Title

A single-cell sensor reveals spatial and temporal regulation of heterochromatin domain formation

Permalink

https://escholarship.org/uc/item/8q61k2fw

Author

Greenstein, R. A.

Publication Date

2020

Peer reviewed|Thesis/dissertation

A single-cell sensor reveals spatial and temporal regulation of heterochromatin domain formation

by
R.A. Greenstein

DISSERTATION
Submitted in partial satisfaction of the requirements for degree of
DOCTOR OF PHILOSOPHY

in

Biochemistry and Molecular Biology

in the

GRADUATE DIVISION
of the
UNIVERSITY OF CALIFORNIA, SAN FRANCISCO

Approved:

DocuSigned by:

_____   Bassem Al-Sady
251AC7A581B94CA...   Chair

DocuSigned by:
Jonathan Weissman   Jonathan Weissman
_____
DocuSigned by:
Barbara Panning   Barbara Panning
_____
DocuSigned by:
                                                     Fred Chang
_____
05042907E5B249D...

_____
                                                     Committee Members

Copyright 2020

by

R.A. Greenstein

~ in loving memory ~

Sheri Abrams

Irving Abrams

Acknowledgements

I would like to thank the following people who through mentorship and/or support have made significant impact on my time in graduate school:

- My thesis advisor Bassem Al-Sady for his mentorship, patience, and unwavering support.

- Members of the Al-Sady lab, past and present, for their support, advice, and many good laughs over the years and especially for fostering such a welcoming environment. In particular I would like to thank Nick Sanchez for being there for me through it all.

- My thesis committee – Jonathan Weissman, Barbara Panning, and Fred Chang – for their support, encouragement, and thoughtful advice on my project.

- The UCSF community, in particular members of the G.W. Hooper Foundation whom I had the pleasure of interacting with daily and two post-doc mentors– Sandra Catania and Jordan Bisanz – who generously dedicated their time to help me learn new skills.

- My undergraduate research mentors Doug Chalker, Rachel Schwope, and Sally Elgin from whom I learned the skills that helped prepare me for graduate school.

- My Tetrad program classmates for their friendship, support, and advice.

- My loved ones, friends, and the cats in my life for supporting and encouraging me.

Contributions

Chapter 2 - this chapter is reproduced from the following publication:

Greenstein RA, Jones SK, Spivey EC, Rybarski JR, Finkelstein IJ, Al-Sady B. 2018. Noncoding RNA-nucleated heterochromatin spreading is intrinsically labile and requires accessory elements for epigenetic stability. *Elife* **7**:e32948.

Chapter 3 - this chapter is reproduced from a preprint indended for publication:

Greenstein RA*, Ng H*****, Barrales RR, Tan C, Braun S, Al-Sady B. 2020. Local chromatin context dictates the genetic determinants of the heterochromatin spreading reaction. *bioRxiv*: 2020.2005.2026.117143. ***equal contribution.**

Chapter 4 - this chapter is reproduced from the following publication:

Greenstein RA, Barrales RR, Sanchez NA, Bisanz JE, Braun S, Al-Sady B. 2019. Set1/COMPASS repels heterochromatin invasion at euchromatic sites by disrupting Suv39/Clr4 activity and nucleosome stability. *Genes Dev* **34**: 99-117.

Chapter 6 – The discussion is partially based upon text from the following publication:

Greenstein RA, Al-Sady B. 2019. Epigenetic fates of gene silencing established by heterochromatin spreading in cell identity and genome stability. *Current Genetics* 65: 423-428.

A single-cell sensor reveals spatial and temporal regulation of heterochromatin domain formation

R.A. Greenstein

Abstract

Cell type specification in multicellular systems is due, at least in part, to differential access to and usage of the genomic DNA which is common to all cells, leading to cell type specific gene expression. This genomic pattern in enacted through the formation of gene-repressive heterochromatin, a nuclear ultrastructure, which silences transcription of genes embedded in the underlying DNA that are orthogonal to the intended cell fate and expands progressively, in a process known as spreading, while cells proceed through lineage commitment. Importantly, once the cell type specific pattern of gene expression has been established, it must be recapitulated over repeated cell division to ensure faithful maintenance of cell identity and avoid disease. Both the mechanisms which direct this process of differential expansion and the features of heterochromatin domains which are critical to its robust inheritance have not been fully elucidated. Utilizing a genetically tractable fission yeast model that recapitulates features of the metazoan chromatin environment, we interrogated the requirements for spatial and temporal regulation of heterochromatin domain formation. We find that at cell identity loci, robust gene silencing and the capacity to remember heterochromatin states over repeated division requires the collaboration of multiple sequence elements with distinct spreading properties and capacities to resist chromatin perturbations. Additionally, we find chromatin context dependent requirements for genetic regulation of the spreading process. Lastly, we described a mechanism for regulating the spatial expansion of heterochromatin domains that relies on the mutually antagonistic signals and properties that differ between hetero- and eu- chromatin domains.

Table of Contents

List of Figures

List of Tables

List of Abbreviations

Abbreviation	Definition
ac	Acetyl / Acetylation
5-FOA	5-Fluoro-orotic acid
bp	Base pairs
CD	Chromodomain
cDNA	Complementary DNA
ChIP	Chromatin Immunoprecipitation
ChIP-seq	Chromatin Immunoprecipitation – Next Generation Sequencing
COMPASS	Complex of Proteins Associated with Set1
DMSO	Dimethyl-sulfoxide
EMD	Earth mover's distance
EMM	Edinburgh minimal media
ER	Estradiol
FACS	Fluorescence-activated cell sorting
FP	Fluorescence polarization
FYLM	Fission yeast lifespan microdissector
H3K27	Histone H3 Lysine 27
H3K4	Histone H3 Lysine 4
H3K9	Histone H3 Lysine 9
HAT	Histone Acetyl-transferase
HDAC	Histone deacetylase
HSS	Heterochromatin Spreading Sensor
HTS	High throughput sampler
HU	Hydroxyurea
IP	Immunoprecipitation
IR-R	Inverted repeat - right
kb	Kilobases, kilobase pairs
LAD	Lamin associated domain
MAD	Median absolute deviation
MAT	Mating type locus
mb	megabases, megabase pairs
me	Methylation
me1	Mono-methylation
me2	Di-methylation

Abbreviation	Definition
me3	Tri-methylation
mRNA	Messenger RNA
ncRNA	non-coding RNA
NFR	Nucleosome free region
OD	Optical density
PCR	Polymerase chain reaction
qPCR	Quantitative polymerase chain reaction
REII	Repression Element II – a DNA element in *S. pombe*
REIII	Repression Element III – a DNA element in *S. pombe*
RITE	Recombination induced tag exchange
RNAi	RNA interference
RT-qPCR	Reverse transcriptase quantitative polymerase chain reaction
SD	Standard deviation
t-SNE	t-distributed stochastic neighbor embedding
TSA	Trichostatin A
TSS	Transcription start site
WCE	Whole cell extract
WT	Wild type
XFP	Fluorescent protein
YES	Yeast Extracts + supplements media

1.   Introduction

One of the key features of complex, multicellular organisms is the functional diversity of the myriad cell types which collaborate to define the whole system. This division of labor is achieved through the expression of distinct proteins and pathways, leading to changes in cellular structure and biochemical activities, despite almost all cells reading from the same genetic DNA blueprint. Decades of prior work have elucidated various mechanisms that can orchestrate the differential interpretation of a static genetic code to create this diversity in functional output. Interestingly, these mechanisms are enacted across a large range of genomic scales from direct, gene-based transcriptional regulation by DNA-binding activator or repressor proteins (Jacob and Monod 1961; Fuda et al. 2009; Graf 2011; Spitz and Furlong 2012), to differential gene expression mediated by structural changes to DNA packaging and accessibility (Li 2002; Bonasio et al. 2010; Ho and Crabtree 2010; Bannister and Kouzarides 2011; Orkin and Hochedlinger 2011; Klemm et al. 2019), to large-scale changes in the three-dimensional organization of DNA within the nucleus (Lieberman-Aiden et al. 2009; van Steensel and Belmont 2017; Yu and Ren 2017; Zheng and Xie 2019). However, these mechanisms are all connected in that their implementation by the cell results in differential access to and interpretation of underlying DNA.

Chromatin context is one of the critical components to determining the functional output of a DNA sequence. Chromatin is defined as the packaging of DNA into the repeating biochemical unit of 147 base pairs (bp) of DNA wrapped around a core octamer of histone proteins that form a long array of nucleosomes (DNA + histones) across the chromosome (Luger et al. 1997). Chromatin is further sub-classified into euchromatin – transcriptionally active, and accessible regions – and heterochromatin, which is typified by transcriptional repression and a

closed or compacted state (Kouzarides 2007; Allshire and Madhani 2018). Although the DNA itself can be chemically modified to alter its expression state, the platform of the protein:DNA nucleosome context within chromatin offers significantly increased modularity for signals that can change expression of or access to the underlying DNA (Margueron et al. 2005; Kouzarides 2007). Many post-translational modifications (PTMs) of the histone proteins comprising chromatin and their role in directing gene expression have been described to date (Berger 2002; Kouzarides 2007; Bonasio et al. 2010; Bannister and Kouzarides 2011). Broadly, these modifications can be classified into activating and repressive categories based on their typical effect on or association with transcriptional access (Kouzarides 2007; Zhao and Garcia 2015). Activating PTMs often act locally at the gene level, appearing either in direct response to transcription or prior to it, facilitating continued gene access and progress of the transcriptional machinery through gene bodies. In contrast, repressive PTMs can mark both vast genomic regions and individual genes and preventing transcriptional access to sequences not needed for a particular cell's function, including genic and intergenic regions. Both types of chromatin modifications play a key role in specifying cellular function through controlling access to and interpretation of genomic DNA in euchromatin and heterochromatin respectively.

Evidence in the literature suggests that differential expansion of heterochromatin domains is a major contributor to defining the genomic pattern of gene expression that determines cell type. Broadly, heterochromatin domains are formed in a two-step processes whereby first the domains are seeded at DNA-encoded nucleation sites which recruit the enzymes responsible for catalyzing chemical modification to nucleosomes and also downstream effector protein required to repress gene expression (Hall et al. 2002; Jia et al. 2004; Reyes-Turcu et al. 2011). Next, the histone modifying enzymes, effector proteins, and the repressed state expand outward from

2

nucleation sites in a sequence-indifferent process termed spreading (Talbert and Henikoff 2006). For example, in development small pre-existing heterochromatin domains marked by methylation of histone H3 Lysine 9 (H3K9me) expand via spreading to repress up to megabases of DNA encoding genes that contribute to orthogonal cell types (Wen et al. 2009; Zhu et al. 2013; Zylicz et al. 2015; Zylicz et al. 2018; Nicetto and Zaret 2019). Interestingly, this process occurs in a lineage-dependent manner suggesting an iterative process that tracks with the cell fates in the lineage tree (Chen et al. 2012). Importantly, once a heterochromatin domain is established via spreading, it has to be precisely retained throughout the cell cycle to achieve a stable cell fate. Failure to properly recapitulate heterochromatin domains over repeated cell division can have critical consequences including cancer (Peters et al. 2001; Chen et al. 2010b; Casciello et al. 2015; Feinberg et al. 2016; Janssen et al. 2018) and aging (Villeponteau 1997; Scaffidi and Misteli 2006; Zhang et al. 2015).

Significant progress has been made in elucidating when and where heterochromatic PTMs (including H3K9me) are placed and has also generated an extensive list of proteins that contribute to gene silencing and heterochromatin domain formation broadly. While the process of nucleation is well studied (Hall et al. 2002; Jia et al. 2004; Verdel et al. 2004; Bayne et al. 2010; Reyes-Turcu et al. 2011), understanding how multicellular organisms generate such complexity in cell types through differential domain expansion followed by clonal inheritance of cell state requires further investigation into the sequence-indifferent process of spreading. Techniques which measure the distribution of PTMs along the chromatin and reporter assays in model systems that typically rely on marker genes which assess the chromatin state when challenged with selective pressure, have yielded only a short list of protein components that are required to propagate these marks beyond nucleation sites (Nakayama et al. 2000; Noma et al.

2004; Zhang et al. 2008a). Similarly, little is known about the requirements for stable inheritance of nucleation-site distal domains, though in some systems heterochromatic signals can be propagated though through retention of modified nucleosomes (Alabert et al. 2015).

One major hurdle in the chromatin field has been the difficulty to specifically separate heterochromatin spreading from the process of nucleation. A second, and related, challenge is obtaining quantitative measurements of the chromatin state of single, isogenic cells. This can inform the relationship between individual nucleation and spreading events to the broader population behaviors of gene silencing. To address these key challenges, we developed a molecular reporter system, which we term the Heterochromatin Spreading Sensor (HSS), that relies on genetically integrated transcriptional reporters that encode fluorescent proteins whose output can be quantitatively assessed without selection by methods including flow cytometry and fluorescence microscopy (Al-Sady et al. 2016; Greenstein et al. 2018; Greenstein et al. 2019). In this method, two separate reporters are placed within or adjacent to heterochromatin loci such that they serve as a readout for nucleation and spreading independently. A third reporter serves as a noise filter and is constitutively expressed from a protected euchromatic locus. We additionally developed a concomitant computational analysis pipeline which can both specifically isolate the contribution of spreading independent of nucleation and also capture the population distribution of gene silencing phenotypes broadly. Leveraging the power of this novel assay, I address two key aspects of heterochromatin spreading in the formation and maintenance of cell identity in this work: (1) the molecular regulation of differential expansion that typifies heterochromatin domains that change with cell state and (2) the requirements for stable inheritance of the repressed state distal to nucleation sites but within established heterochromatin

domains. These correspond to spatial and temporal mechanisms of heterochromatin spreading regulation, respectively.

Once a heterochromatin domain is established by spreading, it must be faithfully recapitulated in daughter cells after cell division. There are in principle two possible ways that this happens – (1) this gene expression pattern is retained when the cell replicates its DNA and divides or (2) the pattern is recreated *de novo* at every cell division, potentially by other cues within the cellular milieu, perhaps faithfully spreading from the DNA-encoded nucleation site to the same extent each cycle. While it is likely that the global pattern of cell-specific gene expression relies on both mechanisms, there is great interest in understanding the non-DNA modalities, such as histone PTMs, that can be inherited across division, and hence are epigenetic in nature, to ensure faithful retention of cell type. Chromatin PTMs have long been targets of interest in the epigenetics field as they can modulate gene expression and exist in the local environment of DNA, perhaps facilitating their inheritance through division. While recent work speculates that many activating PTMs may not be epigenetic in nature (Reinberg and Vales 2018), it has long been understood that epigenetic gene repression can be enacted through post translational modification to DNA itself via methylation at CpG dinucleotides and a semi-conservative mechanism of propagation through DNA replication (Jones and Liang 2009; Law and Jacobsen 2010). While the known link between DNA methylation and repressive chromatin modifications (Estève et al. 2006; Bostick et al. 2007; Du et al. 2015) likely accounts for at least some of amount of potential epigenetic inheritance, some biological systems do not rely on DNA methylation for the memory of the cellular state across division (Capuano et al. 2014). In this context repressive heterochromatin PTMs appear to the be drivers of epigenetic memory and it is also likely that this capacity is retained in the more complex systems that additionally exhibit

DNA methylation (see Discussion and (Greenstein and Al-Sady 2019)). Understanding the contribution of heterochromatin PTMs to the epigenetic stability of cell fate is critical to our understanding of development and the processes that fail in aging and disease.

We aimed to investigate the roles of heterochromatin spreading and epigenetic inheritance of heterochromatin PTMs in the formation and maintenance of cell fate in the fission yeast *Schizosaccharomyces pombe*. Fission yeast is a unicellular eukaryotic organism that shares key aspects of genomic patterning through heterochromatin domain formation with higher order metazoans and is also highly genetically tractable with limited genetic redundancy. As in more complex systems, fission yeast heterochromatin is marked by di- and tri-methylation of H3K9 (Nakayama et al. 2001). There are three major constitutive heterochromatin loci in *S. pombe* – the pericentromere, the subtelomere, and the silent mating type (MAT) locus, which is critical to cell identity. In addition, cell type-specific (facultative) heterochromatin forms at developmentally regulated meiotic genes dependent on extracellular signals (Zofall et al. 2012). Silencing at these loci depends on both 'writer' enzymes and 'reader' proteins, including HP1, that create and sense the heterochromatic H3K9me mark (Nakayama et al. 2000; Nakayama et al. 2001; Zhang et al. 2008a; Al-Sady et al. 2013). However, unlike metazoans, this organism does not feature DNA methylation (Capuano et al. 2014), providing an ideal system in which to assess the contribution of chromatin PTMs to epigenetic inheritance of cell fate. By deploying the HSS at various heterochromatic loci in the fission yeast genome, along with molecular genetics, genomic, and biochemical approaches, we have uncovered novel mechanisms for spatial and temporal regulation of heterochromatin domain formation in shaping cell identity.

To gain a better understanding of the requirements for robust intergenerational inheritance of heterochromatin domains and memory of cellular state, we employed the HSS

system to report on heterochromatin spreading and domain formation in fission yeast (Chapter 1, reproduced from (Greenstein et al. 2018)). Using this assay, we first assessed the intrinsic capacity of heterochromatin to spread at an engineered ectopic heterochromatin locus. We then implemented the HSS in dissecting the requirements for the robust spreading and inheritance of repression at the fission yeast MAT cell identity locus. Through this analysis we find that different classes of nucleation elements yield distinct spreading behaviors both within a population (steady state) and over generational time (temporal dynamics). We further elucidate the contributions of these different nucleators to the stability of the heterochromatin locus when challenged with chemical, environmental, and genetic perturbations and identify differential reduction of histone turnover as a potential explanation for variation in temporal epigenetic stability.

Having characterized the contributions of different nucleators to heterochromatin spreading efficiency and epigenetic stability, we sought to determine *trans*-acting factors and protein complexes that regulate spreading, both broadly across multiple heterochromatin contexts and also in a context specific manner (Chapter 2, reproduced from (Greenstein et al. 2020)). To address this, we utilized our previously described HSS backgrounds to perform a series of reverse genetic screens querying a ~400 gene nuclear function deletion library in a variety of chromatin contexts. From analysis of this dataset, we find that the genetic regulators of heterochromatin spreading differ both between chromatin context and to some extent by type of nucleator. We identify both positive and negative regulators of spreading, independent of nucleation, at both the individual protein and protein complex levels.

Next, leveraging results from a related genetic screen, we explore mechanisms for regulation of cell type-specific heterochromatin domain expansion (Chapter 3, reproduced from

(Greenstein et al. 2019)). In this study we provide evidence for the spatial regulation of heterochromatin spreading by euchromatin and its associated activities. Specifically, we uncover a role for active genes in containment of the spreading reaction when heterochromatin invades transcriptionally active euchromatic regions, serving as a model for domain expansion characteristic of early development. This mechanism is mediated not through the act of transcription itself, but instead via direct and indirect effects of methylation of histone H3 Lysine 4 (H3K4me) by the Set1/COMPASS complex, supporting the long-observed mutual exclusion of H3K4me- and H3K9me-marked chromatin in many organisms. We also provide preliminary evidence for mechanisms by which gene orientation could sculpt the positional extent of heterochromatin domains, possibly explaining the gene orientation bias observed at heterochromatin domain borders in mammalian systems.

Finally, I present unpublished work towards an understanding of the requirements of the *S. pombe* H3K9 methylase in regards to heterochromatin spreading and gene silencing and as well as a preliminary system to measure the kinetics of spreading *in vivo* (Chapter 4). Future elaboration on these initial results will hopefully yield important insights on the biochemical and biophysical constraints of the spreading reaction at the individual protein level, as well as provide a platform for building and testing models of spreading *in vitro* and *in silico*.

Together the results described in this document constitute significant advances in our understanding of the spatial and temporal regulation of heterochromatin spreading and their contributions to defining cell identity through the formation and maintenance of an epigenetic, chromatin-driven pattern of gene expression.

2.   Noncoding RNA-nucleated heterochromatin spreading is intrinsically labile and requires accessory elements for epigenetic stability.

FOREWORD

The following is a reproduced from a previously published manuscript (Greenstein et al. 2018). This chapter describes the implementation and validation of a novel method to capture the population distribution and temporal dynamics of heterochromatin spreading in single cells while specifically controlling for the separate process of nucleation. This spreading specific assay is then utilized to dissect both the intrinsic capacity of heterochromatin spreading and the origins of the regulated properties achieved at robustly silenced cell identity loci where repression is inherited with high fidelity throughout the population and overtime.

ABSTRACT

The heterochromatin spreading reaction is a central contributor to the formation of gene-repressive structures, which are re-established with high positional precision, or fidelity, following replication. How the spreading reaction contributes to this fidelity is not clear. To resolve the origins of stable inheritance of repression, we probed the intrinsic character of spreading events in fission yeast using a system that quantitatively describes the spreading reaction in live single cells. We show that spreading triggered by noncoding RNA-nucleated elements is stochastic, multimodal, and fluctuates dynamically across time. This lack of stability correlates with high histone turnover. At the mating type locus, this unstable behavior is restrained by an accessory *cis*-acting element *REIII*, which represses histone turnover. Further, *REIII* safeguards epigenetic memory against environmental perturbations. Our results suggest

that the most prevalent type of spreading, driven by noncoding RNA-nucleators, is epigenetically unstable and requires collaboration with accessory elements to achieve high fidelity.

INTRODUCTION

The formation of gene-repressive heterochromatin domains is critical for genome integrity and for the establishment and maintenance of cell identity. Most heterochromatin formation occurs by a sequence-indifferent spreading reaction that propagates heterochromatic marks, structural proteins, and associated effector proteins outwards from nucleation sites. The precise extent of the spreading reaction has critical heritable consequences for cell identity. For example, in early pluripotent precursors, pre-existing heterochromatin domains spread, sometimes over megabases, to repress specifiers of inappropriate fates. Importantly, the final extent of spreading from a locus depends on the lineage pathway, hence it varies across different precursors (Wen et al. 2009; Zhu et al. 2013), and has to be precise to achieve a stable cell fate and avoid disease (Ceol et al. 2011). Similarly, spreading also specifies cell type in yeasts, where the cell type is maintained by repressing the mating cassettes at the mating type loci (Ekwall et al. 1991; Rusche et al. 2003). Despite the centrality of the spreading reaction in shaping cell identity, its native and intrinsic cellular characteristics, as well as mechanisms for its inter-generational propagation, have remained opaque.

We have some understanding of how cells inherit silencing at nucleation sites, i.e. the DNA-sequence driven component of heterochromatin. Recent results in heterochromatin systems signaled by Histone 3 Lysine 9 and Lysine 27 methylation (H3K9me and H3K27me) indicate that several mechanisms act together to ensure intergenerational inheritance: continuous DNA-mediated recruitment of the histone methylase (Jia et al. 2004; Audergon et al. 2015;

Ragunathan et al. 2015; Laprell et al. 2017; Wang and Moazed 2017), low histone turnover (Aygun et al. 2013; Taneja et al. 2017), as well as the positive "read-write" feedback loop for histone methylases (Zhang et al. 2008a; Al-Sady et al. 2013). Additionally, studies suggest that either the histone mark (Gaydos et al. 2014) or the histone methylases (Petruk et al. 2012) can persist trans-generationally.

These insights concerning nucleation sites do not necessarily account for how regions of heterochromatin distal to these sites are maintained. Unlike nucleation, which depends on DNA based enzyme recruitment (Verdel et al. 2004; Bayne et al. 2010), spreading depends on the ability of the system to propagate along the chromosome, independent of the underlying DNA sequence. Such propagation requires the "read-write" positive feedback function of the system (Noma et al. 2004; Zhang et al. 2008a; Margueron et al. 2009; Al-Sady et al. 2013; Muller et al. 2016).

To determine how the spreading reaction acts in the maintenance of cell fate, it is central to understand the native behavior of two interconnected but separable phases of spreading: The initial spreading event, and the propagation of the states formed by this initial event through cell divisions. There is evidence that the initial spreading, at least in contexts outside the native chromosomal position, is stochastic, i.e. only some nucleation events result in a spreading event. This was first demonstrated by observing position effect variegation (PEV) in flies (Muller 1930; Elgin and Reuter 2013). Such stochastic behavior would have to be mitigated across cells to achieve a coherent specification outcome.

Intergenerational propagation of spreading is straightforward to conceptualize when epigenetic information is strongly reinforced, but more challenging in situations where modified nucleosomes are less likely to persist. This is the case for H3K9me-signaled heterochromatin in

the fission yeast system, which lacks DNA methylation that can reinforce the epigenetic state. Persistence of the modified state is opposed by an anti-silencing protein Epe1 (Ayoub et al. 2003; Zofall and Grewal 2006), which acts by antagonizing retention of H3K9me histones (Aygun et al. 2013; Ragunathan et al. 2015), and passage through S-phase, which significantly weakens heterochromatin domains (Chen et al. 2008). For fission yeast, there is evidence in favor of both high fidelity and stochastic propagation of the state formed by spreading. In support of a high fidelity model, theoretical work suggests that heterochromatin can display fundamentally bistable behavior, indicating that the 'ON' and 'OFF' states are intrinsically highly stable (Dodd et al. 2007). Similar bistable behavior has also been experimentally observed in plants (Angel et al. 2011; Angel et al. 2015). Conversely, the telomere position effect (TPE) observed in budding and fission yeast supports a model where intergenerational inheritance is fundamentally stochastic. In TPE the heterochromatic state is switched at high frequencies in the inheriting generations (Gottschling et al. 1990; Nimmo et al. 1994).

To distinguish whether spreading shapes and enables epigenetic maintenance of a cell identity locus via either of those modes, or combinations thereof, we focused on one of the most well understood heterochromatin loci, the fission yeast MAT locus, as a model. This locus remains tightly repressed to avoid simultaneous expression of both mating cassettes (Ekwall et al. 1991; Noma et al. 2001). The MAT locus contains two *cis* elements that directly recruit H3K9me. (1) *cenH,* which is related to the *dg* and *dh* repeats at the pericentromere and *tlh2* at the subtelomere (Grewal and Klar 1997; Hansen et al. 2006). These sequences nucleate H3K9me by at least two pathways, which depend on transcription of noncoding RNAs (ncRNAs): the RNAi pathway (Hall et al. 2002; Volpe et al. 2002), and at least one separate pathway dependent on nascent RNA polymerase II transcripts, which requires the budding yeast Nrd1 homolog Seb1

(Marina et al. 2013)(collectively "ncRNA-nucleation"). Separately, and unique to the MAT locus, (2) a region downstream of *cenH* including the *REIII* element, which recruits the H3K9 histone methylase, HP1 proteins, and histone deacetylases (HDACs). This is dependent on *REIII*-bound transcription factors (Jia et al. 2004; Kim et al. 2004; Yamada et al. 2005), but independent of RNA processes. Heterochromatin formation within the MAT locus is confined by boundary elements (Noma et al. 2001; Noma et al. 2006).

In this work, we probe heterochromatin spreading nucleated both at the MAT locus as well as ectopically in the genome with a "heterochromatin spreading sensor" (HSS), which enables quantitative examination of spreading separately from nucleation in single *S. pombe* cells. Using the HSS, we show that ncRNA- dependent elements trigger epigenetically unstable spreading that is stabilized by an accessory RNA-independent *cis*-acting element. Both elements collaborate to form a high-fidelity domain. The strategy we uncover has important implications for how heterochromatin spreading achieves and maintains "epigenetic" character and can safeguard cell identity against environmental perturbations.

RESULTS

A single cell heterochromatin spreading sensor (HSS) controls for nucleation and cellular noise.

To assess the intrinsic behavior of heterochromatin spreading and what shapes its precise re-establishment with respect to position and extent of repression ("fidelity"), we employed transcriptionally encoded fluorescent reporters to read silencing by heterochromatin at a given locus, as previously reported. Several critical improvements over prior systems enable documentation of the spreading reaction at high sensitivity (Xu et al. 2006; Osborne et al. 2009;

Hathaway et al. 2012; Bintu et al. 2016; Obersriebnig et al. 2016). First, our system has high

signal to noise and minimized delay from epigenetic changes to fluorescent output. We

accomplish this using the weak, well-characterized *ade6* gene promoter (*ade6p*) (Allshire et al.

1994; Kagansky et al. 2009) to drive production of bright, fast-folding fluorescent proteins

(XFPs) (Al-Sady et al. 2016). Second, our system provides separate sensors for nucleation,

spreading, and cellular noise. We used *ade6p*-driven recoded super-folder GFP (Pedelacq et al.

2006) ("green") and monomeric Kusabira Orange (Sakaue-Sawano et al. 2008) ("orange") to

report on nucleation and spreading, respectively (**Figure 2.1A**). A third XFP, *ade6p*-driven

triple fusion of E2Crimson (Strack et al. 2009) ("red", noise filter), is fully uncoupled from

heterochromatin and inserted in a euchromatic locus. Here it reports on intrinsic or extrinsic

noise that arises from cell-to-cell variation in the content of specific and general transcription

factors and also translational efficiency (**Figure 2.1A**). To validate this reporter system, we

characterized the non-heterochromatic state, via null mutation of *clr4* (*Δclr4*), encoding the only

S. pombe H3K9 methyltransferase. We show that in the absence of heterochromatin, expression

of the noise reporter ("red") correlates well with that of reporters for both nucleation ("green")

and spreading ("orange") (**Figure 2.1 Supplement A, B**), especially when all cells in the

population are considered without applying a size gate (**Figure 2.1 Supplement B**, ρ~0.83-0.93).

This analysis mode is required when cell number is limiting. When a smaller subset is considered

where all the cells are of similar size and stage of the cell cycle, the correlation still provides

useful noise filtering (**Figure 2.1 Supplement A**), which becomes evident when the

normalization is applied to *clr4+* cells that fall in the size gate (**Figure 2.1 Supplement C**).

Thus, cellular noise is mitigated by dividing the signals from the proximal "green" and distal

"orange" heterochromatic reporters by the signal of the "red", euchromatic reporter

("green"/"red"; "orange"/"red"). Together, these elements constitute our heterochromatin spreading sensor (HSS).
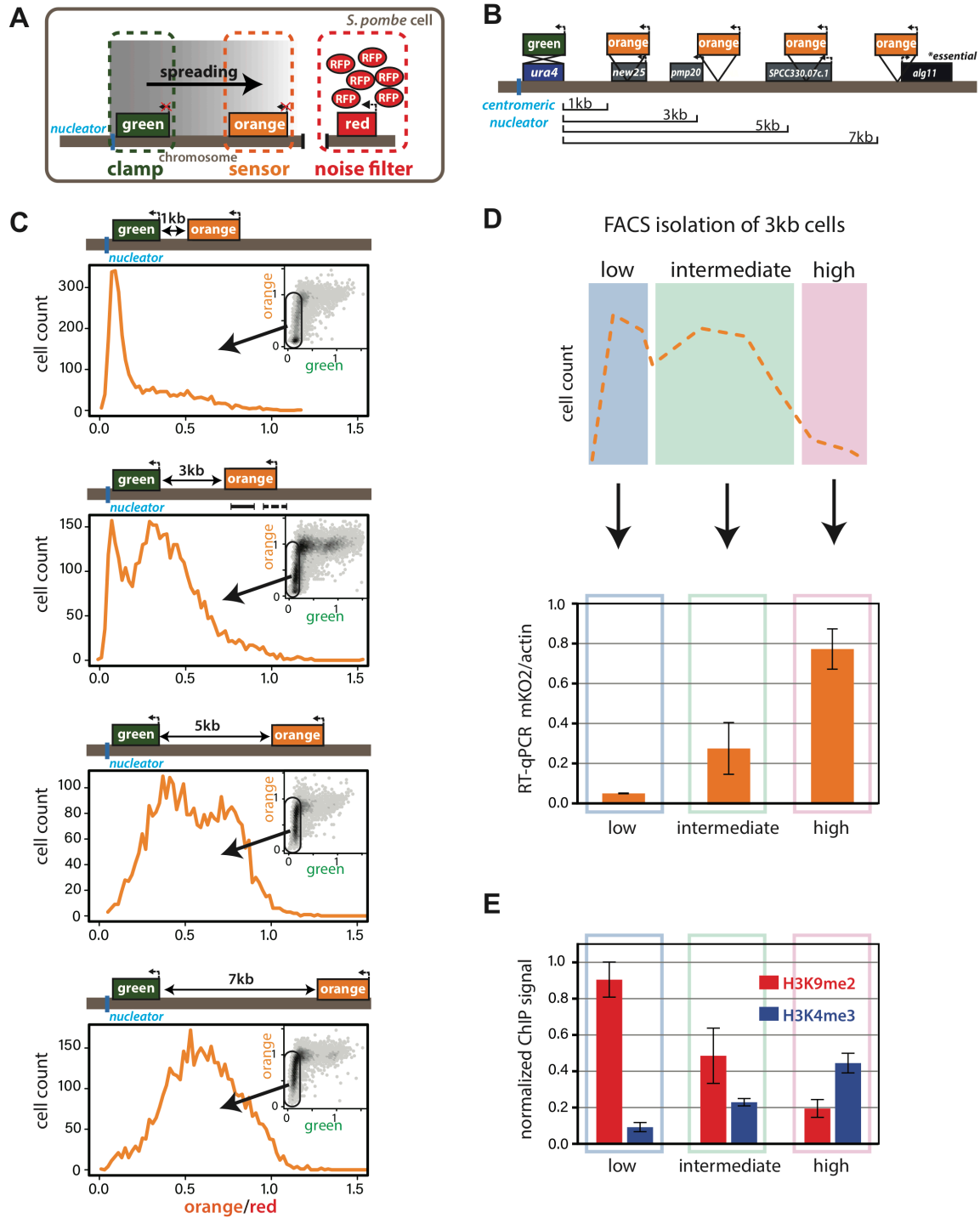

Figure 2.1: Heterochromatin spreading from ncRNA-nucleated elements is stochastic and produces intermediate states.

A. Overview of heterochromatin spreading sensor. Three transcriptionally encoded fluorescent proteins are inserted in the genome: The "clamp" site enables isolation of successful nucleation events, the "sensor" reports on spreading events and the "noise filter" normalizes for cell-to-cell noise. **B.** Overview of the $ura4::dh$HSS$^{1\text{-}7kb}$ strains. Genes downstream of the "green" nucleation color are annotated. **C.** Spreading from $ura4::dh$ visualized by the HSS with "orange" inserted at different distances shown in B. The "red"-normalized "orange" fluorescence distribution of "green"^{OFF} cells plotted on a histogram. Inset: 2D-density hexbin plot showing red-normalized "green" and "orange" fluorescence within the size gate, with no "green" or "orange" filtering. The "green"^{OFF} population is schematically circled. The fluorescence values are normalized to =1 for the $\Delta clr4$ derivate of each strain. **D.** TOP: cartoon overview of the FACS experiment for D. and E. "green"^{OFF} cells collected from the $ura4::dh$HSS^{3kb} were separated in three populations ("Low", "Intermediate" and "High") as shown schematically based on the "orange" fluorescence. BOTTOM: "orange" RT-qPCR signal for the indicated populations. The y-axis is scaled to =1 based on the "orange" signal in $\Delta clr4$. Error bars indicate standard deviation of two replicate RNA isolations. **E.** ChIP for H3K9me2 and H3K4me3 in the same populations as D. Each ChIP is normalized over input and scaled to =1 for a positive control locus (dh repeat for H3K9me2 and $act1$ promoter for H3K4me3). Error bars indicate standard deviation of two technical ChIP replicates. Primer pairs for RT-qPCR and ChIP are indicated by solid and dashed line, respectively, in the C. $ura4::dh$HSS^{3kb} diagram.

Spreading from ectopic ncRNA nucleators is stochastic and produces intermediate states.

We first examined the intrinsic behavior of the heterochromatin spreading reaction in an ectopic context. We constructed the initial ectopic HSS based on a strain where a part of the centromeric ncRNA-driven nucleation element (dh) is inserted proximal to the endogenous $ura4$ gene (Canzio et al. 2011; Marina et al. 2013). We replaced the $ura4+$ open reading frame (ORF) with "green" to track nucleation element-proximal events. Then, to track distal events, we inserted "orange" at one of several sites downstream from "green" ($ura4::dh$HSS^{1kb}, $ura4::dh$HSS^{3kb}, $ura4::dh$HSS^{5kb} $ura4::dh$HSS^{7kb}, **Figure 2.1B**). The noise filter ("red") was inserted between $SPBC1711.11$ and $SPBC1711.12$, a $bona\ fide$ euchromatic region (Garcia et al. 2015). All strains were initially constructed in a $\Delta clr4$ background, and we initiated heterochromatin formation by crossing in $clr4+$. We assessed heterochromatin formation after $\sim 80\text{-}100$ generations by quantifying the production of "green" and "orange". This period is significantly longer than ~ 25 generation timeframe required for full formation of a heterochromatic domain (Obersriebnig et al. 2016), ensuring that the population is at equilibrium.

To quantitatively assess the products of heterochromatin formation, we performed steady-state flow cytometry on log-phase cells, which were size-gated for small, recently divided cells (~91% G2, **Figure 2.1 Supplement D** and SUPPLEMENTAL MATERIALS AND METHODS) to remove size- and cell cycle-related effects. At this stage, we only normalize the cells by the "red" noise filter and scale the signal in each channel to *Δclr4*, giving us a broad overview of the possible expression states of "green" and "orange". We observe no cells that express "green" but repress "orange" (insets, **Figure 2.1C**), instead, all cells that are fully or partially "orange" repressed are also robustly "green" repressed. This observation, together with our finding that "green" repression kinetically anticipates "orange" repression (**Figure 2.3 Supplement 1**), is consistent with heterochromatin spreading outward from the *ura4::dh* nucleator. Considering "green" repression thus a proxy for nucleation, we observed that cells populate a wide range of nucleation states rather than a single state, with the distribution of repressed states varying among the HSS distance sensor strains (*ura4::dh*HSS[1-7kb], **Figure 2.1C**). To specifically examine cells that have fully nucleated, we applied a computational "nucleation clamp" that isolates cells with a "green" signal that is lower than the median plus two standard deviations of wild-type cells containing no XFPs (see SUPPLEMENTAL MATERIALS AND METHODS). Using "orange" as a spreading proxy, we found spreading to be stochastic in nucleated cells, with some cells exhibiting full repression, but others partial repression or full de-repression (*Δclr4*, x=1) of the "orange" spreading sensor. The proportion of cells that are fully repressed by spreading declines linearly with distance (scheme, **Figure 2.1B;** data **Figure 2.1C**). Intriguingly, cells that are not fully repressed mostly exhibit intermediate levels of repression, which are neither at values of full repression or de-repression.

We next assessed the nature of these intermediate states in the 3kb distance reporter strain, where ~30% of cells had maximal repression at the "orange" locus and the remainder had intermediate states ranging from strongly to weakly repressed. Using Fluorescence Activated Cell Sorting (FACS), we gated for successful nucleation in the "green" channel and then binned the "orange" channel for fully repressed (low), intermediate and de-repressed (high) populations (**Figure 2.1D**, cartoon). We queried each bin for molecular events associated with heterochromatin formation, using RT-qPCR to determine the expression levels of "orange", and Chromatin Immunoprecipitation (ChIP) to query the presence of the marks H3K9me2 and H3K4me3. These marks are thought to be mutually exclusive, associating with repressed heterochromatin and active promoters, respectively (Noma et al. 2001). The message level of "orange" is tightly repressed in the "low" population (0.05 of max), partially repressed in the intermediate population (0.3 of max), and nearly fully "de-repressed" (0.8 of max) in the "high" population. Thus, cells with intermediate fluorescence also exhibit partial gene repression, demonstrating that fluorescence accurately reports on gene expression (**Figure 2.1D,** RT primers indicated in diagram in 1C, solid line). Histone modification levels also correlated well with the HSS signals (**Figure 2.1E,** ChIP primers indicated in diagram in 1C, dashed line). The "low" fluorescence population has high H3K9me2 (0.9 of *dh*, positive control) and low H3K4me3 (0.09 of actin, positive control); the intermediate population had intermediate H3K9me2 (0.49 of *dh*) and H3K4me3 (0.23 of actin), and the high population had low H3K9me2 (0.2 of *dh*) and higher H3K4me3 (0.44 of actin). Hence, successfully nucleated cells with intermediate fluorescence also exhibit intermediate amounts of the mRNA for "orange" and histone marks reflecting heterochromatin (H3K9me2) and transcriptional activity (H3K4me3). These results

support the notion that intermediate states of repression observed by cytometry represent intermediate states of spreading.

These observations are not due to the particularities of the ectopic site chosen or the behavior of the XFPs, as our results are recapitulated at the *his1* locus (*his1::dh*HSS^{3kb}, **Figure 2.1 Supplement E**), which contains only one gene (*rec10*) in the "spreading zone", rather than several transcriptional units. Additionally, switching the nucleation and spreading reporter fluorophores produced similar results (**Figure 2.1 Supplement E**). These results suggest that ncRNA-driven heterochromatin spreading at ectopic sites is intrinsically stochastic and multimodal, producing intermediate states of repression.

Distinct forms of heterochromatin spreading at MAT.

We next examined spreading behavior at the endogenous mating type locus (MAT), which is tightly repressed (Grewal and Klar 1997; Thon et al. 2002) and a *bona fide* high-fidelity locus, as it can behave in a bistable manner with stable epigenetic inheritance even when disrupted (Grewal and Klar 1996). The MAT locus has two known elements shown to recruit the H3K9 methylase Clr4: the *cenH* element, homologous to the ncRNA-nucleated *dh* fragment we inserted at *ura4* and *his1*, and the RNA-independent element termed *REIII* (Thon et al. 1999; Jia et al. 2004). At *REIII*, two stress-responsive transcription factors, Atf1 and Pcr1, which form a heterodimer (Wahls and Smith 1994), recognize two DNA binding sites within *REIII*, directly recruit Clr4, Swi6/HP1 and histone deacetylases (HDACs) (Jia et al. 2004; Kim et al. 2004) and are required for heterochromatin formation at MAT when *cenH* is compromised (Noma et al. 2004). We validated that MAT retains its well-documented tight repression following insertion of the HSS, placing the "green" reporter within the *cenH* nucleator, and the "orange" reporter

proximal to the *REIII* nucleator. Both colors were fully repressed in the large majority of cells (**Figure 2.2B**), which is reproduced when the color orientations are reversed (**Figure 2.2 Supplement 1A**). However, for both reporter configurations, the *REIII* proximal color showed a small proportion of cells that are slightly de-repressed compared to the *cenH* internal color, consistent with previous findings (Thon and Friis 1997). We conclude that the HSS can be used to dissect spreading at the MAT locus.

We then examined spreading in cells nucleated solely by the *cenH* element. The *REIII* nucleator was inactivated by deleting the critical *cis*-acting Atf1/Pcr1 binding sites, to create a strain designated *ΔREIII*[HSS] (**Figure 2.2C**). To our surprise, the high fidelity that the MAT locus exhibits in the repressed state (Grewal and Klar 1996) disappeared. Instead, *cenH* nucleated spreading in the *ΔREIII* strain behaved similarly to spreading from the ectopic ncRNA-nucleated strains, showing high stochasticity and predominantly intermediate repression states (**Figure 2.2C**). We wanted to address if this stochastic silencing is reflected in weakened heterochromatin assembly. We preformed ChIP for H3K9me2 and H3K9me3, marks signaling heterochromatin assembly (Nakayama et al. 2001) and repression or spreading (Zhang et al. 2008a; Al-Sady et al. 2013; Jih et al. 2017), respectively. We found that these marks decline progressively towards the distal "orange" reporter in *ΔREIII*[HSS] (**Figure 2.2D**), compared to the wild-type (WT) MAT[HSS]. This is consistent with the observed tight repression for WT MAT[HSS] (**Figure 2.2B**) and weakened silencing at the distal "orange" in *ΔREIII*[HSS] (**Figure 2.2D**). It is possible that this difference in spreading results from an altered heterochromatin structure at *cenH* in *ΔREIII*[HSS]. However, H3K9me2 and me3 accumulation does not differ between *ΔREIII*[HSS] and WT MAT[HSS] at the *cenH* nucleator, or the leftward *REII* locus (**Figure 2.2D**). Thus, the observed behavior of

$\Delta REIII^{\text{HSS}}$ is consistent with stochastic and multimodal spreading, rather than compromised
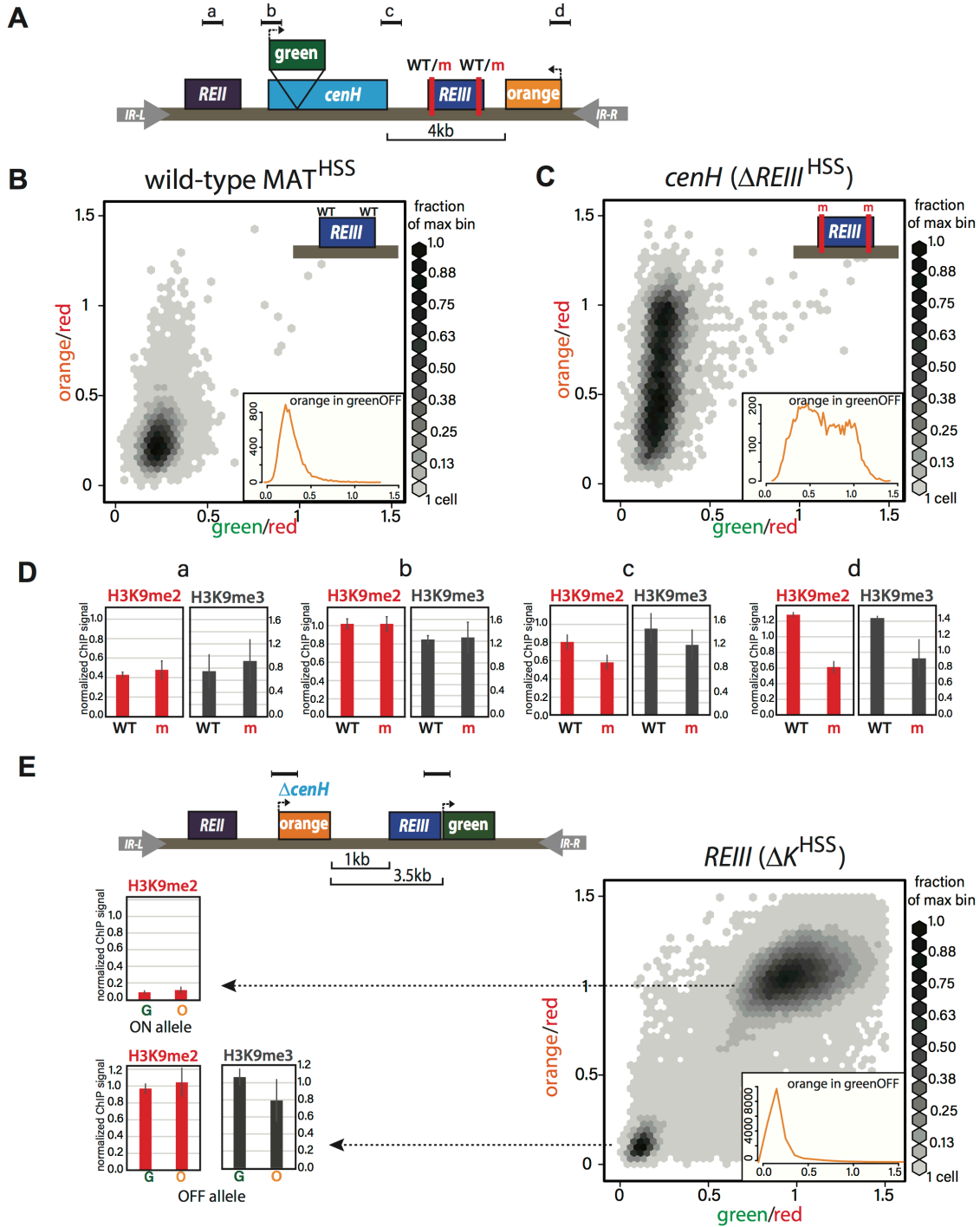
nucleation at *cenH*.

Figure 2.2: ncRNA- dependent and independent nucleation yields qualitatively different spreading reactions in the MAT locus.

A. Diagram of the reporters within MAT^HSS and *ΔREIII*^HSS. WT and m for *REIII* indicate the presence or deletion of the Atf1/Pcr1 binding sites, respectively. **B.** 2D-density hexbin plot showing the "red"-normalized "green" and "orange" fluorescence for wild-type MAT^HSS cells. Scale bar shows every other bin cutoff as a fraction of the bin with the most cells. Inset: histogram of the "red"-normalized "orange" fluorescence distribution of "green"^OFF cells. **C.** 2D-density hexbin plot and inset as above for *ΔREIII*^HSS, which contains two 7bp Atf1/Pcr1 binding site deletions (m) within the *REIII* element. **D.** ChIP for H3K9me2 (red) and H3K9me3 (grey) for amplicons indicated in **A**. normalized to *dh.* WT, wild-type MAT^HSS, m, *ΔREIII*^HSS. **E.** TOP: diagram of the reporters within *ΔK*^HSS. The *cenH* nucleator and additional 5' sequence are deleted and replaced by "orange". "green" is located directly proximal to *REIII* and serves as the nucleation clamp. BOTTOM: 2D- density hexbin plot and inset as above. LEFT: ChIP for H3K9me2 (red) and H3K9me3 (grey) for amplicons for "green" and "orange" indicated in TOP in isolated *ΔK*^HSS-ON or *ΔK*^HSS-OFF alleles. In hexbin plots, the *Δclr4* derivative of each strain was used to normalize the X- and Y-axes to =1. Error bars indicate standard deviation of technical replicates.

To examine heterochromatin formation independent of *cenH*, we used the historical *ΔK* strain, where the entire *cenH* nucleation element is deleted and replaced with a *ura4+* reporter (Grewal and Klar 1996). We introduced the HSS into this context (*ΔK*^HSS, **Figure 2.2E**), placing the "green" reporter proximal to *REIII* and the "orange" reporter distally, replacing *ura4*. We then introduced *clr4+* by cross and directly cultured colonies derived from germinated *clr4+* spores. We found that although *ΔK*^HSS has very weak nucleation compared to strains with intact ncRNA nucleators, the distribution of cells is sharply bimodal: Cells were either repressed at both reporters ('OFF', lower left corner) or de-repressed at both reporters ('ON', upper right corner; **Figure 2.2E**). We note that isolation of single colonies on nonselective media from original spores of the cross yields mostly ON (*ΔK*^HSS-ON) or OFF (*ΔK*^HSS-OFF) colonies, consistent with each state being metastable (Grewal and Klar 1996; Thon and Friis 1997; Grewal et al. 1998). This heterochromatin formation pattern requires *REIII,* as in 34/34 strains tested, no silencing can be established if Atf1/Pcr1 binding sites are deleted before *clr4+* is introduced (**Figure 2.2 Supplement 2A,B**). Additionally, the bimodal behavior does not require the H3K9me-independent gene-repressive *REII* element (Hansen et al. 2011), as *ΔK*^HSS *REII::LEU2*, containing a deletion of *REII*, behaved similarly to *ΔK*^HSS (**Figure 2.2 Supplement 1C**), and is

further independent of reporter placement (**Figure 2.2 Supplement 2C, D**). We next characterized the molecular signature of the locus. While in our two-color plots cells that were repressed in "green" did not show any de-repression in "orange" (**Figure 2.2E,** cells in bottom left corner), we wanted to test if the heterochromatic state at these loci correlated with this silencing pattern. Since we can isolate $\Delta K^{\text{HSS-ON}}$ and $\Delta K^{\text{HSS-OFF}}$ alleles by simple plating of ΔK^{HSS} cells, we performed H3K9me2 ChIP on both and H3K9me3 ChIP for $\Delta K^{\text{HSS-OFF}}$ cells (not detectable for $\Delta K^{\text{HSS-ON}}$). We found that methylation correlates with the repression state (**Figure 2.2E**) and importantly, not does not significantly differ between "green" and "orange". Together, these results indicate that in $\Delta K^{\text{HSS-OFF}}$ cells heterochromatin spreading is continuous across the locus and does not, unlike *cenH*-triggered spreading, accumulate any intermediates.

Multi-generational single cell imaging reveals ncRNA-driven spreading to be unstable.

Our measurements thus far cannot reveal the dynamics of transitions between states. This requires long-term imaging of cells over a substantial number of generations (>20), which is difficult with traditional microscopy because of cell crowding effects. To deal with this issue, we used the Fission Yeast Lifespan Micro-dissector (FYLM) microfluidic device (Spivey et al. 2014; Spivey et al. 2017), which traps the old pole of a rod-shaped *S. pombe* cell at the bottom of a chamber well for its entire lifetime. Sibling cells generated at the new pole by medial fission eventually exit the chamber. We continuously image the old-pole cell with fluorescence microscopy for up to 60hrs (**Figure 2.3A**). We note that unlike *Saccharomyces cerevisiae*, *S. pombe* does not execute an aging program but rather dies stochastically (Coelho et al. 2013; Nakaoka and Wakamoto 2017; Spivey et al. 2017). Thus, imaging *S. pombe* over long timescales avoids the confounding effects of aging on epigenetic behavior (Guarente 2000; Li et al. 2017).

To capture the long-range dynamics of spreading, we imaged approximately one hundred cells of each strain concurrently (see **Figure 2.3 Supplement 2B** for a summary of cell fates in all experiments). For each cell, we imaged all three channels continuously, and performed similar normalizations as for the flow cytometry data (see SUPPLEMENTAL MATERIALS AND METHODS). We first imaged the HSS distance sensor strain (ectopic *ura4::dh*HSS^{3kb}). Our ability to observe cells that were initially fully de-repressed allowed us to trace "green" and "orange" repression kinetically. Consistent with linear heterochromatin spread outward of the *dh* nucleator, we find that "orange" repression is anticipated by repression at "green" (**Figure 2.3 Supplement 1**). While nucleation in this strain is not stable (likely due to "green" being adjacent to, rather than within *dh*), over time intervals where nucleation does persist, we observed dynamic fluctuations in the distal "orange" color without a fixed temporal pattern (**Figure 2.3 Supplement 2A**), which is not due to the repression state of "green" (**Figure 2.3 Supplement 2F**).
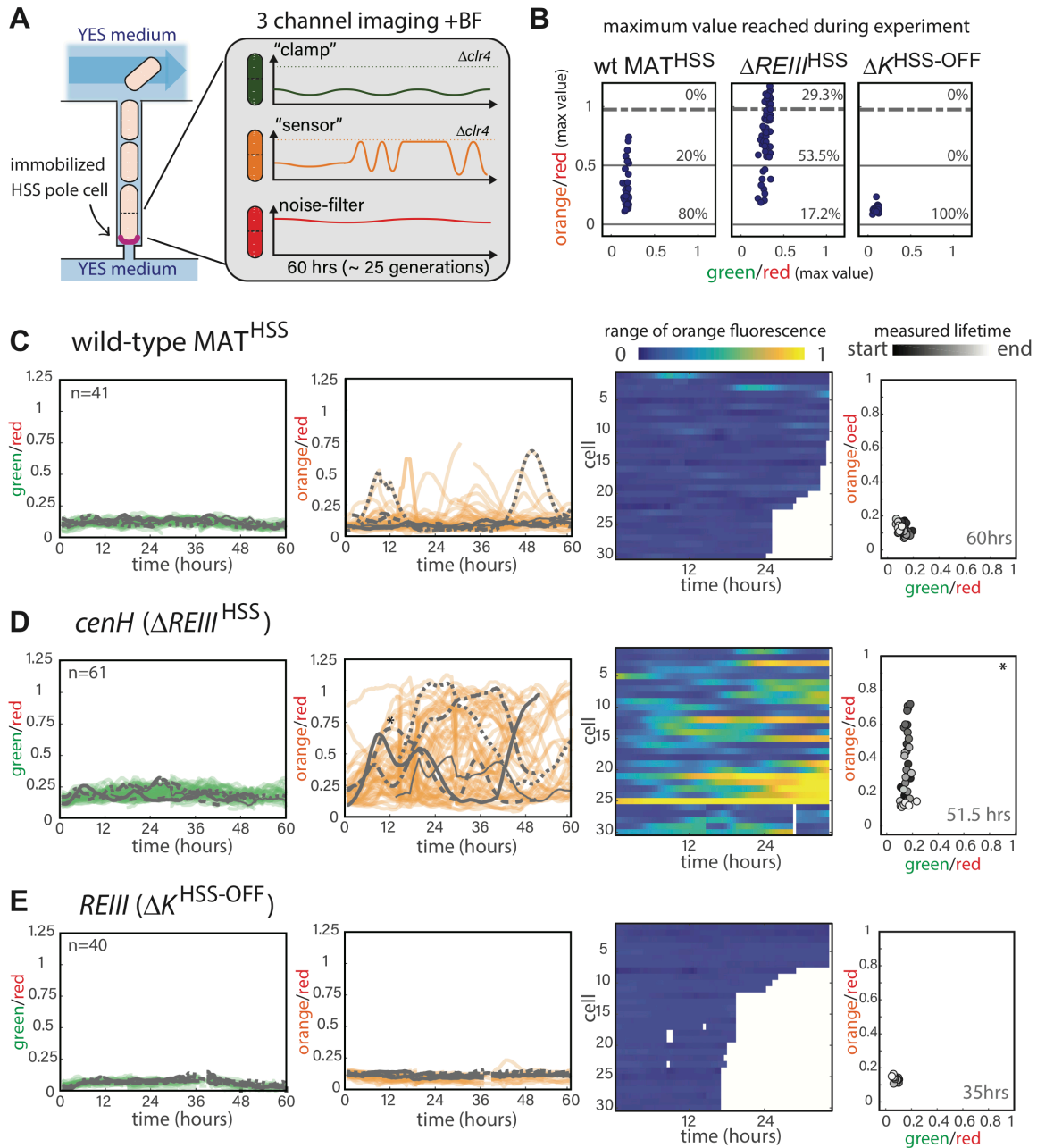
Figure 2.3: Single cell analysis of nucleation and spreading using a Fission Yeast Lifespan Micro-dissector (FYLM).

A. Overview of the FYLM-based heterochromatin spreading assay. The old-pole cell is trapped at the bottom of one of hundreds of wells in the FYLM microfluidic device and is continuously imaged in brightfield (to enable cell annotation), green, orange and red channels. Hypothetical example traces are shown. **B**. Maximum values attained by each nucleated cell for normalized "orange" plotted against normalized "green". Solid horizontal lines correspond to y=0 and y=0.5. Dashed line corresponds to an ON cutoff determined by mean less 3 standard deviations for each strain's matched Δclr4 strain. Percentage of cells between each line was calculated. **C.** FYLM analysis of wild type MAT^HSS cells. CELL TRACES: 60hrs of normalized "green" (left) and

"orange" (right) fluorescence in cells that maintained nucleation with the same 5 cells overlaid in different gray line styles in both plots. Gaps indicate loss of focus. HEATMAP: Up to 36 hours of normalized "orange" fluorescence for 30 cells that maintained nucleation is represented from blue (0) to yellow (1). X-Y FLUORESCENCE PLOT: for one representative sample cell, plot of normalized "green" and "orange" fluorescence across its measured lifetime (grayscale). **D.** FYLM analysis of *ΔREIII*^{HSS} cells as in C. The example cell in the X-Y dot plot is marked with an asterisk (*) on the orange traces **E.** FYLM analysis of *ΔK*$^{HSS-OFF}$ isolate, as in C., D. All cells were normalized to *Δclr4* (max, 1).

Next, we analyzed the MAT locus strains and selected cells that maintained nucleation for their entire measured lifespan (SUPPLEMENTAL MATERIALS AND METHODS). Under this constraint, the three strains exhibit vastly different behaviors (**Figure 2.3B**). WT MAT^{HSS} cells maintained "orange" repression for the majority of their measured lifespans (**Figure 2.3C**, **Figure 2.3 Supplement 2C**). However, we documented transient loss of "orange" silencing for 20% of the cells. (**Figure 2.3B** and **3C**). In contrast, while most cells stay similarly nucleated in *ΔREIII*^{HSS} (**Figure 2.3D, Figure 2.3 Supplement 2D**) 83% of the cells imaged experienced at least half-maximal "orange" de-repression at some time points (**Figure 2.3B**). For this strain, 30% of the cells transited through the fully ON state (**Figure 2.3B**, **Figure 2.3D, Figure 2.3 Supplement 2D**). In fact, cells sampled a wide range of values from OFF to fully ON, indicating that cells do not occupy ON or OFF states exclusively, but adopt intermediate values across time (**Figure 2.3D**). Importantly, *ΔREIII*^{HSS} cells, just as *ura4::dh*HSS^{3kb} cells, fluctuate in their "orange" values, indicating that spreading is unstable and adopts a random walk type behavior. To analyze *ΔK*^{HSS} cells, which exist predominantly in fully "green" and "orange" ON state (**Figure 2.2C**), we analyzed *ΔK*$^{HSS-OFF}$ cells (see above). *ΔK*$^{HSS-OFF}$ behaved markedly differently from *ΔREIII*^{HSS}: in all of the cells analyzed, "green" and "orange" reporters remained OFF throughout the time course (**Figure 2.3B**, **E**), up to 25 generations, revealing a fundamentally different dynamic behavior of *cenH*- or *REIII*- dependent heterochromatin. We note it remains possible that isolation of a *ΔK*$^{HSS-OFF}$ colony may bias our analysis against potentially more

frequent OFF-ON switching events in the primary mixed population derived from continuous propagation of the germinated spore (**Figure 2.2E**). However, since the mixed population resolves spontaneously into ON and OFF states once plated, and OFF cells behave similarly in either the mixed ΔK^{HSS} or $\Delta K^{HSS-OFF}$ isolated populations (compare **Figure 2.2E** and **Figure 2.5C**), we believe the stability of $\Delta K^{HSS-OFF}$ is intrinsic to the ΔK MAT locus.

Epigenetic stability at MAT is dependent on *REIII*.

To probe memory capacity (i.e., the ability of cells to retain information of an ancestral state established many generations prior) we compared cells containing an intact MAT locus to those lacking either ncRNA- or *REIII*-dependent heterochromatin. We established two ancestral states (**Figure 2.4A**); one with unperturbed heterochromatin, and a second treated with the HDAC inhibitor trichostatin A (TSA), known to fully disrupt the heterochromatin state ((Hall et al. 2002) and **Figure 2.4 Supplement 1**). Following production of the ancestral states, we grew cells either in rich media alone or in a TSA concentration gradient (0-50μM) for 25 generations and then measured the fraction of fully nucleated cells that effectively silence the "orange" spreading marker (**Figure 2.4A**). Cells exhibit memory if the fraction of the population with full spreading ("orange"^{OFF}) depends on the ancestral state, which would be indicated by separation of the unperturbed (light orange) and perturbed (red) lines. In contrast, no memory is indicated by convergence of the two lines (graphs in **Figure 2.4B-D**). We further measure a second parameter we term relative "resistance", which is defined as the TSA concentration at which the fraction of cells with "orange"^{OFF} declines to 50% of the no TSA pretreatment value. This value reports on the intrinsic sensitivity to perturbation of the locus formed by spreading.
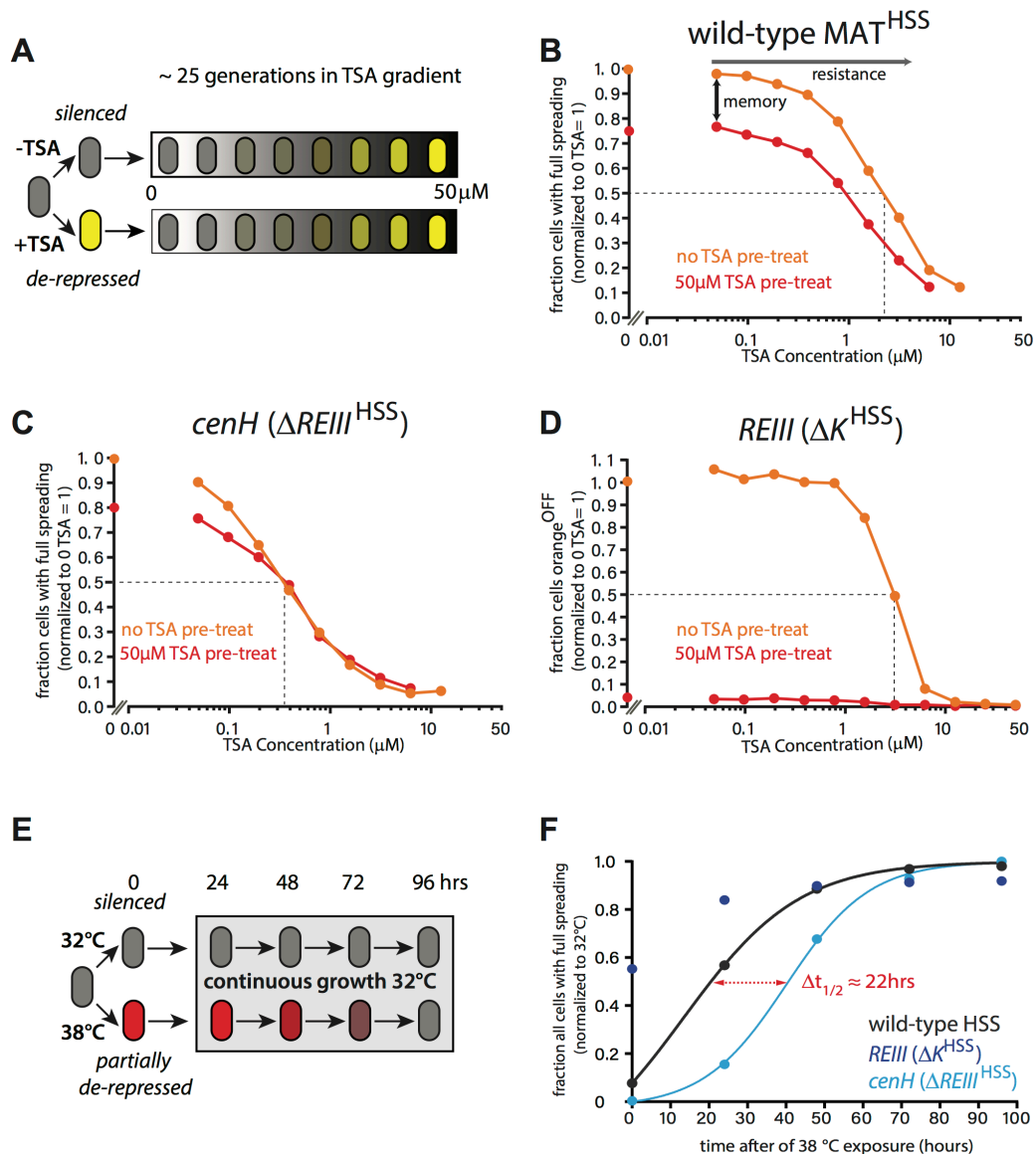
Figure 2.4: ncRNA-nucleated spreading exhibits weak memory and resistance in the absence of *REIII*.

A. Experimental schematic for memory and resistance measurements. Cells in log phase were treated with TSA (50 μM) for 10 generations to erase all heterochromatin (de-repressed, yellow) or kept untreated (repressed, gray). Both populations are then grown in a gradient of TSA concentration from 0 to 50 μM for 25 generations. **B.** The wild-type MAT locus exhibits memory in silencing "orange" throughout the TSA gradient. The fraction of "green"[OFF] cells that fully silence "orange" normalized to the no TSA pre-treatment, 0 μM TSA point are plotted for each TSA concentration. Red line: cell ancestrally TSA pre-treated; light orange line: cells without pre-treatment. **C.** Spreading from *cenH* exhibits weak memory and low resistance. Cell populations as above. **D.** ncRNA-independent spreading exhibits high resistance. The fraction of "orange"[OFF] for all cells is plotted, because in the TSA pre-treatment almost no "green"[OFF] cells can be detected. Dotted lines indicate the half-resistance points: TSA concentration at which 50% of non-pretreated cells fail to form heterochromatin at "orange". Memory is the

difference between orange and red lines. One of two full biological repeats of the experiment is shown. **E.** Experimental schematic for heat stress and recovery. Cells were grown at either 32°C or 38°C for 10 generations and strains subsequently grown continuously for 96 hours at 32°C. **F.** The fraction of cells with full spreading ("green"^{OFF} and "orange"^{OFF}) after 38°C exposure and recovery normalized to the fraction of cells with full spreading at 32°C for each strain is plotted over time. For wild-type MAT^{HSS} and $\Delta REIII^{HSS}$ strains, we fit a simple sigmoidal dose response curve and determined a $t_{1/2}$ value. The difference in $t_{1/2}$ values or $\Delta t_{1/2}$ is ~22hrs or ~9-10 generations.

As expected, wild-type MAT^{HSS} exhibited clear memory at 25 generations (**Figure 2.4B**), which was still weakly evident even at 35 generations (**Figure 2.4 Supplement 1C**). Among fully nucleated ("green"^{OFF}) cells, those that derived from untreated ancestral cells showed a greater fraction of silencing ("orange"^{OFF}) than those derived from treated cells throughout the entire TSA gradient, with a half-resistance point of ~2 μM (**Figure 2.4B**). Thus, wild-type MAT^{HSS} memory is robust in the face of perturbations of the heterochromatic state.

In sharp contrast, when spreading exclusively nucleates from *cenH* ($\Delta REIII^{HSS}$ strain), memory of silencing ("orange"^{OFF}) is significantly weaker. Memory collapsed beyond low TSA concentrations (> 0.2 μM TSA), with the red and light orange lines coinciding for much of the gradient. Even at 0 μM TSA, history dependence was erased at 35 generations (**Figure 2.4 Supplement 2C**). Interestingly, the half-resistance point was ~0.2 μM, 10-fold lower than that of wild-type MAT (**Figure 2.4C**). As *cenH*-nucleated spreading in *REIII*^{HSS} produces little memory capacity and lacks resistance, the memory capacity at MAT does not derive from ncRNA-nucleated spreading. These results are consistent with *REIII* being required for the memory behavior of WT-MAT.

The ΔK^{HSS} strain at face value had the widest separation in the behavior of the progeny of TSA pretreated and untreated cells. However, ascribing this behavior directly to memory is complicated by the fact that ΔK^{HSS} cells are no longer able to re-nucleate if they were ancestrally TSA treated, consistent with previous findings indicating that RNAi factors are required for

heterochromatin establishment at MAT (Hall et al. 2002). However, when examining resistance, i.e. the behavior of cells not ancestrally TSA pretreated, we observe that the *REIII* dependent ΔK^{HSS} strain has a half-resistance point of ~ 3 µM TSA (**Figure 2.4D**), similar to the intact locus. This indicates that the increased resistance of the wild-type over $\Delta REIII^{\mathrm{HSS}}$ is conferred by *REIII*. Together these results indicate that *REIII* is required for epigenetic stability at MAT.

REIII **imposes epigenetic behavior under environmental stress conditions.**

We next studied how *REIII* contributes to epigenetic stability in the context of a physiological perturbation, such as change in ambient temperature. Consistent with previous reports, we found that ncRNA-nucleated spreading is sensitive to continuous growth at high temperature, likely due to the cytosolic shuttling of RNAi-components ((Woolcock et al. 2012) and **Figure 2.4 Supplement 2A**). WT MAT behaved in a similarly sensitive manner. In contrast, heterochromatin in ΔK^{HSS} cells was highly resistant to elevated temperature (**Figure 2.4 Supplement 2A**).

We next probed the ability to remember the heterochromatin state after a transient exposure to elevated temperature, by exposing cells to 38˚C for 10 doublings, followed by return to growth at 32˚C (**Figure 2.4E**). As expected from our steady state experiments above, *REIII*-dependent heterochromatin (ΔK^{HSS} cells) is only minimally affected by the perturbation and regains full spreading rapidly (**Figure 2.4F, Figure 2.4 Supplement 2F**), whereas WT MAT and ncRNA-nucleated ($\Delta REIII^{\mathrm{HSS}}$) strains lose a significant amount of spreading (**Figure 2.4F, Figure 2.4 Supplement 2D,E**) and nucleation (Figure insets). Both strains regain nucleation at *cenH* rapidly (1 day after return to 32˚C; **Figure 2.4 Supplement 2B,C**). However, they are discrepant in their kinetics of restoration to the 32˚C extent of spreading, with WT MAT

recovering much more rapidly than the strain nucleated exclusively by ncRNA ($\Delta REIII^{HSS}$)

(**Figure 2.4F**). Indeed, plot fitting reveals a half-life ($t_{1/2,}$ time to reach 50% of initial state)

difference of ~22hrs, or ~9-10 generations between WT MAT and $\Delta REIII^{HSS}$ (**Figure 2.4F**).

Therefore, *REIII-* is required for efficient recovery to the fully repressed state after heat

perturbation. These data suggest that a central role of *REIII* is to ensure that epigenetic stability

at MAT is maintained in the face of environmental perturbations in the wild.

Stability of heterochromatin in the absence of *cenH* and *REIII trans*-acting factors.

To address dependence of the epigenetic maintenance of spreading on nucleation following

heterochromatin establishment, we examined the behavior of cells following the removal of

trans-acting factors required for the initial recruitment of nucleation factors such as Clr4,

Swi6/HP1 and HDACs. This experiment is similar to the induced removal of the *cis*-acting sites

in *S. cerevisiae* (Cheng and Gartenberg 2000). $\Delta REIII^{HSS}$ and $\Delta K^{HSS\text{-}OFF}$ isolate cells (see above,

derived from nonselective plating of ΔK^{HSS}) with established heterochromatin were crossed to

mutants disrupting recruitment of nucleation factors at each element (**Figure 2.5A**). To impair

REIII, we crossed the $\Delta K^{HSS\text{-}OFF}$ reporter strain to $\Delta pcr1$ (Noma et al., 2004). To impair ncRNA

nucleation, we crossed the $\Delta REIII^{HSS}$ reporter strain to *seb1-1,* a mutant allele of the Seb1 RNA

binding protein. Seb1 functions redundantly with the RNAi pathway in ncRNA nucleation,

including binding *cenH* transcripts, and the mutant allele *seb1-1* has defects in triggering

nucleation at *dh* and *dg* pericentromeric elements (Marina et al. 2013). We focus on Seb1, as

RNAi pathway mutants have little discernable effect on MAT when introduced after

establishment (our unpublished data and (Hall et al. 2002)), indicating a stronger role for Seb1.

Identifiable $\Delta REIII^{HSS}seb1-1$ and $\Delta K^{HSS\text{-}OFF}\Delta pcr1$ colonies were grown for flow cytometry

analysis immediately following mating and selection. The control cross mutant strains

$\Delta REIII^{HSS}\Delta pcr1$ and $\Delta K^{HSS\text{-}OFF}seb1\text{-}1\Delta dcr1$ (loss of all ncRNA-nucleation (Marina et al. 2013))

allowed us to assess any effects the *trans*-factor may have even in the absence of its cognate site
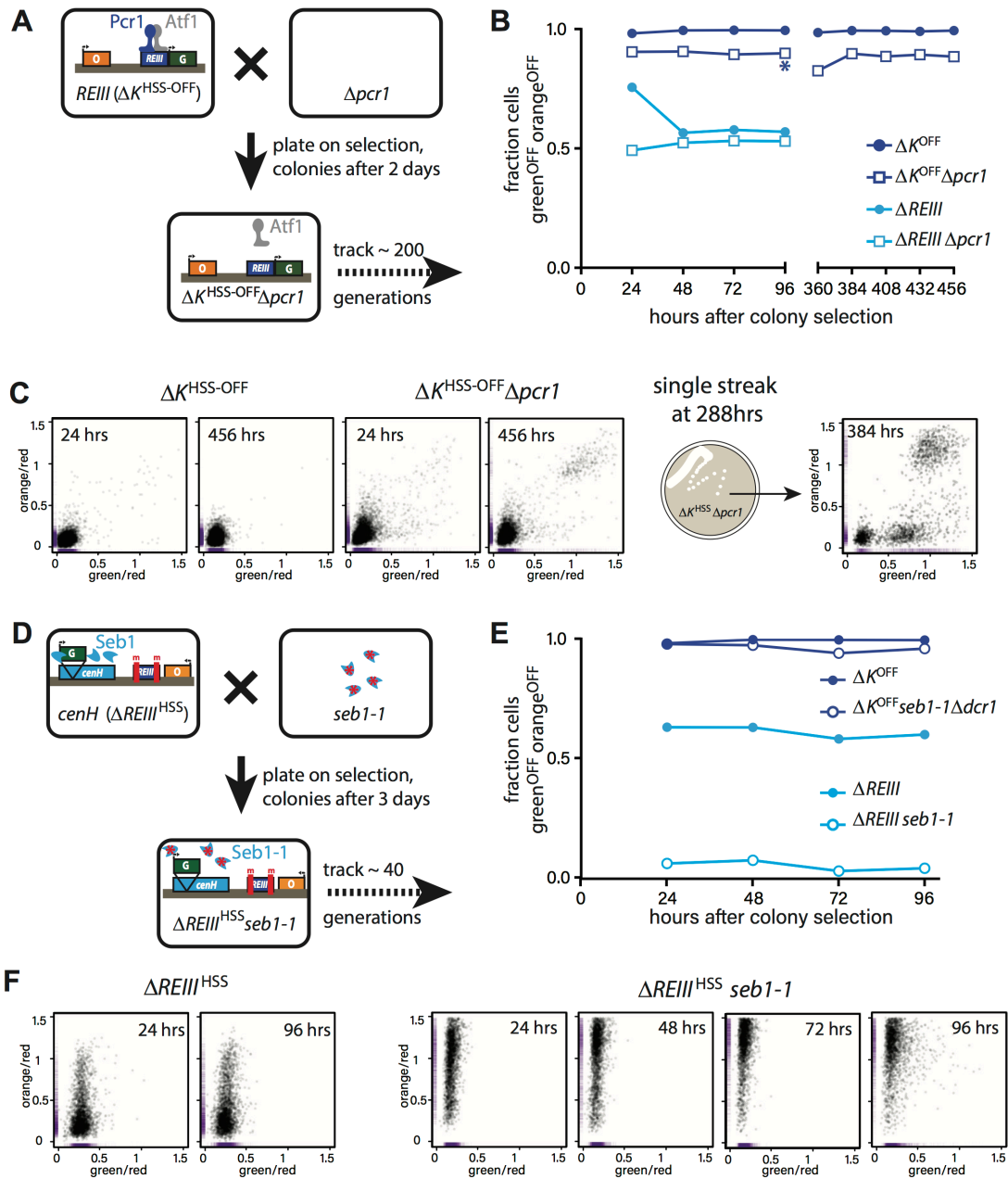
of action.

Figure 2.5: Differential inheritance of ncRNA- dependent and independent spreading in the absence of nucleation factors.

A. Scheme for removal of Pcr1 (*REIII* binding factor) in the *ΔK*^{HSS} strain OFF isolate (*ΔK*^{HSS-OFF}). Progeny of the cross was selected for *ΔK*^{HSS-OFF}*Δpcr1* genotype and identifiable colonies immediately grown for cytometry, and passaged for 456hrs. **B.** Stable inheritance of repression in *ΔK*^{HSS-OFF}*Δpcr1*. *ΔK*^{HSS-OFF}*Δpcr1* or *ΔK*^{HSS-OFF} cells (dark blue lines) where analyzed by flow cytometry over consecutive days, the break indicating passaging without analysis. *Δpcr1* had no significant effect on *ΔREIII*^{HSS} (light blue lines). **C.** LEFT: scatter plots with partial point transparency of *ΔK*^{HSS-OFF} or *ΔK*^{HSS-OFF}*Δpcr1* early and late in the time course. RIGHT: In the middle of the time course *ΔK*^{HSS-OFF}*Δpcr1* were struck for single colonies. The scatter plots for one of the isolates is shown. **D.** Scheme for removal of functional Seb1 in *ΔREIII*^{HSS} strain. Selection and growth as in **A.**, total passaging time 96hrs. **E.** Weak inheritance of repression in *ΔREIII*^{HSS}*seb1-1* (light blue lines). Analysis as above, total time course 96hrs. Removal of both Seb1 and RNAi pathways (*ΔK*^{HSS-OFF}*seb1-1Δpcr1*) does not affect maintenance of silencing (dark blue lines). **F.** Scatter plots of *ΔREIII*^{HSS} at 24 and 96hrs and through the entire time course for *ΔREIII*^{HSS}*seb1-1.* In these scatter plots, X and Y values of each cell are represented by purple dashes along the corresponding axis.

Strikingly, most $\Delta K^{\text{HSS-OFF}}\Delta pcr1$ cells remains robustly repressed over 456 hours, around 200 generations (**Figure 2.5B**). However, removal of Pcr1 does have a small discernable effect, as the $\Delta K^{\text{HSS-OFF}}\Delta pcr1$ strain showed a small population of cells not completely in the OFF state compared to the $\Delta K^{\text{HSS-OFF}}$ parent (**Figure 2.5C LEFT**). Further, by ~ 400 hours we detected a small fully ON population absent in the parent. This behavior is broadly consistent with the reported stability of intact ΔK^{OFF} (switch rate of ~10^{-4} generation, (Thon and Friis 1997; Grewal et al. 1998)), even though our assay appears to show even smaller ON populations. Very small ON populations are more apparent in a growth selection-based assay as only the targeted population survives, as opposed to our assay, which captures all cells. We note a formal possibility remains that selection of OFF colonies yields higher apparent stability. To get a closer view of the behavior of individual isolates from the population, after 288 hours of continuous passage, we streaked for single colonies and measured the resulting populations. While 5/6 isolates behaved like the broader population, we found 1/6 isolates that experienced more severe breakdown in its heterochromatic state (**Figure 2.5 C RIGHT**). In this isolate heterochromatin collapsed in a manner not ordered with respect to *REIII* proximity and exhibited "green"^{ON}/"orange"^{OFF} cells. In contrast, $\Delta REIII^{\text{HSS}}seb1-1$ lost most spreading at the first

measurement point (24hrs, **Figure 2.5E**) with progressively increasing de-repression of

"orange", but also some loss of "green", over the next 72 hrs (**Figure 2.5F**). This suggests that

the epigenetic inheritance *cenH-* spreading requires continuous nucleation, at least via the Seb1

pathway, consistent with the behavior at synthetic nucleators (Audergon et al. 2015; Ragunathan

et al. 2015).

***REIII-,* but not *cenH-* dependent heterochromatin suppresses histone turnover.**

It is known that *REIII* recruits the HDAC Clr3 (Yamada et al. 2005), which was later shown to

repress the turnover of histones (Aygun et al. 2013). This suggested the intriguing possibility that

unstable epigenetic inheritance in the absence of *REIII* is linked to elevated histone turnover. To

test this idea, we adopted the Recombination Induced Tag Exchange (RITE) system

(Verzijlbergen et al. 2010) to assay replication-independent turnover of H3 in $\Delta REIII^{HSS}$ and

ΔK^{HSS} strains (**Figure 2.6A**). Tag switching (T7 for HA tag) in log phase growth was induced by

administering β-estradiol (βER) concurrently with stalling replication with 15mM hydroxyurea

(HU) for 4 hours, during which time cells remain in early S phase (**Figure 2.6 Supplement**). We

compared the incorporation of T7 at 4hrs vs. 0hrs between $\Delta REIII^{HSS}$, $\Delta K^{HSS-OFF}$ and ΔK^{HSS-ON}

strains. First, we examined two euchromatic genes, *pyk1* on chromosome 1, and *mtd1,* which is

just outside the MAT locus. H3 turnover at these at regions does not differ between the strains

(**Figure 2.6B**) and is highest in in the strongly expressed *pyk1* gene. We next examined sites in

the MAT locus that are shared in sequence and genomic position between $\Delta REIII^{HSS}$ and ΔK^{HSS}

(probes indicated in diagram, **Figure 2.6B**). We note this includes also *REIII*, since in $\Delta REIII^{HSS}$

the sequence differs between strains only by 14bp, i.e. the two Atf1/Pcr1 binding sites. In

contrast to euchromatic loci, we observed that $\Delta K^{HSS-OFF}$ experiences very low or no histone

turnover at MAT targets by 4hrs HU compared to $\Delta K^{\text{HSS-ON}}$ and $\Delta REIII^{\text{HSS}}$, which experienced

levels of H3 turnover more consistent with our euchromatic controls. This in not unexpected for

$\Delta K^{\text{HSS-ON}}$, as is it effectively not heterochromatic (**Figure 2.2E**), and is consistent with previous

results (Aygun et al. 2013). However, the observation that $\Delta REIII^{\text{HSS}}$ displays H3 exchange at

levels similar to $\Delta K^{\text{HSS-ON}}$ and euchromatin suggests that it is memory, rather than

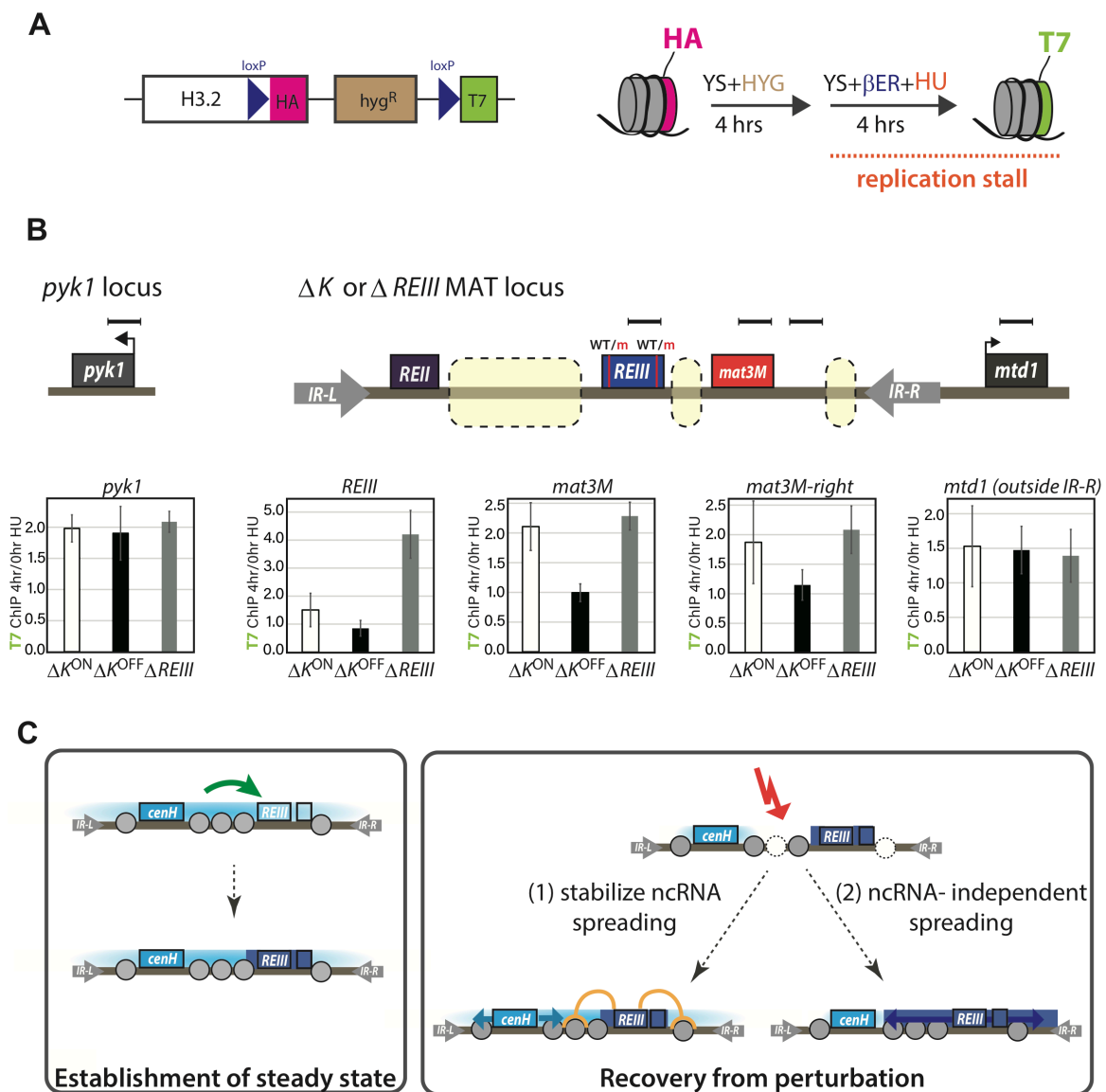heterochromatin formation itself, that requires repressed histone turnover.

Figure 2.6: Histone turnover correlates with epigenetic stability in ncRNA-dependent and REIII-dependent heterochromatin.

A. LEFT: Overview of the RITE system for histone 3.2. Cre recombinase allows tag exchange from HA to T7. RIGHT: experimental scheme for detecting replication-independent H3 turnover. Cells were grown to log phase and then grown for 4hrs in the presence of β-estradiol and 15mM hydroxyurea. **B.** Enrichment for H3-T7 at indicated loci in ΔK^{HSS-ON}, $\Delta K^{HSS-OFF}$ or $\Delta REIII^{HSS}$ strain. TOP: Location of amplicons for T7-ChIP indicated by bars. Dashed boxes in MAT indicated regions of genomic difference between ΔK^{HSS} and $\Delta REIII^{HSS}$. WT and m for *REIII* indicate presence or deletion of Atf1/Pcr1 binding sites respectively. BOTTOM: Enrichment of T7 tag by ChIP at 4hrs in HU over 0hrs for indicated strains. 1 indicates no enrichment over 0 hrs. Error bars indicate standard deviation of technical replicates. **C.** Model for collaboration of *cenH* and *REIII* in establishing and maintaining the high-fidelity MAT locus. (LEFT) During initial establishment, *cenH* heterochromatin raises the nucleation frequency at *REIII* (green arrow). A box right of *REIII* represents a putative additional nucleation element. (RIGHT) Labile *cenH*-nucleated spreading is disrupted, in part by de-stabilized nucleosomes, in an environmental perturbation or a stochastic event. *REIII* promotes reestablishment of the initial state by repressing histone turnover, limiting nucleosome loss (orange) and thus aiding spreading from *cenH* (light blue arrows, (1)), or promoting heterochromatin spreading from surrounding elements (dark blue arrows, (2)).

DISCUSSION

The patterning of the genome into regions of activity and inactivity underlies the formation of cellular identity. In many systems, heterochromatin spreading is the dominant contributor to the pattern (Schultz 1939; Schwartz et al. 2006; Wen et al. 2009). Maintaining identity requires the capacity to "remember" the positional extent of heterochromatic spreading. Yet, how precise epigenetic memory is linked to the intrinsic properties of the spreading reaction itself has remained opaque. In this work, we were able to directly measure the heterochromatin spreading reaction in single *S. pombe* cells, separate from DNA-directed events at nucleation elements, and probe its behaviors and memory characteristics. The central principle that emerges from this work is that heterochromatin spreading in fission yeast, driven predominantly by ncRNA elements, is epigenetically unstable and requires stabilization by accessory elements for high fidelity epigenetic inheritance. At the MAT locus, which carries cell identity information, a separate type of heterochromatin, independent from nc-RNA elements and dependent on the *REIII* element, safeguards epigenetic propagation by repressing histone turnover.

ncRNA-triggered spreading is epigenetically unstable and labile in the face of perturbations.

The dominant form of heterochromatin in *S. pombe,* triggered by ncRNA nucleators, leads to stochastic spreading of both silencing and H3K9 methylation that only occurs in some cells, and forms intermediate states (**Figure 2.1, Figure 2.1 Supplement E, Figure 2.2C and 3D**). This is consistent with position effect variegation in genetically disrupted systems (Nimmo et al. 1994; Elgin and Reuter 2013). Additionally, the linear, distance-dependent behavior we observe (**Figure 2.1C**) is reminiscent of the continuous spreading model in *S. cerevisiae* telomeres (Renauld et al. 1993; Talbert and Henikoff 2006). This behavior of ncRNA spreading is not due to weak nucleation, as repressive histone marks accumulate to the same high extent at *cenH* in wild-type and *ΔREIII* and at Atf1/Pcr1 proximal region in *ΔK* cells. (**Figure 2.1E, 2D and E**). In a key result, we find ncRNA-triggered spreading to be epigenetically unstable. This is evidenced by highly dynamic behaviors over time and across generations, little discernable memory, and low resistance to chemical or environmental perturbations (**Figures 3-5**). Those behaviors are not necessarily predicted by the stochastic induction of spreading, given that PEV in flies results in clonally inherited patches (Elgin and Reuter 2013). This result opens the question of how high fidelity can be achieved with ncRNA nucleators at loci that carry critical cell type specification information. The most likely cause for this instability is elevated and near-euchromatic levels of histone turnover (**Figure 2.6B**). This implies that while elevated histone turnover is compatible with heterochromatin formation *per se,* it is incompatible with epigenetic memory.

In contrast to the behavior of *ΔREIII*, *ΔK* cells, dependent on *REIII* for heterochromatin formation (**Figure 2.2 Supplement 2A&B**), do not display stochasticity in spreading (**Figure 2.2E and 1.3E**), and instead repress MAT uniformly across nucleated cells in the population

(**Figure 2.2E**). Under environmental perturbation, *ΔK* heterochromatin is extraordinarily

resistant (**Figure 2.4F** and **Figure 2.4 Supplement 2A,F**) and capable of high memory retention,

even in the absence of the *REIII*-targeted Pcr1 protein, which attracts HDACs and Clr4/Swi6 (Jia

et al. 2004; Kim et al. 2004) (**Figure 2.5B, C**). This is consistent with previously documented

bistable behaviors ascribed to the overall locus (Grewal and Klar 1996; Dodd et al. 2007). It

however remains to be resolved whether heterochromatin in *ΔK* cells is nucleated by *REIII* and

spreads outwards, or is nucleated at multiple sites, yielding apparent uniform heterochromatin

formation. A *REIII* nucleated spreading model is favored by results presented here and by others

(Jia et al. 2004; Wang and Moazed 2017) that demonstrate that Atf1 and Pcr1 proteins or their

binding sites are absolutely required for established of heterochromatin in *ΔK* cells, yet not for

ΔREIII cells. However, unlike for *cenH,* where sufficiency has been clearly shown (Hall et al.

2002), we and others (Wang and Moazed 2017) do not document significant heterochromatin

formation by *REIII* when it is placed ectopically (**Figure 2.2 Supplement 1D**). Thus, it remains

possible that the uniform heterochromatin formation in *ΔK* is the result of cooperation between

different yet-to-be identified *cis*-acting elements, or a special property of *REIII*-driven spreading.

Single site driven spreading of this "all or none" type could be the result of looping, invoked for

the polycomb system (Bantignies and Cavalli 2011), and predicted to improve spreading

efficiency and memory in fission yeast (Erdel and Greene 2016), or a unique molecular signature

at *REIII.* For example, *REIII* recruits the HDAC Clr3 (Yamada et al. 2005), which promotes

accumulation of the H3K9 trimethyl state, required for efficient spreading by Clr4 (Zhang et al.

2008a; Al-Sady et al. 2013; Jih et al. 2017).

REIII **stabilizes heterochromatin spreading by repressing histone turnover.**

Regulation of histone turnover has been linked to epigenetic memory in fission yeast (Taneja et al. 2017) and has been previously shown to be low at wild-type MAT (Aygun et al. 2013). Hence, the high histone turnover we observe in *ΔREIII* cells results from unaided *cenH*-spreading. *REIII* recruits the HDAC Clr3 (Yamada et al. 2005), which represses histone turnover (Aygun et al. 2013). Our finding that the ΔK^{OFF} allele features very low histone turnover (**Figure 2.6B**), similar to the wild-type locus (Aygun et al. 2013), is thus consistent with *REIII* acting to repress histone turnover, when in a heterochromatic state. The extraordinary memory of repression we observe in ΔK^{OFF} likely is explained by this repressed turnover, although we should note it is possible that isolation of ΔK^{OFF} alleles, while consistent with the literature (for example (Grewal and Klar 1996; Thon and Friis 1997)), could bias the population to enhanced inheritance of repression. We speculate that reduced turnover increases retention of H3K9me3 nucleosomes, promoting methylation across nucleosomes by Clr4 via its H3K9me-dependent "read-write" functionality (Zhang et al. 2008a; Al-Sady et al. 2013; Ragunathan et al. 2015; Jih et al. 2017), thus facilitating re-establishment in the next generation. H3K9me3 is also directly promoted by Clr3, which is recruited to *REIII* (Yamada et al. 2005), further favoring reestablishment of methylation.

Collaboration of ncRNA dependent and independent mechanisms in the maintenance of MAT heterochromatin.

Repression of histone turnover and resulting epigenetic stability in *ΔK* strains requires cells to first nucleate and adopt a heterochromatic state (ΔK^{OFF}, **Figure 2.6B**). However, since *ΔK* cells only nucleate infrequently (**Figure 2.2E**), how is *REIII* able to stabilize heterochromatin in most

wild-type MAT cells (**Figure 2.2B, 3B,C** and **4B**)? The independent action of *cenH* and *REIII*

elements cannot account for this behavior, hence they must collaborate. We propose that in the

context of wild-type MAT, *cenH* stimulates *REIII* nucleation (model, **Figure 2.6C**). Recent

findings indicate that Atf1/Pcr1 are present at *REIII* even in non-silenced *ΔK*-type cells (Wang

and Moazed 2017). We speculate that since Atf1/Pcr1 recruits silencing factors such as Clr4 and

HDACs (Jia et al. 2004; Kim et al. 2004; Yamada et al. 2005) heterochromatin originating from

cenH might stabilize this recruitment. This hypothesis is supported by our observation for

nucleation during TSA recovery. Whereas *ΔK*[HSS] cells very rarely renucleate (**Figure 2.4D)**,

REIII at the intact MAT locus must be active in most cells, as the heterochromatin reformed after

erasure has much higher resistance to perturbation than that nucleated from *cenH* alone (red lines

in **Figure 2.4B** vs **C**).

Activated *REIII* in turn stabilizes the MAT locus most prominently when the

heterochromatin state is perturbed. We infer this from the difference between the initial

challenge and recovery from growth at high temperatures. When initially challenged,

heterochromatin spreading at the wild-type MAT locus resembles that of ncRNA-nucleated

heterochromatin (**Figure 2.4 Supplement 2A**), suggesting that *REIII* or other nearby elements

play a minor role under normal circumstances at MAT. However, the heat recovery experiment

suggests that changes in the *REIII*–dependent heterochromatin stabilization or assembly, not

cenH nucleation (**Figure 2.4 Supplement 2B** vs. **C**), takes on a major role in the accelerated

recovery of the collapsed heterochromatin (**Figure 2.4F**). Thus, *REIII* is required under

perturbation conditions to protect or quickly re-establish the heterochromatin state (**Figure 2.4B,

F** and **Figure 2.6C** (model)). The relatively transient distal de-repression events experienced by

wild-type MAT cells, which are much more pronounced in *ΔREIII* cells (**Figure 2.3C vs. D**),

further points to *REIII* acting after stochastic loss of *cenH* spreading in steady state. It is possible that *REIII* does so by stabilizing existing heterochromatin via repression of histone turnover when the loss of these structures is sensed, or alternatively, that *REIII*-dependent structures expand or "fill-in" the void left by collapse of *cenH*-spreading. In either case, we propose that *REIII* acts as a failsafe, ensuring the integrity, and ultimately epigenetic memory, of heterochromatin at MAT through perturbations.

In summary, we propose a model whereby the division of labor between *cenH* and *REIII* is uniquely suited for a cell type specification locus such as MAT, which requires silencing that is both robust and intergenerationally stable. ncRNA-nucleation is extremely robust but intrinsically too labile and stochastic to reliably control the cell type specification locus, thus requires support from an accessory element. The need for reliable control of cell type specification loci is shared in more complex systems. The nature of accessory elements equivalent to *REIII* and how they act in these cases remains to be determined.

generated in the UCSF Parnassus Flow Cytometry Core which is supported by the Diabetes Research Center (DRC) grant, NIH P30 DK063720

MATERIALS AND METHODS

Strain Construction:

Plasmid/construct construction:

Plasmids to generate constructs for genomic integration were generated by standard methods including Gibson assembly and *in vivo* recombination. *S. pombe* transformants were selected directly on dropout media for auxotrophic markers or onto rich media (YES) for 24 hours followed by selective media YES+ G418, YES+hygromycin or YES+nourseothricin).

Ura4 replacement method:

To avoid interference of selection cassettes with heterochromatin function in our HSS, we produced "scarless" genomic integrations, lacking selection markers. To do so we marked the insertion site first with a *ura4* cassette by genomic integration and then replaced this cassette either with an XFP cassette or altered genomic sequence for site mutations. *ura4* replacements were isolated by 5-FOA counter-selection and confirmed by genomic PCR. This method was used to generate the atf/creb site deletions and sequence insertions. *ura4* was targeted to the region between Mat3M and *cenH*, specifically including the two 7 base atf/creb binding sites (s1 and s2, and (Wang and Moazed 2017)). The entire *ura4* cassette was then replaced with a construct containing the two 7 base pair deletions of s1 and s2 or a deletion of s1 with additional 700bp of sequence from the *sib1* open reading frame. Point mutations and restoration of the pre-substitution locus was confirmed by PCR and sequencing.

Flow Cytometry and FACS sorting:

For standard flow cytometry experiments, cells were grown overnight in rich media (YES) and then diluted in the morning to OD=0.1 in minimal media plus supplements (EMM complete) and grown 4-6 hours before analysis by flow cytometry. Flow cytometry was performed using Fortessa X20 Dual or LSRII instruments (Becton Dickinson, San Jose, CA, U.S.A). Samples sizes ranged from ~2,000-100,000 cells depending on strain growth. Compensation was performed using cells expressing no XFPs and single-color controls expressing 1 XFP each. Compensated data was used for all downstream analysis. Fluorescence was detected for each color as described (Al-Sady et al., 2016).

For FACS sorting experiments, cells were grown overnight from OD=0.025 in YES and then in the morning concentrated into a smaller volume to achieve a flow rate of ~5000 events/second on the cytometer. Sorting was performed using either Aria2 or Aria3u machines (Becton Dickinson). Prior to sorting cells were strained through a 35-40 μm mesh (Corning) to reduce clogs. Sorting criteria included a gate for size (forward (FSC) and side (SSC) scatter), removal of doublets, a gate for "green"[OFF] ("green" signal within the range of an unstained control) and then gated into Low, Intermediate, High "orange" signal defined by the following: Low encompassed signal overlapping that of an unstained control and High encompassed signal overlapping that of the *Δclr4* no heterochromatin control strain PAS355. Intermediate gate was set in between Low and High with about 100 fluorescence units of a gap (representing ~2% of the full range of captured fluorescence) to ensure reliable separation. The entire range of fluorescence detected was ~2.5 orders of magnitude. At least 8×10^6 cells were collected for each population for Chromatin Immunoprecipitation and 2×10^6 cells for RT-qPCR. Immediately after

sorting, the final populations were subjected to the appropriate treatment for either Chromatin Immunoprecipitation or RT-qPCR.

Sytox Green Staining and Cell Cycle Analysis:

Cell cycle analyses were performed essentially as described (Knutsen et al. 2011). Briefly, cells were fixed with 70% ethanol, washed with 20 mM EDTA pH 8.0, and treated with RNaseA for 3 hours at 37°C. Immediately before analysis by flow cytometry, 2 μM Sytox Green (Invitrogen) in 20 mM EDTA pH 8.0 was used to resuspend pelleted cells. Cells were excited with a 488 nm laser and Sytox Green signal was detected with a 505 nm longpass filter and a 530/30 bandpass filter. Cell cycle analysis was performed in the FlowJo Software (Tree Star Inc, Ashland, Oregon, U.S.A.) The identification of cell populations and fraction of cells in each cell cycle phase (G2, S, and G1+M) were determined as described (Knutsen et al. 2011).

Trichostatin A (TSA) gradient experiment:

Cells were taken from fresh plates, and then grown overnight with shaking (Elmi) in 96-well plates containing 150 μL YES (Day -1). The next day (Day 0), cells were diluted into YES and measured by cytometry. At the end of Day 0, cells were passaged into YES+ DMSO (0μM TSA) or YES+ 50 μM TSA overnight. The next day (Day 1), cells were diluted and grown briefly into the same pretreatment conditions and the 50 μM TSA pre-treated cells were checked for complete de-repression by flow cytometry. Complete de-repression was defined as a qualitative overlap of WT and *Δclr4* profiles, with no evidence of repression. Both 0μM and 50μM TSA pretreated cells were then diluted into a gradient of TSA of eleven two-fold dilutions from 50

μM along with a twelfth 0 μM (DMSO) point. Cells were measured after ~6hrs and then passaged into the same TSA gradient conditions to continue growth.

The next day (Day 2) cells were diluted from overnight growth into the same gradient as above, measured ~ 6hrs later by flow cytometry and passaged into the same gradient again overnight. The same protocol was followed for Days 3 and 4. The full experiment was performed twice at different times (biological replicate). Given the lengthy continuous growth, contamination was occasionally observed in <1% of wells. The replicate shown was chosen based on lacking contamination.

Heat recovery experiment:

Cells were taken from fresh plates, and then grown overnight with shaking (Elmi) at either 32°C or 38°C (Day-1) in 96-well plates containing 200 μL YES medium per well. In the morning, cells were diluted into 200μL YES and grown ~6hrs at the same temperature before measurement by flow cytometry (Day 0). At the end of Day 0 all cells were all diluted again into YES and grown at 32°C. The next day (Day 1) cells were diluted from overnight growth into YES at 32°C, measured ~ 6hrs later by flow cytometry and passaged into the same temperature overnight. The same protocol was followed for Days 2, 3, and 4.

Nucleation factor removal experiment:

HSS strains were crossed to parent strains lacking functional nucleation factors for *REIII* (*Δpcr1*) or *cenH* (*Δdcr1 seb1-1*). Cross progeny were identified via a random spore approach by growth on selective media two or three days after plating. Absence of *pcr1* or *dcr1* open reading frames was confirmed by PCR. Presence of *seb1-1* allele was confirmed by sequencing. Single colonies

were grown in 96-well plates at 32°C containing 200µL YES medium per well. In the morning, cells were diluted into 200µL EMM and grown ~6hrs at the same temperature before measurement by flow cytometry. Cells were again diluted into 200 µL YES for overnight growth at 32°C and grown and measured similarly the subsequent days. For *Δdcr1* and/or *seb1-1* strains and their controls, this was continued for four days. For *Δpcr1* strains and their controls this was continued for five days then resulting cells were plated onto selective media and allowed to grow 48 hours at 32°C. Patches were then passaged in bulk on selective plates every 36-48 hours for 7 additional days. On the 6th day the passaged *ΔK*^{HSS}*Δpcr1* cells were additionally struck for singles. On the 8th day, patches of passaged cells and 6 single colonies of *ΔK*^{HSS}*Δpcr1* cells were grown in 96-well plates as above and measured by flow cytometry for five additional days.

Nucleosome turnover assay:

Recombination Induced Tag Exchange (RITE) parent strain (HU2549) was crossed into HSS reporter strains. Resulting isolates were verified by growth on selective media. The *cdc-25ts* allele was crossed out. RITE was performed essentially as described (Audergon et al. 2015; Svensson et al. 2015) with the following exceptions. Given the labile nature of heterochromatin at elevated temperatures, replication stalling was performed with hydroxyurea as published (Aygun et al. 2013). Cells were grown to saturation overnight in YES supplemented with Hygromycin. In the morning cells were diluted to OD=0.1 in 50mL YES+Hygromycin and grown for 4 hours at 30°C, 225rpm. After 4 hours of growth 13mL of cells were pelleted and processed for ChIP as the 0 hour time point. The remaining cells were washed twice in media devoid of Hygromycin and finally resuspended in YES supplemented with 15mM Hydroxyurea

(HU) and 1.5μM β-Estradiol (ER) and incubated for 4 additional hours at 30°C, 225rpm. After 4 hours incubation with HU and ER, 10mL of cells were pelleted and processed for ChIP.

Chromatin Immunoprecipitation (ChIP) and quantification:

We found that sonication of a small number of cells such as can be collected by FACS leads to a marked increase in background signal from negative control regions that was absent when ChIP was performed with larger log phase cultures ($>50 \times 10^6$ cells). To address this, ChIP in Fig 1E was performed on each of the FACS sorted populations with the addition of 42×10^6 formaldehyde fixed cells of *S. cerevisiae* W303 strain as a carrier. ChIP in Fig 2D was performed with 15×10^6 cells of each fission yeast strain and 50×10^6 additional W303. ChIPs for Fig 2E and Fig 2 Supplement 2B were performed with 80×10^6 cells and no added W303. ChIPs for Figure 2.6B were performed with no added W303. ChIP was additionally performed on a sample of W303 alone, which only produced signal equivalent to background. *S.pombe* ChIP samples and W303 cells were fixed and pre-processed for ChIP separately, then mixed together immediately prior to lysis. Cells were cross-linked and lysates prepared for ChIP as described (Canzio et al. 2011) with the following exceptions: After lysis, the chromatin fraction was resuspended in 350μL lysis buffer and sonication performed using a Diagenode Bioruptor Pico machine at 4°C, with 16-20 rounds of 30 seconds ON, 30 seconds rest. ChIP was essentially as described, with the total lysate split into 2-6 equal volumes (after ~8% set aside as input fraction) and ChIP performed in 600-800 μL per sample. Two or three technical replicates were performed across experiments. 1 μL of each of the following antibodies was added per ChIP replicate: anti-H3K9me2 (Abcam ab1220); anti-H3K4me3 (Active Motif 39159); anti-H3K9me3 (Millipore 07-442); anti-T7 (Novagen 69522-3). ChIP samples were agitated on a Nutator overnight at 4°C.

Immune complexes were collected for 3 hrs with 15-20 μL washed protein A Dynabead slurry (Invitrogen). Washing and downstream processing steps were essentially as described, except "wash buffer" wash was performed once. Samples were purified using a Machery-Nagel PCR purification kit and NTB buffer for SDS containing samples. DNAs were quantified by RT-qPCR (see below). Enrichments were calculated as follows: For Fig1E, 2D, 2E IP/input values for amplicons of interest were calculated and normalized to the IP/Input values for positive controls for each antibody, *dh* for H3K9me2 and H3K9me3 and the actin promoter for H3K4me3. For Fig 2 Supplement 2B, ChIP signal was normalized to signal from a matched background *Δclr4* strain. For Figure 2.6B IP/input values for the 4h time points were normalized to the IP/input values for the 0h time point.

RNA Extraction and mRNA quantification:

After sorting, samples were spun at 5000xg, supernatant decanted, and pellets flash frozen in liquid nitrogen and stored at -80°C. For the *Δclr4* strain PAS335, cells were grown into log phase and then cell pellets were isolated in the same fashion. Total RNA was extracted in technical duplicates from the same cell pellets using the "MasterPure- Yeast RNA Purification Kit" (Epicentre), including a 30-minute DNAseI treatment step post-RNA isolation. Reverse Transcription was performed with SuperScript III RT (Invitrogen), using the supplied protocol and 1.5-2μg of RNA and an oligo dT primer. Following cDNA synthesis, the reaction was treated with RNAse H (New England Biolabs). cDNA samples were quantified by RT-qPCR. For each sorted sample mKO2 cDNA values were normalized to actin and then divided by the max value calculated similarly from PAS355 (*Δclr4*).

RT-qPCR:

Real time quantitative PCR was performed using a BioRad CFX-384 machine. 15μL reactions were prepared, each containing 7.5μL of Applied Biosystems SYBR Select Master Mix, 4.5μL 3.3M betaine, 1.2μL of 2.5μM oligo mix, 0.8μL water, and 1μL template. The thermocycler protocol was: 2min at 50°C then 2min at 95°C followed by 40 cycles of 15sec at 95°C and then 1min at 60°C followed by a plate read. Lastly a melt curve was generated. Standards were generated with 5-fold dilutions of genomic DNA containing templates for all PCR products.

Single-cell Microscopy:

Single cells of strains PAS 387, 389, 391 and 244 (see Table 2.1; E2Crimson under *act1* promoter) were captured in microfluidic devices as described (Spivey et al. 2017). Multi-channel fission yeast lifespan microdissectors (multFYLM) contained six independent devices (channels), each of which is capable of capturing up to 392 cells (Jones et al. 2018). In brief, the devices were cast in polydimethylsiloxane (PDMS, Sylgard 184, Dow Corning) using conventional soft lithography methods. Master structures were fabricated from P-doped silicon wafers (ID#452, University Wafers) and SU-8 photoresists 3005 and 2010 (Microchem, Westborough, MA). MultFYLMs were cleaned and adhered to glass coverslips (48 x 65 mm #1, Gold Seal), and then connected to syringes (60 mL, Becton-Dickson) containing YES 225 liquid media (Sunrise Science) via PFA tubing and microfluidic fittings (IDEX Health and Science). The multFYLM was maintained at 30˚C in a custom staged-mounted environmental chamber on an inverted microscope (Eclipse Ti, Nikon) equipped with NIS Elements software (Nikon), a 60X air objective (CFI Plan Apo λ, 0.95 NA, Nikon) fitted with an objective heater (Bioptechs), a motorized stage (Proscan III, Prior), and an active feedback-based focusing system (Perfect

Focus System, Nikon). An LED lamp (Sola II, Lumencorp) and a scientific-grade CMOS camera (Zyla 5.5, Andor) were used for fluorescent imaging. Multi-color fluorescent imaging of sfGFP, mKO2 and E2Crimson fluorophores was carried out by alternating between three filter sets mounted in a computer-controlled filter ring (Chroma 49002, 49010 and 49015, respectively). To help with the semi-automated cell identification, each channel was imaged every ten minutes via brightfield imaging (100 ms exposure, both in focus and 4 μm below the focal plane). Fluorescent images of each of the three fluorophores were taken every thirty minutes (150 ms exposure). This illumination scheme was well below the phototoxicity limit, as described previously (Al-Sady et al., 2016). Raw images were saved as uncompressed 16 bit ND2 files and further analyzed using a custom-written image analysis pipeline (see below).

Cells were grown overnight (30°C with 225 rpm shaking) to saturation in YES media, then diluted in YES to an optical density at 600 nm (OD_{600}) of 0.1 and allowed to grow for approximately 5 hours to reach an OD_{600} of 0.5. Cells (60 μL at OD 0.5 in YES+2% Bovine Serum Albumin, BSA) were loaded at the entry port of the multFYLM. After cells entered individual channels, media lines were reattached and YES media was pumped through on a pulse cycle (14 min: 5 μLmin^{-1}, 1 min: 55 μLmin^{-1}) for the entire experiment. This flow regime was optimized to flush out occasional cell clumps that grew at the device inlets and other fluidic interfaces. Four genotypes were imaged simultaneously for 60 hours in each channel of a multFYLM device to ensure identical imaging and growth conditions. In all cases, we only analyze the innermost cell, which was the oldest cell pole (see below). Cells that were ejected or died within the first 12 hours after loading were not included in the downstream analysis.

Single-cell image analysis:

Single-cell imaging data was processed using an updated version of the custom-written FYLM Critic analysis package (Spivey et al. 2017). The source-code is available via GitHub (https://github.com/finkelsteinlab/fylm). FYLM Critic performs the following automated processing on the raw images: (1) rotation; (2) jitter removal via a cross-correlation algorithm; and (3) generation of kymograph and individual cell images. The latter were used to create videos of individual cells in Fiji (Schindelin et al. 2012). The final outputs of FYLM critic are the position and contour of each dividing cell, as well as the time-dependent fluorescence intensities for each cell. These fluorescence intensities are obtained by averaging the intensity across all pixels that fall within the cell volume, as defined by the bright field images. This normalization also ensures that the fluorescence intensity is corrected for the size of the rapidly dividing cells. Time-dependent fluorescent intensities were analyzed via custom-written MATLAB scripts (version 2017a Mathworks, available upon request). Background fluorescence from the PDMS device was subtracted using catch tubes that did not receive a cell. The maximum heterochromatin reporter (GFP, mKO2) fluorescence intensity was calculated using Δ*clr4* cells in the same reporter construct background. To control for expression variation across the cell cycle, the fluorescence from heterochromatin reporters was also reported as a ratio of the control fluorophore, E2Crimson. Similarly, cells fluorescing in the clamp channel were removed from analysis for MAT locus derived strains (SUPPLEMENTAL MATERIALS AND METHODS).

Single cell images generated by the FYLM Critic analysis were compiled into stacked movies using Fiji. Images in bright field and for each color channel were processed separately in batch and then later combined into a vertical stack. For each channel, 0.2% of pixels were

allowed to become saturated and pixel values were normalized to the maximum range for the whole sequence in that channel. For bright field, every third image was included to match the imaging frequency of the fluorescent channels. Movies were edited for length to include contiguous imaging sequences without loss of focus and for size to remove non-cellular debris and cells from the opposite side of the channel that entered the field of view. After combining all color channels and bright field, the brightness and contrast were increased for cell 407 to match the red channel brightness of the other strains. Image sequences were saved as uncompressed .avi files with a rate of 15 frames per second.

Table 2.1: Yeast strains used in this study

Strain	Genotype
PAS075	Locus2*::ade6p::3xE2C:hygMX* at Locus2 (between SPBC1711.11 andSPBC1711.12)
PM03	Wild type strain: h(+); ura4-D18; leu1-32; ade6-M216; his7-366
PM1035	*ura4::natMX:dh* fragment 1, *clr4::KAN* as in *(Marina et al., 2013)*
PAS111	*ura4::natMX:dh*:*ade6p:*SF-GFP, *ade6p:*mKO2 7kb, *ade6p:*3xE2C: *hygMX* at Locus2
PAS112	*ura4::natMX:dh*:*ade6p:*SF-GFP, *ade6p:*mKO2 7kb, *ade6p:*3xE2C: *hygMX* at Locus2; *clr4::kanMX*
PAS133	*ura4::natMX:dh*:*ade6p:*SF-GFP, *ade6p:*mKO2 1kb, *ade6p:*3xE2C: *hygMX* at Locus2; *clr4::kanMX*
PAS134	*ura4::natMX:dh*:*ade6p:*SF-GFP, *ade6p:*mKO2 1kb, *ade6p::*3xE2C: *hygMX* at Locus2
PAS135	*ura4::natMX:dh*:*ade6p:*SF-GFP, *ade6p:*mKO2 3kb, *ade6p::*3xE2C: *hygMX* at Locus2; *clr4::kanMX*
PAS136	*ura4::natMX:dh*:*ade6p:*SF-GFP, *ade6p:*mKO2 3kb, *ade6p::*3xE2C: *hygMX* at Locus2

Strain	Genotype
PAS141	*ura4::natMX:dh:ade6p:*SF-GFP , *ade6p:*mKO2 5kb, *ade6p::*3xE2C: *hygMX* at Locus2
PAS142	*ura4::natMX:dh:ade6p:* SF-GFP, *ade6p:*mKO2 5kb; *ade6p::*3xE2C: *hygMX* at Locus2; *clr4::kanMX*
PAS192	Δ*K::ade6p:*mKO2; *ade6p:* SF-GFP between *REIII* and mat3M; *ade6p:*3xE2C: *hygMX* at Locus2, h(-)
PAS193	Δ*K::ade6p:*mKO2; *ade6p:*SF-GFP between *REIII* and mat3M; *ade6p:*3xE2C: *hygMX* at Locus2; *clr4::kanMX,* h(-)
PAS214	Δ*K::ade6p:*mKO2*:ura4t; mat3m(Eco*RV*)::* ade6p:SF-GFP; *ade6p:*3xE2C: *hygMX* at Locus2; *clr4::kanMX,* h(-)
PAS215	Δ*K::ura4t:*mKO2*:ade6p; mat3m(Eco*RV*)::* ade6p:SF-GFP; *ade6p:*3xE2C: *hygMX* at Locus2; *clr4::kanMX,* h(-)
PAS216	*cenH::ade6p:*SF-GFP (Kint2); *mat3m(Eco*RV*)::* ade6p:mKO2; *ade6p:*3xE2C: *hygMX* at Locus2; *clr4::kanMX,* h90
PAS217	*cenH: ade6p:*SF-GFP (Kint2); *mat3m(Eco*RV*)::* ade6p:mKO2; *ade6p:*3xE2C: *hygMX* at Locus2*,* h90
PAS218	*cenH::ade6p:*mKO2 (Kint2); *mat3m(Eco*RV*)::* ade6p:SF-GFP; *ade6p:*3xE2C: *hygMX* at Locus2; in *clr4::kanMX,* h90
PAS219	*cenH: ade6p:*mKO2 (Kint2); *mat3m(Eco*RV*)::* ade6p:SF-GFP; *ade6p:3xE2C: hygMX* at Locus2*,* h90
PAS231	*ura4::natMX:dh:ade6p:*SF-GFP, *ade6p:*mKO2 3kb, *leu1::ade6p:*3xE2C: *hygMX*
PAS237	*ura4::natMX:dh:ade6p:*SF-GFP, *ade6p:*mKO2 3kb, *act1p::q*xE2C: *hygMX* at Locus2; *clr4::kanMX*
PAS243	*ura4::natMX:dh:ade6p:*SF-GFP, *ade6p:*mKO2 3kb, *act1p::*1xE2C: *hygMX* at Locus2; *clr4::kanMX*
PAS244	*ura4::natMX:dh:ade6p:*SF-GFP, *ade6p:*mKO2 3kb, *act1p::*1xE2C: *hygMX* at Locus2
PAS264	*cenH:: ade6p:*SF-GFP (Kint2); *mat3m(Eco*RV*)::* ade6p:mKO2; *ade6p:*3xE2C: *hygMX* at Locus2, *pcr1::kanMX,* h90
PAS268	Δ*K:: ade6p:*mKO2; *ade6p:*SF-GFP between REIII and mat3M; *ade6p:*3xE2C: *hygMX* at Locus2, *REII::LEU2,* h(-)

Strain	Genotype
PAS269	*ΔK*:: *ade6p:*mKO2; *ade6p:*SF-GFP between REIII and mat3M; *ade6p:*3xE2C:*hygMX* at Locus2; *clr4::kanMX, REII::LEU2,* h(-)
PAS331	*cenH*:: *ade6p:*SF-GFP (Kint2); *mat3m(Eco*RV*)*:: *ade6p:*mKO2; *ade6p:*3xE2C:*hygMX* at Locus2; *ΔREIII*::REIII(Δs1, Δs2) in *clr4::kanMX,* h90
PAS332	*cenH*:: *ade6p:*SF-GFP (Kint2); *mat3m(Eco*RV*)*:: *ade6p:*mKO2; *ade6p:*3xE2C:*hygMX* at Locus2; *ΔREIII*::REIII(Δs1, Δs2) , h90
PAS348	*ura4::hygMX:REIII:ade6p:*SF-GFP; *ade6p:*mKO2 5kb, *ade6p:*3xE2C:*natMX* at Locus2
PAS350	*ura4::hygMX:REIII:ade6p:*SF-GFP; *ade6p:*mKO2 5kb, *ade6p:*3xE2C:*natMX* at Locus2 *dcr1::kanMX*
PAS355	*ura4::natMX:dh:ade6p:*SF-GFP, *ade6p:*mKO2 3kb, *leu1::ade6p:*3xE2C:*hygMX*; *clr4::kanMX*
PAS385	*ΔK*:: *ade6p:*mKO2; *ade6p:*SF-GFP between *REIII* and mat3M; *act1p:*1xE2C:*hygMX* at Locus2; *clr4::kanMX,* h(-)
PAS387	*ΔK*:: *ade6p:*mKO2; *ade6p:* SF-GFP between *REIII* and mat3M; *act1p:*1xE2C*: hygMX* at Locus2, h(-)
PAS388	*cenH:: ade6p:*SF-GFP (Kint2); *mat3m(Eco*RV*)*:: *ade6p:*mKO2; *ura4* at Locus2; *leu1::act1p:*1xE2C:*hygMX, clr4::kanMX,* h90
PAS389	*cenH:: ade6p:*SF-GFP (Kint2); *mat3m(Eco*RV*)*:: *ade6p:*mKO2; *ura4* at Locus2; *leu1::act1p:*1xE2C:*hygMX,* h90
PAS390	*cenH:: ade6p:*SF-GFP (Kint2); *mat3m(Eco*RV*)*:: *ade6p:*mKO2; *ura4* at Locus2; *ΔREIII*::REIII(Δs1, Δs2), *leu1::act1p:*1xE2C:*hygMX,* in *clr4::kanMX,* h90
PAS391	*cenH:: ade6p:*SF-GFP (Kint2); *mat3m(Eco*RV*)*:: *ade6p:*mKO2; *ura4* at Locus2; *ΔREIII*::REIII(Δs1, Δs2), *leu1::act1p:*1xE2C:*hygMX,* h90
PAS398	*his1::natMX:dh:ade6p:*mKO2*; ade6p:*SF-GFP 3kb, *ade6p::*3xE2C:*hygMX* at Locus2, *clr4::kanMX, ura4::phyB.*
PAS399	*his1::natMX:dh:ade6p:*mKO2*; ade6p:*SF-GFP 3kb, *ade6p::*3xE2C:*hygMX* at Locus2, *ura4::phyB.*
PAS410	*ΔK*:: *ade6p:*mKO2; *ade6p:* SF-GFP between *REIII* and mat3M; *ade6p::*3xE2C:*hygMX* at Locus2, *natMX:clr4+,* h(-); 'OFF' allele
PAS411	*ΔK*:: *ade6p:*mKO2; *ade6p:* SF-GFP between *REIII* and mat3M; *ade6p::*3xE2C:*hygMX* at Locus2, *natMX:clr4+,* h(-); 'ON' allele

| Strain	Genotype
PAS464	ΔK::*ade6p*:mKO2:*ura4t*; *mat3m(Eco*RV*)*:: *ade6p*:SF-GFP; *ade6p*:3xE2C: *hygMX* at Locus2; *natMX:clr4+,* h(-)
PAS465	ΔK::*ura4t*:mKO2:*ade6p*; *mat3m(Eco*RV*)*:: *ade6p*:SF-GFP; *ade6p*:3xE2C: *hygMX* at Locus2; *natMX:clr4+,* h(-)
PAS473	ΔK:: *ade6p*:mKO2; *ade6p*: SF-GFP between *REIII* and mat3M; 700bp *sib1* ORF between REIII-s1 and mKO2; *ade6p*:3xE2C: *hygMX* at Locus2, *clr4::kanMX,* h(-);
PAS474	ΔK:: *ade6p*:mKO2; *ade6p*: SF-GFP between *REIII* and mat3M; ΔREIII::REIII(Δs1, Δs2), *ade6p*:3xE2C: *hygMX* at Locus2, *clr4::kanMX,* h(-);
PAS478	ΔK:: *ade6p*:mKO2; *ade6p*: SF-GFP between *REIII* and mat3M; 700bp *sib1* ORF between REIII-s1 and mKO2 *ade6p*:3xE2C: *hygMX*, *natMX:clr4+,* h(-);
PAS482	ΔK::*ade6p*:mKO2; *ade6p*: SF-GFP between *REIII* and mat3M; *ade6p*:3xE2C: *hygMX* at Locus2, h(-); 'OFF' allele
PAS483	ΔK:: *ade6p*:mKO2; *ade6p*: SF-GFP between *REIII* and mat3M; ΔREIII::REIII(Δs1, Δs2), *ade6p*:3xE2C: *hygMX* at Locus2, *natMX:clr4+,* h(-);
PAS496	*cenH*:: *ade6p*:SF-GFP (Kint2); *mat3m(Eco*RV*)*:: *ade6p*:mKO2; *ade6p*:3xE2C:*hygMX* at Locus2; ΔREIII::REIII(Δs1, Δs2), *ars1::prad15:cre-EBD:LEU2*; *h3.2:lox:HA:hygMX:lox:T7;* h90
PAS497	ΔK::*ade6p*:mKO2; *ade6p*: SF-GFP between *REIII* and mat3M; *ade6p*:3xE2C: *hygMX* at Locus2; *ars1::prad15:cre-EBD:LEU2*; *h3.2:lox:HA:hygMX:lox:T7;* 'OFF' allele, h(-)
PAS498	ΔK::*ade6p*:mKO2; *ade6p*: SF-GFP between *REIII* and mat3M; *ade6p*:3xE2C: *hygMX* at Locus2; *ars1::prad15:cre-EBD:LEU2*; *h3.2:lox:HA:hygMX:lox:T7;* 'ON' allele; h(-)
PAS508	ΔK::*ade6p*:mKO2; *ade6p*: SF-GFP between *REIII* and mat3M; *ade6p*:3xE2C: *hygMX* at Locus2, 'OFF' allele; *pcr1::kanMX*
PAS510	*cenH*:: *ade6p*:SF-GFP (Kint2); *mat3m(Eco*RV*)*:: *ade6p*:mKO2; *ade6p*:3xE2C:*hygMX* at Locus2; ΔREIII::REIII(Δs1, Δs2), *pcr1::kanMX*
PAS514	ΔK::*ade6p*:mKO2; *ade6p*: SF-GFP between *REIII* and mat3M; *ade6p*:3xE2C: *hygMX* at Locus2, 'OFF' allele; *dcr1::kanMX*; *seb1-1:natMX*
PAS515	*cenH*:: *ade6p*:SF-GFP (Kint2); *mat3m(Eco*RV*)*:: *ade6p*:mKO2; *ade6p*:3xE2C:*hygMX* at Locus2; ΔREIII::REIII(Δs1, Δs2), *seb1-1:natMX*
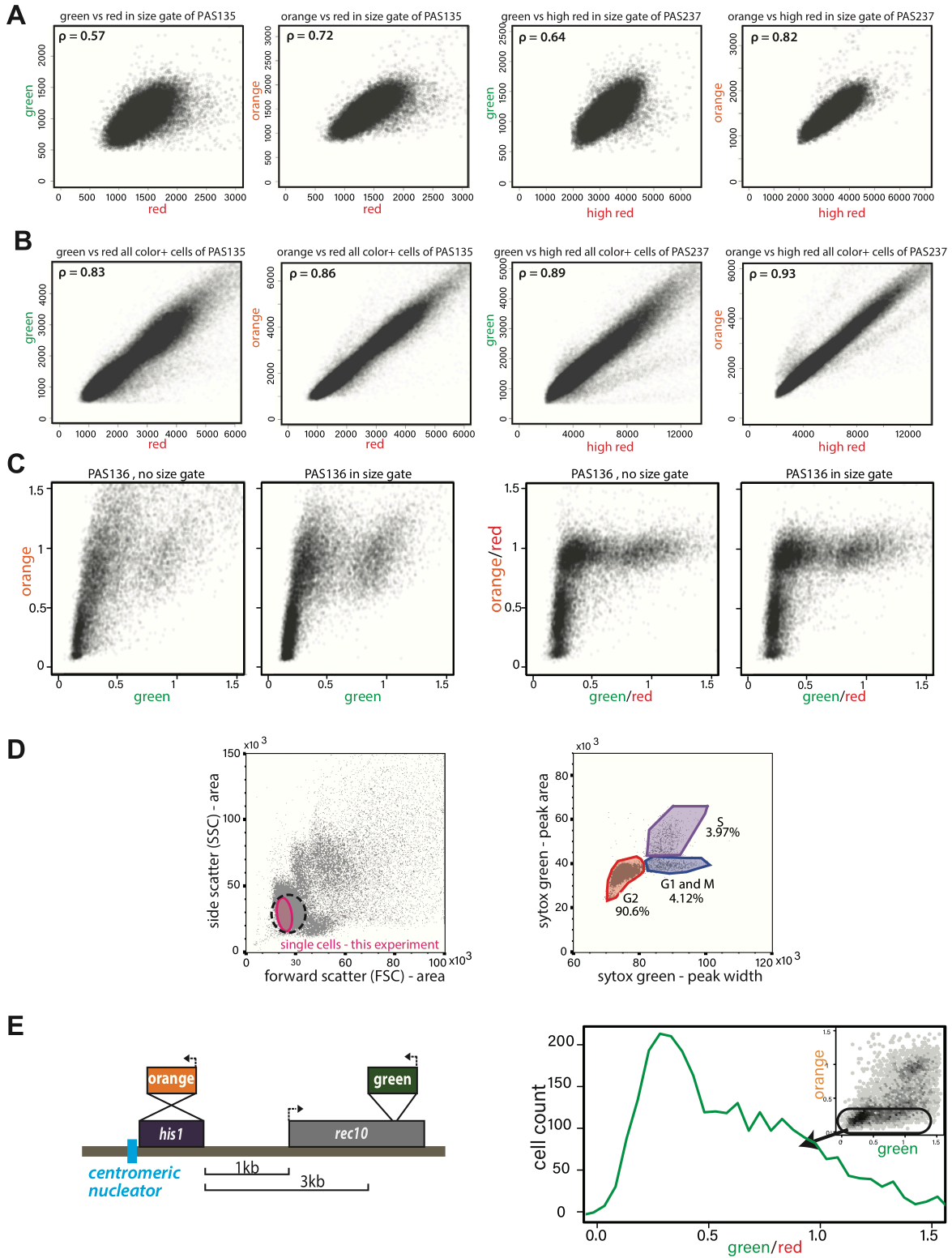
Figure 2.1 Supplement: Validation of ectopic heterochromatin spreading sensor.
A. Correlation of *ade6p:SFGFP* or *ade6p:mKO2* with *ade6p:3XE2C* (Red) or *act1p:1XE2C*

(High Red) in *Δclr4* HSS size-gated cells. LEFT: Plots of green and orange vs. red channel signals of size-gated PAS 135 (*Δclr4*, "red"). RIGHT: Plots of green and orange vs. red channel signals of size-gated PAS 237 (*Δclr4*, "high-red"). The Pearson correlation between "green" and "red"/"high-red" or "orange" and "red"/"high-red" is shown. **B.** Correlation of *ade6p:SFGFP* or *ade6p:mKO2* with *ade6p:3XE2C* (Red) or *act1p:1XE2C* (High Red) in *Δclr4* HSS in cells without size gate. Plots and Pearson correlation as above. **C.** Effect of red-normalization on distribution of *clr4+* HSS cells. Plots of green and orange vs. red channel signals of PAS 136, which contains the ectopic HSS (Figure 1C). LEFT: effect of using only size gate, without red normalization. RIGHT: effect of red-normalization with and without additional size gate. The distribution of cells is tightened by red-normalization. **D.** Cell cycle stage of HSS and wild type cells by flow cytometry. Wild-type cells (PM03, see strain table) were fixed, stained with Sytox green DNA stain, and analyzed by flow cytometry. LEFT: side vs. forward scatter plot. Dotted line: The approximate size gate encompassing all experiments reported. Pink area: cells analyzed in the experiment shown. RIGHT: Plot of area vs. width parameter for the Sytox green channel, gates are drawn to denote cell cycle phases, G2 (red), G1 and M (Blue), S (purple) as described (Knutsen et al., 2011). **E.** Stochastic spreading and intermediate states produced by ncRNA-driven nucleators are replicated at a second ectopic site. LEFT: Overview of the *his1::dh*HSS^{3kb}. The colors are reversed relative to the *ura4::dh*HSS^{1-7kb} with "orange" as the "nucleation clamp" and "green" as the "sensor". "Orange" replaces the *his1* gene and "green" is located 3kb downstream within the *rec10* open reading frame. RIGHT: histogram of "red"-normalized "green" fluorescence distribution of "orange"^{OFF} cells. Inset: 2D density hexbin plot.
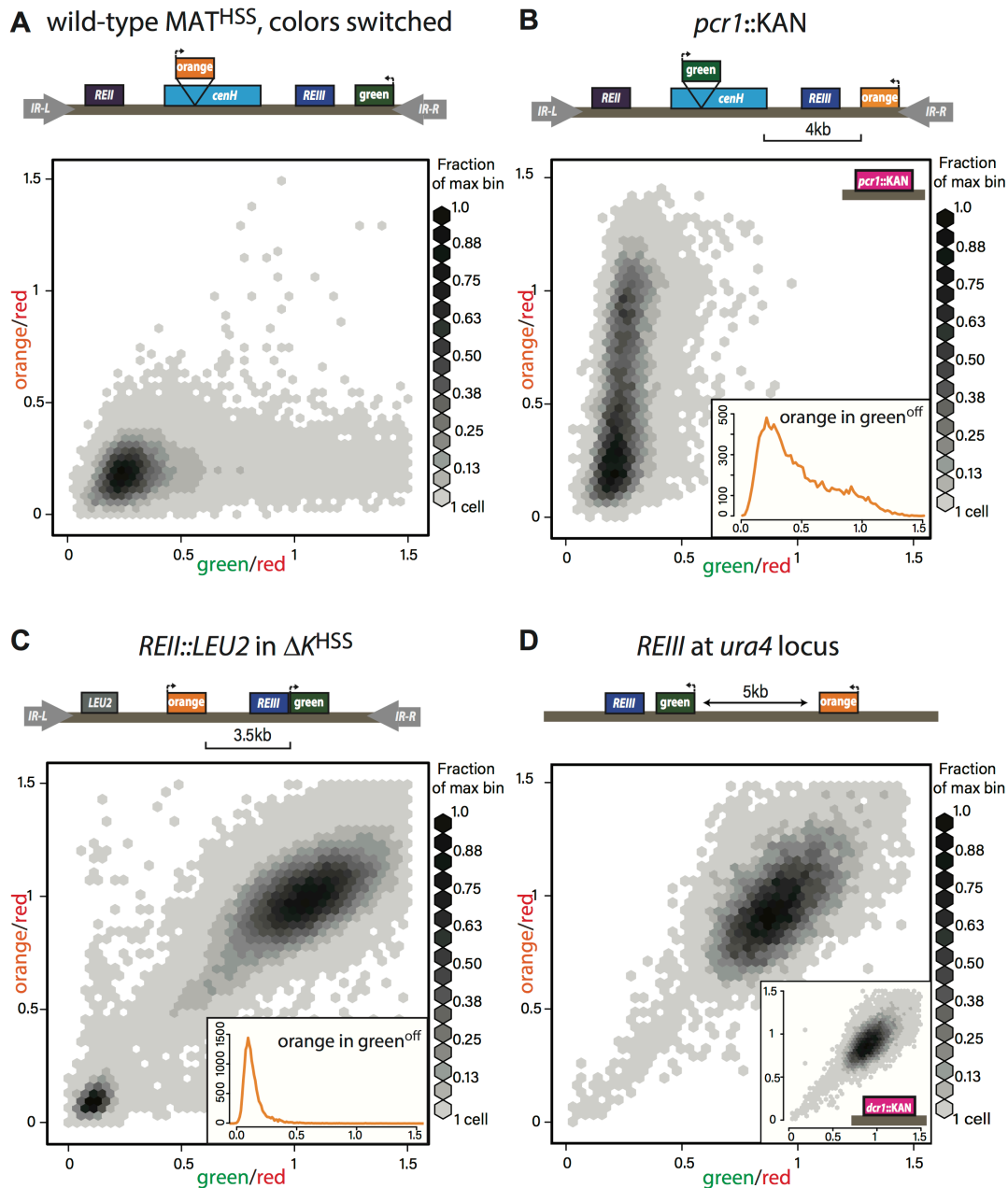
Figure 2.2 Supplement 1: Heterochromatin spreading characteristics of *cis*-acting elements at the tightly repressed MAT locus. A. The MAT[HSS] documents tight repression of the wild type MAT locus. As in Figure 2.2A and B, with "green" and "orange" switched. **B.** stochastic spreading with intermediate states in *pcr1::KAN*. *pcr1* transcription factor was knocked-out in the PAS217 wild-type MAT[HSS]. Plot and inset as in Figure 2.2B. **C.** *REII* does not contribute to bimodal distribution seen for *ΔK*[HSS]. The *REII* locus (1kb) was replaced with the *LEU2* gene before *clr4+* was introduced by cross. **D:** *REIII* is unable to establish spreading at an ectopic site. 2D density hexbin plots of *ura4::REIII*HSS[5kb]. Normalized green and orange are near 1.0, indicating a failure to repress both reporters. Inset: 2D density hexbin plots of *ura4::REIII*HSS[5kb] *dcr1::KAN. dcr1* was deleted to release extra heterochromatin factors from RNAi- repressed loci. No additional silencing is detected.
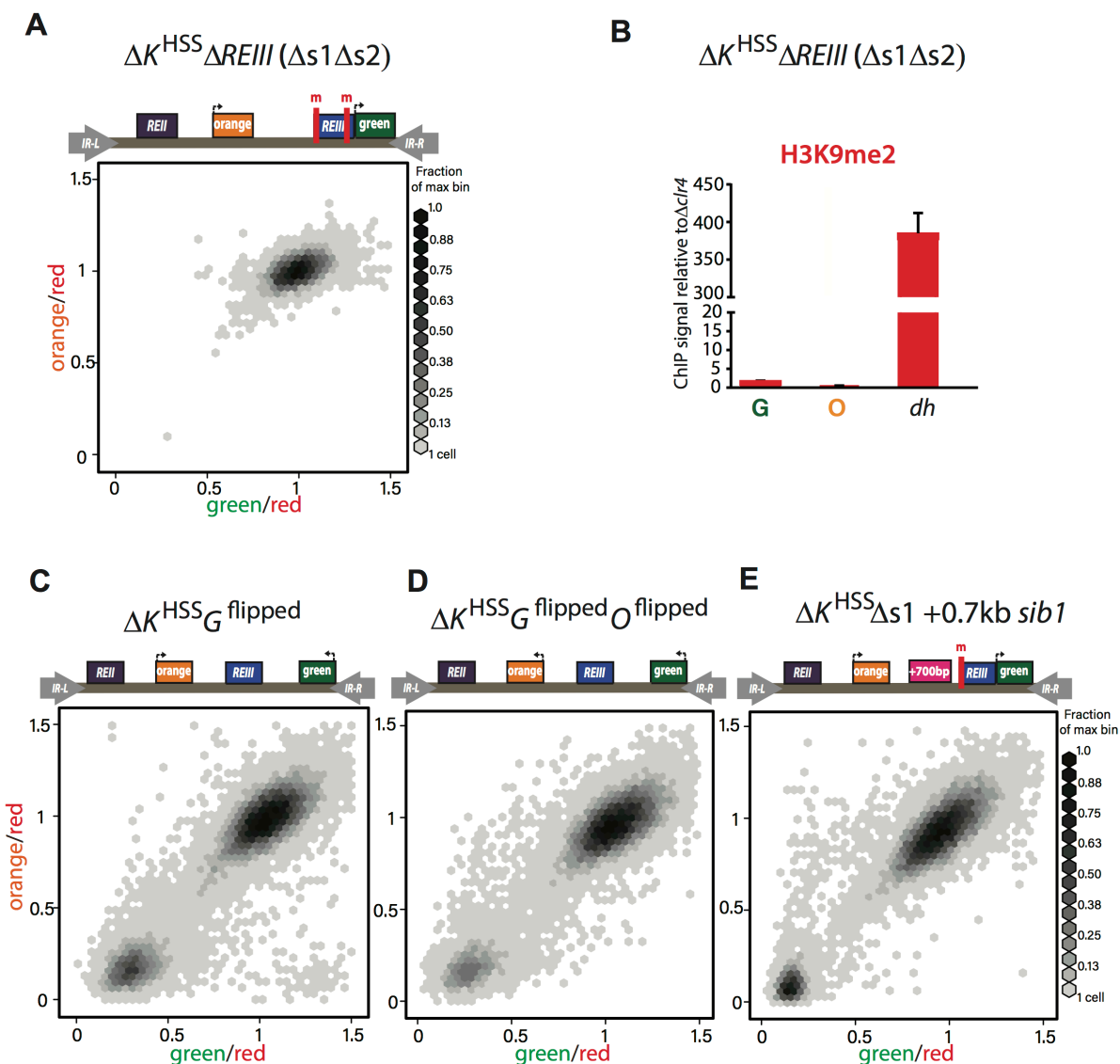
Figure 2.2 Supplement 2: *REIII* is required for heterochromatin formation in *ΔK*^HSS.
A. Deletion of both Atf1/Pcr1 binding sites before introduction of *clr4+* in *ΔK*^HSS blocks gene silencing. In 34/34 strains tested (one representative shown), *ΔK*^HSS *Δs1Δs2* cannot form repressed states. **B.** H3K9me2 does not accumulate when both Atf1/Pcr1 binding sites are deleted in *ΔK*^HSS. H3K9me2 ChIP in *ΔK*^HSS *Δs1Δs2* at "green", "orange" and *dh. ΔK*^HSS -OFF accumulates H3K9me2 to similar extent as *dh,* Figure 2.2E). Error bars indicate standard deviation of technical replicates. **C.** "green" orientation and position does not substantially affect *ΔK*^HSS behavior. In *ΔK*^HSS G^flipped "green" is flipped in orientation with respect to *ΔK*^HSS. **D.** "green" and "orange" orientations do not substantially affect *ΔK*^HSS behavior. In *ΔK*^HSS G^flipped O^flipped "green" is located as in C and "orange" is flipped in orientation with respect to *ΔK*^HSS. "green" in **C**. and **D**. is 2.1kb downstream from its location in *ΔK*^HSS now on the distal side of the *mat3m* cassette. **E.** Increasing distance between *REIII* and "orange" does not substantially affect *ΔK*^HSS behavior. The Atf1/Pcr1 binding site proximal to "orange" was deleted (Δs1) and 700bp of the *sib1* ORF inserted to the left of the Δs1 site. 2D-hexbin plots as in Figure 2.2.

59

Figure 2.3 Supplement 1: Single cell analysis of nucleation and spreading using a Fission Yeast Lifespan Micro-dissector (FYLM). A. For *ura4::dh*HSS[3kb] FYLM experiments, counts of

60

cells in each of seven categories. Diagrams indicate the time-dependent silencing behaviors of cells in each category. Category 1-3 are consistent with proximal to distal silencing, whereas category 4-6 are consistent with a distal to proximal silencing. **B.** Time-dependent traces showing cells from Category 1 where the normalized "green" and "orange" values at each time point are plotted color-coded by time where blue and pink represent the start and end of the measurement respectively. LEFT: Traces for all Category 1 cells, which begin at the start of the silencing event with both colors fully expressed and end when both colors have reached their local minimum. RIGHT: Four example cells where points represent 30-minute time points colored from the start to end of the event. The duration of the time represented is indicated in the lower right corner. **C.** Traces for Category 2 cells during their entire measured lifespan. **D.** Traces for Category 3 cells during their entire measured lifespan. **E.** Time-dependent traces for the one cell in Category 4. Lines are plotted and time is curated as in **B**.
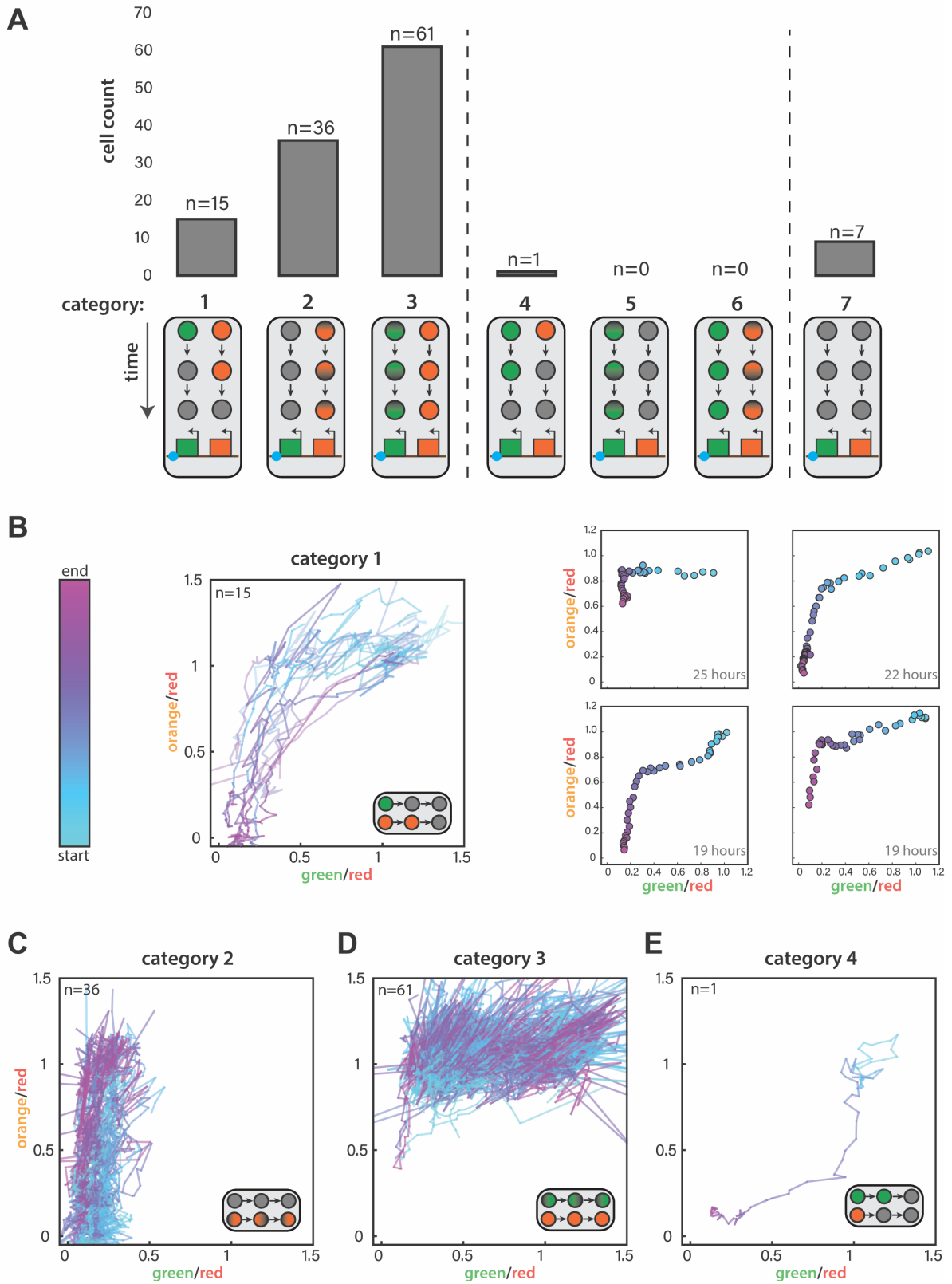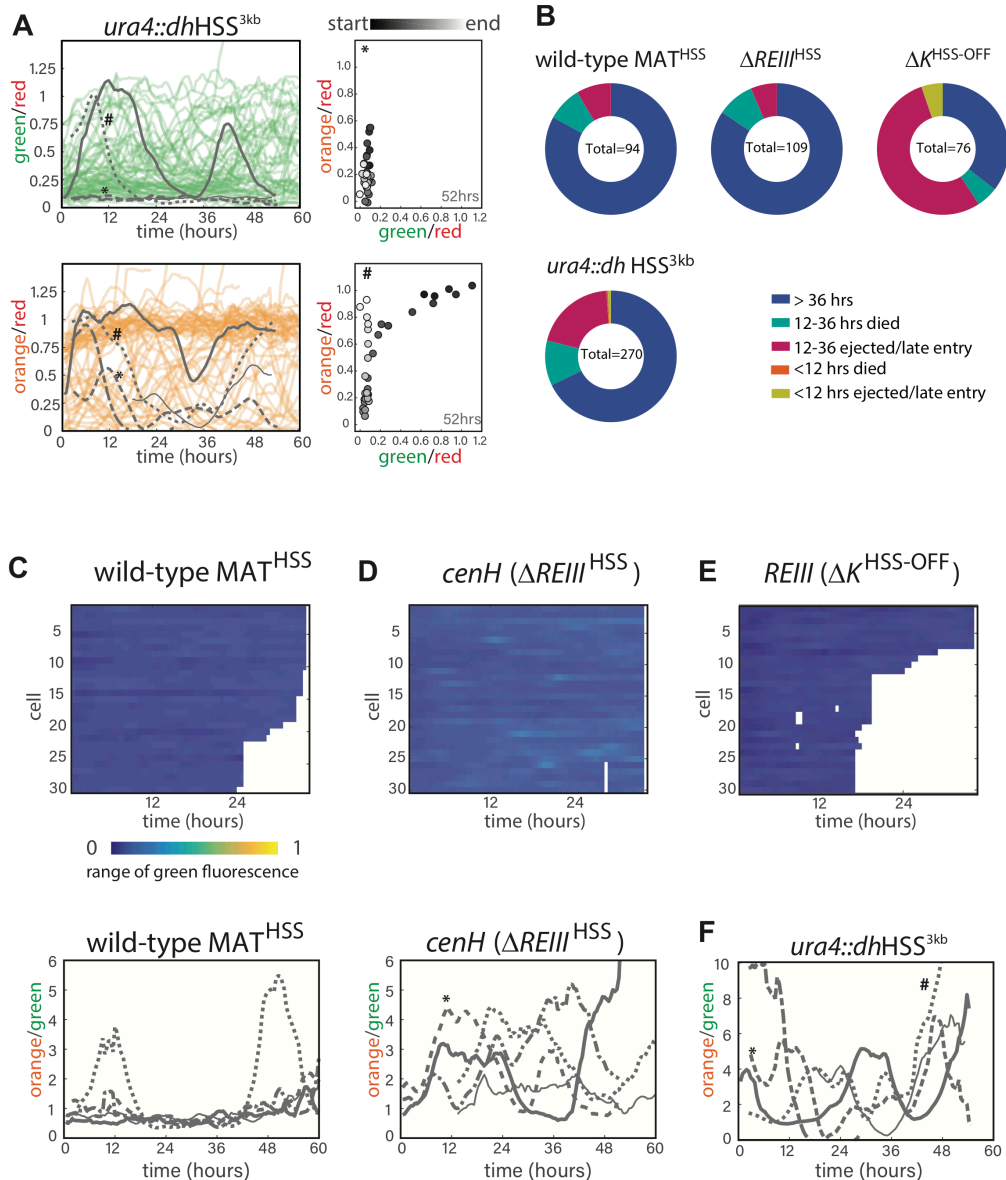
Figure 2.3 Supplement 2: Single cell analysis of nucleation and spreading using a Fission Yeast Lifespan Micro-dissector (FYLM). A. FYLM analysis of *ura4::dh*HSS[3kb] cells. TOP LEFT: 60 hours of normalized "green" fluorescence, a subset of cells is shown for clarity. 5 example cells are overlaid in gray each with different line types. BOTTOM LEFT: 60 hours of normalized "orange" fluorescence in the matching subset of cells with the same 5 overlaid in gray. *, # represent two example cells. RIGHT: for two representative sample cells imaged, plots of normalized "green" and "orange" across its measured lifetime (grayscale). The corresponding cells are marked in the orange traces on LEFT. **B.** Categorization of cell longevity of all cells analyzed in the FLYM experiment. Measured lifespan ends when a cell dies or is ejected from its capture channel. **C.** For wild-type MAT[HSS] TOP: "green" fluorescence heatmap (blue (0) to yellow (1)) for the same 30 cells as in 3C. BOTTOM: 60 hours of traces for "orange" divided by "green" for the five example cells indicated in 3C. **D.** "green" fluorescence heatmap and "orange"/"green" traces for *ΔREIII*[HSS] as in C. **E.** "green" fluorescence heatmap *ΔK*[HSS] as in C. **F.** "orange"/"green" traces for *ura4::dh*HSS[3kb] as in C. *,# indicate the same cells as in A.
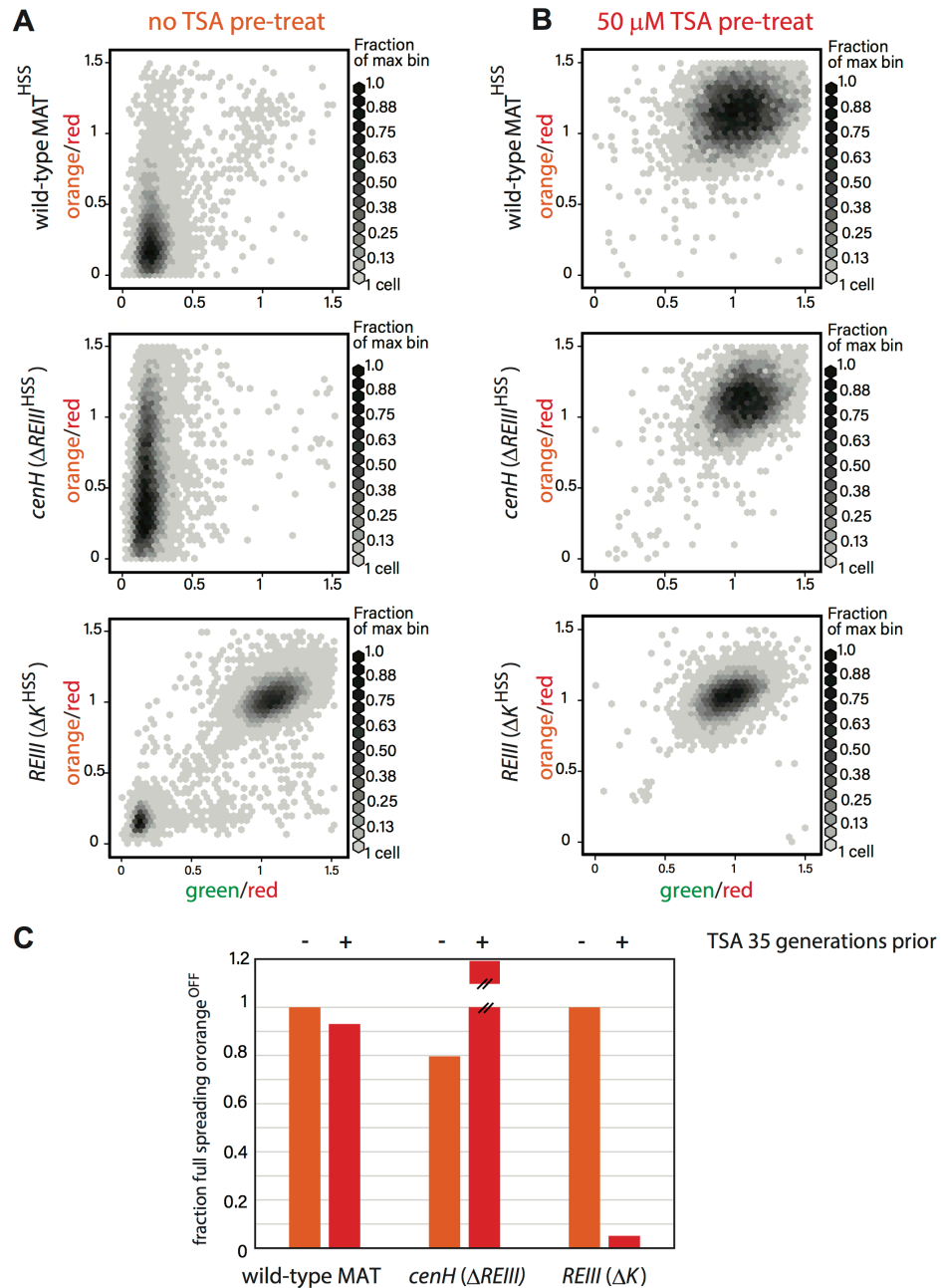
Figure 2.4 Supplement 1: Heterochromatin behaviors during TSA treatment and after 35 generations. A. 2D density hexbin plots of wild-type MAT^{HSS}, $\Delta REIII^{HSS}$, and ΔK^{HSS} strains grown 10 generations without TSA. **B.** 2D density hexbin plots of wild-type MAT locus^{HSS}, $\Delta REIII^{HSS}$, and ΔK^{HSS} strains grown 10 generations in 50 µM TSA. The density distributions are near 1.0 in all strains indicating complete erasure of heterochromatin. **C.** History dependence at 35 generations after pretreatments. The fraction of cells with full spreading (wild-type MAT and $\Delta REIII$) or fraction of cells with orange^{OFF} (ΔK) normalized to the highest value for ancestrally untreated cells (=1) is shown for the 0 µM TSA point. TSA pretreated cells for $\Delta REIII^{HSS}$ show higher repression than untreated cells. We interpret this to indicate experimental variations in silencing in the absence of memory. This is because for all other circumstances, TSA treatment results in reduced spreading, including for $\Delta REIII^{HSS}$ at 25 generations post treatment.
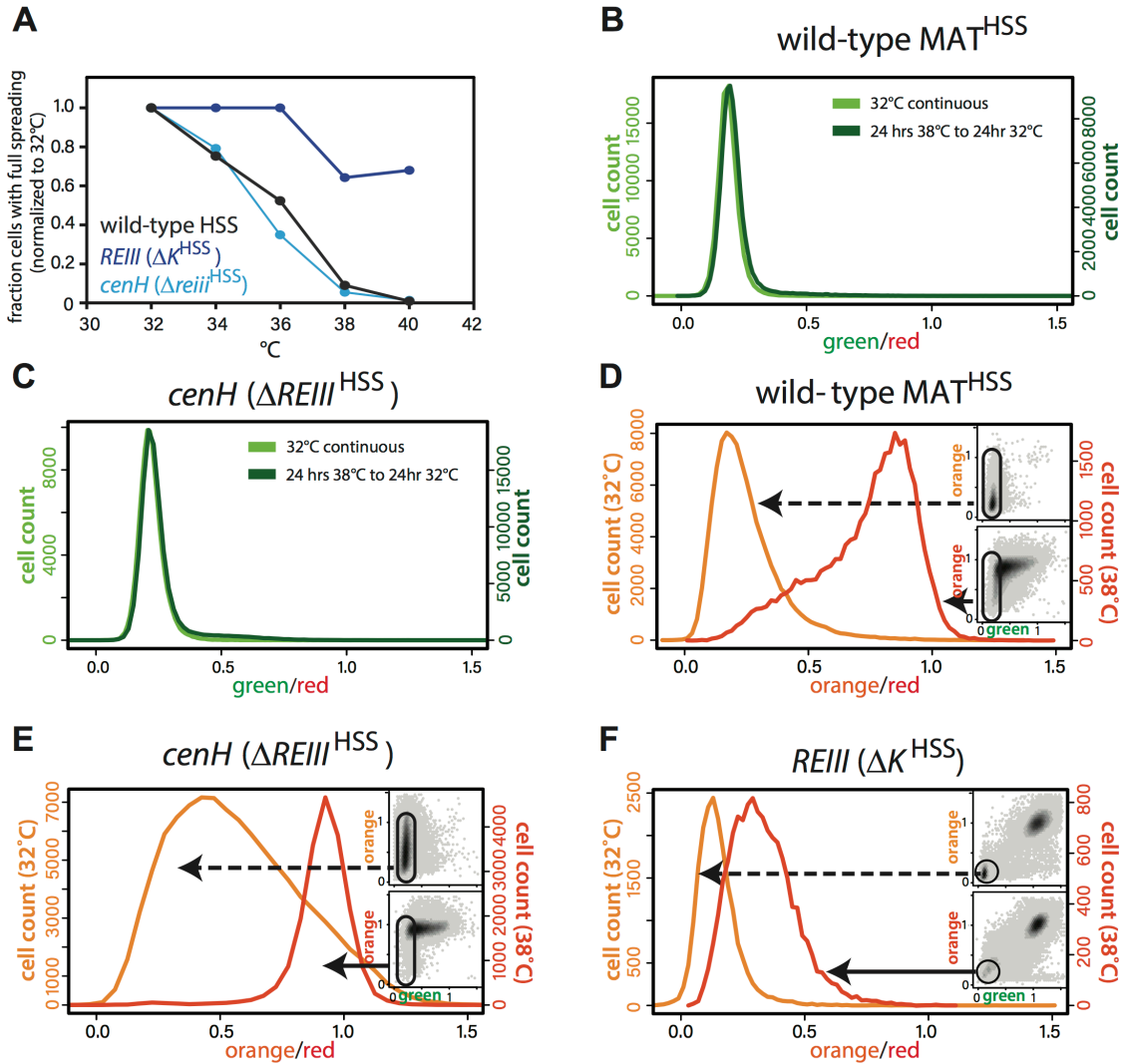
Figure 2.4 Supplement 2: Behavior of MAT heterochromatin at elevated temperature. A. The resistance of the heterochromatin state from 32-40°C in wild-type MAT[HSS], ΔK[HSS], and ΔREIII[HSS]. The fraction of cells that fully repress both "orange" and "green" (full spreading) at each temperature is plotted normalized to the given strains value at 32°C. **B.&C.** nucleation is recovered within 24 hours at 32°C. 1-D histogram showing the distribution of green fluorescence in wild-type MAT locus[HSS] (**B**.) or ΔREIII[HSS] (**C**.) cells grown either for 48hrs continuously at 32°C (left y-axis, light green) or heat stressed for 24hrs at 38°C followed by 24hrs growth at 32°C (right y-axis, dark green). **D.-F.** Histograms of "red"-normalized "orange" fluorescence distribution in "green"[OFF] cells are shown for cells grown at both 32°C (light orange) and 38°C (dark orange). Insets: 2D density hexbin plots, "green"[OFF] cells are schematically circled. C.-E. represent t=0 in Figure 4F.
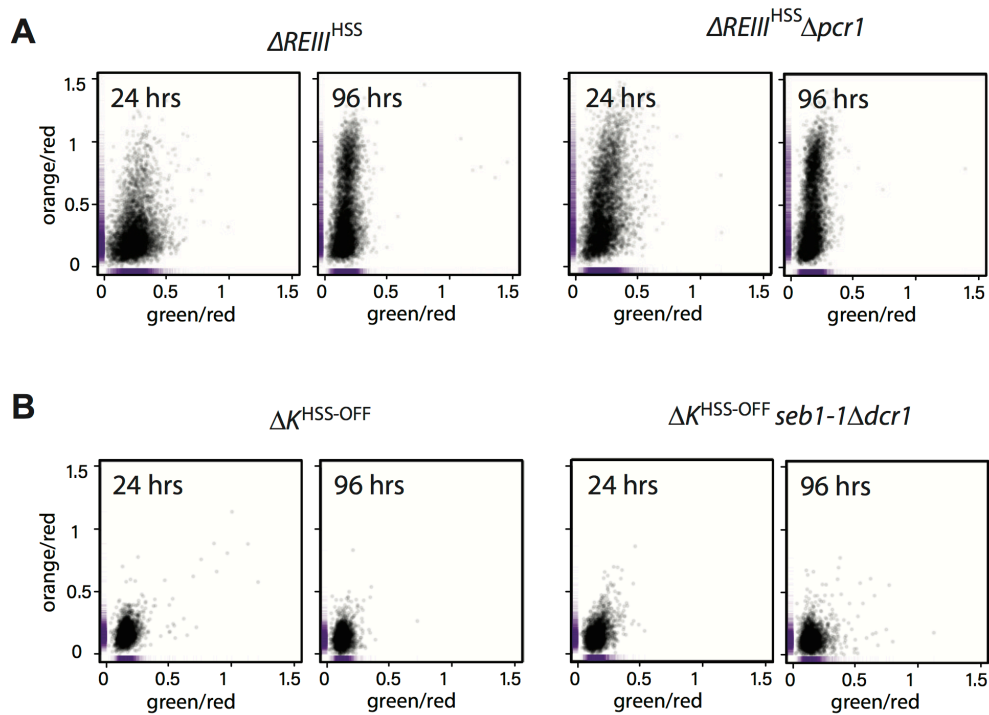
Figure 2.5 Supplement: *Trans*-factor mutants do not substantially affect spreading when their cognate *cis*-acting element is inactivated. **A.** Scatter plots of *ΔREIII*[HSS] and *ΔREIII*[HSS]*Δpcr1* at 24 and 96hrs. **B.** Scatter plots of *ΔK*[HSS] and *ΔK*[HSS]*seb1-1Δdcr1* at 24 and 96hrs. The *seb1-1* and *Δdcr1* double mutant should abolish all RNA- dependent nucleation (Marina et al., 2013). The X and Y values of each cell are represented by purple dashes along the axis.
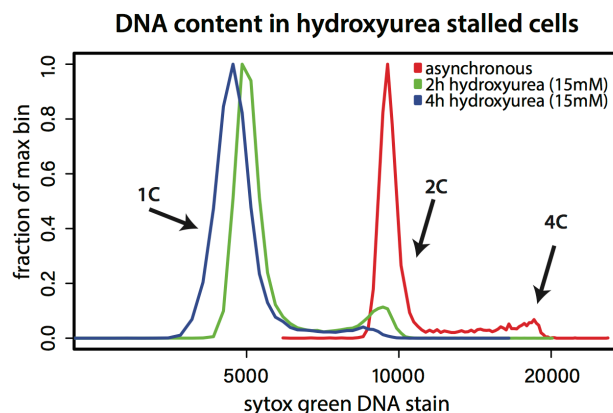
Figure 2.6 Supplement: Hydroxyurea induced cell cycle arrest. Cells were grown without (asynchronous) or with 15mM hydroxyurea for 2hrs or 4hrs and DNA content was determined by Sytox green staining and flow cytometry. Hydroxyurea treatment stalls cells in early S phase, evident from loss of 2C and 4C peaks.

SUPPLEMENTAL DISCUSSION

Alternative formal model for *cenH* and *REIII* interaction

In the main discussion, we propose that *cenH* stimulates *REIII* nucleation to account for the high proportion of the spreading in wild-type MAT cells. Another formal possibility remains that non-nucleated, Atf1/Pcr1-bound, *REIII* raises *cenH* spreading efficiency. In the *ΔREIII* strain, Atf1/Pcr1 binding sites have been fully deleted, it is possible that when bound but not in a heterochromatin state, Atf1/Pcr1 acts to encourage more efficient spreading out of *cenH,* possibly by directing the locus to a more spreading competent location or via its recruitment of HDACs.

SUPPLEMENTAL MATERIALS AND METHODS

Basic 3 color HSS Analysis in R:

Reading in the data:

Standard flow cytometry data files (.fcs) were read in with the R package flowCORE (Bioconductor, https://www.bioconductor.org/packages/release/bioc/html/flowCore.html).

Isolating successfully nucleated cells (Fig1&2):

A strain closely matched in genetic background to HSS strains but containing no XFPs was analyzed under the same flow cytometry conditions in each experiment. This "unstained" control was gated for cell size in the same manner as analysis strains and both the median fluorescence and standard deviation determined in green or orange channels (the signal from the *ade6p:*SF-GFP, *ade6p:*mKO2 or *ade6p:*E2C transcriptional units is referred to here as "green", "orange" or "red"). A nucleation cutoff was set for a value corresponding to the median of fluorescence units

in the clamp channel plus two standard deviations from this unstained control. Only cells with signal less than this value were considered for post nucleation analysis.

Normalizing to max fluorescence values from Δclr4 strains (Fig1&2):

Max values in *Δclr4* strains were determined by calculating the median raw fluorescence in each color channel after gating for cell size. For each cell of each strain for analysis, the signal in each channel was divided by this max value for the corresponding *Δclr4* strain.

Normalizing to constitutive red signal:

For each cell of each strain, the "green" and "orange" values were divided by the "red" value. The red- and *Δclr4*-normalized values range from 0 to ~1.5, where 1 corresponds to the *Δclr4* (max) value. As this value is derived from the mean of a cell distribution, with *Δclr4* cells falling above and below the mean, we plotted out to 1.5 to capture cells with ratio values above 1.0.

Hexbin (2-D Histogram) Analysis:

Normalized "green" and "orange" values (without any nucleation cutoff applied) were plotted as hexbin (or 2-D histogram) plot where density is color-coded in grayscale. Data points within x=0-1.5 and y=0-1.5 were isolated and a hexbin plot was generated using n=40-45 bins along each axis. Hexbin plots were generated using the R package hexbin (https://CRAN.R-project.org/package=hexbin)

Spreading Analysis with Nucleation Clamp:

Cells within the FSC/SSC gate and with signal below the cutoff value for the nucleation color were plotted in a 1-D histogram with n=50-200 bins where the points in the middle of the histogram bins were plotted connected by a line.

Modifications for Figure 2.4:

For every TSA concentration and pre-condition (no TSA or 50μM) each strain was normalized to the median fluorescence of a *Δclr4* strain grown under the same treatment. The cytometer settings were adjusted so an unstained control had mean fluorescence in the green channel of 10^2. A "green" cutoff for nucleation was assigned to be 400 fluorescence units. "orange" cutoff values for each analysis strain were generated by determining mean and two standard deviations in PAS217 at 0μM TSA from the no TSA precondition normalized to the appropriate *Δclr4* strain for each analysis strain. Previous analysis demonstrated both colors in PAS217 to be fully repressed, as evident in the mean for each channel. For Figure 2.4B,C we calculated at each TSA concentration the fraction of cells with green signal below the "green" cutoff that have an orange signal below the "orange" cutoff. These values were normalized to the fraction calculated for cells of that strain at 0μM TSA from the no TSA precondition. For Figure 2.4D we calculated at each concentration of TSA the fraction of all cells with orange signal below the "orange" cutoff, because in the 50μM TSA pre-condition, insufficient cells exist that are below the cutoff for "green" to perform above analysis. These values were normalized to the fraction calculated for cells of that strain at 0μM TSA from the no TSA precondition.

In Figure 2.4F, for each strain at each temperature, we calculated the fraction of cells that had "green" signal less than the "green" cutoff AND "orange" signal less than the "orange"

cutoff. These values were normalized to the fraction calculated for cells of that strain at 32°C. The cutoff values were based on calculations of two standard deviations from the mean of a red only strain normalized to *Δclr4* controls. For "green" and "orange" these values were approximately 0.35 and 0.4 respectively so these values were used for all strains to standardize the analysis.

Fitting $t_{1/2}$ for 38°C spreading recovery:

To derive a $t_{1/2}$, which is the time required to recover to 50% of the full spreading observed at 32°C, we fit the data to a simple sigmoidal dose-response variable slope model:

$$\text{fraction all cells with full spreading} = \text{Bottom} + \frac{t^n * (\text{Top} - \text{Bottom})}{t^n + t_{1/2}^{\,n}}$$

where Bottom is the starting fraction of cells with full spreading at t=0, Top =1, t is time in hrs. *n* represents a Hillslope.

In Figure 4 Supplement 2A, we calculated the fraction of cells that had "green" signal less than the "green" cutoff AND "orange" signal less than the "orange" cutoff. The cutoff values were based on two standard deviations from the mean of a red only strain normalized to *Δclr4* controls. For green and orange these values were approximately 0.4 and 0.4 respectively so these values were used for all strains to standardize the analysis. These cells were gated for FSC and SSC to isolate live cells because at the highest temperature many cells had died. In Figure 4 Supplement 2B,C the 1-D histograms for all "green" values were calculated as above with the modification that no size gate was called**.** In Fig4 Supplement 2C-E the 1-D histograms and

hexbin plot insets were calculated as above with the modification that no size gate was called and the "green" cutoff values were generated as in Figure 2.4F.

Modifications for Figure 2.5:

For Figure 2.5B,E for each strain we calculated at each time point the fraction of cells that had "green" signal less than the "green" cutoff AND "orange" signal less than the "orange" cutoff. The cutoff values for "green" and "orange" were each 0.4 as in Figure 2.4 Supplement 2A. Samples in Figure 2.5B,C were gated for size because we were able to measure 10^5 cells as the strains grew well. Samples in Figure 2.5E,F were not gated for size because cell growth was poor in *seb1*-1 isolates. In Figure 2.5C 5000 cells were plotted in a scatter plot with point transparency. In Figure 2.5D ~1400 cells were plotted in a scatter plot with point transparency. For Figure 2.5F ~2500 cells were plotted in a scatter plot with point transparency. Data in Figure 2.5 Supplement A were plotted as in Figure 2.5C and in Fig5 Supplement B as in Figure 2.5F.

FYLM data analysis:

Initial data calculations:

Loss of focus was identified by red fluorescence measurements below a cutoff of one standard deviation from the mean of all collected values of red for all cells. This loss of focus data was removed from analysis. Background fluorescence from the PDMS device at each time point was then subtracted using catch tubes that did not receive a cell. For each MAT strain, its matched *Δclr4* strain was also imaged and a mean and standard deviation were calculated. In each strain cells were normalized to this mean *Δclr4* value (defined as "max") and to their own red values as in the flow cytometry data analysis. An ON gate (used in **Figure 2.3**B) for cells that reached

maximal de-repression was calculated for each strain from the *Δclr4* strain mean less 3 standard deviations.

Calculating nucleation gates:

As seen by flow cytometry and visual inspection of collected movies, the vast majority of MAT locus cells have a repressed nucleation reporter ("green"), which allowed us to formulate a very strict nucleation cutoff from the collected FYLM data itself. This cutoff was the mean plus two standard deviations of all measured values of all cells. Only cells that maintained a green signal less than this cutoff for their entire measured lifespan were included for further analysis in **Figure 2.3**. We did not apply this nucleation gate to the ectopic strain, as few cells maintained "green" tightly repressed throughout their measured lifespan. Instead, we show all the cells in the traces plot and highlight in grey example cells that remain nucleated or mostly nucleated throughout their measure lifespans.

Rescaling orange to fix negative values:

Due to background subtraction (see above) a significant fraction of cells experienced negative values for their adjusted fluorescence in the orange channel. To account for this, the data for the MAT strains in **Figure 2.3** was rescaled by determining the lowest value measured (minVal) and adding the difference between that value and 0 to every time point of every cell. Values for all timepoints were then divided by 1+minVal to rescale back to 1 = max. The ectopic strain was not rescaled.

Data smoothing:

For trace plots and heatmaps data was smoothed using a moving average of the two-nearest neighbor data points before and after. This number was chosen as it represents the timeframe of

one cell division and is on the timescale of the expression and maturation of the XFPs used in these strains.

Traces:

Individual cell traces represent the red normalized and smoothed, green and orange fluorescence data plotted over time. Traces begin and end at whatever time a cell entered or exited the channel or died. Therefore, not all traces start at x = 0 or end at x = 60. Curated example cells were also plotted as overlays using gray lines. For these curated cells similar trace plots for orange divided by green was plotted in Figure 2.3 Supplement 2 C-BOTTOM, D-BOTTOM, and F.

Heatmaps:

Points with red values greater than 50% of the mean were removed. For cells that remained nucleated throughout their measured lifespan, up to 36 hours of measurements of normalized green or orange was plotted as a heat map from blue (0) to yellow (1) for 30 of the longest imaged cells. White gaps indicate transient loss of focus of less than 2 hours (4 time points). Heatmaps were no longer plotted for any cell that had a loss of focus event for more than 4 time points.

X-Y fluorescence line plots:

For the cells in each behavior category, an X-Y plot was generated that plots the unsmoothed, red-normalized "orange" vs "green" values for each cell as a line. Values were normalized to the mean in "green" and "orange" from a matched background *Δclr4* strain to set the value of 1, while background fluorescence from empty channels were set to the value of 0. Line segments are colored from blue to pink generated based on the measurement time of that cell while the behavior was observed. The first measured point is represented in blue, the last in pink, and the color values in between are divided into increments by the total measured time for the cell.

X-Y fluorescence dot plots:

For one or two selected cells per strain an X-Y fluorescence dot plot was generated that plots the smoothed "orange" vs "green" values for every third time point imaged over its entire measured lifespan. Points are colored in a grayscale that is generated based on the measured lifespan of that cell. The first measured point is represented in black, the last in white and the number of remaining points set by the total measured lifespan of the cell. In Figure 2.3 Supplement 1B X-Y fluorescence dot plots were generated for four example category 1 cells. In this figure, time is colored from blue to pink as in the X-Y line plots and points represent 30-minute intervals beginning when both colors are fully expressed and ending when both reach a local minimum.

Cell fate pie charts:

The number of measured time points for each cell was determined and converted to hours (1 fluorescence image every 30 minutes). Cells were then binned into lifetime groups of <12hrs, 12-36hrs, or >36hours. Within these bins the cells were separated based on whether they died as annotated in FYLM Critic or if their traces were cut short due to late entry into or ejection from the catch channel.

3. Local chromatin context dictates the genetic determinants of the heterochromatin spreading reaction.

FOREWORD

The following is a reproduced from a manuscript submitted for publication and publicly available as a preprint (Greenstein et al. 2020). In this chapter, a customized nuclear function deletion library was screened against each of the four heterochromatin genomic contexts generated in the previous chapter and the resulting dataset analyzed for gene silencing and spreading specific defects. This chapter expands upon the findings in Chapter 1 that ascribe different properties of spreading stability to chromatin domains formed by distinct nucleation elements and identifies novel spreading regulators that will be the topic of future investigations.

ABSTRACT

Heterochromatin spreading, the expansion of gene-silencing structures from DNA-encoded nucleation sites, occurs in distinct chromatin contexts. Spreading re-establishes gene-poor constitutive heterochromatin every cell cycle, but also invades gene-rich euchromatin *de novo* to steer fate decisions. Unlike heterochromatin nucleation and assembly, the determinants of the spreading process remain poorly understood. Our heterochromatin spreading sensor separately records nucleation site-proximal, and distal, heterochromatin gene silencing. By screening a nuclear function gene deletion library in fission yeast, we identified regulators that alter the propensity, both positively and negatively, of a nucleation site to spread heterochromatin. Critically, the involvement of many regulators is conditioned by the chromatin context within which spreading occurs. We find spreading, but not nucleation, within constitutive heterochromatin, requires distinct Clr6 histone deacetylase complexes. However, spreading is

universally antagonized by a suite of chromatin remodelers. Our results disentangle the machineries that control lateral heterochromatin spreading from those that instruct DNA-directed assembly.

INTRODUCTION

Cellular specification requires the genome to be partitioned into regions of activity and inactivity such that only genes appropriate for a given cellular state are available for expression. This requires the formation and propagation, in time and space, of gene-repressive heterochromatin structures. Heterochromatin is most commonly seeded by DNA-directed nucleation (Hall et al. 2002; Reyes-Turcu et al. 2011), and then propagates across the chromosome by a DNA-sequence indifferent process termed spreading, to repress genes expression in the underlying regions. Silencing is instructed by assembly factors, such as HP1, that recognize heterochromatic chemical modifications (Lachner et al. 2001; Jacobs and Khorasanizadeh 2002). The spreading of silencing structures occurs in very different chromatin contexts, which may intrinsically promote or antagonize this process: (1) Constitutive heterochromatin, which is generally gene-poor and as such depleted of activities associated with active genes known to antagonize heterochromatin (Scott et al. 2006; Greenstein et al. 2019). Its maintenance through replication is aided by the inheritance of nucleosomes bearing heterochromatic marks (Alabert et al. 2015). This inheritance promotes modification of nearby nucleosomes due to the "read-write" positive feedback intrinsic to heterochromatin histone modifiers (Zhang et al. 2008a; Al-Sady et al. 2013; Ragunathan et al. 2015). (2) Conversely, during differentiation, heterochromatin is either seeded at new nucleation sites or invades gene-rich euchromatin *de-novo* from existing nucleation sites (Wen et al. 2009; Zhu et al. 2013; Zylicz et al. 2015; Zylicz et al. 2018; Nicetto and Zaret 2019).

In either case, it encounters active chromatin that lacks repressive marks and can specifically antagonize heterochromatin (Greenstein et al. 2019). Thus, during this initial invasion process, heterochromatin spreading cannot benefit from the inheritance through replication of pre-existing marked nucleosomes. Beyond the differences between active and inactive chromatin, it remains unclear whether distinct nucleation elements require different regulators to enact efficient spreading outward from those sites. Nonetheless, we recently described that distinct nucleation elements trigger divergent types of spreading behavior (Greenstein et al. 2018).

Over the past four decades, forward and reverse genetic screens in fission yeast have established an exhaustive list of factors required for heterochromatin silencing and heterochromatin nucleation. Some of these nucleation mechanisms include repeat sequences that instruct specialized RNAi-machinery to process noncoding (nc) RNAs involved in targeting the histone methyl transferase Clr4 (Moazed 2009); signals within nascent transcripts that trigger heterochromatin island formation (Zofall et al. 2012); and shelterin-dependent pathways (Wang et al. 2016; Zofall et al. 2016). However, those genetic screens have neither directly addressed factors that specifically regulate the spreading process, nor whether such factors act in general or context-specific settings. With our previously established fluorescent reporter-based heterochromatin spreading sensor (HSS) we can segregate the central output of heterochromatin (gene silencing) from the spatial control of the reaction (spreading) (Greenstein et al. 2018; Greenstein et al. 2019). This allows us to address the following questions: (1) Are there known or novel regulators of heterochromatin that primarily regulate spreading, versus nucleation? (2) Does spreading over chromatin with distinct characteristics, such as gene density or nucleosome arrangement, require different sets of regulators? (3) Does the set of regulators required for

efficient spreading depend on the type of nucleator that seeds it - for example nucleators using transcription and ncRNA pathways versus direct tethering of heterochromatic factors?

To address these questions, we conducted a series of reverse genetic screens in fission yeast, using a custom nuclear function gene deletion library in four heterochromatin contexts. These include derivations of the fission yeast mating type (MAT) locus, a gene-poor constitutive heterochromatin region contained by *IR-L* and *IR-R* boundaries and nucleated by at least two elements: (1) *cenH,* homologous to pericentromeric *dh* and *dg* elements, which rely on ncRNA pathways, including RNAi, and (2) *REIII*, a sequence element that directly recruits heterochromatin factors via the stress-response transcription factors Atf1 and Pcr1 (Jia et al. 2004; Kim et al. 2004). We also queried an ectopic heterochromatin domain that is embedded in gene rich euchromatin. This domain is nucleated by an ectopically inserted *dh* element fragment proximal to the *ura4* locus (Canzio et al. 2013; Marina et al. 2013; Greenstein et al. 2018).

We find that the genetic requirements for promotion and containment of heterochromatin spreading diverge significantly between different chromatin contexts, and to some degree also between different types of nucleators. However, despite these context-dependent differences for spreading, common themes also emerge from this work: (1) Sub-complexes of the Clr6 histone deacetylase (HDAC) complex, in particular the Fkh2-associated Clr6 complex I″, appear to promote spreading, but not nucleation, at multiple heterochromatin loci. (2) At both euchromatic and heterochromatic loci, spreading is antagonized by a diverse set of nucleosome remodelers, in particular Ino80 and Swr1C. (3) Unexpectedly, members of an AP3 adaptor complex, normally involved in membrane traffic, are moderate positive regulators of spreading in hetero- and euchromatin. Together, these results specifically pinpoint the machineries required to regulate

the spatial expansion of heterochromatin domains independent of the initial seeding by DNA-encoded nucleation sites.

RESULTS

Our previously developed HSS relies on three transcriptionally-encoded fluorescent protein-coding genes that collectively report single cell measurements of heterochromatin formation via flow cytometry, while normalizing for transcriptional and translational noise (Al-Sady et al. 2013; Greenstein et al. 2018). It provides separate, quantitative recordings of nucleation-proximal ("green") and distal ("orange") events at a heterochromatin site over large populations of isogenic cells (typically N >20,000) (**Figure 3.1A**). In contrast to the singular readout employed by auxotrophy-dependent reporter gene silencing assays, the HSS assay provides a "multidimensional snapshot" that distinguishes changes in heterochromatin nucleation and spreading, and additionally permits tracking of emerging multimodal cell populations and unique population distributions. This is an important conceptual advance compared to traditional methods.

With this ability to observe the entire population distribution pattern from any isogenic background, we sought first to investigate the requirements for heterochromatin activity for both nucleation and distal spreading, within either constitutive heterochromatin or gene-rich euchromatic regions. Further, we queried whether different types of heterochromatic nucleation sequences utilize similar or different sets of regulators to nucleate and/or spread the resulting heterochromatin structures outwards. To address the latter, we explored three different chromatin contexts, derived from the constitutive heterochromatic mating type (MAT) locus, each containing an embedded HSS (**Figure 3.1A**) (Greenstein et al. 2018): wild type, with the *cenH*

and *REIII* nucleating DNA elements uncompromised, and two *MAT* variants that contained mutations in either the *cenH* or *REIII* elements (**Figure 3.1A**). Mutations in these DNA elements limit initiation of heterochromatin spreading to one nucleator (Greenstein et al. 2018). To address differences in chromatin context in addition to variants of the constitutive MAT locus, we examined heterochromatin formation at the euchromatic *ura4* locus*,* where heterochromatin spreading is ectopically driven by the upstream insertion of a pericentromeric *dh* DNA element (Marina et al. 2013; Greenstein et al. 2018). We refer to this chromatin context as *ECT* (<u>ect</u>opic). When analyzed by flow cytometry, *WT MAT* and *MAT ΔcenH* populations appear largely fully nucleated with near-complete local spreading, as evidenced by population density in the bottom left in the 2D density histogram (**Figure 3.1C**,**D** and (Greenstein et al. 2018)). *MAT ΔREIII* and *ECT* cell populations, while mostly nucleated, display a stochastic distribution of spreading states, evidenced by a vertical distribution on the left of the 2D density histogram (**Figure 3.1E,F** and (Greenstein et al. 2018)). The distribution of cells in a strong loss of silencing mutant, *Δclr3,* is shown for all chromatin contexts next to the wild-type parents (**Figure 3.1C-F**). To identify potential regulators of the spreading reaction, we conducted a genetic screen by crossing a deletion library of ~400 nuclear function genes (**Figure 3.1 Supplement 1**) to reporter strains with the HSS integrated in the four chromatin contexts described above. We quantified the fluorescence signal from the three reporters in each mutant by flow cytometry, with approximately 20k-30k cells per mutant, depending on growth conditions.
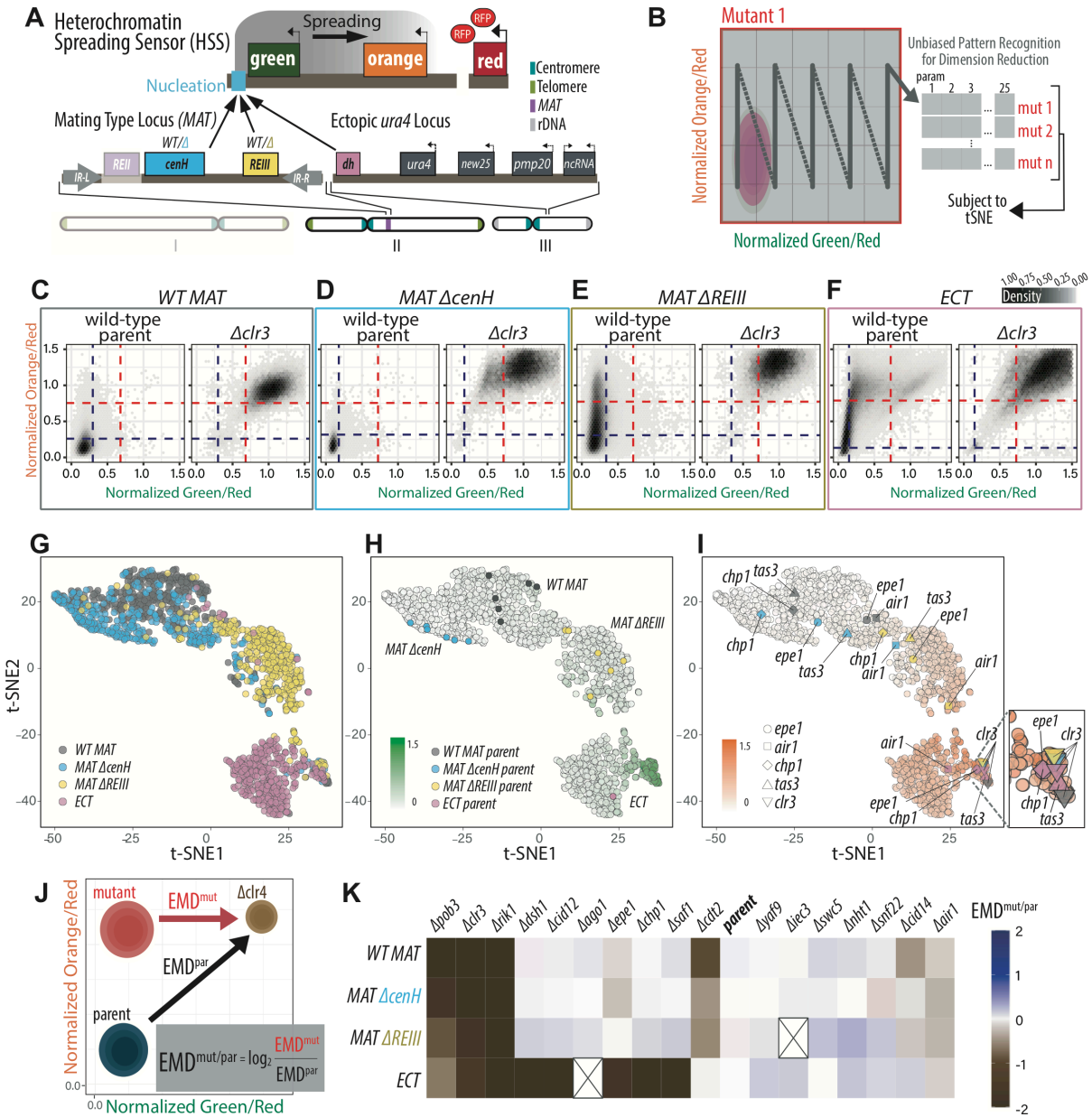
Figure 3.1: A genetic screen based on a suite of fluorescent reporters identifies context-dependent positive and negative regulators of heterochromatin function.

A. TOP: Overview of heterochromatin spreading sensor (HSS, (Greenstein et al. 2018)). Three transcriptionally encoded fluorescent protein genes are integrated in the genome. *SFGFP* ("green") proximal to the nucleation site allows identification of heterochromatin nucleation; *mKO2* ("orange") distal to the nucleation site allows identification of heterochromatin spreading. *3xE2C* ("red") in a euchromatin region normalizes cell-to-cell noise. BOTTOM: The endogenous mating type locus (*MAT*) and ectopically heterochromatic *ura4* locus (Greenstein et al. 2018) were examined with the HSS in the screen. *Bona fide* mutations of the nucleators, *cenH* and *REIII*, in *MAT* were made to limit nucleation to occur from one site. **B.** Overview of data processing for t-SNE. The multidimensional fluorescence data is linearized before subjection to

t-SNE. Two-dimensional "orange"/"red" *v.* "green"/"red" density plots are broken down into 25 grids in an unbiased manner, and normalized cell counts of each grid are used as parameters for each mutant in the t-SNE processing. The 25 parameters of all mutants from all chromatin contexts were embedded into a model two-dimensional t-SNE space. **C.-F.** 2D-density hexbin plots of the wild-type parent or *Δclr3* mutant in the (**C**) *WT MAT,* (**D**) *MAT ΔcenH,* (**E**) *MAT ΔREIII,* and (**F**) *ECT* background. Hexagonal bins are colored from light grey to black indicating low to high density of cells per bin. Blue lines indicate boundary guidelines for the fully repressed state and red lines indicate boundary guidelines values for fully expressed state (see methods for treatment of MAT and ECT strains) **G-I.** t-SNE visualization of all mutants across all four chromatin contexts. Each data point represents a mutant, the fill color represents the (**G**) chromatin context of the mutant, or median (**H**) "green" or (**I**) "orange" fluorescence of the entire mutant population. In (**H**) the parent isolates of each background are depicted in individual colors indicated in the key. In (**I**), selected mutants are shown, with the chromatin contexts highlighted with the same colors as (C-F). An enlarged region to highlight the *Δclr3* mutants is shown to the right. **J.** To linearize the multidimensional fluorescence data, the earth mover's distance (EMD) between each mutant and a *Δclr4* mutant is calculated. EMD$^{mut/par}$ is computed by dividing each mutant EMD (EMD^{mut}) by the respective parent EMD (EMD^{par}) and transforming the quotient by log2. **K.** Heatmaps depicting EMD$^{mut/par}$ of indicated mutants in each chromatin context. Any values less than -2 were converted to -2 and interpreted as such. Crossed-out boxes indicate mutants excluded from the analysis due to growth defect or low sample size.

To visualize the pattern of heterochromatin domain activity across the isogenic mutant populations, we first employed t-distributed stochastic neighbor embedding (t-SNE). For all the mutants in each chromatin context, we converted the normalized 2D distribution of cells into a numerical vector representing cell density in 25 isometric plot regions as different parameters (**Figure 3.1B**). The t-SNE model was built by calculating the similarity between different cell distributions and comparing these 25 parameters from all mutants and wild type controls in the four chromatin contexts (**Figure 3.1B**).

We plotted the t-SNE model with all the four chromatin contexts and respective mutants, coloring each mutant by its parental chromatin context (**Figure 3.1G**). We also colored the mutants by the median nucleation ("green", **Figure 3.1H**) or spreading ("orange", **Figure 3.1I**) reporter expression values. As expected, the majority of mutants did not strongly deviate from their parent and broadly clustered together into a "neighborhood" by chromatin context. This is evident for *ECT*, for which the parent strain has "green" and "orange" in a less repressed state,

particularly compared to *WT MAT* and *MAT ΔcenH*. Within each neighborhood, the distribution

of "orange" expression, especially for *MAT ΔREIII* and *ECT*, is graded from above to below the

expression level of the parent(s), revealing a continuum of mutants with enhanced or abrogated

spreading. We could not find mutants that display more repression than the parental strains of

MAT ΔcenH, which are highly repressed in the OFF state, as previously described (Grewal et al.

1998; Greenstein et al. 2018). However, we did observe mutants located out of the area of their

chromatin context-driven "neighborhood". First, the major known heterochromatin mutants,

Δclr4, Δswi6, Δclr3 among others, from each chromatin context formed a cluster with high

expression of "green" and "orange" (**Figure 3.1I** enlarged region, exemplified by *Δclr3*),

segregating from the rest of the population. Second, we observed mutants, such as *Δcdt2, Δepe1*

and *Δchp1,* that segregate out of neighborhood only for selected chromatin contexts, indicating

specificity (highlighted in **Figure 3.1I**). The t-SNE analysis visualized the relationship of all four

chromatin contexts, and mutants therein, with respect to their nucleation and spreading behavior,

directly revealing the graded nature of mutant phenotypes. This is particularly the case with

respect to spreading in *ECT* and *MAT ΔREIII* neighborhoods (**Figure 3.1I**)

However, in the t-SNE analysis, the phenotype patterns are weighted by the intrinsic

behavior of each parent's chromatin context. To quantify how much each mutation impacted the

heterochromatin state in each chromatin context, we performed Earth Mover's Distance analysis

(EMD, **Figure 3.1J** see also materials and methods and (Orlova et al. 2016)). We express the

contribution of each mutant relative to the parental isolates by a quotient of their respective

EMDs to *Δclr4*, which is completely deficient in heterochromatin assembly and serves as a fixed

reference point for all chromatin contexts. We represent a subset of the mutants' EMD values in

a heatmap (**Figure 3.1K**). Some mutants contributed similarly in different chromatin contexts,

such as *Δclr3* and *Δrik1,* which display strong de-repression, and *Δair1,* which displays an

intermediate de-repression (Thon and Friis 1997; Keller et al. 2010). However, many mutants

had differential contributions in each chromatin context. As an example, *Δpob3,* a mutation in

the FACT complex (Lejeune et al. 2007), had weaker phenotypes where nucleation is solely

ncRNA-driven (*MAT ΔREIII* and *ECT*) and stronger phenotypes when *REIII,* an ncRNA-

independent element, is intact (*WT MAT* and *MAT ΔcenH*) (**Figure 3.1K**). Interestingly, the

TRAMP subunit mutant *Δcid14*, had more pronounced roles in *MAT* contexts over *ECT* (**Figure

3.1K**). Beyond these subtler differences, we noticed a large divergence between *ECT* and other

contexts. This is evident from very low pairwise correlations of EMD$^{mut/par}$ values between all the

mutants for *ECT* against the *MAT* contexts **(Figure 3.1 Supplement 2C,E,F)**. Conversely, we

observe much higher correlations between all pairwise combinations of the three *MAT* chromatin

contexts (Figure 3.1 Supplement 2**A,B,D**). The divergence between *MAT* and *ECT* contexts is

especially true for the RNAi pathway. *Δchp1* and *Δtas3* (members of RITS complex), *Δdsh1*

(RNAi factor)*, Δcid12* (member of RDRC complex)*,* and *Δsaf1* (RNAi associated splicing factor,

(Bayne et al. 2014)) all have very strong effects only in *ECT*.

The above results are consistent with previous reports of *WT MAT* and ectopic reporters

(Hall et al. 2002), which are respectively independent and dependent on RNAi for

heterochromatin maintenance. Surprisingly, *MAT ΔREIII,* which similar to *ECT,* solely relies on

a *dh*-homologous *cenH* element for H3K9me nucleation, also behaved independent of RITS for

maintenance (**Figure 3.1K**). *MAT* has another H3K9me-independent silencing element, *REII*

(Hansen et al. 2011), which is located upstream of "green". However, this element appears to act

quite locally (Ayoub et al. 2000; Hansen et al. 2011). Moreover, we showed removal of *REII* had

no effect on expression of HSS reporters in *MAT ΔcenH* (Greenstein et al. 2018). Thus, we do not believe *REII* accounts for the difference observed between *MAT ΔREIII* and *ECT*.

Similar to RNAi factors, the heterochromatin-antagonizing *epe1* pathway had strikingly different impacts on the *MAT* chromatin contexts compared to *ECT* (**Figure 3.1G**). *Δepe1* results in destabilized heterochromatin at *ECT,* but the deletion has little effect at *MAT* chromatin contexts. Heterochromatin in these contexts instead is sensitive to mutations in the Epe1-turnover pathway, exemplified by *cdt2*, the specificity factor for the Epe1-targeting E3 ubiquitin ligase (Braun et al. 2011). In contrast to loss of silencing mutants, we also identified mutations that induce hyper-repressed states (EMD$^{mut/par}$ >1) in multiple chromatin contexts with some differences in contribution to hyper-repression. These included *Δyap9, Δsnf22, Δswc5* and *Δiec3* (see below).

Overall, the EMD analysis enabled us to systematically and quantitatively discern how each mutant contributes to the combined nucleation and spreading heterochromatin state within each chromatin context. The results suggest that the genetic circuitry for heterochromatin maintenance differs significantly between naïve sites in euchromatin and constitutive loci. This is consistent with the tightly repressed *MAT ΔcenH* or the redundantly nucleated *WT MAT*. But the fact that *MAT ΔREIII* also diverged strongly from *ECT* in its dependence on genes for heterochromatin nucleation and spreading was more surprising, because both contexts are driven exclusively by a similar ncRNA nucleators, *cenH* and *dh* (Hansen et al. 2006), and display stochastic spreading over similar ranges (**Figure 3.1C, F**). Further we previously showed that they both exhibit dynamically unstable heterochromatin over time (Greenstein et al. 2018). The finding that even heterochromatin domains with very similar nucleation and spreading dynamics

rely on different regulators for their maintenance, suggests that such context-specific

requirements for heterochromatin assembly are likely common across the genome.

The above analyses addressed the genetic requirements for heterochromatin silencing,

including both nucleation and distal spreading, in different chromatin contexts. Next, we aimed

to identify regulators that are specific to heterochromatin spreading but function independent of

nucleation. To do so, we employed the capability of the HSS to segregate those activities. We

isolated cell populations that reside within a "green"-off gate which represents cells with

heterochromatin fully assembled at the nucleation site (see methods, and (Greenstein et al.

2018)). To quantify increased or decreased spreading in a given mutant, we calculated a

$\text{Grid}_n^{mut/par}$ metric (described in methods), which tracks the changes of cell distributions in

"orange" expression within the "green"-off gate (**Figure 3.2A**). To isolate gain or loss of

spreading mutants (hits) for further analysis, we only considered the top 15% of mutants in

which $\text{Grid}_n^{mut/par}$ values were also above 2 standard deviations from the mean of the replicate

parent isolates. (**Figure 3.2 Supplement 1**). With these gene hits isolated, we proceeded to

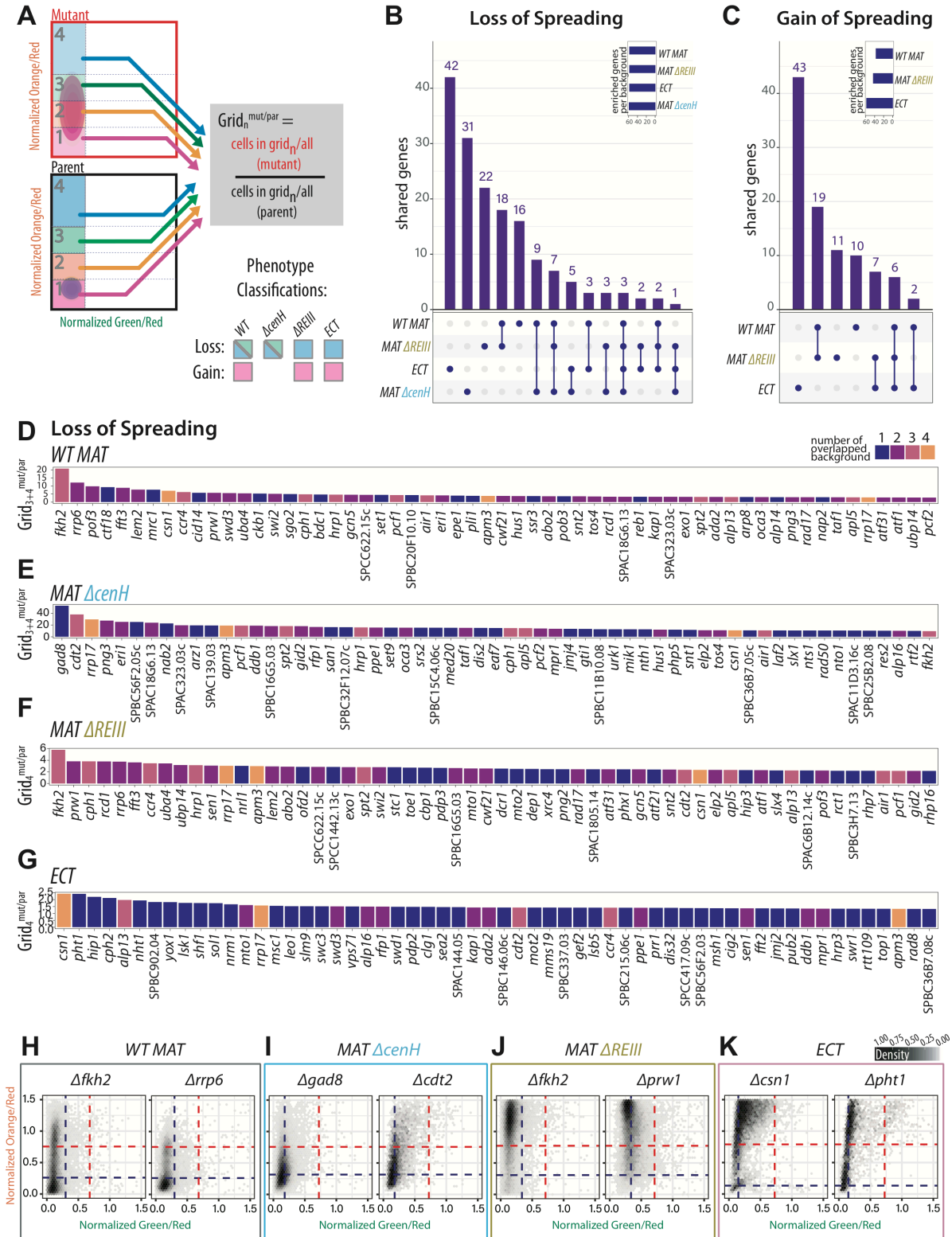analyze their relationships within and across chromatin contexts.

Figure 3.2: Identification of heterochromatin spreading regulators in different chromatin contexts.

A. Overview of the spreading-specific analysis with mock distributions of cells and grids indicated. To segregate spreading from nucleation or silencing phenotypes, "green"-off populations (successful nucleation events) are isolated. Within these populations, enrichment of cell populations in particular "orange" fluorescence ranges ($Grid_n$) are calculated as $Grid_n^{mut/par}$. *E.g.* to identify loss of spreading mutants in *WT MAT*, $Grid_{3+4}^{mut/par}$ is calculated as percentage of mutant population divided by percentage of parent population in $Grid_{3+4}$ (green, blue). The Grids used for analysis of gain and loss of spreading in the four chromatin contexts are indicated. $Grid_{3+4}$ was used for *WT MAT* and *MAT ΔcenH,* given the highly repressed nature of this chromatin context. **B.-C.** Upset plots indicating the frequency of (**B**) Loss of Spreading, or (**C**) Gain of Spreading gene hits appearing in one or multiple chromatin contexts. For each bar, the chromatin context(s) with shared phenotypes for the underlying gene hits is indicated below the plot. The inset indicates the total number gene hits in each chromatin context of the same phenotype. **D.-G.** Bar graphs representing the $Grid_n^{mut/par}$ and number of chromatin context(s) of gene hits with Loss of Spreading phenotype from (**D**) *WT MAT*, (**E**) *MAT ΔcenH,* (**F**) *MAT ΔREIII* and (**G**) *ECT* respectively. The genes are ranked in descending order of $Grid_n^{mut/par}$ and the color of each bar represents the number of backgrounds the mutant show the same phenotype. **H.-I.** 2D-density hexbin plots of the top two loss of spreading gene hits for all 4 chromatin contexts, (**H**) *WT MAT,* (**I**) *MAT ΔcenH,* (**J**) *MAT ΔREIII,* and (**K**) *ECT*. Plotting as in Figure 3.1, C-F. For wild-type comparison, see Figure 3.1 C-F.

We first examined the degree to which spreading modulators are shared between chromatin contexts via upset plots (**Figure 3.2B&C**). Conceptually similar to a Venn diagram, this analysis allows rapid visualization of the degree of overlap between sets, with the number of shared genes plotted as a bar graph and the sets each bar represents annotated below the plot. The upset plot for loss of spreading phenotypes (*i.e.* genes that promote spreading, **Figure 3.2B**) showed that exceedingly few genes, three out of 164 unique genes found as hits, are shared across all chromatin contexts (*csn1, rrp17, apm3*). This result emphasizes the specific impact that each chromatin context has on heterochromatin spreading. Seven genes were shared across all the MAT locus chromatin contexts. In contrast, 111 genes contributed only to one chromatin context. The degree to which genes contributed positively towards spreading, and the degree of overlap across chromatin contexts is shown in bar-graphs of $Grid_n^{mut/par}$ in **Figure 3.2D-G**. We additionally show the 2D density histogram of the screen mutants for the top two loss of spreading hits for each chromatin context in **Figure 3.2H-K**. The top two hits for *MAT ΔcenH* were *gad8* and *cdt2* (**Figure 3.2E&I**) The significant contribution of *cdt2* suggests that

Epe1specifically limits spreading in *MAT ΔcenH* compared to other chromatin contexts.

Spreading in *WT MAT* and *MATΔREIII* was most dependent on *fkh2,* while the second strongest

hits were *rrp6* and *prw1,* respectively (**Figure 3.2D,F,H,J** and see below). Prw1 and Fkh2 form

parts of Clr6 complexes (see below), while Rrp6 is a central part of the exosome, required for

gene silencing in multiple ncRNA-dependent pathways (Bühler et al. 2007). The top hit in *ECT*

was the Csn1 COP9 signalosome subunit, which has been implicated in neddylation of cullin-

based E3 ligases and may operate in a similar pathway as Cdt2 (Bayne et al. 2014). *ECT* was

also highly sensitive to *Δpht1* (**Figure 3.2G**) which codes for H2A.Z, normally associated with

antagonizing spreading (Meneghini et al. 2003).

We also observed a significant number of genes that showed gain of spreading in *WT*

MAT, MAT ΔREIII and *ECT* (**Figure 3.2C**). We could not examine *MAT ΔcenH* for this

phenotype because this chromatin context is highly repressed in the OFF state as reported

previously (Grewal and Klar 1996; Greenstein et al. 2018). Six out of 98 genes found as hits are

shared across these three chromatin contexts (**Figure 3.2C**, **Figure 3.2 Supplement 2**). The 2D

density histograms of the screen mutants are shown for top two hits per chromatin context

(**Figure 3.2 Supplement 2D-F**). Validating the approach, we found *leo1,* a gene previously

impacted in spreading control across boundaries (Verrier et al. 2015), as a moderate gain of

spreading hit. *ECT* displayed the largest fraction, 43 out of 58 of spreading-antagonizing genes

that are unique to one context. In contrast, we found 11 out of 43 genes unique to *MAT ΔREIII*,

which displays a very similar spatio-temporal spreading behavior to *ECT*. This likely reflects that

even though heterochromatin can assemble in a euchromatin context, spreading is under multiple

layers of constraint in this setting.

To understand which nuclear pathways, as opposed to individual genes, direct or antagonize heterochromatin spreading across backgrounds, we performed a protein complex-level analysis. Using the Gene Ontology (GO) protein complex annotations from Pombase (Lock et al. 2019), we annotated the GO complex membership of each screen hit and tabulated the frequency ("counts") of each GO complex per background in both loss of spreading (loss) and gain of spreading (gain) categories. Using these GO complex counts, we generated a heatmap for all the protein complexes represented by hits in our screen (**Figure 3.3A**). In order to assess the similarities between chromatin contexts (*WT MAT, MAT ΔcenH, MAT ΔREIII, ECT*) and categories (loss or gain of spreading) we performed unsupervised clustering on the heatmap columns. Broadly, the hit categories (loss or gain of spreading) clustered together, (**Figure 3.3A**, bottom), revealing relative similarity at the GO complex level. The exception however was the *ECT* "loss" set of hits. Consistent with our gene-level analysis (**Figure 3.2B-G, Figure 3.2 Supplement 2**), complexes that contribute most strongly to "loss" and "gain" categories are different in *ECT* compared to the *MAT* variants.

Overall, we identified three common trends within and between categories: (1) a role for antagonizing spreading by chromatin remodelers, (2) a role of histone deacetylase complexes (HDACs), in particular Clr6 sub-complexes, in promoting spreading, and (3) a role for a small number of additional known and novel spreading regulators, including the AP-3 adaptor and COP9 signalosome complex, which promoted spreading across all backgrounds.
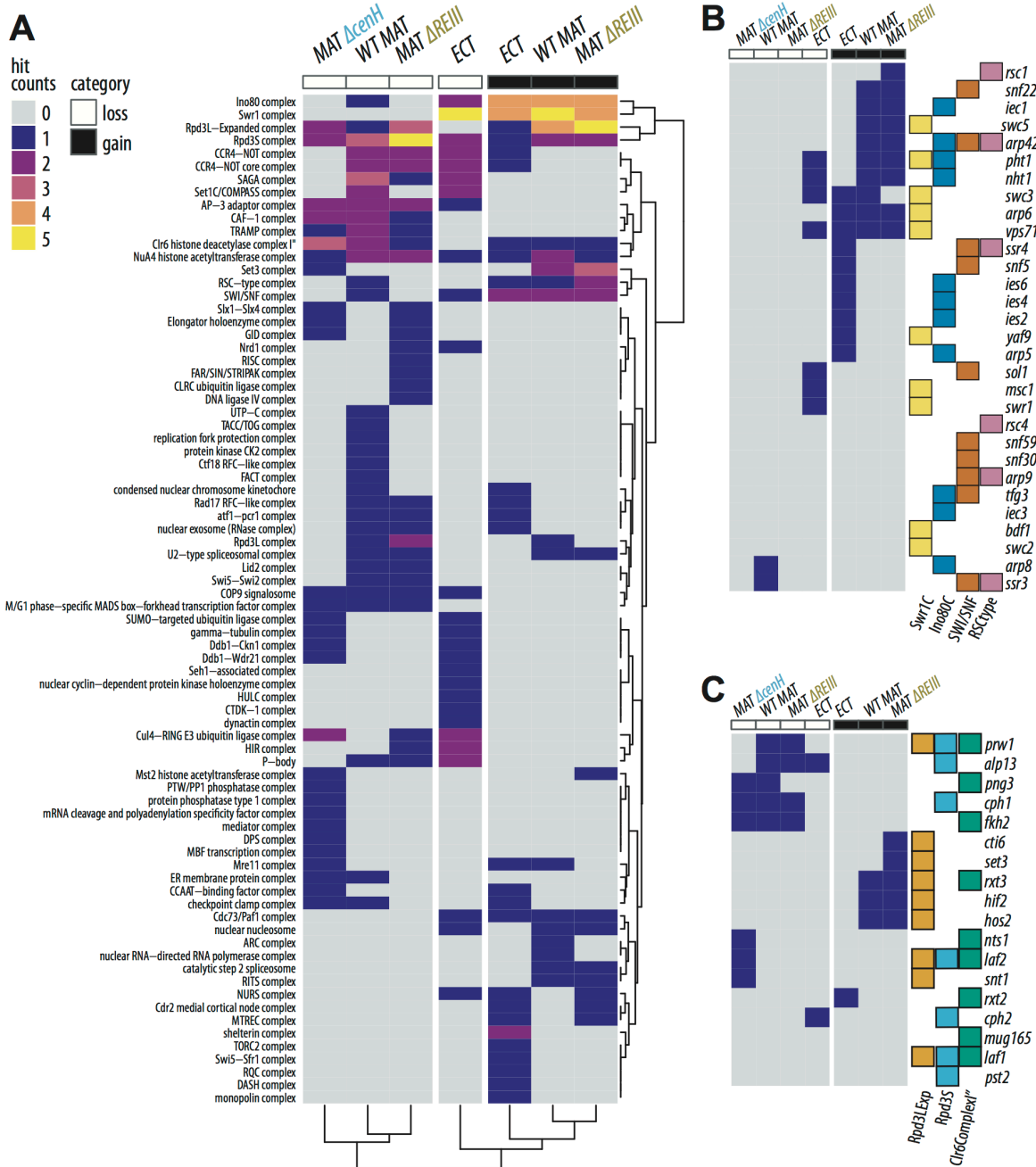
Figure 3.3: Heterochromatin spreading is regulated by sets of unique and common protein complexes across different chromatin contexts.

A. Heatmap of GO complex annotations for hits in each category and strain. Rows, representing GO complexes annotated to genes within the screen that were identified as hits, are arranged via hierarchical clustering. Columns are defined by the hit category (loss of spreading – white; gain of spreading – black), and each screen chromatin context is indicated at the top. The columns were clustered by hierarchical clustering and the tree was cut to define 3 clusters. **B.** Hit table of complex members for Swr1C, Ino80, SWI/SNF, and RSC-type complexes. Components identified as a hit in either the "gain" or "loss" category for each background are

marked blue. Columns are defined and ordered as in (**A**). The proteins present in each complex are annotated at the right with the presence of color indicating membership of that protein in the complex. **C.** Hit table of complex members for Rpd3L-Expanded, Rpd3S, and Clr6 complex I" as in (**B**). Columns are defined and ordered as in (**A**) Proteins present in each complex are annotated as in (**B**).

As evidenced by the heatmap (**Figure 3.3A, top right**), chromatin remodeling complexes are strongly represented in the gain of spreading hit category, including the Swr1C, Ino80, SWI/SNF, and RSC-type complexes. To explore this further, we assessed which protein components were contributing to these GO complex counts. For all the genes annotated to a given chromatin remodeling complex and present in our screen, we displayed whether they were identified as a hit (blue) or not (grey) (**Figure 3.3B**). Indeed, we found that the large majority of the gene hits annotated fall within the "gain" category across backgrounds. The manner in which the cell distributions change in these screen hits is evident from the 2D density histograms (**Figure 3.3 Supplement 1**). While complex-specific heterochromatin antagonizing activities have previously been ascribed to chromatin remodelers, such as Swr1C (Meneghini et al. 2003)and Ino80 (Xue et al. 2015), we observe a broad involvement across remodelers. The specific contributions of these remodeling complexes to destabilizing heterochromatin spreading are interesting candidates for future studies.

Most of the GO complexes we found to be required for spreading have been implicated in chromatin regulation or silencing pathways specifically (**Figure 3.3A**). One notable exception was two predicted subunits of the AP-3 adaptor complex, Apm3 and Apl5, with roles either in all chromatin contexts (Apm3) or only in *MAT* contexts (Apl5). We were intrigued by the discovery of an AP-3 adaptor protein complex normally associated with vesicular traffic. While the spreading phenotype is moderate, *apm3* is one of only three genes that display a loss of spreading phenotype across all contexts, along with *csn1* and *rrp17* encoding a key member of

the COP9 signalosome and a putative rRNA exonuclease, respectively. Thus, we sought to examine whether this phenotype correlates with a H3K9me2 spreading defect and assessed H3K9me2 levels at the constitutive and facultative heterochromatin loci by chromatin immunoprecipitation followed by qPCR (ChIP-qPCR). Consistent with its moderate phenotype, *Δapm3* had mildly reduced H3K9me2 within the MAT locus, but not at pericentromeres and subtelomeres (**Figure 3.3 Supplement 2A-C**). Importantly we find a role for Apm3 in H3K9me2 accumulation at heterochromatin islands, including *mei4, ssn4* and *mcp7* (**Figure 3.3 Supplement 2D**), indicating a role in facultative heterochromatin. Since AP-3 adaptor proteins are cytoplasmic, we tested whether a fraction of Apm3 and Apl5 is also targeted to the nucleus. We expressed Apm3 and Apl5 as fusions with SF-GFP and visualized them together with Swi6:E2C, which marks heterochromatin foci in the nucleus (**Figure 3.3 Supplement 2E**). Notably, we found that Apm3:SF-GFP is distributed broadly in both the cytosol and nucleus, whereas the related Apl5 protein appeared to be excluded from the nucleus (**Figure 3.3 Supplement 2F**). Together, these data indicate that Apm3 may represent a novel regulator of heterochromatin spreading, which is also supported by the finding that it physically interacts with the heterochromatin regulators Fft3 (Lee et al. 2017) and Epe1 (Wang et al. 2013).

The relationship between histone deacetylation and gene silencing is well described. Three classes of HDACs exist, which have partially redundant and non-overlapping functions in the formation of heterochromatin domains and gene silencing. Sir2 belongs to the class III HDAC of the sirtuin family (Shankaranarayana et al. 2003). Clr3 belongs to class II and is a member of the SHREC complex (Sugiyama et al. 2007). Clr6 belongs to class I and is part of several sub-complexes, contributing to both heterochromatic and euchromatic gene regulation (Grewal et al. 1998; Nicolas et al. 2007). We find here that unlike class II and III HDACs, sub-

complexes of the Clr6 family, including the Rpd3S, Rpd3L-Expanded, and Clr6 I″, contribute exclusively to spreading and not nucleation (**Figure 3.1, Figure 3.2** and **Figure 3.3A**, top of heatmap). In particular, we noticed that the forkhead transcription factor Fkh2, was identified as a common spreading regulator in all *MAT* HSS strains. Despite not being formally annotated to the Clr6 I″ complex by GO terms, Fkh2 was previously described as a member of this sub-complex (Zilio et al. 2014). We included Fkh2 as a member of Clr6 I″ in further analysis for this reason. Analogous to the analysis for nucleosome remodeling complexes, we represented the HDAC components that were identified as hits (**Figure 3.3C**). The population distributions of HSS reporter fluorescence in these Clr6 complex mutants are shown in **Figure 3.3 Supplement 2**. We find an interesting bifurcation in the contribution of Clr6 sub-complexes towards "loss" or "gain" categories. Clr6 I″ and Clr6S (Complex II) positively contributed to spreading ("loss"), while several members of the Rpd3L-Expanded complex antagonized spreading and were found as hits in the "gain" category. This includes a subset belonging to the Set3 Complex (Set3, Hif2, Hos2, Snt1). Overall, this suggests that Clr6 I″ and Clr6S HDAC complexes specifically promote heterochromatin spreading in addition to their described roles in transcriptional gene silencing.

Given the strong initial phenotype of Crl6 I″ subunits in spreading (**Figure 3.2**), we further explored their role here. Amongst the known members of Clr6 sub-complexes (**Figure 3.4A**), we chose three subunits to validate our results: two Clr6 I″-specific members, Fkh2 and Png3, and the shared Clr6 core subunit Prw1. We validated the phenotypes by *de novo* single and double deletions for these genes in the three *MAT* HSS reporter backgrounds (**Figure 3.4B-C, Figure 3.4 Supplement 1A,B**) and performed three-color flow cytometry to record fluorescence.

We first sought to validate the phenotype of these deletions in the *WT MAT* and *MAT ΔREIII* HSS*,* which were most prominent in the screen. We found a similar phenotype for the single *Δprw1* and the *Δfkh2 Δprw1* double mutant at *MAT ΔREIII,* corroborating the notion that Fkh2 is a Clr6 component and acts in the same pathway as Prw1 (**Figure 3.4C**, bottom panels). However, we noted that the phenotype of *Δprw1* was weaker than *Δfkh2* in the *WT MAT* background; nonetheless, also here the double mutant showed a non-additive phenotype (**Figure 3.4C**, top panels). We additionally aimed to validate these three deletions in the *MAT ΔcenH* HSS reporter background (**Figure 3.4 Supplement 1C**). While this strain remains strongly repressed in the majority of the population, we detected a noticeable increase in reporter signal in *Δfkh2* and *Δprw1*.
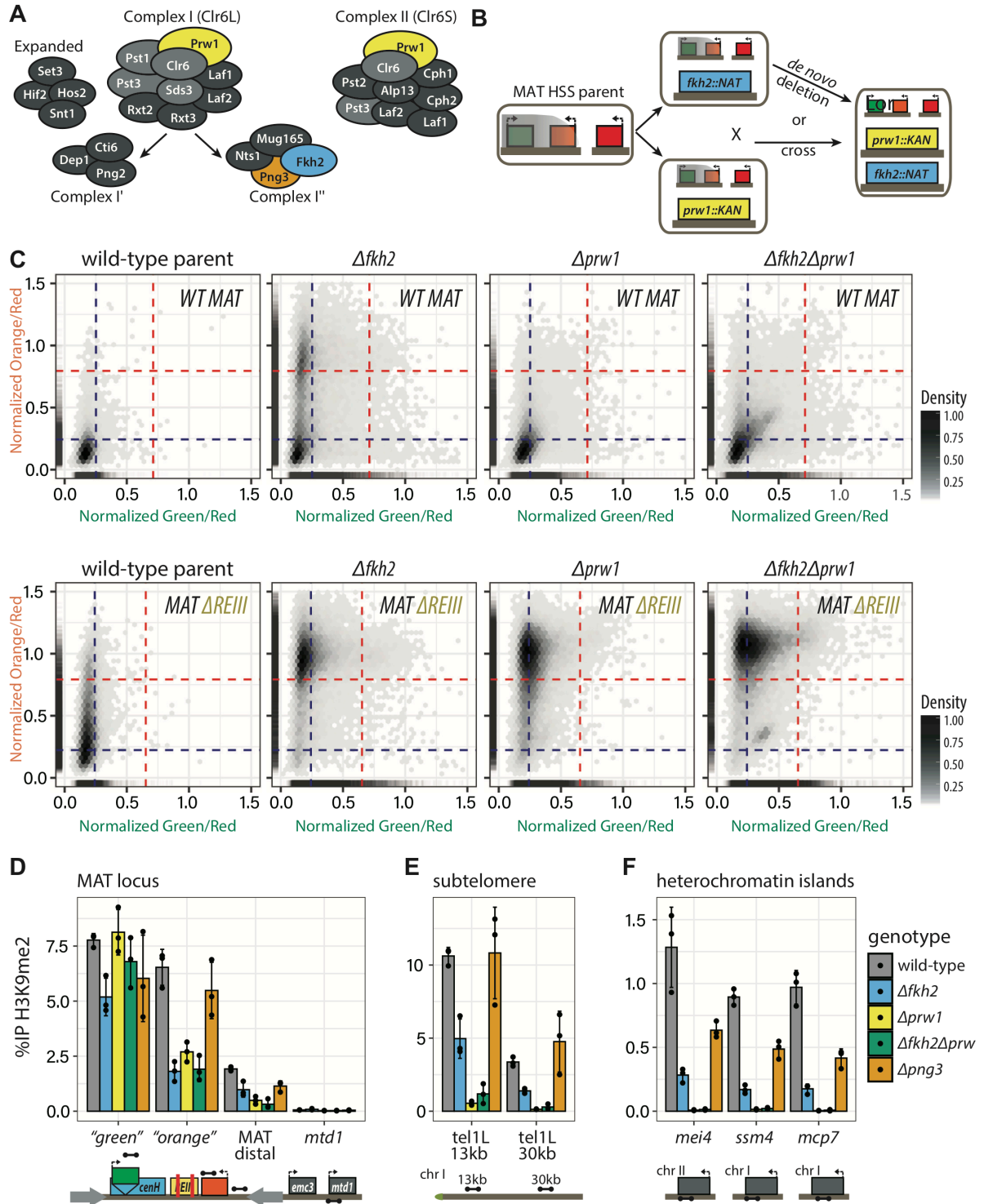
Figure 3.4: Clr6 Complex I" regulates heterochromatin spreading at constitutive and facultative heterochromatin loci.

A. Cartoon of Clr6 complexes (Clr6L, the Clr6LExpanded modules, Clr6 subcomplex 1", and Clr6S). Subunits not in the screen are shown in light grey. **B.** Scheme for generation of deletion

mutants. Double mutant of fkh2 and prw1 was generated by both cross and *de novo* deletion of *prw1* in *Δfkh2*. **C.** TOP PANELS: 2D-density hexbin plots for WT and Clr6 I″ mutants in the *WT MAT* strain background. Hexagonal bins are colored from light grey to black indicating low to high density of cells per bin. A rug plot is included on the X and Y axes indicating the 1D density for each color. Rug lines are colored with partial transparency to assist with visualization of density changes. Blue lines indicate boundary guidelines for the fully repressed state and red lines indicate boundary guidelines values for fully expressed state as in Figure 3.1. BOTTOM PANELS: 2D-density hexbin plots with rug as above for WT and Clr6 I″ mutants in the *MAT ΔREIII* strain background. **D.-F.** ChIP-qPCR quantification of H3K9me2 signal in the *MAT ΔREIII* strain at constitutive and facultative heterochromatin regions. Error bars represent 1SD from three biological replicate isolates. Individual values are plotted for each isolate. The WT data is additionally replicated in Figure 3.3 Supplement 2 as these experiments were performed together. SD; standard deviation. Telomere 1L primer distances are from the end of the assembled genomic sequence.

The above results, evidencing an impact on silencing at the spreading ("orange"), but not nucleation ("green") reporter, predicted that the chromatin state may also be affected primarily at nucleation-distal sites in Clr6 I″ complex mutants. We therefore next examined the chromatin state by ChIP-qPCR analysis of the H3K9me2 mark at *MATΔREIII* and other heterochromatin loci. First, we examined the MAT locus. Here we observed strong reductions of H3K9me2 signal at the "orange" spreading reporter and also more nucleation-distal targets, in *Δfkh2, Δprw1*, and the *Δfkh2Δprw1* double mutant (**Figure 3.4D**). There is no increase in severity of the spreading defect in the *Δfkh2Δprw1* double mutant relative to the two single mutations, although further decreases in H3K9me2 should have been detectable (see euchromatic target *mtd1,* **Figure 3.4D**). This further supports that Fkh2 and Prw1 act together to promote spreading. However, consistent with the flow data, we did not observe strong H3K9me2 reductions at the nucleation "green" reporter (**Figure 3.4D**), which is embedded in the *cenH* nucleation element, except for a mild reduction in *Δfkh2*. We obtained a similar result at the related pericentromeric *dh* and *dg* nucleation elements (**Figure 3.4 Supplement 1D**). We also find defects in H3K9me2 spreading in Clr6 1″ subunit mutants at loci in addition to *MAT*. In particular, we observed reduced H3K9me2 accumulation at subtelomeric targets in chromosomes I and II in *Δprw1* and *fkh2*, with

a stronger effect in *Δprw1* (**Figure 3.1E**). At the facultative heterochromatin islands *mei4, ssm4,* and *mcp7, Δfkh2, Δprw1*, and *Δfkh2Δprw1* have similar effects as at subtelomeres (**Figure 3.4F**).

The *Δpng3* has a moderate to mild phenotype in *WT-MAT* and *MAT ΔREIII*, respectively (**Figure 3.4 Supplement 1A,B**). Consistent with this phenotype, we can observe a small change in H3K9me2 accumulation only distal to the "orange" reporter (**Figure 3.4D**) and a moderate effect at all heterochromatin islands tested (**Figure 3.4F**), suggesting that Png3 plays a less prominent role than Fkh2 within Clr6 I″. These results evidence a defect of heterochromatin assembly in Clr6 I″ subunit mutants primarily at distal, but not nucleation sites, such as *cenH* or *REIII*, indicating a surprising heterochromatin-spreading specific role of Clr6 I″.

DISCUSSION

The formation of a heterochromatin domain requires three interconnected steps, nucleation, assembly of silencing structures, and the lateral spreading from DNA-sequence driven nucleation sites. While the nucleation reaction has been well described, it has not been fully resolved which parts of the genetic circuitry discovered to date are required for spreading. Prior studies on spreading have focused primarily on factors that restrain heterochromatin formation across boundaries, which include Leo1, Mst2, Epe1, Bdf2, among others (Ayoub et al. 2003; Wang et al. 2013; Verrier et al. 2015; Wang et al. 2015). Here, our ability to separate requirements for nucleation and distal spreading within heterochromatin domains allowed us to pinpoint which factors are necessary to enable the spreading reaction and factors that constrain it genome-wide. A key finding from this work is the requirement of variants of the Clr6 HDAC complex specifically in the spreading reaction, in addition to the antagonism by a broad class of chromatin remodelers, including Ino80, Swr1C, SWI/SNF and RSC (**Figure 3.3**).

HDACs have long been implicated in heterochromatin function generally, and here we were able to distinguish which major HDACs regulate heterochromatin broadly, versus spreading specifically. The Clr3 HDAC and associated SHREC complex is required for silencing, likely via its ability to repress nucleosome turnover (Aygun et al. 2013), maintain nucleosome occupancy (Sugiyama et al. 2007; Garcia et al. 2010), and remove H3K14 acetylation known to antagonize heterochromatin assembly (Wirén et al. 2005). We find that SHREC mutants completely lose heterochromatic silencing (**Figure 3.1C-F, K**). Similarly, the Sir2 HDAC is broadly implicated in heterochromatin nucleation and assembly (Shankaranarayana et al. 2003; Alper et al. 2013), and we also observe near-complete silencing loss in *Δsir2* (**Figure 3.1 Supplement 2**). Clr6 complexes have been implicated in suppression of antisense transcription globally, particularly complex II/Rpd3S (Nicolas et al. 2007; Yamane et al. 2011) and the maintenance of heterochromatic gene silencing at major constitutive sites (Grewal et al. 1998), via the recruitment by HP1/Swi6 (Fischer et al. 2009). Remarkably, we find here that particular Clr6 subcomplexes are specifically required for distal spreading within constitutive heterochromatin and heterochromatin islands (**Figure 3.3, Figure 3.4**). Clr6 sub-complex activity is only marginally required for maintaining repression at nucleation-proximal sites. This is consistent with the finding that the *clr6-1* allele has only small impacts on transcription of the *cenH* nucleator-encoded ncRNAs (Yamane et al. 2011). We find the recently characterized Fkh2-associated complex I″ (Zilio et al. 2014) to be a central player in promoting distal spread (**Figure 3.3, Figure 3.4**). This complex contains the core of complex I, Nts1, Png3, Mug165 and associates with Fkh2. Not all these subunits contributed equally to regulating spreading in all chromatin contexts, with Fkh2 having the strongest effect (**Figure 3.3, Figure 3.4, Figure 3.3 Supplement 3,** and **Figure 3.4 Supplement 1**). The double mutant analysis

indicates that Fkh2 and Prw1 act together in promoting spreading (**Figure 3.4C**), which can be inferred from non-additive phenotype in *MAT ΔREIII*. The mutants are also non-additive with respect to silencing in *WT MAT*, however *Δprw1* and *Δfkh2Δprw1* phenotypes are weaker than *Δfkh2*. This type of behavior, i.e. the partial suppression of phenotype by additional deletion of complex components, has been observed for silencing defects of protein complexes (Barrales et al. 2016). In addition to Clr6 I″, the Alp13 and Cph1 subunits, which are assigned to the separate complex II/Rpd3S (Nicolas et al. 2007) associated with global deacetylation, also had positive impacts on spreading. This may indicate that (1) the composition of different Clr6 subcomplexes *in vivo* is either dynamic, or (2) a version of Rpd3S, jointly with a complex typified by Fkh2, cooperate in promoting distal spreading of heterochromatin. Interestingly, the Set3-submodule that typifies the Rpd3L-Expanded complex (Shevchenko et al. 2008), has a distinct spreading-antagonizing behavior (**Figure 3.3C**). This contrasts with a mild positive role of the Set3 complex at pericentromeres, which was proposed to be mediated indirectly, via transcriptional regulation of ClrC H3K9 methylase complex genes (Yu et al. 2016).

We do not believe that the spreading-specific role of Clr6 complexes at the MAT locus is mediated primarily by Asf/HIRA (Yamane et al. 2011), since Asf/HIRA subunits Hip1, Hip3 and Slm9 have mild to no phenotype for spreading in *MAT* contexts. Asf/HIRA mutant phenotypes were more pronounced in *ECT*, which implies less reliance on Clr6 for spreading in *ECT* (**Figure 3.2**). The spreading-specific role of specific Clr6 complexes may be encoded by their recruitment mechanisms. HP1/Swi6, which recruits Clr6 (Fischer et al. 2009), is a known spreading regulator (Hall et al. 2002; Canzio et al. 2011). Further, the Fkh2 transcription factor, which plays roles in origin coordination and clustering in budding yeast (Knott et al. 2012), may play a key role in either directing Clr6 to distal sites or ensuring continued association through the spreading

process, which would explain its dominant roles versus Png3 and other Clr6 I″ subcomplex-specific subunits (**Figure 3.4**, **Figure 3.4 Supplement 1**). The precise role of Fkh2 will be the subject of further study.

Chromatin remodelers across several classes antagonize spreading. While Ino80 and Swr1C have been linked to heterochromatin containment by specific mechanisms in budding yeast either by H2A.Z exchange or preventing the invasion of euchromatin (Meneghini et al. 2003; Xue et al. 2015), our results appear to indicate a more widespread effect of remodeling activities on spreading. This is because in addition to Ino80 and Swr1C, the major remodeling complexes RSC and SWI/SNF also contributed to spreading antagonism. A more general way remodelers have been implicated in heterochromatin function is creating nucleosome free regions (NFRs, (Lorch and Kornberg 2017)) that antagonize heterochromatin. Since NFRs may be roadblocks to spreading (Garcia et al. 2010; Lantermann et al. 2010), it is possible that remodelers employ this mechanism to restrain heterochromatin spreading globally. In addition, remodelers such as SWI/SNF and RSC destabilize nucleosomes generally (Narlikar et al. 2001; Rowe and Narlikar 2010), leading to increased turnover (Rawal et al. 2018), which would antagonize heterochromatin formation. This increased turnover may be tolerated at ncRNA nucleation sites, where it is at near euchromatic levels (Greenstein et al. 2018), likely due to ncRNA transcription (Volpe et al. 2002; Noma et al. 2004). This would indicate that regulation of nucleosome stability has a particular significance at distal, but not nucleation sites.

Further, we found that related ncRNA nucleators in different chromatin environments require distinct factors for spreading. Similarly, spreading from qualitatively different nucleators within the same environment, namely *REIII* and *cenH,* also differ in their sensitivity to different mutants. The significant overlap in factors between *WT MAT* and *MAT ΔREIII* indicates that

heterochromatin formation at MAT is dominated by the ncRNA nucleator *cenH,* in agreement

with our previous findings (Greenstein et al. 2018)*.* The *REIII* element, which nucleates

heterochromatin independent of ncRNA (Jia et al. 2004), had different requirements. ncRNA-

independent spreading is strongly antagonized by the Epe1 pathway, and uniquely promoted by

the MTOR pathway Gad8 kinase, partially consistent with a previous report implicating Gad8 for

MAT silencing (Cohen et al. 2018). The biggest difference was between the MAT contexts and

ECT. Some of the factors unique to *ECT* have been implicated in inhibiting spreading across

boundaries, such as Leo1 (Verrier et al. 2015), or heterochromatin stability and spreading

generally, such as HIRA (Yamane et al. 2011). However, these factors do not significantly

contribute to spreading within constitutive heterochromatin. The significant vulnerability of

ECT, compared to the similarly behaving *MAT ΔREIII* context, could be accounted by the

following possibilities: (1) The loss of Epe1 may impair RNAi specifically at *ECT,* nucleated by

a pericentromeric *dh* nucleator*,* but not *MAT ΔREIII,* which relies on the *dh*-homologous *cenH*

nucleator (Trewick et al. 2007; Braun et al. 2011). (2) Alternatively, this Epe1-dependence may

indicate that a naïve euchromatic context is less able to compete for heterochromatin factors.

Δepe1 is known to induce heterochromatin domain expansion at constitutive heterochromatin

sites (Ayoub et al. 2003; Trewick et al. 2007; Braun et al. 2011; Wang et al. 2015) and also

heterochromatin islands (Zofall et al. 2012; Wang et al. 2015; Greenstein et al. 2019), likely

depleting factors available for *ECT.* This alternative would hold important implications for

heterochromatin formation in euchromatic contexts, which occurs throughout differentiation in

animal and plant systems (Ringrose and Paro 2007; Schmitz and Amasino 2007; Zylicz et al.

2018). Of note, only the *ECT* context appeared to strongly require Hip1, and moderately Slm9,

for efficient spreading, which code for a key subunits of the HIRA H3/H4 chaperone. HIRA has

been implicated in stabilizing heterochromatic nucleosomes (Yamane et al. 2011) and it is possible this requirement additionally highlights the challenge faced by heterochromatic domains expanding within gene-rich chromatin, known to destabilize nucleosomes via transcription-associated processes.

In this work, we defined how regulation of heterochromatin silencing and nucleation differ in fundamental ways from distal spreading. While similar nucleation elements likely rely on a common set of machinery, the success of heterochromatin spreading appears much more sensitive to the chromatin context, particularly if euchromatic regions are targeted for *de novo* silencing. This finding has important implications for directing gene silencing of new loci appropriately, as cells change states in differentiation. In these situations, regions that are previously in a transcriptionally active state are invaded by heterochromatin and will have to compete for factors in a dosage limited system (Eissenberg et al. 1992; Nakayama et al. 2000; Al-Sady et al. 2016).

MATERIALS AND METHODS

Strains and strain construction

Mutant Generation for Genetic Screen

For the ectopic locus HSS reporter strain, the screen was performed essentially as described (Greenstein et al. 2019). Briefly, the parent HSS reporter strain was crossed to a ~400 gene deletion mini-library (see table 1) primarily consisting of subset of the Bioneer haploid deletion and several independently validated mutants. Crosses were performed as described (Verrier et al. 2015; Barrales et al. 2016; Greenstein et al. 2019) using a RoToR HDA colony pinning robot (Singer) for ECT while for the MAT HSS reporter strains, crosses were generated using a 96 well manual pinner. In addition, for the MAT HSS strains three Δ*clr4* mutant isolates and six individual parent isolates from each genomic context were included as controls. Crosses for the MAT HSS strains were performed on ME media for 3d at 27°C, while for the ectopic HSS strain crosses were performed using SPAS media for 4d at room temperature. For all strains, crosses were incubated for 4-5d at 42°C to retain spores, while removing unmated haploid and diploid cells. For MAT HSS strains, spores were germinated on YES medium supplemented with G418 and hygromycin B. The ectopic locus HSS spores were germinated on YES medium supplemented with G418, hygromycin B, and nourseothricin. The resulting colonies were pinned

into YES liquid medium for overnight growth and then prepared for flow cytometry as described below.

Validation Strain and Plasmid Construction

Plasmid constructs for gene knockout validation were generated by *in vivo* recombination as described (Greenstein et al. 2018; Greenstein et al. 2019). *S. pombe* transformants were selected as described (Greenstein et al. 2018). For microscopy, *hygMX* super-folder GFP (SFGFP) constructs for C-terminal tagging that we described previously (Al-Sady et al. 2016) were amplified with 175bp ultramer primers with homology to *apm3* or *apl5* and transformed into a Swi6:E2C *kanMX* strain. Apm3:SFGFP;Swi6:E2C and Apl5:SFGFP;Swi6:E2C strains were selected on hygromycin B and G418. Integrations and gene knockout were confirmed by PCR. All strains used for this study beyond the individual deletion library mutants are listed in Table 2.

Flow Cytometry

Flow cytometry data collection and normalization for genetic screen

In preparation for flow cytometry, overnight cultures were diluted to OD = 0.1 (approximately a 1:40 dilution) in rich media (YES) and incubated at 32°C with shaking of rpm for 4–6 hours. For the ectopic locus HSS strains, flow cytometry was performed essentially as described (Greenstein et al. 2018; Greenstein et al. 2019). For the MAT locus HSS strains, flow cytometry was performed using a Fortessa X20 Dual instrument (Becton Dickinson) attached with high throughput sampler (HTS) module. With a threshold of 30,000 events, samples sizes ranged from ~1000 to 30,000 cells depending on strain growth. Fluorescence detection and compensation, and data analysis were as described (Al-Sady et al. 2016; Greenstein et al. 2018; Greenstein et al. 2019).

Flow cytometry data collection and normalization for validation

For validation flow cytometry experiments, cells were grown as described (Greenstein et al. 2018; Greenstein et al. 2019) with the exception that cells were diluted into YES medium and grown 5-8 hours before measurement. Flow cytometry was performed as above. Depending on strain growth and the volume collected per experiment, fluorescence values were measured for ~20,000-100,000 cells per replicate. Fluorescence detection, compensation, and data analysis were as described (Al-Sady et al. 2016; Greenstein et al. 2018; Greenstein et al. 2019).

2D-density histogram plots with ON and OFF boundary guidelines.

2D-density histogram plots (Greenstein et al. 2018; Greenstein et al. 2019) were generated as described previously, (see (Greenstein et al. 2018) methods for Figure 3.4) with the following exceptions: For MAT locus strains, the guide-lines for boundary values of "off" and "on" states were determined using median of a Red-Only control plus 3 times the median absolute deviation (MAD) and median of Δ*clr4* minus 2 times the MAD value respectively. For the ECT strain, the "on" boundary guideline was calculated by median of Δ*clr4* minus one MAD value. The boundary guideline for the ECT "off" state was determined by adjusting the raw red-channel values for cells from of a no-color control strain analyzed on the same flow cytometry run to simulate a Red-Only control strain by adding the median value of the respective Δ*clr4* strain to the red value of each cell. The resulting adjusted data was used to calculate the "off" cutoff by median +3MAD as above. Validation flow cytometry plots were generated using the ggplot2 R package (Wickham 2016).

t-distributed stochastic neighbor embedding (t-SNE) analysis

Since our data varied in sample size, we transformed normalized fluorescent data into 5x5 density grids. The percentage of population residing in each of the 25 grids is represented by one variable. Increasing the number of grids to 8x8 or 10x10 did not significantly alter results. The following input settings were used before t-SNE reduction: concatenated distributions, 1000 iterations, 60 perplexity, 0 Theta. t-SNE reduction were conducted using the Rtsne package (Krijthe 2015) and the generated t-SNE model was plotted using the R packages ggplot2 (Wickham 2016).

Earth Mover's Distance (EMD) Analysis

EMD is a distance measure between two multi-dimensional distributions to evaluate the dissimilarity. EMD calculates the minimal amount of work to match the two equisized distribution (Rubner et al. 1998). The normalized fluorescent data are transformed into a 15x15 density grid as above. To calculate the fraction of change between mutant and parent population, we calculated a metric of $EMD^{mut/par}$ by taking the log2 transformation of the quotient of the EMD between mutant to $\Delta clr4$ distributions (EMD^{mut}) by the EMD between parent to $\Delta clr4$ distributions (EMD^{par}). For MAT HSS strains, the three $\Delta clr4$ and six parent control isolates were respectively combined and transformed into one reference population matrix. Two-dimensional EMD between any pair of distributions in Euclidean distance are calculated using functions from the R package emdist (Urbanek and Rubner 2012).

We validated the 1000 cell cutoff for the EMD analysis by iteratively down-sampling parent strains, calculating the EMD value to the $\Delta clr4$ distribution, and comparing the resulting values across iterations. For each HSS parent context, all wells containing parent isolates (N=6 for MAT locus backgrounds, N=1 for ECT) were combined into one pool from which a random

sample was drawn 100 times for each N = 1000, and 5000-30,000 in 5000 cell increments. For each iteration of each N number of cells, the EMD value of the down-sampled population was calculated relative to the respective Δ*clr4* population. We calculated the mean and SD for the 100 iterations of for each N and determined that even at an N of 1000 cells in the most broadly distributed background (ECT) the SD of EMD values was <2% of the EMD value of the parent calculated from its entire distribution. Given this analysis, we conclude that while the EMD values for wells with lower cell counts will likely have higher error than those with more cells, this error is not likely to have major effects on the calculated metric.

EMD$^{mut/par}$ data were plotted in a heatmap using functions from the R package superheat (Barter and Yu 2018). Any values less than -2 were converted to -2 for interpreting the heatmap.

Correlation Analysis

The correlation between the EMD$^{mut/par}$ data of different pairs of genomic contexts were calculated using a linear regression model. Sir2, ClrC and SHREC are essential heterochromatin assembly factor, and their mutants have very low EMD$^{mut/par}$. These mutants were disregarded from the model fitting because they would drive the correlation. Pearson correlation coefficients and 95% confidence interval were calculated using ggscatter functions from R package ggplot2 (Wickham 2016) and ggpubr (Kassambara 2020).

Spreading Analysis

Nucleated cells were extracted using a "green"-off gate, using median of a "red"-Only control plus 2 times the SD. Enrichment of cell populations in particular "orange" fluorescence ranges (Grid$_n$) are calculated as Grid$_n$$^{mut/par}$: fraction of mutant population is divided by the fraction of parent population in one grid. The intervals of "orange" fluorescence used in grids are

determined by: median+2SD of "orange"-OFF cells, median-1SD of "orange"-ON cells and the median of the two. To evaluate gain of spreading phenotype, enrichment in Grid 1 in *WT MAT, MAT ΔREIII and ECT* were calculated. To evaluate loss of spreading phenotype, enrichment in Grid 3 and 4 in *WT MAT* and *MAT ΔcenH,* as well as Grid 4 in *MAT ΔREIII and ECT* were calculated. The distribution of the $Grid_n^{mut/par}$ were plotted as histogram with annotation of 85th percentile and median+2SD from parent isolates. Gene hit lists comprised mutations above median and 2SD within the 85th percentile. Upset plots were generated using the R package UpSetR (Conway et al. 2017). Barplots were plotted using the R packages ggplot2 (Wickham 2016).

GO Complex and Sub-Complex analysis

Generating the heatmap count data

GO Complexes – Based on the GO Complex annotations [link] (2019) retrieved from pombase (Lock et al. 2019), GO complex membership was determined for genes identified as hits for each strain background and hit category (gain/loss). Using functions from the R package dplyr (Wickham et al. 2020), gene names were converted to systematic ID numbers and these systematic IDs were queried against the GO complex annotation table. The number of times a GO complex appeared per background and hit category was tabulated. Genes can be associated with any number of GO complexes depending on their annotations. However, any particular gene was only counted once per GO complex despite potentially being annotated to that GO complex by more than one evidence code. The unique list of GO complexes for all hits was determined and a matrix was computed representing the number of times each GO complex (row) was identified per strain/hit category (column). This counts matrix was used to generate the GO complex heatmap in Figure 3.3A, described below.

Sub-complex analysis – Genes annotated to the seven complexes in Figure 3.3B,C were obtained from pombase (Lock et al. 2019). *fkh2* was added to the Clr6 I″ complex given the protein contacts described previously (Zilio et al. 2014). For the unique set of genes per panel it was determined if each gene was identified as a hit in each strain background/hit category combination. The data was summarized in a counts matrix where rows represent the unique list of genes per panel and columns represent the strain background / hit category. The counts matrix for each set of genes was used to generate the heatmaps in Figure 3.3B, C as described below.

Generating the heatmap clustering

Using the R package ComplexHeatmap (Gu et al. 2016), both row and column dendrogram and clustering were generated using hierarchical clustering. Based on an optimal Silhouette score, the strain background / hit category (columns) were clustered into 3 (Figure 3.3A). The dendrogram representing complexes (Figure 3.3A) in rows were not separated because validations of the clustering by connectivity, Dunn index or Silhouette score were inconclusive. Clustering validations were conducted using the R package clValid (Brock et al. 2008).

Chromatin Immunoprecipitation and Quantification

Chromatin Immunoprecipitation (ChIP) was performed essentially as described (Greenstein et al. 2018; Greenstein et al. 2019). Bulk populations of cells for were grown overnight to saturation in YES medium. The following morning, cultures were diluted to OD 0.1 in 25mL YES and grown for 8h at 32°C and 225rpm. Based on OD measurements, 60×10^6 cells were fixed and processed for ChIP as previously described (Greenstein et al. 2018; Greenstein et al. 2019) without the addition of W303 carrier. Cleared chromatin for each ChIP sample was incubated with 1μL of anti-H3K9me2 antibody (Abcam, ab1220) overnight after a small fraction was retained as

Input/WCE. DNAs were quantified by RT-qPCR and percent immunoprecipitation (%IP, ChIP DNA/Input DNA*100) was calculated as described (Greenstein et al. 2018; Greenstein et al. 2019). Data for %IP was plotted using the R packages ggplot2 (Wickham 2016) and dplyr (Wickham et al. 2020).

Microscopy

Swi6:E2C; Apl5:SFGFP and Swi6:E2C; Apm3:SFGFP cells were grown is YS media as described. Slides (ibidi, Cat. No. 80606) were pre-coated with 100 mg/mL lectin (Sigma-Aldrich, Cat. No. L1395) diluted in water by adding lectin solution to slide for 1 min. and removing supernatant. Log-phase growing cells were applied to the slide and excess cells were rinsed off with YS. Cells were immediately imaged with a 60x objective (CFI Plan Apochromat VC 60XC WI) on a Nikon TI-E equipped with a spinning-disk confocal head (CSU10, Yokogawa) and an EM-CCD camera (Hammamatsu). Cells were imaged in brightfield and additionally excited with 488nm (SFGFP) and 561nm (E2C) lasers. Emission was collected using a 510/50 band-pass filter for GFP emission and a 600/50 band-pass filter for E2C emission. For the SFGFP and E2C channels, z-stacks were obtained at 0.3μm/slice for 11 slices total. An overlay of the maximum z-projections for SFGFP and E2C channels are shown separately from the brightfield images. Brightness and contrast were adjusted in ImageJ to clearly show both Swi6 and Apl5/Apm5 signals in the overlay. At least 2 isolates were imaged to confirm localization patterns.

Data Availability

Screen flow cytometry data and analysis scripts will be made available upon publication of the manuscript.

Table 3.1: Nuclear function deletion library

Systematic ID	Symbol	Description
SPAC1002.05c	*jmj2*	histone demethylase Jmj2
SPAC1006.03c	*red1*	RNA elimination defective protein Red1
SPAC1039.05c	*klf1*	transcription factor, zf-fungal binuclear cluster type Klf1
SPAC1071.02	*mms19*	CIA machinery protein Mms19
SPAC1071.06	*arp9*	SWI/SNF and RSC complex subunit Arp9
SPAC10F6.08c	*nht1*	Ino80 complex HMG box subunit Nht1
SPAC10F6.11c	*atg17*	autophagy associated protein kinase activator Atg17
SPAC1142.03c	*swi2*	Swi5 complex subunit Swi2
SPAC1142.08	*fhl1*	forkhead transcription factor Fhl1
SPAC11D3.07c	*toe4*	transcription factor, zf-fungal binuclear cluster type(predicted)
SPAC11D3.16c		Schizosaccharomyces specific protein
SPAC11E3.01c	*swr1*	SNF2 family ATP-dependent DNA helicase Swr1
SPAC11H11.01	*sst6*	ESCRT I complex subunit Vps23
SPAC11H11.05c	*fta6*	Mis6-Sim4 complex Fta6
SPAC12B10.10	*nod1*	medial cortical node Gef2-related protein protein Nod1
SPAC12G12.13c	*cid14*	TRAMP complex poly(A) polymerase subunit Cid14
SPAC139.03	*toe2*	transcription factor, zf-fungal binuclear cluster type (predicted)
SPAC139.06	*hat1*	histone acetyltransferase Hat1
SPAC1399.05c	*toe1*	transcription factor, zf-fungal binuclear cluster type
SPAC13A11.04c	*ubp8*	SAGA complex ubiquitin C-terminal hydrolase Ubp8
SPAC13D6.02c	*byr3*	translational activator, zf-CCHC type zinc finger protein (predicted)
SPAC13G6.01c	*rad8*	ubiquitin-protein ligase E3/ ATP-dependent DNA helicase Rad8
SPAC144.02	*iec1*	Ino80 complex subunit Iec1
SPAC144.05		DNA-dependent ATPase/ ubiquitin-protein ligase E3 (predicted)
SPAC144.06	*apl5*	AP-3 adaptor complex subunit Apl5 (predicted)
SPAC144.14	*klp8*	kinesin-like protein Klp8
SPAC14C4.06c	*nab2*	poly(A) binding protein Nab2 (predicted)
SPAC14C4.12c	*laf1*	Clr6 L associated factor 1 Laf1
SPAC14C4.13	*rad17*	RFC related checkpoint protein Rad17
SPAC1556.01c	*rad50*	DNA repair protein Rad50
SPAC15A10.11	*ubr11*	UBR ubiquitin-protein ligase E3 Ubr11
SPAC15A10.15	*sgo2*	inner centromere protein, shugoshin Sgo2
SPAC1610.01	*saf5*	splicing factor Saf5
SPAC1610.02c	*mrpl1*	mitochondrial ribosomal protein subunit L1 (predicted)
SPAC1687.05	*pli1*	SUMO E3 ligase Pli1

Systematic ID	Symbol	Description
SPAC1687.09	*irs4*	autophagy/CVT pathway ENTH/VHS domain protein Irs4 (predicted)
SPAC16A10.03c		ubiquitin-protein ligase E3 involved in vesicle docking Pep5/Vps11-like (predicted)
SPAC16A10.07c	*taz1*	shelterin complex subunit Taz1
SPAC16C9.04c	*mot2*	CCR4-Not complex ubiquitin-protein ligase E3 subunit Mot2
SPAC16C9.05	*cph1*	Clr6 histone deacetylase associated PHD protein-1 Cph1
SPAC16E8.12c	*png3*	ING family homolog Png3 (predicted)
SPAC1751.01c	*gti1*	gluconate transmembrane transporter inducer Gti1
SPAC1782.05	*ypa2*	protein phosphatase type 2A regulator, PTPA family Ypa2
SPAC1782.08c	*rex3*	exonuclease Rex3 (predicted)
SPAC1782.09c	*clp1*	Cdc14-related protein phosphatase Clp1/Flp1
SPAC1783.05	*hrp1*	ATP-dependent DNA helicase Hrp1
SPAC17A2.12	*rrp1*	ATP-dependent DNA helicase/ ubiquitin-protein ligase E3 (predicted)
SPAC17G8.05	*med20*	mediator complex subunit Med20
SPAC17G8.07	*yaf9*	YEATS family histone acetyltransferase subunit Yaf9
SPAC17G8.09	*shg1*	Set1C complex subunit Shg1
SPAC17G8.10c	*dma1*	mitotic spindle checkpoint ubiquitin ligase Dma1
SPAC17G8.13c	*mst2*	histone acetyltransferase Mst2
SPAC17H9.10c	*ddb1*	Cul4-RING E3 adaptor Ddb1
SPAC17H9.19c	*cdt2*	WD repeat protein Cdt2
SPAC1805.14		Schizosaccharomyces specific protein
SPAC1805.15c	*pub2*	HECT-type ubiquitin-protein ligase E3 Pub2
SPAC1851.03	*ckb1*	CK2 family regulatory subunit Ckb1
SPAC18G6.02c	*chp1*	heterochromatin (HP1) family chromodomain protein Chp1
SPAC18G6.10	*lem2*	LEM domain nuclear inner membrane protein Heh1/Lem2
SPAC18G6.13		Schizosaccharomyces specific protein
SPAC1952.05	*gcn5*	SAGA complex histone acetyltransferase catalytic subunit Gcn5
SPAC19A8.10	*rfp1*	SUMO-targeted ubiquitin-protein ligase subunit Rfp1
SPAC19D5.06c	*din1*	RNA pyrophosphohydrolase Din1
SPAC19D5.11c	*ctf8*	Ctf18 RFC-like complex subunit Ctf8
SPAC19E9.02	*fin1*	serine/threonine protein kinase, NIMA related Fin1
SPAC19G12.13c	*poz1*	shelterin complex subunit Poz1
SPAC19G12.17	*erh1*	enhancer of rudimentary homolog Erh1
SPAC1B3.17	*clr2*	chromatin silencing protein Clr2
SPAC1D4.09c	*rtf2*	replication termination factor Rtf2

Systematic ID	Symbol	Description
SPAC1D4.11c	*lkh1*	dual specificity protein kinase Lkh1
SPAC1F3.01	*rrp6*	exosome 3'-5' exoribonuclease subunit Rrp6
SPAC1F3.06c	*spo15*	mitotic and mieotic spindle pole body protein Spo15
SPAC1F7.01c	*spt6*	nucleosome remodeling protein Spt6
SPAC20G4.04c	*hus1*	checkpoint clamp complex protein Hus1
SPAC20G8.08c	*fft1*	SMARCAD1 family ATP-dependent DNA helicase Fft1 (predicted)
SPAC20H4.03c	*tfs1*	general transcription elongation factor TFIIS
SPAC20H4.10	*ufd2*	ubiquitin-protein ligase E4 Ufd2 (predicted)
SPAC21E11.03c	*pcr1*	transcription factor Pcr1
SPAC21E11.05c	*cyp8*	cyclophilin family peptidyl-prolyl cis-trans isomerase Cyp8
SPAC222.04c	*ies6*	Ino80 complex subunit Ies6
SPAC222.15	*meu13*	Tat binding protein 1(TBP-1)-interacting protein (TBPIP) homolog (predicted)
SPAC22A12.01c	*pso2*	DNA 5' exonuclease (predicted)
SPAC22E12.11c	*set3*	histone lysine methyltransferase Set3
SPAC22E12.19	*snt1*	Set3 complex subunit Snt1
SPAC22F3.02	*atf31*	transcription factor Atf31
SPAC22F3.09c	*res2*	MBF transcription factor complex subunit Res2
SPAC22F8.12c	*shf1*	small histone ubiquitination factor Shf1
SPAC22H12.02	*tfg3*	TFIID, TFIIF, Ino80, SWI/SNF, and NuA3 complex subunit Tfg3
SPAC23A1.07		ubiquitin-protein ligase E3 (predicted)
SPAC23C11.08	*php3*	CCAAT-binding factor complex subunit Php3
SPAC23C11.15	*pst2*	Clr6 histone deacetylase complex subunit Pst2
SPAC23C4.03	*hrk1*	haspin related kinase Hrk1
SPAC23D3.01	*pdp3*	PWWP domain protein, involved in chromatin remodeling (predicted)
SPAC23D3.09	*arp42*	SWI/SNF and RSC complex subunit Arp42
SPAC23E2.01	*fep1*	iron-sensing transcription factor Fep1
SPAC23E2.03c	*ste7*	arrestin family meiotic suppressor protein Ste7
SPAC23G3.04	*ies4*	Ino80 complex subunit Ies4
SPAC23G3.07c	*snf30*	SWI/SNF complex subunit Snf30
SPAC23G3.08c	*ubp7*	ubiquitin C-terminal hydrolase Ubp7
SPAC23G3.10c	*ssr3*	SWI/SNF and RSC complex subunit Ssr3
SPAC23H3.05c	*swd1*	Set1C complex subunit Swd1
SPAC23H4.12	*alp13*	MRG family Clr6 histone deacetylase complex subunit Alp13
SPAC24B11.10c	*cfh1*	SEL1/TPR repeat protein Cfh1 (predicted)
SPAC25A8.01c	*fft3*	SMARCAD1 family ATP-dependent DNA helicase Fft3

Systematic ID	Symbol	Description
SPAC25H1.02	*jmj1*	histone demethylase Jmj1 (predicted)
SPAC26H5.03	*pcf2*	CAF assembly factor (CAF-1) complex subunit B, Pcf2
SPAC29A4.09		rRNA exonuclease Rrp17 (predicted)
SPAC29A4.18	*prw1*	Clr6 histone deacetylase complex subunit Prw1
SPAC29B12.02c	*set2*	histone lysine H3-K36 methyltransferase Set2
SPAC29B12.03	*spd1*	ribonucleotide reductase (RNR) inhibitor
SPAC29B12.06c	*rcd1*	CCR4-Not complex RNA-binding protein subunit Rcd1
SPAC29B12.08	*clr5*	Clr5 protein
SPAC2C4.07c	*dis32*	3'-5'-exoribonuclease activity Dis3L2
SPAC2F3.15	*lsk1*	P-TEFb-associated cyclin-dependent protein kinase Lsk1
SPAC2F3.16		ubiquitin-protein ligase E3, implicated in DNA repair (predicted)
SPAC2F7.07c	*cph2*	Clr6 histone deacetylase associated PHD protein Cph2
SPAC2F7.08c	*snf5*	SWI/SNF complex subunit Snf5
SPAC2G11.05c	*rim20*	BRO1 domain protein Rim20
SPAC2G11.10c	*uba42*	thiosulfate sulfurtransferase, URM1 activating enzyme E1-type Uba42 (predicted)
SPAC30D11.07	*nth1*	DNA endonuclease III
SPAC31A2.09c	*apm4*	AP-2 adaptor complex mu subunit Apm4 (predicted)
SPAC31A2.16	*gef2*	RhoGEF Gef2
SPAC31G5.09c	*spk1*	MAP kinase Spk1
SPAC31G5.19	*abo1*	ATPase with bromodomain protein
SPAC323.03c		Schizosaccharomyces specific protein
SPAC328.05	*hrb1*	RNA-binding protein involved in export of mRNAs Hrb1 (predicted)
SPAC32A11.03c	*phx1*	stationary phase-specific homeobox transcription factor Phx1
SPAC343.04c	*gid7*	GID complex subunit Gid7 (predicted)
SPAC343.11c	*msc1*	Swr1 complex subunit Msc1
SPAC343.18	*rfp2*	SUMO-targeted ubiquitin-protein ligase subunit Rfp2
SPAC3A11.05c	*kms1*	meiotic spindle pole body KASH domain protein Kms1
SPAC3C7.08c	*elf1*	AAA family ATPase Elf1
SPAC3F10.10c	*map3*	pheromone M-factor receptor Map3
SPAC3F10.12c		transcription factor (predicted)
SPAC3G6.01	*hrp3*	ATP-dependent DNA helicase Hrp3
SPAC3G6.06c	*rad2*	FEN-1 endonuclease Rad2
SPAC3G6.11	*chl1*	ATP-dependent DNA helicase Chl1 (predicted)
SPAC3G9.07c	*hos2*	histone deacetylase (class I) Hos2
SPAC3H1.11	*hsr1*	transcription factor Hsr1
SPAC3H1.12c	*snt2*	Lid2 complex PHD finger subunit Snt2
SPAC3H8.08c		transcription factor (predicted)

Systematic ID	Symbol	Description
SPAC4A8.09c	cwf21	complexed with Cdc5 protein Cwf21
SPAC4F8.11	sea2	SEA complex WD repeat subunit Sea2 (predicted)
SPAC4G9.06c	chz1	histone H2A-H2B dimer chaperone Chz1 (predicted)
SPAC4H3.02c	swc3	Swr1 complex subunit Swc3
SPAC4H3.05	srs2	ATP-dependent DNA helicase, UvrD subfamily
SPAC56F8.16	esc1	transcription factor Esc1 (predicted)
SPAC57A10.09c	nhp6	High-mobility group non-histone chromatin protein (predicted)
SPAC5D6.02c	mug165	Clr6 histone deacetylase complex subunit Mug165
SPAC5D6.08c	mes1	meiotic APC inhibitor Mes1
SPAC630.14c	tup12	transcriptional corepressor Tup12
SPAC631.02	bdf2	BET family double bromodomain protein Bdf2
SPAC637.09	rex1	3'-5'- exoribonuclease Rex1 (predicted)
SPAC644.14c	rad51	RecA family recombinase Rad51/Rhp51
SPAC664.01c	swi6	heterochromatin (HP1) family chromodomain protein Swi6
SPAC664.02c	arp8	Ino80 complex actin-like protein Arp8
SPAC664.07c	rad9	checkpoint clamp complex protein Rad9
SPAC664.15	caf4	CCR4-Not complex subunit Caf4/Mdv1 (predicted)
SPAC688.06c	slx4	structure-specific endonuclease subunit Slx4
SPAC694.06c	mrc1	claspin, Mrc1
SPAC6B12.05c	ies2	Ino80 complex subunit Ies2
SPAC6B12.07c		ubiquitin-protein ligase E3 with SPX domain, human LORNRF1 ortholog (predicted)
SPAC6B12.14c		conserved fungal protein
SPAC6B12.16	meu26	DUF4451 family conserved fungal protein
SPAC6F12.09	rdp1	RNA-directed RNA polymerase Rdp1
SPAC6F6.09	eaf6	Mst2/NuA4 histone acetyltransferase complex subunit Eaf6
SPAC6G9.03c	mug183	histone H3.3 H4 heterotetramer chaperone Rtt106-like (predicted)
SPAC6G9.10c	sen1	ATP-dependent 5' to 3' DNA/RNA helicase Sen1
SPAC6G9.16c	xrc4	XRCC4 nonhomologous end joining factor Xrc4
SPAC7D4.04	atg11	autophagy associated protein Atg11
SPAC7D4.14c	iss10	NURS complex subunit Iss10
SPAC821.07c	moc3	transcription factor Moc3
SPAC823.03	ppk15	serine/threonine protein kinase Ppk15 (predicted)
SPAC824.04	swd22	mRNA cleavage and polyadenylation specificity factor complex, WD repeat protein Swd22
SPAC890.07c	rmt1	type I protein arginine N-methyltransferase Rmt1
SPAC8C9.14	prr1	transcription factor Prr1

Systematic ID	Symbol	Description
SPAC8C9.17c	*spc34*	DASH complex subunit Spc34
SPAC8F11.03	*msh3*	MutS protein homolog 3
SPAC9E9.08	*rad26*	ATRIP, ATR checkpoint kinase regulatory subunit Rad26
SPAC9E9.10c	*cbh1*	CENP-B homolog Cbh1
SPAP14E8.02	*tos4*	chromatin binding FHA domain protein Tos4 (predicted)
SPAP27G11.15	*slx1*	structure-specific endonuclease catalytic subunit Slx1
SPAP32A8.03c	*bop1*	ubiquitin-protein ligase E3, human RNF126 ortholog (predicted)
SPAP8A3.02c	*ofd2*	histone H2A dioxygenase Ofd2
SPAPB1E7.02c	*mcl1*	DNA polymerase alpha accessory factor Mcl1
SPAPB24D3.01	*toe3*	transcription factor (predicted)
SPAPB2B4.03	*cig2*	G1/S-specific B-type cyclin Cig2
SPBC1105.04c	*cbp1*	CENP-B homolog
SPBC119.08	*pmk1*	MAP kinase Pmk1
SPBC119.14	*rti1*	Rad22 homolog Rti1
SPBC1198.11c	*reb1*	RNA polymerase I transcription termination factor/ RNA polymerase II transcription factor Reb1
SPBC11B10.05c	*rsp1*	random septum position protein, DNAJ domain protein Rsp1
SPBC11B10.08		WW domain containing conserved fungal protein
SPBC11B10.10c	*pht1*	histone H2A variant H2A.Z, Pht1
SPBC1347.07	*rex2*	RNA exonuclease (predicted)
SPBC13E7.08c	*leo1*	RNA polymerase II associated Paf1 complex subunit Leo1
SPBC13G1.08c	*ash2*	Ash2-trithorax family protein
SPBC146.06c	*fan1*	DNA repair protein Fan1
SPBC14C8.17c	*spt8*	SAGA complex subunit Spt8
SPBC14F5.07	*doa10*	ER ubiquitin-protein ligase E3 Doa10 (predicted)
SPBC15C4.01c	*oca3*	TPR repeat protein Oca3/ ER membrane protein complex Ecm2 (predicted)
SPBC15C4.06c		ubiquitin-protein ligase E3 Meu34, human RNF13 family homolog, unknown biological role (predicted)
SPBC15D4.03	*slm9*	histone H3.3 H4 chaperone, hira family Slm9
SPBC1604.09c	*rex4*	exoribonuclease Rex4 (predicted)
SPBC1604.16c		RNA-binding protein, G-patch type, human GPANK1 ortholog
SPBC1685.08	*cti6*	histone deacetylase complex ubiquitin-like protein ligase subunit Cti6
SPBC16A3.07c	*nrm1*	MBF complex corepressor Nrm1
SPBC16A3.19	*eaf7*	histone acetyltransferase complex subunit Eaf7
SPBC16D10.07c	*sir2*	Sirtuin family histone deacetylase Sir2
SPBC16E9.11c	*pub3*	HECT-type ubiquitin-protein ligase E3 Pub3 (predicted)

Systematic ID	Symbol	Description
SPBC16E9.12c	*pab2*	poly(A) binding protein Pab2
SPBC16G5.03	*mrz1*	ubiquitin-protein ligase E3/SUMO transferase, Topors, possibly associated with DNA damage (predicted)
SPBC16G5.15c	*fkh2*	forkhead transcription factor Fkh2
SPBC16G5.17		transcription factor, zf-fungal binuclear cluster type (predicted)
SPBC1703.04	*mlh1*	MutL family protein Mlh1 (predicted)
SPBC1703.14c	*top1*	DNA topoisomerase I
SPBC1709.11c	*png2*	ING family histone acetyltransferase complex PHD-type zinc finger subunit Png2
SPBC1711.14	*rec15*	meiotic recombination protein Rec15
SPBC1718.02	*hop1*	linear element associated protein Hop1
SPBC1734.06	*rhp18*	Rad18 homolog ubiquitin protein ligase E3, Rhp18
SPBC1734.15	*rsc4*	RSC complex subunit Rsc4
SPBC1773.16c		transcription factor, zf-fungal binuclear cluster type(predicted)
SPBC1778.10c	*ppk21*	serine/threonine protein kinase Ppk21 (predicted)
SPBC17D11.04c	*nto1*	histone acetyltransferase complex PHD finger subunit Nto1 (predicted)
SPBC17G9.05	*rct1*	cyclophilin family peptidyl-prolyl cis-trans isomerase, RRM-containing Rct1
SPBC18H10.06c	*swd2*	Set1C complex subunit Swd2.1
SPBC18H10.15	*cdk11*	serine/threonine protein kinase Cdk11
SPBC19C7.02	*ubr1*	N-end-recognizing protein, UBR ubiquitin-protein ligase E3 Ubr1
SPBC1A4.03c	*top2*	DNA topoisomerase II
SPBC1D7.03	*clg1*	cyclin-like protein involved in autophagy Clg1 (predicted)
SPBC1D7.04	*mlo3*	RNA binding protein Mlo3
SPBC20F10.05	*nrl1*	RNAi-mediated silencing protein, human NRDE2 ortholog Nrl1
SPBC20F10.10	*psl1*	cyclin pho85 family Psl1 (predicted)
SPBC215.03c	*csn1*	COP9/signalosome complex subunit Csn1
SPBC215.06c		nucleolar RNA-binding protein, human LYAR homolog, implicated in transcriptional regulation
SPBC215.07c	*pdp2*	PWWP domain protein Pdp2 (predicted)
SPBC216.05	*rad3*	ATR checkpoint kinase Rad3
SPBC216.06c	*swi1*	replication fork protection complex subunit Swi1
SPBC21B10.13c	*yox1*	MBF complex corepressor Yox1
SPBC21C3.02c	*dep1*	Sds3-like family protein Dep1
SPBC21C3.20c	*git1*	C2 domain protein Git1
SPBC21D10.09c	*rkr1*	RQC complex ubiquitin-protein ligase E3 Rkr1 (predicted)

Systematic ID	Symbol	Description
SPBC21D10.10	*bdc1*	bromodomain protein Bdc1
SPBC23E6.02	*rrp2*	ATP-dependent DNA helicase/ ubiquitin-protein ligase E3 (predicted)
SPBC23E6.09	*ssn6*	transcriptional corepressor Ssn6
SPBC23G7.13c		plasma membrane urea transmembrane transporter (predicted)
SPBC24C6.05	*sec28*	coatomer epsilon subunit (predicted)
SPBC25B2.08		Schizosaccharomyces pombe specific protein
SPBC26H8.09c	*snf59*	SWI/SNF complex subunit Snf59
SPBC28E12.02		RNA-binding protein
SPBC28F2.07	*sfr1*	Swi five-dependent recombination mediator Sfr1
SPBC28F2.10c	*ngg1*	SAGA complex subunit Ngg1/Ada3
SPBC29A10.03c	*pcf1*	CAF assembly factor (CAF-1) complex large subunit Pcf1
SPBC29A10.05	*exo1*	exonuclease I Exo1
SPBC29A10.14	*rec8*	meiotic cohesin complex subunit Rec8
SPBC29A3.03c	*gid2*	GID complex ubiquitin-protein ligase E3 subunit Gid2/Rmd5 (predicted)
SPBC29A3.05	*vps71*	Swr1 complex subunit Vps71
SPBC29A3.13	*pdp1*	PWWP domain protein Pdp1
SPBC29B5.01	*atf1*	transcription factor, Atf-CREB family Atf1
SPBC2A9.04c	*san1*	sir antagonist, ubiquitin-protein ligase E3
SPBC2D10.11c	*nap2*	histone H2A-H2B chaperone Nap2
SPBC2D10.17	*clr1*	SHREC complex intermodule linker subunit Clr1
SPBC2F12.09c	*atf21*	transcription factor, Atf-CREB family Atf21
SPBC2F12.12c	*cay1*	cactin, spliceosome complex subunit
SPBC2G2.06c	*apl1*	AP-2 adaptor complex beta subunit Apl1 (predicted)
SPBC2G2.14	*csi1*	mitotic centromere-SPB clustering protein Csi1
SPBC2G5.02c	*ckb2*	CK2 family regulatory subunit Ckb2 (predicted)
SPBC30B4.04c	*sol1*	SWI/SNF complex subunit Sol1
SPBC30D10.10c	*tor1*	serine/threonine protein kinase Tor1
SPBC31F10.07	*lsb5*	actin cortical patch component Lsb5 (predicted)
SPBC31F10.10c	*mub1*	Armadillo-type fold protein, zf-MYND type zinc finger protein, Mub1-Rad6-Ubr2 ubiquitin ligase complex Mub1 (predicted)
SPBC31F10.13c	*hip1*	histone H3.3 H4 chaperone, hira family Hip1
SPBC31F10.14c	*hip3*	HIRA interacting protein Hip3
SPBC32F12.07c		membrane associated ubiquitin-protein ligase E3, MARCH family (predicted)
SPBC32H8.06	*mug93*	TPR repeat protein, meiotically spliced
SPBC337.03	*rhn1*	RNA polymerase II transcription termination factor homolog
SPBC342.05	*crb2*	DNA repair protein Rad9 homolog Crb2

Systematic ID	Symbol	Description
SPBC342.06c	*rtt109*	RTT109 family histone lysine acetyltransferase
SPBC354.03	*swd3*	WD repeat protein Swd3
SPBC354.05c	*sre2*	membrane-tethered transcription factor Sre2
SPBC365.10	*arp5*	Ino80 complex actin-like protein Arp5
SPBC36B7.05c	*pib1*	endosomal and vacuolar ubiquitin-protein ligase E3/phosphatidylinositol(3)-phosphate binding protein Pib1
SPBC36B7.08c	*ccp1*	histone chaperone, CENP-A nucleosome disassembly Ccp1
SPBC3B8.02	*php5*	CCAAT-binding factor complex subunit Php5
SPBC3D6.04c	*mad1*	mitotic spindle checkpoint protein Mad1
SPBC3D6.09	*dpb4*	DNA polymerase epsilon subunit Dpb4
SPBC3H7.13	*far10*	SIP/FAR complex FHA domain subunit Far10/Csc1
SPBC4.05	*mlo2*	ubiquitin protein ligase E3 component human N-recognin 7 homolog Mlo2
SPBC428.06c	*rxt2*	histone deacetylase complex subunit Rxt2
SPBC428.08c	*clr4*	histone lysine H3 methyltransferase Clr4
SPBC4B4.03	*rsc1*	RSC complex subunit Rsc1
SPBC4C3.12	*Sep1*	forkhead transcription factor Sep1
SPBC530.08		membrane-tethered transcription factor (predicted)
SPBC530.14c	*dsk1*	SR protein-specific kinase Dsk1
SPBC543.07	*pek1*	MAP kinase kinase Pek1
SPBC56F2.03	*arp10*	dynactin complex actin-like protein Arp10 (predicted)
SPBC56F2.05c		transcription factor (predicted)
SPBC582.04c	*dsh1*	RNAi protein, Dsh1
SPBC582.06c	*mcp6*	horsetail movement protein Hrs1/Mcp6
SPBC609.05	*pob3*	histone H2A-H2B chaperone, FACT complex subunit Pob3
SPBC651.11c	*apm3*	AP-3 adaptor complex subunit Apm3 (predicted)
SPBC660.06		WW domain containing conserved fungal protein
SPBC660.14	*mik1*	mitotic inhibitor kinase Mik1
SPBC6B1.04	*mde4*	microtubule-site clamp monopolin complex subunit Mde4
SPBC6B1.06c	*ubp14*	Lys48-specific deubiquitinase Ubp14
SPBC725.02	*mpr1*	histidine-containing response regulator phosphotransferase Mpr1
SPBC725.11c	*php2*	CCAAT-binding factor complex subunit Php2
SPBC776.02c	*dis2*	serine/threonine protein phosphatase PP1, Dis2
SPBC776.16	*mis20*	centromere protein Mis20/Eic2
SPBC800.03	*clr3*	histone deacetylase (class II) Clr3
SPBC83.03c	*tas3*	RITS complex subunit 3
SPBC902.02c	*ctf18*	Ctf18 RFC-like complex subunit Ctf18

Systematic ID	Symbol	Description
SPBC902.04	*rmn1*	RNA-binding protein
SPBC902.06	*mto2*	gamma tubulin complex linker Mto2
SPBP16F5.03c	*tra1*	SAGA complex phosphatidylinositol pseudokinase Tra1
SPBP22H7.05c	*abo2*	ATPase with bromodomain protein (predicted)
SPBP23A10.05	*ssr4*	SWI/SNF and RSC complex subunit Ssr4
SPBP35G2.08c	*air1*	TRAMP complex zinc knuckle subunit Air1
SPBP35G2.10	*mit1*	SHREC complex ATP-dependent DNA helicase subunit Mit1
SPBP35G2.13c	*swc2*	Swr1 complex subunit Swc2
SPBP8B7.07c	*set6*	histone lysine methyltransferase Set6 (predicted)
SPBP8B7.23	*rnf10*	ubiquitin-protein ligase E3 (predicted)
SPBP8B7.28c	*stc1*	CLRC ubiquitin ligase complex linker protein, LIM-like Stc1
SPCC1020.12c	*xap5*	xap-5-like protein
SPCC11E10.08	*rik1*	CLRC ubiquitin ligase complex WD repeat protein Rik1
SPCC1223.13	*cbf12*	CBF1/Su(H)/LAG-1 family transcription factor Cbf12
SPCC1235.05c	*fft2*	SMARCAD1 family ATP-dependent DNA helicase Fft2 (predicted)
SPCC1235.09	*hif2*	Set3 complex subunit Hif2
SPCC1235.12c	*mug146*	Schizosaccharomyces specific protein Mug46
SPCC1259.04	*iec3*	Ino80 complex subunit Iec3
SPCC1259.07	*rxt3*	transcriptional regulatory protein Rxt3
SPCC126.02c	*pku70*	Ku domain protein Pku70
SPCC126.04c	*sgf73*	SAGA complex deubiquitinating submodule subunit Sgf73
SPCC126.07c	*asr1*	ubiquitin-protein ligase E3 Asr1 (predicted)
SPCC126.11c		RNA-binding protein, rrm type
SPCC126.13c	*sap18*	splicing factor Sap18 (predicted)
SPCC132.02	*hst2*	Sirtuin family histone deacetylase Hst2
SPCC1393.02c	*spt2*	non-specific DNA binding protein Spt2 (predicted)
SPCC1393.05	*ers1*	RNA-silencing factor Ers1
SPCC1442.13c	*sqs2*	R3H and G-patch domain protein Sqs2
SPCC1450.02	*bdf1*	Swr1 complex bromodomain subunit Bdf1
SPCC1450.03	*utp502*	ribonucleoprotein (RNP) complex Utp502 (predicted)
SPCC1494.03	*arz1*	human RAP1 GTPase-GDP dissociation stimulator ortholog, Zfs1 target number 1
SPCC162.11c	*urk1*	uridine kinase/uracil phosphoribosyltransferase (predicted)
SPCC1620.14c	*snf22*	ATP-dependent DNA helicase Snf22
SPCC1682.13	*laf2*	Clr6 associated factor 2, Laf2
SPCC16C4.11	*pef1*	Pho85/PhoA-like cyclin-dependent kinase Pef1

Systematic ID	Symbol	Description
SPCC1739.03	*hrr1*	Helicase Required for RNAi-mediated heterochromatin assembly Hrr1
SPCC1739.05	*set5*	histone lysine methyltransferase Set5 (predicted)
SPCC1739.07	*cti1*	exosome C1D family subunit Cti1
SPCC1739.12	*ppe1*	serine/threonine protein phosphatase Ppe1
SPCC1753.03c	*rec7*	meiotic recombination protein Rec7
SPCC1840.04	*pca1*	metacaspase Pca1
SPCC188.07	*ccq1*	shelterin complex HEAT repeat subunit Ccq1
SPCC188.13c	*dcr1*	dicer
SPCC18B5.03	*wee1*	M phase inhibitor protein kinase Wee1
SPCC18B5.07c	*nup61*	nucleoporin Nup61
SPCC24B10.07	*gad8*	AGC family protein kinase Gad8
SPCC24B10.08c	*ada2*	SAGA complex subunit Ada2
SPCC24B10.14c	*xlf1*	XRCC4-like nonhomologous end joining factor, Cernunnon Xlf1/Nej1
SPCC24B10.19c	*nts1*	Clr6 histone deacetylase complex subunit Nts1
SPCC297.03	*ssp1*	Ca2+/calmodulin-dependent (CaMMK)-like protein kinase Ssp1
SPCC297.04c	*set7*	histone lysine H3-K37 methyltransferase Set7
SPCC2H8.05c	*dbl1*	double strand break localizing Dbl1
SPCC306.04c	*set1*	histone lysine H3-K4 methyltransferase Set1
SPCC31H12.08c	*ccr4*	CCR4-Not complex 3'-5'-exoribonuclease subunit Ccr4
SPCC330.01c	*rhp16*	Rad16 homolog ATP-dependent DNA helicase/ ubiquitin protein ligase E3 Rhp16
SPCC330.02	*rhp7*	Rad7 homolog Rhp7
SPCC338.16	*pof3*	F-box protein Pof3
SPCC364.02c	*bis1*	splicing factor Bis1
SPCC364.06	*nap1*	histone H2A-H2B chaperone Nap1
SPCC417.07c	*mto1*	gamma tubulin complex linker Mto1
SPCC417.09c		transcription factor (predicted)
SPCC4B3.12	*set9*	histone lysine H4-K20 methyltransferase Set9
SPCC4G3.15c	*not2*	CCR4-Not complex NOT box subunit Not2
SPCC4G3.19	*alp16*	gamma tubulin complex subunit Alp16
SPCC548.05c	*dbl5*	ubiquitin-protein ligase E3 Dbl5
SPCC550.12	*arp6*	actin-like protein Arp6
SPCC550.15c	*rei1*	ribosome biogenesis protein Rei1 (predicted)
SPCC553.04	*cyp9*	WD repeat containing cyclophilin family peptidyl-prolyl cis-trans isomerase Cyp9 (predicted)
SPCC576.13	*swc5*	Swr1 complex subunit Swc5
SPCC594.05c	*spf1*	Set1C ubiquitin-protein ligase E3 subunit Spf1
SPCC61.02	*spt3*	SAGA complex subunit Spt3

Systematic ID	Symbol	Description
SPCC613.12c	*raf1*	CLRC ubiquitin ligase complex WD repeat subunit Raf1/Dos1
SPCC622.15c		Schizosaccharomyces specific protein
SPCC622.16c	*epe1*	Jmjc domain chromatin associated protein Epe1
SPCC622.19	*jmj4*	peptidyl-lysine 3-dioxygenase activity jmj4 (predicted)
SPCC645.13	*bye1*	transcription elongation regulator Bye1 (predicted)
SPCC663.11	*saf1*	splicing associated factor Saf1
SPCC663.12	*cid12*	poly(A) polymerase Cid12
SPCC736.08	*cbf11*	CBF1/Su(H)/LAG-1 family transcription factor Cbf11
SPCC736.11	*ago1*	argonaute
SPCC757.09c	*rnc1*	KH domain RNA-binding protein Rnc1
SPCC895.06	*elp2*	elongator complex WD repeat protein Elp2 (predicted)
SPCC895.07	*alp14*	TOG/XMAP215 microtubule plus end tracking polymerase Alp14
SPCC970.07c	*raf2*	CLRC ubiquitin ligase complex subunit Raf2

Table 3.2: Yeast strains used in this study

Strain	Genotype
PAS075	Locus2*::ade6p::3xE2C:hygMX* at Locus2 (between SPBC1711.11 and SPBC1711.12)
PM003	Wild-type strain: h(+); ura4-D18; leu1-32; ade6-M216; his7-366
PM006	972 h- wild-type
PAS193	Δ*K*::*ade6p*:mKO2; *ade6p*:SF-GFP between *REIII* and mat3M; *ade6p*:3xE2C*:hygMX* at Locus2; *clr4::kanMX,* h(-)
PAS216	*cenH*::*ade6p*:SFGFP(Kint2); *mat3m(Eco*RV):: *ade6p*:mKO2; *ade6p*:3xE2C:*hygMX* at Locus2; *clr4::kanMX,* h90
PAS217	cenH: ade6p:SFGFP(Kint2); mat3m(EcoRV):: ade6p:mKO2; ade6p:3xE2C:hygMX at Locus2, h90
PAS231	*ura4::natMX:dh:ade6p:*SF-GFP, *ade6p*:mKO2 3 kb, *leu1*::*ade6p*:3xE2C: *hygMX*
PAS331	*cenH*:: *ade6p*:SF-GFP (Kint2); *mat3m(Eco*RV):: *ade6p*:mKO2; *ade6p*:3xE2C:*hygMX* at Locus2; Δ*REIII*::REIII(Δs1, Δs2) in *clr4::kanMX,* h90
PAS332	*cenH*:: *ade6p*:SFGFP(Kint2); *mat3m(Eco*RV):: *ade6p*:mKO2; *ade6p*:3xE2C:*hygMX* at Locus2; Δ*REIII*::REIII(Δs1, Δs2), h90
PAS482	Δ*K*::*ade6p*:mKO2; *ade6p:* SF-GFP between *REIII* and mat3M; *ade6p*:3xE2C*:hygMX* at Locus2, h(-); 'OFF' allele
PAS795	*cenH*: *ade6p*:SFGFP(Kint2); *mat3m(Eco*RV):: *ade6p*:mKO2; *ade6p*:3xE2C: *hygMX* at Locus2; *fkh2::natMX*
PAS796	*cenH*: *ade6p*:SFGFP(Kint2); *mat3m(Eco*RV):: *ade6p*:mKO2; *ade6p*:3xE2C: *hygMX* at Locus2; *prw1::kanMX*
PAS797	*cenH*: *ade6p*:SFGFP(Kint2); *mat3m(Eco*RV):: *ade6p*:mKO2; *ade6p*:3xE2C: *hygMX* at Locus2; *png3::kanMX*

| Strain	Genotype
PAS798	*cenH*:: *ade6p:*SFGFP(Kint2); *mat3m(Eco*RV):: *ade6p:*mKO2; *ade6p:*3xE2C:*hygMX* at Locus2; Δ*REIII*::REIII(Δs1, Δs2); *fkh2::natMX*
PAS799	*cenH*:: *ade6p:*SFGFP(Kint2); *mat3m(Eco*RV):: *ade6p:*mKO2; *ade6p:*3xE2C:*hygMX* at Locus2; Δ*REIII*::REIII(Δs1, Δs2); *prw1::kanMX*
PAS800	*cenH*:: *ade6p:*SFGFP(Kint2); *mat3m(Eco*RV):: *ade6p:*mKO2; *ade6p:*3xE2C:*hygMX* at Locus2; Δ*REIII*::REIII(Δs1, Δs2); *png3::kanMX*
PAS803	Δ*K*::*ade6p:*mKO2; *ade6p:* SF-GFP between *REIII* and mat3M; *ade6p:*3xE2C*: hygMX* at Locus2; 'OFF' allele; *fkh2::natMX*
PAS804	Δ*K*::*ade6p:*mKO2; *ade6p:* SF-GFP between *REIII* and mat3M; *ade6p:*3xE2C*: hygMX* at Locus2; 'OFF' allele; *prw1::kanMX*
PAS805	Δ*K*::*ade6p:*mKO2; *ade6p:* SF-GFP between *REIII* and mat3M; *ade6p:*3xE2C*: hygMX* at Locus2; 'OFF' allele; *png3::kanMX*
PAS808	*cenH*: *ade6p:*SFGFP(Kint2); *mat3m(Eco*RV):: *ade6p:*mKO2; *ade6p:*3xE2C: *hygMX* at Locus2; *fkh2::natMX; prw1::kanMX* by cross
PAS809	*cenH*: *ade6p:*SFGFP(Kint2); *mat3m(Eco*RV):: *ade6p:*mKO2; *ade6p:*3xE2C: *hygMX* at Locus2; *fkh2::natMX; prw1::kanMX* by sequential knockout
PAS810	*cenH*:: *ade6p:*SFGFP(Kint2); *mat3m(Eco*RV):: *ade6p:*mKO2; *ade6p:*3xE2C:*hygMX* at Locus2; Δ*REIII*::REIII(Δs1, Δs2); *fkh2::natMX; prw1::kanMX* by cross
PAS811	*cenH*:: *ade6p:*SFGFP(Kint2); *mat3m(Eco*RV):: *ade6p:*mKO2; *ade6p:*3xE2C:*hygMX* at Locus2; Δ*REIII*::REIII(Δs1, Δs2); *fkh2::natMX; prw1::kanMX* by sequential knockout
PAS813	*cenH*:: *ade6p:*SFGFP(Kint2); *mat3m(Eco*RV):: *ade6p:*mKO2; *ade6p:*3xE2C:*hygMX* at Locus2; Δ*REIII*::REIII(Δs1, Δs2); *apm3::natMX*
PAS816	apl5:SF-GFP:*hygMX*; Swi6:E2C:*kanMX*
PAS817	apm3:SF-GFP:*hygMX;* Swi6:E2C:*kanMX*

SUPPLEMENTAL FIGURES
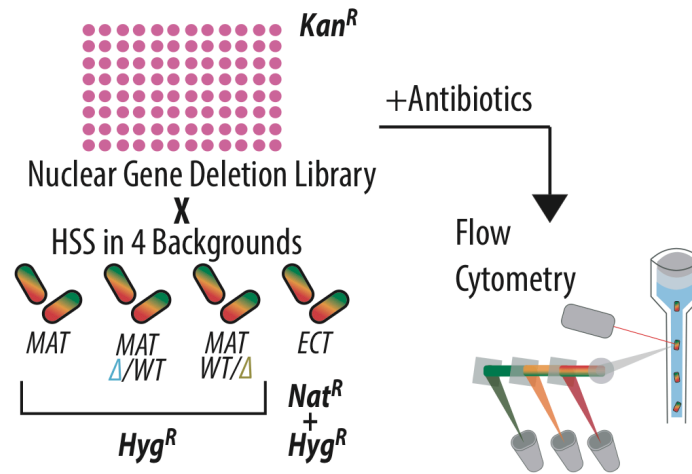
Figure 3.1 Supplement 1: Screen overview.

Schematic of the screen to identify genes that contribute to heterochromatin nucleation and spreading. A custom nuclear function deletion library (Table 3.1) was mated with four different reporter strains (*WT MAT, MAT ΔcenH, MAT ΔREIII and ECT*). The fluorescence of "green", "orange" and "red" for each mutant cell within each background are recorded by flow cytometry.
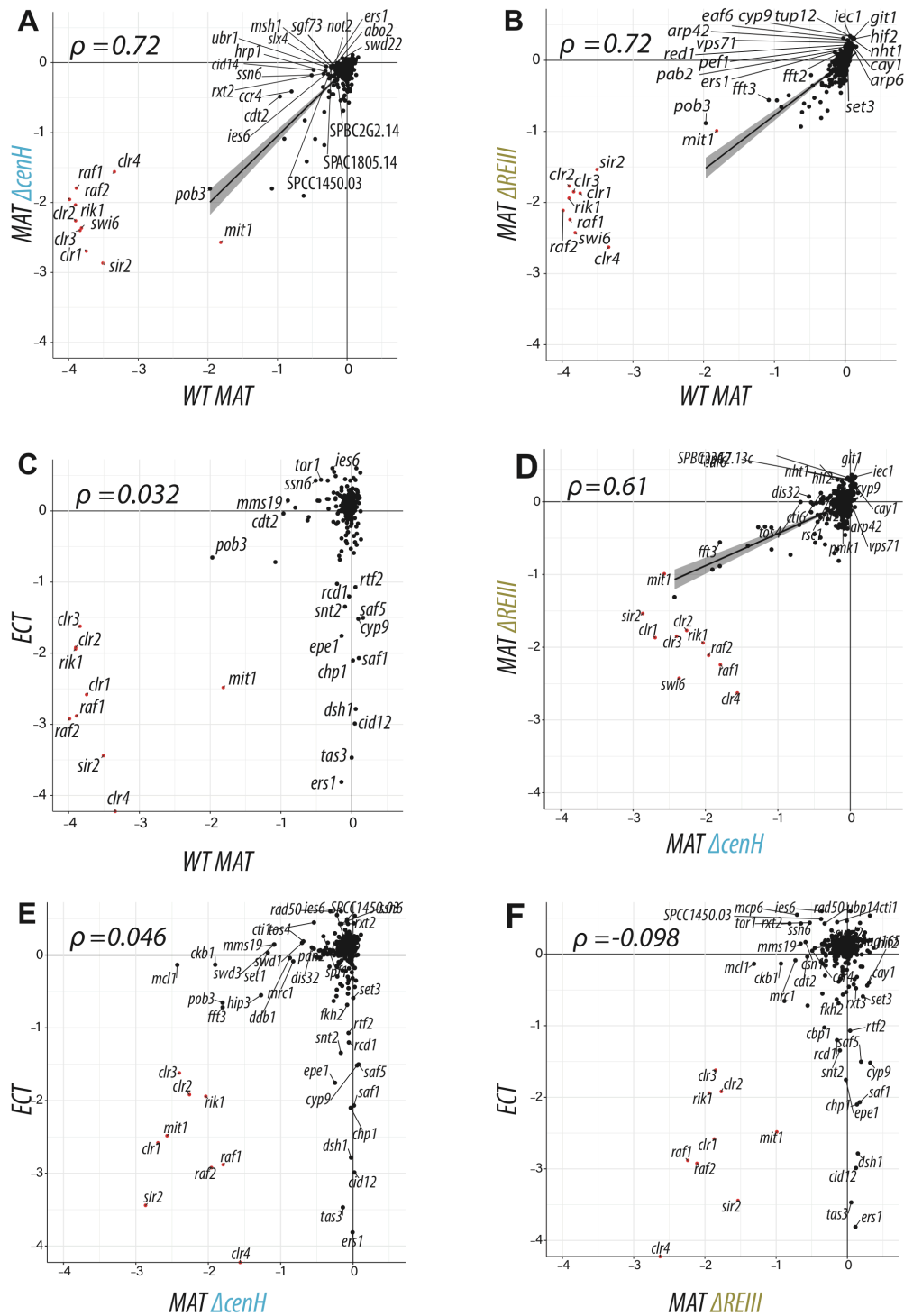
Figure 3.1 Supplement 2: Pairwise plots depicting the comparisons of mutant EMD[mut/par].

Pairwise plots depicting the EMD[mut/par] comparisons of (**A**) *MAT ΔcenH* with *WT MAT,* (**B**) *MAT ΔREIII* with *WT MAT,* (**C**) *ECT* with *WT MAT,* (**D**) *MAT ΔREIII* with *MAT ΔcenH,* (**E**) *ECT* with *MAT ΔcenH* and (**F**) *ECT* with *MAT ΔREIII*. Pearson correlation coefficients (ρ), determined without SHREC, ClrC mutants and *Δsir2*, were calculated and correlations plotted for A, B, D.
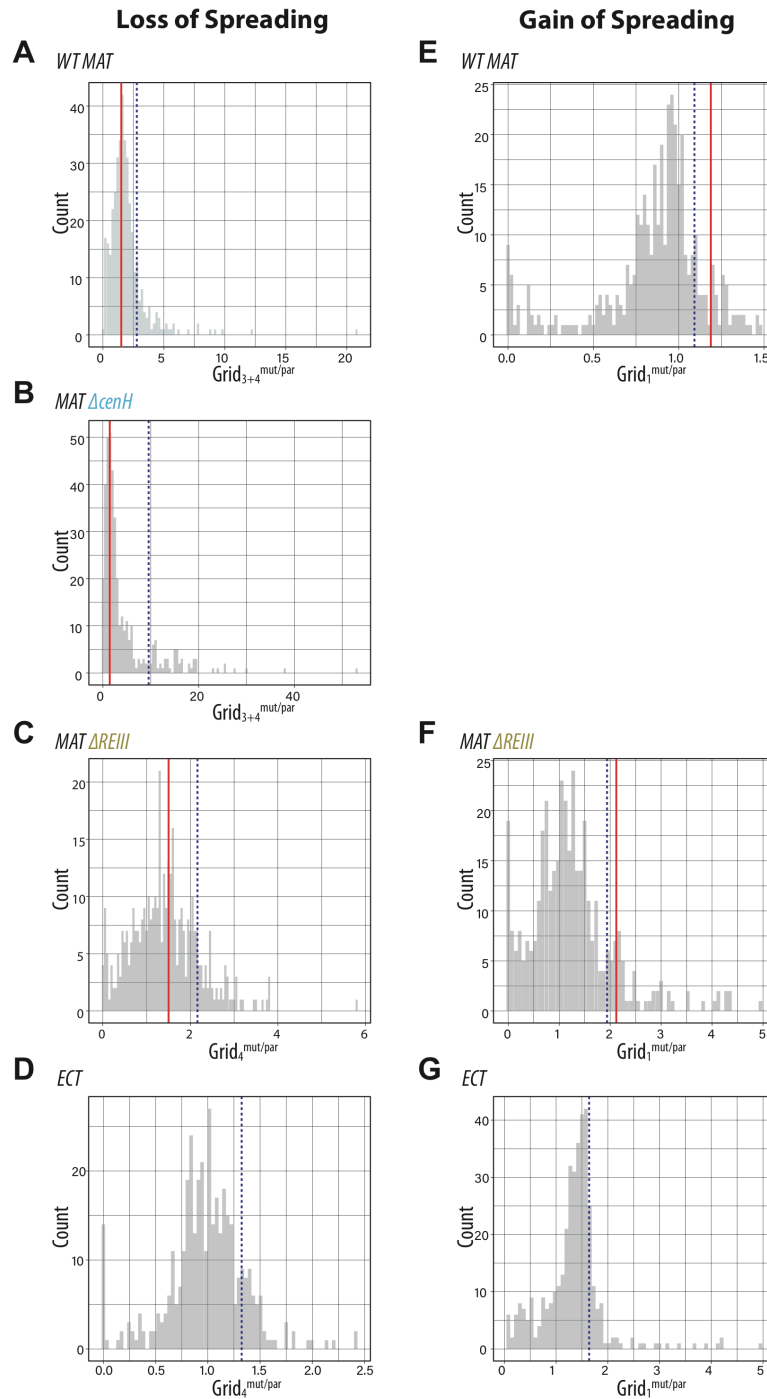
Figure 3.2 Supplement 1: The distribution of $Grid_n^{mut/par}$ of all mutants in the loss or gain of spreading hit categories.

A.-D. Histograms representing the distribution of $Grid_n^{mut/par}$ of (**A**) *WT MAT*, (**B**) *MAT ΔcenH*, (**C**) *MAT ΔREIII* and (**D**) *ECT* in the loss of spreading hit category. The red line indicates two standard deviation above median of parent isolates. The dashed blue line indicates the 85th percentile. **E.-G.** Histograms representing the distribution of $Grid_n^{mut/par}$ of (**E**) *WT MAT,* (**F**) *MAT ΔREIII* and (**G**) *ECT* in the gain of spreading hit category. The red line indicates two standard deviation above median of parent isolates. The dashed blue line indicates the 85th percentile.
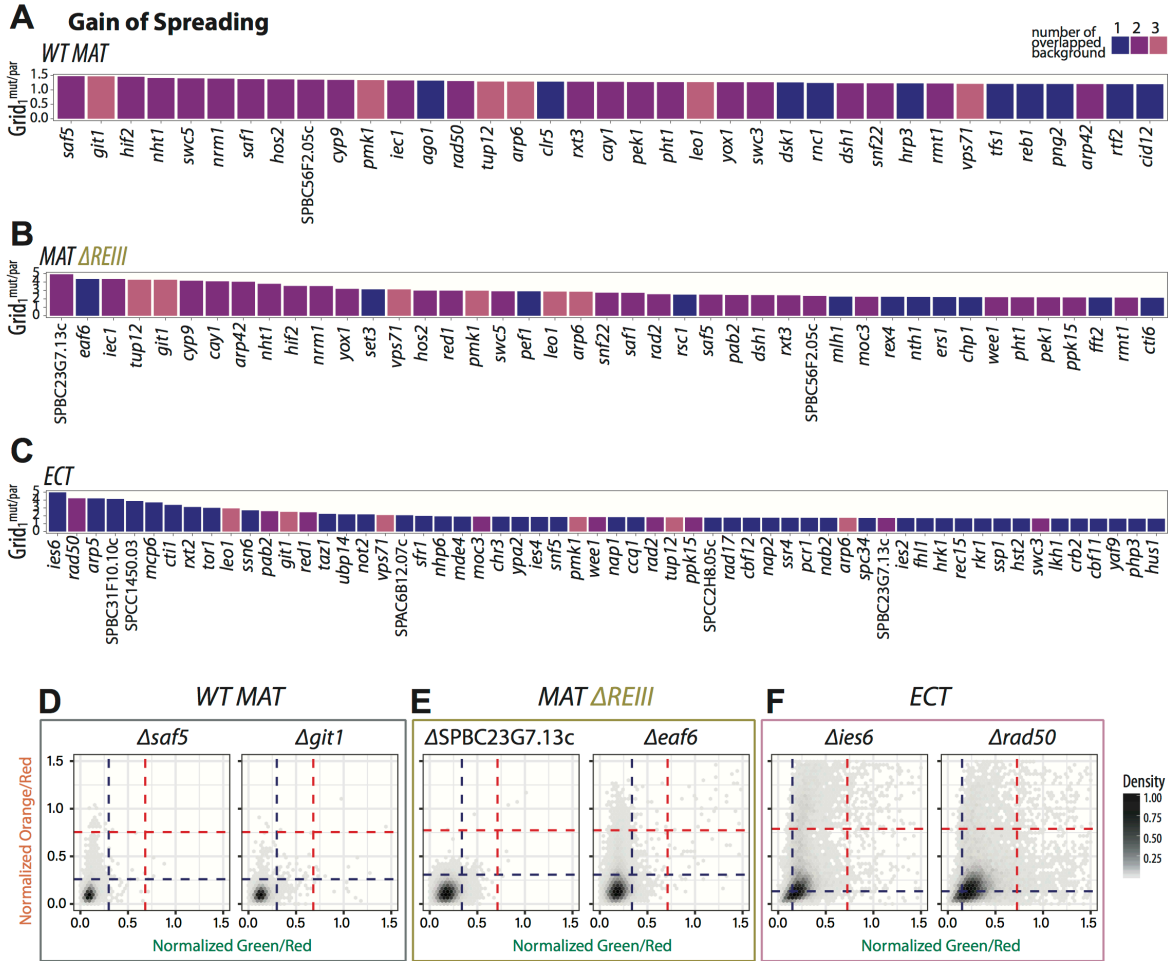
126

Figure 3.2 Supplement 2: Grid$_n^{mut/par}$ and overlapped chromatin contexts of identified gain of spreading mutants

A.-C. Bar graphs representing the Grid$_n^{mut/par}$ and number of overlapped chromatin context(s) of gene hits with Gain of Spreading phenotype in (**A**) *WT MAT*, (**B**) *MAT ΔREIII* and (**C**) *ECT* chromatin contexts respectively. The genes are ranked in descending order of Grid$_n^{mut/par}$ and the color of each bar represents the number of backgrounds the mutant show the same phenotype. **D.-F.** 2D-density hexbin plots of the top two gain of spreading gene hits for all 3 chromatin contexts, (**D**) *WT MAT*, (**E**) *MAT ΔREIII,* and (**F**) *ECT*. Plotting as in Figure 3.1, C-F. For wild-type comparison, see Figure 3.1 C-F.
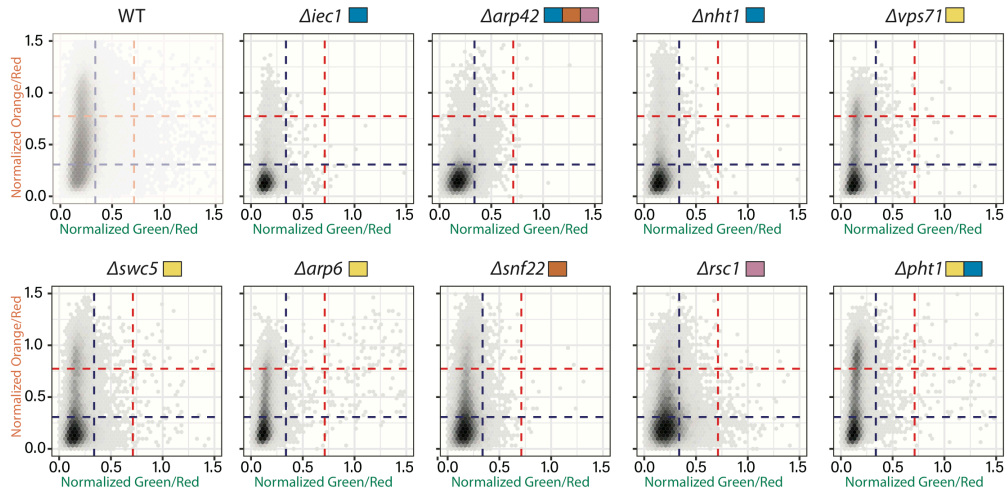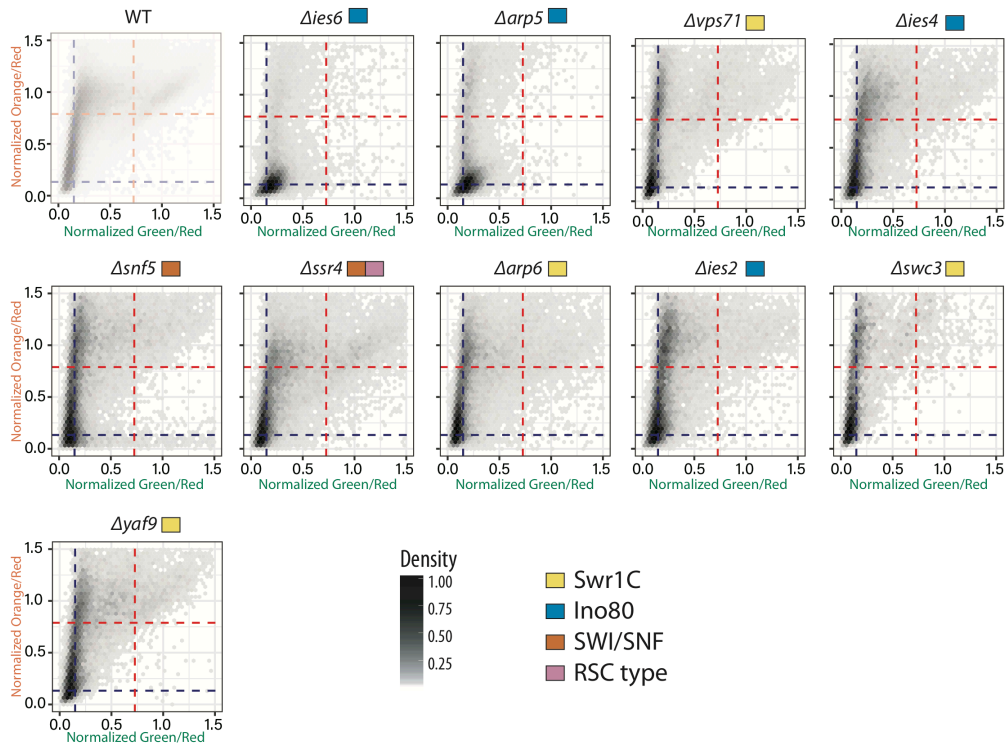
Figure 3.3 Supplement 1: Original 2D density histograms for chromatin remodeler gain of spreading hits.

A. Chromatin remodeler gain of spreading gene hits for *MAT ΔREIII.* Original 2D density histograms for all the gain of spreading gene hits from Figure 3.3B are shown. **B.** Chromatin remodeler gain of spreading gene hits for *ECT.* Original 2D density histograms for all the gain of spreading gene hits from Figure 3.3B are shown. GO complex annotations are indicated next to each mutant. 2D density histograms as in Figure 3.1 C-F. Original *MAT ΔREIII* and *ECT* parents shown in Figure 3.1 are reproduced here again (with transparency) for comparison. Mutants within each chromatin next are shown in descending order of their $Grid_n^{mut/par}$ values.
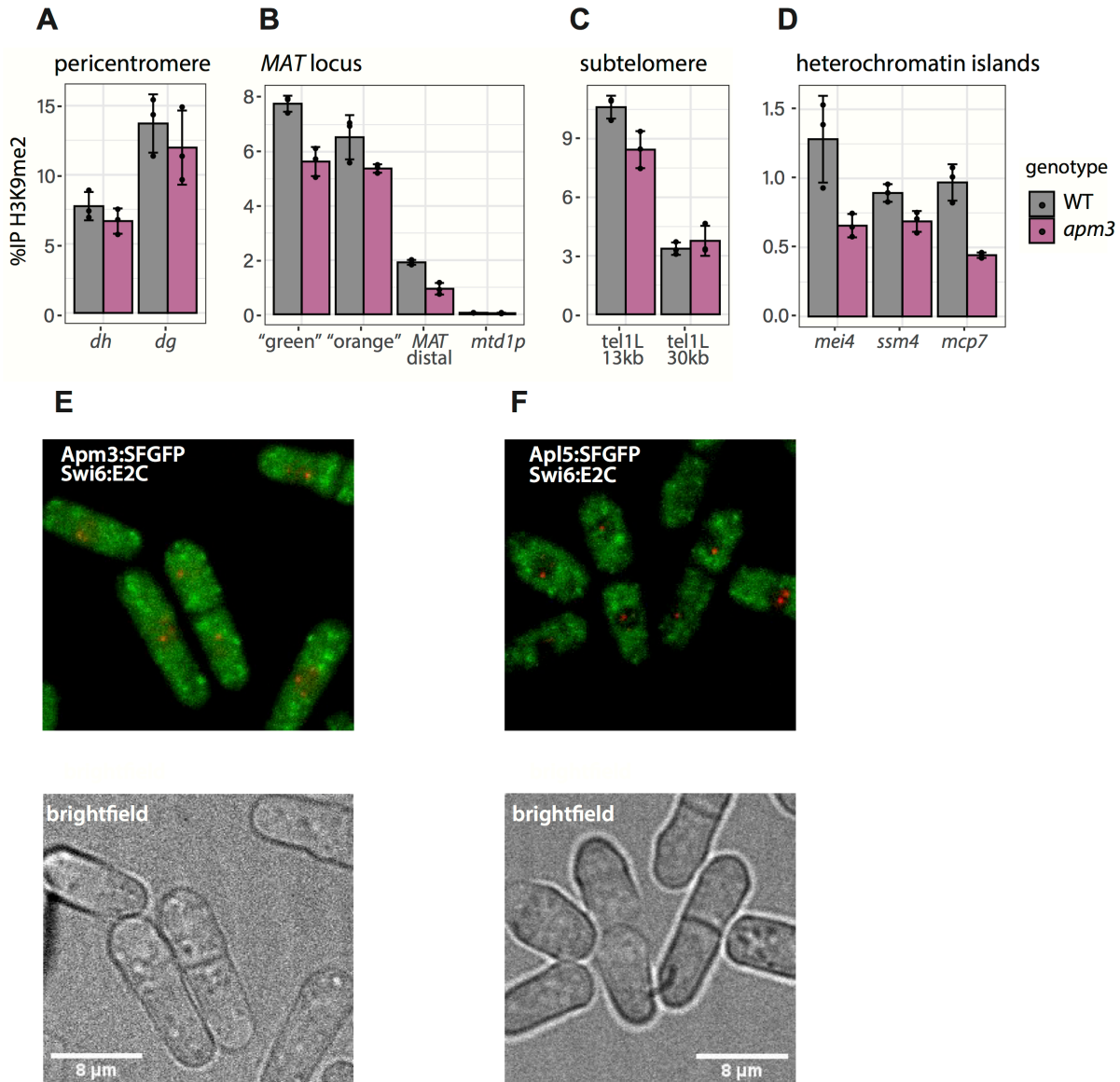
Figure 3.3 Supplement 2: *Δapm3* has a mild heterochromatin spreading phenotype and the protein is distributed in the cytosol and nucleus.

A.-D. ChIP-qPCR quantification of H3K9me2 signal in the *MAT ΔREIII* strain at constitutive and facultative heterochromatin regions in wild-type and an independently generated *Δapm3* alleles. Error bars represent 1SD from three biological replicate isolates. Individual values are plotted for each isolate. The WT data is additionally replicated in Figure 3.4 as these experiments were performed together. **E.** Apm3:SFGFP is distributed in the cytosol and nucleus. Apm3:SFGFP was expressed from its native locus and co-expressed with Swi6:E2C. Swi6:E2C labels nuclear heterochromatin. Z-projection overlays of the Apm3:SFGFP and Swi6:E2C on top, and a brightfield image on the bottom. **F.** Apl5:SFGFP is largely nuclear excluded. Apl5:SFGFP was expressed from its native locus and co-expressed with Swi6:E2C. Swi6:E2C labels nuclear heterochromatin. Z-projection overlays of the Apl5:SFGFP and Swi6:E2C on top, and a brightfield image on the bottom.
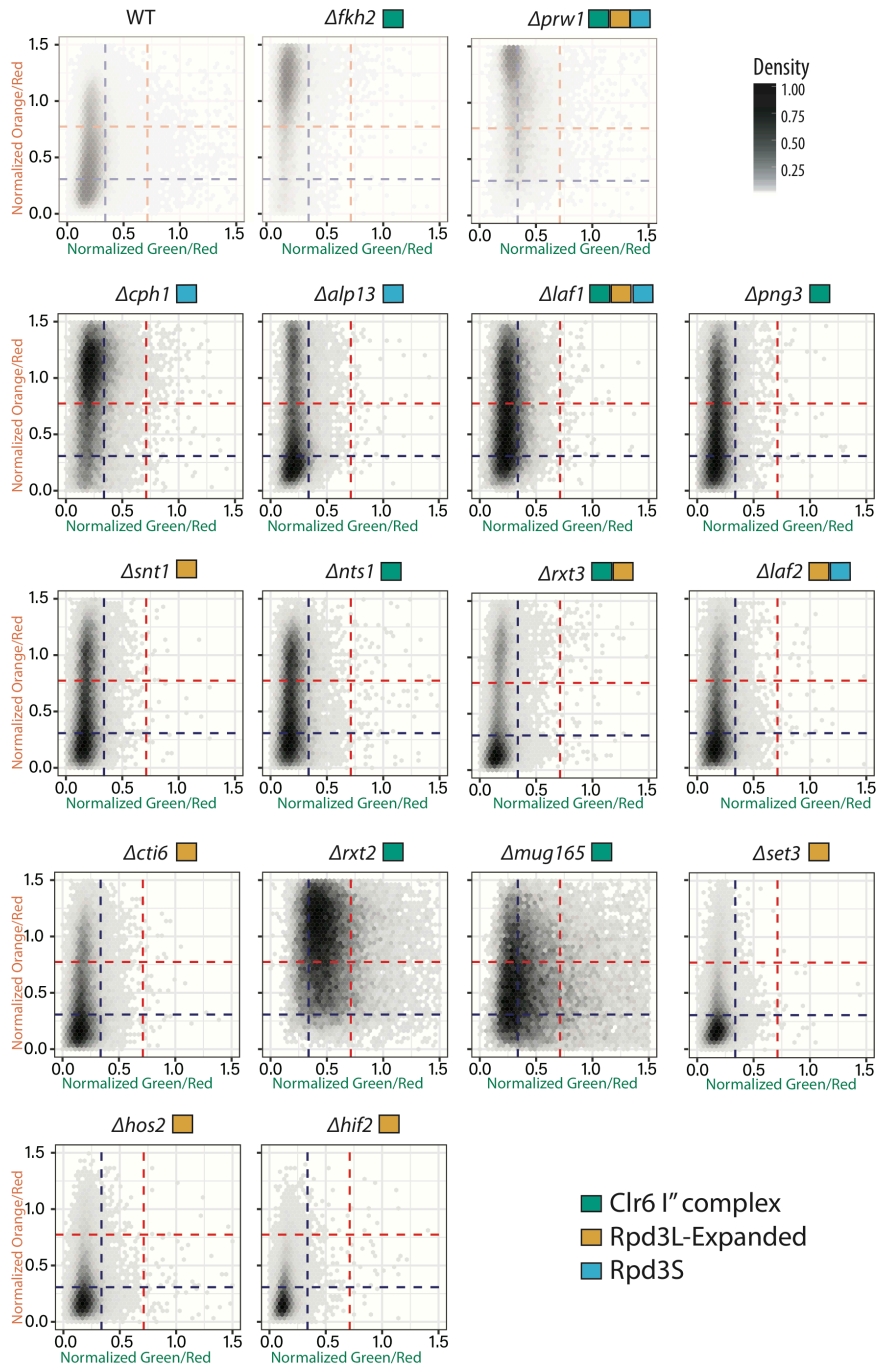
Figure 3.3 Supplement 3: Original 2D density histogram for all Clr6 complex subunit mutants in *MAT ΔREIII*.

Original 2D density histograms of all Clr6 complexes gene mutants corresponding to Figure 3.3C in *MAT ΔREIII* context. Original *MAT ΔREIII* wild type parent and *Δfkh2* and *Δprw1* mutants shown in Figures 3.1 and 3.2 are reproduced here again (with transparency) for comparison.
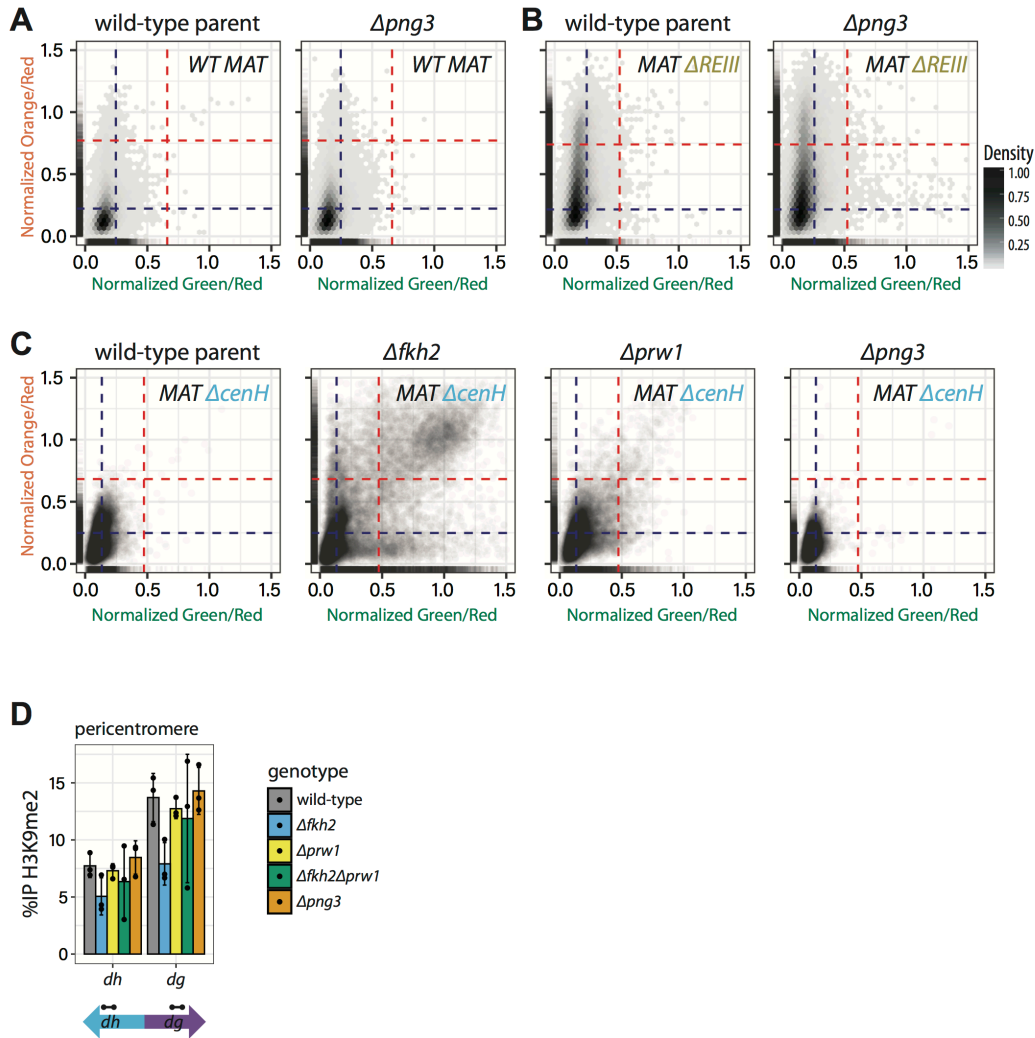
Figure 3.4 Supplement 1: The effect of Clr6 I" in the *MAT ΔcenH* background.

A. 2D-density hexbin plots with rug as in Figure 3.4C for *Δpng3* and WT control run on the same day in the *WT MAT* strain background. **B.** 2D-density hexbin plots with rug as in Figure 3.4C for *Δpng3* and WT control run on the same day in the *MATΔREIII* strain background. **C.** Scatter plot with rug for WT and Clr6 I" mutants in the *MAT ΔcenH* strain background. Cell are plotted as individual points versus summarized in 2D density hexbin plots for increased resolution. Points and rug lines are colored with partial transparency to assist with visualization of density changes. **D.** ChIP-qPCR quantification of H3K9me2 signal in the *MAT ΔREIII* strain at the pericentromeric heterochromatin region. Error bars represent 1SD from three biological replicate isolates. Individual values are plotted for each isolate. SD; standard deviation.

4.   Set1/COMPASS repels heterochromatin invasion at euchromatic sites by disrupting Suv39/Clr4 activity and nucleosome stability.

FOREWORD

The following is a reproduced from a previously published manuscript (Greenstein et al. 2019).

This chapter explores the biological mechanisms that constrain the spatial extent of

heterochromatin spreading during euchromatin invasion as a model for beginning to understand

how the differential spreading that patterns genomes during development to direct cell fate

decisions is regulated. Specifically, this chapter addresses the long-understood phenomenon of

mutually exclusive eu- and heterochromatic signals and apparently mutually antagonistic

activities and suggests a mechanism for gene-based mechanisms of spreading containment.

ABSTRACT

Protection of euchromatin from invasion by gene-repressive heterochromatin is critical for

cellular health and viability. In addition to constitutive loci such as pericentromeres and

subtelomeres, heterochromatin can be found interspersed in gene-rich euchromatin, where it

regulates gene expression pertinent to cell fate. While hetero- and euchromatin are globally

poised for mutual antagonism, the mechanisms underlying precise spatial encoding of

heterochromatin containment within euchromatic sites remain opaque. We investigated ectopic

heterochromatin invasion by manipulating the fission yeast mating type locus boundary, using a

single-cell spreading reporter system. We found that heterochromatin repulsion is locally

encoded by Set1/COMPASS on certain actively transcribed genes and that this protective role is

most prominent at heterochromatin islands, small domains interspersed in euchromatin that

regulate cell fate specifiers. Sensitivity to invasion by heterochromatin, surprisingly, is not

dependent on Set1 altering overall gene expression levels. Rather, the gene protective effect is strictly dependent on Set1's catalytic activity. H3K4 methylation, the Set1 product, antagonizes spreading in two ways: Directly inhibiting catalysis by Suv39/Clr4, and locally disrupting nucleosome stability. Taken together, these results describe a mechanism for spatial encoding of euchromatic signals that repel heterochromatin invasion.

INTRODUCTION

Heterochromatin is a conserved nuclear ultrastructure (Rea et al. 2000), which enacts genome partitioning by repressing transcription and recombination at repetitive sequences and structural elements, as well as genetic information not pertaining to the specified cell fate. Once seeded at specific sequences (Hall et al. 2002; Jia et al. 2004; Reyes-Turcu et al. 2011), heterochromatin is subsequently propagated in cis over qualitatively distinct regions of the chromosome in a process termed spreading. Positional regulation of heterochromatin is key to determining and remembering cell fate decisions. Boundary regions often separate adjacent heterochromatin and euchromatin domains, reinforcing the distinct signals and functional environments on each side and countering the intrinsic propensity for heterochromatin to invade and silence genes. Major mechanisms of boundary formation fall into three broad classes: (1) recruitment of factors that directly antagonize the opposite state, for example by removal of state-specific signals on chromatin (Ayoub et al. 2003; Schlichter and Cairns 2005; Lan et al. 2007; Trewick et al. 2007; Braun et al. 2011). (2) Promotion of the original state by either depositing or protecting such signals (Wang et al. 2013; Sadeghi et al. 2015; Verrier et al. 2015; Wang et al. 2015). (3) Structural constraint via recruitment of DNA binding proteins that tether heterochromatin regions to the nuclear periphery (Bell and Felsenfeld 1999; Kurukuti et al. 2006; Noma et al.

2006). Despite the varied modalities employed in boundary formation, containment is not absolute. This is evidenced by the observation that boundaries can be overcome by modest dosage changes in heterochromatin factors (Noma et al. 2006; Ceol et al. 2011), which leads to the silencing of genes critical to normal cellular function.

In addition to constitutive heterochromatin found at centromeres, telomeres, and other repetitive sequences, repressed domains also form at additional genomic locations in response to developmental and environmental signals (Wen et al. 2009; Zofall et al. 2012; Zhu et al. 2013). These facultative heterochromatin domains are often embedded in euchromatic regions and silence developmental genes in a lineage-specific manner (Wen et al. 2009). Resulting from response to changing stimuli, the final extent of facultative domains can change over time, expanding to different degrees (Wen et al. 2009) and even contracting (McDonald et al. 2011) in genomic space, though how this is achieved is not well understood. Facultative domain size may be tuned at the level of the heterochromatin spreading reaction (Hathaway et al. 2012) and/or the activities promoting its containment or disassembly. While little is known about the former, several models, beyond those known to operate at constitutive boundaries (Guelen et al. 2008; Zofall et al. 2012), could be invoked to explain the latter.

How might euchromatin regulate heterochromatin spreading at facultative sites or respond to its expansion beyond constitutive domains? One of the defining features of euchromatin is the presence of active genes. It is thought that transcription from active genes is incompatible with heterochromatin formation (Scott et al. 2006). Multiple direct effects of transcription have been proposed to interfere with heterochromatin assembly. These include nucleosome turnover (eviction) by transcribing polymerase, formation of nucleosome-depleted regions at transcriptional units, or steric interference by transcription-associated complexes

(Noma et al. 2006; Garcia et al. 2010; Aygun et al. 2013). Furthermore, we understand that unique molecular signatures characterize eu- and heterochromatin states and are critical to their formation. Heterochromatin is marked by methylation of histone 3 at lysine 9 or lysine 27 (H3K9me and H3K27me, respectively) and hypoacetylation of various histone lysine residues. In contrast, euchromatin features H3K4me, H3K36me and histone hyperacetylation (Nielsen et al. 2001; Guelen et al. 2008). Multiple studies have documented the apparent mutual exclusion of H3K9me- and H3K4me- marked regions (Litt et al. 2001; Noma et al. 2001; Cam et al. 2005; Guelen et al. 2008) and the requirement for removal of signals associated with the opposite state (Lan et al. 2007; Li et al. 2008). While we are beginning to understand how this dichotomy is formed, it still remains unclear whether this is a cause or consequence of separating heterochromatin and euchromatin.

We aimed to investigate the role of euchromatic signals in regulating the extent of spreading in fission yeast, a well-characterized model system for the study of heterochromatin formation, which shares critical features with the processes found in metazoans. Fission yeast form constitutive heterochromatin marked by H3K9me at centromeres, telomeres, and the mating type (MAT) locus. Boundary formation occurs at peri-centromeric regions and the MAT locus via at least two mechanisms – tethering to the nuclear periphery through binding of TFIIIC proteins to B-box element sequences in boundary regions (Noma et al. 2006) as well as specific enrichment of a JmjC domain-containing protein, Epe1 (Ayoub et al. 2003; Zofall and Grewal 2006; Trewick et al. 2007; Braun et al. 2011), which recruits additional downstream boundary effectors. In addition to these constitutive sites, facultative heterochromatin forms at developmentally regulated meiotic genes in regions surrounded by canonical euchromatin, which are partially dependent on Epe1 for containment (Zofall et al. 2012; Wang et al. 2015). Utilizing

135

the well-characterized MAT locus boundary as a model for euchromatic invasion, we found that active genes units could repel spreading and that this function depends on the H3K4 methylase complex Set1/COMPASS. Set1 is the catalytic subunit of COMPASS and is responsible for mono-, di-, and tri-methylation of H3K4 *in vivo*. It is recruited by RNA polymerase and forms a characteristic pattern of H3K4 methylation states over genes, with H3K4me3 near the transcription start site (TSS) and H3K4me2 in the gene body (reviewed in (Shilatifard 2012)). We show that rather than acting as a global antagonist of spreading, like Epe1 or the histone acetyltransferase Mst2 (Wang et al. 2015), Set1 regulates spreading at gene-rich environments such as heterochromatin islands. Set1 does not exert its euchromatin protective function by modulating steady state transcript levels. Rather, it acts via two separate mechanisms, both dependent on its catalytic activity: (1) the disruption of nucleosome stability and (2) catalytic inhibition of the sole fission yeast H3K9 methylase Suv39/Clr4, by the Set1 product H3K4me. This study provides a mechanism for the encoding of spatial cues within euchromatin that contain heterochromatin expansion.

RESULTS

Genes can function as a barrier to heterochromatin spreading.

To investigate heterochromatin invasion into euchromatin, we employed our previously described heterochromatin spreading sensor (HSS) (Al-Sady et al. 2016; Greenstein et al. 2018) in the euchromatic region proximal to the MAT Inverted Repeat Right (IR-R) boundary (Ayoub et al. 2000). This HSS system contains two central components: (1) the spreading sensor, a monomeric Kusabira-Orange 2 fluorescent protein driven by the validated *ade6* promoter, hereafter referred to as "orange", integrated 0.7kb outside IR-R, and (2) the control, a E2Crimson

fluorescent protein driven by the same promoter, hereafter referred to as "red", integrated at a constitutive euchromatic locus (Greenstein et al. 2018) (**Figure 4.1A** and **Table 4.1**). The IR-R is a well-described boundary system that can be easily manipulated (Garcia et al. 2015). Precisely controlling its disruption leads to an excellent model system for identifying determinants within euchromatin that regulate heterochromatin spreading. With the HSS system, we use flow cytometry to capture information from tens of thousands of single cells. We divide "orange" by "red" for each cell to normalize for cell-to-cell transcription and translation noise, allowing us to quantify heterochromatin-specific gene silencing at the "orange" reporter over the population.

We first examined the normalized orange fluorescence of a strain with a WT boundary (*epe1+, B-box+*) and detect no silencing in the population distribution (**Figure 4.1B**), as expected (see legend to Figure 4.1 and (Greenstein et al. 2018)). We define a threshold for silencing as the mean of the appropriate WT (*epe1+*) strain less two standard deviations (see dashed red line). We next compromised one or both of the pathways required for containment of spreading at IR-R (Ayoub et al. 2003; Trewick et al. 2007; Garcia et al. 2015) and assessed the effect on "orange" silencing. Consistent with previous results (Garcia et al. 2015), little silencing is detected in *Δepe1* isolates harboring a partially compromised boundary (referred to hereafter as boundary[C], solid line histograms) (**Figure 4.1C**). In a fully compromised boundary, absent both *epe1* and the 5 *B-box* sequence elements contained within IR-R (Noma et al. 2006) (referred to hereafter as Δboundary, dashed line histograms), we detected increased silencing (**Figure 4.1D**). Yet, even in the Δboundary background, greater than 80% of cells in the population fully express "orange". Given this result, and the observation that H3K9me2 spreading declines sharply over endogenous IR-R bordering genes (Garcia et al. 2015), we wondered whether other activities beyond boundaries, possibly centered on active genes, repel spreading.
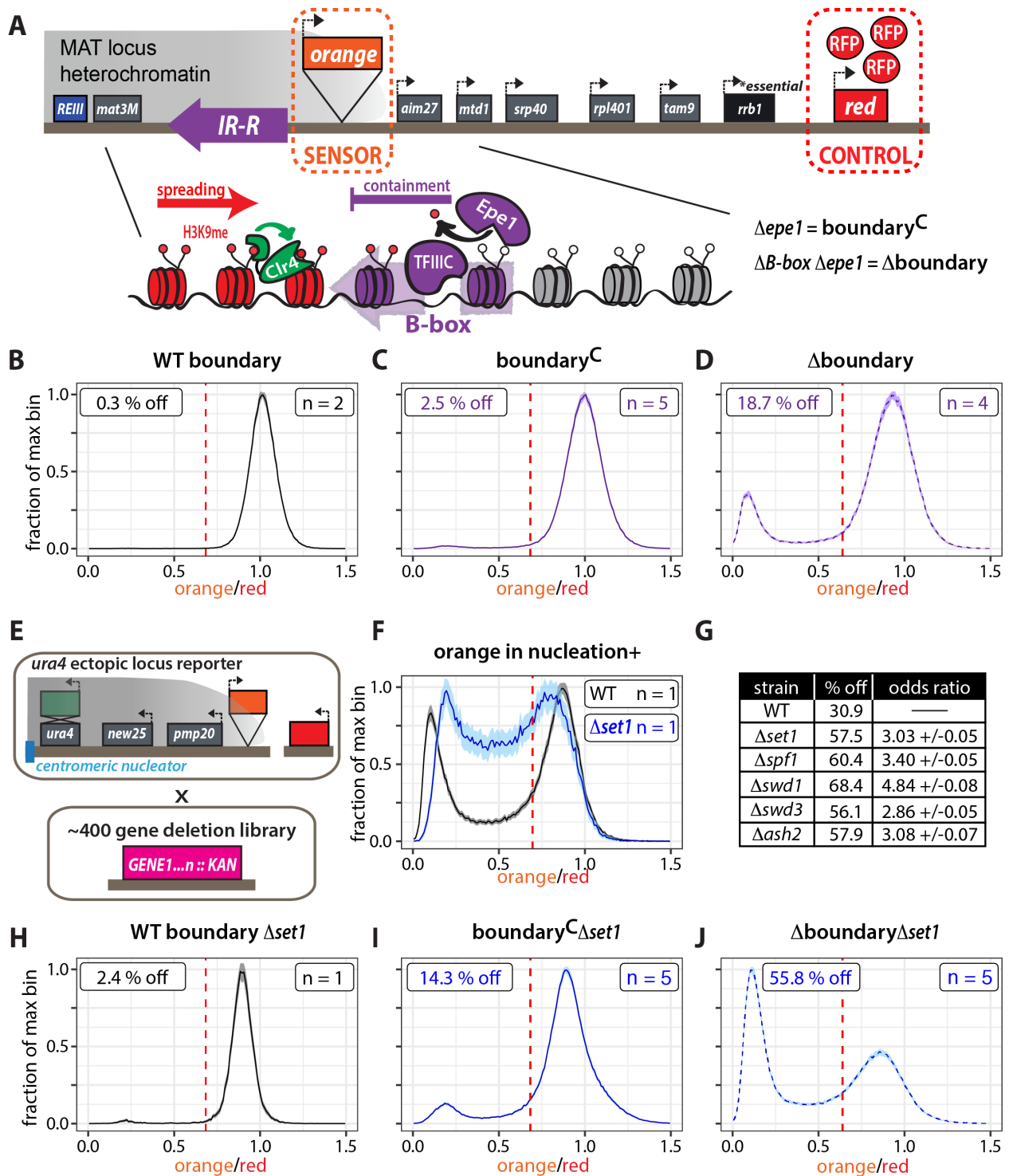
Figure 4.1: Genes repel heterochromatin across boundaries in a manner dependent on Set1/COMPASS.

A. An overview of the heterochromatin spreading sensor (HSS) outside the MAT locus IR-R boundary with transcriptional reporters encoding fluorescent proteins as sensor ("orange") and control ("red"). IR-R (depicted as purple arrow) employs at least two independent pathways

138

dependent on Epe1 and TFIIIC, respectively, to contain spreading of H3K9 methylation via Suv39/Clr4. IR-R function can be abrogated by deletion of *epe1* and removing the *B-box* binding sequences for TFIIIC. **B.** Histogram of "orange" signal in a WT boundary background normalized to Δ*clr4*. **C.** Histogram of "orange" signal in boundary[C] (Δ*epe1*) background normalized to the corresponding WT (*epe1+*) strain. **D.** Histogram of normalized "orange" signal in Δboundary (Δ*epe1* Δ*B-box*) background as in C. **E.** Illustration depicting genetic screen for modulators of gene-mediated heterochromatin repulsion. An HSS variant at the *ura4* locus was crossed to approximately 400 gene deletions. The resulting strains were analyzed by flow cytometry. **F.** Histograms plotted as in C. of normalized "orange" signal in nucleation-gated cells in WT and Δ*set1*. **G.** Data table, LEFT: fraction of cells that experienced silencing at "orange". Two thresholds were applied, a cutoff for nucleation at "green" and a cutoff for silencing at "orange". Cells that met both criteria were counted as repressed. RIGHT: Odds ratio, calculated by Fisher's Exact Test, comparing the odds of being the silenced "off" state for a cell in the Set1C mutant relative to wild-type populations. **H.** Histogram of normalized "orange" signal Δ*set1* in the WT boundary background as in C. **I.** Histogram of normalized "orange" signal in boundary[C]Δ*set1* background as in C. **J.** Histogram of normalized "orange" signal in WT Δboundary Δ*set1* background as in C. All 1D histograms are plotted as the mean +/- 3SD of 300 bootstrap iterations for combined data from the indicated number of biological isolates (n). Signal is normalized to the median signal from a Δ*clr4* or corresponding WT (*epe1+*) strain control to represent the maximum fluorescence in the absence of heterochromatin (x=1). A threshold for silencing (dashed red line) represents the mean signal of the WT strain less 2SD with the exception of F. where the threshold for silencing in nucleation+ cells was determined as mean less 1SD of the "orange" signal from the Δ*clr4* strain. The faction of cells below this cutoff was calculated (%off).

Set1/COMPASS regulates genic protection from heterochromatin spreading.

In order to identify potential factors that regulate gene-mediated repulsion of heterochromatin

spreading, we designed a genetic screen to query the effect of gene deletions on silencing

measured via our reporters. We conducted the screen in the context of the HSS embedded at the

euchromatic *ura4* locus (Greenstein et al. 2018), downstream of an ectopically placed RNAi-

based heterochromatin nucleator (Marina et al. 2013) **(Figure 4.1E)**. This construct can generate

spreading up to 8kb downstream over two endogenous and two reporter genes (**Figure 4.1**

Supplement A, (Greenstein et al. 2018)), representing about one third the size of the MAT

locus. Importantly, this 8kb region is not protected by a natural *cis*-encoded boundary,

eliminating the need to remove any boundary factors, which avoids confounding global effects

on growth in the screen. Since at this locus nucleation is less robust than endogenous

heterochromatin domains (Greenstein et al. 2018), we exploited the presence of a nucleator-proximal third reporter cassette encoding "green" at this locus. Based on this reporter, we can apply a computational gate to isolate successfully nucleated cells (greenOFF, described in (Greenstein et al. 2018)) and assess their spreading state at the "orange" reporter, 3kb downstream from "green". In the WT background, the nucleation gated "orange" signal in this strain resembles the behavior seen in the Δboundary IR-R HSS strain (compare **Figure 4.1F,** black line, and **Figure 4.1D**), exhibiting both gene silencing and fully expressed states.

We crossed this *ura4*-HSS background strain to a curated ~400 gene subset of the *S. pombe* deletion library enriched for nuclear factors (**Figure 4.1E**) and measured reporter fluorescence from the resultant strains via flow cytometry. For each strain, we plotted a 2D histogram of red-normalized orange versus green fluorescence (**Figure 4.1 Supplement B**) and calculated the fraction of cells that experienced silencing at "orange". Silencing in this context is defined as the fraction of all cells that met both the greenOFF criteria for nucleation (blue line) and had orange signal below the mean less 1 standard deviation of the matched *Δclr4* strain (red line).

Upon analysis of this dataset, we noticed 5 genes whose absence had the same characteristic effect of increased silencing at the spreading reporter – *ash2, swd1, swd3, spf1,* and *set1* (**Figure 4.1F, Figure 4.1G, Figure 4.1 Supplement B**). To probe the significance of increased silencing in these mutants, we performed a Fisher's Exact Test and found the odds of being in the "off" state for the mutants to be 3-4 times higher than for wild type. (**Figure 4.1G**). In contrast, this odds ratio comparing the other mutants to *Δset1* was close to 1 (**Figure 4.1 Supplement B**), indicating a similar likelihood of silencing. These genes are five members of the Set1/COMPASS complex, which catalyzes H3K4me and deposits H3K4me3 at active gene

promoters (Miller et al. 2001; Noma and Grewal 2002; Santos-Rosa et al. 2002; Roguev et al. 2003). Of the remaining complex members, *Δswd2* did not grow and *Δsdc1* was not in the screen, while *Δshg1* showed no phenotype, consistent with other studies that denote it as marginally associated with the complex (Roguev et al. 2003). All five gene deletions were validated by independent knockout in the parental reporter background (**Figure 4.1 Supplement B**).

Given this result, we sought to test whether the removal of Set1C might have a similar effect at the boundary proximal locus. While there was not a major effect of *Δset1* on reporter strains with a WT boundary (**Figure 4.1H**), both boundary[C] (**Figure 4.1I**) and Δboundary (**Figure 4.1J**) proximal reporters experienced a significant increase in silencing in *Δset1*, supporting the hypothesis that Set1/COMPASS enacts a heterochromatin-protective function.

IR-R endogenous genes regulate H3K9me2 spreading and silencing.

In order to probe the effect of *Δset1* on euchromatic invasion at heterochromatic sites genome-wide, we performed Chromatin Immunoprecipitation followed by Next Generation Sequencing (ChIP-Seq) with antibodies against H3K4me3 and H3K9me2 in WT, *Δepe1*, and *Δepe1Δset1* strains that contained no reporters (**Figure 4.2A**). We did not perform H3K4me3 ChIP-Seq for *Δset1* isolates due to the absence of H3K4me, which we validated by ChIP-qPCR (**Figure 4.2 Supplement A**). Signal tracks for each genotype are plotted as mean and 95% confidence interval of 2-4 replicates.

Given that our above results show *set1*-dependent heterochromatin containment at our reporter gene, we asked whether the removal of *set1* would affect H3K9me2 spreading beyond IR-R (**Figure 4.2B**). Unlike WT (black line), both *Δepe1* (purple line), and *Δepe1Δset1* (blue

line) display similar and significant enrichment for H3K9me2 immediately next to IR-R, as seen by their closely superimposed means and confidence intervals (for clarity, we did not plot MAT-internal traces). As distance increases from IR-R, the traces begin to separate, with H3K9me2 signal from *Δepe1Δset1* strains exceeding that from *Δepe1* and WT. This separation is most evident over the open reading frame of *rpl401* (**Figure 4.2B,** inset) and is statistically significant as indicated by the separation of the 95% confidence bounds and the p-value analysis below the traces. Interestingly, this gene is also highly enriched for H3K4me3. This increase in H3K9me2 spreading significantly affects the transcript levels of genes proximal to the separation of the H3K9me2 traces in *Δepe1Δset1* versus *Δepe1* strains (**Figure 4.2C**), but not genes either immediately by the compromised boundary or beyond *rpl401*. We wanted to test the role of endogenous gene promoters in effecting the Set1-dependent decline in H3K9me2 spreading and chose two genes, *mtd1,* and *rpl401*, around which spreading is most strongly impaired. To do so, we first modified the original *ade6p*:HSS to express "orange" from the *rpl401* promoter at the same locus (**Figure 4.2D**). The *rpl401* gene promoter effectively repels spreading in the context of a compromised (boundary[C], **Figure 4.2D,** middle) or fully abrogated (Δboundary, **Figure 4.2D,** bottom) IR-R boundary. However, the removal of *set1* (*Δset1*) resulted in complete *rpl401p*:HSS repression in a Δboundary context (**Figure 4.2D**). In the case of *mtd1p,* instead of inserting it at the original reporter locus, we replaced the endogenous *mtd1* open reading frame with "orange" to generate an *mtd1p*:HSS (**Figure 4.2E**), which is located 2.5kb from the edge of IR-R. Just like *ade6p*:HSS and *rpl401p*:HSS at the IR-R proximal locus, the *mtd1p*:HSS also displays genic barrier function that is *set1*-dependent (**Figure 4.2E**). Thus, for all the promoters tested, formation of a spreading barrier is highly sensitive to the presence of Set1. Given these

results, we conclude that Set1 contributes to the containment of spreading into the euchromatic region outside of IR-R in the case of boundary failure.
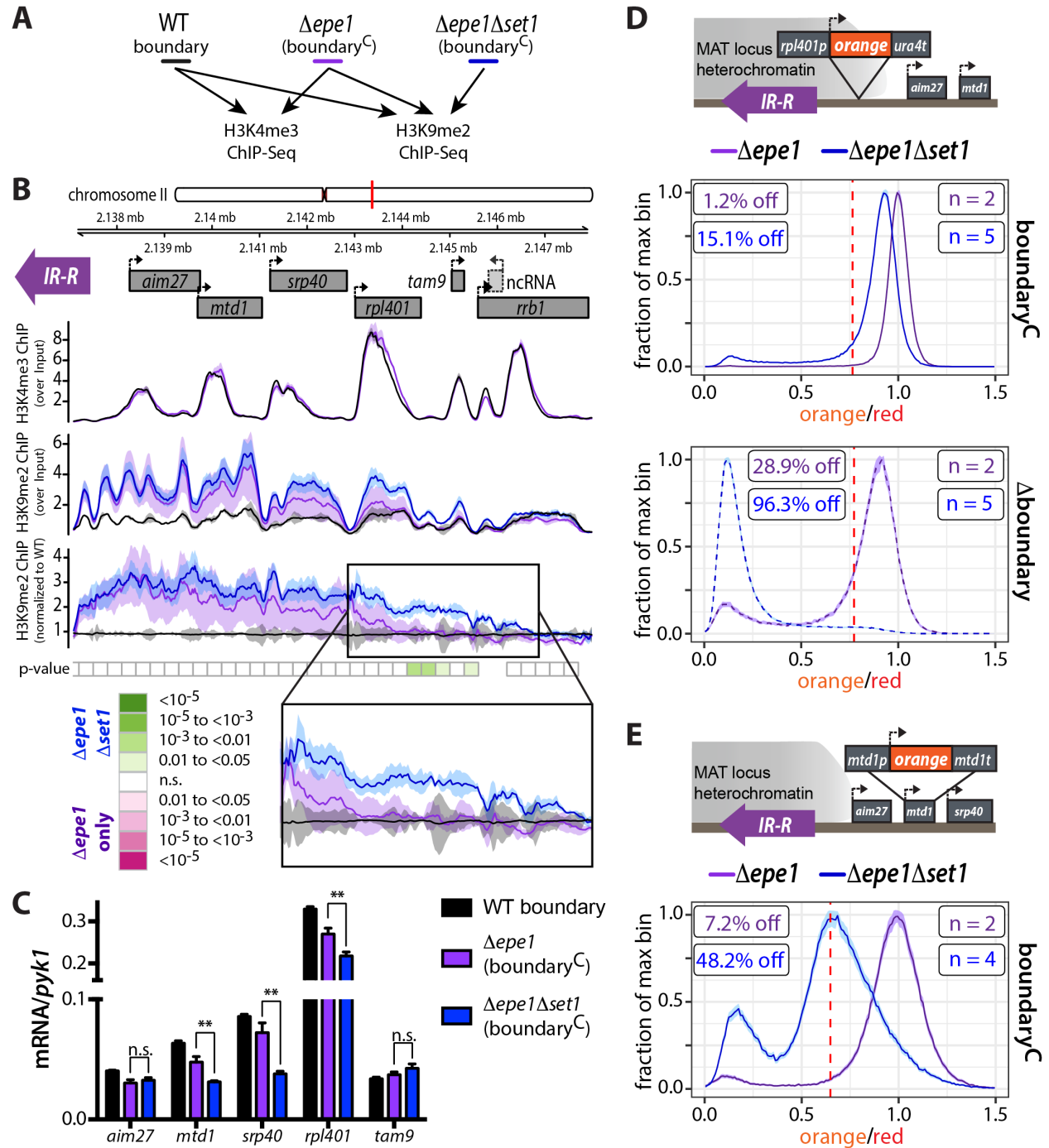

Figure 4.2: Set1 regulates H3K9me spreading at the IR-R proximal region.

A. Overview of ChIP-Seq experiment. **B.** Input normalized (top and middle tracks, see methods) ChIP-Seq signal tracks and gene annotations for the IR-R proximal region. H3K9me2 ChIP-Seq

datasets were independently normalized to signal from a sample containing merged data from both WT isolates (bottom track). Tracks are represented as mean (line) and 95% confidence interval (shaded region) per genotype (WT n=2, $\Delta epe1$ n=3, $\Delta epe1 \Delta set1$ n=4; each n represents a single colony deriving from an original knockout for each genotype). P-value track represents regions above a threshold for H3K9me2 enrichment over background (grey boxes, 300bp bins; absence of box indicates values below threshold). P-values for differences between genotypes are indicated in colors according to the scale. **C.** RT-qPCR analysis for genes in the IR-R proximal region. Error bars represent 1SD of three replicate cultures from single colonies deriving from one parent isolate. n.s. represents P>0.05 and ** represents P<0.01 (t-test). **D.** TOP: Overview of the *rpl401p*:HSS. MIDDLE: Histogram plots as in Figure 4.1 of normalized "orange" signal from *set1+* (purple) and *Δset1* (blue) *rpl401p*:HSS boundary[C] isolates. BOTTOM: Histogram plots of normalized "orange" signal from *rpl401p*:HSS Δboundary isolates. **E.** TOP: Overview of the *mtd1p*:HSS. BOTTOM: Histogram plots of normalized "orange" signal from *mtd1p*:HSS boundary[C] isolates.

Set1 contributes to spreading containment at facultative but not constitutive

heterochromatin.

We next examined other constitutive heterochromatin loci, centromeres and telomeres, for *set1*-

mediated spreading effects. Broadly, *Δset1* did not significantly increase the extent of spreading

already evident in *Δepe1* at such loci. Marginally increased spreading was detected in

Δepe1Δset1 beyond the boundaries of pericentromeric heterochromatin on chromosome II and III

(**Figure 4.3A, Figure 4.2 Supplement B**), while at the right subtelomere I and at the

pericentromere of chromosome I spreading was in fact reduced in *Δepe1Δset1* relative to *Δepe1*

(**Figure 4.3A, Figure 4.2 Supplement B,C**).

Given the major role of Set1/COMPASS at genes and the enrichment of H3K4me in

canonical euchromatin (Noma et al. 2001), we wondered if *set1* might regulate spreading at

facultative heterochromatin sites, islands of H3K9me embedded in gene-rich euchromatin

(Zofall et al. 2012; Gallagher et al. 2018). In our relatively stringent ChIP-Seq analysis (see

methods), we detected around 10 heterochromatin islands (8 and 13 in both WT replicates).

These islands display TSS-proximal H3K4me and low to intermediate levels of H3K9me2

(Figure 4.3B) (Zofall et al. 2012). Our analysis found an increase in the number of known

heterochromatin islands and novel ectopic H3K9me2 peaks (sites where WT shows no significant H3K9me2 enrichment) in both *Δepe1* and *Δepe1Δset1* mutants (**Figure 4.3A** and **Figure 4.3 Supplement A**). However, we found that heterochromatin spreading is exacerbated significantly in *Δepe1Δset1* compared to *Δepe1* at several sites, which include several known heterochromatin islands (**Figure 4.3B**). Importantly, the ability of Set1 to antagonize H3K9me heterochromatin does not strictly depend on Epe1, as we can observe enhanced H3K9me2 enrichment at heterochromatin islands and ectopic sites in *Δset1* alone (**Figure 4.3C**). Since transcription at islands and island- proximal genes is already extremely low in wild-type (compare **Figure 4.2C** and **Figure 4.3 Supplement B**), partly due to locally-acting RNA processing pathways (Lee et al. 2013; Egan et al. 2014; Sugiyama et al. 2016), it is not surprising that we only observe a mild further effect on transcript levels in *Δepe1Δset1* (**Figure 4.3 Supplement B**). These results describe a critical role for *set1* in spreading containment at gene-rich euchromatin with prominent H3K4me3 peaks, but not at gene-poor constitutive heterochromatin regions.
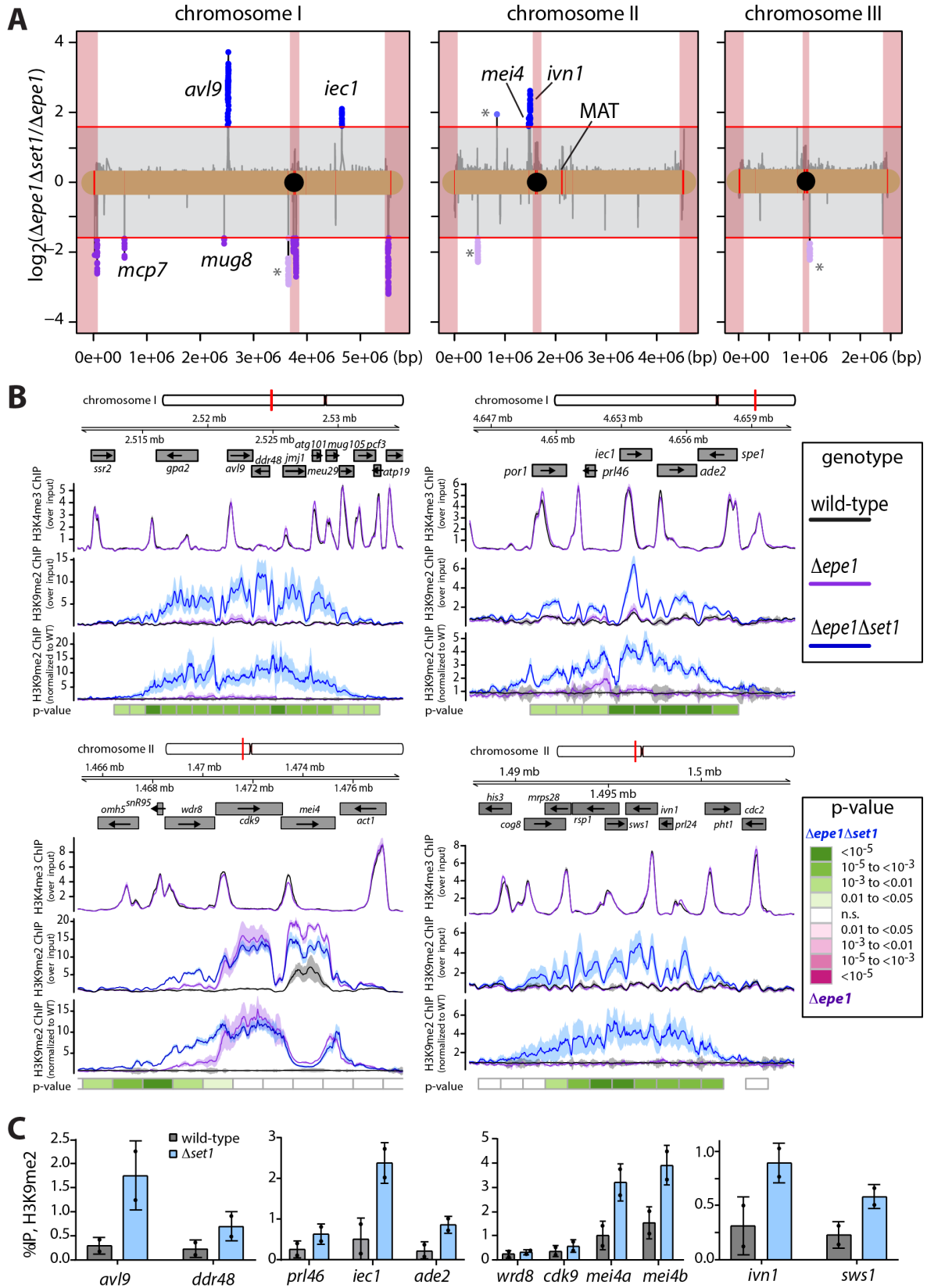
Figure 4.3: Set1 regulates spreading at euchromatic heterochromatin islands.

146

A. A global analysis comparing H3K9me2 accumulation measured by ChIP-Seq in *Δepe1Δset1* and *Δepe1* genotypes. The mean value of Input-normalized H3K9me2 ChIP signal per 300bp bin was calculated for each genotype. For bins containing H3K9me2 signal above 1.5 times the global background, the log2 ratio of *Δepe1Δset1* over *Δepe1* is plotted by chromosome (black line). Bins where this ratio exceeds a cutoff of 3 times enriched (red lines at y= +/- log2(3)) in *Δepe1Δset1* (blue) or *Δepe1* (purple) are plotted as individual points. Pericentromeres, centromeres, and telomeres are demarcated by red shaded boxes. H3K9me2 merged peak calls from the WT strains are annotated in red along the chromosome. * denotes regions where mean signal is increased due disproportionate enrichment in a single isolate. **B.** Signal tracks analysis at euchromatic heterochromatin islands and ectopic domains for H3K4me3 and H3K9me2 ChIP-Seq as in Figure 4.2B. P-values calculated as in Figure 4.2B except with 1200bp bins. **C.** H3K9me2 ChIP-qPCR measured at heterochromatin islands and ectopic domains in wild-type (black) and *Δset1* (backgrounds). Error bars represent 1SD from two technical replicate ChIPs. Replicate values are plotted as individual points.

Set1 functions in spreading containment independent of regulating steady state transcription.

What might be the mechanisms by which Set1 confers barrier activity to genes? We considered three pathways that could account for this activity: (1) Regulation of steady state transcription by Set1, where altered frequency of RNA polymerase II passage would disrupt spreading; (2) interference with heterochromatin spreading by the Set1/COMPASS enzymatic product, H3K4me, or; (3) a non-enzymatic effect of chromatin-bound Set1, consistent with prior reports (Lorenz et al. 2014; Mikheyeva et al. 2014). We summarize the possible mechanisms in **Figure 4.4A** and in what follows, we test whether and how mechanisms (1)-(3) contribute to the observed Set1-dependent barrier activity.

Previous reports have described both transcription- activating and repressive roles for Set1/COMPASS (Buratowski and Kim 2010; Mikheyeva et al. 2014; D'Urso et al. 2016). To directly test the involvement of mechanism (1), we examined the "orange" signal expressed from *rpl401p, mtd1p,* and *ade6p* in a *set1+* or *Δset1* backgrounds in a WT boundary context (**Figure 4.4B**). "orange" signal was normalized to forward scatter (fsc). This parameter tracks with the size of a cell and has been used extensively to estimate the cell volume, which is a central

parameter to normalize RNA and protein between single cells. The use of fsc bypasses any confounding effect *Δset1* might have on our *ade6p*-driven "red" control. We did not detect any major decrease in "orange" in *Δset1* isolates (**Figure 4.4B**). We confirmed this result by RT-qPCR analysis, where we normalized *ade6p* "orange", *mtd1*, and *rpl401* transcripts to an *act1* control (**Figure 4.4 Supplement 1A,** left). In the normalization, we adjusted for the *Δset1* effect on this *act1* control (**Figure 4.4 Supplement 1A**, right, methods). Together, these results argue against Set1 regulating the mean level of RNA polymerase II mediated transcription at these genes. Thus, mechanism (1), is an unlikely avenue for the *Δset1*-dependent phenotype in reporter silencing.
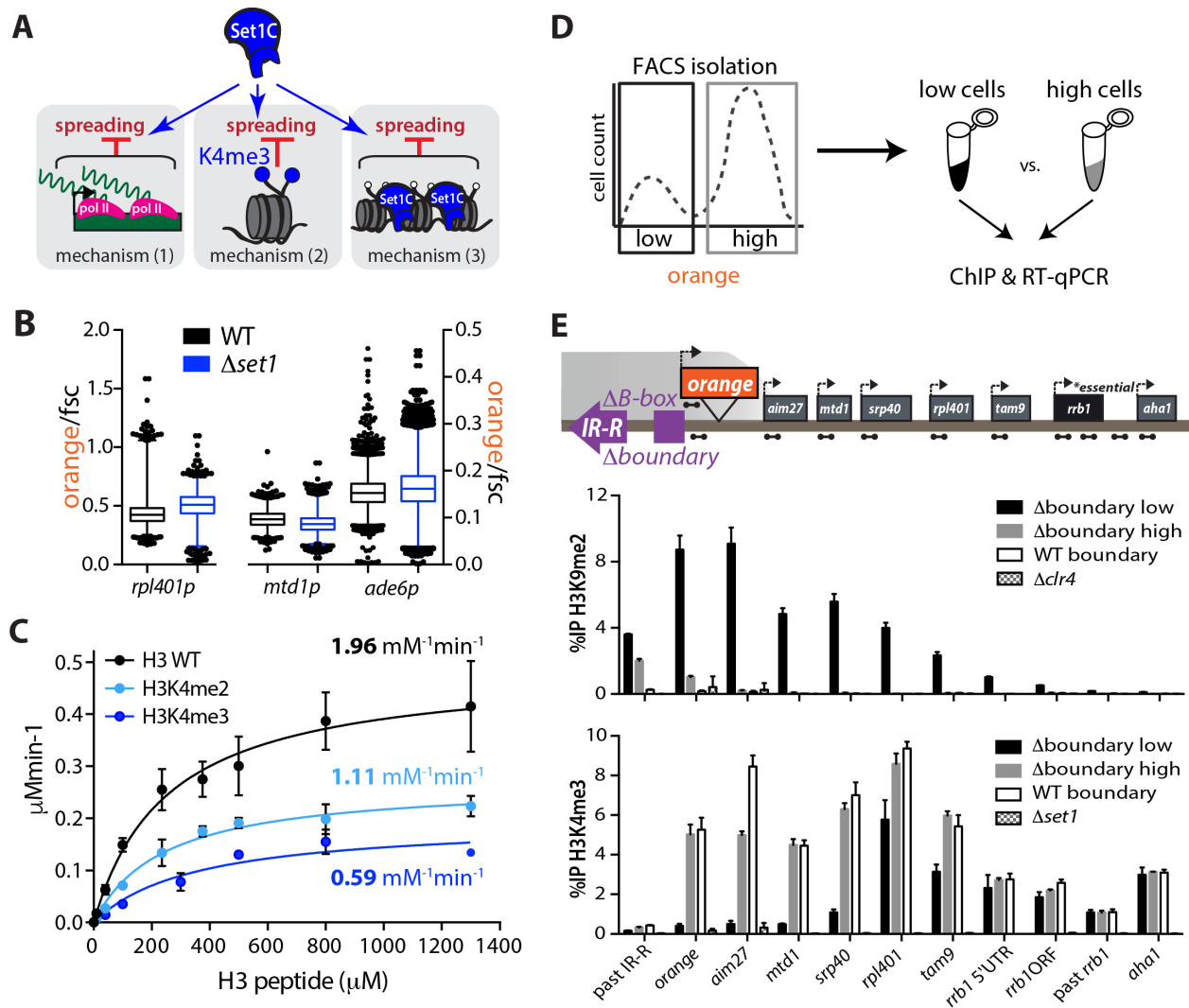
Figure 4.4: The gene-protective activity of Set1 is independent of mean transcription levels and is rooted in catalytic inhibition of Suv39/Clr4 by H3K4me2/3.

A. Possible mechanisms by which Set1 repels heterochromatin spreading: (1) Maintaining a level of transcription that is refractory to heterochromatin invasion due to local RNA polymerase activity and associated cycles of nucleosome eviction; (2) interference of H3K4me3, the Set1-product, with heterochromatin spreading; (3) non-catalytic effect of Set/COMPASS including its occupancy on chromatin. **B.** Box and whisker plots of "orange" signal normalized to forward scatter (fsc) for *rpl401p:, mtd1p:,* and *ade6p:*HSS in *set1+* (black) and *Δset1* (blue) backgrounds. 1%-99% of the data is included within the whiskers. Outliers are plotted as individual points. **C.** Histone methyltransferase assay with Clr4-SET and H3(1-20) peptides with modifications as indicated. Error bars represent 1SD from three replicate experiments. k_{cat}/K_M (specificity constant) values are derived from measurement of the k_{cat} and K_M (see Figure 4.4 Supplement 1F). **D.** Cartoon overview depicting FACS isolation of "low" and "high" Δboundary 5′ *ade6p-*"orange" cells followed by ChIP and RT-qPCR. **E.** ChIP-qPCR data for FACS sorted cells - H3K9me2 (top) and H3K4me3 (bottom). Amplicons for each qPCR are depicted as dumbbells on cartoon locus. Error bars represent 1SD from three technical replicate ChIPs.

149

H3K4me directly interferes with Suv39/Clr4 catalysis.

Set1 is the only H3K4 methylase in fission yeast (Noma and Grewal 2002) and H3K4me and H3K9me appear mutually exclusive (Noma et al. 2001). Hence, we hypothesized one implementation of mechanism (2) could be direct interference with Suv39/Clr4 activity. This could potentially occur via two mechanisms – either by directly impacting catalysis of H3K9 methylation by Suv39/Clr4 or by disrupting the 'read-write' positive feedback characteristic of histone methyl transferases. This spreading feedback mechanism is mediated by the binding of Suv39/Clr4 enzyme to its own product via the chromodomain (CD), which stimulates the catalysis of H3K9 methylation on proximal nucleosomes via the SET domain (Zhang et al. 2008a; Margueron et al. 2009; Al-Sady et al. 2013; Muller et al. 2016). Clr4-CD recognition of H3K9me has been shown to be sensitive to acetylation (ac) of the H3K4 residue (Xhemalce and Kouzarides 2010).

We tested whether the Clr4-CD's ability to recognize H3K9me is impacted by H3K4me3. We purified the Clr4-CD (**Figure 4.4 Supplement 1B**) and performed fluorescence polarization with modified histone tail peptides. We found that the Clr4-CD has a similar binding affinity for H3K9me3 and H3K4me3K9me3 tail peptides (**Figure 4.4 Supplement 1C**), which is recapitulated by the full-length Clr4 enzyme (**Figure 4.4 Supplement 1D**). Thus, the presence of H3K4me3, unlike H3K4ac, does not disrupt the 'read-write' feedback mechanism.

Our understanding of H3K4me effects on H3K9me catalysis by various enzymes are based mostly on endpoint analysis and have yielded conflicting results. Two previous studies employing endpoint analysis indicated no obvious effect of H3K4me2 or a K4A mutation on Suv39/Clr4 activity (Nakayama et al. 2001; Kusevic et al. 2017), yet a number of other studies document a range of effects of H3K4me2 or H3K4me3 on H3K9 methyltransferases, and these

results do not always agree (Wang and Zhang 2001; Nishioka et al. 2002; Chin et al. 2005; Binda et al. 2010). To definitively determine any effect of H3K4me2 or H3K4me3 may have on Suv39/Clr4 catalysis, we performed multiple turnover Michaelis-Menten kinetic analysis using N-terminal truncation of Clr4 comprising residues 192-490 (Collazo et al. 2005; Dirk et al. 2007) which includes the catalytic SET domain (**Figure 4.4C** and **Figure 4.4 Supplement 1E**). The masses of the H3K4me0, 2 and 3 peptides used were verified by MADLI-TOF analysis (**Figure 4.4 Supplement 2**). We determined k_{cat}, K_M, and specificity constant (k_{cat}/K_M) values (**Figure 4.4C, Figure 4.4 Supplement 1F**) and interestingly, found that H3K4me3 and H3K4me2 reduce Clr4's k_{cat}/K_M by 3.3 times and 1.8 times, respectively, relative to an H3K4me0 (WT) peptide. This derives mostly from an adverse effect on Suv39/Clr4's k_{cat} rather than on the K_M (see **Figure 4.4 Supplement 1F**). We confirmed that this effect is reflected in the full-length enzyme under k_{cat}/K_M conditions (**Figure 4.4 Supplement 1G**), where we find H3K4me3 to reduce k_{cat}/K_M by 4.6 times, in good agreement with the Michaelis-Menten parameters extracted with the SET domain. These results were confirmed with an independently produced set of H3K4me0 and H3K4me3 peptides (not shown). In conclusion, these results demonstrate that Suv39/Clr4 catalysis (**Figure 4.4C**, **Figure 4.4 Supplement 1F,G**), but not its product recognition (**Figure 4.4 Supplement 1C,D**), is inhibited by the presence of H3K4me3, to and a milder extent, by H3K4me2. These results support a role for mechanism (2) and make the following two predictions: First, if H3K4me3 is directly involved in repelling spreading, genes downstream from an H3K4me3 peak are protected from heterochromatin invasion if the peak correlates with effective disruption of silencing. Second, chromatin recruitment of Set1 is insufficient for barrier activity, which requires Set1's catalytic activity.

Protection of downstream genes by H3K4me.

To test the first prediction, we used Fluorescence Assisted Cell Sorting (FACS) to isolate both repressed ("low") and expressed ("high") populations of 5′ *ade6p*:HSS Δboundary cells (**Figure 4.4D**) and then assessed their chromatin and transcriptional state via ChIP and RT-qPCR, respectively (**Figure 4.4E, Figure 4.4 Supplement 1H**). While both populations evidenced H3K9me2 accumulation upstream of the reporter, H3K9me2 signal cannot be detected at any point beyond "orange" in the "high" cells (grey bars). This immediate drop coincides with the *ade6p* H3K4me3 peak in the "high" cells, and H3K4me3 is enriched at the downstream gene promoters comparable to WT levels. Consistent with this H3K4me3 distribution, transcription levels are similar to the no heterochromatin (*Δclr4*) state. This result, in conjunction with our above findings (**Figure 4.1I,J** and **Figure 4.4C**) suggest that H3K4me3 accumulation at *ade6p* protects downstream transcriptional units. On the other hand, the "low" population (black bars) displays high levels of H3K9me2 at and beyond "orange", while H3K4me3 is severely reduced (**Figure 4.4E**). H3K9me2 levels eventually decline towards the essential *rrb1* gene, concomitant with a rise in H3K4me3 enrichment. The discrepancy between the H3K4me3 signal in the "low" and "high" populations thus eventually decreases with distance. In cells where *ade6p*-localized H3K4me3 is overcome, downstream transcriptional units, therefore, appear to succumb to repressive H3K9me2. Thus, these data are consistent with a model where encounter of a substantial and/or persistent H3K4me3 peak disrupts spreading, protecting downstream gene units.

Catalytic activity of Set1, and not chromatin recruitment alone, underpins heterochromatin containment.

To test the second prediction concerning catalytic activity of Set1, we constructed an allelic series of H3K4 methylation Set1 hypomorphs, based on sequence alignments with the *Saccharomyces cerevisiae* Set1 ortholog, and published catalytic mutants within this gene (Schlichter and Cairns 2005). We introduced tagless C862A (Set1^{C862A}) and G852S (Set1^{G852S}) (**Figure 4.5A**) into *S. pombe* Set1 within its native gene context, marked with a nourseothricin resistance (NAT^R) gene, and produced a corresponding wild-type (wt-Set1) control. We produced a separate set of strains where Set1^{C862A} and Set1^{G852S} and wt-Set1 version were N-terminally 2xFLAG tagged and inserted at the native *set1* locus, to test for expression by western blot. N-terminally FLAG tagged Set1 has been shown to retain function (Mikheyeva et al. 2014). We find mutants and wt-Set1 to accumulate to similar levels by two independent extraction methods (**Figure 4.5 Supplement A**). Next, we probed for H3K4me3 accumulation by western blot and find that both mutants show a defect in H3K4me3 accumulation, with the C862A mutant showing almost no H3K4me3 by Western blot and G852S accumulating significantly reduced amounts (**Figure 4.5B**). We next moved to directly test above prediction, and found that both Set1^{C862A} and Set1^{G852S} mutants are significantly impaired in their ability to protect *rpl401p*:HSS from invasion by heterochromatin compared to wt-Set1 (**Figure 4.5C**). Set1/COMPASS, similar to Suv39/Clr4, contains a positive feedback loop with the enzyme recognizing its product (Roguev et al. 2003; Kirmizis et al. 2007), and it is possible that abrogation of catalytic function leads to reduction of Set1 recruitment to chromatin. We used our 2xFLAG Set1 constructs to test whether Set1^{C862A} and Set1^{G852S} are still normally recruited to chromatin at the TSS (**Figure 4.5D**) and indeed find no major difference in enrichment of

Set1^{C862A} or Set1^{G852S} versus wt-Set1. Importantly, this data allows us to exclude that the recruitment of Set1, part of a megadalton complex (Miller et al. 2001), itself or associated H3K4me-independent functions, repel heterochromatin spreading (mechanism (3), **Figure 4.4A**). Instead, these data offer further support for mechanism (2), showing *in vivo* that Set1 catalytic activity is required for containment of heterochromatin spreading, likely in part via direct interference with Suv39/Clr4 catalysis.
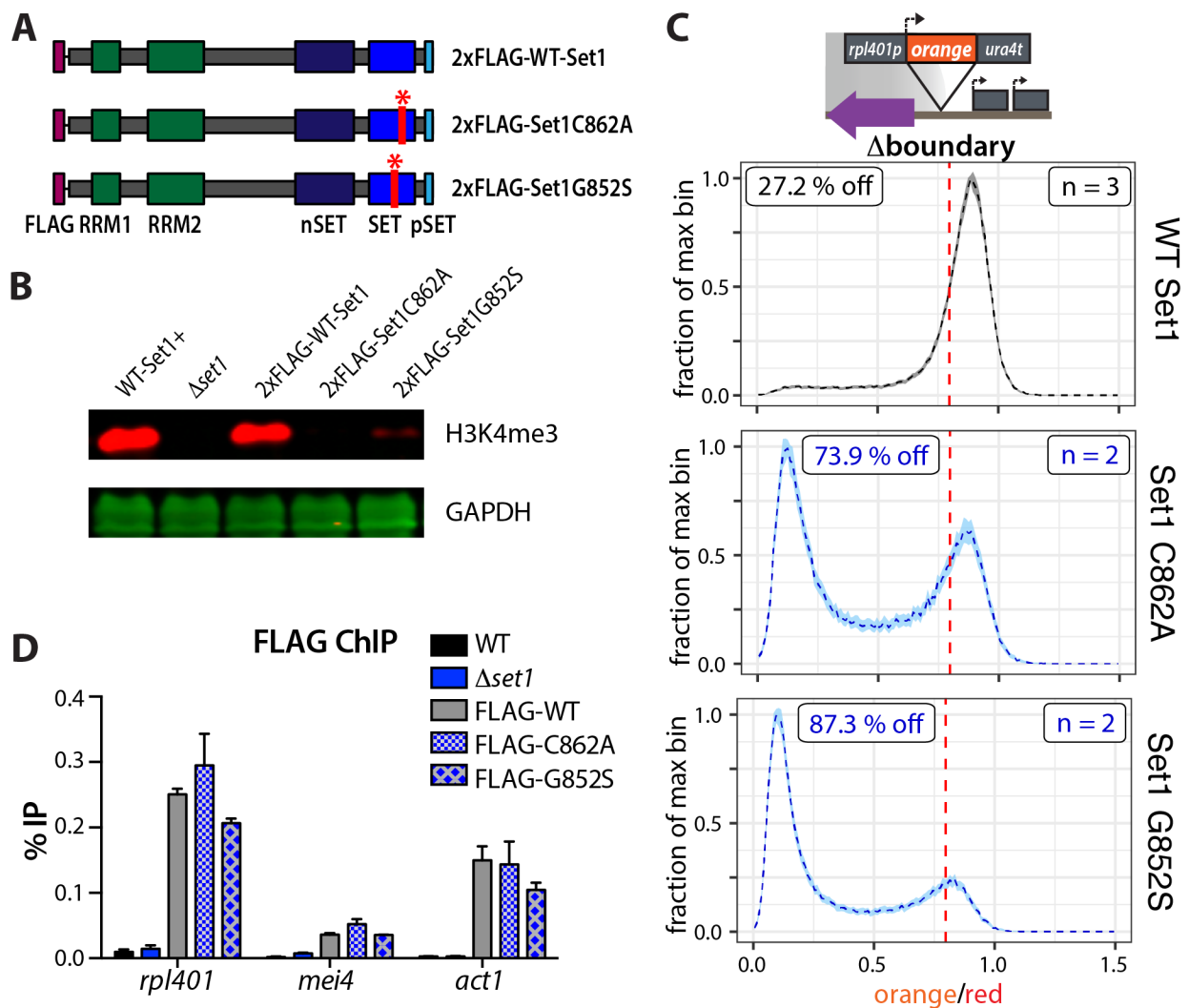

Figure 4.5: Set1 catalytic activity but not its recruitment to chromatin is required for its gene protective function.

A. Diagram of Set1 constructs including WT-Set1 and two point mutations in the catalytic SET domain – C862A and G852S. Constructs are expressed from the *set1* promoter at the native

set1 locus with an N-terminal 2xFLAG tag. **B.** Licor Western blot for H3K4me3 with GAPDH loading control of whole cell extracts from wild-type untagged Set1, the Δ*set1* parent of the 2xFLAG constructs, 2xFLAG-wtSet1, 2xFLAG-Set1C862A, and 2xFLAG-Set1G852S. **C.** Histogram plots as in Figure 4.1 of normalized "orange" signal from *rpl401p*:HSS Δboundary isolates that were transformed with either untagged wt-Set1, Set1C862A, or Set1G852S. **D.** Anti-FLAG ChIP-qPCR data in genetic backgrounds as in B. Error bars represent 1SD from two technical replicate ChIPs.

The distribution of H3K4me3 and nucleosome occupancy over genes correlate with orientation dependence of genic heterochromatin boundary function.

H3K4me3 is enriched near the TSS of genes (Santos-Rosa et al. 2002; Pokholok et al. 2005), and in fission yeast, heterochromatin silencing can proceed in a co-transcriptional manner (Buhler et al. 2006; Buhler et al. 2008). This led us to hypothesize that encountering a gene first at the promoter (5′ end) versus the terminator (3′ end) will more effectively protect against gene silencing, since heterochromatin will be antagonized before co-transcriptional silencing mechanisms can proceed. To our surprise, while our *ade6p*:HSS is clearly more effective in the 5′ proximal orientation (**Figure 4.6A, Figure 4.6 Supplement 1A**), *rpl401p*:HSS shows much less bias (**Figure 4.6B, Figure 4.6 Supplement 1B**). In both 5′ and 3′ proximal orientations, *rpl401p*:HSS effectively disrupts spreading in a *set1*-dependent manner (**Figure 4.6B**). We wondered whether this discrepancy can be explained by the profile of H3K4me3 over the native gene. Indeed, we find that *rpl401* has significantly elevated H3K4me3 over the middle and, critically, 3′ of the gene, while it was strongly diminished at the 3′ of *ade6* (**Figure 4.6C**). We find a similar distribution pattern for the respective HSS cassettes (**Figure 4.6 Supplement 1D**). The difference in orientation bias between *ade6p*:HSS and *rpl401p*:HSS can thus be partially accounted for by the H3K4me3 profile. However, we wondered if additional, nonetheless H3K4me-dependent, mechanisms beyond the direct catalytic interference we document above, underlie the striking difference in gene orientation effect.

We focused on regulation of nucleosome occupancy, known to adversely affect spreading (Garcia et al. 2010; Aygun et al. 2013). We first assessed nucleosome occupancy in *set1+* and *Δset1* strains by H3 ChIP in log phase cultures (**Figure 4.6D-F, Figure 4.6 Supplement 1E**), but also in G2 stalled cells to exclude cell cycle passage effects (**Figure 4.6 Supplement 1F**). Intriguingly, nucleosome occupancy is highly elevated at the 3′ of *ade6* but remains low throughout *rpl401* (**Figure 4.6D, Figure 4.6 Supplement 1E**). Low nucleosome occupancy is strongly antagonistic to spreading (Garcia et al. 2010; Aygun et al. 2013). Therefore, the data showing that *ade6* retains high nucleosome occupancy at its 3′ provide an additional explanation why *ade6* and not *rpl401* is vulnerable to heterochromatin invasion from the 3′. More broadly, we observed increases in nucleosome occupancy at heterochromatin islands and the IR-R proximal genes in *Δset1* (**Figure 4.6E**, **F**) and across active genes distributed on the three *S. pombe* chromosomes (**Figure 4.6 Supplement 1E**), but not on heterochromatin targets (**Figure 4.6E,** grey box). This effect therefore likely represents a general feature of Set1 activity. Importantly, we found that the catalytic activity of Set1 is required for this regulation of nucleosome occupancy, as the catalytic Set1^{C862A} and Set1^{G852S} mutants partially or fully mirror the *Δset1* phenotype at *ade6, rpl401* and heterochromatin islands (**Figure 4.6D, F**). The extent to which the catalytic mutants recapitulate the *Δset1* phenotype correlates both with the global H3K4me3 accumulation defect in each hypomorph (**Figure 4.5B**), as well as the residual H3K4me3 at any given gene-internal location relative to the wt-Set1 (**Figure 4.6C, D**).
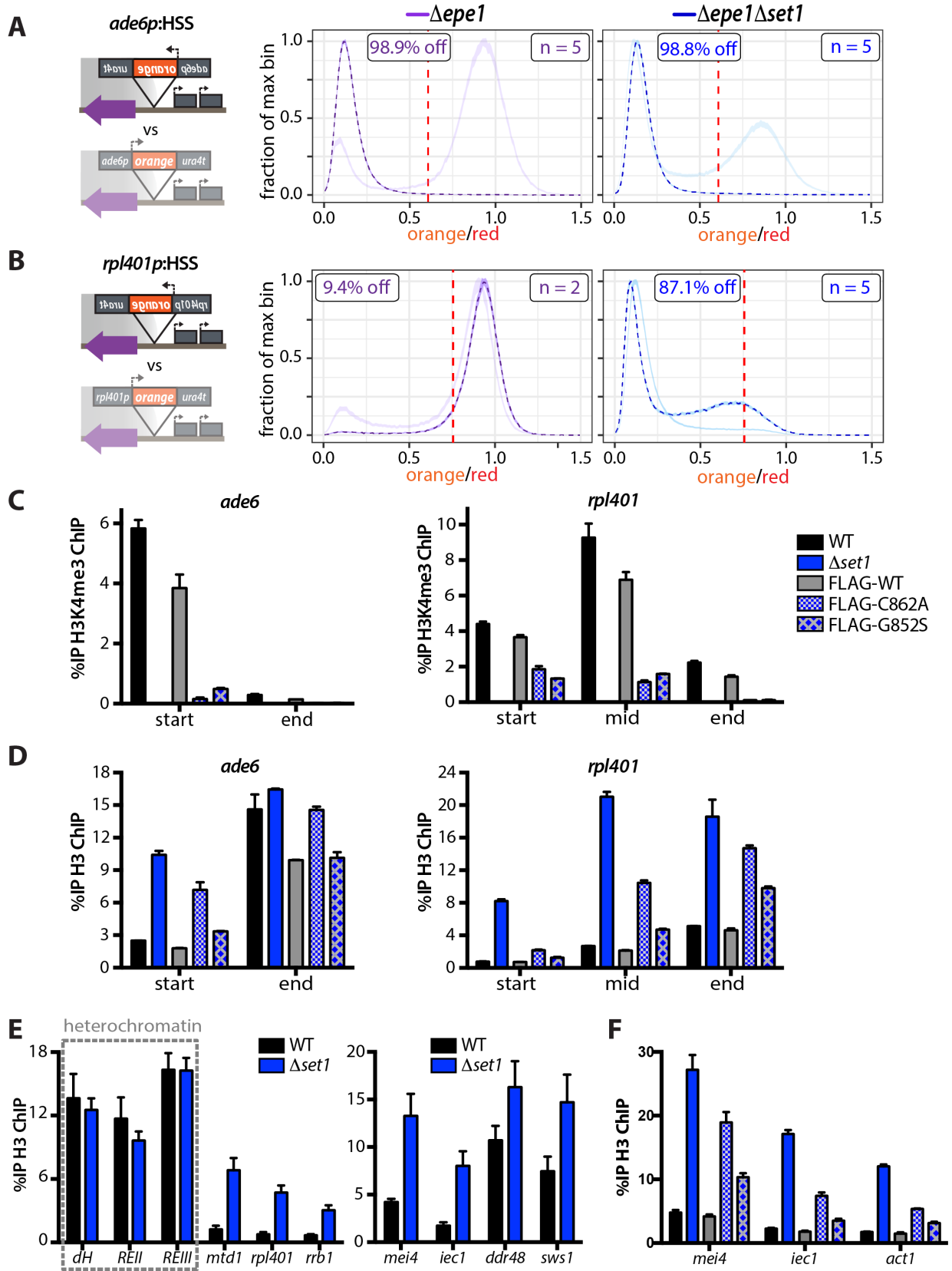
Figure 4.6: H3K4me3 disrupts local nucleosome occupancy.

A. Locus cartoon for *ade6p*-driven "orange" reporters in either the 3′ or 5′ (shaded) orientation with respect to IR-R (cartoon). Histograms for normalized "orange" signal in Δboundary context as in Figure 4.1. 3′ *ade6p*:HSS is plotted in full color while 5′ *ade6p*:HSS is shaded (data depicted in Figure 4.1, re-drawn for comparison). *epe1* and *set1* genotypes are as indicated. **B.** Locus cartoons and histogram plots as in A. for 3′ and 5′ *rpl401p*:HSS. Shaded lines reproduced from Figure 2D. **C.** H3K4me3 ChIP-qPCR over the gene body of *ade6* and *rpl401* open reading frames in wild-type untagged Set1, *Δset1*, 2xFLAG-wtSet1, 2xFLAG-Set1C862A, and 2xFLAG-Set1G852S. **D.** H3 ChIP-qPCR over the gene body of *ade6* and *rpl401* open reading frames in genotypes as in C. **E.** H3 ChIP-qPCR in WT (black) and *Δset1* (blue). Constitutive heterochromatin targets (boxed in grey). **F.** H3 ChIP-qPCR at *mei4*, *iec1*, and *act1* in genotypes as in C. In E., H3 ChIP-qPCR Error bars represent 1SD from four replicates, each representing a single colony deriving from each genotype. For ChIPs in C.,D., and F. error bars represent 1SD from two technical replicate ChIPs.

However, these data do not explain how heterochromatin can overcome the TSS-localized H3K4me3 peak when invading a gene like *ade6* from 3′ and then enact stable repression. We hypothesized that for this to occur, 3′-invading heterochromatin would need to be able to 1. Partially invade the gene, 2. Downregulate transcription without fully reaching the promoter, consistent with co-transcriptional gene silencing, and finally, 3. Reduce H3K4me3, the key spreading antagonizing signal, likely via a reduction in transcription (Shilatifard 2012). To address these hypotheses, we built a variant of the 3′*ade6p*:HSS reporter construct that would permit spreading to proceed into the gene unit but hinder its ability to reach the promoter. To achieve this, we fused the "orange" and "green" coding sequences by an in-frame linker containing 5 B-box elements (**Figure 4.6 Supplement 2A**), multimers of which have been shown to confer synthetic boundary activity (Noma et al. 2006). Signal from "green" and "orange" in WT, boundary[C], and Δboundary contexts, as well as their RNA levels (**Figure 4.6 Supplement 2A,B**), were well correlated in each isolate. This indicates that the entire transcriptional unit is uniformly regulated, despite presence of the synthetic B-box boundary midway through the tandem gene unit. We next assessed the chromatin state at "green" and "orange" by ChIP. H3K9me2 is significantly reduced at "orange" compared to "green" across all

isolates from both boundary[C] and Δboundary contexts (**Figure 4.6 Supplement 2C**) supporting

that the 5x B-box sequence was functioning as a synthetic roadblock to spreading. The difficulty

of separating nucleosomes by shearing within heterochromatin likely prevented us from

documenting any potentially sharper drops across the synthetic barrier. Surprisingly, H3K4me3

ChIP revealed that boundary[C] and Δboundary had significantly reduced methylation levels

compared to WT at the "orange" TSS (**Figure 4.6 Supplement 2D**). These results demonstrate

that invasion of a gene from the 3′ end can reduce both inhibitory H3K4me levels and

transcription, despite not fully reaching the gene promoter. This mechanism would presumably

not operate in the 5′ orientation, since H3K4me3 would be encountered first.

Histone acetylation links H3K4me3 to Set1- dependent regulation of nucleosome

occupancy.

Our data show both that the catalytic activity of Set1, hence production of H3K4me, is required

for containment of heterochromatin spreading, and that regulation of nucleosome turnover tightly

correlates with this containment function. However, the question remains of how regulation of

nucleosome turnover is tied to H3K4me. Previous studies have identified a role for

Set1/COMPASS and H3K4me in promoting global histone acetylation at various residues

(Noma and Grewal 2002; Taverna et al. 2006; Ginsburg et al. 2014). To validate this finding in

our system, we performed ChIP against H3K9ac, as well as H3 and H4 acetylation broadly, and

found indeed that in *Δset1,* acetylation was similarly reduced (**Figure 4.7A-C**) at all the genes

tested, whether at heterochromatin islands or canonical euchromatin. The fact that we found a

robust decrease in H3 and H4 acetylation as well as H3K9ac specifically, indicates the

involvement of multiple histone acetyltransferase (HAT) complexes (Buratowski and Kim 2010;

Woo et al. 2017). Since HAT mutants or knock-downs have broad effects on heterochromatin (Gomez et al. 2005; Tong et al. 2012; Wang et al. 2013; Wang et al. 2015), we chose not to pursue mutational analysis of the catalytic subunits. *S. pombe* contains two genes that are orthologs of H3K4me3-specific PHD reader modules within HAT complexes: *png1*, which associates with Mst1 in *S. pombe* (Chen et al. 2010a)*,* and *png2*, which does not impact H4 acetylation or have known HAT associations in *S. pombe* (Chen et al. 2010a). To test whether recruitment of HATs to H3K4me3 specifically is involved in providing protection against heterochromatin invasion, we deleted either the H3K4me3-reading PHD fingers of *png1* or *png2* in the context of the 3′ -oriented *rpl401p*:HSS. While *png2ΔPHD* did not have an effect on *rpl401p*:HSS in the Δboundary context, we found *png1ΔPHD* to have a subtle but highly reproducible effect, resulting in elevated silencing at the otherwise highly efficient barrier forming 3′- *rpl401p*:HSS (**Figure 4.7D**). This phenotype was recovered after outcrossing to wild-type and re-testing 9 *bone fide Δepe1 png1ΔPHD* resulting progeny (**Figure 4.7 Supplement A**). Additionally, the phenotype was recovered upon re-introduction of *Δepe1* into *epe1+ png1ΔPHD* isolates, indicating it is a stable phenotype (**Figure 4.7D**). *png1ΔPHD* importantly does not affect basal expression of *rpl401p:*HSS in the presence of *epe1+* (**Figure 4.7D,** dashed grey line). The fact that the phenotype is significantly weaker than *Δset1* is expected, since it appears that more than one HAT is involved in maintaining elevated acetylation in response to Set1 activity (**Figure 4.7A-C**). The *Δset1 png1ΔPHD* double mutant has very similar degree of silencing as *Δset1*, 86% vs 82% of cells, respectively, (**Figure 4.7 Supplement B**)*,* indicating that the effect of *png1ΔPHD* on silencing is likely not additive with Set1. These data provide evidence for Png1's involvement in H3K4me-dependent

heterochromatin containment, possibly recruiting Mst1, which likely acts redundantly with other
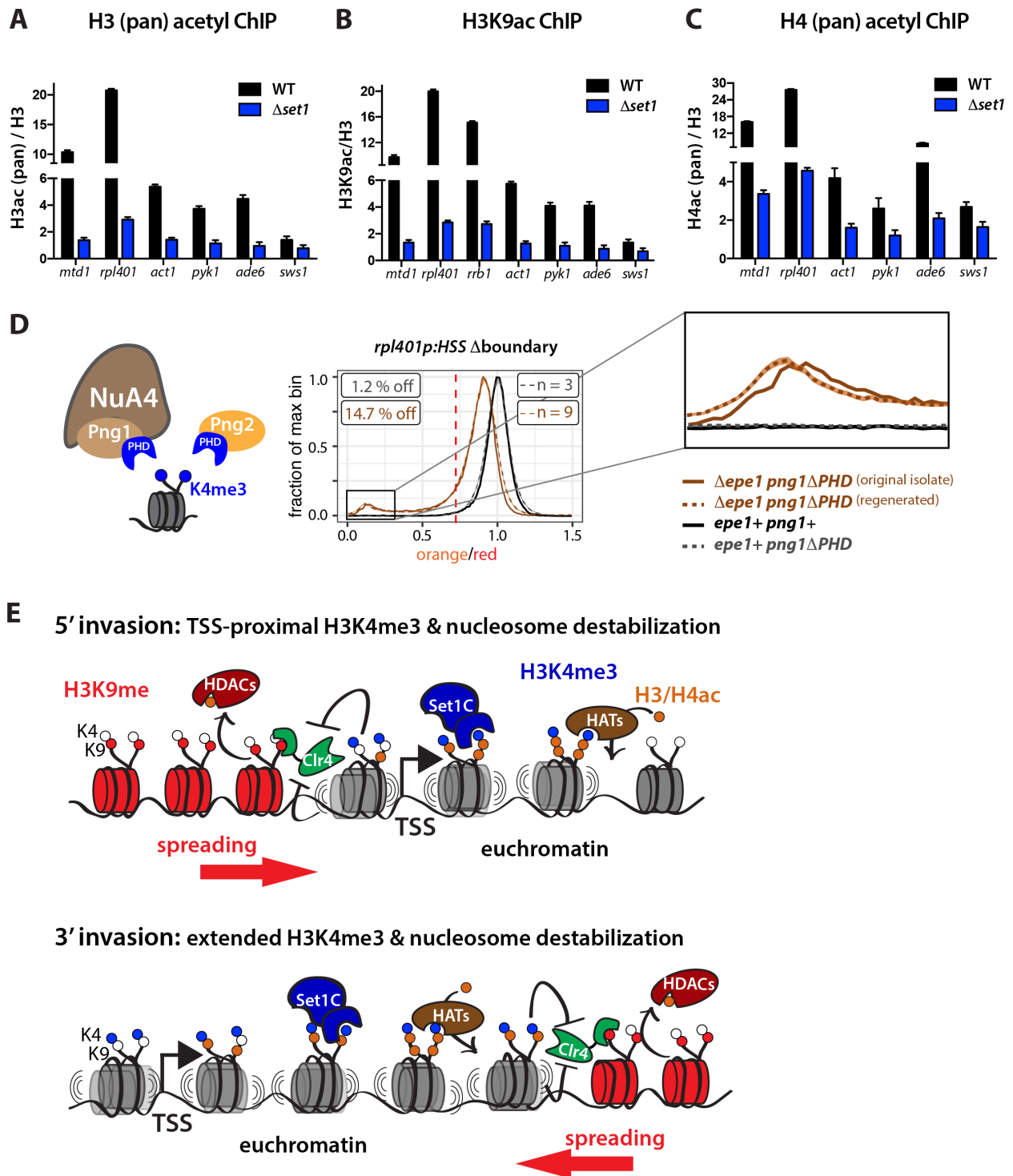
HAT complexes (Buratowski and Kim 2010).

Figure 4.7: Role of histone acetylation in heterochromatin containment.

A. H3 (pan) acetyl ChIP-qPCR in WT and *Δset1*. ChIP is normalized to H3 signal to account for differences in nucleosome occupancy. **B.** H3K9ac ChIP-qPCR plotted as in A. **C.** H4 (pan) acetyl ChIP-qPCR plotted as in A. For A.-C.) error bars represent 1SD from four replicates, each representing a single colony deriving from each genotype. **D.** Cartoon depicting Png1 and Png2 containing H3K4me3 reading PHD finger domains. Png1 is associated with NuA4. Histograms as in Figure 4.1 of normalized "orange" signal from 3′ *rpl401p*:HSS Δboundary isolates from *png1ΔPHD* (brown), or WT boundary with *png1ΔPHD* (grey) and *png1+* (black). **E.** Model for the contribution of Set1/COMPASS to gene-mediated heterochromatin repulsion. In 5′ invasion (TOP), Set1-dependent TSS-proximal H3K4me3 repels heterochromatin spreading via direct Suv39/Clr4 inhibition and nucleosome destabilization. (BOTTOM) Broader distributions of Set1-dependent H3K4me3 in bodies of some genes and the ensuing increased nucleosome destabilization repels 3′ heterochromatin invasion. Histone Acetyl Transferase complexes (HATs) attracted to the H3K4me3 via reader proteins acetylate locally and contribute to nucleosome destabilization.

DISCUSSION

Two paradigms have emerged for heterochromatin domain regulation, which when taken together present an intriguing paradox. On one hand is the ability for heterochromatin domains to expand beyond their borders when containment mechanisms are compromised (Noma et al. 2006; Zofall and Grewal 2006; Trewick et al. 2007; Zofall et al. 2012; Wang et al. 2013; Wang et al. 2014; Garcia et al. 2015). On the other, is the widespread dispersion of factors, activities, and posttranslational modifications embedded in euchromatin, which are known to antagonize the establishment and maintenance of heterochromatic domains (Lan et al. 2007; Sugiyama et al. 2007; Garcia et al. 2010; Aygun et al. 2013; Wang et al. 2013; Wang et al. 2015). Why then is heterochromatin spreading able to overcome these negative regulators and expand into euchromatin? Part of the answer may lie in the activities inherently associated with the spreading machinery, including HDACs (Grewal et al. 1998; Shankaranarayana et al. 2003; Yamada et al. 2005; Sugiyama et al. 2007), nucleosome remodelers (Sugiyama et al. 2007; Taneja et al. 2017), and H3K4 – demethylase complexes (Li et al. 2008), which apparently can overpower euchromatin. Yet, how and why heterochromatin spreading is halted at specific euchromatic locations is not understood.

In this work we investigated the signals within local active euchromatin that define spatial limits to heterochromatin spreading in fission yeast. The key principles that derive from this work are (1) Euchromatic barrier signals depend on Set1/COMPASS activity at active genes. (2) High gene transcript levels are not intrinsically refractory to heterochromatin invasion. (3) Set1-dependent repulsion of heterochromatin acts via two pathways downstream of H3K4 methylation: direct catalytic inhibition of Clr4/Suv39 and nucleosome mobilization. (4) The ability to repel heterochromatin can be gene orientation-specific, directed by the distribution of H3K4me over the gene.

Mechanisms regulating facultative heterochromatin domain size.

We find Set1/COMPASS enacts a heterochromatin containment signal at gene-rich regions, including facultative heterochromatin in fission yeast that responds to environmental conditions (Zofall et al. 2012; Sugiyama et al. 2016; Gallagher et al. 2018) (**Figure 4.2 & Figure 4.3**). These findings are in contrast to previously identified spreading regulators that function globally, such as Epe1, Leo1, Paf1 and Mst2 (Trewick et al. 2007; Zofall et al. 2012; Kowalik et al. 2015; Sadeghi et al. 2015; Verrier et al. 2015; Wang et al. 2015; Flury et al. 2017). The containment function of Set1 is localized to specific euchromatic regions, heterochromatin islands, and euchromatin exposed to boundary failure at IR-R (**Figure 4.2 & Figure 4.3**), but is not prominent at constitutive heterochromatin (**Figure 4.2 Supplement C**).

Critically, containment of heterochromatin spreading does not require a change in mean transcript level (**Figure 4.4B**), but specifically its ability to methylate H3K4. This is remarkable as active transcription in fission yeast and other systems leads to formation of nucleosome free regions (NFRs) at the TSS (Lantermann et al. 2010), and NFRs are thought to be refractory to

heterochromatin spreading (Garcia et al. 2010; Lantermann et al. 2010). Our data point to NFRs still being intact in *Δset1*, as is evident from the coinciding dips in the H3K9me2 tracks in *Δepe1* and *Δepe1Δset1* (**Figure 4.2B**) This is consistent with findings that formation of NFRs alone is insufficient to block spreading (Oki and Kamakaka 2005). Just as heterochromatin overcomes the TSS-proximal NFR it can overcome the presence of Set1 on chromatin, even though it is part of a megadalton complex. This conclusion is supported by the normal chromatin localization, but defective H3K4 methylation and heterochromatin repulsion, of the Set catalytic hypomorphs (**Figure 4.5D**). The crucial heterochromatin repelling signal is, therefore, the H3K4me mark. This modification takes two parallel tracks to push back against encroaching heterochromatin: 1. Catalytic interference. Contrasting with prior findings obtained by endpoint analysis (Nakayama et al. 2001; Kusevic et al. 2017) we find that H3K9me catalysis by Suv39/Clr4 H3K9 is directly inhibited by Set1 products, most strongly by H3K4me3. This finding represents a rare example of direct regulation of the Suv39/Clr4 SET domain active site, beyond auto-inhibition (Iglesias et al. 2018), but is consistent with the effect H3K4me can have on other H3K9 methylases (Wang and Zhang 2001; Nishioka et al. 2002; Binda et al. 2010). 2. Locally decreased nucleosome occupancy. We find that the distribution of H3K4me3 tracks with a Set1-dependent decrease in nucleosome occupancy. It is known that specific nucleosome stabilizing factors are required for constitutive heterochromatin assembly (Yamane et al. 2011; Taneja et al. 2017) and we recently showed that repression of turnover is critical to stable spreading (Greenstein et al. 2018). Thus, the disruption of nucleosome occupancy by Set1 will antagonize heterochromatin formation. Increased mobilization is dependent on Set1's catalytic activity (**Figure 4.6D, F**), raising the question of how increased mobilization is instructed. In principle, this could occur via direct recruitment of nucleosome remodelers, or via changes in the chromatin landscape that increase

nucleosome turnover. Set1 has been shown to increase histone acetylation (Noma and Grewal 2002; Ginsburg et al. 2014), which has long been linked to decreased nucleosome occupancy (Reinke and Horz 2003; Wirén et al. 2005) and stability (Ausio and van Holde 1986; Brower-Toland et al. 2005). We observed strong, Set1-dependent increases in pan-H3ac and H4ac, as well as H3K9ac in most genes tested, including all those acting as spreading boundaries in our system (**Figure 4.7A-C**). The H3K4me3 HAT-targeting pathways in *S. pombe* are not well understood. However the involvement of HAT targeting downstream of H3K4me3 is evidenced by the moderate loss of heterochromatin containment at *rpl401p*:HSS in the in-frame deletion of the Png1, but not Png2, PHD finger, a conserved H3K4me3 targeting module (**Figure 4.7D,E**). This result is consistent with the observation that in *S. pombe,* only Png1 and not Png2 associates with a HAT (Chen et al. 2010a). We believe this phenotype indicates significant contribution of Png1 in containment given that 1. *rpl401* features a very large H3K4me3 peak (**Figure 4.2B**) and the 3′- oriented HSS we used is our strongest barrier construct, and 2. HATs likely act additively in implementing the H3K4me3 signal, as we observed increases in H3ac and H4ac, which are known to be mediated by a number of HATs including SAGA, Mst1, Mst2, and Hat1. Further, it remains possible that direct recruitment of chromatin remodelers by H3K4me works in concert with histone acetylation. Collectively, our data point to catalytic interference and reduced nucleosome occupancy working synergistically in the containment of heterochromatin spread downstream of Set1. Of note, unlike in fission yeast as documented here, in budding yeast, Set1 has a more global heterochromatin-antagonizing role, in concert with H2A.Z (Venkatasubrahmanyam et al. 2007). This suggests that Set1's role in constraining heterochromatin in euchromatin specifically may have co-evolved with H3K9me marked heterochromatin systems, with other factors regulating constitutive domains (see above).

Regulation of active and repressed chromatin states by Set1 and COMPASS.

How do the mechanisms of heterochromatin regulation we describe for Set1/COMPASS relate to its known roles in transcriptional regulation? The recruitment of Set1/COMPASS to chromatin requires H2B monoubiquitination mediated by Rad6 and Bre1 as well as interaction with the Paf1 elongation complex (Paf1C), which engages RNA polymerase and is additionally responsible for activation of Rad6 and Bre1 function on chromatin. Set1/COMPASS also associates with elongating RNA polymerase, giving rise to a characteristic pattern of H3K4 methylation states (see above). Interestingly, previous studies in fission yeast have described a role for Paf1C components Paf1 and Leo1 in antagonizing heterochromatin spreading through promoting increased histone turnover and H4K16 acetylation (Sadeghi et al. 2015; Verrier et al. 2015). Both studies tested, but did not identify, a role for Set1 in their respective systems at loci (IRC1L of the centromere, and IR-L of MAT) where we also do not detect an effect of *Δset1* even in the sensitized *Δepe1* genetic background (**Figure 4.2 Supplement**).

Several additional data support a model where Set1 and Paf1/Leo1 act in separate pathways to regulate heterochromatin spreading: 1. *set1* was not found to be epistatic to *leo1* in genome-wide genetic interaction study for heterochromatin spreading using an IRC1L reporter (Verrier et al. 2015). 2. Global H4K16 acetylation levels did not change in response to *Δset1* (Noma and Grewal 2002), whereas acetyl marks such as H3K9 and H3K14, were reduced in this background (**Figure 4.7B** and (Noma and Grewal 2002)). 3. In our repelling factor screen, *Δleo1* did not result in the characteristic spreading phenotype seen for Set1/COMPASS complex deletions (data not shown). Taken together these results describe separate mechanisms for spreading regulation by Paf1/Leo1 and Set1/COMPASS. Additionally, since the observations on

heterochromatin containment are dependent on Set1's catalytic activity (**Figure 4.5**), they are

unlikely to be related to the gene-repressive functions of Set1 that are independent of its

H3K4me catalytic activity (Lorenz et al. 2014).

The role of gene orientation in heterochromatin repulsion.

The nucleosome mobilizing effect of Set1 we document is generally strongest close to the TSS,

as is evident from *ade6,* as well as the very long *sib1* gene, where we observe lowest occupancy

and greatest Set1-dependence at the 5′ end (**Figure 4.6 Supplement 1E**). However, in the case of

rpl401, both H3K4me3 and the concomitant decrease in occupancy is much more broadly

distributed. This phenomenon is especially true for H3K4me2 which is evenly distributed

throughout the *rpl401p*:HSS reporter (**Figure 4.6 Supplement 1D**). While recruitment of HATs

via PHD- fingers is H3K4me3 specific (Li et al. 2006; Taverna et al. 2006), Suv39/Clr4 catalysis

is still impacted by H3K4me2 (**Figure 4.4C**), implying that both methylation states could work

in concert through both catalytic and nucleosome mobilization pathways to repel spreading. The

differential distributions of H3K4me and nucleosome occupancy changes we observe across

genes give rise to an orientation bias in the ability of a gene to repel heterochromatin (**Figure

4.6A**). If gene orientation can influence containment effectiveness, an orientation bias may

emerge at genomic sites where containment of silencing is critical. Such a case has indeed been

documented in mammals. Lamina Associated Domains (LADs) are gene-repressive chromatin

domains associated with the nuclear periphery that contain both H3K9 and H3K27 methylation

(reviewed in (van Steensel and Belmont 2017)), and regions immediately flanking LADs are

enriched for 5′ oriented genes and concomitant H3K4 methylation (Guelen et al. 2008). It is

unsurprising that mammalian genomes may require use of 5′ orientation more than fission yeasts,

which lack such a bias at boundaries of constitutive heterochromatin domains (**Figure 4.7 Supplement C**): yeast genes are very small, at a median length of about 1.8kb, with the H3K4me3 peak comprising on average 25% of the gene, while mouse and human genes have a median length of 16 and 20kb, respectively (**Figure 4.7 Supplement D**), yet preserve a similar TSS-localized H3K4me3 peak (Guenther et al. 2007). Thus, it is plausible that H3K4me3 signals distribute far enough across a gene to make effective boundaries in either orientation for fission yeast genes, but explains a 5′ bias for mammalian genes, where the H3K4me3 peak is restricted to a narrow fraction of the gene.

Our above results lead to a model (**Figure 4.7E**) for how facultative heterochromatin domains can be delimited in a manner that is specific in genomic space. It remains to be determined why only some, but not other euchromatically-embedded heterochromatin domains require Set1 for their containment, and we believe this may be encoded in the relative rates of local heterochromatin spreading and availability of limiting factors (Nakayama et al. 2000; Noma et al. 2006; Kagansky et al. 2009). Regardless, the gene-centered role of Set1/COMPASS we document here in constraining heterochromatin spreading gives insight into the mechanisms of locus encoded, and potentially cell-type specific, restriction of facultative gene-repressive domains, as opposed to the global means of delimiting heterochromatin that have been described to date.

MATERIALS AND METHODS

Strain and plasmid construction

Plasmids used to generate genomic integration constructs were assembled using *in vivo* recombination. *S. pombe* transformants were selected as described (Greenstein et al. 2018). XFP reporters were targeted to specific genomic locations as described (Greenstein et al. 2018). Direct gene knockout constructs were generated using long primer PCR to amplify resistance cassettes with homology to the regions surrounding the open reading frame of the target. Genomic integrations were confirmed by PCR.

Flow cytometry and FACS sorting

Cells were grown for flow cytometry experiments as described (Greenstein et al. 2018). Flow cytometry was performed using a Fortessa X20 Dual machine (Becton Dickinson, San Jose, CA) and High Throughput Sampler (HTS) module. Approximately 20,000 to 100,000 cells were collected, dependent on strain growth and volume collected. Fluorescence detection, compensation, and data analysis were as described (Al-Sady et al. 2016; Greenstein et al. 2018).

For the FACS experiment, cells were grown overnight from OD = 0.05 in YES and in the morning concentrated into a smaller volume (~3-5x) and filtered with 35–40μm mesh (Corning) to achieve 5-7k events/second on the cytometer and reduce potential for clogs. Cells were first gated for size (forward and side scatter), removal of doublet cells, the presence of the control "red" signal and then sorted into Low and High populations for "orange". Low "orange" population was defined by signal overlapping a control with no fluors. High "orange" population was defined by signal overlapping the matched background Δ*clr4* control. For each population, 16-18x10^6 cells were collected for Chromatin Immunoprecipitation and 3x10^6 cells were collected for RT-qPCR. Cells were processed for downstream analysis immediately following sorting.

Repelling Factor Screen

An h- reporter strain with "green" and "orange" at the *ura4* locus (*natMX* marked) and "red" at the *leu1* locus (*hygMX* marked) was crossed to a 408-strain subset of the *Bioneer* haploid deletion library (*kanMX* marked). Crosses were performed as described (Verrier et al. 2015; Barrales et al. 2016) with limited modifications. Briefly, crosses were arrayed onto SPAS plates using a RoToR HDA colony pinning robot (Singer) and mated for 4 days at room temperature.

The plates were incubated at 42°C for 4 days following mating to remove haploid and diploid cells, retaining spores. Resultant spores were germinated on YES medium with added Hygromycin B, G418, and nourseothricin for selection of both reporter loci and the appropriate gene deletion. The resultant colonies were passaged into liquid YES and grown overnight for flow cytometry as described above. In the morning, cells were diluted again into YES medium and grown 4-6 hours at 32°C prior to analysis via flow cytometry.

RNA extraction and quantification

Cells from log phase cultures or FACS sorted cells were pelleted supernatant was decanted, and flash frozen in liquid nitrogen. Pellets were stored at -80°C. RNA extraction was performed as described (Greenstein et al. 2018). cDNA synthesis was performed with either SuperScript RTIII or IV (Invitrogen) and an oligo dT primer (**Figure 4.2C, Figure 4.1 Supplement A, Figure 4.3 Supplement B, Figure 4.4 Supplement 1H**) or SuperScript RTIV (Invitrogen) and random hexamers (**Figure 4.4 Supplement 1A, Figure 4.6 Supplement 2C**) via the manufacturer's protocol. cDNA samples were quantified by RT-qPCR as described (Greenstein et al. 2018). Values from cDNA targets were normalized to *act1* or *pyk1*. Samples in **Figure 4.1 Supplement A, Figure 4.4 Supplement 1H, and Figure 4.6 Supplement 2C,** were normalized to the target/actin value for the Δ*clr4* strain of a matched background. For **Figure 4.4 Supplement 1A** LEFT, given that signal from *act1p* driven "red" increases by ~50% in Δ*set1* backgrounds, the target/actin values in Δ*set1* samples were multiplied by the mean ratio Δ*set1*/WT of *act1p* driven "red" signal from the 4 WT and mutant pairs in **Figure 4.4 Supplement 1A**, RIGHT. This adjusts the normalization for the up-regulation of actin observed in this background.

Chromatin immunoprecipitation

Chromatin Immunoprecipitation (ChIP) followed by qPCR was performed essentially as described (Greenstein et al. 2018) with the following modifications. For **Figure 4.4E** $16\text{-}18\times10^6$ cells of both "low" and "high" FACS populations, as well as controls, were collected and processed for ChIP. Prior to lysis, 50×10^6 cells of independently fixed *S. cerevisiae* W303 strain were added to each population as carrier. ChIP experiments with bulk populations of log phase cells were performed as described (Greenstein et al. 2018) without the addition of W303 carrier.

In **Figure 4.6 Supplement 1F**, Hht2-HA cells were grown at 25°C, 225rpm in YES+Hygromycin B from OD=0.05. After cells reached OD=0.2, G2 stall was induced by shifting the temperature to 37°C for 3 hours prior to fixation. Following lysis, sonication was performed using a Diagenode BioRuptor Pico for 20-28 rounds of 30s ON/30s rest or Diagenode BioRuptor Standard on High for 30-40 for rounds of 30s ON/30s rest. Cleared chromatin was split into equal volumes per IP after a small fraction (5-10%) was set aside as Input/WCE. 1μL of the following antibodies were added per ChIP sample: H3K9me2 (Abcam ab1220), H3K4me3 (Active Motif 39159), H3K4me2 (Active Motif 39141), H3K9ac (Active Motif 39137), H3(pan)ac (Active Motif 39064), H4(pan)ac (Active Motif 39140), HA (Abcam ab9110). 1.5uL ANTI-FLAG M2 antibody (Sigma) was added per ChIP sample. 1.4μg of H3 antibody (Active Motif 39064) was added per ChIP sample. Immune complexes were collected with Protein A Dynabeads (Thermofisher) for all ChIP samples except for the anti-FLAG ChIP samples which were collected with Protein G Dynabeads (Thermofisher). DNA was quantified by RT-qPCR and %IP (ChIP DNA / Input DNA) was calculated as described (Greenstein et al. 2018).

ChIP-Seq Sample and Library Preparation

Sample preparation and ChIP prior to sequencing was performed essentially as described (Greenstein et al. 2018) with the following modifications. 50mL of cells were grown to OD=0.6-0.8 overnight from OD=0.025. Biological duplicate samples were generated for WT, biological triplicate samples were generated for Δ*epe1*, and four biological samples were generated for Δ*epe1*Δ*set1* genotypes. Based on OD measurements, 300×10^6 cells per sample were fixed and processed for ChIP. Shearing was performed with 20 cycles of 30s ON/30s rest. Samples were not pre-cleared. Sonication efficiency was determined for each sample and only samples where DNAs averaged 200-300bp were used. Chromatin was split into two samples after 8% was set aside as input. 3μL of H3K9me2 (abcam1220) or H3K4me3 (Active Motif 39159) antibodies were added per tube and incubated overnight at 4°C with rotation. (Only H3K9me2 ChIP was performed for Δ*set1* strains. The absence of H3K4me3 was validated by ChIP qPCR in **Figure 4.2 Supplement A**). Immune complexes were collected with 30μL twice-washed Protein A Dynabeads (Invitrogen) for 3 hours at 4°C. Beads were washed as above with the exception that the Wash Buffer step was performed twice. Following incubation at 70°C for 20 minutes, DNA was eluted in 100μl of TE + 1%SDS and the beads were washed and eluted a second time with 100μl of TE + 1%SDS + 5 μl of 20mg/mL Proteinase K (Roche). Following overnight incubation at 65°C, ChIP and Input samples were purified using Machery Nagel PCR clean up kit. Library preparation for sequencing was performed as described (Inada et al. 2016; Parsa et al. 2018). Samples were sequenced on a HiSeq 4000 platform (Illumina) with a Single End 50 run. Data is available via GEO accession GSE140067.

ChIP-seq data analysis

Sliding window quality filtering and adapter trimming were carried out using Trimmomatic 0.38 (Bolger et al. 2014) before the reads were aligned to the *S. pombe* genome (Wood et al. 2002) with Bowtie2 2.3.4.2 (Langmead and Salzberg 2012) using standard end-to-end sensitive alignment. Indexed bam files were generated using SAMtools 1.9 (Li et al. 2009) "view", "sort", and "index" functions. Combined Input files and WT H3K9me2 ChIP files were generated with SAMtools "merge" function for use in normalization. Input or WT normalized signal tracks were generated using the MACS2 version 2.1.1.20160309 (Zhang et al. 2008b; Feng et al. 2012) callpeak function to generate reads per million normalized bedGraph files with the following flags: -g 1.26e7 --nomodel --extsize 200 --keep-dup auto -B --SPMR -q 0.01. The resulting pileup was normalized with the bdgcmp function via the fold enrichment method (m –FE). The resulting normalized signal track files were trimmed back to the length of the genome and converted to bigwig format using UCSCtools bedClip and bedGraphToBigWig functions. BigWig files were imported into R 3.5.1 with rtracklayer 1.40.6 (Lawrence et al. 2009). The genome was divided into 25bp bins and the average enrichment value per bin was calculated using the tileGenome and binnedAverage functions of GenomicRanges 1.32.7 (Lawrence et al. 2013). Gene annotations were imported from PomBase (Lock et al. 2019) and converted to genomic coordinates with the makeTxDbFromGFF function from GenomicFeatures 1.32.3 (Lawrence et al. 2013). Finally mean and confidence interval per each genotype were generated during signal track plotting using the DataTrack command from Gviz 1.24.0 (Hahne and Ivanek 2016). For the p-value track, reads for H3K9me2 ChIP-Seq in each isolate of each strain were extended to 200bp and counted into sliding 150bp windows beginning every 30bp using the windowCounts function from R package csaw 1.18.0 (Lun and Smyth 2016). Global background

was determined from 5kb bins and a filter of 1.7 times the global average was applied with the filterWindows function and subsetting. Composition bias was corrected using the TMM method via the normFactors and asDGEList functions and then dispersion was calculated via estimateDisp function before a generalized linear model based on genotype was fit with glmQLfit. P values result from testing a contrast between *Δepe1Δset1* and *Δepe1* based on the fitted model and summarizing the per window p values over 300bp or 1200bp bins. P values were interpreted colors based on the specified ranges and added to the signal track plots with the AnnotationTrack command from Gviz. Peaks were called with epic2 0.0.14 (Stovner and Saetrom 2019) with the following flags: --effective-genome-fraction 0.999968 -bin 200 -g 3 -fs 200 -fdr 0.05. Regions of known heterochromatin formation were imported from a previously curated list (Parsa et al. 2018). Regions were extended by 10kb on each side to account for differences in coordinates that may exist for different genome assemblies, as well as variable spreading. Peaks and known regions were plotted using Gviz (Hahne and Ivanek 2016). For the global analysis comparison between *Δepe1Δset1* and *Δepe1* genotypes, the average value per 300bp window for the input normalized H3K9me2 ChIP-Seq was computed using deeptools2 3.1.3 (Ramirez et al. 2016) function multiBigWigSummary. The counts per bin output file was read into R 3.6.0 and the mean value for each genotype was computed per bin. The log$_2$ ratio of the *Δepe1Δset1* genotype average over the *Δepe1* genotype average was computed for each bin that had H3K9me2 signal above a threshold of 1.5x the global average calculated in the same manner as for p value track.

Clr4 Purification

The chromodomain of Clr4 (residues 6-64, Clr4-CD) and SET domain (residues 192-490, Clr4-SET) were each cloned into MacroLab vector 14C containing N-terminal 6xHis and Maltose Binding Protein (MBP) tags. Full-length Clr4 was expressed from a previously described vector (Al-Sady et al. 2013). Proteins were expressed as described (Al-Sady et al. 2013) except that for Clr4-SET and full-length Clr4, LB was substituted for 2XYT medium supplemented with 10μM $ZnSO_4$. Lysis and Talon affinity resin purification (Takara Bio) and size exclusion chromatography was essentially as described (Al-Sady et al. 2013). Lysis Buffer was 100mM HEPES pH 7.5, 300mM NaCl, 10% glycerol, 7.5mM imidazole, 0.5% Triton-X100, 1mM β-mercaptoethanol, and protease inhibitors. For Clr4-SET and full-length Clr4, Triton was substituted for 0.01% Igepal NP-40. After final size exclusion chromatography, Clr4-CD was eluted into FP storage buffer (20mM HEPES pH 7.5, 100mM KCl, 10% glycerol, and 5mM β-mercaptoethanol). Clr4-SET and full-length Clr4 were eluted into Clr4 Storage Buffer (100 mM Tris pH 8.5, 100 mM KCl, 10% glycerol, 1 mM $MgCl_2$, 20μM $ZnSO_4$, and 10 mM β-mercaptoethanol). All proteins were flash frozen and stored at -80°C. Protein concentrations were determined by Sypro Ruby (Biorad) gel staining against a BSA standard curve and for Clr4-CD and Clr4-SET were verified by UV absorption at 280 nm using the theoretical extinction coefficient (ExPasy ProtParam) $88810 \text{cm}^{-1}\text{M}^{-1}$ and $98210 \text{cm}^{-1}\text{M}^{-1}$ for Clr4-CD and Clr4-SET, respectively.

Fluorescence Polarization Assay

Fluorescence polarization assay for binding of Clr4-CD or full-length Clr4 to H3 tail peptides was performed as described (Canzio et al. 2013). 10nM of H3 tail peptide with K4me0K9me0

(unmodified), K4me0K9me3, or K4me3K9me3 modifications (GenScript) was used as probe. Reactions were performed in FP buffer (20mM HEPES pH 7.5, 100mM KCl, 10% glycerol, and 0.1-0.01% NP-40 substitute), and incubated for 20 minutes at room temparature prior to measurement. Fluorescence polarization measurements and data analysis including fitting of curves were performed as described (Canzio et al. 2013).

Histone Methyltransferase Assay

Multiple turnover kinetic assays were performed as described (Al-Sady et al. 2013) with the following modifications. Reactions contained 100μM cold SAM (disulfate tosylate, Abcam) and 10-15μM ^{3H}SAM tracer (55–75 Ci/mmol, PerkinElmer) and were incubated with 1μM Suv39/Clr4-SET or full-length and varying amounts of biotinylated H3(1–20) peptide with K4me0 (unmodified), K4me2, or K4me3 (GenScript). Reactions were performed at 30°C in Clr4 Reaction Buffer (100-120mM Tris pH 8.5, 100mM KCl, 10% glycerol, 1mM MgCl2, 20μM ZnSO4, and 10mM β -mercaptoethanol).

Licor Western Blot

For western blot lysate method 2, whole cell total protein extracts were prepared as described (Al-Sady et al. 2016). For anti-FLAG western blot lysate method 1, pellets from 1mL of saturated overnight cultures were flash-frozen and then resuspended in 10% Trichloro-acetic Acid, mixed by vortexing and then incubated on ice for 10'. The precipitate was washed once with cold acetone and the pellet air-dryed and then resuspended in 40μL Tris/HCl pH 8 with 200μL 2x Laemmli sample buffer. 400μL 0.5mM glass beads were added per tube and each sample was mixed in a platform vortexer for 2x60". The bottom of the tube was then pierced

with a 26G needle and the supernatant was recovered into another tube by centrifugation. Prior to loading the gel all samples were boiled for 10' and then centrifuged at >10,000xg for 2' to remove insoluble material. Western blot was performed as described (Al-Sady et al. 2016) and the following primary antibodies were used H3K4me2 (Active Motif 39141), H3K4me3 (Active Motif 39159), anti-GAPDH (Thermo Scientific MA5-15738) and anti-FLAG M2, (Sigma). H3K4me2/3 blots were co-incubated with anti-GAPDH anti-sera and followed by both secondary antibodies as described (Al-Sady et al. 2016). For the anti-FLAG Western blot, given the size difference between GAPDH and Set1, the membrane was cut between the 50 and 75kDa bands on the ladder. The larger half was incubated with anti-FLAG 1° and then anti-mouse 2° while the smaller half was incubated separately with anti-GAPDH 1° and then anti-mouse 2°.

Table 4.1: Yeast strains used in this study

Strain	Genotype
PAS075	Locus2::*ade6p::3xE2C:hygMX* at Locus2 (between SPBC1711.11 andSPBC1711.12)
PM003	Wild-type strain: h(+); ura4-D18; leu1-32; ade6-M216; his7-366
PM004	Wild-type strain: h(-); *mat1 smt0*
PAS230	*ura4::natMX:dh:ade6p:*SF-GFP, *ade6p:*mKO2 3kb, *leu1::ade6p:*3xE2C: *hygMX*
PAS315	*ura4::natMX:dh:ade6p:*SF-GFP, *ade6p:*mKO2 3kb, *leu1::ade6p:*3xE2C: *hygMX; ash2::kanMX*
PAS316	*ura4::natMX:dh:ade6p:*SF-GFP, *ade6p:*mKO2 3kb, *leu1::ade6p:*3xE2C: *hygMX; spf1::kanMX*
PAS317	*ura4::natMX:dh:ade6p:*SF-GFP, *ade6p:*mKO2 3kb, *leu1::ade6p:*3xE2C: *hygMX; swd1::kanMX*
PAS318	*ura4::natMX:dh:ade6p:*SF-GFP, *ade6p:*mKO2 3kb, *leu1::ade6p:*3xE2C: *hygMX; swd1::kanMX*
PAS330	*ura4::natMX:dh:ade6p:*SF-GFP, *ade6p:*mKO2 3kb, *leu1::ade6p:*3xE2C: *hygMX; set1::kanMX*
PAS356	*ura4::natMX:dh:ade6p:*SF-GFP, *ade6p:*mKO2 3kb, *leu1::ade6p:*3xE2C: *hygMX; clr4::kanMX*
PAS448	IR-Rproximal::*ade6p:mKO2* (3′) ; Locus2::*act1p::1xE2C:hygMX*
PAS450	IR-Rproximal::*ade6p:mKO2* (5′) ; Locus2:: *act1p::1xE2C:hygMX*

Strain	Genotype
PAS452	*IR-R::IR-RΔB-box*; IR-Rproximal::*ade6p:mKO2* (3′) ; Locus2:: *act1p::1xE2C:hygMX*
PAS454	*IR-R::IR-RΔB-box;* IR-Rproximal::*ade6p:mKO2* (5′) ; Locus2:: *act1p::1xE2C:hygMX*
PAS459	h(-); *set11::natMX*
PAS463	*IR-R::IR-RΔB-box*; IR-Rproximal::*ade6p:mKO2* (3′) ; Locus2:: *act1p::1xE2C:hygMX; epe1::kanMX*
PAS525	h(-); *mat1* Smt0; *epe1::kanMX*
PAS528	*epe1::kanMX; set1::natMX*
PAS534	IR-Rproximal::*ade6p:mKO2* (3′) ; Locus2::*act1p::1xE2C:hygMX; set1::natMX*
PAS535	IR-Rproximal::*ade6p:mKO2* (5′) ; Locus2:: *act1p::1xE2C:hygMX; set1::natMX*
PAS536	*IR-R::IR-RΔB-box*; IR-Rproximal::*ade6p:mKO2* (3′) ; Locus2:: *act1p::1xE2C:hygMX; set1::natMX*
PAS537	*IR-R::IR-RΔB-box;* IR-Rproximal::*ade6p:mKO2* (5′) ; Locus2:: *act1p::1xE2C:hygMX; set1::natMX*
PAS583	IR-Rproximal::*rpl401p:mKO2* (3′) ; Locus2::*ade6p::3xE2C:hygMX*
PAS584	IR-Rproximal::*rpl401p:mKO2* (5′) ; Locus2::*ade6p::3xE2C:hygMX*
PAS585	*IR-R::IR-RΔB-box*; IR-Rproximal::*rpl401p:mKO2* (3′) ; Locus2::*ade6p::3xE2C:*
PAS586	*IR-R::IR-RΔB-box;* IR-Rproximal::*rpl401p:mKO2* (5′) ; Locus2::*ade6p::3xE2C:hygMX*
PAS587	IR-Rproximal::*ade6p:mKO2* (3′) ; Locus2::*ade6p::3xE2C:hygMX*
PAS588	IR-Rproximal::*ade6p:mKO2* (5′) ; Locus2::*ade6p::3xE2C:hygMX*
PAS589	*IR-R::IR-RΔB-box*; IR-Rproximal::*ade6p:mKO2* (3′) ; Locus2::*ade6p::3xE2C:hygMX,*
PAS590	*IR-R::IR-RΔB-box;* IR-Rproximal::*ade6p:mKO2* (5′) ; Locus2::*ade6p::3xE2C:hygMX*
PAS591	IR-Rproximal::*rpl401p:mKO2* (3′) ; Locus2::*ade6p::3xE2C:hygMX; epe1::kanMX*
PAS592	IR-Rproximal::*rpl401p:mKO2* (5′) ; Locus2::*ade6p::3xE2C:hygMX; epe1::kanMX*
PAS593	*IR-R::IR-RΔB-box*; IR-Rproximal::*rpl401p:mKO2* (3′) ; Locus2::*ade6p::3xE2C:hygMX,; epe1::kanMX*
PAS594	*IR-R::IR-RΔB-box;* IR-Rproximal::*rpl401p:mKO2* (5′) ; Locus2::*ade6p::3xE2C:hygMX; epe1::kanMX*
PAS595, PAS704	IR-Rproximal::*ade6p:mKO2* (3′) ; Locus2::*ade6p::3xE2C:hygMX; epe1::kanMX*
PAS596, PAS707	IR-Rproximal::*ade6p:mKO2* (5′) ; Locus2::*ade6p::3xE2C:hygMX; epe1::kanMX*
PAS597, PAS710	*IR-R::IR-RΔB-box*; IR-Rproximal::*ade6p:mKO2* (3′); Locus2::*ade6p::3xE2C:hygMX,; epe1::kanMX*

Strain	Genotype
PAS598, PAS713	*IR-R::IR-RΔB-box;* IR-Rproximal::*ade6p:mKO2* (5′) ; Locus2::*ade6p::3xE2C:hygMX; epe1::kanMX*
PAS622	IR-Rproximal::*ade6p:mKO2* (5′) ; Locus2::*ade6p::3xE2C:hygMX; set1::natMX*
PAS624	*R-R::IR-RΔB-box;* IR-Rproximal::*ade6p:mKO2* (5′) ; Locus2::*ade6p::3xE2C:hygMX; set1::natMX*
PAS625, PAS703	IR-Rproximal::*ade6p:mKO2* (3′) ; Locus2::*ade6p::3xE2C:hygMX; epe1::kanMX; set1::natMX*
PAS626, PAS708	IR-Rproximal::*ade6p:mKO2* (5′) ; Locus2::*ade6p::3xE2C:hygMX; epe1::kanMX; set1::natMX*
PAS627, PAS711	*IR-R::IR-RΔB-box*; IR-Rproximal::*ade6p:mKO2* (3′); Locus2::*ade6p::3xE2C:hygMX,; epe1::kanMX; set1::natMX*
PAS628, PAS714	*IR-R::IR-RΔB-box;* IR-Rproximal::*ade6p:mKO2* (5′) ; Locus2::*ade6p::3xE2C:hygMX; epe1::kanMX; set1::natMX*
PAS634	*mtd1::mtd1t:mKO2:mtd1t* (3′); Locus2::*ade6p::3xE2C:hygMX*
PAS635	*IR-R::IR-RΔB-box; mtd1::mtd1t:mKO2:mtd1t* (3′); Locus2::*ade6p::3xE2C:hygMX*
PAS636	*mtd1::mtd1p:mKO2:mtd1t;* Locus2::*ade6p::3xE2C:hygMX*
PAS637	*IR-R::IR-RΔB-box; mtd1::mtd1p:mKO2:mtd1t;* Locus2::*ade6p::3xE2C:hygMX*
PAS649	*mtd1::mtd1t:mKO2:mtd1t* (3′); Locus2::*ade6p::3xE2C:hygMX; epe1::kanMX*
PAS650	*IR-R::IR-RΔB-box; mtd1::mtd1t:mKO2:mtd1t* (3′); Locus2::*ade6p::3xE2C:hygMX; epe1::kanMX*
PAS651	*mtd1::mtd1p:mKO2:mtd1t;* Locus2::*ade6p::3xE2C:hygMX; epe1::kanMX*
PAS652	*IR-R::IR-RΔB-box; mtd1::mtd1p:mKO2:mtd1t;* Locus2::*ade6p::3xE2C:hygMX; epe1::kanMX*
PAS657	IR-Rproximal::*ade6p:SF-GFP:5xB-box:mKO2* (3′) ; Locus2::*act1p::1xE2C:hygMX;*
PAS667	IR-Rproximal::*ade6p:SF-GFP:5xB-box:mKO2* (3′) ; Locus2::*act1p::1xE2C:hygMX; epe1::kanMX*
PAS668	*IR-R::IR-RΔB-box;* IR-Rproximal::*ade6p:SF-GFP:5xB-box:mKO2* (3′) ; Locus2::*act1p::1xE2C:hygMX; epe1::kanMX*
PAS671	IR-Rproximal::*ade6p:SF-GFP:5xB-box:mKO2* (3′) ; Locus2::*act1p::1xE2C:hygMX; clr4::kanMX*
PAS673	IR-Rproximal::*ade6p:SF-GFP:5xB-box:mKO2* (3′) ; Locus2::*act1p::1xE2C:hygMX; set1::natMX*
PAS679	*IR-R::IR-RΔB-box;* IR-Rproximal::*rpl401p:mKO2* (3′) ; Locus2::*ade6p::3xE2C:hygMX; epe1::kanMX; set1::natMX*
PAS680	*IR-R::IR-RΔB-box;* IR-Rproximal::*rpl401p:mKO2* (5′) ; Locus2::*ade6p::3xE2C:hygMX; epe1::kanMX; set1::natMX*
PAS683	*mtd1::mtd1t:mKO2:mtd1t* (3′); Locus2::*ade6p::3xE2C:hygMX; epe1::kanMX; set1::natMX*

Strain	Genotype
PAS684	*mtd1::mtd1p:mKO2:mtd1t;* Locus2*::ade6p::3xE2C:hygMX; epe1::kanMX; set1::natMX*
PAS694	*IR-R::IR-RΔB-box;* IR-Rproximal::*rpl401p:mKO2* (5′) ; Locus2*::ade6p::3xE2C:hygMX; set1::natMX*
PAS699	*mtd1::mtd1p:mKO2:mtd1t;* Locus2*::ade6p::3xE2C:hygMX; set1::natMX*
PAS706	IR-Rproximal::*ade6p:mKO2* (5′) ; Locus2*::ade6p::3xE2C:hygMX; clr4::kanMX*
PAS715	IR-Rproximal::*rpl401p:mKO2* (3′) ; Locus2*::ade6p::3xE2C:hygMX; epe1::kanMX; set1::natMX*
PAS716	IR-Rproximal::*rpl401p:mKO2* (5′) ; Locus2*::ade6p::3xE2C:hygMX; epe1::kanMX; set1::natMX*
PAS719	*ars1::prad15:cre-EBD:LEU2; h3.2:lox:HA:hygMX:lox:T7; cdc25-22ts*
PAS720	*ars1::prad15:cre-EBD:LEU2; h3.2:lox:HA:hygMX:lox:T7; cdc25-22ts; set1::natMX*
PAS748	*set1::2xFlAG:wt-set1:3'utr:ura4*
PAS749	*set1::2xFlAG:set1C862A:3'utr:ura4*
PAS750	*set1::2xFlAG:set1G852S:3'utr:ura4*
PAS753	*IR-R::IR-RΔB-box*; IR-Rproximal::*rpl401p:mKO2* (3′) ; Locus2*::ade6p::3xE2C:hygMX,; epe1::kanMX; png1ΔPHD:ura4*
PAS764	*IR-R::IR-RΔB-box;* IR-Rproximal::*rpl401p:mKO2* (5′) ; Locus2*::ade6p::3xE2C:hygMX; epe1::kanMX; wt-set1:3'utr:natMX*
PAS765	*IR-R::IR-RΔB-box;* IR-Rproximal::*rpl401p:mKO2* (5′) ; Locus2*::ade6p::3xE2C:hygMX; epe1::kanMX; set1C862A:3'utr:natMX*
PAS766	*IR-R::IR-RΔB-box;* IR-Rproximal::*rpl401p:mKO2* (5′) ; Locus2*::ade6p::3xE2C:hygMX; epe1::kanMX; set1G852S:3'utr:natMX*
PAS772	*IR-R::IR-RΔB-box*; IR-Rproximal::*rpl401p:mKO2* (3′) ; Locus2*::ade6p::3xE2C:hygMX,; epe1+; png1ΔPHD:ura4*
PAS782	*IR-R::IR-RΔB-box*; IR-Rproximal::*rpl401p:mKO2* (3′) ; Locus2*::ade6p::3xE2C:hygMX,; epe1::kanMX; png1ΔPHD:ura4*
PAS783	*IR-R::IR-RΔB-box*; IR-Rproximal::*rpl401p:mKO2* (3′) ; Locus2*::ade6p::3xE2C:hygMX,; epe1::kanMX; png1ΔPHD:ura4; set1::natMX*

SUPPLEMENTAL FIGURES
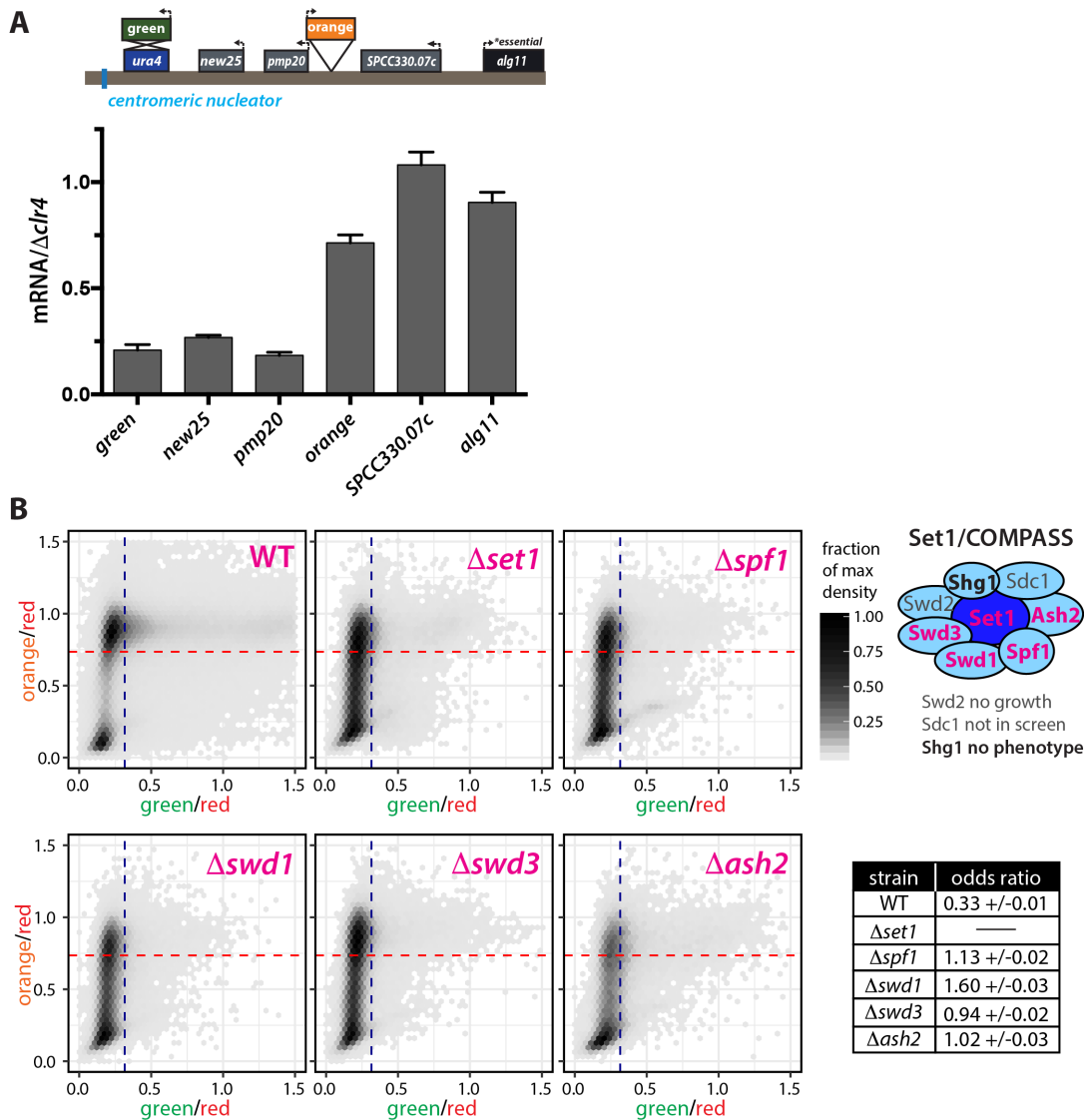
Figure 4.1 Supplement: Set1/COMPASS regulates gene-mediated heterochromatin barriers.

A. RT-qPCR analysis of gene expression at the *ura4*[dh3kb] HSS reporter (Greenstein et al 2018). Leftward of the centromeric nucleator is an inserted boundary element and expressed *natMX* selectable marker (Marina et al 2013). Expression of native genes and inserted "green" and "orange" reporters are normalized to values form a no heterochromatin *Δclr4* control. Error bars represent 1SD from three replicates, each representing a single colony deriving from each genotype. **B.** Two-dimensional density hexbin plots of red-normalized "green" and "orange" signal from WT parental strain and deletions of five Set1/COMPASS complex members. Cells with values below the nucleation cutoff (blue line) and spreading cutoff (red line) are considered repressed. (TABLE) The odds ratio calculated by Fisher's Exact Test representing the ratio of the odds of a cell occupying the silenced "off" state for Set1C mutants and WT in each strain compared to that for *Δset1*.
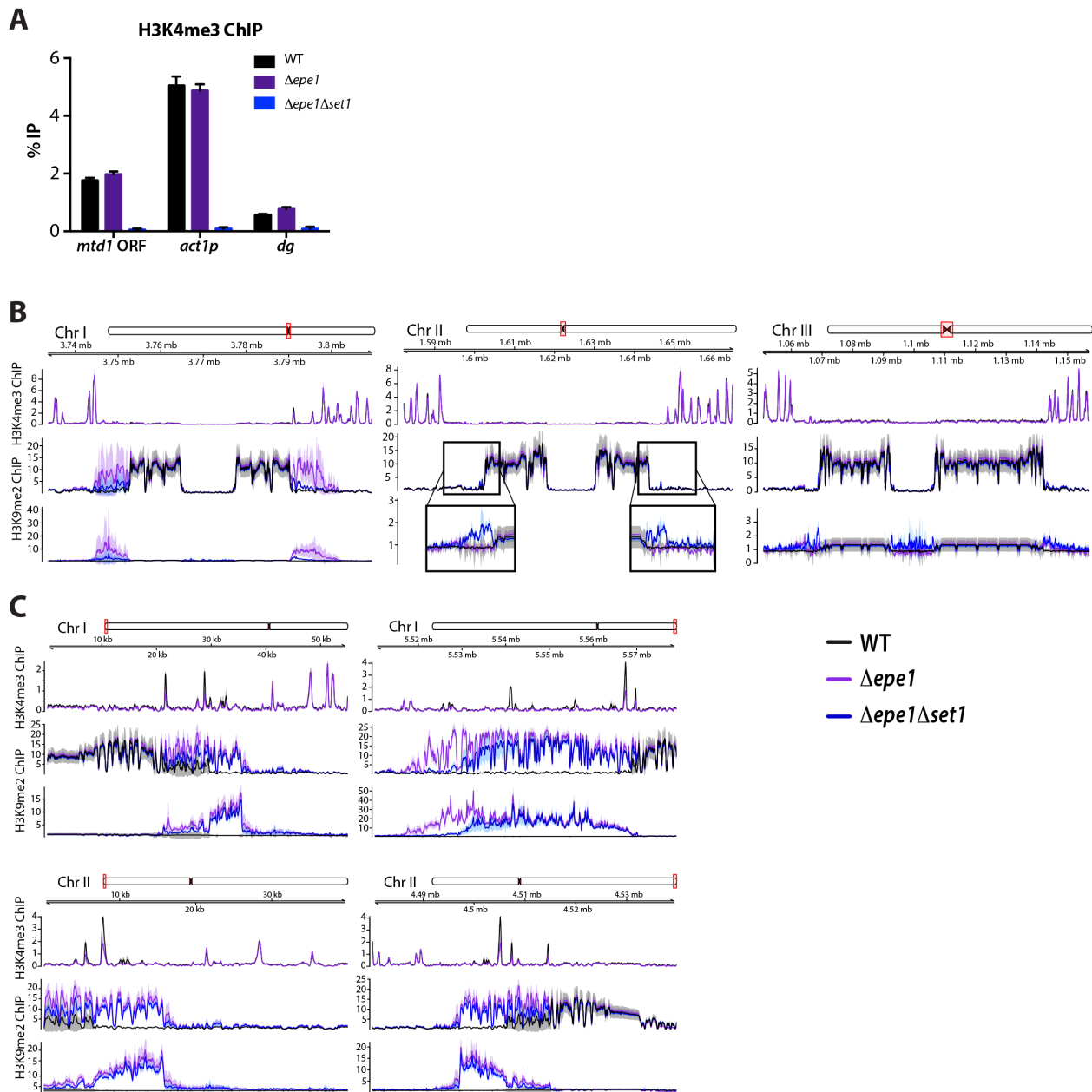
Figure 4.2 Supplement: The effect of Epe1 and Set1 at constitutive heterochromatin loci.

A. ChIP-qPCR data to validate the absence of H3K4 methylation in *Δset1;* Error bars represent 1SD from three technical replicate ChIPs. **B.** ChIP-Seq signal tracks plotted as in Figure 4.2 for centromere proximal regions. In each panel, the top track represents Input normalized H3K4me3 signal, the middle track represents Input normalized H3K9me2 signal, and the bottom track represents WT-normalized H3K9me2 signal. For centromere II, the H3K9me2 ChIP signal normalized to WT is cropped and expanded for a closer view. **C.** ChIP-Seq signal tracks for telomere proximal regions of chromosomes I and II as in B.
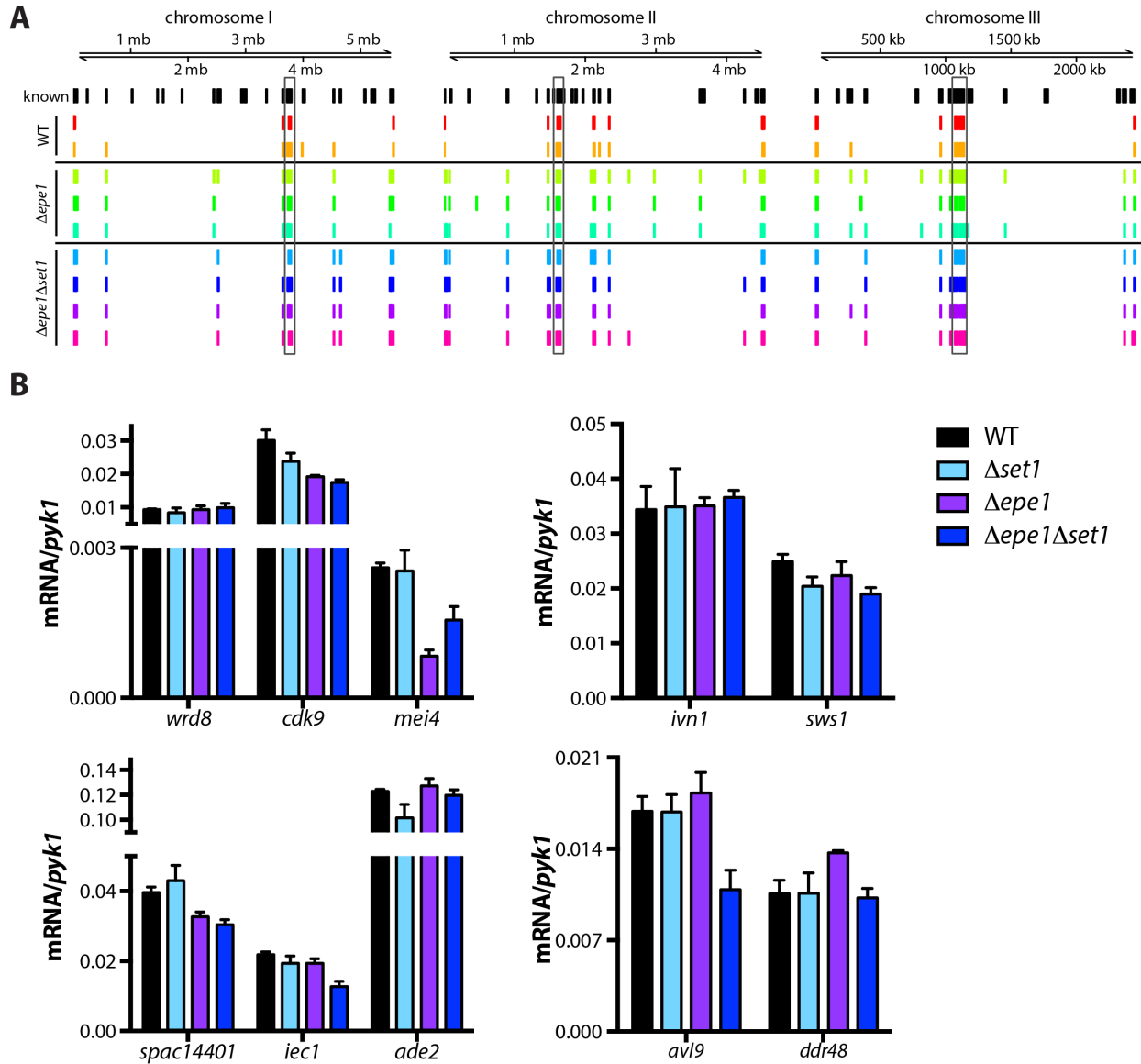
Figure 4.3 Supplement: The effect of Epe1 and Set1 at facultative heterochromatin loci.

A. The location of H3K9me2 peaks in each strain on each chromosome. Previously identified H3K9me peaks (Zofall et al. 2012; Yamanaka et al. 2013; Wang et al. 2015; Parsa et al. 2018) were included as reference (black). Known regions were extended by 10kb on each side to account for differences in coordinates that may exist for different genome assemblies, as well as variable spreading. Centromeric regions are boxed in grey. **B.** RT-qPCR analysis for heterochromatin island genes. Error bars represent 1SD of three replicate cultures each representing a single colony deriving from each genotype.

A

RT-qPCR
act1p:**red** signal

B His-MPB-Clr4-CD

C fluorescence polarization, Clr4-CD

D fluorescence polarization, full-length Clr4

E His-MBP-Clr4$_{192-490}$

peptide	K_d (μM)
H3K9me3	1.97+/-0.05
H3K4me3K9me3	1.85+/-0.16

F

H3(1-20)	WT H3	H3K4me2	H3K4me3
k_{cat} (min^{-1})	0.49+/-0.11	0.27+/-0.02	0.18+/-0.03
K_M (μM)	251+/-46	243+/-11	300+/-81
specificity (mM^{-1}min^{-1})	**1.96**	**1.11**	**0.59**

G k_{cat}/K_M full-length Clr4

H3(1-20)	WT H3	H3K4me3
specificity (mM^{-1}min^{-1})	**5.5**	**1.2**

H
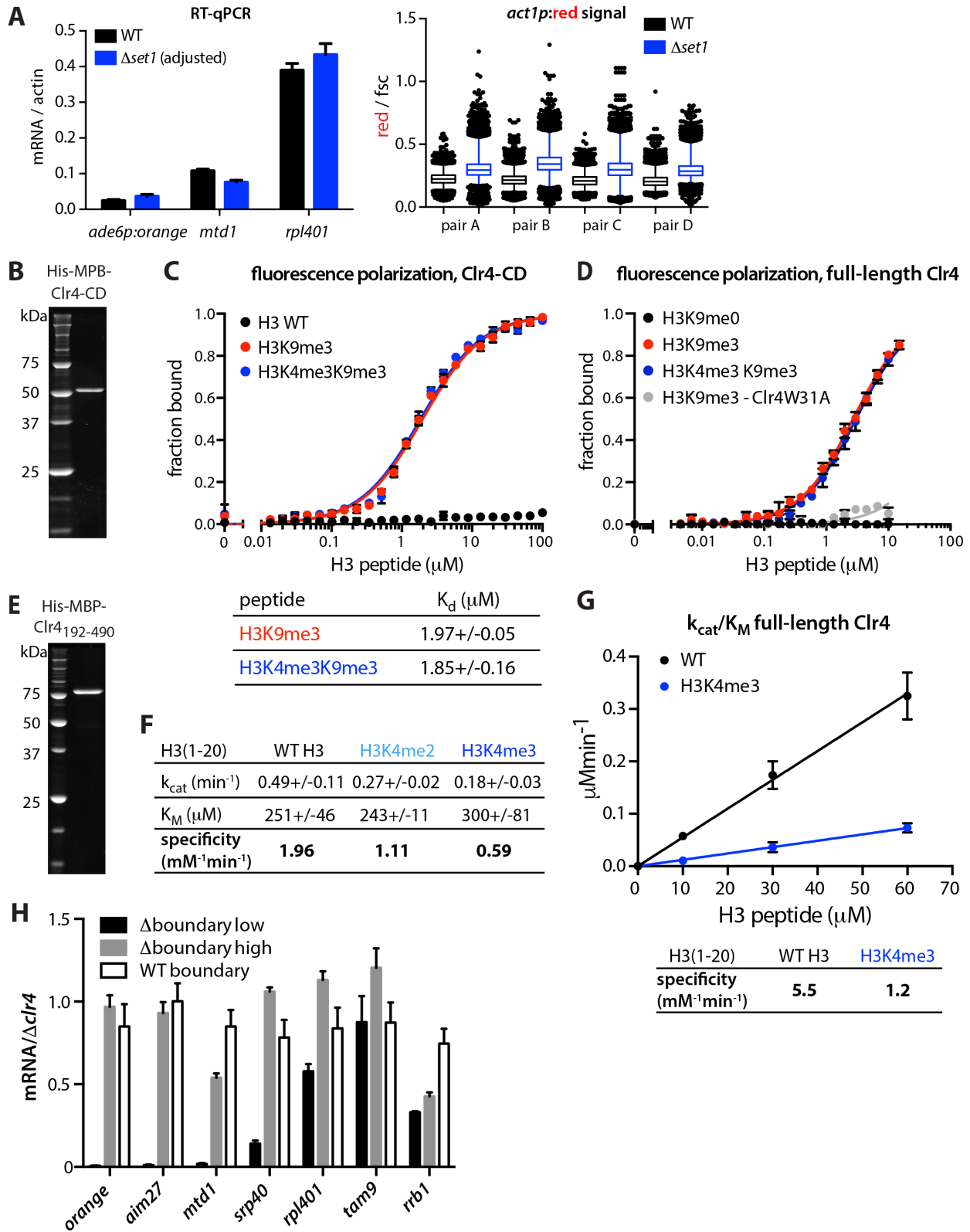
Figure 4.4 Supplement 1: Gene-protective activity of Set1/COMPASS is rooted in catalytic inhibition of Suv39/Clr4 by H3K4me2/3 and not reduction in mean transcript levels.

185

A. LEFT: RT-qPCR signal for the *ade6p*-driven "orange" transcript, as well as native *mtd1* and *rpl401* transcripts. Values are normalized to signal from *act1*/actin. In *Δset1* the *act1* signal was adjusted for the mean ratio of *act1* in *Δset1* to *set1+* as seen at RIGHT. Error bars indicate standard deviation of 3 technical replicate RNA isolations. RIGHT: Box and whisker plots of fsc-normalized "red" signal plotted as in Figure 4.3B for four independent pairs of *set1+* (black) and *Δset1* (blue). **B.** SDS-PAGE gel with His-MBP-Clr4-CD used in Figure 4.4A. **C.** Fluorescence polarization for the Suv39/Clr4 Chromodomain and fluoresceinated H3(1-20) peptides with modifications as indicated. **D.** Fluorescence polarization for full-length wild-type Suv39/Clr4 (black, red, and blue curves), full-length Clr4[W31A] (Al-Sady et al. 2013), a point mutant which disrupts the function of the chromodomain (grey curve), and fluoresceinated H3(1-20) peptides with modifications as indicated. **E.** SDS-PAGE gel with His-MBP-Clr4-192-490 (Clr4-SET) used in Figure 4.4C. **F.** Calculated kinetic parameters for Clr4-SET activity on peptides (Figure 4.4C). Values represent mean and 1SD of three independent curve fits. **G.** Histone methyltransferase assay with full-length wild-type Clr4 and H3(1-20) peptides with modifications as indicated under k_{cat}/K_M conditions. Error bars represent 1SD from three replicate experiments. Specificity constants are denoted below the plot. **H.** RT-qPCR data from "low" and "high" sorted populations and one WT isolate from Figure 4.4E. Signal is normalized to values from corresponding *Δclr4* strain. Error bars represent 1SD of three technical replicates.

Biotin-{PEG3}- ART**K**QTAR**K**STGGKAPRKQL

2612.58

Isotope Number	m/z	Percent Total	Percent Maximum
0	2611.46987	21.78	70.37
1	2612.47265	30.94	100.00
2	2613.47496	24.26	78.40
3	2614.47703	13.64	44.08
4	2615.47898	6.09	19.67
5	2616.48088	2.27	7.35
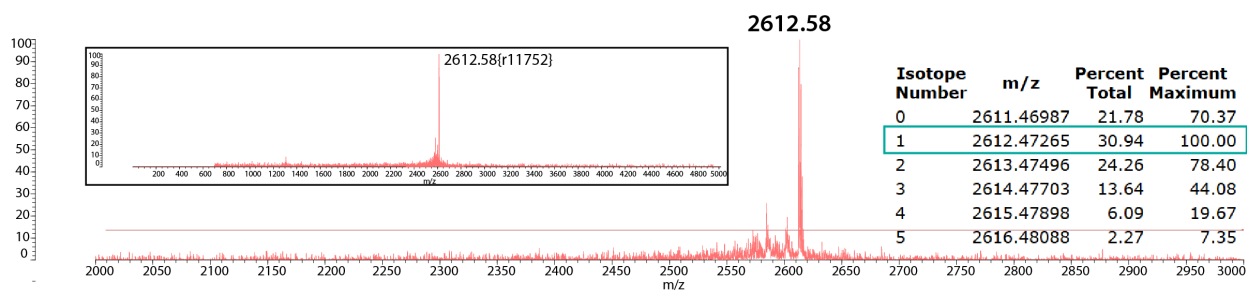
Biotin-{PEG3}- ART**K(me2)**QTAR**K**STGGKAPRKQL

2640.52

Isotope Number	m/z	Percent Total	Percent Maximum
0	2639.50117	21.28	69.25
1	2640.50396	30.73	100.00
2	2641.50629	24.41	79.43
3	2642.50838	13.88	45.16
4	2643.51035	6.26	20.37
5	2644.51227	2.36	7.69
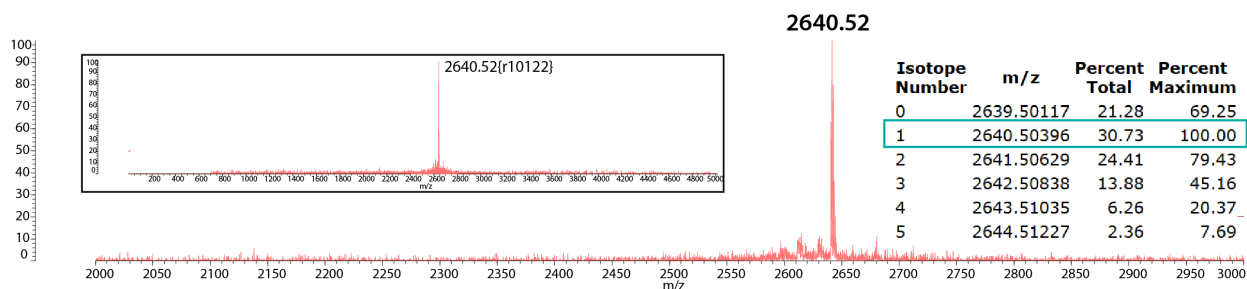
Biotin-{PEG3}- ART**K(me3)**QTAR**K**STGGKAPRKQL

2654.51

Isotope Number	m/z	Percent Total	Percent Maximum
0	2653.51682	21.04	68.71
1	2654.51962	30.63	100.00
2	2655.52196	24.48	79.94
3	2656.52406	14.00	45.71
4	2657.52604	6.35	20.72
5	2658.52797	2.41	7.86
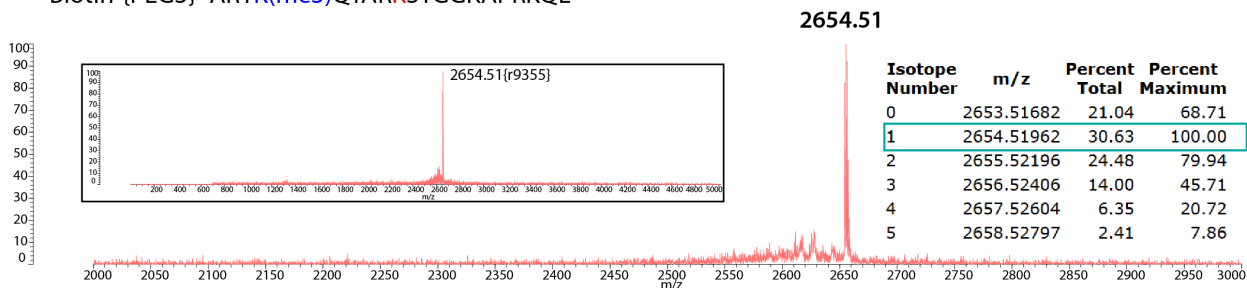
Figure 4.4 Supplement 2: MALDI-TOF validation of peptides.

Biotin(mini-PEG3)–H3 1-20 peptides modified at the K4 residue with me2, me3, or unmodified, were produced to 98% purity and HPLC purified by the vendor. For MADLI-TOF analysis, peptides were resuspended in 0.1X Clr4 reaction buffer and spotted on a MADLI-grid 1:1 with saturated α-cyano-4-hydroxy-*cinnamic* acid. Predicted masses form UCSF peptide prospector, "MS-isotope" function.
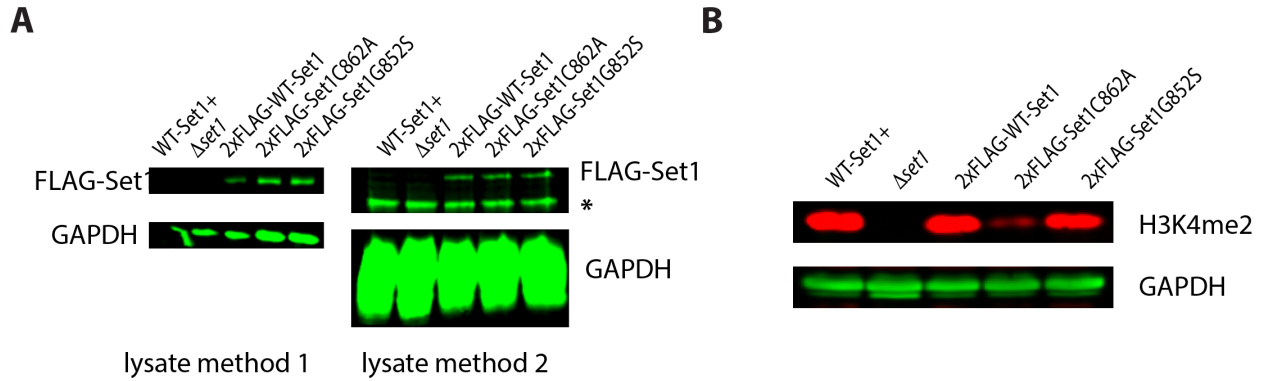
Figure 4.5 Supplement: Set1 hypomorphic alleles are expressed at wild-type levels but differ in their reduction of H3K4me2.

A. Licor Western blots for FLAG with GAPDH loading control of whole cell extracts from wild-type untagged Set1, the *Δset1* parent, 2xFLAG-wtSet1, 2xFLAG-Set1[C862A], and 2xFLAG-Set1[G852S]. Whole cell extracts were prepared by two different methods from independent cell pellets. * represents a background band. **B.** Licor Western blot for H3K4me2 with GAPDH loading control of whole cell extracts generated by lysate method 2 from strains as in A.
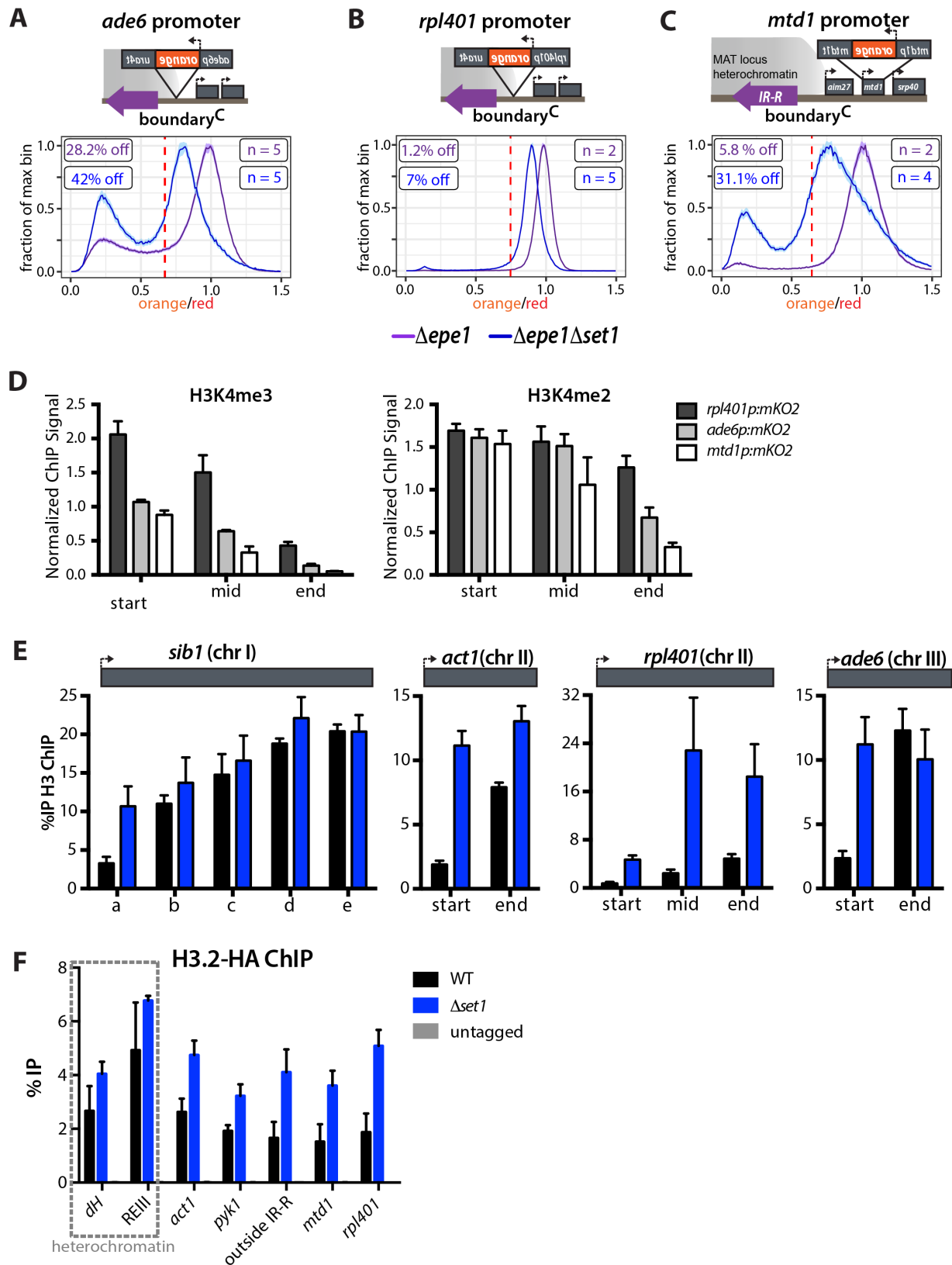
Figure 4.6 Supplement 1: Differential distributions of H3K4me3 across gene bodies can explain orientation bias in gene-mediated protection from spreading.

A. Locus cartoon for 3′ *ade6p*-driven "orange" and histogram of normalized *ade6p*- "orange" signal in boundary^C in *set1+* (purple) and *Δset1* (blue) background as in Figure 4.1. **B.** Locus cartoon for 3′ *rpl401p*-driven "orange" and histogram of normalized *rpl401p*- "orange" signal in boundary^C strains. **C.** Locus cartoon for 3′ *mtd1p*-driven "orange" and histogram of normalized *mtd1p*- "orange" signal in boundary^C strains. **D.** H3K4me3 ChIP-qPCR for targets in the start, middle, and end of the "orange" (mKO2) open reading frame in the context of 3′ *rpl401p:HSS,* 3′ *ade6p:HSS,* 3′ *mtd1p:HSS* in a WT boundary. Error bars represent 1SD of three replicate cultures from single colonies deriving from one parent isolate. For each replicate, signal at each amplicon was normalized to signal from *act1+* as an internal positive control. The resulting value was normalized to a corresponding value determined from an H3 ChIP of the same sheared chromatin to account for differences in nucleosome occupancy that could occur from differential nucleosome phasing under different promoters. **E.** H3 ChIP as in Figure 4.6F. Multiple targets per open reading frame for genes at distinct chromosomal loci were measured. Error bars represent 1SD from four replicates, each representing a single colony deriving from each genotype. **F.** HA ChIP signal from WT (black) and *Δset1* (blue) strains. Heterochromatin targets are boxed in grey. In these cells, the *hht2* was tagged with an HA epitope and they expressed the *cdc25-22ts* allele. Cells were stalled in G2 phase by shifting the temperature to 37°C for three hours prior to fixation. Error bars represent 1SD from three biological replicates.
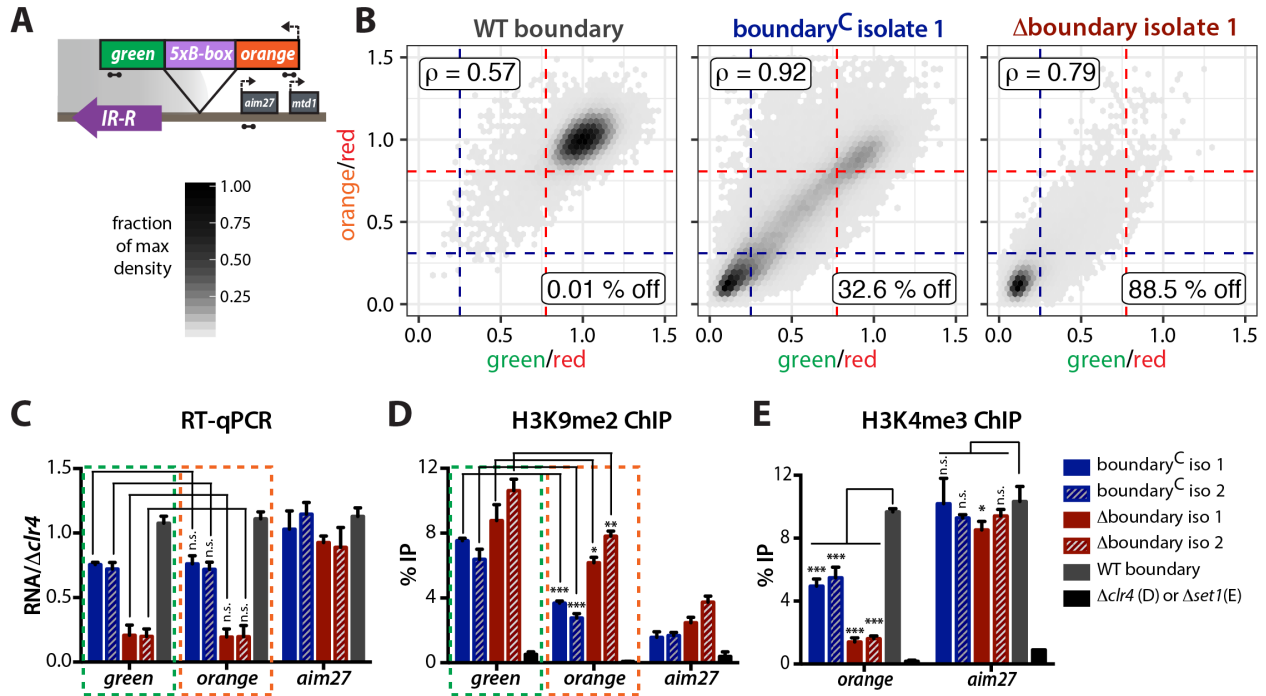
Figure 4.6 Supplement 2: 3′- invasion of a reporter gene results in silencing and reduction of TSS proximal H3K4me3.

A. Locus cartoon for 3′ *ade6p* construct which expresses "orange" and "green" ORFs joined by an in-frame linker containing 5 *B-box* sequences. **B.** 2D density hexbin plots of normalized green and orange signal for WT, boundary[C], and Δboundary isolates. All plots are normalized to the median signal from the WT boundary strain. Lines represent "on" (red) and "off" (blue) cutoffs. "On" is defined as mean of WT less 2SD while "off" is defined as the mean plus 2SD of a "red"-only control strain. Pearson correlation (ρ) for normalized "green" and "orange" and percentage of values less than the "off" cutoff for both colors are annotated. **C.** RT-qPCR data from two boundary[C], two Δboundary, and one WT isolate. **D.** H3K9me2 ChIP-qPCR data from two boundary[C] and two Δboundary isolates with *Δclr4* negative control. **E.** H3K4me3 ChIP-qPCR data from two boundary[C], two Δboundary and one WT isolate with *Δset1* negative control. Amplicons for each qPCR are depicted as dumbbells on the cartoon locus. Error bars represent 1SD of three technical replicate ChIPs. n.s. represents P>0.05, * represents P<0.05, ** represents P<0.01, *** represents P<0.001 (t-test).
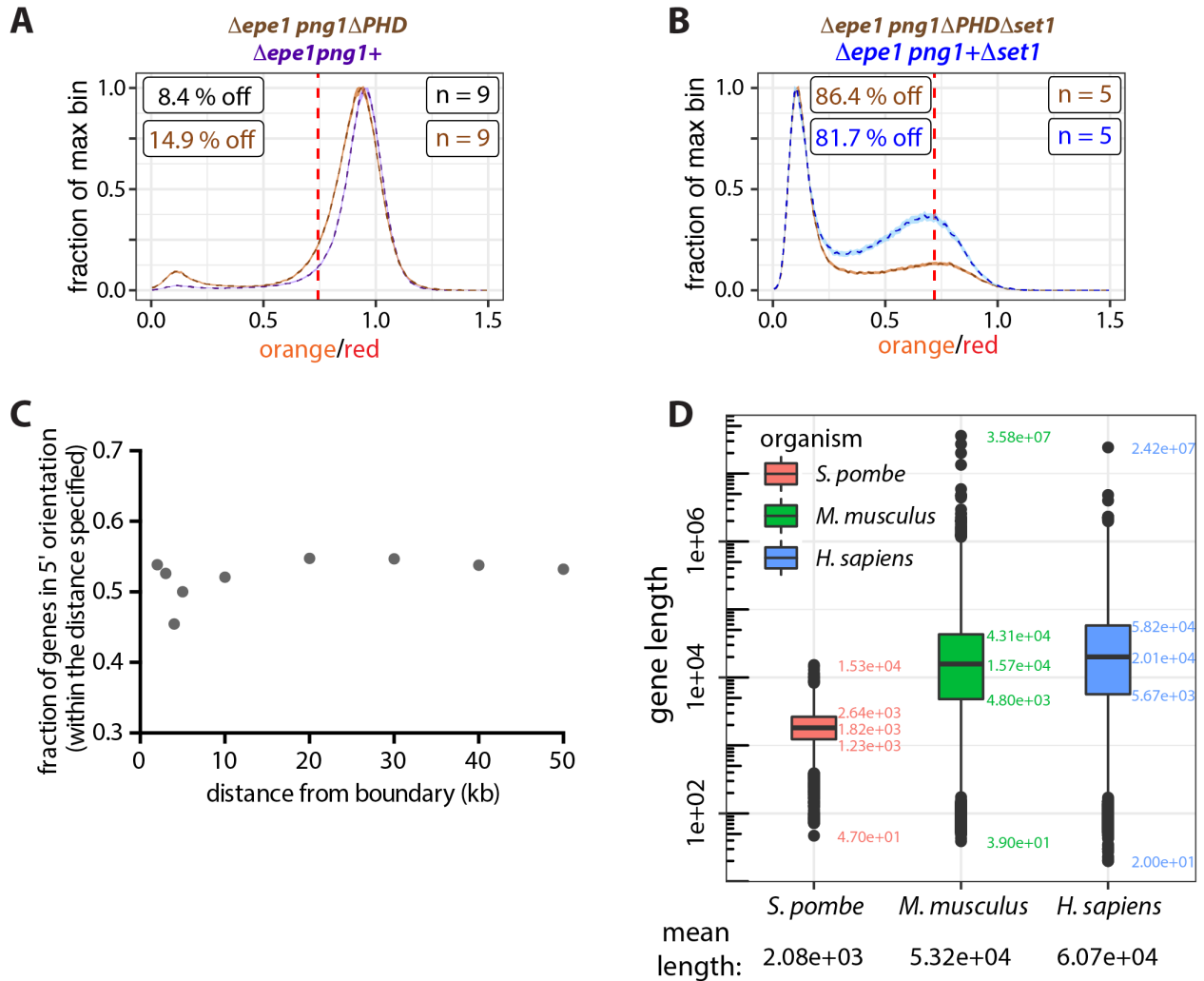
Figure 4.7 Supplement: Backcrossing and epistasis analysis of png1ΔPHD, and genome-wide analysis of gene sizes and heterochromatin-proximal orientation bias.

A. Histogram plots of normalized "orange" fluorescence as in Figure 4.1 for *Δepe1png1ΔPHD* (brown) and *Δepe1png1+* (purple) in 3′ Δboundary *rpl401p:HSS*. Isolates shown were recovered from a backcross of *Δepe1png1ΔPHD* to a wild-type strain. **B.** In the Δboundary 3′ *rpl401p:HSS* background *set1+* was concordantly knocked out (*Δset1*) in *Δepe1png1ΔPHD* (brown) and *Δepe1png1+* (blue, repeated independently from Figure 4.6B). Histogram plots of normalized "orange" fluorescence as in Figure 4.1 **C.** Analysis of gene orientation adjacent to canonical heterochromatin boundaries (centromeres and MAT locus). The fraction of genes within the distance specific (x-axis) in the 5′ orientation with respect to boundaries is plotted. **D.** Box and whisker plot of gene lengths in *S. pombe* (fission yeast), *Mus musculus* (mouse), and *Homo sapiens* (human). Values adjacent to each box plot from low to high represent the smallest value, the 25th percentile value, the median, the 75th percentile value, and the largest value. Whiskers extend from the box to the largest (upper) or smallest (lower) values within 1.5 times the interquartile range. Outlier values beyond that threshold are plotted as individual points. The mean value is denoted below the plot.

5. Unpublished Work

FOREWORD

The following chapter includes unpublished data relevant to the work described in this dissertation. It includes validation of previously described genetic phenotypes via the HSS reporter system and describes a preliminary method that could be adapted to measure the kinetic parameters of spreading *in vivo*. These data are included for completeness and so future experiments can build on this framework to address key questions in the heterochromatin field.

INTRODUCTION

While we have some understanding of the genetic requirements for H3K9 methylation spreading, much of our knowledge about mechanism of this critical cellular process relies on *in vitro* kinetic assays with of the H3K9 methylase. Biochemical experiments, though critical to elucidating biophysical and kinetic parameters of the minimal system, reveal a disconnect between what we measure *in vitro* and what behaviors we observe for spreading via static *in vivo* analyses. The HSS provides an ideal system for dissection of identified spreading regulators (ex: Suv39/Clr4, HP1/Swi6) *in vivo* by methods such as unbiased mutagenesis or targeted point mutations. Using methods described in earlier chapters of this work, we sought to validate and expand on the requirements for two known residues within the Clr4 H3K9 methylase that were previously described to alter its functions *in vivo* and *in vitro*. Specifically, we interrogated the roles of the Clr4 chromodomain (CD) and the enzyme's ability to catalyze the H3K9 tri-methyl state in H3K9me spreading and gene silencing.

One defining feature of many histone methyltransferase enzymes and/or their associated complexes is the ability to bind or "read" the chemical product of their own catalysis and in

many cases, this product recognition feeds back positively onto the system and results in stimulation of catalytic activity (Zhang et al. 2008a; Margueron et al. 2009; Al-Sady et al. 2013; Muller et al. 2016). This product recognition stimulation has been described for the *S. pombe* H3K9 methylase Clr4, whereby binding of H3K9me by the chromodomain (CD) results in accelerated catalysis via the SET domain (Al-Sady et al. 2013). Chromodomains are a conserved protein domain family that bind methylated lysines (Bannister et al. 2001; Lachner et al. 2001) via cation-Π interactions between their aromatic cage and the positively charged and methylated histone lysine residue (Hughes et al. 2007). Mutations to residues in this domain can abrogate binding via the CD without disrupting catalytic function (Nakayama et al. 2001; Zhang et al. 2008a). In addition to its role in binding H3K9me, the CD has been shown to be required for spreading of H3K9 methylation beyond nucleation sites (Noma et al. 2004). For example, in the context of a Clr4 W31G or W31A chromodomain mutant, H3K9me is found to reside only along the *cenH* ncRNA nucleation element within the MAT locus (Noma et al. 2004; Al-Sady et al. 2013) instead of spreading to coat the entire 20kb domain.

In addition to the importance of product recognition stimulation, recent work has ascribed varying roles to the methylation states of the resulting products of the Clr4 methylase. Specifically, Jih *et al* found tri-methylation of H3K9 to be required for epigenetic inheritance and transcriptional repression by heterochromatin in the context of a variant of Clr4 that could not catalyze H3K9me3 (but was competent to catalyze H3K9me2) (Jih et al. 2017). While these two functions of Clr4 are interesting on their own, studies have also described the co-dependent relationship between H3K9me3 and binding via the CD – loss of CD binding function significantly reduces H3K9me3 accumulation (Al-Sady et al. 2013), while loss of H3K9me3 catalytic function reduces Clr4 occupancy on chromatin (Jih et al. 2017). Dissecting the

interaction between these two functions of Clr4 *in vivo* will be critical to understanding how spreading works at the molecular level.

Spreading kinetics have primarily been addressed by biochemical methods, however these experimental systems are overly reductionist – typically featuring a very limited set of the components known to be involved in heterochromatin domain formation. For example, the Clr4 HMT is known to be a member of the ClrC complex (Hong et al. 2005; Horn et al. 2005; Jia et al. 2005; Li et al. 2005; Thon et al. 2005), however this is not typically accounted for in *in vitro* systems. Given that there may be many components involved in spreading, likely including unknown factors, in order to better understand cellular regulation, an *in vivo* system to measure kinetics is needed. Current analysis methods, including previous implementations of the HSS, cannot robustly address this question of spreading in living cells. In order to begin to resolve these questions, we built a pilot system that facilitates the measurement of *in vivo* kinetics. While further optimization of this assay is needed to reliably capture spreading rates from live cells, it will hopefully provide an exciting opportunity to ask key questions about how spreading dynamics can be tuned *in vivo.*

In what follows, I describe preliminary results validating the roles of H3K9me3 and the Clr4 CD in spreading as a framework for future targeted investigations of specific components of the spreading reaction. Similarly, we introduce a pilot method to capture the kinetics of spreading in live cells, which will provide a new platform to test hypotheses about requirements for this critical cellular process and better inform future *in silico* models.

RESULTS

The chromodomain of Clr4/Suv39 is required for spreading and gene silencing of HSS reporters.

We were first interested in querying the contribution of the Clr4 chromodomain to nucleation and spreading in our HSS reporter system. We generated isogenic pairs of strains with a Clr4 allele that was either wild-type or contained a W31A chromodomain point mutation that has been demonstrated does not bind H3K9me (Al-Sady et al. 2013) and measured the HSS reporter signal by flow cytometry. At each of the three reporter contexts assayed both the nucleation reporter ("green") and spreading reporter ("orange") were fully expressed (**Figure 5.1**). This result is consistent with previous reports that found the chromodomain to be required for spreading of H3K9me2 beyond nucleation sites (Noma et al. 2004; Al-Sady et al. 2013). On the other hand, it differs in that we do not detect heterochromatin formation at the nucleation site with this HSS silencing assay. However, we did not directly assess the chromatin state, for example via ChIP, at these gene units, the surrounding locus, or other genomic heterochromatin loci. It remains possible that H3K9me is present at the *cenH* ncRNA nucleator and that while our "green" nucleation reporter is integrated within this element, silencing of the reporter is still subject to a limited amount of local spreading to over the integration homology and promoter/terminator sequences. Similarly, given the relationship between H3K9me3 and the Clr4-CD as well as the requirement for H3K9me3 for gene silencing in this system (Al-Sady et al. 2013; Jih et al. 2017), it also remains possible that H3K9me2 is present at these loci. Further experiments will be needed to address these questions.
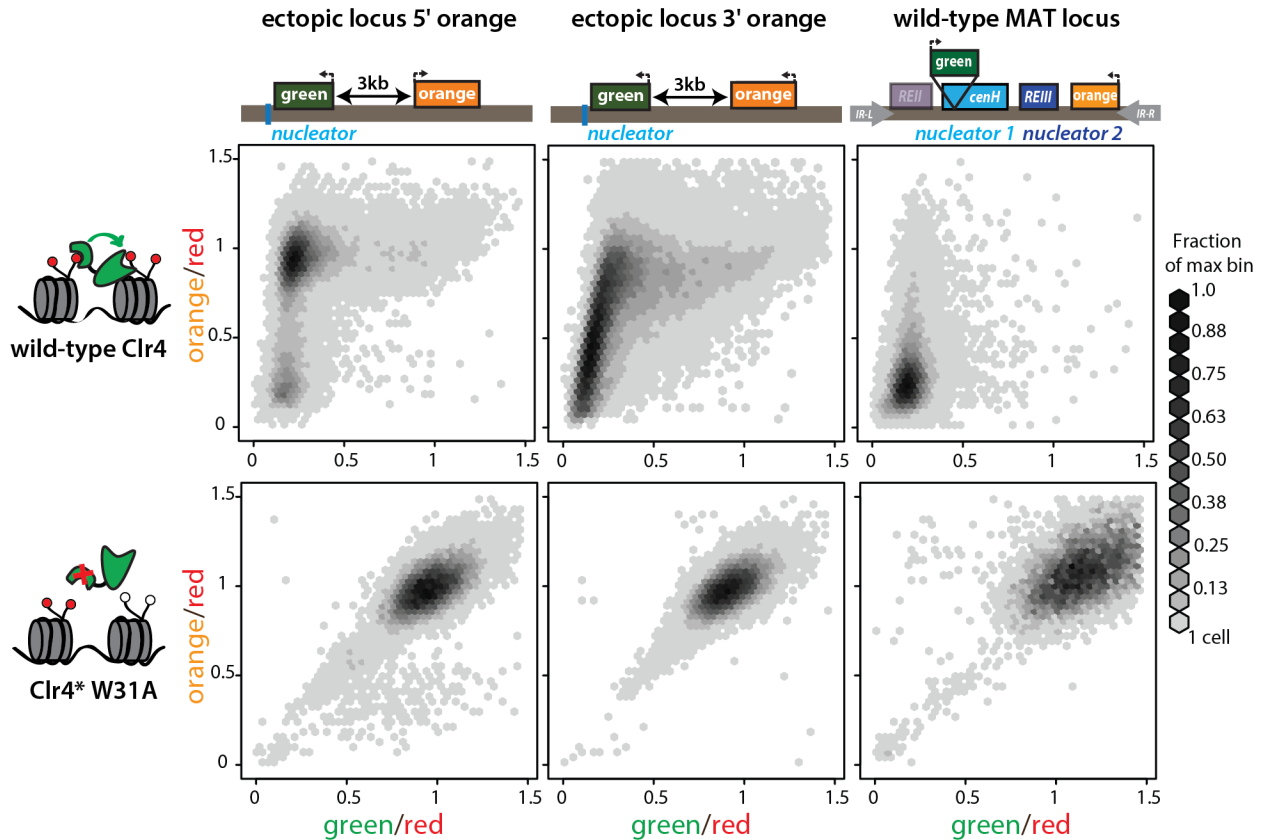
Figure 5.1: The chromodomain of Clr4 is required for spreading and gene silencing of HSS reporters.

2-D density hexbin plots of red-normalized "green" and "orange" signal for HSS strains expressing wild-type Clr4 (top) and a Clr4W31A point mutant of within the chromodomain (bottom). Reporter locations as indicated by cartoon loci. Any methylation of H3K9 depicted by red circles.

A Suv39/Clr4 hypomorph that cannot catalyze H3K9me3 disrupts gene silencing and H3K9me2 spreading.

Similarly, we were interested in assessing the contribution of H3K9me3, which is understood to be required for gene silencing (Jih et al. 2017), in spreading at the MAT locus. To query the role of H3K9me3 in these backgrounds, we generated isogenic pairs of strains with a Clr4 allele that was either wild-type or contained a F449Y point mutation in the catalytic SET domain that was previously shown to catalyze only mono- and di-methylation (Jih et al. 2017) (**Figure 5.2A**). We

measured the both "green" and "orange" signal from reporters in the three MAT locus contexts – wild-type MAT, *ΔREIII* MAT, and *ΔK* (*ΔcenH)* MAT. In all cases, the Clr4-F449Y strains failed to silence either reporter (**Figure 5.2C**), while the wild-type Clr4 isolates recapitulated the parental phenotype described previously (Greenstein et al. 2018).

To query the effect of this mutation at the chromatin level, we assessed the heterochromatin assembly mark H3K9me2 by ChIP. While we find H3K9me2 levels in the wild-type strains to be comparable to our previously published data (Greenstein et al. 2018), we cannot detect H3K9me2 at "green" or "orange" in the Clr4-F449Y strains (**Figure 5.2B**). This is at least partially consistent with the strong reduction of H3K9me2 detected at MAT in the original publication of this allele (Jih et al. 2017). We did not assess H3K9me2 at other regions where the authors detect it in this genetic context. Similar to the chromodomain mutation, it remains possible that H3K9me2 is present at the *cenH* ncRNA nucleator (but not within "green") since it is present at other ncRNA nucleators in this organism (Jih et al. 2017). However, due to the repetitive nature of this type of nucleator, it is difficult to ascertain the genomic origin of ChIP signal at these sequences. On the other hand, given the role of H3K9me3 in spreading, it could also be that even our nucleation reporter cannot be silenced due to the need for local spreading at this site. The intriguing relationship between H3K9me3, the Clr4 chromodomain, and spreading will be discussed below.
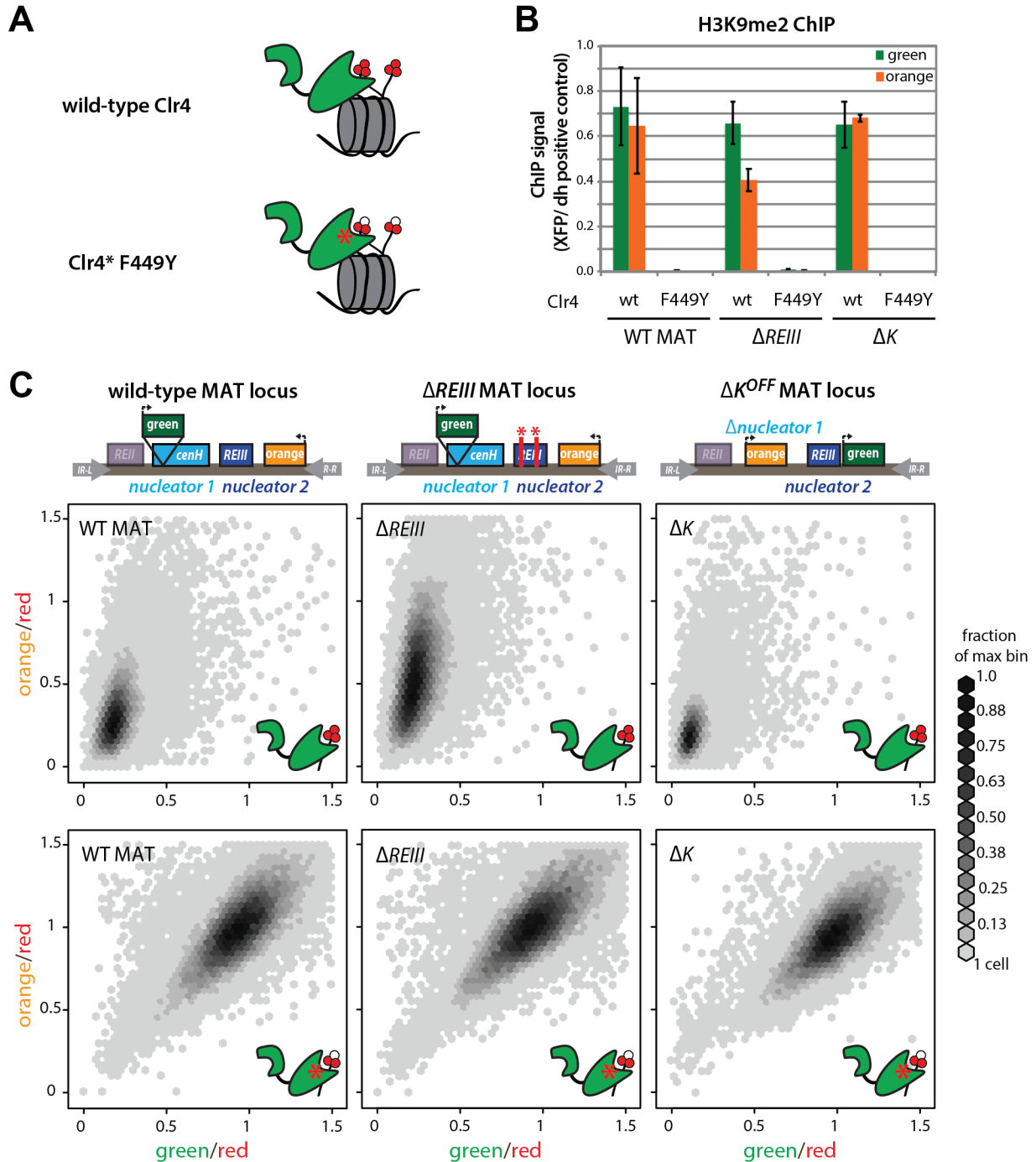
Figure 5.2: A Clr4 hypomorph that cannot catalyze H3K9me3 disrupts gene silencing and H3K9me2 spreading.

A. Cartoon depicting wild-type Clr4 which can catalyze methylation of H3K9 to the tri-methyl (me3) state and Clr4F449Y hypomorph which can catalyze H3K9me1 and H3K9me2 but not H3K9me3. Individual methylation states of H3K9 depicted by small circles – methylated (red) and unmodified (white). **B.** ChIP-qPCR measurement of H3K9me2 levels at "green" and "orange" in wild-type Clr4 and Clr4F449Y variants of three MAT locus HSS reporter

backgrounds. ChIP signal is normalized to the centromeric *dh* element as a positive control. Error bars represent 1SD from 3 technical replicates. **C.** 2-D density hexbin plots of red-normalized "green" and "orange" signal for HSS strains expressing wild-type Clr4 (top) and Clr4F449Y point mutant (bottom). Reporter locations as indicated by cartoon loci.

An inducible system for measuring heterochromatin spreading rates *in vivo*.

Much of our understanding about the structure-function relationship and kinetic parameters of spreading come from *in vitro* studies of H3K9 methyltransferase enzymes. These minimal-system assays have elucidated key features of the system including product recognition stimulation (Zhang et al. 2008a; Margueron et al. 2009; Al-Sady et al. 2013; Muller et al. 2016), non-productive binding modes (Al-Sady et al. 2013), and substrate effects (Rea et al. 2000; Schultz et al. 2002; Fischle et al. 2005; Duan et al. 2008; Binda et al. 2010; Xhemalce and Kouzarides 2010; Greenstein et al. 2019). However, the kinetic rates measured in these assays do not appear to be compatible with our understanding of the *in vivo* biology or agree with *in silico* models of the system. To address these concerns, we are interested in measuring the kinetics of spreading *in vivo,* however at the present, we lack a sensitive and robust experimental system to measure these types of changes. Here I introduce preliminary results from a pilot system that can measure the kinetics of spreading in living cells. This pilot system relies on a DNA-binding protein blockade comprised of *tetO* motifs bound by TetR proteins that significantly impede spreading from silencing a reporter (**Figure 5.3A** LEFT). With the addition of tetracycline, the TetR proteins disassociate from the DNA within five minutes (Audergon et al. 2015) and spreading can proceed to silence the reporter (**Figure 5.3A** RIGHT). We can measure the time-dependent changes in "orange" signal by flow cytometry after the addition of tetracycline as a proxy for spreading. For this pilot experiment, the TetR blockade and "orange" reporter are located ~0.8kb outside the boundary of the MAT locus and spreading beyond the MAT boundary is triggered by deletion of *epe1*. We first measured the reporter fluorescence of cells that contain

tetO binding sites but do not express TetR (**Figure 5.3B**). These samples represent the fully repressed state as found in the absence of the blockade and do not demonstrate silencing changes with tetracycline administration. Next, we measured the reporter fluorescence of cells that contain *tetO* binding sites prior to "orange" and express TetR from the low copy *spf1* promoter (**Figure 5.3C**). Prior to the administration of tetracycline, some amount of silencing is detected, however the reporter is significantly more expressed that in the absence of TetR (compare purple *Δepe1* lines in **Figure 5.3B,C**). However, when the population fluorescence is measured after tetracycline administration, time dependent changes in signal are detected (green lines).

Next, we sought to quantify the time-dependent population changes of reporter fluorescence after tetracycline administration. To do this we calculated the earth mover's distance (EMD) of *Δepe1* strains relative the their *epe1+* parent of matched tetracycline administration status (**Figure 5.3D**). In this context, the EMD value quantifies the difference between the two distributions, with lower values representing a de-repressed state and higher values indicative of reporter silencing. EMD to parent values for two different TetR expression constructs and the no TetR control at 0, 8, 24, and 34 hours of tetracycline administration are plotted (**Figure 5.3E**). From this result we can see that full spreading is achieved by 34hr of tetracycline administration and is nearly complete at 24hr. This provides an upper limit to the quantification of *in vivo* spreading kinetics as several aspects of this system need to be improved in order to increase kinetic resolution. In particular it would benefit from optimizing the turnover of the fluorescent protein encoded by "orange", perhaps through a degron system. Fluorescence decay after promoter shutoff can be measured (Al-Sady et al. 2016) and included in kinetic modeling to determine spreading rates. With finer time scale measurements and an optimized reporter system, this type of assay could yield important kinetic parameters of *in vivo* spreading

which would allow us to build *in silico* spreading models using appropriate rate constants, test genetic determinants of spreading both within the Clr4 methylase or other protein factors, and provide a framework in which to contextualize our biochemical spreading assays.
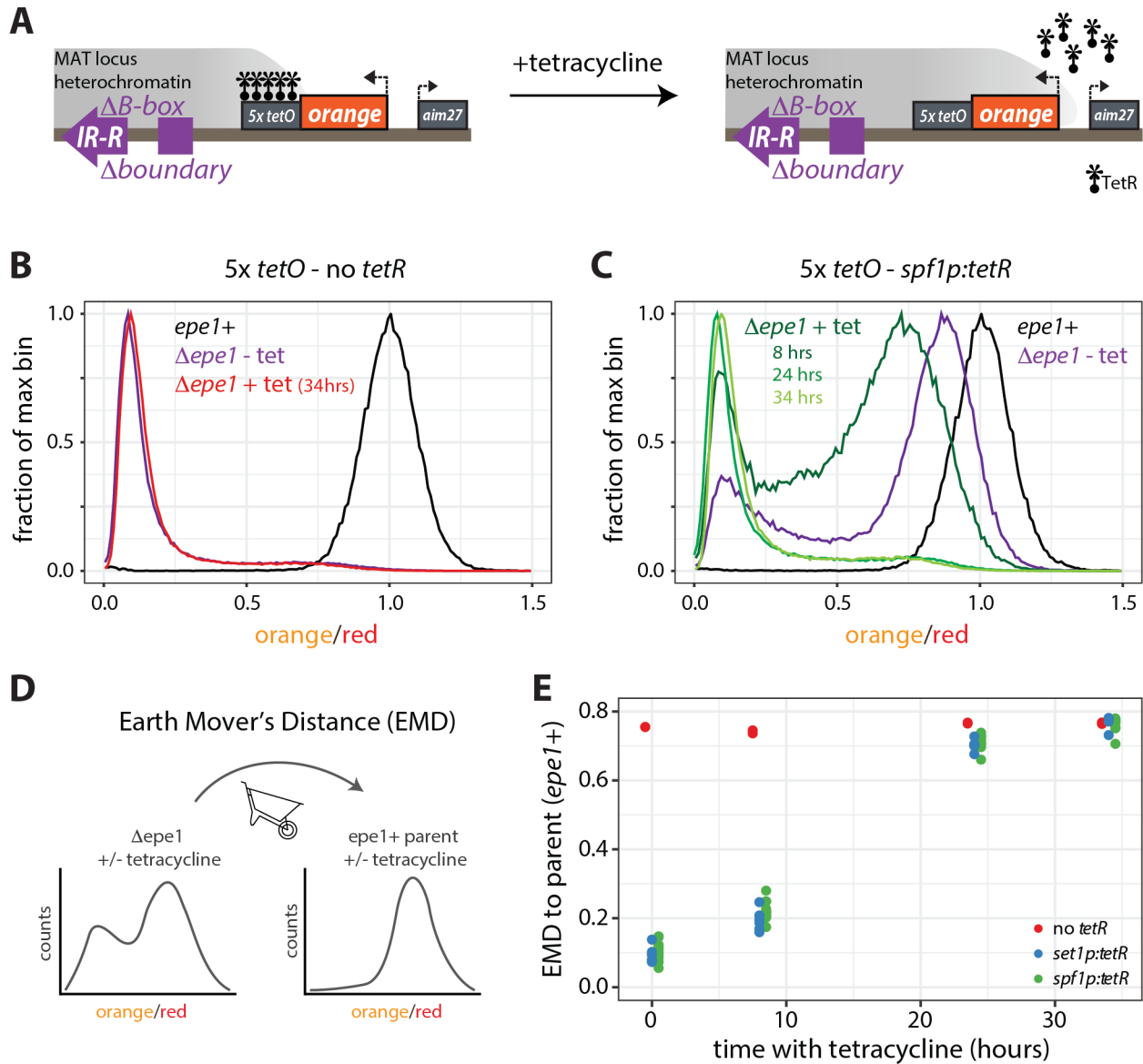
Figure 5.3: An inducible system for measuring heterochromatin spreading rates in vivo.

A. Cartoon depicting experimental set-up for tetracycline spreading experiment. The "orange" reporter is located outside a Δboundary MAT locus in the 3' orientation. 5 *tetO* binding sites are placed at the end of the terminator for "orange" such that they are located between the *IR-R* MAT locus boundary and "orange". TetR, expressed from a separate genomic location under a low copy promoter (*set1p* or *spf1p*), is bound at *tetO* in the absence of tetracycline (left) and spreading is partially blocked. When tetracycline is administered, TetR dissociates from DNA (right) and spreading proceeds into "orange". **B.** 1-D histogram of red-normalized "orange"

signal from a series of strains with the depicted reporter construct that do not contain *tetR*. Signal is normalized to the median signal from the wild-type (*epe1+*) control strain to represent the maximum fluorescence in the absence of heterochromatin (x=1). Genotype and tetracycline administration as indicated. **C.** 1-D histograms as in B for strains expressing TetR from the low copy *spf1* promoter. Genotype and tetracycline administration time course as indicated. **D.** Cartoon depiction of Earth Mover's Distance (EMD) analysis. EMD between each *Δepe1* isolate and its corresponding *epe1+* parent of matched tetracycline administration status is calculated. **E.** EMD to parent values are plotted against duration of tetracycline administration. EMD values from strains with no *tetR* in *Δepe1* (red dots, N = 2 biological replicates) represent the maximum amount of silencing expected upon long term administration of tetracycline, while EMD = 0 would indicate no silencing (as defined by *epe1+* parent). TetR is expressed in *Δepe1* from two different promoters: *spf1p* (green dots, N = 12 biological replicates) and *set1p* (blue dots, N = 6 biological replicates). Points from each genetic background are plotted with jitter on the x-axis to improve visualization of replicates, however the blocks of points represent 0hr (no tetracycline), 8hr, 24hr, and 34hr respectively.

DISCUSSION

As described throughout this dissertation, the HSS as a powerful tool for genetic dissection of spreading both at the level of genes and pathways (ex: Chapter 3 and Chapter 4) and in addition to exploring the requirement for individual amino acid residues and subdomains of protein components (ex: Chapter 4 and this chapter). Here we described an example of using the HSS to assess the spreading outcomes of targeted mutations to the H3K9 methylase Clr4 in validation of previously described results (Noma et al. 2004; Al-Sady et al. 2013; Jih et al. 2017). However, one can also imagine deploying the HSS in additional contexts such as unbiased screening of residues (ex: Error Prone PCR, Deep Mutational Scan) or whole genome approaches to add further complexity and address additional questions.

In the work described in this chapter, we validate previously described residues within the Clr4 protein that are required for both positive feedback in spreading and H3K9me3 (**Figure 5.1, Figure 5.2**). We do not detect silencing at nucleation or spreading reporters in either context. For both experiments, additional controls will be needed to assess the whole picture - namely a more complete assessment of the chromatin state via H3K9me2/3 ChIP and validation of previous findings in these alleles at other heterochromatin loci. Additionally, the lack of

H3K9me3 and H3K9me2 at MAT in the Clr4-F449Y H3K9me3 hypomorph as compared to at ncRNA nucleators like the peri-centromere (Jih et al. 2017) should be validated. Why MAT is differentially sensitive to this mutation at nucleation regions, in addition to spreading sites within the locus, remains unknown. This is interesting as it could prove a critical defining feature for a robustly inherited cell-identity locus like MAT, as opposed to a structural heterochromatin region like the centromere.

In addition, the relationship between the CD and H3K9me3 is intriguing and has implications for spreading, silencing, and epigenetic stability. The CD was first described to be required for binding of H3K9me and spreading beyond nucleation sites but not catalysis of H3K9me (Nakayama et al. 2001; Noma et al. 2004; Zhang et al. 2008a). Later, it was then shown that this may be due to the avoidance of nonproductive binding modes and product recognition stimulation which particularly effects the catalytic step from di- to tri-methylation which is much slower than from mono- to di-methylation (Al-Sady et al. 2013). Interestingly, the Clr4-CD has binding preference for H3K9me3 over H3K9me2 although H3K9me3 is the least abundant of the three PTMs *in vivo*. This has led to a linear spreading model in the field, where Clr4 binds the methyl product it catalyzes, facilitating modification of proximal nucleosomes. The precise definition of "proximal" and "linear" are active topics of investigation, though minimally we have demonstrated spreading to be distance-dependent over limited regions *in vivo*.

Recently it has been described that H3K9me3 is required for transcriptional repression (Jih et al. 2017) and that while H3K9me2 can accumulate at ncRNA centromeric nucleators, transcription is still occurring. In fact, in both the Clr4-F449Y and Clr4-W31G (similar to W31A) mutations, H3K9me2 accumulates to higher levels than in wild-type while H3K9me3 is strongly reduced. Both the CD and the ability to catalyze H3K9me3 appear to be required for

stabile epigenetic inheritance. This raises the possibility that spreading, silencing, and inheritance are interconnected through molecular regulation of Clr4 and its associated activities. The details of the mechanism remain unknown, but will undoubtedly be targets of future studies.

In this context, recent work has described autoregulation of Clr4 through methylation of an auto-inhibitory loop located between the SET and Post-SET domains (Iglesias et al. 2018). This auto-regulation appears to be required for epigenetic stability as mutants that constitutively release this auto-inhibition exhibit heterochromatin invasion beyond boundaries. The authors found the non-autoregulated variants to be more active via *in vitro* methylation assays than wild-type, though a full kinetic analysis is needed to ascribe this change to product binding or catalysis steps. On the other hand, in order to understand the role of this process in spreading we will need to measure spreading kinetics *in vivo* and in the context of other native complex components to determine how this autoregulation of Clr4 interacts with the other spreading machinery and cellular processes such as the cell cycle.

To address questions such as this, we describe here a framework for a molecular system that can be used to measure spreading rates *in vivo* (**Figure 5.3**). Though it needs some improvements to generate a more comprehensive picture – minimally: destabilized fluors, a finer time course, and fluorescence decay modelling – results from implementing this method will have important implications for our ability to understand spreading. Once optimized, this type assay will facilitate measuring the parameters of the system in their native state, potentially enabling the comparison of spreading kinetics between different loci and even assessing rates in single cells via FYLM, which in addition to steady state analysis, will aid in interrogating key aspects of this process including regulation of spreading speed, auto-methylation of Clr4, HP1 oligomerization, and mutations to ClrC members, among others. Importantly this would allow us

to build a synergistic way to generate and test hypotheses between standard steady state HSS experiments, enzymology, and *in vivo* kinetics to get a better picture of how spreading works mechanistically. Lastly, looking towards the future of computational modeling for spreading and epigenetic inheritance, it is critical to generate realistic rate constants so that we can test hypotheses *in silico* and build computational models that help us understand the relationships between spreading, silencing, and inheritance.

MATERIALS AND METHODS

Strain and plasmid construction

Plasmids used to generate genomic integration constructs were assembled using *in vivo* recombination. *S. pombe* transformants were selected as described (Greenstein et al. 2018). XFP reporters were targeted to specific genomic locations using the *ura4* FOA replacement method as described (Greenstein et al. 2018). Point mutants were generated at the native locus under the endogenous promoter. Genomic integrations and point mutations were confirmed by PCR.

Flow cytometry and data analysis

Cells were grown for flow cytometry experiments as described (Greenstein et al. 2018; Greenstein et al. 2019). Flow cytometry was performed using a Fortessa X20 Dual machine (Becton Dickinson, San Jose, CA) and High Throughput Sampler (HTS) module. Approximately 20,000 to 100,000 cells were collected, dependent on strain growth and volume collected. Fluorescence detection, compensation, and data normalization were as described (Al-Sady et al. 2016; Greenstein et al. 2018). 2-D density hexbin plots were generated as described (Greenstein

et al. 2018). 1-D histogram plots were generated as described (Greenstein et al. 2018; Greenstein et al. 2019) and plotted with ggplot2 (Wickham 2016).

Tetracycline Experiment

5 *tetO* binding sites were inserted downstream of the terminator sequence for *mKO2* "orange" reporter, between this reporter and the MAT locus boundary. *tetR* was expressed from either the *spf1* locus (N=12), *set1* locus (N=6), or not integrated into the genome (N=2). Cells were grown overnight to saturation in YES in the absence of tetracycline. In the morning cells were diluted either into YES without tetracycline or YES supplemented with 2.5 μg/mL tetracycline as described (Ragunathan et al. 2015). Measurement by flow cytometry was conducted at 8hrs post dilution. Cells were diluted into the same media conditions and then grown again overnight. The next morning cells were measured by flow cytometry (24hrs) and diluted again into the same media conditions. The final measurement was taken at 34hrs.

Earth Mover's Distance Analysis

Red-normalized orange fluorescence values for individual samples were transformed into sample population percentages for bins comprising the space between 0-1.5 in increments of 0.05 to account for variation in sample size due to sample collection. One-dimensional EMD was calculated between each *Δepe1* isolate and its corresponding *epe1+* parent isolate of matched tetracycline administration status. The EMD value between each pair of distributions was calculated using functions from the R package emdist (Urbanek and Rubner 2012). EMD values were plotted against tetracycline administration time using functions from R packages dplyr

(Wickham et al. 2020) and ggplot2 (Wickham 2016). The 0hr (no tetracycline) timepoint

represents the average EMD values of the no tetracycline strains measured at 8hrs and 24hrs.

Table 5.1: Yeast strains used in this study.

Strain	Genotype
PAS261	*ura4::natMX:dh:ade6p:*SF-GFP, *ade6p:*mKO2 3kb, *leu1*::*ade6p:*3xE2C: *hygMX; clr4::kanMX-clr4W31A*
PAS260	*ura4::natMX:dh:ade6p:*SF-GFP, *ade6p:*mKO2 3kb, *leu1*::*ade6p:*3xE2C: *hygMX; clr4::kanMX-clr4+*
PAS286	*cenH*: *ade6p:*SF-GFP (Kint2); *mat3m(Eco*RV)*::* *ade6p:*mKO2; *ade6p:*3xE2C: *hygMX* at Locus2; *clr4::kanMX-clr4+*
PAS287	*cenH*: *ade6p:*SF-GFP (Kint2); *mat3m(Eco*RV)*::* *ade6p:*mKO2; *ade6p:*3xE2C: *hygMX* at Locus2; *clr4::kanMX-clr4W31A*
PAS408	*cenH*: *ade6p:*SF-GFP (Kint2); *mat3m(Eco*RV)*::* *ade6p:*mKO2; *ade6p:*3xE2C: *hygMX* at Locus2; *clr4::natMX-clr4+*
PAS409	*cenH*: *ade6p:*SF-GFP (Kint2); *mat3m(Eco*RV)*::* *ade6p:*mKO2; *ade6p:*3xE2C: *hygMX* at Locus2; *clr4::natMX-clr4F449Y*
PAS413	*cenH*:: *ade6p:*SF-GFP (Kint2); *mat3m(Eco*RV)*::* *ade6p:*mKO2; *ade6p:*3xE2C:*hygMX* at Locus2; Δ*REIII*::REIII(Δs1, Δs2); *clr4::natMX-clr4+*
PAS414	*cenH*:: *ade6p:*SF-GFP (Kint2); *mat3m(Eco*RV)*::* *ade6p:*mKO2; *ade6p:*3xE2C:*hygMX* at Locus2; Δ*REIII*::REIII(Δs1, Δs2); *clr4::natMX-clr4F449Y*
PAS415	Δ*K*::*ade6p:*mKO2; *ade6p:* SF-GFP between *REIII* and mat3M; *act1p:1*xE2C: *hygMX* at Locus2, h(-); 'OFF' allele; *clr4::natMX-clr4+*
PAS416	Δ*K*::*ade6p:*mKO2; *ade6p:* SF-GFP between *REIII* and mat3M; *act1p:1*xE2C: *hygMX* at Locus2, h(-); *clr4::natMX-clr4F449Y*
PAS656	IR-Rproximal::*5xtetO:ura4t:mKO2:ade6p* (3′) ; Locus2*::act1p::3xE2C:hygMX*
PAS665	IR-Rproximal::*5xtetO:ura4t:mKO2:ade6p* (3′) ; Locus2*::act1p::3xE2C:hygMX; epe1::kanMX*
PAS689	IR-Rproximal::*5xtetO:ura4t:mKO2:ade6p* (3′) ; Locus2*::act1p::3xE2C:hygMX; leu1::spf1p:tetR*:3xFlag:*natMX;*
PAS691	IR-Rproximal::*5xtetO:ura4t:mKO2:ade6p* (3′) ; Locus2*::act1p::3xE2C:hygMX; leu1::set1p:tetR*:3xFlag:*natMX;*
PAS700	IR-Rproximal::*5xtetO:ura4t:mKO2:ade6p* (3′) ; Locus2*::act1p::3xE2C:hygMX; leu1::spf1p:tetR*:3xFlag:*natMX; epe1::kanMX*
PAS701	IR-Rproximal::*5xtetO:ura4t:mKO2:ade6p* (3′) ; Locus2*::act1p::3xE2C:hygMX; leu1::set1p:tetR*:3xFlag:*natMX; epe1::kanMX*

6.   Discussion

In this dissertation, I describe the results from three studies investigating the spatial and temporal regulation of heterochromatin spreading, furthering our understanding of the requirements for the formation and maintenance of epigenetic domains that delimit transcriptional programs to define cell type.

Potential mechanisms for epigenetic memory and heterochromatin inheritance

In a fission yeast model system, we find that robust inheritance of repressed states requires combinatorial action of nucleators with distinct properties (Chapter 2). In this system, spreading from ncRNA elements, which nucleate the majority of heterochromatin, is inherently labile and dynamically unstable. We find that accessory elements are required to impart epigenetic stability to these heterochromatin domains over generational time and additionally confer resistance to perturbations of the epigenetic state. This epigenetic memory is likely enacted in part through reduction of histone turnover, which retains epigenetic signals to reinforce the state.

The results from this study have important implications for our understanding of how genome partitioning, once enacted by heterochromatin spreading, is maintained with high fidelity through repeated cell division. Epigenetic states imparted through heterochromatin formation are typically conceived to be monostable or functionally bistable given that the repressed or expressed state is maintained over repeated divisions with a limited switching rate. This bistability, or epigenetic memory, realizes the inheritance of the ancestral signals that lead to repression of chromatin regions. While bistability and robust inheritance of epigenetic states have been documented across a variety of cell identity loci, it remains unclear whether this is an intrinsic feature of heterochromatin formed by the spreading process. In addition to the

variegated inheritance of epigenetic domains documented in a variety of systems, our work demonstrates that epigenetic memory is not necessarily a feature of spreading (Chapter 2 and (Greenstein et al. 2018)). In fact, the majority of heterochromatin in this organism is formed by nucleators that yield intrinsically labile spreading. Only with the addition of an accessory element that strongly reduces histone turnover can the epigenetic state be maintained at a cell identity locus. Whether the intrinsically labile nature of spreading and requirement for accessory elements to impart epigenetic stability is similarly true in higher order metazoans is a critical future avenue of investigation.

By what mechanisms is this epigenetic stability achieved in biological systems? Our work would suggest that auxiliary functions need to be built into heterochromatin domains to achieve this stability at nucleation-distal sites. In the case of the fission yeast MAT locus, the *REIII* accessory element imparts memory to the system at least in part via the reduction of histone turnover which has previously been implicated in inheritance of epigenetic states (Aygun et al. 2013; Taneja et al. 2017). Retaining modified nucleosomes through DNA replication would then promote reestablishment of heterochromatin domains following cell division by leveraging the read-write capacity and product recognition stimulation of histone methyltransferase complexes to fill in the gaps (Zhang et al. 2008a; Al-Sady et al. 2013; Ragunathan et al. 2015; Jih et al. 2017).

In addition to inheriting nucleosome signals, chromatin structure could be another potential mechanism by which memory might be enacted. For example, long range chromatin loops have been proposed to facilitate epigenetic memory in the polycomb H3K27me pathway (Bantignies and Cavalli 2011). Distally located genomic regions could potentially form long range interactions that are dependent on chromosomal context (Dekker and Heard 2015; Bonev

and Cavalli 2016) and looping is one parameter that has been found to promote memory in *in silico* models of fission yeast heterochromatin (Erdel and Greene 2016). We favor a looping or structural constraint model to help explain the extreme stability of the ΔK^{OFF} MAT locus driven by *REIII* accessory element. Unlike ncRNA nucleators which can drive heterochromatin formation at both native and ectopic locations, the function of REIII is not recapitulated when inserted at locations outside MAT (Wang and Moazed 2017; Greenstein et al. 2018). This suggest that in addition to reducing histone turnover REIII might promote memory through facilitating higher order chromatin structures. However, we cannot formally exclude that additional elements within MAT are necessary, but not sufficient, for REIII function. How chromatin structures change throughout the cell cycle (Kakui and Uhlmann 2018) and perhaps could facilitate epigenetic inheritance through S-Phase is a critical topic of future study.

Another potential mechanism that could improve retention of heterochromatic states would be direct or indirect linkage to the central cellular process of DNA replication, which occurs with high fidelity every cell cycle. It has been observed that activating and repressive chromatin domains are accurately reestablished after division (Petryk et al. 2018; Reverón-Gómez et al. 2018). Inheriting these modified nucleosomes might rely on coordination with histone chaperones (Hammond et al. 2017; Serra-Cardona and Zhang 2018) or occur indirectly through feedback with the DNA methylation pathway. In metazoans, DNA methylation is explicitly tied to DNA replication and proceeds in a well-described semi-conservative manner with high fidelity (Jones and Liang 2009; Law and Jacobsen 2010). Methylation of H3K9 is enhanced by DNA methylation (Sarraf and Stancheva 2004; Estève et al. 2006) and in some systems found to be required for the stable maintenance of H3K9me patterns (Mathieu et al. 2007). Additionally, continual presence of chromatin reader proteins has been documented to be

critical for stable heterochromatin domains (Tchasovnikarova et al. 2015). Regardless, given that the epigenomic signals can be altered by active processes such as chromatin remodeling or targeted erasure of PTMs (Ayoub et al. 2003; Lan et al. 2007; Trewick et al. 2007; Li et al. 2008; Braun et al. 2011; Geisler and Paro 2015; Zukowski and Johnson 2018) or passively though nucleosome loss via histone turnover, it is likely that systems beyond merely "copying" the epigenetic signals are required to ensure maintenance of the repressed state. Our findings in the fission yeast MAT locus suggest that collaboration between nucleation elements with distinct properties may facilitate heritable repressive heterochromatin domains by ensuring robust establishment of the silenced state and by defending the locus against antagonizing activities and promoting reestablishment through the cell cycle. Whether this holds in other systems remains to be tested.

Spatial regulation of heterochromatin domain expansion in cell type specification

Via a genetic screen, we identified SET1/COMPASS as a negative spreading regulator (Chapter 4 and (Greenstein et al. 2019)). Specifically, we found that Set1 limits the spatial expansion of facultative H3K9me domains embedded in canonical euchromatin which regulate cell-type specifying genes. Set1 constrains heterochromatin spreading both though catalysis of an enzymatic product that inhibits the H3K9 methylase and indirectly through promoting nucleosome mobilization. The catalytic function of Set1, but not its occupancy on chromatin, is required for this role in spreading containment. Consistent with these results, we found that the extent of H3K4 methylation over the body of a gene correlated with the ability of this gene to function as an effective spreading barrier independent of orientation. While this study was conducted in a fission yeast model system, it addresses long-held questions in the chromatin field about the anti-correlation of H3K9 and H3K4 methylation and provides mechanistic insight as to

how these domains might regulate each other within the cell to properly delimit hetero- and eu-chromatin.

One important aspect of chromatin biology that remains unaddressed is the regulation of spreading in development. We have little understanding of the spreading mechanism at the molecular level, the drivers of the iterative expansion process, and how limits on spatial extent are imposed in a cell-type specific manner. The mutually antagonistic relationship between heterochromatin- and euchromatin- associated activities suggests that the formation of epigenetic domains is likely a constantly occurring process, whereby activities in each domain function to continuously exclude those from the opposing state. While in principle this iterative expansion could be tuned at the level of the forward rate of the spreading reaction, given our findings, we favor that notion that differential expansion could be additionally be regulated by the selective triggering of gene-based barriers. In support of this hypothesis, preliminary analyses from other groups have documented an inhibitory effect of H3K4me on at least three mammalian H3K9 methylases (G9a, SETDB1, Suv39h1 - (Nishioka et al. 2002; Chin et al. 2005; Binda et al. 2010)) though a thorough kinetic analysis is warranted to understand both the effect of various H3K4 methylation states and the biochemical nature of this effect. Under these conditions, domain expansion could be contained in a lineage specific manner by H3K4me-decorated genes that are activated in specific cell types. From our work, it is clear that H3K4me-marked regions are not concrete boundaries to spreading so it is likely this mechanism functions in coordinating with insulators or boundary elements at least at some loci.

One interesting case that bears further study is the enrichment of 5′ proximally oriented genes at LAD boundaries in mammalian cells (Guelen et al. 2008). It has previously been suggested that the orientation of these genes and the presence of their H3K4me-marked

promoters might facilitate boundary function in this setting (Guelen et al. 2008) due to the mutual exclusion of these PTMs, though to our knowledge no direct evidence of this has been demonstrated. In our study, we provide direct evidence for gene orientation bias in spreading containment and its association with the breadth of the H3K4me3 signal relative to the gene body. While fission yeast does not harbor a similar enrichment of 5′ proximal co-oriented genes at their heterochromatin boundaries, the ratio of the mean H3K4me3 peak width to typical gene length to mean is quite large (~25%) whereas in mammals, H3K4me3 peaks are of similar width and genes are an order of magnitude longer in median length. As such, the potent H3K9 methylase-inhibitory H3K4me3 signal and concomitant destabilized nucleosomes are much more restricted around the TSS region in mammalian cells potentially limiting the ability of genes for form strong barriers from the 3′ orientation, allowing evolution to select for 5′ genes at boundaries vulnerable to heterochromatin invasion. Careful molecular genetic dissection will be critical to deciphering the relationship between gene orientation, H3K4 methylation, and spreading repulsion in mammalian systems and their potential roles in shaping developmental genome patterning.

Distinct heterochromatin for different biological needs

Our investigation of the properties of fission yeast heterochromatin suggests a model in which the ability to impart epigenetic memory to a domain is restricted to a minority of loci specialized in cell type control. In contrast, at the centromeric and telomeric regions, which the harbor the majority of H3K9me-marked domains in the fission yeast system (Grewal and Klar 1997), heterochromatin functions in genome defense and structural integrity in chromosome segregation similar to its role at these regions in metazoans (Bernard et al. 2001; Saksouk et al. 2015). At

centromeric regions, nucleation proceeds via multiple ncRNA elements that are spaced in close proximity. As the distances between these ncRNA elements is limited, efficient nucleation may supersede the need for epigenetic memory at these regions, allowing them to tolerate labile spreading. Subtelomeric regions are similarly redundant in their nucleation (Kanoh et al. 2005; Hansen et al. 2006), however spreading propagates a much further distance (Kanoh et al. 2005) though its extent is variable (Nimmo et al. 1994). The major role of telomeric heterochromatin is thought to be guarding against genome instability by repressing recombination of highly homologous subtelomere sequences (Cooper et al. 1997; Nimmo et al. 1998). However, it is likely that some stochasticity can be tolerated at these domains as long as spreading covers enough sequence to avoid this deleterious recombination.

In fission yeast, the MAT locus encodes spare copies of genes encoding cell (mating) type information that are heterochromatinized to prevent both recombination of homologous sequences and expression of both mating type cassettes which can lead to haploid meiosis and cell death (Kelly et al. 1988). This robust and stable repression depends on the collaboration of two nucleator elements that demonstrate distinct nucleation and spreading properties (Chapter 2 and (Greenstein et al. 2018)). We propose that the fitness defect incurred from low-fidelity inheritance of repression at the fission yeast MAT locus or similar loci in simple eukaryotes, could have been an important evolutionary pressure that led to the emergence of epigenetic fidelity. Additionally, we speculate that cell type specification in multicellular organisms was enabled by these fidelity-ensuring mechanisms to stably pattern the epigenome. Cell type stabilization through cell division is likely connected to the ability to protect epigenetic information from environmental variation (D'Urso and Brickner 2017). In fission yeast, we found REIII to confer strong resistance to environmental perturbations, which is a critical

function at the MAT cell identity locus as heterochromatin has been shown to be highly sensitive to ambient changes in temperature in multiple systems (Gowen and Gay 1933; Woolcock et al. 2012; Elgin and Reuter 2013; Greenstein et al. 2018). To combat this sensitivity, we propose that accessory elements might have evolved to protect against this vulnerability to a variable environment, in addition to ensuring epigenetic stability in typical conditions.

The future of heterochromatin spreading

In vitro reconstitution and *in silico* modelling are important methods to improve our understanding of biological processes. The former offers the opportunity to deconstruct a process down to the molecular level and permits us to determine the minimal parameters that define the system, while the latter allows for the opportunity to test hypotheses that may be less accessible by lab-based experimentation. Computational modelling, however, requires the investigator to choose parameters that represent biological reality. *In vitro* analysis of spreading with most minimal system, just the histone methylase alone, reveals parameters that do not appear to be compatible with *in vivo* biology (Al-Sady et al. 2013). Additionally, genetic analyses are needed to reveal additional components required to make up the difference and a careful measurement of *in vivo* kinetics is critical to understand how far off our reconstitutions are. In this dissertation, I propose a preliminary method to gain such an *in vivo* kinetic measurement and utilize the HSS in molecular dissection of spreading components both at the protein and residue level.

Using tools described in this work, we can address the genetic requirements for heterochromatin spreading independent of nucleation and expand on the functions of previously identified components of heterochromatic gene silencing as a whole. Through analysis described in this dissertation, we identified protein factors involved in regulation of spreading that are

shared amongst chromatin contexts triggered by varying nucleators or specifically required for distinct spreading environments (Chapter 3 and (Greenstein et al. 2020)). Analysis of this screen is an important first step to building a greater understanding of the genetic components required for spreading and how they function in the cell.

While further investigation is needed to validate many of these hits and uncover their mechanism of action within the spreading process, one long term goal of this work is to help define the minimal set of requirements and parameters for spreading *in vivo* which would permit us to assess the finer details of mechanism via *in vitro* analysis. Through a framework based on integrating *in vivo, in vitro,* and *in silico* measurements we can address important open questions in the chromatin field, and in particular gain a greater understanding of the functions of spreading in development and disease that may differ across genomic contexts and biological systems. Future studies will undoubtedly build on the work described here to expand our knowledge of heterochromatin spreading, in both its molecular mechanisms and role in creating and maintaining the epigenetic patterns that coordinate control of complex eukaryotic genomes.

References

Al-Sady B, Greenstein RA, El-Samad HJ, Braun S, Madhani HD. 2016. Sensitive and
    Quantitative Three-Color Protein Imaging in Fission Yeast Using Spectrally Diverse,
    Recoded Fluorescent Proteins with Experimentally-Characterized In Vivo Maturation
    Kinetics. *PLoS One* **11**: e0159292.

Al-Sady B, Madhani HD, Narlikar GJ. 2013. Division of labor between the chromodomains of
    HP1 and Suv39 methylase enables coordination of heterochromatin spread. *Molecular
    cell* **51**: 80-91.

Alabert C, Barth TK, Reveron-Gomez N, Sidoli S, Schmidt A, Jensen ON, Imhof A, Groth A.
    2015. Two distinct modes for propagation of histone PTMs across the cell cycle. *Genes
    Dev* **29**: 585-590.

Allshire RC, Javerzat JP, Redhead NJ, Cranston G. 1994. Position effect variegation at fission
    yeast centromeres. *Cell* **76**: 157-169.

Allshire RC, Madhani HD. 2018. Ten principles of heterochromatin formation and function.
    Nature Reviews Molecular Cell Biology **19**: 229-244.

Alper BJ, Job G, Yadav RK, Shanker S, Lowe BR, Partridge JF. 2013. Sir2 is required for Clr4
    to initiate centromeric heterochromatin assembly in fission yeast. *The EMBO journal* **32**:
    2321-2335.

Angel A, Song J, Dean C, Howard M. 2011. A Polycomb-based switch underlying quantitative
    epigenetic memory. *Nature* **476**: 105-108.

Angel A, Song J, Yang H, Questa JI, Dean C, Howard M. 2015. Vernalizing cold is registered
    digitally at FLC. *Proc Natl Acad Sci U S A* **112**: 4146-4151.

Audergon PN, Catania S, Kagansky A, Tong P, Shukla M, Pidoux AL, Allshire RC. 2015. Epigenetics. Restricted epigenetic inheritance of H3K9 methylation. *Science* **348**: 132-135.

Ausio J, van Holde KE. 1986. Histone hyperacetylation: its effects on nucleosome conformation and stability. *Biochemistry* **25**: 1421-1428.

Aygun O, Mehta S, Grewal SI. 2013. HDAC-mediated suppression of histone turnover promotes epigenetic stability of heterochromatin. *Nat Struct Mol Biol* **20**: 547-554.

Ayoub N, Goldshmidt I, Lyakhovetsky R, Cohen A. 2000. A fission yeast repression element cooperates with centromere-like sequences and defines a mat silent domain boundary. *Genetics* **156**: 983-994.

Ayoub N, Noma K-i, Isaac S, Kahan T, Grewal SIS, Cohen A. 2003. A novel jmjC domain protein modulates heterochromatization in fission yeast. *Mol Cell Biol* **23**: 4356-4370.

Bannister AJ, Kouzarides T. 2011. Regulation of chromatin by histone modifications. *Cell Research* **21**: 381-395.

Bannister AJ, Zegerman P, Partridge JF, Miska EA, Thomas JO, Allshire RC, Kouzarides T. 2001. Selective recognition of methylated lysine 9 on histone H3 by the HP1 chromo domain. *Nature* **410**: 120-124.

Bantignies F, Cavalli G. 2011. Polycomb group proteins: repression in 3D. *Trends in genetics : TIG* **27**: 454-464.

Barrales RR, Forn M, Georgescu PR, Sarkadi Z, Braun S. 2016. Control of heterochromatin localization and silencing by the nuclear membrane protein Lem2. *Genes Dev* **30**: 133-148.

Barter RL, Yu B. 2018. Superheat: An R package for creating beautiful and extendable heatmaps for visualizing complex data. *J Comput Graph Stat* **27**: 910-922.

Bayne EH, Bijos DA, White SA, de Lima Alves F, Rappsilber J, Allshire RC. 2014. A systematic genetic screen identifies new factors influencing centromeric heterochromatin integrity in fission yeast. *Genome biology* **15**: 481.

Bayne EH, White SA, Kagansky A, Bijos DA, Sanchez-Pulido L, Hoe KL, Kim DU, Park HO, Ponting CP, Rappsilber J et al. 2010. Stc1: a critical link between RNAi and chromatin modification required for heterochromatin integrity. *Cell* **140**: 666-677.

Bell AC, Felsenfeld G. 1999. Stopped at the border: boundaries and insulators. *Curr Opin Genet Dev* **9**: 191-198.

Berger SL. 2002. Histone modifications in transcriptional regulation. *Curr Opin Genet Dev* **12**: 142-148.

Bernard P, Maure JF, Partridge JF, Genier S, Javerzat JP, Allshire RC. 2001. Requirement of heterochromatin for cohesion at centromeres. *Science* **294**: 2539-2542.

Binda O, LeRoy G, Bua DJ, Garcia BA, Gozani O, Richard S. 2010. Trimethylation of histone H3 lysine 4 impairs methylation of histone H3 lysine 9: regulation of lysine methyltransferases by physical interaction with their substrates. *Epigenetics* **5**: 767-775.

Bintu L, Yong J, Antebi YE, McCue K, Kazuki Y, Uno N, Oshimura M, Elowitz MB. 2016. Dynamics of epigenetic regulation at the single-cell level. *Science* **351**: 720-724.

Bolger AM, Lohse M, Usadel B. 2014. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* **30**: 2114-2120.

Bonasio R, Tu S, Reinberg D. 2010. Molecular Signals of Epigenetic States. *Science* **330**: 612.

Bonev B, Cavalli G. 2016. Organization and function of the 3D genome. *Nature reviews Genetics* **17**: 661-678.

Bostick M, Kim JK, Estève P-O, Clark A, Pradhan S, Jacobsen SE. 2007. UHRF1 Plays a Role in Maintaining DNA Methylation in Mammalian Cells. *Science* **317**: 1760.

Braun S, Garcia JF, Rowley M, Rougemaille M, Shankar S, Madhani HD. 2011. The Cul4-Ddb1(Cdt)(2) ubiquitin ligase inhibits invasion of a boundary-associated antisilencing factor into heterochromatin. *Cell* **144**: 41-54.

Brock G, Pihur V, Datta S, Datta S. 2008. clValid: An R Package for Cluster Validation. *Journal of Statistical Software; Vol 1, Issue 4 (2008)*.

Brower-Toland B, Wacker DA, Fulbright RM, Lis JT, Kraus WL, Wang MD. 2005. Specific contributions of histone tails and their acetylation to the mechanical stability of nucleosomes. *J Mol Biol* **346**: 135-146.

Bühler M, Haas W, Gygi SP, Moazed D. 2007. RNAi-dependent and -independent RNA turnover mechanisms contribute to heterochromatic gene silencing. *Cell* **129**: 707-721.

Buhler M, Spies N, Bartel DP, Moazed D. 2008. TRAMP-mediated RNA surveillance prevents spurious entry of RNAs into the Schizosaccharomyces pombe siRNA pathway. *Nat Struct Mol Biol* **15**: 1015-1023.

Buhler M, Verdel A, Moazed D. 2006. Tethering RITS to a nascent transcript initiates RNAi- and heterochromatin-dependent gene silencing. *Cell* **125**: 873-886.

Buratowski S, Kim T. 2010. The role of cotranscriptional histone methylations. *Cold Spring Harb Symp Quant Biol* **75**: 95-102.

Cam HP, Sugiyama T, Chen ES, Chen X, FitzGerald PC, Grewal SIS. 2005. Comprehensive

    analysis of heterochromatin- and RNAi-mediated epigenetic control of the fission yeast

    genome. *Nature genetics* **37**: 809-819.

Canzio D, Chang EY, Shankar S, Kuchenbecker KM, Simon MD, Madhani HD, Narlikar GJ, Al-

    Sady B. 2011. Chromodomain-mediated oligomerization of HP1 suggests a nucleosome-

    bridging mechanism for heterochromatin assembly. *Molecular cell* **41**: 67-81.

Canzio D, Liao M, Naber N, Pate E, Larson A, Wu S, Marina DB, Garcia JF, Madhani HD,

    Cooke R et al. 2013. A conformational switch in HP1 releases auto-inhibition to drive

    heterochromatin assembly. *Nature* **496**: 377-381.

Capuano F, Mülleder M, Kok R, Blom HJ, Ralser M. 2014. Cytosine DNA Methylation Is Found

    in Drosophila melanogaster but Absent in Saccharomyces cerevisiae,

    Schizosaccharomyces pombe, and Other Yeast Species. *Analytical chemistry* **86**: 3697-

    3702.

Casciello F, Windloch K, Gannon F, Lee JS. 2015. Functional Role of G9a Histone

    Methyltransferase in Cancer. *Frontiers in Immunology* **6**.

Ceol CJ, Houvras Y, Jane-Valbuena J, Bilodeau S, Orlando DA, Battisti V, Fritsch L, Lin WM,

    Hollmann TJ, Ferre F et al. 2011. The histone methyltransferase SETDB1 is recurrently

    amplified in melanoma and accelerates its onset. *Nature* **471**: 513-517.

Chen ES, Zhang K, Nicolas E, Cam HP, Zofall M, Grewal SI. 2008. Cell cycle control of

    centromeric repeat transcription and heterochromatin assembly. *Nature* **451**: 734-737.

Chen JQ, Li Y, Pan X, Lei BK, Chang C, Liu ZX, Lu H. 2010a. The fission yeast inhibitor of

    growth (ING) protein Png1p functions in response to DNA damage. *The Journal of

    biological chemistry* **285**: 15786-15793.

Chen M-W, Hua K-T, Kao H-J, Chi C-C, Wei L-H, Johansson G, Shiah S-G, Chen P-S, Jeng Y-M, Cheng T-Y et al. 2010b. H3K9 Histone Methyltransferase G9a Promotes Lung Cancer Invasion and Metastasis by Silencing the Cell Adhesion Molecule Ep-CAM. *Cancer Research* **70**: 7830.

Chen X, Skutt-Kakaria K, Davison J, Ou YL, Choi E, Malik P, Loeb K, Wood B, Georges G, Torok-Storb B et al. 2012. G9a/GLP-dependent histone H3K9me2 patterning during human hematopoietic stem cell lineage commitment. *Genes Dev* **26**: 2499-2511.

Cheng TH, Gartenberg MR. 2000. Yeast heterochromatin is a dynamic structure that requires silencers continuously. *Genes Dev* **14**: 452-463.

Chin HG, Pradhan M, Esteve PO, Patnaik D, Evans TC, Jr., Pradhan S. 2005. Sequence specificity and role of proximal amino acids of the histone H3 tail on catalysis of murine G9A lysine 9 histone H3 methyltransferase. *Biochemistry* **44**: 12998-13006.

Coelho M, Dereli A, Haese A, Kuhn S, Malinovska L, DeSantis ME, Shorter J, Alberti S, Gross T, Tolic-Norrelykke IM. 2013. Fission yeast does not age under favorable conditions, but does so after stress. *Current biology : CB* **23**: 1844-1852.

Cohen A, Habib A, Laor D, Yadav S, Kupiec M, Weisman R. 2018. TOR complex 2 in fission yeast is required for chromatin-mediated gene silencing and assembly of heterochromatic domains at subtelomeres. *The Journal of biological chemistry* **293**: 8138-8150.

Collazo E, Couture JF, Bulfer S, Trievel RC. 2005. A coupled fluorescent assay for histone methyltransferases. *Anal Biochem* **342**: 86-92.

Conway JR, Lex A, Gehlenborg N. 2017. UpSetR: an R package for the visualization of intersecting sets and their properties. *Bioinformatics* **33**: 2938-2940.

Cooper JP, Nimmo ER, Allshire RC, Cech TR. 1997. Regulation of telomere length and function by a Myb-domain protein in fission yeast. *Nature* **385**: 744-747.

D'Urso A, Brickner JH. 2017. Epigenetic transcriptional memory. *Current genetics* **63**: 435-439.

D'Urso A, Takahashi YH, Xiong B, Marone J, Coukos R, Randise-Hinchliff C, Wang JP, Shilatifard A, Brickner JH. 2016. Set1/COMPASS and Mediator are repurposed to promote epigenetic transcriptional memory. *Elife* **5**.

Dekker J, Heard E. 2015. Structural and functional diversity of Topologically Associating Domains. *FEBS letters* **589**: 2877-2884.

Dirk LM, Flynn EM, Dietzel K, Couture JF, Trievel RC, Houtz RL. 2007. Kinetic manifestation of processivity during multiple methylations catalyzed by SET domain protein methyltransferases. *Biochemistry* **46**: 3905-3915.

Dodd IB, Micheelsen MA, Sneppen K, Thon G. 2007. Theoretical analysis of epigenetic cell memory by nucleosome modification. *Cell* **129**: 813-822.

Du J, Johnson LM, Jacobsen SE, Patel DJ. 2015. DNA methylation pathways and their crosstalk with histone methylation. *Nature Reviews Molecular Cell Biology* **16**: 519-532.

Duan Q, Chen H, Costa M, Dai W. 2008. Phosphorylation of H3S10 blocks the access of H3K9 by specific antibodies and histone methyltransferase. Implication in regulating chromatin dynamics and epigenetic inheritance during mitosis. *The Journal of biological chemistry* **283**: 33585-33590.

Egan ED, Braun CR, Gygi SP, Moazed D. 2014. Post-transcriptional regulation of meiotic genes by a nuclear RNA silencing complex. *Rna* **20**: 867-881.

Eissenberg JC, Morris GD, Reuter G, Hartnett T. 1992. The heterochromatin-associated protein HP-1 is an essential protein in Drosophila with dosage-dependent effects on position-effect variegation. *Genetics* **131**: 345-352.

Ekwall K, Nielsen O, Ruusala T. 1991. Repression of a mating type cassette in the fission yeast by four DNA elements. *Yeast (Chichester, England)* **7**: 745-755.

Elgin SC, Reuter G. 2013. Position-effect variegation, heterochromatin formation, and gene silencing in Drosophila. *Cold Spring Harbor perspectives in biology* **5**: a017780.

Erdel F, Greene EC. 2016. Generalized nucleation and looping model for epigenetic memory of histone modifications. *Proc Natl Acad Sci U S A* **113**: E4180-4189.

Estève PO, Chin HG, Smallwood A, Feehery GR, Gangisetty O, Karpf AR, Carey MF, Pradhan S. 2006. Direct interaction between DNMT1 and G9a coordinates DNA and histone methylation during replication. *Genes Dev* **20**: 3089-3103.

Feinberg AP, Koldobskiy MA, Göndör A. 2016. Epigenetic modulators, modifiers and mediators in cancer aetiology and progression. *Nature Reviews Genetics* **17**: 284-299.

Feng J, Liu T, Qin B, Zhang Y, Liu XS. 2012. Identifying ChIP-seq enrichment using MACS. *Nature protocols* **7**: 1728-1740.

Fischer T, Cui B, Dhakshnamoorthy J, Zhou M, Rubin C, Zofall M, Veenstra TD, Grewal SIS. 2009. Diverse roles of HP1 proteins in heterochromatin assembly and functions in fission yeast. *Proceedings of the National Academy of Sciences* **106**: 8998.

Fischle W, Tseng BS, Dormann HL, Ueberheide BM, Garcia BA, Shabanowitz J, Hunt DF, Funabiki H, Allis CD. 2005. Regulation of HP1-chromatin binding by histone H3 methylation and phosphorylation. *Nature* **438**: 1116-1122.

Flury V, Georgescu PR, Iesmantavicius V, Shimada Y, Kuzdere T, Braun S, Buhler M. 2017. The Histone Acetyltransferase Mst2 Protects Active Chromatin from Epigenetic Silencing by Acetylating the Ubiquitin Ligase Brl1. *Molecular cell* **67**: 294-307.e299.

Fuda NJ, Ardehali MB, Lis JT. 2009. Defining mechanisms that regulate RNA polymerase II transcription in vivo. *Nature* **461**: 186-192.

Gallagher PS, Larkin M, Thillainadesan G, Dhakshnamoorthy J, Balachandran V, Xiao H, Wellman C, Chatterjee R, Wheeler D, Grewal SIS. 2018. Iron homeostasis regulates facultative heterochromatin assembly in adaptive genome control. *Nat Struct Mol Biol* **25**: 372-383.

Garcia JF, Al-Sady B, Madhani HD. 2015. Intrinsic Toxicity of Unchecked Heterochromatin Spread Is Suppressed by Redundant Chromatin Boundary Functions in Schizosacchromyces pombe. *G3 (Bethesda)* **5**: 1453-1461.

Garcia JF, Dumesic PA, Hartley PD, El-Samad H, Madhani HD. 2010. Combinatorial, site-specific requirement for heterochromatic silencing factors in the elimination of nucleosome-free regions. *Genes Dev* **24**: 1758-1771.

Gaydos LJ, Wang W, Strome S. 2014. Gene repression. H3K27me and PRC2 transmit a memory of repression across generations and during development. *Science* **345**: 1515-1518.

Geisler SJ, Paro R. 2015. Trithorax and Polycomb group-dependent regulation: a tale of opposing activities. *Development (Cambridge, England)* **142**: 2876-2887.

Ginsburg DS, Anlembom TE, Wang J, Patel SR, Li B, Hinnebusch AG. 2014. NuA4 links methylation of histone H3 lysines 4 and 36 to acetylation of histones H4 and H3. *The Journal of biological chemistry* **289**: 32656-32670.

Gomez EB, Espinosa JM, Forsburg SL. 2005. Schizosaccharomyces pombe mst2+ encodes a
MYST family histone acetyltransferase that negatively regulates telomere silencing. *Mol
Cell Biol* **25**: 8887-8903.

Gottschling DE, Aparicio OM, Billington BL, Zakian VA. 1990. Position effect at S. cerevisiae
telomeres: reversible repression of Pol II transcription. *Cell* **63**: 751-762.

Gowen JW, Gay EH. 1933. EFFECT OF TEMPERATURE ON EVERSPORTING EYE
COLOR IN DROSOPHILA MELANOGASTER. *Science* **77**: 312.

Graf T. 2011. Historical origins of transdifferentiation and reprogramming. *Cell stem cell* **9**: 504-
516.

Greenstein RA, Al-Sady B. 2019. Epigenetic fates of gene silencing established by
heterochromatin spreading in cell identity and genome stability. *Current genetics* **65**:
423-428.

Greenstein RA, Barrales RR, Sanchez NA, Bisanz JE, Braun S, Al-Sady B. 2019.
Set1/COMPASS repels heterochromatin invasion at euchromatic sites by disrupting
Suv39/Clr4 activity and nucleosome stability. *Genes Dev* **34**: 99-117.

Greenstein RA, Jones SK, Spivey EC, Rybarski JR, Finkelstein IJ, Al-Sady B. 2018. Noncoding
RNA-nucleated heterochromatin spreading is intrinsically labile and requires accessory
elements for epigenetic stability. *Elife* **7:e32948**.

Greenstein RA, Ng H, Barrales RR, Tan C, Braun S, Al-Sady B. 2020. Local chromatin context
dictates the genetic determinants of the heterochromatin spreading reaction. *bioRxiv*:
2020.2005.2026.117143.

Grewal SI, Klar AJ. 1997. A recombinationally repressed region between mat2 and mat3 loci

 shares homology to centromeric repeats and regulates directionality of mating-type

 switching in fission yeast. *Genetics* **146**: 1221-1238.

Grewal SIS, Bonaduce MJ, Klar AJS. 1998. Histone deacetylase homologs regulate epigenetic

 inheritance of transcriptional silencing and chromosome segregation in fission yeast.

 Genetics **150**: 563-576.

Grewal SIS, Klar AJS. 1996. Chromosomal inheritance of epigenetic states in fission yeast

 during mitosis and meiosis. *Cell* **86**: 95-101.

Gu Z, Eils R, Schlesner M. 2016. Complex heatmaps reveal patterns and correlations in

 multidimensional genomic data. *Bioinformatics* **32**: 2847-2849.

Guarente L. 2000. Sir2 links chromatin silencing, metabolism, and aging. *Genes Dev* **14**: 1021-

 1026.

Guelen L, Pagie L, Brasset E, Meuleman W, Faza MB, Talhout W, Eussen BH, de Klein A,

 Wessels L, de Laat W et al. 2008. Domain organization of human chromosomes revealed

 by mapping of nuclear lamina interactions. *Nature* **453**: 948-951.

Guenther MG, Levine SS, Boyer LA, Jaenisch R, Young RA. 2007. A chromatin landmark and

 transcription initiation at most promoters in human cells. *Cell* **130**: 77-88.

Hahne F, Ivanek R. 2016. Visualizing Genomic Data Using Gviz and Bioconductor. *Methods in

 molecular biology (Clifton, NJ)* **1418**: 335-351.

Hall IM, Shankaranarayana GD, Noma KI, Ayoub N, Cohen A, Grewal SIS. 2002.

 Establishment and maintenance of a heterochromatin domain. *Science* **297**: 2232-2237.

Hammond CM, Strømme CB, Huang H, Patel DJ, Groth A. 2017. Histone chaperone networks

 shaping chromatin function. *Nat Rev Mol Cell Biol* **18**: 141-158.

Hansen KR, Hazan I, Shanker S, Watt S, Verhein-Hansen J, Bahler J, Martienssen RA, Partridge JF, Cohen A, Thon G. 2011. H3K9me-independent gene silencing in fission yeast heterochromatin by Clr5 and histone deacetylases. *PLoS Genet* **7**: e1001268.

Hansen KR, Ibarra PT, Thon G. 2006. Evolutionary-conserved telomere-linked helicase genes of fission yeast are repressed by silencing factors, RNAi components and the telomere-binding protein Taz1. *Nucleic Acids Res* **34**: 78-88.

Hathaway NA, Bell O, Hodges C, Miller EL, Neel DS, Crabtree GR. 2012. Dynamics and memory of heterochromatin in living cells. *Cell* **149**: 1447-1460.

Ho L, Crabtree GR. 2010. Chromatin remodelling during development. *Nature* **463**: 474-484.

Hong EJ, Villén J, Gerace EL, Gygi SP, Moazed D. 2005. A cullin E3 ubiquitin ligase complex associates with Rik1 and the Clr4 histone H3-K9 methyltransferase and is required for RNAi-mediated heterochromatin formation. *RNA biology* **2**: 106-111.

Horn PJ, Bastie JN, Peterson CL. 2005. A Rik1-associated, cullin-dependent E3 ubiquitin ligase is essential for heterochromatin formation. *Genes Dev* **19**: 1705-1714.

Hughes RM, Wiggins KR, Khorasanizadeh S, Waters ML. 2007. Recognition of trimethyllysine by a chromodomain is not driven by the hydrophobic effect. *Proceedings of the National Academy of Sciences* **104**: 11184.

Iglesias N, Currie MA, Jih G, Paulo JA, Siuti N, Kalocsay M, Gygi SP, Moazed D. 2018. Automethylation-induced conformational switch in Clr4 (Suv39h) maintains epigenetic stability. *Nature* **560**: 504-508.

Inada M, Nichols RJ, Parsa JY, Homer CM, Benn RA, Hoxie RS, Madhani HD, Shuman S, Schwer B, Pleiss JA. 2016. Phospho-site mutants of the RNA Polymerase II C-terminal

domain alter subtelomeric gene expression and chromatin modification state in fission yeast. *Nucleic Acids Res* **44**: 9180-9189.

Jacob F, Monod J. 1961. Genetic regulatory mechanisms in the synthesis of proteins. *Journal of Molecular Biology* **3**: 318-356.

Jacobs SA, Khorasanizadeh S. 2002. Structure of HP1 Chromodomain Bound to a Lysine 9-Methylated Histone H3 Tail. *Science* **295**: 2080.

Janssen A, Colmenares SU, Karpen GH. 2018. Heterochromatin: Guardian of the Genome. *Annual Review of Cell and Developmental Biology* **34**: 265-288.

Jia S, Kobayashi R, Grewal SI. 2005. Ubiquitin ligase component Cul4 associates with Clr4 histone methyltransferase to assemble heterochromatin. *Nature cell biology* **7**: 1007-1013.

Jia S, Noma K, Grewal SI. 2004. RNAi-independent heterochromatin nucleation by the stress-activated ATF/CREB family proteins. *Science* **304**: 1971-1976.

Jih G, Iglesias N, Currie MA, Bhanu NV, Paulo JA, Gygi SP, Garcia BA, Moazed D. 2017. Unique roles for histone H3K9me states in RNAi and heritable silencing of transcription. *Nature* **547**: 463-467.

Jones PA, Liang G. 2009. Rethinking how DNA methylation patterns are maintained. *Nature Reviews Genetics* **10**: 805-811.

Jones SK, Jr., Spivey EC, Rybarski JR, Finkelstein IJ. 2018. A Microfluidic Device for Massively Parallel, Whole-lifespan Imaging of Single Fission Yeast Cells. *Bio-protocol* **8**.

Kagansky A, Folco HD, Almeida R, Pidoux AL, Boukaba A, Simmer F, Urano T, Hamilton GL, Allshire RC. 2009. Synthetic heterochromatin bypasses RNAi and centromeric repeats to establish functional centromeres. *Science* **324**: 1716-1719.

Kakui Y, Uhlmann F. 2018. SMC complexes orchestrate the mitotic chromatin interaction landscape. *Current genetics* **64**: 335-339.

Kanoh J, Sadaie M, Urano T, Ishikawa F. 2005. Telomere binding protein Taz1 establishes Swi6 heterochromatin independently of RNAi at telomeres. *Current biology : CB* **15**: 1808-1819.

Kassambara A. 2020. ggpubr: 'ggplot2' Based Publication Ready Plots. https://CRAN.R-project.org/package=ggpubr.

Keller C, Woolcock K, Hess D, Bühler M. 2010. Proteomic and functional analysis of the noncanonical poly(A) polymerase Cid14. *RNA (New York, NY)* **16**: 1124-1129.

Kelly M, Burke J, Smith M, Klar A, Beach D. 1988. Four mating-type genes control sexual differentiation in the fission yeast. *Embo j* **7**: 1537-1547.

Kim HS, Choi ES, Shin JA, Jang YK, Park SD. 2004. Regulation of Swi6/HP1-dependent heterochromatin assembly by cooperation of components of the mitogen-activated protein kinase pathway and a histone deacetylase Clr6. *The Journal of biological chemistry* **279**: 42850-42859.

Kirmizis A, Santos-Rosa H, Penkett CJ, Singer MA, Vermeulen M, Mann M, Bahler J, Green RD, Kouzarides T. 2007. Arginine methylation at histone H3R2 controls deposition of H3K4 trimethylation. *Nature* **449**: 928-932.

Klemm SL, Shipony Z, Greenleaf WJ. 2019. Chromatin accessibility and the regulatory epigenome. *Nature Reviews Genetics* **20**: 207-220.

Knott SR, Peace JM, Ostrow AZ, Gan Y, Rex AE, Viggiani CJ, Tavaré S, Aparicio OM. 2012.
Forkhead transcription factors establish origin timing and long-range clustering in S.
cerevisiae. *Cell* **148**: 99-111.

Knutsen JH, Rein ID, Rothe C, Stokke T, Grallert B, Boye E. 2011. Cell-cycle analysis of fission
yeast cells by flow cytometry. *PLoS One* **6**: e17175.

Kouzarides T. 2007. Chromatin modifications and their function. *Cell* **128**: 693-705.

Kowalik KM, Shimada Y, Flury V, Stadler MB, Batki J, Buhler M. 2015. The Paf1 complex
represses small-RNA-mediated epigenetic gene silencing. *Nature* **520**: 248-252.

Krijthe J. 2015. Rtsne: T-Distributed Stochastic Neighbor Embedding using a Barnes-Hut
Implementation. https://github.com/jkrijthe/Rtsne.

Kurukuti S, Tiwari VK, Tavoosidana G, Pugacheva E, Murrell A, Zhao Z, Lobanenkov V, Reik
W, Ohlsson R. 2006. CTCF binding at the H19 imprinting control region mediates
maternally inherited higher-order chromatin conformation to restrict enhancer access to
Igf2. *Proc Natl Acad Sci U S A* **103**: 10684-10689.

Kusevic D, Kudithipudi S, Iglesias N, Moazed D, Jeltsch A. 2017. Clr4 specificity and catalytic
activity beyond H3K9 methylation. *Biochimie* **135**: 83-88.

Lachner M, O'Carroll D, Rea S, Mechtler K, Jenuwein T. 2001. Methylation of histone H3 lysine
9 creates a binding site for HP1 proteins. *Nature* **410**: 116-120.

Lan F, Zaratiegui M, Villen J, Vaughn MW, Verdel A, Huarte M, Shi Y, Gygi SP, Moazed D,
Martienssen RA et al. 2007. S. pombe LSD1 homologs regulate heterochromatin
propagation and euchromatic gene transcription. *Molecular cell* **26**: 89-101.

Langmead B, Salzberg SL. 2012. Fast gapped-read alignment with Bowtie 2. *Nat Methods* **9**:
357-359.

Lantermann AB, Straub T, Stralfors A, Yuan GC, Ekwall K, Korber P. 2010.
Schizosaccharomyces pombe genome-wide nucleosome mapping reveals positioning
mechanisms distinct from those of Saccharomyces cerevisiae. *Nat Struct Mol Biol* **17**:
251-U215.

Laprell F, Finkl K, Muller J. 2017. Propagation of Polycomb-repressed chromatin requires
sequence-specific recruitment to DNA. *Science* **356**: 85-88.

Law JA, Jacobsen SE. 2010. Establishing, maintaining and modifying DNA methylation patterns
in plants and animals. *Nature Reviews Genetics* **11**: 204-220.

Lawrence M, Gentleman R, Carey V. 2009. rtracklayer: an R package for interfacing with
genome browsers. *Bioinformatics* **25**: 1841-1842.

Lawrence M, Huber W, Pages H, Aboyoun P, Carlson M, Gentleman R, Morgan MT, Carey VJ.
2013. Software for computing and annotating genomic ranges. *PLoS computational
biology* **9**: e1003118.

Lee J, Choi ES, Seo HD, Kang K, Gilmore JM, Florens L, Washburn MP, Choe J, Workman JL,
Lee D. 2017. Chromatin remodeller Fun30(Fft3) induces nucleosome disassembly to
facilitate RNA polymerase II elongation. *Nature communications* **8**: 14527.

Lee NN, Chalamcharla VR, Reyes-Turcu F, Mehta S, Zofall M, Balachandran V,
Dhakshnamoorthy J, Taneja N, Yamanaka S, Zhou M et al. 2013. Mtr4-like protein
coordinates nuclear RNA processing for heterochromatin assembly and for telomere
maintenance. *Cell* **155**: 1061-1074.

Lejeune E, Bortfeld M, White SA, Pidoux AL, Ekwall K, Allshire RC, Ladurner AG. 2007. The
chromatin-remodeling factor FACT contributes to centromeric heterochromatin
independently of RNAi. *Current biology : CB* **17**: 1219-1224.

Li E. 2002. Chromatin modification and epigenetic reprogramming in mammalian development. *Nature Reviews Genetics* **3**: 662-673.

Li F, Goto DB, Zaratiegui M, Tang X, Martienssen R, Cande WZ. 2005. Two novel proteins, dos1 and dos2, interact with rik1 to regulate heterochromatic RNA interference and histone modification. *Current biology : CB* **15**: 1448-1457.

Li F, Huarte M, Zaratiegui M, Vaughn MW, Shi Y, Martienssen R, Cande WZ. 2008. Lid2 is required for coordinating H3K4 and H3K9 methylation of heterochromatin and euchromatin. *Cell* **135**: 272-283.

Li H, Ilin S, Wang W, Duncan EM, Wysocka J, Allis CD, Patel DJ. 2006. Molecular basis for site-specific read-out of histone H3K4me3 by the BPTF PHD finger of NURF. *Nature* **442**: 91-95.

Li H, Motamedi MR, Yip CK, Wang Z, Walz T, Patel DJ, Moazed D. 2009. An alpha motif at Tas3 C terminus mediates RITS cis spreading and promotes heterochromatic gene silencing. *Molecular cell* **34**: 155-167.

Li Y, Jin M, O'Laughlin R, Bittihn P, Tsimring LS, Pillus L, Hasty J, Hao N. 2017. Multigenerational silencing dynamics control cell aging. *Proc Natl Acad Sci U S A* **114**: 11253-11258.

Lieberman-Aiden E, van Berkum NL, Williams L, Imakaev M, Ragoczy T, Telling A, Amit I, Lajoie BR, Sabo PJ, Dorschner MO et al. 2009. Comprehensive Mapping of Long-Range Interactions Reveals Folding Principles of the Human Genome. *Science* **326**: 289.

Litt MD, Simpson M, Gaszner M, Allis CD, Felsenfeld G. 2001. Correlation between histone lysine methylation and developmental changes at the chicken beta-globin locus. *Science* **293**: 2453-2455.

Lock A, Rutherford K, Harris MA, Hayles J, Oliver SG, Bahler J, Wood V. 2019. PomBase

    2018: user-driven reimplementation of the fission yeast database provides rapid and

    intuitive access to diverse, interconnected information. *Nucleic Acids Res* **47**: D821-d827.

Lorch Y, Kornberg RD. 2017. Chromatin-remodeling for transcription. *Quarterly Reviews of*

    Biophysics **50**: e5.

Lorenz DR, Meyer LF, Grady PJ, Meyer MM, Cam HP. 2014. Heterochromatin assembly and

    transcriptome repression by Set1 in coordination with a class II histone deacetylase. *Elife*

    3: e04506.

Luger K, Mäder AW, Richmond RK, Sargent DF, Richmond TJ. 1997. Crystal structure of the

    nucleosome core particle at 2.8 A resolution. *Nature* **389**: 251-260.

Lun AT, Smyth GK. 2016. csaw: a Bioconductor package for differential binding analysis of

    ChIP-seq data using sliding windows. *Nucleic Acids Res* **44**: e45.

Margueron R, Justin N, Ohno K, Sharpe ML, Son J, Drury WJ, 3rd, Voigt P, Martin SR, Taylor

    WR, De Marco V et al. 2009. Role of the polycomb protein EED in the propagation of

    repressive histone marks. *Nature* **461**: 762-767.

Margueron R, Trojer P, Reinberg D. 2005. The key to development: interpreting the histone

    code? *Curr Opin Genet Dev* **15**: 163-176.

Marina DB, Shankar S, Natarajan P, Finn KJ, Madhani HD. 2013. A conserved ncRNA-binding

    protein recruits silencing factors to heterochromatin through an RNAi-independent

    mechanism. *Genes Dev* **27**: 1851-1856.

Mathieu O, Reinders J, Caikovski M, Smathajitt C, Paszkowski J. 2007. Transgenerational

    stability of the Arabidopsis epigenome is coordinated by CG methylation. *Cell* **130**: 851-

    862.

McDonald OG, Wu H, Timp W, Doi A, Feinberg AP. 2011. Genome-scale epigenetic reprogramming during epithelial-to-mesenchymal transition. *Nat Struct Mol Biol* **18**: 867-874.

Meneghini MD, Wu M, Madhani HD. 2003. Conserved histone variant H2A.Z protects euchromatin from the ectopic spread of silent heterochromatin. *Cell* **112**: 725-736.

Mikheyeva IV, Grady PJ, Tamburini FB, Lorenz DR, Cam HP. 2014. Multifaceted genome control by Set1 Dependent and Independent of H3K4 methylation and the Set1C/COMPASS complex. *PLoS Genet* **10**: e1004740.

Miller T, Krogan NJ, Dover J, Erdjument-Bromage H, Tempst P, Johnston M, Greenblatt JF, Shilatifard A. 2001. COMPASS: a complex of proteins associated with a trithorax-related SET domain protein. *Proc Natl Acad Sci U S A* **98**: 12902-12907.

Moazed D. 2009. Small RNAs in transcriptional gene silencing and genome defence. *Nature* **457**: 413-420.

Muller HJ. 1930. Types of visible variations induced by X-rays inDrosophila. *Journal of Genetics* **22**: 299-334.

Muller MM, Fierz B, Bittova L, Liszczak G, Muir TW. 2016. A two-state activation mechanism controls the histone methyltransferase Suv39h1. *Nature chemical biology* **12**: 188-193.

Nakaoka H, Wakamoto Y. 2017. Aging, mortality, and the fast growth trade-off of Schizosaccharomyces pombe. *PLoS biology* **15**: e2001109.

Nakayama J, Klar AJ, Grewal SI. 2000. A chromodomain protein, Swi6, performs imprinting functions in fission yeast during mitosis and meiosis. *Cell* **101**: 307-317.

Nakayama J, Rice JC, Strahl BD, Allis CD, Grewal SI. 2001. Role of histone H3 lysine 9 methylation in epigenetic control of heterochromatin assembly. *Science* **292**: 110-113.

Narlikar GJ, Phelan ML, Kingston RE. 2001. Generation and Interconversion of Multiple
Distinct Nucleosomal States as a Mechanism for Catalyzing Chromatin Fluidity.
Molecular cell **8**: 1219-1230.

Nicetto D, Zaret KS. 2019. Role of H3K9me3 heterochromatin in cell identity establishment and
maintenance. *Curr Opin Genet Dev* **55**: 1-10.

Nicolas E, Yamada T, Cam HP, Fitzgerald PC, Kobayashi R, Grewal SI. 2007. Distinct roles of
HDAC complexes in promoter silencing, antisense suppression and DNA damage
protection. *Nat Struct Mol Biol* **14**: 372-380.

Nielsen AL, Oulad-Abdelghani M, Ortiz JA, Remboutsika E, Chambon P, Losson R. 2001.
Heterochromatin formation in mammalian cells: interaction between histones and HP1
proteins. *Molecular cell* **7**: 729-739.

Nimmo ER, Cranston G, Allshire RC. 1994. Telomere-associated chromosome breakage in
fission yeast results in variegated expression of adjacent genes. *Embo j* **13**: 3801-3811.

Nimmo ER, Pidoux AL, Perry PE, Allshire RC. 1998. Defective meiosis in telomere-silencing
mutants of Schizosaccharomyces pombe. *Nature* **392**: 825-828.

Nishioka K, Chuikov S, Sarma K, Erdjument-Bromage H, Allis CD, Tempst P, Reinberg D.
2002. Set9, a novel histone H3 methyltransferase that facilitates transcription by
precluding histone tail modifications required for heterochromatin formation. *Genes Dev*
16: 479-489.

Noma K, Allis CD, Grewal SI. 2001. Transitions in distinct histone H3 methylation patterns at
the heterochromatin domain boundaries. *Science* **293**: 1150-1155.

Noma K, Cam HP, Maraia RJ, Grewal SI. 2006. A role for TFIIIC transcription factor complex
in genome organization. *Cell* **125**: 859-872.

Noma K, Grewal SIS. 2002. Histone H3 lysine 4 methylation is mediated by Set1 and promotes maintenance of active chromatin states in fission yeast. *Proceedings of the National Academy of Sciences of the United States of America* **99**: 16438-16445.

Noma K, Sugiyama T, Cam H, Verdel A, Zofall M, Jia S, Moazed D, Grewal SI. 2004. RITS acts in cis to promote RNA interference-mediated transcriptional and post-transcriptional silencing. *Nature genetics* **36**: 1174-1180.

Obersriebnig MJ, Pallesen EM, Sneppen K, Trusina A, Thon G. 2016. Nucleation and spreading of a heterochromatic domain in fission yeast. *Nature communications* **7**: 11518.

Oki M, Kamakaka RT. 2005. Barrier function at HMR. *Molecular cell* **19**: 707-716.

Orkin SH, Hochedlinger K. 2011. Chromatin connections to pluripotency and cellular reprogramming. *Cell* **145**: 835-850.

Orlova DY, Zimmerman N, Meehan S, Meehan C, Waters J, Ghosn EE, Filatenkov A, Kolyagin GA, Gernez Y, Tsuda S et al. 2016. Earth Mover's Distance (EMD): A True Metric for Comparing Biomarker Expression Levels in Cell Populations. *PLoS One* **11**: e0151859.

Osborne EA, Dudoit S, Rine J. 2009. The establishment of gene silencing at single-cell resolution. *Nature genetics* **41**: 800-806.

Parsa JY, Boudoukha S, Burke J, Homer C, Madhani HD. 2018. Polymerase pausing induced by sequence-specific RNA-binding protein drives heterochromatin assembly. *Genes Dev* **32**: 953-964.

Pedelacq JD, Cabantous S, Tran T, Terwilliger TC, Waldo GS. 2006. Engineering and characterization of a superfolder green fluorescent protein. *Nature biotechnology* **24**: 79-88.

Peters AH, O'Carroll D, Scherthan H, Mechtler K, Sauer S, Schöfer C, Weipoltshammer K, Pagani M, Lachner M, Kohlmaier A et al. 2001. Loss of the Suv39h histone methyltransferases impairs mammalian heterochromatin and genome stability. *Cell* **107**: 323-337.

Petruk S, Sedkov Y, Johnston DM, Hodgson JW, Black KL, Kovermann SK, Beck S, Canaani E, Brock HW, Mazo A. 2012. TrxG and PcG proteins but not methylated histones remain associated with DNA through replication. *Cell* **150**: 922-933.

Petryk N, Dalby M, Wenger A, Stromme CB, Strandsby A, Andersson R, Groth A. 2018. MCM2 promotes symmetric inheritance of modified histones during DNA replication. *Science* **361**: 1389-1392.

Pokholok DK, Harbison CT, Levine S, Cole M, Hannett NM, Lee TI, Bell GW, Walker K, Rolfe PA, Herbolsheimer E et al. 2005. Genome-wide map of nucleosome acetylation and methylation in yeast. *Cell* **122**: 517-527.

Ragunathan K, Jih G, Moazed D. 2015. Epigenetics. Epigenetic inheritance uncoupled from sequence-specific recruitment. *Science* **348**: 1258699.

Ramirez F, Ryan DP, Gruning B, Bhardwaj V, Kilpert F, Richter AS, Heyne S, Dundar F, Manke T. 2016. deepTools2: a next generation web server for deep-sequencing data analysis. *Nucleic Acids Res* **44**: W160-165.

Rawal Y, Chereji RV, Qiu H, Ananthakrishnan S, Govind CK, Clark DJ, Hinnebusch AG. 2018. SWI/SNF and RSC cooperate to reposition and evict promoter nucleosomes at highly expressed genes in yeast. *Genes Dev* **32**: 695-710.

Rea S, Eisenhaber F, O'Carroll D, Strahl BD, Sun Z-W, Schmid M, Opravil S, Mechtler K, Ponting CP, Allis CD et al. 2000. Regulation of chromatin structure by site-specific histone H3 methyltransferases. *Nature* **406**: 593.

Reinberg D, Vales LD. 2018. Chromatin domains rich in inheritance. *Science* **361**: 33.

Reinke H, Horz W. 2003. Histones are first hyperacetylated and then lose contact with the activated PHO5 promoter. *Molecular cell* **11**: 1599-1607.

Renauld H, Aparicio OM, Zierath PD, Billington BL, Chhablani SK, Gottschling DE. 1993. Silent domains are assembled continuously from the telomere and are defined by promoter distance and strength, and by SIR3 dosage. *Genes Dev* **7**: 1133-1145.

Reverón-Gómez N, González-Aguilera C, Stewart-Morgan KR, Petryk N, Flury V, Graziano S, Johansen JV, Jakobsen JS, Alabert C, Groth A. 2018. Accurate Recycling of Parental Histones Reproduces the Histone Modification Landscape during DNA Replication. *Molecular cell* **72**: 239-249.e235.

Reyes-Turcu FE, Zhang K, Zofall M, Chen E, Grewal SI. 2011. Defects in RNA quality control factors reveal RNAi-independent nucleation of heterochromatin. *Nat Struct Mol Biol* **18**: 1132-1138.

Ringrose L, Paro R. 2007. Polycomb/Trithorax response elements and epigenetic memory of cell identity. *Development (Cambridge, England)* **134**: 223-232.

Roguev A, Schaft D, Shevchenko A, Aasland R, Shevchenko A, Stewart AF. 2003. High conservation of the Set1/Rad6 axis of histone 3 lysine 4 methylation in budding and fission yeasts. *The Journal of biological chemistry* **278**: 8487-8493.

Rowe CE, Narlikar GJ. 2010. The ATP-dependent remodeler RSC transfers histone dimers and

    octamers through the rapid formation of an unstable encounter intermediate.

    Biochemistry **49**: 9882-9890.

Rubner Y, Tomasi C, Guibas L. 1998. A metric for distributions with applications to image

    databases. . in *IEEE International Conference on Computer Vision*, pp. pp. 59-66.

Rusche LN, Kirchmaier AL, Rine J. 2003. The establishment, inheritance, and function of

    silenced chromatin in Saccharomyces cerevisiae. *Annu Rev Biochem* **72**: 481-516.

Sadeghi L, Prasad P, Ekwall K, Cohen A, Svensson JP. 2015. The Paf1 complex factors Leo1

    and Paf1 promote local histone turnover to modulate chromatin states in fission yeast.

    EMBO Rep **16**: 1673-1687.

Sakaue-Sawano A, Kurokawa H, Morimura T, Hanyu A, Hama H, Osawa H, Kashiwagi S,

    Fukami K, Miyata T, Miyoshi H et al. 2008. Visualizing spatiotemporal dynamics of

    multicellular cell-cycle progression. *Cell* **132**: 487-498.

Saksouk N, Simboeck E, Déjardin J. 2015. Constitutive heterochromatin formation and

    transcription in mammals. *Epigenetics & Chromatin* **8**: 3.

Santos-Rosa H, Schneider R, Bannister AJ, Sherriff J, Bernstein BE, Emre NC, Schreiber SL,

    Mellor J, Kouzarides T. 2002. Active genes are tri-methylated at K4 of histone H3.

    Nature **419**: 407-411.

Sarraf SA, Stancheva I. 2004. Methyl-CpG binding protein MBD1 couples histone H3

    methylation at lysine 9 by SETDB1 to DNA replication and chromatin assembly.

    Molecular cell **15**: 595-605.

Scaffidi P, Misteli T. 2006. Lamin A-Dependent Nuclear Defects in Human Aging. *Science* **312**:

    1059.

Schindelin J, Arganda-Carreras I, Frise E, Kaynig V, Longair M, Pietzsch T, Preibisch S,
Rueden C, Saalfeld S, Schmid B et al. 2012. Fiji: an open-source platform for biological-
image analysis. *Nat Methods* **9**: 676-682.

Schlichter A, Cairns BR. 2005. Histone trimethylation by Set1 is coordinated by the RRM,
autoinhibitory, and catalytic domains. *Embo j* **24**: 1222-1231.

Schmitz RJ, Amasino RM. 2007. Vernalization: a model for investigating epigenetics and
eukaryotic gene regulation in plants. *Biochimica et biophysica acta* **1769**: 269-275.

Schultz DC, Ayyanathan K, Negorev D, Maul GG, Rauscher FJ, 3rd. 2002. SETDB1: a novel
KAP-1-associated histone H3, lysine 9-specific methyltransferase that contributes to
HP1-mediated silencing of euchromatic genes by KRAB zinc-finger proteins. *Genes Dev*
16: 919-932.

Schultz J. 1939. The function of heterochromatin. in *Proc Int Congr Genet* Cambridge
University Press, Edinburgh.

Schwartz YB, Kahn TG, Nix DA, Li XY, Bourgon R, Biggin M, Pirrotta V. 2006. Genome-wide
analysis of Polycomb targets in Drosophila melanogaster. *Nature genetics* **38**: 700-705.

Scott KC, Merrett SL, Willard HF. 2006. A heterochromatin barrier partitions the fission yeast
centromere into discrete chromatin domains. *Current biology : CB* **16**: 119-129.

Serra-Cardona A, Zhang Z. 2018. Replication-Coupled Nucleosome Assembly in the Passage of
Epigenetic Information and Cell Identity. *Trends in biochemical sciences* **43**: 136-148.

Shankaranarayana GD, Motamedi MR, Moazed D, Grewal SIS. 2003. Sir2 regulates histone H3
lysine 9 methylation and heterochromatin assembly in fission yeast. *Current Biology* **13**:
1240-1246.

Shevchenko A, Roguev A, Schaft D, Buchanan L, Habermann B, Sakalar C, Thomas H, Krogan NJ, Shevchenko A, Stewart AF. 2008. Chromatin Central: towards the comparative proteome by accurate mapping of the yeast proteomic environment. *Genome biology* **9**: R167.

Shilatifard A. 2012. The COMPASS family of histone H3K4 methylases: mechanisms of regulation in development and disease pathogenesis. *Annu Rev Biochem* **81**: 65-95.

Spitz F, Furlong EEM. 2012. Transcription factors: from enhancer binding to developmental control. *Nature Reviews Genetics* **13**: 613-626.

Spivey EC, Jones SK, Jr., Rybarski JR, Saifuddin FA, Finkelstein IJ. 2017. An aging-independent replicative lifespan in a symmetrically dividing eukaryote. *Elife* **6**.

Spivey EC, Xhemalce B, Shear JB, Finkelstein IJ. 2014. 3D-printed microfluidic microdissector for high-throughput studies of cellular aging. *Analytical chemistry* **86**: 7406-7412.

Stovner EB, Saetrom P. 2019. epic2 efficiently finds diffuse domains in ChIP-seq data. *Bioinformatics*.

Strack RL, Hein B, Bhattacharyya D, Hell SW, Keenan RJ, Glick BS. 2009. A rapidly maturing far-red derivative of DsRed-Express2 for whole-cell labeling. *Biochemistry* **48**: 8279-8281.

Sugiyama T, Cam HP, Sugiyama R, Noma K-i, Zofall M, Kobayashi R, Grewal SIS. 2007. SHREC, an effector complex for heterochromatic transcriptional silencing. *Cell* **128**: 491-504.

Sugiyama T, Thillainadesan G, Chalamcharla VR, Meng Z, Balachandran V, Dhakshnamoorthy J, Zhou M, Grewal SIS. 2016. Enhancer of Rudimentary Cooperates with Conserved

RNA-Processing Factors to Promote Meiotic mRNA Decay and Facultative

Heterochromatin Assembly. *Molecular cell* **61**: 747-759.

Svensson JP, Shukla M, Menendez-Benito V, Norman-Axelsson U, Audergon P, Sinha I, Tanny

JC, Allshire RC, Ekwall K. 2015. A nucleosome turnover map reveals that the stability of

histone H4 Lys20 methylation depends on histone recycling in transcribed chromatin.

Genome research **25**: 872-883.

Talbert PB, Henikoff S. 2006. Spreading of silent chromatin: inaction at a distance. *Nature*

reviews Genetics **7**: 793-803.

Taneja N, Zofall M, Balachandran V, Thillainadesan G, Sugiyama T, Wheeler D, Zhou M,

Grewal SI. 2017. SNF2 Family Protein Fft3 Suppresses Nucleosome Turnover to

Promote Epigenetic Inheritance and Proper Replication. *Molecular cell* **66**: 50-62.e56.

Taverna SD, Ilin S, Rogers RS, Tanny JC, Lavender H, Li H, Baker L, Boyle J, Blair LP, Chait

BT et al. 2006. Yng1 PHD Finger Binding to H3 Trimethylated at K4 Promotes NuA3

HAT Activity at K14 of H3 and Transcription at a Subset of Targeted ORFs. *Molecular*

cell **24**: 785-796.

Tchasovnikarova IA, Timms RT, Matheson NJ, Wals K, Antrobus R, Göttgens B, Dougan G,

Dawson MA, Lehner PJ. 2015. GENE SILENCING. Epigenetic silencing by the HUSH

complex mediates position-effect variegation in human cells. *Science (New York, NY)*

348: 1481-1485.

Thon G, Bjerling KP, Nielsen IS. 1999. Localization and properties of a silencing element near

the mat3-M mating-type cassette of Schizosaccharomyces pombe. *Genetics* **151**: 945-

963.

Thon G, Bjerling P, Bunner CM, Verhein-Hansen J. 2002. Expression-state boundaries in the mating-type region of fission yeast. *Genetics* **161**: 611-622.

Thon G, Friis T. 1997. Epigenetic inheritance of transcriptional silencing and switching competence in fission yeast. *Genetics* **145**: 685-696.

Thon G, Hansen KR, Altes SP, Sidhu D, Singh G, Verhein-Hansen J, Bonaduce MJ, Klar AJ. 2005. The Clr7 and Clr8 directionality factors and the Pcu4 cullin mediate heterochromatin formation in the fission yeast Schizosaccharomyces pombe. *Genetics* **171**: 1583-1595.

Tong K, Keller T, Hoffman CS, Annunziato AT. 2012. Schizosaccharomyces pombe Hat1 (Kat1) is associated with Mis16 and is required for telomeric silencing. *Eukaryot Cell* **11**: 1095-1103.

Trewick SC, Minc E, Antonelli R, Urano T, Allshire RC. 2007. The JmjC domain protein Epe1 prevents unregulated assembly and disassembly of heterochromatin. *Embo j* **26**: 4670-4682.

Urbanek S, Rubner Y. 2012. emdist: earth mover's distance.

van Steensel B, Belmont AS. 2017. Lamina-Associated Domains: Links with Chromosome Architecture, Heterochromatin, and Gene Repression. *Cell* **169**: 780-791.

Venkatasubrahmanyam S, Hwang WW, Meneghini MD, Tong AH, Madhani HD. 2007. Genome-wide, as opposed to local, antisilencing is mediated redundantly by the euchromatic factors Set1 and H2A.Z. *Proc Natl Acad Sci U S A* **104**: 16609-16614.

Verdel A, Jia S, Gerber S, Sugiyama T, Gygi S, Grewal SI, Moazed D. 2004. RNAi-mediated targeting of heterochromatin by the RITS complex. *Science* **303**: 672-676.

Verrier L, Taglini F, Barrales RR, Webb S, Urano T, Braun S, Bayne EH. 2015. Global regulation of heterochromatin spreading by Leo1. *Open biology* **5**.

Verzijlbergen KF, Menendez-Benito V, van Welsem T, van Deventer SJ, Lindstrom DL, Ovaa H, Neefjes J, Gottschling DE, van Leeuwen F. 2010. Recombination-induced tag exchange to track old and new proteins. *Proc Natl Acad Sci U S A* **107**: 64-68.

Villeponteau B. 1997. The heterochromatin loss model of aging. *Experimental Gerontology* **32**: 383-394.

Volpe TA, Kidner C, Hall IM, Teng G, Grewal SI, Martienssen RA. 2002. Regulation of heterochromatic silencing and histone H3 lysine-9 methylation by RNAi. *Science* **297**: 1833-1837.

Wahls WP, Smith GR. 1994. A heteromeric protein that binds to a meiotic homologous recombination hot spot: correlation of binding and hot spot activity. *Genes Dev* **8**: 1693-1702.

Wang HB, Zhang Y. 2001. Mi2, an auto-antigen for dermatomyositis, is an ATP-dependent nucleosome remodeling factor. *Nucleic Acids Res* **29**: 2517-2521.

Wang J, Cohen AL, Letian A, Tadeo X, Moresco JJ, Liu J, Yates JR, 3rd, Qiao F, Jia S. 2016. The proper connection between shelterin components is required for telomeric heterochromatin assembly. *Genes Dev* **30**: 827-839.

Wang J, Lawry ST, Cohen AL, Jia S. 2014. Chromosome boundary elements and regulation of heterochromatin spreading. *Cell Mol Life Sci* **71**: 4841-4852.

Wang J, Reddy BD, Jia S. 2015. Rapid epigenetic adaptation to uncontrolled heterochromatin spreading. *Elife* **4**.

Wang J, Tadeo X, Hou H, Tu PG, Thompson J, Yates JR, 3rd, Jia S. 2013. Epe1 recruits BET

family bromodomain protein Bdf2 to establish heterochromatin boundaries. *Genes Dev*

27: 1886-1902.

Wang X, Moazed D. 2017. DNA sequence-dependent epigenetic inheritance of gene silencing

and histone H3K9 methylation. *Science* **356**: 88-91.

Wen B, Wu H, Shinkai Y, Irizarry RA, Feinberg AP. 2009. Large histone H3 lysine 9

dimethylated chromatin blocks distinguish differentiated from embryonic stem cells.

Nature genetics **41**: 246-250.

Wickham H. 2016. *ggplot2: Elegant Graphics for Data Analysis*. Springer

Wickham H, François R, Henry L, Müller K. 2020. dplyr: A Grammar of Data Manipulation. R

package version 0.8.4. .

Wirén M, Silverstein RA, Sinha I, Walfridsson J, Lee H-M, Laurenson P, Pillus L, Robyr D,

Grunstein M, Ekwall K. 2005. Genomewide analysis of nucleosome density histone

acetylation and HDAC function in fission yeast. *EMBO J* **24**: 2906-2918.

Woo H, Dam Ha S, Lee SB, Buratowski S, Kim T. 2017. Modulation of gene expression

dynamics by co-transcriptional histone methylations. *Exp Mol Med* **49**: e326.

Wood V Gwilliam R Rajandream MA Lyne M Lyne R Stewart A Sgouros J Peat N Hayles J

Baker S et al. 2002. The genome sequence of Schizosaccharomyces pombe. *Nature* **415**:

871-880.

Woolcock KJ, Stunnenberg R, Gaidatzis D, Hotz HR, Emmerth S, Barraud P, Buhler M. 2012.

RNAi keeps Atf1-bound stress response genes in check at nuclear pores. *Genes Dev* **26**:

683-692.

Xhemalce B, Kouzarides T. 2010. A chromodomain switch mediated by histone H3 Lys 4 acetylation regulates heterochromatin assembly. *Genes Dev* **24**: 647-652.

Xu EY, Zawadzki KA, Broach JR. 2006. Single-cell observations reveal intermediate transcriptional silencing states. *Molecular cell* **23**: 219-229.

Xue Y, Van C, Pradhan SK, Su T, Gehrke J, Kuryan BG, Kitada T, Vashisht A, Tran N, Wohlschlegel J et al. 2015. The Ino80 complex prevents invasion of euchromatin into silent chromatin. *Genes Dev* **29**: 350-355.

Yamada T, Fischle W, Sugiyama T, Allis CD, Grewal SIS. 2005. The nucleation and maintenance of heterochromatin by a histone deacetylase in fission yeast. *Molecular cell* **20**: 173-185.

Yamanaka S, Mehta S, Reyes-Turcu FE, Zhuang F, Fuchs RT, Rong Y, Robb GB, Grewal SI. 2013. RNAi triggered by specialized machinery silences developmental genes and retrotransposons. *Nature* **493**: 557-560.

Yamane K, Mizuguchi T, Cui B, Zofall M, Noma K, Grewal SI. 2011. Asf1/HIRA facilitate global histone deacetylation and associate with HP1 to promote nucleosome occupancy at heterochromatic loci. *Molecular cell* **41**: 56-66.

Yu M, Ren B. 2017. The Three-Dimensional Organization of Mammalian Genomes. *Annual Review of Cell and Developmental Biology* **33**: 265-289.

Yu Y, Zhou H, Deng X, Wang W, Lu H. 2016. Set3 contributes to heterochromatin integrity by promoting transcription of subunits of Clr4-Rik1-Cul4 histone methyltransferase complex in fission yeast. *Scientific reports* **6**: 31752.

Zhang K, Mosch K, Fischle W, Grewal SIS. 2008a. Roles of the Clr4 methyltransferase complex in nucleation, spreading and maintenance of heterochromatin. *Nat Struct Mol Biol* **15**: 381-388.

Zhang W, Li J, Suzuki K, Qu J, Wang P, Zhou J, Liu X, Ren R, Xu X, Ocampo A et al. 2015. A Werner syndrome stem cell model unveils heterochromatin alterations as a driver of human aging. *Science* **348**: 1160.

Zhang Y, Liu T, Meyer CA, Eeckhoute J, Johnson DS, Bernstein BE, Nusbaum C, Myers RM, Brown M, Li W et al. 2008b. Model-based analysis of ChIP-Seq (MACS). *Genome biology* **9**: R137.

Zhao Y, Garcia BA. 2015. Comprehensive Catalog of Currently Documented Histone Modifications. *Cold Spring Harbor perspectives in biology* **7**: a025064-a025064.

Zheng H, Xie W. 2019. The role of 3D genome organization in development and cell differentiation. *Nature Reviews Molecular Cell Biology* **20**: 535-550.

Zhu J, Adli M, Zou JY, Verstappen G, Coyne M, Zhang X, Durham T, Miri M, Deshpande V, De Jager PL et al. 2013. Genome-wide chromatin state transitions associated with developmental and environmental cues. *Cell* **152**: 642-654.

Zilio N, Codlin S, Vashisht AA, Bitton DA, Head SR, Wohlschlegel JA, Bähler J, Boddy MN. 2014. A novel histone deacetylase complex in the control of transcription and genome stability. *Molecular and cellular biology* **34**: 3500-3514.

Zofall M, Grewal SI. 2006. Swi6/HP1 recruits a JmjC domain protein to facilitate transcription of heterochromatic repeats. *Molecular cell* **22**: 681-692.

Zofall M, Smith DR, Mizuguchi T, Dhakshnamoorthy J, Grewal SIS. 2016. Taz1-Shelterin

Promotes Facultative Heterochromatin Assembly at Chromosome-Internal Sites

Containing Late Replication Origins. *Molecular cell* **62**: 862-874.

Zofall M, Yamanaka S, Reyes-Turcu FE, Zhang K, Rubin C, Grewal SI. 2012. RNA elimination

machinery targeting meiotic mRNAs promotes facultative heterochromatin formation.

Science **335**: 96-100.

Zukowski A, Johnson AM. 2018. The interplay of histone H2B ubiquitination with budding and

fission yeast heterochromatin. *Current genetics* **64**: 799-806.

Zylicz JJ, Borensztein M, Wong FC, Huang Y, Lee C, Dietmann S, Surani MA. 2018. G9a

regulates temporal preimplantation developmental program and lineage segregation in

blastocyst. *Elife* **7**.

Zylicz JJ, Dietmann S, Gunesdogan U, Hackett JA, Cougot D, Lee C, Surani MA. 2015.

Chromatin dynamics and the role of G9a in gene regulation and enhancer silencing

during early mouse development. *Elife* **4**.

Publishing Agreement

It is the policy of the University to encourage open access and broad distribution of all theses, dissertations, and manuscripts. The Graduate Division will facilitate the distribution of UCSF theses, dissertations, and manuscripts to the UCSF Library for open access and distribution. UCSF will make such theses, dissertations, and manuscripts accessible to the public and will take reasonable steps to preserve these works in perpetuity.

I hereby grant the non-exclusive, perpetual right to The Regents of the University of California to reproduce, publicly display, distribute, preserve, and publish copies of my thesis, dissertation, or manuscript in any form or media, now existing or later derived, including access online for teaching, research, and public service purposes.

Author Signature

8/20/2020

Date