Ndc80 protein levels in control and *set2Δset3Δ* mutant cells harboring 0 or 8 copies of lexO sites. The growth defects observed in Figure 2.28A were reflected in the Ndc80 protein levels (Figure 2.28B-C). These data suggest that high levels of $NDC80^{LUTI}$ transcription could bypass the requirement for Set2 and Set3 in $NDC80^{PROX}$ repression.

Since increased expression of $NDC80^{LUTI}$ leads to stronger repression of $NDC80^{PROX}$, we tested whether the strength of the $NDC80^{PROX}$ promoter influenced the effectiveness of $NDC80^{LUTI}$-mediated repression. To examine this, we increased the levels of $NDC80^{PROX}$ by ectopically expressing the meiotic transcription factor Ndt80 in mitotic cells. Ndt80 induces the expression of $NDC80^{PROX}$ via the middle sporulation element (MSE) in the $NDC80^{PROX}$ promoter. In the presence of Ndt80 expression, the growth defect caused by $NDC80^{LUTI}$ transcription was suppressed (Figure 2.29A). This suppression is dependent on the presence of the MSE site in the $NDC80^{PROX}$ promoter (Figure 2.29A, compare MSE positive versus negative in the presence of *pGAL-NDT80* and *pGAL-NDC80^{LUTI}*). In addition, the growth changes as observed in the spot assays correlated well with Ndc80 protein levels (Figure 2.29B-C). Thus, increased transcription from the $NDC80^{PROX}$ promoter can bypass $NDC80^{LUTI}$ mediated repression. Taken together, transcriptional interference by expression of a 5'-extended transcript can be tuned by adjusting the relative strengths of the distal and proximal promoters. Hence, this mechanism can be adapted as a regulatory module to generate a range of gene expression outputs.

2.3 Discussion

In this study, we have identified an integrated regulatory circuit that controls the inactivation and subsequent reactivation of the meiotic kinetochore (Figure 2.30). This circuit controls the synthesis of a kinetochore subunit, Ndc80, and relies on the regulated expression of two distinct *NDC80* mRNAs. A meiosis-specific switch in promoter usage induces the expression of a 5' extended transcript isoform, $NDC80^{LUTI}$, which itself cannot produce Ndc80 protein. Rather, its function is purely regulatory. Transcription of this alternate
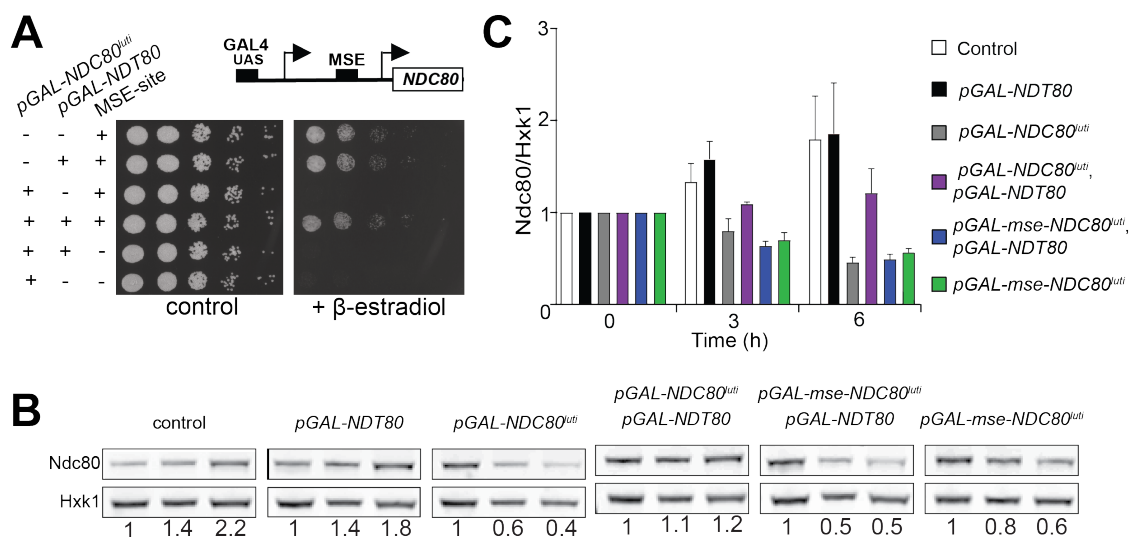
Figure 2.29: Increased $NDC80^{PROX}$ promoter activity bypasses $NDC80^{LUTI}$ mediated repression. A) Cells were spotted on YP + raffinose + galactose plates in the absence or presence of β-estradiol (1μM). For the analyses, we used three sets of strains: 1) Cells with a wild-type $NDC80$ and with a functional MSE site (+ MSE), but with either a wild-type $NDT80$ (UB3351) or a $pGAL1$-10 driven $NDT80$ ($pGAL$-$NDT80$, UB3370); 2) cells with $pGAL$-$NDC80^{LUTI}$and with a functional MSE site, along with either a wild-type $NDT80$ (UB5154) or $pGAL$-$NDT80$ (UB9181); 3) cells with $pGAL$-$NDC80^{LUTI}$and a non-functional MSE site, along with either $pGAL$-$NDT80$ (UB9921) or wild-type $NDT80$ (UB9923). These cells also expressed Gal4-ER to activate $pGAL$ driven expression. B) Exponentially growing cells of the strains in A were treated with ethanol or 1μM β-estradiol. Samples were taken at 3 hours or 6 hours after β-estradiol induction. The amount of samples loaded corresponded to the same OD_{600} across all the cultures. Ndc80 level was normalized to Hxk1 level, and the number under each lane shows the Ndc80/Hxk1 ratio normalized to that in the pre-induction condition. C) Quantification of Ndc80 protein levels from B. The Ndc80 protein levels were normalized to Hxk1 protein abundance. The relative levels with respect to the 0 hour time point are displayed. The mean from three independent experiments plus the standard error of the mean are displayed.
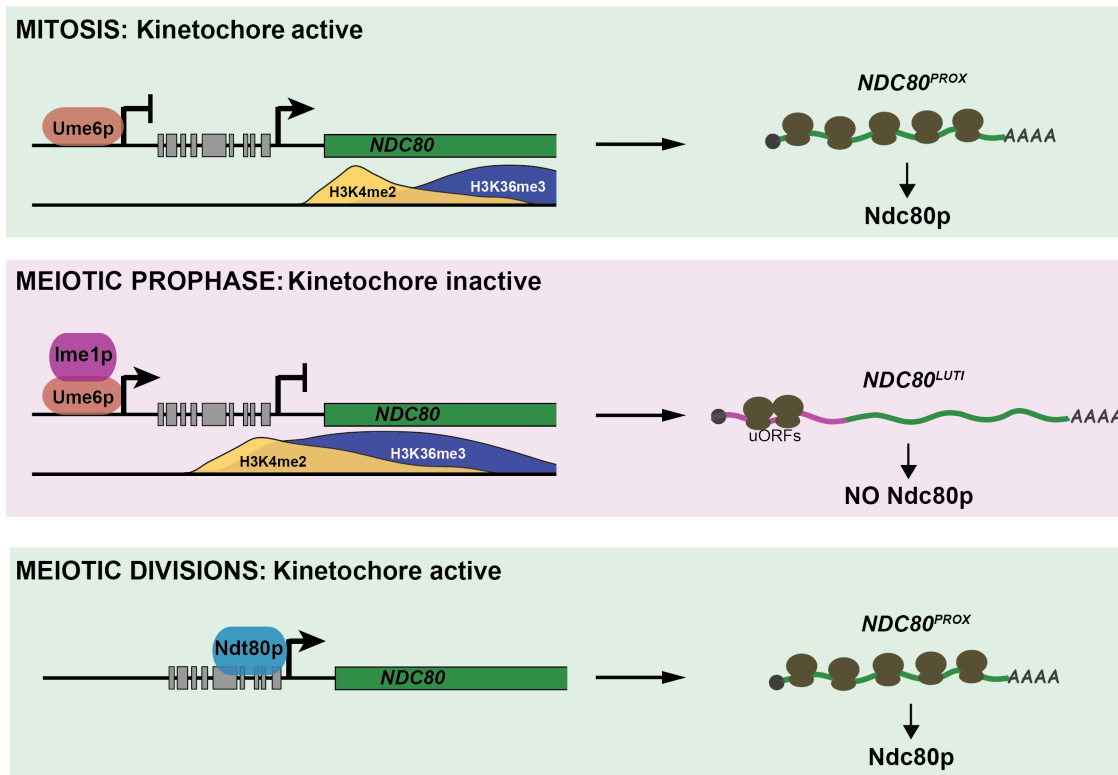
Figure 2.30: Model of *NDC80* gene regulation in budding yeast. During vegetative growth, a stage in which kinetochores are active, a short *NDC80* mRNA isoform $NDC80^{PROX}$ is expressed, and the 5' extended isoform $NDC80^{LUTI}$ is repressed by Ume6. Translation of $NDC80^{PROX}$ results in Ndc80 protein synthesis (top panel). At meiotic entry, the master transcription factor Ime1 induces expression of $NDC80^{LUTI}$. Transcription from this distal $NDC80^{LUTI}$ promoter silences the proximal $NDC80^{PROX}$ promoter through a mechanism that increases H3K4me2 and H3K36me3 marks over the $NDC80^{PROX}$ promoter. $NDC80^{LUTI}$ does not support Ndc80 synthesis due to translation of the uORFs. The overall synthesis of Ndc80 is repressed in meiotic prophase, and the kinetochores are inactive (middle panel). As cells enter the meiotic divisions, the transcription factor Ndt80 induces $NDC80^{PROX}$ re-expression, allowing for Ndc80 re-synthesis and formation of active kinetochores (bottom panel).

---

isoform leads to repression of the protein-translating $NDC80^{PROX}$ isoform in *cis*. This results in inhibition of Ndc80 protein synthesis and ultimately the inactivation of kinetochore function in meiotic prophase. Reactivation of the kinetochore is achieved by the transcription of $NDC80^{PROX}$ upon exiting meiotic prophase. Temporally coordinated by two master transcription factors, the timely expression of these two mRNA isoforms is essential for kinetochore function, accurate chromosome segregation, and gamete viability. Altogether, our study describes a new gene regulatory mechanism and provides insight into its biological purpose.

$NDC80^{LUTI}$ is an mRNA that does not produce protein

A key aspect of the work presented here is the surprising finding that an mRNA can serve a purely regulatory function. Indeed, $NDC80^{LUTI}$ is a bona fide mRNA. Despite being poly-adenylated and engaged by the ribosome, the ORF within $NDC80^{LUTI}$ cannot be decoded due to the presence of AUG-uORFs in its extended 5'-leader. By competitively engaging with the ribosome, these uORFs prevent translation of Ndc80 protein. The polypeptides that the uORFs encode are unlikely to play a role in the repression of kinetochore function as the uORFs can be minimized to 2-codon units while maintaining $NDC80^{LUTI}$-based re-pression (Figure 2.5B). Interestingly, upstream AUG codons are also present in the putative $NDC80^{LUTI}$ mRNAs predicted from the other fungal species. Three regions were enriched for the presence of such AUGs (Figure 2.31 and 2.32), but the sequences and the length of these putative uORFs did not seem to be conserved. This observation is consistent with the idea that the act of uORF translation, rather than the identity of the uORF peptides, serves as a conserved feature in evolution.

The repressive nature of the uORFs contained in $NDC80^{LUTI}$ mirrors those found in the uORF-containing prototype transcript, *GCN4* (Mueller & Hinnebusch 1986). However, in the case of *GCN4*, changes in nutrient availability can relieve the uORF-mediated trans-lational repression, whereas for $NDC80^{LUTI}$, the uORF-mediated repression appears to be constitutive. In both cases, *GCN4* and *NDC80* can exist in on and off states. For *GCN4*, this switch is manifested in the two translational states of the same mRNA molecule. For *NDC80*, the switch is manifested instead by two distinct transcripts, one, which results in protein synthesis and one, which represses protein synthesis. It is important to note that for other potential LUTI-mRNAs, the precise mechanism of translational repression may not be conserved and could instead involve other means such as RNA hairpins or binding sites for translational repressors.

The function of $NDC80^{LUTI}$ mRNA is purely regulatory

Why do meiotic cells express an mRNA that does not encode any functional polypep-tides? We propose that the biological purpose of $NDC80^{LUTI}$ is to shut down Ndc80 protein synthesis by repressing $NDC80^{PROX}$ in *cis*, thereby inactivating kinetochore function dur-ing meiotic prophase. Multiple lines of evidence support this model. First, disruption of $NDC80^{LUTI}$ expression in meiosis results in elevated levels of $NDC80^{PROX}$ and Ndc80 pro-tein in meiotic prophase, leading to premature kinetochore activation. Second, induction of $NDC80^{LUTI}$ transcription in *cis* is sufficient to repress $NDC80^{PROX}$ and inactivate kine-tochore function in mitotic cells. Third, transcription of $NDC80^{LUTI}$ introduces repressive chromatin marks at the $NDC80^{PROX}$ promoter that are necessary for the downregulation of $NDC80^{PROX}$ and Ndc80 protein. Altogether, these findings strongly suggest that the primary function of the $NDC80^{LUTI}$ mRNA is to turn off the *NDC80* gene.

```
Sbay: TTTAGTTTCTGGGAACGTATAGTAGTAGTACTTAAATAGCGCAGAGCAATTTCAATCCAG
Skud: -TTATTACGTATGAATGCTCAGTAAAGGCACCTT---AAAGAAGACTTCAATGTTAGCAC
Smik: -TGATTGTGTATGAACATAAAATGTAAGCACCTATATAAAAAGGTGTGATATTAACGCAA
Spar: -CAATTGTATACGAATATTCAGTATAAACGCCTGTATAGGATGGCGTACTATCAACACAA
Scer: -CAATTCTATATGGATATGCAGTATAAAAACTTGTATAAAACAGTGCAGTATCAACACAC
             *  *   *  * *    * *    *   * *    *    *     *      **
Sbay: GC--TCAACTCTTTACCATATGATGAATTTCGATTCACCAAAATGATCTAGAAGCTATTC
Skud: AAGCTCACTTCTTATCGGTATAAATATATGCGATACATTA-AATAATGATCTCATTTCTT
Smik: GCTCACTCCTATACTAGGAATAGGTATATGCATTCTATCG-AATAATGAACATTATTCTC
Spar: GC---CAATTGTTCTACACATAGGTATATATAATATATCG-AAAAATGCTCATTATTCTC
Scer: AC---CACTTCTTGTACACATAAATATGTACGCTACATCG-AGAAATGTTCATTAATCTC
             *   * *       **    *  *    *   *    *  **         *
Sbay: GTATTGAAAAAAACGGTAAAAAAAAGAAAAAAAAAATCAAAACATACATTACCCGCACAT
Skud: TTGTTGAAATATGC---GACTGAAAAAAAATGACCGTCAGAATAACATTTACCCGCACTT
Smik: TTGTTGGAAGATTT---GGTAATGAAAAATTTCAC--CTACATATCATTTACCCGCAC-A
Spar: TTTGTTGGAAAGTA---TGACAGA---AAAATCTCATCAAAATAACATTTACCCGGATTT
Scer: TTGTTGGAAAATAT---GACAAAAGAAAAAAACTCGTTAAAGTAACATTTACCCGGATA-
           *  *    * *        **              *    ******* *
                                         URS1
Sbay: T-TGGCT-TCTCTTGCCCTTCAAAATAGTTGGGCGGTTAGTTTGTTTCTTTCTC-TTAAT
Skud: T-TCACTGTTTAATATCTTCAAGAAAAGATGGGCGCTTAATTTGTTTTCTTCTTTCTTGT
Smik: TTAGTTTTCGTCCAATATTAAAAAAAACATGGGCGCTTAAGTTATTCTTTTACACTTCTT
Spar: TCTAATTTAGTCTAATATTTTAAAAAAGATGGGCGCTTAATTTTTACTTTTGCTGTTACT
Scer: --TCTGTTCAGCCAACATTATAAAAAGATGGGCGCTTAATTTCCACTCTTACA-GATAT
           *       *  * ** *  ****** *** **   **        *
Sbay: CCTTTATTTTGTTG-TCATCGACTACAACACGATGGCAAGAAGAACGGAACAGGATAATA
Skud: TTACTC-TGTTTATTGAACAAACATCAGTAAGGTAGTAAGACTAAGACACAGGA--CAAA
Smik: ACTCACTTCTTTCATTCGGCAACATCAAGAAGATATTAAGAAGAGCAATAGCAC---ACT
Spar: CACTTCTTCT--AATCAACATAAATCAGTAAGATAACAAGAAGAGCAATAACAC---ACT
Scer: CACTCCTTCT--AATTAACTTACATCAGTAAGAAAGTAAAAAGAGCAACAATTC---GCC
           * *     **  * *       ** *  *
Sbay: AGAAGAAGGTAACAAGACAAACAAATGCGGGCTGATTTAATGCCATTTGATTCGCTATGG
Skud: AGAGGTTTGGTA---------------CTATGCGATTAGTACTGGAGTAATTGTGTATG
Smik: ATGATTCAGGTT---------------CTACTAGTTTATTATCACCAGATTGTTGTTAG
Spar: CCAACAAAGTTT---------------TCGTTGGCTCAATAGTAACAGAGTTTTGCCGG
Scer: ACAAGAAGGTCT---------------CTGTAGGGTCAATAGTAGCAGGGCTCTTGTCA
           *                        * * *
Sbay: TGATGGT---ATTTAAATACAGAAAGGAACACCAGCGGTGCAAATATGTGTCTAGATTTG
Skud: AGAA------------CAAAAAGACTGAGGGTTTTGAATCAGTTTCGATTTATTGGCGCT
Smik: AGAGTAA---TGTTGCGTAAAGCCAAGAAAACCGGAGGTCCGAAAAGTATTGGTGT---T
Spar: AGAAAGATTGTGGACAAAAAGGAAAGAAAAATTGGAGATCTGAAAATTATTCTTGGTATT
Scer: GGGAAAGACTGTGAATGCAAAGGAAAGACGGTTGGAGGTCCGAAAAGCATTACTGATGTT
           *             *         *        *         *   *
Sbay: TCGGTACACGGTGGCTAAGGCCAGTTTTGACATGATATGAGTGATCATGGTGGTAATAAA
Skud: AAAATAGAGCAGAGACT-------TTGTTGTAGTGCATGCTAGAGCCAATAGCTATTGTA
Smik: AGAATAGGAGAGGAATT-------ATACTACACTACTGGGCACATCTAAAAGCATTTATA
Spar: AGAATAAAAGAGAAACT-------CA-CTGTACTATGAGCTAGAGTTAGTAGTAACTTTA
Scer: AGGATTGAAGAGAAATT-------CA-CCGTACTATGTGCCAGATG--GTTATAACTGAA
           *                      *    *   *        * *
Sbay: CATACAGCTAGATCGGGTTGGCTCTG---CTCGGCGGTTAAGGGGCTTTCAATGGCACTT
Skud: TTATGAGCTGAAATGGCTCTATCTGATTCCTAGAAGGTTTCA-TCTAG-TACTGGCAAAT
Smik: GTTAGGA--TTAATCTATGACCACATCTCGTTGGGACCCTTG-TTTGGTAACGTTAAATA
Spar: CATTGGAATAGAGTCGTGATGCATATC--GGTTGGGAGCCAT-GTTAACTATGGTAAGGA
Scer: CATTGAAATAATATCGTGATTATGGGCTGGTTGAGAGCCCCG-TTAAGTATTGTTAAAGA
                                                                  *
```

Figure 2.31: Clustal analysis for the upstream intergenic region of the *NDC80* locus across five *Saccharomyces* species: Part 1. * indicates sequence identity; - indicates a skipped base pair. The displayed sequences are on the same strand as the *NDC80* ORF sense strand, in 5' to 3' direction (from left to right). Figure 2.32 is a continuation of this figure. Putative Ume6 (URS1) and Ndt80 (MSE) binding sites are marked as pink and blue regions, respectively. Colored boxes (one color for each species) denote the putative upstream ATGs that locate at least 50 bp downstream of the URS1 site and up to the start of the *NDC80* ORF. The number of putative AUG uORFs for each species is summarized in the table in Figure 2.32.

```
                                                                    MSE
Sbay: GCTACCTAATCTTCTGCCCTCGGCGGCTAAACTC-----------CGTTACCCAGCTTCA
Skud: TCTACCCA-CTCTGCTGCCTCGGCGGCTAATATACTGAGAGAATGGGTTACCCATCCCAA
Smik: TATAGCCAACTCATTGCCGTCGGCGGCTAAAAAATACCATAAAGAGGTTACCCATCCTGA
Spar: TGTAGCCAGTTTACTGCCATTGGCGGTTAAAATATCATTCAGAA-GGTTACCCTTTCTGA
Scer: TGTAAGCAGTCTATTGCGATTAGCGGTTAAAATACCATTTCAAG-GGTTACCCATACTGA
       **   *          *  **** ***              *******     *

Sbay: AA-TACCCAAAAAGACGCTTTCTTGCGTACAAAATGATAAT--ACACAAAACAAAAATTT
Skud: GA-GATTCAATTTAACGCTTCACCACGTACAAAATGATATATCCACAAAACGAAAATACC
Smik: AAAGATTAAAGAGGAAAGTCTACGACGTATAAAGTGATACATCCACAAAACGAATGTTTT
Spar: AA-GTATAAAAAAATG-GATCATAACGTATAGAATGATATA--CACAAAATTGAATTCCC
Scer: AA-GTGCCAAAAGAAAAAAAGATAACGTATGGAATGATATAGTCACAAAACTAATATTTC
       *       **           ****   * *****      ***   *

Sbay: CTAAAATAAACAAACGACACCAAAAATGGCAAAAAACAAAAGAAATAGATAGCATAGA--
Skud: C-AAAATAAACAAACAACACCAAAAATGGTAAAACAGA--AGAGATAGATAATCTACA--
Smik: CCAAAGTAAACAAAGAACACCAAAAATAGCGAAAT-AAT-GGAAATAGATAACCTGGGAG
Spar: CCAAAATAAACAAACAACAACACCAAAAATGGCAAGAATTAT-TGGAATAGATTGCGTAG--G
Scer: CTAAAATAAACAAACAACACCAAAAATGGTAGAAATAAT-TGGAATACATTCACAGG---
       * *** ********  *********** *    *    *   *  *** **

Sbay: AGAGGTCGGGATCTACCATTGGCTCTTCAGA--GACCGCACTGTCAAAAAATAACCGAAT
Skud: GGAGGTTAAAATCTGCACTAGCTCTTGAGA---ACCCAAACTGACTAAAA-AAAAGAAAA
Smik: GGGGGTAAAATCCATCTCTGACGTTTACGAAAGACTAGAACAAACCAACCGTAAAAATAA
Spar: AGAGGTACAATCCGTCCCTAGCTTCTGTCAAAAACTAAAACAGAGCAAGAAAAAT----A
Scer: AGAGGTAGAA------TCGTCCCTGTAAAAAAAATTGAAACGGAACAAGAAAAAT----A
       * ***              *         **      **    **

Sbay: TCCTTTCCATCCTTTCTCCCCATTTAGAGTGTACTATAGAAATTTTTTATCAGAGAGATA
Skud: AATCAGCCTCTTCTCCTGCAAGCCTATAGTGTACTATAGAAAGAT-------------A
Smik: AATTTAACCCCTCCTTTTCATATCTAGAGTGTACTATAGAATATCTACATTAGAGAGATA
Spar: ACTGAACTTTTTTCCTTTCAAATTTAGAGTGTACTATATAGTATATTC-TATAGGAGAGATA
Scer: ACTGACGTTCCTTTCTTTAAACGTTAGGGTGTACTATAATATATTC-CATAAGAGATACC
       *        **  ***** ****  *

Sbay: CCCGTTTCTGTAAA
Skud: CCTGTTTCTGTAAG
Smik: CCCGTTTCTGCAGA
Spar: CCCGTTTCTATAAA
Scer: CGTTT--CTATAA-
       *   *  **  *
```

Summary:

Species	Sbay	Skud	Smik	Spar	Scer
Number of putative uORFs	10	8	5	6	9

Figure 2.32: Clustal analysis for the upstream intergenic region of the *NDC80* locus across five *Saccharomyces* species: Part 2. * indicates sequence identity; - indicates a skipped base pair. The displayed sequences are on the same strand as the *NDC80* ORF sense strand, in 5' to 3' direction (from left to right). This is a continuation of Figure 2.31. Putative Ume6 (URS1) and Ndt80 (MSE) binding sites are marked as pink and blue regions, respectively. Colored boxes (one color for each species) denote the putative upstream ATGs that locate at least 50 bp downstream of the URS1 site and up to the start of the *NDC80* ORF. The number of putative AUG uORFs for each species is summarized in the table.

The transcriptional mechanism of *NDC80* repression

How does $NDC80^{LUTI}$ interfere with $NDC80^{PROX}$ transcription? Our data show that Set2 and Set3 are essential for establishing a repressive chromatin state and for inhibiting $NDC80^{PROX}$ transcription. We propose a two-step mechanism. First, transcription of $NDC80^{LUTI}$ deposits Set1 mediated H3K4me2 and Set2 mediated H3K36me3 in the $NDC80^{PROX}$ promoter. Second, co-transcriptional deposition of these marks facilitates the recruitment of the histone deacetylase complexes Set3C and Rpd3S (Carrozza *et al.* 2005; Keogh *et al.* 2005; Kim & Buratowski 2009). Notably, the two marks localize to the same area of the $NDC80^{PROX}$ promoter, perhaps indicating that there is redundancy between the two pathways (see following section in the discussion). Previous work established a role for Set2 in suppressing histone exchange and promoting nucleosome stability through chromatin remodelers (Venkatesh *et al.* 2012; Smolle *et al.* 2012). Hypo-acetylated histones are also associated with increased nucleosome stability (Venkatesh & Workman 2015). Hence, cells lacking both Set2 and Set3 show reduced nucleosome occupancy in the $NDC80^{PROX}$ promoter, and $NDC80^{PROX}$ transcription is unimpeded despite active $NDC80^{LUTI}$ transcription (Figure 2.14). Taken together, $NDC80^{LUTI}$ repression of $NDC80^{PROX}$ is mediated by co-transcriptional chromatin reorganization of the $NDC80^{PROX}$ promoter.

The Set1/Set3C and Set2/Rpd3S pathways have well characterized roles in preventing cryptic transcription and regulating gene expression via long noncoding RNA transcription (Carrozza *et al.* 2005; Keogh *et al.* 2005; Kim & Buratowski 2009; Kim *et al.* 2012; van Werven *et al.* 2012; Ard & Allshire 2016; Venkatesh *et al.* 2016). It has been reported that Set2 and Set3 modulate the expression of different genes based on the length of adjacent ncRNAs, which overlap with their promoters. (Kim *et al.* 2012; Kim *et al.* 2017). Kim *et al.* (2017) demonstrate that during a series of carbon source shifts, genes whose promoters overlap with longer transcripts (2.0 kb) are repressed by Set2/Rpd3S whereas those with shorter overlapping transcripts (0.9 kb), are repressed by Set1/Set3C (Kim *et al.* 2017). According to their classification, $NDC80^{LUTI}$ is a short overlapping transcript. Strikingly, $NDC80^{LUTI}$ mediated repression of $NDC80^{PROX}$ was compromised in the *set2Δset3Δ* double mutant cells, but not in the single mutants (Figures 2.12, 2.13, 2.26). We propose that Set1/Set3C and Set2/Rpd3S act redundantly during $NDC80^{LUTI}$ mediated repression of $NDC80^{PROX}$.

The mechanism that we have described here has clear similarities with transcriptional interference mediated by intergenic or promoter transcription. Like $NDC80^{LUTI}$, transcription of the intergenic/promoter ncRNA, *SRG1*, leads to increased nucleosome occupancy and lower binding of activators in the downstream *SER3* promoter (Martens *et al.* 2004; Hainer & Martens 2011). Transcriptional interference is also important for mating-type control of sporulation. In cells with a single mating type, transcription of a long noncoding RNA *IRT1*, in the promoter of the master regulatory transcription factor Ime1, is critical for *IME1* repression (van Werven *et al.* 2012). Similar to $NDC80^{LUTI}$, *IRT1* transcription establishes repressive chromatin throughout the *IME1* promoter via a mechanism requiring both Set2

and Set3. Interestingly, while Set1 and Set2 act on different parts of the *IME1* promoter, the H3K4me2 and H3K36me3 marks overlap in the $NDC80^{PROX}$ promoter suggesting that both modifications control the same promoter region. Perhaps, H3K4me2 and H3K36me3 occur on the same nucleosome as part of a repressive combinatorial histone code.

Transcription factor-driven gene repression by LUTI-mRNA: an evolutionary perspective

Why do budding yeast cells use this seemingly complex mechanism, which relies on the transcription of an undecoded mRNA isoform, to repress a kinetochore gene during meiosis? We would argue from an evolutionary point of view that this solution could be both economical and highly flexible. First, the meiotic cell is co-opting two existing transcription factors, Ime1 and Ndt80, for roles in activating and repressing gene expression, obviating the need to evolve novel trans-acting factors. This mechanism also ensures temporal coordination of gene activation and inactivation using the same transcription factor. In the case of *NDC80*, the LUTI-mRNA rides the Ime1 wave of gene expression to shutoff kinetochore function while the protein-coding mRNA rides the subsequent Ndt80 wave to reactivate the kinetochore for the division phases. While transcription factors have previously been implicated in the repression of downstream promoters (Martens *et al.* 2004; van Werven *et al.* 2012; Bird *et al.* 2006; Shearwin *et al.* 2005), our study is the first clear demonstration that it is the choice of promoter and the identity of the resulting mRNA isoform that governs whether a gene is turned on or turned off by a given transcription factor.

This mode of gene repression relies on two sets of *cis*-regulatory sequences, which are evolutionarily flexible (Carroll 2008; Stern & Orgogozo 2008; Wittkopp & Kalay 2012). The first *cis*-acting sequence is the distal transcription factor-binding site, which induces transcription of $NDC80^{LUTI}$, and, in concert with co-transcriptional chromatin modifications, silences the downstream canonical promoter activity. The second *cis*-acting sequence is the AUG-uORFs within the extended 5'-leader of the LUTI-mRNA, which prevents downstream PROX translation. Inherent to a mechanism that is so heavily reliant on *cis*-regulatory elements is the notion that minor changes in the DNA sequence can impact gene expression at a multitude of levels, thus tuning gene output. This tuning can be manifested at the level of nucleosome spacing, strength of transcription factor binding and translational regulation. Therefore, the cell has a vast evolutionary space, which can be explored through small changes in DNA sequence.

Synthetic tuning of gene expression by LUTI mRNAs

By demonstrating that the leader of $NDC80^{LUTI}$ can be placed in front of *NUF2* to repress Nuf2 gene expression and that LUTI-mediated repression can occur in mitosis as well as meiosis, it opens the door for use of LUTI mRNAs to synthetically repress gene expression in additional contexts. Further, using a scalable expression system, we showed

that modulating the levels of $NDC80^{LUTI}$ affects the efficiency of $NDC80^{PROX}$ repression (Figure 5). The higher the levels of $NDC80^{LUTI}$ transcription, the better the repression of $NDC80^{PROX}$ becomes. This could be used to fine-tune gene expression systems. Notably, Set2 and Set3 are no longer required for repressing $NDC80^{PROX}$ when $NDC80^{LUTI}$ is highly expressed. One possible explanation is that the rate of nucleosome deposition at the $NDC80^{PROX}$ promoter is increased during higher levels of $NDC80^{LUTI}$ transcription. In this situation, the requirement for histone deacetylase complexes to stabilize nucleosomes becomes obsolete. Alternatively, elongating RNA polymerase might physically interfere with the $NDC80^{PROX}$ promoter when $NDC80^{LUTI}$ is highly expressed.

Whereas most studies have reported a binary switch for transcription interference mechanisms (Martens *et al.* 2004; Hongay *et al.* 2006; van Werven *et al.* 2012; Camblong *et al.* 2007; Bumgarner *et al.* 2009), we propose that transcriptional interference by expression of a 5'-extended transcript is tunable. This principle could be further adapted and used in synthetic genetic circuits to modulate gene expression levels. Indeed, mechanisms of transcriptional interference have been applied to coordinate activities of adjacent genes in both E. coli and budding yeast (Buetti-Dinh *et al.* 2009; Bordoy *et al.* 2016; Hao *et al.* 2016; Hoffmann *et al.* 2016).

Pervasiveness of LUTI-mRNA biology in yeast meiosis and beyond

The defining sequence features of the *NDC80* LUTI-mRNA are a 5'-extended mRNA leader coupled with repressive uORFs contained in this extended leader. Analysis of the mRNA-seq and ribosome profiling datasets of meiotic yeast revealed hundreds of transcripts with potential LUTI-like signatures (Brar *et al.* 2012). In support of this idea, two other genes, *ORC1* and *BOI1*, have been shown to express meiosis-specific transcript isoforms with uORF-containing leader extensions (Liu *et al.* 2015; Xie *et al.* 2016). Rather than dissecting each candidate LUTI-mRNA on a case by case basis, future studies that integrate additional genome-wide datasets to measure stage-specific transcription factor binding sites, transcription-coupled chromatin modification states, mRNA translation status with isoform specificity and protein abundance would result in a high-confidence map of LUTI-mRNAs and aid in the dissection of their cellular functions.

Beyond budding yeast meiosis, can the regulatory circuit described in our study be present in other developmental programs and in other organisms? We would argue so, because various organisms also possess the three principles of this module, namely, alternative promoter usage, transcription-coupled repression, and uORF-mediated translational repression. Alternative promoter usage is widespread in development and among different cell types. For example, in the fruit fly, more than 40% of developmentally expressed genes have at least two promoters with distinct regulatory programs (Batut *et al.* 2013). Half of human genes have more than one promoter, resulting in the expression of mRNA isoforms

with 5' heterogeneity (Kimura *et al.* 2006). Furthermore, transcription-based interference mechanisms, as well as transcription-coupled histone modifications, have been described in a variety of organisms (Corbin & Maniatis 1989; Eissenberg & Shilatifard 2010; Shearwin *et al.* 2005; Wagner & Carpenter 2012). Finally, recent studies have shown that uORF translation is much more widespread than traditionally believed and acts in a regulatory manner (Calvo *et al.* 2009; Chew *et al.* 2016; Johnstone *et al.* 2016). Therefore, we envision that the regulatory circuit described here can be used as a roadmap in future studies to uncover transcription-coupled gene repression during cell fate transitions across multiple species.

Interpreting genome-wide data in the context of LUTI-mRNA biology

A key implication of this model of gene regulation is a blurring of the line between coding and non-coding RNAs. Seminal work has uncovered multiple classes of non-coding RNAs that play regulatory functions in the cell, such as long non-coding RNAs, microRNAs, small interfering RNAs, and piwiRNAs (Batista & Chang 2013; Guttman *et al.* 2009; Cech & Steitz 2014; Ambros 2001). Our study demonstrates that mRNAs, which are deemed protein coding units, can themselves be direct regulators of gene expression by at least two simultaneous means: they can induce transcription-coupled silencing of a downstream promoter, and features in their 5'-leaders, such as the presence of uORFs or secondary structures, could directly impact translation efficiency in a positive or negative manner (Arribere & Gilbert 2013; Brar *et al.* 2012; Rojas-Duran & Gilbert 2012). Notably, multiple studies have reported poor correlation between mRNA and protein abundance (Maier *et al.* 2009). For those mRNAs that anti-correlate with their protein levels, this apparent contradiction might be due to a LUTI-mRNA being misattributed as a canonical protein-coding transcript. Our study could dramatically transform the way we understand the function of alternate mRNA isoforms and aid in the proper biological interpretation of genome-wide transcription studies.

Chapter 3

Key Determinants of Long Undecoded Transcript Isoform-based Gene Repression

3.1   Introduction

During processes of cell differentiation and development, changes in gene expression are coordinated through both space and time. If gene regulation goes awry during key developmental stages, it can be detrimental to an organism. Thus, gene expression, and its role in development, has been studied intensely for decades. Through this, it has come to light that the most critical drivers of changes in cell state are transcription factors. They are the first genes to turn on at the start of a developmental program, and they are responsible for initiating cascades of regulatory events. Traditionally, transcription factors are most associated with the activation of genes, so little is known about how, especially during times of vast and rapid transcriptional changes, gene repression is coordinated with the transcription-factor dependent waves of gene activation.

Ironically, new insights into a mechanism of transcription factor-mediated gene repression came from studying transcription and translation simultaneously. Advances in technology such as ribosome profiling opened the door for genome-wide investigations at multiple steps during gene regulation (Ingolia *et al.* 2009; Brar *et al.* 2012). These techniques revealed that transcript abundance correlates poorly with translational activity at many gene loci (Ingolia *et al.* 2009; Brar *et al.* 2012; Cheng *et al.* 2018). In-depth characterization of one such locus came from our study of how *NDC80*, a gene encoding an essential and conserved kinetochore protein, is regulated (Chapter 2). We showed that upon induction of the meiotic transcription factor Ime1, a long undecoded transcript isoform (LUTI) is produced from a promoter upstream of the canonical *NDC80* promoter (Chapter 2). This mRNA ($NDC80^{LUTI}$) cannot be translated due to 9 uORFs in its 5'-leader. On top of the translational repression,
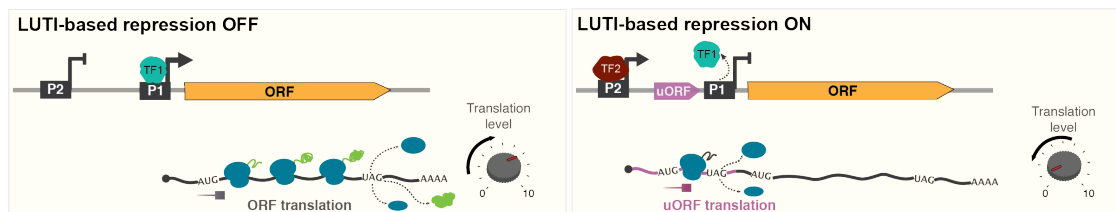
Figure 3.1: A schematic of LUTI-mediated gene regulation. On the left is a depiction of a gene when the LUTI mRNA is off and protein production is high. The right depicts features of LUTI-mRNAs that lead to low protein production.

transcription of $NDC80^{LUTI}$ decreases downstream transcription initiation from the gene proximal canonical promoter, the promoter that is able to produce a well-translated mRNA. This discovery marked the first description of an mRNA, which contains an entire known ORF, but plays a purely regulatory role (Chapter 2) (Figure 3.1). It became critical to further investigate other instances of possible LUTI-mediated repression to better understand the extent of its biological significance.

Budding yeast meiosis provides an ideal system to better understand LUTI mRNAs. Meiosis is the highly conserved and specialized cell division required for the production of gametes in all sexual organisms. More specifically, in budding yeast it is the single developmental program with the most densely sampled measurements of transcript abundance, translational activity, and protein abundance (Brar *et al.* 2012; Cheng *et al.* 2018). These measurements revealed dynamic patterns of gene expression which lead to almost 2/3 of genes experiencing > 10-fold changes in their expression and almost every gene being expressed; despite there being only two known meiosis-specific transcription factors to orchestrate the processes: Ime1 and Ndt80 (Chu & Herskowitz 1998; Kassir *et al.* 1988). Could it be that LUTIs are commonly targets of Ime1 and Ndt80 in order to better coordinate up- and downregulation of genes in a rapid developmental program without having to evolve additional *trans* regulators? A previously identified 5'-extended meiotic transcript at the *ORC1* locus is Ndt80 dependent (Brar *et al.* 2012; Xie *et al.* 2016), and both $NDC80^{LUTI}$ and an extended *BOI1* transcript are known to be repressed by Ume6 in mitosis, and activated by Ime1 in the case of $NDC80^{LUTI}$ (Chapter 2)(Liu *et al.* 2015). These examples indicated that meiotic transcription factors are tightly linked to some individual LUTIs, but more candidates must be investigated before we can conclude how frequently.

Since the discovery of $NDC80^{LUTI}$, other LUTIs have been identified, in budding yeast and in humans (Cheng *et al.* 2018; Van Dalfsen *et al.* 2018; Hollerer *et al.* 2019). 380 were found in budding yeast meiosis, a sizeable subset of which are most likely dependent on Ndt80 (Cheng *et al.* 2018). During the unfolded protein response (UPR) in yeast, the conserved transcription factor Hac1 was implicated in the regulation of 15 LUTIs (Van Dalfsen *et al.* 2018). Additionally, LUTI-based repression at the MDM2 locus in human cells has

recently been reported (Hollerer *et al.* 2019). The prevalence of LUTIs in budding yeast, their discovery in humans, the abundance of alternative TSS usage, and the frequency of uORFs in other organisms all point to the conservation of LUTI-based gene repression across vast evolutionary time (Batut *et al.* 2013; Calvo *et al.* 2009; Chew *et al.* 2016; Johnstone *et al.* 2016; Kimura *et al.* 2006).

Previous studies of LUTIs either provided a deep investigation of a single LUTI mRNA, or they analyzed genome-wide datasets to identify LUTIs functionally by looking for instances in which mRNA and protein abundances anti-correlate. Both approaches have proven successful in helping to further characterize the phenomenon of LUTI-based gene repression, however they have left a number of questions unanswered. First, some 5'-extended transcripts do not appear to repress gene expression, why? Second, if gene repression is observed, does it occur by a common mechanism? Lastly, what are the key features of LUTI-based repression?

In this study, we developed a pipeline using transcript leader sequencing (TL-seq) and direct RNA long read Nanopore sequencing to discover 74 5'-extended putative LUTI mRNAs present in budding yeast meiotic prophase but not under starvation conditions (Arribere & Gilbert 2013; Garalde *et al.* 2018). We find that 60% of the candidates are regulated by the meiosis-specific transcription factors Ime1-Ume6 at consensus binding motifs that are conserved across the sensu stricto genus. Of the 74 transcripts, 72 of them contain uORFs, and evidence indicates that very little, if any, translation occurs from the LUTI. Transcriptionally, the outlook is more complex. We show that in some cases, but not all, LUTI transcription leads to downregulation of the proximal transcript. Similar to the *NDC80* LUTI case, chromatin modifications and nucleosome position do play a role in LUTI-mediated repression. Additionally, a higher LUTI transcript abundance is more likely to correlate with repression of the proximal promoter. We conclude that LUTI mRNAs, both Ume6 dependent and otherwise, can facilitate gene repression in a manner that integrates translational repression, chromatin state, and LUTI expression. Further investigation is needed to determine how the PROX transcript expression levels and PROX promoter sequence may be involved.

3.2 Results

The combined use of transcript leader and Nanopore sequencing identifies 74 potential LUTIs in meiotic prophase

In order to uncover the prominent features of LUTI-based repression, we first developed a pipeline to identify meiotic mRNAs with 5' extensions (Figure 3.2A). For cell synchronization, we adopted a previously established protocol in which the expression of two early meiotic regulators, Ime1 and Ime4, are controlled by the copper inducible *CUP1* promoter (Berchowitz *et al.* 2013). Using this system, we focused on the 5'-extended mRNAs that are expressed in meiotic prophase (2 hours after induction of *IME1* and *IME4*), but not in
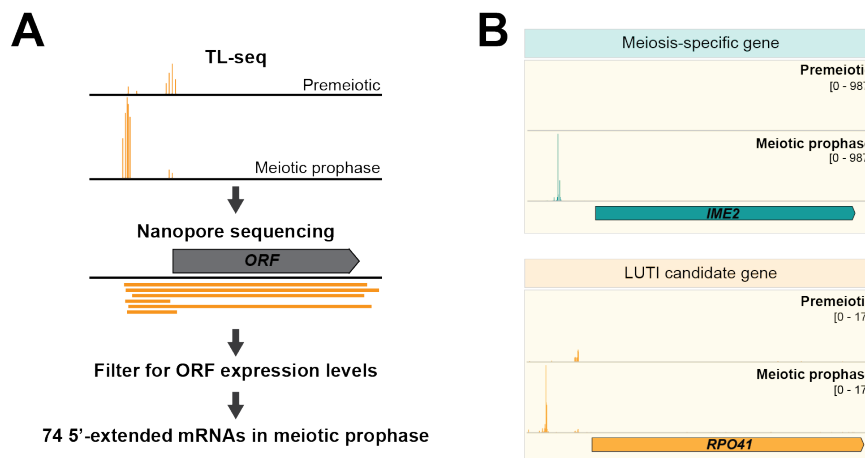
Figure 3.2: A pipeline for discovering LUTIs. A) An overview of the pipeline used to discover 5'-extended transcripts. Transcription starts sites (TSSs) that increase by at least log2 fold change > 2 were identified by transcript leader sequencing (TL-seq). Long read nanopore sequencing was used to confirm that transcripts produced from these loci spanned the entire open reading frame of the downstream gene. Instances in which no proximal promoter was identified were removed. B) Genome browser views of TL-seq for the meiosis-specific gene *IME2* and the *RPO41* locus which contains a LUTI mRNA.

premeiotic phase (no induction of *IME1* and *IME4*) (Berchowitz *et al.* 2013). Identification of the 5'-extended isoforms relied on data from two orthogonal sequencing techniques. The first method, transcript leader sequencing (TL-seq), involves sequencing just the most 5' end of a transcript (Arribere & Gilbert 2013; Malabat *et al.* 2015). TL-seq allowed for the identification of transcription start sites (TSSs) that robustly increased in expression as cells transitioned from premeiotic phase to meiotic prophase (Figure 3.2A-B). The second technique, Nanopore sequencing, can directly sequence an entire RNA transcript as a single read (Garalde *et al.* 2018). Nanopore sequencing confirmed the instances in which TSSs identified by TL-seq produced transcripts that elongate across an entire open reading frame (ORF) (Figure 3.2A) rather than being early terminated. Lastly, we were specifically interested in distinguishing LUTIs from the canonical, meiosis-specific mRNAs. Therefore only loci in which the promoter identified by TL-seq was upstream of a second promoter located more proximal to the corresponding ORF were selected (Figure 3.2A-B). This resulted in the identification of 74 potential LUTIs that were induced in meiotic prophase (Figure 3.2A, Table 3.1).

Table 3.1: Genes with LUTIs

Systematic Name	Gene	Ume6 Enrichced	URS1	Conserved
YBL011W	*SCT1*	Yes	Yes	Yes
YBL085W	*BOI1*	Yes	Yes	Yes

Table 3.1: continued

Systematic Name	Gene	Ume6 Enrichced	URS1	Conserved
YBR112C	CYC8	Yes	Yes	No
YBR114W	RAD16	Yes	Yes	No
YBR185C	MBA1	Yes	Yes	Yes
YBR257W	POP4	Yes	Yes	Yes
YBR281C	DUG2	Yes	Yes	Yes
YCL050C	APA1	No	No	na
YCL057C-A	MIC10	Yes	Yes	Yes
YDL070W	BDF2	Yes	No	na
YDL174C	DLD1	Yes	Yes	Yes
YDR060W	MAK21	Yes	No	na
YDR222W	YDR222W	Yes	Yes	No
YDR501W	PLM2	No	No	na
YEL015W	EDC3	Yes	No	na
YEL025C	YEL025C	Yes	Yes	Yes
YER111C	SWI4	Yes	Yes	No
YFL036W	RPO41	Yes	Yes	No
YGL097W	SRM1	Yes	No	na
YGL103W	RPL28	No	Yes	na
YGR194C	XKS1	Yes	No	na
YGR260W	TNA1	No	No	na
YGR264C	MES1	Yes	Yes	Yes
YGR266W	YGR266W	Yes	Yes	Yes
YGR281W	YOR1	Yes	Yes	No
YHR006W	STP2	Yes	Yes	No
YHR013C	ARD1	Yes	Yes	Yes
YHR071W	PCL5	Yes	Yes	No
YIL031W	ULP2	Yes	Yes	Yes
YIL144W	NDC80	Yes	Yes	Yes
YJL003W	COX16	No	No	na
YJL083W	TAX4	Yes	Yes	Yes
YJL084C	ALY2	Yes	Yes	Yes
YJL196C	ELO1	No	No	na
YJR077C	MIR1	Yes	No	na
YJR138W	IML1	Yes	Yes	Yes
YKL122C	SRP21	No	Yes	na
YLL031C	GPI13	Yes	Yes	Yes
YLR114C	AVL9	Yes	No	na
YLR115W	CFT2	Yes	Yes	No
YLR259C	HSP60	Yes	No	na

Table 3.1: continued

Systematic Name	Gene	Ume6 Enrichced	URS1	Conserved
YLR260W	LCB5	Yes	Yes	Yes
YLR274W	MCM5	Yes	Yes	Yes
YLR325C	RPL38	No	No	na
YML075C	HMG1	Yes	Yes	Yes
YML111W	BUL2	Yes	Yes	Yes
YMR100W	MUB1	No	Yes	na
YMR122W-A	NCW1	Yes	Yes	No
YMR208W	ERG12	Yes	Yes	Yes
YNL015W	PBI2	Yes	No	na
YNL305C	BXI1	Yes	Yes	Yes
YNR017W	TIM23	Yes	No	na
YOL002C	IZH2	No	No	na
YOL100W	PKH2	Yes	Yes	No
YOL103W	ITR2	Yes	Yes	Yes
YOL135C	MED7	No	Yes	na
YOR020W-A	MCO10	Yes	Yes	No
YOR076C	SKI7	Yes	Yes	Yes
YOR246C	ENV9	No	No	na
YOR290C	SNF2	Yes	No	na
YOR320C	GNT1	Yes	No	na
YOR321W	PMT3	Yes	Yes	Yes
YOR350C	MNE1	No	No	na
YOR352W	TFB6	Yes	Yes	Yes
YOR354C	MSC6	Yes	Yes	Yes
YPL088W	YPL088W	Yes	No	na
YPL120W	VPS30	Yes	Yes	Yes
YPL152W	RRD2	Yes	Yes	Yes
YPL160W	CDC60	Yes	Yes	Yes
YPL163C	SVS1	Yes	No	na
YPL274W	SAM3	Yes	Yes	No
YPR013C	CMR3	Yes	No	na
YPR029C	APL4	Yes	Yes	Yes
YPR164W	MMS1	Yes	Yes	Yes

URS1 Consensus Motif
TCGGCGGCTA

Motif in 5'-extended mRNA promoters
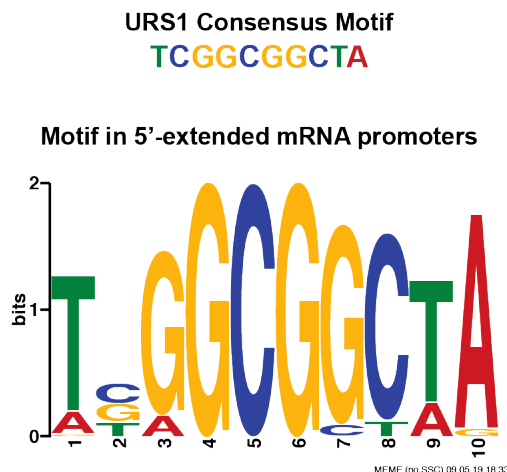
Figure 3.3: A URS1 motif in LUTI promoters. The consensus binding motif observed in the 300 bp +/- the identified distal promoters as identified by MEME. In 50/74 instances, a significant (combined match p-value < 0.05) URS1 binding motif match was found.

A subset of the putative LUTIs are regulated by the same meiotic transcription factor

With a list of 74 candidates, we sought to determine those that were regulated by a common transcription factor. A search for enriched regulatory motifs in the promoters of 5'-extended transcripts was performed. The one and only significant hit matched the URS1 consensus motif (Figure 3.3, Table 3.1, 50/74 sequences, combined match p-value < 0.05) (Sumrada & Cooper 1987). In mitosis, this motif is bound by the repressor Ume6 (Park *et al.* 1992). Upon entry into meiosis however, Ume6 interacts with the meiotic transcription factor Ime1 to induce expression of the genes necessary for entry into meiosis (Bowdish *et al.* 1995). We previously showed that Ime1/Ume6 regulate the expression of the LUTI at the *NDC80* locus, $NDC80^{LUTI}2$, and others have demonstrated that Ume6 represses 5'-extended transcripts during mitosis at the BOI1, CFT2, and RTT10 loci (Lardenois *et al.* 2015; Liu *et al.* 2015). The presence of a URS1 motif upstream of these 5'-extended mRNAs led us to hypothesize that Ime1-Ume6 may play a regulatory role at many of the loci producing 5' extensions.

To investigate how many of the motifs are in fact bound by Ume6, chromatin immuno-precipitation followed by sequencing (ChIP-seq) was performed. Enrichment of Ume6 was present at 61 of the 74 candidate LUTI promoters (Table 3.1, q-value < 0.001, FE > 4). The rate of enrichment was similar in this set of genes compared to a set of previously identified Ume6 targets. Both of these groups were far more enriched with Ume6 than genes not in these lists (Figure 3.4A-B, Figure 3.5).
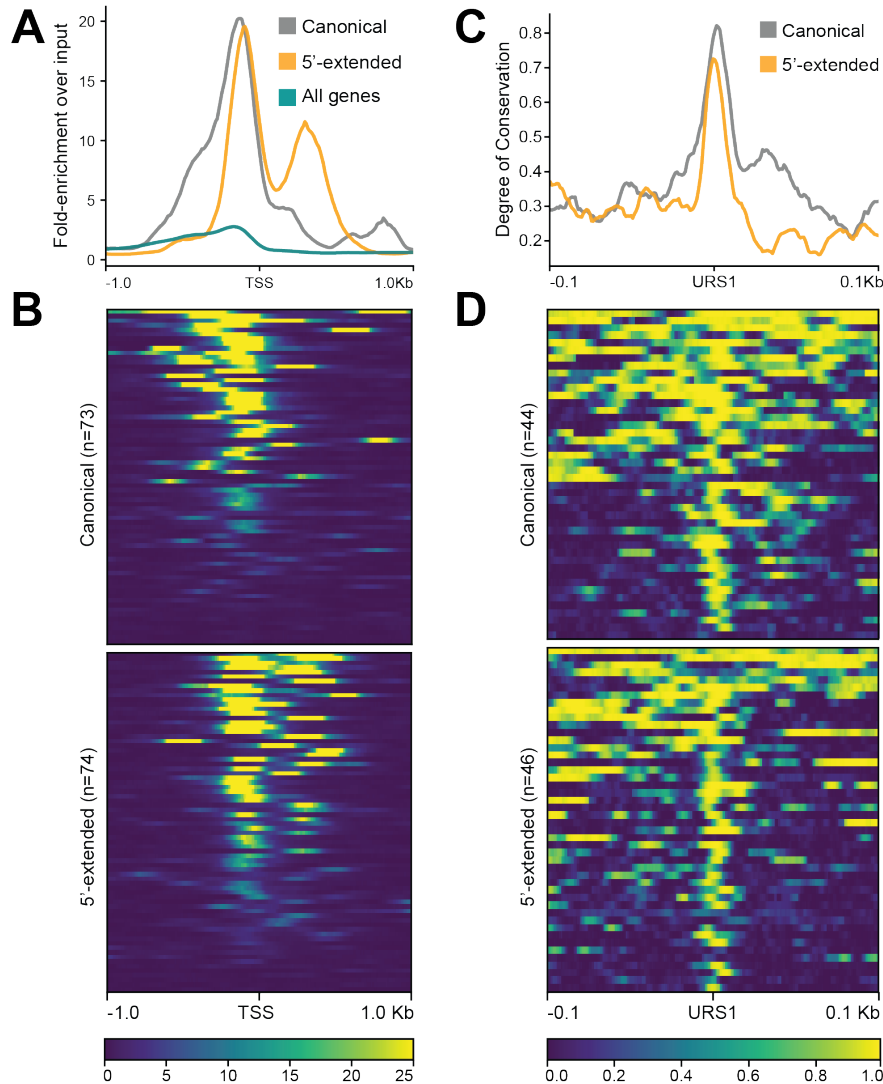
Figure 3.4: Ume6 enrichment and URS1 conservation in promoters of 5'-extended transcripts. A-B) Ume6 ChIP was performed (UB3301) on cells grown in BYTA to saturation. A) Metagene analysis of Ume6 fold enrichment over input in the promoters of all genes compared to the promoters of a previously identified Ume6 targets and to the promoters of the 5'-extended transcripts identified in this study. A representative image from one of three replicates. B) Heatmap of Ume6 fold enrichment over input in the promoters of previously identified Ume6 targets (top) and in the promoters of the 5'-extended transcripts identified in this study (bottom). Representative images from one of three replicates. C-D) For all genes with both Ume6 enrichment and a URS1 motif +/- 300 bp from their TSS, the degree of conservation within the sensu stricto genus was determined by phastcons. C) Metagene analysis for degree of conservation in the promoters of previously identified Ume6 targets compared to the promoters of the 5'-extended transcripts identified in this study. D) Heatmap of degree of conservation in the promoters of previously identified Ume6 targets (top) and in the promoters of the 5'-extended transcripts identified in this study (bottom).
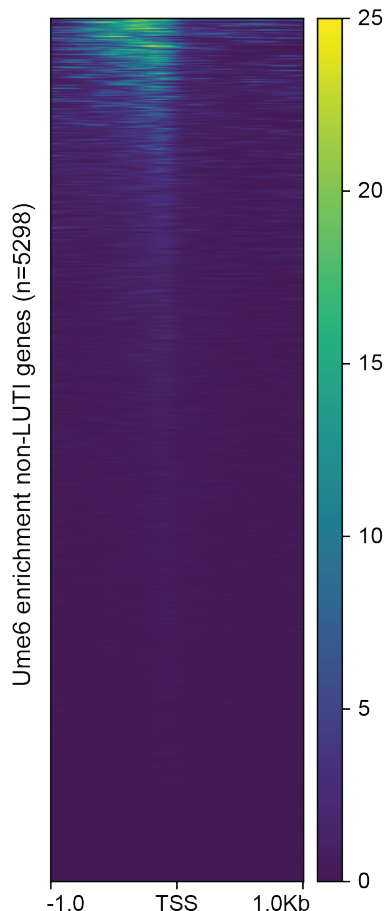
Figure 3.5: Heatmap of Ume6 enrichment over input for all of the genes without 5'-extensions. Representative image from one of three replicates

We further examined URS1 motif conservation in putative LUTI promoters as a means to assess functional significance. Using an alignment of 5 yeast species in the sensu stricto clade (*S. cerevisiae, S. paradoxus, S. mikatae, S. kudriavzevii, and S. bayanus*), conservation of the region around the URS1 motif was calculated for all sites that both had a URS1 motif and were bound by Ume6. In both sets of genes, conservation sharply increased around the URS1 motif (Figure 3.4C-D); providing strong evidence that 33 of the identified 5'-extended transcript isoforms have strongly conserved URS1 binding sites, and their regulation by Ume6-Ime1 is likely functional.

The role of transcript isoform diversity and uORFs in translational regulation of 5'-extended isoforms

LUTI mRNAs are by definition translationally impaired. They were originally identified when a discrepancy between transcript abundance and level of translation was observed at
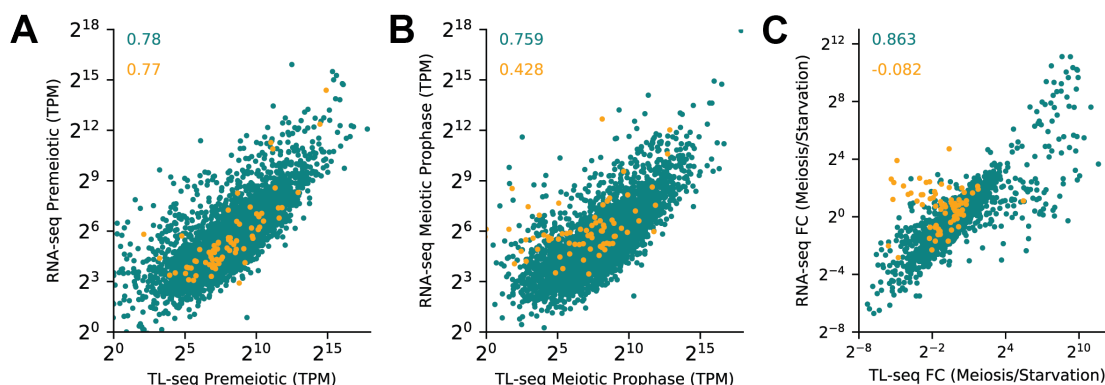
Figure 3.6: The correlation between RNA-seq and TL-seq. A-B) Scatterplot of RNA-seq and TL-seq for genes with (orange) and without (teal) 5'-extensions. Cells were induced to undergo meiosis and collected during meiotic prophase (*UB14584*). Experiments were performed in duplicate. Transcripts per million were quantified by salmon for RNA-seq and cageR for TL-seq. Spearman rank correlations are in the upper left corner. A) Cells with the *pCUP1-IME1/pCUP1-IME4* meiotic induction system (*UB14584*) were collected for TL-seq and RNA-seq after 2 hours in sporulation medium (SPO), before induction of meiosis (Premeiotic). B) Cells with the *pCUP1-IME1/pCUP1-IME4* meiotic induction system (*UB14584*) were collected for TL-seq and RNA-seq after 4 hours in sporulation medium (SPO), 2 hours after induction of meiosis by 50 μM $CuSO_4$ (Meiotic Prophase). C) The fold-change by which gene expression changes as the cells enter early meiosis from starvation as quantified by DESeq2. Spearman rank correlations were calculated for genes with extensions (orange) and without 5'-extensions (teal).

the *NDC80* locus (Chapter 2) (Brar *et al.* 2012). Since then, 379 additional LUTIs were identified in meiosis by looking for an anti-correlation between protein and RNA abundance (Cheng *et al.* 2018). Intriguingly, only 32 of the 74 targets in our study were identified by the previous study. 10 of the genes were missed due to lack of a quantifiable protein measurement. However, this is not the only reason for the discrepancy. By selecting for instances in which the presence of an extended isoform correlates with decreased protein abundance, examples in which a 5'-extended leader is present, but does not lead to an apparent change in protein abundance are missed. By using all 5'-extended transcripts present in meiotic prophase, we now have the opportunity to determine how frequently they repress gene expression and under what circumstances this repression occurs.

It was previously difficult to answer these questions because, by mRNA-seq, the transcript produced from the proximal promoter, hereon referred to as the PROX isoform, does not have any unique identifying sequence compared to the transcript produced from the distal promoter. By sequencing the most 5' end of a transcript with TL-seq, that was no longer a limitation. We performed TL-seq and mRNA-seq on samples collected from a premeiotic stage and meiotic prophase. Interestingly, while the relationship between RNA-seq and TL-
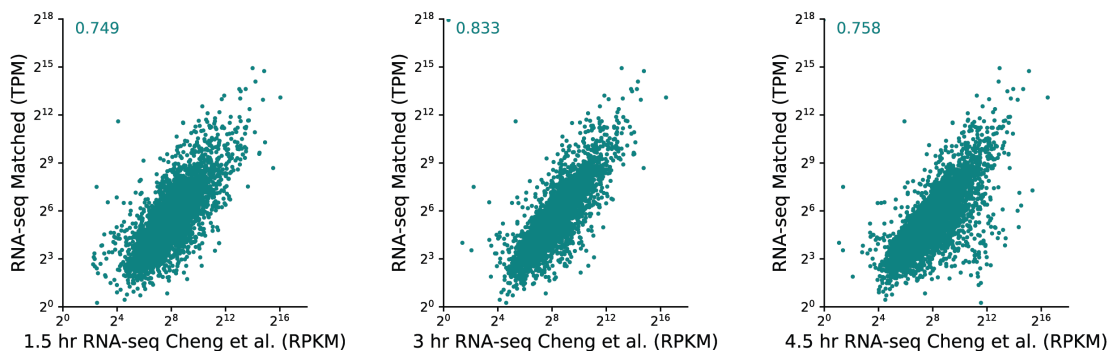
Figure 3.7: Determining best matched time points in Cheng *et al.* (2018). Scatterplot comparing RNA-seq performed in this study to RNA-seq performed in Cheng *et al.* (2018). In this study, cells with the *pCUP1-IME1/pCUP1-IME4* meiotic induction system (*UB14584*) were collected for RNA-seq (in duplicate) after 4 hours in sporulation medium (SPO), 2 hours after induction of meiosis by 50 µM (CuSO4). In Cheng et al., a prototrophic strain was used. The spearman rank correlation was calculated.

seq was good for most genes, it was quite poor for the 5'-extended subset in meiotic prophase (Figure 3.6A-B). Even more strikingly, when the fold-change of expression was taken into account as cells transition from premeiotic to meiotic prophase, there was no correlation between TL-seq and mRNA-seq for genes with 5' extensions (Figure 3.6C). This reflects the fact that TL-seq, unlike mRNA-seq, can individually quantify 5'-extended and PROX isoforms.

Given that previous LUTI mRNAs were characterized by a negative correlation between mRNA abundance and levels of translation, we next set out to compare the correlation between translation and transcript abundance, with the added ability to quantify the PROX transcript in meiotic prophase. The 3-hour time point from the previously published matched meiotic ribosome profiling and mRNA-seq dataset Cheng *et al.* (2018) best matched our meiotic prophase time point after comparing the two mRNA-seq datasets (Figure 3.7). We compared those ribosome profiling measurements to our RNA-seq and TL-seq measurements (Cheng *et al.* 2018). For genes without 5'-extensions, the distributions as a whole looked quite similar (Figure 3.8A-B). However, a large number of genes had lower than expected footprints in the subset of genes with 5'-extensions (Figure 3.8C-D). Remarkably, this bias disappeared completely when only PROX transcripts were quantified (Figure 3.8D). This analysis provided further evidence that the identified 5'-extended transcripts as a whole did not productively translate the ORFs contained within them. For this reason, our identified targets will be referred to as LUTIs going forward.
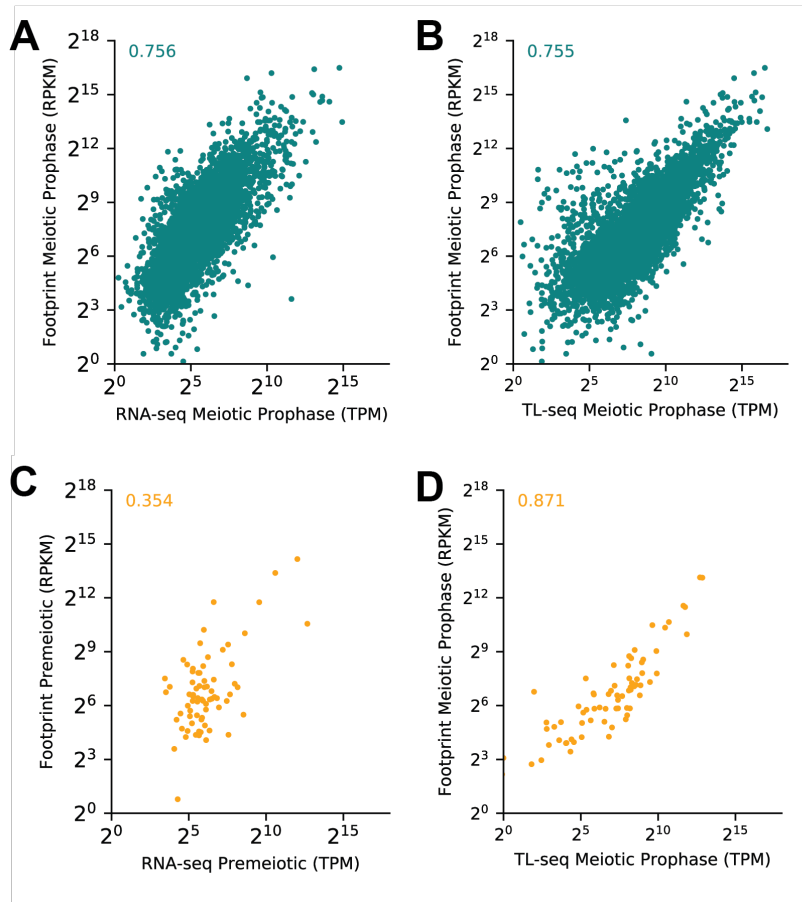
Figure 3.8: Translational regulation and transcript isoform diversity A-B) Scatterplot of translation as measured by ribosome footprints to A) RNA-seq (in duplicate) for all genes without 5'-extensions and B) TL-seq (in duplicate) for all genes with 5'-extensions. The ribosome profiling data comes from the 3 hour time point in Cheng *et al.* (2018). Cells with the *pCUP1-IME1/pCUP1-IME4* meiotic induction system (UB14584) were collected for TL-seq and RNA-seq after 4 hours in sporulation medium (SPO), 2 hours after induction of meiosis by 50 μM CuSO$_4$. The spearman rank correlation was calculated. C-D) as in A and B, but for genes with 5-extensions.

Prevalence of uORFs in meiotic LUTI mRNAs

We next investigated how LUTI mRNAs were translationally impaired. While 5'-leader mediated translational repression occurs through multiple pathways including secondary RNA structure, the presence of RNA modification, or the binding of a protein to the RNA (Hinnebusch *et al.* 2016), based on the uORF abundances in meiosis, and the case of *NDC80*, in which uORFs are essential for the translational repression of $NDC80^{LUTI}$, we hypothesized that they would play a dominant role in dampening translation from other LUTIs (Chapter 2). The number of ATG codons in the region between the LUTI TSS and the PROX TSS was used to determine uORF abundance in LUTI mRNAs. Only ATG uORF start codons were counted because although translation from uORFs with near-cognate start sites is frequent, current evidence does not indicate that they prevent translation of the downstream ORF (Brar *et al.* 2012).

Table 3.2: TE of LUTIs with < 2 uORFs

Gene	uORF number	TE fold-change
YCL057C-A	0	4.894
PLM2	0	3.411
ELO1	1	1.078
COX16	1	0.995
ULP1	1	0.057
MNE1	1	0.005

The vast majority of LUTIs had between 4 and 17 ATGs (Figure 3.9A). However, 2 genes had no ATG uORF, and 4 genes had only a single ATG initiated uORF (Table 3.2). Both genes lacking an uORF (*PLM2* and *YCL057C-A*), experienced an increase in TE (Translational Efficiency: ribosome footprints RPKM / transcript abundance RPKM) upon entry into meiosis, providing evidence that only a minority of extended transcripts were more efficiently translated (Table 3.2). In the cases when an extension contained a single uORF, the effect on translational efficiency was variable. The TE changed very little for 2 genes (*ELO1* and *COX16*), possibly because the LUTI was the minor isoform. The LUTI was the major isoform in the other 2 cases, and the TE decreased dramatically by ~200-fold for *MNE1* and ~20-fold for *ULP1* (Table 3.2), suggesting that in these 2 examples, a single uORF was sufficient to inhibit translation initiation at downstream AUGs.

The ability of a single uORF to robustly repress translation of the downstream ORF was directly and conclusively observed at the *NDC80* locus. The wild-type $NDC80^{LUTI}$ contains 9 uORFs in its leader. Our previous work demonstrated that if all 9 ATGs are mutated to ATCs (*Δ9AUG*), *NDC80* protein is translated from the LUTI mRNA (Chapter 2). Strains were constructed in which either the first, fifth, or ninth ATG was the only one left in the leader. The first uORF alone led to *NDC80* protein levels similar to what was observed in
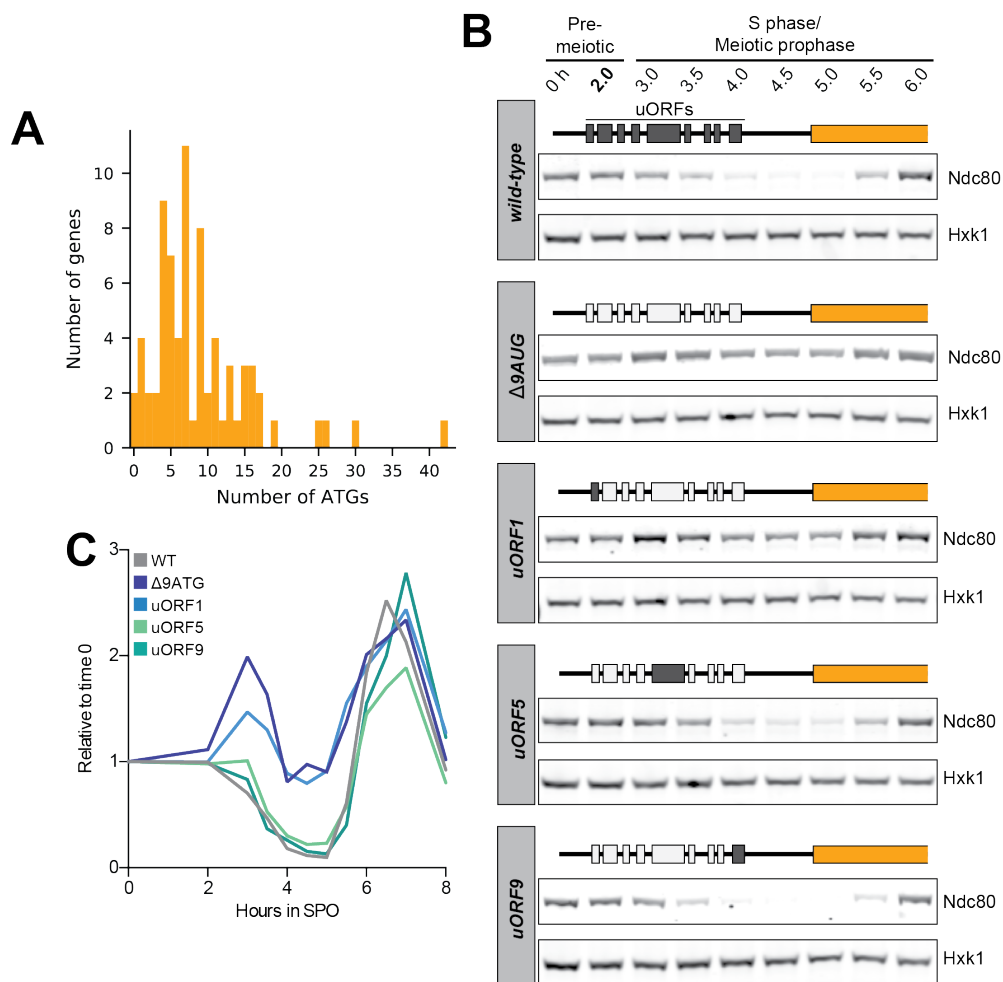
Figure 3.9: uORFs in the leaders of 5'-extended genes. A) A histogram of the number of ATGs found in the region between the proximal and the distal TSSs at loci with LUTI mRNAs. B) uORF presence affects protein production at the *NDC80* locus. Cells with the *pCUP1-IME1/pCUP1-IME4* meiotic induction system and 3V5-tagged *NDC80* were induced to undergo meiosis by 50 μM $CuSO_4$ after 2 hours in SPO. Strains used included wild-type (*UB6190*), a strain in which all of the ATGs in the $NDC80^{LUTI}$leader were mutated to ATC (*Δ9AUG: UB6183*), and strains in which only the first (*uORF1: UB10579*), fifth (*uORF5: UB10581*), or ninth (*uORF9: UB10583*) ATG was left intact. Immunoblots were performed on samples collected between 0-6 hours in SPO. Immunoblots were performed with an α-V5 antibody to recognize 3V5-tagged Ndc80. Hxk1 was the loading control. These blots represent one of two replicates. C) Quantification of the western blots in B. Signal at each time point was first normalized to the Hxk1 loading control and then to the first time point. This quantification represents one of two replicates.
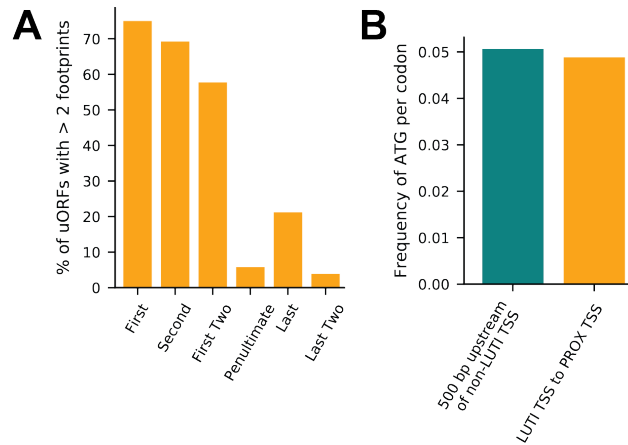
Figure 3.10: Translation of uORFs in 5'-extended gene leaders. A) Translational activity found at the first and last two uORFs in genes with at least one uORF. Translation was determined to be occurring if at least 2 reads were counted in the first 6 codon of the uORF. B) The frequency of ATGs in the region between the LUTI TSS and the proximal promoter TSS compared to the region 500 bp upstream of TSS that don't express LUTIs in early meiosis.

the *Δ9AUG* strain, but the fifth and the ninth uORFs repressed translation of *NDC80* just as well as *wild type* (Figure 3.9B-C). We concluded that a single uORF can be sufficient to cause translational repression.

Because the presence of uORFs does not always lead to repression, we set out to find another way to determine whether LUTI mRNAs are translationally repressed. As uORF number in a 5'-leader increases, the likelihood of repression at the downstream ORF also increases (Calvo *et al.* 2009; Chew *et al.* 2016; Johnstone *et al.* 2016). We predicted that if translational repression of a transcript occurs, there would be more translation over the most 5' uORFs and less translation over the uORFs closest to the annotated gene's coding region. Because uORFs are short and frequently overlapping, it can be difficult to accurately quantify their translation. Instead we determined a threshold of at least 2 footprint counts within the third to the sixth codons of a uORF to call it as translated. The call of 2 footprint was used because the background rate of reads per every 4 codons in the regions between the PROX TSS and a gene's ORF was 2.23. The first two codons were not used in this quantification in order to remove any bias due to cyclohexamide treatment of the cells. With this metric, 75% of the first uORFs in transcripts with at least 4 uORFs were translated. Compellingly, only 10% of the last 2 uORFs in those same transcripts were translated, and in less than 4% of cases was translation at both of the final uORFs observed (Figure 3.10A). Thus, the ribosomes frequently get caught up before scanning across all uORFs. This is consistent with the observation that ATG frequency was not higher in the 5' extensions compared to the 500 bp upstream of genes not expressing LUTI mRNAs in early meiosis (Figure 3.10B). If indeed LUTI mRNAs do play important and functional roles in mediating meiotic gene
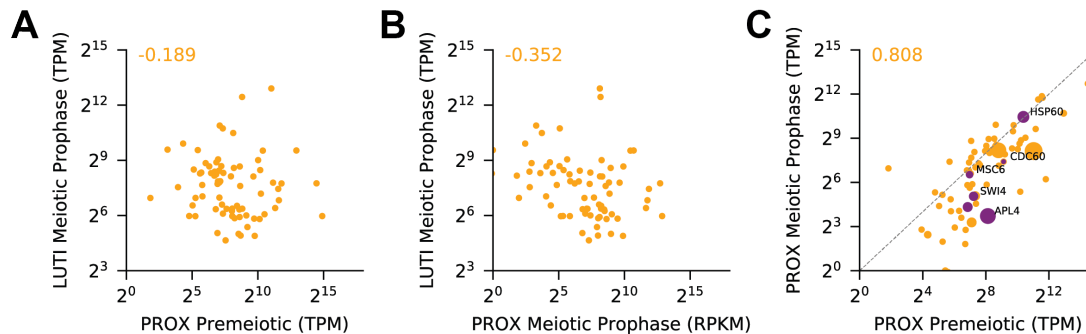
Figure 3.11: The relationship between LUTI and PROX transcripts. A) Scatterplot of LUTI abundance in meiotic prophase vs PROX abundance in premeiotic cells with the *pCUP1-IME1/pCUP1-IME4* meiotic induction system (UB14584). The premeiotic time point was taken after 2 hours in SPO, and the meiotic prophase timepoint was taken after 4 hours in SPO, that is 2 hours after 50 µM $CuSO_4$ addition. The values plotted are from the mean of 3 replicates. They are displayed in transcripts per million as determined by CAGEr. The spearman correlation coefficient is in the upper left corner. B) The same as A except the PROX transcript abundance was calculated from the early meiotic timepoint. C) Visualization of the change in PROX transcript abundance as the cells enter meiosis. Scatterplot of PROX abundance in meiotic prophase or before meiotic entry as transcripts per million (TPM). The size of each point correlates to the LUTI abundance in early meiosis. The spearman correlation coefficient is in the upper left corner. Purple points indicate genes of interest for more in-depth analysis.

expression, the lack of uORF selection would indicate that the natural frequency of ATGs in intergenic regions is sufficient to result in the necessary degree of translational inhibition. Given that translation is not occurring over most of the uORFs closest to a gene's coding region, we conclude that uORFs are found in abundance in the leaders of most LUTIs.

Variable transcriptional repression by LUTI mRNAs

In addition to being translationally impaired, transcription of previously identified LUTI mRNAs prevents transcription initiation at PROX promoters. To assess the degree of transcriptional repression in the set of LUTIs identified by this study, the abundances of LUTI and PROX transcripts were measured by TL-seq. Though LUTI levels in meiotic prophase had no correlation with the abundance of PROX transcripts in a premeiotic stage (Figure 3.11A), a significant negative relationship (spearman r= -0.352, p-value = 2.08 x 10-3) was observed meiotic prophase, associating LUTI expression with a decrease in PROX isoform level (Figure 3.11B). In addition, the PROX transcript abundance in meiotic prophase was less than in premeiotic stage for a large number of genes, further demonstrating their transcriptional repression by LUTIs (Figure 3.11C).
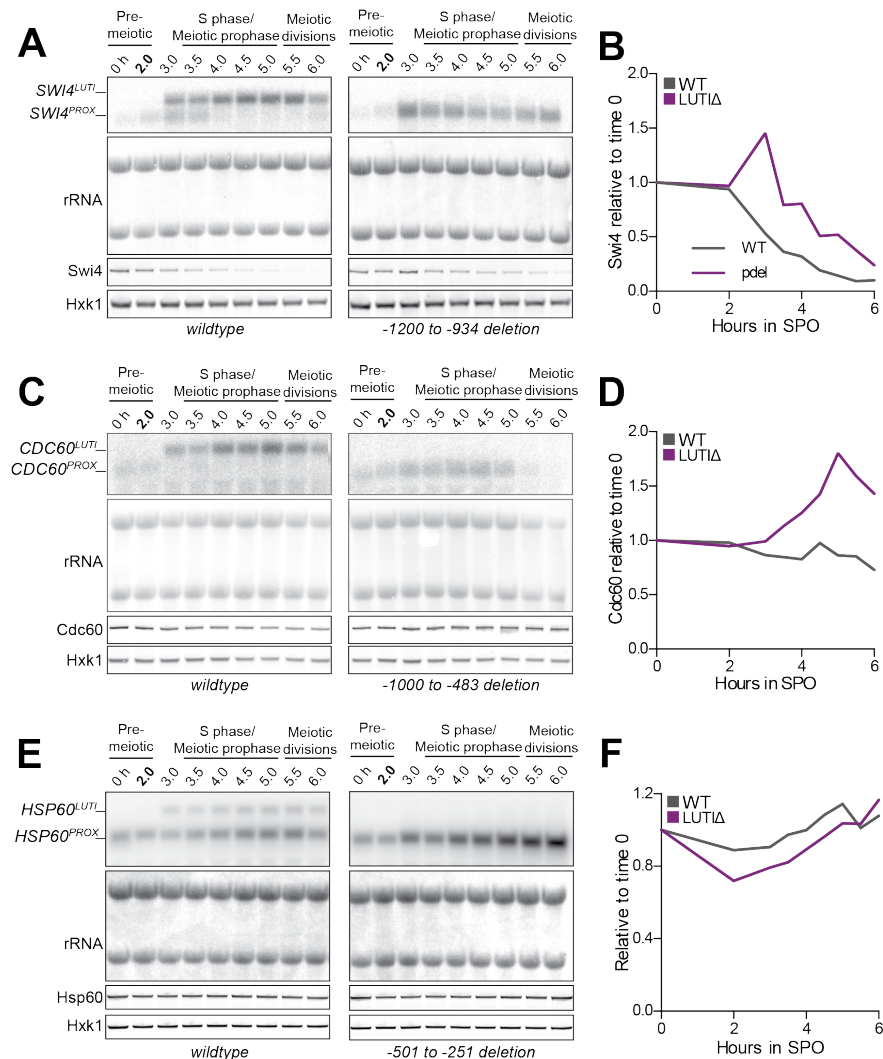
Figure 3.12: Investigation of LUTI deletion in candidate genes: Part 1. A-F) Cells with the *pCUP1-IME1/pCUP1-IME4* meiotic induction system were induced to enter meiosis with 50 μM CuSO$_4$ after 2 hours in SPO. A-B) Were performed with a *SWI4-3V5* tagged strain with either *wild-type* (*UB18175*) or deleted (*UB18176*) LUTI promoters, C-D) Were performed with a *CDC60-3V5* tagged strain with either *wild-type* (*UB18185*) or deleted (*UB18186*) LUTI promoters, E-F) Were performed with a *HSP60-3V5* tagged strain with either *wild-type* (*UB18336*) or deleted (*UB18188*) LUTI promoters. A, C, E) RNA and Immuno-blots were performed on samples collected between 0-6 hours in SPO. RNA blots were performed with a probe specific for 3V5 and it's linker. Methylene blue detection of rRNA bands was the loading control. Immunoblots were performed with an α-V5 antibody to recognize 3V5-tagged proteins. Hxk1 was the loading control. These blots represent one of two replicates for *CDC60* and *SWI4*, and a single replicate for *HSP60*. B, D, F) Quantification of A, C, E. 3V5 signal was quantified relative to Hxk1 and then relative to 0 hr. Represents one of two replicates for *CDC60* and *SWI4*, and a single replicate for *HSP60*
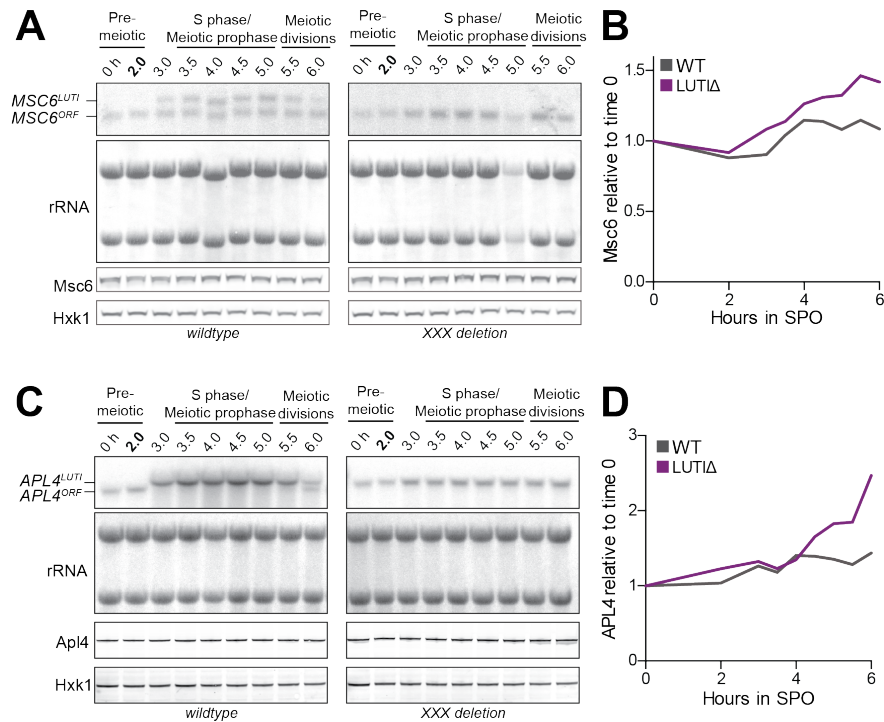
Figure 3.13: Investigation of LUTI deletion in candidate genes: Part 2. A-D) Cells with the *pCUP1-IME1/pCUP1-IME4* meiotic induction system were induced to enter meiosis with 50 μM CuSO$_4$ after 2 hours in SPO. A-B) Were performed with a *MSC6-3V5* tagged strain with either *wild-type* (*UB18238*) or deleted (*UB18190*) LUTI promoters, C-D) Were performed with a *APL4-3V5* tagged strain with either *wild-type* (*UB18539*) or deleted (*UB18181*) LUTI promoters. A, C) RNA and Immuno-blots were performed on samples collected between 0-6 hours in SPO. RNA blots were performed with a probe specific for 3V5 and it's linker. Methylene blue detection of rRNA bands was the loading control. Immunoblots were performed with an α-V5 antibody to recognize 3V5-tagged proteins. Hxk1 was the loading control. These blots represent one of two replicates. B, D) Quantification of A and C. 3V5 signal was quantified relative to Hxk1 and then relative to 0 hr. Represents one of two replicates

To parse out how causative the relationship was between LUTI and PROX expression is, five genes were selected for more in-depth analysis. Two of them (*SWI4* and *APL4*) had very strongly repressed PROX transcripts in early meiosis, two (*MSC6* and *HSP60*) had PROX transcripts present at the same level in both time points, and one (*CDC60*) had an intermediate amount of repression (Figure 3.11C). For each gene, 3V5-tagged wild type and LUTI deletion strains were constructed. Cells were synchronized throughout meiosis. RNA blots and immunoblots were performed to track transcript isoforms and protein abundance. In all instances, the LUTI deletion correlated with an increase in PROX transcript abundance, but not necessarily protein abundance (Figure 3.12A-C, Figure 3.13A-B). For *SWI4* and *CDC60*, deletion of the LUTI led to an increase in the corresponding protein's abundance. Despite strong LUTI-dependent transcriptional repression, deletion of $APL4^{LUTI}$ does not lead to higher Apl4 levels, possibly because $APL4^{LUTI}$ contains only 3 ATG uORFs, all of which are translated, and/or because Apl4 protein levels are regulated post-translationally. For *MSC6* and *HSP60* the deletion of the LUTI transcript has no effect on Msc6 or Hsp60 levels, respectively. It's apparent that LUTI transcription does lead to repression of the PROX transcript in all of the above cases, including *HSP60* and *MSC6*. However, even if transcriptional repression occurs, additional layers of gene regulation in coordination determine the final protein output.

The role of chromatin in LUTI-mediated transcriptional repression

With the knowledge that transcription of a LUTI mRNA can play a role in repressing downstream transcription at some loci, but not others, we set out to understand what differentiates the two classes. It is established that in both yeast and humans, the presence of a LUTI mRNA leads to increased H3K36me3 over the PROX transcript promoter (Chapter 2)(Hollerer *et al.* 2019). In yeast, an increase in H3K4me2 was also observed at the *NDC80* locus upon $NDC80^{LUTI}$ induction (Chapter 2). Both marks were necessary for LUTI-mediated repression of $NDC80^{PROX}$ by $NDC80^{LUTI}$ (Chapter 2). H3K4me2 and H3K36me3 are found at sites of active transcription, with H3K4me2 present in more abundance towards the 5' end of a transcribed region (Kirmizis *et al.* 2007; Liu *et al.* 2005; Pokholok *et al.* 2005). Despite localizing to regions undergoing active transcription, they are not necessarily associated with gene activation. Particularly in budding yeast, H3K36me3 is known to play a role in repressing transcription initiation within gene bodies (Carroll 2008; Keogh *et al.* 2005; Li *et al.* 2007). The role of H3K4me2 is less well understood, but reports indicate that it plays a role in the induction kinetics of genes at loci with overlapping transcription (Kim & Buratowski 2009; Kim *et al.* 2012; Pijnappel *et al.* 2001).

The relationship between LUTI-mediated changes to the H3K36me3 and H3K4me2 levels over a PROX promoter and repression of PROX transcription was investigated by ChIP-seq in the premeiotic stage and meiotic prophase. In premeiotic cells, the LUTI genes appeared quite similar to non-LUTI genes for both modifications with a dis-enrichment just upstream
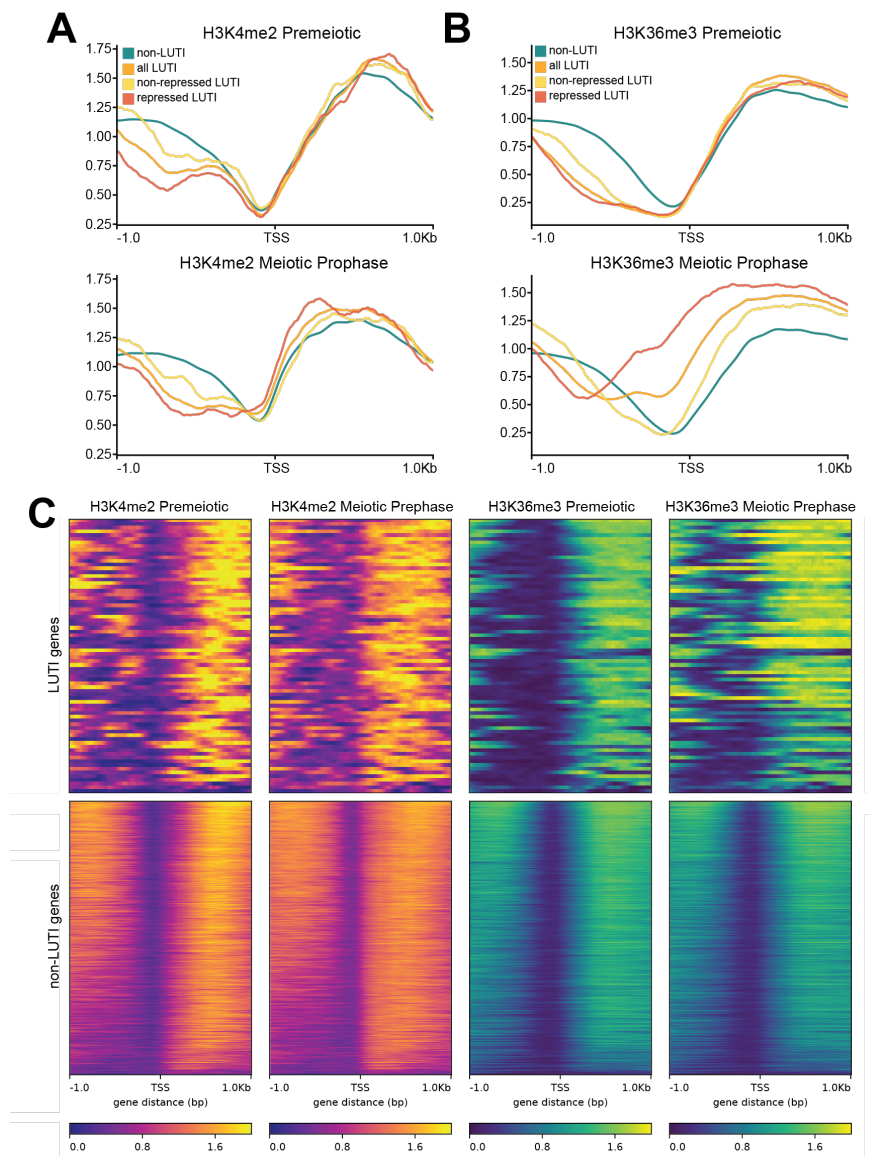
Figure 3.14: H3K4me2, H3K36me3 and repression of the PROX transcript. A) Metagene analysis
ChIP-seq of H3K4me2. The ChIP was performed in cells with the *pCUP1-IME1/pCUP1-IME4*
meiotic induction system (UB14584). The premeiotic time point was taken after 2 hours in SPO,
and the meiotic prophase time point was taken after 4 hours in SPO, 2 hours after 50 μM $CuSO_4$
addition. Repressed LUTI genes include genes with a fold-change in PROX abundance of < 0.25
between the early meiotic and starvation timepoints (n=22). Non-repressed LUTI genes are those
with a fold-change in PROX abundance of > 1 (n=19). The images are from 1 of 3 replicates. B)
Same as A but for H3K36me3. C) Heatmap of ChIP-seq from A) H3K4me2 (left) and B) H3K36me3
(right) in premeiotic cells and cells during meiotic prophase. The plot is centered around the TSS
of the proximal promoter. The images are from of 1 of 3 replicates.
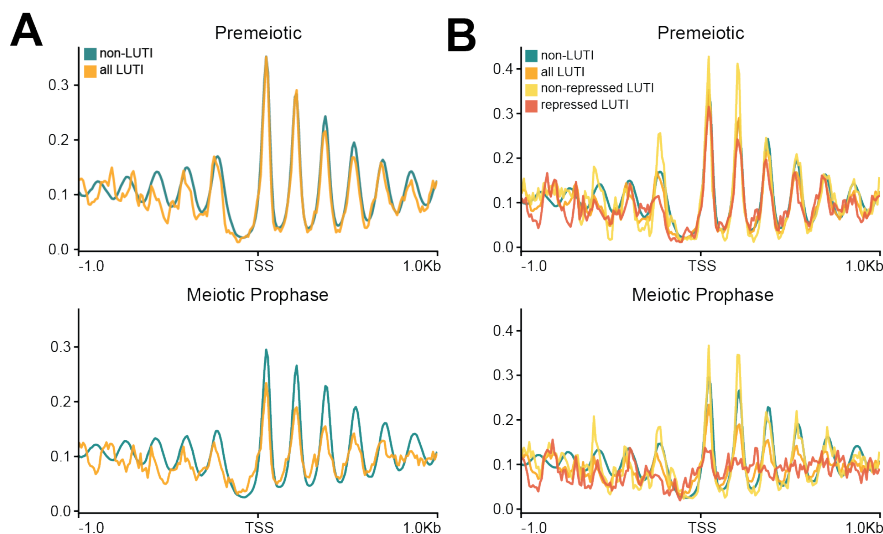
Figure 3.15: Nucleosome position and repression of the PROX transcript. A) Metagene analysis MNase-seq signal in the premeiotic stage and meiotic prophase. The MNase DNA digestion was performed on cells with the *pCUP1-IME1/pCUP1-IME4* meiotic induction system(UB14584). The premeiotic time point was taken after 2 hours in SPO, and the meiotic prophase time point was taken after 4 hours in SPO, that is 2 hours after 50 μM $CuSO_4$ addition. The plot is centered around the TSS of the proximal promoter. The images are from 1 of 3 replicates. B) Same as A but including a repressed subset of LUTI genes with a fold-change in PROX abundance of < 0.25 between the early meiotic and starvation timepoints (n=22), and a non-repressed subset of LUTI genes with a fold-change in PROX abundance of > 1 (n=19).

of the PROX TSS (Figure 3.14A-C). This dis-enrichment of H3K36me3 was lost in meiotic prophase specifically over the promoters of those genes expressing LUTI mRNAs. Most convincingly, those genes whose PROX transcript wass most repressed (log2FoldChange $<$ -2, n=22) had the highest levels of H3K36me3 over their proximal TSS. Further, genes that had associated LUTI mRNAs, but did not experience a decrease in the abundance of the PROX transcript (log2FoldChange > 0, n=19), had only a minor increase in H3K36me3 levels (Figure 3.14B-C). For H3K4me2, a moderate increase in the chromatin modification was observed over the PROX promoters in meiotic prophase, and the increase was stronger in those genes that were most repressed (Figure 3.14A,C). Thus, it appears that H3K4me2 is important only in some instances, such as in the case for *NDC80*. However unlike H3K36me3, it is not a major player in LUTI-based repression.

In addition to methylation states of the histones around the PROX TSS, the presence of nucleosomes can occlude the binding of transcription factors and other machinery required for transcription initiation (Klemm *et al.* 2019; Venkatesh & Workman 2015). Evidence from the case of $NDC80^{LUTI}$ indicates that upstream transcription leads to nucleosome repositioning around the proximal promoter thereby shrinking the nucleosome-depleted region
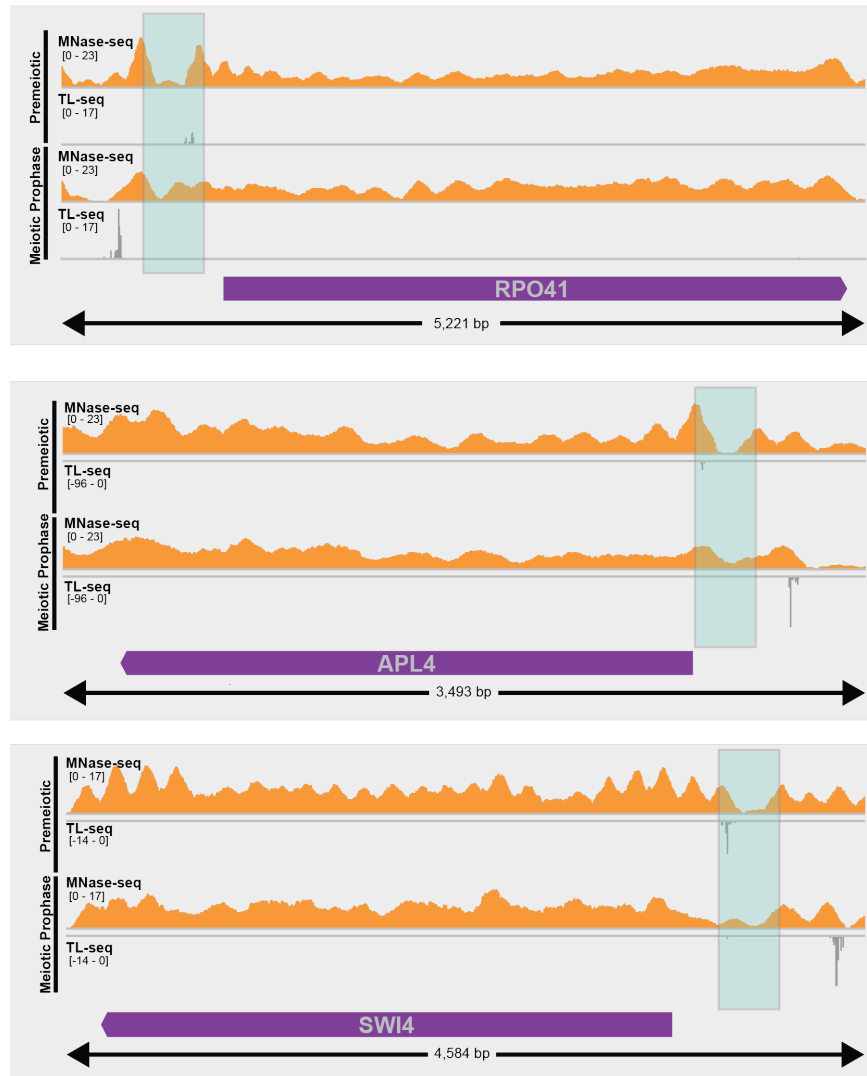
Figure 3.16: Genome browser views of MNase-seq and TL-seq. The TL-seq and MNase DNA digestion was performed on cells with the *pCUP1-IME1/pCUP1-IME4* meiotic induction system(UB14584). The premeiotic time point was taken after 2 hours in SPO, and the meiotic prophase time point was taken after 4 hours in SPO, that is 2 hours after 50 μM $CuSO_4$ addition. Three candidate genes, all of which are robustly transcriptionally repressed are shown. The green box denotes the rough location of the PROX transcript's promoter.

(NDR) (Chapter 2). By performing micrococcal nuclease sequencing (MNase-seq), changes to nucleosome positions were tracked between premeiotic stage and meiotic prophase. As the cells enter meiosis, the nucleosome peaks decreased and signal in the valleys increased specifically for the LUTI subset of genes, thus increasing the fuzziness of nucleosome at loci with LUTIs (Figure 3.15A-B). The effect was strongest for loci with the greatest degree of PROX transcript repression (Figure 3.15B). In that subset of 22 genes, the nucleosome position was so disrupted that a consensus nucleosome periodicity could not be identified (Figure 3.15B). Examining such a small subset of genes might make it hard to observe strong periodicity, however, in a similarly small subset of genes (19), loci with non-repressive LUTIs, periodic nucleosome positioning was still observed (Figure 3.15B). Therefore the complete lack of periodicity at loci with repressive LUTIs, was most likely due to variability in the extent of repositioning at each locus (Figure 3.16). For this reason, we concluded that robust nucleosome repositioning occurs over the promoters of genes with repressive LUTIs.

The features defining LUTI-based transcriptional repression

The observation that increases in H3K36me3 and nucleosome occupancy over proximal promoters were associated with greater repression of PROX transcription fit our proposed model, but we expected that other variables could play a role in LUTI-based transcriptional repression. Stemming from previous work showing that the distance between promoters with overlapping transcription affects the mechanism of transcriptional repression in cells undergoing carbon source shifts, we considered the importance of the distance between the PROX and the LUTI TSSs in mediating PROX repression (Kim *et al.* 2017). The LUTI abundance and the length of the gene were also considered, as were changes to the +1 and 1 nucleosome positions and fuzziness. We found that an increase in H3K36me3 (spearman r=-0.451, p-value=5.44 x 10-5), high LUTI levels (spearman r=-0.446, p-value=6.90 x 10-5), +1 nucleosome peak moving toward the NDR (spearman r= -0.412, p-value=2.64 x 10-4), and an increase in +1 nucleosome fuzziness (spearman r= -0.328, p-value=4.44 x 10-3) all significantly anti-correlated with the log2 fold-change of PROX transcript abundance. However, changes at the −1 nucleosome had no significant correlation, neither did changes in H3K4me2. Altogether, these analyses helped distinguish some important factors involved in LUTI-based transcriptional repression.

Interestingly, we found that the longer a gene's coding sequence was, the more likely it was to be repressed by a LUTI (Figure 3.17A). Coincidentally, upon further investigation, it was observed that in the set of genes with early meiotic LUTIs, the shorter genes had higher PROX abundances than did longer genes (Figure 3.17B). Could it be that the promoters of strongly expressed genes are better able to continue transcribing their gene products even in the presence of LUTI mRNAs? This will be a line of future investigation to better understand all features involved in the transcriptional aspects of LUTI-mediated gene repression.
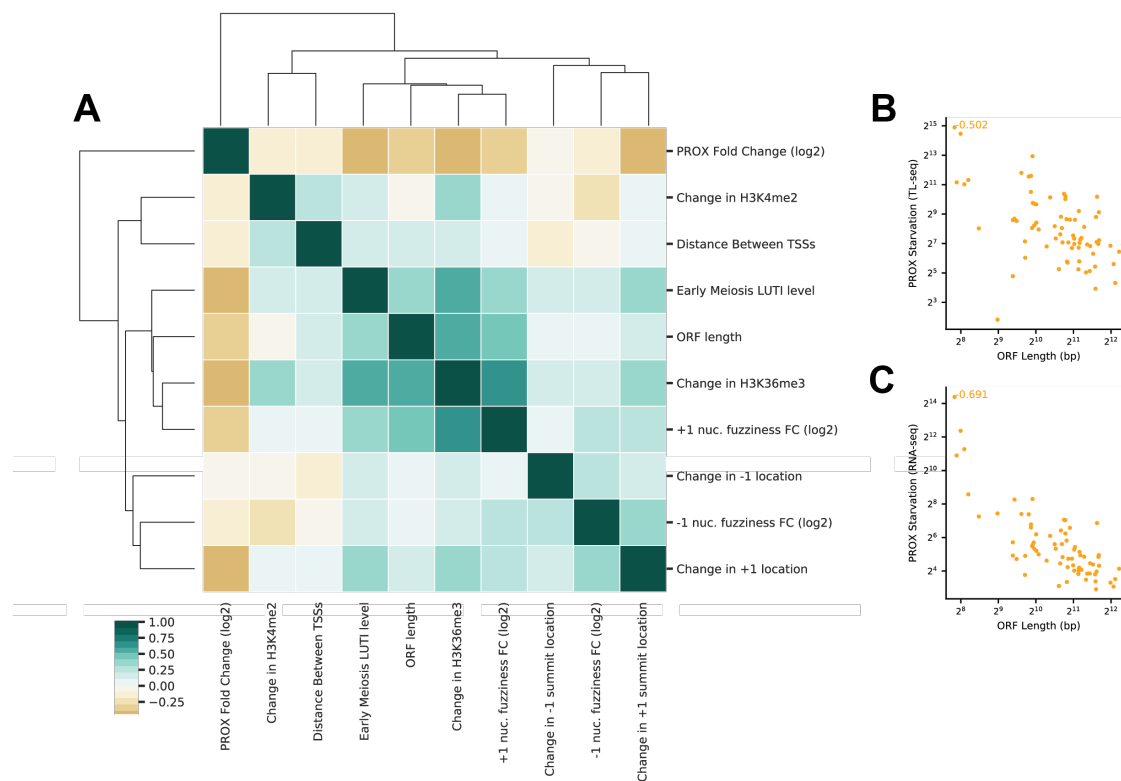
Figure 3.17: Features most associated with LUTI-based repression. A) Cluster map of features associated with LUTI-mediated repression. Features were clustered by spearman correlation coefficient. B) Scatter plot of the correlation between ORF length and PROX abundance in premeiotic cells as quantified by TL-seq. The spearman correlation coefficient is displayed. C) Same as B, but the PROX abundance was calculated from RNA-seq

3.3   Discussion

LUTIs are a recently identified class of mRNAs (Chapter 2) (Cheng *et al.* 2018; Hollerer *et al.* 2019; Van Dalfsen *et al.* 2018). Their 5'-extended leaders contain features, namely uORFs, that prevent the translation of the downstream ORF, and transcription of the LUTI represses transcription from the proximal gene promoter. Consequently, LUTIs are associated with a decrease in protein abundance. Here however, we show that LUTIs occur more frequently than does LUTI-based gene repression. Although we find strong evidence for the translational inhibition of the vast majority of LUTIs, the level of transcriptional PROX transcript repression by a LUTI is variable. We discovered that whether a LUTI leads to the repression of downstream promoters depends on how abundant the LUTI is, the degree of increase in H3K36me3 over the PROX promoter, and changes to the +1 nucleosome position. Our study provides the first genome-wide investigation of the key features involved

in LUTI-based repression.

Developmental regulation of LUTIs

The combined use of two sequencing techniques, TL-seq and Nanopore sequencing, allowed us to make the first genome-wide claims as to which observed 5'-extensions were part of full-length mRNAs, not just short intergenic transcripts (Figure 1B). Direct RNA and cDNA long-read sequencing hold immense promise for future studies of LUTIs and other transcript isoform biology (Garalde *et al.* 2018; Sharon *et al.* 2013). In addition to assisting in the identification new splice isoforms, TSSs, and sites of transcription termination, some reports have indicated accurate transcript isoform quantification using mixtures of spike-in RNAs (Byrne *et al.* 2017; Garalde *et al.* 2018; Oikonomopoulos *et al.* 2016). Unfortunately, these studies were performed with transcripts less than 2.5 kb. In our hands, a 3' bias was observed for most transcripts, but it was strongest in transcripts greater than 2.5 kb of which there are numerous LUTIs. This 3' bias was significant enough to prevent the accurate quantification of 5'-extended isoforms since it would have completely relied on accurate TSS identification. The use of TL-seq for quantification allowed us to overcome this hurdle. Through the systematic identification of 5'-extended, ORF-containing transcripts by direct RNA sequencing and quantification of TSSs by sequencing just the most 5'-ends of transcripts, we found 74 genes with early meiotic LUTI mRNAs.

The DNA binding protein Ume6 has a URS1 binding motif and is enriched in the promoter for 46 of these LUTIs, 33 of which have a conserved Ume6 binding motif present across the *sensu stricto* genus (Figure1D-H). We posit, that the conservation of the regulatory binding site in such a large subset of LUTI mRNA promoters indicates that they alter gene expression in a manner functionally important for meiotic progression and gamete fitness. Though Ume6 acts as a transcription repressor during vegetative growth, it interacts with the master meiotic transcription factor Ime1 in order to activate the early meiotic transcriptome (Bowdish *et al.* 1995). It remains to be seen which of the above-identified sites recruit Ime1 directly, but this subset of LUTIs depends entirely on the induction of *IME1*, making it a probable regulator. The linking of many highly conserved LUTIs to a meiosis-specific regulator inherently ties them to the progression of a developmental program: budding yeast meiosis. We predict that LUTIs play roles during other developmental processes because it is during these times that a cell demands large and precise changes in gene expression. By employing LUTIs, a single transcription factor key to a developmental program can both turn on the expression of genes and turn off gene expression at loci with LUTIs in a coordinated and timely manner. In addition to the Ume6-regulated promoters, we identified other LUTIs that are likely regulated by other transcription factors. We predict that no other motifs were found because the other regulators may each only play a role in turning on one or two LUTIs at this time in meiosis.

uORF-mediated translational inhibition of LUTI mRNAs

It is accepted that uORFs are ubiquitously found in walks of life from yeast to humans. They are functionally associated with translational inhibition by sequestering the ribosome at the 5' end of a transcript (Zhang *et al.* 2019). Various features including initiation sequence context, the distance between a uORF and the coding sequence, and uORF number in a 5' transcript leader can all affect the degree of translational repression, but reports as to the significance of each feature vary (Calvo *et al.* 2009; Chew *et al.* 2016; Johnstone *et al.* 2016). What is clear is that not all uORFs are created equal. We confirm that a single uORF can lead to repression at the *NDC80* locus, but that it depends on which uORF. The uORF closest to the LUTI TSS (uORF1) doesn't repress translation of the *NDC80* ORF despite evidence from ribosome profiling that uORF1 is well translated. uORFs 5 and 9 both lead to robust repression even though no translation of uORF 9 is observed in meiotic cells (Figure 3C) (Brar *et al.* 2012). This is consistent with the genome-wide observation that greater distances between a uORF and a coding sequence correlate with greater translation efficiency of the ORF and reduced translational repression (Chew *et al.* 2016; Johnstone *et al.* 2016). In the context of the most well-known case of uORF-mediated repression, *GCN4*, the distances between the 4 uORFs and the ORF start sequence matters greatly. That is because upon amino acid starvation, the concentration of the ternary complex, a factor required for ribosome re-initiation is decreased. This results in an extended scanning time after the uORF1 before re-initiation can occur. Ultimately, the repressive uORF4 is skipped and *GCN4* is translated (Hinnebusch 1997). Interestingly, the position of the $NDC80^{LUTI}$ uORF9 starts 151 bps upstream of the *NDC80* coding domain, almost exactly the same position as uORF4 with respect to the *GCN4* coding domain. Harnessing the power of yeast genetics, it could be worthwhile to further study the role of ternary complex or other trans-factors in ensuring which uORFs are integral to the translational repression of LUTIs.

Until those features are more fully developed, one cannot simply equate the presence of uORFs with translational repression. The presence of translation over uORFs and a corresponding decrease in TE upon said translation, as used previously, is a good indication that the uORFs are playing a functional role, but it does not indicate the extent of the translational regulation. Further, TE measurements at loci with LUTI mRNAs are further complicated by the presence of PROX transcripts (Figure 4C). Observing that the final 2 uORFs in LUTI mRNAs with 4 uORFs or more are devoid of ribosome footprints 5.8% and 21.2% of the time for the penultimate and the last uORF, respectively, supporting a strong shutdown of translation (Figure 3D). The high correlation between PROX abundance and footprints over the ORF was also a strong indication that the LUTIs provide minor if any contributions to translation observed for ORFs in meiotic prophase. Future use of TL-seq in combination with ribosome profiling will allow for identification of more instances in which apparent translational regulation is due to transcript isoform toggling and not true translational regulation of a single isoform. Combined with a lack of uORF selection in the

5'-extensions of LUTIs, our study has provided the most comprehensive demonstration that uORFs in LUTIs do not just dampen ORF translation, they almost entirely repress it in the majority of cases.

A path towards predicting LUTI-based transcriptional repression

In the first genome-wide investigation to uncover abundant 5'-extended transcript isoforms during budding yeast meiosis, it was observed that only $\sim$ half of them correlated poorly with translation levels (Brar *et al.* 2012). If, as discussed above, almost all 5'-extended transcripts do not translate their ORFs, why does translation not decrease upon LUTI transcription more frequently? As it turns out, robust transcriptional repression of a PROX transcript by a LUTI mRNA occurs far less frequently than does translational repression of the LUTI itself (Figure 4C). In fact, only 22/74 early meiotic LUTIs correlate with a PROX transcript fold-change < 0.25. We determined that high LUTI abundance, increased H3K36me3, increased fuzziness of the $+1$ nucleosome, and longer ORFs are all significantly associated with gene repression.

Of the features found to be important, most related in some way to changes in the chromatin landscape at the PROX promoter. Unexpectedly, increased H3K4me2 was not associated with repression of the PROX promoter genome-wide even though it is necessary for repression at the *NDC80* locus. We predict that only a subset of loci are dependent on H3K4me2 for repression or H3K4me2 may help to delay the kinetics of PROX re-expression later in meiosis, as it does at sites of overlapping transcription upon shifts in carbon source (Kim *et al.* 2012; Kim *et al.* 2017). Both H3K4me2 and H3K36me3 are implicated in gene repression at sites of overlapping transcription. Repression due to overlapping transcription was reported to occur on average when the TSSs were 0.9 kb apart if mediated by the Set1-H3K4me2-Set3C pathway and 2.0 kb apart when mediated by the Set2-H3K36me3 pathway (Kim *et al.* 2017; Woo *et al.* 2017). We had expected to find that LUTIs starting further from the PROX TSS would rely on H3K36me3-mediated repression and those starting closer would rely on H3K4me2. Given that H3K4me2 had no significant relationship with repression, this was not testable. However, the average distances reported above are both much longer than the 536 bp mean distance between LUTI and PROX TSSs identified in this study. This led us to hypothesize that longer distances would be associated with greater PROX down regulation. We observed no evidence of such a relationship (Figure 3.17). We expect that, it is quite possible for a LUTI of the same strength to affect PROX repression differently as its TSS is moved with respect to the PROX TSS, but in our dataset, the level of LUTI expression may mask any of those effects as it plays a far more deterministic role.

Even with features correlating significantly with repression, there were still instances in which an abundant LUTI, increases in H3K36me3, and changes to the $+1$ nucleosome were observed with no repression of the PROX transcript. We hypothesize that something about the PROX promoter is different for these genes. Inspired by the surprise correlation between

repression and ORF length, which possibly stems from the trend for shorter transcripts to be more abundant, we predict that more robustly expressed PROX transcript become more difficult to repress by LUTI mRNAs. This will be investigated further in the future.

Conclusion

In summary, we identified a group of LUTIs expressed in a coordinated manner during the early stages of budding yeast meiosis. Almost all of them contain uORFs and are severely translationally inhibited. However, they do not ubiquitously lead to repression of PROX promoters. Rather, an interplay between the strength of the promoters, the chromatin landscape, and possibly other yet to be discovered features interact to decide the transcriptional output. Building up more datasets with increased numbers of assays applied to the same developmental program will provide more and more LUTI mRNA targets to study. It is possible that in the future enough features will be defined to allow for the building of a model that would allow prediction of repressive LUTI mRNAs where the data are sparse.

In addition to a model to predict LUTI-based repression, it still remains to be seen, what happens if a LUTIs function is disturbed. The LUTIs present during the UPR in budding yeast together dial down expression of genes involved in cellular respiration thus improving the fitness of cells undergoing such a stress (Van Dalfsen *et al.* 2018). $NDC80^{LUTI}$ is essential for proper meiotic chromosome segregation (Chapter 2). However, the vast numbers of other LUTIs identified in budding yeast are rife for further investigation. Additionally, the presence of a LUTI at the MDM2 locus (Hollerer *et al.* 2019) in human cells lines, opens the door for more studies into how significant LUTIs are in human development and disease.

Chapter 4

Chapter 4: Conclusions and future of the field

The following chapter contains published material from a publications that I am the first author on (Tresenrider & Ünal 2018). The article is distributed under the terms of the Creative Commons Attribution License (CC BY 4.0), which permits unrestricted use and redistribution provided that the original author and source are credited.

4.1   A new wrinkle in the central dogma

The teaching of central dogma states that one gene serves as a template for one type of mRNA. This mRNA in turn acts as a template for the production of a single protein product. It is widely accepted that this view is simplistic and does not capture the biological complexity of gene regulation. We now know that alternative splicing can produce hundreds of mRNA isoforms from a single gene. We know that non-coding RNAs make up a greater portion of the transcriptome than coding RNAs, and that these untranslated RNAs, such as micro-RNAs or long non-coding RNAs play essential roles in organisms across eukaryotes (Lee *et al.* 1993; Brown *et al.* 1992; Wightman *et al.* 1993; Mercer *et al.* 2009). The mechanisms by which long non-coding RNAs affect gene regulation are varied. They can act in *cis* or *trans*, as antisense or intergenic transcripts, they can be either repressive or activating, be structural, alter chromatin states, or cause gene looping (Marchese *et al.* 2017; Solé *et al.* 2015; Wu *et al.* 2017). The unifying characteristic is that they are not translated. However, we are still very much attached to the notion that mRNA molecules (i.e. those that are capped, polyadenylated, and engaged with the ribosome) are translated. Their translation may be temporally delayed, as is the case for *CLB3* and *SSP2* in budding yeast meiosis, or their translation may only occur under specific circumstances as is the case for the uORF repressed *GCN4* and ATF4 transcripts (Mueller & Hinnebusch 1986; Harding *et al.* 2000; Jin & Neiman 2016). In other words, translational repression is widely viewed as a switch-like

mechanism, where translation of the ORF is repressed under certain conditions, but this repression can be bypassed under other conditions.

In this dissertation, we uncovered a novel mechanism, where a developmental transcription factor induces the expression of mRNAs that serve a purely regulatory function. These mRNAs are *never* translated into functional proteins due to the uORFs in their 5'-leader regions (Chapter 2, Chapter 3). Instead, they serve to inactivate a gene through an integrated transcriptional and translational mechanism (Chapter 2, Chapter 3). This new insight challenges the assumption that mRNA molecules must produce the gene product encoded in their open reading frames and provides a fresh perspective on gene regulation.

4.2   A 5'-extended mRNA represses kinetochore function

Our studies of kinetochore regulation during meiosis in budding yeast, initially confirmed what we and others had seen previously: the essential kinetochore protein Ndc80 is downregulated during meiotic S-phase and prophase (Chapter 2) (Asakawa *et al.* 2005; Kim *et al.* 2013; Meyer *et al.* 2015; Miller *et al.* 2012; Sun *et al.* 2011). This assists in kinetochore remodeling which allows homologous chromosomes to be segregated in meiosis I. A deeper investigation into the mechanism by which the Ndc80 protein level decreases led us to discover an initially counterintuitive mechanism by which cells can downregulate gene expression.

At the *NDC80* locus, a 5'-extended transcript isoform is expressed exclusively during meiosis. It is developmentally regulated by the master meiotic transcription factor Ime1 and its binding partner Ume6 (Chapter 2) (Bowdish *et al.* 1995; Park *et al.* 1992; Washburn & Esposito 2001). The extended transcript contains 9 uORFs in addition to the entire *NDC80* coding sequence. Its translation status, determined by ribosome profiling, confirms that the extended transcript is engaged with the ribosome sequence (Chapter 2) (Brar *et al.* 2012; Miller *et al.* 2012). By all accounts, this RNA molecule would be considered an mRNA. Intriguingly, due to the uORFs, its ORF cannot be translated. For this reason, we have termed this new class of RNAs, LUTI mRNAs, for long undecoded transcript isoform mRNAs and for which *NDC80* is the founding member. The undecoded *NDC80* isoform is referred to as $NDC80^{LUTI}$, while the transcript produced from the more promoter, which is translated into Ndc80 protein, is referred to as $NDC80^{PROX}$.

In addition to being translationally repressed, $NDC80^{LUTI}$ expression blocks transcription initiation at the $NDC80^{PROX}$ promoter. When Ime1 induces $NDC80^{LUTI}$, RNA polymerase II (pol II) complex is recruited to the distal promoter and transcription elongation occurs across the proximal $NDC80^{PROX}$ promoter. This causes an increase in repressive histone modifications, namely H3 lysine 4 dimethylation (H3K4me2) and H3 lysine 36 trimethylation (H3K36me3), over what had previously been an active promoter (Chapter 2). Both

marks are co-transcriptionally deposited in a pol II-dependent manner and are involved in repressing cryptic transcription initiation within gene bodies (Carrozza *et al.* 2005; Keogh *et al.* 2005; Kim & Buratowski 2009; Ng *et al.* 2003; Li *et al.* 2003; Xiao *et al.* 2003). Their deposition at the $NDC80^{PROX}$ promoter however correlates with the repression of a previously active promoter. Furthermore, an increase in nucleosome occupancy upon $NDC80^{LUTI}$ expression is dependent on the conserved chromatin associated proteins: Set3, a member of the Set3C histone deacetylase complex, and Set2, an H3K36me3 methyltransferase (Strahl *et al.* 2002; Pijnappel *et al.* 2001). When deposited over the canonical *NDC80* promoter, H3K4me2 and H3K36me3 together increase nucleosome occupancy, which in turn inhibits transcription initiation at the $NDC80^{PROX}$ promoter (Chapter 2).

When we consider both the transcriptional and translational methods utilized by the cell to affect Ndc80 protein levels together, a fuller picture of gene regulation at this locus begins to emerge. An mRNA transcript which the ribosome does not translate into full-length functional protein still serves a key function. The act of its transcription leads to the recruitment of chromatin modifying enzymesthought to prevent spurious initiation within gene bodies. In the case of *NDC80*, initiation of transcription upstream of active promoters serves as a *cis*-mediated mechanism to prevent transcription of downstream mRNA isoforms during a developmental program. This unexpected one-two punch of both transcriptional and translational repression in such a non-canonical manner called for further investigation for other instances by which the integration of multiple steps in gene regulation determines the final gene output (Chapter 3) (Cheng *et al.* 2018; Van Dalfsen *et al.* 2018; Hollerer *et al.* 2019).

Later in meiosis, kinetochores must be active in order to segregate chromosomes. This occurs through the re-expression of Ndc80 protein. Instead of changing the translational status of $NDC80^{LUTI}$ as in the case of *GCN4*, the cells evolved a different solution to produce Ndc80 protein. This strategy relies on a switch in promoter usage, from distal to proximal, which is dictated by at least two events: The first is the reduction in Ime1 levels as cells progress through meiosis (Brar *et al.* 2012). This leads to reduced expression of $NDC80^{LUTI}$, which likely dampens LUTI-mediated repression of $NDC80^{PROX}$. Second, and more importantly, the increased activity of Ndt80, a transcription factor known to be responsible for the subsequent major wave of gene expression, induces the protein coding *NDC80* mRNA isoform, $NDC80^{PROX}$(Chapter 2) (Chu & Herskowitz 1998; Xu *et al.* 1995). As a result, Ndc80 protein is expressed, permitting re-activation of kinetochores to segregate meiotic chromosomes. This finding further highlights how a developmental switch in promoter usage can cause coordinated expression of two disparate mRNA isoforms to achieve precise temporal control of protein translation during cellular differentiation.

4.3   Identification and characterization of early meiotic LUTI mRNAs

A LUTI mRNA as we have defined it has three essential features: 1) it has a 5'-extended leader sequence, 2) it is not decoded by the ribosome and therefore does not produce a functional full length protein, and 3) it is regulated by condition-specific transcription factors. Additionally, for a LUTI mRNA to repress gene expression, it must prevent transcription initiation from the downstream gene promoter. This final feature is not necessary for a transcript to be called a LUTI mRNA, but it is necessary for it to have a repressive effect on gene expression. Since the discovery of $NDC80^{LUTI}$, both genome-wide and single gene analyses have identified additional LUTI candidates in budding yeast meiosis. For example, *BOI1*, which is involved in polarized growth, has an extended meiosis-specific transcript that, similar to $NDC80^{LUTI}$, is repressed by the mitotic repressor of meiotic genes Ume6 (Liu *et al.* 2015). In another case, it was demonstrated that the origin of recognition gene *ORC1* has an extended transcript that is regulated by the mid-meiotic transcription factor Ndt80 (Xie *et al.* 2016). While these transcripts were both identified before $NDC80^{LUTI}$ they were only studied at the transcriptional level. By ribosome profiling it is apparent that uORFs in their 5'-extended leaders are translated. On a more genome-wide scale, hundreds more transcripts with 5'-extended leaders are expressed in meiosis; the majority of which contain translated AUG initiated uORFs, as observed by ribosome profiling (Brar *et al.* 2012; Cheng *et al.* 2018).

Here, by performing an integrated in-depth investigation of the 5'-extended transcripts present during meiotic prophase, we demonstrated the key features required for LUTI mRNA-mediated repression (Chapter 3). Of 74 identified 5'-extended transcripts, 46 are developmentally regulated by the meiosis-specific Ime1/Ume6 complex. Our analysis found that the overwhelming majority of said transcripts are devoid of ORF translation, while the remaining 5'-extended transcripts either have greatly reduced translation or are not repressed due to a lack of ATG uORFs. This set of translationally repressed mRNAs, LUTI mRNAs, can lead to transcriptional repression as observed in Chapter 2 at the *NDC80* locus, but such transcriptonal down regulation is not universal. It relies on the LUTI mRNA abundance and the change in both nucleosome positioning and H3K36me3 over the PROX transcript's promoter. These features will be ideal to search for when trying to understand which LUTI mRNAs cause transcriptional down regulation in future studies. Further investigation as to how important the PROX promoter is will yield even greater insight into how the LUTI and ORF promoters interact with each other to regulate gene expression. The integrated investigation of the mechanism of LUTI-mediated gene repression during meiotic prophase deepens our understanding of this mode of repression and should be utilized when considering LUTI mRNAs throughout budding yeast meioisis and beyond.

4.4 LUTI mRNAs in additional contexts

Considering that alternative start site usage (Batut *et al.* 2013; Kimura *et al.* 2006; Aanes *et al.* 2013), translation of repressive uORFs (Calvo *et al.* 2009; Chew *et al.* 2016; Johnstone *et al.* 2016), conditional regulation of gene expression by transcription factors, and inhibition of transcription initiation at sites of overlapping transcription have all been individually observed in organisms ranging from yeast to humans (Corbin & Maniatis 1989; Eissenberg & Shilatifard 2010; Wagner & Carpenter 2012; Shearwin *et al.* 2005), we speculate that the LUTI-based gene repression mechanism is widespread.

Outside of meiosis, 5'-extended transcripts appear when budding yeast cells are shifted between carbon sources such as from dextrose to galactose. This appears to rely on the same chromatin associated proteins Set2 and Set3 that are necessary for $NDC80^{PROX}$ transcriptional repression (Kim *et al.* 2012; Kim *et al.* 2017). The extent to which the carbon source-dependent extended transcripts are translationally inhibited is still unknown. Implementation of techniques such as ribosome profiling may uncover a mechanism of translational repression in addition to transcriptional repression just as for the case of *NDC80*.

In addition, a recent paper described a phenomenon where, upon zinc depletion, budding yeast cells induce the expression of 5'-extended transcript isoform at the *RTC4* gene locus. This is mediated by the transcription factor Zap1 and produces a transcript that is not translated into Rtc4 protein (Taggart *et al.* 2017). Their observation of a decrease in the coding *RTC4* transcript and Rtc4 protein upon zinc starvation, points towards a yet to be determined mechanism by which the extended *RTC4* transcript represses downstream transcription initiation. Increases in H3K4me2 and H3K36me3 across the promoter of the shorter transcript may very well prove to be involved again.

Based on the LUTI-mediated repression of the *NDC80* gene, our partner lab, the Brar lab, sought to identify additional LUTIs. They found that in response to the unfolded protein response (UPR), 15 LUTIs are induced by the conserved transcription factor Hac1. These LUTIs appear to down regulate the expression of genes required for cellular respiration, thereby generating a protective effect on cells experiencing stress due to treatment with DTT (Van Dalfsen *et al.* 2018). They also outlined a method to find functionally repressive LUTI mRNAs by looking for instances in which mRNA abundance and protein levels correlate poorly. Through this method 380 LUTIs that repress gene expression were identified in meiosis alone (Cheng *et al.* 2018).

These examples, all represent rapidly changing conditions where gene expression is highly dynamic. We predict that cases of LUTIs in other organisms are likely to be observed in analogous contexts such during differentiation, in response to environmental stress, infection and changes in metabolism. Yeast meiosis, with its dynamic gene expression pattern and tractability, has enabled us to dissect each step at which a LUTI mRNA affects gene reg-

ulation. We can now use this information as a roadmap to explore LUTI mRNAs in more complex genomes.

Approximately 50% of mouse, >40% of drosophila, and 30-50% of human genes have alternative start site usage during development (Batut *et al.* 2013; Kimura *et al.* 2006). In those same organisms, uORF translation is prevalent with an estimated 50% of human genes harboring translated uORFs (Calvo *et al.* 2009; Chew *et al.* 2016; Johnstone *et al.* 2016; Dunn *et al.* 2013). Transcription-coupled chromatin modifications are also highly conserved across evolution (Eissenberg & Shilatifard 2010; Wagner & Carpenter 2012). Therefore, it does not seem far-fetched to propose that the form of regulation controlling so many yeast genes could be responsible for fine-tuning gene expression in other organisms.

In fact, there is a confirmed example of a LUTI in human cells at the MDM2 locus (Hollerer *et al.* 2019). MDM2 is a known oncogene and repressor of p53 (Rayburn *et al.* 2005). It has two previously annotated promoters, one of which produces a poorly translated product due to the presence of 2 uORFs (Brown *et al.* 1999; Barak *et al.* 1994). Hollerer *et al.* (2019) demonstrated that transcription from the distal promoter represses the proximal promoter. They also provided evidence to indicate that the same chromatin modification present at promoters affected by LUTI mRNAs in yeast, H3K36me3, may be involved (Hollerer *et al.* 2019). Strikingly, toggling between the promoters occurs naturally during both neuronal and endodermal differentiation of human embryonic stem cells (Hollerer *et al.* 2019). What role this transcriptional hand-off plays in development remains to be seen, but knowing that overexpression of MDM2 is associated with cancer, improper regulation of the proximal promoter by the $MDM2^{LUTI}$ expression during development could be highly detrimental.

Further examples of possible LUTI-mediated gene regulation were observed during neuronal differentiation of human embryonic stem cells by Blair *et al.* (2017). The translational efficiency (TE) of hundreds of genes varies during this developmental process. It they showed that the genes with variable TE during differentiation also experience changes in ribosome occupancy over their 5'-leaders. Furthermore, the authors concluded that the changes in TE are dominated not by changes in the ribosome composition but by transcript isoform usage (Blair *et al.* 2017). Could these be LUTI mRNAs? Investigation of the transcriptional regulation controlling isoform usage would shed more light on the role that LUTI mRNAs may play in neuronal differentiation.

4.5 Why LUTI mRNAs?

Upon first glance, this mode of gene regulation appears counterintuitive. Why would a cell utilize transcription activators to inhibit gene expression? Why would a cell transcribe mRNAs that cannot produce protein? Why would a cell repress gene promoters by distal transcription of mRNA isoforms? Although the why questions are extremely difficult

to answer in biology, we will describe our perspective on the logic of this gene repression mechanism.

Over evolutionary time, it is thought that *cis*-regulatory mechanisms are the dominant drivers of changes in gene expression (Carroll 2008; Stern & Orgogozo 2008; Wittkopp & Kalay 2012). Most developmental processes use a small number of transcription factors to turn on hundreds or even thousands of targets. If the transcription factor were to change in function, it would affect the levels of all of its targets, likely to the detriment of the organism. However, small changes to the regulatory regions of target genes allow for the tuning of gene expression (Carroll 2008). We propose that this is how LUTI-mediated regulation arose. Small changes in the regulatory sequence that co-opted existing transcription factors to produce 5' extended transcript isoforms enabled these gene expression at these loci to now be influenced by LUTI mRNAs. The extended isoforms most likely already had uORFs, as short open reading frames are pervasive outside of coding regions (Chapter 3) (Chew *et al.* 2016). Thus, mutations in the regulatory regions over time may have evolved to allow transcription factors to behave as both activators and repressors. Through this pathway, transcription factors would gain a wider range of regulatory potential without having to evolve novel trans-acting factors.

Having the same transcription factor tuning gene expression both up and down, could also provide regulatory advantages in development. As stated above, a few transcription factors usually regulate a large number of targets. In the case of budding yeast meiosis, the relay between two transcription factors, Ime1 and Ndt80, control the landmark events in meiotic differentiation (Chu & Herskowitz 1998; van Werven & Amon 2011; Xu *et al.* 1995). Contrast this with the scenario where a transcriptional activator must be coordinated with the activity of a transcriptional repressor to produce the same gene expression output. This situation requires more parts and more co-regulation. Furthermore, inhibiting the expression of genes, whose functions are no longer required or can even be detrimental for a given developmental stage is as important as activating gene expression. Therefore LUTIs provide a clever solution to temporally coordinate gene repression with the transcription-factor driven waves of gene activation.

The prototype view of uORF-mediated translational regulation comes from the elegant studies of *GCN4*, the transcriptional activator of amino acid biosynthetic genes (Mueller & Hinnebusch 1986). *GCN4* expression is controlled by a switch-like mechanism where the uORFs repress translation of the *GCN4* mRNA in nutrient rich conditions, but this translational repression is relieved upon nutrient starvation (Mueller & Hinnebusch 1986). In other words, whether or not a cell expresses Gcn4 protein under different conditions is primarily determined by the translational status of a single *GCN4* mRNA isoform. In the case of LUTI mRNAs, the uORF-mediated translational repression renders the LUTI mRNA permanently non-coding. Because of the perpetual nature of this translational repression, whether or not a cell expresses protein under different conditions is instead determined by a

switch in promoter usage. The promoter choice for LUTI versus proximal mRNA ultimately determines whether a gene is turned on or turned off by a given transcription factor. In addition to tuning the translational capacity of a given mRNA, we propose that the function of some uORFs is to translationally silence what is in fact a regulatory LUTI mRNA.

Finally, LUTI-mediated gene repression is highly dynamic and tunable. While such plasticity may be especially critical during cell fate transitions to enable rapid adaptation to internal as well as external changes, it could also impact gene expression programs in a broader biological context including signaling and metabolic pathways. The repressive chromatin marks, which are established at proximal gene promoters as a result of LUTI mRNA transcription, can be rapidly reversed. In the case of $NDC80^{LUTI}$ when its expression is halted, the chromatin modifications are removed within 15 minutes, followed by the return of $NDC80^{PROX}$ and Ndc80 protein expression (Chia *et al.* 2017). This, along with recent studies of chromatin remodelers and histone demethylases prompt us to revisit the question of how stable chromatin modifications are (Perino & Veenstra 2016). However, it is worth noting that in organisms with DNA methylation, LUTI mRNA regulation could serve to recruit H3K36me3-dependent DNA methyltransferases, such as DNMT3B, to permanently silence gene expression (Jeziorska *et al.* 2017; Morselli *et al.* 2015; Neri *et al.* 2017).

4.6   A future for LUTI mRNA biology

To identify LUTI mRNAs, both transcription and translation must be studied hand in hand. RNA-seq has long been a staple of gene expression studies, but as ribosome profiling has demonstrated, studying RNA abundance alone frequently does not reflect protein abundance in the cell (Ingolia *et al.* 2009). Codon optimality, 3' UTR features, and 5' UTR features are all expected to affect how well a transcript is translated (Barrett *et al.* 2012; Pop *et al.* 2014; Floor & Doudna 2016). This body of work has conclusively demonstrated that uORFs in the 5'-leaders of LUTI mRNAs affect how well a transcript is translated, but that is not the only reason why gene expression decreases when a LUTI is expressed. The repression of proximal transcripts due to increased repressive chromation modifications and nucleosome occupancy additively decreases translation of Ndc80 protein (Chapter 2, Chapter 3). Without considering both the transcriptional and translational effects that a single transcript isoform could have on gene regulation, our mechanism would have been overlooked.

One hurdle in identifying LUTI mRNAs in other organisms will be identifying potential LUTI mRNA transcript isoforms with high confidence. We took advantage of both TL-seq and direct RNA Nanopore sequencing in this work to identify LUTIs in early meiosis, but our methods still relied on a degree of "by eye" analysis (Arribere & Gilbert 2013; Garalde *et al.* 2018). For organisms with much larger genomes, a more streamlined approach will need to be implemented. As both PacBio and Nanopore long-read sequencing technologies advance,

it is possible that the 3' bias we observed will decrease and 5' transcript starts sites will be able to be called more accurately and possibly quantitatively. Technologies such as these are already invaluable in identifying true full-length LUTI mRNAs, which contain the entire ORF of a gene from truncated transcripts that only partially overlap with a gene's ORF. Once putative LUTI mRNA isoforms are identified, techniques such as ribosome profiling can be utilized to call which transcripts have translated uORFs (Ingolia *et al.* 2009). The translational status of the LUTI mRNAs can be further dissected by Transcript Isoforms in Polysomes sequencing (TrIP-seq) (Floor & Doudna 2016).

The final aspect of LUTI-based repression that must be further characterized is its ability to affect protein output. This will depend on how stable the protein is. Classical methods such as immunoblotting and reporter assays can reveal changes in protein expression. However, recent advances in quantitative mass spectrometry have allowed this third layer of gene regulation to be investigated in the context of LUTI mRNAs on a proteome-wide scale (Cheng *et al.* 2018). Without considering effects on protein levels, the impact of LUTI-based regulation cannot be fully understood. This is yet another reason why we should continue to expand the number of studies that integrate analyses at multiple steps of gene expression. Without such comprehensive studies, we have only a limited view of gene regulation and can merely speculate as to what other mechanisms cells may be employing to ensure that gene expression is accurate and reproducible throughout development.

With the current know existence of hundreds of LUTI and a reasonable understanding of their mechanism of action, investigations of LUTI and their affects on biological function will determine how robustly cells rely on this mode of regulation.

Chapter 5

Appendix A

5.1   Set2-AID and Set3-AID depletion in meiosis

Introduction

As discussed in Chapter 2, deletion of *SET2* and *SET3* in combination leads to a slower and less synchronous meiotic progression. This can be partially ameliorated by using reduced YPD medium instead of BYTA to prepare cells for meiosis (Chia & van Werven 2016), but such cells are still not as synchronous as otherwise *wild-type* cells. We proposed that creating conditional alleles of *SET2* and *SET3* could prove helpful for future investigation of these proteins and their role in LUTI-based gene repression. Because meiosis is sensitive to changes in temperature, it's been impossible to use many traditional conditional alleles during this differentiation program. More recently, degron alleles have allowed us to induce programmed degradation of a protein of interest in a temperature-independent manner. One such degron is the auxin degron.

Auxin is a plant hormone. Its presence brings together two other proteins: an auxin-responsive element and an F-box protein. The recognition of the auxin-responsive element by an F-box protein will recruit the protein complex to an E3 ubiquitin ligase, which will then trigger poly-ubiquitination and eventual proteasomal degradation. In plants, this is used naturally to rapidly degrade repressors of genes required, most frequently, for proper plant development (Teale *et al.* 2006). This system has been co-opted for use in the programmed degradation of proteins first and most prominently in budding yeast, but it has also been adapted in other organisms (Nishimura *et al.* 2009; Zhang *et al.* 2015; Natsume *et al.* 2016).

This system of degradation was applied to Set2 and Set3, and the degradation of these two proteins was assessed in both mitosis and meiosis. Progression through meiosis was also assessed in strains with AID-depleted Set2 and Set3 compared to *wild-type* and *set2Δset3Δ* cells.
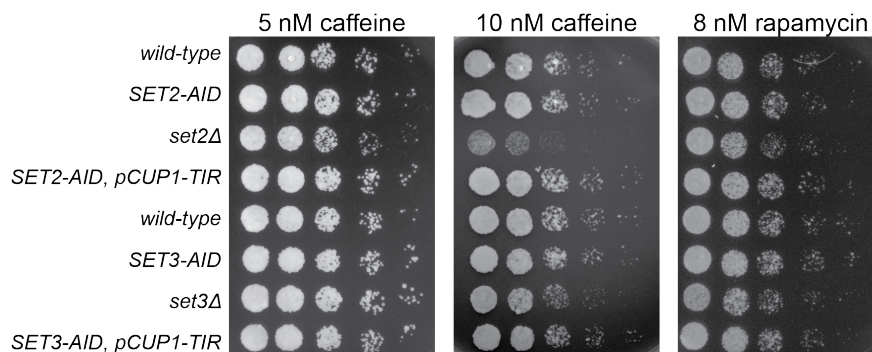
Figure 5.1: The functions of Set2 and Set3 are not disrupted by 3V5-AID tagging. *Wild-type* (UB13), *SET2-AID* (UB12098), *set2Δ* (UB4056), *SET2-AID* with *pCUP1-TIR* (UB12097), *SET3-AID* (UB12099), *set3Δ* (UB4062), and *SET3-AID* with *pCUP1-TIR* (UB12101) strains were each spotted on either a 5 nM caffeine, 10 nM caffeine, or 8 nM rapamycin YPD plates and grown at 30 °C for 1-2 days.

Results

Both *SET2* and *SET3* were individually tagged with a 3xV5 epitope and the auxin-responsive protein IAA7, referred to as AID from now on, and crossed into strains with a copper-inducible and codon-optimized OsTIR1. To assess that the tagged proteins were still functional despite the tag, the cells were serially diluted and spotted onto plates with various concentrations of caffeine and rapamycin, both of which are known to be especially toxic to *set2Δ* cells (Figure 5.1) (McDaniel *et al.* 2017). The deletion strains did appear to have a growth defect that was not apparent in any of the strains with the AID tag. It was much stronger for *set2Δ* than *set3Δ*.

Knowing that the alleles were functional, we next tested how well the protein could be degraded in mitosis. We tested out two different modes of inducing the osTIR domain before addition of auxin: with a LexO/LexA-ER system, and a copper inducible promoter. This was important to test as we have noticed that the presence of a TIR domain, even in the absence of auxin, can lead to substantial instability of proteins with AID tags. Upon addition of auxin and induction of osTIR domains, exponentially growing cells with either the LexO/LexA-ER system (Figure 5.2A-B) or copper induced TIR (Figure 5.2C-D) showed down-regulation of Set2 in a manner that relied on the AID tag. The degradation of Set3 in mitosis under the regulation of the LexO/LexA-ER TIR, had a more tempered down-regulation (Figure 5.3A-B).

Once it was established that the degrons were working in mitosis, we moved forward to testing how auxin-mediated degradation of Set2 and Set3 affects $NDC80^{PROX}$ levels are the progression of cells through meiosis compared to cells with wild-type *SET2* and *SET3* or *set2Δset3Δ*. We observed very little difference in $NDC80^{PROX}$ levels in all tested strains,
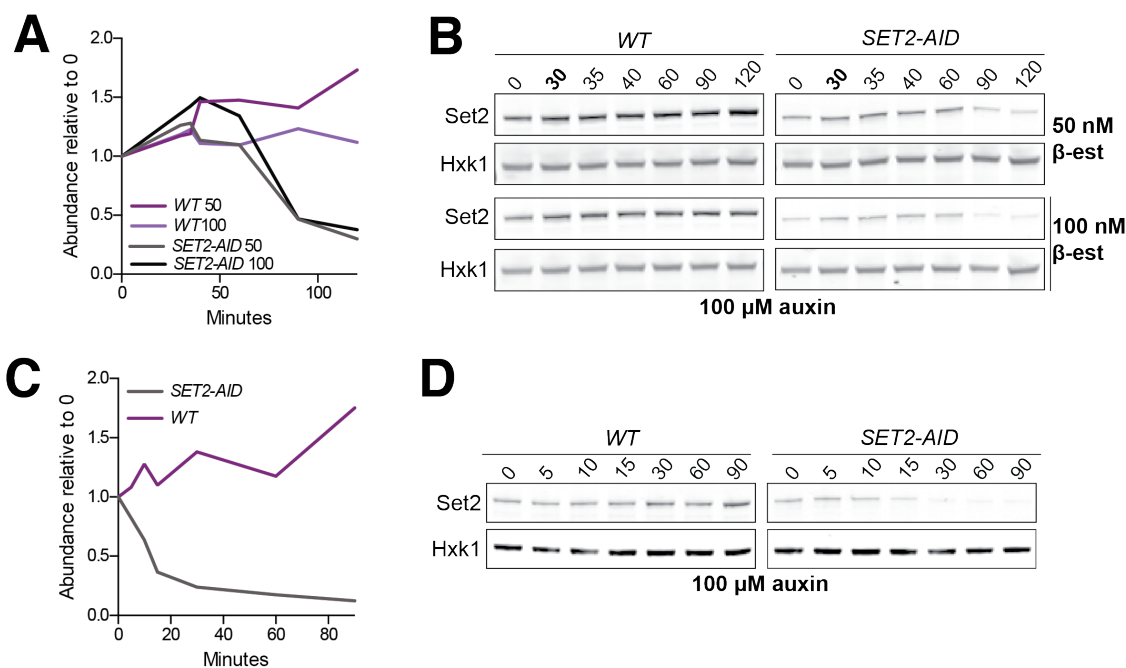
Figure 5.2: Mitotic auxin depletion of Set2 with varying inducible TIR promoters. A-B) In YPD, exponentially growing cells harboring a copy of osTIR under the regulation of a 4xLex operator sequence and a LexA-ER fusion protein with either a 3V5-tagged *SET2* (UB11525) or a 3V5-AID tagged *SET2* (UB11520) were first induced to express osTIR with either 50 nM or 100 nM of Bat time 0. After 30 minutes, the cells were treated with 100 μM auxin and protein degradation was monitored. A) Quantification of B. Set2 abundance was calculated relative to the 0 minute time point. B) Set2 level was determined by anti-V5 immunoblot at the indicated time points. Hxk1, loading control for immunoblot. C-D) In YPD, exponentially growing cells harboring a copy of osTIR under the regulation of the copper-inducible *CUP1* promoter and either a 3V5-tagged *SET2* (UB11879) or a 3V5-AID tagged *SET2* (UB11878) were treated with 100 μM auxin at time 0 and protein degradation was monitored. The osTIR was not specifically induced due to the copper already present in YPD.
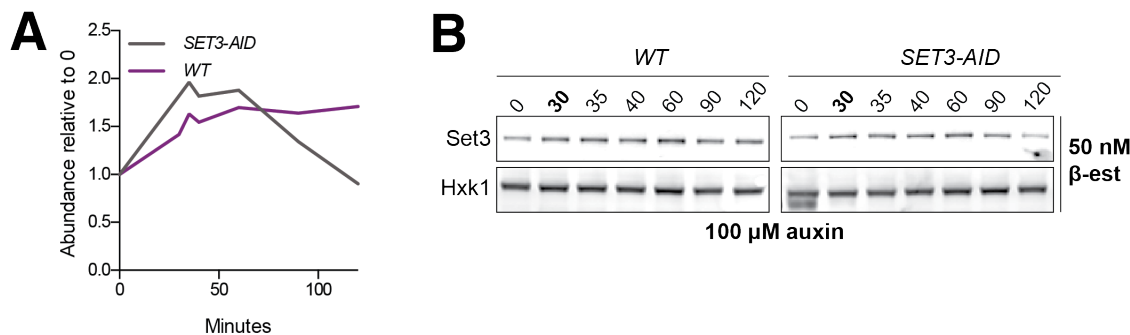
Figure 5.3: Mitotic auxin depletion of Set3. A-B) In YPD, exponentially growing cells harboring a copy of osTIR under the regulation of a 4xLex operator sequence and a LexA-ER fusion protein with either a 3V5-tagged *SET3* (UB11531) or a 3V5-AID tagged *SET3* (UB11530) were first induced to express osTIR with either 50 nM or 100 nM of Bat time 0. After 30 minutes, the cells were treated with 100 µM auxin and protein degradation was monitored. A) Quantification of B. Set3 abundance was calculated relative to the 0 minute time point. B) Set3 level was determined by anti-V5 immunoblot at the indicated time points. Hxk1, loading control for immunoblot.

including the *set2Δset3Δ* which we had previously found suppresses $NDC80^{LUTI}$ mediated repression. This may be due to the poor quality of the northerns, but should be investigated further at a later date. When observing progression through meiosis, the *SET2-AID SET3-AID* strains enter and complete meiotic S-phase with kinetics indistinguishable from *wild-type* compared to the *set2Δset3Δ* cells which enter S-phase later and take longer to complete S-phase even with the protocol optimized in Chia & van Werven (2016). Unsurprisingly, the *set2Δset3Δ* cells completed meiosis at a lower rate than any of the other strains tested. Degradation of Set2 and Set3 also led to a decreased spore packaging rate.

Finally, we assessed the ideal concentration of auxin required for the best conditional depletion of Set2 and Set3. Depletion was more rapid and complete with a 2 mM concentration of auxin than with the 500 uM concentration which had been used in all previous experiments. Future experiments should use the higher concentration as it doesn't affect completion of meiosis as determined by the rate of proper spore packaging.

Discussion

Set2 and Set3 were successfully depleted in both mitosis and meiosis using the auxin degron system. The goal of this depletion was to obtain a population of cells that had synchronously entered meiosis but that lacked Set2 and Set3 in order to best understand the role Set2 and Set3 play in mediating LUTI repression. Under the conditions tested, depletion of Set2 and Set3 in early meiosis did not affect the kinetics of meiotic progression through S-phase, nor did it affect the frequency of packaged tetrads or dyads at the end of meiosis. While these conditions allow for the investigation of cells that are progressing
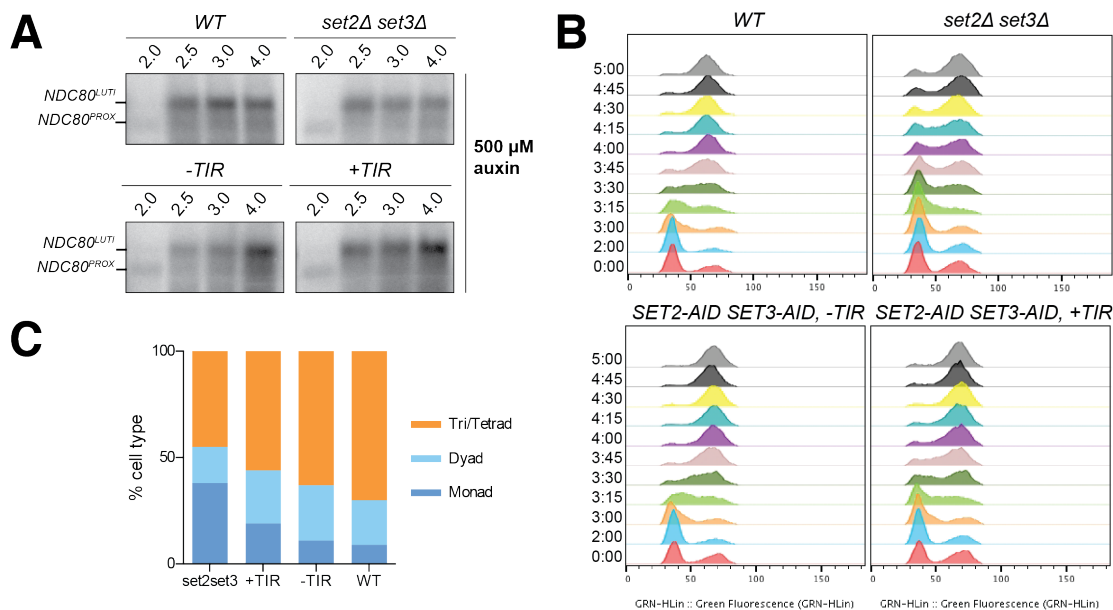
Figure 5.4: Meiotic characterization of *set2Δset3Δ* and *SET2-AID SET3-AID* strains. A-C) Cells harboring the *pCUP1-IME1 pCUP1-IME4* meiotic induction system and an *AMN1* deletion were induced to go through meiosis as in Chia & van Werven (2016). *Wild-type* (UB14584), *set2Δset3Δ* (UB14259), and *SET2-3V5-AID SET3-3V5-AID* strains either with (UB14261) or without (UB14263) a copper inducible OsTIR were used. Auxin was added to 500 µM at 2 hours in SPO. A) $NDC80^{LUTI}$ and $NDC80^{PROX}$ where detected by RNA blotting with a probe specific to the body of *NDC80*. B) Fixed cells were stained with SYTOX green and DNA content was determined by Flow cytometry. C) After 24 hours, spore packaging was assessed.

through meiosis synchronously even in the absence of Set2 and Set3, it remains unclear how robustly LUTI-based repression is affected by the depletion of Set2 and Set3. In the future, the RNA blotting could be optimized or an unstable reporter gene such as ubiquitin-fused GFP could be used as an indirect reading of the suppression of LUTI-mediated repression.
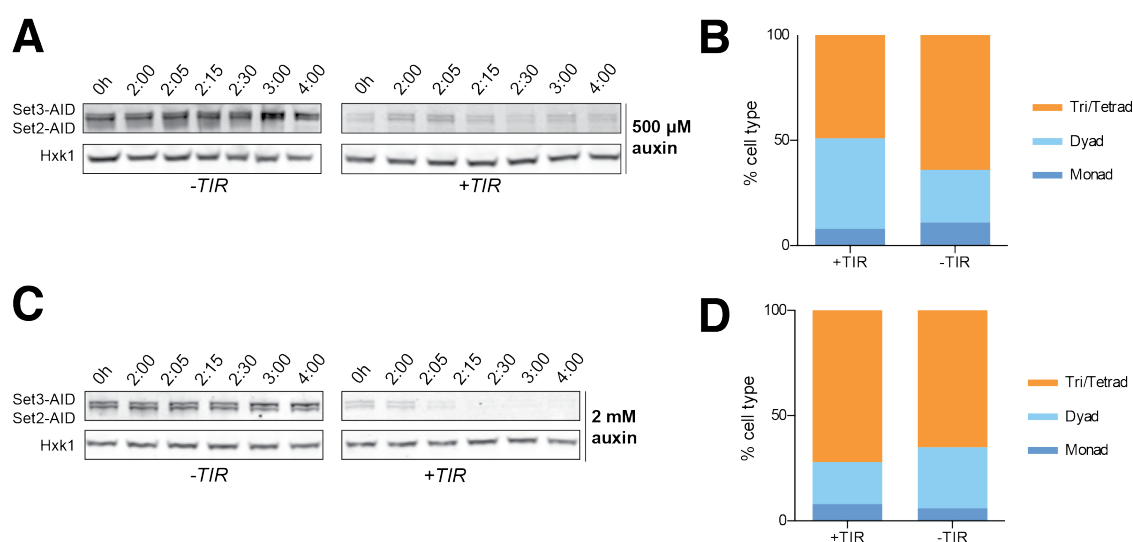
Figure 5.5: Meiotic auxin mediated degradation of Set2 and Set3. A-D) Cells harboring the *pCUP1-IME1 pCUP1-IME4* meiotic induction system and an *AMN1* deletion were induced to go through meiosis as in Chia & van Werven (2016). *SET2-3V5-AID SET3-3V5-AID* strains either with (UB14261) or without (UB14263) a copper inducible OsTIR were used. Auxin was added to 500 μM (A-B) or 2 mM (C-D) at 2 hours in SPO. A and C) Set2 and Set3 levels were determined by anti-V5 immunoblot at the indicated time points. Hxk1, loading control for immunoblot. B and D) After 24 hours, spore packaging was assessed.

Chapter 6

Appendix B: Materials and Methods

6.1   Plasmid and strain construction with tables

Yeast strain and plasmid construction

All the strains used in this study are described in Table 6.2 and are derivatives of SK1, except for the strains that harbored the LexA/lexO system for which the W303 strain background was used. The *pGAL-NDT80* Gal4-ER and the *pCUP-IME1 pCUP-IME4* synchronization systems have been described previously (Benjamin *et al.* 2003; Berchowitz *et al.* 2013). The centromeric TetR/TetO GFP dot assay is described in (Michaelis *et al.* 1997). The *ndc80-1* temperature-sensitive mutant was first described in (Wigge *et al.* 1998), the Zip1::GFP (700) described in (Scherthan *et al.* 2007), and *pCUP-NDC80 pCUP-CLB3* described in (Miller *et al.* 2012). The LexA/lexO system was described previously (Ottoz *et al.* 2014). *NDC80-3V5*, *NUF2-3V5*, *pGAL-NDC80^{LUTI}*, *pGAL-Δ9AUG*, *ndc80Δ*, *nuf2Δ*, *(Δ-600 to -300)-NDC80*, *(Δ-600 to -400)-NDC80*, and *SUA7-3V5* were generated at the endogenous gene loci using PCR-based methods (Longtine *et al.* 1998). The V5 plasmid is kind gift from Vincent Guacci. Single integration plasmids carrying either *NDC80*, *NUF2*, or the other LUTI genes of interest (*SWI4*, *ALP4*, *CDC60*, *HSP60*, and *MSC6*) were constructed by Gibson Assembly (Gibson *et al.* 2009). For *NDC80*, the *LEU2* integration plasmid included the SK1 genomic sequence spanning from 1000 bps upstream to 357 bps downstream of the *NDC80* coding region; and for *NUF2*, spanning from 1000 bps upstream to 473 bps downstream of the *NUF2* coding region. Both constructs included a C-terminal fusion of the 3V5 epitope to *NDC80* and *NUF2*, and both completely rescued the full deletion of *NDC80* or *NUF2*, respectively. For LUTI genes of interest, the WT gene with between 800 and 1200 bp upstream of the ORF, depending on the length of the 5-extension, and a C-terminal 3xV5 epitope tag with the *NDC80* terminator were cloned into a *LEU2* single integration vector by Gibson Assembly (Gibson *et al.* 2009). LUTI promoter deletion strains were similarly constructed for each gene with the cloned upstream region only containing the sequence downstream from the LUTI TSS as determined by TL-seq. In all strains, the WT copies of

these 6 genes remained untouched. Deletions (*ndc80-urs1Δ* and *(Δ-600 to -479)-NDC80)* and point mutations (*ndc80-mse*) were generated from the *NDC80 LEU2* single integration plasmid using the site-directed mutagenesis kit (Q5 Site-Directed Mutagenesis Kit, NEB, Ipswitch, MA). The entire URS1 site and the A right upstream of the site were deleted in the *ndc80-urs1Δ* strain. The *ndc80-mse* construct has two C to A mutations, marked using black diamonds in 2.17. The *Δ6AUG*, *Δ9AUG*, mini uORF, *NDC80^{LUTI}*-Ter, and *NDC80^{LUTI}*-*NUF2* constructs were generated by Gibson assembly (Gibson *et al.* 2009) using the *NDC80* and *NUF2 LEU2* integration plasmids, as well as gBlockső gene fragments (IDT, Redwood City, CA) for the *Δ9AUG* and mini uORF constructs. The uORF1, uORF5, and uORF9 plasmids were constructed from the *Δ9AUG* plasmid using the site-directed mutagenesis kit (Q5 Site-Directed Mutagenesis Kit, NEB) by reverting the designated ATC back to ATG. All single integration plasmids were digested with PmeI to integrate at the *LEU2* locus.

Primers

Table 6.1: Primers

Primer Name	Oligonucleotide sequence from 5 to 3
NDC80_probe_F	GGAGAGGTAGAATCGTCCCTG
NDC80_probe_R	CTCCTCTTGAATAGCGCTTTGG
NUF2_probe_F	AACAGGGGATGGTCACTTACAGG
NUF2_probe_R	CCCACAAGTTCCGTTTCAGTTCG
SCR1_probe_F	GAAGTGTCCCGGCTATAATAAA
SCR1_probe_R	GACGCTGGATAAAACTCCCC
CIT1_probe_F	CCGTGTTAGACCCCGAAGAAG
CIT1_probe_R	GGGCAGAAACGTTACCACCTTC
3V5_probe_F	CTAGTGGATCCAGGTAAACCTAT
3V5_probe_R	TAATACGACTCACTATAGGCCAGTCCTAATAGAGGATTAGG
NDC80_1_FW	GCTCCTGTGTTCTCCATT
NDC80_1_RV	GTGTGTTGATACTGCACTG
NDC80_2_FW	ACCCGGATATCTGTTCAGCC
NDC80_2_RV	TGTGGCGAATTGTTGCTCTT
NDC80_3_FW	CGCCACAAGAAGGTCTC
NDC80_3_RV	GCTTTTCGGACCTCCAAC
NDC80_4_FW	GTTGGAGGTCCGAAAAGC
NDC80_4_RV	GTTCAGTTATAACCATCTGGCAC
NDC80_5_FW	GTGCCAGATGGTTATAACTGAAC
NDC80_5_RV	CCGCTAATCGCAATAGACTG
NDC80_6_FW	GGTTGAGAGCCCCGTTAAGT
NDC80_6_RV	TTGGCACTTTCAGTATGGGT
NDC80_7_FW	CCCATACTGAAAGTGCCAAAAGA
NDC80_7_RV	GGGACGATTCTACCTCTCCTGTG

Table 6.1: continued

NDC80_8_FW	GGAATACATTCACAGGAGAGG
NDC80_8_RV	GGAATATATTATAGTACACCCTAACG
NDC80_9_FW	TGCAAAGCTCAACAAGTACTGA
NDC80_9_RV	TGCAGTTGGTATTTGGGACG
NDC80_10_FW	CAAGGTCTAACCGACATGATC
NDC80_10_RV	CATTTGTACCTCCTGCAAC
PHO5-TATA_FW	CCATTTGGGATAAGGGTAAACATC
PHO5-TATA_RV	AGAGATGAAGCCATACTAACCTCG
pNUF2_F	GTCGCTGCGTATTCAGCGTA
pNUF2_R	GAACGCTGATATACTCGACTAAC
ACT1_F	GTACCACCATGTTCCCAGGTATT
ACT1_R	AGATGGACCACTTTCGTCGT
HMR_FW	ACGATCCCCGTCCAAGTTATG
HMR_RV	CTTCAAAGGAGTCTTAATTTCCCTG
NDC80_ORF_F	ATCCGAGTGTGAACTGAAAGAAG
NDC80_ORF_R	GAACTGCTCAGTTGAAATTCCC
IME2_URS1_F	CCAAATACGCTTTTTAAACTTGG
IME2_URS1_R	CTCAAATAGCCGCCGTAAC
MAM1_MSE_F	CACAATTGAAATCCGAGCTGT
MAM1_MSE_R	CATCTGAATTTTGAATGGCTTT

Strains

Table 6.2: Strains

Strain Name	*Genotype*
SK1 wild-type	ho::LYS2 lys2 ura3 leu2::hisG his3::hisG trp1::hisG
UB91	MATa, ura3::pGPD1-GAL4(848).ER::URA3
UB494	MATa ndc80-1
UB877	MATa/MATαura3::pGPD1-GAL4(848).ER::URA3/ura3::pGPD1-GAL4(848).ER::URA3 GAL-NDT80::TRP1/GAL-NDT80::TRP1 leu2::pURA3-TetR-GFP::LEU2 CENV::TetOx224::HIS3 Ndc80-3V5:KanMX pCUP-CLB3::KANMX

Table 6.2: continued

UB880	MATa/MATαura3::pGPD1-GAL4(848).ER::URA3/ura3::pGPD1-GAL4(848).ER::URA3 GAL-NDT80::TRP1/GAL-NDT80::TRP1 leu2::pURA3-TetR-GFP::LEU2 CENV::TetOx224::HIS3 KanMX:pCUP-Ndc80-3V5:CNAT pCUP-CLB3::KANMX
UB1217	MATa HISMX:pGAL-Ndc80-3V5:KanMX ura3::pGPD1-GAL4(848).ER::URA3 pGAL integrated 536 bp upstream of NDC80 AUG
UB1218	MATαHISMX:pGAL-Ndc80-3V5:KanMX ura3::pGPD1-GAL4(848).ER::URA3 pGAL integrated 536 bp upstream of Ndc80 AUG
UB1235	MATαset2::hismx set3::his3mx HISMX:pGAL-Ndc80-3V5:KanMX ura3::pGPD1-GAL4(848).ER::URA3 pGAL integrated 536 bp upstream of Ndc80 AUG
UB1236	MATαset2::hismx HISMX:pGAL-Ndc80-3V5:KanMX ura3::pGPD1-GAL4(848).ER::URA3 pGAL integrated 536 bp upstream of Ndc80 AUG
UB1237	MATαset3::his3mx, HISMX:pGAL-Ndc80-3V5:KanMX, ura3::pGPD1-GAL4(848).ER::URA3 pGAL integrated 536 bp upstream of Ndc80 AUG
UB1240	MATa Ndc80-3V5:KanMX ura3::pGPD1-GAL4(848).ER::URA3
UB1252	MATαNdc80-3V5:KanMX, ura3::pGPD1-GAL4(848).ER::URA3
UB1323	MATa KanMX:pGAL-Δ9AUG-5'UTR-Ndc80-3V5:CNAT ura3::pGPD1-GAL4(848).ER::URA3 pGAL integrated 536 bp upstream of NDC80 AUG ATG-ATC mutation in 9 of the 9 potential upstream start codons within NDC80-5'UTR. The leader sequence contains 2 SNPs from S288C introduced by a gene block and a third mutation (T to C) 8 bp after the 6th ATG.
UB1337	MATa/MATαpCUP-IME1::NAT/pCUP-IME1::NAT pCUP-IME4::NAT/pCUP-IME4::NAT Ndc80-3V5:KanMX/Ndc80-3V5:KanMX

Table 6.2: continued

UB2388	MATa amn1::KanMX6 ura3::pGPD1-GAL4(848).ER::URA3 HISMX:pGAL-Ndc80-3V5:KanMX pGAL integrated 536 bp upstream of NDC80 AUG
UB2389	MATa amn1::KanMX6 Ndc80-3V5:KanMX ura3::pGPD1-GAL4(848).ER::URA3
UB2531	MATa/MATαirt1:cup1::Hphmx/irt1:cup1::Hphmx ime4::cup1::NAT/ime4::cup1::NAT
UB2936	MATa/MATαura3::pGPD1-GAL4(848).ER::URA3/ura3::pGPD1-GAL4(848).ER::URA3 GAL-NDT80::TRP1/GAL-NDT80::TRP1 leu2::pURA3-TetR-GFP::LEU2 CENV::TetOx224::HIS3 HisMX:Δ9AUG-5'UTR-Ndc80-3V5:CNAT/HisMX:Δ9AUG-5'UTR-Ndc80-3V5:CNAT pCUP-CLB3::KANMX ATG-ATC mutation in 9 of the 9 potential upstream start codons within NDC80-5'UTR. The leader sequence contains 2 other point mutations, in addition to the ATCs.
UB2940	MATa/MATαura3::pGPD1-GAL4(848).ER::URA3/ura3::pGPD1-GAL4(848).ER::URA3 GAL-NDT80::TRP1/GAL-NDT80::TRP1 leu2::pURA3-TetR-GFP::LEU2 CENV::TetOx224::HIS3 pCUP-CLB3::KANMX HIS3MX::(Δ-600 to -400)-NDC80-3V5::KanMX/HIS3MX::(Δ-600 to -400 bp)-NDC80-3V5::KanMX
UB2942	MATa/MATαura3::pGPD1-GAL4(848).ER::URA3/ura3::pGPD1-GAL4(848).ER::URA3 GAL-NDT80::TRP1/GAL-NDT80::TRP1 leu2::pURA3-TetR-GFP::LEU2 CENV::TetOx224::HIS3 Ndc80-3V5:KanMX/Ndc80-3V5:KanMX
UB3262	MATa ndc80Δ:KanMX4 leu2::NDC80-3V5:LEU2
UB3301	MATa/MATαUME6-3V5::His3MX/UME6-3V5::His3MX irt1:cup1::Hphmx/irt1:cup1::Hphmx ime4::cup1::NAT/ime4::cup1::NAT
UB3338	MATa ura3::pGPD1-GAL4(848).ER::URA3 HISMX:pGAL-Ndc80

Table 6.2: continued

	pGAL integrated 536 bp upstream of Ndc80 AUG
UB3351	MATαura3::pGPD1-GAL4(848).ER::URA3, ndc80Δ:KanMX4, leu2::NDC80-3V5:LEU2
UB3370	MATαGAL-NDT80::TRP1 ura3::pGPD1-GAL4(848).ER::URA3 ndc80Δ:KanMX4 leu2::NDC80-3V5:LEU2
UB3392	MATa/MATαGAL-NDT80::TRP1/GAL-NDT80::TRP1 ura3::pGPD1-GAL4(848).ER::URA3/ura3::pGPD1-GAL4(848).ER::URA3 ndc80Δ(-1000 and ORF):KanMX4/ndc80Δ(-1000 and ORF):KanMX4 leu2::mse-NDC80-3V5:LEU2/leu2::mse-NDC80-3V5:LEU2
UB3545	MATαset2::his3mx Ndc80-3V5:KanMX ura3::pGPD1-GAL4(848).ER::URA3
UB3547	MATαset3::his3mx Ndc80-3V5:KanMX ura3::pGPD1-GAL4(848).ER::URA3
UB3549	MATαset2::hismx set3::his3mx Ndc80-3V5:KanMX ura3::pGPD1-GAL4(848).ER::URA3
UB4074	MATa/MATαGAL-NDT80::TRP1/GAL-NDT80::TRP1 ura3::pGPD1-GAL4(848).ER::URA3/ura3::pGPD1-GAL4(848).ER::URA3 ndc80Δ(-1000 and ORF):KanMX4/ndc80Δ(-1000 and ORF):KanMX4 leu2::NDC80-3V5:LEU2/leu2::NDC80-3V5:LEU2
UB4212	MATa leu2::urs1Δ-NDC80-3V5:LEU2 ndc80Δ::KanMX
UB5101	MATa/MATαpCUP-IME1::NAT/pCUP-IME1::NAT pCUP-IME4::NAT/pCUP-IME4::NAT nuf2::KanMX/nuf2::KanMX leu2::NDC80(-1000 to -1)-NUF2-3V5:LEU2/leu2::NDC80(-1000 to -1)-NUF2-3V5:LEU2
UB5103	MATa/MATαpCUP-IME1::NAT/pCUP-IME1::NAT pCUP-IME4::NAT/pCUP-IME4::NAT nuf2::KanMX/nuf2::KanMX leu2::NUF2-3V5:LEU2/leu2::NUF2-3V5:LEU2

Table 6.2: continued

UB5154	MATαura3::pGPD1-GAL4(848).ER::URA3 ndc80Δ:KanMX4 leu2::pGAL-NDC80-3V5:LEU2 pGAL integrated 536 bp upstream of Ndc80 AUG
UB5437	MATa/MATαndc80Δ(-1000 and ORF):KanMX4/ndc80Δ(-1000 and ORF):KanMX4 leu2::mse-NDC80-3V5:LEU2/leu2::mse-NDC80-3V5:LEU2 CENV::tetOx224::HIS3/CENV::tetOx224::HIS3 his3::pURA3-TetR-GFP::HIS3/his3::pURA3-TetR-GFP::HIS3
UB5876	MATa/MATαndc80Δ(-1000 and ORF):KanMX4/ndc80Δ(-1000 and ORF):KanMX4 leu2::NDC80-3V5:LEU2/leu2::NDC80-3V5:LEU2 CENV::tetOx224::HIS3/CENV::tetOx224::HIS3 his3::pURA3-TetR-GFP::HIS3/his3::pURA3-TetR-GFP::HIS3
UB6075	MATa/Matαirt1:cup1::Hphmx/irt1:cup1::Hphmx ime4::cup1::NAT/ime4::cup1::NAT leu2::urs1Δ-NDC80-3V5:LEU2/leu2::urs1Δ-NDC80-3V5:LEU2 ndc80Δ(-1000 and ORF):KanMX4/ndc80Δ(-1000 and ORF):KanMX4
UB6077	MATa/Matαirt1:cup1::Hphmx/irt1:cup1::Hphmx ime4::cup1::NAT/ime4::cup1::NAT leu2::(-295::ADH1)-NDC80-3V5:LEU2/leu2::(-295::ADH1)-NDC80-3V5:LEU2 ndc80Δ(-1000 and ORF):KanMX4/ndc80Δ(-1000 and ORF):KanMX4
UB6079	MATa/Matαirt1:cup1::Hphmx/irt1:cup1::Hphmx ime4::cup1::NAT/ime4::cup1::NAT leu2::(Δ-600 to -479)-NDC80-3V5:LEU2/leu2::(Δ-600 to -479)-NDC80-3V5:LEU2 ndc80Δ(-1000 and ORF):KanMX4/ndc80Δ(-1000 and ORF):KanMX4
UB6181	MATa/Matαirt1:cup1::Hphmx/irt1:cup1::Hphmx ime4::cup1::NAT/ime4::cup1::NAT leu2::Δ6AUG-NDC80-3V5:LEU2/leu2::Δ6AUG-NDC80-3V5:LEU2 ndc80Δ(-1000 and ORF):KanMX4/ndc80Δ(-1000 and ORF):KanMX4 ATG-ATC mutation in 6 of the 9 potential upstream start codons within NDC80-5'UTR

Table 6.2: continued

UB6183	MATa/MATαirt1:cup1::Hphmx/irt1:cup1::Hphmx ime4::cup1::NAT/ime4::cup1::NAT leu2::Δ9AUG-NDC80-3V5:LEU2/leu2::Δ9AUG-NDC80-3V5:LEU2 ndc80Δ(-1000 and ORF):KanMX4/ndc80Δ(-1000 and ORF):KanMX4 ATG-ATC mutation in 9 of the 9 potential upstream start codons within NDC80-5'UTR
UB6190	MATa/MATαirt1:cup1::Hphmx/irt1:cup1::Hphmx ime4::cup1::NAT/ime4::cup1::NAT ndc80Δ(-1000 and ORF):KanMX4/ndc80Δ(-1000 and ORF):KanMX4 leu2::NDC80-3V5:LEU2/leu2::NDC80-3V5:LEU2
UB6760	MATa/MATαUme6-3V5::His3MX/Ume6-3V5::His3MX irt1:cup1::Hphmx/irt1:cup1::Hphmx ime4::cup1::NAT/ime4::cup1::NAT leu2::urs1Δ-NDC80:LEU2/leu2::urs1Δ-NDC80:LEU2 ndc80Δ(-1000 and ORF):KanMX4/ndc80Δ(-1000 and ORF):KanMX4
UB7496	MATa/MATαTRP1::GAL-NDT80-3V5::KanMX/TRP1::GAL-NDT80-3V5::KanMX ura3::pGPD1-GAL4(848).ER::URA3/ura3::pGPD1-GAL4(848).ER::URA3 ndc80Δ(-1000 and ORF):KanMX4/ndc80Δ(-1000 and ORF):KanMX4 leu2::mse-NDC80:LEU2/leu2::mse-NDC80:LEU2
UB7997	MATa/MATαGAL-NDT80::TRP1/GAL-NDT80::TRP1 ura3::pGPD1-GAL4(848).ER::URA3/ura3::pGPD1-GAL4(848).ER::URA3 leu2::NDC80:LEU2/leu2::NDC80:LEU2 ndc80Δ(-1000 and ORF):KanMX4/ndc80Δ(-1000 and ORF):KanMX4
UB7999	MATa/MATαTRP1::GAL-NDT80-3V5::KanMX/TRP1::GAL-NDT80-3V5::KanMX ura3::pGPD1-GAL4(848).ER::URA3/ura3::pGPD1-GAL4(848).ER::URA3 leu2::NDC80:LEU2/leu2::NDC80:LEU2 ndc80Δ(-1000 and ORF):KanMX4/ndc80Δ(-1000 and ORF):KanMX4

Table 6.2: continued

UB8001	MATa HISMX:pGAL-Ndc80-3V5:KanMX ura3::pGPD1-GAL4(848).ER::URA3 leu2::NDC80-3V5:LEU2 pGAL integrated 536 bp upstream of Ndc80 AUG
UB8110	MATa ura3::pGPD1-GAL4(848).ER::URA3 Ndc80-3V5:KanMX set2::HygB set3::CNAT
UB8114	MATa ura3::pGPD1-GAL4(848).ER::URA3 HISMX:pGAL-Ndc80-3V5:KanMX set2::HygB set3::CNAT pGAL integrated 536 bp upstream of Ndc80 AUG
UB8358	MATa, ADE2, leu2-3, ura3, trp1-1, his3-11,15, can1-100, GAL, phi+ KanMX:p1X-LexO-pCyc1-Ndc80luti trp1::pGPD1-LexA-ER-HA-B112::TRP1 W303 1X-LexO-pCyc1 integrated 536 bp upstream of Ndc80 AUG, thus replacing the Ndc80luti promoter
UB8362	MATa, ADE2, leu2-3, ura3, trp1-1, his3-11,15, can1-100, GAL, phi+ KanMX:p2X-LexO-pCyc1-Ndc80luti, trp1::pGPD1-LexA-ER-HA-B112::TRP1 W303 2X-LexO-pCyc1 integrated 536 bp upstream of Ndc80 AUG, thus replacing the Ndc80luti promoter
UB8366	MATa, ADE2, leu2-3, ura3, trp1-1, his3-11,15, can1-100, GAL, phi+ KanMX:p3X-LexO-pCyc1-Ndc80luti trp1::pGPD1-LexA-ER-HA-B112::TRP1 W303 3X-LexO-pCyc1 integrated 536 bp upstream of Ndc80 AUG, thus replacing the Ndc80luti promoter
UB8370	MATa, ADE2, leu2-3, ura3, trp1-1, his3-11,15, can1-100, GAL, phi+ KanMX:p8X-LexO-pCyc1-Ndc80luti trp1::pGPD1-LexA-ER-HA-B112::TRP1 W303 8X-LexO-pCyc1 integrated 536 bp upstream of Ndc80 AUG, thus replacing the Ndc80luti promoter
UB8374	MATa, ADE2, leu2-3, ura3, trp1-1, his3-11,15, can1-100, GAL, phi+

Table 6.2: continued

	trp1::pGPD1-LexA-ER-HA-B112::TRP1 W303
UB8682	MATa ura3::pGPD1-GAL4(848).ER::URA3 amn1::HygB CENV::TetOx224::HIS3 leu2::pURA3-TetR-GFP::LEU2 SPC42-mCherry::NAT
UB8684	MATa ura3::pGPD1-GAL4(848).ER::URA3 HISMX:pGAL-Ndc80-3V5:KanMX amn1::HygB CENV::TetOx224::HIS3 leu2::pURA3-TetR-GFP::LEU2 SPC42-mCherry::NAT pGAL integrated 536 bp upstream of Ndc80 AUG
UB8686	MATa, ADE2, leu2-3, ura3, trp1-1, his3-11,15, can1-100, GAL, psi+ KanMX:p3X-LexO-pCyc1-Ndc80luti trp1::pGPD1-LexA-ER-HA-B112::TRP1 set2::HygB set3::CNAT W303 3X-LexO-pCyc1 integrated 536 bp upstream of Ndc80 AUG, thus replacing the Ndc80luti promoter
UB8691	MATa, ADE2, leu2-3, ura3, trp1-1, his3-11,15, can1-100, GAL, psi+ trp1::pGPD1-LexA-ER-HA-B112::TRP1 set2::HygB set3::CNAT W303
UB8693	MATa, ADE2, leu2-3, ura3, trp1-1, his3-11,15, can1-100, GAL, phi+ KanMX:p8X-LexO-pCyc1-Ndc80luti trp1::pGPD1-LexA-ER-HA-B112::TRP1 set2::HygB set3::CNAT W303 8X-LexO-pCyc1 integrated 536 bp upstream of Ndc80 AUG, thus replacing the Ndc80luti promoter
UB9181	MATαura3::pGPD1-GAL4(848).ER::URA3 ndc80Δ:KanMX4 leu2::pGAL-NDC80-3V5:LEU2 pGAL-NDT80::TRP1 pGAL integrated 536 bp upstream of NDC80 AUG
UB9243	MATa/MATαirt1:cup1::Hphmx/irt1:cup1::Hphmx ime4::cup1::NAT/ime4::cup1::NAT leu2::uORF(mini)-NDC80- 3V5::LEU2/leu2::uORF(mini)-NDC80-3V5::LEU2 ndc80Δ(-1000 and ORF):KanMX4/ndc80Δ(-1000 and ORF):KanMX4

Table 6.2: continued

UB9921	MATαGAL-NDT80::TRP1 ura3::pGPD1-GAL4(848).ER::URA3 leu2::pGAL-mse-NDC80-3V5:LEU2 ndc80Δ:KanMX4 pGAL integrated 536 bp upstream of NDC80 AUG
UB9923	MATαura3::pGPD1-GAL4(848).ER::URA3 leu2::pGAL-mse-NDC80-3V5:LEU2 ndc80Δ:KanMX4 pGAL integrated 536 bp upstream of NDC80 AUG
UB10579	MATa/MATαirt1:cup1::Hphmx/irt1:cup1::Hphmx ime4::cup1::NAT/ime4::cup1::NAT leu2::pUB88-Δ8AUG-uORF1-revert-5'UTR-NDC80-3V5:LEU2/leu2::pUB88-Δ8AUG-uORF1-revert-5'UTR-NDC80-3V5:LEU2 ndc80Δ(-1000 and ORF):KanMX4/ndc80Δ(-1000 and ORF):KanMX4 All but uORF1s ATGs are mutated to ATC
UB10581	MATa/MATαirt1:cup1::Hphmx/irt1:cup1::Hphmx ime4::cup1::NAT/ime4::cup1::NAT leu2::pUB88-Δ8AUG-uORF5-revert-5'UTR-NDC80-3V5:LEU2/leu2::pUB88-Δ8AUG-uORF5-revert-5'UTR-NDC80-3V5:LEU2 ndc80Δ(-1000 and ORF):KanMX4/ndc80Δ(-1000 and ORF):KanMX4 All but uORF5s ATGs are mutated to ATC
UB10583	MATa/MATαirt1:cup1::Hphmx/irt1:cup1::Hphmx ime4::cup1::NAT/ime4::cup1::NAT leu2::pUB88-Δ8AUG-uORF9-revert-5'UTR-NDC80-3V5:LEU2/leu2::pUB88-Δ8AUG-uORF9-revert-5'UTR-NDC80-3V5:LEU2 ndc80Δ(-1000 and ORF):KanMX4/ndc80Δ(-1000 and ORF):KanMX4 All but uORF9s ATGs are mutated to ATC
UB12945	MATa, ADE2, leu2-3, ura3, trp1-1, his3-11,15, can1-100, GAL, phi+ NDC80-3V5:HisMX trp1::pGPD1-LexA-ER-HA-B112::TRP1 W303
UB12947	MATa, ADE2, leu2-3, ura3, trp1-1, his3-11,15, can1-100, GAL, phi+ NDC80-3V5:HisMX trp1::pGPD1-LexA-ER-HA-B112::TRP1 set2::HygB set3::CNAT W303
UB12949	MATa, ADE2, leu2-3, ura3, trp1-1, his3-11,15, can1-100, GAL, phi+

Table 6.2: continued

	KanMX:p8X-LexO-pCyc1-Ndc80luti-NDC80-3V5:HisMX trp1::pGPD1-LexA-ER-HA-B112::TRP1 W303 8X-LexO-pCyc1 integrated 536 bp upstream of Ndc80 AUG, thus replacing the Ndc80luti promoter
UB12951	MATa, ADE2, leu2-3, ura3, trp1-1, his3-11,15, can1-100, GAL, phi+ KanMX:p8X-LexO-pCyc1-Ndc80luti-NDC80-3V5:HisMX trp1::pGPD1-LexA-ER-HA-B112::TRP1 set2::HygB set3::CNAT W303 8X-LexO-pCyc1 integrated 536 bp upstream of Ndc80 AUG, thus replacing the Ndc80luti promoter
UB14584	MATa/MATαpCUP-IME1::NAT/pCUP-IME1::NAT pCUP-IME4::NAT/pCUP-IME4::NAT amn1(BY4741 allele)unmarked/amn1(BY4741 allele)unmarked
UB18175	MATa/MATαirt1:cup1::Hphmx/irt1:cup1::Hphmx ime4::cup1::NAT/ime4::cup1::NAT leu2::SWI4-3V5- NDC80term::LEU2/leu2::SWI4-3V5-NDC80term::LEU2 Contains the 1200 bp upstream of the ORF
UB18176	MATa/MATαirt1:cup1::Hphmx/irt1:cup1::Hphmx ime4::cup1::NAT/ime4::cup1::NAT leu2::pSWI4LUTI(-1200 to -934)Δ-SWI4-3V5-NDC80term::LEU2/leu2::pSWI4LUTI(-1200 to -934)Δ-SWI4-3V5-NDC80term::LEU2 Contains the 933 bp upstream of the ORF
UB18181	MATa/MATαirt1:cup1::Hphmx/irt1:cup1::Hphmx ime4::cup1::NAT/ime4::cup1::NAT leu2::pAPL4Δ(-415 to -800)-APL4-3V5-NDC80term::LEU2/leu2::pAPL4Δ(-415 to -800)-APL4-3V5-NDC80term::LEU2 Contains the 414 bp upstream of the ORF
UB18185	MATa/MATαirt1:cup1::Hphmx/irt1:cup1::Hphmx ime4::cup1::NAT/ime4::cup1::NAT leu2::CDC60-3V5- NDC80term::LEU2/leu2::CDC60-3V5-NDC80term::LEU2 Contains the 1000 bp upstream of the ORF

Table 6.2: continued

UB18186	MATa/MATαirt1:cup1::Hphmx/irt1:cup1::Hphmx ime4::cup1::NAT/ime4::cup1::NAT leu2::pCDC60LUTI(-1000 to -483)Δ-CDC60-3V5-NDC80term::LEU2/leu2::pCDC60LUTI(-1000 to -483)Δ-CDC60-3V5-NDC80term::LEU2 Contains the 482 bp upstream of the ORF
UB18188	MATa/MATαirt1:cup1::Hphmx/irt1:cup1::Hphmx ime4::cup1::NAT/ime4::cup1::NAT leu2::pHSP60Δ(-501 to -251)-HSP60-3V5-NDC80term::LEU2/leu2::pHSP60Δ(-501 to -251)-HSP60-3V5-NDC80term::LEU2 Contains the 250 bp upstream of the ORF
UB18190	MATa/MATαirt1:cup1::Hphmx/irt1:cup1::Hphmx ime4::cup1::NAT/ime4::cup1::NAT leu2::pMSC6LUTI(-800 to -335)Δ-MSC6-3V5-NDC80term::LEU2/leu2::pMSC6LUTI(-800 to -335)Δ-MSC6-3V5-NDC80term::LEU2 Contains the 334 bp upstream of the ORF
UB18238	MATa/MATαirt1:cup1::Hphmx/irt1:cup1::Hphmx ime4::cup1::NAT/ime4::cup1::NAT leu2::MSC6-3V5-NDC80term::LEU2/leu2::MSC6-3V5-NDC80term::LEU2 Contains the 800 bp upstream of the ORF
UB18336	MATa/MATαirt1:cup1::Hphmx/irt1:cup1::Hphmx ime4::cup1::NAT/ime4::cup1::NAT leu2::HSP60-3V5-NDC80term::LEU2/leu2::HSP60-3V5-NDC80term::LEU2 Contains the 1000 bp upstream of the ORF
UB18539	MATa/MATαirt1:cup1::Hphmx/irt1:cup1::Hphmx ime4::cup1::NAT/ime4::cup1::NAT leu2::APL4-3V5-NDC80term::LEU2/leu2::APL4-3V5-NDC80term::LEU2 Contains the 1000 bp upstream of the ORF
FW1208	MATa/MATαUME6-3V5::His3MX/UME6-3V5::His3MX
FW1472	MATa/MATαirt1::pCUP-3HA-IME1::KanMX/irt1::pCUP-3HA-IME1:: KanMX ime4::pCUP-3HA-IME4::KanMX/ime4::pCUP-3HA-IME4:: KanMX NDC80-3V5::KanMX/NDC80-3V5::KanMX set2::His3MX/set2::His3MX

Table 6.2: continued

FW1511	MATa/MATα
FW1868	MATa/MATαirt1::pCUP-3HA-IME1::NatMX/irt1::pCUP-3HA-IME1:: NatMX ime4::pCUP-3HA-IME4::NatMX/ime4::pCUP-3HA-IME4:: NatMX NDC80-3V5::KanMX/NDC80-3V5::KanMX NDC80::pndc80(600-300)::His3MX/NDC80::pndc80(600-300)::His3MX
FW1871	MATa/MATαime1::pCUP-IME1::NatMX/ime1::pCUP-IME1::NatMX ime4::pCUP-IME4::NatMX/ime4::pCUP-3HA-IME4::NatMX NDC80-3V5::KanMX/NDC80-3V5::KanMX (Δ-600 to -300)-NDC80::His3MX/(Δ-600 to -300)-NDC80::His3MX
FW1899	MATa/MATαime1::pCUP-IME1::NatMX/ime1::pCUP-IME1::NatMX ime4::pCUP-IME4::NatMX/ime4::pCUP-IME4::NatMX NDC80-3V5:KanMX/(Δ-600 to -300)-NDC80::His3MX
FW1900	MATa/MATαime1::pCUP-IME1::NatMX/ime1::pCUP-IME1::NatMX ime4::pCUP-IME4::NatMX/ime4::pCUP-IME4::NatMX NDC80-3V5:KanMX
FW1902	MATa/MATαime1::pCUP-IME1::HphMX/ime1::pCUP-IME1:: HphMX ime4::pCUP-IME4::NatMX/ime4::pCUP-IME4::NatMX NDC80-3V5::KanMX/NDC80-3V5::KanMX
FW1922	MATa/MATαirt1::pCUP-3HA-IME1::HphMX/irt1::pCUP-3HA-IME1:: HphMX ime4::pCUP-3HA-IME4::NatMX/ime4::pCUP-3HA-IME4::NatMX NDC80-3V5::KanMX/NDC80-3V5::KanMX set2::His3MX/set2::His3MX set3::His3MX/set3::His3MX
FW1923	MATa/MATαime1::pCUP-IME1::NatMX/ime1::pCUP-IME1::NatMX ime4::pCUP-IME4::NatMX/ime4::pCUP-IME4::NatMX HisMX::(Δ-600 to -300)-NDC80-3V5:KanMX/NDC80

Table 6.2: continued

FW2928	MATa/MATαirt1::pCUP-3HA-IME1::HphMX/irt1::pCUP-3HA-IME1:: HphMX ime4::pCUP-3HA-IME4::NatMX/ime4::pCUP-3HA-IME4::NatMX NDC80-3V5::KanMX/NDC80-3V5::KanMX set3::His3MX/set3::His3MX
FW2929	MATa/MATαirt1::pCUP-3HA-IME1::HphMX/irt1::pCUP-3HA-IME1:: HphMX ime4::pCUP-3HA-IME4::NatMX/ime4::pCUP-3HA-IME4::NatMX NDC80-3V5::KanMX NDC80-3V5::KanMX set2::His3MX/set2::His3MX
FW2957	MATa/MATαirt1::pCUP-3HA-IME1::HphMX/irt1::pCUP-3HA-IME1:: HphMX ime4::pCUP-3HA-IME4::NatMX/ime4::pCUP-3HA-IME4::NatMX SUA7-3V5::KanMX/SUA7-3V5::KanMX
FW3033	MATa/MATαirt1::pCUP-3HA-IME1::HphMX/irt1::pCUP-3HA-IME1:: HphMX ime4::pCUP-3HA-IME4::NatMX/ime4::pCUP-3HA-IME4:: NatMX set1::KanMX/set1::KanMX
FW3058	MATa/MATαime1::His3MX/ime1::His3MX ime4::pCUP-IME4::NatMX/ime4::pCUP-IME4::NatMX NDC80-3V5::KanMX/NDC80-3V5::KanMX
FW5530	MATa/MATαSUA7-3V5::Kanmx/SUA7-3V5::Kanmx irt1::pCUP-3HA-IME1::HphMX/irt1::pCUP-3HA-IME1::HphMX ime4::pCUP-3HA-IME4::NatMX/ime4::pCUP-3HA-IME4::NatMX NDC80::pndc80(600-300)::His3MX/NDC80::pndc80(600-300)::His3MX

6.2   Time courses and growth assays

pCUP-IME1 pCUP-IME4 synchronous sporulation

For meioses in Chapter 2 synchronously sporulating cell cultures were prepared as in (Berchowitz *et al.* 2013). In short, the endogenous promoters of *IME1* and *IME4* were replaced with the inducible *CUP1* promoter. Diploid cells were grown in YPD (1% yeast

extract, 2% peptone, 2% glucose, and supplemented with 22.4 mg/L uracil and 80 mg/L tryptophan) for 20-24 hours at room temperature. For optimal aeration, the total volume of the flask exceeded the volume of the medium by 10-fold. Subsequently, cells were transferred to BYTA (1% yeast extract, 2% bacto tryptone, 1% potassium acetate, 50mM potassium phthalate) and grown for another 16-18 hours at $30\,^{\circ}$C. The cells were then pelleted, washed with sterile milliQ water, and resuspended at 1.85 OD_{600} in sporulation (SPO) media (0.5% (w/v) potassium acetate [pH 7], 0.02% (w/v) raffinose) at $30\,^{\circ}$C. To initiate synchronous sporulation, expression of *IME1* and *IME4* was induced 2 hours after cells were transferred to SPO by adding copper (II) sulphate to a final concentration of 50 μM.

For genome-wide cell collections in Chapter 3, cells were prepared to progress synchronously through meiosis as in (Chia & van Werven 2016). Briefly, liquid YPD (1% yeast extract, 2% peptone, 2% dextrose, tryptophan (96mg/L), uracil (24 mg/L), and adenine (12 mg/L)) cultures were started and grown for 6 hours at $30\,^{\circ}$C until they reached an OD_{600} between 0.5 and 2.0. They were then diluted to an OD_{600} of 0.05 in reduced YPD (1% yeast extract, 2% peptone, 1% dextrose, uracil (24 mg/L, and adenine (12 mg/L)) and allowed to grow for 16-18 hours at $30\,^{\circ}$C until they reached an OD_{600} > 6. Cells were transferred to supplemented sporulation media or SPO (1% potassium acetate at pH 7.0 supplemented with adenine and uracil to 40 mg/L and histidine, leucine, and tryptophan to 20 mg/L, and 0.02% raffinose) with a final OD_{600} of 2.5 for 2 hours at $30\,^{\circ}$C before inducing *pCUP1-IME1* and *pCUP1-IME4* with 50 μM $CuSO_4$.

In all other meiotic experiments in Chapter 3, cells were prepared as in (Berchowitz *et al.* 2013) but with 2% potassium acetate and supplements as for the genome-wide experiments. Briefly, after 24 hours of growth in YPD at RT, saturated cells (OD_{600} > 10) were diluted to an OD_{600} of 0.2-0.3 and placed in BYTA (1% yeast extract, 2% bacto tryptone, 1% potassium acetate, and 1.02% potassium phthalate) for 16-18 hours of growth at $30\,^{\circ}$C (ideally to an OD_{600} of > 10). Enough cells to give a final OD_{600} of 1.85 were transferred to SPO with 2% acetate at $30\,^{\circ}$C. After 2 hours *IME1* and *IME4* were induced with 50 μM $CuSO_4$.

pGAL-NDT80 synchronous meiotic divisions

The *pGAL-NDT80* Gal4-ER system was used to generate populations of cells synchronously undergoing the meiotic divisions (Carlile and Amon, 2008). Cells were prepared for meiosis as in the *pCUP-IME1 pCUP-IME4* protocol, and resuspended at 1.85 OD_{600} in SPO. The flasks were placed at $30\,^{\circ}$C for 5 hours to block cells in meiotic prophase. To release cells from pachytene, *NDT80* expression was induced with 1 μM β-estradiol. Subsequently, cells progressed through meiosis synchronously.

Mitotic time courses with $pGAL$ induction

For Figure 2.25, cells were grown for 6-8 hours in YPD at 30 °C, diluted to an OD_{600} of 0.002, transferred to YP (1.0% (w/v) yeast extract and 2.0% (w/v) peptone) + 2% raffinose + 2% galactose (YP-RG) and grown for another 16-18 hours. The cells were diluted to an OD_{600} of 0.2, grown for another 2.5-3 hours, diluted back to OD 0.2 and induced to express $NDC80^{LUTI}$ by the addition of 1 μM β-estradiol. 25-30 OD_{600} units of cells were collected for ChIP analyses at 0 hours and at 3 hours after induction.

For the time courses in Figure 2.26, cells were grown for 6-8 hours in YPD at 30 °C, diluted to an OD_{600} of 0.002, transferred to YP-RG and grown for another 16-18 hours. Cells were diluted to an OD_{600} of 0.2 in YP-RG and $NDC80^{LUTI}$ expression was induced by the addition of 1 μM β-estradiol. Samples were taken at 0, 3, 4.5, and 6 hours after β-estradiol addition.

For Figure 2.27B, cells were grown for 16-18 hours in YPD at 30 °C and then diluted to an OD_{600} of 0.2. Subsequently, the cells were grown for 2.5-3 hours to reach exponential phase. The cells were diluted again to an OD_{600} of 0.2 and induced to express $NDC80^{LUTI}$ with either 10 or 20 nM of β-estradiol. Cells were collected for qPCR analysis at 0 hour and at 3 hours after induction.

For Figure 2.28B-C, cells were grown in YPD at 30 °C overnight to saturation, diluted to OD_{600} of 0.1, and then grown to OD_{600} of 0.3 - 0.5 at 30 °C. Three OD_{600} of cells were taken as the pre-induction samples. Cells were then diluted to an OD_{600} of 0.1 in YPD and split into three flasks. Subsequently, βestradiol was added to the cells with a final concentration of either 15 nM or 25 nM. Ethanol was added as to the uninduced cells. 2 hours after βestradiol induction, ~ 3 OD_{600} of cells were taken for western analysis, and at 4 hours, another ~ 3 OD_{600} of cells were collected. All the samples were processed according to the western blotting protocol. The OD_{600} of each culture was also measured when the samples were taken. Equal OD_{600} units of samples was loaded during the gel electrophoresis.

α-factor arrest-release mitotic time course

MATa cells were first grown to an OD_{600} of 1-2 at 30 °C in YPD, diluted back to OD_{600} 0.005 in YEP-RG (2% raffinose and 2% galactose in YEP supplemented with 22.4 mg/L uracil and 80 mg/L tryptophan), and then grown at room temperature for 15-17 hours. Exponentially growing cells were diluted again to an OD_{600} of 0.19 in YEP-RG, and arrested in G1 with 4.15 μg/mL α-factor, and 1.5 hours later, an additional 2.05 μg/mL of α-factor was added to the cells. After 2 hours in α-factor, 1 μM β-estradiol was added to cultures to induce $pGAL$ expression. One hour after the β-estradiol addition, cells were filtered, rinsed with YEP (10 times volume of the culture volume) to remove the α-factor, and placed into a receiving flask containing YEP-RG with 1 μM β-estradiol. Time points were taken before β-estradiol induction, before release, and every 15 minutes after release, for 3 hours.

Spot growth assay

For strains harboring $NDC80^{LUTI}$ under control of the *GAL1-10* promoter, cells were first grown on YP plus 2% glycerol (YPG) plates overnight, and then re-suspended in milliQ H_2O to an OD_{600} of 0.2. Next, 5-fold serial dilutions were performed and diluted cells were spotted onto either YP-RG plates with no β-estradiol or YP-RG plates supplemented with1 μM β-estradiol. The cells were incubated at 30 °C for 1-2 days. Note that the *GAL1-10* promoter in the SK1 strain background does not directly respond to galactose. At least two independent biological experiments were performed for each spot assay experiment.

For strains harboring constructs in which $NDC80^{LUTI}$ expression is driven by LexA/lexO, cells were grown on YPD plates, re-suspended in milliQ water to an OD_{600} of 0.2, serially diluted as above, and then spotted onto either YPD plates with no β-estradiol or YPD plates with different concentrations of β-estradiol (10, 15, 20, 25, or 30 nM). The cells were incubated at 30 °C for 1 day before imaging. At least two independent biological experiments were performed for each spot assay experiment.

6.3 Conservation analyses

NDC80 clustal alignment

Clustal analysis (Goujon *et al.* 2010; Sievers *et al.* 2011) was performed using the genomic sequences of *S. bayanus*, *S. kudriavzevii*, *S. mikatae*, *S. cerevisiae* and *S. paradoxus* from *Saccharomyces sensu stricto* genus (Scannell *et al.* 2011), and imported into the Webpage of the Clustal Omega Multiple Sequence Alignment tool.

URS1 motif discovery

Meme was applied to the sequences 300 bp up and 300 bp downstream from each LUTI TSS with options *w 10 dna revcomp* (Bailey & Elkan 1994). A motif was considered significant in an individual sequence if it had a combined match p-value < 0.05.

Conservation of URS1 binding sites

The 5-extended and canonical Ume6 targets enriched with both Ume6 and a URS1 binding site in their promoters were analyzed for degree of conservation. Meme was run on the sequences +/- 300 bp from their TSSs with options *w 10 dna revcomp* was run to identify the location of the URS1 binding sites with regard to the TSS. Using a custom python script, the midpoint of each URS1 binding motif was determined relative to the ORF associated with either the LUTI or canonical target. The chromosome locations of the URS1 midpoint in the sacCer3 S288C reference genome were then found. To assess conservation of the regions around URS1 motifs at 5-extended and canonical targets, phastCons (Siepel *et al.*

2005) was performed with options *target-coverage 0.025 –expected-length 12 rho 0.4*. The tree phylogeny model and the genome alignments of *S. cerevisiae*, *S. paradoxus*, *S. mikatae*, *S. kudriavzevii*, *S. bayanus*, *S. castelli*, and *S. kluyveri* were from Siepel *et al.* (2005). However, the alignments of *S. castelli* and *S. kluyveri* were excluded in all analyses here because these yeast species have lost the *IME1* and *UME6* genes. Metagene plots and heatmaps of conservation were generated with deeptools2 (v3.0.1, (Ramírez *et al.* 2016)).

6.4 Chromatin immunoprecipitation

Chromatin immunoprecipitation qPCR

The Ume6-3V5 chromatin immunoprecipitation experiments were performed as described previously with the following modifications (van Werven *et al.* 2012). Cells were fixed with formaldehyde (1% v/v) for 15 min at room temperature and quenched with 100-125 mM glycine. Frozen cell pellets were disrupted 4 times (5 min each) using a Beadbeater (Mini-Beadbeater-96, Biospec Products, Bartlesville, OK). Chromatin was sheared 5-7 x 30 seconds ON/30 seconds OFF with a Bioruptorő Pico (Diagenode, Denville, NJ) to a fragment size of 200 bp. Chromatin extracts with 3V5-tagged DNA-binding proteins were incubated overnight with 20 μL of anti-V5 agarose beads (A7345, Sigma, St. Louis, MO) at 4 °C. The Ndt80-3V5 chromatin immunoprecipitation experiments were performed as described previously with the same modifications as used for Ume6-3V5 except for the sonication conditions (Strahl-Bolsinger *et al.* 1997). Chromatin was sheared 5 x 10 seconds ON/30 seconds OFF with a Bioruptorő Pico (Diagenode) to a fragment size of 500 bp. Extracts for chromatin modification and histone immunoprecipitation were incubated for 2 hours or overnight at 4 °C with magnetic Prot A beads (Sigma) coupled to a polyclonal antibody against Histone H3 tri methyl lysine 36 (Ab9050, Abcam), Histone H3 di methyl lysine 4 (Ab32356, Abcam) or Histone H3 (Ab1791, Abcam). Reverse cross-linking was done in Tris-EDTA buffer (100 mM Tris pH 8.0, 10 mM EDTA, 1.0% v/v SDS) at 65 °C overnight. After 2 hours of proteinase K treatment, samples were cleaned up and input DNA and immunoprecipitated DNA fragments were amplified with Absolute SYBR green (AB4163/A, Thermo Fisher, Waltham, MA). They were quantified with a 7500 Fast Real-Time PCR machine (Thermo fisher) using the relevant primer pairs. The oligonucleotide sequences used are listed in Table 6.1.

ChIP qPCR on Micrococcal nuclease (MNase) treated chromatin extracts

To determine the chromatin structure at the *NDC80* locus, we extracted mononucleosomes using a MNase digestion protocol that was described previously followed by ChIP for histone H3 (J Rando 2011). Approximately 250 OD_{600} units of cells were crosslinked for 15 min with formaldehyde (1% v/v) and the reaction was quenched with glycine (125

mM). Subsequently, cells were resuspended in 20 ml of buffer Z (1 M sorbitol, 50 mM Tris-HCl pH 7.4) plus β-mercaptoethanol (10mM) and treated with 250 μg of T100 Zymolase (MP Biomedicals) for 60 min. Next, cells were resuspended in 2.5 ml NP buffer (0.5 mM spermidine, 1 mM β-mercaptoethanol (β-ME), 0.075% (w/v) Tergitol solution-type NP-40 detergent (NP-40), 50 mM NaCl, 10 mM Tris-HCl pH 7.4, 5 mM $MgCl_2$, 1 mM $CaCl_2$), and extract was treated with 5, 0.625, 0.2 or 0.04 μl of MNase (2 mg/ml, NEB) for 30 min at 37°C, the reaction was quenched with EDTA (10 mM). The extract was adjusted to 0.1 M Hepes-KOH pH 7.5, 150 mM NaCl, 0.1% w/v sodium deoxycholate, and 1% w/v Triton X-100. To check for the extent of MNase digestion, 60 μl of MNase treated and untreated extracts were reverse crosslinked overnight in SDS-TE (1% (w/v) SDS, 10 mM Tris pH 8, 1 mM EDTA), treated with RNase A, purified DNA fragments were separated by gel electrophoresis. The extracts which showed a mono-nucleosome pattern were used for ChIP with histone H3 antibodies. The ChIP was performed with 600 μl of extract as described in the chromatin immunoprecipitation section of the materials and methods. ChIP samples were quantified by qPCR on a 7500 FAST Real-Time PCR machine (Applied Biosystems). Scanning primer pairs covering the *NDC80* locus and upstream region were used for the analysis. Signals were quantified relative to untreated genomic DNA, and normalized over a primer pair directed against the *PHO5* core promoter (Chang & Vancura 2012). The oligonucleotide sequences are available in Table 3.

Preparation of ChIP-seq libraries

For Ume6-3V5 ChIP, 300 OD_{600} units of stationary phase cells in BYTA. In histone modification experiments, 112.5 OD_{600} units of cells were collected during both starvation and early meiosis. In all instances, cells were fixed with 1% formaldehyde. The formaldehyde was quenched with 125 mM glycine, cells were pelleted, washed with PBS, and then lysed by Beadbeater (Mini-Beadbeater-96, Biospec Products, Bartlesville, OK) with zirconia beads 4 x 5 minutes in FA Buffer (50 mM Hepes pH 7.5, 150 mM NaCl, 1 mM EDTA, 1% Triton, 0.1% sodium deoxycholate, 0.1% SDS supplemented with cOmplete protease inhibitor tablets (11873580001, Roche, Indianapolis, IN)). Note that for the Ume6-3V5 ChIPs, due to the number of cells collected, lysates were prepared in 3 separate tubes. They were processed separately until after the IP. Lysates were collected and centrifuged for 3 minutes at 2000 g. The supernatants were transferred to fresh tubes and centrifuged for 15 minutes at 20,000 g. The supernatant was discarded, and the pellet of chromatin was resuspended in 1 mL FA Buffer. Samples were sonicated: 30 seconds on, 30 second off, for 5 minutes with a Bioruptor Pico (Diagenode, Denville, NJ) to an average fragment size of 200 bp. The supernatant from a 1 minute 20,000 g centrifugation was carried forward to the IP.

From the isolated chromatin, 30 μL were set aside as input. For Ume6-3V5 ChIPs, to each of 3 600 μL chromatin aliquots, 1 μL of a mouse α-V5 antibody (46-0705, Invitrogen) was added. For histone modification ChIPs, 3 μL of an antibody specific to either H3K36me3 (ab9050, abcam) or H3K4me2 (ab32356) was added to the chromatin. Samples were incu-

bated for 2 hours with nutation. Protein A Dynabeads (10001D, Invitrogen) were blocked with 0.1% BSA in FA Buffer for at least 2 hours at 4 °C. They were washed 2 x with FA Buffer, resuspended with FA Buffer to their original volume, and 10-20 µL (10 for V5 and 20 for histone modification IP) of the resuspended beads were added to each tube of chromatin. The chromatin-bead mixture was nutated at 4 °C overnight. The IP was washed 6 times: 2x FA Buffer, 2x Buffer 1 (FA Buffer, 260 mM NaCl), and 2x Buffer 2 (10 mM Tris pH 8.0, 250 mM LiCl, 0.5% NP-40, 0.5% sodium deoxycholate, 1 mM EDTA). Between each wash, samples were nutated for 5 minutes at 4 °C. To IP and input samples, 150 µL or 120 µL of TE (10 mM Tris pH 8.0, 1 mM EDTA) with 1% SDS was added, respectively. The precipitate was eluted from the beads by shaking at 1200 RPM in a Thermomixer (Thermo Fisher) at 65 °C. Eluates were treated with 0.33 mg/mL RNase A (12091021, Invitrogen) for 30 minutes at 37 °C and then with 1.2 mg/mL Proteinase K (3115879001, Roche) overnight at 65 °C. Samples were cleaned up with Qiagen QiaQuick PCR Purification Kit (28106, Qiagen) and eluted in EB. DNA was quantified by Qubit with the dsDNA HS Assay Kit (Q32854, Invitrogen). Libraries were prepared as instructed by the ThruPLEX DNA-seq Kit (R400427, Takara). For input and H3K36me3 IP samples, 15 ng of starting material was used with 7 rounds of PCR. For all other IP samples, 0.5-1 ng of starting material was used with 11 rounds of PCR. AMPure XP beads (A63881, Beckman Coulter, Brea, CA) were used to select fragments between 200-500 bp. Samples were submitted for 50 bp SE sequencing by the Vincent J. Coates Genomics Sequencing Laboratory with a HiSeq4000.

6.5   Flow cytometry

Flow cytometry was used to monitor meiotic DNA replication as described previously (Chia & van Werven 2016). Cells were fixed in 80% (v/v) ethanol and re-suspended in 50 mM Tris-HCl pH 7.5. Cells were sonicated for a few seconds and were treated with 0.2 mg/ml ribonuclease A in 50 mM Tris-HCl pH 7.5 at 37 °C overnight. Cells were stained with 50 µg/ml propidium iodide in FACs buffer (200 mM Tris-HCl pH 7.5, 211 mM NaCl and 78 mM MgCl$_2$) for one hour at room temperature before flow cytometry analysis (BD LSRFortessa, BD Biosciences). Propidium iodide stained cells were excited with a 561 nm yellow-green laser and signals were detected using a 610/20 yellow filter. Pulse shape analysis (pulse height against pulse area) was used to exclude clumps and doublets. DNA content from single cells was estimated with a histogram of counts against pulse area. At least 50,000 cells were used for the analysis.

6.6   Microscopy and image quantification

Fluorescence microscopy (CENV-GFP dots and Spc42-mCherry)

Cells were fixed with 3.7% formaldehyde at room temperature for 15 min, washed once with potassium phosphate/sorbitol buffer (100 mM potassium phosphate [pH 7.5], 1.2 M

sorbitol), and then permeabilized with 1% Triton X-100 with 0.05 μg/ml DAPI in potassium phosphate/sorbitol buffer. Cells were imaged using a DeltaVision microscope with a 100x/1.40 oil-immersion objective (DeltaVision, GE Healthcare, Sunnyvale, CA) and filters: DAPI (EX390/18, EM435/48), GFP/FITC (EX475/28, EM525/48), and mCherry (EX575/25, EM625/45). Images were acquired using the softWoRx software (softWoRx, GE Healthcare).

Nuclei/DAPI counting

DAPI staining was used to monitor meiotic divisions throughout meiotic time courses. Cells were fixed in 80% (v/v) ethanol, pelleted by centrifugation and re-suspended in PBS with DAPI (1 μg/ml). Cells were sonicated for a few seconds and left in the dark at room temperature for at least 5 minutes. The proportion of cells containing one, two, three, or four DAPI masses was counted using a fluorescence microscope. At least two independent biological experiments were performed for each meiotic time-course experiment.

Quantification of spindle length and CENV-GFP dots in mitosis

For Figure 2.23 and Figure 2.24, diploid cells were first grown to an OD_{600} of 1-2 at 30 °C in YPD. They were then diluted to an OD_{600} of 0.002 in YEP-RG and grown at 30 °C for 16 hours. Exponentially growing cells were diluted back to an OD_{600} of 0.2 in YEP-RG and induced to express $NDC80^{LUTI}$ with 1 μM of β-estradiol. Samples were taken before induction and 6 hours after induction. Images were acquired as described in the fluorescence microscopy method section, and analysed using the FIJI image processing software (RRID:SCR_002285, (Schindelin *et al.* 2012)). First, maximum-intensity projection was performed. Second, projected spindle length (defined as the distance between Spc42-mCherry foci) was measured using the measure plugin. The distribution of the projected spindle length was graphed as violin plots using BoxPLotR (RRID:SCR_015629, Spitzer et al., 2014). Third, in cells with separated spindle poles, the status of the Spc42-mCherry association with CENV-GFP dots was categorized as 1) each Spc42-mCherry focus is associated with a CENV-GFP dot, 2) only one Spc42-mCherry focus is associated with CENV-GFP dots (either one or both of the GFP dots), or 3) neither Spc42-mCherry focus is associated with a CENV-GFP dot. After categorizing the localization of the CENV-GFP dots, the projected spindle length was measured for spindles in category 2 and 3, and the spindle length distributions were graphed as violin plots using BoxPLotR (RRID:SCR_015629, (Spitzer *et al.* 2014)). Finally, in cells with separated spindle poles, the location of the spindle was recorded as 1) in the mother, 2) across the bud neck, or 3) in the bud. The percentage of spindles that were both less than 2.0 μm and abnormally localized (across the bud neck or in the bud) was calculated. For each analysis, 100 cells were counted.

6.7   Molecular Biology

Northern blotting

In Chapter 2, previously described northern blot protocol was modified as below (Koster *et al.* 2014). RNA was extracted with acid phenol:chloroform:isoamyl alcohol (125:24:1; pH 4.7) and then isopropanol precipitated. RNA samples (8-10 µg) were denatured in a glyoxal/DMSO mix (1M deionized glyoxal, 50% v/v DMSO, 10 mM sodium phosphate buffer pH 6.5-6.8) at 70 °C for 10 min and then separated on a 1.1% agarose gel for 3 hours at 80 V. RNAs were transferred onto nylon membranes overnight by capillary transfer. rRNA bands were visualized by methylene blue staining. The membranes were blocked for at least 3 hours at 42 °C in ULTRAhybő Ultrasensitive Hybridization Buffer (Thermo Fisher) before hybridization. Radioactive probes were synthesized using a Prime-It II Random Primer Labelling Kit (Agilent, Santa Clara, CA). Membranes were washed twice in Low Stringency Buffer (2x SSC, 0.1%SDS) and three times in High Stringency Buffer (0.1X SSC, 0.1% SDS). All hybridization and wash steps were done at 42 °C. The oligonucleotide sequences of the primers used to amplify the *NDC80*, *NUF2*, *SCR1*, and *CIT1* DNA template primers are displayed in Table 6.1.

Quantification was performed with FIJI (RRID:SCR_002285, (Schindelin *et al.* 2012)). For all the images, the LUT (lookup table) was inverted. Then, a rectangular box was drawn around a band of interest. The mean signal intensity (gray-scale) within the box area was calculated using the measure plugin. For background subtraction, the same box was moved directly above and below the band, the signal intensity of these two regions was measured, and the average background intensity (top and bottom) was calculated. After subtracting the average background intensity of a given lane from the signal intensity of the band in that lane, this corrected value for each time point was then normalized to the initial time point. The same-sized box was used for all the time points in one experiment.

In Chapter 3, RNA blot analysis protocol was performed similarly but with minor modifications. 8 ug of total RNA was denatured in a glyoxal/DMSO mix at 70 °C for 10 min. Denatured samples were mixed with loading buffer and separated on an agarose gel (1.1% w/v agarose, 0.01 M NaPi buffer) for 3 hr at 100 V. The gels were then soaked for 25 minutes in denaturation buffer (0.05 N NaOH, 0.15 M NaCl) followed by 20 minutes in neutralization buffer (0.1 M Tris-HCl pH 7.5, 0.15 M NaCl). RNA was transferred to nitrocellulose membrane for 1 hour via vacuum transfer as described in Stratagenes Membranes Instruction Manual (Stratagene, La Jolla, CA). All remaining steps were performed as in Chapter 2. The oligonucleotide sequences of the primers used to amplify the 3V5 epitope DNA template primers are displayed in Table 6.1.

RT-qPCR

For the RT-PCR, RNA was isolated by acid phenol-chloroform extraction, treated with DNase (TURBO DNA-free kit, Thermo Fisher), and reverse transcribed into cDNA (Superscript III Supermix, Thermo Fisher). The cDNA was quantified using the Absolute Blue qPCR Mix (Thermo Fisher). The $NDC80^{LUTI}$ and $NUF2^{LUIT}$ signals were normalized to *ACT1* transcript levels. The oligonucleotide sequences used for RT-PCR experiments are displayed in Table 6.1.

Immunoblot

Protein extracts were prepared using a trichloroacetic acid (TCA) extraction protocol. Briefly, $\sim 4\,\mathrm{OD}_{600}$ units of cells were treated with 5% trichloroacetic acid for at least 15 min at $4\,°\mathrm{C}$. For vegetative samples, pellets were washed with TE50 buffer (50 mM Tris pH 7.5, 1 mM EDTA), then with acetone, and completely dried. For meiotic samples, the pellets were only washed with acetone and dried. The cell pellet was lysed with glass beads in lysis buffer (50 mM TrisHCl [pH 7.5], 1 mM EDTA, 2.75 mM DTT, protease inhibitor cocktail (cOmplete EDTA-free, Roche, Basel, Switzerland) using a Mini-Beadbeater-96 (Biospec Products). Next, 3x SDS sample buffer (187.5 mM Tris [pH 6.8], 6% β-mercaptoethanol, 30% glycerol, 9% SDS, 0.05% bromophenol blue) was added and the cell lysate was boiled for 5 min. Proteins were separated by PAGE using 4%-12% Bis-Tris Bolt gels (Thermo Fisher) and transferred onto nitrocellulose membranes (0.45 μm, Bio-rad, Hercules, CA) using a semi-dry transfer apparatus (Trans-Blotő TurboTM Transfer System, Bio-rad). The membranes were blocked for at least 30 min with Odysseyő Blocking Buffer (PBS) (LI-COR Biosciences, Lincoln, NE) before incubation overnight at $4\,°\mathrm{C}$ with a mouse anti-V5 antibody (RRID:AB_2556564, R960-25, Thermo Fisher) at a 1:2,000 dilution. We monitored Hxk1 levels using a rabbit anti-hexokinase antibody (RRID:AB_2629457, H2035, US Biological, Salem, MA) at a 1:10,000-20,000 dilution, Pgk1 levels with a 1:10,000 diluted mouse anti-Pgk1 antibody (RRID:AB_2687965, SC7167, Molecular Probes, Carlsbad, CA), and Kar2 levels with a 1:200,000 rabbit anti-Kar2 antibody (provided by Mark Rose). Membranes were washed in PBST (phosphate buffered saline with 0.01% tween-20) and incubated with an anti-mouse secondary antibody conjugated to IRDyeő 800CW at a 1:15,000 dilution (RRID:AB_621847, 926-32212, LI-COR Biosciences) and an anti-rabbit antibody conjugated to IRDyeő 680RD at a 1:15,000 dilution (RRID:AB_10956166, 926-68071, LI-COR Biosciences). Immunoblot images were generated and quantified using the Odysseyő system (LI-COR Biosciences). Intensities of 3V5 and Hxk1 bands on western blots were quantified using Image Studio Lite (LI-COR). Ndc80 levels were first normalized to Hxk1 levels and further normalized to that of the first time point on the same membrane.

6.8   Preparation of genome-wide sequencing libraries

Common RNA preparation for TL-seq, Nanopore sequencing, and RNA-seq

At the indicated timepoints, 50 OD_{600}units of cells were collected by vacuum filtration and snap frozen in liquid nitrogen. Cells were thawed on ice and resuspended in TES (10 mM Tris pH 7.5, 10 mM EDTA, 0.5% SDS) to 20 OD_{600}. An equal volume of Acid Phenol:Chloroform:Isoamyl alcohol (125:24:1; pH 4.7) was added to cells, and they were incubated at 65 °C for 45 minutes in a Thermomixer C (Eppendorf) shaking at 1400 RPM. The aqueous phase was transferred to a second tube of acid phenol. Samples were incubated at RT for 5 minutes while shaking at 1400 RPM in a Thermomixer. A final extraction with chloroform was performed. The aqueous phase was vortexed with the chloroform for 30 seconds, separated by centrifugation, and then precipitated in isopropanol and sodium acetate overnight at −20 °C. Pellets were washed with 80% ethanol and resuspended in DEPC water for 10 min at 37 °C. Total RNA was quantified using the Qubit RNA BR Assay Kit (Q10211, Thermo Fisher).

Transcript Leader Sequencing (TL-seq)

The 5 end sequencing approach was performed as in Wu *et al.* (2018). At least 5 μg of mRNA was purified from total RNA using the Poly(A)Purist MAG kit (AM1922, Ambion). mRNAs were fragmented for 3 minutes at 70 °C using a Zinc-based alkaline fragmentation reagent (AM8740, Ambion). RNAs were cleaned up using RNeasy MinElute Cleanup Kits (74204, Qiagen) to enrich for 200-300 nt fragments. These fragments were dephosphorylated with 30 units of recombinant shrimp alkaline phosphatase (M0371, NEB) for 1 hour at 37 °C with RNasin Plus (N2611, Promega). The RNA was extracted with Acid Phenol:Chloroform:Isoamyl alcohol (125:24:1, pH 4.7) and precipitated in ethanol with 0.3M sodium acetate and 1 μl linear acrylamide (AM9520, Ambion). RNA was then subjected to a decapping reaction with 2 units of Cap-Clip acid pyrophosphatase (C-CC15011H, Tebu-Bio) and with RNasin Plus. RNA was then again extracted using acid Phenol:chloroform:isoamyl alcohol (125:24:1) and precipitated in ethanol. Some RNA from a starvation time point was set apart without the decapping reaction as a non-decapping control. Subsequently, the RNA was mixed with 10 μM of custom 5 adapter and the ligation reaction was done using T4 RNA ligase 1 (M0437M, NEB) and with RNasin Plus. The ligation reaction was cleaned up with the RNeasy MinElute Cleanup Kit (74204, Qiagen) and RNAs were mixed with 2.5 μM random hexamers (N8080127, ThermoFisher Scientific) and RNasin Plus, denatured at 65 °C for 5 minutes and cooled on ice. Reverse transcription reactions were carried out using SuperScript IV reverse transcriptase (18090010, Invitrogen). The RNA templates were degraded by incubating reactions with 5 units of RNase H (M0297, NEB) and 1.0 μL of RNase cocktail enzyme mix (AM2286, Ambion). DNA products were purified using 1.8x volume of HighPrep PCR beads (AC-60050, MagBio). Purified products were subjected to

second strand synthesis using 0.3 µM of second strand biotinylated primer and the KAPA Hi-Fi hot start ready mix (KK2601, Roche). The second strand reaction was carried out at 95 °C for 3 minutes, 98 °C for 15 seconds, 50 °C for 2 minutes, 65 °C for 15 minutes and held at 4 °C. Double stranded product (dsDNA) was purified with 1.8x volume HighPrep PCR beads and concentration was quantified using the Qubit dsDNA HS assay kit (Q32851, Invitrogen). 25 ng of dsDNA was then used as input for the KAPA Hyper Prep Kit (KK8504, Roche) and ligated to KAPA single indexed adapters Set B (KK8702, Roche). Samples were processed according to manufacturers instructions with one exception: just prior to the library amplification step, samples were bound to MyOne Streptavidin C1 Dynabeads (65001, ThermoFisher Scientific) to capture biotinylated dsDNA. Library amplification over 14 PCR cycles was done on the biotinylated dsDNA fraction bound to the beads. Amplified libraries were quantified by Qubit, and adapter-dimers were removed by electrophoresing libraries on Novex 6% TBE gels (EC62655BOX, Invitrogen) at 120 V for 1 hour, and excising the smear above 150 bp. Gel slices containing libraries were shredded by centrifugation at 13000 g for 3 minutes. Gel shreds were re-suspended in 500 µL crush and soak buffer (500 mM NaCl, 1.0 mM EDTA and 0.05% v/v SDS) and incubated at 65 °C for 2 hours on a thermomixer (1400 rpm for 15 seconds, rest for 45 seconds). Subsequently, the buffer was transferred into a Costar SpinX column (8161, Corning Incorporated) with two 1 cm glass pre-filters (1823010, Whatman). Columns were centrifuged at 13000 g for 1 minute. DNA libraries in the flow through were precipitated at −20 °C overnight in ethanol with 0.3 M sodium acetate and 1 µL linear acrylamide (AM9520, Ambion). Purified libraries were further quantified and inspected on a Tapestation (Agilent Technologies). They were sent for 100 bp SE sequencing on an Illumina HiSeq 4000 at the Vincent J. Coates Genomics Sequencing Laboratory.

polyA selection for Nanopore sequencing and RNA-seq

PolyA selection was performed on 100 µg of RNA using 150 µL of oligo-dT DynaBeads (61002, Thermo Fisher). RNA was denatured at 80 °C for 2 minutes in binding buffer (10 mM Tris-HCl pH 7.5, 0.5 M LiCl, 3.35 mM EDTA) before being placed on ice. At RT the oligo-dT beads were added to the sample and together they were incubated at room temperature of 5 minutes. Beads were washed 2x in Buffer B (10 mM Tris-HCl pH 7.5, 0.15 M LiCl, 1 mM EDTA). PolyA-selected RNA was eluted from the beads by heating at 80 °C for 2 minutes in 10 mM Tris 7.0. It was quantified with a Qubit using the RNA HS assay kit (Q32852, Thermo Fisher).

Nanopore sequencing

500 ng of polyA-selected RNA was used as directed in for the Direct RNA Sequencing Kit (SQK-RNA001, Oxford Nanopore Technologies, Oxford Science Park, UK). The library was loaded onto a minION (MIN-101B, Oxford Nanopore Technologies) with an R9.4.1 flow cell (FLO-MIN106, Oxford Nanopore Technologies). MinKNOW (v1.10.23, Oxford Nanopore Technologies) was run without live basecalling for 48 hours.

RNA-seq library preparation

The RNA-seq libraries were prepared using the NEXTflexTM Rapid Directional mRNA-Seq Kit (Bioo Scientific, Austin, Texas) according to manufacturer instructions. 100ng of poly-A selected mRNA was used for all libraries. Libraries were quantified using the Agilent 4200 TapeStation (Agilent Technologies, Inc., Santa Clara, CA). AMPure XP beads (A63881, Beckman Coulter, Brea, CA) were used to select fragments between 200-500 bp. Samples were submitted for 150 bp SE sequencing by the Vincent J. Coates Genomics Sequencing Laboratory with a HiSeq4000.

Micrococcal nuclease sequencing library preparation

The protocol was adapted from Basic Protocol 1 in (Rodriguez *et al.* 2014) with the following changes. In starvation and early meiosis, 112.5 OD_{600} units of cells were fixed in 1% formaldehyde with light shaking at RT for 15 minutes. Crosslinking was quenched by 125 mM of glycine for 5 minutes at RT. Cells were pelleted and washed twice with ice cold milliQ water. Cells were spheroblasted in 20 mL of Spheroblast Solution (1 M Sorbitol, 50 mM Tris pH 7.5, 10 mM β-ME) with 100 μL of 10 mg/mL zymolase until they appeared non-refractive and shadow-like (20-30 minutes). Spheroblasted cells were resuspended in 2 mL MNase Digestion Buffer (1 M Sorbitol, 50 mM NaCl, 10 mM Tris pH 7.5, 5 mM $MgCl_2$, 1 mM $CaCl_2$, 0.075% NP-40, 0.5 mM spermidine, 1 mM β-ME) if collected during starvation and 4 mL if collected during early meiosis, after completion of S-phase. Digestions were performed with 600 uL of spheroblasts, 30 units of Exonucleasae III (M0206S, New England Biolabs), and either 10, 20, or 40 units of MNase (LS004797, Worthington). Crosslinks were reversed, protein was degraded by Proteinase K (3115879001, Roche), and a phenol/chloroform/isoamyl alcohol DNA extraction, ethanol precipitation, RNase A (12091021, Invitrogen) treatment, and phosphatase treatment were performed as described previously (Rodriguez *et al.* 2014). Size selection was performed by running samples on a 1.8% LMT agarose gel at 80 V for 40 minutes at room temperature and gel extracting the mononucleosome band with a Monarch Gel Extraction Kit (T1020S, New England Biolabs). Note that of the samples digested with 10, 20, and 40 units of MNase, only the sample with a ratio of mononucleosomes to dinucleosomes closest to 80/20 were size selected and carried forward for library preparation. Gel extracted samples were quantified by Qubit with the dsDNA HS Assay Kit (Q32854, Invitrogen). Libraries were prepared with 50 ng starting material as instructed by the ThruPLEX DNA-seq kit (R400427, Takara). Amplification was performed with 5 rounds of PCR. AMPure XP beads (A63881, Beckman Coulter) were used to select fragments between 150-500 bp. Samples were submitted for 100 bp PE sequencing to the Vincent J. Coates Genomics Sequencing Laboratory with a HiSeq4000.

6.9 Genomic analyses

TL-seq analysis

From the sequencing reads, the 3 Illumina adapter (AGATCGGAAGAGC) was trimmed using cutadapt with the `-minimum-length` option set to 20 bp (v2.3,(Compeau *et al.* 2013)). From the 3 trimmed output, the 5 Illumina adapter (CACTCTGAGCAATACC) was trimmed from reads by cutadapt. To select for reads with the most 5 end of a transcript, only reads in which the 5 adapter was recognized and then trimmed were carried forward. Reads were aligned by STAR (v2.5.3a, (Dobin *et al.* 2013)) using indices generated from an SK1 genome assembled by combined PacBio and Illumina sequencing (Yue *et al.* 2017). A custom SK1 genome was forged with BSgenome (v1.50.0, (Pagè 2018)) using the above assembly (Yue *et al.* 2017). Bam files were imported into CAGEr and the CAGEr pipeline was applied to define TSSs and quantify transcript abundances as follows (v1.24.0, (Haberle *et al.* 2015)). Reads at TSSs were counted (getCTSS) and normalized by `"simpleTpm"` (normalizeTagCount). An initial clustering was performed (clusterCTSS with `threshold = 2, thresholdIsTpm =TRUE, method = "distclu", maxDist = 5, removeSingletons = TRUE`, and `keepSingletonsAbove = 3`), and the output was aggregated into larger clusters representative of all the activity expected from a single promoter (aggregateTagClusters with `tpmThreshold = 1` and `maxDist = 50`). Clustered TSSs were exported as bedGraph files for visualization in IGV (exportCTStoBedGraph with `values = "normalized"`). Clusters counts were exported to DESeq2 by timepoint (concensusClustersDESeq2), and fold-changes were calculated by DESeq2 with default settings (Love *et al.* 2014). Output from this clustering was used to define TSSs coordinates of 5-extended transcripts in the pipeline below. A secondary and more permissive clustering (`threshold = 1, thresholdIsTpm =TRUE, method = "distclu", maxDist = 5, removeSingletons = FALSE`) was performed after LUTI-mRNA genes were defined. Output from the secondary clustering was used for quantification and in all presented TL-seq scatterplots. The TL-seq used to define LUTI promoters was performed in triplicate. All TL-seq comparisons to RNA-seq were performed in duplicate.

Nanopore sequencing analysis

Bases were called from fast5 files with the Albacore script `read_fast5_basecaller.py` (v2.1.10, Oxford Nanopore Technologies). 491,142 reads were sequenced. Reads were aligned to the genome with minimap2 (v2.9-r720, (Li 2018)) using options `-ax splice -k14 -uf`. Bam files were visualized directly in IGV.

Pipeline for 5'-extended transcript discovery

Using the output from DESeq2 after CAGEr, TSS clusters were filtered for coordinates in which the mean over both time points was > 2 transcripts per million and the log2 fold-

change as cells entered early meiosis compared to starvation was > 2. After applying these filters, the coordinates for each peak were manually inputted into IGV. The TL-seq peak was compared to Nanopore sequencing reads from a sample taken during early meiosis (4 hours). If at least one Nanopore read extended from a region near the TSS coordinates and continued uninterrupted across the entirety of a neighboring ORF, the coordinates were marked for continued investigation. Purely intergenic and either 5 or 3 truncated transcripts were removed in this way. From the remaining subset of peaks, a 5-extension was only called if a second promoter, downstream, but on the same strand, was closer to the ORF. Through this criterion, canonical meiosis-specific genes were eliminated from the analysis. It resulted in 74 candidate LUTI-mRNAs with 5-extensions. For downstream analyses, the single most dominant bp in each TSS cluster was determined by a custom python script.

uORF analysis

ATGs were counted and the codon frequency was determined with a custom python script. For genes with LUTIs, the counts and codon frequencies were determined for the region between the PROX TSS and the LUTI TSS. For all other genes, sequences from the 500 bp upstream of the TSS were used.

LUTI mRNAs with > 3 uORFs were analyzed to determine which of the uORFs were translated. Footprints were quantified for the first 6 codons of each uORF using the tools in Brar et al. (2012) and Ingolia et al. (2012) written by NIcholas Ingolia. The ribosome footprinting data was taken from the 3 hr timepoint in Cheng *et al.* (2018). Any uORFs with at least 2 footprint reads found across the third through sixth codons of the gene were considered to be actively translated. The cutoff of 2 footprint reads was used as the background in the UTR of PROX transcripts (the region from the PROX TSS until the ORF was 2.23 footprint per 4 codons.

RNA-seq analysis

Quantification of RNA as transcripts per million was done using salmon in the mapping-based mode with mapping validation (v0.13.1, (Patro *et al.* 2017)). Fold-change quantification was performed by DESeq2 with counts generated by summarizeOverlaps using default options (v1.22.2, (Love *et al.* 2014)). Scatterplots were generated with matplotlib (Hunter 2007).

Chromatin immunoprecipitation sequencing (ChIP-seq) analysis

Ume6 ChIP

Reads were aligned to the SK1 genome with bowtie2 (v2.3.4.3, (Langmead & Salzberg 2012; Langmead *et al.* 2019)). Using randsample from macs2, all libraries were down-sampled

to 2 million reads (v2.1.1.20160309, (Zhang *et al.* 2008)). Macs2 callpeak was used to call peaks in IP samples over input samples with options *B q 0.001 –keep-dup all –call-summits nomodel extsize 147*. Bigwig files for viewing in IGV were generated by macs2 bdgcmp with option *m FE* followed by bedGraphToBigWig (v4, (Kent *et al.* 2010)). Heatmaps and metagene plots centered around TSSs as defined by TL-seq were constructed with deeptools2 (v3.0.1, (Ramírez *et al.* 2016)). A 5-extended or canonical target promoter was considered to be enriched by Ume6 if, in at least 2 of 3 ChIP replicates, a peak was called (log2 fold-change > 2 over input) within 300 bp of the transcripts TSS.

Chromatin modifications

Reads were aligned to the SK1 PacBio genome assembly with bowtie2 (v2.3.4.3, (Langmead & Salzberg 2012)). Macs2 callpeak was used to call peaks in IP samples over input samples with options *B q 0.01 nomodel extsize 147*. Bigwig files for viewing in IGV and for further quantification were generated by macs2 bdgcmp with option *m FE* followed by bedGraphToBigWig (v4, (Kent *et al.* 2010; Zhang *et al.* 2008)). To quantify the change in H3K36me3 and H3K4me2 enrichment over the promoters of PROX transcripts, fold enrichment scores were extracted from regions 50 bp upstream and 500 bp downstream from the PROX TSS with bedtools (Quinlan & Hall 2010). With custom python scripts, the scores from each bp of the upstream regions and each bp of the downstream regions were summed for each gene. The ratio of the upstream and the downstream region enrichments were quantified and the change in the score from starvation to early meiosis was determined. Ultimately, the mean of the fold-change was calculated from samples in triplicate. Heatmaps and metagene plots were prepared with deeptools2 (Ramírez *et al.* 2016).

Microccocal Nuclease Sequencing (MNase-seq) analysis

Reads were aligned to the SK1 genome with bowtie2 (v2.3.4.3, (Langmead & Salzberg 2012)). To select for only fragments of between 130 and 170 bps, alignmentSieve from deeptools2 was performed (Ramírez *et al.* 2016). BigWig files were generated by bamCoverage with options *–MNase bs 1 –normalizeUsing CPM* (Ramírez *et al.* 2016). DANPOS (v2.2.2) was run to determine various aspects of nucleosome location and occupany and fuzziness (Chen *et al.* 2013; Chen *et al.* 2015). A custom python script was used to assign the locations of +1 and 1 nucleosome with respect to PROX TSSs.

Bibliography

1. Aanes, H., Østrup, O., Andersen, I. S., Moen, L. F., Mathavan, S., Collas, P. & Alestrom, P. Differential transcript isoform usage pre- and post-zygotic genome activation in zebrafish. *BMC Genomics* **14.** doi:`10.1186/1471-2164-14-331` (2013).

2. Acquaviva, L., Székvölgyi, L., Dichtl, B., Dichtl, B. S., de La Roche Saint Andre, C., Nicolas, A. & Géli, V. The COMPASS Subunit Spp1 Links Histone Methylation to Initiation of Meiotic Recombination. *Science* **339** (2013).

3. Ambros, V. microRNAs: Tiny Regulators with Great Potential. *Cell* **107,** 823–826 (2001).

4. Anderson, E. L., Baltus, A. E., Roepers-Gajadien, H. L., Hassold, T. J., de Rooij, D. G., van Pelt, A. M. M. & Page, D. C. Stra8 and its inducer, retinoic acid, regulate meiotic initiation in both spermatogenesis and oogenesis in mice. *Proceedings of the National Academy of Sciences* **105,** 14976–14980 (2008).

5. Ard, R. & Allshire, R. C. Transcription-coupled changes to chromatin underpin gene silencing by transcriptional interference. *Nucleic Acids Research* **44,** 10619–10630 (2016).

6. Arribere, J. A. & Gilbert, W. V. Roles for Transcript Leaders in Translation and mRNA Decay Revealed by Transcript Leader Sequencing. *Genome Research,* 977–987 (2013).

7. Asakawa, H., Hayashi, A., Haraguchi, T. & Hiraoka, Y. Dissociation of the Nuf2-Ndc80 Complex Releases Centromeres fron the Spindle-Pole Body during MEiotic Prophase in Fission Yeast. *Molecular biology of the cell* **16.** doi:`10.1091/mbc.E04` (2005).

8. Bailey, T. L. & Elkan, C. Fitting a mixture model by expectation maximization to discover motifs in biopolymers. *Proceedings. International Conference on Intelligent Systems for Molecular Biology* **2,** 28–36 (1994).

9. Bannister, A. J. & Kouzarides, T. The CBP co-activator is a histone acetyltransferase. *Nature* **384,** 641–643 (1996).

10. Bannister, A. J., Zegerman, P., Partridge, J. F., Miska, E. A., Thomas, J. O., Allshire, R. C. & Kouzarides, T. Selective recognition of methylated lysine 9 on histone H3 by the HP1 chromo domain. *Nature* **410,** 120–124 (2001).

11. Barak, Y., Gottlieb, E., Juven-Gershon, T. & Oren, M. Regulation of mdm2 expression by p53: Alternative promoters produce transcripts with nonidentical translation potential. *Genes and Development* **8,** 1739–1749 (1994).

12. Barrett, L. W., Fletcher, S. & Wilton, S. D. Regulation of eukaryotic gene expression by the untranslated gene regions and other non-coding elements. *Cellular and Molecular Life Sciences* **69,** 3613–3634 (2012).

13. Batista, P. J. & Chang, H. Y. Long noncoding RNAs: Cellular address codes in development and disease. *Cell* **152,** 1298–1307 (2013).

14. Batut, P., Carninci, P., Dobin, A., Gingeras, T. R. & Plessy, C. High-fidelity promoter profiling reveals widespread alternative promoter usage and transposon-driven developmental gene expression. *Genome Research* **23,** 169–180 (2013).

15. Benjamin, K. R., Zhang, C., Shokat, K. M. & Herskowitz, I. Control of landmark events in meiosis by the CDK Cdc28 and the meiosis-specific kinase Ime2. *Genes and Development* **17,** 1524–1539 (2003).

16. Berchowitz, L. E., Gajadhar, A. S., van Werven, F. J., De Rosa, A. a., Samoylova, M. L., Brar, G. a., Xu, Y., Xiao, C., Futcher, B., Weissman, J. S., White, F. M. & Amon, A. A developmentally regulated translational control pathway establishes the meiotic chromosome segregation pattern. *Genes & development* **27,** 2147–63 (2013).

17. Bird, A. J., Gordon, M., Eide, D. J. & Winge, D. R. Repression of ADH1 and ADH3 during zinc deficiency by Zap1-induced intergenic RNA transcripts. *EMBO Journal* **25,** 5726–5734 (2006).

18. Blair, J. D., Hockemeyer, D., Doudna, J. A., Bateup, H. S. & Floor, S. N. Widespread Translational Remodeling during Human Neuronal Differentiation. *Cell Reports* **21,** 2005–2016 (2017).

19. Borde, V., Robine, N., Lin, W., Bonfils, S., Géli, V. & Nicolas, A. Histone H3 lysine 4 trimethylation marks meiotic recombination initiation sites. *EMBO Journal* **28,** 99–111 (2009).

20. Bordoy, A. E., Varanasi, U. S., Courtney, C. M. & Chatterjee, A. Transcriptional Interference in Convergent Promoters as a Means for Tunable Gene Expression. *ACS Synthetic Biology* **5,** 1331–1341 (2016).

21. Bowdish, K. S., Yuan, H. E. & Mitchell, A. P. Positive control of yeast meiotic genes by the negative regulator UME6. *Molecular and Cellular Biology* **15,** 2955–2961 (1995).

22. Brar, G. A., Yassour, M., Friedman, N., Regev, A., Ingolia, N. T. & Weissman, J. S. High-resolution view of the yeast meiotic program revealed by ribosome profiling. *Science (New York, N.Y.)* **335,** 552–7 (2012).

23. Brown, C. J., Hendrich, B. D., Rupert, J. L., Lafrenière, R. G., Xing, Y., Lawrence, J. & Willard, H. F. The human XIST gene: Analysis of a 17 kb inactive X-specific RNA that contains conserved repeats and is highly localized within the nucleus. *Cell* **71,** 527–542 (1992).

24. Brown, C. Y., Mize, G. J., Pineda, M., George, D. L. & Morris, D. R. Role of two upstream open reading frames in the translational control of oncogene mdm2. *Oncogene* **18,** 5631–5637 (1999).

25. Brownell, J. E. & Allis, C. D. An activity gel assay detects a single, catalytically active histone acetyltransferase subunit in Tetrahymena macronuclei. *Proceedings of the National Academy of Sciences* **92,** 6364–6368 (1995).

26. Brownell, J. E., Zhou, J., Ranalli, T., Kobayashi, R., Edmondson, D. G., Roth, S. Y. & Allis, C. D. Tetrahymena histone acetyltransferase A: A homolog to yeast Gcn5p linking histone acetylation to gene activation. *Cell* **84,** 843–851 (1996).

27. Buetti-Dinh, A., Ungricht, R., Kelemen, J. Z., Shetty, C., Ratna, P. & Becskei, A. Control and signal processing by transcriptional interference. *Molecular Systems Biology* **5,** 1–8 (2009).

28. Bumgarner, S. L., Dowell, R. D., Grisafi, P., Gifford, D. K. & Fink, G. R. Toggle involving cis-interfering noncoding RNAs controls variegated gene expression in yeast. *Proceedings of the National Academy of Sciences* **106,** 1–6 (2009).

29. Byrne, A., Beaudin, A. E., Olsen, H. E., Jain, M., Cole, C., Palmer, T., DuBois, R. M., Forsberg, E. C., Akeson, M. & Vollmers, C. Nanopore long-read RNAseq reveals widespread transcriptional variation among the surface receptors of individual B cells. *Nature Communications* **8,** 1–11 (2017).

30. Cairns, B. R. The logic of chromatin architecture and remodelling at promoters. *Nature* **461,** 193–8 (2009).

31. Calvo, S. E., Pagliarini, D. J. & Mootha, V. K. Upstream open reading frames cause widespread reduction of protein expression and are. *Proceedings of the National Academy of Sciences* **106,** 7507–7512 (2009).

32. Camblong, J., Iglesias, N., Fickentscher, C., Dieppois, G. & Stutz, F. Antisense RNA Stabilization Induces Transcriptional Gene Silencing via Histone Deacetylation in S. cerevisiae. *Cell* **131,** 706–717 (2007).

33. Carroll, S. B. Evo-Devo and an Expanding Evolutionary Synthesis: A Genetic Theory of Morphological Evolution. *Cell* **134,** 25–36 (2008).

34. Carrozza, M. J., Li, B., Florens, L., Suganuma, T., Swanson, S. K., Lee, K. K., Shia, W. J., Anderson, S., Yates, J., Washburn, M. P. & Workman, J. L. Histone H3 methylation by Set2 directs deacetylation of coding regions by Rpd3S to suppress spurious intragenic transcription. *Cell* **123,** 581–592 (2005).

35. Cech, T. R. & Steitz, J. A. The Noncoding RNA RevolutionTrashing Old Rules to Forge New Ones. *Cell* **157,** 77–94 (2014).

36. Chang, J. & Vancura, A. in *Transcriptional Regulation: Methods and Protocols* (ed Vancura, A.) 321–333 (Springer New York, New York, NY, 2012). ISBN: 978-1-61779-376-9. doi:`10.1007/978-1-61779-376-9_22`. `https://doi.org/10.1007/978-1-61779-376-9%7B%5C_%7D22`.

37. Chen, J., Tresenrider, A., Chia, M., McSwiggen, D. T., Spedale, G., Jorgensen, V., Liao, H., van Werven, F. J. & Ünal, E. Kinetochore inactivation by expression of a repressive mRNA. *eLife* **6,** 1–31 (2017).

38. Chen, K., Chen, Z., Wu, D., Zhang, L., Lin, X., Su, J., Rodriguez, B., Xi, Y., Xia, Z., Chen, X., Shi, X., Wang, Q. & Li, W. Broad H3K4me3 is associated with increased transcription elongation and enhancer activity at tumor-suppressor genes. *Nature Genetics* **47,** 1149–1157 (2015).

39. Chen, K., Xi, Y., Pan, X., Li, Z., Kaestner, K., Tyler, J., Dent, S., He, X. & Li, W. DANPOS: Dynamic analysis of nucleosome position and occupancy by sequencing. *Genome research* **23,** 341–351 (2013).

40. Cheng, Z., Otto, G. M., Powers, E. N., Keskin, A., Mertins, P., Carr, S. A., Jovanovic, M. & Brar, G. A. Pervasive, Coordinated Protein-Level Changes Driven by Transcript Isoform Switching during Meiosis. *Cell* **172,** 910–923.e16 (2018).

41. Chew, G. L., Pauli, A. & Schier, A. F. Conservation of uORF repressiveness and sequence features in mouse, human and zebrafish. *Nature Communications* **7,** 1–10 (2016).

42. Chia, M., Tresenrider, A., Chen, J., Spedale, G., Jorgensen, V., Ünal, E. & van Werven, F. J. Transcription of a 5' extended mRNA isoform directs dynamic chromatin changes and interference of a downstream promoter. *eLife* **6,** 1–23 (2017).

43. Chia, M. & van Werven, F. J. Temporal Expression of a Master Regulator Drives Synchronous Sporulation in Budding Yeast. *G3; Genes|Genomes|Genetics Genes|Genomes|Genetics* **6,** 3553–3560 (2016).

44. Chu, S. & Herskowitz, I. Gametogenesis in yeast is regulated by a transcriptional cascade dependent on Ndt80. *Molecular cell* **1,** 685–696 (1998).

45. Cirillo, L. A., McPherson, C. E., Bossard, P., Stevens, K., Cherian, S., Shim, E. Y., Clark, K. L., Burley, S. K. & Zaret, K. S. Binding of the winged-helix transcription factor HNF3 to a linker histone site on the nucleosome. *EMBO Journal* **17,** 244–254 (1998).

46. Cirillo, L. A., Lin, F. R., Cuesta, I., Friedman, D., Jarnik, M. & Zaret, K. S. Opening of compacted chromatin by early developmental transcription factors HNF3 (FoxA) and GATA-4. *Molecular cell* (2002).

47. Compeau, P. E. C., Pevzner, P. A., Tesler, G., Papoutsoglou, G., Roscito, J. G., Dahl, A., Myers, G., Winkler, S., Pippel, M., Sameith, K., Hiller, M., Francoijs, K.-J., Gurevich, A., Saveliev, V., Vyahhi, N., Tesler, G., Nagarajan, N., Pop, M., Abbas, M. M., Malluhi, Q. M., Balakrishnan, P., Lantz, H., Notredame, C., Soler, L., Hjerde, E., Klopp, C., Dominguez Del Angel, V., Bocs, S., Binzer-Panchal, M., Gibrat, J.-F., Leskosek, B. L., Sterck, L., Vinnere Pettersson, O., Bouri, L., Capella-Gutierrez, S., Amselem, J., Vlasova, A., Buermans, H. P., den Dunnen, J. T., Davis, J. M., Bolton, R. E., Garrett, J., Ekblom, R., Wolf, J. B. & Martin, M. Cutadapt removes adapter sequences from high-throughput sequencing reads kenkyuhi hojokin gan rinsho kenkyu jigyo. *EMBnet.journal* **17,** pp. 10–12 (2013).

48. Corbin, V. & Maniatis, T. Role of transcriptional interference in the Drosophila melanogaster Adh promoter switch. *Nature* **337,** 279–282 (1989).

49. Cutter, A. R. & Hayes, J. J. A brief review of nucleosome structure. *FEBS Letters* **589,** 2914–2922 (2015).

50. Dehé, P.-M. & Géli, V. The multiple faces of Set1. *Biochemistry and cell biology = Biochimie et biologie cellulaire* **84,** 536–48 (2006).

51. Dobin, A., Davis, C. A., Schlesinger, F., Drenkow, J., Zaleski, C., Jha, S., Batut, P., Chaisson, M. & Gingeras, T. R. STAR: Ultrafast universal RNA-seq aligner. *Bioinformatics* **29,** 15–21 (2013).

52. Dunbar, T. L., Yan, Z., Balla, K. M., Smelkinson, M. G. & Troemel, E. R. C. elegans detects pathogen-induced translational inhibition to activate immune signaling. *Cell Host and Microbe* **11,** 375–386 (2012).

53. Dunn, J. G., Foo, C. K., Belletier, N. G., Gavis, E. R. & Weissman, J. S. Ribosome profiling reveals pervasive and regulated stop codon readthrough in Drosophila melanogaster. *eLife* **2,** 1–32 (2013).

54. Eissenberg, J. C. & Shilatifard, A. Histone H3 lysine 4 (H3K4) methylation in development and differentiation. *Developmental Biology* **339,** 240–249 (2010).

55. Errington, J. Regulation of endospore formation in Bacillus subtilis. *Nature Reviews Microbiology* **1,** 117–126 (2003).

56. Field, Y., Kaplan, N., Fondufe-Mittendorf, Y., Moore, I. K., Sharon, E., Lubling, Y., Widom, J. & Segal, E. Distinct modes of regulation by chromatin encoded through nucleosome positioning signals. *PLoS Computational Biology* **4.** doi:`10.1371/journal.pcbi.1000216` (2008).

57. Floor, S. N. & Doudna, J. A. Tunable protein synthesis by transcript isoforms in human cells. *eLife* **5,** 1–25 (2016).

58. Garalde, D. R., Snell, E. A., Jachimowicz, D., Sipos, B., Lloyd, J. H., Bruce, M., Pantic, N., Admassu, T., James, P., Warland, A., Jordan, M., Ciccone, J., Serra, S., Keenan, J., Martin, S., McNeill, L., Wallace, E. J., Jayasinghe, L., Wright, C., Blasco, J., Young, S., Brocklebank, D., Juul, S., Clarke, J., Heron, A. J. & Turner, D. J. Highly parallel direct RN A sequencing on an array of nanopores. *Nature Methods* **15,** 201–206 (2018).

59. Garraway, L. A. & Lander, E. S. Lessons from the cancer genome. *Cell* **153,** 17–37 (2013).

60. Gibson, D. G., Young, L., Chuang, R.-Y., Venter, J. C., Hutchison, C. A. & Smith, H. O. Enzymatic assembly of DNA molecules up to several hundred kilobases. *Nature methods* **6,** 343–5 (2009).

61. Gkikopoulos, T., Gkikopoulos, T., Schofield, P., Schofield, P., Singh, V., Singh, V., Pinskaya, M., Pinskaya, M., Mellor, J., Mellor, J., Smolle, M., Smolle, M., Workman, J. L., Workman, J. L., Barton, G. J., Barton, G. J., Owen-Hughes, T. & Owen-Hughes, T. A role for Snf2-related nucleosome-spacing enzymes in genome-wide nucleosome organization. *Science* **333,** 1758–1760 (2011).

62. Goujon, M., McWilliam, H., Li, W., Valentin, F., Squizzato, S., Paern, J. & Lopez, R. A new bioinformatics analysis tools framework at EMBL-EBI. *Nucleic Acids Research* **38,** 695–699 (2010).

63. Govind, C. K., Qiu, H., Ginsburg, D. S., Ruan, C., Hofmeyer, K., Hu, C., Swaminathan, V., Workman, J. L., Li, B. & Hinnebusch, A. G. Phosphorylated Pol II CTD recruits multiple HDACs, including Rpd3C(S), for methylation-dependent deacetylation of ORF nucleosomes. *Molecular cell* **39,** 234–46 (2010).

64. Guttman, M., Amit, I., Garber, M., French, C., Lin, M. F., Feldser, D., Huarte, M., Zuk, O., Carey, B. W., Cassady, J. P., Cabili, M. N., Jaenisch, R., Mikkelsen, T. S., Jacks, T., Hacohen, N., Bernstein, B. E., Kellis, M., Regev, A., Rinn, J. L. & Lander, E. S. Chromatin signature reveals over a thousand highly conserved large non-coding RNAs in mammals. *Nature* **458,** 223–227 (2009).

65. Haberle, V., Forrest, A. R., Hayashizaki, Y., Carninci, P. & Lenhard, B. CAGEr: Precise TSS data retrieval and high-resolution promoterome mining for integrative analyses. *Nucleic Acids Research* **43.** doi:`10.1093/nar/gkv054` (2015).

66. Hainer, S. J. & Martens, J. A. Transcription of ncDNA: Many roads lead to local gene regulation. *Transcription* **2,** 120–123 (2011).

67. Hampsey, M. & Reinberg, D. Tails of intrigue: Phosphorylation of RNA polymerase II mediates histone methylation. *Cell* **113,** 429–432 (2003).

68. Hansen, J. C., Connolly, M., McDonald, C. J., Pan, A., Pryamkova, A., Ray, K., Seidel, E., Tamura, S., Rogge, R. & Maeshima, K. The 10-nm chromatin fiber and its relationship to interphase chromosome organization. *Biochemical Society Transactions* **46,** 67–76 (2018).

69. Hao, N., Palmer, A. C., Ahlgren-Berg, A., Shearwin, K. E. & Dodd, I. B. The role of repressor kinetics in relief of transcriptional interference between convergent promoters. *Nucleic Acids Research* **44,** 6625–6638 (2016).

70. Harding, H. P., Novoa, I., Zhang, Y., Zeng, H., Wek, R., Schapira, M. & Ron, D. Regulated Translation Initiation Controls Stress-Induced Gene Expression in Mammalian Cells. *Molecular Cell* **6,** 1099–1108 (2000).

71. Hassold, T. & Hunt, P. To err (meiotically) is human: the genesis of human aneuploidy. *Nature reviews. Genetics* **2,** 280–291 (2001).

72. Henikoff, S. Heterochromatin function in complex genomes. *Biochimica et biophysica acta* **1470,** 1–8 (2000).

73. Herbert, M., Kalleas, D., Cooney, D., Lamb, M. & Lister, L. Meiosis and maternal aging: Insights from aneuploid oocytes and trisomy births. *Cold Spring Harbor Perspectives in Biology* **7.** doi:`10.1101/cshperspect.a017970` (2015).

74. Hinnebusch, A. G. Translational Regulation of Yeast GCN4. *Journal of Biological Chemistry* **272,** 21661–21664 (1997).

75. Hinnebusch, A. G., Ivanov, I. P. & Sonenberg, N. Translational control by 5-untranslated regions of eukaryotic mRNAs. *Science* **352,** 1413–1416 (2016).

76. Hirschhorn, J. N., Brown, S. A., Clark, C. D. & Winston, F. Evidence that SNF2/SWI2 and SNF5 activate transcription in yeast by altering chromatin structure. *Genes and Development* **6,** 2288–2298 (1992).

77. Hoffmann, S. A., Kruse, S. M. & Arndt, K. M. Long-range transcriptional interference in E. coli used to construct a dual positive selection system for genetic switches. *Nucleic Acids Research* **44.** doi:`10.1093/nar/gkw125` (2016).

78. Hollerer, I., Barker, J. C., Jorgensen, V., Tresenrider, A., Dugast-Darzacq, C., Chan, L. Y., Darzacq, X., Tjian, R., Ünal, E. & Brar, G. A. Evidence for an Integrated Gene Repression Mechanism Based on mRNA Isoform Toggling in Human Cells. *G3: Genes|Genomes|Genetics,* g3.200802.2018 (2019).

79. Hongay, C. F., Grisafi, P. L., Galitski, T. & Fink, G. R. Antisense transcription controls cell fate in Saccharomyces cerevisiae. *Cell* **127,** 735–45 (2006).

80. Houseley, J., Rubbi, L., Grunstein, M., Tollervey, D. & Vogelauer, M. A ncRNA Modulates Histone Modification and mRNA Induction in the Yeast GAL Gene Cluster. *Molecular Cell,* 685–695 (2008).

81. Hunter, J. D. Matplotlib: A 2D graphics environment. *Computing in Science and Engineering* **9,** 99–104 (2007).

82. Ingolia, N. T., Brar, G. A., Stern-Ginossar, N., Harris, M. S., Talhouarne, G. J., Jackson, S. E., Wills, M. R. & Weissman, J. S. Ribosome Profiling Reveals Pervasive Translation Outside of Annotated Protein-Coding Genes. *Cell Reports,* 1365–1379 (2014).

83. Ingolia, N. T., Ghaemmaghami, S., Newman, J. R. S. & Weissman, J. Genome-Wide Analysis in Vivo of Translation with NUcleotide Resolution Usinf Ribosome Profiling. *Science* **1168978,** 218–324 (2009).

84. Ingolia, N. T., Lareau, L. F. & Weissman, J. S. Ribosome profiling of mouse embryonic stem cells reveals the complexity and dynamics of mammalian proteomes. *Cell* **147,** 789–802 (2011).

85. J Rando, O. Genome-Wide Measurement of Histone H3 Replacement Dynamics in Yeast. *Methods in molecular biology (Clifton, N.J.)* **759,** 41–60 (2011).

86. Jeziorska, D. M., Murray, R. J. S., De Gobbi, M., Gaentzsch, R., Garrick, D., Ayyub, H., Chen, T., Li, E., Telenius, J., Lynch, M., Graham, B., Smith, A. J. H., Lund, J. N., Hughes, J. R., Higgs, D. R. & Tufarelli, C. DNA methylation of intragenic CpG islands depends on their transcriptional activity during differentiation and disease. *Proceedings of the National Academy of Sciences* **114,** E7526–E7535 (2017).

87. Jin, L. & Neiman, A. M. Post-transcriptional regulation in budding yeast meiosis. *Current Genetics* **62,** 313–315 (2016).

88. Johnstone, T. G., Bazzini, A. A. & Giraldez, A. J. Upstream ORFs are prevalent translational repressors in vertebrates. *The EMBO Journal* **35,** 706–723 (2016).

89. Kadoch, C., Hargreaves, D. C., Hodges, C., Elias, L., Ho, L., Ranish, J. & Crabtree, G. R. Proteomic and bioinformatic analysis of mammalian SWI/SNF complexes identifies extensive roles in human malignancy. *Nature Genetics* **45,** 592–601 (2013).

90. Kadosh, D. & Struhl, K. Repression by Ume6 involves recruitment of a complex containing Sin3 corepressor and Rpd3 histone deacetylase to target promoters. *Cell* **89,** 365–371 (1997).

91. Kaplan, N., Moore, I. K., Fondufe-Mittendorf, Y., Gossett, A. J., Tillo, D., Field, Y., LeProust, E. M., Hughes, T. R., Lieb, J. D., Widom, J. & Segal, E. The DNA-encoded nucleosome organization of a eukaryotic genome. *Nature* **458,** 362–366 (2009).

92. Kassir, Y., Granot, D. & Simchen, G. IME1, a positive regulator gene of meiosis in S. cerevisiae. *Cell* **52,** 853–862 (1988).

93. Kebede, A. F., Schneider, R. & Daujat, S. Novel types and sites of histone modifications emerge as players in the transcriptional regulation contest. *FEBS Journal* **282,** 1658–1674 (2015).

94. Kent, W. J., Zweig, A. S., Barber, G., Hinrichs, A. S. & Karolchik, D. BigWig and BigBed: Enabling browsing of large distributed datasets. *Bioinformatics* **26,** 2204–2207 (2010).

95. Keogh, M. C., Kurdistani, S. K., Morris, S. a., Ahn, S. H., Podolny, V., Collins, S. R., Schuldiner, M., Chin, K., Punna, T., Thompson, N. J., Boone, C., Emili, A., Weissman, J. S., Hughes, T. R., Strahl, B. D., Grunstein, M., Greenblatt, J. F., Buratowski, S. & Krogan, N. J. Cotranscriptional set2 methylation of histone H3 lysine 36 recruits a repressive Rpd3 complex. *Cell* **123,** 593–605 (2005).

96. Kim, J. H., Lee, B. B., Oh, Y. M., Zhu, C., Steinmetz, L. M., Lee, Y., Kim, W. K., Lee, S. B., Buratowski, S. & Kim, T. S. Modulation of mRNA and lncRNA expression dynamics by the Set2-Rpd3S pathway. *Nature communications* **8,** 16122 (2017).

97. Kim, S., Meyer, R., Chuong, H. & Dawson, D. S. Dual mechanisms prevent premature chromosome segregation during meiosis. *Genes and Development* **27,** 2139–2146 (2013).

98. Kim, T. & Buratowski, S. Dimethylation of H3K4 by Set1 Recruits the Set3 Histone Deacetylase Complex to 5' Transcribed Regions. *Cell* **137,** 259–272 (2009).

99. Kim, T., Xu, Z., Clauder-Münster, S., Steinmetz, L. M. & Buratowski, S. Set3 HDAC mediates effects of overlapping noncoding transcription on gene induction kinetics. *Cell* **150,** 1158–69 (2012).

100. Kimura, K., Wakamatsu, A., Suzuki, Y., Ota, T., Nishikawa, T., Yamashita, R., Yamamoto, J.-i., Sekine, M., Tsuritani, K., Wakaguri, H., Ishii, S., Sugiyama, T., Saito, K., Isono, Y., Irie, R., Kushida, N., Yoneyama, T., Otsuka, R., Kanda, K., Yokoi, T., Kondo, H., Wagatsuma, M., Murakawa, K., Ishida, S., Ishibashi, T., Takahashi-fujii, A., Tanase, T., Nagai, K., Kikuchi, H., Nakai, K., Isogai, T. & Sugano, S. Diversification of transcriptional modulation: Large-scale identification and characterization of putative alternative promoters of human genes. *Genome research* **16,** 55–65 (2006).

101. Kirmizis, A., Santos-Rosa, H., Penkett, C. J., Singer, M. A., Vermeulen, M., Mann, M., Bähler, J., Green, R. D. & Kouzarides, T. Arginine methylation at histone H3R2 controls deposition of H3K4 trimethylation. *Nature* **449,** 928–932 (2007).

102. Kizer, K. O., Phatnani, H. P., Shibata, Y., Hall, H., Greenleaf, A. L. & Strahl, B. D. A novel domain in Set2 mediates RNA polymerase II interaction and couples histone H3 K36 methylation with transcript elongation. *Molecular and cellular biology* **25,** 3305–3316 (2005).

103. Klemm, S. L., Shipony, Z. & Greenleaf, W. J. Chromatin accessibility and the regulatory epigenome. *Nature reviews. Genetics* **20,** 207–220 (2019).

104. Kojima, M. L., de Rooij, D. G. & Page, D. C. Amplification of a broad transcriptional program by a common factor triggers the meiotic cell cycle in mice. *eLife* **8,** 1–32 (2019).

105. Koster, M. J., Yildirim, A. D., Weil, P. A., Holstege, F. C. & Timmers, H. T. M. Suppression of intragenic transcription requires the MOT1 and NC2 regulators of TATA-binding protein. *Nucleic Acids Research* **42,** 4220–4229 (2014).

106. Kouzarides, T. Chromatin modifications and their function. *Cell* **128,** 693–705 (2007).

107. Krogan, N. J., Dover, J., Wood, A., Schneider, J., Heidt, J., Boateng, M. A., Dean, K., Ryan, O. W., Golshani, A., Johnston, M., Greenblatt, J. F. & Shilatifard, A. The Paf1 complex is required for histone H3 methylation by COMPASS and Dot1p: Linking transcriptional elongation to histone methylation. *Molecular Cell* **11,** 721–729 (2003).

108. Krumlauf, R. Hox genes in vertebrate development. *Cell* **78,** 191–201 (1994).

109. Lachner, M., O'Carroll, D., Rea, S., Mechtler, K. & Jenuwein, T. Methylation of histone H3 lysine 9 creates a binding site for HP1 proteins. *Nature* **410,** 116–120 (2002).

110. Langmead, B. & Salzberg, S. L. Fast gapped-read alignment with Bowtie 2. *Nature methods* **9,** 357–9 (2012).

111. Langmead, B., Wilks, C., Antonescu, V. & Charles, R. Scaling read aligners to hundreds of threads on general-purpose processors. *Bioinformatics* **35,** 421–432 (2019).

112. Lardenois, A., Becker, E., Walther, T., Law, M. J., Xie, B., Demougin, P., Strich, R. & Primig, M. Global alterations of the transcriptional landscape during yeast growth and development in the absence of Ume6-dependent chromatin modification. *Molecular Genetics and Genomics* **290,** 2031–2046 (2015).

113. Laurent, B. C., Yang, X. & Carlson, M. An essential Saccharomyces cerevisiae gene homologous to SNF2 encodes a helicase-related protein in a new family. *Molecular and Cellular Biology* **12,** 1893–1902 (1992).

114. Laurent, B. C., Treich, I. & Carlson, M. The yeast SNF2/SWI2 protein has DNA-stimulated ATPase activity required for transcriptional activation. *Genes and Development* **7,** 583–591 (1993).

115. Lee, C. K., Shibata, Y., Rao, B., Strahl, B. D. & Lieb, J. D. Evidence for nucleosome depletion at active regulatory regions genome-wide. *Nature Genetics* **36,** 900–905 (2004).

116. Lee, R. C., Feinbaum, R. L. & Ambros, V. The C. elegans Heterochronic Gene lin-4 Encodes Small RNAs with Antisense Complementarity to lin-14. *Cell* **75: 84385,** 843–854 (1993).

117. Lee, W., Tillo, D., Bray, N., Morse, R. H., Davis, R. W., Hughes, T. R. & Nislow, C. A high-resolution atlas of nucleosome occupancy in yeast. *Nature Genetics* **39,** 1235–1244 (2007).

118. Li, B., Gogol, M., Carey, M., Lee, D., Seidel, C. & Workman, J. L. Combined action of PHD and chromo domains directs the Rpd3S HDAC to transcribed chromatin. *Science (New York, N.Y.)* **316,** 1050–1054 (2007).

119. Li, B., Howe, L. A., Anderson, S., Yates, J. R. & Workman, J. L. The Set2 histone methyltransferase functions through the phosphorylated carboxyl-terminal domain of RNA polymerase II. *Journal of Biological Chemistry* **278,** 8897–8903 (2003).

120. Li, H. Minimap2: Pairwise alignment for nucleotide sequences. *Bioinformatics* **34,** 3094–3100 (2018).

121. Liu, C. L., Kaplan, T., Kim, M., Buratowski, S., Schreiber, S. L., Friedman, N. & Rando, O. J. Single-nucleosome mapping of histone modifications in S. cerevisiae. *PLoS Biology* **3.** doi:`10.1371/journal.pbio.0030328` (2005).

122. Liu, Y., Stuparevic, I., Xie, B., Becker, E., Law, M. J. & Primig, M. The conserved histone deacetylase Rpd3 and the DNA binding regulator Ume6 repress BOI1's meiotic transcript isoform during vegetative growth in Saccharomyces cerevisiae. *Molecular Microbiology* **96,** 861–874 (2015).

123. Longtine, M. S., McKenzie, a., Demarini, D. J., Shah, N. G., Wach, a., Brachat, a., Philippsen, P. & Pringle, J. R. Additional modules for versatile and economical PCR-based gene deletion and modification in Saccharomyces cerevisiae. *Yeast (Chichester, England)* **14,** 953–61 (1998).

124. Love, M. I., Huber, W. & Anders, S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome biology* **15,** 550 (2014).

125. Maier, T., Güell, M. & Serrano, L. Correlation of mRNA and protein in complex biological samples. *FEBS letters* **583,** 3966–73 (2009).

126. Malabat, C., Feuerbach, F., Ma, L., Saveanu, C. & Jacquier, A. Quality control of transcription start site selection by Nonsense-Mediated-mRNA Decay. *eLIFE* (2015).

127. Marchese, F. P., Raimondi, I. & Huarte, M. The multidimensional mechanisms of long noncoding RNA function. *Genome Biology* **18,** 1–13 (2017).

128. Martens, J. a., Laprade, L. & Winston, F. Intergenic transcription is required to repress the Saccharomyces cerevisiae SER3 gene. *Nature* **429,** 571–4 (2004).

129. McCord, R., Pierce, M., Xie, J., Wonkatal, S., Mickel, C. & Vershon, A. K. Rfm1, a Novel Tethering Factor Required To Recruit the Hst1 Histone Deacetylase for Repression of Middle Sporulation Genes. *Molecular and Cellular Biology* **23,** 2009–2016 (2003).

130. McDaniel, S. L., Hepperla, A. J., Huang, J., Dronamraju, R., Adams, A. T., Kulkarni, V. G., Davis, I. J. & Strahl, B. D. H3K36 Methylation Regulates Nutrient Stress Response in Saccharomyces cerevisiae by Enforcing Transcriptional Fidelity. *Cell Reports* **19,** 2371–2382 (2017).

131. Mercer, T. R., Dinger, M. E. & Mattick, J. S. Long non-coding RNAs: insights into functions. *Nature Reviews Genetics* **10.** doi:`10.1038/nrg2521` (2009).

132. Meyer, R. E., Chuong, H. H., Hild, M., Hansen, C. L., Kinter, M. & Dawson, D. S. Ipl1/Aurora-B is necessary for kinetochore restructuring in meiosis I in Saccharomyces cerevisiae. *Molecular Biology of the Cell* **26,** 2986–3000 (2015).

133. Michaelis, C., Ciosk, R. & Nasmyth, K. Cohesins: Chromosomal proteins that prevent premature separation of sister chromatids. *Cell* **91,** 35–45 (1997).

134. Miller, M. P., Unal, E., Brar, G. a. & Amon, A. Meiosis I chromosome segregation is established through regulation of microtubule-kinetochore interactions. *eLife* **1,** e00117 (2012).

135. Mizzen, C. A., Yang, X. J., Kokubo, T., Brownell, J. E., Bannister, A. J., Owen-Hughes, T., Workman, J., Wang, L., Berger, S. L., Kouzarides, T., Nakatani, Y. & Allis, C. D. The TAF(II)250 subunit of TFIID has histone acetyltransferase activity. *Cell* **87,** 1261–1270 (1996).

136. Moretto, F., Wood, N. E., Kelly, G., Doncic, A. & Van Werven, F. J. A regulatory circuit of two lncRNAs and a master regulator directs cell fate in yeast. *Nature Communications* **9.** doi:10.1038/s41467-018-03213-z. http://dx.doi.org/10.1038/s41467-018-03213-z (2018).

137. Morselli, M., Pastor, W. A., Montanini, B., Nee, K., Ferrari, R., Fu, K., Bonora, G., Rubbi, L., Clark, A. T., Ottonello, S., Jacobsen, S. E. & Pellegrini, M. In vivo targeting of de novo DNA methylation by histone modifications in yeast and mouse. *eLife* **4,** 1–21 (2015).

138. Mueller, P. P. & Hinnebusch, A. G. Multiple upstream AUG codons mediate translational control of GCN4. *Cell* **45,** 201–207 (1986).

139. Musacchio, A. & Desai, A. A Molecular View of Kinetochore Assembly and Function. *Biology* **6,** 5 (2017).

140. Nagaoka, S. I., Hassold, T. J. & Hunt, P. A. Human aneuploidy: Mechanisms and new insights into an age-old problem. *Nature Reviews Genetics* **13,** 493–504 (2012).

141. Nakayama, J.-i., Rice, J. C., Strahl, B. D., Allis, C. D. & Grewal, S. I. S. Role of histone H3 lysine 9 methylation in epigenetic control of heterochromatin assembly. *Science* **292,** 110–113 (2001).

142. Narlikar, G. J., Sundaramoorthy, R. & Owen-Hughes, T. Mechanisms and functions of ATP-dependent chromatin-remodeling enzymes. *Cell* **154,** 490–503 (2013).

143. Natsume, T., Kiyomitsu, T., Saga, Y. & Kanemaki, M. T. Rapid Protein Depletion in Human Cells by Auxin-Inducible Degron Tagging with Short Homology Donors. *Cell Reports* **15,** 210–218 (2016).

144. Neigeborn, L. & Carlson, M. Genes Affecting The Regulation of SUC2 Gene Expression by Glucose Repression in Saccharomyces Cerevisiae. *Genetics* **108,** 845–858 (1984).

145. Nelson, H. C., Finch, J. T., Luisi, B. F. & Klug, A. The structure of an oligo(dA):oligo(dT) tract and its biological implications. *Nature* **330,** 221–226 (1987).

146. Neri, F., Rapelli, S., Krepelova, A., Incarnato, D., Parlato, C., Basile, G., Maldotti, M., Anselmi, F. & Oliviero, S. Intragenic DNA methylation prevents spurious transcription initiation. *Nature* **543,** 72–77 (2017).

147. Ng, H. H., Robert, F., Young, R. a. & Struhl, K. Targeted recruitment of Set1 histone methylase by elongating Pol II provides a localized mark and memory of recent transcriptional activity. *Molecular Cell* **11,** 709–719 (2003).

148. Nishimura, K., Fukagawa, T., Takisawa, H., Kakimoto, T. & Kanemaki, M. An auxin-based degron system for the rapid depletion of proteins in nonplant cells. *Nature Methods* **6,** 917–922 (2009).

149. Nislow, C., Ray, E. & Pillus, L. SET1, A Yeast Member of the Trithorax Family, Functions in Transcriptional Silencing and Diverse Cellular Processes. *Molecular Biology of the Cell* **8,** 2421–2436 (1997).

150. Ogryzko, V. V., Schiltz, R. L., Russanova, V., Howard, B. H. & Nakatani, Y. The transcriptional coactivators p300 and CBP are histone acetyltransferases. *Cell* **87,** 953–959 (1996).

151. Oikonomopoulos, S., Wang, Y. C., Djambazian, H., Badescu, D. & Ragoussis, J. Benchmarking of the Oxford Nanopore MinION sequencing for quantitative and qualitative assessment of cDNA populations. *Scientific reports* **6,** 31602 (2016).

152. Ottoz, D. S., Rudolf, F. & Stelling, J. Inducible, tightly regulated and growth condition-independent transcription factor in Saccharomyces cerevisiae. *Nucleic Acids Research* **42.** doi:`10.1093/nar/gku616` (2014).

153. Pagè, H. *BSgenome: Software infrastructure for efficient representation of full genomes and their SNPs* 2018.

154. Pajerowska-Mukhtar, K. M., Wang, W., Tada, Y., Oka, N., Tucker, C. L., Fonseca, J. P. & Dong, X. The HSF-like transcription factor TBF1 is a major molecular switch for plant growth-to-defense transition. *Current Biology* **22,** 103–112 (2012).

155. Park, H. D., Luche, R. M. & Cooper, T. G. The yeast UME6 gene product is required for transcriptional repression mediated by the CAR1 URS1 repressor binding site. *Nucleic Acids Research* **20,** 1909–1915 (1992).

156. Patro, R., Duggal, G., Love, M. I., Irizarry, R. A. & Kingsford, C. Salmon provides fast and bias-aware quantification of transcript expression. *Nature Methods* **14,** 417–419 (2017).

157. Perino, M. & Veenstra, G. J. C. Chromatin Control of Developmental Dynamics and Plasticity. *Developmental Cell* **38,** 610–620 (2016).

158. Pierce, M., Benjamin, K. R., Montano, S. P., Georgiadis, M. M., Winter, E. & Vershon, A. K. Sum1 and Ndt80 Proteins Compete for Binding to Middle Sporulation Element Sequences That Control Meiotic Gene Expression. *Molecular and Cellular Biology* **23,** 4814–4825 (2003).

159. Pierce, M., Wagner, M., Xie, J., Gailus-Durner, V., Six, J., Vershon, A. K. & Winter, E. Transcriptional Regulation of the SMK1 Mitogen-Activated Protein Kinase Gene during Meiotic Development in Saccharomyces cerevisiae. *Molecular and Cellular Biology* **18,** 5970–5980 (1998).

160. Pijnappel, W. W. M. P., Schaft, D., Roguev, A., Shevchenko, A., Tekotte, H., Wilm, M., Rigaut, G., Séraphin, B., Aasland, R. & Stewart, A. F. The S . cerevisiae SET3 complex includes two histone deacetylases , Hos2 and Hst1 , and is a meiotic-specific repressor of the sporulation gene program. *Genes and Development* **15,** 2991–3004 (2001).

161. Pokholok, D. K., Harbison, C. T., Levine, S., Cole, M., Hannett, N. M., Tong, I. L., Bell, G. W., Walker, K., Rolfe, P. A., Herbolsheimer, E., Zeitlinger, J., Lewitter, F., Gifford, D. K. & Young, R. A. Genome-wide map of nucleosome acetylation and methylation in yeast. *Cell* **122,** 517–527 (2005).

162. Pop, C., Rouskin, S., Ingolia, N. T., Han, L., Phizicky, E. M., Weissman, J. S. & Koller, D. Causal signals between codon bias, mRNA structure, and the efficiency of translation and elongation. *Molecular Systems Biology* **10,** 770–770 (2014).

163. Prieto, E. I. & Maeshima, K. Dynamic chromatin organization in the cell. *Essays In Biochemistry* **63,** 133–145 (2019).

164. Quinlan, A. R. & Hall, I. M. BEDTools: A flexible suite of utilities for comparing genomic features. *Bioinformatics* **26,** 841–842 (2010).

165. Ramírez, F., Ryan, D. P., Grüning, B., Bhardwaj, V., Kilpert, F., Richter, A. S., Heyne, S., Dündar, F. & Manke, T. deepTools2: a next generation web server for deep-sequencing data analysis. *Nucleic acids research* **44,** W160–W165 (2016).

166. Rayburn, E., Zhang, R., He, J. & Wang, H. MDM2 and Human Malignancies: Expression, Clinical Pathology, Prognostic Markers, and Implications for Chemotherapy. *Current Cancer Drug Targets* **5,** 27–41 (2005).

167. Rea, S., Eisenhaber, F., O'Carroll, D., Strahl, B. D., Sun, Z.-W., Schmid, M., Opravil, S., Mechtler, K., Ponting, C. P., Allis, C. D. & Jenuwein, T. Regulation of chromatin structure by site-specific histone H3 methyltransferases. *Nature* **406,** 593–599 (2000).

168. Rodriguez, J., McKnight, J. N. & Tsukiyama, T. Genome-wide analysis of nucleosome positions, occupancy, and accessibility in yeast: Nucleosome mapping, high-resolution histone chIP, and NCAM. *Current Protocols in Molecular Biology* **2014,** 21.28.1–21.28.16 (2014).

169. Roguev, A., Schaft, D., Shevchenko, A., Pijnappel, W. W. M. P., Wilm, M., Aasland, R. & Stewart, A. F. The Saccharomyces cerevisiae Set1 complex includes an Ash2 homologue and methylates histone 3 lysine 4. *EMBO Journal* **20,** 7137–7148 (2001).

170. Rojas-Duran, M. F. & Gilbert, W. V. Alternative transcription start site selection leads to large differences in translation activity in yeast. *RNA* **18,** 2299–2305 (2012).

171. Rubin-Bejerano, I., Mandel, S., Robzyk, K. & Kassir, Y. Induction of meiosis in Saccharomyces cerevisiae depends on conversion of the transcriptional represssor Ume6 to a positive regulator by its regulated association with the transcriptional activator Ime1. *Molecular and cellular biology* **16,** 2518–2526 (1996).

172. Rundlett, S. E., Carmen, A. A., Kobayashi, R., Bavykin, S., Turner, B. M. & Grunstein, M. HDA1 and RPD3 are members of distinct yeast histone deacetylase complexes that regulate silencing and transcription (chromatinacetylationtelomeresPHO5CUP1). *Biochemistry* **93,** 14503–14508 (1996).

173. Scannell, D. R., Zill, O. A., Rokas, A., Payen, C., Dunham, M. J., Eisen, M. B., Rine, J., Johnston, M. & Hittinger, C. T. The Awesome Power of Yeast Evolutionary Genetics: New Genome Sequences and Strain Resources for the Saccharomyces sensu stricto Genus. *G3; Genes|Genomes|Genetics Genes|Genomes|Genetics* **1,** 11–25 (2011).

174. Scherthan, H., Wang, H., Adelfalk, C., White, E. J., Cowan, C., Cande, W. Z. & Kaback, D. B. Chromosome mobility during meiotic prophase in Saccharomyces cerevisiae. *Proceedings of the National Academy of Sciences* **104,** 16934–16939 (2007).

175. Schindelin, J., Arganda-Carreras, I., Frise, E., Kaynig, V., Longair, M., Pietzsch, T., Preibisch, S., Rueden, C., Saalfeld, S., Schmid, B., Tinevez, J.-Y., White, D. J., Hartenstein, V., Eliceiri, K., Tomancak, P. & Cardona, A. Fiji: an open-source platform for biological-image analysis. *Nature methods* **9,** 676–82 (2012).

176. Sharon, D., Tilgner, H., Grubert, F. & Snyder, M. A single-molecule long-read survey of the human transcriptome. *Nature Biotechnology* **31,** 1009–1014 (2013).

177. Shearwin, K. E., Callen, B. P. & Egan, J. B. Transcriptional interference a crash course. *Trends in Genetics* **21,** 339–345 (2005).

178. Shi, X., Hong, T., Walter, K. L., Ewalt, M., Michishita, E., Hung, T., Carney, D., Peña, P., Lan, F., Kaadige, M. R., Lacoste, N., Cayrou, C., Davrazou, F., Saha, A., Cairns, B. R., Ayer, D. E., Kutateladze, T. G., Shi, Y., Côté, J., Chua, K. F. & Gozani, O. ING2 PHD domain links histone H3 lysine 4 methylation to active gene repression. *Nature* **442,** 96–99 (2006).

179. Siepel, A., Bejerano, G., Pedersen, J. S., Hinrichs, A. S., Hou, M., Rosenbloom, K., Clawson, H., Spieth, J., Hillier, L. W., Richards, S., Weinstock, G. M., Wilson, R. K., Gibbs, R. A., Kent, W. J., Miller, W. & Haussler, D. Evolutionarily conserved elements in vertebrate, insect, worm, and yeast genomes. *Genome research* **15,** 1034–50 (2005).

180. Sievers, F., Wilm, A., Dineen, D., Gibson, T. J., Karplus, K., Li, W., Lopez, R., McWilliam, H., Remmert, M., Söding, J., Thompson, J. D. & Higgins, D. G. Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. *Molecular Systems Biology* **7.** doi:`10.1038/msb.2011.75` (2011).

181. Smith, H. E., Driscoll, S. E., Sia, R. a. L., Yuan, H. E. & Mitchell, a. P. Genetic evidence for transcriptional activation by the yeast IME1 gene product. *Genetics* **133,** 775–784 (1993).

182. Smolle, M., Venkatesh, S., Gogol, M. M., Li, H., Zhang, Y., Florens, L., Washburn, M. P. & Workman, J. L. Chromatin remodelers Isw1 and Chd1 maintain chromatin structure during transcription by preventing histone exchange. *Nature Structural and Molecular Biology* **19,** 884–892 (2012).

183. Solé, C., Nadal-Ribelles, M., de Nadal, E. & Posas, F. A novel role for lncRNAs in cell cycle control during stress adaptation. *Current Genetics* **61,** 299–308 (2015).

184. Sommermeyer, V., Béneut, C., Chaplais, E., Serrentino, M. E. & Borde, V. Spp1, a member of the Set1 Complex, promotes meiotic DSB formation in promoters by tethering histone H3K4 methylation sites to chromosome axes. *Molecular cell* **49,** 43–54 (2013).

185. Soufi, A., Garcia, M. F., Jaroszewicz, A., Osman, N., Pellegrini, M. & Zaret, K. S. Pioneer transcription factors target partial DNA motifs on nucleosomes to initiate reprogramming. *Cell* **161,** 555–568 (2015).

186. Spitzer, M., Wildenhain, J., Rappsilber, J. & Tyers, M. BoxPlotR: a web tool for generation of box plots. *Nature Methods* **11,** 121–122 (2014).

187. Stern, D. L. & Orgogozo, V. The loci of evolution: How predictable is genetic evolution? *Evolution* **62,** 2155–2177 (2008).

188. Strahl-Bolsinger, S., Hecht, A., Luo, K. & Grunstein, M. SIR2 and SIR4 interactions differ in core and extended telomeric heterochromatin in yeast. *Genes and Development* **11,** 83–93 (1997).

189. Strahl, B. D. & Allis, C. D. The language of covalent histone modifications. *Nature* **403,** 41–5 (2000).

190. Strahl, B. D., Grant, P. a., Briggs, S. D., Sun, Z.-W., Bone, J. R., Caldwell, J. a., Mollah, S., Cook, R. G., Shabanowitz, J., Hunt, D. F. & Allis, C. D. Set2 is a Nucleosomal Histone H3-Selective Methyltransferase That Mediates Transcriptional Repression. *Molecular and cellular biology* **22,** 1298–1306 (2002).

191. Strich, R., Surosky, R. T., Steber, C., Dubois, E., Messenguy, F. & Esposito, R. E. UME6 is a key regulator of nitrogen repression and meiotic development. *Genes and Development* **8,** 796–810 (1994).

192. Struhl, K. Histone acetylation and transcriptional regulatory mechanisms. *Genes and Development* **12,** 599–606 (1998).

193. Struhl, K. & Segal, E. Determinants of nucleosome positioning. *Nature Structural and Molecular Biology* **20,** 267–273 (2013).

194. Sumrada, R. A. & Cooper, T. G. Ubiquitous upstream repression sequences control activation of the inducible arginase gene in yeast. *Proceedings of the National Academy of Sciences* **84,** 3997–4001 (1987).

195. Sun, S.-C., Zhang, D.-X., Lee, S.-E., Xu, Y.-N. & Kim, N.-H. Ndc80 Regulates Meiotic Spindle Organization, Chromosome Alignment, and Cell Cycle Progression in Mouse Oocytes. *Microscopy and Microanalysis* **17,** 431–439 (2011).

196. Suter, B., Schnappauf, G. & Thoma, F. Poly (dAůdT) sequences exist as rigid DNA structures in nucleosome-free yeast promoters in vivo. *Nucleic Acids Research* **28,** 4083–4089 (2000).

197. Taggart, J., MacDiarmid, C. W., Haws, S. & Eide, D. J. Zap1-dependent transcription from an alternative upstream promoter controls translation of RTC4 mRNA in zinc-deficient Saccharomyces cerevisiae. *Molecular Microbiology* **106,** 678–689 (2017).

198. Takahashi, K. & Yamanaka, S. Induction of pluripotent stem cells from mouse embryonic and adult fibroblast cultures by defined factors. *Cell.* doi:`10.1016/j.cell.2006.07.024` (2006).

199. Taunton, J., Hassig, C. A., Schreiber, S. L., Taunton, J., Hassig, C. A. & Schreiber, S. L. A mammalian histone deacetylase related to the yeast regulator Rpd3p. *Science* **272** (1996).

200. Teale, W. D., Paponov, I. A. & Palme, K. Auxin in action: Signalling, transport and the control of plant growth and development. *Nature Reviews Molecular Cell Biology* **7,** 847–859 (2006).

201. Tedesco, M., La Sala, G., Barbagallo, F., De Felici, M. & Farini, D. STRA8 shuttles between nucleus and cytoplasm and displays transcriptional activity. *Journal of Biological Chemistry* **284,** 35781–35793 (2009).

202. Tirosh, I. & Barkai, N. Two strategies for gene regulation by promoter nucleosomes. *Genome Research* **18,** 1084–1091 (2008).

203. Tooley, J. & Stukenberg, P. T. The Ndc80 complex: Integrating the kinetochore's many movements. *Chromosome Research* **19,** 377–391 (2011).

204. Tresenrider, A. & Ünal, E. One-two punch mechanism of gene repression: a fresh perspective on gene regulation. *Current Genetics* **64,** 581–588 (2018).

205. Van Dalfsen, K. M., Hodapp, S., Keskin, A., Otto, G. M., Berdan, C. A., Higdon, A., Cheunkarndee, T., Nomura, D. K., Jovanovic, M. & Brar, G. A. Global Proteome Remodeling during ER Stress Involves Hac1-Driven Expression of Long Undecoded Transcript Isoforms. *Developmental Cell* **46,** 219–235.e8 (2018).

206. Van Werven, F. J. & Amon, A. Regulation of entry into gametogenesis. *Philosophical transactions of the Royal Society of London. Series B, Biological sciences* **366,** 3521–31 (2011).

207. Van Werven, F. J., Neuert, G., Hendrick, N., Lardenois, A., Buratowski, S., van Oudenaarden, A., Primig, M. & Amon, A. Transcription of two long noncoding RNAs mediates mating-type control of gametogenesis in budding yeast. *Cell* **150,** 1170–81 (2012).

208. Vattem, K. M. & Wek, R. C. Reinitiation involving upstream ORFs regulates ATF4 mRNA translation in mammalian cells. *Proceedings of the National Academy of Sciences* **101,** 11269–11274 (2004).

209. Venkatesh, S., Li, H., Gogol, M. M. & Workman, J. L. Selective suppression of antisense transcription by Set2-mediated H3K36 methylation. *Nature Communications* **7,** 1–14 (2016).

210. Venkatesh, S., Smolle, M., Li, H., Gogol, M. M., Saint, M., Kumar, S., Natarajan, K. & Workman, J. L. Set2 methylation of histone H3 lysine 36 suppresses histone exchange on transcribed genes. *Nature* **489,** 452–455 (2012).

211. Venkatesh, S. & Workman, J. L. Histone exchange, chromatin structure and the regulation of transcription. *Nature Reviews Molecular Cell Biology* **16,** 178–189 (2015).

212. Wagner, E. J. & Carpenter, P. B. Understanding the language of Lys36 methylation at histone H3. *Nature Reviews Molecular Cell Biology* **13,** 115–126 (2012).

213. Wang, W., Cherry, J. M., Nochomovitz, Y., Jolly, E., Botstein, D. & Li, H. Inference of combinatorial regulation in yeast transcriptional networks: A case study of sporulation. *Proceedings of the National Academy of Sciences* **102,** 1998–2003 (2005).

214. Washburn, B. K. & Esposito, R. E. Identification of the Sin3-binding site in Ume6 defines a two-step process for conversion of Ume6 from a transcriptional repressor to an activator in yeast. *Molecular and cellular biology* **21,** 2057–2069 (2001).

215. Weidberg, H., Moretto, F., Spedale, G., Amon, A. & van Werven, F. J. Nutrient Control of Yeast Gametogenesis Is Mediated by TORC1, PKA and Energy Availability. *PLoS Genetics* **12,** 1–26 (2016).

216. Wigge, P. A., Jensen, O. N., Holmes, S., Souès, S., Mann, M. & Kilmartin, J. V. Analysis of the Saccharomyces Spindle POle by Matrix-assisted Laser Desorption/Ionization (MALDI) Mass Spectrometry. *The Journal of cell biology* **141,** 967–977 (1998).

217. Wightman, B., Ha, I. & Ruvkun, G. Posttranscriptional regulation of the heterochronic gene lin-14 by lin-4 mediates temporal pattern formation in C. elegans. *Cell* **75,** 855–862 (1993).

218. Williams, R. M., Primig, M., Washburn, B. K., Winzeler, E. A., Bellis, M., Sarrauste de Menthiere, C., Davis, R. W. & Esposito, R. E. The Ume6 regulon coordinates metabolic and meiotic gene expression in yeast. *Proceedings of the National Academy of Sciences* **99,** 13431–13436 (2002).

219. Winter, E. The Sum1/Ndt80 Transcriptional Switch and Commitment to Meiosis in Saccharomyces cerevisiae. *Microbiology and Molecular Biology Reviews* **76,** 1–15 (2012).

220. Wittkopp, P. J. & Kalay, G. Cis-regulatory elements: Molecular mechanisms and evolutionary processes underlying divergence. *Nature Reviews Genetics* **13,** 59–69 (2012).

221. Woo, H., Ha, S. D., Lee, S. B., Buratowski, S. & Kim, T. S. Modulation of gene expression dynamics by co-Transcriptional histone methylations. *Experimental and Molecular Medicine* **49,** e326–9 (2017).

222. Woznica, A., Gerdt, J. P., Hulett, R. E., Clardy, J. & King, N. Mating in the Closest Living Relatives of Animals Is Induced by a Bacterial Chondroitinase. *Cell* **170,** 1175–1183.e11 (2017).

223. Wu, A. C., Patel, H., Chia, M., Moretto, F., Frith, D., Snijders, A. P. & van Werven, F. J. Repression of Divergent Noncoding Transcription by a Sequence-Specific Transcription Factor. *Molecular Cell* **72,** 942–954.e7 (2018).

224. Wu, H., Yang, L. & Chen, L. L. The Diversity of Long Noncoding RNAs and Their Generation. *Trends in Genetics* **33,** 540–552 (2017).

225. Xiao, T., Hall, H., Kizer, K. O., Shibata, Y., Hall, M. C., Borchers, C. H. & Strahl, B. D. Phosphorylation of RNA polymerase II CTD regulates H3 methylation in yeast. *Genes and Development,* 654–663 (2003).

226. Xie, B., Horecka, J., Chu, A., Davis, R. W., Becker, E. & Primig, M. Ndt80 activates the meiotic ORC1 transcript isoform and SMA2 via a bi-directional middle sporulation element in Saccharomyces cerevisiae. *RNA Biology* **13,** 772–782 (2016).

227. Xie, J., Pierce, M., Gailus-Durner, V., Wagner, M., Winter, E. & Vershon, A. K. Sum1 and Hst1 repress middle sporulation-specific gene expression during mitosis in Saccharomyces cerevisiae. *EMBO Journal* **18,** 6448–6454 (1999).

228. Xu, L., Ajimura, M., Padmore, R., Klein, C. & Kleckner, N. NDT80, a meiosis-specific gene required for exit from pachytene in Saccharomyces cerevisiae. *Molecular and cellular biology* **15,** 6572–6581 (1995).

229. Yan, C., Chen, H. & Bai, L. Systematic Study of Nucleosome-Displacing Factors in Budding Yeast. *Molecular Cell* **71,** 294–305.e4 (2018).

230. Yen, K., Vinayachandran, V., Batta, K., Koerber, R. T. & Pugh, B. F. Genome-wide nucleosome specificity and directionality of chromatin remodelers. *Cell* **149,** 1461–1473 (2012).

231. Yuan, G.-C., Liu, Y.-J., Dion, M. F., Slack, M. D., Wu, L. F., Altschuler, S. J. & Rando, O. J. Genome-scale identification of nucleosome positions in S. cerevisiae. *Science (New York, N.Y.)* **309,** 626–30 (2005).

232. Yue, J. X., Li, J., Aigrain, L., Hallin, J., Persson, K., Oliver, K., Bergström, A., Coupland, P., Warringer, J., Lagomarsino, M. C., Fischer, G., Durbin, R. & Liti, G. Contrasting evolutionary genome dynamics between domesticated and wild yeasts. *Nature Genetics* **49,** 913–924 (2017).

233. Zernicka-Goetz, M., Morris, S. A. & Bruce, A. W. Making a firm decision: Multifaceted regulation of cell fate in the early mouse embryo. *Nature Reviews Genetics* **10,** 467–477 (2009).

234. Zhang, H., Wang, Y. & Lu, J. Function and Evolution of Upstream ORFs in Eukaryotes. *Trends in Biochemical Sciences,* 1–13 (2019).

235. Zhang, L., Ward, J. D., Cheng, Z. & Dernburg, A. F. The auxin-inducible degradation (AID) system enables versatile conditional protein depletion in C. elegans. *Development* **142,** 4374–4384 (2015).

236. Zhang, Y., Liu, T., Meyer, C. A., Eeckhoute, J., Johnson, D. S., Bernstein, B. E., Nusbaum, C., Myers, R. M., Brown, M., Li, W. & Liu, X. S. Model-based analysis of ChIP-Seq (MACS). *Genome biology* **9,** R137 (2008).

237. Zhang, Y., Moqtaderi, Z., Rattner, B. P., Euskirchen, G., Snyder, M., Kadonaga, J. T., Liu, X. S. & Struhl, K. Intrinsic histone-DNA interactions are not the major determinant of nucleosome positions in vivo. *Nature Structural and Molecular Biology* **16,** 847–852 (2009).