Chromatin Remodeling Around Nucleosome Free Regions Represses Non-Coding RNA By DNA Looping and Transcription Factor Dependent Targeting of Isw2

Adam Nicholas Yadon

A dissertation
submitted in partial fulfillment of the
requirements for the degree of

Doctor of Philosophy

University of Washington

2012

Reading Committee:

Toshio Tsukiyama, Chair
Susan Biggins
Steven M. Hahn

Program Authorized to Offer Degree:

Molecular and Cellular Biology

University of Washington

Abstract

Chromatin Remodeling Around Nucleosome Free Regions Represses Non-Coding RNA By DNA Looping and Transcription Factor Dependent Targeting of Isw2

Adam Nicholas Yadon

Chair of the Supervisory Committee:
Adjunct Professor, Toshio Tsukiyama
Department of Biochemistry

The efficient three-dimensional packaging of DNA into eukaryotic nuclei is accomplished through spatially organizing and compacting DNA into chromatin. Maintaining proper access to DNA, by modulation of either the location of DNA within the nucleus or the positions of nucleosomes, is essential for regulating DNA-dependent processes. The work in this dissertation focuses on elucidating the functions and molecular mechanisms by which the three-dimensional packaging of eukaryotic DNA affects DNA-dependent processes. Here I show that the ATP-dependent chromatin remodeling enzyme Isw2 is a global repressor of non-coding RNA (ncRNA) transcription that initiates from the edges of nucleosome free regions (NFRs) genome-wide. Isw2-dependent chromatin remodeling activity is required to reduce accessibility to DNA by sliding nucleosomes toward NFRs and occluding transcription start sites. This work

establishes Isw2 as the first factor that functions to reduce the size of NFRs *in vivo*. My evidence also suggests that proper repression of ncRNA by Isw2 prevents transcriptional interference of mRNA, providing an important biological role for Isw2-dependent chromatin remodeling. Analysis of the targeting mechanisms of Isw2 to NFRs identified the sequence-specific transcription factors (TFs) Ume6, Nrg1, Cin5, and Sok2 as globally required for Isw2 recruitment to many target loci genome-wide. This establishes the first comprehensive genome-wide map for TF-dependent targeting of a chromatin remodeling enzyme. The observation that Isw2 is targeted in a TF-dependent fashion to a large number of loci not containing an annotated TF binding site led to the discovery that Isw2 can also be targeted to specific loci via Ume6- and TFIIB-dependent DNA looping. Both Ume6 and TFIIB-dependent DNA looping are required to maintain transcriptional repression at target loci. I have thus identified DNA looping as a previously unknown mechanism to target a chromatin remodeling enzyme and uncovered a novel physiological role for DNA looping. My work has led to a better understanding of how the three-dimensional packaging of DNA into eukaryotic nuclei affects a DNA-dependent process in which transcriptional repression is facilitated by DNA-looping mediated TF-dependent targeting of a chromatin remodeling enzyme.

TABLE OF CONTENTS

Page

LIST OF FIGURES

Page

LIST OF TABLES

Page

ACKNOWLEDGEMENTS

There are numerous people whom I would like to genuinely thank for their generosity and support during my graduate studies. First, I want to acknowledge my advisor and mentor Toshi. His insight and dedication created an invaluable resource and environment for which I could learn and develop as scientist. I especially want to thank him for his guidance and support in overseeing and advising me on my projects from the first day to the last. Thank you also to my many colleagues, past and present, within the Tsukiyama lab: Iestyn Whitehouse, Jack Vincent, Courtney Boulton, Daniel Van de Mark, Molly Hogan, Ashwin Unnikrishnan, Vishu Anekonda, Tracey Au, Kim Barney, Eric Alcid, Laura Lee, Naomi Bogenschutz, Tanya Cunningham, and Jairo Rodriguez. In particular, I want to recognize Iestyn and Dan for blazing the initial RNA trail; Eric, Laura, Naomi, Tracey, and Tanya for all the insightful scientific dialogues and humor; and Jairo for being an outstanding, fun bay-mate full of long scientific (and too often photographic) discussions. In addition, I want to acknowledge the labs of Sue Biggins, Susan Parkhurst, and Steve Hahn for advice, help, and a fun work environment.

The work in this dissertation could not have been completed without technical help and advice from both Jeff Delrow and Ryan Basom; the support and guidance from my thesis committee, Sue Biggins, Steve Hahn, Steve Tapscott, and Barbara Wakimoto; administrative assistance and support from Michelle Karantsavalos, Jeremy Mseitif, and the MCB team at UW, including Terry Duffy, Diane Darling, and MaryEllin

Robinson; and the support of the Developmental Biology Training Grant, in particular the leadership and mentorship of Barbara Wakimoto.

A heartfelt appreciation and thank you belongs to my beautiful and loving wife Marisa-Claire, whose unfailing love and patience provided the foundation of support for which this dissertation could not have been completed otherwise.

Finally, I desire to express my sincere appreciation to all the many countless individuals, from teachers and coaches to family and friends (too numerous to list!), who have directly and/or indirectly shaped my life by believing in me and always encouraging me to aim big.

Thank you!!!

DEDICATION

To Mom and Dad

For your constant and unwavering support, guidance, and never ending love!

and

To the Lord Jesus Christ

For the beauty and complexity of the creation for which You have led me to explore!

INTRODUCTION

DNA-dependent processes are essential for normal cellular and organismal growth and differentiation. As a result, all eukaryotes must efficiently regulate access to DNA. However, access to DNA is complicated by the fact that relatively large amounts of DNA are packaged into comparatively small nuclei. For example, it has been estimated that a typical human cell must enclose ~2 meters of DNA into a nucleus less than 20 μM in diameter. Therefore, the total length of DNA is over 100,000 times the diameter of the nucleus in which it is contained. This packaging of DNA thus presents major structural challenges for eukaryotic cells in properly organizing and regulating access to DNA that is condensed with the nucleus.

To this end, eukaryotic cells package their DNA within the confines of the three-dimensional space of the nucleus at two distinct levels: by spatially positioning DNA within the nucleus and by compacting DNA into chromatin. Access to DNA can then be modulated by a variety of mechanisms that act by changing either the location of DNA with the nucleus, for example by DNA looping, and/or chromatin structure, for example through ATP-dependent chromatin remodeling enzymes.

The work in this dissertation focuses on elucidating the *in vivo* molecular mechanisms and functions that an ATP-dependent chromatin remodeling enzyme and DNA looping have on DNA-dependent processes.

The Spatial Positioning of Eukaryotic Genomes

The three-dimensional packaging of DNA is spatially positioned within eukaryotic nuclei at two distinct levels: the location of entire chromosomes within the nucleus (known as chromosome territories, CTs) and the physical co-localization of individual loci that are located far on a linear template (known as DNA looping). Disruption of either of these levels of DNA organization within the eukaryotic nucleus is associated with many complex diseases (Matarazzo et al., 2007; Misteli, 2010). The mechanisms by which eukaryotic DNA is spatially positioned within nuclei or how mis-regulating the location of DNA results in disease are very poorly understood. Elucidating these mechanisms is therefore critical to understanding basic cellular processes and complex diseases.

As early as the late 19th century, numerous microscopic-based studies visualized the nuclear structure of mitotic chromosomes. In 1909 Theodor Boveri, using rudimentary light microscopy, first coined the term “chromosome territories” (CTs) as a way to describe how individual mitotic chromosomes, in blastomere stages of the horse roundworm, appeared to retain their individuality during interphase and occupy distinct parts of the nuclear space (Boveri 1909). Despite the early description of CTs, direct evidence for CTs was not established until the mid-1980s when the development of fluorescence *in situ* hybridization (FISH) techniques paved the way for the visualization of individual chromosomes (Manuelidis, 1985; Schardin et al., 1985). The results from these early studies, while only capable of visualizing small regions of several chromosomes or a single entire chromosome, were consistent with Boveri’s CT observations.

Eventually, advances in both FISH (for example, the ability to individually “paint” every chromosome in the same cell with different “colors”) and microscopic techniques (such as the optical serial sectioning of nuclei using laser confocal microscopy and 3D image reconstitution) (Cremer et al., 2008) had begun to unequivocally show that chromosomes occupy discrete non-random territories, CTs, within the nuclei of animals (Cremer and Cremer, 2001), plants (Berr et al., 2006; Pecinka et al., 2004; Shaw et al., 2002), and yeast (Bystricky et al., 2005; Cremer and Cremer, 2010). However, it was later shown that the spatial separation of chromosomes into CTs is not necessarily absolute, as data revealed that chromosomal intermingling does occur around the periphery of the territories (Branco and Pombo, 2006).

More detailed studies of CTs has further demonstrated that chromosomes preferentially adopt a radial position within the nucleus, with large chromosomes found more often at the nuclear periphery and small chromosomes found more interiorly (Bolzer et al., 2005; Cremer and Cremer, 2006a, b). More interestingly, however, was the observation that within individual CTs, gene-poor and gene-rich regions are spatially separated (Shopland et al., 2006). It was suggested that this segregation of transcriptionally active and inactive regions within CTs reflected transcriptional regulation by the discrete positioning of DNA within the nucleus. This has since been supported by the observations that certain genes loop out of their CT upon activation and are repositioned into distinct transcription factories (Bickmore et al., 2004; Chambeyron and Bickmore, 2004a, b; Ferrai et al., 2010a; Ferrai et al., 2010b). Additionally, a correlation between transcriptional activation and the location of genes within the nucleus has been observed, such that inactive genes are preferentially

located closer to the nuclear periphery than transcriptionally active genes (Kosak et al., 2002; Meister et al., 2010; Ragoczy et al., 2006). Together, these data suggested that the spatial positioning of genomes within eukaryotic nuclei has strong regulatory effects on transcription (Saez-Vasquez and Gadal, 2010; Sexton et al., 2009).

The development of chromosome conformation capture (3C) (Dekker et al., 2002), and the subsequent 3C-derived genomic techniques 4C (Simonis et al., 2006), 5C (Dostie et al., 2006), HI-C (Lieberman-Aiden et al., 2009), ChIP-loop (Horike et al., 2005), and ChIA-PET (Fullwood et al., 2009), have begun to provide unprecedented insight into the molecular details governing how eukaryotic genomes are spatially packaged within nuclei. These techniques, with a resolution of a few hundred base pairs (compared to thousands of base pairs for FISH), allow for the quantitative detection and/or *de novo* identification of chromosomal loci that are far away on a linear template, but co-localize in three-dimensional space and thus are located in close proximity within the nucleus. This phenomenon has been termed “DNA looping” because intervening DNA between two interacting loci is “looped” out to allow the physical co-localization of the loci. The 3C techniques have subsequently confirmed FISH studies by demonstrating that chromosomes do indeed generally occupy defined territories, CTs, which can be further subdivided into transcriptionally active and inactive regions (Noordermeer et al., 2011; Simonis et al., 2006). More importantly, these techniques have begun to elucidate a more detailed picture of the spatial positioning of eukaryotic genomes within the nucleus by providing high resolution mapping of distinct chromosomal interactions and identifying the molecular components required for these interactions.

3C based technologies have discovered two DNA looping confirmations (Figure 1): chromatin looping and gene looping. Chromatin loops have traditionally been defined as the juxtaposition of a distal regulatory region, such as an enhancer or silencer, with a gene promoter (often over 10s of KBs) (Figure 1A). Gene loops are described as the physical association of promoter and terminator regions of the same gene (generally between 1 and 10 KBs) (Figure 1B). To date, both chromatin loops and gene loops have been described at numerous loci in metazoan and yeast cells (Ansari and Hampsey, 2005; Hampsey et al., 2011; Laine et al., 2009; Murrell et al., 2004; Nemeth et al., 2008; O'Reilly and Greaves, 2007; O'Sullivan et al., 2004; Perkins et al., 2008; Singh and Hampsey, 2007; Spilianakis et al., 2005; Tan-Wong et al., 2008; Tolhuis et al., 2002; Vernimmen et al., 2007).

Extensive molecular characterization of DNA loops has led to the identification of numerous interactions between structural proteins (Comet et al., 2011; Hadjur et al., 2009; Parelho et al., 2008; Wendt et al., 2008), transcription factors (TFs) (Drissen et al., 2004; Phillips and Corces, 2009; Splinter et al., 2006; Vakoc et al., 2005), general transcription factors (Singh and Hampsey, 2007), and RNA 3'-end processing factors (Singh and Hampsey, 2007) that are required for the *in vivo* formation and/or maintenance of DNA loops at numerous loci in various organisms. Additionally, chromatin loops have been shown to play roles in transcription regulation (Comet et al., 2011; Nemeth et al., 2008; Perkins et al., 2008; Schoenfelder et al., 2010a; Schoenfelder et al., 2010b; Wang et al., 2011) while gene loops have been implicated in transcriptional memory (Laine et al., 2009; Tan-Wong et al., 2009). While much work has described the molecular framework governing DNA loop formation and/or

maintenance, the molecular mechanisms by which DNA loops effect transcription regulation or memory are very poorly understood. Discerning these mechanisms would provide essential details linking the roles for the spatial positioning of the genome within nuclei with DNA-dependent processes.

The work in this dissertation demonstrates a novel mechanism by which Ume6- and TFIIB-dependent DNA looping mediates the targeting of the ATP-dependent chromatin remodeling enzyme Isw2 to loci genome-wide. This identifies Ume6 as the first transcriptional repressor required for DNA looping. Both Ume6- and TFIIB-dependent DNA looping is shown to be required to maintain proper transcriptional repression at target loci. These results establish a novel physiological role for DNA looping in which the three-dimensional spatial positioning of the genome in the nucleus of a eukaryotic organism affects chromatin structure.

Chromatin Structure

The packaging of DNA into eukaryotic nuclei is facilitated, at least in part, by its compaction into chromatin. The most basic repeating unit of chromatin is the nucleosome, composed of ~147 base pairs (bp) of DNA wrapped 1.65 times around an octamer of histone proteins, two each of H3, H4, H2A, and H2B (Arents et al., 1991; Kornberg and Lorch, 1999; Luger et al., 1997). Nucleosome arrays appear as “beads on a string” by electron microscopy, where the beads are nucleosomes and the intervening linker DNA is the string (Olins and Olins, 1974). Since the basic structure and composition of nucleosomes was first discovered in 1974 (Kornberg, 1974; Kornberg and Thomas, 1974; Sahasrabuddhe and Van Holde, 1974), much interest has

focused on determining the positions of nucleosomes on DNA and their effects on DNA-dependent processes.

Prior to 2004, mapping nucleosome positions was performed using indirect end-labeling. This rather laborious methodology resulted in the mapping of only a small handful of nucleosome positions at a select few, intensively studied, single gene loci, such as at the chicken β-globin locus, and the yeast *PHO5*, *GAL1-10*, and *HIS3* promoters. By 2004, technological advances led to the invention of the DNA microarray, which, for the first time, began providing a more global assessment of chromatin structure across the *S. cerevisiae* genome. These studies (Bernstein et al., 2004; Lee et al., 2004; Sekinger et al., 2005), combined with the locus specific studies, led to the general conclusions that transcriptionally active promoters are relatively nucleosome deficient while transcriptionally inactive promoters are highly enriched with nucleosomes. These early microarrays however, were extremely low-resolution (~500-2000 bp resolution). As a result, they could only distinguish the relative enrichment of nucleosomes across large chromosomal domains and could not adequately map the locations of single nucleosomes.

Eventually microarray technology advanced to “high-resolution” single nucleosome mapping capabilities. It was thus in 2005 that Yuan and colleagues (Yuan et al., 2005) mapped nucleosome positions at single nucleosome resolution across chromosome III of the *S. cerevisiae* genome. This seminal study identified a ~150 bp nucleosome free region (NFR) that resided within the promoters of many transcriptionally active genes. Subsequently, microarray resolution improved to ~4 bp resolution and is capable of entirely tiling the relatively small *S. cerevisiae* genome.

Advances in high-throughput sequencing technologies have further enabled a theoretical resolution of 1 bp and been successfully utilized to map nucleosome positions in several eukaryotic organisms. To date, high-resolution (single nucleosome) maps of nucleosome positions have been generated in a number of eukaryotic organisms including *S. cerevisiae* (Albert et al., 2007; Field et al., 2008; Kaplan et al., 2009; Lee et al., 2007; Mavrich et al., 2008a; Raisner et al., 2005; Shivaswamy et al., 2008; Whitehouse et al., 2007; Yuan et al., 2005; Zhang et al., 2009), *D. melanogaster* (Mavrich et al., 2008b), *C. elegans* (Johnson et al., 2006; Valouev et al., 2008), *O. latipes* (Sasaki et al., 2009), and humans (Barski et al., 2007; Jin et al., 2009; Schones et al., 2008).

Surprisingly, all of these eukaryotes display a remarkably similar chromatin structure (Figure 2), revealing a stereotypical pattern of nucleosome positions that spans both gene coding regions and transcriptional regulatory regions (Rando and Chang, 2009). In general, gene coding regions have a characteristically high nucleosome occupancy, containing arrays of progressively decreasing phased nucleosomes extending from the transcription start sites (TSSs) (Figure 2A). In contrast, transcriptional regulatory regions are strongly nucleosome depleted (containing what is often referred to as a NFR) immediately upstream and downstream of the transcription start sites (TSSs) (Figure 2A) and transcription termination sites (TTSs) (Figure 2B), respectively. It should be noted that the use of the term NFR simply refers to the absence of canonical nucleosomes, as assayed by the nuclease digestion pattern of MNase which is known to preferentially cleave “linker” DNA between two canonical nucleosomes. The ability of MNase to digest DNA within non-canonical

nucleosomes has not been well established, however recent data suggests that the isolation procedure (i.e. salt concentration during extraction) of nucleosomes may select for canonical nucleosomes and the inadvertent removal of non-canonical nucleosomes (Jin et al., 2009).

As early as 1980, both *in vivo* and *in vitro* studies concluded that histones are generally repressive to transcription (Ehrenhofer-Murray, 2004; Han and Grunstein, 1988; Han et al., 1988; Kim et al., 1988; Knezetic and Luse, 1986; Lorch et al., 1987; Svaren and Horz, 1997; Weisbrod et al., 1980; Wu, 1980). The biochemical characterization and crystal structure of nucleosomes has reinforced these conclusions by demonstrating that the many histone-DNA contacts made within nucleosomes make the nucleosome one of the most stable protein DNA complexes (Luger et al., 1997) which can effectively occlude access to the underlying DNA or impede passage of proteins along DNA. As a result, processes such as transcription, which rely on access to and processive passage of proteins along DNA, are significantly inhibited by the presence of nucleosomes. These results are consistent with subsequent more detailed studies of NFRs located at the 5'- and 3'-end of genes (5'-NRFs and 3'-NFRs, respectively) and have led to the general conclusions that NFRs are the primary sites of transcriptional regulation.

5'-NFRs have been shown to be significantly enriched with sequence-specific TF binding sites (Bernstein et al., 2004; Lee et al., 2007; Whitehouse et al., 2007; Yuan et al., 2005) (Figure 2A) and are predominately associated with transcriptionally active promoters. Studies in *S. cerevisiae* (Mavrich et al., 2008a) and *D. melanogaster* (Mavrich et al., 2008b) found that the sizes of 5'-NFRs were positively correlated with

transcriptional frequency. On the other hand, the promoters of transcriptionally repressed genes are generally compacted into arrays of nucleosomes. For example, in *S. cerevisiae* (Jiang and Pugh, 2009a; Mavrich et al., 2008a; Shivaswamy et al., 2008) and *D. melanogaster* (Mavrich et al., 2008b), TATA-containing gene promoters, which are typically associated with transcriptionally inducible or regulated genes (Basehoar et al., 2004), generally lack NFRs and are predominately found condensed into arrays of nucleosomes. Upon inducing conditions, many of these promoters become decondensed, forming a 5'-NFR prior to transcription initiation (Almer and Horz, 1986; Almer et al., 1986; Lam et al., 2008). It should be noted that 5'-NFRs can be found at some transcriptionally inactive gene promoters in *S. cerevisiae* (Mavrich et al., 2008a), *D. melanogaster* (Mavrich et al., 2008b), and humans (Schones et al., 2008). Together these studies suggest that the presence of a NFR may be necessary but not sufficient for transcription initiation. Understanding the mechanisms by which the presence or size of NFRs regulate transcription would be invaluable to the further elucidation of transcriptional regulatory mechanisms.

3'-NFRs, first identified in *S. cerevisiae* (Kaplan et al., 2009; Mavrich et al., 2008a; Shivaswamy et al., 2008) but subsequently found in *D. melanogaster* and humans (Jin et al., 2009; Mavrich et al., 2008b), are generally located just downstream of the translation termination site and are highly enriched with TTSs (Jin et al., 2009; Mavrich et al., 2008a; Mavrich et al., 2008b) (Figure 2B). It has been hypothesized that 3'-NFRs may function in the assembly of the transcription termination machinery or their regulators (Mavrich et al., 2008a; Mavrich et al., 2008b), however the mechanisms for this are not known. Therefore, understanding the role and function of 3'-NFRs in

mediating or contributing to transcription regulation will be an important step in understanding how chromatin structure may regulate this process.

Interestingly, the locations of TSS are also highly stereotypical with respect to NFRs. For example in *S. cerevisiae,* mRNA TSSs typically reside ~10-15 bp downstream from the edge of 5'-NFRs (Albert et al., 2007). Similarly, antisense TSSs, initiating bidirectionally from *S. cerevisiae* promoter or terminator regions, were similarly mapped to the upstream edge of 5'-NFRs and 3'-NFRs, respectively (Neil et al., 2009; Xu et al., 2009) (Figure 2). In humans (Schones et al., 2008) and *D. melanogaster* (Mavrich et al., 2008b) TSSs tend to reside ~40-60 bp upstream of the edge of 5'-NFRs (Mavrich et al., 2008b; Schones et al., 2008). However, these studies also revealed the presence of a paused RNA polymerase II (RNAPII) immediately downstream of the edges of 5'-NFRs at a number of transcriptionally inactive gene promoters (Mavrich et al., 2008b; Schones et al., 2008). At least in *D. melanogaster*, this transcriptionally engaged, but paused, RNAPII appears to physically interact with the first nucleosome downstream of the NFR (Mavrich et al., 2008b). These results suggest the nucleosomes immediately flanking NFRs may have important regulatory roles in mediating transcription initiation or elongation in yeast and flies, respectively. In support of this, studies have shown that the physical properties of DNA (Anderson and Widom, 2001; Iyer and Struhl, 1995; Kaplan et al., 2009; Mavrich et al., 2008a; Yuan et al., 2005; Zhang et al., 2009), sequence-specific TFs (Badis et al., 2008; Hartley and Madhani, 2009), and chromatin regulators (Badis et al., 2008; Hartley and Madhani, 2009) can positively regulate the formation and/or size of NFRs *in vivo*. The activities of these factors to establish larger NFRs were thought to facilitate the initiation of

transcription by allowing the transcription machinery greater access to DNA. Whether there are mechanisms to negatively regulate the size of NFRs *in vivo* is not known.

The work in this dissertation describes the systematic annotation of NFRs across the yeast genome from multiple nucleosome mapping studies. I identify, in addition to 5'-NFRs and 3'-NFRs, two novel classes of NFRs located within open reading frames (ORF-NFRs) and far from open reading frames (Other-NFRs). The ATP-dependent chromatin remodeling enzyme Isw2 was found to target a subset of each of the four classes of NFRs where it functions to reduce the accessibility of these NFRs by sliding the nucleosomes immediately flanking the NFRs toward the center. This establishes Isw2 as the first chromatin regulator that functions to negatively regulate the size of NFRs. The Isw2-dependent reduction in NFR size was required to globally repress cryptic ncRNA transcription from initiating at the edges of NFRs by positioning a nucleosome to occlude the cryptic transcription start site. These results thus describe mechanisms by which modulation of chromatin structure is directly required for the repression of cryptic non-coding RNA.

ATP-Dependent Chromatin Remodeling Enzymes

Due to the inhibitory nature of nucleosomes on DNA-dependent processes, eukaryotic cells have evolved highly conserved protein complexes to regulate the accessibility of chromatin by modifying the structure, composition, and/or positions of nucleosomes on DNA. The need to understand the regulation, functions, and mechanisms of action of these chromatin regulators is underscored by their tight link to the regulation of all DNA-dependent processes and the development of cancer (Dey,

2006). One major class of chromatin regulators is the ATP-dependent chromatin remodeling enzymes. This class utilizes the energy of ATP-hydrolysis to slide nucleosomes, evict or exchange histones from nucleosomes, or destabilize nucleosomes by altering histone-DNA contacts (Ehrenhofer-Murray, 2004; Mellor and Morillon, 2004; Tsukiyama, 2002) to affect numerous cellular processes including transcription, replication, DNA repair, and recombination. All ATP-dependent chromatin remodeling enzymes identified to date are multisubunit protein complexes belonging to the Swi2/Snf2 superfamily of ATPases and are further subclassified by the homology of their ATPase domains into the SWI/SNF (Switching/Sucrose Non-Fermenting), ISWI (Imitation Switch), CHD (Chromodomain, Helicase, and DNA binding), and INO80 (Inositol80) families (Eisen et al., 1995; Flaus et al., 2006; Langst and Becker, 2004).

The ISWI family is highly conserved from yeast to humans (Flaus et al., 2006; Yadon and Tsukiyama, 2011) and has been shown to be necessary for normal development in *D. melanogaster* (Deuring et al., 2000), *C. elegans* (Andersen et al., 2006), *X. laevis* (Dirscherl et al., 2005; Wysocka et al., 2006), and mice (Stopka and Skoultchi, 2003) (Figure 3). While the requirement of the ISWI ATPases in development has largely been attributed to their roles in transcriptional regulation of developmentally controlled genes (Badenhorst et al., 2002; Barak et al., 2003; Deuring et al., 2000; Lazzaro et al., 2006), ISWI containing complexes have also been shown to play key roles in a variety of biological processes, including transcription (Badenhorst et al., 2002; Fazzio et al., 2001; Goldmark et al., 2000; Kent et al., 2001; Yasui et al., 2002), global chromatin structure (Deuring et al., 2000), DNA replication (Collins et al., 2002; Poot et al., 2004; Vincent et al., 2008), cell cycle progression (Fyodorov et al.,

2004), ribosomal DNA silencing (Zhou et al., 2002), and cohesin loading (Hakimi et al., 2002). It is therefore not surprising that mutations of ISWI complexes have been associated with disease, such as melanotic tumors in *D. melanogaster* (Badenhorst et al., 2002) or malignancy of acute leukemia cells in humans (Stopka et al., 2000). *In vitro*, ISWI complexes are capable of sliding (Corona et al., 1999; Hamiche et al., 1999; Langst et al., 1999), exchanging histones within (Bruno et al., 2003), altering the spacing between (Tsukiyama et al., 1999; Varga-Weisz et al., 1997), and assembling nucleosomes (Ito et al., 1997; Loyola et al., 2001). How these *in vitro* activities are used *in vivo* is largely unknown and is essential to understanding the biological processes associated with ISWI complexes.

S. cerevisiae contains two ISWI homologues, Isw1 and Isw2 (Tsukiyama et al., 1999). Isw2 has been shown to form two distinct complexes *in vivo*, a two-subunit (Isw2 and Itc1) (Tsukiyama et al., 1999) and a four-subunit (Isw2, Itc1, Dls1, and Dbp4) (Iida and Araki, 2003; McConnell et al., 2004). A functional distinction between these complexes has not been elucidated. *In vitro,* the Isw2 complex interacts with DNA and nucleosome arrays in an ATP-independent manner (Gelbart et al., 2001), facilitates regular spacing of nucleosomes (Gelbart et al., 2001; Tsukiyama et al., 1999), and slides nucleosomes toward the center of DNA fragments without disrupting nucleosome integrity (Kassabov et al., 2002). *In vivo,* Isw2 affects integration site selection of the Ty1 retrotransposon upstream of tRNA genes (Bachman et al., 2005; Gelbart et al., 2005), facilitates DNA replication in the presence of methyl methanesulfonate (MMS) (Vincent et al., 2008), and represses coding transcription initiation by overriding the

underlying nucleosome positioning signals of DNA (Fazzio et al., 2001; Gelbart et al., 2005; Goldmark et al., 2000; Kent et al., 2001; Whitehouse and Tsukiyama, 2006).

Most recently, global Isw2 function was elucidated using high-resolution tiled microarrays across the yeast genome (Whitehouse et al., 2007). This study identified ~2100 target loci that encompass three primary classes of targets at which Isw2 is recruited and remodels chromatin: the 5'-end of genes, the 3'-end of genes (hereafter denoted as 5'-end targets and 3'-end targets, respectively), and upstream of tRNA genes. At 5'- and 3'-end targets, Isw2 was shown to increase nucleosome occupancy within intergenic regions by sliding nucleosomes away from genic regions. Surprisingly, the identification of the 3'-end of genes as a novel class of Isw2 targets led to the discovery that Isw2 suppresses non-coding RNA (ncRNA) transcription at three of these loci, implicating Isw2 in transcriptional regulation of ncRNA at 3'-end targets (Whitehouse et al., 2007). Subsequent mapping of these ncRNAs transcription start sites onto nucleosome positions revealed initiation occurring at the edges of NFRs, similar to that observed for coding transcription start sites. These results have raised the intriguing possibility that Isw2 functions to alter nucleosome density around NFRs to suppress ncRNA transcription initiation. However, whether all Isw2 targets are adjacent to NFRs or whether Isw2 functions to globally suppress ncRNA at 3'-end targets is currently unknown.

In addition to understanding the biological functions of ATP-dependent chromatin remodeling enzymes *in vivo*, a major challenge is to elucidate the mechanisms by which these enzymes are targeted to specific genomic loci. ISWI ATPases are extremely abundant in eukaryotic organisms, ranging from ~4,500 copies in haploid yeast nuclei to

~100,000 copies in diploid *D. melanogaster* embryonic cells. Therefore the targeting of these complexes to specific genomic loci is likely important for proper regulation of the enzyme. An understanding of the targeting mechanisms for ISWI homologues is just beginning to be investigated. For example, in yeast, Isw2 has been shown to be targeted to some loci by the sequence-specific TFs Ume6 (Goldmark et al., 2000), α2-Mcm1 (Bachman et al., 2005; Gelbart et al., 2005), and Bdp1 (Bachman et al., 2005; Gelbart et al., 2005). However, the high-resolution identification of Isw2 targets across the genome (Whitehouse et al., 2007) has revealed a far greater number of Isw2 targets than can be accounted for by these TFs alone, indicating much is still to be discovered underlying the targeting mechanisms of Isw2. Elucidation of these mechanisms will further our mechanistic understanding of the regulation and function of the ISWI chromatin remodeling enzymes *in vivo*.

The work in this dissertation identifies the ATP-dependent chromatin remodeling enzyme Isw2 as the first factor that is required to reduce the accessibility of NFRs by sliding nucleosomes toward the centers of NFRs. The Isw2-dependent chromatin remodeling of NFRs is required to repress cryptic ncRNA from initiating bi-directionally from 5'-NFRs or antisense from 3'-NFRs by positioning the nucleosomes flanking NFRs such that they occlude the TSSs. The TFs Ume6, Nrg1, Cin5, and Sok2 were found to be required for the targeting of Isw2 to many loci genome-wide, providing the first comprehensive genome-wide map of TF-dependent targeting of an ATP-dependent chromatin remodeling enzyme. Finally, DNA looping is shown to be a novel mechanism for the targeting of Isw2 to many loci genome-wide. These results provide a genome-

wide evaluation of the functions and mechanisms for the targeting and transcriptional repression of the ATP-dependent chromatin remodeling enzyme Isw2.

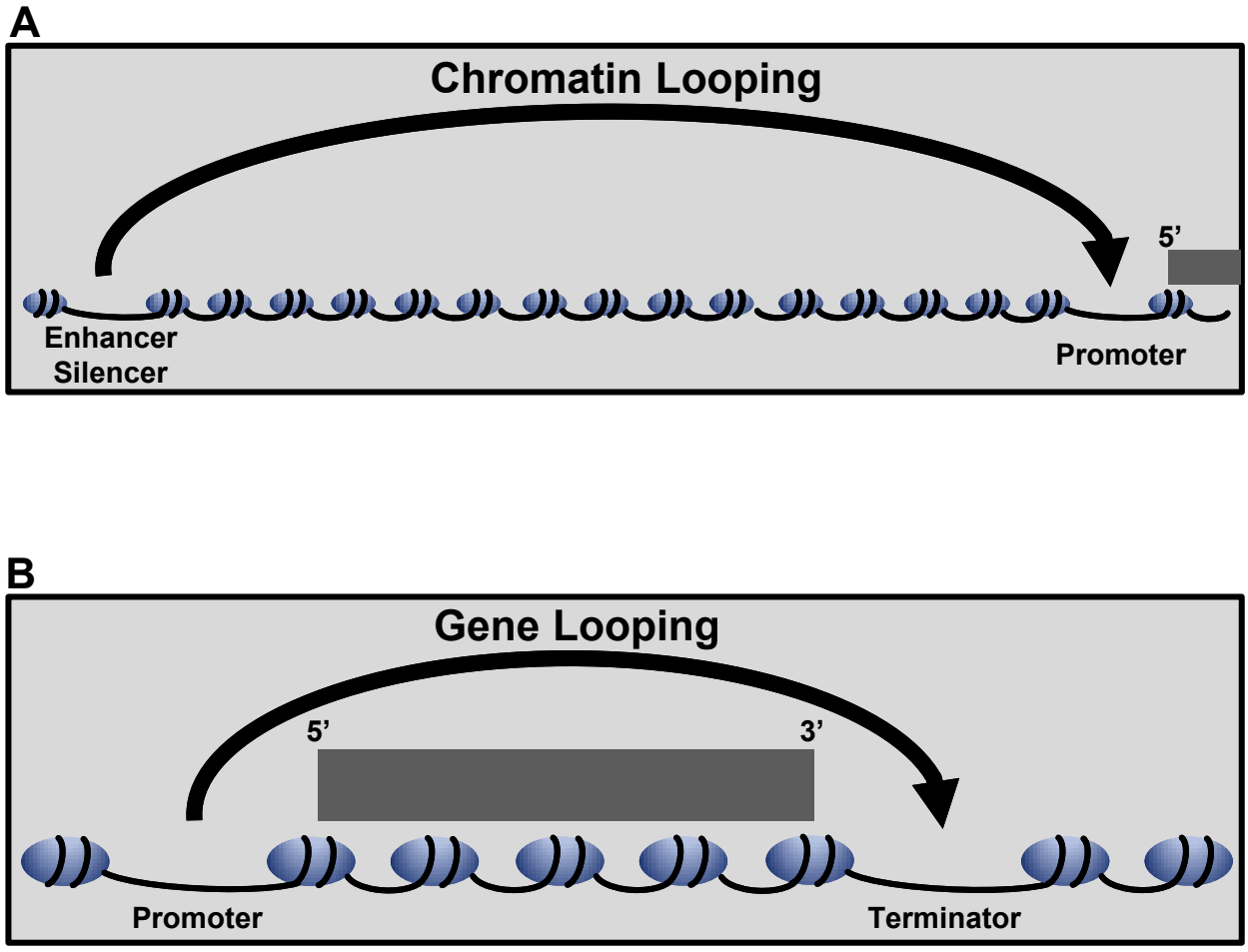


Figure 1: Schematic Diagram of Identified DNA Looping Conformations: Chromatin Looping and Gene Looping
(A and B) Nucleosomes and DNA are depicted as blue ovals and black lines, respectively. Genes are marked as grey rectangles. Arrow depicts physical interactions between two distally located regions on the linear chromosome.
(A) Chromatin Looping
(B) Gene Looping

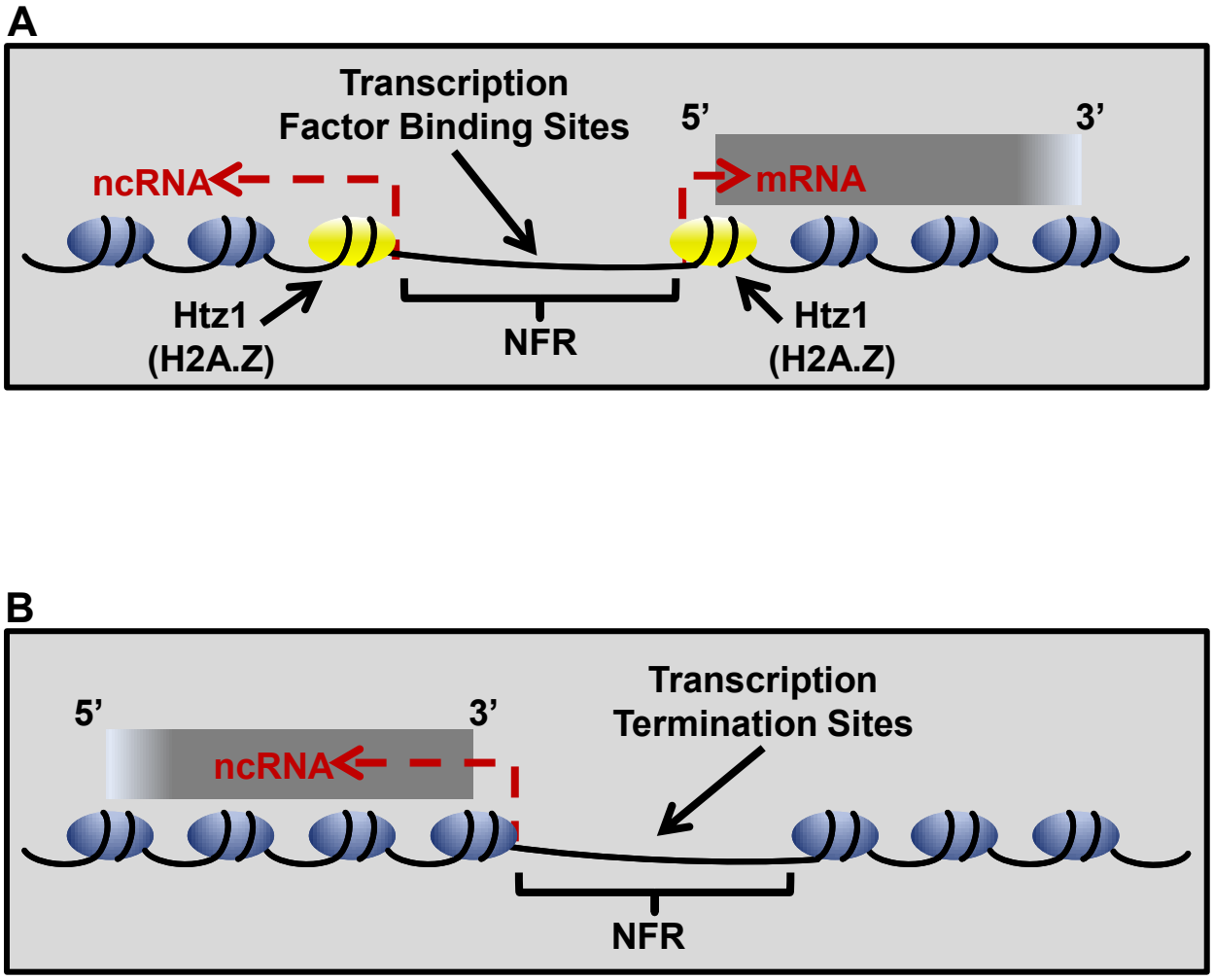


Figure 2: Schematic Diagram of Stereotypical Chromatin Structure in Yeast
(A and B) Canonical and Htz1 containing nucleosomes are depicted as blue and yellow ovals, respectively. DNA is black lines and genes are marked as grey rectangles. Transcription start sites for mRNA and ncRNA is depicted as dashed arrows.
(A) Chromatin structure surrounding the promoter of a typical yeast gene
(B) Chromatin structure surrounding the terminator of a typical yeast gene

Figure 3: Cell SnapShot – Chromatin Remodeling: ISWI

The ISWI (Imitation SWItch) family of ATP-dependent chromatin remodeling enzymes are a set of highly conserved protein complexes which utilize the energy released from ATP hydrolysis to slide nucleosomes along DNA and/or replace histones within nucleosomes. All ATP-dependent chromatin remodeling complexes, including the ISWI family, contain a conserved catalytic DEXD ATPase domain and helicase domain. However, unique to ISWI family members is the combination of three C-terminally located domains, known as HAND, SANT, and SLIDE. The SANT domain (structurally related to the c-Myb DNA binding domains) binds unmodified histone tails, the SLIDE (SANT-Like ISWI Domain) domain binds nucleosomal DNA near the dyad axis, and the HAND domain is implicated in both histone and DNA binding/recognition. Representative ISWI containing protein complexes from multiple species are depicted in this SnapShot with specific *in vivo* biological functions for each listed below (indicated by numbers).

Biological Functions of ISWI Containing Complexes

(1) Replication, Transcription, and Ty Integration Regulation: Facilitates replication fork progression through late replicating regions. Represses mRNA and cryptic non-coding RNA transcription by negatively regulation NFR size. Required for the periodic integration pattern of the Ty1 retrotransposon.

(2) Transcription Regulation: Represses and activates transcription at a small number of loci and is implicated in transcription elongation and termination regulation.

(3) Cell Fate Determination: Promotes the expression of vulval cell fates by antagonizing the transcriptional and chromatin-remodeling activities of complexes similar to Myb-MuvB/dREAM, NuRD, and Tip60/NuA4.

(4) Transcription Regulation: Downregulates VSV expression sites.

(5) Nucleosome Assembly, Transcription, and Replication Regulation: Required for the establishment and/or maintenance of periodic nucleosome arrays that contributes to pericentric position-effect variegation (PEV) and heterochromatic *Polycomb*-mediated transcriptional gene silencing. Implicated in the regulation of S-phase length/progression.

(6) Chromatin Structure and Transcription Regulation: Maintains higher-order chromatin structure by mediating chromatin compaction. Disrupts the enhancer-blocking function of Fab7 and SF1 while augmenting the function of Fab8. Activates transcription of GAGA target genes, ecdysone responsive genes, and is a co-activator of Armadillo. Regulates innate immunity by repressing transcription of JAK/STAT target genes.

(7) Chromatin Structure and Transcription Regulation: Involved in formation of silent heterochromatin by incorporating the histone variant H2Av, thus suppressing position-effect variegation (PEV).

(8) Developmental Cellular Expansion and Nuclear Divisions: Necessary for cell expansion during late diploid (sporophytic) embryogenesis and mitotic nuclear divisions during haploid (gametophytic) phase.

(9) Replication and Transcription Regulation: Required for S-phase progression and facilitates pericentromeric heterochromatin DNA replication. Represses transcription of the vitamin D3 receptor-regulated genes in humans.

(10) Chromatin Structure and Centromere Maintenance: Implicated in chromatin assembly and actively supports the assembly of CENP-A chromatin in humans.

(11) Development and Transcription Regulation: Functions in neural tube formation and terminal differentiation of ovarian granulose cells through regulation of StAR gene expression in mice.

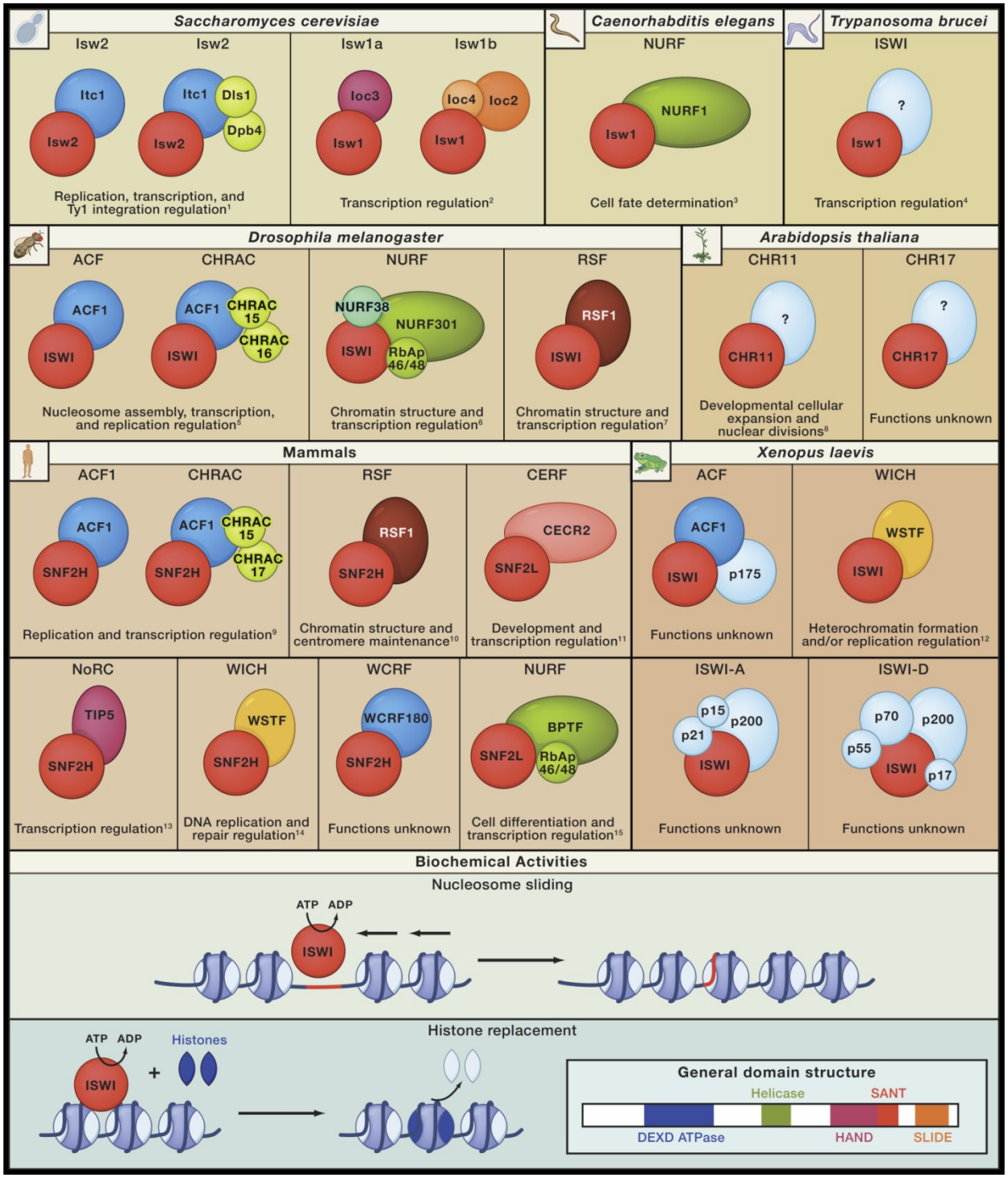
(12) Heterochromatin Formation and/or Replication Regulation: Targeted to pericentromeric heterochromatin during early stages of chromosome condensation and DNA replication.

(13) Transcription Regulation: Involved in the transcriptional repression of ribosomal RNA genes.

(14) DNA Replication and Repair Regulation: Implicated in heterochromatin DNA replication by binding PCNA in mice. Necessary for cell survival following DNA damage by methyl methanesulfate (MMS) and facilitates a DNA damage response pathway by controlling histone H2A.Z function in mice.

(15) Cell Differentiation and Transcription Regulation: Promotes neurite outgrowth and transcription of engrailed 1 and 2 in humans.

Figure and text adapted from Yadon and Tsukiyama, 2011


Saccharomyces cerevisiae
Isw2
Isw2
Itc1
Isw2
Itc1
Dls1
Dpb4
Isw2
Replication, transcription, and Ty1 integration regulation[1]
Isw1a
Isw1b
Ioc3
Isw1
Ioc4
Ioc2
Isw1
Transcription regulation[2]
Caenorhabditis elegans
NURF
NURF1
Isw1
Cell fate determination[3]
Trypanosoma brucei
ISWI
?
Isw1
Transcription regulation[4]
Drosophila melanogaster
ACF
ACF1
ISWI
CHRAC
ACF1
CHRAC 15
CHRAC 16
ISWI
Nucleosome assembly, transcription, and replication regulation[5]
NURF
NURF38
NURF301
ISWI
RbAp 46/48
Chromatin structure and transcription regulation[6]
RSF
RSF1
ISWI
Chromatin structure and transcription regulation[7]
Arabidopsis thaliana
CHR11
?
CHR11
Developmental cellular expansion and nuclear divisions[8]
CHR17
?
CHR17
Functions unknown
Mammals
ACF1
ACF1
SNF2H
CHRAC
ACF1
CHRAC 15
CHRAC 17
SNF2H
Replication and transcription regulation[9]
RSF
RSF1
SNF2H
Chromatin structure and centromere maintenance[10]
CERF
CECR2
SNF2L
Development and transcription regulation[11]
Xenopus laevis
ACF
ACF1
ISWI
p175
Functions unknown
WICH
WSTF
ISWI
Heterochromatin formation and/or replication regulation[12]
NoRC
TIP5
SNF2H
Transcription regulation[13]
WICH
WSTF
SNF2H
DNA replication and repair regulation[14]
WCRF
WCRF180
SNF2H
Functions unknown
NURF
BPTF
SNF2L
RbAp 46/48
Cell differentiation and transcription regulation[15]
ISWI-A
p15
p200
p21
ISWI
Functions unknown
ISWI-D
p70
p200
p55
ISWI
p17
Functions unknown
Biochemical Activities
Nucleosome sliding
ATP ADP
ISWI
Histone replacement
ATP ADP
Histones
ISWI
+
General domain structure
Helicase
SANT
DEXD ATPase
HAND
SLIDE

CHROMATIN REMODELING AROUND NUCLEOSOME FREE REGIONS LEADS TO REPRESSION OF NON-CODING RNA TRANSCRIPTION

Abstract

Nucleosome free regions (NFRs) at the 5'- and 3'-end of genes are general sites of transcription initiation for mRNA and non-coding RNA (ncRNA). The presence of NFRs within transcriptional regulatory regions and the conserved location of transcription start sites at NFRs strongly suggest the regulation of NFRs profoundly affects transcription initiation. To date, multiple factors are known to facilitate transcription initiation by positively regulating the formation and/or size of NFRs *in vivo*. However, mechanisms to repress transcription by negatively regulating the size of NFRs have not been identified. We identified four distinct classes of NFRs located at the 5'-and 3'-end of genes, within ORFs, and far from ORFs. The ATP-dependent chromatin remodeling enzyme Isw2 was enriched at all classes of NFRs. Analysis of RNA levels also demonstrated Isw2 is required to repress ncRNA transcription from many of these NFRs. Thus by the systematic annotation of NFRs across the yeast genome and analysis of ncRNA transcription, we establish, for the first time, a mechanism by which NFR size is negatively regulated to repress ncRNA transcription from NFRs. Finally, we provide evidence suggesting one biological consequence for

repression of ncRNA, by Isw2 or the exosome, prevents transcriptional interference of mRNA.

Introduction

Eukaryotic cells compact their DNA into a nucleoprotein complex known as chromatin. The most basic repeating unit of chromatin is the nucleosome, consisting of ~147 base pairs (bp) of DNA wrapped around an octamer of histone proteins (Luger et al., 1997). Nucleosomes are one of the most stable protein-DNA complexes known (Luger et al., 1997) and can effectively inhibit all DNA-dependent processes, including transcription, replication, repair, and recombination, by limiting the access of proteins to DNA (Ehrenhofer-Murray, 2004). As a result, the mechanisms by which chromatin structure and nucleosome positions are specified and maintained *in vivo* are critical for the regulation of all DNA-dependent processes.

Genome-wide maps of nucleosome positions have recently been generated in a number of organisms, including *S. cerevisiae* (Albert et al., 2007; Bernstein et al., 2004; Field et al., 2008; Jiang and Pugh, 2009b; Kaplan et al., 2003; Lee et al., 2004; Lee et al., 2007; Mavrich et al., 2008a; Raisner et al., 2005; Shivaswamy et al., 2008; Whitehouse et al., 2007; Yuan et al., 2005; Zhang et al., 2009), *D. melanogaster* (Mavrich et al., 2008b), *C. elegans* (Johnson et al., 2006; Valouev et al., 2008), *O. latipes* (Sasaki et al., 2009), and humans (Barski et al., 2007; Schones et al., 2008). Each of these organisms displays a characteristic chromatin structure spanning gene coding regions and transcriptional regulatory regions. Gene coding regions generally have high nucleosome occupancy with arrays of well-phased nucleosomes extending

from the 5'-end of genes. In contrast, transcriptional regulatory regions, such as promoters, enhancers, and terminators, have low nucleosome occupancy and often contain a nucleosome free region (NFR). NFRs, also known as nucleosome depleted regions (NDRs), typically represent regions with an increased accessibility to micrococcal nuclease (MNase) digestion. Thus, the term NFR refers to a deficiency in experimentally determined canonical nucleosomes and does not necessarily imply a complete lack of histones.

To date, predominately two major classes of NFRs, 5'-NFRs and 3'-NFRs, have been characterized. In *S. cerevisiae*, these NFRs are typically ~80-300 bp in length and flanked by two well-positioned nucleosomes that often contain the histone variant Htz1 (Albert et al., 2007; Raisner et al., 2005). 5'-NFRs, associated with the promoters of many genes, are highly enriched for sequence-specific transcription factor binding sites (Bernstein et al., 2004; Lee et al., 2007; Whitehouse et al., 2007; Yuan et al., 2005), and demarcate mRNA transcription start sites (TSSs) to their downstream edges (Albert et al., 2007; Lee et al., 2007; Whitehouse et al., 2007). 3'-NFRs are located at the 3'-end of genes and are enriched for transcription termination sites (TTSs) (Mavrich et al., 2008a).

Recent genome-wide expression analyses have demonstrated the majority of eukaryotic genomes are transcribed (Core et al., 2008; David et al., 2006; Davis and Ares, 2006; Dutrow et al., 2008; He et al., 2008; Miura et al., 2006; Nagalakshmi et al., 2008; Neil et al., 2009; Preker et al., 2008; Samanta et al., 2006; Seila et al., 2008; Wilhelm et al., 2008; Xu et al., 2009), resulting in the identification of numerous non-coding RNA (ncRNA) transcripts. In *S. cerevisiae*, many of these ncRNA transcripts

were found to initiate at the upstream edge of 5'-NFRs between tandemly oriented genes or 3'-NFRs (Neil et al., 2009; Xu et al., 2009). However, whether these transcripts are subjected to active regulation is not known. The conserved locations of ncRNA TSSs around NFRs have strongly suggested that NFRs are general locations of transcription initiation and that the mechanisms controlling NFR accessibility are critical to transcriptional regulation of ncRNA.

Multiple factors, including the physical properties of DNA (Anderson and Widom, 2001; Iyer and Struhl, 1995; Kaplan et al., 2009; Mavrich et al., 2008a; Yuan et al., 2005; Zhang et al., 2009), transcription factors (Badis et al., 2008; Hartley and Madhani, 2009), and chromatin regulators (Badis et al., 2008; Hartley and Madhani, 2009), are known to positively regulate the formation and/or size of NFRs *in vivo*. The activities of these factors to establish larger NFRs are thought to facilitate the initiation of transcription by allowing transcription factors greater access to DNA. Whether there are mechanisms to negatively regulate the size of NFRs *in vivo* is not known. However, we have recently shown in *S. cerevisiae* that the ATP-dependent chromatin remodeling enzyme Isw2 functions at the 5'- and 3'-end of genes to increase nucleosome occupancy within intergenic regions by sliding nucleosomes away from coding regions. Interestingly, Isw2 was also required to repress non-coding antisense transcripts from the 3'-end of three genes tested (Whitehouse et al., 2007). Whether Isw2-dependent chromatin remodeling generally affects chromatin structure and ncRNA transcription around NFRs has not been established. We hypothesized that Isw2 may generally function to repress ncRNA transcription by negatively regulating the size of NFRs *in vivo*.

To test this model, we first analyzed data from multiple nucleosome mapping studies to systematically annotate a consensus set of NFRs across the *S. cerevisiae* genome. Our work identified two additional classes of NFRs apart from 5'- and 3'-NFRs that were located within ORFs (ORF-NFRs) and far from ORFs (Other-NFRs). Isw2 targets were found significantly enriched at all classes of NFRs, thus identifying a previously unknown target of Isw2, ORF-NFRs. In addition, we employed custom strand-specific tiled microarrays to analyze ncRNA transcripts and found that Isw2 is globally required to repress initiation of cryptic RNA transcripts from NFRs by sliding nucleosomes toward NFRs to restrict their size. Finally, we provide evidence that a potential biological function for Isw2-dependent repression of some cryptic transcripts is to prevent transcriptional interference. To our knowledge, this is the first example in which the negative regulation of NFR size by a chromatin remodeling enzyme is actively required to repress transcription of ncRNA from NFRs.

Results

Genome-Wide Annotation of NFRs

It is generally accepted that mRNA transcription initiates at the edges of NFRs and can be facilitated by multiple factors that positively regulate the formation and/or size of NFRs. Recent studies have shown that ncRNA also initiates at the edges of NFRs (Neil et al., 2009; Xu et al., 2009). However the mechanisms for regulation of these transcripts are not well understood. Given the widespread prevalence of ncRNA detected throughout the *S. cerevisiae* genome (David et al., 2006; Davis and Ares,

2006; Miura et al., 2006; Nagalakshmi et al., 2008; Neil et al., 2009; Xu et al., 2009), it is critical to understand the mechanisms for regulation of ncRNAs. Isw2 is known to target the 3'-end of ~250 genes and to repress non-coding transcription at three of these loci tested (Whitehouse et al., 2007). These results suggest a possibility that Isw2 may be preferentially targeted to NFRs where it functions as a unique chromatin regulator to repress ncRNA transcription by negatively restricting the size of NFRs. To test our model, we first systematically annotated NFRs genome-wide using data from multiple independent nucleosome mapping datasets.

Global identification of NFRs was first done at the 5'-end of genes by Yuan et al (Yuan et al., 2005). Subsequently, numerous studies recapitulated this finding and further showed NFRs are also commonly present at the 3'-end of genes (Albert et al., 2007; Jiang and Pugh, 2009a; Mavrich et al., 2008a; Neil et al., 2009; Sun et al., 2009; Xu et al., 2009). Generally, these studies based their analyses on a single nucleosome mapping dataset and identified NFRs as regions with a wider than average linker length or by fitting an idealized model to smoothed nucleosome signals. However, all nucleosome mapping datasets contain a large number of loci where nucleosomes are poorly defined, and while the overall agreement in nucleosome positions between studies is good, significant differences between all studies exist (Jiang and Pugh, 2009a) (Figure 4A and B). Furthermore, some NFRs were identified only by their proximity to the ends of genes, thus precluding the annotation of NFRs located elsewhere in the genome. These reasons thus necessitated an unbiased and systematic annotation of NFRs.

Therefore, we developed an algorithm using multiple criteria to systematically identify NFRs commonly found in multiple nucleosome mapping datasets (see Materials and Methods for details) (Figure 5). These criteria were designed to eliminate NFRs from regions with poorly defined nucleosomes and identify NFRs present in several nucleosome mapping datasets, thus mitigating the identification of spurious NFRs and enriching for NFRs with high-confidence (Figure 6). This algorithm was applied to four independent nucleosome mapping datasets, two using high-resolution microarrays and two using high-throughput sequencing (Field et al., 2008; Lee et al., 2007; Mavrich et al., 2008a; Whitehouse et al., 2007). This algorithm identified a reference set of 6589 core NFRs (Figure 6A). Because the annotated NFRs represent the overlapping regions identified from each dataset and not necessarily the definitive edges of each NFR, the mean length and total genomic coverage (99 bp and ~5.4%, respectively) are likely underestimated.

It should be noted that during the course of this study, Jiang and Pugh (Jiang and Pugh, 2009a) independently annotated NFRs genome-wide. In their study, NFRs were annotated based on the length of the linker regions from a compiled reference set of nucleosome positions, resulting in the identification of 14,467 NFRs. A direct comparison of both sets of NFRs revealed a statistically significant overlap (4548 overlapping NFR regions, p-value<10e-300). However, many of the NFRs identified by Jiang and Pugh and not in our dataset correspond to regions of poorly defined nucleosomes or regions deficient of nucleosomes in only a single dataset. Thus for the purpose of this study, NFRs defined by our algorithm were more suitable.

Classification of NFRs

NFRs were first identified around the TSSs at the 5'-end of genes and, more recently, around TTSs at the 3'-end of genes. Of the annotated NFRs, a total of 3127 (averaging ~111 bp in size and associated with ~64% of non-dubious ORFs) and 2440 (averaging ~109 bp in size and associated with ~50% of non-dubious ORFs) were classified as 5'-NFRs and 3'-NFRs, respectively (Figure 6B). Because of the compact nature of the yeast genome, some 5'- and 3'-NFRs are shared between two neighboring genes in tandem, divergent, or convergent orientations. Therefore, we further classified 5'-and 3'-NFRs and identified 1312, 555, and 484 shared tandem, divergent, or convergent NFRs, respectively (Figure 6C). Thus 2438 (~42%) non-dubious genes contain a shared NFR at their 5'-end and 2291 (~40%) contain a shared NFR at their 3'-end.

Our unbiased genome-wide annotation of NFRs allowed us to further identify two additional classes of NFRs, ORF-NFRs and Other-NFRs (Figure 6B). The ORF-NFR class comprises 2114 long linkers (averaging ~91 bp in size) that are located completely within the open reading frame of 1639 genes. The remaining 758 NFRs that are not associated with any non-dubious ORF were classified as Other-NFRs (averaging ~88 bp in size) and represent NFRs located within long intergenic regions or upstream of tRNA genes and retrotransposons.

General Properties of NFRs

Previously, others have shown that nucleosomes immediately adjacent to 5'- and 3'-NFRs are well-positioned and that the phasing of nucleosomes progressively

decreases with distance from NFRs (Mavrich et al., 2008a). More recently, NFRs were also shown to have a sequence-intrinsic tendency to exclude nucleosomes (Kaplan et al., 2009; Zhang et al., 2009). Whether these properties are shared by ORF- or Other-NFRs is not known. We therefore compared the phasing and sequence-dependent exclusion of nucleosomes around each class of NFR.

To this end, the nucleosome signals from two *in vivo* wild-type (WT) and one *in vitro*-reconstituted nucleosome mapping datasets were analyzed around all NFRs (Figure 7A). As expected, the *in vivo* nucleosome maps display a prominent NFR that is flanked on both sides by an array of well-positioned nucleosomes whose phasing progressively decreases with distance from the NFR midpoint. The *in vitro*-reconstituted nucleosomes also exhibit a general depletion of nucleosomes at NFRs (Figure 7A).

Next, we individually compared the average *in vivo* and *in vitro* nucleosome profiles around each class of NFR (Figure 7B). This analysis revealed that all classes of NFRs are flanked by an array of highly positioned nucleosomes with phasing that progressively decreases with distance from the NFR *in vivo*. *In vitro*, all classes of NFRs exhibit a sequence-intrinsic property to exclude nucleosomes. However, the level of phasing or sequence-intrinsic exclusion of nucleosomes varies between classes. For example, 5'- and 3'-NFRs are generally bordered by the most highly phased nucleosomes and display the most prominent *in vitro* nucleosome exclusion, while ORF- and Other-NFRs have significantly reduced levels of phasing and sequence-intrinsic nucleosome exclusion.

Because the *in vivo* and *in vitro* nucleosome profiles at each class of NFRs are different, we speculated additional factors, beyond the sequence-dependent exclusion of nucleosomes, are required to establish the *in vivo* nucleosome architecture surrounding NFRs. As such, we analyzed the distribution of various histone modifications, transcriptional machineries, and chromatin remodeling enzymes surrounding all NFRs. k-means clustering of this profile revealed distinct classes of NFRs that are enriched within different chromatin environments (Figure 8). This result is consistent with our model that the chromatin architecture surrounding distinct classes of NFRs is differentially influenced by chromatin and transcription regulators *in vivo*. The observation that the annotated NFRs define boundaries for some histone modifications, such as H3K4me3, H4ac, and H3K14ac within clusters 1 and 8, as well as a histone variant, Htz1 within clusters 1, 4, 5, 6, and 8, provides additional support of the quality of the NFR annotation.

Isw2 Association with NFRs

We next examined whether Isw2 is preferentially targeted to and functions around NFRs. To address these possibilities, we determined the total number of Isw2 targets, defined as regions where both enrichment of Isw2-ChIP signals and Isw2-dependent chromatin remodeling (Whitehouse et al., 2007) takes place, that are in close proximity to one of the annotated NFRs. Because we utilized a strict definition for Isw2 targets and restricted our analysis to the edges of the annotated core NFRs, our calculation for the association of Isw2 targets with NFRs is likely a significant underestimation. Nonetheless, our analysis revealed a striking association of Isw2

targets with NFRs. A total of 406 Isw2 targets (p-value=9.5e-279), representing ~48% of all Isw2 targets, were found within a single nucleosome distance (150 bp) of an annotated NFR. Classification of these NFRs targeted by Isw2 shows a statistical enrichment in all classes of NFRs: 5'- (217 Isw2 targets, p-value=1.3e-148), 3'- (164 Isw2 targets, p-value=2.1e-108), ORF- (50 Isw2 targets, p-value=9.0e-14) and Other-NFRs (124 Isw2 targets, p-value=4.1e-129). This analysis identified ORF-NFRs as a previously unknown class of Isw2 targets located within genes.

To further understand a role for Isw2 targeting NFRs, we first used self-organizing maps (SOMs) of Isw2-ChIP signals and Isw2-dependent chromatin remodeling around each class of Isw2 target NFR (Figure 9A to D, top panels). This analysis revealed that at each class of NFR, Isw2 tends to be enriched and remodels nucleosomes immediately adjacent to NFRs, consistent with the fact Isw2 preferentially targets NFRs. Second, plotting the distribution of the change in NFR size between WT and *Δisw2* strains (Whitehouse et al., 2007) revealed an increase in the size of many, but not all, NFRs in *Δisw2* strains compared to WT (Figure 9A to D, bottom panels). These results suggest that Isw2-dependent chromatin remodeling around target NFRs functions to restrict the size of many, but not all, target NFRs.

Isw2-Dependent Repression of Cryptic Transcripts

We next sought to determine a functional consequence for Isw2-dependent chromatin remodeling around target NFRs. We have previously shown that Isw2 represses cryptic antisense transcripts from the 3'-end of three genes (Whitehouse et al., 2007). However, whether Isw2 generally functions to repress cryptic RNA

transcription is currently unknown. Given that Isw2 is targeted to and remodels chromatin around NFRs and that the TSSs of ncRNA are enriched around the edges of NFRs, we speculated that Isw2-dependent chromatin remodeling may result in repression of ncRNA transcription at NFRs.

To test our model, we hybridized total RNA from Isw2 deletion strains to high-resolution, strand-specific microarrays tiling chromosomes III, VI, and XII, covering ~14% of the genome. Because the exosome complex efficiently degrades cryptic transcripts, deletion of components in the exosome pathway, either TRF4 or RRP6, in combination with ISW2 (*Δisw2 trf4* and *Δisw2 rrp6*, respectively) is required to stabilize some cryptic transcripts (Whitehouse et al., 2007). These mutations are not expected to alter the frequency of ncRNA transcription (Houseley et al., 2006; Houseley and Tollervey, 2008, 2009; Schmid and Jensen, 2008) but are required for detection of cryptic RNA transcripts. Furthermore, because it is unclear how cryptic RNA transcripts are processed *in vivo*, especially in the absence of Trf4 or Rrp6, we avoided any selection or amplification of RNA (see Materials and Methods for details). To this end, isolated RNA was reverse transcribed, labeled with Cy5, and competitively hybridized against Cy3 labeled genomic DNA. The log2 ratio of Cy5 to Cy3 hybridization signals was then determined for each probe in each experiment. To determine regions with statistically different expression patterns between strains, probes whose signals were significantly (p-value≤0.05) different between strains (*Δisw2* vs. WT, *Δtrf4* vs. WT, *Δrrp6* vs. WT, *Δisw2 trf4* vs. *Δtrf4*, and *Δisw2 rrp6* vs. *Δrrp6*) were identified. Regions containing multiple probes with significantly increased signals were identified as loci expressing cryptic transcripts (Figure 10). Therefore, transcripts specifically suppressed

by Isw2 were identified by comparing signals from a *Δisw2 trf4* or a *Δisw2 rrp6* strain directly to a *Δtrf4* or a *Δrrp6* strain, respectively (see Materials and Methods for details).

While the *Δisw2* single mutant shows relatively few changes compared to WT (*Δisw2* vs. WT, 35 total Isw2-dependent cryptic transcripts), a total of 80 (mean length of 604 bases) and 141 (mean length of 411 bases) Isw2-dependent cryptic transcripts were identified in the *Δtrf4* (*Δisw2 trf4* vs. *Δtrf4*) and *Δrrp6* (*Δisw2 rrp6* vs. *Δrrp6*) backgrounds, respectively. Interestingly, a comparison of cryptic transcripts identified in either *Δtrf4* or *Δrrp6* mutants revealed 74% and 95%, respectively, directly overlap (Figure 11A). The same analysis in the double mutants similarly revealed that 74% and 79% of the identified cryptic transcripts overlapped in *Δisw2 trf4* or *Δisw2 rrp6* cells, respectively (Figure 11A). These data suggest that Trf4 and Rrp6 have overlapping but distinct functions in cryptic RNA regulation.

We next classified each identified cryptic transcript as either cryptic sense, antisense, or other (Figure 11B), with respect to the transcriptional direction of an overlapping ORF (mean overlap with OFRs is 491 bp). The major class of Isw2-repressed cryptic transcripts in both *Δtrf4* and *Δrrp6* backgrounds are cryptic antisense, representing 54% (43 transcripts) and 34% (48 transcripts) of identified cryptic transcripts, respectively (Figure 11B). A comparison of the cryptic RNA levels between each class of transcripts further revealed differences between cryptic sense and antisense transcripts (Fig. 12 to 16). For example, as expected, the exosome components have little contribution to Isw2-repressed cryptic sense transcripts, as shown by the same RNA levels in *Δisw2* and *Δisw2 trf4* or *Δisw2 rrp6* strains (compare red and purple cryptic sense RNA profiles in Figure 14 or 16). In contrast, Isw2-

repressed cryptic antisense transcripts generally require loss of both Isw2 and an exosome component for maximum derepression (compare red and purple cryptic antisense RNA profiles in Figure 14 or 16), consistent with the known role of Trf4 and Rrp6 in the selective degradation of cryptic antisense transcripts.

We next examined the relationships between Isw2-dependent chromatin remodeling and cryptic RNA transcription. We found 22 (28%) and 41 (29%) Isw2-repressed cryptic RNA TSSs in *Δisw2 trf4* and *Δisw2 rrp6*, respectively, are located within 300 bp of an Isw2 target. It should be noted that this is likely an underestimation, as many cryptic transcripts have multiple TSSs, resulting in blurring of microarray signals around the TSSs (see below). Strikingly, at *Δisw2 trf4* loci nucleosomes are preferentially shifted downstream of the cryptic TSSs in the absence of Isw2, compared to WT (Figure 17). In contrast, at *Δisw2 rrp6* loci nucleosome positions appeared relatively unchanged compared to WT (Figure 17). For unknown reasons, nucleosomes were found to be more highly phased around cryptic TSSs in *Δisw2 trf4* compared to *Δisw2 rrp6* (Figure 17). It is possible that this phasing masks the shift in nucleosome positions observed in *Δisw2 rrp6* at individual loci and is also the cause for the lower level of cryptic transcription observed at *Δisw2 rrp6* loci compared to *Δisw2 trf4* loci (Figure 17). In any case, these results show that Isw2-dependent chromatin remodeling is often, but not always, associated with the repression of cryptic transcripts. Finally, an increased level of cryptic transcripts in *Δisw2 trf4* is also observed in non-Isw2 targets (Figure 17). However, the fold change in transcription, from the baseline to the peak of the cryptic transcript, is lower in non-targets than in targets. These data suggests that

indirect effects of Isw2 on ncRNA transcription at non-targets do exist, but these effects tend to be smaller than the direct effects at Isw2 targets.

5'-RACE was then performed at select loci to verify the cryptic transcripts identified by microarray analysis and to map TSS locations (Figure 18). All three loci tested from a *Δisw2 trf4* strain revealed capped transcripts with multiple TSSs that initiated from the edges of a NFR (within 150 bp) specifically targeted by Isw2. This data is consistent with previous studies mapping the TSSs of cryptic transcripts (Neil et al., 2009; Xu et al., 2009) and further confirms the role of Isw2-dependent chromatin remodeling in the repression of cryptic ncRNA transcription around NFRs.

Transcriptional Interference

There are several potential biological reasons for the degradation or repression of cryptic transcripts by the exosome or Isw2. These include, but are not limited to, conservation of resources for transcription and translation, prevention of abnormal protein synthesis, and alleviation of transcriptional interference. In particular, recent reports in *S. cerevisiae* have demonstrated examples of transcriptional interference in the regulation of coding transcription (Hongay et al., 2006; Martens et al., 2004; Martens et al., 2005). It is currently unknown how frequently transcriptional interference occurs on a global scale. Thus to identify a potential biological role for the repression of ncRNA by Isw2 or the exosome, we identified genes whose mRNA levels are decreased when the levels of directly overlapping cryptic sense (sense-sense transcriptional interference) or antisense (antisense-sense transcriptional interference) transcripts are increased in each mutant strain analyzed.

Potential antisense-sense transcriptional interference loci were identified in all strains analyzed, totaling 1 locus in *Δisw2* vs. WT, 36 loci in *Δtrf4* vs. WT, 19 loci in *Δrrp6* vs. WT, 5 loci in *Δisw2 trf4* vs. *Δtrf4*, and 3 loci in *Δisw2 rrp6* vs. *Δrrp6* (one example locus is displayed in Figure 19A and B). In contrast, potential sense-sense transcriptional interference loci were less frequent, totaling 9 loci in *Δtrf4* vs. WT and 1 locus in *Δisw2Δrrp6* vs. *Δrrp6* (one example locus is displayed in Figure 19C and D). Closer inspection revealed that many of the increased cryptic transcripts directly overlap the TSS of the repressed mRNA, totaling 1 locus in *Δisw2* vs. WT, 12 loci in *Δtrf4* vs. WT, 5 loci in *Δrrp6* vs. WT, 3 loci in *Δisw2 trf4* vs. *Δtrf4*, and 1 locus in *Δisw2 rrp6* vs. *Δrrp6* for antisense-sense transcriptional interference and 9 loci in *Δtrf4* vs. WT and 1 locus in *Δisw2 rrp6* vs. *Δrrp6* for sense-sense transcriptional interference. These data suggest the possibility that transcriptional interference may be caused by cryptic RNA transcribing through the promoter of a gene, underscoring the importance of Isw2 and the exosome in the repression of ncRNA. Considering our microarrays represent ~14% of the genome, both sense-sense and antisense-sense transcriptional interference likely both occur at a high frequency throughout the genome. For example, assuming that chromosomes III, VI, and XII accurately represent the global picture, we estimate that ~60 sense-sense and ~250 antisense-sense transcriptional interference loci are present genome-wide in a Δtrf4 mutant alone.

Discussion

Genome-Wide Annotation of NFRs

In this manuscript we annotated a consensus set of core NFRs across the yeast genome. This led to the identification of four distinct classes of NFRs, located at the 5'-end of genes (5'-NFRs), 3'-end of genes (3'-NFRs), within ORFs (ORF-NFRs), or far from ORFs (Other-NFRs). This annotation allowed the direct comparison of each class of NFRs. Examination of nucleosome positions surrounding NFRs have shown that all classes of NFRs have a sequence-intrinsic tendency to exclude nucleosomes *in vitro* and are generally flanked by an array of highly-positioned nucleosomes whose phasing progressively decreases with distance from the NFR *in vivo*. Additionally, k-means clustering of the distribution of chromatin and transcriptional regulators around all NFRs revealed that distinct classes of NFRs are surrounded by different chromatin environments *in vivo* (Fig. 8). Enrichment of TSSs around the edges of 5'- NFRs and TTSs toward the middle of 3'-NFRs has led to speculation that these NFRs play crucial roles in transcriptional regulation. In contrast, the functions of ORF- and Other-NFRs are not known. We speculate that similar to 5'- and 3'-NFRs, ORF- and Other-NFRs may become sites of transcription initiation and/or termination under certain conditions. In fact, several recent studies have shown that deletion of a number of chromatin and transcription regulators, including Rpd3, Set2, Spt6, and Spt16, or a nutritional shift of WT cells, leads to initiation of cryptic RNA transcription within ORFs (Carrozza et al., 2005; Cheung et al., 2008; Kaplan et al., 2009). While the TSS of these transcripts have not been mapped on a global scale, it is possible that some of these cryptic

transcripts initiate at the edges of ORF-NFRs. Thus our annotation of NFRs lays the foundation for better understanding the relationship between chromatin architecture and any DNA-dependent processes that are affected by NFRs.

The yeast genome is very compact, with intergenic regions averaging ~500 bp. As a result, promoters and terminators are generally within close proximity. Consistently, we found that a significant fraction of tandem, divergent, and convergent gene pairs contain a shared NFR within their transcriptional regulatory regions at the 5'- or 3'- end of the gene. How cells prevent RNA polymerases from colliding or interfering with each other at shared NFRs is unknown. We speculate that if the genes associated with shared NFRs were transcribed at different times, it would mitigate the likelihood of colliding or interfering RNA polymerases. Alternatively, it is possible that the collision of RNA polymerases is not a frequent event and does not pose significant problems for cells.

In Vivo Functions of Isw2

In this study, we show Isw2 targets are significantly enriched at all classes of NFRs and identify a novel class of Isw2 targets within ORFs (ORF-NFRs). Strikingly, at all classes of NFRs targeted by Isw2, Isw2-dependent chromatin remodeling was found to restrict the size of many, but not all, NFRs by sliding nucleosomes toward the middle of NFRs. To our knowledge, this is the first example in which a functional role for a chromatin remodeling enzyme has been shown to decrease the size of NFRs *in vivo*. For example, previous reports have demonstrated that the RSC complex is required at a subset of gene promoters, or 5'-NFRs, to exclude nucleosomes and increase the size

of NFRs (Badis et al., 2008; Hartley and Madhani, 2009). These results demonstrate that the accessibility of NFRs is under dynamic control by multiple chromatin regulators, which is consistent with the idea that NFRs play highly important roles *in vivo*.

It was recently found that ncRNA is transcribed throughout the *S. cerevisiae* genome (David et al., 2006; Davis and Ares, 2006; Miura et al., 2006; Nagalakshmi et al., 2008; Neil et al., 2009; Xu et al., 2009). However, how these transcripts are regulated is not well understood. Several studies have shown that a number of other chromatin regulators and transcription factors appear to function co-transcriptionally with the elongating RNA polymerase to repress cryptic transcripts from initiating within ORFs (Carrozza et al., 2005; Cheung et al., 2008; Kaplan et al., 2003). In contrast, we have found a large fraction of cryptic transcripts repressed by Isw2 initiate around NFRs and are antisense to known ORFs. These results revealed a mechanism by which transcription of cryptic ncRNA can be repressed by restricting the size of NFRs. In addition, our results and those of others (Carrozza et al., 2005; Cheung et al., 2008; Kaplan et al., 2003) collectively establish that a large number of chromatin regulators are used to control cryptic RNA transcription *in vivo*. This suggests the possibility that there may be a significant number of unidentified mechanisms used to control ncRNA transcription *in vivo*.

Considering the widespread cryptic RNA transcription that occurs in all eukaryotes (Core et al., 2008; David et al., 2006; Davis and Ares, 2006; Dutrow et al., 2008; He et al., 2008; Miura et al., 2006; Nagalakshmi et al., 2008; Neil et al., 2009; Preker et al., 2008; Samanta et al., 2006; Seila et al., 2008; Wilhelm et al., 2008; Xu et al., 2009), it is likely that many of the mechanisms for repressing cryptic RNA

transcription are conserved in eukaryotes. ISWI homologues are necessary for normal development in *D. melanogaster* (Deuring et al., 2000), *C. elegans* (Andersen et al., 2006), *X. laevis* (Dirscherl et al., 2005; Wysocka et al., 2006), and mice (Stopka and Skoultchi, 2003). In fact, the ISWI family is known to play key roles in a variety of essential biological processes, including transcription (Andersen et al., 2006; Badenhorst et al., 2002; Deuring et al., 2000; Fazzio et al., 2005; Fazzio et al., 2001; Goldmark et al., 2000; Kent et al., 2001; Lazzaro et al., 2006; Whitehouse et al., 2007; Yasui et al., 2002), global chromatin structure (Deuring et al., 2000), DNA replication (Collins et al., 2002; Poot et al., 2004; Vincent et al., 2008), cell cycle progression (Fyodorov et al., 2004), ribosomal DNA silencing (Zhou and Grummt, 2005; Zhou et al., 2002), and cohesin loading (Hakimi et al., 2002). However, the underlying mechanisms by which these ISWI homologues are required for these processes are not fully understood. Because the subunits of the ISWI complex are highly conserved (Flaus et al., 2006), we speculate that ISWI may be required to repress cryptic RNA transcription in other eukaryotes, similar to Isw2 in *S. cerevisiae*. If this is the case, mis-regulation of cryptic RNA transcripts caused by loss of ISWI could trigger widespread RNA interference (RNAi), which might lead to the observed developmental defects. While *S. cerevisiae* have no functional RNAi pathway, our work revealed a large number of loci potentially exhibiting transcriptional interference upon loss of Isw2. These results underscore the importance for future investigations into Isw2 function and establish this remodeling enzyme as a model for elucidating the functions and regulatory mechanisms of cryptic RNA transcription and the modulation of chromatin structure.

Materials and Methods

Nucleosome Positions

Nucleosome positions used for NFR annotations are as follows:

(Whitehouse et al., 2007): Pearson correlation coefficient (r-value) of ≥0.5.

(Lee et al., 2007): Positions were empirically determined as previously described (Whitehouse et al., 2007) using the normalized log2 microarray signal and only nucleosomes with a Pearson correlation coefficient (r-value) of ≥0.5.

(Mavrich et al., 2008a): A read count of ≥3 from the combined Watson and Crick data.

(Field et al., 2008): All uniquely mapped reads.

NFR Annotation

For (Whitehouse et al., 2007), (Lee et al., 2007), and (Mavrich et al., 2008a) a Gaussian distribution of linker lengths was fit to the frequency of linker lengths from each data set. Linkers with lengths less than or equal to the average plus two standard deviations were discarded. NFRs were then defined as those remaining linkers meeting at least one of the following criteria: 1) the average Pearson correlation coefficient (r-value) (Lee et al., 2007; Whitehouse et al., 2007) or standard deviation (Mavrich et al., 2008a) of the six surrounding nucleosomes (three immediately flanking nucleosomes on either side of the linker) is greater than or equal to the genome-wide average of all linkers from each dataset, 2) both of the immediately adjacent linkers (one on either side) have a length that was less than or equal to the average plus two standard deviations, or 3) the midpoint of a Htz1 containing nucleosome (rank order of greater

than or equal to three from (Albert et al., 2007)) is present anywhere within one nucleosome distance (150 bp) upstream or downstream of the boundaries of the NFR. For (Field et al., 2008) a Gaussian distribution of linker lengths was fit to the frequency of region lengths sequenced fewer than twice. All linkers with lengths less than or equal to the average plus two standard deviations were discarded. All remaining linkers with at least one nucleosome distance (150 bp) flanking one side that was sequenced at least twice were annotated as NFRs. The regions identified as NFRs in at least three of the four data sets were annotated as core NFRs. NFRs separated by less than one nucleosome distance (150 bp) were combined into a single NFR. Annotated NFR locations are available for download at http://labs.fhcrc.org/tsukiyama.

Classification of NFRs

Core NFRs were classified as 5'-NFRs, 3'-NFRs, ORF-NFRs, or Other-NFRs based on their location with respect to all non-dubious ORFs. 5'- and 3'-NFRs were required to be within a single nucleosome distance (150 bp) of a non-dubious ORF's TSS or TTS, respectively, as experimentally determined by (Nagalakshmi et al., 2008). If no experimentally defined TSS or TTS was available for any particular ORF, the average length of all experimentally determined 5'- or 3'-UTRs (82 and 135 bp, respectively) was added to the translation start or termination site, respectively. NFRs contained completely within the coding region of a non-dubious ORF were classified as ORF-NFRs. A single NFR may meet more than one criteria listed above and will thus have multiple classifications (5'-, 3'-, or ORF-NFR). NFRs not classified as 5'-, 3'-, or ORF-NFR were classified as Other-NFRs. Tandem, divergent, and convergent NFRs

were defined as individual NFRs classified as both a 5'- and 3'-NFR of tandemly oriented gene pairs, a 5'-NFR of two divergently oriented gene pairs, or a 3'-NFR of two convergently oriented gene pairs, respectively. Location of annotated shared NFRs are available for download at http://labs.fhcrc.org/tsukiyama.

Genomic DNA Isolation, Fragmentation, and Labeling

Genomic DNA was isolated from WT cells with Qiagen Genomic DNA columns according to the manufacturer's protocols. Purified DNA was fragmented with DNaseI to an average size ~25-75 bp. Fragmented DNA (30 ng) was heated at 70°C for 5 min. Then, 5 µl Buffer #4 (NEB), 5 µl of 2.5 mM CoCl2 (NEB), 5 µl of 1 mM Cy3-dUTP (GE Healthcare), and 5 µl of 20,000 U ml-1 terminal deoxytransferase (NEB) were added to a total volume of 50 µl. The reaction was incubated at 37°C for 3 hours. 50 µl H_2O was added and labeled DNA was purified with a gel filtration spin column and ethanol precipitated. The pellet was suspended in a final volume of 50 µl H_2O.

RNA Isolation, Fragmentation, and cDNA Labeling

Yeast cells were grown to an OD660=0.70 (+/-0.05) and standard hot acid phenol extraction was used to isolate total RNA. 25 µg RNA and 12 µg random hexamers were placed in 40 µl H_2O. The reaction was incubated at 70°C for 5 min, 25°C for 5 min, and 4°C for 5 min. 0.5 µL of 1 mg ml-1 Actinomycin D, 20 µl of 5X SuperScript III Buffer (Invitrogen), 8 µl of 0.1 M DTT, 1 µl of 40 U µl-1 RNase Inhibitors (Roche), 4 µL of dNTP mix (10 mM dATP, 10 mM dGTP, 10 mM dCTP, 8.5 mM dTTP, and 1.5 mM dUTP), 4 µl of 200 U µl-1 SuperScript III (Invitrogen), and 20 µL H_2O were

added, and the reaction was incubated at 25°C for 10 min, 48°C for 90 min, and 70°C for 10 min. 2 µl of 0.1 mg ml-1 RNaseA and 2 µl of 5,000 U µl-1 RNaseH (NEB) were added and incubated at 37°C for 30 min. The reaction was stopped by phenol:chloroform extraction and unincorporated random hexamers were removed with a gel filtration spin column. cDNA was ethanol precipitated and the pellet was suspended in 80 µl H_2O.

cDNA was fragmented with 5 µl of 10 U µl-1 APE1 (NEB), 5 µl of 2 U µl-1 uracil deglycosylase (NEB), and 10 µl Buffer #4 (NEB) at 37°C for 60 min. The reaction was stopped by phenol:chloroform extraction and was purified with a gel filtration spin column. The fragmented cDNA was ethanol precipitated and the pellet was suspended in 20 µl H_2O.

10 µg fragmented cDNA was heated at 70°C for 5 min. 3 µl Buffer #4 (NEB), 3 µl of 2.5 mM CoCl2 (NEB), 5 µl of 1 mM Cy5-dUTP (GE Healthcare), and 3 µl of 20,000 U µl-1 terminal deoxytransferase (NEB) were added to the cDNA to a total volume of 30 µl. The reaction was incubated at 37°C for 3 hours. 50 µl H_2O was added and the reaction was purified with a gel filtration spin column and ethanol precipitated. The pellet was suspended in a final volume of 20 µl H_2O.

Microarray Hybridizations

5 µg labeled cDNA and 2.5 µg labeled genomic DNA were competitively hybridized to custom strand-specific microarrays from NimbleGen. Microarrays tile both strands of chromosomes III, VI, and XII with 50-mer probes overlapping, on average, by ~42 bp with a ~4 bp offset between strands. Log2 ratios (Cy5/Cy3, generated using

NimbleScan software) were biweight mean adjusted. Raw and analyzed microarray hybridization results files are available for download at http://labs.fhcrc.org/tsukiyama. The sequence of probes on microarrays is available upon request.

Identification of Cryptic Transcription Units

To identify probes with significantly different signals between strains, we used LIMMA (Smyth, 2004; Wettenhall and Smyth, 2004), utilizing an object containing the log2 ratio data from all hybridizations as input, to generate p-values (adjusted for multiple testing using the method of Benjamini and Hochberg) for each probe between strains. A 10 probe sliding window (~80 bp) was used to identify regions that had at least 5 probes with a p-value≤0.05. All probes within that window had to be less than or equal to 12 bp apart to account for missing or non-tiled regions of the genome. Each region's start and endpoints were further trimmed so as to begin and end with a statistically significant probe. A Gaussian distribution was fit to a frequency plot of all the distances between adjacent regions. Regions that were closer to or equal in distance than the mean (~60 bp) plus two standard deviations (~30 bp) were combined into a single region. Finally, for each pair-wise comparison, a Gaussian distribution curve was fit to a frequency plot of the lengths of the regions and mean and standard deviations were calculated. Only regions that were greater than or equal to the mean plus two standard deviations were used, thus eliminating transcriptional units that were very small. Annotated cryptic transcriptional units are available for download at http://labs.fhcrc.org/tsukiyama.

Annotation of Transcriptional Interference Loci

Genes containing both a decreased sense coding transcript (mRNA) and an increased cryptic sense or antisense transcript directly overlapping the ORF (defined from the annotated TSS to TTS as described above for NFR annotation) were annotated as sense-sense or antisense-sense transcriptional interference loci. Annotated transcriptional interference loci are available for download at http://labs.fhcrc.org/tsukiyama.

Heat Maps, Graphs, and Screenshots

For all heat maps and average profiles data were aligned as described in Figure Legends and the signals were averaged into 20 bp non-overlapping bins. The log2 signal, taken before averaging and binning, was required for profiles of Swr1 (Venters and Pugh, 2009), Rpb3 (Venters and Pugh, 2009), Rpo21 (Venters and Pugh, 2009), Sua7 (Venters and Pugh, 2009), TBP (Venters and Pugh, 2009), Rsc9 (Venters and Pugh, 2009), and Reb1 (Venters and Pugh, 2009). The log2 signal, taken before averaging and binning, and an 11 bin moving average was required for profiles of H3K9Ac (Pokholok et al., 2005), H3K14Ac (Pokholok et al., 2005), H4Ac (Pokholok et al., 2005), H3K4Me1 (Pokholok et al., 2005), H3K4Me2 (Pokholok et al., 2005), H3K4Me3 (Pokholok et al., 2005), H3K36Me3 (Pokholok et al., 2005), H3K79Me3 (Pokholok et al., 2005), Esa1 (Pokholok et al., 2005), Gcn5 (Pokholok et al., 2005), and Rsc8 (Badis et al., 2008) (3 bin moving average). Additionally, chromosome X was omitted from analysis due to a misalignment of data from Venters and Pugh (Venters and Pugh, 2009). Each transcription factor binding site represents binding p-value<0.05

and conserved in at least one other yeast from Harbison et al 2004 (Harbison et al., 2004).

All screenshots were taken using Integrated Genome Browser (Helt et al., 2009; Nicol et al., 2009).

Microarray Data Accession

All raw and processed microarray data is available for download at Gene Expression Omnibus (GEO), http://www.ncbi.nlm.nih.gov/projects/geo/, using accession number GSE23108.

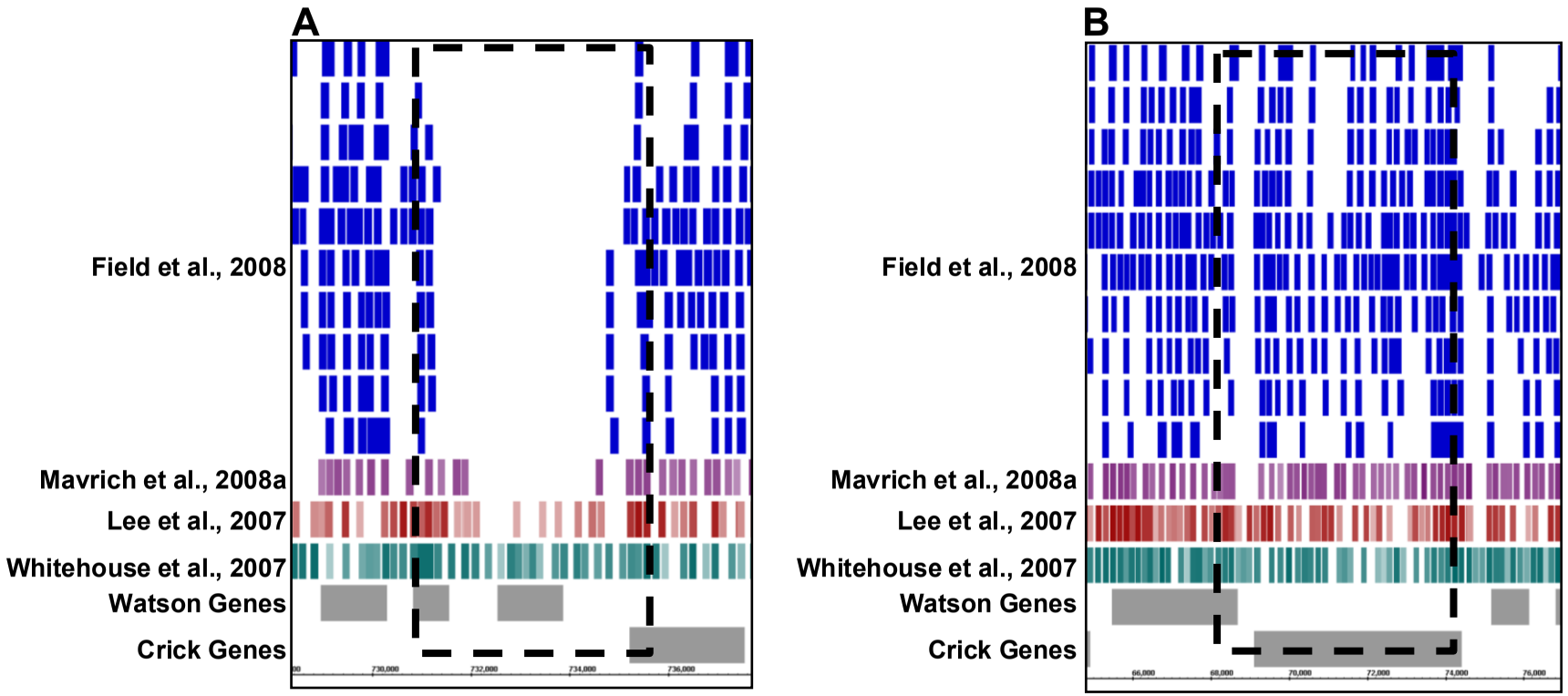
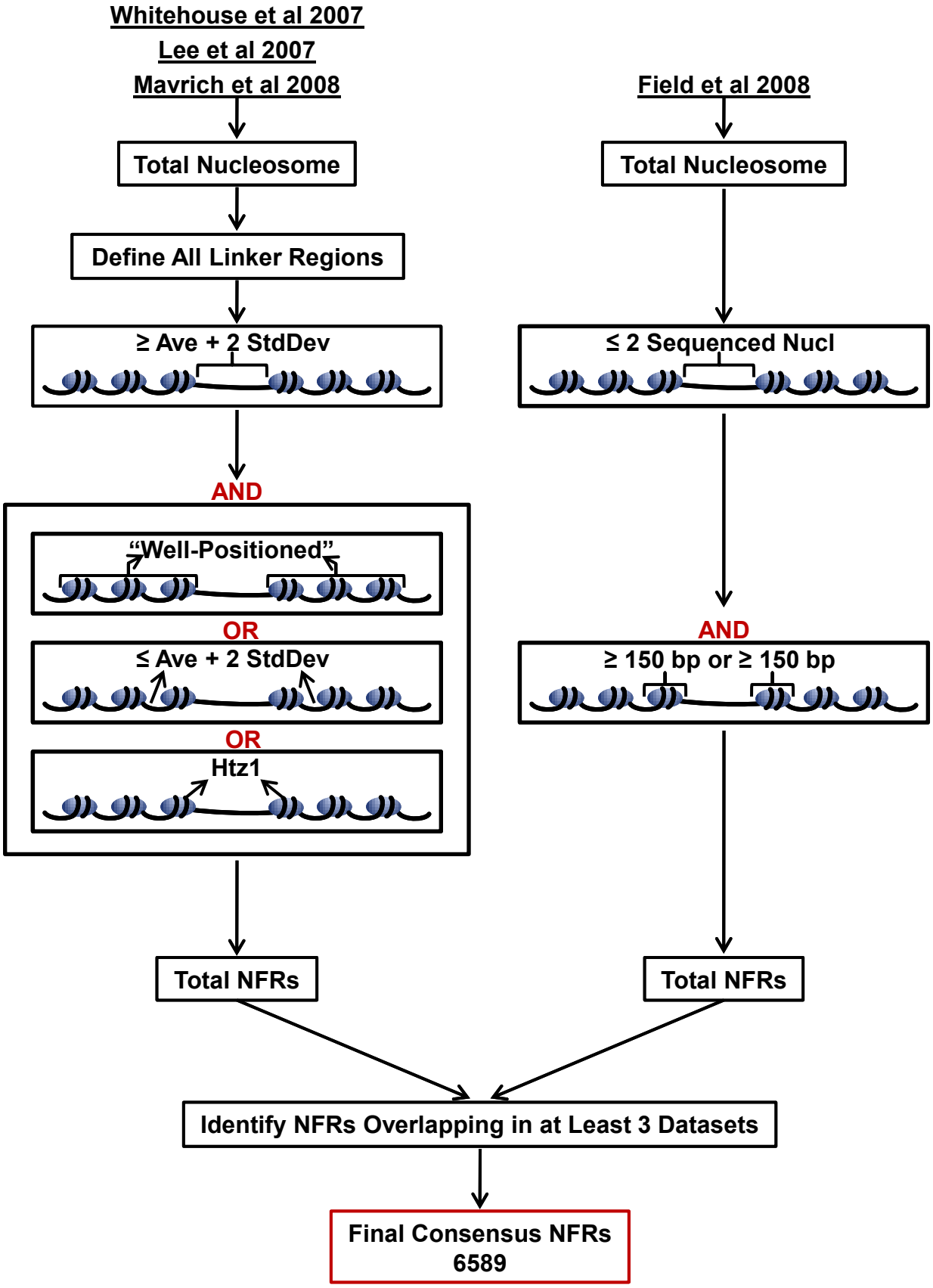


Figure 4: Differences in Nucleosome Positioning Maps
(A and B) Two representative screenshots displaying the mapped nucleosomes (depicted as colored boxes with lighter boxes representing delocalized nucleosomes) from four independent nucleosome mapping studies. Regions of interest are boxed with a dashed line for emphasis. ORFs are displayed as grey boxes.
(A) Chr12: 728,000-738,000
(B) Chr6: 65,000-77,000

Figure 5: NFR Annotation Scheme
A schematic description of our NFR annotation algorithm. Nucleosomes are represented by cyan ovals. For Whitehouse et al., 2007, Lee et al., 2007, and Mavrich et al., 2008a a Gaussian distribution of linker lengths was fit to the frequency of linker lengths from each data set. Linkers with lengths less than or equal to the average plus two standard deviations were discarded. NFRs were then defined as those remaining linkers meeting at least one of the following criteria: 1) the average Pearson correlation coefficient (r-value) (Whitehouse et al. 2007; Lee et al., 2007) or standard deviation (Mavrich et al., 2008a) of the six surrounding nucleosomes (three immediately flanking nucleosomes on either side of the linker) is greater than or equal to the genome-wide average of all linkers from each dataset, 2) both of the immediately adjacent linkers (one on either side) have a length that was less than or equal to the average plus two standard deviations, or 3) the midpoint of a Htz1 containing nucleosome (rank order of greater than or equal to three from Albert et al., 2007) is present anywhere within one nucleosome distance (150 bp) upstream or downstream of the boundaries of the NFR. For Field et al., 2008 a Gaussian distribution of linker lengths was fit to the frequency of region lengths sequenced fewer than twice. All linkers with lengths less than or equal to the average plus two standard deviations were discarded. All remaining linkers with at least one nucleosome distance (150 bp) flanking one side that was sequenced at least twice were annotated as NFRs. The regions identified as NFRs in at least three of the four data sets were annotated as core NFRs. NFRs separated by less than one nucleosome distance (150 bp) were combined into a single NFR. Annotated NFR locations are available for download at http://labs.fhcrc.org/tsukiyama.


Whitehouse et al 2007
Lee et al 2007
Mavrich et al 2008
Total Nucleosome
Define All Linker Regions
≥ Ave + 2 StdDev
AND
"Well-Positioned"
OR
≤ Ave + 2 StdDev
OR
Htz1
Total NFRs
Field et al 2008
Total Nucleosome
≤ 2 Sequenced Nucl
AND
≥ 150 bp or ≥ 150 bp
Total NFRs
Identify NFRs Overlapping in at Least 3 Datasets
Final Consensus NFRs
6589

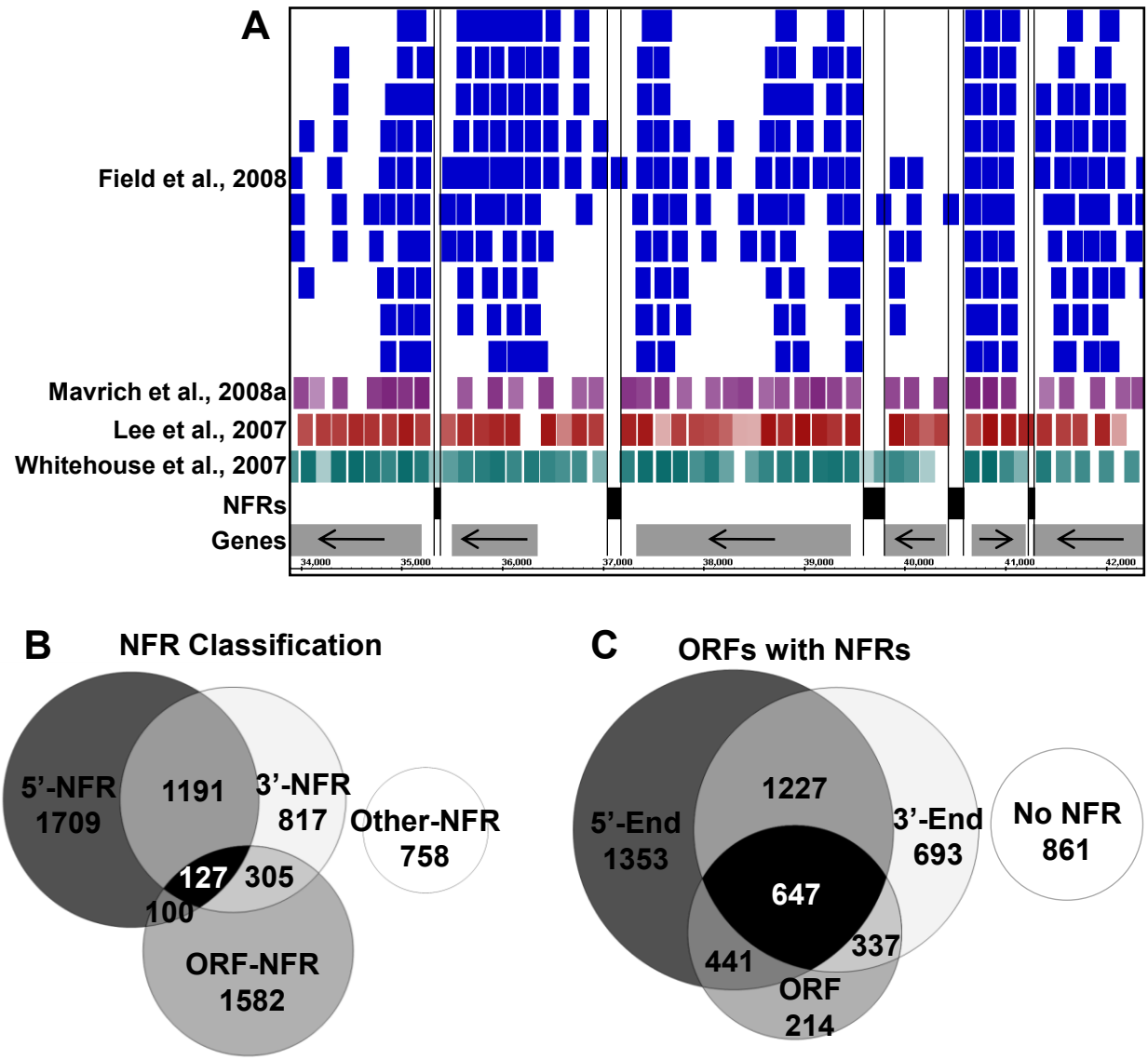


Figure 6: Systematic Annotation of NFRs
(A) Representative screenshot (Chr12: 34,000-42,000) displaying the annotated NFRs with respect to nucleosome positions. NFRs are shown as black boxes and nucleosomes are depicted as colored boxes with lighter boxes representing delocalized nucleosomes (nucleosomes with a lower r-value or higher standard deviation). Genes are shown as grey rectangles with arrows indicating the direction of coding transcription and black horizontal lines represent the boundaries of each core NFR.
(B) Venn diagram displaying the classification, 5'-, 3'-, ORF-, and Other-NFRs (as described in the Materials and Methods), of all 6589 annotated NFRs.
(C) Venn diagram of all 5773 non-dubious ORFs associated with at least one annotated NFR at the 5'-end, 3'-end, or within the middle of the ORF. Overlapping regions indicate ORFs associated with multiple NFRs. No NFR refers to non-dubious ORFs not associated with any NFR.

Figure 7: All Classes of NFRs are Surrounded by Phased Nucleosomes and Have a Sequence-Intrinsic Tendency to Exclude Nucleosomes
(A) Top: Nucleosome signals from two *in vivo* (Field et al., 2008; Whitehouse et al., 2007) and one *in vitro* (Field et al., 2008) nucleosome mapping datasets are shown around the annotated NFRs. NFRs for each dataset are aligned by their midpoint and then sorted by size, the smallest at the top and the largest at the bottom. "Outline" marks the edges of each NFR in black. The nucleosome signals from 500 bp on either side of a NFR midpoint are shown. Yellow denotes high nucleosome occupancy and cyan denotes lack of nucleosomes. Bottom: A graph displaying the average nucleosome signals across each NFR from the datasets presented above. **(B)** NFRs were broken down into 5'-NFRs, 3'-NFRs, ORF-NFRs, and Other-NFRs and the nucleosome signals were averaged for all NFRs within the class. 5'- and 3'-NFRs were realigned to the downstream or upstream edge of each NFR, respectively, such that transcription of the associated coding gene proceeds from left to right. ORF-NFRs are aligned by their midpoint such that transcription of the associated coding gene proceeds from left to right. Other-NFRs are aligned as in (A).

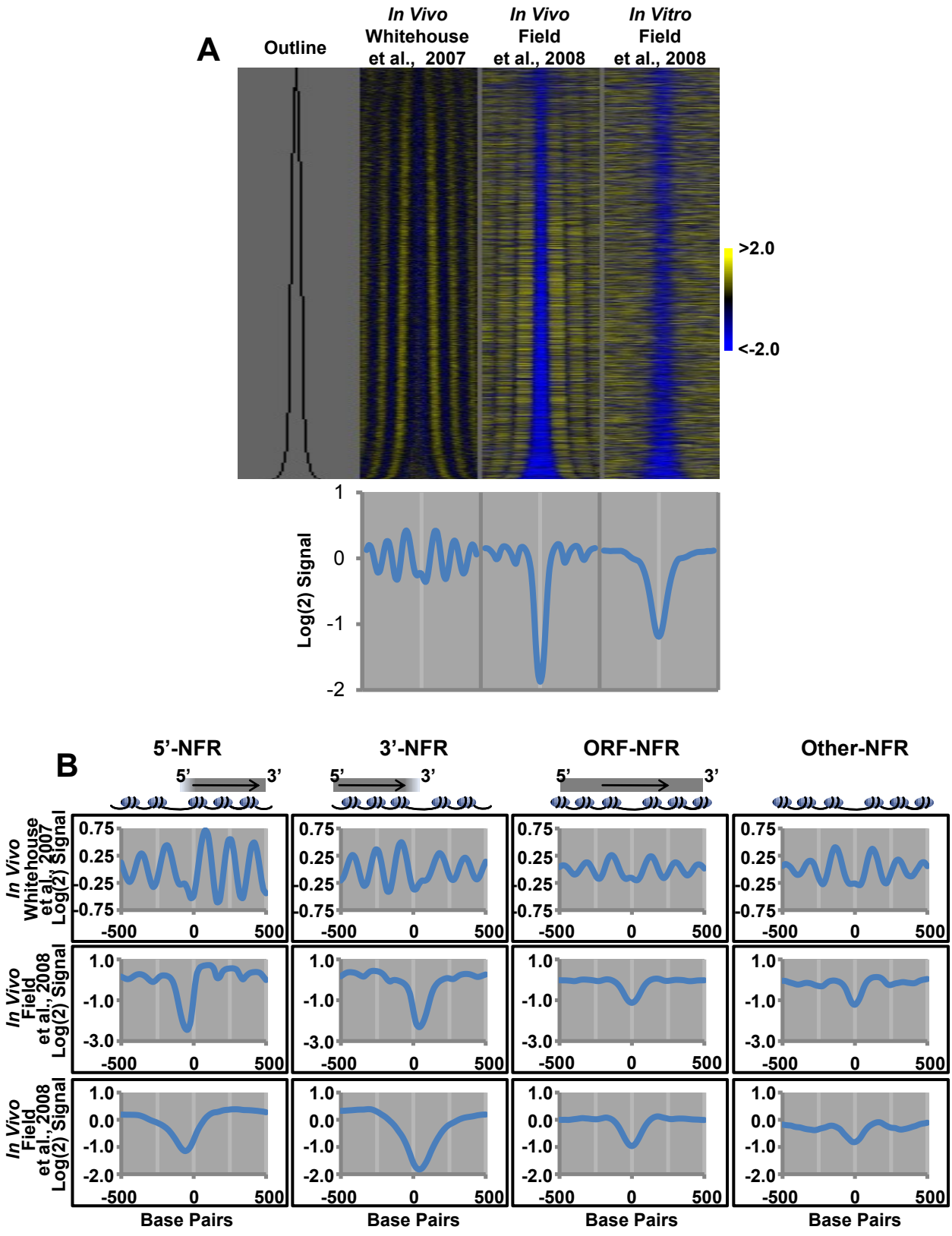

A
Outline
In Vivo
Whitehouse
et al., 2007
In Vivo
Field
et al., 2008
In Vitro
Field
et al., 2008
>2.0
<-2.0
Log(2) Signal
1
0
-1
-2
B
5'-NFR
3'-NFR
ORF-NFR
Other-NFR
5'
3'
In Vivo
Whitehouse
et al., 2007
Log(2) Signal
In Vivo
Field
et al., 2008
Log(2) Signal
In Vivo
Field
et al., 2008
Log(2) Signal
0.75
0.25
-0.25
-0.75
1.0
-1.0
-3.0
0.0
-2.0
-500
0
500
Base Pairs

Figure 8: k-Means Clustering Revealed Distinct Classes of NFRs That Are Enriched Within Different Chromatin Environments
Left: Displays the k-means clustering result (k=10) of nucleosomes, ChIP signals for histone modifications, transcription factors, and chromatin remodeling factors around the annotated NFRs. For each factor, NFRs were aligned by their midpoints, and the signals from 500 bp on either side of a NFR midpoint are shown. Identified clusters are delineated by brackets. Right: The total number of NFRs, and the respective enrichment (p-value, hypergeometric distribution) of NFR classes and subclasses, within each of the ten identified clusters. Yellow indicates a positive enrichment and cyan indicates negative enrichment. ● Nucleosomes, ● histone modifications, ● chromatin modifying, ● general transcription factor, and ● other (transcription factors or sequence motifs)
Outline – marks the edges of the annotated NFRs from this study in black
In Vivo Nucl #1 (Whitehouse et al., 2007) – average log2 signal
In Vivo Nucl #2 (Field et al., 2008) – average normalized log2 signal
In Vitro Nucl (Field et al., 2008) – average log2 signal
In Vivo Nucl #3 (Mavrich et al., 2008a) – annotated sequenced nucleosomes are displayed in black
H2A.Z Nucl (Albert et al., 2007) – annotated sequenced nucleosomes are displayed in black
Swr1 (Venters and Pugh, 2009) – average log2 signal
H3K9Ac (Pokholok et al., 2005) – average log2 signal
H3K14Ac (Pokholok et al., 2005) – average log2 signal
H4Ac (Pokholok et al., 2005) – average log2 signal
H3K4Me1 (Pokholok et al., 2005) – average log2 signal
H3K4Me2 (Pokholok et al., 2005) – average log2 signal
H3K4Me3 (Pokholok et al., 2005) – average log2 signal
H3K36Me3 (Pokholok et al., 2005) – average log2 signal
H3K79Me3 (Pokholok et al., 2005) – average log2 signal
Esa1 (Pokholok et al., 2005) – average log2 signal
Gcn5 (Pokholok et al., 2005) – average log2 signal
Rpb3 (Venters and Pugh, 2009) – average log2 signal
Rpo21 (Venters and Pugh, 2009) – average log2 signal
Sua7 (Venters and Pugh, 2009) – average log2 signal
TBP (Venters and Pugh, 2009) – average log2 signal
Isw2-ChIP (Whitehouse et al., 2007) – average log2 signal
Isw2-Remod (Whitehouse et al., 2007) – average log2 signal
Rsc9 (Venters and Pugh, 2009) – average log2 signal
Rsc8 (Badis et al., 2008) – average log2 signal
Rsc3-1 (Badis et al., 2008) – average log2 signal
Transcription Factors (Harbison et al., 2004) –TF binding sites is displayed in black
Reb1 (Venters and Pugh, 2009) – average log2 signal
Boundary Elements (Field et al., 2008) – strength of nucleosome exclusion sequence

Outline
In Vivo Nucl #1
In Vivo Nucl #2
In Vitro Nucl
In Vivo Nucl #3
H2A.Z Nucl
Swr1
H3K9Ac
H3K14Ac
H4Ac
H3K4Me1
H3K4Me2
H3K4Me3
H3K36Me3
H3K79Me3
Esa1
Gcn5
Rpb3
Rpo21
Sua7
TBP
Isw2-ChIP
Isw2-Remod
Rsc9
Rsc8
Rsc3-1 Remod
Transcription Factors
Reb1
Boundary Elements
<-2.5
>2.5
Total
5'-NFR
3'-NFR
ORF-NFR
Other-NFR
Tandem
Divergent
Convergent
Cluster #
1
2
3
4
5
6
7
8
9
10
648
547
454
457
807
504
498
686
550
1035
426 (1.5E-23)
16 (<-100)
139 (1.7E-14)
416 (2.6E-94)
744 (<-100)
368 (2.9E-34)
287 (4.5E-07)
418 (1.6E-14)
103 (7.2E-49)
16 (<-100)
446 (1.7E-67)
259 (7.0E-08)
108 (1.2E-10)
182 (1.8E-02)
219 (4.1E-11)
141 (1.6E-06)
197 (1.9E-02)
416 (1.8E-40)
196 (2.8E-02)
131 (3.3E-81)
141 (9.9E-10)
324 (2.7E-44)
64 (2.2E-19)
26 (3.2E-45)
59 (1.4E-70)
59 (8.3E-28)
90 (4.2E-13)
157 (1.9E-08)
114 (6.9E-10)
926 (<-100)
19 (7.3E-17)
42 (5.7E-04)
187 (1.5E-65)
17 (6.0E-10)
25 (4.1E-20)
59 (5.7E-02)
80 (3.2E-04)
47 (6.0E-06)
213 (8.3E-70)
22 (1.2E-33)
323 (1.6E-74)
9 (3.2E-42)
22 (1.0E-21)
161 (6.9E-16)
198 (9.4E-05)
90 (2.3E-02)
104 (3.9E-02)
284 (1.8E-43)
37 (1.6E-19)
4 (<-100)
12 (3.2E-14)
0 (4.3E-23)
3 (6.4E-15)
86 (2.8E-13)
304 (<-100)
73 (1.6E-06)
29 (4.3E-03)
19 (4.6E-11)
5 (1.6E-16)
0 (4.5E-45)
57 (2.1E-02)
171 (1.4E-71)
8 (2.6E-08)
1 (6.6E-15)
4 (2.1E-23)
7 (1.9E-10)
24 (4.9E-03)
50 (6.2E-02)
47 (3.5E-02)
86 (2.2E-02)

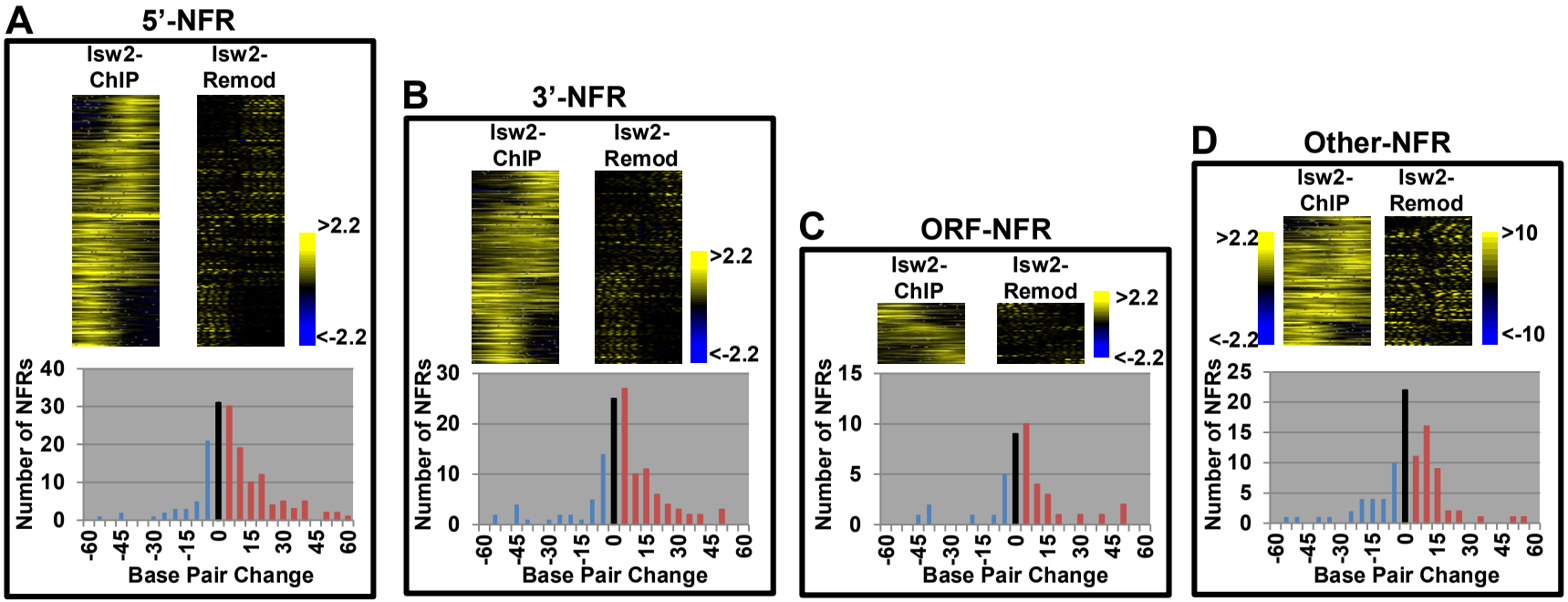


Figure 9: Isw2-Dependent Chromatin Remodeling Functions to Restrict the Size of NFRs

(A to D) Top: NFRs targeted by Isw2 were grouped into 5'-NFRs, 3'-NFRs, ORF-NFRs, or Other-NFRs, and aligned by their midpoint. 500 bp on either side of NFRs are shown. Isw2-ChIP: Self organizing maps (SOMs) of Isw2-ChIP signals (Whitehouse et al., 2007) were used to rearrange NFRs (top to bottom) with similar patterns adjacent to each other. Isw2-Remod: Isw2-Remodeling signals (Whitehouse et al., 2007) are shown for NFRs aligned by SOMs of Isw2 ChIP signals as above. Bottom: The distribution of the change in NFR size between WT and *Δisw2* strains (Whitehouse et al., 2007). Red bars indicate an increase, black bars indicate no change, and cyan bars indicate a decrease in NFR size in *Δisw2* strains compared to WT.

(A) 203 5'-NFRs that are Isw2 targets

(B) 153 3'-NFRs that are Isw2 targets

(C) 51 ORF-NFRs that are Isw2 targets

(D) 114 Other-NFRs that are Isw2 targets

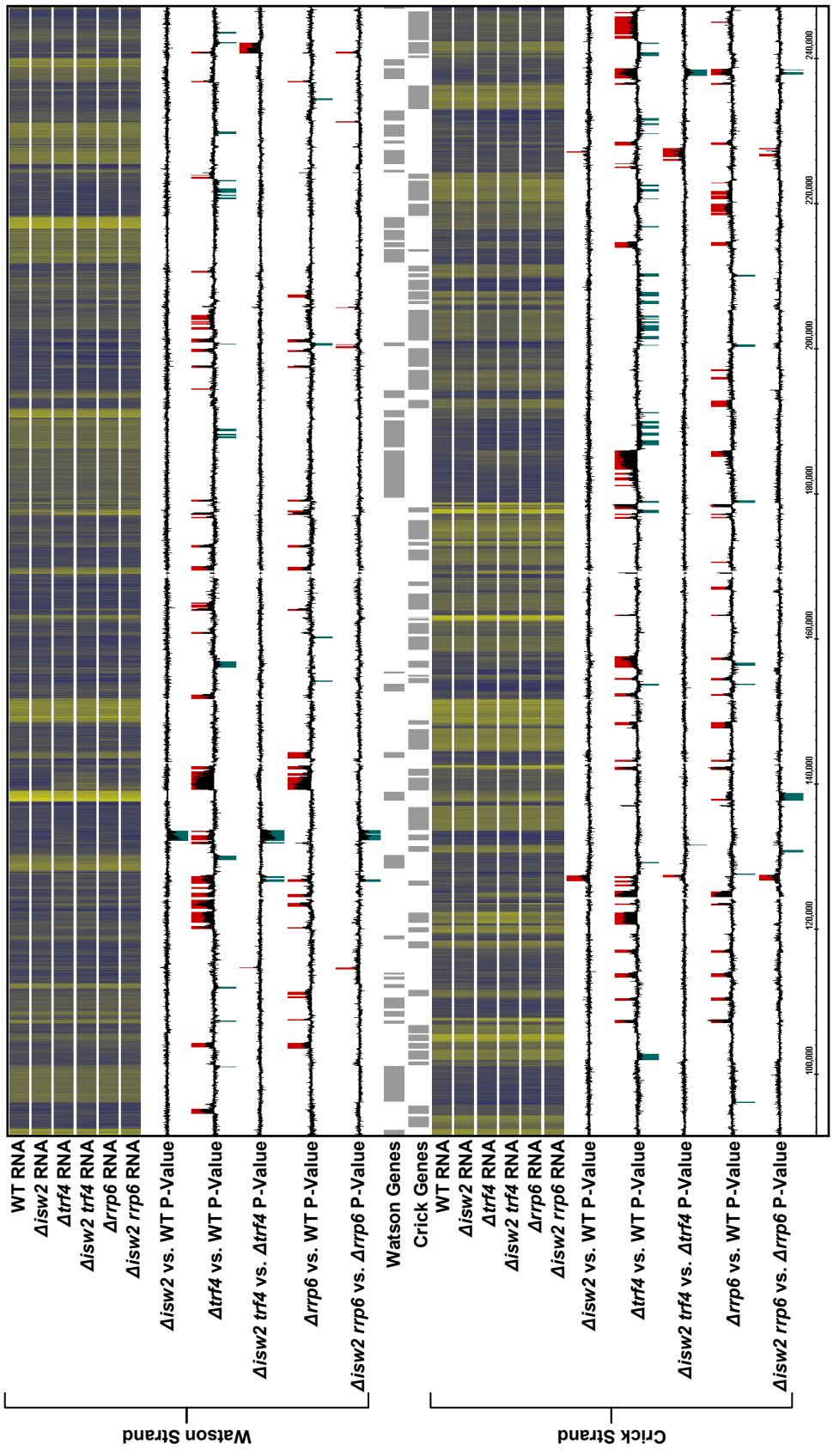


Figure 10: Systematic Annotation of Cryptic Transcripts
Representative screenshot of chr3: 100,000-240,000. The RNA signals (average of each replicate) from all the strains examined on high-density microarrays are displayed as heat maps. The black lines for each strain pair show the p-values for the differences in RNA signals, with the lines above the midline indicating increased signals and those below the midline indicating decreased signals. Annotated cryptic transcripts are displayed in colored boxes, with red indicating increase and green indicating decrease in transcripts. ORFs are shown as grey boxes.

A

	Δisw2 vs. WT	*Δtrf4* vs. WT	*Δrrp6* vs. WT	*Δisw2 trf4* vs. *Δtrf4*	*Δisw2 rrp6* vs. *Δrrp6*
***Δisw2* vs. WT**	*N/A*	15 (79%)	8 (72%)	24 (83%)	22 (94%)
***Δtrf4* vs. WT**	20 (82%)	*N/A*	266 (74%)	41 (75%)	40 (68%)
***Δrrp6* vs. WT**	8 (83%)	294 (95%)	*N/A*	20 (69%)	10 (63%)
Δisw2 trf4* vs. *Δtrf4	24 (76%)	38 (79%)	19 (70%)	*N/A*	47 (74%)
Δisw2 rrp6* vs. *Δrrp6	20 (78%)	34 (78%)	10 (72%)	46 (79%)	*N/A*

B

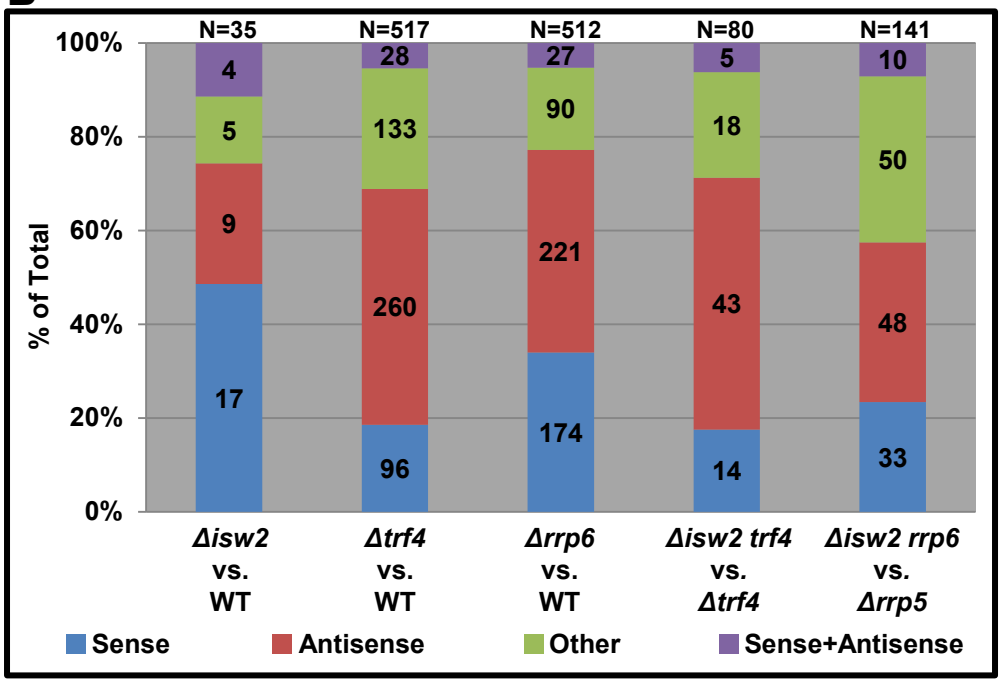


Figure 11: Classification of Cryptic Transcripts
(A) Displayed is the total number of cryptic transcripts identified from comparison of each strain pair that directly overlaps with a cryptic transcript identified from a different strain pair. The average percentage (%) of overlapping bp of cryptic transcript is shown in parenthesis.
(B) Cryptic transcripts identified were classified as cryptic sense (cyan), antisense (red), other (green), and sense+antisense (purple, represents cryptic transcripts that overlap two ORFs in both the sense and antisense direction) as described in the Materials and Methods. The numbers within the columns denote the number of transcripts from each strain identified in a particular class. "N" denotes the total number of cryptic transcripts identified from comparison of each strain pairs.

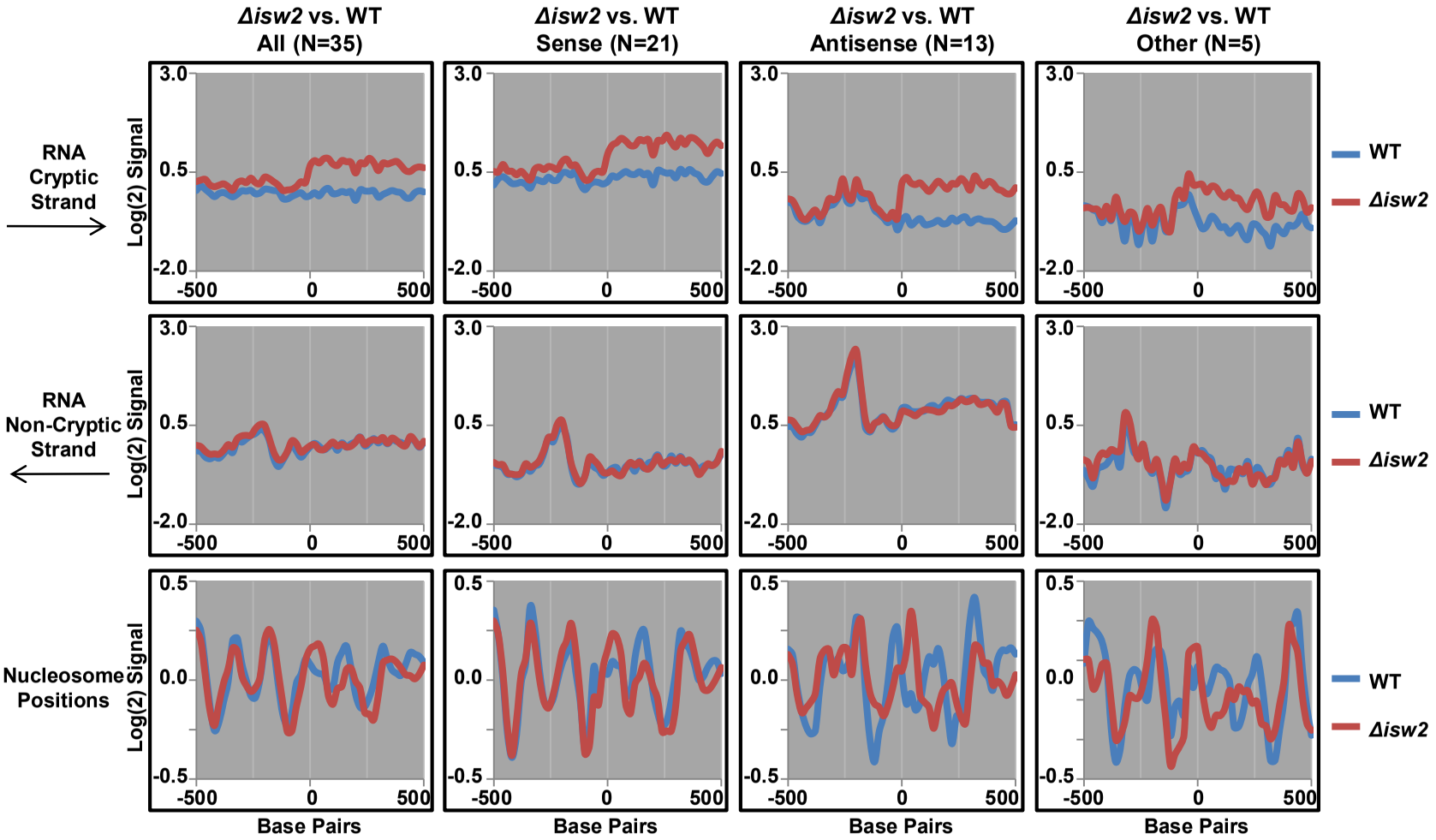


Figure 12: Comparison of RNA and Nucleosome Signals Around TSSs of Cryptic Transcript in *Δisw2* vs. WT
Cryptic transcripts identified from comparison of each strain pair were broken down by classification ("All" denotes total transcripts). Cryptic transcripts were aligned by their 5'-ends (0 bp denotes the 5'-end of each transcript) and oriented such that cryptic transcription occurs to the right. The average signals of RNA (top two panels) and nucleosomes (bottom panels, WT and *Δisw2* (Whitehouse et al., 2007)) around cryptic RNA TSSs are displayed.

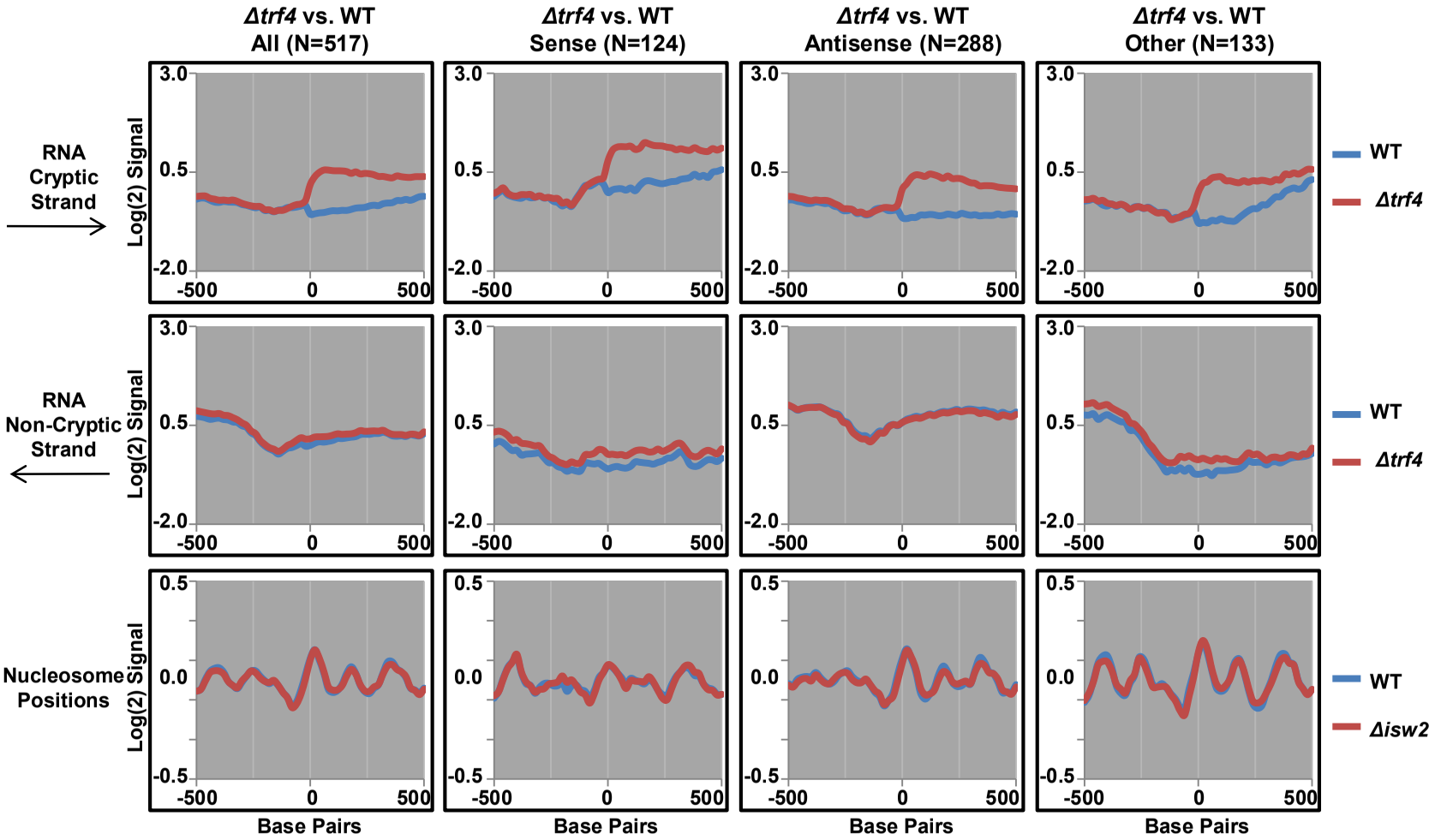


Figure 13: Comparison of RNA and Nucleosome Signals Around TSSs of Cryptic Transcript in *Δtrf4* vs. WT
Cryptic transcripts identified from comparison of each strain pair were broken down by classification ("All" denotes total transcripts). Cryptic transcripts were aligned by their 5'-ends (0 bp denotes the 5'-end of each transcript) and oriented such that cryptic transcription occurs to the right. The average signals of RNA (top two panels) and nucleosomes (bottom panels, WT and *Δisw2* (Whitehouse et al., 2007)) around cryptic RNA TSSs are displayed.

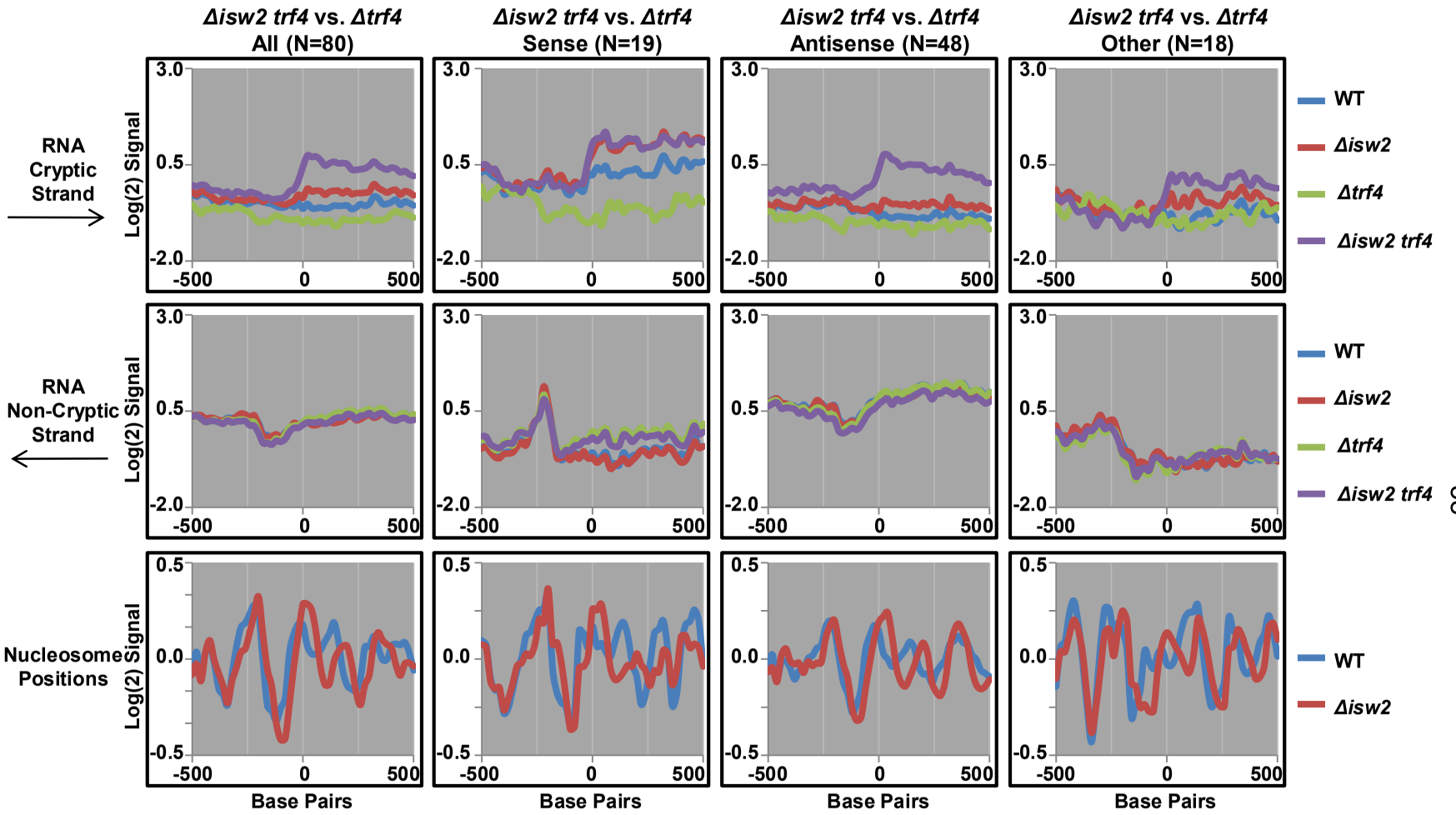


Figure 14: Comparison of RNA and Nucleosome Signals Around TSSs of Cryptic Transcript in *Δisw2 trf4* vs. *Δtrf4*

Cryptic transcripts identified from comparison of each strain pair were broken down by classification ("All" denotes total transcripts). Cryptic transcripts were aligned by their 5'-ends (0 bp denotes the 5'-end of each transcript) and oriented such that cryptic transcription occurs to the right. The average signals of RNA (top two panels) and nucleosomes (bottom panels, WT and *Δisw2* (Whitehouse et al., 2007)) around cryptic RNA TSSs are displayed.

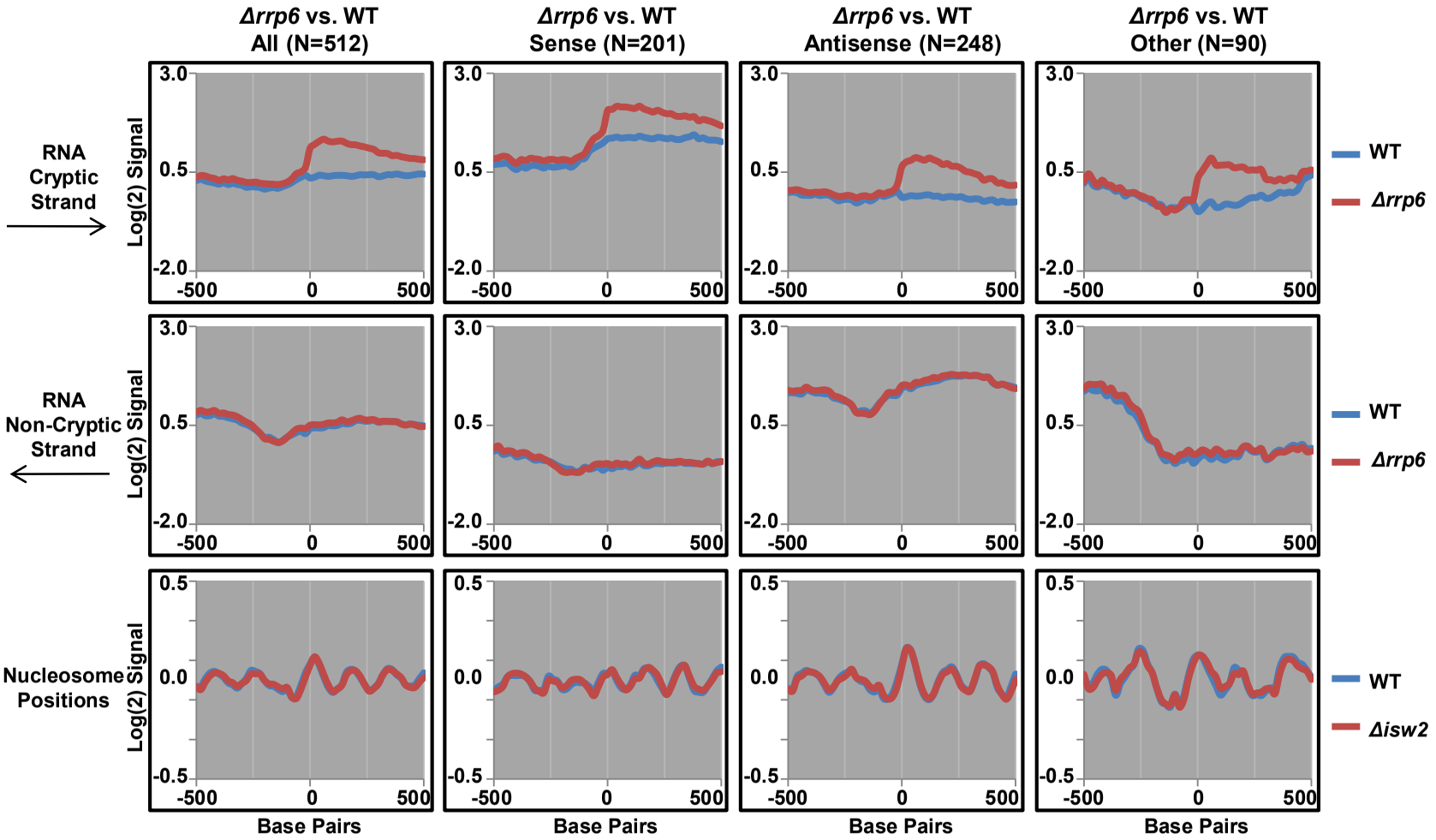


Figure 15: Comparison of RNA and Nucleosome Signals Around TSSs of Cryptic Transcript in *Δrrp6* vs. WT
Cryptic transcripts identified from comparison of each strain pair were broken down by classification ("All" denotes total transcripts). Cryptic transcripts were aligned by their 5'-ends (0 bp denotes the 5'-end of each transcript) and oriented such that cryptic transcription occurs to the right. The average signals of RNA (top two panels) and nucleosomes (bottom panels, WT and *Δisw2* (Whitehouse et al., 2007)) around cryptic RNA TSSs are displayed.

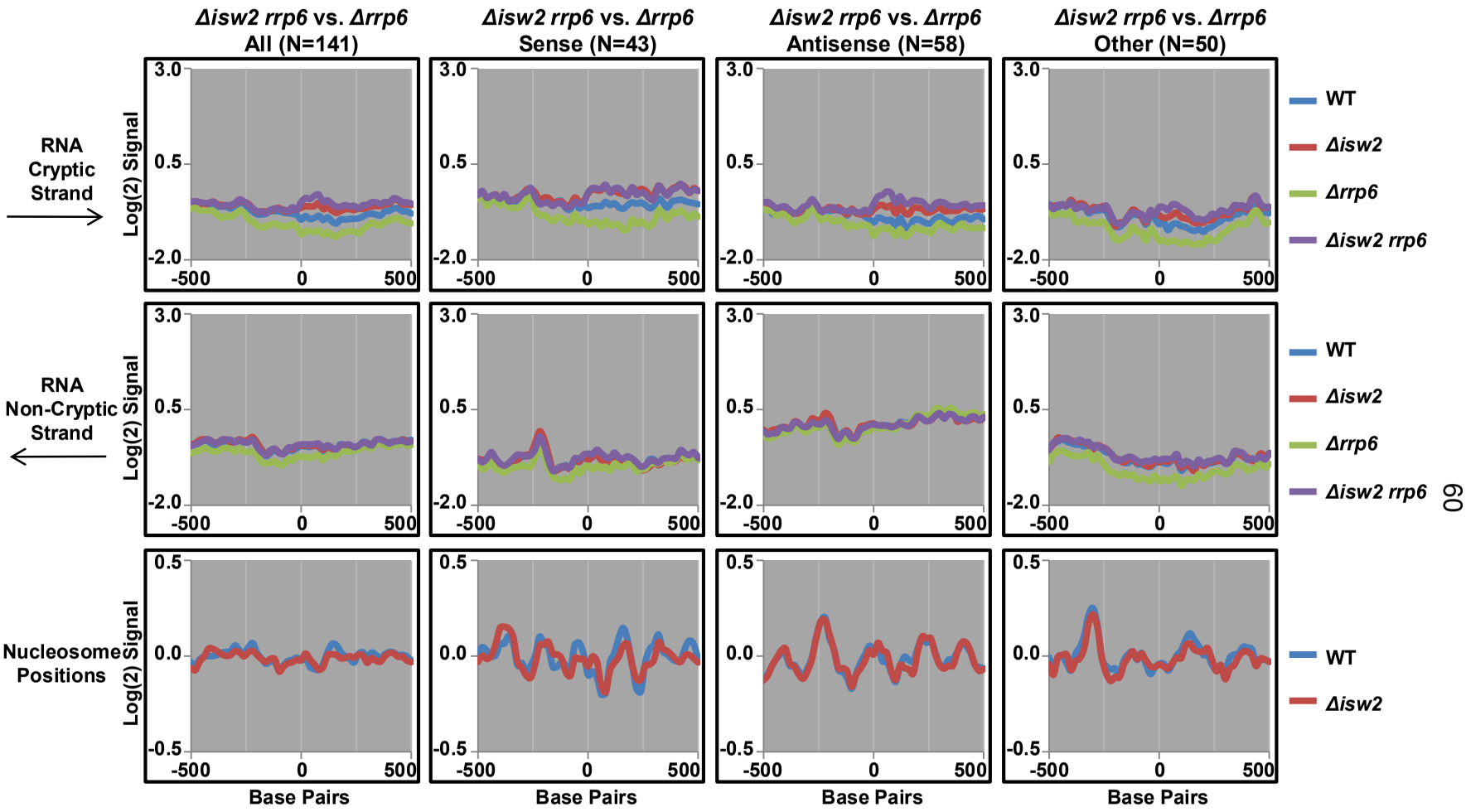


Figure 16: Comparison of RNA and Nucleosome Signals Around TSSs of Cryptic Transcript in *Δisw2 rrp6* vs. *Δrrp6*

Cryptic transcripts identified from comparison of each strain pair were broken down by classification ("All" denotes total transcripts). Cryptic transcripts were aligned by their 5'-ends (0 bp denotes the 5'-end of each transcript) and oriented such that cryptic transcription occurs to the right. The average signals of RNA (top two panels) and nucleosomes (bottom panels, WT and *Δisw2* (Whitehouse et al., 2007)) around cryptic RNA TSSs are displayed.

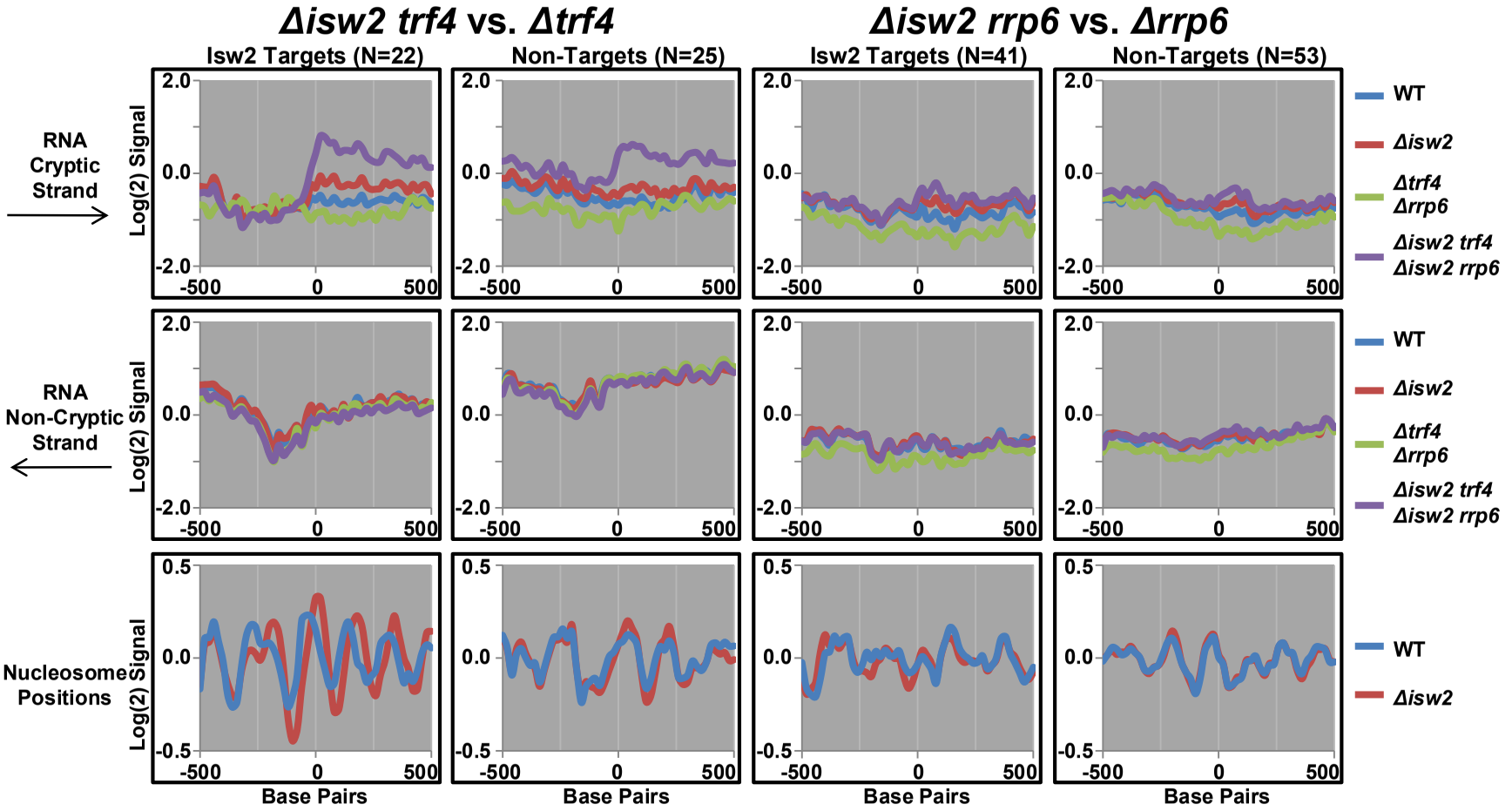


Figure 17: Isw2-Dependent Chromatin Remodeling is Associated with Suppression of Cryptic Transcription
Cryptic transcripts identified in either *Δisw2 trf4* vs. *Δtrf4* or were *Δisw2 rrp6* vs. *Δrrp6* broken down into those detected at "Targets" (defined as cryptic transcripts with TSSs within 300 bp of regions with both enriched Isw2-ChIP signals and Isw2-dependent chromatin remodeling (Whitehouse et al., 2007)) and "Non-Targets" (defined as cryptic transcripts with TSSs farther than 300 bp from either enriched Isw2-ChIP signals or Isw2-dependent chromatin remodeling (Whitehouse et al., 2007)). All cryptic transcripts were then aligned by their 5'-ends (0 bp denotes the 5'-end of each transcript) and oriented such that cryptic transcription occurs to the right. The average signals of RNA and nucleosomes (WT and *Δisw2* (Whitehouse et al., 2007)) around cryptic RNA TSSs are displayed.

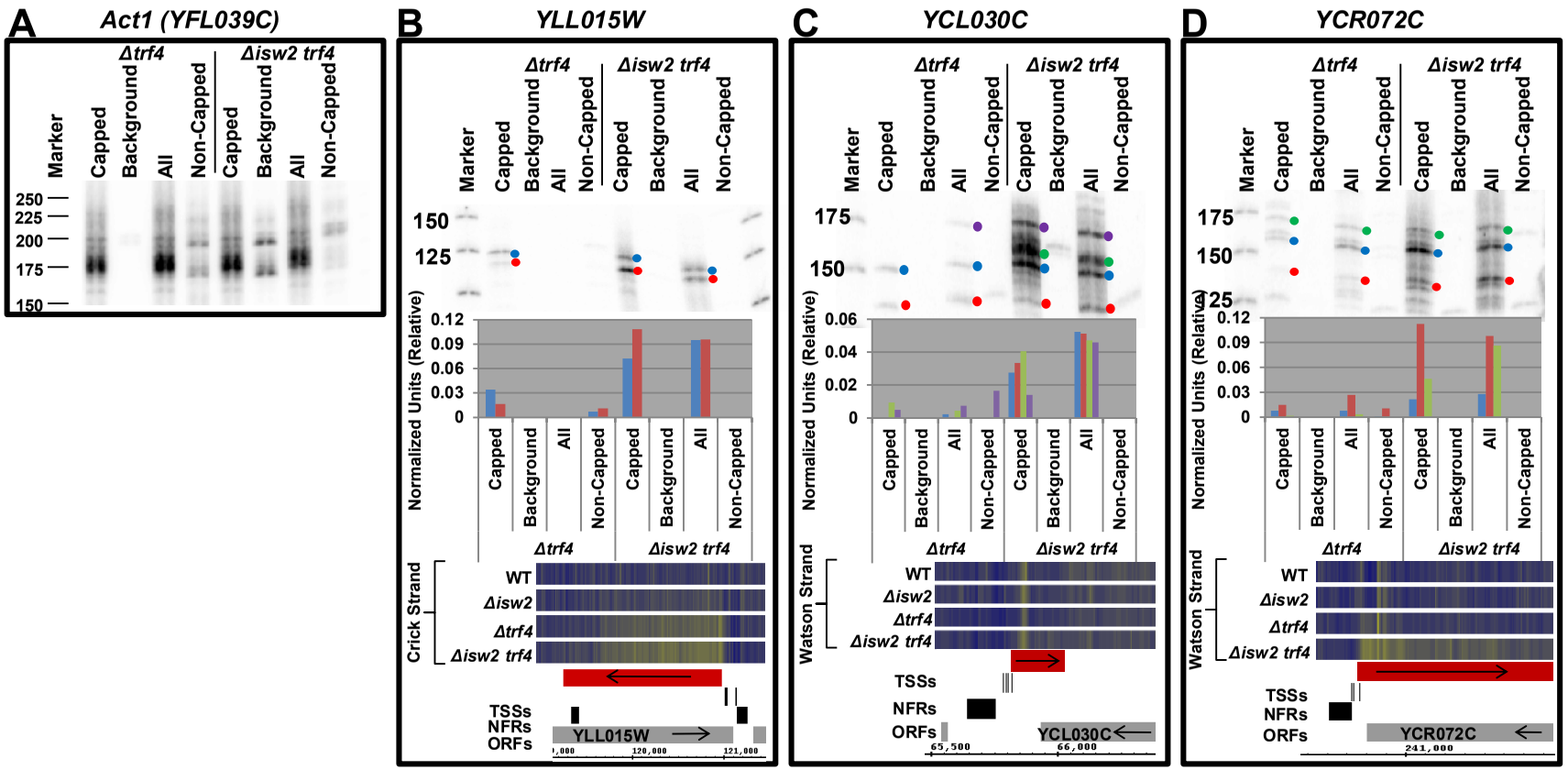


Figure 18: 5'-RACE Confirms Isw2 Represses Cryptic Transcription from NFRs

(A to D) Results of 5'-RACE using *Δtrf4* and *Δisw2 trf4* strains. Capped, total ("All"), and non-capped transcripts were detected according to manufactures (Ambion) protocol. Colored dots correspond to the quantified transcripts (graph, middle panel), which were normalized to the loading control *ACT1*. Bottom panel displays the mapped 5'-RACE TSS locations (denoted as "TSSs" track) of each quantified transcript in relation to ORFs (designated by grey boxes in "ORFs" tracks, arrows denote direction of transcription) and NFRs. The RNA signals (average of each replicate) from microarray data at corresponding loci are displayed as heat maps, with yellow denoting higher RNA signals and cyan denoting lower RNA signals. The cryptic transcripts (*Δisw2 trf4* vs. *Δtrf4*) identified on microarrays are shown by red boxes with the arrows designating the direction of the transcripts.

(A) *ACT1 - YFL039C* (loading control) **(B)** *YLL015W* **(C)** *YCL030C* **(D)** *YCR072C*

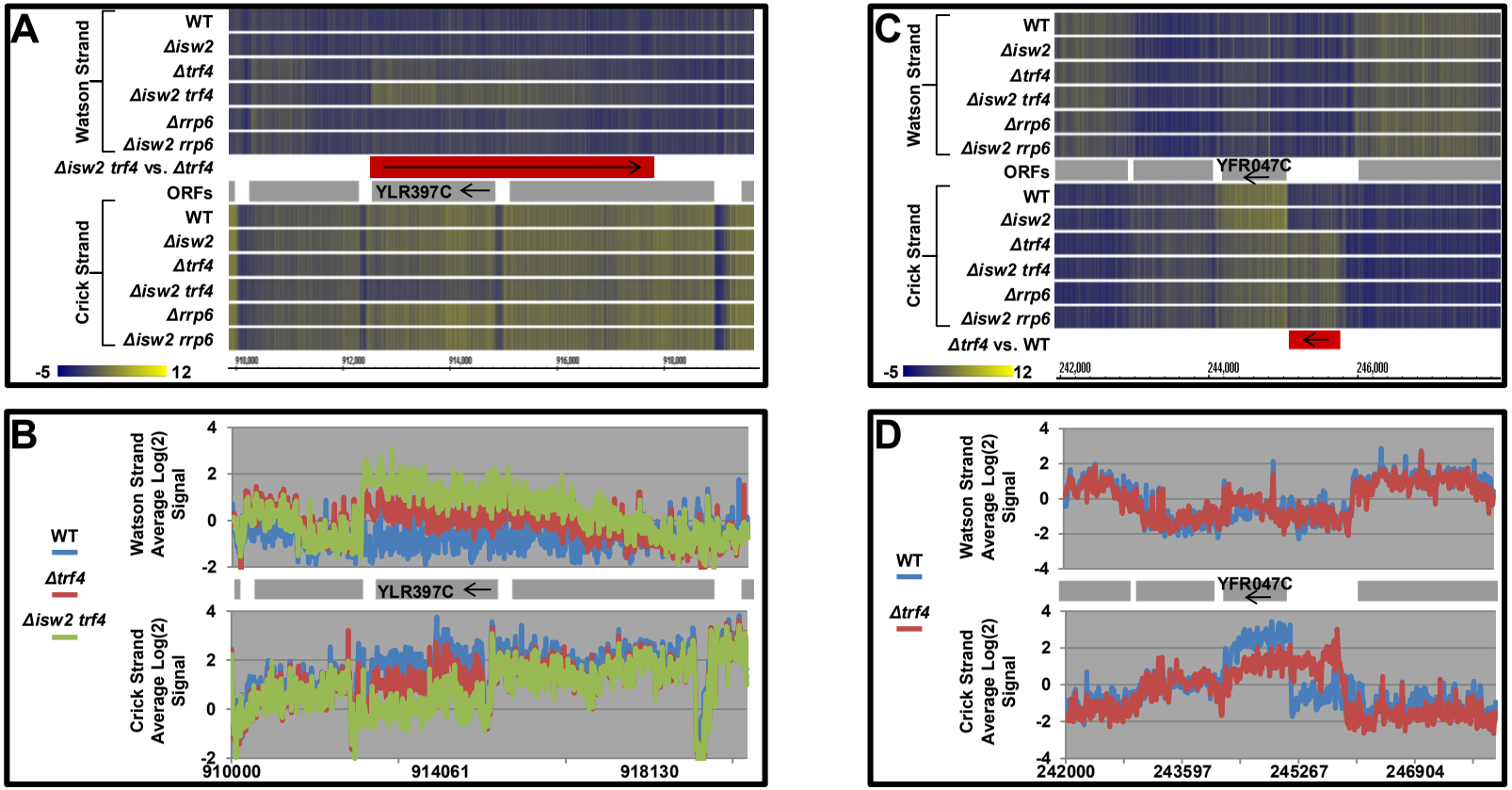


Figure 19: Evidence for Transcriptional Interference

(A and C) The RNA signals (average of each replicate) for all the strains examined on high-density microarrays are displayed as heat maps. Arrows indicate the direction of transcription for either ORFs (denoted as grey rectangles) or annotated cryptic transcripts (denoted as red boxes). ORFs displaying transcriptional interference are labeled. Yellow indicates higher RNA signals and cyan denotes lower RNA signals.

(B and D) The RNA signals (average of each replicate) for strains relevant to the transcriptional interference locus are displayed as line graphs. Arrows indicate the direction of transcription for ORFs (denoted as grey rectangles). ORFs displaying transcriptional interference are labeled.

(A and B) Example of antisense-sense transcription interference at *YLR397C*.

(C and D) Example of sense-sense transcription interference at *YFR047C*.

DNA LOOPING MEDIATES TARGETING OF THE ISW2 CHROMATIN REMODEING ENZYME

Abstract

ATP-dependent chromatin remodeling enzymes are highly abundant protein complexes that play pivotal roles in the regulation of DNA-dependent processes. The mechanisms by which they are recruited to target loci are not well understood. Here we obtained the first genome-wide view for sequence-specific transcription factor (TF)-dependent targeting of Isw2 and identified Ume6, Nrg1, Cin5, and Sok2 as global mediators of Isw2 targeting. Unexpectedly, Isw2 is targeted in a TF-dependent fashion to a large number of loci not containing an annotated TF binding site. Using 3C we discovered that Isw2 can be targeted to specific loci via both Ume6- and TFIIB-dependent DNA looping, suggesting a model for DNA looping in mediating TF-dependent targeting of a chromatin remodeling enzyme. We thus established a novel role for repressor mediated DNA looping and identified DNA looping as a previously unknown mechanism for the recruitment of a chromatin remodeling enzyme. Finally, abrogation of DNA looping using *Δume6* or *sua7-1* mutants also results in transcriptional derepression of target genes. Our work has characterized the genome-wide role for TF-dependent targeting of a chromatin remodeling enzyme and uncovered a novel physiological role for DNA looping in which the three-dimensional spatial organization of genomes in the nucleus affects chromatin structure.

Inroduction

Over the past two decades, an unprecedented amount of information has accumulated on both the structure and function of eukaryotic genomes. DNA sequences, their evolutionary conservation, transcription factor binding sites, nucleosome positions, DNA and histone modification patterns, and transcription start and end sites have been determined at high resolution across many eukaryotic genomes. These studies established linear maps of genomic information that affects the regulation of DNA-dependent processes. However, eukaryotic genomes are packaged and function within the three-dimensional space of the nucleus. How this packaging of DNA affects DNA-dependent processes is not well understood.

The efficient three-dimensional packaging of genomes into the relatively small nuclei of eukaryotic cells is achieved at two distinct levels: the compaction of DNA into chromatin and the folding of the chromatin fiber. Both proper compaction of DNA into chromatin and the folding of the chromatin fiber are required for normal cellular and developmental processes (Cremer and Cremer, 2001; Cremer and Cremer, 2010; Eisen et al., 1995; Radman-Livaja and Rando, 2010; Rando and Chang, 2009) and changes in either chromatin structure or the folding of the chromatin fiber are associated with many complex diseases (Matarazzo et al., 2007; Misteli, 2010; Timme et al., 2011; Wiech et al., 2009; Zardo et al., 2008). Elucidating the roles that the three-dimensional packaging of DNA has on DNA-dependent processes would therefore reveal previously unknown layers of regulation for these DNA-dependent processes.

Using microscopic approaches, chromosomes within the nuclei of animals (Cremer and Cremer, 2001), plants (Berr et al., 2006; Pecinka et al., 2004; Shaw et al., 2002), and yeast (Bystricky et al., 2005; Cremer and Cremer, 2010; Duan et al., 2010) have been shown to occupy highly organized nonrandom "territories" (chromosome territories, CTs). These discrete chromosome conformations have been postulated to reflect and/or regulate DNA-dependent processes. In support of this, studies have shown that abnormal folding of the chromatin fiber is associated with many complex diseases, including cancer (Matarazzo et al., 2007; Misteli, 2010). It is therefore important to fully understand the roles by which the spatial organization of the genome affects DNA-dependent process.

The chromosome conformation capture (3C) assay (Dekker et al., 2002) identifies DNA looping events by detecting the frequency of interactions between two distantly located chromosomal loci, enabling the discovery of loci that are in close proximity within three-dimensional space. This assay has enabled the identification and functional characterization of DNA looping within eukaryotic nuclei. Using 3C, two general classes of DNA loops have been identified: chromatin loops and gene loops. Chromatin loops have traditionally been defined as the juxtaposition of a distal regulatory region, such as an enhancer or silencer, with a gene promoter, while gene loops are defined as the physical association of the promoter and terminator regions of same gene. To date, chromatin loops and gene loops have been described in human, fly, and yeast cells (Ansari and Hampsey, 2005; Duan et al., 2010; Hampsey et al., 2011; Laine et al., 2009; Murrell et al., 2004; Nemeth et al., 2008; O'Reilly and Greaves, 2007; O'Sullivan et al., 2004; Perkins et al., 2008; Singh and Hampsey, 2007;

Spilianakis et al., 2005; Tan-Wong et al., 2008; Tan-Wong et al., 2009; Tolhuis et al., 2002; Vernimmen et al., 2007).

Extensive characterization of these DNA loops using 3C has led to the identification of numerous interactions between structural proteins (Comet et al., 2011; Hadjur et al., 2009; Parelho et al., 2008; Wendt et al., 2008), sequence-specific transcription factors (TFs) (Drissen et al., 2004; Phillips and Corces, 2009; Splinter et al., 2006; Vakoc et al., 2005), general transcription factors (Singh and Hampsey, 2007), and RNA 3'-end processing factors (Singh and Hampsey, 2007) that are required for the formation and/or maintenance of DNA looping. Functionally, chromatin loops have been shown to play roles in transcription (Comet et al., 2011; Nemeth et al., 2008; Perkins et al., 2008; Schoenfelder et al., 2010a; Schoenfelder et al., 2010b; Wang et al., 2011), while gene loops have been implicated in transcriptional memory (Laine et al., 2009; Tan-Wong et al., 2009). However, the molecular mechanisms by which DNA loops affect transcription regulation or memory are not understood.

The physical compaction of DNA within nuclei is achieved by wrapping 147 bp of DNA around an octamer of histone proteins to form the nucleosome (Arents et al., 1991; Kornberg and Lorch, 1999; Luger et al., 1997), the most basic repeating unit of chromatin. The physical compaction of DNA into nucleosomes effectively inhibits access of proteins to the genome by occluding the underlying DNA sequence or preventing translocation of proteins along DNA (Ehrenhofer-Murray, 2004). As a result, all DNA-dependent processes, including transcription, replication, repair, and recombination are significantly affected by the positions of nucleosomes across the genome.

To modulate chromatin structure and gain access to the underlying DNA, eukaryotic cells have evolved highly conserved ATP-dependent chromatin remodeling enzymes that utilize the energy of ATP hydrolysis to slide, evict, or replace histones within nucleosomes (Clapier and Cairns, 2009). Because their biochemical activities alter chromatin structure, chromatin remodeling enzymes are key modulators of DNA-dependent processes (Ehrenhofer-Murray, 2004). As a result, much effort has been put forth to identify targets of ATP-dependent chromatin remodeling enzyme family members in a genome-wide scale (Damelin et al., 2002; Floer et al., 2010; Gkikopoulos et al., 2011; Hartley and Madhani, 2009; Ng et al., 2002; Parnell et al., 2008; Tirosh et al., 2010; Whitehouse et al., 2007; Yen et al., 2012). For example, in yeast, Isw2 targets nucleosome free regions (NFRs) to repress transcription by reducing NFR size (Whitehouse et al., 2007; Yadon et al., 2010) while RSC increases NFR size to activate transcription (Hartley and Madhani, 2009). Ino80 also targets NFRs to enhance transcription activation and functions in the localization of the histone variant Htz1 (H2A.Z) (Papamichos-Chronakis et al., 2011). In contrast, Chd1 and Isw1 predominately target the body of genes and are involved in positioning nucleosomes toward the 3'-end (Tirosh et al., 2010). However, the underlying mechanisms for the targeting of chromatin remodeling activities to these specific loci remain a major unanswered question.

To date, two distinct mechanisms have been implicated in the targeting of chromatin remodeling enzymes to specific loci: the direct interactions with the covalently modified N-terminal tails of histones and the physical interactions with TFs. In yeast, both RSC and Chd1 contain bromo- and chromodomains, respectively, which are

capable of recognizing acetylated or methylated lysine residues (Clapier and Cairns, 2009). However, the extent by which these interactions contribute to targeting of these factors across the yeast genome remain to be determined. On the other hand, Isw2 can physically interact with and be targeted directly by the TF Ume6 to its cognate binding sites at small number of loci (Gelbart et al., 2005; Goldmark et al., 2000). The genome-wide requirement for Ume6 in Isw2 targeting has not been determined. Similarly, the TFs Abf1 and Reb1 have been implicated in targeting RSC to a subset of loci (Hartley and Madhani, 2009), although direct proof for either of these TFs in RSC targeting has not been shown. Indeed, no systematic genome-wide analysis of the targeting mechanisms for any chromatin remodeling enzyme has been undertaken. This is a fundamental issue, as chromatin remodeling enzymes are very abundant proteins, estimated at one remodeler per ~14 nucleosomes (Erdel et al., 2010) in human cells, yet, they alter chromatin structure at select loci.

In this study, we sought to determine the mechanisms for the targeting of the chromatin remodeling enzyme Isw2 to specific loci across the *S. cerevisiae* genome. To this end, we obtained the first comprehensive genome-wide view of TF-mediated targeting of Isw2. Unexpectedly, Isw2 is targeted in a TF-dependent fashion to a large number of loci that lack an annotated TF binding site. This led us to the discovery that Isw2 can be targeted to specific loci via DNA looping. Finally, we show that both Ume6- and TFIIB-dependent DNA looping are required for maintaining transcriptional repression at gene looping and chromatin looping targets, respectively. Therefore, our work has identified a previously unknown mechanism to target a chromatin remodeling

enzyme and uncovered a novel physiological role for DNA looping, in which the three-dimensional spatial organization of genomes in the nucleus affects chromatin structure.

Results

Enrichment of TFs at Isw2 Targets

In the budding yeast *S. cerevisiae*, the highly conserved ATP-dependent chromatin remodeling enzyme Isw2 has been found to target ~2,100 regions genome-wide (Whitehouse et al., 2007). Previous studies from our lab have demonstrated that these targets can be categorized into three major classes: the 5'-end of genes, the 3'-end of genes, and upstream of tRNA genes (Gelbart et al., 2005; Goldmark et al., 2000; Whitehouse et al., 2007). At the 5'- and 3'-end of genes, Isw2 utilizes the energy of ATP hydrolysis to slide nucleosomes along DNA and reduce the size of NFRs to repress both coding and non-coding RNA transcription (Whitehouse et al., 2007; Yadon et al., 2010). Upstream of tRNA genes, Isw2 is required for the periodic integration pattern of the Ty1 retrotransposon (Bachman et al., 2005; Gelbart et al., 2005). As such, TF-dependent Isw2 targeting could explain enrichment of Isw2 at some loci: Ume6 for the promoters of early meiotic genes during mitotic growth (~193 genes) (Goldmark et al., 2000), α2-MCM1 for the promoters of MATa-specific genes in MATα cells (~5 genes) (Bachman et al., 2005; Gelbart et al., 2005), and Bdp1 for tRNA genes (~275 genes) (Bachman et al., 2005; Gelbart et al., 2005). It is therefore apparent that these three TFs account for the targeting of Isw2 to only a small fraction of loci genome-wide. Considering that other TFs, including Abf1, Reb1, Pho4, and Gal4, have also

been implicated in the recruitment of other classes of chromatin remodeling enzymes (Adkins et al., 2007; Badis et al., 2008; Bryant et al., 2008; Hartley and Madhani, 2009), we postulated that additional TFs may be involved in Isw2 targeting.

To address this possibility, we performed a statistical analysis to curate a list of annotated sequence-specific TF binding sites that were preferentially enriched upstream of Isw2 targets located at the 5'-end of genes (see Materials and Methods for details). A total of 15 TFs with a p-value≤0.01 were identified (Figure 20). Ume6 was found to be the most highly enriched TF, validating our strategy for identifying TFs involved in Isw2 recruitment.

TF-Dependent Isw2 Targeting

Next, we assessed the role of these TFs in the targeting of Isw2 genome-wide. To this end, chromatin immunoprecipation of Isw2 on genome-wide tilling microarrays (Isw2 ChIP-chip) in strains containing a null mutation in the TFs, *Δume6*, *Δnrg1*, *Δcin5*, or *Δsok2,* was performed. Each TF was chosen because their binding sites are highly enriched at Isw2 5'-end targets (Figure 20) and to bind a large number of genes genome-wide. The *Δume6* mutant was included as a positive control for Isw2 targeting to select loci.

Previous studies from our lab have demonstrated that a catalytically inactive Isw2 mutant, Isw2-K215R, is preferentially enriched at Isw2 target sites while wild type (WT) Isw2 is nonspecifically bound across the genome (Gelbart et al., 2005). As a result, Isw2 targets have been defined as sites at which Isw2-K215R is enriched relative to WT Isw2 (Fazzio et al., 2005; Gelbart et al., 2005; Whitehouse et al., 2007). Therefore,

each Isw2 ChIP-chip experiment was performed by competitively hybridizing immunoprecipitated Isw2-K215R against WT Isw2 from the same parental strain background. The direct hybridization of Isw2-K215R and WT Isw2 to the same microarray reduced the number of microarray hybridizations and the amount of data analysis and normalization. Isw2 ChIP-chip was performed for at least 2 biological replicates of each strain (see Materials and Methods for details). Visualization of the average log2 signal in each strain revealed defined peaks of Isw2 enrichment at many loci in all strains (Figure 21). The enrichment of Isw2 ChIP signals observed around the Ume6 binding sites at both *INO1* (Figure 21A) and *SIP4* (Figure 21B) in *UME6* strains is selectively lost in the *Δume6* strains, validating our strategy (Goldmark et al., 2000).

Annotation of TF-Dependent Isw2 Targets

We next sought to systematically identify regions in which Isw2 targeting is affected in each of the TF mutants. To this end, regions containing consecutive probes, totaling at least 250 base pairs (bp) in length, each with a statistically significant (p-value≤0.05) reduction of Isw2 signal in the mutants relative to WT and containing an average of at least 1.65 fold-change were annotated (see Materials and Methods for details). Manual evaluation of the results (Figure 21) confirms that our annotation accurately marks regions in which Isw2 targeting is reduced in the mutants relative to WT.

A total of 563, 194, 341, and 226 regions with decreased Isw2 signals were identified in *Δume6*, *Δnrg1*, *Δcin5*, and *Δsok2* strains, respectively (hereafter referred to as Ume6-, Nrg1-, Cin5-, and Sok2-dependent Isw2 targets). It should be noted that the

number of regions annotated in each mutant is an underestimate, as our algorithm utilized conservative parameters and chromosomes II, V, and XII were removed from the analysis (see Materials and Methods for details). Nonetheless, these data revealed, for the first time, a role for TFs in the targeting of a chromatin remodeling enzyme on a genome-wide scale.

At Ume6-dependent Isw2 targets, we identified 49% (70 of 142) of the annotated Ume6 binding sites, representing a significant fraction of the total. In contrast, only 21% (30 of 167) of the Nrg1 binding sites, 10% (36 of 355) of the Cin5 binding sites, and 29% (160 of 544) of the Sok2 binding sites were identified within their respective Isw2 target regions. This could be partly explained by the stringent criteria used to identify regions with decreased Isw2 ChIP signal. On the other hand, only a small fraction of the total TF-dependent Isw2 targets contained annotated binding sites for the corresponding TF (Figure 22): 10% (58 of 563) of the Ume6-dependent targets, 11% (21 of 194) of the Nrg1-dependent targets, 7% (25 of 341) of the Cin5-dependent targets, and 19% (42 of 226) of the Sok2-dependent targets. Interestingly, the average Isw2 ChIP signal in *UME6* strains is much higher around Ume6-dependent Isw2 targets containing a Ume6 binding site than without (Figure 22A, solid blue and red lines, respectively). Nevertheless, the average Isw2 ChIP signals were strongly reduced in *Δume6* strains regardless of the presence or absence of a Ume6 binding site (Figure 22A, dashed lines). Similar trends were seen around Nrg1- (Figure 22B), Cin5- (Figure 22C), and Sok2-dependent regions (Figure 22D), although the average Isw2 ChIP signals between regions with or without the TF binding site were more similar. These results are consistent with a model in which TFs are required for the targeting of

chromatin remodeling factors to specific loci. Importantly, they also reveal that there are many Isw2 targets across the *S. cerevisiae* genome which cannot be explained by a simple model of TF-dependent targeting to its binding site. We therefore sought to determine the mechanisms for TF-dependent Isw2 targeting to regions without a corresponding TF binding site. In the following analysis we used Ume6-dependent targeting of Isw2 as a model, as its role in Isw2 targeting has been the best characterized (Goldmark et al., 2000).

Genome-Wide Ume6 ChIP-chip

As shown above, our analyses revealed that only 58 out of 563 Ume6-dependent Isw2 targets contained an annotated Ume6 binding site. One possible explanation is miss-annotation of Ume6 binding sites. To address this possibility, genome-wide Ume6 ChIP-chip was performed. As expected, the majority of annotated Ume6 binding sites have high Ume6 ChIP signals surrounding the binding sites (examples are shown in Figure 23A and B, red arrows). However, at a small number of loci either Ume6 binding sites were missed (Figure 23C) or no enrichment of Ume6 is observed at annotated binding sites (Figure 23B, black arrow). To address the effect of Ume6 binding on Isw2 targeting, the average signal at Ume6-dependent Isw2 targets either containing or lacking an annotated binding site was determined (Figure 23D, blue and red lines, respectively). As clearly seen, Isw2 targets with an annotated binding site exhibit on average much higher levels of Ume6 ChIP signal compared to those without an annotated binding site (Figure 23D). Together, these data establish that there are a large number of Ume6-dependent Isw2 targets where there is no sign of Ume6 binding.

TFIIB-Dependent Isw2 Targeting

During analysis of the genome-wide targeting of Isw2, we noted that Isw2 targets both the 5'- and 3'-end of the same gene at a highly statistically significant frequency (p-value<10^{-50}) (Figure 24A). Recent studies examining the spatial organization of the yeast genome using chromosome confirmation capture (3C) have shown that many genes juxtapose their promoter and terminator regions via gene looping (Ansari and Hampsey, 2005; Hampsey et al., 2011; Laine et al., 2009; O'Sullivan et al., 2004; Singh and Hampsey, 2007; Tan-Wong et al., 2009). Based on these data, we hypothesized that gene looping between the 5'- and 3'-end of genes that do and do not have either annotated Ume6 binding sites or Ume6 ChIP enrichment, respectively, mediates Ume6-dependent targeting of Isw2.

To test this hypothesis, we performed Isw2 ChIP-chip using the *sua7-1* mutant. This mutant contains a single point mutation, E62K, in the general transcription factor TFIIB that abrogates gene looping in yeast (Singh and Hampsey, 2007). Supporting our model, Isw2 ChIP signals in the *sua7-1* mutant were strongly decreased at the 3'-end of a few genes containing a Ume6 binding site at the 5'-end (Figure 24B). Unexpectedly, however, it was far more common to observe a decrease in both Sua7- and Ume6-dependent Isw2 signals far from Ume6 binding sites and across many genes (Figure 24C and D). These observations suggested that both gene looping and chromatin looping facilitate TF-dependent targeting of Isw2.

The Isw2 ChIP signals in WT and *sua7-1* strains were compared genome-wide (see Material and Methods for details). This analysis identified a total of 454 regions as

Sua7-dependent Isw2 targets, which, for the first time, revealed that the general transcription factor TFIIB has a global effect on the targeting of Isw2. If DNA looping indeed mediates Ume6-dependent targeting of Isw2 between two distant loci, one with and another without a Ume6 binding site, we would expect Isw2 ChIP signals at the Ume6 binding site to be dependent only on Ume6, while Isw2 ChIP signals at sites without a binding site to be dependent on both Ume6 and TFIIB. We found that 255 Isw2 targets satisfy these criteria (Figure 25A). We named Ume6-dependent Isw2 targets containing a Ume6 binding site "canonical" targets and Ume6- and Sua7-dependent Isw2 targets without a Ume6 binding site "ectopic" targets. It is noteworthy that the overlaps between Ume6- and Sua7-dependent Isw2 target regions are highly statistically significant (p-value<10^{-50}), showing that ectopic targets represent a major class among both Ume6- and TFIIB-dependent Isw2 targets. Analysis of all Ume6-dependent Isw2 targets (Figure 25B) further confirmed our conclusion that DNA looping facilitates TF-dependent targeting of Isw2 and also shows that there are a large number of uncharacterized Ume6-dependent Isw2 targets that do not have Ume6 binding or are TFIIB-dependent.

To rule out the possibility that the *sua7-1* mutation directly or indirectly alters Ume6 binding, Ume6 ChIP-chip was performed in a *sua7-1* mutant. The average log2 signal in the *sua7-1* mutant was indistinguishable from *SUA7* (data not shown). In fact, there was not a single locus in the yeast genome where Ume6 binding was significantly different between the *sua7-1* mutant and *SUA7*. This result excludes the possibility that TFIIB indirectly affects Isw2 targeting through Ume6.

DNA Looping-Dependent Isw2 Targeting

To directly test our model that DNA looping mediates Isw2 targeting, our collaborators, Dr. M. Hampsey and Dr. B.N. Singh, performed 3C assay to examine whether DNA looping takes place between canonical and ectopic Isw2 targets (see Materials and Methods for details). To this end, three loci (depicted in Figure 24B-D), representing one gene looping locus (Figure 24B) and two chromatin looping loci (Figure 24C and D) were selected. At all loci tested, DNA loops were specifically formed between the canonical and ectopic loci in WT cells (Figure 26). Importantly, quantification of the PCR signal using primer pairs that spanned the entire region of each locus revealed maximal intensity corresponded specifically with the region directly around the canonical and ectopic loci (primer pairs F9-F4, F5-F3, and F9-F2 for Figure 26A-C, respectively), not the intervening DNA. For each canonical and ectopic locus (F9-F4, F5-F3, and F9-F2), the PCR products were cross-linking, ligation, and restriction digestion dependent (Figure 27). These results are consistent with our model that DNA looping mediates Isw2 targeting to ectopic loci.

To confirm that the decreased Isw2 ChIP signals observed at ectopic loci was the result of impaired/loss of DNA looping, 3C was then performed in the *sua7-1* strain. Consistent with the known role of TFIIB in gene looping (Singh and Hampsey, 2007), a statistically significant loss of PCR signal was observed for the canonical and ectopic F9-F4 primer pair of the gene looping locus in the *sua7-1* strain relative to WT (Figure 28A). Strikingly, a similar loss of PCR signal in the *sua7-1* strain was also observed for the F5-F3 and F9-F2 primer pair of the chromatin looping loci (Figure 28B and C, respectively). These results demonstrate for the first time at two loci, that chromatin

looping, analogous to gene looping, is partially dependent on TFIIB. Furthermore, because Ume6 was also required for Isw2 targeting to ectopic loci we then evaluated DNA looping by 3C assay in *Δume6* strains. Strikingly, at each of the three loci tested, DNA looping between the canonical and ectopic loci was nearly completely abrogated (primer pairs F9-F4, F5-F3, and F9-F2 for Figure 29A-C, respectively), suggesting for the first time, a repressor-dependent role in DNA looping. In contrast to the TFIIB- and Ume6-dependent role in DNA looping at these three loci, the catalytically inactive Isw2-K215R mutant had no effect on either gene looping or chromatin looping at the three loci tested (Figure 29A-C). Together, our 3C results confirm our model that DNA looping is required for Isw2 targeting to three ectopic loci and suggests a co-dependence for both TFIIB and Ume6, but not the catalytic activity of Isw2, in mediating DNA looping at these loci.

DNA Looping-Dependent Transcription Regulation

To determine a biological role for DNA looping-dependent Isw2 targeting, we next assessed the transcriptional status of canonical and ectopic loci in the presence or absence of DNA loops. To this end, RT-qPCR (see Materials and Methods for details) was performed at three DNA looping loci in WT, *Δume6*, *Δisw2*, and *sua7-1* strains. As expected the absence of Ume6 resulted in transcriptional derepression at the canonical gene targets *YMR018W*, *ACS1*, and *YGR067C* (Figure 30A-C, respectively). This is consistent with the known role for Ume6 in transcriptional repression at its binding sites (Anderson et al., 1995; Goldmark et al., 2000; Jackson and Lopes, 1996; Kadosh and Struhl, 1997; Steber and Esposito, 1995; Strich et al., 1994; Sweet et al., 1997) and

confirms that our assay is working as expected. Additionally, transcriptional derepression in the *Δume6* mutant was also observed at the gene looping associated ectopic target, *SPO20* (Figure 30A), but not at the chromatin looping associated ectopic targets *BDH1* (Figure 30B) or *ERG25* (Figure 30C). These results suggest that the effect of Ume6-dependent gene looping on transcription regulation may be different at gene looping versus chromatin looping loci.

On the other hand, the *sua7-1* mutant displayed a statistically significantly increase in transcription only at the chromatin looping associated ectopic target *BDH1* (Figure 30B), but not at *ERG25* (Figure 30C) or the gene looping associated ectopic target *SPO20* (Figure 30A). Previous studies have shown using the *sua7-1* mutant, that gene loops have no effect on either basal or activated transcription (Ansari and Hampsey, 2005; Hampsey et al., 2011; Laine et al., 2009; Singh and Hampsey, 2007; Tan-Wong et al., 2009). Our results for *SPO20* are consistent with these studies. However, our results do suggest that, at least for one locus, TFIIB-dependent chromatin looping affects transcription regulation. Interestingly, at the canonical target *YGR067C*, the *sua7-1* mutant resulted in a dramatic decrease in transcription, suggesting a novel role for TFIIB-dependent DNA looping for the activation of this gene.

In contrast to either Ume6 or TFIIB, deletion of Isw2 had no statistically significant change in transcription at canonical or ectopic targets of gene looping or chromatin looping loci. We speculate that this could be the result of Rpd3 working in a parallel pathway as Isw2 (see Discussion for details). Together these results suggest that Ume6- and TFIIB-dependent DNA looping differentially regulates transcription at gene looping versus chromatin looping dependent ectopic targets, respectively.

Discussion

Targeting of ATP-Dependent Chromatin Remodeling Enzymes

ATP-dependent chromatin remodeling enzymes are highly conserved and abundant protein complexes that play essential roles in numerous cellular and developmental processes. Characterization of the genome-wide targets for many of these chromatin remodeling enzymes has often demonstrated both redundant and opposing roles for multiple remodelers at a single locus (Yen et al., 2012). It has thus become apparent that elucidating the mechanisms for targeting these enzymes is critical in understanding their mechanisms of action and functions. Previous studies from our lab had demonstrated that Ume6 targets Isw2 to its cognate binding site via physical interactions at a few loci. Similarly, Abf1, Reb1, Pho4, and Gal4 have also been implicated in the recruitment of other classes of chromatin remodeling enzymes, including SWI/SNF and RSC (Adkins et al., 2007; Badis et al., 2008; Bryant et al., 2008; Hartley and Madhani, 2009). However, to date, no comprehensive evaluation of the roles for TFs in chromatin remodeling factor targeting has been performed. In this study, we obtained the first genome-wide view of TF-mediated targeting of the ATP-dependent chromatin remodeling enzyme Isw2. We identified Ume6, Nrg1, Cin5, and Sok2 as global mediators of Isw2 targeting and established that TF-dependent targeting is a primary mechanism for the recruitment of a chromatin remodeling enzyme.

Previously, studies from our lab have shown that the physical interaction of Ume6 and Isw2 is required for the targeting of Isw2 to Ume6 binding sites. The

mechanisms for Nrg1, Cin5, and Sok2-dependent targeting of Isw2 are not known. Interestingly, a recent study identified ~20 sequence-specific TFs that are likely involved in targeting of the transcriptional co-repressor complex Tup1-Ssn6 (Hanlon et al., 2011). Strikingly, 11 of the 15 sequence-specific TFs (Nrg1, Sut1, Skn7, Phd1, Sok2, Cin5, Sko1, Swi6, Rox1, Swi4, and Rfx1) that were enriched at the 5'-end of Isw2 target genes (Figure 20) were implicated in Tup1-Ssn6 targeting. Studies from our lab have further shown that the MATα-specific transcriptional repressor α2 is required for Isw2 targeting to a-specific gene promoters (Gelbart et al., 2005). α2 physically interacts with Tup1 and is required for the targeting of Tup1 to MATa-specific gene promoters (Komachi and Johnson, 1997). It is therefore likely that Nrg1, Cin5, and Sok2, and possibly also Sut1, Skn7, Phd1, Sko1, Swi6, Rox1, Swi4, and Rfx1, are involved in Isw2 targeting via their interactions with Tup1. Hanlon et al. (Hanlon et al., 2011) further suggested that Tup1 can be simultaneously recruited to a single locus by multiple cofactors. This is consistent with our observation that many of the Isw2 target regions with decreased Isw2 ChIP signal in *Δnrg1*, *Δcin5*, and *Δsok2* strains directly overlap (data not shown).

Physiological Role for DNA Looping

The physiological roles for DNA looping are not well understood. In mammalian cells, chromatin and DNA looping are correlated with transcriptional regulation, both activation and repression (Fullwood et al., 2009; Gheldof et al., 2010; Horike et al., 2005; Kagey et al., 2010; Krivega and Dean, 2012; O'Reilly and Greaves, 2007; Perkins et al., 2008; Vernimmen et al., 2007). On the other hand, in yeast, gene looping

appears to function in transcriptional memory to enable a more rapid transcriptional response to previously seen stimuli (Laine et al., 2009; Tan-Wong et al., 2009). Our results demonstrate a novel physiological role for DNA looping in facilitating repressor-mediated targeting of an ATP-dependent chromatin remodeling enzyme. Abrogation of either DNA loops, using the previously characterized *sua7-1* mutant (Singh and Hampsey, 2007), or deletion of Ume6, prevented the targeting of Isw2 to ectopic loci, which do not contain a Ume6 binding site or Ume6 ChIP enrichment. 3C assay was used to demonstrate that DNA loops are specifically formed between Ume6 binding sites and three ectopic loci. These results support the model that DNA looping facilitates Ume6-dependent targeting of Isw2.

The fact that 255 ectopic Isw2 targets were identified suggests a major role for DNA looping in facilitating the transcription factor dependent targeting of Isw2. Unfortunately, the low throughput and laborious nature of 3C prevented the genome-wide characterization of DNA looping-dependent Isw2 targeting. Recently, a chromosome interaction map in yeast using a modified 4C (chromosome conformation capture-on-chip) assay combined with massively parallel sequencing has been published (Duan et al., 2010). However, the low resolution (on the order of ~3-4 kilobases) of this data set precluded its use for comparison to ectopic Isw2 targets. Future studies assessing the direct genome-wide role of DNA looping in Isw2 targeting will require genome-wide, high-resolution chromatin interaction maps.

Gene looping in yeast has been shown to depend only on activator-mediated transcription (Ansari and Hampsey, 2005; El Kaderi et al., 2009; Hampsey et al., 2011; Laine et al., 2009; Singh and Hampsey, 2007). Transcription itself, either basal or

activated, is unaffected by gene loops. In contrast, in mammalian cells, gene looping at both the BRCA1 and CD68 genes in breast tumor cell lines and B- and T-lymphocytes, respectively, is associated with repression of transcription (O'Reilly and Greaves, 2007; Tan-Wong et al., 2008). Here we show that similar to mammalian gene looping functions, both gene looping and chromatin looping are involved in transcriptional repression. At ectopic gene looping and chromatin looping targets, which do not contain a Ume6 binding site or Ume6 ChIP enrichment, transcriptional deregulation was observed in the *Δume6* and *sua7-1* mutants, suggesting for the first time in yeast that DNA looping facilitates repressor mediated transcriptional repression. Transcriptional repression in the absence of Isw2 was not observed. However, previous studies from our lab have shown that the histone deacetylase Rpd3 is recruited by Ume6 and functions to transcriptionally repress genes in a parallel pathway to Isw2 (Fazzio et al., 2001; Goldmark et al., 2000). Loss of both Isw2 and Rpd3 has been shown to be required for full transcriptional derepression of some Ume6 target genes (Fazzio et al., 2001; Goldmark et al., 2000). It is therefore likely Rpd3 is functioning in a similar capacity to maintain transcriptional repression at ectopic loci even in the absence of Isw2.

A major unanswered question is how DNA looping specifies the ectopic targets. For gene looping, the act of transcription may facilitate the interactions between the 5'- and 3'-end of the same gene through the physical translocation of RNA polymerase II along DNA from the promoter to the terminator during the initial/pioneering round of transcription (Hampsey et al., 2011). On the other hand, how chromatin looping takes place between two specific loci separated by 10s of kilobases is completely unknown.

In mammalian cells, it has been shown that some sequence-specific TFs bind DNA at both the promoter and ectopic locus and that their physical interaction is sufficient to maintain DNA looping. However, we have found that Ume6 does cross-link to DNA at the ectopic Isw2 targets. It is possible that Ume6 or TFIIB interacts with unidentified sequence-specific DNA binding proteins that bind both the promoter and ectopic locus to facilitate DNA looping. However, Isw2 is unlikely to be one such protein. Previous studies indicate that Isw2 does not bind DNA in any sequence-specific manner (Fazzio et al., 2005; Gelbart et al., 2005; Gelbart et al., 2001). Furthermore, our results show that Isw2 is not required for DNA looping as demonstrated by the maintenance of DNA loops in the absence of Isw2.

Strikingly, our analysis further demonstrated that a number of Ume6-dependent Isw2 targets neither contain an annotated Ume6 binding site nor are dependent on TFIIB-mediated DNA looping for targeting (Figure 25). These results suggest that there are still uncharacterized mechanisms for the recruitment of Isw2 to these loci. One possibility is TFIIB-independent DNA looping. While all characterized DNA looping in yeast to date has been dependent on TFIIB, this does not exclude the possibility that TFIIB-independent DNA looping mechanisms exist. In fact no genome-wide identification of TFIIB-dependent DNA looping has been done. It is thus conceivable that other mechanisms, beyond TFIIB, can mediate DNA looping, which facilitates Ume6-dependent targeting to these loci.

In summary, our work has established the first comprehensive genome-wide map for TF-dependent recruitment of a highly conserved chromatin remodeling enzyme. We further report a previously unknown mechanism of chromatin remodeling enzyme

targeting, and uncovered a novel physiological role for DNA looping in facilitating the TF-dependent targeting of an ATP-dependent chromatin remodeling enzyme. This work thus establishes functions for which the three-dimensional packaging of a eukaryotic genome into the nucleus affects chromatin structure.

Materials and Methods

Isw2 Targets and TF Enrichment

Isw2 target genes are defined as Isw2-K215R enriched and classified as 5'-end or 3'-end targets as previously reported for non-dubious genes (Whitehouse et al., 2007). Transcription factor binding sites are as previously reported (Harbison et al., 2004) using the binding threshold of $p<0.005$ and no conservation criteria. The population of 5642 intergenic regions upstream of non-dubious genes was used to determine the hypergeometric distribution of 1020 intergenic regions upstream of genes with Isw2 enrichment at the 5'-end and intergenic regions containing at least one annotated binding site for a particular transcription factor.

Isw2 ChIP-chip

Chromatin from strains with 3X-Flag-tagged Isw2 was crosslinked for 20 min, fragmented by sonication, and immunoprecipitated using approximately 6 mL (of OD660=0.70 +/- 0.05) of cell equivalent as described (Gelbart et al., 2005). Purified DNA was amplified using linker-mediated PCR amplification. dUTP incorporation was achieved by boiling of 3 µg DNA, 55 µM Tris-HCl pH6.8, 5.55 mM $MgCl_2$, 15 µg

Random Hexamers, and H_2O to final volume of 45 µL for 5 min followed by 5 min incubation on ice. 5 µL dNTP mix (1.2 mM dATP, 1.2 mM dCTP, 1.2 mM dGTP, 0.95 mM dTTP, and 0.25 mM dUTP) and 1 µL 50 U/µL Klenow (exo-) was incubated 10 min at 22°C, 30 min at 37°C, 5 min at 95°C, and 5 min at 4°C. Incubation was repeated following the addition of 1 µL Klenow (exo-). After phenol-chloroform extraction and ethanol precipitation, DNA was fragmented and Cy-dye labeled (WT Isw2=Cy3, and Isw2-K215R=Cy5) as described (Yadon et al., 2010). Equal concentrations of WT Isw2 and Isw2-K215R were competitively hybridized onto custom Nimblegen microarrays according to manufactures protocol. Log2 ratios and normalization was performed individually for forward and reverse strands using Ringo (Toedling et al., 2007) and chipchipnorm (nlag=30) (Peng et al., 2007) R packages followed by median adjustment and pseudomedian smoothing (150 bp window) (Royce et al., 2007). Figures display the average log2 ratio of forward and reverse strands for all replicates combined. Visualization of the average log2 signal of chromomsomes II, V, and XII revealed gross abnormal increase and/or decrease of Isw2 enrichment chromosome wide in the *Δume6* strain. We speculate this is partially the result of chromosomal duplications and/or deletions, as has been previously reported (Fazzio et al., 2001). As a result these chromosomes were removed from the analysis in all strains. The sequence of the probes on the microarrays is available upon request. Raw and normalized data are available for download at http://labs.fhcrc.org/tsukiyama

Changes in Isw2 Targeting

Changes in Isw2 enrichment was performed using LIMMA (Smyth, 2004) by identifying probes with significantly different signals between WT and mutant strains, utilizing an object containing the normalized log2 ratio from each hybridization (forward and reverse strands for each comparison was performed separately). Consecutive probes, totaling at least 250 base pairs (bp) in length, each with a statistically significant (p-value≤0.05) reduction of Isw2 signal in each mutant relative to WT and averaging at least 1.65 fold-change were marked. Reported regions represent the directly overlapping areas of forward and reverse comparisons. Annotated regions are available for download at http://labs.fhcrc.org/tsukiyama

RT-qPCR

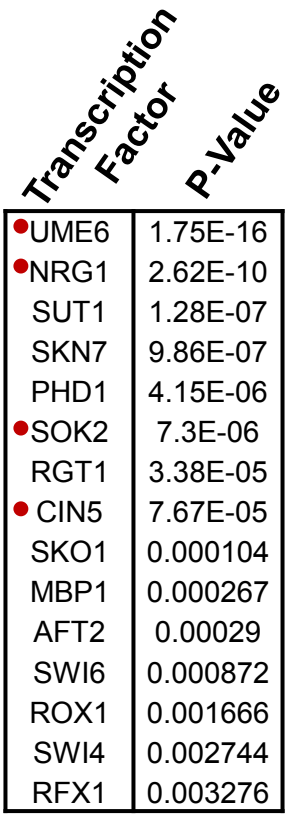
RNA harvest and reverse transcription was performed as previously described (Yadon et al., 2010). qPCR was performed using an ABI 7900HT Real Time PCR System and 384 well plates in triplicate for each cDNA sample. The relative quantification of the average triplicate Ct values for each gene was performed using the paired average Ct values of Act1 as the calibrator and then normalized to WT. The average and standard deviations displayed are the results from two independent qPCR reactions performed on two dilutions of cDNA derived from two biological replicates for each strain.

Chromosome Confirmation Capture

DNA loops were analyzed by a modified version of 3C (Dekker, 2006; Dekker et al., 2002), as described elsewhere (Singh et al., 2009). PCR reactions (40 cycles) were performed using the indicated tandem or convergent primer pairs (Table 1). Control PCR products were generated using a convergent primer pair corresponding to an intergenic region of chromosome V (Ahn et al., 2004). Control template DNA was generated by digesting yeast genomic DNA with *MspI* and ligating in high concentration and used to check primer pair efficiency (Dekker, 2006). PCR products were fractionated on a 1.5% agarose gel, visualized, and quantified by ethidium bromide staining using an AlphaImager 2000. Control reactions were also performed to establish that PCR products are dependent on formaldehyde cross-linking, restriction digestion, and subsequent ligation (Figure 27). PCR primer sequences are listed in (Table 1).

Microarray Data Accession

All raw and processed microarray data is available for download at Gene Expression Omnibus (GEO), http://www.ncbi.nlm.nih.gov/projects/geo/, using accession number GSE39542.



Transcription Factor	P-Value
● UME6	1.75E-16
● NRG1	2.62E-10
SUT1	1.28E-07
SKN7	9.86E-07
PHD1	4.15E-06
● SOK2	7.3E-06
RGT1	3.38E-05
● CIN5	7.67E-05
SKO1	0.000104
MBP1	0.000267
AFT2	0.00029
SWI6	0.000872
ROX1	0.001666
SWI4	0.002744
RFX1	0.003276

Figure 20: Transcription Factors Enriched at 5'-End Isw2 Targets
Transcription factors whose annotated binding sites were found to be statistically enriched (P-Value, hypergeometric distribution) upstream of Isw2 targets located at the 5'-end of a gene. Red dots indicate transcription factor mutants used in this study.

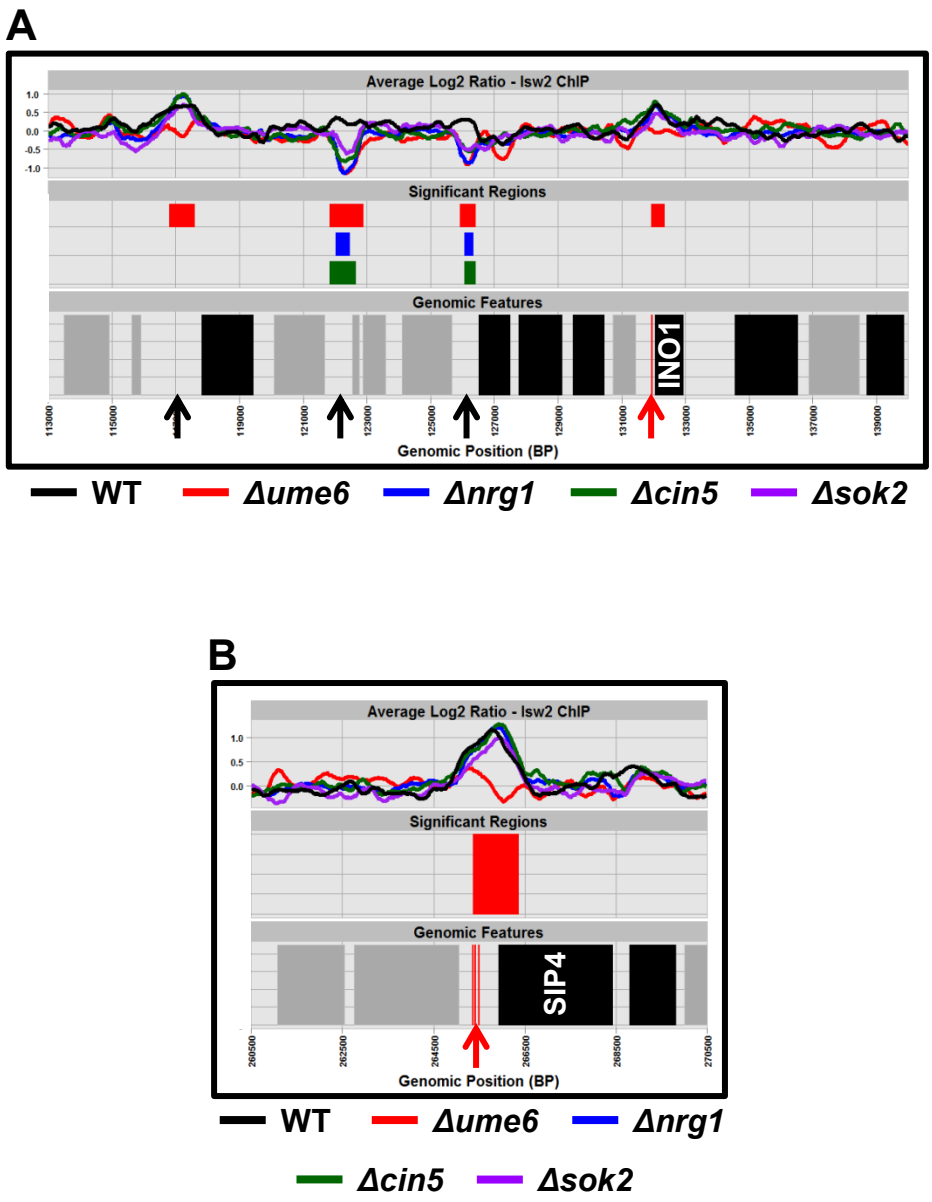


Figure 21: Ume6, Nrg1, Cin5, and Sok2 Targeting Isw2 Genome-Wide
(A) and **(B)** The average log2 Isw2 ChIP signal for WT (black), *Δume6* (red), *Δnrg1* (blue), *Δcin5* (green), and *Δsok2* (purple) strains is displayed in the "Average Log2 Ratio – Isw2 ChIP" track. In the "Significant Regions" track colored boxes indicate annotated regions having significantly different Isw2 ChIP signal between WT and mutant strains (box color is the same as ChIP signal color). Non-dubious genes are displayed as black (Watson strand) and grey (Crick strand) boxes in "Genomic Features" track. Loss of Isw2 ChIP enrichment is highlighted by arrows (red, Ume6 binding sites, and black, no Ume6 binding site) **(A)** ChrVII:113000-140000 **(B)** ChrX:260500-270500

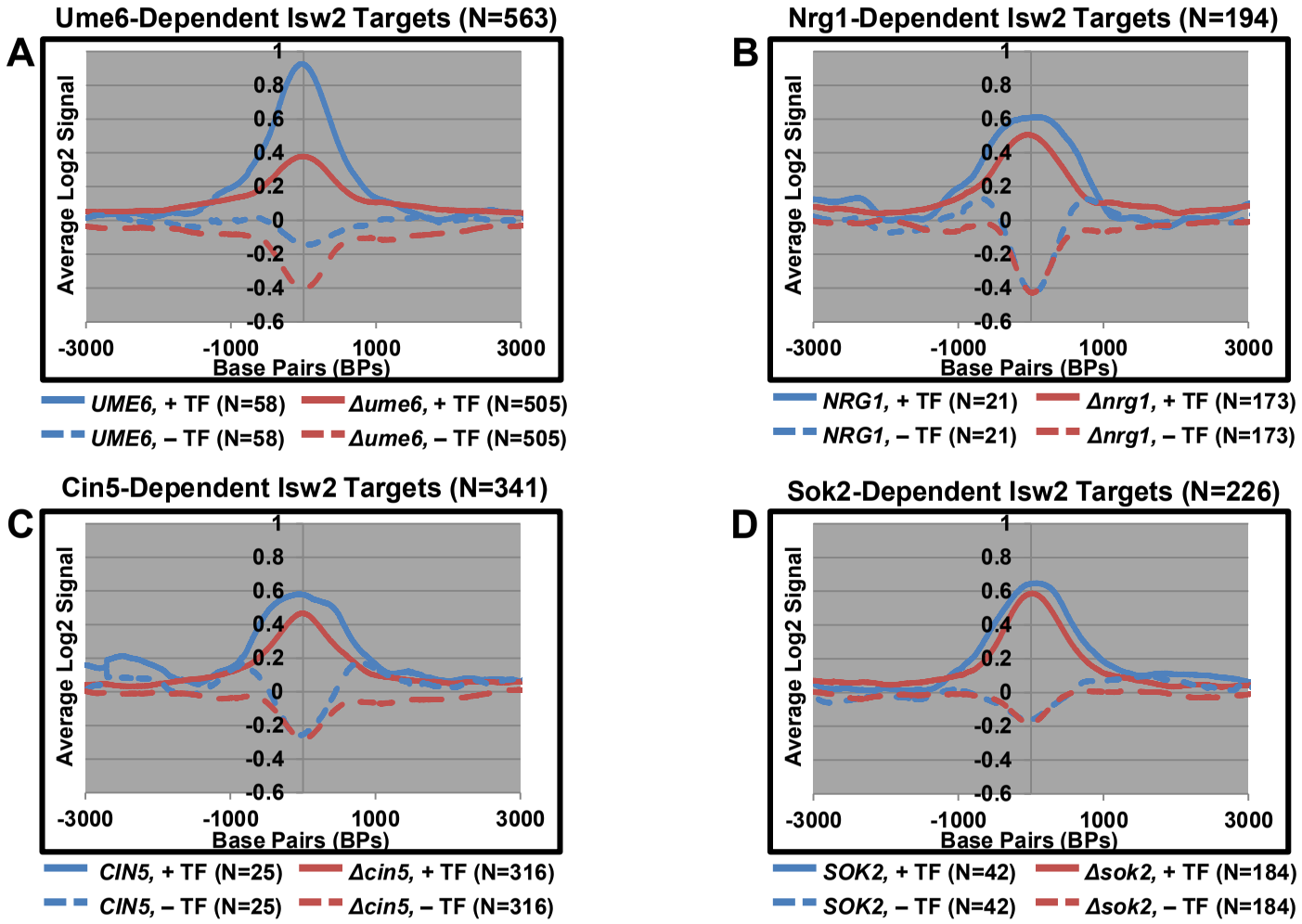


Figure 22: Isw2 ChIP Signals at TF-Dependent Isw2 Targets

(A), **(B)**, **(C)**, and **(D)** Annotated regions from each mutant were aligned by their midpoint and the average log2 Isw2 ChIP signal in WT (solid lines) or transcription factor null mutants (dashed lines) is displayed across 6,000 bp. Regions containing an annotated binding site are displayed in blue and regions without an annotated binding site displayed in red. **(A)** *Δume6* vs. WT annotated regions. **(B)** *Δnrg1* vs. WT annotated regions. **(C)** *Δcin5* vs. WT annotated regions. **(D)** *Δsok2* vs. WT annotated regions.

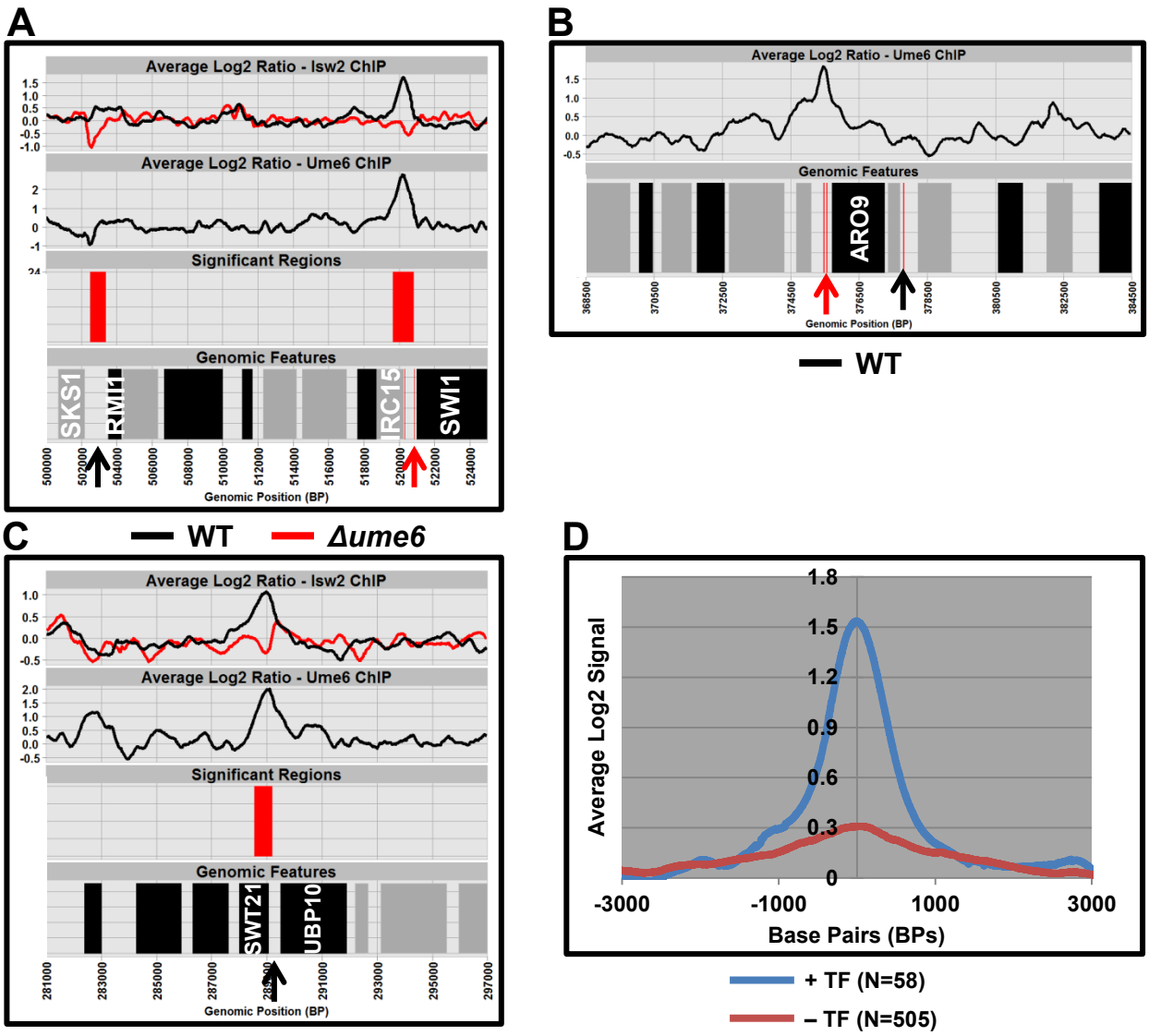


Figure 23: Ume6 Binding Does Not Account for All Ume6-Dependent Isw2 Targets

(A), (B), and **(C)** The average log2 Isw2 ChIP signal for *UME6* (black) and *Δume6* (red) strains is displayed in the "Average Log2 Ratio – Isw2 ChIP" track . The average log2 Ume6 ChIP signal for *UME6* (black) strains is displayed in the "Average Log2 Ratio – Ume6 ChIP" track. Red boxes (in "Significant Regions" track) indicate annotated regions having significantly different Isw2 ChIP signal between *UME6* and *Δume6* strains. Non-dubious genes are displayed as black (Watson strand) and grey (Crick strand) boxes in "Genomic Features" track. Loss of Isw2 ChIP enrichment is highlighted by arrows (red, Ume6 binding sites, and black, no Ume6 binding site) **(D)** The average log2 Ume6 ChIP signal in *Ume6* strains is plotted for a 6000 bp window centered around Ume6-dependent Isw2 targets. Targets containing an annotated Ume6 binding site (+TF, N=58, blue line) and not containing an annotated Ume6 binding site (– TF, N=505, red line). **(A)** ChrXVI:500000-525000 **(B)** ChrVIII:368500-384500 **(C)** ChrXIV:281000-297000

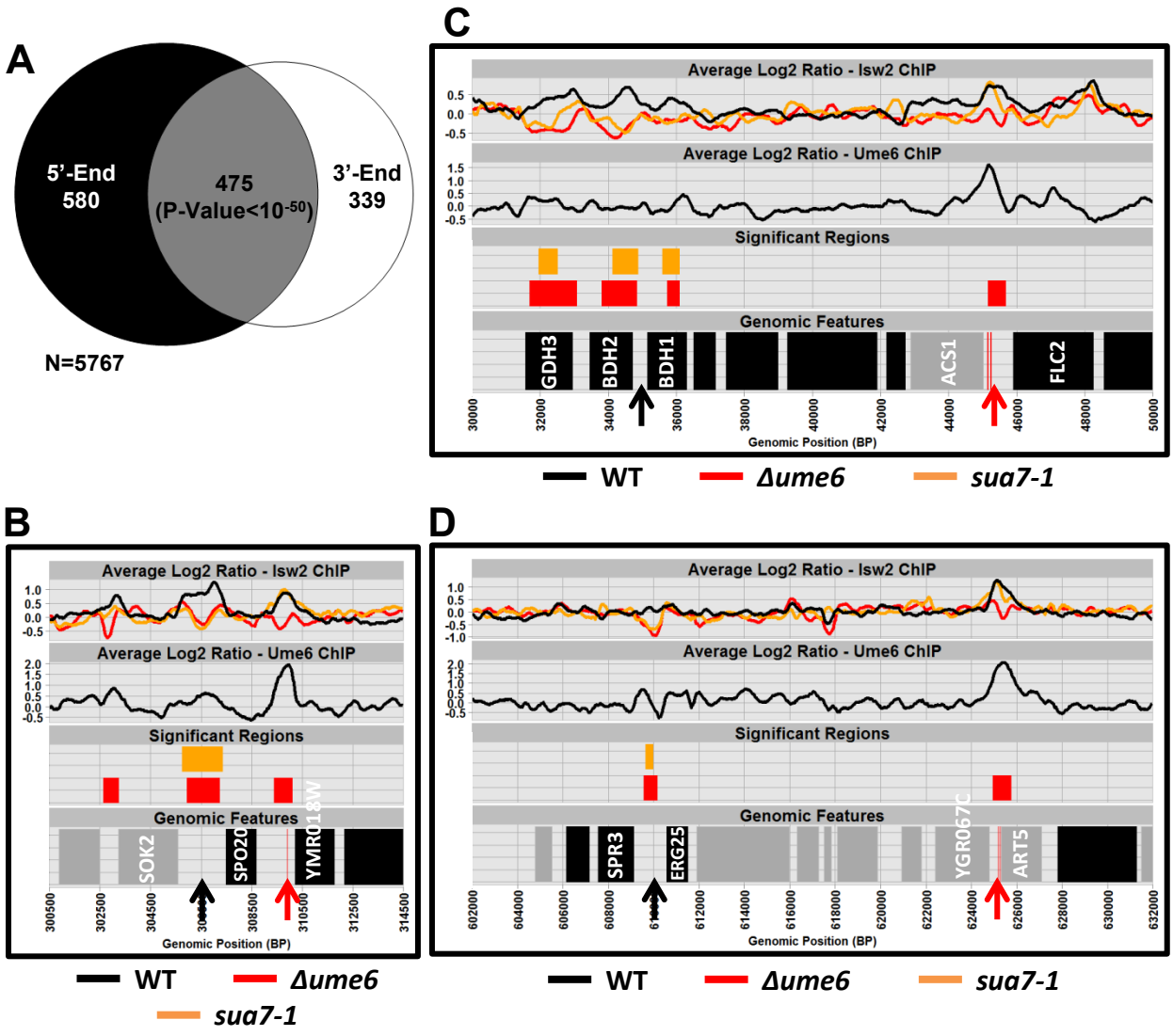


Figure 24: Evidence for DNA Looping-Dependent Isw2 Targeting
(A) Venn diagram displaying the number of genes with annotated Isw2 targets located only at the 5'-end of the gene (5'-End), only at the 3'-end of the gene (3'-End), or at both the 5'- and 3'-end of the same gene (overlapping region). P-value was calculated with a hypergeometric distribution using non-dubious genes only. **(B)**, **(C)**, and **(D)** The average log2 Isw2 ChIP signal for *UME6* (WT, black), *Δume6* (red), and *sua7-1* (orange) strains is displayed in the "Average Log2 Ratio – Isw2 ChIP" track . The average log2 Ume6 ChIP signal for *UME6* (WT, black) strains is displayed in the "Average Log2 Ratio – Ume6 ChIP" track. Red and orange boxes (in "Significant Regions" track) indicate annotated regions having significantly different Isw2 ChIP signal between *UME6* and *Δume6* or *sua7-1* strains, respectively. Non-dubious genes are displayed as black (Watson strand) and grey (Crick strand) boxes in "Genomic Features" track. Loss of Isw2 ChIP enrichment is highlighted by arrows (red, Ume6 binding sites, and black, no Ume6 binding site) **(B)** ChrXIII:300500-314500 **(C)** ChrI:30000-50000 **(D)** ChrVII:602000-632000

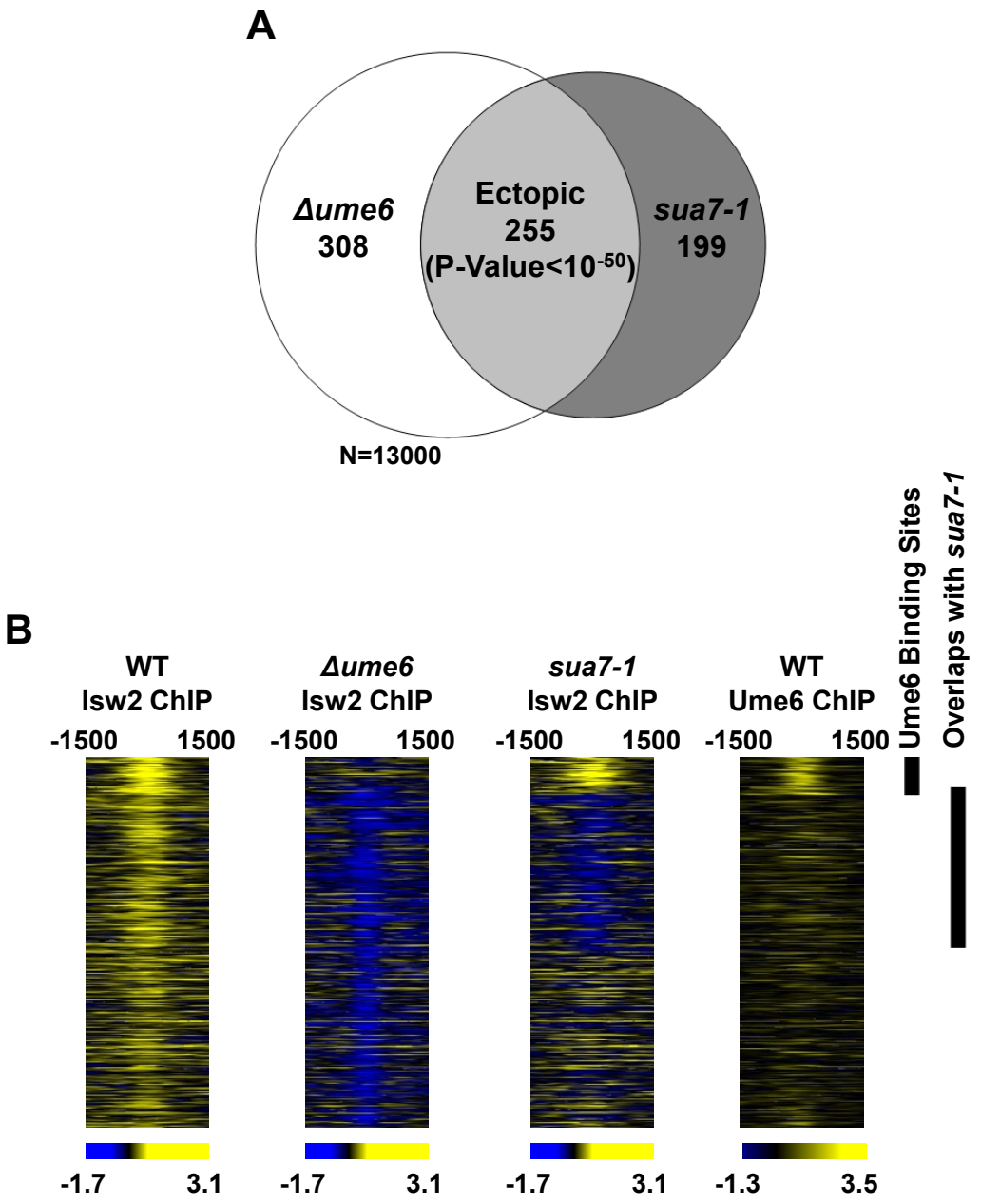


Figure 25: Overlapping Ume6- and TFIIB-Dependent Isw2 Targets Reveal Many Ectopic Isw2 Targets Genome-Wide

(A) Venn diagram displaying the number of Ume6- (*Δume6*) and TFIIB-dependent (*sua7-1*) Isw2 targets that directly overlap (Ectopic regions). P-value was calculated using a hypergeometric distribution. **(B)** The average log2 Isw2 ChIP signal in *UME6* (WT Isw2 ChIP), *Δume6* (*Δume6* Isw2 ChIP), or *sua7-1* (*sua7-1* Isw2 ChIP) strains or the average log2 Ume6 ChIP signal in *UME6* (WT Ume6 ChIP) strains for Ume6-dependent Isw2 targets is displayed for a 3000 bp window centered around the midpoint. Regions containing an annotated Ume6 binding site and/or that directly overlap with an annotated *sua7-1* region are marked.

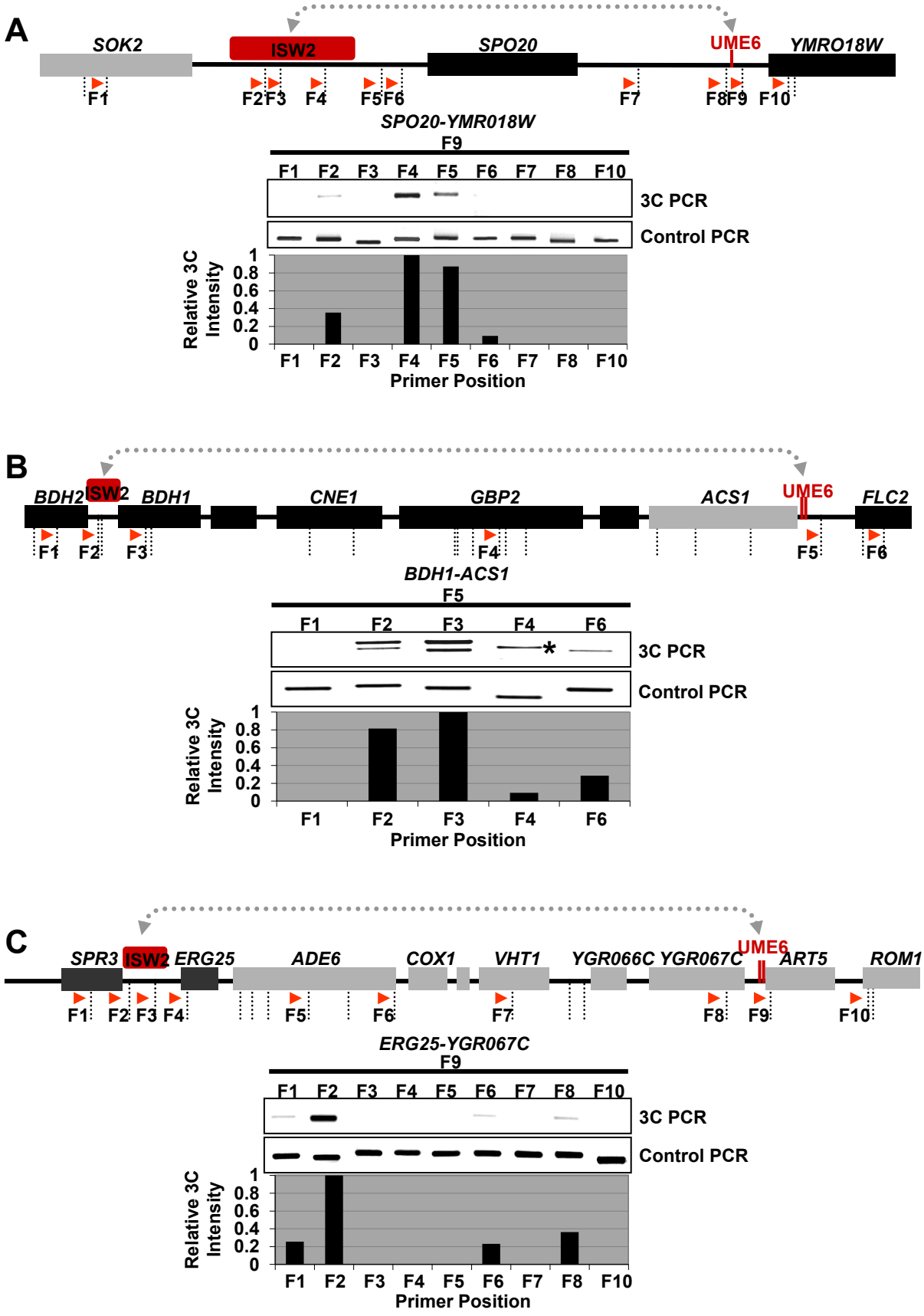
Figure 26: DNA Loops are Formed Between Canonical and Ectopic Isw2 Targets in WT Strains

(A) (B) and (**C)** Schematic diagram of each locus analyzed by 3C is displayed at the top. Watson and Crick genes are displayed as black and grey boxes. Ectopic Isw2 targets and Ume6 binding sites are marked in red. Arrows depict primer locations and directionalities and are numbered below. Double headed, dashed arrows indicate looping between canonical and ectopic Isw2 targets. Ethidium bromide staining of 3C and Control PCRs are displayed and quantitated using the indicated primer pairs for each locus. Asterisk indicates non-specific background band and was not included in the quantitation.

(A) Gene looping locus *SPO20-YMR018W*

(B) Chromatin looping locus *BDH1-ACS1*

(C) Chromatin looping locus *ERG25-YGR067C*


A
SOK2
ISW2
SPO20
UME6
YMRO18W
F1
F2 F3
F4
F5 F6
F7
F8 F9
F10
SPO20-YMR018W
F9
F1 F2 F3 F4 F5 F6 F7 F8 F10
3C PCR
Control PCR
Relative 3C Intensity
1
0.8
0.6
0.4
0.2
0
F1 F2 F3 F4 F5 F6 F7 F8 F10
Primer Position
B
BDH2
ISW2
BDH1
CNE1
GBP2
ACS1
UME6
FLC2
F1
F2
F3
F4
F5
F6
BDH1-ACS1
F5
F1 F2 F3 F4 F6
3C PCR
Control PCR
Relative 3C Intensity
1
0.8
0.6
0.4
0.2
0
F1 F2 F3 F4 F6
Primer Position
C
SPR3
ISW2
ERG25
ADE6
COX1
VHT1
YGR066C
YGR067C
UME6
ART5
ROM1
F1
F2
F3
F4
F5
F6
F7
F8
F9
F10
ERG25-YGR067C
F9
F1 F2 F3 F4 F5 F6 F7 F8 F10
3C PCR
Control PCR
Relative 3C Intensity
1
0.8
0.6
0.4
0.2
0
F1 F2 F3 F4 F5 F6 F7 F8 F10
Primer Position

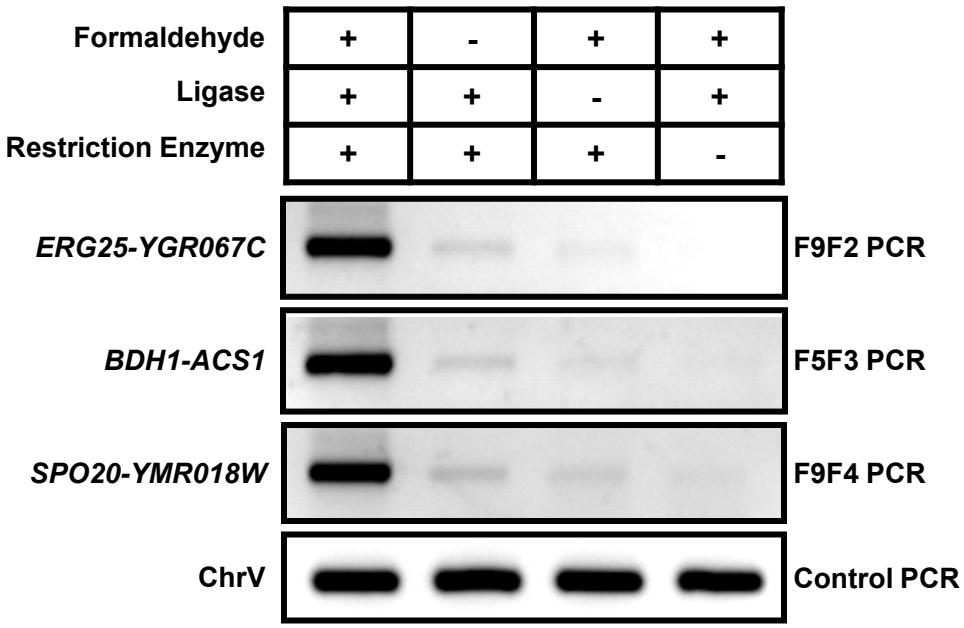


Figure 27: 3C PCR Products are Cross-Linking, Ligation, and Restriction Digestion Dependent

3C PCR products from the canonical and ectopic or control primer pairs at the indicated locus are shown for samples with (+) or without (-) either formaldehyde, ligase, or restriction enzyme.

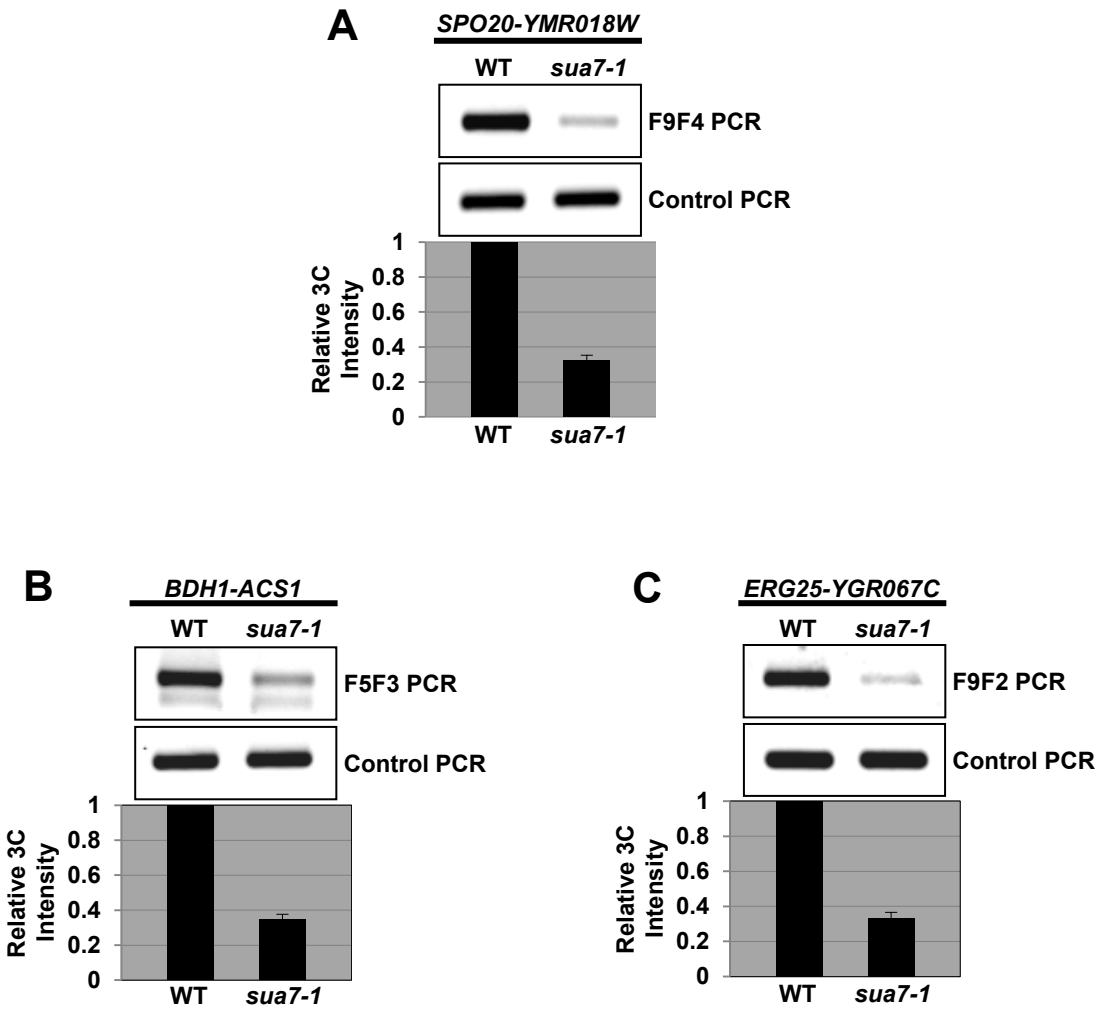


Figure 28: DNA Looping Between Canonical and Ectopic Isw2 Targets are Significantly Reduced in *sua7-1* Strains

(A) (B) and (**C**) 3C PCR products (top) and quantitation (bottom) using the canonical and ectopic or control primer pairs at the indicated locus are shown for WT and sua7-1 strains

(A) Gene looping locus *SPO20-YMR018W*

(B) Chromatin looping locus *BDH1-ACS1*

(C) Chromatin looping locus *ERG25-YGR067C*

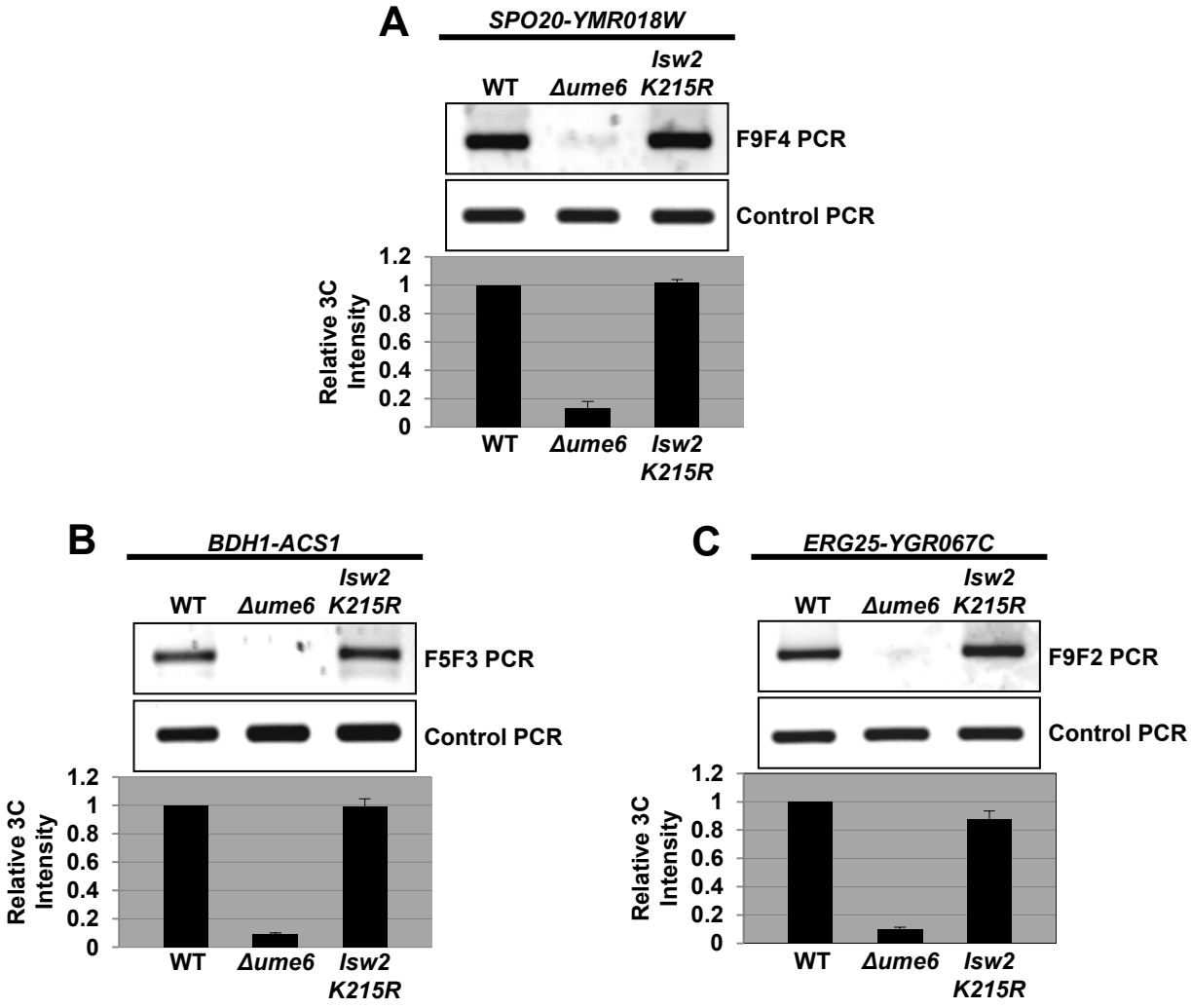


Figure 29: DNA Looping Between Canonical and Ectopic Isw2 Targets are Significantly Reduced in *Δume6,* but not Isw2-K215R Strains

(A) (B) and (**C)** 3C PCR products (top) and quantitation (bottom) using the canonical and ectopic or control primer pairs at the indicated locus are shown for WT, *Δume6*, and Isw2-K215R strains

(A) Gene looping locus *SPO20-YMR018W*

(B) Chromatin looping locus *BDH1-ACS1*

(C) Chromatin looping locus *ERG25-YGR067C*

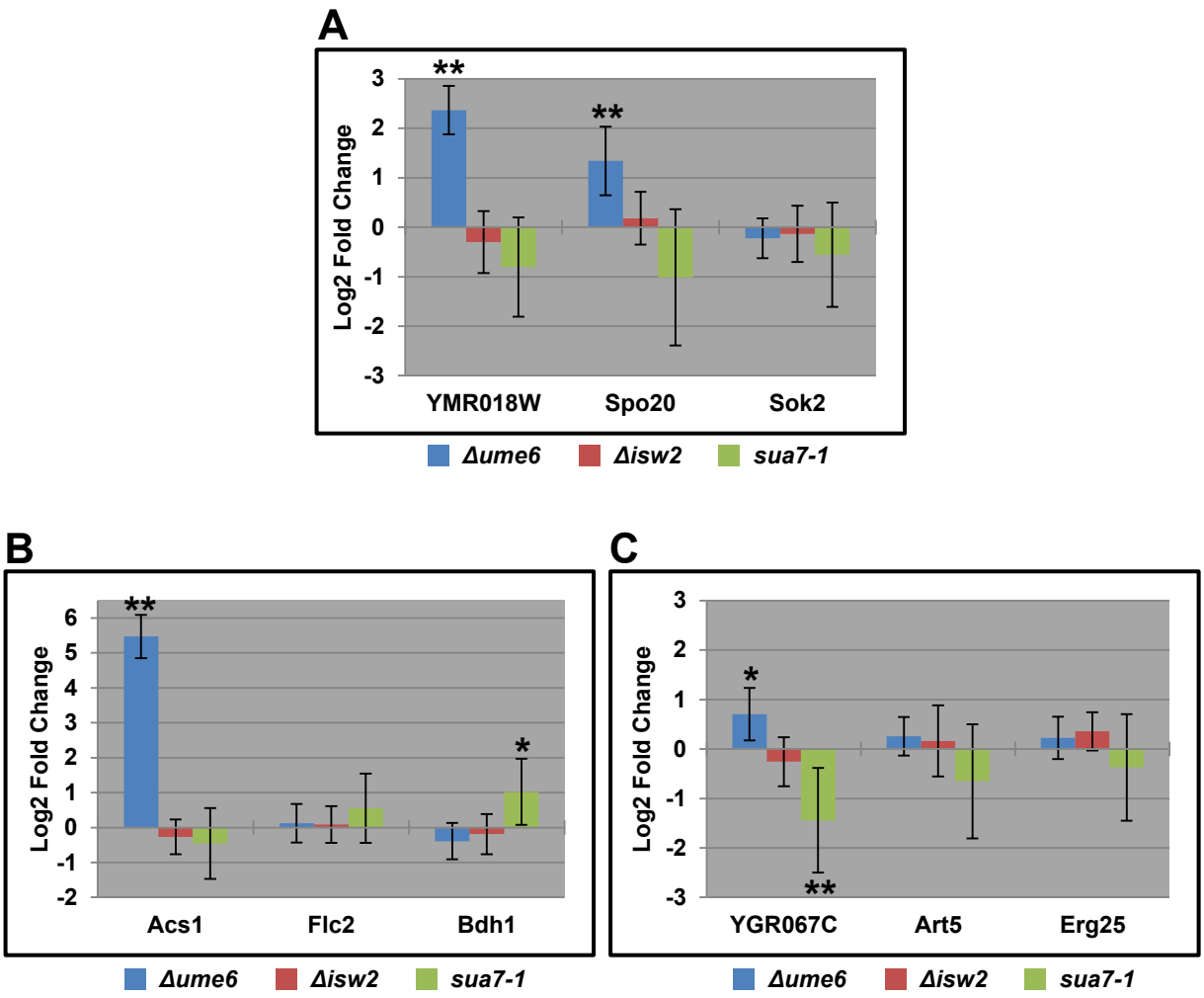


Figure 30: Gene Expression of Canonical and Ectopic Isw2 Targets by qPCR
(A) (B) and (**C**) Normalized log2 fold change of the genes immediately adjacent to canonical and ectopic Isw2 targets in the *Δume6* (blue), *Δisw2* (red), and *sua7-1* (green) relative to WT control (*, p-value≤0.05, **, p-value≤0.01)
(A) Gene looping locus *SPO20-YMR018W*
(B) Chromatin looping locus *BDH1-ACS1*
(C) Chromatin looping locus *ERG25-YGR067C*

Table 1: List of 3C DNA Primers

Name - Tandem Primers	Sequence (5'-3')
SPO20-YMR018W_F1	GGAAGGTTGTGGATAGTTTGAC
SPO20-YMR018W_F2	CGAAGAACTGAGTGAAGGC
SPO20-YMR018W_F3	CGAAGAACTGAGTGAAGGC
SPO20-YMR018W_F4	GTTTTGCCCATTCAGAAACAATATC
SPO20-YMR018W_F5	GCCATTGTTGAAGTACGGAATC
SPO20-YMR018W_F6	CCGGTTAAGGCTGTTAGTCAG
SPO20-YMR018W_F7	AAGAAAGATAGATACCGCTGTC
SPO20-YMR018W_F8	CCTACAAGAAATGTACACGAC
SPO20-YMR018W_F9	TTTTGATGTGCACCTGGGAAC
SPO20-YMR018W_F10	ACTTGTTCCATAACTGGTGAC
BDH1-ACS1_F1	GAAGATGGACACACACATGAG
BDH1-ACS1_F2	AGGCTAAACTGCGTCACAC
BDH1-ACS1_F3	GGTCCTAAAGTGACAAAGGTG
BDH1-ACS1_F4	CAAATAGGAGAGCTAGCGTTC
BDH1-ACS1_F5	GAATAATGGCATCGGAGGCTC
BDH1-ACS1_F6	AGGCTGAGCTGCTTACTTAC
ERG25-YGR067C_F1	TCCAAGAAGAACAACCTGATAGG
ERG25-YGR067C_F2	AGCCAATAATTTCGCGCAAG
ERG25-YGR067C_F3	GGGCCTACGTAGAAGATGATG
ERG25-YGR067C_F4	GCCGTTTTCAACAACGCTAC
ERG25-YGR067C_F5	GATTTAAAGATTTCACATCAGCAGC
ERG25-YGR067C_F6	GCAATTTAGTGTCCTGTTCAGAC
ERG25-YGR067C_F7	GCTGGAGTACAAAATACTTTCTTCC
ERG25-YGR067C_F8	GTAGTGATGAGAACAGGTCTGG
ERG25-YGR067C_F9	GGGATCAGTCATCGTCGCAG
ERG25-YGR067C_F10	GTCGTTGGTAAGAATTTGGTTAGG

Name - Convergent Primers	Sequence (5'-3')
ChrV F	GGCTGTCAGAATATGGGGCCGTAGTA
ChrV R	CACCCCGAAGCTGCTTTCACAATAC

CONCLUSIONS AND FUTURE DIRECTIONS

As described in this dissertation, I have utilized genome-wide molecular biology techniques to elucidate and characterize the *in vivo* functions for an ATP-dependent chromatin remodeling enzyme and DNA looping. This work has further extended our understanding of the molecular mechanisms underlying both Isw2-dependent chromatin remodeling and transcriptional regulation and describes novel physiological roles for DNA looping in mediating Isw2 targeting genome-wide. These results reveal new insights into the relationships that link our knowledge of the three-dimensional packaging of genomes into eukaryotic nuclei and our understanding of the linear maps of genomic information to the regulatory mechanisms governing chromatin structure and DNA-dependent processes.

Major Findings and Conclusions

A consensus set of nucleosome free regions (NFRs) was systematically annotated across the genome of the budding yeast *S. cerevisiae* from multiple nucleosome mapping studies. In addition to the previously identified NFRs located within promoters (5'-NFRs) and transcription termination regions (3'-NFRs), two novel classes of NFRs located within open reading frames (ORF-NFRs) and far from open reading frames (Other-NFRs) were characterized. The targets of the ATP-dependent chromatin remodeling enzyme Isw2 were found significantly enriched at all classes of

NFRs, thus identifying a previously unknown target of Isw2, ORF-NFRs. Importantly, at all classes of NFRs Isw2 function is required to restrict access to the underlying DNA by sliding nucleosomes toward the centers of NFRs to reduce their size. This is the first example in which a chromatin remodeling enzyme has been shown to be required for decreasing the size of NFRs *in vivo*.

Transcriptional profiling using strand-specific tiled microarrays further revealed that Isw2 is globally required to repress cryptic non-coding RNA (ncRNA) transcription from initiating around NFRs. Isw2-dependent chromatin remodeling reduced the size of these target NFRs, effectively positioning nucleosomes to occlude access to the cryptic transcription start site. Isw2-dependent repression of some cryptic transcripts was required to prevent transcriptional interference, providing an important biological consequence for proper transcriptional regulation by Isw2-depedent chromatin remodeling.

Numerous sequence-specific transcription factors (TFs) were found to be statistically enriched at Isw2 targets located at the 5'-end of genes. Genome-wide chromatin immunoprecipitation of Isw2 (Isw2 ChIP-chip) in the TF null mutants, *Δume6*, *Δnrg1*, *Δcin5*, and *Δsok2,* revealed a widespread role for each of these factors in the targeting of Isw2 to numerous loci genome-wide. This result provides the first systematic genome-wide evaluation of TF-mediated targeting for any chromatin remodeling enzyme.

Not all Isw2 targets could be accounted for solely by enriched TF binding sites. Additionally, a highly statistically significant number of genes have Isw2 targeted to both the 5'- and 3'-end of the same gene. These results led us to evaluate the role of DNA

looping in mediating targeting of Isw2. Isw2 ChIP-chip in the DNA looping deficient mutant, *sua7-1*, identified numerous Sua7-dependent targets of Isw2. This revealed a novel role for the general transcription factor TFIIB (Sua7) in facilitating the targeting of Isw2 genome-wide. At many loci a decrease in Isw2 ChIP signals was observed in both *sua7-1* and *Δume6*. These loci, termed ectopic targets, lacked both an annotated Ume6 binding site and Ume6 ChIP enrichment. These results suggested that DNA looping facilitates targeting of Isw2 from a Ume6 binding site to an ectopic target.

Chromosome confirmation capture (3C) demonstrated that DNA loops are formed at three loci between a Ume6 binding site and an ectopic target in WT strains. Importantly, the frequency of DNA loops were significantly reduced in both *Δume6* and *sua7-1* strains at all loci tested. Transcriptional analysis in *Δume6* and *sua7-1* strains revealed a dependence on both Ume6- and TFIIB-mediated DNA looping in maintaining proper transcription regulation at the ectopic locus. These results identified Ume6 as the first repressor required for mediating DNA looping and demonstrated a novel physiological role for DNA looping in mediating the transcription factor dependent targeting of Isw2.

In Vivo Model of Isw2 Function

The results presented in this dissertation combined with previous studies support a primary model for the *in vivo* functions, mechanisms of action, and targeting of the ATP-dependent chromatin remodeling enzyme Isw2. In this model, Isw2 functions as a general transcriptional repressor of both coding and ncRNA that initiate from the edges of NFRs located at the 5'- and 3'-end of target genes. Transcriptional repression is

achieved through the Isw2-dependent sliding of nucleosomes toward the centers of NFRs, which effectively reduces the size of NFRs and occludes access to the underlying transcription start sites. The targeting of Isw2 is mediated directly through the physical interaction of Isw2 with the sequence-specific transcription factor Ume6 which binds its cognate binding sites within promoters. TFIIB- and Ume6-mediated DNA looping further facilitate Ume6-dependent targeting of Isw2 between a Ume6 binding site and a distally located locus.

Future Directions

In light of this model, much is still to be understood regarding the *in vivo* functions, mechanisms of action, and targeting of Isw2. While conceivably there are many potentially very interesting directions with which to further explore, I propose to focus on elucidating additional mechanisms and details regulating the targeting of Isw2.

ChIA-PET: Elucidating Both DNA Looping-Dependent and Independent Isw2 Targets Genome-Wide

The work in this dissertation has suggested a genome-wide role for DNA Looping in facilitating the TF-dependent targeting of Isw2. In fact, based on our analysis, we estimate that at least 255 Isw2 targets (Figure 25) are dependent on DNA looping. This however is likely an underestimation, which is based only on the overlap of Ume6- and TFIIB-dependent Isw2 targets. How can one unbiasedly and systematically identify all Isw2 and DNA looping-dependent targets genome-wide independently of either Ume6 or TFIIB? While, the use of ChIP based techniques are highly sensitive assays for the

identification of protein targets, they only provide linear information about protein binding sites along chromosomes. On the other hand, the 3C assay is capable of analyzing long-distance chromatin interactions, but is extremely low throughput and laborious to perform, preventing the genome-wide characterization of or protein-specific interactions associated with DNA looping.

I propose that combining Isw2 ChIP, 3C, and paired-end tag deep sequencing, using ChIA-PET (Chromatin Interaction Analysis by Paired-End Tag Sequencing) (Fullwood et al., 2009; Fullwood and Ruan, 2009) one could simultaneously identify and characterize both the genome-wide targets of Isw2 and the DNA looping interactions between these sites. In summary, this assay would utilize traditional ChIP techniques by fixing cells using formaldehyde to covalently cross-link and stabilize long-range chromatin interactions with bound DNA-protein interactions. Sonicated chromatin is then enriched for a protein of interest using immunoprecipitation. Next, using 3C based approaches, enriched chromatin is connected with DNA linkers by proximity ligation and finally paired-end tags (PETs) are extracted for deep sequencing. This technique thus allows for the identification of both traditional ChIP fragments, by self-ligation of the same DNA fragments, and DNA looping fragments, by inter-ligation between different DNA fragments, which are all bound by the protein of interest. Additionally, the use of sonication, rather than restriction enzyme digestion, would increase the resolution of the assay and provide better signal to noise ratio (Fullwood and Ruan, 2009). Finally, mapping of sequenced fragments back to the genome and utilization of bioinformatics approaches, would in theory allow for the simultaneous

identification of both DNA looping-dependent and independent Isw2 targets genome-wide.

Furthermore, by performing Isw2 ChIA-PET in TF null mutants, such as *Δume6*, or DNA looping-deficient mutants, such as *sua7-1*, the contribution of Ume6- and/or TFIIB-dependent DNA looping on Isw2 targeting genome-wide could be established. In fact, this analysis may help to explain the uncharacterized mechanisms of Ume6-dependent Isw2 targeting to loci that do not have either a Ume6 binding site nor are TFIIB-dependent (Figure 25). For example, it is possible that TFIIB-independent DNA looping facilitates targeting of Isw2 to these loci. We have already demonstrated (Figure 29) that Ume6 itself is partially required for DNA looping at some loci. The genome-wide contribution of neither Ume6 nor TFIIB in DNA looping has been established. Understanding their global role in DNA looping genome-wide would provide an unprecedented glimpse into how the three-dimensional packing of DNA into eukaryotic nuclei affects chromatin structure.

Tup1-Dependent Isw2 Targeting

Previous studies from the Tsukiyama lab (Gelbart et al., 2005; Goldmark et al., 2000) lab have shown that the physical interactions of Ume6 and Isw2 are required for the targeting of Isw2 to Ume6 binding sites. The mechanisms for Nrg1-, Cin5-, and Sok2-dependent targeting of Isw2 are not known. As mentioned in a previous chapter, we hypothesized that the transcriptional co-repressor complex Tup1-Ssn6 mediates Nrg1-, Cin5-, and Sok2-dependent targeting of Isw2. Because neither Tup1 nor Ssn6 has sequence-specific DNA binding activity their role in transcription depends on their

targeting by sequence-specific DNA binding proteins (Keleher et al., 1992; Komachi et al., 1994). It is therefore possible that the role for Nrg1, Cin5, and Sok2 in mediating Isw2 targeting is dependent on their ability to actively recruit Tup1 to promoters (Hanlon et al., 2011). It is interesting to further note that Tup1 functionally represses target genes in part through interactions with multiple histone deacetylase complexes, including Rpd3 (Davie et al., 2003; Watson, 2000). Isw2 is known to function in a parallel pathway as Rpd3 in a Ume6-dependent manner (Goldmark et al., 2000). These findings further suggest that the co-repressor complex Tup1-Ssn6 is likely directly involved in Isw2 targeting.

I propose to test this hypothesis by performing genome-wide Isw2 ChIP-chip in Tup1 null mutants. By comparing *Δtup1* to *TUP1* the genome-wide contribution of Tup1 to Isw2 targeting can be determined. Based on our hypothesis, I would expect a very large statistically significant overlap between Tup1-dependent Isw2 targets and the previously characterized Nrg1-, Cin5-, and Sok2-dependent Isw2 targets. Co-immunoprecipitation experiments of Isw2 with the Tup1-Ssn6 complex could then be performed to verify that the physical interactions with Tup1 are required for Isw2 targeting, similar to those observed for Ume6 (Goldmark et al., 2000).

The targeting mechanism for ISWI homologues in higher eukaryotes is extremely poorly understood. Intriguingly, Ume6 is not conserved in plants or animals, however both Tup1 (Braun and Johnson, 1997; Fisher and Caudy, 1998; Grbavec et al., 1999; Levanon et al., 1998; Liu and Karmarkar, 2008; Pflugrad et al., 1997; Todd et al., 2003) and Rpd3 are highly conserved (Yang and Seto, 2008). It is interesting to speculate that perhaps Tup1 homologues may also be required for the targeting of ISWI

complexes in higher eukaryotes, just as they are for Rpd3 homologues (Chen et al., 1999; Choi et al., 1999; Mannervik and Levine, 1999; Sekiya and Zaret, 2007; Winkler et al., 2010). Therefore, the results obtained from these experiments would likely have far reaching implications, not just in S. *cerevisiae*, but also in plant and animal cells.

REFERENCES

Adkins, M.W., Williams, S.K., Linger, J., and Tyler, J.K. (2007). Chromatin disassembly from the PHO5 promoter is essential for the recruitment of the general transcription machinery and coactivators. Molecular and cellular biology *27*, 6372-6382.

Ahn, S.H., Kim, M., and Buratowski, S. (2004). Phosphorylation of serine 2 within the RNA polymerase II C-terminal domain couples transcription and 3' end processing. Molecular cell *13*, 67-76.

Albert, I., Mavrich, T.N., Tomsho, L.P., Qi, J., Zanton, S.J., Schuster, S.C., and Pugh, B.F. (2007). Translational and rotational settings of H2A.Z nucleosomes across the Saccharomyces cerevisiae genome. Nature *446*, 572-576.

Almer, A., and Horz, W. (1986). Nuclease hypersensitive regions with adjacent positioned nucleosomes mark the gene boundaries of the PHO5/PHO3 locus in yeast. The EMBO journal *5*, 2681-2687.

Almer, A., Rudolph, H., Hinnen, A., and Horz, W. (1986). Removal of positioned nucleosomes from the yeast PHO5 promoter upon PHO5 induction releases additional upstream activating DNA elements. The EMBO journal *5*, 2689-2696.

Andersen, E.C., Lu, X., and Horvitz, H.R. (2006). C. elegans ISWI and NURF301 antagonize an Rb-like pathway in the determination of multiple cell fates. Development *133*, 2695-2704.

Anderson, J.D., and Widom, J. (2001). Poly(dA-dT) promoter elements increase the equilibrium accessibility of nucleosomal DNA target sites. Molecular and cellular biology *21*, 3830-3839.

Anderson, S.F., Steber, C.M., Esposito, R.E., and Coleman, J.E. (1995). UME6, a negative regulator of meiosis in Saccharomyces cerevisiae, contains a C-terminal Zn2Cys6 binuclear cluster that binds the URS1 DNA sequence in a zinc-dependent manner. Protein science : a publication of the Protein Society *4*, 1832-1843.

Ansari, A., and Hampsey, M. (2005). A role for the CPF 3'-end processing machinery in RNAP II-dependent gene looping. Genes & development *19*, 2969-2978.

Arents, G., Burlingame, R.W., Wang, B.C., Love, W.E., and Moudrianakis, E.N. (1991). The nucleosomal core histone octamer at 3.1 A resolution: a tripartite protein assembly and a left-handed superhelix. Proceedings of the National Academy of Sciences of the United States of America *88*, 10148-10152.

Bachman, N., Gelbart, M.E., Tsukiyama, T., and Boeke, J.D. (2005). TFIIIB subunit Bdp1p is required for periodic integration of the Ty1 retrotransposon and targeting of Isw2p to S. cerevisiae tDNAs. Genes & development *19*, 955-964.

Badenhorst, P., Voas, M., Rebay, I., and Wu, C. (2002). Biological functions of the ISWI chromatin remodeling complex NURF. Genes & development *16*, 3186-3198.

Badis, G., Chan, E.T., van Bakel, H., Pena-Castillo, L., Tillo, D., Tsui, K., Carlson, C.D., Gossett, A.J., Hasinoff, M.J., Warren, C.L.*, et al.* (2008). A library of yeast transcription factor motifs reveals a widespread function for Rsc3 in targeting nucleosome exclusion at promoters. Molecular cell *32*, 878-887.

Barak, O., Lazzaro, M.A., Lane, W.S., Speicher, D.W., Picketts, D.J., and Shiekhattar, R. (2003). Isolation of human NURF: a regulator of Engrailed gene expression. The EMBO journal *22*, 6089-6100.

Barski, A., Cuddapah, S., Cui, K., Roh, T.Y., Schones, D.E., Wang, Z., Wei, G., Chepelev, I., and Zhao, K. (2007). High-resolution profiling of histone methylations in the human genome. Cell *129*, 823-837.

Basehoar, A.D., Zanton, S.J., and Pugh, B.F. (2004). Identification and distinct regulation of yeast TATA box-containing genes. Cell *116*, 699-709.

Bernstein, B.E., Liu, C.L., Humphrey, E.L., Perlstein, E.O., and Schreiber, S.L. (2004). Global nucleosome occupancy in yeast. Genome biology *5*, R62.

Berr, A., Pecinka, A., Meister, A., Kreth, G., Fuchs, J., Blattner, F.R., Lysak, M.A., and Schubert, I. (2006). Chromosome arrangement and nuclear architecture but not centromeric sequences are conserved between Arabidopsis thaliana and Arabidopsis lyrata. The Plant journal : for cell and molecular biology *48*, 771-783.

Bickmore, W.A., Mahy, N.L., and Chambeyron, S. (2004). Do higher-order chromatin structure and nuclear reorganization play a role in regulating Hox gene expression during development? Cold Spring Harbor symposia on quantitative biology *69*, 251-257.

Bolzer, A., Kreth, G., Solovei, I., Koehler, D., Saracoglu, K., Fauth, C., Muller, S., Eils, R., Cremer, C., Speicher, M.R.*, et al.* (2005). Three-dimensional maps of all chromosomes in human male fibroblast nuclei and prometaphase rosettes. PLoS biology *3*, e157.

Branco, M.R., and Pombo, A. (2006). Intermingling of chromosome territories in interphase suggests role in translocations and transcription-dependent associations. PLoS biology *4*, e138.

Braun, B.R., and Johnson, A.D. (1997). Control of filament formation in Candida albicans by the transcriptional repressor TUP1. Science *277*, 105-109.

Bruno, M., Flaus, A., Stockdale, C., Rencurel, C., Ferreira, H., and Owen-Hughes, T. (2003). Histone H2A/H2B dimer exchange by ATP-dependent chromatin remodeling activities. Molecular cell *12*, 1599-1606.

Bryant, G.O., Prabhu, V., Floer, M., Wang, X., Spagna, D., Schreiber, D., and Ptashne, M. (2008). Activator control of nucleosome occupancy in activation and repression of transcription. PLoS biology *6*, 2928-2939.

Bystricky, K., Laroche, T., van Houwe, G., Blaszczyk, M., and Gasser, S.M. (2005). Chromosome looping in yeast: telomere pairing and coordinated movement

reflect anchoring efficiency and territorial organization. The Journal of cell biology *168*, 375-387.

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.*, et al.* (2005). Histone H3 methylation by Set2 directs deacetylation of coding regions by Rpd3S to suppress spurious intragenic transcription. Cell *123*, 581-592.

Chambeyron, S., and Bickmore, W.A. (2004a). Chromatin decondensation and nuclear reorganization of the HoxB locus upon induction of transcription. Genes & development *18*, 1119-1130.

Chambeyron, S., and Bickmore, W.A. (2004b). Does looping and clustering in the nucleus regulate gene expression? Current opinion in cell biology *16*, 256-262.

Chen, G., Fernandez, J., Mische, S., and Courey, A.J. (1999). A functional interaction between the histone deacetylase Rpd3 and the corepressor groucho in Drosophila development. Genes & development *13*, 2218-2230.

Cheung, V., Chua, G., Batada, N.N., Landry, C.R., Michnick, S.W., Hughes, T.R., and Winston, F. (2008). Chromatin- and transcription-related factors repress transcription from within coding regions throughout the Saccharomyces cerevisiae genome. PLoS biology *6*, e277.

Choi, C.Y., Kim, Y.H., Kwon, H.J., and Kim, Y. (1999). The homeodomain protein NK-3 recruits Groucho and a histone deacetylase complex to repress transcription. The Journal of biological chemistry *274*, 33194-33197.

Clapier, C.R., and Cairns, B.R. (2009). The biology of chromatin remodeling complexes. Annual review of biochemistry *78*, 273-304.

Collins, N., Poot, R.A., Kukimoto, I., Garcia-Jimenez, C., Dellaire, G., and Varga-Weisz, P.D. (2002). An ACF1-ISWI chromatin-remodeling complex is required for DNA replication through heterochromatin. Nature genetics *32*, 627-632.

Comet, I., Schuettengruber, B., Sexton, T., and Cavalli, G. (2011). A chromatin insulator driving three-dimensional Polycomb response element (PRE) contacts and Polycomb association with the chromatin fiber. Proceedings of the National Academy of Sciences of the United States of America *108*, 2294-2299.

Core, L.J., Waterfall, J.J., and Lis, J.T. (2008). Nascent RNA sequencing reveals widespread pausing and divergent initiation at human promoters. Science *322*, 1845-1848.

Corona, D.F., Langst, G., Clapier, C.R., Bonte, E.J., Ferrari, S., Tamkun, J.W., and Becker, P.B. (1999). ISWI is an ATP-dependent nucleosome remodeling factor. Molecular cell *3*, 239-245.

Cremer, M., Grasser, F., Lanctot, C., Muller, S., Neusser, M., Zinner, R., Solovei, I., and Cremer, T. (2008). Multicolor 3D fluorescence in situ hybridization for imaging interphase chromosomes. Methods in molecular biology *463*, 205-239.

Cremer, T., and Cremer, C. (2001). Chromosome territories, nuclear architecture and gene regulation in mammalian cells. Nature reviews Genetics *2*, 292-301.

Cremer, T., and Cremer, C. (2006a). Rise, fall and resurrection of chromosome territories: a historical perspective. Part I. The rise of chromosome territories. European journal of histochemistry : EJH *50*, 161-176.

Cremer, T., and Cremer, C. (2006b). Rise, fall and resurrection of chromosome territories: a historical perspective. Part II. Fall and resurrection of chromosome territories during the 1950s to 1980s. Part III. Chromosome territories and the functional nuclear architecture: experiments and models from the 1990s to the present. European journal of histochemistry : EJH *50*, 223-272.

Cremer, T., and Cremer, M. (2010). Chromosome territories. Cold Spring Harbor perspectives in biology *2*, a003889.

Damelin, M., Simon, I., Moy, T.I., Wilson, B., Komili, S., Tempst, P., Roth, F.P., Young, R.A., Cairns, B.R., and Silver, P.A. (2002). The genome-wide localization of Rsc9, a component of the RSC chromatin-remodeling complex, changes in response to stress. Molecular cell *9*, 563-573.

David, L., Huber, W., Granovskaia, M., Toedling, J., Palm, C.J., Bofkin, L., Jones, T., Davis, R.W., and Steinmetz, L.M. (2006). A high-resolution map of transcription in the yeast genome. Proceedings of the National Academy of Sciences of the United States of America *103*, 5320-5325.

Davie, J.K., Edmondson, D.G., Coco, C.B., and Dent, S.Y. (2003). Tup1-Ssn6 interacts with multiple class I histone deacetylases in vivo. The Journal of biological chemistry *278*, 50158-50162.

Davis, C.A., and Ares, M., Jr. (2006). Accumulation of unstable promoter-associated transcripts upon loss of the nuclear exosome subunit Rrp6p in Saccharomyces cerevisiae. Proceedings of the National Academy of Sciences of the United States of America *103*, 3262-3267.

Dekker, J. (2006). The three 'C' s of chromosome conformation capture: controls, controls, controls. Nature methods *3*, 17-21.

Dekker, J., Rippe, K., Dekker, M., and Kleckner, N. (2002). Capturing chromosome conformation. Science *295*, 1306-1311.

Deuring, R., Fanti, L., Armstrong, J.A., Sarte, M., Papoulas, O., Prestel, M., Daubresse, G., Verardo, M., Moseley, S.L., Berloco, M.*, et al.* (2000). The ISWI chromatin-remodeling protein is required for gene expression and the maintenance of higher order chromatin structure in vivo. Molecular cell *5*, 355-365.

Dey, P. (2006). Chromatin remodeling, cancer and chemotherapy. Current medicinal chemistry *13*, 2909-2919.

Dirscherl, S.S., Henry, J.J., and Krebs, J.E. (2005). Neural and eye-specific defects associated with loss of the imitation switch (ISWI) chromatin remodeler in Xenopus laevis. Mechanisms of development *122*, 1157-1170.

Dostie, J., Richmond, T.A., Arnaout, R.A., Selzer, R.R., Lee, W.L., Honan, T.A., Rubio, E.D., Krumm, A., Lamb, J., Nusbaum, C.*, et al.* (2006). Chromosome Conformation Capture Carbon Copy (5C): a massively parallel solution for

mapping interactions between genomic elements. Genome research *16*, 1299-1309.

Drissen, R., Palstra, R.J., Gillemans, N., Splinter, E., Grosveld, F., Philipsen, S., and de Laat, W. (2004). The active spatial organization of the beta-globin locus requires the transcription factor EKLF. Genes & development *18*, 2485-2490.

Duan, Z., Andronescu, M., Schutz, K., McIlwain, S., Kim, Y.J., Lee, C., Shendure, J., Fields, S., Blau, C.A., and Noble, W.S. (2010). A three-dimensional model of the yeast genome. Nature *465*, 363-367.

Dutrow, N., Nix, D.A., Holt, D., Milash, B., Dalley, B., Westbroek, E., Parnell, T.J., and Cairns, B.R. (2008). Dynamic transcriptome of Schizosaccharomyces pombe shown by RNA-DNA hybrid mapping. Nature genetics *40*, 977-986.

Ehrenhofer-Murray, A.E. (2004). Chromatin dynamics at DNA replication, transcription and repair. European journal of biochemistry / FEBS *271*, 2335-2349.

Eisen, J.A., Sweder, K.S., and Hanawalt, P.C. (1995). Evolution of the SNF2 family of proteins: subfamilies with distinct sequences and functions. Nucleic acids research *23*, 2715-2723.

El Kaderi, B., Medler, S., Raghunayakula, S., and Ansari, A. (2009). Gene looping is conferred by activator-dependent interaction of transcription initiation and termination machineries. The Journal of biological chemistry *284*, 25015-25025.

Erdel, F., Schubert, T., Marth, C., Langst, G., and Rippe, K. (2010). Human ISWI chromatin-remodeling complexes sample nucleosomes via transient binding reactions and become immobilized at active sites. Proceedings of the National Academy of Sciences of the United States of America *107*, 19873-19878.

Fazzio, T.G., Gelbart, M.E., and Tsukiyama, T. (2005). Two distinct mechanisms of chromatin interaction by the Isw2 chromatin remodeling complex in vivo. Molecular and cellular biology *25*, 9165-9174.

Fazzio, T.G., Kooperberg, C., Goldmark, J.P., Neal, C., Basom, R., Delrow, J., and Tsukiyama, T. (2001). Widespread Collaboration of Isw2 and Sin3-Rpd3 Chromatin Remodeling Complexes in Transcriptional Repression. Molecular and cellular biology *21*, 6450-6460.

Ferrai, C., de Castro, I.J., Lavitas, L., Chotalia, M., and Pombo, A. (2010a). Gene positioning. Cold Spring Harbor perspectives in biology *2*, a000588.

Ferrai, C., Xie, S.Q., Luraghi, P., Munari, D., Ramirez, F., Branco, M.R., Pombo, A., and Crippa, M.P. (2010b). Poised transcription factories prime silent uPA gene prior to activation. PLoS biology *8*, e1000270.

Field, Y., Kaplan, N., Fondufe-Mittendorf, Y., Moore, I.K., Sharon, E., Lubling, Y., Widom, J., and Segal, E. (2008). Distinct modes of regulation by chromatin encoded through nucleosome positioning signals. PLoS computational biology *4*, e1000216.

Fisher, A.L., and Caudy, M. (1998). Groucho proteins: transcriptional corepressors for specific subsets of DNA-binding transcription factors in vertebrates and invertebrates. Genes & development *12*, 1931-1940.

Flaus, A., Martin, D.M., Barton, G.J., and Owen-Hughes, T. (2006). Identification of multiple distinct Snf2 subfamilies with conserved structural motifs. Nucleic acids research *34*, 2887-2905.

Floer, M., Wang, X., Prabhu, V., Berrozpe, G., Narayan, S., Spagna, D., Alvarez, D., Kendall, J., Krasnitz, A., Stepansky, A.*, et al.* (2010). A RSC/nucleosome complex determines chromatin architecture and facilitates activator binding. Cell *141*, 407-418.

Fullwood, M.J., Liu, M.H., Pan, Y.F., Liu, J., Xu, H., Mohamed, Y.B., Orlov, Y.L., Velkov, S., Ho, A., Mei, P.H.*, et al.* (2009). An oestrogen-receptor-alpha-bound human chromatin interactome. Nature *462*, 58-64.

Fullwood, M.J., and Ruan, Y. (2009). ChIP-based methods for the identification of long-range chromatin interactions. Journal of cellular biochemistry *107*, 30-39.

Fyodorov, D.V., Blower, M.D., Karpen, G.H., and Kadonaga, J.T. (2004). Acf1 confers unique activities to ACF/CHRAC and promotes the formation rather than disruption of chromatin in vivo. Genes & development *18*, 170-183.

Gelbart, M.E., Bachman, N., Delrow, J., Boeke, J.D., and Tsukiyama, T. (2005). Genome-wide identification of Isw2 chromatin-remodeling targets by localization of a catalytically inactive mutant. Genes & development *19*, 942-954.

Gelbart, M.E., Rechsteiner, T., Richmond, T.J., and Tsukiyama, T. (2001). Interactions of Isw2 chromatin remodeling complex with nucleosomal arrays: analyses using recombinant yeast histones and immobilized templates. Molecular and cellular biology *21*, 2098-2106.

Gheldof, N., Smith, E.M., Tabuchi, T.M., Koch, C.M., Dunham, I., Stamatoyannopoulos, J.A., and Dekker, J. (2010). Cell-type-specific long-range looping interactions identify distant regulatory elements of the CFTR gene. Nucleic acids research *38*, 4325-4336.

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

Goldmark, J.P., Fazzio, T.G., Estep, P.W., Church, G.M., and Tsukiyama, T. (2000). The Isw2 chromatin remodeling complex represses early meiotic genes upon recruitment by Ume6p. Cell *103*, 423-433.

Grbavec, D., Lo, R., Liu, Y., Greenfield, A., and Stifani, S. (1999). Groucho/transducin-like enhancer of split (TLE) family members interact with the yeast transcriptional co-repressor SSN6 and mammalian SSN6-related proteins: implications for evolutionary conservation of transcription repression mechanisms. The Biochemical journal *337 (Pt 1)*, 13-17.

Hadjur, S., Williams, L.M., Ryan, N.K., Cobb, B.S., Sexton, T., Fraser, P., Fisher, A.G., and Merkenschlager, M. (2009). Cohesins form chromosomal cis-interactions at the developmentally regulated IFNG locus. Nature *460*, 410-413.

Hakimi, M.A., Bochar, D.A., Schmiesing, J.A., Dong, Y., Barak, O.G., Speicher, D.W., Yokomori, K., and Shiekhattar, R. (2002). A chromatin remodelling complex that loads cohesin onto human chromosomes. Nature *418*, 994-998.

Hamiche, A., Sandaltzopoulos, R., Gdula, D.A., and Wu, C. (1999). ATP-dependent histone octamer sliding mediated by the chromatin remodeling complex NURF. Cell *97*, 833-842.

Hampsey, M., Singh, B.N., Ansari, A., Laine, J.P., and Krishnamurthy, S. (2011). Control of eukaryotic gene expression: gene loops and transcriptional memory. Advances in enzyme regulation *51*, 118-125.

Han, M., and Grunstein, M. (1988). Nucleosome loss activates yeast downstream promoters in vivo. Cell *55*, 1137-1145.

Han, M., Kim, U.J., Kayne, P., and Grunstein, M. (1988). Depletion of histone H4 and nucleosomes activates the PHO5 gene in Saccharomyces cerevisiae. The EMBO journal *7*, 2221-2228.

Hanlon, S.E., Rizzo, J.M., Tatomer, D.C., Lieb, J.D., and Buck, M.J. (2011). The stress response factors Yap6, Cin5, Phd1, and Skn7 direct targeting of the conserved co-repressor Tup1-Ssn6 in S. cerevisiae. PloS one *6*, e19060.

Harbison, C.T., Gordon, D.B., Lee, T.I., Rinaldi, N.J., Macisaac, K.D., Danford, T.W., Hannett, N.M., Tagne, J.B., Reynolds, D.B., Yoo, J.*, et al.* (2004). Transcriptional regulatory code of a eukaryotic genome. Nature *431*, 99-104.

Hartley, P.D., and Madhani, H.D. (2009). Mechanisms that specify promoter nucleosome location and identity. Cell *137*, 445-458.

He, Y., Vogelstein, B., Velculescu, V.E., Papadopoulos, N., and Kinzler, K.W. (2008). The antisense transcriptomes of human cells. Science *322*, 1855-1857.

Helt, G.A., Nicol, J.W., Erwin, E., Blossom, E., Blanchard, S.G., Jr., Chervitz, S.A., Harmon, C., and Loraine, A.E. (2009). Genoviz Software Development Kit: Java tool kit for building genomics visualization applications. BMC bioinformatics *10*, 266.

Hongay, C.F., Grisafi, P.L., Galitski, T., and Fink, G.R. (2006). Antisense transcription controls cell fate in Saccharomyces cerevisiae. Cell *127*, 735-745.

Horike, S., Cai, S., Miyano, M., Cheng, J.F., and Kohwi-Shigematsu, T. (2005). Loss of silent-chromatin looping and impaired imprinting of DLX5 in Rett syndrome. Nature genetics *37*, 31-40.

Houseley, J., LaCava, J., and Tollervey, D. (2006). RNA-quality control by the exosome. Nature reviews Molecular cell biology *7*, 529-539.

Houseley, J., and Tollervey, D. (2008). The nuclear RNA surveillance machinery: the link between ncRNAs and genome structure in budding yeast? Biochimica et biophysica acta *1779*, 239-246.

Houseley, J., and Tollervey, D. (2009). The many pathways of RNA degradation. Cell *136*, 763-776.

Iida, T., and Araki, H. (2003). Noncompetitive Counteractions of DNA Polymerase and ISW2/yCHRAC for Epigenetic Inheritance of Telomere Position Effect in Saccharomyces cerevisiae. Molecular and cellular biology *24*, 217-227.

Ito, T., Bulger, M., Pazin, M.J., Kobayashi, R., and Kadonaga, J.T. (1997). ACF, an ISWI-containing and ATP-utilizing chromatin assembly and remodeling factor. Cell *90*, 145-155.

Iyer, V., and Struhl, K. (1995). Poly(dA:dT), a ubiquitous promoter element that stimulates transcription via its intrinsic DNA structure. The EMBO journal *14*, 2570-2579.

Jackson, J.C., and Lopes, J.M. (1996). The yeast UME6 gene is required for both negative and positive transcriptional regulation of phospholipid biosynthetic gene expression. Nucleic acids research *24*, 1322-1329.

Jiang, C., and Pugh, B.F. (2009a). A compiled and systematic reference map of nucleosome positions across the Saccharomyces cerevisiae genome. Genome biology *10*, R109.

Jiang, C., and Pugh, B.F. (2009b). Nucleosome positioning and gene regulation: advances through genomics. Nature reviews Genetics *10*, 161-172.

Jin, C., Zang, C., Wei, G., Cui, K., Peng, W., Zhao, K., and Felsenfeld, G. (2009). H3.3/H2A.Z double variant-containing nucleosomes mark 'nucleosome-free regions' of active promoters and other regulatory regions. Nature genetics *41*, 941-945.

Johnson, S.M., Tan, F.J., McCullough, H.L., Riordan, D.P., and Fire, A.Z. (2006). Flexibility and constraint in the nucleosome core landscape of Caenorhabditis elegans chromatin. Genome research *16*, 1505-1516.

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

Kagey, M.H., Newman, J.J., Bilodeau, S., Zhan, Y., Orlando, D.A., van Berkum, N.L., Ebmeier, C.C., Goossens, J., Rahl, P.B., Levine, S.S.*, et al.* (2010). Mediator and cohesin connect gene expression and chromatin architecture. Nature *467*, 430-435.

Kaplan, C.D., Laprade, L., and Winston, F. (2003). Transcription elongation factors repress transcription initiation from cryptic sites. Science *301*, 1096-1099.

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.*, et al.* (2009). The DNA-encoded nucleosome organization of a eukaryotic genome. Nature *458*, 362-366.

Kassabov, S.R., Henry, N.M., Zofall, M., Tsukiyama, T., and Bartholomew, B. (2002). High-Resolution Mapping of Changes in Histone-DNA Contacts of Nucleosomes Remodeled by ISW2. Molecular and cellular biology *22*, 7524-7534.

Keleher, C.A., Redd, M.J., Schultz, J., Carlson, M., and Johnson, A.D. (1992). Ssn6-Tup1 is a general repressor of transcription in yeast. Cell *68*, 709-719.

Kent, N.A., Karabetsou, N., Politis, P.K., and Mellor, J. (2001). In vivo chromatin remodeling by yeast ISWI homologs Isw1p and Isw2p. Genes & development *15*, 619-626.

Kim, U.J., Han, M., Kayne, P., and Grunstein, M. (1988). Effects of histone H4 depletion on the cell cycle and transcription of Saccharomyces cerevisiae. The EMBO journal *7*, 2211-2219.

Knezetic, J.A., and Luse, D.S. (1986). The presence of nucleosomes on a DNA template prevents initiation by RNA polymerase II in vitro. Cell *45*, 95-104.

Komachi, K., and Johnson, A.D. (1997). Residues in the WD repeats of Tup1 required for interaction with alpha2. Molecular and cellular biology *17*, 6023-6028.

Komachi, K., Redd, M.J., and Johnson, A.D. (1994). The WD repeats of Tup1 interact with the homeo domain protein alpha 2. Genes & development *8*, 2857-2867.

Kornberg, R.D. (1974). Chromatin structure: a repeating unit of histones and DNA. Science *184*, 868-871.

Kornberg, R.D., and Lorch, Y. (1999). Twenty-five years of the nucleosome, fundamental particle of the eukaryote chromosome. Cell *98*, 285-294.

Kornberg, R.D., and Thomas, J.O. (1974). Chromatin structure; oligomers of the histones. Science *184*, 865-868.

Kosak, S.T., Skok, J.A., Medina, K.L., Riblet, R., Le Beau, M.M., Fisher, A.G., and Singh, H. (2002). Subnuclear compartmentalization of immunoglobulin loci during lymphocyte development. Science *296*, 158-162.

Krivega, I., and Dean, A. (2012). Enhancer and promoter interactions-long distance calls. Current opinion in genetics & development *22*, 79-85.

Laine, J.P., Singh, B.N., Krishnamurthy, S., and Hampsey, M. (2009). A physiological role for gene loops in yeast. Genes & development *23*, 2604-2609.

Lam, F.H., Steger, D.J., and O'Shea, E.K. (2008). Chromatin decouples promoter threshold from dynamic range. Nature *453*, 246-250.

Langst, G., and Becker, P.B. (2004). Nucleosome remodeling: one mechanism, many phenomena? Biochimica et biophysica acta *1677*, 58-63.

Langst, G., Bonte, E.J., Corona, D.F., and Becker, P.B. (1999). Nucleosome movement by CHRAC and ISWI without disruption or trans-displacement of the histone octamer. Cell *97*, 843-852.

Lazzaro, M.A., Pepin, D., Pescador, N., Murphy, B.D., Vanderhyden, B.C., and Picketts, D.J. (2006). The imitation switch protein SNF2L regulates

steroidogenic acute regulatory protein expression during terminal differentiation of ovarian granulosa cells. Molecular endocrinology *20*, 2406-2417.

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

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

Levanon, D., Goldstein, R.E., Bernstein, Y., Tang, H., Goldenberg, D., Stifani, S., Paroush, Z., and Groner, Y. (1998). Transcriptional repression by AML1 and LEF-1 is mediated by the TLE/Groucho corepressors. Proceedings of the National Academy of Sciences of the United States of America *95*, 11590-11595.

Lieberman-Aiden, E., van Berkum, N.L., Williams, L., Imakaev, M., Ragoczy, T., Telling, A., Amit, I., Lajoie, B.R., Sabo, P.J., Dorschner, M.O.*, et al.* (2009). Comprehensive mapping of long-range interactions reveals folding principles of the human genome. Science *326*, 289-293.

Liu, Z., and Karmarkar, V. (2008). Groucho/Tup1 family co-repressors in plant development. Trends in plant science *13*, 137-144.

Lorch, Y., LaPointe, J.W., and Kornberg, R.D. (1987). Nucleosomes inhibit the initiation of transcription but allow chain elongation with the displacement of histones. Cell *49*, 203-210.

Loyola, A., LeRoy, G., Wang, Y.H., and Reinberg, D. (2001). Reconstitution of recombinant chromatin establishes a requirement for histone-tail modifications during chromatin assembly and transcription. Genes & development *15*, 2837-2851.

Luger, K., Mader, A.W., Richmond, R.K., Sargent, D.F., and Richmond, T.J. (1997). Crystal structure of the nucleosome core particle at 2.8 A resolution. Nature *389*, 251-260.

Mannervik, M., and Levine, M. (1999). The Rpd3 histone deacetylase is required for segmentation of the Drosophila embryo. Proceedings of the National Academy of Sciences of the United States of America *96*, 6797-6801.

Manuelidis, L. (1985). Individual interphase chromosome domains revealed by in situ hybridization. Human genetics *71*, 288-293.

Martens, J.A., Laprade, L., and Winston, F. (2004). Intergenic transcription is required to repress the Saccharomyces cerevisiae SER3 gene. Nature *429*, 571-574.

Martens, J.A., Wu, P.Y., and Winston, F. (2005). Regulation of an intergenic transcript controls adjacent gene transcription in Saccharomyces cerevisiae. Genes & development *19*, 2695-2704.

Matarazzo, M.R., Boyle, S., D'Esposito, M., and Bickmore, W.A. (2007). Chromosome territory reorganization in a human disease with altered DNA methylation. Proceedings of the National Academy of Sciences of the United States of America *104*, 16546-16551.

Mavrich, T.N., Ioshikhes, I.P., Venters, B.J., Jiang, C., Tomsho, L.P., Qi, J., Schuster, S.C., Albert, I., and Pugh, B.F. (2008a). A barrier nucleosome model for statistical positioning of nucleosomes throughout the yeast genome. Genome research *18*, 1073-1083.

Mavrich, T.N., Jiang, C., Ioshikhes, I.P., Li, X., Venters, B.J., Zanton, S.J., Tomsho, L.P., Qi, J., Glaser, R.L., Schuster, S.C.*, et al.* (2008b). Nucleosome organization in the Drosophila genome. Nature *453*, 358-362.

McConnell, A.D., Gelbart, M.E., and Tsukiyama, T. (2004). Histone Fold Protein Dls1p Is Required for Isw2-Dependent Chromatin Remodeling In Vivo. Molecular and cellular biology *24*, 2605-2613.

Meister, P., Towbin, B.D., Pike, B.L., Ponti, A., and Gasser, S.M. (2010). The spatial dynamics of tissue-specific promoters during C. elegans development. Genes & development *24*, 766-782.

Mellor, J., and Morillon, A. (2004). ISWI complexes in Saccharomyces cerevisiae. Biochimica et biophysica acta *1677*, 100-112.

Misteli, T. (2010). Higher-order genome organization in human disease. Cold Spring Harbor perspectives in biology *2*, a000794.

Miura, F., Kawaguchi, N., Sese, J., Toyoda, A., Hattori, M., Morishita, S., and Ito, T. (2006). A large-scale full-length cDNA analysis to explore the budding yeast transcriptome. Proceedings of the National Academy of Sciences of the United States of America *103*, 17846-17851.

Murrell, A., Heeson, S., and Reik, W. (2004). Interaction between differentially methylated regions partitions the imprinted genes Igf2 and H19 into parent-specific chromatin loops. Nature genetics *36*, 889-893.

Nagalakshmi, U., Wang, Z., Waern, K., Shou, C., Raha, D., Gerstein, M., and Snyder, M. (2008). The transcriptional landscape of the yeast genome defined by RNA sequencing. Science *320*, 1344-1349.

Neil, H., Malabat, C., d'Aubenton-Carafa, Y., Xu, Z., Steinmetz, L.M., and Jacquier, A. (2009). Widespread bidirectional promoters are the major source of cryptic transcripts in yeast. Nature *457*, 1038-1042.

Nemeth, A., Guibert, S., Tiwari, V.K., Ohlsson, R., and Langst, G. (2008). Epigenetic regulation of TTF-I-mediated promoter-terminator interactions of rRNA genes. The EMBO journal *27*, 1255-1265.

Ng, H.H., Robert, F., Young, R.A., and Struhl, K. (2002). Genome-wide location and regulated recruitment of the RSC nucleosome-remodeling complex. Genes & development *16*, 806-819.

Nicol, J.W., Helt, G.A., Blanchard, S.G., Jr., Raja, A., and Loraine, A.E. (2009). The Integrated Genome Browser: free software for distribution and exploration of genome-scale datasets. Bioinformatics *25*, 2730-2731.

Noordermeer, D., de Wit, E., Klous, P., van de Werken, H., Simonis, M., Lopez-Jones, M., Eussen, B., de Klein, A., Singer, R.H., and de Laat, W. (2011).

Variegated gene expression caused by cell-specific long-range DNA interactions. Nature cell biology *13*, 944-951.

O'Reilly, D., and Greaves, D.R. (2007). Cell-type-specific expression of the human CD68 gene is associated with changes in Pol II phosphorylation and short-range intrachromosomal gene looping. Genomics *90*, 407-415.

O'Sullivan, J.M., Tan-Wong, S.M., Morillon, A., Lee, B., Coles, J., Mellor, J., and Proudfoot, N.J. (2004). Gene loops juxtapose promoters and terminators in yeast. Nature genetics *36*, 1014-1018.

Olins, A.L., and Olins, D.E. (1974). Spheroid chromatin units (v bodies). Science *183*, 330-332.

Papamichos-Chronakis, M., Watanabe, S., Rando, O.J., and Peterson, C.L. (2011). Global regulation of H2A.Z localization by the INO80 chromatin-remodeling enzyme is essential for genome integrity. Cell *144*, 200-213.

Parelho, V., Hadjur, S., Spivakov, M., Leleu, M., Sauer, S., Gregson, H.C., Jarmuz, A., Canzonetta, C., Webster, Z., Nesterova, T.*, et al.* (2008). Cohesins functionally associate with CTCF on mammalian chromosome arms. Cell *132*, 422-433.

Parnell, T.J., Huff, J.T., and Cairns, B.R. (2008). RSC regulates nucleosome positioning at Pol II genes and density at Pol III genes. The EMBO journal *27*, 100-110.

Pecinka, A., Schubert, V., Meister, A., Kreth, G., Klatte, M., Lysak, M.A., Fuchs, J., and Schubert, I. (2004). Chromosome territory arrangement and homologous pairing in nuclei of Arabidopsis thaliana are predominantly random except for NOR-bearing chromosomes. Chromosoma *113*, 258-269.

Peng, S., Alekseyenko, A.A., Larschan, E., Kuroda, M.I., and Park, P.J. (2007). Normalization and experimental design for ChIP-chip data. BMC bioinformatics *8*, 219.

Perkins, K.J., Lusic, M., Mitar, I., Giacca, M., and Proudfoot, N.J. (2008). Transcription-dependent gene looping of the HIV-1 provirus is dictated by recognition of pre-mRNA processing signals. Molecular cell *29*, 56-68.

Pflugrad, A., Meir, J.Y., Barnes, T.M., and Miller, D.M., 3rd (1997). The Groucho-like transcription factor UNC-37 functions with the neural specificity gene unc-4 to govern motor neuron identity in C. elegans. Development *124*, 1699-1709.

Phillips, J.E., and Corces, V.G. (2009). CTCF: master weaver of the genome. Cell *137*, 1194-1211.

Pokholok, D.K., Harbison, C.T., Levine, S., Cole, M., Hannett, N.M., Lee, T.I., Bell, G.W., Walker, K., Rolfe, P.A., Herbolsheimer, E.*, et al.* (2005). Genome-wide map of nucleosome acetylation and methylation in yeast. Cell *122*, 517-527.

Poot, R.A., Bozhenok, L., van den Berg, D.L., Steffensen, S., Ferreira, F., Grimaldi, M., Gilbert, N., Ferreira, J., and Varga-Weisz, P.D. (2004). The Williams

syndrome transcription factor interacts with PCNA to target chromatin remodelling by ISWI to replication foci. Nature cell biology *6*, 1236-1244.

Preker, P., Nielsen, J., Kammler, S., Lykke-Andersen, S., Christensen, M.S., Mapendano, C.K., Schierup, M.H., and Jensen, T.H. (2008). RNA exosome depletion reveals transcription upstream of active human promoters. Science *322*, 1851-1854.

Radman-Livaja, M., and Rando, O.J. (2010). Nucleosome positioning: how is it established, and why does it matter? Developmental biology *339*, 258-266.

Ragoczy, T., Bender, M.A., Telling, A., Byron, R., and Groudine, M. (2006). The locus control region is required for association of the murine beta-globin locus with engaged transcription factories during erythroid maturation. Genes & development *20*, 1447-1457.

Raisner, R.M., Hartley, P.D., Meneghini, M.D., Bao, M.Z., Liu, C.L., Schreiber, S.L., Rando, O.J., and Madhani, H.D. (2005). Histone variant H2A.Z marks the 5' ends of both active and inactive genes in euchromatin. Cell *123*, 233-248.

Rando, O.J., and Chang, H.Y. (2009). Genome-wide views of chromatin structure. Annual review of biochemistry *78*, 245-271.

Royce, T.E., Carriero, N.J., and Gerstein, M.B. (2007). An efficient pseudomedian filter for tiling microrrays. BMC bioinformatics *8*, 186.

Saez-Vasquez, J., and Gadal, O. (2010). Genome organization and function: a view from yeast and Arabidopsis. Molecular plant *3*, 678-690.

Sahasrabuddhe, C.G., and Van Holde, K.E. (1974). The effect of trypsin on nuclease-resistant chromatin fragments. The Journal of biological chemistry *249*, 152-156.

Samanta, M.P., Tongprasit, W., Sethi, H., Chin, C.S., and Stolc, V. (2006). Global identification of noncoding RNAs in Saccharomyces cerevisiae by modulating an essential RNA processing pathway. Proceedings of the National Academy of Sciences of the United States of America *103*, 4192-4197.

Sasaki, S., Mello, C.C., Shimada, A., Nakatani, Y., Hashimoto, S., Ogawa, M., Matsushima, K., Gu, S.G., Kasahara, M., Ahsan, B.*, et al.* (2009). Chromatin-associated periodicity in genetic variation downstream of transcriptional start sites. Science *323*, 401-404.

Schardin, M., Cremer, T., Hager, H.D., and Lang, M. (1985). Specific staining of human chromosomes in Chinese hamster x man hybrid cell lines demonstrates interphase chromosome territories. Human genetics *71*, 281-287.

Schmid, M., and Jensen, T.H. (2008). The exosome: a multipurpose RNA-decay machine. Trends in biochemical sciences *33*, 501-510.

Schoenfelder, S., Clay, I., and Fraser, P. (2010a). The transcriptional interactome: gene expression in 3D. Current opinion in genetics & development *20*, 127-133.

Schoenfelder, S., Sexton, T., Chakalova, L., Cope, N.F., Horton, A., Andrews, S., Kurukuti, S., Mitchell, J.A., Umlauf, D., Dimitrova, D.S.*, et al.* (2010b). Preferential associations between co-regulated genes reveal a transcriptional interactome in erythroid cells. Nature genetics *42*, 53-61.

Schones, D.E., Cui, K., Cuddapah, S., Roh, T.Y., Barski, A., Wang, Z., Wei, G., and Zhao, K. (2008). Dynamic regulation of nucleosome positioning in the human genome. Cell *132*, 887-898.

Seila, A.C., Calabrese, J.M., Levine, S.S., Yeo, G.W., Rahl, P.B., Flynn, R.A., Young, R.A., and Sharp, P.A. (2008). Divergent transcription from active promoters. Science *322*, 1849-1851.

Sekinger, E.A., Moqtaderi, Z., and Struhl, K. (2005). Intrinsic histone-DNA interactions and low nucleosome density are important for preferential accessibility of promoter regions in yeast. Molecular cell *18*, 735-748.

Sekiya, T., and Zaret, K.S. (2007). Repression by Groucho/TLE/Grg proteins: genomic site recruitment generates compacted chromatin in vitro and impairs activator binding in vivo. Molecular cell *28*, 291-303.

Sexton, T., Bantignies, F., and Cavalli, G. (2009). Genomic interactions: chromatin loops and gene meeting points in transcriptional regulation. Seminars in cell & developmental biology *20*, 849-855.

Shaw, P.J., Abranches, R., Paula Santos, A., Beven, A.F., Stoger, E., Wegel, E., and Gonzalez-Melendi, P. (2002). The architecture of interphase chromosomes

and nucleolar transcription sites in plants. Journal of structural biology *140*, 31-38.

Shivaswamy, S., Bhinge, A., Zhao, Y., Jones, S., Hirst, M., and Iyer, V.R. (2008). Dynamic remodeling of individual nucleosomes across a eukaryotic genome in response to transcriptional perturbation. PLoS biology *6*, e65.

Shopland, L.S., Lynch, C.R., Peterson, K.A., Thornton, K., Kepper, N., Hase, J., Stein, S., Vincent, S., Molloy, K.R., Kreth, G.*, et al.* (2006). Folding and organization of a contiguous chromosome region according to the gene distribution pattern in primary genomic sequence. The Journal of cell biology *174*, 27-38.

Simonis, M., Klous, P., Splinter, E., Moshkin, Y., Willemsen, R., de Wit, E., van Steensel, B., and de Laat, W. (2006). Nuclear organization of active and inactive chromatin domains uncovered by chromosome conformation capture-on-chip (4C). Nature genetics *38*, 1348-1354.

Singh, B.N., Ansari, A., and Hampsey, M. (2009). Detection of gene loops by 3C in yeast. Methods *48*, 361-367.

Singh, B.N., and Hampsey, M. (2007). A transcription-independent role for TFIIB in gene looping. Molecular cell *27*, 806-816.

Smyth, G.K. (2004). Linear models and empirical bayes methods for assessing differential expression in microarray experiments. Statistical applications in genetics and molecular biology *3*, Article3.

Spilianakis, C.G., Lalioti, M.D., Town, T., Lee, G.R., and Flavell, R.A. (2005). Interchromosomal associations between alternatively expressed loci. Nature *435*, 637-645.

Splinter, E., Heath, H., Kooren, J., Palstra, R.J., Klous, P., Grosveld, F., Galjart, N., and de Laat, W. (2006). CTCF mediates long-range chromatin looping and local histone modification in the beta-globin locus. Genes & development *20*, 2349-2354.

Steber, C.M., and Esposito, R.E. (1995). UME6 is a central component of a developmental regulatory switch controlling meiosis-specific gene expression. Proceedings of the National Academy of Sciences of the United States of America *92*, 12490-12494.

Stopka, T., and Skoultchi, A.I. (2003). The ISWI ATPase Snf2h is required for early mouse development. Proceedings of the National Academy of Sciences of the United States of America *100*, 14097-14102.

Stopka, T., Zakova, D., Fuchs, O., Kubrova, O., Blafkova, J., Jelinek, J., Necas, E., and Zivny, J. (2000). Chromatin remodeling gene SMARCA5 is dysregulated in primitive hematopoietic cells of acute leukemia. Leukemia : official journal of the Leukemia Society of America, Leukemia Research Fund, UK *14*, 1247-1252.

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

Sun, W., Xie, W., Xu, F., Grunstein, M., and Li, K.C. (2009). Dissecting nucleosome free regions by a segmental semi-Markov model. PloS one *4*, e4721.

Svaren, J., and Horz, W. (1997). Transcription factors vs nucleosomes: regulation of the PHO5 promoter in yeast. Trends in biochemical sciences *22*, 93-97.

Sweet, D.H., Jang, Y.K., and Sancar, G.B. (1997). Role of UME6 in transcriptional regulation of a DNA repair gene in Saccharomyces cerevisiae. Molecular and cellular biology *17*, 6223-6235.

Tan-Wong, S.M., French, J.D., Proudfoot, N.J., and Brown, M.A. (2008). Dynamic interactions between the promoter and terminator regions of the mammalian BRCA1 gene. Proceedings of the National Academy of Sciences of the United States of America *105*, 5160-5165.

Tan-Wong, S.M., Wijayatilake, H.D., and Proudfoot, N.J. (2009). Gene loops function to maintain transcriptional memory through interaction with the nuclear pore complex. Genes & development *23*, 2610-2624.

Timme, S., Schmitt, E., Stein, S., Schwarz-Finsterle, J., Wagner, J., Walch, A., Werner, M., Hausmann, M., and Wiech, T. (2011). Nuclear position and shape deformation of chromosome 8 territories in pancreatic ductal adenocarcinoma. Analytical cellular pathology *34*, 21-33.

Tirosh, I., Sigal, N., and Barkai, N. (2010). Widespread remodeling of mid-coding sequence nucleosomes by Isw1. Genome biology *11*, R49.

Todd, R.B., Greenhalgh, J.R., Hynes, M.J., and Andrianopoulos, A. (2003). TupA, the Penicillium marneffei Tup1p homologue, represses both yeast and spore development. Molecular microbiology *48*, 85-94.

Toedling, J., Skylar, O., Krueger, T., Fischer, J.J., Sperling, S., and Huber, W. (2007). Ringo--an R/Bioconductor package for analyzing ChIP-chip readouts. BMC bioinformatics *8*, 221.

Tolhuis, B., Palstra, R.J., Splinter, E., Grosveld, F., and de Laat, W. (2002). Looping and interaction between hypersensitive sites in the active beta-globin locus. Molecular cell *10*, 1453-1465.

Tsukiyama, T. (2002). The in vivo functions of ATP-dependent chromatin-remodelling factors. Nature reviews Molecular cell biology *3*, 422-429.

Tsukiyama, T., Palmer, J., Landel, C.C., Shiloach, J., and Wu, C. (1999). Characterization of the imitation switch subfamily of ATP-dependent chromatin-remodeling factors in Saccharomyces cerevisiae. Genes & development *13*, 686-697.

Vakoc, C.R., Letting, D.L., Gheldof, N., Sawado, T., Bender, M.A., Groudine, M., Weiss, M.J., Dekker, J., and Blobel, G.A. (2005). Proximity among distant regulatory elements at the beta-globin locus requires GATA-1 and FOG-1. Molecular cell *17*, 453-462.

Valouev, A., Ichikawa, J., Tonthat, T., Stuart, J., Ranade, S., Peckham, H., Zeng, K., Malek, J.A., Costa, G., McKernan, K.*, et al.* (2008). A high-resolution,

nucleosome position map of C. elegans reveals a lack of universal sequence-dictated positioning. Genome research *18*, 1051-1063.

Varga-Weisz, P.D., Wilm, M., Bonte, E., Dumas, K., Mann, M., and Becker, P.B. (1997). Chromatin-remodelling factor CHRAC contains the ATPases ISWI and topoisomerase II. Nature *388*, 598-602.

Venters, B.J., and Pugh, B.F. (2009). A canonical promoter organization of the transcription machinery and its regulators in the Saccharomyces genome. Genome research *19*, 360-371.

Vernimmen, D., De Gobbi, M., Sloane-Stanley, J.A., Wood, W.G., and Higgs, D.R. (2007). Long-range chromosomal interactions regulate the timing of the transition between poised and active gene expression. The EMBO journal *26*, 2041-2051.

Vincent, J.A., Kwong, T.J., and Tsukiyama, T. (2008). ATP-dependent chromatin remodeling shapes the DNA replication landscape. Nature structural & molecular biology *15*, 477-484.

Wang, K.C., Yang, Y.W., Liu, B., Sanyal, A., Corces-Zimmerman, R., Chen, Y., Lajoie, B.R., Protacio, A., Flynn, R.A., Gupta, R.A.*, et al.* (2011). A long noncoding RNA maintains active chromatin to coordinate homeotic gene expression. Nature *472*, 120-124.

Watson, A.D. (2000). Ssn6-Tup1 interacts with class I histone deacetylases required for repression. Genes & development *14*, 2737-2744.

Weisbrod, S., Groudine, M., and Weintraub, H. (1980). Interaction of HMG 14 and 17 with actively transcribed genes. Cell *19*, 289-301.

Wendt, K.S., Yoshida, K., Itoh, T., Bando, M., Koch, B., Schirghuber, E., Tsutsumi, S., Nagae, G., Ishihara, K., Mishiro, T.*, et al.* (2008). Cohesin mediates transcriptional insulation by CCCTC-binding factor. Nature *451*, 796-801.

Wettenhall, J.M., and Smyth, G.K. (2004). limmaGUI: a graphical user interface for linear modeling of microarray data. Bioinformatics *20*, 3705-3706.

Whitehouse, I., Rando, O.J., Delrow, J., and Tsukiyama, T. (2007). Chromatin remodelling at promoters suppresses antisense transcription. Nature *450*, 1031-1035.

Whitehouse, I., and Tsukiyama, T. (2006). Antagonistic forces that position nucleosomes in vivo. Nature structural & molecular biology *13*, 633-640.

Wiech, T., Stein, S., Lachenmaier, V., Schmitt, E., Schwarz-Finsterle, J., Wiech, E., Hildenbrand, G., Werner, M., and Hausmann, M. (2009). Spatial allelic imbalance of BCL2 genes and chromosome 18 territories in nonneoplastic and neoplastic cervical squamous epithelium. European biophysics journal : EBJ *38*, 793-806.

Wilhelm, B.T., Marguerat, S., Watt, S., Schubert, F., Wood, V., Goodhead, I., Penkett, C.J., Rogers, J., and Bahler, J. (2008). Dynamic repertoire of a eukaryotic transcriptome surveyed at single-nucleotide resolution. Nature *453*, 1239-1243.

Winkler, C.J., Ponce, A., and Courey, A.J. (2010). Groucho-mediated repression may result from a histone deacetylase-dependent increase in nucleosome density. PloS one *5*, e10166.

Wu, C. (1980). The 5' ends of Drosophila heat shock genes in chromatin are hypersensitive to DNase I. Nature *286*, 854-860.

Wysocka, J., Swigut, T., Xiao, H., Milne, T.A., Kwon, S.Y., Landry, J., Kauer, M., Tackett, A.J., Chait, B.T., Badenhorst, P.*, et al.* (2006). A PHD finger of NURF couples histone H3 lysine 4 trimethylation with chromatin remodelling. Nature *442*, 86-90.

Xu, Z., Wei, W., Gagneur, J., Perocchi, F., Clauder-Munster, S., Camblong, J., Guffanti, E., Stutz, F., Huber, W., and Steinmetz, L.M. (2009). Bidirectional promoters generate pervasive transcription in yeast. Nature *457*, 1033-1037.

Yadon, A.N., and Tsukiyama, T. (2011). SnapShot: Chromatin remodeling: ISWI. Cell *144*, 453-453 e451.

Yadon, A.N., Van De Mark, D., Basom, R., Delrow, J., Whitehouse, I., and Tsukiyama, T. (2010). Chromatin Remodeling Around NFRs Leads to

Yang, X.J., and Seto, E. (2008). The Rpd3/Hda1 family of lysine deacetylases: from bacteria and yeast to mice and men. Nature reviews Molecular cell biology *9*, 206-218.

Yasui, D., Miyano, M., Cai, S., Varga-Weisz, P., and Kohwi-Shigematsu, T. (2002). SATB1 targets chromatin remodelling to regulate genes over long distances. Nature *419*, 641-645.

Yen, K., Vinayachandran, V., Batta, K., Koerber, R.T., and Pugh, B.F. (2012). Genome-wide Nucleosome Specificity and Directionality of Chromatin Remodelers. Cell *149*, 1461-1473.

Yuan, G.C., Liu, Y.J., Dion, M.F., Slack, M.D., Wu, L.F., Altschuler, S.J., and Rando, O.J. (2005). Genome-scale identification of nucleosome positions in S. cerevisiae. Science *309*, 626-630.

Zardo, G., Cimino, G., and Nervi, C. (2008). Epigenetic plasticity of chromatin in embryonic and hematopoietic stem/progenitor cells: therapeutic potential of cell reprogramming. Leukemia : official journal of the Leukemia Society of America, Leukemia Research Fund, UK *22*, 1503-1518.

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

Zhou, Y., and Grummt, I. (2005). The PHD finger/bromodomain of NoRC interacts with acetylated histone H4K16 and is sufficient for rDNA silencing. Current biology : CB *15*, 1434-1438.

Zhou, Y., Santoro, R., and Grummt, I. (2002). The chromatin remodeling complex NoRC targets HDAC1 to the ribosomal gene promoter and represses RNA polymerase I transcription. The EMBO journal *21*, 4632-4640.

CURRICULUM VITAE

Education

University of Washington and Fred Hutchison Cancer Research Center
Seattle, WA | August 2012
Ph.D. Molecular and Cellular Biology

Whitman College
Walla Walla, WA | May 2005
B.A. Biochemistry, Biophysics, Molecular Biology (interdisciplinary major)

Research/Work Experience

Fred Hutchison Cancer Research Center
Seattle, WA | January 2007-Present
Dissertation Laboratory – *Dr. Toshio Tsukiyama*

- Understand the functions and targeting mechanism of the ATP-dependent chromatin remodeling enzyme Isw2 in the budding yeast *Saccharomyces cerevisiae*.
- Annotated nucleosome free regions across the yeast genome by computationally and statistically comparing multiple genome-wide nucleosome positioning maps
- Determined the prevalence of non-coding RNA transcription resulting from the absence of Isw2 using custom, high-density, tiled microarrays

University of Washington
Seattle, WA | August 2006-December 2006
Laboratory Rotation – *Dr. Jaisri Lingappa*

- Biochemically determined the *in vivo* binding and dimerization interactions of HIV1-GAG and ABCE1 using co-immunoprecipication and biochemical techniques.
- Immunopurified HIV1-GAG and ABCE1 from mammalian cell culture lines

Seattle Biomedical Research Institute
Seattle, WA | June 2006-August 2006
Laboratory Rotation – *Dr. Kenneth Stuart*

- Determining the *in vivo* global interaction map of *Trypanosome brucei* mitochondrial proteins through TAP-tag purification and mass spectroscopy

- Maintained, passaged, and transfected *Trypanosome brucei* in cell culture

National Institutes of Health – NICHD
Bethesda, MD | June 2005-June 2006
Post-Baccalaureate Intramural Research Training Award Program – *Dr. Catherine L. Jackson*

- Fluorescently tagged WT and mutant yeast vesicular trafficking Arf-GEF proteins through cloning
- Phenotypically classified and scored mutant yeast vesicular trafficking Arf-GEF proteins by confocal fluorescent microscopy with endoplasmic reticulum, Golgi, and cell membrane makers

National Institutes of Health – NICHD
Bethesda, MD | June 2004-August 2004
Biomedical Research Summer Internship Program – *Dr. Catherine L. Jackson*

- Determined Arf-GEF protein interaction with coatamer and adapter protein complexes using yeast two-hybrid analysis

Teaching and Outreach Activities

Fred Hutchison Cancer Research Center
Seattle, WA | January 2008-August 2008
Student Education Partnership (SEP)

- Mentored a high school science teacher in curriculum development and hands-on biomedical research experience
- Completed 40 hours of community science outreach

University of Washington
Seattle, WA | January 2007-March 2007
Biology 200 Teaching Assistant

- Instructed and facilitated lab for 48 students
- Assisted in development and grading of exams

Whitman College
Walla Walla, WA | January 2005-May 2005
Biochemistry Laboratory Assistant

- Prepared lab equipment and solutions for each week
- Assisted students during lab excersises

Whitman College
Walla Walla, WA | September 2005-May 2005
Chemistry and Biochemistry Tutor

- Assisted two students in daily assignments and test preparation

Fellowships/Funding

Developmental Biology Predoctoral Training Grant T32HD007183
Seattle, WA | September 2008-Present
National Institute of Child Health and Human Development of the National Institutes of Health

Publications

Yadon AN, Tsukiyama T. SnapShot: Chromatin Remodeling: ISWI. Cell. (in press)

Yadon AN, Van de Mark D, Basom R, Delrow J, Whitehouse I, Tsukiyama T. Chromatin remodeling around nucleosome-free regions leads to repression of noncoding RNA transcription. Mol Cell Biol. 2010 Nov;30(21):5110-22 PMC2953044

Tsukiyama T, Bogenschutz N, Kwong T, Rodriguez J, Unnikrishnan A, and **Yadon AN**. ISWI Chromatin Remodeling Complexes. In: Bradshaw RA and Dennis EA, ed., Handbook of Cell Signaling 2nd Edition. Oxford: Academic Press, 2009, pp. 2357-2362.

Presentations

Systems Biology: Global Regulation of Gene Expression
Cold Spring Harbor Laboratory, NY | March 2010
Poster: **A. N. Yadon,** D. Van de Mark, I. Whitehouse, and T. Tsukiyama. “Genome-wide reduction of NFR size and suppression of cryptic RNA transcription by an ATP-dependent chromatin remodeling enzyme.”

Northwest Regional Developmental Biology Conference
Friday Harbor Laboratories, WA | March 2010
Poster: **A. N. Yadon,** D. Van de Mark, I. Whitehouse, and T. Tsukiyama. “Genome-wide reduction of NFR size and suppression of cryptic RNA transcription by an ATP-dependent chromatin remodeling enzyme.”

Northwest Regional Developmental Biology Conference
Friday Harbor Laboratories, WA | March 2009
Poster: **A. N. Yadon,** I. Whitehouse, D. Van de Mark, and T. Tsukiyama, “Isw2 Suppresses Non-Coding RNA from Nucleosome Free Regions Globally.”