Epigenome control by chromatin modifiers: roles for histone H3 lysine modifiers in the regulation of repetitive elements

Author: Patrick James Robert Grady

Boston College

The Graduate School of Arts and Sciences

Department of Biology

EPIGENOME CONTROL BY CHROMATIN MODIFIERS: ROLES FOR HISTONE H3 LYSINE MODIFIERS IN THE REGULATION OF REPETITIVE ELEMENTS

a dissertation

by

PATRICK JAMES ROBERT GRADY

Submitted in partial fulfillment of the requirements

for the degree of

Doctor of Philosophy

December 2015

Epigenome control by chromatin modifiers: roles for histone H3 lysine modifiers in the regulation of repetitive elements. Patrick Grady. Advisor: Hugh P. Cam

Chromatin is the site of numerous structural features that contribute to the regulation of the genome. Although numerous posttranslational modifications to the histone proteins that make up chromatin have been identified, it remains unclear whether and to what extent these modifications might regulate transposons and other repetitive sequences. One such modification is methylation of histone H3 lysine 4 (H3K4me), which is catalyzed by Set1 and its associated complex Set1C/COMPASS. Although H3K4me is associated with actively transcribed regions in euchromatin, an emerging body of evidence suggests that Set1-mediated transcriptional control is often repressive. This thesis work describes expanded functions for Set1C/COMPASS as a regulatory module with roles throughout the genome. We identify novel locus-dependent repressive functions for Set1 at repetitive genomic regions. Interestingly, Set1 has multiple repressive modes that are dependent and independent of H3K4me. Additionally, we show that Set1 controls the nuclear organization of *Tf2* retrotransposons by antagonizing H3K4 acetylation. We describe how the roles of Set1 in the nuclear organization and transcriptional repression of *Tf2* cooperate to restrict *Tf2* transposition. Finally, we identify an H3K4-dependent role in countering the reduced dosage of histone H3 genes to help maintain genome stability and silencing of *Tf2*s and pericentromeric heterochromatin. Our study considerably expands the regulatory repertoire of an important histone modifier and highlights the multifaceted function by a highly conserved chromatin-modifying complex with diverse roles in genome control.

This thesis is dedicated to my parents, who taught me to value knowledge and gave me the strength to face many challenges along the way.

Acknowledgements

The work described in this thesis would have been impossible without the help of a generous community of mentors, colleagues, family, and friends. Everyone on this list has made considerable sacrifices of their time, skills, and resources to help bring this work to fruition.

I'd first like to acknowledge my colleagues who performed the experiments referenced in the supplemental figures of this dissertation and who painstakingly built the stories of which this work is a part.

David Lorenz provided extensive technical advice and recommendations especially regarding the analysis of qPCR data and the use of database resources, and contributed many helpful suggestions on the experiments described in Chapter two.

A huge thank you to my colleague and friend Irina Mikheyeva, who generated mutants of Set1 and Set1C subunits and performed the experiments characterizing expression of repetitive element in these strains. She also provided extensive laboratory and experimental support.

Fiona Tamburini performed western blot and immunofluorescence experiments to characterize Set1 stability and methyltransferase activity.

Heather Murtona and Simon Whitehall performed extensive experimental work on *Tf2* mobilization control, including generating strains and performing expression and mobilization analyses.

Bret Judson provided extensive microscopy support and provided many helpful suggestions regarding sample preparation, imaging, and data analysis.

Thank you to my fellow graduate students in the Cam lab, Lauren Meyer, Peter Johansen, and David Layman for providing moral support, technical advice, and experimental assistance.

Additional thanks to Daniel Shams, Fiona Tamburini, Arman Mohammad, and the other undergraduate students who helped to prepare media and keep the lab running, as well as providing helpful comments and discussions on much of the data presented here.

Thank you to the members of my thesis committee, MJ Gubbels, Michelle Meyer, and Anthony Annunziato, for your attentive guidance of this thesis work from its proposal through its completion.

Thank you to Laura Anne Lowery for joining the committee in evaluating this dissertation and its defense.

Additional thanks to Anthony Annunziato for providing the α-H4 antibody, for allowing me access to his personal library, and for wide-ranging and enjoyable conversations regarding chromatin biology and many other topics.

Thanks to Peter Marino for solving a vast array of problems so I could focus on my work.

A special thank you to Collette McLaughlin, who went above and beyond to solve a host of problems from the most mundane to the most time-sensitive and whose wise words gave me perspective and kept me in good spirits.

Mike Piatelli, Meghan Rice, and Andrew Farrell saw potential in me and could always find the right way to support and encourage. Mike, your mentorship and generosity supported me as a teacher, a graduate student, and so many other ways besides.

Thank you Meghan Rice for her steadfast friendship and love throughout these years.

Thank you to Lisanne D'Andrea-Winslow for teaching me to read scientific literature and for inspiring me to continue to expand my abilities.

A very special thanks to Tim Darr and Charlotte Davis, who took me into their family and made sure I was well-fed and healthy in body and mind.

I am grateful to my beautiful wife Randi Lynn. Without her perseverance, sense of adventure, and loving support this work would have not been possible.

Many thanks to my parents David and Amy Grady to whom I dedicate this thesis.

Finally I wish to thank my advisor Hugh Cam. His joy for science and the generosity of his mentorship have inspired me every day. No word or gesture could express the immensity of gratitude I feel for his guidance.

Table of Contents

List of Figures

Abbreviations used

cDNA	Reverse transcribed messenger RNA
CDS	Coding sequence
cen	Centromere
CENP-B	Centromeric protein B
cfu	Colony forming unit
ChIP	Chromatin Immunoprecipitation
cnt	Centromeric central core
COMPASS	Complex of proteins associated with Set1
CPT	Camptothecin
CTD	RNA polymerase II C-terminal domain
Cy3-dCTP	Fluoresecently tagged nucleotide used for labelling FISH probes
DAPI	4'-6-diamidino-2-phenylindole, a fluorescent DNA stain
dg	Centromeric otr repeats. From Japanese *dogentai* meaning kinetochore
dh	Centromeric otr repeats
DNA	Deoxyribonucleic acid
dsRNA	Double stranded RNA
EMM	Edinburgh minimal media
ER	Endoplasmic reticulum
FISH	Fluorescence *in situ* hybridization
FLAG	Protein epitope tag
FOA	Flouroorotic acid. Used for negative selection against ura marker
H2BK119ub	Ubiquitylated form of lysine at 119th position on histone H2B
H2Bub	Ubiquitylated form of lysine at 119th position on histone H2B
H3K4	Lysine at 4th position on histone H3 N-terminus
H3K4A	Alanine replacing lysine at 4th position on histone H3 N-terminus
H3K4ac	Acetylated form of lysine at 4th position on histone H3 N-terminus
H3K4me	Methylation of lysine at 4th position on histone H3 N-terminus
H3K4me1	Monomethylated form of lysine at 4th position on histone H3 N-terminus
H3K4me2	Dimethylated form of lysine at 4th position on histone H3 N-terminus
H3K4me3	Trimethylated form of lysine at 4th position on histone H3 N-terminus
H3K4R	Arginine replacing lysine at 4th position on histone H3 N-terminus
H3K56ac	Acetylated form of lysine at 56th position on histone H3 N-terminus
H3K9ac	Acetylated form of lysine at 9th position on histone H3 N-terminus
H3K9me	Methylated form of lysine at 9th position on histone H3 N-terminus
HAT	Histone acetyltransferase
HDAC	Histone deacetylase
hhf	*S. pombe* gene coding for histone H4
hht	*S. pombe* gene coding for histone H3
HIRA	Histone chaperone complex

HMT	Histone methyltransferase
HU	Hydroxyurea
LTR	Long terminal repeat
mat	Mating type locus
MMS	Methyl methanesulfonate
mRNA	Messenger RNA
mst1-L344S	Acetyltransferase-deficient temperature sensitive allele of mst1
natR	Resistanat to nourseothricin
NIB	Nuclear isolation buffer
ORF	Open reading frame
PCR	Polymerase chain reaction
PEV	Position effect variegation
PTM	Posttranslational modification
qPCR	Quantitative real-time PCR
RNA	Ribonucleic acid
RNAi	RNA interference
RRM	RNA-recognition motif
Set1C	Set1 complex. Synonymous with COMPASS
set1F$^{\text{-H3K4me}}$	Methyltransferase-deficient set1 allele with FLAG epitope at the C-terminus
siRNA	Short interfering RNA
sre1-N	Constitutively active form of hypoxia transcription factor Sre1
subtel	Subtelomere
TE	Transposable element
tf2-12::Nat-AI	Allele of *Tf2* LTR retrotransposon with artificial intron-containing nourseothricin resistance marker
tRNA	Transfer RNA
UTR	Untranslated region of mRNA
YEA	Yeast extract media with supplemental adenine

Chapter 1: Introduction

History of chromatin

The rapid expansion of molecular biology during the 20th century has revolutionized the understanding of the relationship between the nucleic acid and protein components of chromatin. Although Gregor Mendel and Charles Darwin established critical principles for the transgenerational inheritance of traits, a molecular basis for the transfer and control of genetic information remained elusive. Early work by Friedrich Miescher and Albrect Kossel characterized the biochemical features of DNA[1] and critical experiments in the first half of the 20th century provided evidence that genetic information was located in the DNA component of the chromosomes.[2-4 5] In 1928, Emil Heitz used light microscopy to describe tightly packed regions of chromatin he called heterochromatin, and contrasted these with lightly staining loosely packed regions he named euchromatin. Notably, heterochromatin remains condensed throughout the cell cycle, in contrast to euchromatin, which cycles through condensed and decondensed states.[6] With James Watson and Francis Crick's elucidation of the DNA structure in 1953,[7] the stage was set for biologists to address fundamental questions in gene regulation.

The significance of the chromatin protein component was expanded considerably when electron micrographs of the chromatin subunit produced by Ada Olins, Don Olins , and Christopher L. Woodcock showed that DNA is bound by protein complexes in a "beads on a string" arrangement **(Figure 1)**.[8-10] Roger Kornberg proposed a model in

which two copies of histone H2A, H2B, H3, and H4 associate to form a heterooctamer around which the negatively charged DNA strand is wrapped, a structure later named the nucleosome **(Figure 2)**.[11] The nucleosome crystal structure has since been solved to 2.8Å resolution,[12] revolutionizing the understanding of higher order DNA packaging, genome compaction, transcription, replication, gene expression, and DNA repair.[13]

Histone modifications and gene expression

Histones are much more than static binders of DNA. Histones are positively charged, allowing for association with the negatively charged DNA molecule. Each nucleosome wraps ~147bp of DNA, contacting the DNA strand at 14 conserved arginines which intercalate the DNA minor groove **(Figure 2)**.[12] Inter-nucleosomal contacts between the H4 N-terminal tail, the H2A acidic patch and the H2B C-terminus contribute to higher order compaction to produce structures including the 30nm fiber and metaphase chromosome.[14-16] The compaction role for histones is ancient, likely evolving before the Archea/Eukaryota divergence.[17] Numerous versions of the nucleosome paradigm exist *in vivo.* Variant histones with unique functions can replace several canonical histones.[18-20] The histones themselves are subject to an array of posttranslational modifications including methylation,[21-24] acetylation,[25] phosphorylation,[21] ADP-ribosylation,[26] ubiquitylation,[27,28] citrullination,[29] β-N-acetylglucosamination,[30] proline isomerization,[31] sumolation,[32,33] and others **(Figure 3)**.[34] An emerging body of research suggests that certain combinations of modifications are associated with specific cellular processes, notably the regulation of gene expression.

Modifications to the nucleosome can modulate interaction with DNA by two distinct mechanisms; by altering the charge-based nucleosome-DNA interaction as by lysine acetylation and by providing unique binding sites for chromatin proteins and transcription factors. Histone acetylation was first reported by Allfrey *et al.* in 1964.[25] Since then, histone acetylation has been shown to be a well-conserved and highly dynamic process, regulated by the opposing actions of histone acetyltransferases (HATs) and histone deacetylases (HDACs).[35] HATs use acetyl-CoA to acetylate the ε-amino group of histone lysines. By decreasing the overall positive charge of the lysine-rich histones, HAT activity can weaken charge-based DNA-nucleosome association, contributing to a more open chromatin state. Indeed, highly acetylated histones are associated with actively transcribed euchromatic regions, promoting access for polymerases and other DNA-binding factors. HATs can be divided into two classes based on sequence homology. Type B HATs are highly conserved and predominately acetylate free histones in the cytoplasm,[36] while type A HATs are implicated nuclear processes including transcription and histone deposition. Type A HATs are often capable of acetylating multiple histone lysines, both on the unstructured tails as well as within the histone core. Like other chromatin-modifying factors, Type A HATs are often found in multi-subunit complexes, which regulate their enzymatic activity and substrate specificity.

Opposing the role of the HATs are the histone deacetylases (HDACs). HDACs are generally transcriptional repressors, and restore the positive charge of acetylated

histone lysines by removal of the amino group, thereby strengthening the charge-based DNA-nucleosomal contact. HDAC activity is associated with densely packed heterochromatin. HATs and HDACs appear to have less stringent substrate specificity than histone methyltransferases, and often modify acetyl groups at several different histone lysines.[34] In this thesis we show a novel role in the regulation of *Tf2* nuclear organization for the KAT5 family HAT Mst1, a member of the NuA4 complex.[37,38]

Transitions between histone modification status are implicated in transcriptional control, with markers such as H3K4me1 and H2Bub contributing to the "primed" state of transcriptionally poised regions,[39] Regions which are transcriptionally inactive lose H3K9me and H3K27me3 modifications as they transition into a transcriptionally poised state, replacing them with H3K27ac and H3K4me1,[39] Full activation replaces H3K4me1 marks at the promoter with H3K4me3 and H3K4ac.[40]

The H3K4 methyltransferase complex Set1C/COMPASS

Studies of position effect variegation (PEV) suggested that transcriptional silencing was mediated by molecular factors that could spread from one locus to the surrounding nuclear neighborhood. Suppressor and enhancer screens for modulators of PEV in *Drosophila melanogaster* identified a host of genes encoding proteins that contain the SET domain, named for the proteins identified in the screen: (Su(var)3-9, Enhancer of *zeste*, and *trx*.[41] Subsequent studies demonstrate that these SET-containing proteins are histone melthyltransferases. [42-44] The substrates for SET domain-containing

enzymes are lysines on the unstructured N-terminal tails of the histones, which extend outward from the body of the nucleosome and can act as binding sites for chromatin factors.[12]

One well-conserved SET-domain containing histone lysine methyltransferase is Set1. Set1 and its associated complex COMPASS are catalytic for all three forms of H3K4me, a widespread mark that is generally associated with regions of active transcription **(Figure 4)**. Set1 is an important chromatin factor throughout the genome, with roles in protein modification, transcriptional regulation, and genome surveillance. Set1 was identified in budding yeast by two groups who showed that *set1*Δ results in loss of H3K4me and that the Set1 protein exists in a conserved complex named COMPASS **(Table 1)**.[42,45] These studies suggested that defects in X-chromosome dosage compensation, gene expression, and development when COMPASS homologs are disrupted in *C. elegans* and *D. melanogaster* arise from the effects of a conserved chromatin machinery associated with H3K4me.[44,45]

The COMPASS complex is comprised of eight subunits ranging in size from 25kDa to 130kDa **(Table 1)**.[42] Each subunit has a specific function in the assembly and distribution of H3K4me1, H3K4me2, and H3K4me3.[46] Several subunits in budding yeast are essential for stability or assembly of COMPASS.[46] Homologous complexes have been described in *S. pombe, Drosophila,* and *Homo sapiens* **(Table 1)**.

Ng *et al*. showed that budding yeast Set1 is recruited to the 5' end of actively transcribed genes by Pol II. Importantly, they showed differences in the genome-wide distribution of Set1-catalyzed lysine methylation, showing that H3K4me3 patterns are associated with Set1 occupancy. H3K4me3 persists after Set1 dissociation, suggesting that H3K4me may function as an epigenetic marker of transcriptionally active regions.[47] In fission yeast, Noma *et al.* showed that H3K4me is a stable modification that exists throughout the cell cycle. *set1*Δ in fission yeast results in loss of H3K4me and correlates with decreased H3 acetylation levels. The authors suggest that H3K4me is associated with transcriptionally poised regions.[48]

In an expansion of the discussion of epigenetic role for Set1, Morillon et al. showed induced transcription at MET16 and examined the changes as this gene became transcriptionally active. Their work suggests that the transition towards Pol II elongation is epigenetically controlled, with Spp1-mediated H3K4me3 and K36me2/3 present during a preiniation phase. H3K4me3 controls transient acetylation of H4K8 by the NuA4 complex, which then promotes the ATPase Isw1 to guide transition Pol II from initiation to elongation. As Pol II transitions towards elongation, H3K4me2 and H3K79me2 were detected.[49]

Dehe *et al*. were able to dissect the interrelated roles for the COMPASS subunits. Deletion of individual subunits can affect Set1 stability, the stability of the complex as a whole, and the distribution of H3K4 methylation both across active genes and throughout the genome. Key players for integrity of the complex are Set1, Swd1, and Swd3. Loss of

Swd1, Swd3, or Spp1 affect levels of Set1 protein. Deletion of the SET domain abolishes Bre2 and Sdc1 association with the complex. These subunits are important for H3K4 trimethylation. In contrast to Ng et al., Dehe et al. show that Set1 associates with both Pol II CTD phosphoserine 5 and phosphoserine 2 forms of Pol II. These results together suggest that COMPASS may have multiple forms and multiple roles at active genes.[50] Using an in vitro methyltransferase assay with reconstituted COMPASS, Kim et al. showed that the n-SET domain of Set1 is involved in crosstalk with H2Bub.[51] Soares et al. suggest that Set1 protein levels are regulated by degradation signals on Set1 in response to blockage of COMPASS recruitment, H3K4 methylation, or transcription.[52]

Screens for changes in H3K4me levels have served as a powerful tool for characterizing the factors upstream of H3K4me at coding sequences. Proteomic screens in *S. cerevisiae* identified factors required for H3K4 methylation by immunoblotting protein extracts from each strain in the *S. cerevisiae* non-essential mutant collection with H3K4me-specific antibodies.[53,54] This screen identified the E2-conjugating Rad6 enzyme, showing that Rad6-mediated monoubiquitination of histone H2B is upstream of H3K4me.[28,55] Notably, this was the first demonstration of trans-histone crosstalk. H2B ubiquitination is also upstream of Dot1-mediated trimethylation of H3K79.[56] Rad6 forms a complex with the E3 ligase Bre1 and Lge1, which are required for the recruitment of Rad6 to chromatin. Strains lacking Rad6, Bre1, or Lge1 are defective for H2B monoubiquitination, H3K4me, and H3K79me3.[57,58] H2Bub is generally distributed across the entire CDS of actively transcribed genes.[59] In budding yeast, association of Swd2 with COMPASS is dependent on H2Bub, and Δ*swd2* strains lack H3K4me3 and

H3K4me2, suggesting that the COMPASS subunit Swd2 is the likely mediator of H2Bub-H3K4me crosstalk.[60,61] The human and *Drosophila* Swd2 homolog Wdr82 is also required for H3K4me3.[62] Rad6 and Bre1 are highly conserved from yeast to humans as regulators of the global H2Bub pattern.[63,64] Additionally, H2Bub is implicated in regulating Pol II elongation properties through cooperation with the elongation factor FACT.[65] Indeed, *S. pombe* cells lacking H2Bub have altered Pol II distribution patterns and histone occupancy on coding sequences, accompanied by defects in nuclear structure, cell growth, and cytokinesis.[66]

The Pol II-associated Paf1 complex also functions in regulation of H3K4me at several levels. Paf1, in complex with Rtf1, Cdc73, Ctr9, and Leo1, associates with Pol II and regulates transcriptional initiation and elongation. In addition to its role as a regulator of Rad6/Bre1-mediated H2Bub, the Paf1 complex functions as a "platform" on elongating Pol II that allows for binding of chromatin factors including COMPASS and the H3K36 methyltransferase Set2.[24] H3K4me is lost in cells lacking Paf1 complex components.[24,67-71] In summary, Bre1 recruits Rad6 to chromatin, stimulating ubiquitylation of H2B. Ubiquitylated H2B is read by Swd2, allowing Set1 to methylate H3K4 as it moves through the coding sequence with Pol II.

Current models suggest that H3K4 methylation at active genes is only one aspect of Set1 activity and extend Set1 roles to the larger context of genome control. Indeed, Set1 and its homologs exert repressive control of many genomic elements, including transposons, ncRNAs, and rRNA genes.[72-75]

The silencing of cryptic unstable transcripts in budding yeast requires Set1-mediated H3K4 methyltransferase activity. One such transcript, which encompasses a Ty1 retrotransposon LTR, represses transcription and mobilization of Ty1 by acting *in trans* on the Ty1 transcript. Set1 may exert its influence on transcriptional gene silencing at Ty1 directly through H3K4 methylation or indirectly through promotion of antisense transcription.[72] Arguing in favor of antisense-transcript mediated silencing are experiments showing that Set1 represses budding yeast PHO84 by stimulating antisense RNA production.[73] However, sense transcripts from PHO84 were also increased in *set1*Δ, suggesting that *trans* silencing cannot fully account for Set1-mediated repression.[73]

Furthermore, a large class of noncoding RNAs in budding yeast are under repressive control by Set1. Many regulatory ncRNAs in *S. cerevisiae* are destabilized by the 5'-3' RNA exonuclease Xrn1. Van Dijk and colleagues show that a subset of these antisense ncRNAs mediate transcriptional gene silencing through Set1-dependent H3K4 methylation.[74] rRNA genes in budding yeast are also under repressive control by Set1, especially during diamide stress. Set1 exerts repressive control of rRNA genes through multiple pathways – a H3K4me3-dependent activation of the RPD3L complex which acts at ribosomal biogenesis genes and a second Set1-dependent pathway repressing ribosomal protein genes which in which H3K4me3 levels increase during repression.[76]

Histone chaperones and histone gene dosage

The high degree of compaction of the nuclear genome presents an impediment to processes that require access to DNA, including replication, transcription, recombination, and DNA repair. Indeed, chromatin is thought to have repressive effects on basic transcription. Facilitating these processes are histone chaperones, which have wide-ranging roles in nucleosome assembly, deposition, positioning, and degradation. The importance of histone chaperones highlights the fundamental significance of the histones themselves in chromatin function. Histone levels must be precisely maintained and their deposition and positioning coordinated. One notable histone chaperone complex is HIRA, a well conserved controller of nucleosome assembly important for cell cycle progression, development, and senescence.[77] Histone chaperones can also contribute to posttranslational modification of histones. HIRA is implicated in chromatin reassembly behind elongating Pol II. Additionally, HIRA regulates expression of histone-coding genes in a cell-cycle dependent manner. In chapter four, we describe roles for histone gene dosage in repetitive element control.

S. pombe as a model for chromatin regulation

S. pombe has become a valuable model organism for chromatin biology. Many genes are homologous to those in higher eukaryotes. *S. pombe* has large repetitive centromeres (40kb-100kb) similar to human centromeres. Critically, *S. pombe* has complex chromatin with important similarities to that in higher eukaryotes, including RNAi machinery, H3K4me-rich euchromatin, H3K9me-rich heterochromatin, large replication origins, and conserved telomere proteins. By contrast, *S. cerevisiae* has point

centromeres, non-conserved silencing proteins, small replication origins, and lacks H3K9me altogether.

Heterochromatic regions in fission yeast include repetitive centromeres on each of three chromosomes, telomeres, and the silent portion of the mating type locus, which encodes information for mating type switching. Fisson yeast chromosomes are monocentric, with repetitive centromeres centered around the *cnt* locus, at which H3 is replaced by the variant histone CENP-A. Surrounding the central core are convergent chromosome-specific *imr* repeats and the *otr* repeats. The *otr* is comprised of *dg* and *dh* repeats, of which the number and orientation is different on each of the three chromosomes. Centromeric heterochromatin in *S. pombe* is a well-studied model for factors important for heterochromatin establishment and maintenance. Histones in these regions are deacetylated on the H3 N-terminus by the HDACs Clr3, Clr6, and Sir2.[78] H3K9 is methylated by Clr4 and bound by the HP1 homologs Swi6, Chp1, and Chp2. Positive feedback between Clr4 and Swi6 allows heterochromatin to spread across silenced regions. Interestingly and perhaps paradoxically, transcription through heterochromatic repeats reinforces their silenced state through an RNAi-dependent positive feedback loop. Transcripts are bound by RNA-dependent RNA polymerase which generates dsRNA. This dsRNA is the substrate for the endonuclease Dicer (*dcr1)*, which generates siRNAs that are then used by the argonaute homolog Ago1 to target homologous mRNAs for degradation. Ago1 forms the RITS complex along with Chp1 and Tas3, which works to maintain silencing and H3K9me at the *otr*.[79-83]

Heterochromatin also exists at the mating type locus on chromosome II, which contains the information for mating factors and receptors that allow *S. pombe* to conjugate during nitrogen starvation. Repressive chromatin structures form at the *mat2* and *mat3* loci, while *mat1* is expressible and contains either M or P information dependent on mating type. A region of centromeric homology termed *cenH* is the site of heterochromatin nucleation. Two flanking repetitive regions *IR-L* and *IR-R* act as barrier elements for heterochromatin spread.[84]

3D organization of the genome

An emerging body of evidence suggests that the compaction of DNA fibers is both dynamic and relevant to nuclear function. Examples of functionally related regions of DNA which are grouped together spatially include the nucleolus, nuclear lamina,[85] PML bodies, Cajal bodies, and snRNP-rich nuclear speckles.[86,87] It has been shown that individual chromosomes occupy distinct areas of the nucleus.[88,89] Within these chromosome regions, the distinct compaction levels of different genomic regions further contributes chromosome topology **(Figure 5)**.[90-92] In many organisms, gene poor regions tend to exist nearer to the nuclear periphery than gene rich regions.[93,94] Chromosomal regions can be moved relative to one another and to other nuclear structures, a process that is implicated in the regulation of gene expression.[95] Long-range chromosomal movements in Chinese hamster cells are lost in ATP-depleted cells and cells lacking actin or myosin, indicating that active ATP-dependent remodeling processes may have roles in the dynamic regulation of DNA position.[96,97] Loci that have been shown to move toward

the nuclear interior in activating conditions include *IgH* in antibody producing B-lymphocytes,[98] *c-maf* in T cells,[99] and *Mash1* in neurons.[100] Several techniques have been developed to assess the higher order structure of the genome, including fluorescence in situ hybridization (FISH), microscopy, DamID, mathematical modeling, and the chromosome conformation capture methods 3C, 4C, 5C, GCC, and Hi-C.[101-103] Through use of these techniques, the emerging picture is of a highly organized genome. The developing concept that certain loci physically interact with others suggests that common regulatory factors positioned near the interaction site can influence the transcriptional state of the regulated loci. The three dimensional organization of the genome is therefore a potentially significant large-scale regulator of genome activities.

Chromosome conformation capture experiments using synchronized *S. pombe* cells suggest that genes with high, low, and differentially regulated expression levels are positioned in distinct subnuclear regions which change in a cell-cycle dependent manner.[104] Additionally, tRNA genes colocalize to the nuclear membrane with the help of the tRNA-specific transcription factor TFIIIC[105] and telomeres are clustered and tethered to the nuclear periphery by the Ku heterodimer and the Taz1/Rap1/Bqt3/Bqt4 pathway.[106,107] Our lab has previously demonstrated that the 13 *Tf2* retrotransposons are subject to higher order organization, clustering into one or two distinct nuclear foci aided by the centromeric CENP-B homolog Abp1, histone deacetylases Clr3 and Clr6, and the Set1C/COMPASS complex.[75,108] Given that both tRNA genes and *Tf2*s are distributed across three chromosomes, the clustering of these loci represent areas of significant interchromosomal organization.

Summary of results by chapter

In this thesis, we frame an exploration of chromatin regulation with focus on a single SET-domain containing histone lysine methyltransferase, Set1 and its catalytic product, H3K4me. We characterize the surprisingly multifaceted role of Set1 and its associated Set1C subunits in histone modification, transcriptional regulation, and nuclear organization. Experiments described in Chapter two focus on identifying mechanisms controlling repression of repetitive elements. Additionally, we address the role of histone post-translational modifications in defining the functional state of these repetitive regions. Specifically, we describe roles for the H3K4 methyltransferase Set1 and its associated complex in genome control that are both dependent and independent of H3K4 methylation. Finally, we describe a novel interplay between methyl and acetyl marks at H3K4 that drives the nuclear organization of retrotransposons. Our study considerably expands the regulatory repertoire of an important histone modifier and highlights the multifaceted function by a highly conserved chromatin-modifying complex with diverse roles in genome control.

Transposable element (TE) sequences comprise significant percentages of eukaryotic genomes, accounting for nearly 44% of the human genome[109] and more than 50% of the maize genome.[110] Mobilization events of these "jumping genes" are mutagenic and can cause deletions, insertions, inversions and large-scale chromosome rearrangements. Despite the evolution of cellular controls to inhibit TE activities and

maintain genome integrity, these sequences remain widespread. Chapter three expands our work on Set1/COMPASS to the control of *Tf2* mobilization. Multiple modes of transcriptional repression of *Tf2s* have been described but it remains unclear whether transcriptional repression is the sole mechanism for controlling TE activities. We explore the contributions of transcriptional repression and genome organization to the mobilization of *Tf2*s. We find that mutations which both compromise transcriptional repression and also *Tf* body formation result in elevated mobilization rates. However, loss of silencing in the presence of intact *Tf* bodies is not sufficient to render high levels of mobilization. Our results therefore highlight that LTR retrotransposon mobility is subject to regulation at multiple levels and indicates that higher order chromatin organization is an important aspect of cell control.

In Chapter four, we describe an unexpected cooperation between H3K4 and histone dosage in chromatin regulation. Appropriate levels of histone expression are critical for transcription, chromosome segregation, repair, and other chromatin-mediated processes.[111] We previously observed that heterochromatic silencing is maintained in mutants in which the lysine 4 codon at all three *hht* genes was mutated to code for arginine.[75,112] Interestingly, while mutation of lysine 4 of histone to arginine (H3K4R) has little effect on the repression of the centromeric *dg* repeats, H3K4R in strains lacking two of the histone H3 genes results in dramatic increased *dg* expression.[113] This result suggests that histone modification and histone gene dosage cooperate to maintain centromeric silencing. We sought to characterize cellular mechanisms that could link H3K4 modification with histone gene dosage. Additionally, we asked whether individual

hht copies have unique functional roles that could explain their unique deletion phenotypes at elevated temperatures and in the presence of DNA-damaging agents.[114] Our results indicate that H3K4 is necessary to maintain silencing of repetitive elements and to respond to DNA damage in strains with decreased H3 copy number. These results indicate that H3K4 and histone dosage cooperate to maintain chromatin function.

Finally, in Chapter five we discuss these results in the larger context of chromatin biology with emphasis on the role of histone posttranslational modification as a key factor in the dynamic regulation of chromatin activities. Additionally, we discuss how work described in this thesis opens new questions concerning locus-dependent transcriptional control, the regulation of transposable element mobilization, and the role of histone dosage in the maintenance of chromatin function.

Figures

Table 1. Subunit composition and biological activities of COMPASS family from yeast to human. Adapted with permission from "The COMPASS family of H3K4 Methylases: Mechanisms of Regulation in Development and Disease Pathogenesis" by Ali Shilatifard *Annu. Rev. Biochem.* 2012. [41]

S. pombe Set1C/ COMPASS	*S. cerevisiae* Set1C/ COMPASS	*D. melanogaster* COMPASS/ COMPASS-like complexes	*Mammalian* COMPASS/COMPASS-like complexes	Biological/biochemical properties in *S.cerevisiae*
Set1	Set1	Set1, Trx, Trr	Set1A/B, MLL1-4	The catalytic subunits
Ash2	Cps60 (Bre2)	Ash2	Ash2L	Required for H3K4me3
Swd1	Cps50 (Swd1)	RbBP5	RbBP5	Required for assembly
Spp1	Cps40 (Spp1)	CXXC1 (dCfp1)	CXXC1 (Cfp1)	Components of Set1 complexes
Swd2	Cps35 (Swd2)	Wdr82	Wdr82	Required for proper H3K4me2/3
Swd3	Cps30 (Swd3)	Wds	Wdr5	Required for assembly
Sdc1	Cps25 (Sdc1)	Dpy30	Dpy30	Required for H3K4me2/3
Shg1	Shg1	-	-	Components of Set1 complexes

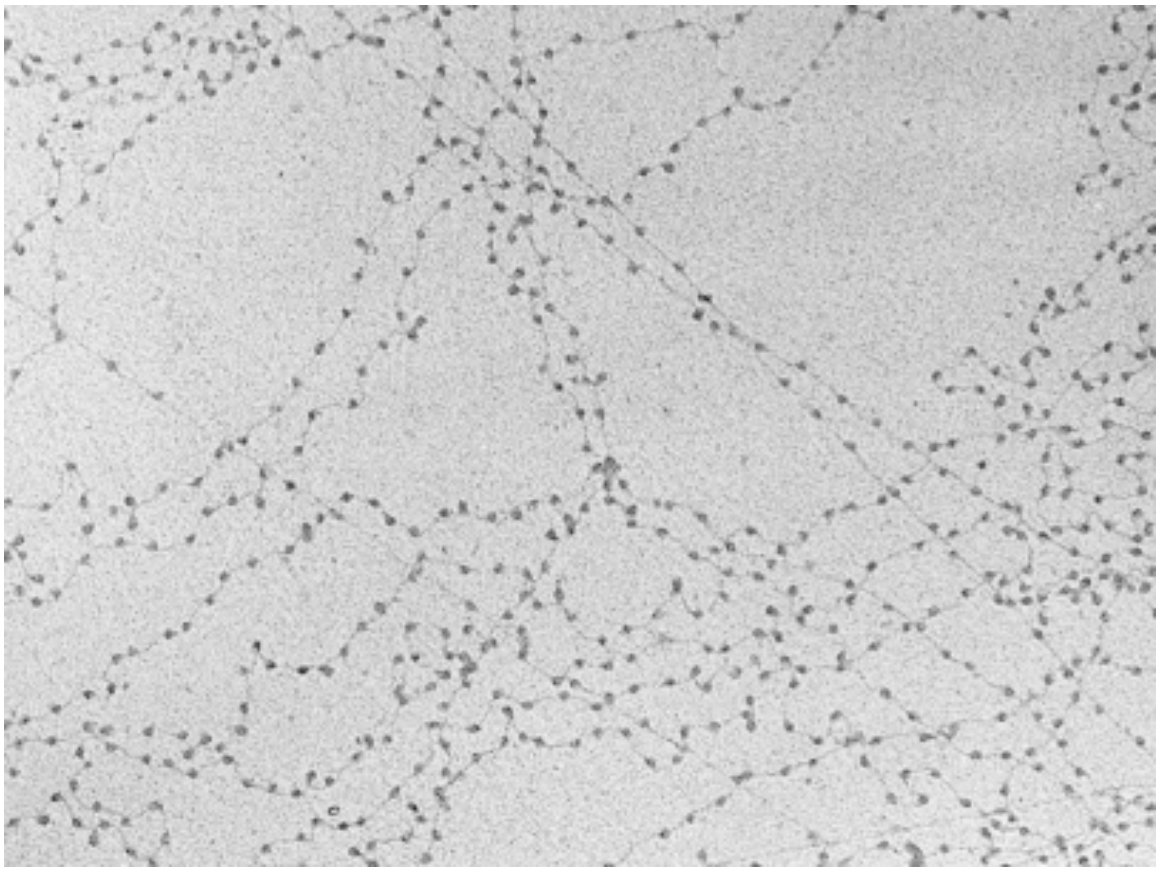


Figure 1. Electron micrograph of decondensed chromatin from chicken erythrocytes. Nucleosomes are visible along the DNA fibers in this seminal image of decondensed chromatin. Image by Chris Woodcock.

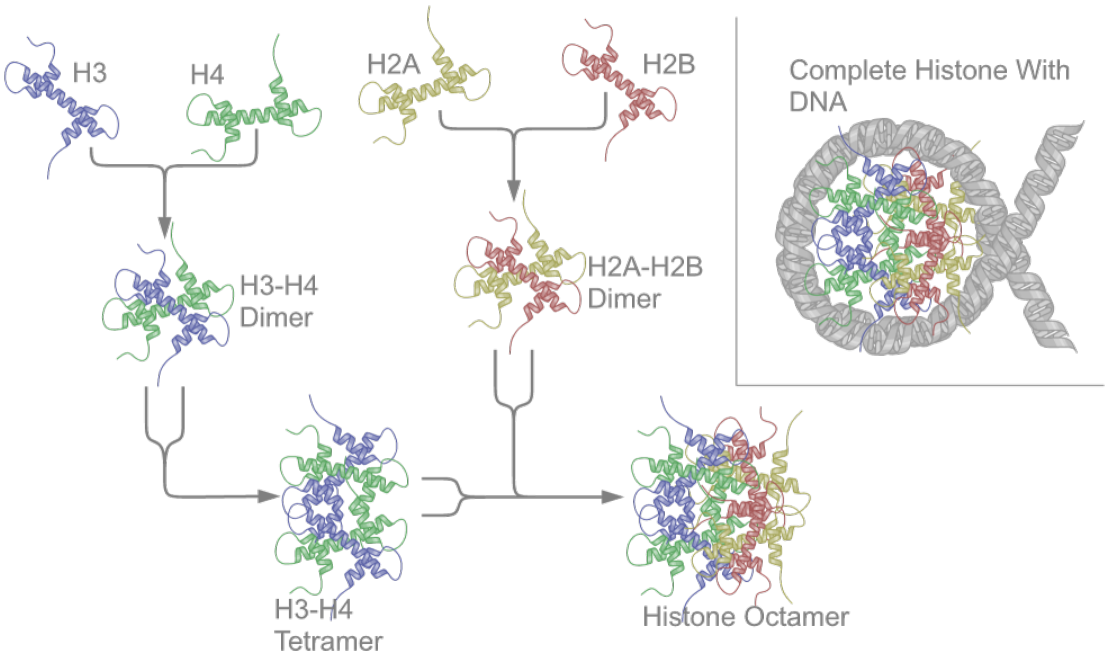


Figure 2. Nucleosome structure. Two copies of each histone form a histone octamer around which ~147bp DNA is wrapped twice. Unstructured polypeptide tails extend from the nucleosome core particle.

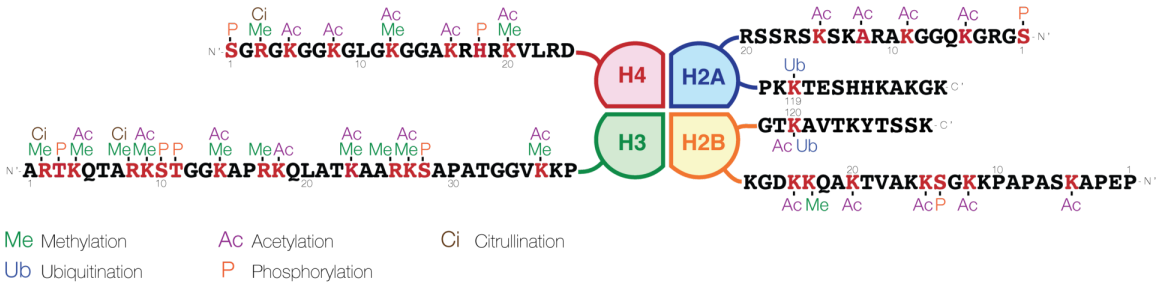


Figure 3. Schematic representation of histone modifications. Unstructured polypeptide tails on each histone are the sites of posttranslational modifications, many of which have demonstrated links to cellular processes. Adapted from Rodriguez-Paredes & Esteller, *Nature*, 2011. [115]

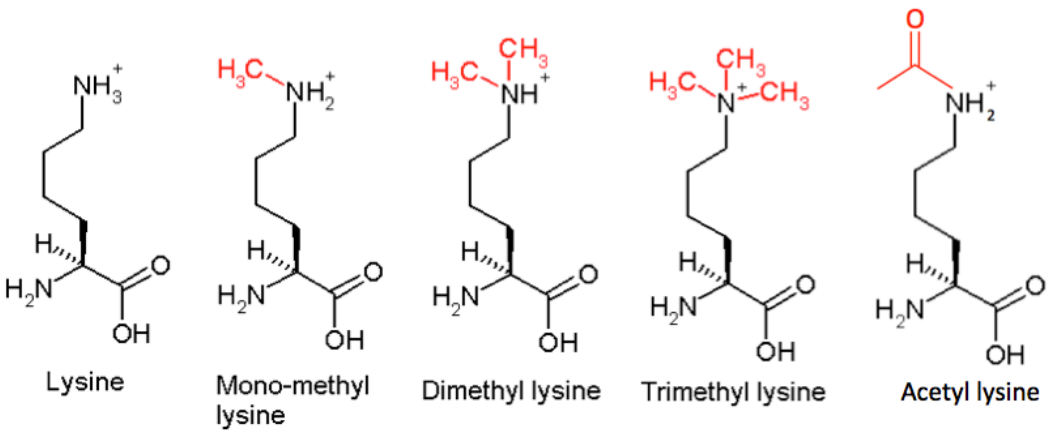


Figure 4. Lysine and its methylated forms. Lysine can be modified posttranslationally by SET-domain methyltransferases such as Set1. The lysine ε-amino group can be modified by the addition of one, two, or three methyl groups or by the addition of an acetyl group.

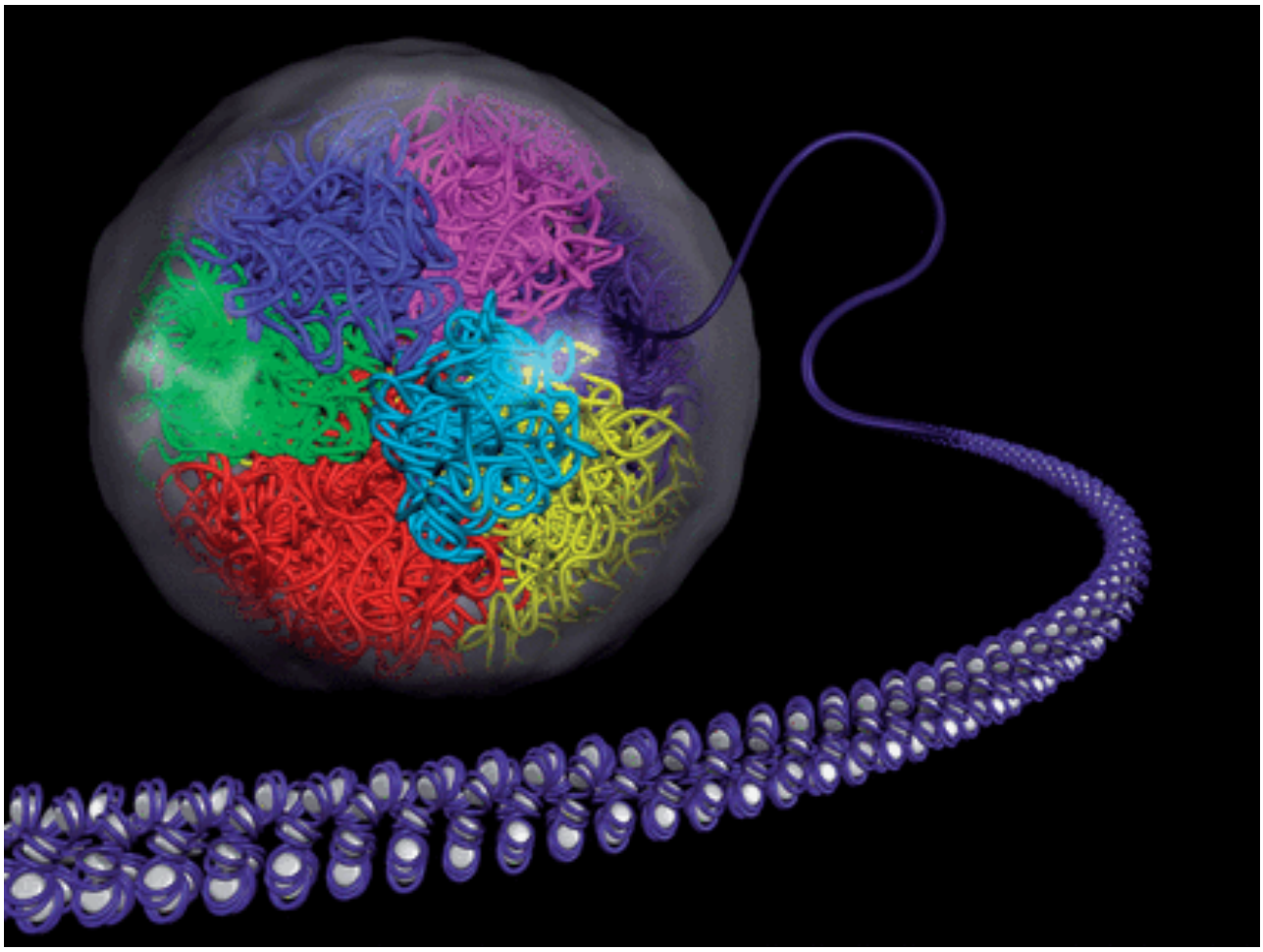


Figure 5. Chromosome territories. Individual chromosomes, represented by colored strands, occupy discrete regions of the nucleus in this artist's representation. Nucleosomes are visible in the 30nm fiber configuration in an enlarged strand in the foreground (purple). Reproduced from with permission from Dorier & Stasiak. *Nucleic Acids Research.* 2009. [116]

Chapter Two : Multifaceted Genome Control by Set1 Dependent and Independent of H3K4 Methylation and the Set1C/COMPASS Complex

Set1 domains and methyltransferase activity play distinct roles in transcriptional silencing of *Tf2* retrotransposons and heterochromatic repeats

Fission yeast Set1 is highly conserved across eukaryotes. At the level of primary peptide sequence it contains two putative RNA-recognition domains RRM1 and RRM2.[48,117-119] An nSET, SET, and post-SET domains form the catalytic core of the enzyme **(Figure 6)**.[118,120,121] Previous studies in *S. cerevisiae* have shown unique contributions to H3K4me for individual Set1C/COMPASS subunits.[19,118,119,122]

To examine the contribution of each of Set1 domains to its methyltransferase function, we generated deletion mutants by site directed mutagenesis, extracted bulk histone by acid extraction and trichloroacetic acid precipitation, and performed western blots using antibodies specific for H3K4me1, H3K4me2, and H3K4me3 **(Figure 7)**. Western blots using an antibody against histone H3 served as a loading control. We determined that loss of the RRM1 domain results in loss of H3K4me3 and H3K4me2, and significantly diminishes H3K4me1 levels. Loss of the RRM2 domain did not significantly affect H3K4me levels. Loss of the nSET, SET, and pSET domains fully abolished H3K4me. Additionally, we found that Set1 tagged with a FLAG epitope at its C-terminus also was unable to methylate H3K4 **(Figure 7)**.

To determine if Set1 protein stability was affected by the disruption of the domains, we performed western blots against N-terminally FLAG-tagged versions of each of the domain deletion strains. In contrast to reported feedback control of Set1

protein levels in *S. cerevisiae,*[52] Set1 was detectable by western blot in all strains tested. Set1-RRM1Δ and Set1-nSETΔ both showed a reduction in Set1 protein levels, suggesting diminished H3K4me levels in these strains may be a result of diminished Set1 availability. Set1 protein levels in RRM2Δ, SETΔ, and pSETΔ were reduced slightly **(Figure S3)**.

Because of H3K4me association with actively transcribing Pol II genes, we sought to determine the role of Set1 domains on transcriptional activity. We extracted RNA from exponentially growing cells, performed reverse transcription and subjected the cDNA to realtime qPCR to determine the relative transcript abundance at several repetitive loci – *Tf2,* the centromeric *dg* repeats, the mating type locus (*mat)*, and the subtelomeric *prl70* gene **(Figure S1)**. We found commonalities between the method of transcriptional repression at these loci. All of the four regions were upregulated in *set1*Δ, suggesting that Set1 acts in a repressive manner at heterochromatic repeats. Genome scale analysis of *set1*Δ cDNA by Agilent tiling microarray supported this finding, with nearly 1,000 of 42,000 array probes upregulated in the absence of *set1*.[123] RRM1 is essential for repression of the mating type locus and subtelomere, but is dispensable for repression of *Tf2* and *dg.* Deletion of nSET and pSET gave comparable results. The RRM2 domain by contrast is essential for repression of *Tf2* and *dg* but is not involved in repression of the mating type locus and subtelomere. The SET domain is necessary for repression of all four loci. These data highlight distinct functional requirements for Set1 in locus-dependent repression.

A catalytically inactive Set1 mutant with a FLAG epitope tag at the C-terminus showed derepression of the mating type locus and subtelomere, but did not affect *Tf2* and *dg* regulation **(Figure S1)**. Taken in context with the results of the H3K4me western blots, these data indicate that any impairment of normal H3K4 methylation results in disregulation of the mating type locus and subtelomere, suggesting a regulatory pathway that is H3K4me-dependent acting at these loci. In contrast, *Tf2* and *dg* are subject to a distinct mode of regulation that is dependent on SET and RRM2 but independent of H3K4 methylation **(Figures 7 and S1)**.

Set1C/COMPASS subunits make unique contributions to Set1 methyltransferase catalysis.

Next we sought to characterize the roles of the Set1C/COMPASS complex subunits. Although the contribution to each subunit to H3K4me has been characterized in budding yeast,[49,50,124-127] the roles of each subunit have not been characterized in *S. pombe* aside from H3K4me2.[117] We generated mutants in which the open reading frames for each Set1C/COMPASS subunit were replaced with a kanamycin resistance marker, extracted bulk histone by acid extraction and trichloroacetic acid precipitation, and performed western blots using antibodies specific for H3K4me1, H3K4me2, and H3K4me3 **(Figure 8)**. Western blots using an antibody against histone H3 served as a loading control. These experiments showed unique profiles of global H3K4 methylation in strains with modified COMPASS. *ash2*Δ and *swd2*Δ abolished H3K4me3 and reduced levels of H3K4me1 and H3K4me2. *spp1*Δ, *swd1*Δ, and *swd3*Δ abolished H3K4me3 and

H3K4me2, and had greatly reduced levels of H3K4me1. *shg1*Δ and *sdc1*Δ retained all three forms of H3K4me, although the abundance of each was distinct from WT strains. *shg1*Δ resulted in decreased H3K4me3 and increased H3K4me1 and *sdc1*Δ resulted in reduced levels of H3K4me3. These data indicate that although strains lacking the catalytic COMPASS subunit Set1 were only strain that resulted in complete loss of all three forms of H3K4me, Spp1, Swd1, and Swd3 are also of critical importance for all three forms of H3K4me. The Ash2 and Swd2 subunits appear to have special roles in H3K4me3 catalysis. The roles of Shg1 and Sdc1 are less critical to Set1 methyltransferase function, but nevertheless have roles in maintaining the relative levels of H3K4me1, H3K4me2, and H3K4me3 **(Figure 8)**. Set1 transcript levels were consistent in Set1C/COMPASS mutant strains **(Figure S4)**.

Notably, the contribution of each COMPASS subunit to H3K4me is different from those in the budding yeast, *Saccharomyces cerevisiae.* In *S. cerevisiae*, *spp1*Δ diminishes or abolishes H3K4me3.[49,124,128] Loss of *swd2* in *S. cerevisiae* is lethal and loss of *ash2* or *sdc1* reduced the levels of all three H3K4me states or abolished H3K4me2 altogether.[44,50,124,126] These results show that although the composition of the COMPASS complex is well conserved across evolutionary distance, there nevertheless is variability in the functional roles of some subunits with regards to global H3K4me levels.

Set1C/COMPASS subunits exhibit distinct effects on transcriptional repression of *Tf2s* and heterochromatic repeats

We then sought to characterize the roles of the Set1C/COMPASS complex subunits in transcriptional regulation of repetitive loci. We extracted total RNA and performed qPCR using primers specific to *Tf2* retrotransposons, centromeric *dg* repeats, the mating type locus and subtelomeric *prl70* **(Figure S2)**. All loci were derepressed in *set1*Δ, again indicating that Set1 has repressive function at these repetitive regions. Although ChIP of H3K4me shows a strong association with actively transcribed regions, this is only one aspect of Set1 influence on genome wide regulation. Repression of *Tf2* and *dg* were consistent with WT in strains lacking any of the other seven COMPASS components, indicating that Set1 itself is the critical regulatory component of the complex at these regions, and exerts repressive control through its SET and RRM2 domains. In contrast, we observed significant derepression of the mating type locus in each COMPASS deletion strain with the exception of *shg1*Δ. *mat* RNA was enriched 8.5-fold over wild type in *ash2*Δ. *mat* RNA was enriched 5-fold over wild type in *set1*Δ, *swd1*Δ, and *sdc1*Δ. Intermediate derepression was observed in *spp1*Δ, *swd2*Δ, and *swd3*Δ. Taking into account the modified levels of H3K4me in each of these strains, we can conclude that repression of the mat locus is correlated with the degree to which wild type levels of H3K4me are present. In other words, the Set1 mechanism of repressive control at the mat locus and subtelomere is dependent on H3K4me, a repressive mechanism distinct from H3K4me-independent repression at the centromere and subtelomere. These

data also suggest an additional repressive role for Sdc1 at the mating type locus and subtelomere **(Figure S2)**.

Set1 regulates the nuclear organization of *Tf2s* distinct from its transcriptional repressor function

Previous work from our group showed that *Tf2*s are clustered within the nucleus into structures termed *Tf* bodies.[75,108] The CENP-B homolog Abp1, HDACs, and Set1 have roles in *Tf* body formation. To determine which domains of Set1 are crucial for *Tf* body formation, we performed fluorescence *in situ* hybridization (FISH) using *Tf2*-specific probes in mutants lacking Set1 domains **(Figure 9)**. Defective *Tf* body formation was observed in strains lacking the RRM1, nSET, SET, and pSET domains as well as *set1*Δ and *set1F*^{H3K4me}, the same strains which lack H3K4me. In RRM2Δ, which maintains H3K4me while modestly disregulating *Tf2* expression, *Tf* body formation was only slightly altered. These data suggest a role for H3K4me in maintaining the nuclear organization of *Tf2*s and suggest that Set1 relies on different domains and potentially different catalytic functions to control various aspects of *Tf* regulation.

Roles for Set1C/COMPASS in *Tf* body maintenance

To further elucidate a possible H3K4me-dependent mechanism for *Tf2* clustering, we expanded our *Tf2* FISH experiments to include strains defective for components of the Set1C/COMPASS complex **(Figure 10)**. These results align with those obtained in FISH experiments for Set1 domain mutants in that mutations that affect H3K4me cause defects in *Tf2* clustering. Specifically, *Tf2* clustering is lost in COMPASS mutants in which H3K4me is lost, including *set1*Δ*, ash2*Δ*, spp1*Δ*, swd1*Δ*, swd2*Δ, and *swd3*Δ. Deletion of *sdc1*, in which H3K4me is not affected, does not disrupt *Tf2* clustering. Deletion of *shg1* is an exception to this pattern, suggesting Shg1 has a separate, H3K4me-independent role in *Tf2* clustering maintenance.

Set1/COMPASS maintains *Tf* body organization by antagonizing the H3K4 acetyltransferase Mst1

Strains with defects in H3K4 methylation are often coorelated with defective *Tf2* clustering. However, *Tf2* clustering is maintained in H3K4R and H3K4A strains lacking both H3K4me and H3K4ac.[75] We asked whether acetylation of H3K4, mediated by the histone acetyltransferase Mst1,[113] could act as a regulator of *Tf2* clustering by competing with Set1-mediated H3K4me. We used an H3K4ac antibody to immunoprecipitate chromatin and observed elevated H3K4ac levels at *Tf2* and the housekeeping gene *act1* in strains lacking H3K4me. Loss of H3K4me in these strains was able to decrease elevated H3K4ac levels in *set1F*$^{-H3K4me}$ strains **(Figure 11B,C)**. We then performed *Tf2* FISH in

strains with the *mst1-L344S* allele, which is defective for H3K4ac at 30°C.[113] Interestingly, *Tf2*s cluster more tightly when H3K4ac is lost, suggesting that loss of *Tf2* clustering in strains lacking H3K4me could be due to inappropriate Mst1 activity such as increased H3K4ac at *Tf2* **(Figure 11A,C)**. Indeed, loss of H3K4ac in cells lacking H3K4me abrogated defects in *Tf2* clustering in *mst1-L344S set1F*$^{-H3K4me}$ strains. These results demonstrate that Set1C/COMPASS-mediated H3K4me controls *Tf* body organization by antagonizing the H3K4ac activity of Mst1.

Loss of *set1* results in increased H3K9 acetylation at pericentromeric repeats.

Heterochromatic repeats and *Tf2*s in certain genetic backgrounds are targeted for RNAi-mediated heterochromatic and exosome-mediated silencing.[129-131] The exosome is a six-member exoribonuclease complex which is involved in RNA processing and degradation in eukaryotes.[132] It has been shown that histone deacetylases (HDACs) cooperate with RNAi to assemble heterochromatin at pericentromeres.[133] Even though loss of set1 does not appear to affect the levels of H3K9 methylation and siRNAs at pericentromeric heterochromatin,[113] there were noticeable increased levels of H3K9 acetylation at that region **(Figure 12)**. Several HDAC mutants including *sir2* and *clr3* are known to retain robust levels of siRNAs and H3K9me and yet exhibit increased levels of certain histone acetylation marks at pericentromeres.[133-135] Thus, it is likely that in the absence of set1, HDACs, RNAi and exosome act in redundant pathways to help maintain heterochromatin.

H3K9 dimethylation (H3K9me2) in strains deficient for set1 and clr3 at the pericentromeric dg repeat.

The class II HDAC Clr3 has been shown to contribute to transcriptional silencing of heterochromatin, *Tf2*s, and stress-response genes. These classes of genetic elements are also regulated by Set1, suggesting a functional link between Clr3 and Set1. Derepression of a reporter gene inserted within the pericentromeric repeats has been observed in mutants deficient for either *set1* or *clr3*. We observed additional derepression in *set1*Δ*clr3*Δ double mutants. We performed ChIP followed by qPCR to assess the status of centromeric heterochromatin in these mutants **(Figure 13)**. H3K9me2 levels are increased at *dg* in *clr3*Δ. In contrast, H3K9me2 levels were dramatically reduced in *set1*Δ*clr3*Δ double mutants. These results reveal compensatory mechanisms by Set1 and Clr3 acting in parallel pathways to maintain centromeric heterochromatin.

Figures

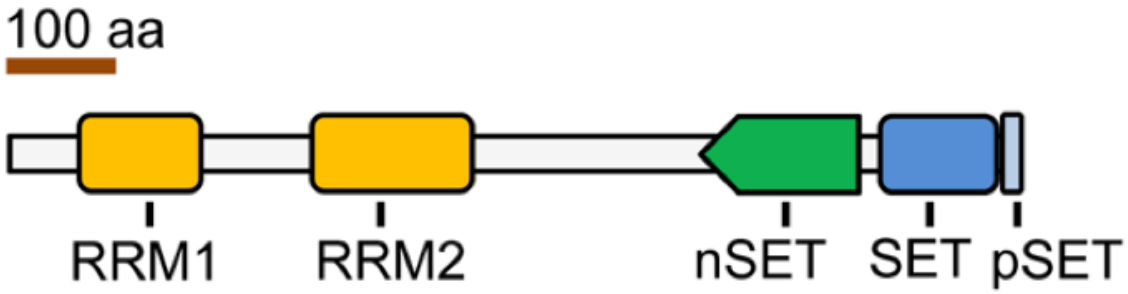


Figure 6. Schematic of Set1 protein architecture.

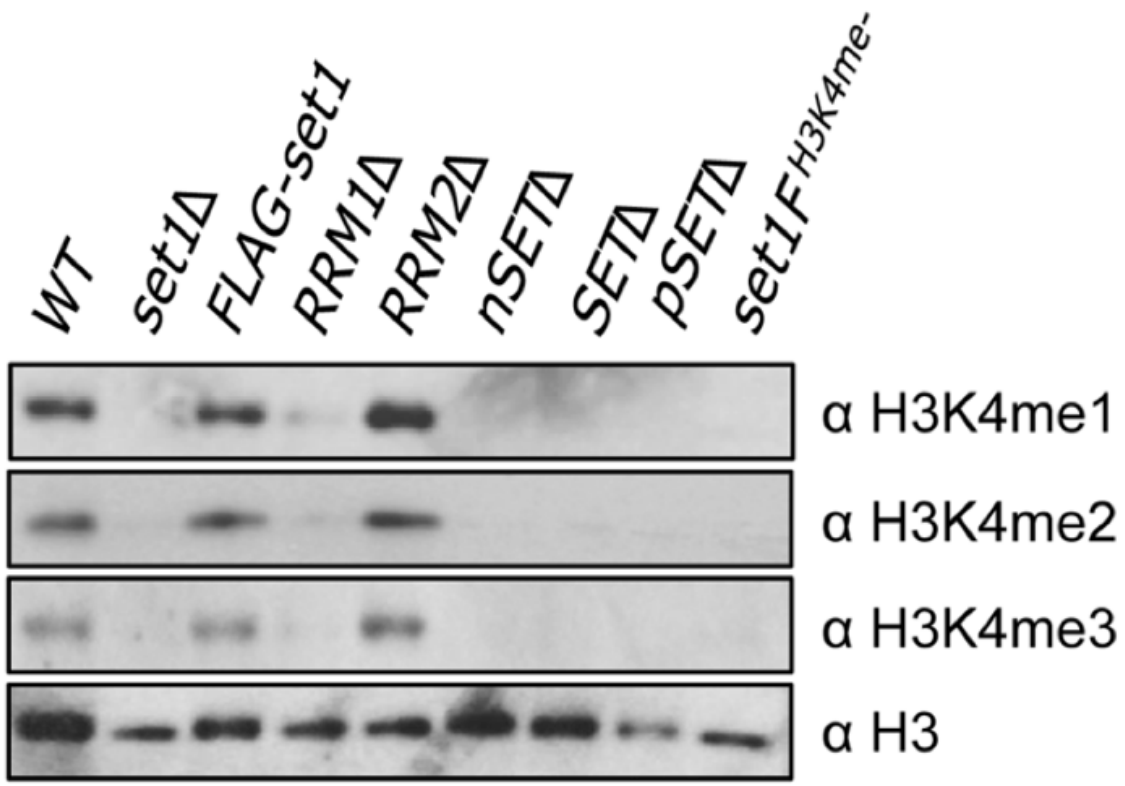


Figure 7. H3K4 methylation (H3K4me) in set1 domain mutants. Mono (H3K4me1), di (H3K4me2), and tri (H3K4me3) methylation of H3K4 was analyzed from histone extracts of indicated set1 mutant strains by western blotting. Full-length FLAG-Set1 and the indicated deleted domain mutants of Set1 contain an N-terminal FLAG epitope. set1F-^{H3K4me} denotes an H3K4me null mutant due to the presence of a FLAG epitope at the Set1 C-terminus.

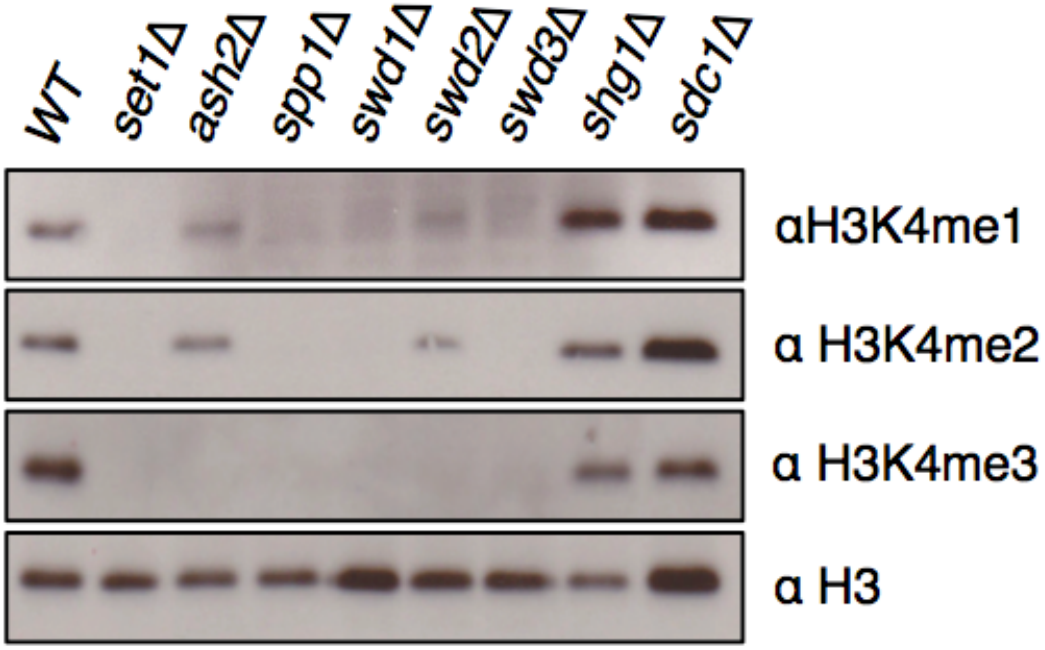


Figure 8. Set1C components contribute differently to H3K4me. H3K4me1, H3K4me2, and H3K4me3 were analyzed from histone extracts of indicated Set1C mutant strains by western blotting.

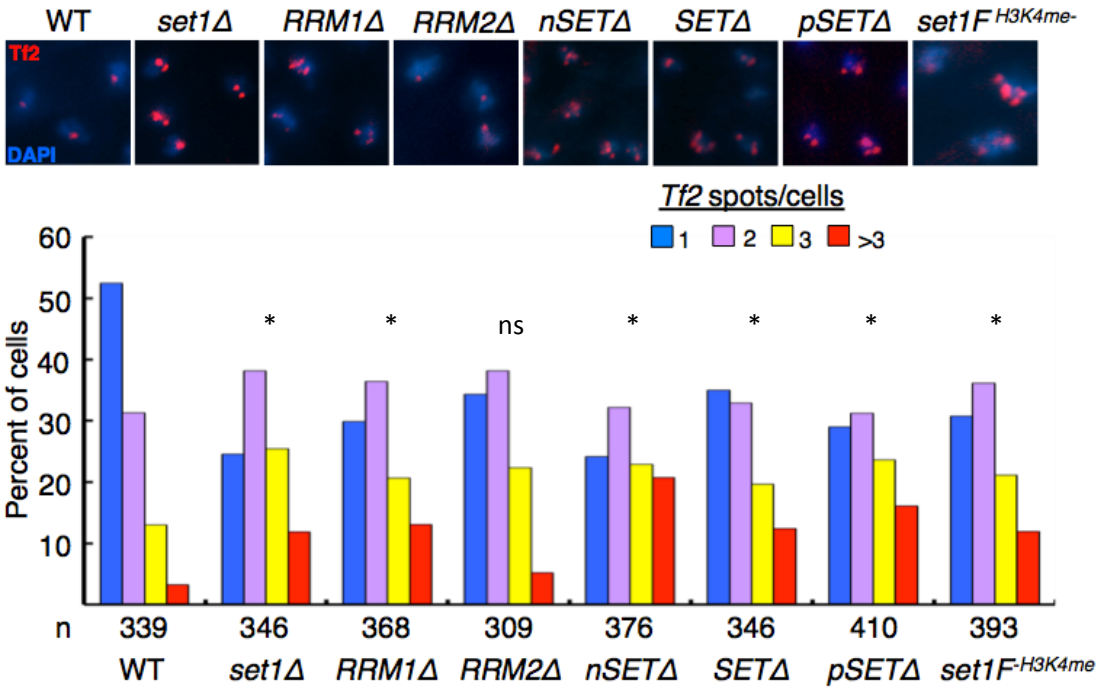


Figure 9. H3K4me is required for the integrity of *Tf* bodies. Fluorescence in situ hybridization (FISH) analysis was performed using a FISH probe corresponding to the ~3.6 kb *Tf2* coding region. Representative FISH images from indicated strains (top panels). Quantitative FISH analysis of observed *Tf2* foci/cell in indicated strains (bar graph; bottom panels). Number of cells analyzed per strain (n). Except for RRM2Δ, *Tf2* declustering in all mutant strains compared to WT was significant (p,0.005, chi-square test). * denotes $p < 0.001$ and ns (not significant) denotes $p > 0.05$ (χ^2).

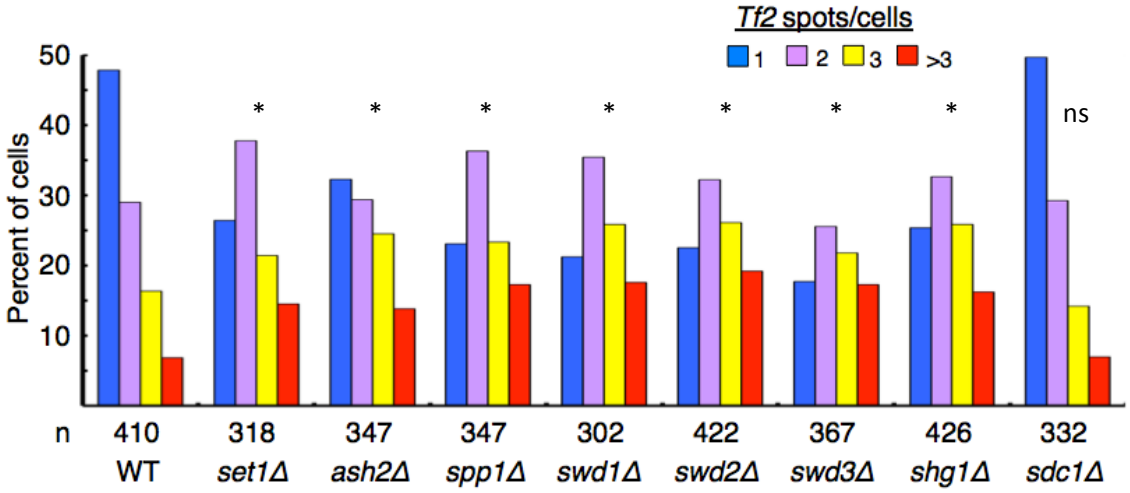
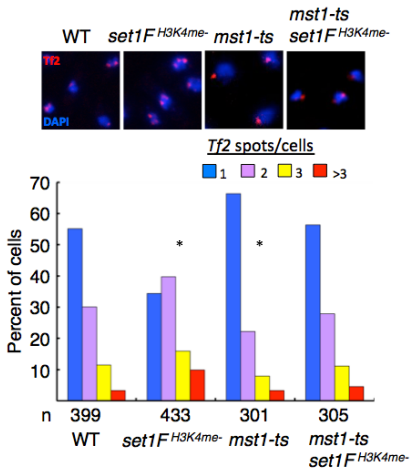
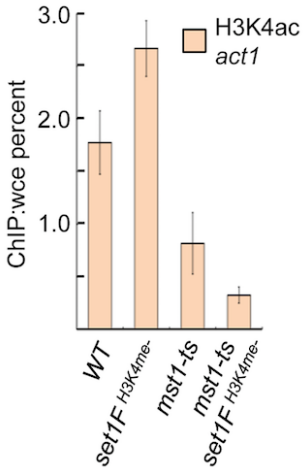


Figure 10. Set1C/COMPASS is required for *Tf* body integrity. Fluorescence *in situ* hybridization (FISH) analysis was performed using a FISH probe corresponding to the ~3.6 kb *Tf2* coding region. Representative FISH images from indicated strains (top panels). Quantitative FISH analysis of observed *Tf2* foci/cell in indicated strains (bar graph; bottom panels). Number of cells analyzed per strain (n). With the exception of *sdc1Δ, Tf2* declustering in all mutant strains compared to WT was significant (p,0.005, chi-square test). * denotes $p < 0.001$ and ns (not significant) denotes $p > 0.05$ (χ^2).

A.



B.



C.

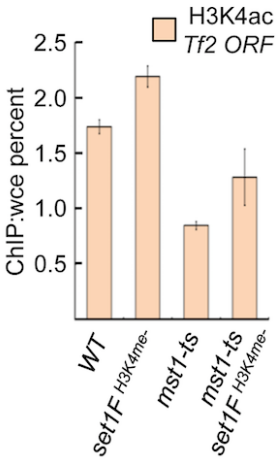


Figure 11. Set1C-mediated H3K4me contributes to the integrity of *Tf* bodies by antagonizing the H3K4 acetyltransferase Mst1. (A) An mst1 mutation (*mst1-ts*) alleviates Tf body defects seen in an H3K4me mutant strain. Representative FISH images from indicated strains (top panels). Quantitative FISH analysis of observed *Tf2* foci/cell in indicated strains (bar graph; bottom panels). *Tf2* declustering for *set1F*$^{H3K4me-}$ strain was significant ($p<0.005$, chi-square test). H3K4ac enrichment at (B) the housekeeping gene *act1* and (C) *Tf2* ORF in indicated mutant strains was analyzed by ChIP followed by qPCR. * denotes $p < 0.001$ and ns (not significant) denotes $p > 0.05$ (χ^2).

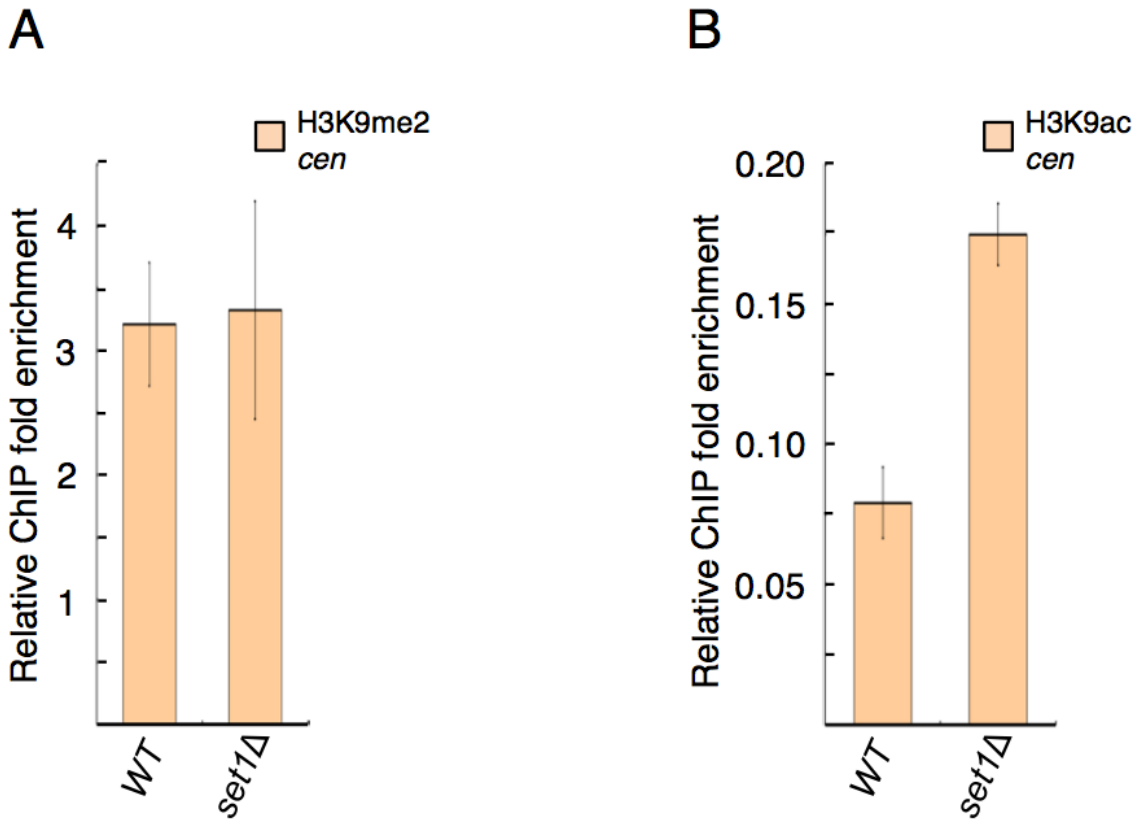


Figure 12. Loss of set1 results in increased H3K9 acetylation at pericentromeric repeats. A) Enrichment of H3K9 methylation (H3K9me2) and (B) H3K9 acetylation (H3K9ac) was determined by chromatin immunoprecipitation (ChIP) followed by quantitative PCR (qPCR) using primers corresponding to the pericentromeric repeat dg region. ChIP fold enrichment of H3K9me2 and H3K9ac at the dg repeat was determined relative to the corresponding enrichment at (A) the act1 promoter and (B) the 3′ region of act1, respectively (s.d., error bars; n=3 triplicates).

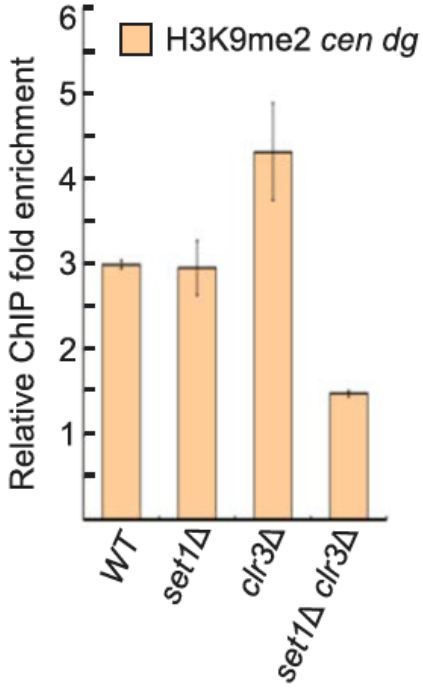


Figure 13. H3K9 dimethylation (H3K9me2) in strains deficient for set1 and clr3 at the pericentromeric *dg* repeat. H3K9me2 enrichment at *dg* in indicated strains was carried out by chromatin immunoprecipitation (ChIP) and quantified by qPCR (s.d., error bars; n=3 triplicates). ChIP fold enrichment is relative to *act1*.

Chapter three: Restriction of retrotransposon mobilization by transcriptional silencing and higher-order chromatin organization

Introduction

Long terminal repeat (LTR) retrotransposons are virtually ubiquitous in eukaryotes and have had major impacts upon host genome evolution, organization and function.[136] They are structurally related to exogenous and endogenous retroviruses and are composed of LTR sequences that flank genes encoding, Gag, protease (PR), reverse transcriptase (RT) and integrase (IN) proteins.[137] Retrotransposon RNA is synthesized by host RNA polymerase II from a promoter in the 5′ LTR. The resulting mRNA serves as a template for the translation of retrotransposon proteins and also for reverse transcription. Reverse transcription occurs within a virus-like particle and the resulting cDNA is inserted into the genome by the element-encoded integrase (IN) or by homologous recombination.[137] The insertion of retrotransposon cDNA is inherently mutagenic with potentially deleterious effects on the host.[138,139] Furthermore, the repetitive nature of retrotransposons renders them substrates for recombination and potential drivers of genome rearrangements. As a result, these elements have traditionally been viewed as harmful genomic parasites.[140] However, there are numerous examples where host cells have domesticated transposon proteins or sequences for their own use, a process termed exaptation.[141,142] As such, retrotransposons provide a reservoir of genetic variability.[143,144]

It is generally accepted that epigenetic controls, which suppress the transcription of retroelements, play a key role in preventing their uncontrolled spread.[145] DNA methylation, RNA interference (RNAi), histone modification and chromatin remodelling

have all been implicated in the suppression of specific families of LTR retrotransposons and endogenous retroviruses.[145,146] However, in many cases the controls that regulate expression of these elements are incompletely understood. Moreover, genetic studies indicate that the mobilization of LTR retrotransposons is subjected to multi-layered regulation.[147]

Analyses of yeast species such as *Saccharomyces cerevisiae* and *Schizosaccharomyces pombe* have provided fundamental insights into LTR retrotransposon biology.[148,149] The genome of the common laboratory strain of *S. pombe* (972) contains a highly homogenous group of 13 *Tf2* LTR retrotransposons which belong to the Ty3/Gypsy family.[150] A closely related element called *Tf1* is present in other wild strains but full length copies of this element are absent in the laboratory strain 972.[150] However, there is an extensive population of ~250 solo LTRs in this strain which includes sequences derived from other *Tf* families including *Tf1*. Sequence analysis indicates that the majority of the *Tf2* elements have the potential to be active. Indeed, *Tf2-12* has been shown to mobilize with a frequency of ~2 new insertions per 10^8 cells.[151] This low rate of mobilization is consistent with low levels of *Tf2* transcription in wild type cells grown under standard conditions. Indeed, a number of studies indicate that the expression of *Tf2* retrotransposons is subjected to chromatin-mediated silencing by a variety of factors including CENP-B proteins,[108] the Set1 histone methyltransferase,[75] multiple histone deacetylases[108,152-154] and the histone chaperones HIRA and Asf1.[155-157] Interestingly, the RNAi machinery plays only an accessory role to the exosome in this process.[83,129,153] The role of the CENP-B homologues represents an interesting instance of

exaptation as these proteins are derived from a transposase derived from an ancient DNA transposon.[158] In addition to their roles in *Tf2* silencing, CENP-B and Set1 also function to cluster *Tf2* elements and solo LTRs into subnuclear structures called *Tf* bodies.[75,108,112] These bodies are not apparently necessary for silencing of *Tf2* elements but they have been proposed to prevent integration via recombination of other *Tf* elements.[108,112] While it is generally accepted that silencing restricts transposable element (TE) mobility, it remains unclear whether transcriptional repression is the sole mechanism for controlling TE activities.[139] The highly tractable genetics of fission yeast combined with the well-characterized biology of *Tf2* retrotransposons represents an excellent system to address this issue.[139] In this study, we have constructed a sensitive reporter assay that enables us to monitor the transposition of an endogenous *Tf2*. Rather than relying on the plasmid-encoded elements expressed from heterologous promoters that could circumvent the transcriptional controls to which endogenous native elements are subjected, this assay enabled us to determine the impact of mutations in key regulatory genes upon the mobilization frequency of an endogenous *Tf2* element. We find that mutations which both compromise transcriptional repression and also *Tf* body formation result in elevated mobilization rates. However, loss of silencing in the presence of intact *Tf* bodies is not sufficient to render high levels of mobilization. Our results therefore, highlight that the mobility of LTR retrotransposons are subjected to regulation at multiple levels and indicates that higher order chromatin organization is an important aspect of host cell control.

Long terminal repeat retrotransposons are widespread in eukaryotes. These elements resemble retroviruses both structurally and in terms of their replication strategy.

Retrotransposons have directional repeat regions at either end (LTRs) and genes coding proteins homolgous to retroviral capsid, protease, reverse transcriptase, and integrase. Additionally, LTR retrotransposons include a primer binding site for reverse transcription, and a polypurine tract where reverse strand cDNA synthesis is primed. These proteins come together to form a virus-like particle where reverse transcription takes place. cDNA produced is integrated back into the genome in a "copy-paste" mechanism.

The effects of retrotransposons on the host organism is a balancing act. Retrotransposons can serve as sources of variation by inducing chromosomal rearrangements, inducing insertions and deletions, and changing the transcriptional status of nearby loci by position effect variegation. It has been hypothesized that changes to the genome resulting from transposable element activities are a major contributor to speciation, especially in individuals that are physically isolated from the rest of the population. At the same time, retrotransposition can have negative impacts on host cell viability if their mobilization disrupts an essential gene.

The laboratory reference strain of S. pombe contains one family of retrotransposons, *Tf2*. *S. pombe* has 13 full length *Tf2* copies. Past LTR-LTR recombination events have left solo LTRs throughout the genome, including 35 *Tf2* solo LTRs and 139 LTRs from other, now extinct, LTR elements.[150] *Tf2*s tend to insert upstream of Pol II promoters. *Tf2* density is highest on chromosome III, with an average density of one *Tf2* insertion per 37.8kB.[150]

Fission yeast has evolved several mechanisms to limit the spread of retrotransposons. Silencing of transposons is achieved in cells by multiple and sometimes overlapping mechanisms. A robust mechanism for transposon silencing in *S. pombe* involves binding of the three CENP-B homologues Abp1, Cbp1, and Cbh1 to LTRs, Tf2s, *Tf2* remnants, *wtf* repeats, and intergenic regions.[108] Interestingly, the CENP-Bs are derive from pogo transposase and have been shown to bind specifically to transposon sequences and aid in the recruitment of other silencing components including the HDACs Clr3 and Clr6. *Tf2*s are enriched in H3K4me and require the H3K4 methyltransferase Set1 for silencing,[75,108] a process that is independent of H3K4me and the COMPASS complex. Genome organization also plays a role in the silencing of *Tf2*s – they are clustered together by CENP-Bs into one or two distinct regions.[108] Although strains lacking RNAi components show mild derepression of *Tf2*s, *Tf2* siRNAs are dectable in strains lacking the exosome component Rrp6, showing that RNAi cooperates with heterochromatin factors and the exosome to silence *Tf2*s.[129]

Transposition assays using *Tf2-12::Nat-AI*

To study the cellular mechanisms controlling mobilization of *Tf2* we utilized an assay developed by the Whitehall group at Newcastle University who generated an allele of *Tf2-12* tagged with a nourseothricin resistance cassette which is disrupted with an artificial intron (*Tf2-12::natAI)* **(Figure 14)**. Because the intron is oriented in the same transcriptional direction as the *Tf2* and the *Nat* cassette is in the reverse orientation, strains become resistant to nourseothricin only after a mobilization event. Specifically, any transcription of the nourseothricin resistance cassette in the endogenous *Tf2-12* will be unable to splice the intron, generating a nonfunctional nourseothricin acetyltransferase. However, transcription of *Tf2-12* generates an mRNA in which the artificial intron is spliced out. Successful reverse transcription and integration to a new locus will then generate functional noursethricin acetyltransferase, allowing for selection on nourseothricin-containing media **(Figure S5)**. Experiments in the Whitehall lab and our own showed that *Tf2-12* mobilizes at an average frequency on 2.06×10^{-8} in wild type strains. Importantly, mobilization was reduced in *rad51*Δ which agrees with results suggesting the majority of *Tf2* mobilization are through homologous recombination rather than via the element-encoded integrase **(Figure S6)**.[159] Rates of mobilization are increased in *abp1*Δ and *set1*Δ **(Figure 15, Figure S8)**.

Constitutive transcriptional activation drives *Tf2* mobilization.

Previous studies demonstrate that LTR retrotransposons growing in quiescent conditions are subject to silencing mechanisms that restrict their mobilization. However, these controls are circumvented in stress conditions.[149,151,160,161] Stress conditions that are associated with increased element mobilization include cadmium stress, hypoxia, MMS, HU and heat stress. *Tf2* response to hypoxic conditions is mediated by the hypoxia transcription factor Sre1, an ortholog of the mammalian sterol element binding protein.[162] In oxygen-rich conditions, Sre1 is bound to the endoplasmic reticulum outer membrane. In low oxygen conditions, proteolytic cleavage allows for the dissociation of Sre1 from the ER membrane and its translocation of the cleaved fragment into the nucleus, where it acts as a transcription factor to activate genes involved in oxygen stress as well as *Tf2*s through SRE elements in the LTRs.[151,162] Cells expressing constitutively active Sre1 (*sre1-N)* induced *Tf2* expression and increased mobilization 20-fold **(Figure S7)**. Therefore, active Sre1 can circumvent *Tf2* mobilization controls.

Loss of HIRA-mediated transcriptional silencing does not result in increased *Tf2* mobilization

Tf2 expression is repressed by the HIRA histone chaperone complex. Loss of any of the four HIRA complex subunits (Hip1, Slm9, Hip3, Hip4) results in increased *Tf2* expression.[155,156,163,164] We hypothesized that increased *Tf2* expression in these strains would correlate with increased *Tf2* mobilization. Surprisingly, despite greatly increased

Tf2 expression in these strains, mobilization was only modestly increased **(Figure 16A,B, Figure S9)**. RT-qPCR analysis of RNA from these strains showed that *Tf2* RNA in *hip1*Δ is more than twice as abundant as in *sre1-N* strains in which *Tf2* mobilization is dramatically increased. Mobilization in *hip1*Δ *sre1-N* double mutants is not reduced from the elevated mobilization frequency observed in *sre1-N*, suggesting that transcriptional regulation of *Tf2* is not sufficient to restrict mobilization **(Figure S9)**.

HIRA is not required for the clustering of *Tf2* elements

The lack of correlation between *Tf2* expression status and mobilization frequency suggests *Tf2* mobilization is subject to additional controls **(Figure 16C)**. Because the nuclear organization of *Tf2*s is disrupted in *abp1* and *set1* mutants, we hypothesized that nuclear clustering of *Tf2*s restricts *Tf2* mobilization. *Tf2*s are spatially organized within the nucleus into *Tf* bodies. *Tf2* clustering is disrupted in *set1* and *abp1* mutants, in which levels of both *Tf2* expression and *Tf2* mobilization are increased. We asked whether mobilization is restricted through a combination of transcriptional repression and *Tf2* clustering. This model predicts that *sre1-N* mutants will decluster *Tf2*s. In contrast, *hip1*Δ strains should retain WT *Tf2* clustering, which would keep mobilization frequencies low despite increased *Tf2* expression. We performed FISH with a *Tf2*-specific Cy3-labelled DNA probe to determine the *Tf2* clustering status in these mutants **(Figure 17A)**. Wild type cells show one or two *Tf2* foci in the nucleus, consistent with prior reports.[75,108,112] We observed a significant increase ($p<0.001$) in the number of *Tf2* foci in *sre1-N* cells, indicating that increased *Tf2* expression in these cells is accompanied by changes in

higher-order chromatin organization that disrupt *Tf* bodies. No significant change in *Tf2* clustering was observed in *hip1*Δ cells, with the majority of cells showing one or two *Tf2* foci **(Figure 17B)**. These results suggest that nuclear organization of *Tf2*s is a regulator of *Tf2* mobilization.

Analysis of *set1* mutants suggests *Tf2* bodies restrict mobilization

To further dissect the roles of transcriptional repression and *Tf2* spatial organization in *Tf2* mobilization control, we performed mobilization assays using separation of function alleles of *set1* **(Figure 18)**. We showed previously that Set1 requires its SET and RRM2 domains to repress *Tf2* expression in an H3K4me-independent manner.[112] In contrast, clustering of *Tf2*s requires H3K4me.[112] We hypothesized that *set1* mutants that are defective for both *Tf2* transcriptional silencing and *Tf2* clustering will have higher mobilization frequencies than mutants with functional silencing or clustering functions. In *Set1-RRM2*Δ*, Tf2* repression is lost but *Tf* body clustering is maintained. Mobilization in *set1-RRM2*Δ was modestly increased **(Figure 18B)**. Next, we assayed for mobilization frequency in $set1F^{H3K4me,}$ which lacks H3K4me and is defective for *Tf2* clustering but maintains normal *Tf2* expression levels. *Tf2* mobilization frequency was modestly increased in $set1F^{H3K4me}$. Combining the $set1F^{H3K4me}$ and *RRM2*Δ mutations further increased the frequency of *Tf2* mobilization events **(Figure 18B)**. Finally, we made use of the *set1-SET*Δ mutant, which is defective for both *Tf2* silencing and clustering. We observed a high mobilization frequency (9-fold over

WT) in this strain. These results indicate that *Tf2* mobilization is subject to multiple levels of regulation that coordinate to control retroelement mobilization.

Figures

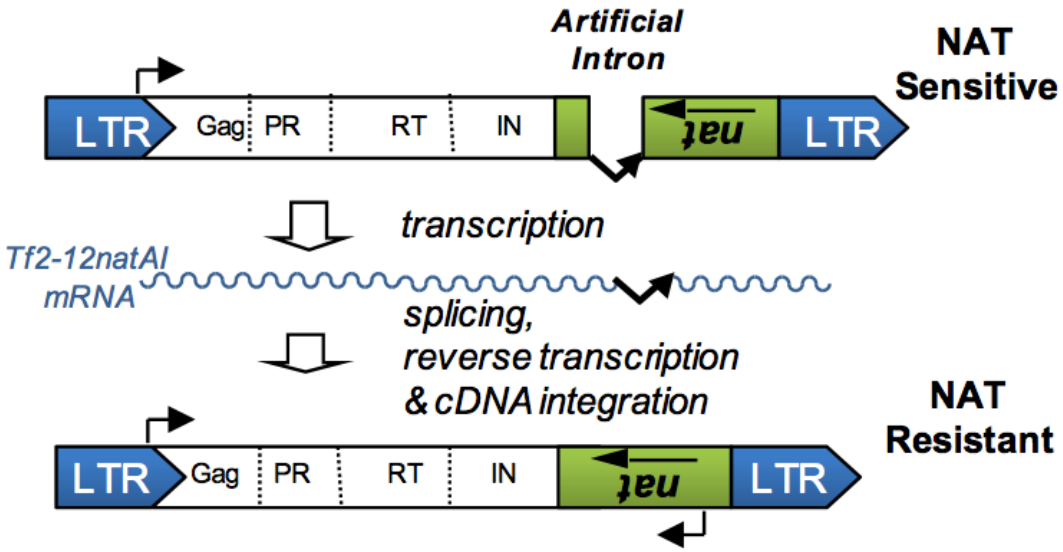


Figure 14. Schematic of the *Tf2-12natAI* mobilization assay. The endogenous *Tf2-12* element was marked with a nourseothricin resistance cassette (*nat)* interrupted with an artificial intron. Mobilization of this element results in the generation of a functional cassette and the acquisition of resistance to nourseothricin.

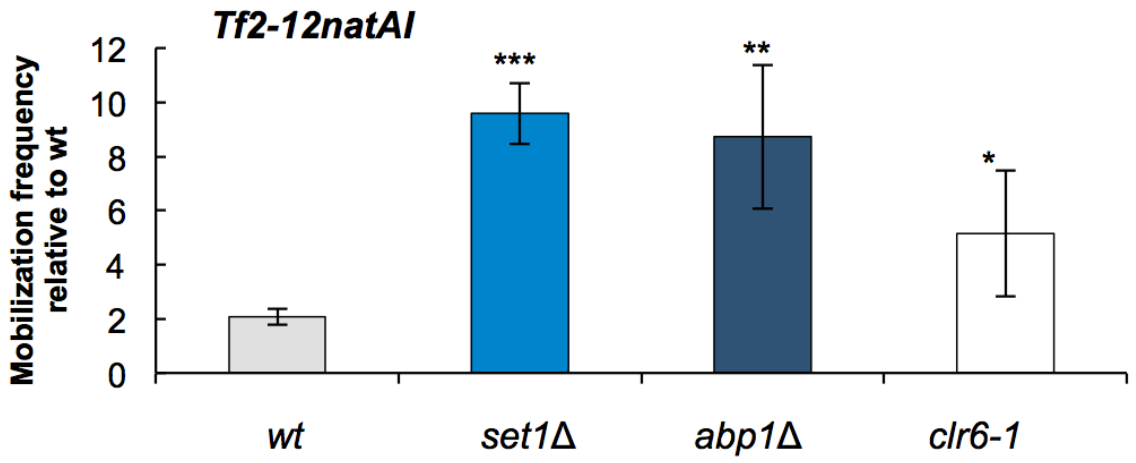


Figure 15. Loss of Set1-mediated silencing induces *Tf2* mobilization. The mobilization frequency of *Tf2-12natAI* in the indicated strain backgrounds was determined by fluctuation analysis as described in materials and methods. Values were scaled relative to the wild type. Error bars indicate ± SEM. *Tf2* mRNA levels in the indicated strains was determined by qRT-PCR. Error bars indicate ± SEM. *** denotes $p < 0.001$, ** $p < 0.01$ and * $p < 0.05$ (t-test).

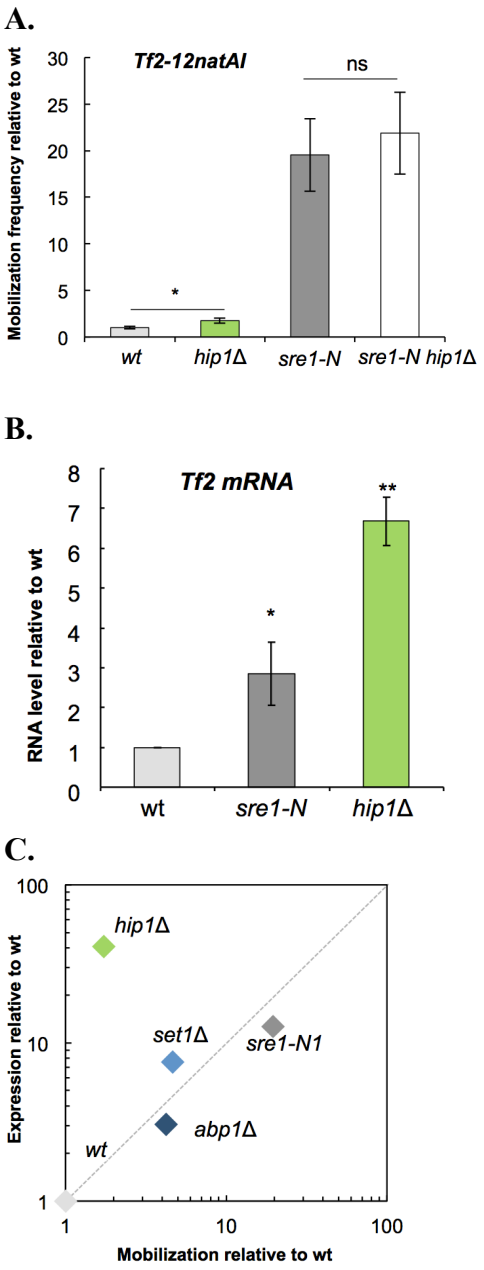


Figure 16. Loss of HIRA-mediated silencing does not result in uncontrolled *Tf2* element mobilization. (A) Deletion of hip1^{+} results in only modest increase in *Tf2* mobilization. The frequency of *Tf2-12natAI* mobilization was determined for the indicated strains by fluctuation analysis using the method of the median. Values were scaled relative to the wild type. Error bars represent ±SEM. Data for sre1-N from is included for comparison (B) Comparison of *Tf2* mRNA levels in *sre1-N* and *hip1*Δ backgrounds. RNA was prepared from the indicated strains and *Tf2* mRNA levels were assayed by qRT-PCR and normalised to act1 mRNA. ** $p < 0.01$, * $p < 0.05$ and ns (not significant) $p > 0.05$ (t-test). (C) Comparison of *Tf2-lacZ* expression with *Tf2-12natAI* mobilization frequency relative to wild type levels in the indicated genetic backgrounds.

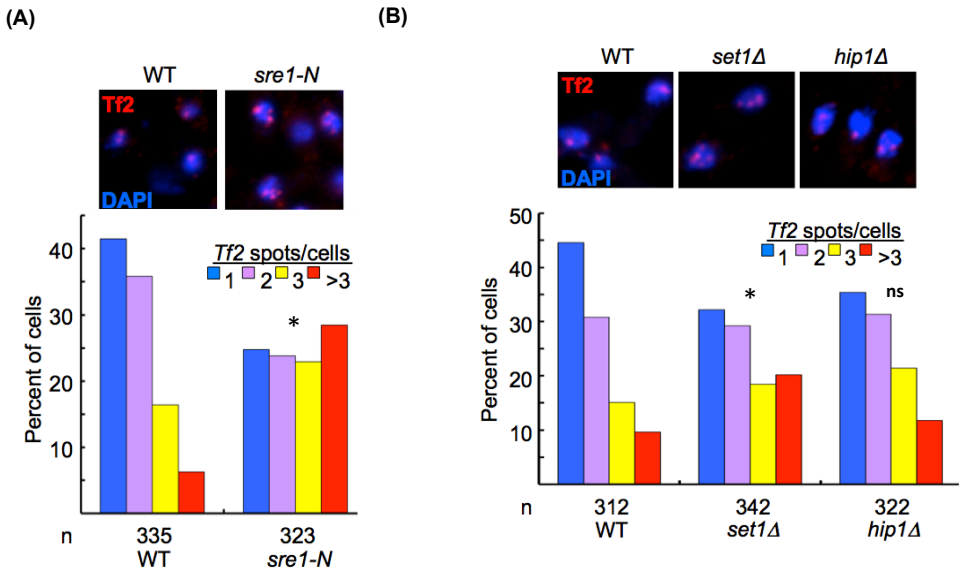
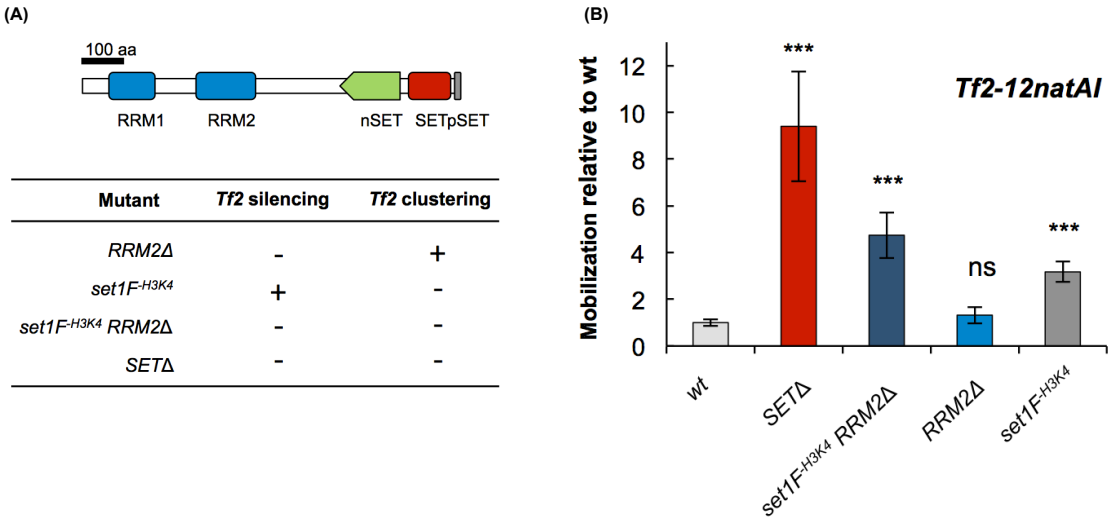


Figure 17. Loss of HIRA does not disrupt *Tf* bodies. (A) FISH (fluorescence in situ hybridization) analysis was performed using a FISH probe corresponding to the ~ 3.6 kb coding region of *Tf2* elements. Representative FISH images from the indicated strains (top panels). Quantitative FISH analysis of observed *Tf2* foci/cell in indicated strains (bar graph; bottom panels). Number of cells analysed per strain (n). (B) Declustering of *Tf2*s assessed by Chi-square test was significant in *sre1-N* and *set1Δ* ($p < 0.001$) but not *hip1Δ* ($p > 0.05$). * denotes $p < 0.001$ and ns (not significant) denotes $p > 0.05$ (χ^2).



Mutant	*Tf2* silencing	*Tf2* clustering
RRM2Δ	-	+
set1F^{-H3K4}	+	-
set1F^{-H3K4} RRM2Δ	-	-
SETΔ	-	-

Figure 18. The transcriptional repression and clustering functions of Set1 suppress *Tf2* mobilization. (A) The top panel shows a schematic of the domain structure of Set1 and the bottom panel a summary of the properties of the set1 mutants. (B) Analysis of *Tf2-12natAI* mobilization frequency was determined in the indicated set1 mutant backgrounds by fluctuation analysis using the method of the median. Values were scaled relative to the wild type. Error bars represent ±SEM. *** denotes $p < 0.001$ and ns (not significant) denotes $p > 0.05$ (t-test).

Chapter four: Histone H3 gene dosage and H3K4 contribute to repression of repetitive elements and protection against genotoxic stress.

Introduction

Appropriate levels of histone expression are critical for transcription, chromosome segregation, repair, and other chromatin-mediated processes.[111] The genes coding for histone proteins are often organized into clusters.[165] *S. pombe* has three copies of histone H3 (*hht1, hht2, hht3*) and H4 (*hhf1, hhf2, hhf3*). These genes are organized in pairs sharing divergent promoters **(Figure 19A)**. In contrast with higher eukaryotes which have multiple H3 variants, the polypeptide products of each fission yeast H3 gene are identical. With the exception of the cell cycle independent gene copy *hht2, hht/hhf* expression is upregulated during S phase to provide increased histone availability for chromatin assembly on newly replicated DNA.[166] To ensure proper chromatin assembly, the production of new histones must be precisely coordinated with new DNA synthesis. Tasked with this coordination are several factors with both activating and repressive effects on expression of the core histone genes. *hht/hhf* expression is restricted to S phase by the histone chaperone HIRA complex.[163] HIRA proteins are highly conserved histone chaperones involved in nucleosome assembly. In budding yeast, the HIRA proteins Hir1 and Hir2 restrict expression of six of the eight core histone genes in a cell cycle dependent manner.[167,168] Additional members of *S. pombe* HIRA complex are Slm9 and Hip1.[156] In addition to its involvement in transcriptional regulation of histone genes, human HIRA is implicated in deposition of H3.3-containing histone octamers.[169] Transcription of histone genes is induced by the GATA-type transcription factor Ams2,[170] which binds to a conserved AAACCT motif within the *hht/hhf* promoter.[166]

We previously observed that heterochromatic silencing is maintained in mutants

in which the lysine 4 at all three H3 genes was mutated to arginine (H3K4R) **(Figure 20)**.[75,112] Interestingly, while mutation of H3K4R has little effect on the repression of the centromeric *dg* repeats, H3K4R in strains lacking two of the histone H3 genes results in dramatic increased *dg* expression **(Figure 19E)**.[113] This result suggests that histone modification and histone gene dosage cooperate to maintain centromeric silencing. We therefore began experiments to characterize cellular mechanisms that could link H3K4 modification with histone gene dosage. Additionally, we asked whether individual H3 copies have unique functional roles that could explain their unique deletion phenotypes at elevated temperatures and in the presence of DNA-damaging agents.[114] Our results indicate H3K4 is necessary to maintain silencing of repetitive elements and to respond to DNA damage in strains with decreased H3 copy number **(Figures 22,24)**. These results indicate that H3K4 and histone dosage cooperate to maintain chromatin function.

Reduction of histone H3 gene copies alters the protein ratio of histone H3 versus H4

To determine if the number of functional H3-coding genes is a factor in maintaining total H3 protein levels, we performed immunoblots for total H3 and H4 levels in strains with single **(Figure 19B-D)** or double deletions of the *hht* genes. These experiments show that cells can maintain their normal ratios of H3/H4 when a single *hht* is deleted **(Figure 21)**. Histone dosage is altered in cells with multiple *hht* deletions.

These data suggest that histone protein level may be subject to feedback regulation that maintains H3 level when the number of H3-coding genes is decreased.

H3K4 is necessary for proper DNA damage response when histone dosage is altered

To determine if proper histone dosage is required for effective DNA repair, we assayed for changes in growth rate in strains grown in media containing methylmethanesulfonate (MMS), camptothecin, or hydroxyurea (HU) **(Figure 22)**. H3K4R strains with double H3 gene deletions had decreased growth rates on MMS-containing media, suggesting that these mutants are unable to effectively mount a DNA repair response. Interestingly, growth rates were normal in double H3 gene mutants with normal H3K4 histones, suggesting that the status of H3K4 is necessary to maintain DNA damage response when histone dosage is decreased.

Reduced histone dosage requires normal H3K4 to maintain silencing of repetitive elements

We then considered how manipulation of H3 level affects the expression of *Tf2s*. We found that while *Tf2* clustering is maintained in H3 deletion mutants, **(Figure 23)** loss of the cell cycle independent H3 gene *hht2* causes significant *Tf2* upregulation **(Figure 24)**. This suggests that some constitutive expression of H3 mRNA is necessary to maintain the silent chromatin at *Tf2s* and pericentromeres.

Expression of H3 and H4-coding genes is upregulated in response to loss of cell-cycle independent H3 gene *hht2*

To address whether H3K4 mediates response to decreased histone dosage by upregulating expression of the core histone genes, we generated primers against unique sequences in the 5'UTR of each H3 and H4 coding gene and performed RT-qPCR to compare H3 and H4 transcript levels in the histone dosage mutants. While H3 and H4 expression were unchanged in *hht1*Δ and *hht3*Δ mutants **(Figure 25A,C)**, *hht2*Δ strains had elevated levels of H3 and H4-coding transcripts, suggesting that regulation of the cell-cycle-dependent H3 copies is linked to the status of *hht2* **(Figure 25B)**.

Transcriptional regulation of core histone genes does not account for ability to respond to decreased histone dosage

Strains lacking two H3 coding genes can effectively respond to MMS-induced DNA damage and maintain silencing of *Tf2*s and pericentromeres in an H3K4-dependent manner. To determine if H3K4-dependent response is mediated through upregulation of the remaining H3 copy, we performed qPCR in H3 double mutants **(Figure 26)**. These data indicate that H3 transcription is not increased in these strains, suggesting that posttranslational factors such as histone deposition, positioning, or regulation of protein degradation may be responsible for maintaining silencing and DNA repair functions in these strains.

Figures

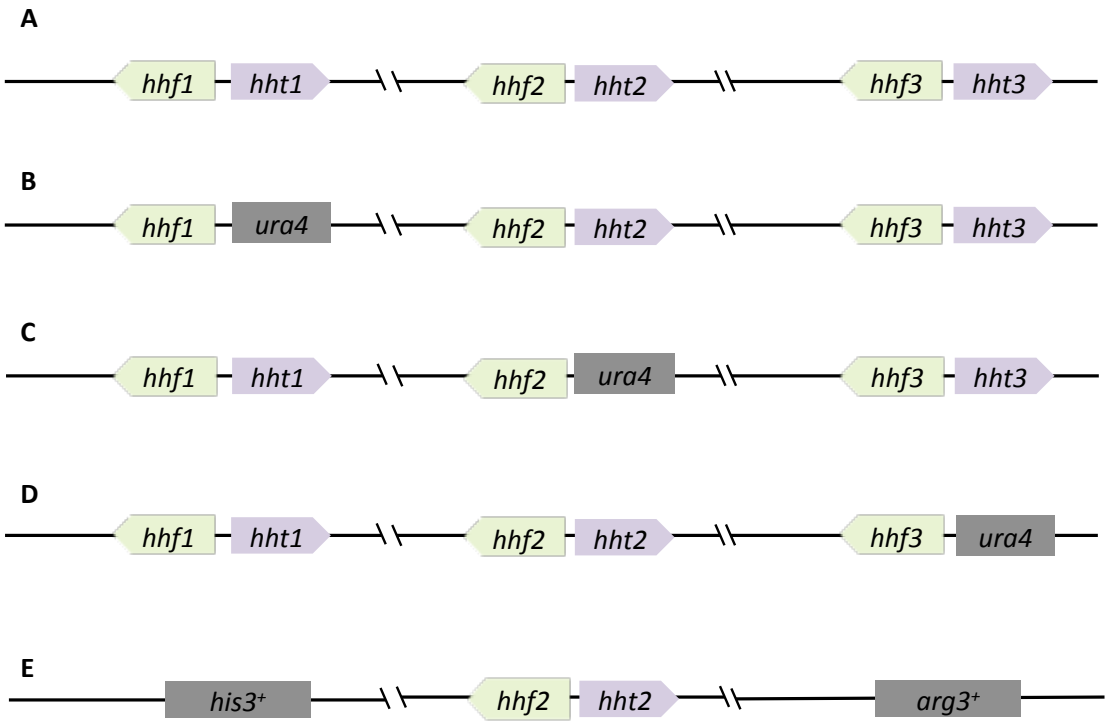


Figure 19. Histone H3 and H4-coding genes in *S. pombe*. Arrangement of H3 and H4-coding genes in *S. pombe.* Each H3/H4 cassette is transcribed divergently from a shared promoter. **A.** Diagram of wild type gene architecture. **B**. Diagram of *hht1*Δ strains. **C.** Diagram of *hht2*Δ strains. **D.** Diagram of *hht3*Δ strains. **E.** Diagram of strains used in Xhemalce et al. *Genes & Development* 2010.[113] The *hht2* lysine four codon was mutagenized to code for arginine. Note full deletion of H3.1/H4.1 and H3.3/H4.3.

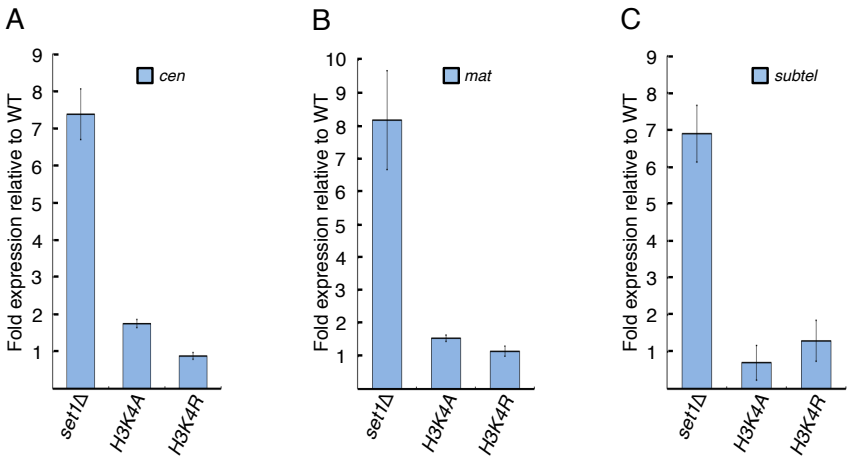


Figure 20. Repression of heterochromatic loci is largely maintained in H3K4 mutants. Expression at (A) pericentromeres, (B) silent mating type locus, and (C) subtelomeres was analyzed using qRT-PCR in indicated mutant strains. Fold changes relative to wildtype were normalized to act1 expression. (s.d., error bars; n=3). Pericentromeric repeat dg (*cen*), silent mating type cenH (*mat*), subtelomeric prl70 (*subtel*).

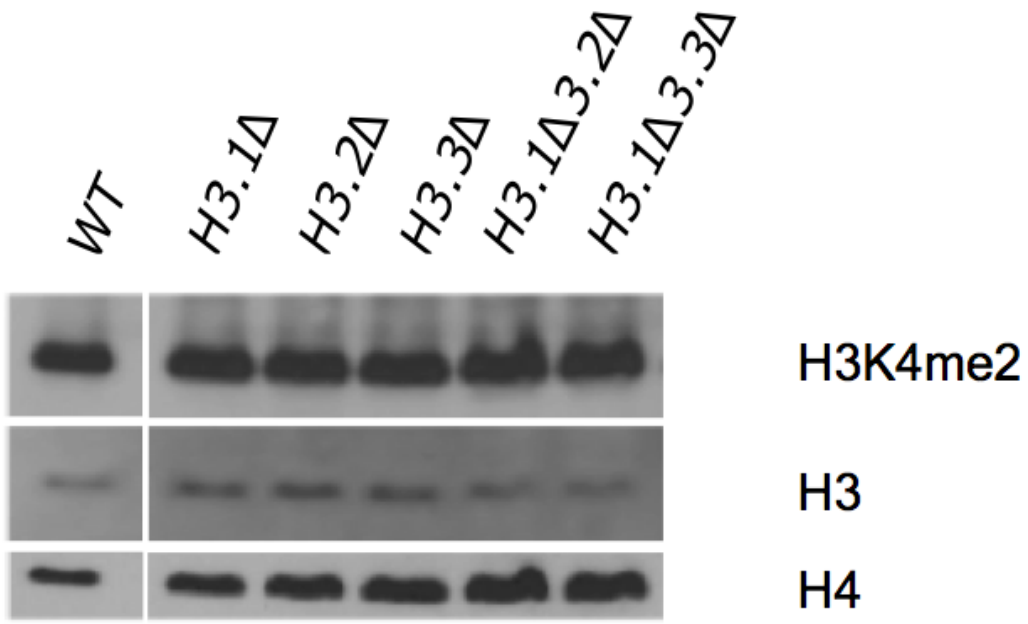


Figure 21. Levels of H3, H4, and H3K4me2 are maintained in histone dosage mutants. Immunoblots on total histone from indicated strains using antibodies for H3, H4, and H3K4me2.

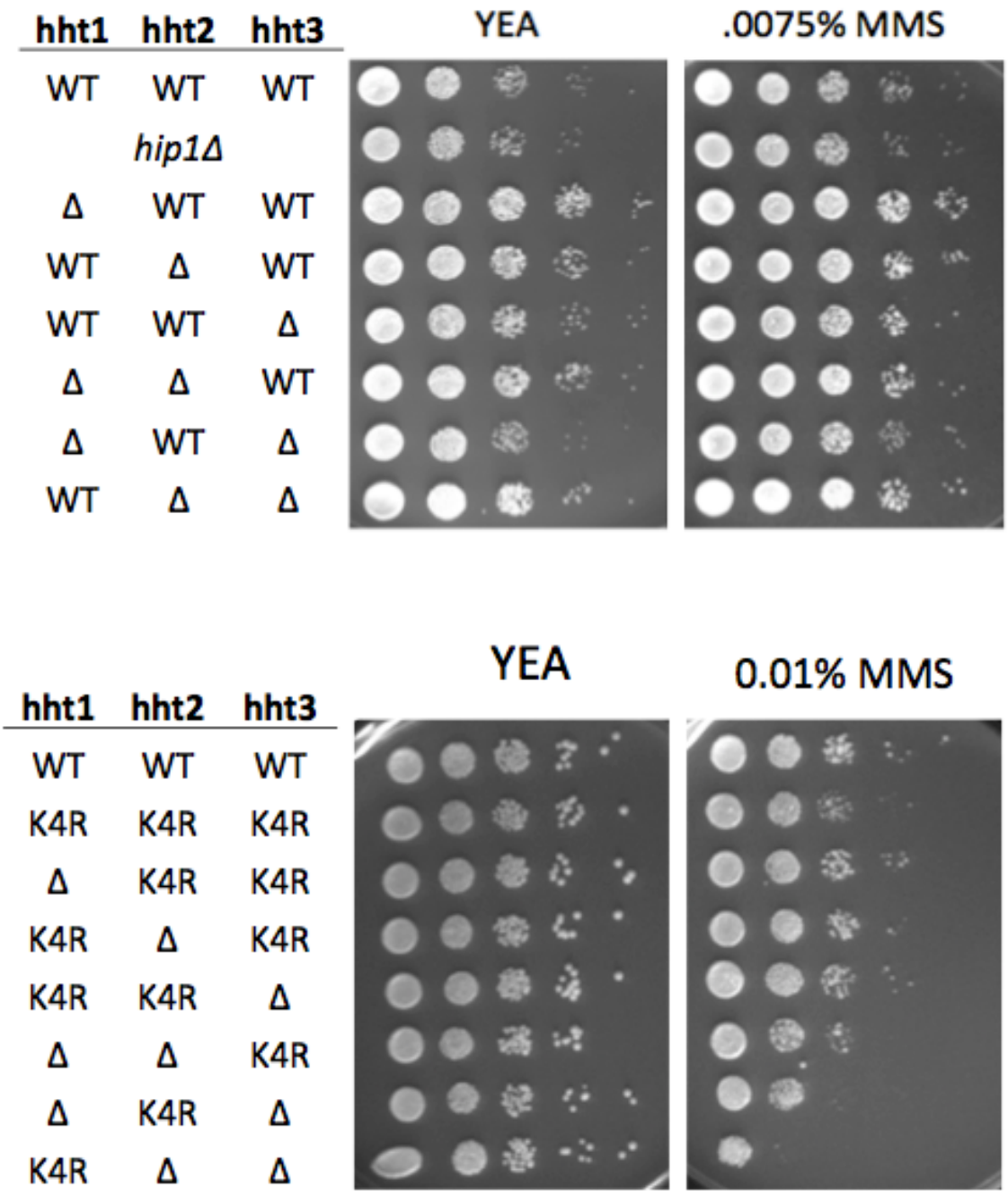


Figure 22. H3K4 is necessary for proper DNA damage response when histone dosage is altered. Serial dilutions of exponential phase cultures of indicated strains on rich media(YEA) and rich media supplemented with MMS. H3 gene status is indicated on the table at left of the image.

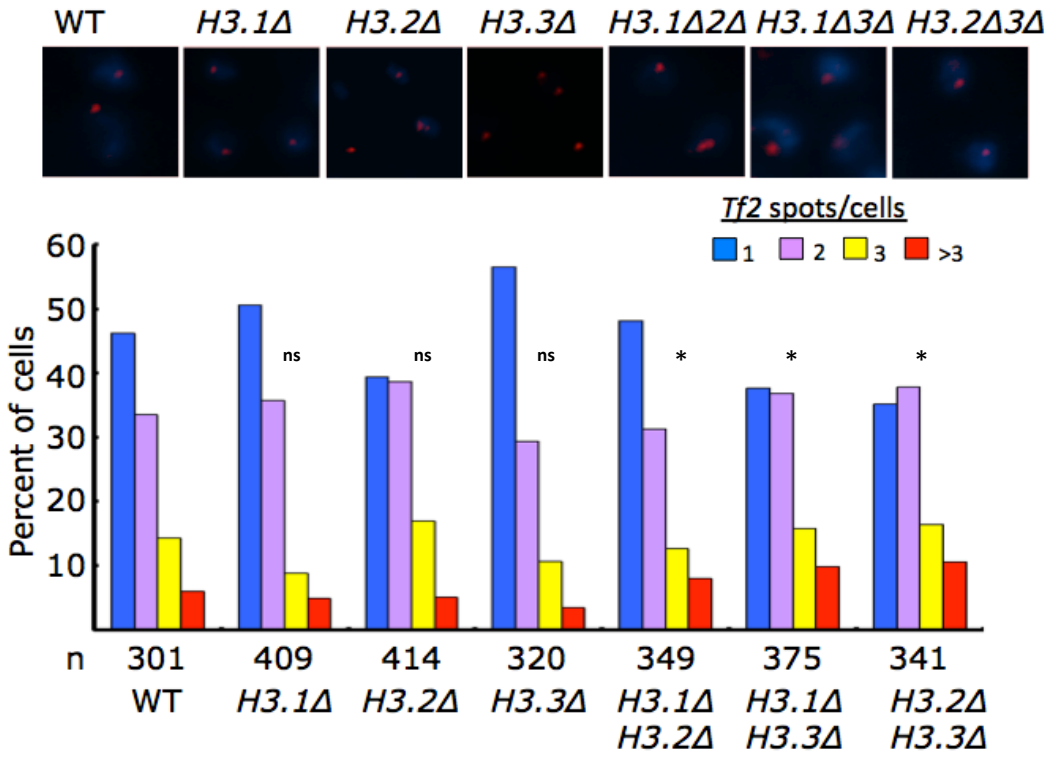


Figure 23. *Tf* body integrity is maintained in single H3 mutants and slightly altered in double H3 mutants. Representative FISH images from indicated strains (top panels). Quantitative FISH analysis of observed *Tf2* foci/cell in indicated strains (bar graph; bottom panels). *Tf2* declustering relative to WT was significant in double H3 mutant strains. ($p<0.005$, chi-square test) H3 gene status is indicated on the table at left of the image. * denotes $p < 0.001$ and ns (not significant) denotes $p > 0.05$ (χ^2).

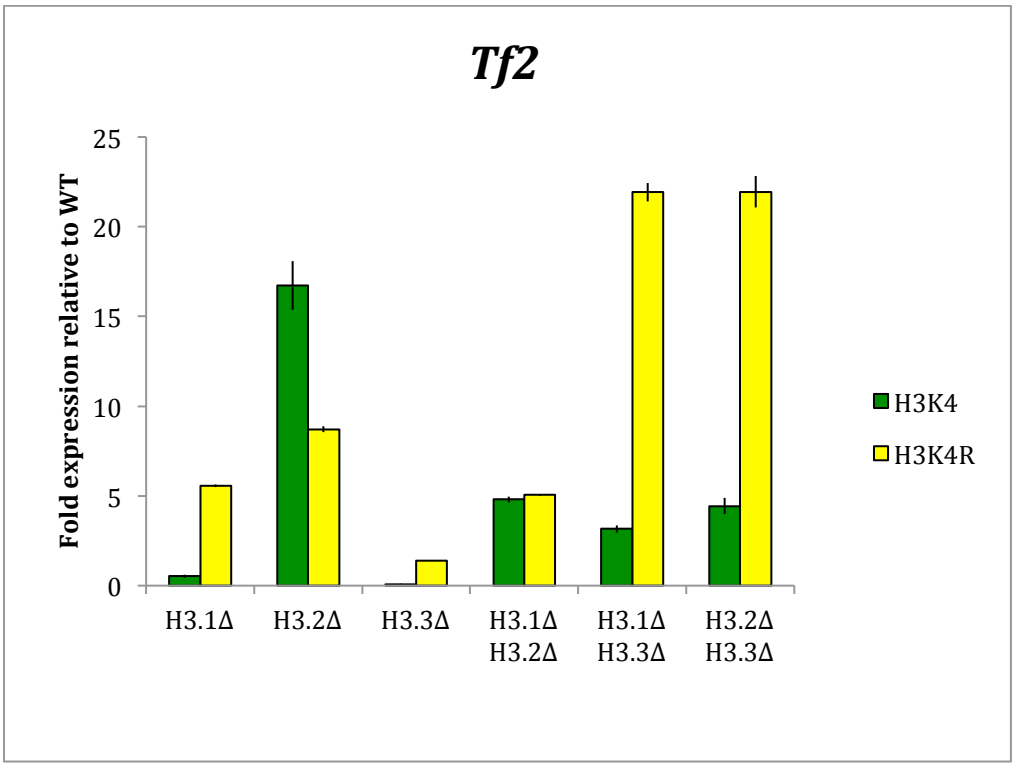


Figure 24A. H3K4 is required for silencing repetitive elements when histone dosage is decreased. *Tf2* Expression was analyzed by qRT-PCR. Fold changes relative to wildtype were normalized by *act1* expression. (s.d., error bars; n = 3 technical RT-qPCR replicates).

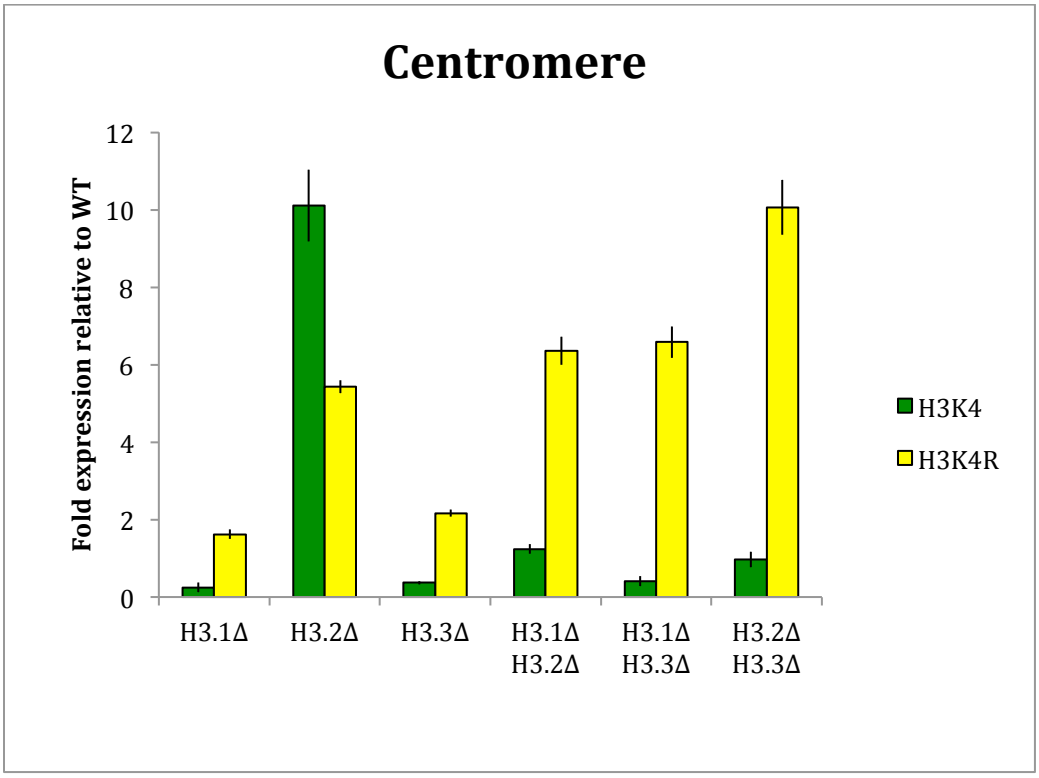


Figure 24B. H3K4 is required for silencing repetitive elements when histone dosage is decreased. Centromere primers represent expression at the *dg* repeats. Expression was analyzed by qRT-PCR. Fold changes relative to wildtype were normalized by *act1* expression. (s.d., error bars; n = 3 technical RT-qPCR replicates).

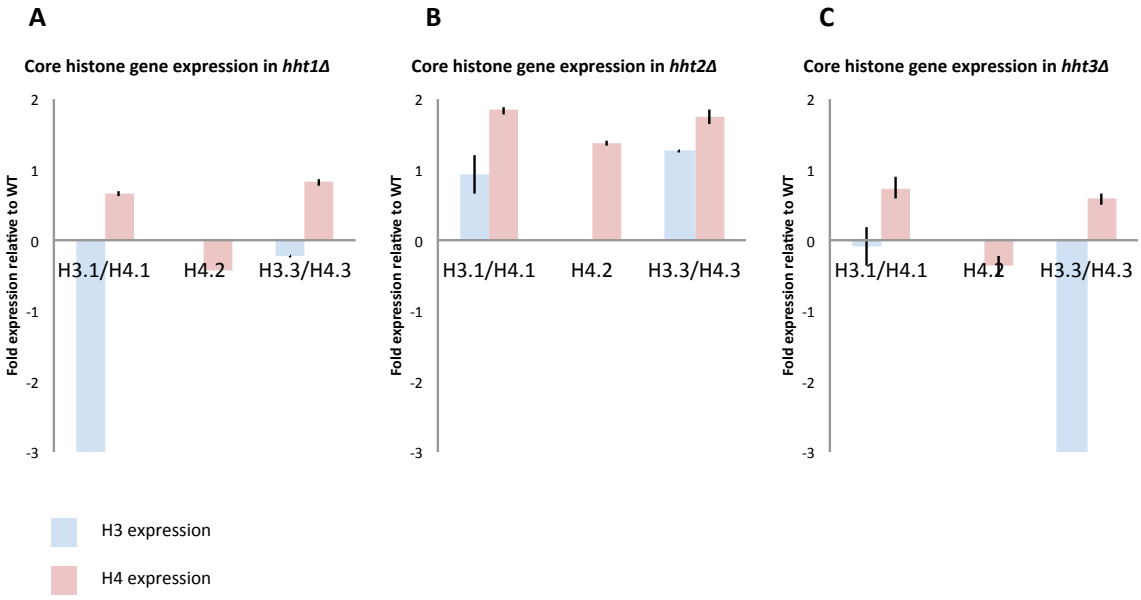


Figure 25. Expression of H3 and H4-coding genes is upregulated in response to loss of cell-cycle independent H3 gene *hht2*. Expression was analyzed by qRT-PCR using primers against unique sequences in the 5'UTRs. Fold changes relative to wildtype were normalized by *act1* expression. Fold enrichment values are log scale. (s.d., error bars; n=3 technical RT-qPCR replicates).

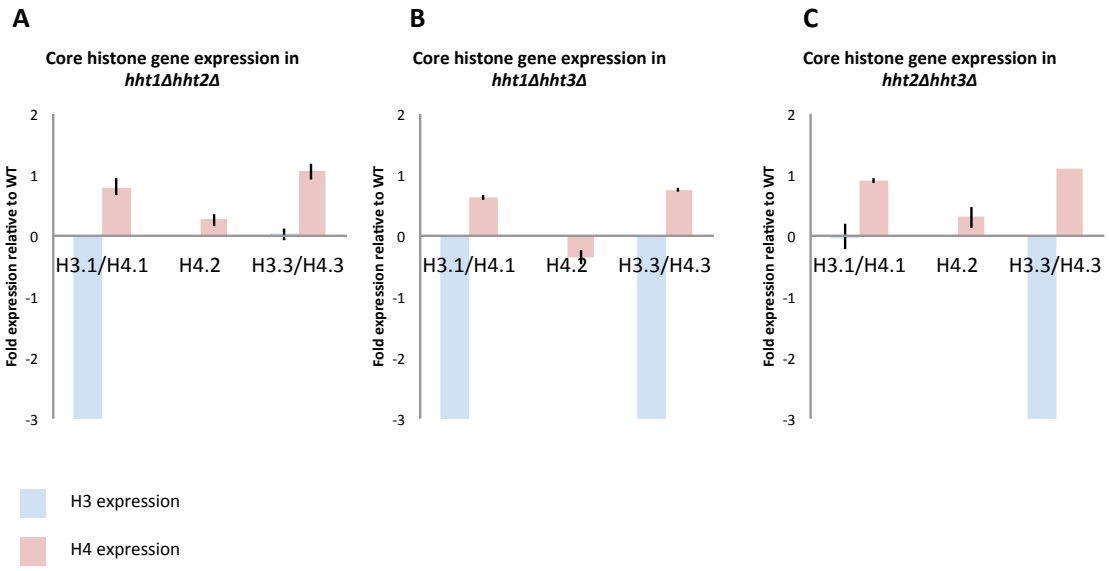


Figure 26. Transcriptional regulation of core histone genes does not account for the ability to respond to decreased histone dosage. Expression was analyzed by qRT-PCR using primers against unique sequences in the 5'UTRs. Fold changes relative to wildtype were normalized by act1 expression. Fold enrichment values are log scale. (s.d., error bars; n=3 technical RT-qPCR replicates).

Chapter Five: Discussion

Summary of Findings

Experiments described in Chapter Two indicate that Set1 and its associated Set1C subunits are surprisingly multifaceted, with roles in histone modification, transcriptional regulation, and nuclear organization. Specifically, we identified roles for the H3K4 methyltransferase Set1 and its associated complex in genome control that are both dependent and independent of H3K4 methylation. Additionally, we identified a novel interplay between methyl and acetyl marks at H3K4 that drives the nuclear organization of retrotransposons. Our study considerably expands the regulatory repertoire of this important histone modifier and highlights the multifaceted function by a well-conserved complex with diverse roles in genome control.

The work described in Chapter three expands our work on the regulation of *Tf2* retrotransposons to the factors regulation of *Tf2* mobilization. We identified cooperation between transcriptional repression and genome organization in the regulation of *Tf2* mobilization. We identified mutations that compromise both transcriptional repression and *Tf* body formation result and determined that mobilization rates in these strains are elevated. Interestingly, loss of silencing in the presence of intact *Tf* bodies is not sufficient to render high levels of mobilization. Our results therefore highlight that that LTR retrotransposon mobility is subject to regulation at multiple levels and indicates that higher order chromatin organization is an important aspect of host cell control.

Experiments described in chapter four describe an unexpected cooperation

between H3K4 and histone dosage in chromatin regulation. Appropriate levels of histone expression are critical for transcription, chromosome segregation, repair, and other chromatin-mediated processes.[111] Our results indicate that histone modification and histone gene dosage cooperate to maintain silencing at *Tf2*s and centromeres. Additionally, H3K4 is required for proper DNA damage response in strains with decreased H3 copy number. These results indicate that H3K4 and histone dosage cooperate to maintain chromatin function and suggest that H3K4 influences posttranscriptional processes that stabilize and maintain chromatin function.

Discussion

Chapter Two: Multifaceted genome control by Set1 dependent and independent of H3K4 methylation and the Set1C/COMPASS complex

This thesis reveals that the chromatin landscape is dynamic and highly regulated. Genome feedback mechanisms are essential for the regulation of the transcriptome and their roles are critical in how cells are able to respond to their external environment.

Set1 and Set1C/COMPASS are highly conserved throughout eukaryotes,[41] and our study reveals additional insights. Early research on Set1C was directed by the observation that actively transcribed regions are enriched in H3K4 methylation. This observation suggested a role for Set1 for the epigenetic maintenance of euchromatic regions. Several further pieces of evidence support this model. First, the characteristic distribution of H3K4me over coding sequences, with the highest levels of H3K4me3 enrichment observed at the promoter, H3K4me2 throughout the open reading frame, and H3K4me1 towards the 3' end of the ORF. Second, Set1 association with actively transcribing Pol II suggests a model for Set1 as an epigenetic factor in which repeated cycles of transcription allow H3K4me to accumulate at regions that are frequently transcribed.[47] Through positive feedback, H3K4me can then recruit or serve as a binding site for factors that further contribute to productive transcription. Indeed, experiments have shown that loci that have previously been activated in response to stress conditions

activate more quickly when the stress condition is experienced a second time.[171] Our studies argue against the model that H3K4me is simply an "activating mark". Instead, we find that H3K4me is often associated with repression, has a range of locus-dependent effects that extend well beyond transcriptional control.

An emerging body of evidence suggests that its role at actively transcribed regions is only one facet of Set1 activities as a genome regulator. Notably, strains lacking H3K4me do not suffer severe phenotypic effects and genome wide analysis of the transcriptome in these strains reveals that many loci are upregulated, suggesting that Set1 is capable of exerting repressive control. Numerous studies have described roles for Set1 in transcriptional repression.[19,24,45,72,73,76,172,173] Our study expands this notion. We show that Set1 requirement for H3K4me to mediate repression is locus-dependent. We show that cells lacking Set1 and H3K4me can maintain transcriptional silencing of *Tf2s* and pericentromeres by means of a methyltransferase deficient Set1 allele (*Set1F*$^{-H3K4me}$), demonstrating that Set1 represses transcription through multiple locus-dependent pathways. In contrast to studies which have ascribed the repressive role of Set1 to H3K4me2 or H3K4me3,[19,74,76,174] we show that repression of *Tf2*s and pericentromeric *dg* repeats is H3K4me-independent and repression of *mat* and subtelomeric *prl70* requires H3K4me. Our experiments using SetC/COMPASS subunit deletions support this model, with defects in *mat* and *prl70* silencing observed in mutants that lack Set1C components important for H3K4me.

The mechanism of Set1-independent silencing has several interesting features **(Figure 27)**. This silencing pathway does not require an intact Set1C complex and is dependent on the action of the RRM2 and SET domains. This suggests that Set1 may mediate centromeric and *Tf2* silencing through a non-H3K4 substrate. Budding yeast Set1 methylates the kinetochore protein Dam1, the first identified case of a chromatin modifier that additionally modifies a non-histone substrate.[126,175] Further experiments on fission yeast Set1 should seek to identify whether a non-histone Set1 substrate exists. If functional conservation of non-histone Set1 targets exists between budding and fission yeast, centromere-associated proteins may be good candidates to explore in *S. pombe*.

Further work is also necessary to characterize the repressive role for the putative RNA binding domain RRM2 of Set1. Because of its multi-domain structure, Set1 may interact with and integrate inputs from multiple factors. RRM2 could facilitate Set1 binding to mRNA or ncRNA. Additionally, the sequence or higher order structure of RNAs bound by the two RRM domains of Set1 could contribute to the allosteric regulation of Set1 targets. Indeed, structural studies on human heterogenous nuclear riboprotein L (hnrbpL) suggest its locus-dependent dual action as an activator or repressor of alternate splicing is dependent on RNA looping mediated the multiple RRM domains of hnrbpL.[176] Structural analysis of RNA binding by the large subunit of human U2 snRNP auxiliary factor suggests that the relative conformation of the RRM1 and RRM2 domains functions as a molecular rheostat that changes conformation based on the nucleotide composition of bound RNA.[177] Further experimentation to identify non-

histone Set1 targets and characterize RRM-RNA interaction will provide mechanistic insight into the multifaceted roles of Set1.

Consistent with results from budding yeast, we show that the nSET, SET, and pSET domains of Set1 are required for all three forms of H3K4me.[52,118] We observed defects in H3K4me2 in the RRM1 mutant, consistent with published results from budding yeast.[48] H3K4me3 was abolished and H3K4me1 reduced in the RRM1 mutant. It is unclear to what degree these changes are a result of reduced Set1 protein levels. By contrast, all three H3K4me forms were slightly increased in *set1* mutants lacking the RRM2 domain. This domain may have an inhibitory role in regulating H3K4me, comparable to experiments in budding yeast that suggest an autoinhibitory role for RRM2.[118]

Several notable differences in H3K4 methylation distinguish the fission yeast Set1C from its counterpart in budding yeast.[50,51,118,119,121] While data from budding yeast implicate Spp1 in control of H3K4me3, [51,117,121] fission yeast *spp1* is essential for all three H3K4 methylation states. This suggests a greater role for Spp1 in maintaining methylation by the COMPASS complex. Whether Spp1 can mediate stabilization of Set1 methyltransferase ability or loss of *spp1* results in destabilization of the COMPASS complex remains to be determined. In contrast to *S. cerevisiae,*[24,42,44] *S. pombe swd2* is not essential. We observed loss of all three H3K4me states in *swd2*Δ deletion strains. *S. cerevisiae* temperature sensitive mutants and *swd2*Δ cells overexpressing a C-terminal Sen1 fragment to suppress the lethality of *swd2*Δ showed reduced levels of H3K4me2

and H3K4me3.[178-180] Reduction of H3K4me2 and H3K4me3 in both species may reflect the requirement of Swd2 to maintain Set1 protein levels.[178,180] Fission yeast Swd2 may be a more dedicated player in Set1C activities compared to budding yeast Swd2, which is also part of the transcription termination factor APT complex and is thought to compete with Set1 for Swd2.[181] Although there is discrepancy between findings from different groups regarding which H3K4me states were defective in budding yeast *sdc1*Δ and *ash2*Δ,[50,51,124,126,180] these data nevertheless support the proposal that Sdc1 and Ash2 form a heterodimer that associates with Set1C.[50,51] Our results indicate that Ash2 has a more extensive role in maintaining H3K4me2 and H3K4me3 than Sdc1 in *S. pombe*, which may reflect divergence in how Ash2 and Sdc1associate with other chromatin modifiers including the H3K4 demethylase Lid2 complex not present in budding yeast.[117]

Our data support an active role for Set1C/COMPASS in maintaining the nuclear organization of *Tf2s* by antagonizing Mst1-mediated H3K4ac. Although *Tf2* clustering is maintained in H3K4A and H3K4R strains,[75] this does not preclude Set1 from acting on *Tf2* organization because these strains lack both H3K4me and H3K4ac. In contrast to *Tf2* FISH experiments in which defects in methyltransferase catalysis caused *Tf2* declustering, we observed tighter clustering in experiments with temperature sensitive *mst1* mutant strains, making Mst1 the first negative regulator of *Tf2* clustering that we have identified. H3K4me can limit the distribution of H3K4ac upstream of transcription start sites in *S. cerevisiae,*[40] suggesting that H3K4me could compete with H3K4ac at *Tf2s* to mediate their genome organization. The idea that both positive and negative regulators control *Tf2* nuclear organization has interesting ramifications. It suggests the possible

existence of mechanisms to actively decluster *Tf2s* by upregulating H3K4ac levels. Indeed, competition between histone modifications has been identified between H3K4me and H3K9me, and between H3K9me and H3K9ac. These mechanisms may be paradigmatic of other internucleosomal modifications and may drive actively declustering and mobilization of *Tf2s* in response to hypoxia or other stress conditions **(Fig S7)**. If this paradigm applies to the control of H3K4me, modifications to neighboring residues such as dimethylation of H3R2 or phosphorylation of H3T3 could function as regulators of H3K4me. In fact, five of the seven lysines on the H3 tail and four of the six H2B N-terminal tail lysines are directly adjacent to one or more residues that are potential substrates for addition of methyl or phosphate groups. Phylogenetic comparison of histone sequences has demonstrated that histones have become enriched in arginines and lysines over time.[182] Similarly, bioinformatic analysis of histone polypeptide sequences could clarify if arginines, threonines, and serines are overrepresented adjacent to histone lysines throughout eukaryotes. Given that *set1*Δ also increases H3K4ac at *act1,* it is possible that changes in *Tf2* clustering in these strains could result from cellular response to elevated levels of H3K4ac genome-wide. *Tf2* organization is affected by the loss of HDACs which are recruited to *Tf2s* by CENP-Bs,[108] suggesting that dynamic interaction between the Set1, HATs, and HDACs could drive changes in genome organization.

Chapter Three: Restriction of retrotransposon mobilization by transcriptional silencing and higher-order chromatin organization

Our results indicate that cellular controls of the *Tf2* life cycle operate at multiple levels and suggest that transcriptional silencing and higher order chromatin organization cooperate to restrict *Tf2* mobility. The fission yeast nucleus is divided into distinct chromosome territories and the subnuclear organization of the genome is an area of growing interest to the field. In fission yeast, several factors cooperate to cluster *Tf2s* to one another and to the centromere, and the centromere to the nuclear periphery.[107] Similarly, the HIV-1 provirus associates with centromeric heterochromatin in infected T lymphocytes.[183] Transcriptional induction of the HIV-1 provirus results in dissociation from centromeric heterochromatin. The HIV-1 provirus has also been shown to associate with PML bodies during viral latency. Dissociation from PML bodies is a requirement for viral expression.[184] However, *Tf2* spatial organization does not necessary correlate with *Tf2* expression status. *Tf2*s in *set1F*$^{-H3K4me}$ strains are transcriptionally silent despite being defective for *Tf2* clustering. Loss of the Nonhomologous end joining (NHEJ) protein Ku (*pku70*Δ or *pku80*Δ) dissociates *Tf2s* from the centromeres and nuclear periphery but does not induce their expression.[107] It is therefore likely that nuclear organization represses mobilization at another point of the retrotransposon lifecycle. Given that *Tf2s* preferentially mobilize via homologous recombination,[159] *Tf2* clustering may act to prevent LTR-LTR recombination events. Although local LTR concentration increases when *Tf2s* cluster, cluster-associated factors could prevent integration via recombination. Indeed, evidence from budding yeast indicates that clustering of Ty1

retrotransposons decreases their recombination via compaction of Ty1 chromatin.[185] Recombination repression by Ty1 is transmissible to adjacent nuclear regions.[185] Additionally, the fission yeast double mutant *abp1*Δ*cbh1*Δ strain which exhibits *Tf2* declustering has elevated levels of Rad22-enriched recombination foci.[186] It is also possible that *Tf2* nuclear organization restricts mobilization at other steps of *Tf2* lifecycle such as reverse transcription, RNA export or RNA processing. Indeed, RNA processing can be influenced by chromatin structure[187] and has been linked with subnuclear structures.[188]

Retrotransposons are widespread in eukaryotes. An ongoing coevolution between transposons and their hosts has been postulated. Nuclear organization of *Tf2s* may reflect an important adaptation by the host to restrict retroelement spread. It would be interesting to explore whether nuclear organization is a conserved retroelement control mechanism. Due to the wide range of total genome content comprised of retrotransposons in diverse organisms, it may be expected that restriction of retroelement mobilization by cellular controls has been successful to varying degrees.

The regulation of retroelement mobilization has evolutionary implications. Given that retroelement mobilization can cause changes at the genomic and transcriptomic levels, upregulation of TE mobilization during stress may provide a reservoir of genetic and epigenetic variability.[136,146,189] At the population level, transposon-mediated genetic variability has been implicated in driving speciation in spatially isolated populations.[190]

Chapter four: Histone H3 gene dosage and H3K4 contribute to repression of repetitive elements and protection against genotoxic stress

Histone levels must be kept in proper balance to ensure proper chromatin function. Overexpression of histones results in diverse defects in silencing, chromatin structure, enhanced sensitivity to DNA damage, and cytotoxicity.[191,192] Additionally, alteration of the H3/H4 ratio can displace CENP-A from the centromeric central core, causing defects in kinetochore assembly and chromosome segregation.[20] Appropriate levels of histone expression are critical for transcription, chromosome segregation, repair, and other chromatin-mediated processes.[111] The work described in chapter four suggests that *S. pombe* is amenable to some degree to changes in histone dosage. However, strains lacking H3K4 are unable to fully repair DNA when growing in mutagenic conditions and lose silencing of repetitive elements in strains with altered histone dosage. This suggests that cellular controls exist that can detect the level of available histone and respond to maintain proper transcriptional silencing. Further experiments are needed to characterize these mechanisms, but our experiments suggest that these processes are likely posttranscriptional, operate during S phase, and require H3K4. To clarify whether unmodified H3K4 or one of its methylated forms are necessary for these controls, we generated double H3 mutant strains with a methyltransferase deficient *set1* allele. Tests for growth rate on MMS-containing media and qPCR analysis of repetitive loci in these strains may shed light on whether methylated H3K4 mediates cellular response to altered H3 dosage.

There are several potential mechanisms for how H3K4 histones could aid in transcriptional repression of repetitive elements under conditions of altered histone dosage. H3K4 could interact with histone chaperones responsible for deposition of histones on newly synthesized DNA and influence nucleosomal distribution to maintain a critical concentration of nucleosomes at repetitive regions so that silencing is maintained. To address this possibility, H3 ChIP arrays or MNase-seq in double H3 mutant strains with and without H3K4 may show significant differences in nucleosome distribution in silenced regions. Because decreasing H3 levels allows for CENP-A chromatin to spread beyond the central core,[20] H3 ChIP experiments could also address whether altered distribution of H3 levels at the centromeres could explain the loss of silencing at the pericentromeric repeats observed in these strains. Additionally, it is possible that cells can maintain the soluble histone pool when histone gene transcription is decreased by downregulating ubiquitin-mediated histone degradation or by regulation of mRNA processing or decay.

Figures

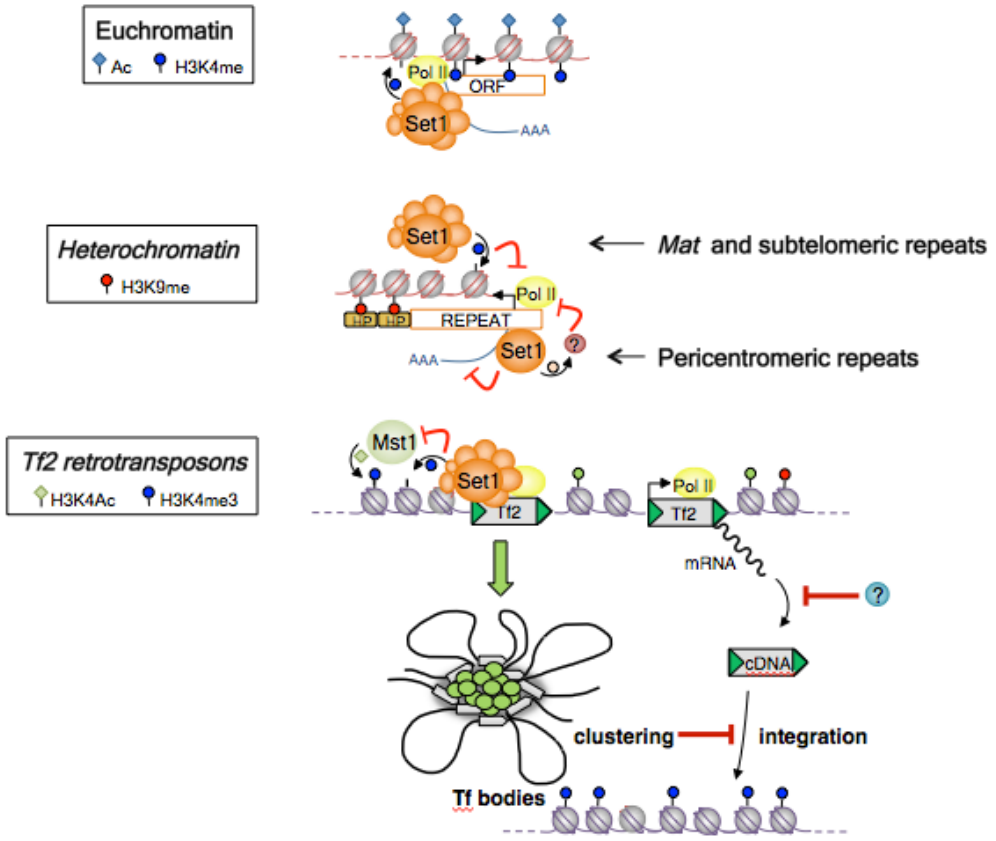


Figure 27. Model for the roles of Set1C in genome control in *S. pombe*. Set1 exerts its multifaceted genome control at euchromatin and heterochromatin dependent and independent of H3K4 methylation and the Set1C complex. At euchromatin, Set1 operates as part of the Set1C/COMPASS complex that is responsible for H3K4me distribution at active RNA polymerase II genes. At interspersed *Tf2s* and heterochromatic loci, Set1 has a repressive role mediated through two distinct pathways: H3K4me/Set1C-dependent repression at the silent mat locus and subtelomeres, and H3K4me/Set1C-independent repression at *Tf2s* and pericentromeric heterochromatin. In addition, Set1C and H3K4me have a distinct nuclear organization role at *Tf2s* by antagonizing the activity of the histone H3K4 acetyltransferase Mst1 to maintain the integrity of *Tf* bodies. Silencing factors such as Set1, Abp1, HIRA and HDACs inhibit *Tf2* transcription and thereby limit cDNA accumulation. The clustering of dispersed *Tf2* elements into *Tf* bodies limits mobilization by restricting cDNA integration by homologous recombination.

Appendix A: Supplementary figures

Data in the supplementary figures was generated by the following researchers:

Figure S1: Irina Mikheyeva
Figure S2: Irina Mikheyeva
Figure S3: Hugh Cam
Figure S4: Fiona Tamburini
Figure S5: Heather Murtona & Simon Whitehall
Figure S6: Heather Murtona & Simon Whitehall
Figure S7: Heather Murtona & Simon Whitehall
Figure S8: Heather Murtona & Simon Whitehall
Figure S9: Heather Murtona & Simon Whitehall

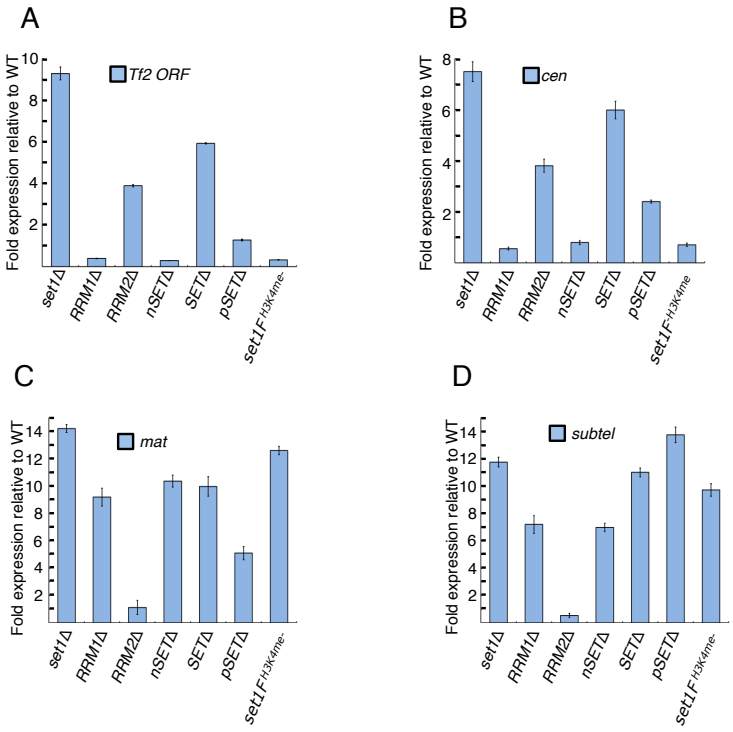


Figure S1. Set1 represses different classes of repetitive elements via distinct functional domains. (A) *Tf2* repression requires intact RRM2 and SET domains. (B) Set1 represses centromeric dg repeats independent of H3K4me. (C,D) H3K4me-dependent repression of mating type locus and subtelomere. Expression of *Tf2* ORF and heterochromatic loci was analyzed by qRT-PCR. Fold changes relative to wildtype were normalized by act1 expression. (s.d., error bars; n = 3 technical RT-qPCR replicates). Pericentromeric repeat *dg* (*cen*), silent mating type cenH (*mat*), subtelomeric prl70 (*subtel*). Experiments performed by Irina Mikheyeva.

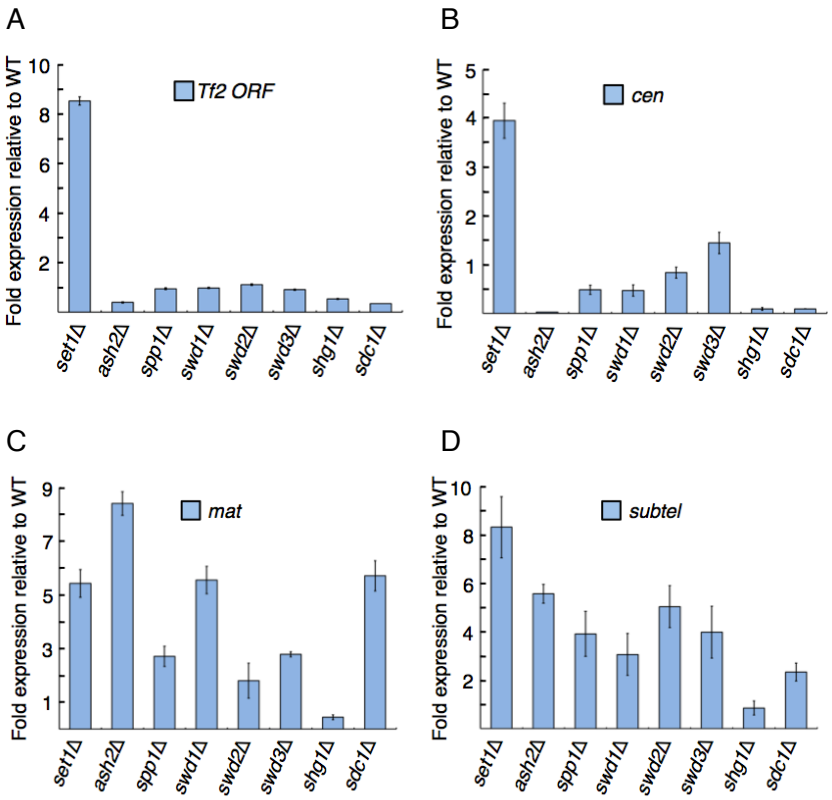


Figure S2. Set1 represses *Tf2* and centromere independent of Set1C/COMPASS. (A-B) Set1-mediated repression of *Tf2* and centromere is minimally affected by loss of other Set1C/COMPASS subunits (C-D) Set1 repression of mating type locus and subtelomere is dependent on Set1C/COMPASS. Expression of *Tf2* ORF and heterochromatic loci was analyzed by qRT-PCR. Fold changes relative to wildtype were normalized by act1 expression. (s.d., error bars; n = 3 technical qPCR replicates). Pericentromeric repeat dg (*cen*), silent mating type cenH (*mat*), subtelomeric prl70 (*subtel*). Experiments performed by Irina Mikheyeva.

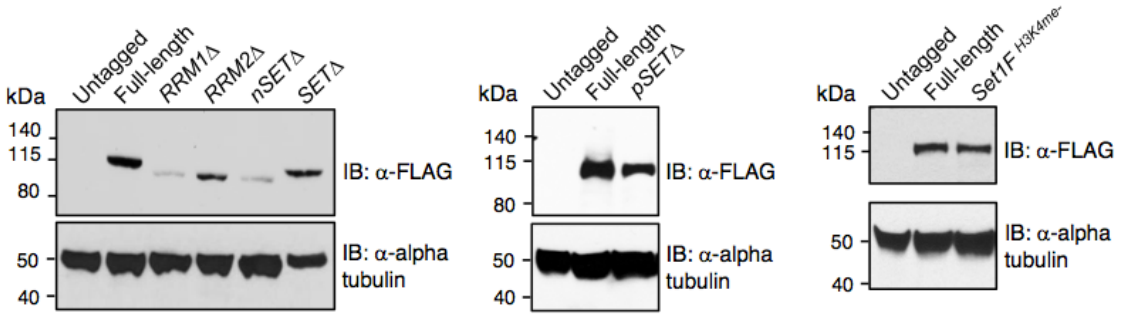


Figure S3. Various domains of Set1 have varied contributions to protein stability. Protein levels of Set1 in various domain deletion mutants. Except for set1FH3K4me-, Set1 proteins from strains expressing an N-terminal FLAG epitope attached to either full-length or domain deletion of set1 were detected by immunoblotting (IB) with an anti-FLAG antibody. Alpha tubulin (loading control) was detected by anti- tubulin antibody (tat-1). Experiments performed by Hugh Cam.

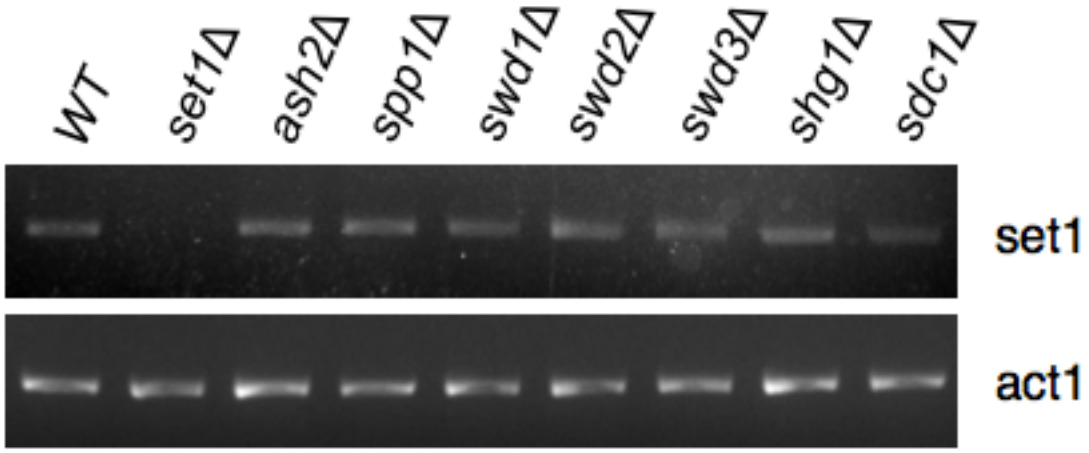


Figure S4. Set1 transcript levels are not noticeably altered in strains deficient for Set1C components. RNA extracted from indicated Set1C mutant strains was converted into cDNA and used in a PCR reaction to assess RNA levels of set1 and the actin gene act1 (control) in corresponding Set1C mutants. Experiments performed by Fiona Tamburini.

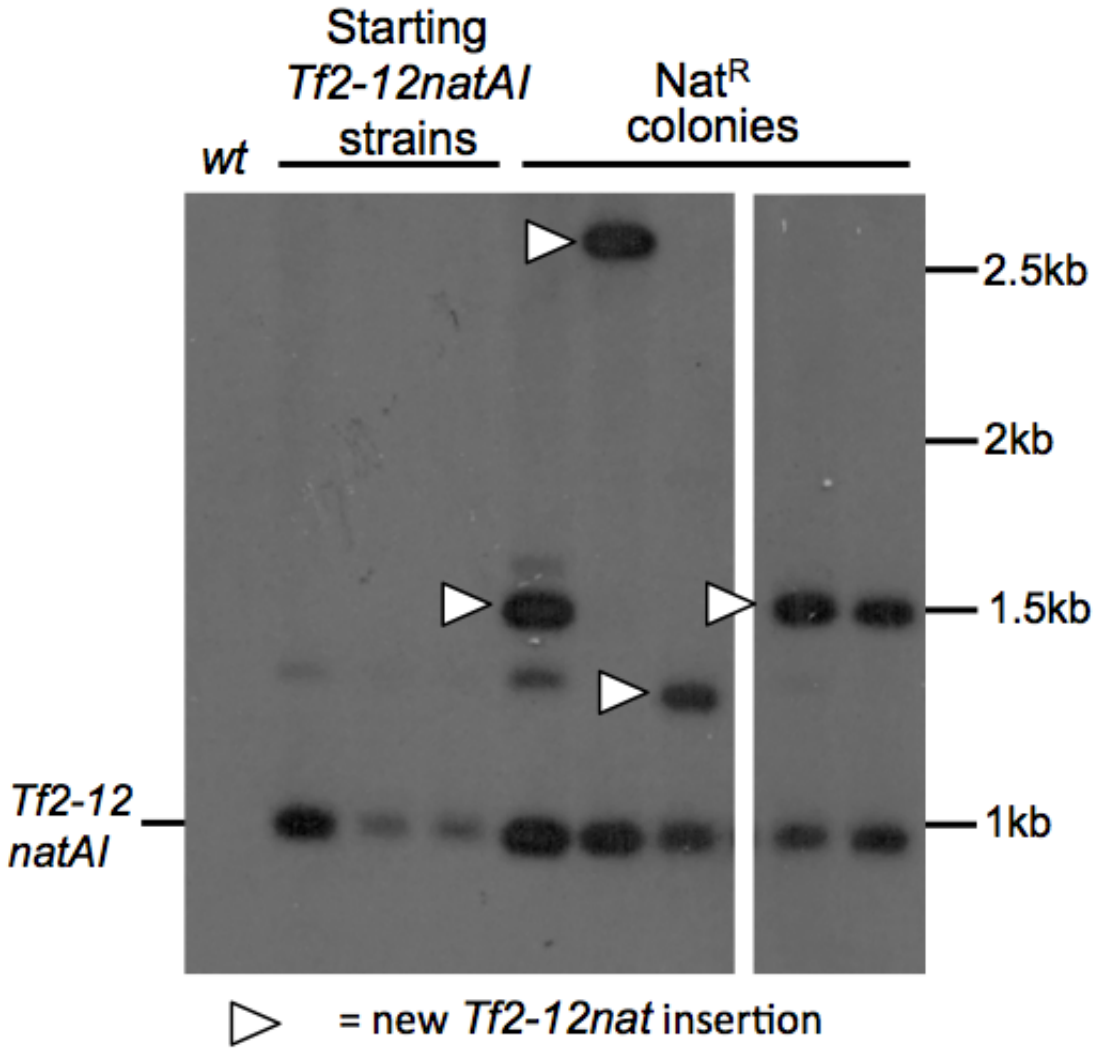


Figure S5. Nat resistance (NatR) arises as a result of *Tf2-12natAI* mobilization. Genomic DNA, isolated from a wild type (untagged) strain, starting *Tf2-natAI* strains and from nourseothricin resistant colonies was analysed by Southern blotting with a probe specific to the *natAI* cassette. Experiments performed by Heather Murtona and Simon Whitehall.

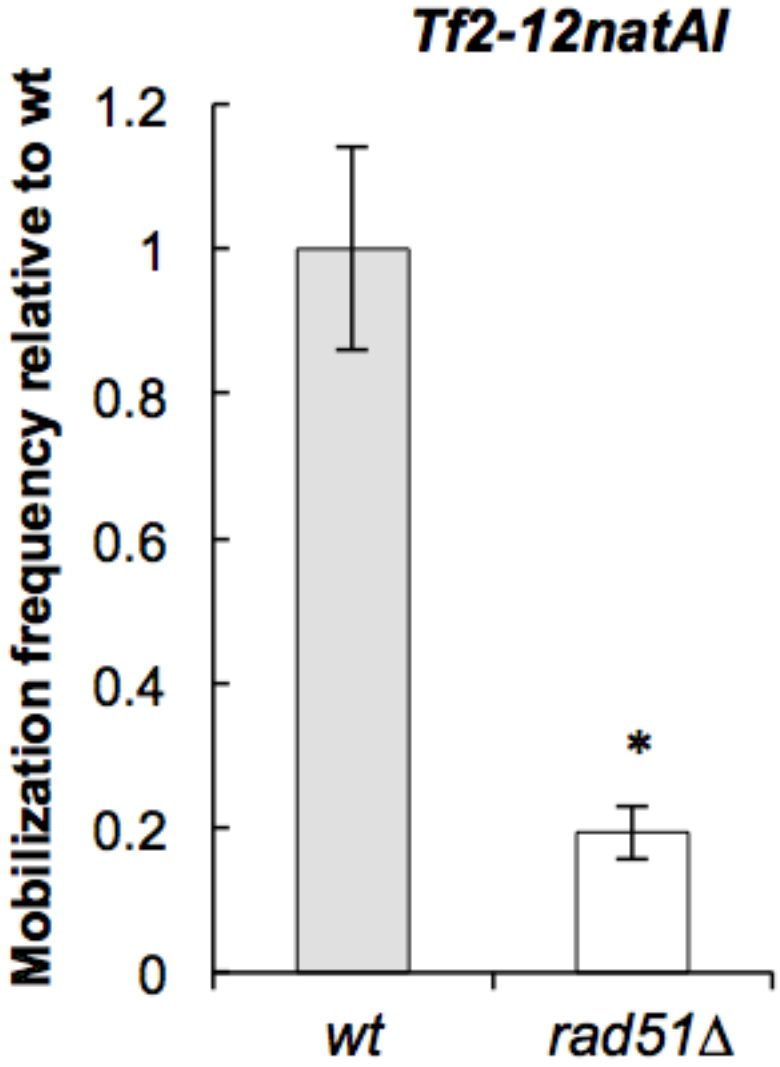


Figure S6. Mobilization is impaired by inactivation homologous recombination. The frequency of *Tf2-12natAI* mobilization was determined for the wild type and *rad51*Δ strains by fluctuation analysis using the method of the median. Values were scaled relative to the wild type. Error bars represent ±SEM. * denotes $p < 0.05$ (t-test). Experiments performed by Heather Murtona and Simon Whitehall.

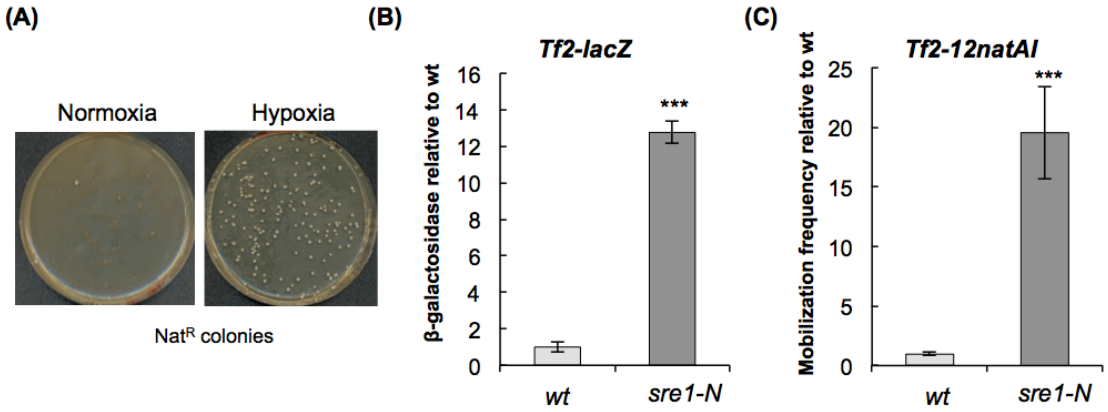


Figure S7. Constitutive activation of the SREBP homolog Sre1 results in high levels of *Tf2* mobilization. (A) Cells were patched onto YES plates and incubated at 30°C for 2 days either under normal oxygen conditions or in an anaerobe jar. Cells were then resuspended in H2O and approximately 1 × 108 cells spread onto YES plates supplemented with Nat. Plates were incubated at 30°C under normal oxygen conditions under colonies appeared. (B) The indicated strains were grown to mid log growth phase at 30°C in YES. Cells were harvested and processed for β-galactosidase assays. Results are the mean of at least 3 independent assays and are scaled relative to the wild type value. Error bars indicate ± SEM. (C) The mobilization frequency of *Tf2-12natAI* was determined by fluctuation analysis as described in materials and methods. Values were scaled relative to the wild type. Error bars indicate ± SEM. *** denotes $p < 0.001$ (t-test). Experiments performed by Heather Murtona and Simon Whitehall.

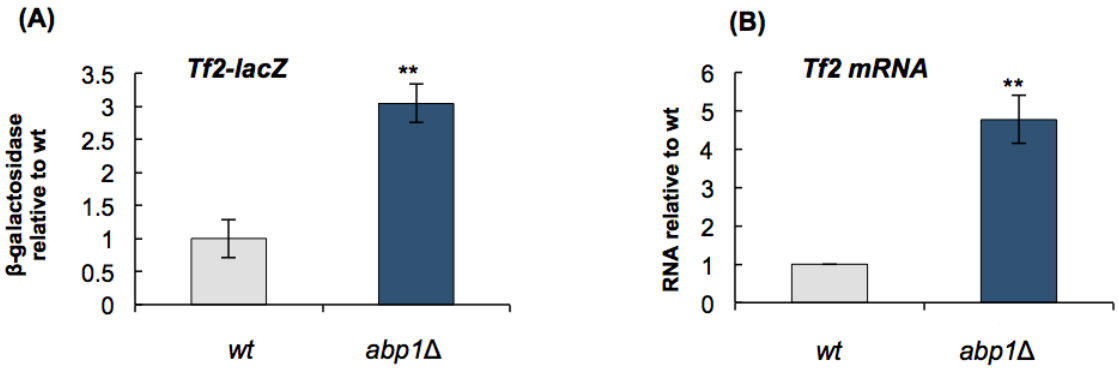


Figure S8. Loss of Abp1-mediated silencing induces *Tf2* mobilization. The indicated strains were grown to mid log growth phase at 30°C in YES. Cells were harvested and processed for β-galactosidase assays. Results are the mean of at least 3 independent assays and are scaled relative to the wild type value. Error bars indicate ± SEM. (B) The mobilization frequency of *Tf2-12natAI* in the indicated strain backgrounds was determined by fluctuation analysis as described in materials and methods. Values were scaled relative to the wild type. Error bars indicate ± SEM. (C) As for (A). (D) *Tf2* mRNA levels in the indicated strains was determined by qRT-PCR. Error bars indicate ± SEM. *** denotes $p < 0.001$, ** $p < 0.01$ and * $p < 0.05$ (t-test). Experiments performed by Heather Murtona and Simon Whitehall.

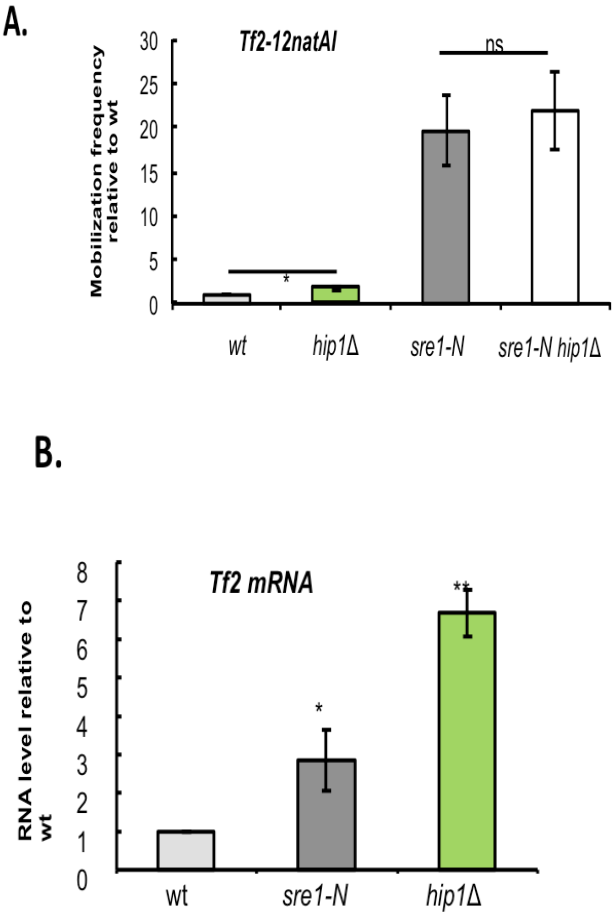


Figure S9. Loss of HIRA-mediated silencing does not result in uncontrolled *Tf2* element mobilization. Mid log phase cells of the indicated strains were subjected to quantitative β-galactosidase assays. Mean values were determined from at least three independent assays and are scaled relative to wild type. Error bars indicate ±SEM. **(A)** Deletion of *hip1*$^{+}$ results in only modest increase in *Tf2* mobilization. The frequency of *Tf2-12natAI* mobilization was determined for the indicated strains by fluctuation analysis using the method of the median. Values were scaled relative to the wild type. Error bars represent ±SEM. Data for *sre1-N* from Fig 2C is included for comparison. **(B)** HIRA suppresses expression of the marked *Tf2-12natAI* element. RNA was prepared from the indicated strains and *Tf2-12natAI* RNA was determined by strand specific RT-PCR. **(C)** Comparison of *Tf2* mRNA levels in *sre1-N* and *hip1*Δ backgrounds. RNA was prepared from the indicated strains and *Tf2* mRNA levels were assayed by qRT-PCR and normalised to *act1*$^{+}$ mRNA. ** $p < 0.01$, * $p < 0.05$ and ns (not significant) $p > 0.05$ (*t*-test). **(F)** Comparison of *Tf2-lacZ* expression with *Tf2-12natAI* mobilization frequency relative to wild type levels in the indicated genetic backgrounds. Experiments performed by Heather Murtona and Simon Whitehall.

Appendix B: Materials and Methods

Strain Construction.

Mutants were generated by standard genetic crosses. Mating was induced by mixing *S. pombe* cells of complementary mating types on EMM media (recipe) for 2-3 days at 26C. Diploids were selected on –adenine media (recipe) for 2 days and streaked to EMM plates to induce spore formation. After 2 days on EMM media at 26C, microscope slides were prepared and cell patches were scored for tetrad spore formation. A small amount of cells were suspended in 1mL sterile dH_20 and a loopful of cell suspension was streaked to a YEA plate. Asci were allowed to rupture and spores were separated using a Singer Instruments MSM-. Spores were allowed to germinate and

Null mutant and C-terminal FLAG (3X) strains were constructed using a Kanamycin cassette with flanking regions homologous to *set1*.[193] Double mutants were generated by standard genetic cross methods.[194] Full-length and domain mutants of *set1* containing an N-terminal FLAG (3X) epitope were generated by a two-step site-directed mutagenesis (SDM). First, the *set1* gene was replaced with a *ura5 lys7* cassette.[195] Second, an SDM PCR fragment containing either full-length or domain deleted FLAG-Set1 was transformed into the above *set1* null strain (*set1*Δ::*ura5 lys7 ura5-14 lys7-2*), and transformants were scored by growth on the uracil counter selective agent 5-Fluoroorotic acid (5-FOA) and sensitivity to lysine minus media.[195] Proper insertions were confirmed by PCR and DNA sequencing. Cells were cultured at 30° C on standard rich media supplemented with 225 mg/L adenine (YEA) agar plates.

Quantitative Reverse Transcription Real-time PCR (qRT-PCR)

RNA was isolated by a hot acid phenol method[196] and converted to cDNA with Superscript III reverse transcriptase and anchored oligo-dT primer (Life Technologies). cDNA was subjected to qPCR analysis using phire polymerase and SYBR green on the Applied Biosystems 7500 Fast Real-Time PCR System. SYBR green signal was normalized to a passive ROX reference dye. Fold expression changes of mutant versus wild-type cells relative to *act1* gene were determined using the $2^{-\Delta\Delta Ct}$ method in Microsoft Excel. Fold expression at heterochromatic regions was normalized to *act1* transcript levels.

Histone Extraction and Detection

Cells from 50 ml culture (OD ~ 0.5) were washed in 10 ml NIB buffer (15 mM PIPES pH 6.8, 0.25M sucrose, 60 mM KCl, 15 mM NaCl, 5mM $MgCl_2$, 1mM $CaCl_2$, 0.8% Triton X-100, 10ng/µl TSA, 1mM PMSF, Roche protease inhibitor mini tablet), lysed with acid-washed glass beads in a bead beater, and centrifuged at 11,000x *g* for 10 min.[48] Cell extract pellets were resuspended in 0.4M H_2SO_4, incubated on ice for 1 h with occasional mixing, and the supernatant was collected following centrifugation at 8,000x *g* for 5 min. Histone extract was concentrated by trichloroacetic acid (TCA) precipitation, washed in acetone and resuspended in 100 µl LDS buffer (Life Technologies) , and

quantitated using the BCA method (Pierce). 5 µg of histone extracts were resolved on 14-22% Tricine SDS PAGE and transferred to a nitrocellulose membrane using the iBlot system (Life Technologies). Histone H3 and modified residues were detected with antibodies against H3 (Abcam, ab1791), H3K4me1 (Abcam, ab8895), H3K4me2 (Fisher, 07030MI), or H3K4me3 (Fisher, 07-473MI).

Protein Extractions and Western Blotting

S. pombe cells (OD 1-2) were lysed in HCS buffer (150mM HEPES pH 7.2, 250mM NaCl, 0.1% NP-40, 1 mM EDTA, 1mM dithiothreitol, 1mM PMSF) and protein inhibitor tablet (Roche)) with acid-washed beads in bead beater (three times 30 s with 2 min interval on ice). 50 µg of protein extracts were run on a PAGE gel (Express Plus 4-20% Bis-tris (MOPS), Genescript) and subjected to overnight western blot transfer at 4°C. Set1 was detected using anti-FLAG antibody (Genescript).

Histone extraction and detection

Cells from 50 ml culture (OD 0.5) were washed in 10 ml NIB buffer (15 mM PIPES pH 6.8, 0.25M sucrose, 60 mM KCl, 15 mM NaCl, 5 mM $MgCl_2$, 1 mM $CaCl_2$, 0.8% Triton X-100, 10 ng/ml TSA, 1 mM PMSF, Roche protease inhibitor mini tablet), lysed with acid-washed glass beads in a bead beater, and centrifuged at 11,000 g for 10 min. Cell

extract pellets were resuspended in 0.4M H_2SO_4, incubated on ice for 1 h with occasional mixing, and the supernatant was collected following centrifugation at 8,000 g for 5 min. Histone extract was concentrated by trichloroacetic acid (TCA) precipitation, washed in acetone and resuspended in 100 ml LDS buffer (Life Technologies) and quantitated using the BCA method (Pierce). 5 mg of histone extracts were resolved on 14–22% Tricine SDS PAGE and transferred to a nitrocellulose membrane using the iBlot system (Life Technologies). Histone H3 and modified residues were detected with antibodies against H3 (Abcam, ab1791), H3K4me1 (Abcam, ab8895), H3K4me2 (Fisher, 07030MI), or H3K4me3 (Fisher, 07-473MI).

Chromatin Immunoprecipitation (ChIP).

Cells were streaked to rich media containg 225ug/L adenine from storage at -80C. Cells were transferred to 5ml liquid media and cultured overnight at 30°C. Cells were transferred to 50ml cultures and grown to 0.8-1.0 OD_{595}. Cells were treated with 1.2M sorbitol and crosslinked with 3% paraformaldehyde for 30 at 18°C in a water bath. Crosslinking reactions were quenched with 6mL 2.5M glycine. Cells pellets were frozen at -80°C and resuspended in 400uL ChIP lysis buffer I (50mM HEPES KOH pH7.5, 140mM NaCl, 1mM EDTA, 1% triton, 0.1%DOC, 1mM PMSF, and proteinase inhibitor tablet) Cells were lysed with glass beads in a bead beater (2x 30s max speed with 2 minutes incubation on ice in between). Lysates were transferred to a microcentrifuge tube and sonicated twice for 20 seconds at 15% amplitude on ice to generate 500-1000bp DNA segments. Samples were then centrifuged 13k rpm 10 minutes at 4°C and chromatin fractions were transferred to new tubes for immunoprecipitation. 11uL aliquots

were removed for whole cell extract reference samples prior to immunoprecipitation. Samples were precleared for non-specific interactions with secondary antibodies by treatment with 2ul washed protein A/G bead mixture at 4C for 3 hours. Precleared chromatin was then immunoprecipitated with 2ug primary antibody overnight at 4°C. The chromatin-antibody complex were transferred to new tubes containing 20uL washed protein A/G beads and rocked at 4C for 3 hours. Chromatin-antibody-bead complex were washed twice in wash buffer I (50mM HEPES KOH pH7.5, 140mM NaCl, 1mM EDTA, 1% triton, 0.1%DOC) for 15 minutes and 5 minutes at 4°C. Chromatin-antibody-bead complex were then washed twice in wash buffer II (50mM HEPES KOH pH7.5, 500mM NaCl, 1mM EDTA, 1% triton, 0.1%DOC) for 15 minutes and 5 minutes at 4°C. Chromatin-antibody-bead complex were then washed twice in wash buffer III (10mM Tris pH8, 0.25M LiCl, 0.5% NP-40, 0.5% DOC, 1mM EDTA) for 15 minutes and 5 minutes at 4°C. Samples were then washed with 1ml TE (50mM Tris pH8, 10mM EDTA) and eluted twice with 50uL of elution buffer (50mM Tris pH8, 10mM EDTA, 1% SDS) at 65°C for 15 minutes. Eluates were then transfered to new tubes containing 100uL TE. ChIP samples and whole cell extracts were incubated overnight at 65°C to reverse crosslinking. Samples were then treated with 2ug RNAse A at 37°C for 1 hour. Following RNase treatment, DNA was extracted with 200uL phenol/chloroform (pH 8.0), followed by extraction with 100% chloroform. Samples were then treated with 10ug glycogen and8uL 5M NaCl and precipitated in 100% ethanol at -80°C for 20 minutes. Samples were then centrifuged for 20 minutes at max speed and washed in 75% ethanol. ChIP samples were suspended in 50uL dH20. WCE samples were resuspended in 100ul dH2O and diluted 1:10 prior to qPCR. qPCR was performed using Phire Hot Start II

DNA Polymerase (Thermo Scientific) supplemented with SYBR green (Life Technologies) on the Applied Biosystems 7500 Fast Real-Time PCR System. Enrichment of ChIP *vs.* input DNA was determined using the $2^{-\Delta\Delta Ct}$ method in Microsoft Excel. Error bars represent the standard deviation of three technical qPCR replicates.

Fluorescence *In Situ* Hybridization (FISH)

S. pombe cells were grown in 10 ml YEA media until OD_{595} ~0.5–1. 10 ml of 2.4 M sorbitol YEA solution was added to culture, and cells were immediately cross-linked with 2.9 ml of freshly made 30% paraformaldehyde/YEA solution for 30 min in a 18°C water bath shaker. Cross-linked reaction was quenched with 1.2 ml of 2.5 M glycine. Cells were transferred to a microcentrifuge tube, subjected to cell wall digestion in 0.5 mg/ml zymolyase solution (Associated of Cape Cod, 100T) at 37°C for 30 min, washed, and blocked with PEMBAL (100 mM PIPES pH 6.9, 1mM EGTA, 1mM MgSO4, 1% BSA, 0.1 M L-lysine) for 1 hr. PEMBAL-treated cells were treated with RNase A (0.1 mg/ml) at 37 °C for 3 h. *Tf2* probes were generated by digestion of *Tf2-orf-4K* plasmid with AluI and DdeI to generate <200bp products wich here then labeled with Cy3-dCTP using a Labeling Kit (Takara). Hybridization was carried out at 40 °C for 12–14 h with 100ng of *Tf2*-orf probes in 100 µl hybridization buffer (50% formamide, 2X SSC, 5X Denhart's solution, 10% dextran sulfate). Cells were washed three times in 100 ml 2 X SSC for 30 min each. Cells were DAPI stained and applied to poly-L-lysine treated microscope slides. Images were obtained using a Zeiss Axioplan 2 microscope. Number of *Tf2* foci were scored by counting the number of discrete Cy3 foci overlapping the

DAPI signal. Volume deconvolution was performed using the Openlab 5.5 software package. The chi-square test of homogeneity was used to determine whether declustering of *Tf2* elements seen in mutant cells relative to wild type was significant.

Tf2-12natAI Mobilization Assays

Strains were plated onto YEA agar to give well-dispersed single colonies. A small (< 1 mm) colony was used to inoculate a 6ml YEA culture which was then incubated at 30°C with shaking until the culture had reached saturation (~48hrs). A 5 ml aliquot was harvested, resuspended in 500 µL H_2O and plated onto two YEA agar plates supplemented with nourseothricin (Nat) (75 µg/mL). An aliquot of the remaining culture was then subjected to 10-fold serial dilution and aliquots of the 10^{-5}, 10^{-6} and 10^{-7} dilutions were plated onto YEA agar plates. Plates were incubated at 30°C for 3-4 days to allow colonies to form. The proportion of Nat resistant cells as fraction of the total viable cells was used to calculate *Tf2-12natAI* mobilization frequency. For each strain under analysis, the mobilization frequency of 5 independent cultures was measured and the median value determined. This process was repeated a minimum of 3 times for each strain under analysis and a mean mobilization frequency was calculated from the median values. For the wild type background the mean mobilization frequency is derived from 13 median values. *p* values were generated by pairwise comparisons using a *t*-test.

Bibliography

1. Dahm, R. From discovering to understanding. Friedrich Miescher's attempts to uncover the function of DNA. *EMBO Rep.* **11,** 153–160 (2010).
2. McCarty, M. & Avery, O. T. Studies on the chemical nature of the substance inducing transformation of pneumococcal types; effect of desoxyribonuclease on the biological activity of the transforming substance. *J. Exp. Med.* **83,** 89–96 (1946).
3. Griffith, F. The Significance of Pneumococcal Types. *J Hyg (Lond)* **27,** 113–159 (1928).
4. Brush, S. G. How theories became knowledge: Morgan's chromosome theory of heredity in America and Britain. *J Hist Biol* **35,** 471–535 (2002).
5. Hershey, A. D. & Chase, M. Independent functions of viral protein and nucleic acid in growth of bacteriophage. *J. Gen. Physiol.* **36,** 39–56 (1952).
6. Passarge, E. Emil Heitz and the concept of heterochromatin: longitudinal chromosome differentiation was recognized fifty years ago. *Am. J. Hum. Genet.* **31,** 106–115 (1979).
7. Watson, J. D. & Crick, F. Molecular structure of nucleic acids. *Nature* (1953).
8. Olins, A. L. & Olins, D. E. Spheroid chromatin units (v bodies). *Science* **183,** 330–332 (1974).
9. Woodcock, C. L. Woodcock: Ultrastructure of inactive chromatin - Google Scholar. *Journal of Cell Biology* (1973).
10. Annunziato, A. T. Assembling chromatin: The long and winding road. *Biochim. Biophys. Acta* **1819,** 196–210 (2012).
11. Kornberg, R. D. Chromatin structure: a repeating unit of histones and DNA. *Science* **184,** 868–871 (1974).
12. Luger, K., Mäder, A. W., Richmond, R. K., Sargent, D. F. & Richmond, T. J. Crystal structure of the nucleosome core particle at 2.8 A resolution. *Nature* **389,** 251–260 (1997).
13. Olins, D. E. & Olins, A. L. Chromatin history: our view from the bridge. *Nat. Rev. Mol. Cell Biol.* **4,** 809–814 (2003).
14. Dorigo, B., Schalch, T., Kulangara, A., Duda, S., Schroeder, R. R., Richmond, T.J. Nucleosome arrays reveal the two-start organization of the chromatin fiber. *Science* **306,** 1571–1573 (2004).
15. Gordon, F., Luger, K. & Hansen, J. C. The core histone N-terminal tail domains function independently and additively during salt-dependent oligomerization of nucleosomal arrays. *J. Biol. Chem.* **280,** 33701–33706 (2005).
16. Schalch, T., Duda, S., Sargent, D. F. & Richmond, T. J. X-ray structure of a tetranucleosome and its implications for the chromatin fibre. *Nature* **436,** 138–141 (2005).
17. Ammar, R. Torti, D., Tsui, K., Gebbia, M., Durbic, T., Bader, G., Giaever, G., Nislow, C. Chromatin is an ancient innovation conserved between Archaea and Eukarya. *Elife* **1,** e00078 (2012).
18. Suto, R. K., Clarkson, M. J., Tremethick, D. J. & Luger, K. Crystal structure of a

nucleosome core particle containing the variant histone H2A.Z. *Nat. Struct. Biol.* **7,** 1121–1124 (2000).
19. Kim, H.S. Vanoosthuyse, V., Fillingham, J., Roguev, A., Watt, S., Kislinger, T., Treyer, A., Carpenter, L.R., Bennett, C.S., Emili, A., Greenblatt, J.F., Hardwick, K.G., Krogan, N.J., Bahler, J., Keogh, M. An acetylated form of histone H2A.Z regulates chromosome architecture in Schizosaccharomyces pombe. *Nat. Struct. Mol. Biol.* **16,** 1286–1293 (2009).
20. Castillo, A. G. Mellone, B.G., Partridge, J.F., Richardson, W., Hamilton, G.L., Allshire, R.C., Pidoux, A.L. Plasticity of fission yeast CENP-A chromatin driven by relative levels of histone H3 and H4. *PLoS Genet.* **3,** e121 (2007).
21. Xhemalce, B., Dawson, M.A., Bannister, A.J. Histone modifications. *Reviews in Cell Biology and Molecular Medicince* (2011).
22. Lachner, M., O'Carroll, D., Rea, S. & Mechtler, K. Methylation of histone H3 lysine 9 creates a binding site for HP1 proteins. *Nature* (2001).
23. Guccione, E., Bassi, C., Casadio, F., Martinato, F., Cesaroni, M., Schuchlautz, H., Luscher, B., Amati, B. Methylation of histone H3R2 by PRMT6 and H3K4 by an MLL complex are mutually exclusive. *Nature* **449,** 933–937 (2007).
24. Krogan, N. J., Kim, M., Tong, A., Golshani, A., Cagney, G., Canadien, V., Richards, D.P., Beattie, B.K., Emili, A., Boone, C., Shilatifard, A., Buratowski, S., Greenblatt, J. Methylation of histone H3 by Set2 in Saccharomyces cerevisiae is linked to transcriptional elongation by RNA polymerase II. *Mol. Cell. Biol.* **23,** 4207–4218 (2003).
25. Allfrey, V.G., Faulkner, R. & Mirsky, A.E. Acetylation and methylation of histones and their possible role in the regulation of RNA synthesis. *Proc. Natl. Acad. Sci. U.S.A.* **51,** 786–794 (1964).
26. Hassa, P. O., Haenni, S. S., Elser, M. & Hottiger, M. O. Nuclear ADP-ribosylation reactions in mammalian cells: where are we today and where are we going? *Microbiol. Mol. Biol. Rev.* **70,** 789–829 (2006).
27. Wang, H., Wang, L., Erdjument-Bromage, H., Vidal, M. Tempst, P., Jones, R.S., Zhang, Y. Role of histone H2A ubiquitination in Polycomb silencing. *Nature* **431,** 873–878 (2004).
28. Sun, Z.W. & Allis, C.D. Ubiquitination of histone H2B regulates H3 methylation and gene silencing in yeast. *Nature* **418,** 104–108 (2002).
29. Cuthbert, G. L., Daugat, S., Snowden, A.W., Erdjument-Bromage, H., Hagiwara, T., Yamada, M., Schneider, R., Gregory, P.D., Tempst, P., Bannister, A.J. Kouzarides, T. Histone deimination antagonizes arginine methylation. *Cell* **118,** 545–553 (2004).
30. Fong, J. J., Nguyen, B. L., Bridger, R. & Medrano, E. E. β-N-Acetylglucosamine (O-GlcNAc) is a novel regulator of mitosis-specific phosphorylations on histone H3. *Journal of Biological Chemistry* (2012).
31. Nelson, C. J., Santos-Rosa, H. & Kouzarides, T. Proline isomerization of histone H3 regulates lysine methylation and gene expression. *Cell* **126,** 905–916 (2006).
32. Nathan, D., Ingvarsdottir, K., Sterner, D.E., Gwendolyn, R.B., Dokmanovic, M., Dorsey, J.A., Whelan, K.A., Krsmanovic, M., Lane, W.S., Meluh, P.B., Johnson, E.S. Berger, S.L. Histone sumoylation is a negative regulator in Saccharomyces cerevisiae and shows dynamic interplay with positive-acting histone

modifications. *Genes Dev.* **20,** 966–976 (2006).
33. Shiio, Y. & Eisenman, R. N. Histone sumoylation is associated with transcriptional repression. *Proc. Natl. Acad. Sci. U.S.A.* **100,** 13225–13230 (2003).
34. Bannister, A. J. & Kouzarides, T. Regulation of chromatin by histone modifications. *Cell Res.* **21,** 381–395 (2011).
35. Grunstein, M. Histone acetylation in chromatin structure and transcription. *Nature* **389,** 349–352 (1997).
36. Benson, L. J. Phillips, J.A., Gu, Y. Parthun, M.R., Hoffman, C.S., Annunziato, A.T. Properties of the type B histone acetyltransferase Hat1: H4 tail interaction, site preference, and involvement in DNA repair. *J. Biol. Chem.* **282,** 836–842 (2007).
37. Allard, S., Utley, R. T., Savard, J. & Clarke, A. NuA4, an essential transcription adaptor/histone H4 acetyltransferase complex containing Esa1p and the ATM-related cofactor Tra1p. *The EMBO Journal* (1999).
38. Doyon, Y. & Côté, J. The highly conserved and multifunctional NuA4 HAT complex. *Curr. Opin. Genet. Dev.* **14,** 147–154 (2004).
39. Levine, M., Cattoglio, C. & Tjian, R. Looping back to leap forward: transcription enters a new era. *Cell* **157,** 13–25 (2014).
40. Guillemette, B., Drogaris, P., Lin, H.S., Armstrong, H., Hiragami-Hamada, K., Imhof, A., Bonneil, E., Thibault, P., Verreault, A., Festensten, R.J. H3 lysine 4 is acetylated at active gene promoters and is regulated by h3 lysine 4 methylation. *PLoS Genet.* **7,** e1001354 (2011).
41. Shilatifard, A. The COMPASS family of histone H3K4 methylases: mechanisms of regulation in development and disease pathogenesis. *Annu. Rev. Biochem.* **81,** 65–95 (2012).
42. Miller, T., Krogan, N.J., Dover, J., Erdjument-Bromage, H., Tempst, P., Johnston, M., Greenblatt, J.F., Shilatifard, A. COMPASS: a complex of proteins associated with a trithorax-related SET domain protein. *Proc. Natl. Acad. Sci. U.S.A.* **98,** 12902–12907 (2001).
43. Min, J., Zhang, X., Cheng, X., Grewal, S. I. S. & Xu, R.-M. Structure of the SET domain histone lysine methyltransferase Clr4. *Nat. Struct. Biol.* **9,** 828–832 (2002).
44. Nagy, P. L., Griesenbeck, J., Kornberg, R. D. & Cleary, M. L. A trithorax-group complex purified from Saccharomyces cerevisiae is required for methylation of histone H3. *Proc. Natl. Acad. Sci. U.S.A.* **99,** 90–94 (2002).
45. Briggs, S. D. Bryk, M., Strahl, B. D., Cheung, W.L. Davie, J. K., Dent, S. Y. R., Winston, F., Allis, C. D. Histone H3 lysine 4 methylation is mediated by Set1 and required for cell growth and rDNA silencing in Saccharomyces cerevisiae. *Genes Dev.* **15,** 3286–3295 (2001).
46. Shilatifard, A. Molecular implementation and physiological roles for histone H3 lysine 4 (H3K4) methylation. *Curr. Opin. Cell Biol.* **20,** 341–348 (2008).
47. Ng, H. H., Robert, F., Young, R. A. & Struhl, K. Targeted recruitment of Set1 histone methylase by elongating Pol II provides a localized mark and memory of recent transcriptional activity. *Mol. Cell* **11,** 709–719 (2003).
48. Noma, K.-I. & Grewal, S. I. S. Histone H3 lysine 4 methylation is mediated by

Set1 and promotes maintenance of active chromatin states in fission yeast. *Proc. Natl. Acad. Sci. U.S.A.* **99 Suppl 4,** 16438–16445 (2002).
49. Morillon, A., Karabetsou, N., Nair, A. & Mellor, J. Dynamic lysine methylation on histone H3 defines the regulatory phase of gene transcription. *Mol. Cell* **18,** 723–734 (2005).
50. Dehe, P. M., Dichtl, B., Schaft, D., Roguev, A., Pamblanco, M., Lebrun, R., Rodriguez-Gil, A., Mkandawire, M., Landsberg, K., Shevchenko, A., Shevchenko, A., Rosaleny, L. E., Tordera, V., Chavez, S., Stewart, A. F., Geli, V. Protein Interactions within the Set1 Complex and Their Roles in the Regulation of Histone 3 Lysine 4 Methylation. *Journal of Biological Chemistry* **281,** 35404–35412 (2006).
51. Kim, J., Kim, J., McGinty, R. K., Nguyen, U. T. T., Muir, T. W., Allis, C. D., Roeder, R. J. The n-SET domain of Set1 regulates H2B ubiquitylation-dependent H3K4 methylation. *Mol. Cell* **49,** 1121–1133 (2013).
52. Soares, L. M., Radman-Livaja, M., Lin, S. G., Rando, O. J. & Buratowski, S. Feedback Control of Set1 Protein Levels Is Important for Proper H3K4 Methylation Patterns. *Cell Rep* **6,** 961–972 (2014).
53. Schneider, J., Dover, J., Johnston, M. & Shilatifard, A. Global proteomic analysis of S. cerevisiae (GPS) to identify proteins required for histone modifications. *Meth. Enzymol.* **377,** 227–234 (2004).
54. Giaever, G. *et al.* Functional profiling of the Saccharomyces cerevisiae genome. *Nature* **418,** 387–391 (2002).
55. Dover, J. *et al.* Methylation of histone H3 by COMPASS requires ubiquitination of histone H2B by Rad6. *J. Biol. Chem.* **277,** 28368–28371 (2002).
56. Ng, H. H., Xu, R.-M., Zhang, Y. & Struhl, K. Ubiquitination of histone H2B by Rad6 is required for efficient Dot1-mediated methylation of histone H3 lysine 79. *J. Biol. Chem.* **277,** 34655–34657 (2002).
57. Wood, A. *et al.* Bre1, an E3 ubiquitin ligase required for recruitment and substrate selection of Rad6 at a promoter. *Mol. Cell* **11,** 267–274 (2003).
58. Hwang, W. W. *et al.* A conserved RING finger protein required for histone H2B monoubiquitination and cell size control. *Mol. Cell* **11,** 261–266 (2003).
59. Schulze, J. M. *et al.* Linking cell cycle to histone modifications: SBF and H2B monoubiquitination machinery and cell-cycle regulation of H3K79 dimethylation. *Mol. Cell* **35,** 626–641 (2009).
60. Lee, J. S. *et al.* Histone Crosstalk between H2B Monoubiquitination and H3 Methylation Mediated by COMPASS. *Cell* **131,** 1084–1096 (2007).
61. Zheng, S., Wyrick, J. J. & Reese, J. C. Novel trans-tail regulation of H2B ubiquitylation and H3K4 methylation by the N terminus of histone H2A. *Mol. Cell. Biol.* **30,** 3635–3645 (2010).
62. Wu, M. *et al.* Molecular regulation of H3K4 trimethylation by Wdr82, a component of human Set1/COMPASS. *Mol. Cell. Biol.* **28,** 7337–7344 (2008).
63. Zhu, B. *et al.* Monoubiquitination of human histone H2B: the factors involved and their roles in HOX gene regulation. *Mol. Cell* **20,** 601–611 (2005).
64. Kim, T. & Buratowski, S. Dimethylation of H3K4 by Set1 recruits the Set3 histone deacetylase complex to 5' transcribed regions. *Cell* **137,** 259–272 (2009).
65. Pavri, R. *et al.* Histone H2B monoubiquitination functions cooperatively with

FACT to regulate elongation by RNA polymerase II. *Cell* **125,** 703–717 (2006).
66. Tanny, J. C., Erdjument-Bromage, H., Tempst, P. & Allis, C. D. Ubiquitylation of histone H2B controls RNA polymerase II transcription elongation independently of histone H3 methylation. *Genes Dev.* **21,** 835–847 (2007).
67. Costa, P. J. & Arndt, K. M. Synthetic lethal interactions suggest a role for the Saccharomyces cerevisiae Rtf1 protein in transcription elongation. *Genetics* **156,** 535–547 (2000).
68. Mueller, C. L. & Jaehning, J. A. Ctr9, Rtf1, and Leo1 are components of the Paf1/RNA polymerase II complex. *Mol. Cell. Biol.* **22,** 1971–1980 (2002).
69. Pokholok, D. K. *et al.* Genome-wide map of nucleosome acetylation and methylation in yeast. *Cell* **122,** 517–527 (2005).
70. Chang, M. *et al.* A complex containing RNA polymerase II, Paf1p, Cdc73p, Hpr1p, and Ccr4p plays a role in protein kinase C signaling. *Mol. Cell. Biol.* **19,** 1056–1067 (1999).
71. Squazzo, S. L. *et al.* The Paf1 complex physically and functionally associates with transcription elongation factors in vivo. *EMBO J* **21,** 1764–1774 (2002).
72. Berretta, J., Pinskaya, M. & Morillon, A. A cryptic unstable transcript mediates transcriptional trans-silencing of the Ty1 retrotransposon in S. cerevisiae. *Genes Dev.* **22,** 615–626 (2008).
73. Camblong, J. *et al.* Trans-acting antisense RNAs mediate transcriptional gene cosuppression in S. cerevisiae. *Genes Dev.* **23,** 1534–1545 (2009).
74. van Dijk, E. L. *et al.* XUTs are a class of Xrn1-sensitive antisense regulatory non-coding RNA in yeast. *Nature* **475,** 114–117 (2011).
75. Lorenz, D. R. *et al.* CENP-B cooperates with Set1 in bidirectional transcriptional silencing and genome organization of retrotransposons. *Mol. Cell. Biol.* **32,** 4215–4225 (2012).
76. Weiner, A. *et al.* Systematic Dissection of Roles for Chromatin Regulators in a Yeast Stress Response. *PLoS Biol* **10,** e1001369 (2012).
77. Gal, C., Moore, K. M., Paszkiewicz, K., Kent, N. A. & Whitehall, S. K. The impact of the HIRA histone chaperone upon global nucleosome architecture. *Cell Cycle* **14,** 123–134 (2015).
78. Wirén, M. *et al.* Genomewide analysis of nucleosome density histone acetylation and HDAC function in fission yeast. *EMBO J* **24,** 2906–2918 (2005).
79. Verdel, A. *et al.* RNAi-mediated targeting of heterochromatin by the RITS complex. *Science* **303,** 672–676 (2004).
80. Grewal, S. I. RNAi-dependent formation of heterochromatin and its diverse functions. *Curr. Opin. Genet. Dev.* **20,** 134–141 (2010).
81. Noma, K.-I. *et al.* RITS acts in cis to promote RNA interference-mediated transcriptional and post-transcriptional silencing. *Nat. Genet.* **36,** 1174–1180 (2004).
82. Sugiyama, T., Cam, H., Verdel, A., Moazed, D. & Grewal, S. I. S. RNA-dependent RNA polymerase is an essential component of a self-enforcing loop coupling heterochromatin assembly to siRNA production. *Proc. Natl. Acad. Sci. U.S.A.* **102,** 152–157 (2005).
83. Cam, H. P. *et al.* Comprehensive analysis of heterochromatin- and RNAi-mediated epigenetic control of the fission yeast genome. *Nat. Genet.* **37,** 809–819

(2005).
84. Hall, I. M. *et al.* Establishment and maintenance of a heterochromatin domain. *Science* **297,** 2232–2237 (2002).
85. Gruenbaum, Y., Margalit, A., Goldman, R. D., Shumaker, D. K. & Wilson, K. L. The nuclear lamina comes of age. *Nat. Rev. Mol. Cell Biol.* **6,** 21–31 (2005).
86. Handwerger, K. E. & Gall, J. G. Subnuclear organelles: new insights into form and function. *Trends Cell Biol.* **16,** 19–26 (2006).
87. Iborra, F. J., Pombo, A., Jackson, D. A. & Cook, P. R. Active RNA polymerases are localized within discrete transcription "factories' in human nuclei. *J. Cell. Sci.* **109 (Pt 6),** 1427–1436 (1996).
88. Bolzer, A. *et al.* Three-dimensional maps of all chromosomes in human male fibroblast nuclei and prometaphase rosettes. *PLoS Biol* **3,** e157 (2005).
89. Cremer, T. & Cremer, C. Chromosome territories, nuclear architecture and gene regulation in mammalian cells. *Nat. Rev. Genet.* **2,** 292–301 (2001).
90. Croft, J. A. *et al.* Differences in the localization and morphology of chromosomes in the human nucleus. *The Journal of Cell Biology* **145,** 1119–1131 (1999).
91. Edelmann, P., Bornfleth, H., Zink, D., Cremer, T. & Cremer, C. Morphology and dynamics of chromosome territories in living cells. *Biochim. Biophys. Acta* **1551,** M29–39 (2001).
92. Zink, D. *et al.* Structure and dynamics of human interphase chromosome territories in vivo. *Hum. Genet.* **102,** 241–251 (1998).
93. Tanabe, H. *et al.* Evolutionary conservation of chromosome territory arrangements in cell nuclei from higher primates. *Proc. Natl. Acad. Sci. U.S.A.* **99,** 4424–4429 (2002).
94. Cremer, M. *et al.* Inheritance of gene density-related higher order chromatin arrangements in normal and tumor cell nuclei. *The Journal of Cell Biology* **162,** 809–820 (2003).
95. Gasser, S. M. Visualizing chromatin dynamics in interphase nuclei. *Science* **296,** 1412–1416 (2002).
96. Levi, V., Ruan, Q., Plutz, M., Belmont, A. S. & Gratton, E. Chromatin dynamics in interphase cells revealed by tracking in a two-photon excitation microscope. *Biophys. J.* **89,** 4275–4285 (2005).
97. Tumbar, T. & Belmont, A. S. Interphase movements of a DNA chromosome region modulated by VP16 transcriptional activator. *Nat. Cell Biol.* **3,** 134–139 (2001).
98. Kosak, S. T. *et al.* Subnuclear compartmentalization of immunoglobulin loci during lymphocyte development. *Science* **296,** 158–162 (2002).
99. Hewitt, S. L., High, F. A., Reiner, S. L., Fisher, A. G. & Merkenschlager, M. Nuclear repositioning marks the selective exclusion of lineage-inappropriate transcription factor loci during T helper cell differentiation. *Eur. J. Immunol.* **34,** 3604–3613 (2004).
100. Williams, R. R. E., Broad, S., Sheer, D. & Ragoussis, J. Subchromosomal positioning of the epidermal differentiation complex (EDC) in keratinocyte and lymphoblast interphase nuclei. *Exp. Cell Res.* **272,** 163–175 (2002).
101. Dekker, J., Rippe, K., Dekker, M. & Kleckner, N. Capturing chromosome

conformation. *Science* **295,** 1306–1311 (2002).
102. Dostie, J., Richmond, T. A., Arnaout, R. A. & Selzer, R. R. Chromosome Conformation Capture Carbon Copy (5C): a massively parallel solution for mapping interactions between genomic elements. *Genome Research* (2006).
103. Zhao, Z. *et al.* Circular chromosome conformation capture (4C) uncovers extensive networks of epigenetically regulated intra- and interchromosomal interactions. *Nat. Genet.* **38,** 1341–1347 (2006).
104. Grand, R. S. *et al.* Chromosome conformation maps in fission yeast reveal cell cycle dependent sub nuclear structure. *Nucleic Acids Res.* **42,** 12585–12599 (2014).
105. Noma K, Cam, H. P., Maraia, R. J. & Grewal, S. A role for TFIIIC transcription factor complex in genome organization. *Cell* (2006).
106. Funabiki, H., Hagan, I. & Uzawa, S. Cell cycle-dependent specific positioning and clustering of centromeres and telomeres in fission yeast. *The Journal of Cell Biology* (1993).
107. Tanaka, A. *et al.* Epigenetic Regulation of Condensin-Mediated Genome Organization during the Cell Cycle and upon DNA Damage through Histone H3 Lysine 56 Acetylation. *Mol. Cell* **48,** 532–546 (2012).
108. Cam, H. P., Noma, K.-I., Ebina, H., Levin, H. L. & Grewal, S. I. S. Host genome surveillance for retrotransposons by transposon-derived proteins. *Nature* **451,** 431–436 (2008).
109. Lander, E., Linton, L., Birren, B. & Nusbaum, C. Initial sequencing and analysis of the human genome. *Nature* (2001).
110. SanMiguel, P. *et al.* Nested retrotransposons in the intergenic regions of the maize genome. *Science* **274,** 765–768 (1996).
111. Gunjan, A., Paik, J. & Verreault, A. Regulation of histone synthesis and nucleosome assembly. *Biochimie* **87,** 625–635 (2005).
112. Mikheyeva, I. V., Grady, P., Tamburini, F. B. & Lorenz, D. R. Multifaceted genome control by Set1 dependent and independent of H3K4 methylation and the Set1C/COMPASS complex. *PLOS Genetics* (2014).
113. Xhemalce, B. & Kouzarides, T. A chromodomain switch mediated by histone H3 Lys 4 acetylation regulates heterochromatin assembly. *Genes Dev.* **24,** 647–652 (2010).
114. Mellone, B. G. *et al.* Centromere silencing and function in fission yeast is governed by the amino terminus of histone H3. *Curr. Biol.* **13,** 1748–1757 (2003).
115. Rodríguez-Paredes, M. & Esteller, M. Cancer epigenetics reaches mainstream oncology. *Nat. Med.* **17,** 330–339 (2011).
116. Dorier, J. & Stasiak, A. Topological origins of chromosomal territories. *Nucleic Acids Res.* **37,** 6316–6322 (2009).
117. Roguev, A. High Conservation of the Set1/Rad6 Axis of Histone 3 Lysine 4 Methylation in Budding and Fission Yeasts. *Journal of Biological Chemistry* **278,** 8487–8493 (2002).
118. Schlichter, A. & Cairns, B. R. Histone trimethylation by Set1 is coordinated by the RRM, autoinhibitory, and catalytic domains. *Nat. Genet.* **24,** 1222–1231 (2005).

119. Trésaugues, L. *et al.* Structural Characterization of Set1 RNA Recognition Motifs and their Role in Histone H3 Lysine 4 Methylation. *Journal of Molecular Biology* **359,** 1170–1181 (2006).
120. Kim, C. *et al.* Regulation of post-translational protein arginine methylation during HeLa cell cycle. *Biochim. Biophys. Acta* **1800,** 977–985 (2010).
121. Thornton, J. L. *et al.* Context dependency of Set1/COMPASS-mediated histone H3 Lys4 trimethylation. *Genes Dev.* **28,** 115–120 (2014).
122. Takahashi, Y.-H. *et al.* Structural analysis of the core COMPASS family of histone H3K4 methylases from yeast to human. *Proc. Natl. Acad. Sci. U.S.A.* **108,** 20526–20531 (2011).
123. Lorenz, D. R., Meyer, L. F., Grady, P. J. R., Meyer, M. M. & Cam, H. P. Heterochromatin assembly and transcriptome repression by Set1 in coordination with a class II histone deacetylase. *Elife* **3,** e04506 (2014).
124. Schneider, J. *et al.* Molecular regulation of histone H3 trimethylation by COMPASS and the regulation of gene expression. *Mol. Cell* **19,** 849–856 (2005).
125. Steward, M. M., Lee, J. S., O'Donovan, A. & Wyatt, M. Molecular regulation of H3K4 trimethylation by ASH2L, a shared subunit of MLL complexes. *Nature structural & molecular biology* (2006).
126. Latham, J. A., Chosed, R. J., Wang, S. & Dent, S. Y. R. Chromatin signaling to kinetochores: transregulation of Dam1 methylation by histone H2B ubiquitination. *Cell* **146,** 709–719 (2011).
127. Mersman, D. P., Du, H.-N., Fingerman, I. M., South, P. F. & Briggs, S. D. A charge-based interaction conserved within the H3K4 methyltransferase complexes is needed for protein stability, histone methylation, and gene expression. *Journal of Biological Chemistry* (2011).
128. Takahashi, Y.-H. *et al.* Regulation of H3K4 trimethylation via Cps40 (Spp1) of COMPASS is monoubiquitination independent: implication for a Phe/Tyr switch by the catalytic domain of Set1. *Mol. Cell. Biol.* **29,** 3478–3486 (2009).
129. Yamanaka, S. *et al.* RNAi triggered by specialized machinery silences developmental genes and retrotransposons. *Nature* **493,** 557–560 (2012).
130. Moazed, D. Mechanisms for the inheritance of chromatin states. *Cell* **146,** 510–518 (2011).
131. Reyes-Turcu, F. E. & Grewal, S. I. Different means, same end-heterochromatin formation by RNAi and RNAi-independent RNA processing factors in fission yeast. *Curr. Opin. Genet. Dev.* **22,** 156–163 (2012).
132. Mitchell, P., Petfalski, E., Shevchenko, A., Mann, M. & Tollervey, D. The exosome: a conserved eukaryotic RNA processing complex containing multiple 3‘-->5’ exoribonucleases. *Cell* **91,** 457–466 (1997).
133. Yamada, T., Fischle, W., Sugiyama, T., Allis, C. D. & Grewal, S. I. S. The nucleation and maintenance of heterochromatin by a histone deacetylase in fission yeast. *Mol. Cell* **20,** 173–185 (2005).
134. Shankaranarayana, G. D., Motamedi, M. R., Moazed, D. & Grewal, S. I. S. Sir2 regulates histone H3 lysine 9 methylation and heterochromatin assembly in fission yeast. *Curr. Biol.* **13,** 1240–1246 (2003).
135. Alper, B. J. *et al.* Sir2 is required for Clr4 to initiate centromeric heterochromatin assembly in fission yeast. *EMBO J* **32,** 2321–2335 (2013).

136. Kazazian, H. H. Mobile elements: drivers of genome evolution. *Science* **303,** 1626–1632 (2004).
137. Beauregard, A., Curcio, M. J. & Belfort, M. The take and give between retrotransposable elements and their hosts. *Annu. Rev. Genet.* **42,** 587–617 (2008).
138. Burns, K. H. & Boeke, J. D. Human transposon tectonics. *Cell* **149,** 740–752 (2012).
139. Levin, H. L. & Moran, J. V. Dynamic interactions between transposable elements and their hosts. *Nat. Rev. Genet.* **12,** 615–627 (2011).
140. Orgel, L. E. & Crick, F. H. Selfish DNA: the ultimate parasite. *Nature* **284,** 604–607 (1980).
141. Shapiro, J. A. Retrotransposons and regulatory suites. *Bioessays* **27,** 122–125 (2005).
142. Feschotte, C. Transposable elements and the evolution of regulatory networks. *Nat. Rev. Genet.* **9,** 397–405 (2008).
143. Hancks, D. C. & Kazazian, H. H. Active human retrotransposons: variation and disease. *Curr. Opin. Genet. Dev.* **22,** 191–203 (2012).
144. Chalopin, D., Naville, M., Plard, F., Galiana, D. & Volff, J.-N. Comparative analysis of transposable elements highlights mobilome diversity and evolution in vertebrates. *Genome Biol Evol* **7,** 567–580 (2015).
145. Maksakova, I. A., Mager, D. L. & Reiss, D. Keeping active endogenous retroviral-like elements in check: the epigenetic perspective. *Cell. Mol. Life Sci.* **65,** 3329–3347 (2008).
146. Slotkin, R. K. & Martienssen, R. Transposable elements and the epigenetic regulation of the genome. *Nat. Rev. Genet.* **8,** 272–285 (2007).
147. Maxwell, P. H. & Curcio, M. J. Host factors that control long terminal repeat retrotransposons in Saccharomyces cerevisiae: implications for regulation of mammalian retroviruses. *Eukaryotic Cell* **6,** 1069–1080 (2007).
148. Kelly, F. D. & Levin, H. L. The evolution of transposons in Schizosaccharomyces pombe. *Cytogenet. Genome Res.* **110,** 566–574 (2005).
149. Lesage, P. & Todeschini, A. L. Happy together: the life and times of Ty retrotransposons and their hosts. *Cytogenet. Genome Res.* **110,** 70–90 (2005).
150. Bowen, N. J., Jordan, I. K., Epstein, J. A., Wood, V. & Levin, H. L. Retrotransposons and their recognition of pol II promoters: a comprehensive survey of the transposable elements from the complete genome sequence of Schizosaccharomyces pombe. *Genome Res.* **13,** 1984–1997 (2003).
151. Sehgal, A., Lee, C.-Y. S. & Espenshade, P. J. SREBP controls oxygen-dependent mobilization of retrotransposons in fission yeast. *PLoS Genet.* **3,** e131 (2007).
152. Durand-Dubief, M. *et al.* Specific functions for the fission yeast Sirtuins Hst2 and Hst4 in gene regulation and retrotransposon silencing. *EMBO J* **26,** 2477–2488 (2007).
153. Hansen, K. R. *et al.* Global effects on gene expression in fission yeast by silencing and RNA interference machineries. *Mol. Cell. Biol.* **25,** 590–601 (2005).
154. Nicolas, E. *et al.* Distinct roles of HDAC complexes in promoter silencing, antisense suppression and DNA damage protection. *Nat. Struct. Mol. Biol.* **14,**

372–380 (2007).
155. Anderson, H. E. *et al.* The fission yeast HIRA histone chaperone is required for promoter silencing and the suppression of cryptic antisense transcripts. *Mol. Cell. Biol.* **29,** 5158–5167 (2009).
156. Greenall, A. *et al.* Hip3 interacts with the HIRA proteins Hip1 and Slm9 and is required for transcriptional silencing and accurate chromosome segregation. *J. Biol. Chem.* **281,** 8732–8739 (2006).
157. Yamane, K. *et al.* Asf1/HIRA facilitate global histone deacetylation and associate with HP1 to promote nucleosome occupancy at heterochromatic loci. *Mol. Cell* **41,** 56–66 (2011).
158. Irelan, J. T., Gutkin, G. I. & Clarke, L. Functional redundancies, distinct localizations and interactions among three fission yeast homologs of centromere protein-B. *Genetics* **157,** 1191–1203 (2001).
159. Hoff, E. F., Levin, H. L. & Boeke, J. D. Schizosaccharomyces pombe retrotransposon Tf2 mobilizes primarily through homologous cDNA recombination. *Mol. Cell. Biol.* **18,** 6839–6852 (1998).
160. Cho, K., Lee, Y.-K. & Greenhalgh, D. G. Endogenous retroviruses in systemic response to stress signals. *Shock* **30,** 105–116 (2008).
161. Grandbastien, M.-A. LTR retrotransposons, handy hitchhikers of plant regulation and stress response. *Biochim. Biophys. Acta* **1849,** 403–416 (2015).
162. Hughes, B. T. & Espenshade, P. J. Oxygen-regulated degradation of fission yeast SREBP by Ofd1, a prolyl hydroxylase family member. *EMBO J* **27,** 1491–1501 (2008).
163. Blackwell, C. *et al.* The Schizosaccharomyces pombe HIRA-like protein Hip1 is required for the periodic expression of histone genes and contributes to the function of complex centromeres. *Mol. Cell. Biol.* **24,** 4309–4320 (2004).
164. Anderson, H. E. *et al.* Silencing mediated by the Schizosaccharomyces pombe HIRA complex is dependent upon the Hpc2-like protein, Hip4. *PLoS ONE* **5,** e13488 (2010).
165. Osley, M. A. The regulation of histone synthesis in the cell cycle. *Annu. Rev. Biochem.* **60,** 827–861 (1991).
166. Takayama, Y. & Takahashi, K. Differential regulation of repeated histone genes during the fission yeast cell cycle. *Nucleic Acids Res.* **35,** 3223–3237 (2007).
167. Sherwood, P. W., Tsang, S. V. & Osley, M. A. Characterization of HIR1 and HIR2, two genes required for regulation of histone gene transcription in Saccharomyces cerevisiae. *Mol. Cell. Biol.* **13,** 28–38 (1993).
168. Spector, M. S., Raff, A., DeSilva, H., Lee, K. & Osley, M. A. Hir1p and Hir2p function as transcriptional corepressors to regulate histone gene transcription in the Saccharomyces cerevisiae cell cycle. *Mol. Cell. Biol.* **17,** 545–552 (1997).
169. Corpet, A. & Almouzni, G. Making copies of chromatin: the challenge of nucleosomal organization and epigenetic information. *Trends Cell Biol.* **19,** 29–41 (2009).
170. Valente, L. P. *et al.* Myb-domain protein Teb1 controls histone levels and centromere assembly in fission yeast. *EMBO J* **32,** 450–460 (2013).
171. Guan, Q., Haroon, S., Bravo, D. G., Will, J. L. & Gasch, A. P. Cellular memory of acquired stress resistance in Saccharomyces cerevisiae. *Genetics* **192,** 495–

505 (2012).
172. Nislow, C., Ray, E. & Pillus, L. SET1, a yeast member of the trithorax family, functions in transcriptional silencing and diverse cellular processes. *Mol. Biol. Cell* **8,** 2421–2436 (1997).
173. Bryk, M. *et al.* Evidence that Set1, a factor required for methylation of histone H3, regulates rDNA silencing in S. cerevisiae by a Sir2-independent mechanism. *Curr. Biol.* **12,** 165–170 (2002).
174. Margaritis, T. *et al.* Two distinct repressive mechanisms for histone 3 lysine 4 methylation through promoting 3'-end antisense transcription. *PLoS Genet.* **8,** e1002952 (2012).
175. Zhang, K. *et al.* The Set1 methyltransferase opposes Ipl1 aurora kinase functions in chromosome segregation. *Cell* **122,** 723–734 (2005).
176. Zhang, W. *et al.* Crystal structures and RNA-binding properties of the RNA recognition motifs of heterogeneous nuclear ribonucleoprotein L: insights into its roles in alternative splicing regulation. *J. Biol. Chem.* **288,** 22636–22649 (2013).
177. Mackereth, C. D. *et al.* Multi-domain conformational selection underlies pre-mRNA splicing regulation by U2AF. *Nature* **475,** 408–411 (2011).
178. Dichtl, B., Aasland, R. & Keller, W. Functions for S. cerevisiae Swd2p in 3' end formation of specific mRNAs and snoRNAs and global histone 3 lysine 4 methylation. *RNA* **10,** 965–977 (2004).
179. Zofall, M. & Grewal, S. I. S. HULC, a histone H2B ubiquitinating complex, modulates heterochromatin independent of histone methylation in fission yeast. *J. Biol. Chem.* **282,** 14065–14072 (2007).
180. Nedea, E. *et al.* The Glc7 phosphatase subunit of the cleavage and polyadenylation factor is essential for transcription termination on snoRNA genes. *Mol. Cell* **29,** 577–587 (2008).
181. Roguev, A., Schaft, D. & Shevchenko, A. The Saccharomyces cerevisiae Set1 complex includes an Ash2 homologue and methylates histone 3 lysine 4. *The EMBO Journal* (2001).
182. Macadangdang, B. R. *et al.* Evolution of histone 2A for chromatin compaction in eukaryotes. *Elife* **3,** (2014).
183. Dieudonné, M. *et al.* Transcriptional competence of the integrated HIV-1 provirus at the nuclear periphery. *EMBO J* **28,** 2231–2243 (2009).
184. Lusic, M. *et al.* Proximity to PML nuclear bodies regulates HIV-1 latency in CD4+ T cells. *Cell Host Microbe* **13,** 665–677 (2013).
185. Ben-Aroya, S., Mieczkowski, P. A., Petes, T. D. & Kupiec, M. The compact chromatin structure of a Ty repeated sequence suppresses recombination hotspot activity in Saccharomyces cerevisiae. *Mol. Cell* **15,** 221–231 (2004).
186. Zaratiegui, M. *et al.* CENP-B preserves genome integrity at replication forks paused by retrotransposon LTR. *Nature* **469,** 112–115 (2011).
187. Mathieu, O. & Bouché, N. Interplay between chromatin and RNA processing. *Curr. Opin. Plant Biol.* **18,** 60–65 (2014).
188. Morimoto, M. & Boerkoel, C. F. The role of nuclear bodies in gene expression and disease. *Biology (Basel)* **2,** 976–1033 (2013).
189. Capy, P., Gasperi, G., Biémont, C. & Bazin, C. Stress and transposable elements: co-evolution or useful parasites? *Heredity* **85,** 101–106 (2000).

190. McDonald, J. F. *Transposable Elements and Genome Evolution*. (Springer Science & Business Media, 2012).
191. Singh, R. K. *et al.* Excess histone levels mediate cytotoxicity via multiple mechanisms. *Cell Cycle* **9,** 4236–4244 (2010).
192. Meeks-Wagner, D. & Hartwell, L. H. Normal stoichiometry of histone dimer sets is necessary for high fidelity of mitotic chromosome transmission. *Cell* **44,** 43–52 (1986).
193. Bähler, J. *et al.* Heterologous modules for efficient and versatile PCR-based gene targeting in Schizosaccharomyces pombe. *Yeast* **14,** 943–951 (1998).
194. Moreno, S., Klar, A. & Nurse, P. Molecular genetic analysis of fission yeast Schizosaccharomyces pombe. *Meth. Enzymol.* **194,** 795–823 (1991).
195. Mudge, D. K., Hoffman, C. A., Lubinski, T. J. & Hoffman, C. S. Use of a ura5+-lys7+ cassette to construct unmarked gene knock-ins in Schizosaccharomyces pombe. *Curr. Genet.* **58,** 59–64 (2012).
196. Lyne, R. *et al.* Whole-genome microarrays of fission yeast: characteristics, accuracy, reproducibility, and processing of array data. *BMC Genomics* **4,** 27 (2003).