Doctoral Thesis

Heterochromatin suppresses gross chromosomal rearrangements at centromeres

ヘテロクロマチンによるセントロメア・リピートを介した染色体異常の抑制機構

Akiko Okita

Department of Biological Sciences

Graduate School of Science

Osaka University

February 2019

CONTENTS

GENERAL INTRODUCTION

1. Inheritance of genetic information requires accurate DNA replication and chromosome segregation

During cell proliferation, chromosomes carrying genetic information in a DNA molecule replicate accurately and the duplicated chromosomes are segregated equally into two daughter cells (Fig. 1A). For accurate replication, the checkpoint factors control the progression of cell cycle to duplicate the entire chromosomes before chromosome segregation (Masai et al. 2010). The licensing factors control the initiation of replication to restrict only once per cell cycle. During cell division, microtubule fibers attach to kinetochores at centromere regions and are organized into spindles that physically pull duplicated chromatids, called sister chromatids, towards opposite poles. Cohesin proteins also bind to centromeres and hold the sister chromatids together (Losada et al. 1998). The sister chromatid cohesion opposes the pulling forces that are generated by microtubules and thereby enables the bi-orientation of chromosomes on the spindle (Tanaka et al. 2000). It creates a tension between the chromatids leading to equal segregation. As each human cell contains 46 chromosomes (23 chromosomes from each parent), duplication and segregation of such a large number of chromosomes require precise mechanisms.

2. Chromosomal abnormality causes cell death and genetic diseases

When there is an error during cell proliferation, the cells suffer chromosomal abnormalities which can result in genetic diseases including cancers or even cell death. In humans, approximately 0.5% of newborn infants have chromosomal abnormalities (Hamerton et al. 1975). There are many types of chromosomal abnormalities, and they are classified into two groups: numerical abnormality and structural abnormality (Gordon et al. 2012; Weckselblatt and Rudd 2015). The numerical abnormality, called aneuploidy, accounts for ~60% of newborns harboring chromosomal abnormality (Hamerton et al. 1975). Errors in chromosomal segregation can lead to the loss or gain of chromosomes in daughter cells. (Fig. 1B). Down syndrome (1/800 births), which is a developmental disorder caused by the presence of an extra chromosome 21, is the most

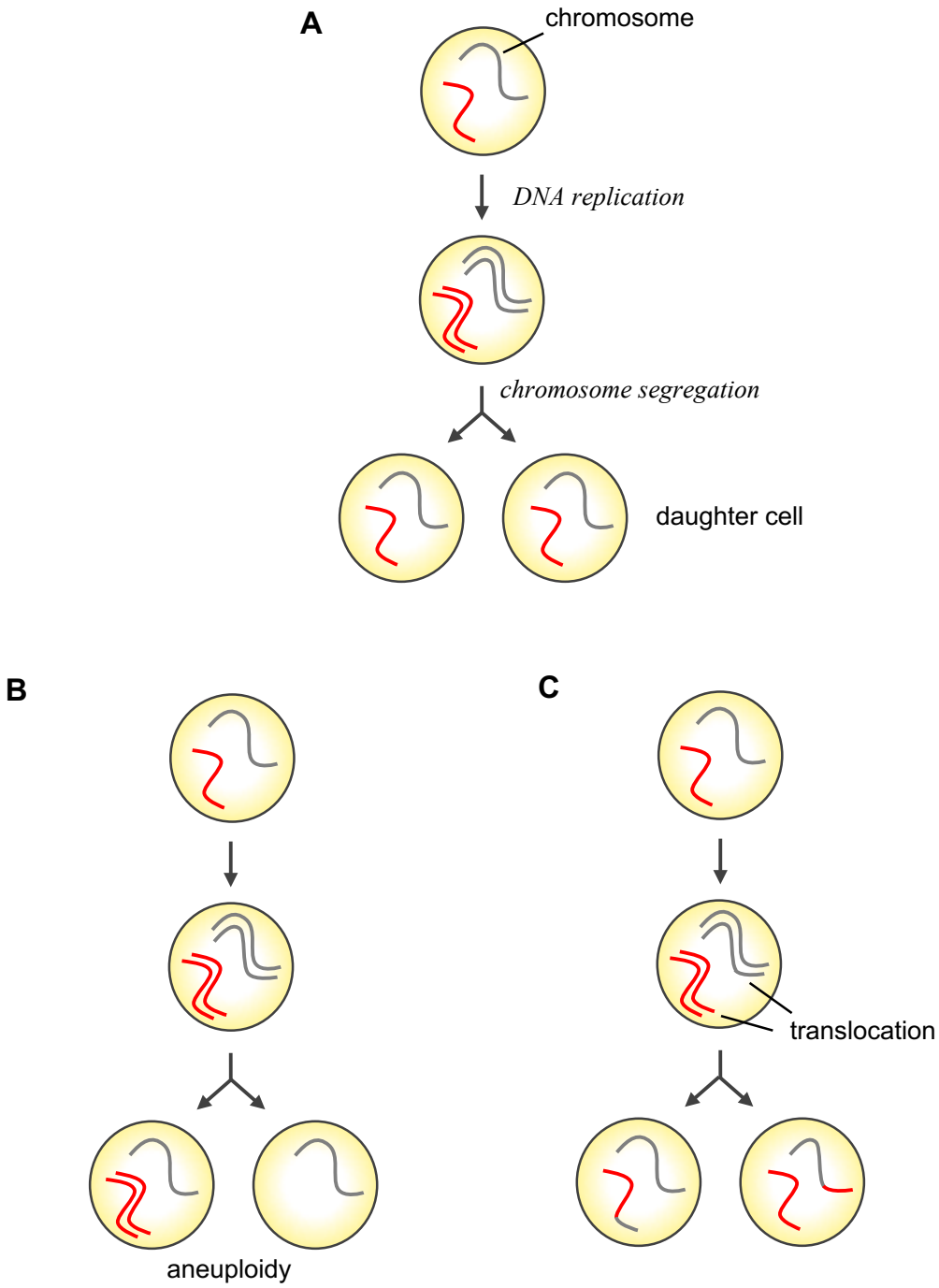


Fig. 1 Chromosomes are stably maintained during cell proliferation.
A. To maintain genome integrity, chromosomes are duplicated accurately and segregated equally to daughter cells during cell proliferation. **B.** When chromosomes are segregated unequally, daughter cells gain or miss the chromosomes (aneuploidy). **C.** Repair of broken replication fork can cause translocation with different chromosomes. Improper repairment of DNA damages can also cause such a structural abnormalities.

common disorder resulting from aneuploidy (Lejeune et al. 1959). On the other hand, the structural abnormalities include rearrangements such as translocation, deletion, or duplication (Fig. 1C). Structural abnormalities involving a change in a large portion of the chromosome are called gross chromosomal rearrangements (GCRs). Robertsonian translocation, a whole-arm translocation between non-homologous chromosomes, is the most common structural abnormality in humans (1/1,000 births) (Hamerton et al. 1975). Robertsonian translocation preferentially occurs around centromeres in the acrocentric chromosome, a chromosome in which the centromere is near one end. The most common translocation event involving acrocentric chromosomes is between chromosomes 13 and 14, which accounts for ~75% of all Robertsonian translocations (Nielsen and Wohlert 1991). Approximately 10% of Patau syndrome (1/10,000), which is caused by the presence of extra chromosome 13 and can cause heart defects and brain and spinal cord abnormalities (Wyllie et al. 1994), have Robertsonian translocation between chromosome 13 and 14. Many clinical cases have shown that chromosomal instability is strongly related with genetic diseases.

3. Repetitive DNA sequences suffer from chromosomal instability

The human genome is made up of approximately 6 billion base pairs of DNA per cell. In 2001, the human genome project revealed that protein-coding sequences occupy only less than 2% of the genome, whereas repetitive DNA sequences, such as transposable elements and tandem repeats, account for more than 50% of the genome (Lander et al. 2001; Padeken et al. 2015). Transposable elements that account for about 45% of the genome and are interspersed across the genome consist of at least 2 copies of DNA segments flanked by specific repeat sequences. Tandem repeats that are preferentially present at specific regions such as centromeres or telomeres consist of multiple copies of DNA segments that are present adjacent to each other in either directly or inverted orientation.

The presence of repetitive sequences is prone to cause replication fork stalling and collapse because of DNA secondary structures, such as hairpin or cruciform (Mirkin and Mirkin 2007). Stalled or collapsed replication forks can induce double-strand breaks (DSB), which is

repaired by non-homologous end joining (NHEJ) or homologous recombination (HR) (Fig. 2). NHEJ involves the re-ligation of the two broken ends of the chromosome and is often accompanied by a gain or loss of a few nucleotides. On the other hand, HR is a high-fidelity repair pathway because it usually copies allelic region on sister chromatids as a repair template. A resected 3' single-stranded DNA (ssDNA) overhang invades into template duplex DNA, forming a displacement loop (D-loop), and synthesizes DNA to copy genetic information from the intact donor duplex to the broken chromosome. After DNA synthesis, the downstream of HR is largely classified into two manners: conservative and non-conservative manners. A conservative HR, which is a potentially error-free system, includes a non-crossover (NCO). On the other hand, a non-conservative HR, which is a potentially mutagenic system, includes a crossover (CO) and a break-induced replication (BIR). The synthesized DNA strand captures the second end and creates a double Holliday Junctions (dHJs). Resolution of dHJs generates either a NCO product that maintains the original sequence or a CO product that exchanges the flanking sequences. Alternatively, dissociation of D-loop before the formation of dHJs leads to synthesis dependent strand annealing (SDSA), which generates only NCO products (Nassif et al. 1994). If one-end of the broken strand was lost, DSBs are repaired by BIR. DNA synthesis extends up to hundreds of kilo-bases (kb) until the telomere.

When DSBs occur within repetitive sequences, they can be repaired by HR between non-allelic region (non-allelic HR: NAHR). Importantly, the combination of NAHR and non-conservative HR causes genome rearrangements. Crossovers between directly oriented DNA repeats on the same chromosomes (intrachromosome) or between the sister chromatids (interchromosome) result in deletion or duplication of DNA segments. Crossovers between DNA repeats present at non-homologous chromosomes produce translocations. BIRs using non-allelic repeats as a donor produce deletions, duplications or non-reciprocal translocations. Although many reports have argued that the choice of recombination pathway is crucial for maintaining genome integrity at repetitive sequences, what regulates recombination remains unclear.

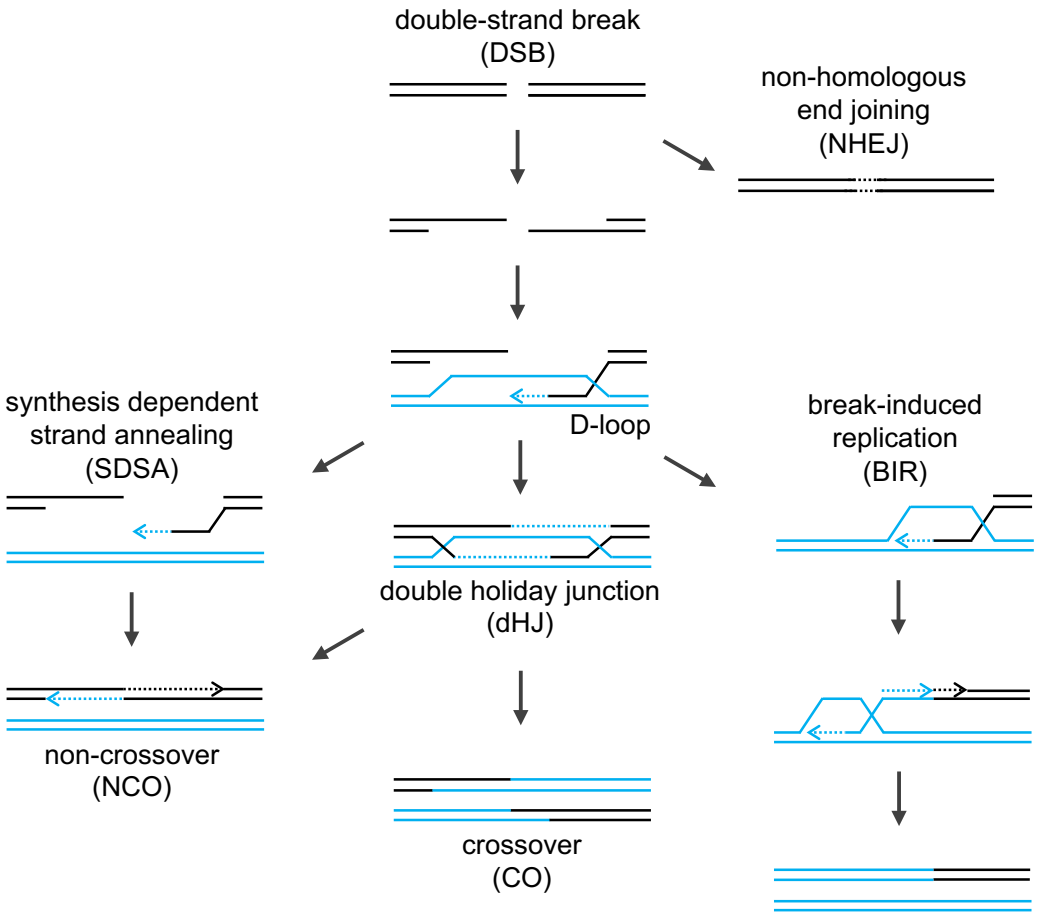


Fig. 2 Pathways of double strand break (DSB) repair.
DSBs are repaired by non-homologous end joining (NHEJ) or homologous recombination (HR). NHEJ involves ligation of the broken ends, with little or no base pairing. In synthesis dependent strand-annealing (SDSA), the newly synthesized strand dissociates from D-loops and results in a non-crossover (NCO) outcome with no change to the template DNA. The double holiday junction (dHJ) involving second end capture can be cleaved by resolvase to produce both NCO or crossover (CO) outcomes. Break-induced replication (BIR) involves DNA strand synthesis and results in loss of heterozygosity.

4. Unique chromatin structures at centromeres

Centromere is essential for accurate segregation of chromosomes. It consists of unique chromatin structures consisting of kinetochore chromatin and flanked by heterochromatin that ensure proper chromosome segregation (Fig. 3A). A centromere-specific histone H3 variant CENP-A, which provides the physical basis for the attachment of microtubules, forms an epigenetic mark of the kinetochore chromatin at all active centromeres. Heterochromatin, which is characterized by the methylation of the 9th lysine of the histone H3 (H3K9), is an important condensed form of chromatin that silences transcription. The H3K9me is important to recruit cohesin that facilitates sister chromatid cohesion. In many eukaryotes, centromeres consist of repetitive sequences (Muller and Almouzni 2017). There is no conserved DNA sequence motif that determines centromere function, but the centromeres in many eukaryotes consist of repetitive DNA sequences, which could be in either inverted or direct configurations. In human cells, centromeres harbor α-satellite DNA repeats that consist of a 171 base-pair (bp) monomer organized into higher-order arrays extending 0.3-5 mega-bases (Mb) in size and are assembled into kinetochore chromatin. Flanking these higher-order arrays are α-satellite monomers that lack both sequence homogeneity and periodicity, and are assembled into heterochromatin (Alexandrov et al. 2001). The centromeres in the fission yeast, *Schizosaccharomyces pombe*, are about 30-110 kb in size. It consists of a single-copy of non-repetitive central core sequence (*cnt*) surrounded by large inverted innermost repeats (*imr*) which are flanked by tandem copies of outermost elements that are composed of dg and dh repeats. The outermost repeats are flanked by inverted repeat centromere (*irc*) sequences. The CENP-A homolog Cnp1 binds to *cnt* and a part of *imr* sequences, while the flanking repeats are enriched with H3K9 methylation. In mice, *Mus musculus domesticus*, all 40 pairs of chromosomes with the exception of the Y sex chromosome are telocentric chromosomes, in which the centromere is located at the terminal end of the chromosome. The centromeres consist of two types of repetitive DNA, minor and major satellite. Minor satellite DNA comprises many copies of a 120 bp repeating unit, organized into tandem arrays of ~600 kb, and is associated with the function of the centromere (Wong and Rattner 1988; Joseph et al. 1989; Guenatri et al. 2004). Major satellite DNA located at the pericentromeric

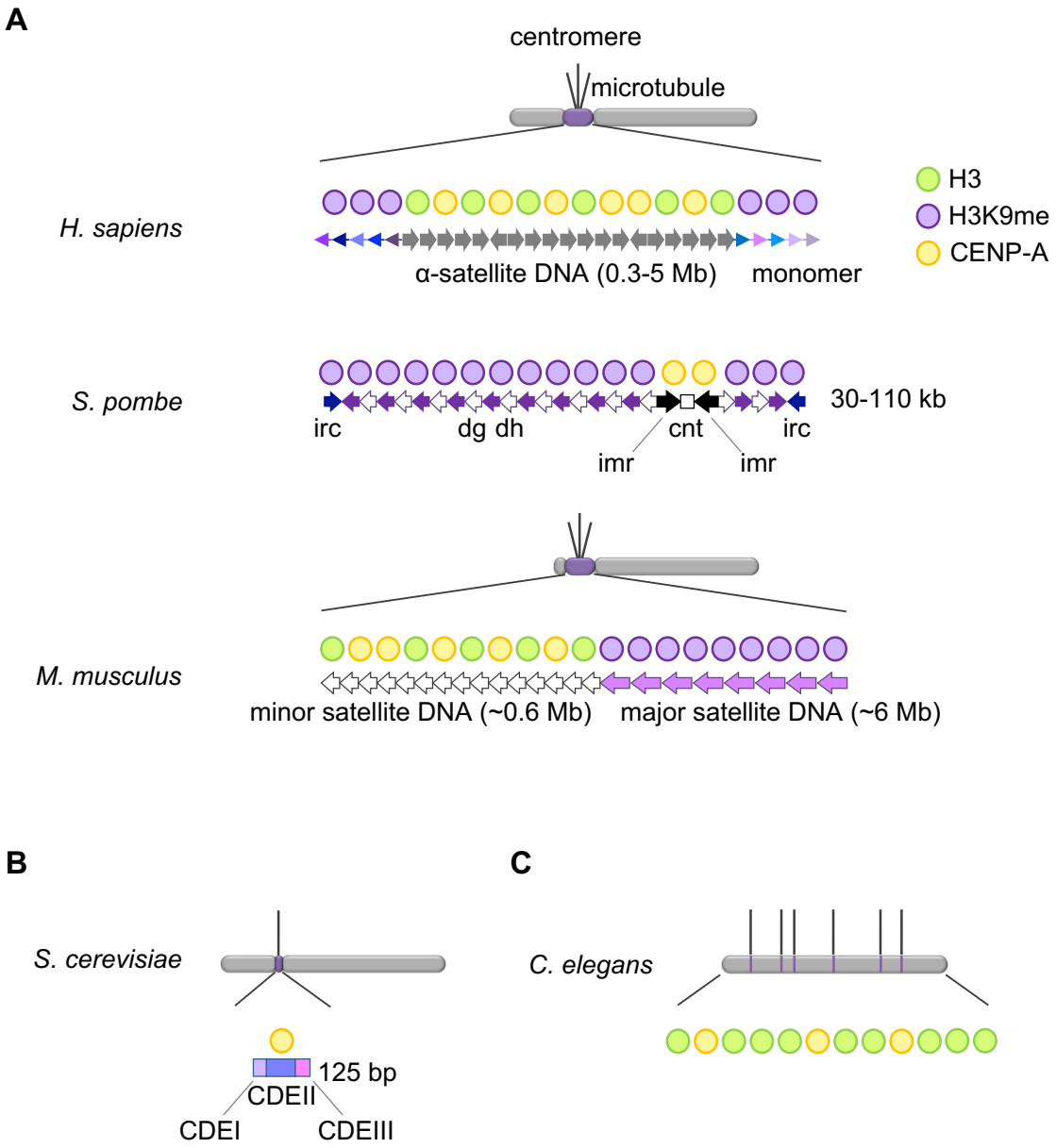


Fig. 3 The unique chromatin structure and sequences of centromeres

A. Illustration is the architecture of centromeres in human, fission yeast, and mice. Spindle microtubules attach to central region of the centromere, which assembles into kinetochore chromatin marked by the binding of CENP-A nucleosome. It is surrounded by heterochromatin marked by H3K9 methylation. In most of the eukaryotes including these organisms, centromeres consist of repetitive sequences. **B.** Budding yeast has a point centromere that encompasses 125 bp DNA and is composed of three unique elements CDEI, CDEII, and CDEIII. This is wrapped around a single nucleosome containing the CENP-A homolog Cse4. Each kinetochore makes only one stable microtubule attachment. **C.** Nematode holocentromere is made of ~700 individual centromeric sites distributed along the length of the chromosomes. Note that heterochromatin assembles on centromeres only when the centromeres consist of repetitive sequences.

regions, is made up of many copies of a 234 bp monomer organized into tandem arrays spanning ~6 Mb and composes the heterochromatin. Thus, centromeres in these organisms contain repetitive sequences and are assembled into kinetochore and heterochromatin. Such centromeres are referred to as "regional centromeres". On the other hand, the "point centromeres" in the budding yeast, *Saccharomyces cerevisiae*, encompass only 125 bp that are composed of three unique elements *CDEI*, *CDEII*, and *CDEIII* (Fig. 3B). This is wrapped around a single nucleosome containing the CENP-A homolog Cse4. Each kinetochore makes only one stable microtubule attachment. The centromere structure in nematode, *Caenorhabditis elegans*, have unique centromere structures distinct from many organisms (Fig. 3C). It contains holocentric chromosomes that recruit and assemble centromeric proteins along their length before cell division. Note that heterochromatin assembly is dispensable for centromere function in these organisms. Consistent with this, heterochromatin is not always formed at the centromeres. In fission yeast strain CBS2777 and pathogenic fungus *Candida lusitaniae*, no heterochromatin or transcriptional silencing was observed at the centromeres that were devoid of repeat sequences (Brown et al. 2014; Kapoor et al. 2015). In chicken DT40 cells, heterochromatin is assembled at the repetitive centromeres but not at the non-repetitive centromeres (Shang et al. 2013). The relevance between repetitive sequences and heterochromatin assembly at centromeres suggests that heterochromatin has an important role especially when the centromeres consist of repetitive elements.

5. Centromeres and genome rearrangements

Several lines of evidence indicate that centromeres are the preferential break sites that result in several kinds of genetic diseases including cancers (Martinez and van Wely 2011; Barra and Fachinetti 2018) (Fig. 4). Over 60% of oral squamous cell carcinoma are caused by GCRs whose breakpoints are present at centromeres (Hermsen et al. 1996). A isochromosome formation of sex chromosome X, a structural abnormality in which the arms of the chromosome are mirror images of each other, is the most common GCR event generated from centromeres. The isochromosome formation of sex chromosome X accounts for more than 20% of the cases of Turner syndrome

chromosome	Type of tumor
18, 8, 7	Breast cancer
1, 18	Melanoma
18	Colorectal
18, 8	Pancreatic
Several	Squamous cell carcinoma
Several	Adenocarcinoma
Several	Oral squamous cell carcinoma
Mostly 3, 5	Head and neck squamous cell carcinoma
1	Non-Hodgkin's lymphoma
1	Multiple myeloma
1	Ovarian cancer
10, 14, 7, 21	Cervical squamous cell carcinoma
Mostly 17	Hematologic malignancies
Several	Prostate cancer
17	Male breast cancer
1, 8	Hepatocellular carcinoma
1, 7	Acute myeloid leukemia

(modified from Barra and Fachinetti et al., 2018)

Fig. 4 List of human tumors that present breakpoints around the centromere region.

(1/2,000 female births), which is a development disorder in females caused by partial or complete monosomy of chromosome X (Palmer and Reichmann 1976). Thus, the presence of tandem repeats at centromere could be subject to NAHR giving rise to GCRs (Wolff et al. 1996).

6. Fission yeast as a model system of GCR and recombination mediated by centromere repeats.

Fission yeast is one of the generally utilized model organisms to study eukaryotic systems. It has several advantages for my studies. First, fission yeast is a powerful tool for using genetic techniques such as making knock out or point mutant strains, because it usually grows as haploid and the cell growth is quite fast as compared to the mammalian cells: ~2 hours compared to ~24 hours in human. Second, the chromatin structure of the fission yeast centromere is comparable to that of metazoan centromere (Fig. 3A). Third, ~70% of fission yeast protein-coding genes are conserved in humans (Hayles and Nurse 2018). Thus, the role of fission yeast centromere chromatin in GCR and recombination is possibly conserved in other eukaryotic organisms. Considering them, using fission yeast as a model system is quite adequate to understand the regulation of GCR and recombination mediated by centromere repeats.

Part I

Suppression of centromeric GCRs through repressing Tfs1/TFIIS-dependent transcription at heterochromatin

ABSTRACT

Most eukaryotic genomes are composed of repetitive sequences, that can potentially cause gross chromosomal rearrangements (GCRs) through recombination between non-allelic homologous sequences. Centromeres that are essential for chromosome segregation, also consist of repetitive sequences. GCRs that occurred around centromeres cause Robertsonian translocation or the formation of isochromosomes, in which the arms of the chromosome are mirror images of each other. Repetitive elements, including centromeres, are preferential sites for the assembly of heterochromatin structure that is characterized by the methylation of the 9th lysine of the histone H3 (H3K9). The chromodomain proteins including heterochromatin protein 1 (HP1) bind to the methylated H3K9 and silence transcription through inhibiting the loading of RNA polymerase II (RNAPII). Heterochromatin is important for proper chromosome segregation by facilitating sister chromatid cohesion. However, a role of heterochromatin in centromere integrity is unclear. Here, using fission yeast, I found that heterochromatin suppresses GCRs that are mediated by centromere repeats. The deletion of Clr4/Suv39, a sole H3K9 methyltransferase in fission yeast, increased the formation of isochromosomes whose breakpoints are present in centromere repeats. Mutations in the SET domain of Clr4 that are required for its catalytic activity and amino acid substitution of H3K9 also increased GCRs, suggesting that Clr4 suppresses GCRs through H3K9 methylation. Remarkably, a mutation in RNAPII reduced RNAPII chromatin binding and GCRs in *clr4*Δ cells, indicating that the repression of RNAPII bypasses the requirement of Clr4 to suppress GCRs. RNAPII frequently pauses when it encounters with a number of proteins such as nucleosomes, and it requires transcription elongation factors for efficient transcription. Tfs1/TFIIS, a one of the transcription elongation factors, specifically rescues arrested and backtracked RNAPII. I identified what kind of transcription triggers GCRs, and found that the deletion of Tfs1 specifically reduces GCRs in *clr4*Δ cells, without changing chromatin binding levels of RNAPII. *tfs1*Δ only slightly reduced non-coding RNA derived from centromere repeats. These results suggest that the restart from arrested and backtracked RNAPII rather than RNAPII

chromatin binding per se trigger GCRs that are medicated by centromere repeats. Thus, I conclude that heterochromatin maintains centromere integrity through repressing Tfs1/TFIIS-mediated transcription.

INTRODUCTION

Heterochromatin in mammals

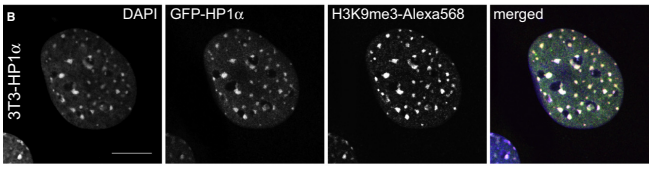
The chromatin structure can be categorized into two, euchromatin and heterochromatin, that are distinguished cytologically by their differences in compaction through the cell cycle (Fig. 5A). Heterochromatin is generally condensed and transcriptionally silenced region of the chromosome. In 1970, it was first demonstrated that mouse satellite DNA repeats are packaged in a condensed heterochromatin (Jones 1970; Pardue and Gall 1970), and subsequent works have indicated that heterochromatic domains are rich in repetitive sequences, including telomeres and ribosomal DNAs. In most organisms, heterochromatin is characterized by di- or tri-methylation of H3K9 (H3K9me2 and H3K9me3, respectively). The heterochromatin protein 1 (HP1) bind to methylated H3K9 using their N-terminal chromodomain and self-associate through their C-terminal chromo shadow domain. HP1 serves as a platform for recruitment of several factors, which lead to the formation of compact domains that silence DNA transactions such as transcription. A recent study showed that phosphorylation of the N-terminal domain of HP1α, one of the three paralogs of human HP1, enables to involve phase-separation of HP1-bound chromatin domains into liquid-like foci with distinct physical properties that are critical for silencing (Larson et al. 2017).

Heterochromatin in fission yeast

Similar to mammalian heterochromatin, the methylation of H3K9 and the localization of HP1 homologs Swi6 and Chp2 are important for heterochromatin assembly and transcriptional silencing in fission yeast (Grewal and Jia 2007) (Fig. 5B). Swi6 is required for transcriptional silencing as shown in human HP1α (Shimada et al. 2009; Larson et al. 2017).

A hypo-acetylated state of histones is another feature of the heterochromatin. Deacetylation of H3K9 provides catalytic substrates for Clr4/Suv39 methyltransferase. As alanine substitution of H3K14 inhibits the recruitment of Clr4, deacetylation of H3K14 is also thought to be essential for heterochromatin nucleation (Mellone et al. 2003; Alper et al. 2013). Swi6 and Chp2 recruit Snf2-histone deacetylase (HDAC) repressor complex (SHREC) containing Clr3

A



(Muller et al., 2009)

B

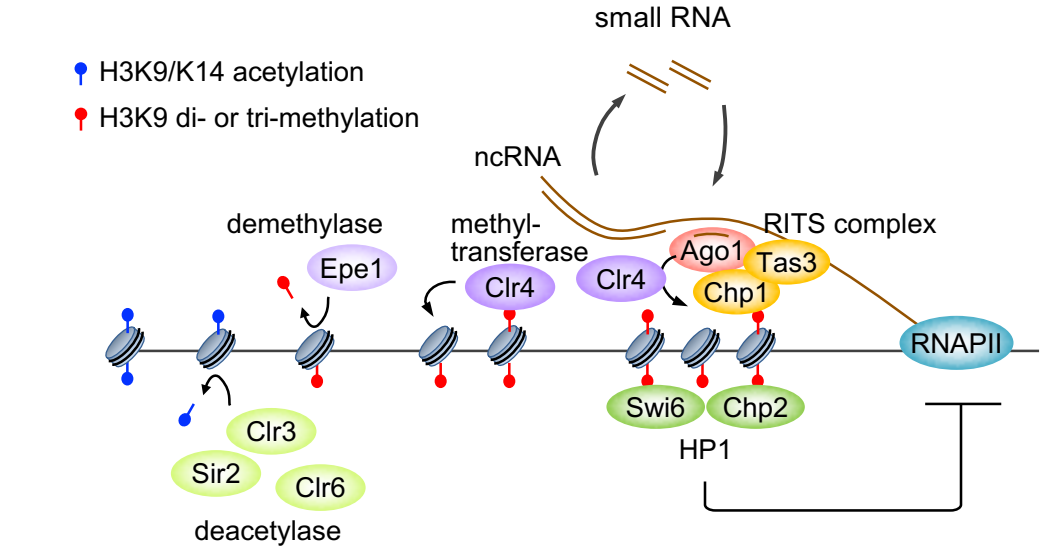


Fig. 5 Assembly of heterochromatin in mammals and fission yeast.
A. Pictures show that GFP-HP1α and tri-methylation of H3K9 are enriched in pericentric heterochromatin foci that are identified by concentrated DAPI staining in mouse NIH3T3 fibroblast cells. **B.** Illustrated is the heterochromatin assembly on centromere repeats in fission yeast. Histone deacetylases such as Sir2, Clr3, and Clr6 catalyze H3K9 or K14, providing methylation sites for Clr4 methyltransferase. The H3K9me2/3 mark is recognized by the proteins that contain the chromodomain, Chp1, Clr4, Swi6, and Chp2. The RITS complex, which consists of Ago1, Chp1, Tas3, and small RNA localizes to centromere repeats through Chp1 and through base–pairing between Ago1-captured small RNAs and nascent transcripts. The RITS complex recruits CLRC complex including Clr4, thereby facilitating H3K9me2. Clr4 binds to H3K9me2 using its chromodomain and make a transition from H3K9me2 to H3K9me3. Swi6 and Chp2 form oligomers and suppress RNAPII recruitment. At heterochromatin boundary with euchromatin, Epe1 demethylase antagonizes H3K9 methylation by Clr4 to prevent from heterochromatin expansion.

deacetylase that targets histone H3K14 (Motamedi et al. 2008). Swi6 also interacts with the other HDAC Clr6, which carries broad substrate specificity for lysine residues on both H3 and H4 (Bjerling et al. 2002). Clr3, Clr6, and the other HDAC Sir2 that deacetylates H3K9 at centromeres, deacetylate H3K9 or K14, thereby facilitating H3K9 methylation (Shankaranarayana et al. 2003; Yamada et al. 2005; Hansen et al. 2011). The H3K9me mark is also recognized by other proteins containing the chromodomain, RNA interference (RNAi) factor Chp1, and Clr4 (Grewal and Jia 2007). The RNAi system is required for the assembly and the maintenance of heterochromatin through a repeat-induced gene silencing and a positive feedback of Clr4 recruitment. Chp1 forms the RNA-induced transcriptional silencing (RITS) complex with Ago1, Tas3, and small RNAs that are derived from repeats, and localizes to heterochromatin through Chp1 and base–pairing between Ago1-captured small RNAs and nascent transcripts (Verdel et al. 2004; Buhler et al. 2006). The RITS complex recruits Clr4-Rik1-Cul4 (CLRC) complex (Verdel et al. 2004; Buhler et al. 2006; Bayne et al. 2010). Clr4 localizes to heterochromatin through its chromodomain and facilitates H3K9 tri-methylation (Zhang et al. 2008; Jih et al. 2017). Thus, each chromodomain protein that bind to H3K9me facilitates H3K9 methylation.

The exosome-dependent RNA degradation can contribute to transcriptional silencing. Cid14 poly(A)$^+$ polymerase, an essential component of the Trf4/Air2/Mtr4 polyadenylation (TRAMP) complex, promotes exosome-dependent degradation of RNAs including centromere transcripts (Bühler et al. 2007). Mlo3 RNA-binding protein, the homolog of budding yeast Yra1 and mammalian Aly/REF, is required for the export of poly(A)$^+$ RNA from the nucleus (Strässer and Hurt 2000; Zhou et al. 2000; Thakurta et al. 2005). Yra1 directly binds to the C-terminal domain of RNA polymerase II (RNAPII) (MacKellar and Greenleaf 2011), facilitating the transcription-coupled loading of RNA export factors. Like RNAPII, Mlo3 localizes to the gene body of the euchromatin, and it binds to centromere repeats in the absence of Clr4 (Zhang et al. 2011). Mlo3 also interacts with Cid14 and facilitates the exosome-dependent RNA degradation (Zhang et al. 2011). Loss of either Mlo3 or Cid14 restores H3K9 methylation in *ago1*Δ cells (Reyes-Turcu et al. 2011), probably via the recruitment of the CLRC complex to non-degraded nascent transcripts at the centromeres. Anti-silencing factor Epe1 contains a JmjC domain, which

is associated with histone demethylase activity, although the enzymatic activity of Epe1 has not been demonstrated *in vitro* (Tsukada et al. 2006). It counteracts H3K9 methylation (Audergon et al. 2015; Ragunathan et al. 2015). Loss of Epe1 also rescues silencing defect in *ago1*Δ cells (Zofall and Grewal 2006). Thus, the several factors are involved in the methylation of H3K9 and the transcriptional silencing at heterochromatin.

Phosphorylation of the C-terminal domain of RNA polymerase II regulates transcription

Defects in HP1 or methylation of H3K9 result in de-repression of transcription at repetitive sequences. The C-terminal domain (CTD) of Rpb1, the largest subunit of RNAPII, consists of repeats of the consensus motif Tyr1-Ser2-Pro3-Thr4-Ser5-Pro6-Ser7 (Fig. 6). The fission yeast carries 29 repeats in the CTD of Rpb1 (Eick and Geyer 2013). The dynamic phosphorylation of CTD on Ser2, Thr4, Ser5, and Ser7 regulates transcription at several stages (Harlen and Churchman 2017). Initiation of transcription requires the assembly of preinitiation complex that contains unphosphorylated RNAPII at a promoter. Then, the phosphorylation of CTD Ser5 triggers RNAPII release to enter a phase of elongation. The CTD Ser7 is also phosphorylated before elongation. After that, RNAPII pauses at the promoter-proximal region, 30-60 nucleotides downstream of the transcription starting site (TSS). RNAPII pausing at promoter-proximal regions involves negative elongation factor (NELF) and DRB-sensitivity-inducing factor (DSIF) to stabilize the paused RNAPII. This is the key rate-limiting step for transcription that can act as a quality checkpoint for transcript 5'-capping and RNAPII modification before productive elongation. Recruitment of capping enzymes requires the phosphorylated CTD Ser5. The release of RNAPII from pausing state requires the phosphorylation of CTD Ser2, NELF, and DSIF. The recruitment of Ser2-specific kinase is dependent on the phosphorylation of CTD Ser5, either in a direct or in an indirect way. The phosphorylated NELF is evicted from RNAPII and the phosphorylated DSIF becomes a positive elongation factor. After RNAPII is released from the promoter-proximal region, it commences productive elongation. Phosphorylation of CTD Ser2 is involved in splicing of nascent transcripts. At the phase of termination, phosphorylation of Ser2 and Thr4 promotes the recruitment of mRNA cleavage, polyadenylation factors, and termination

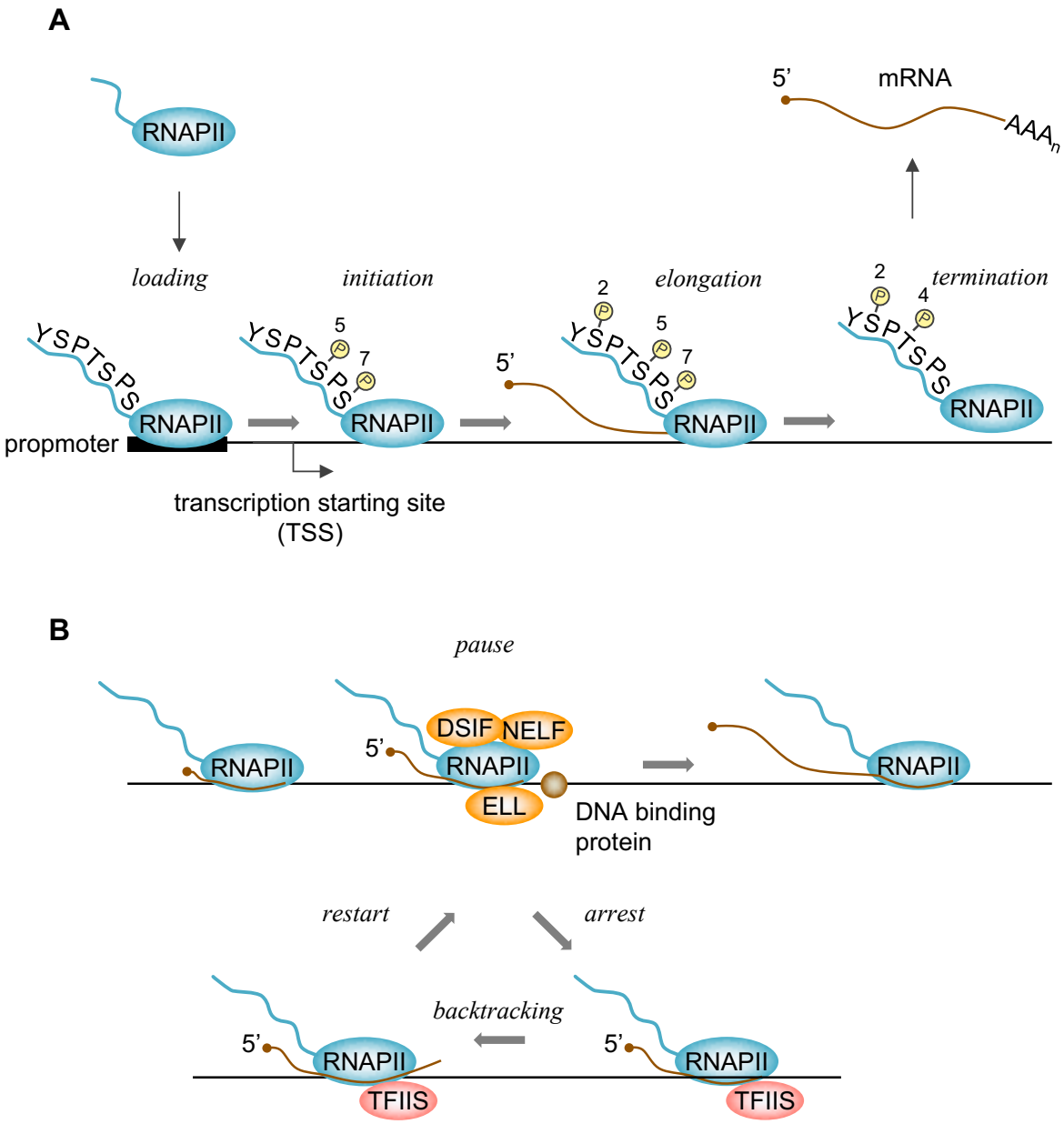


Fig. 6 Regulation of transcription by phosphorylation of RNAPII C-terminal domain (CTD) and by elongation factors.

A. Illustration is the phosphorylation of RNAPII CTD at several stages of transcription. After a loading of RNAPII harboring the unphosphorylated CTD to the promoter, Ser5 and Ser7 are phosphorylated. The phosphorylation of Ser5 allows the release of RNAPII from the promoter. During elongation, the phosphorylation of Ser5 helps recruiting enzymes that caps the 5' end of the nascent transcript. The phosphorylation of Ser2 activates elongation and splicing. After transcription, phosphorylation of Ser2 and Thr4 promotes the recruitment of mRNA cleavage, polyadenylation factors, and termination factors that release RNAPII from DNA. **B.** Illustration is the factors associated with restart from paused or arrested RNAPII. RNAPII often pauses and arrests during transcription when it encounters with DNA binding proteins such as nucleosomes. Transcriptional pause is self-reversible and is regulated by elongation factors ELL, NELF, and DSIF to alleviate the pause. Whereas, an arrested RNAPII backtracks and resumes elongation with the aid of RNA cleavage factor TFIIS.

factors that release RNAPII from DNA. A recent study has suggested that Ser7 facilitates transcription elongation at pausing sites (Sanchez et al. 2018).

Regulation of transcription elongation

Efficient transcriptional elongation must overcome several impediments, such as A/T-rich sequences, supercoiled DNA structure, DNA damaged sites, nucleosomes, replication factors, RNAPII (Gomez-Herreros et al. 2012). When RNAPII encounters with such a impediment, transcription enter into paused or arrested state. Transcriptional pausing occurs when the RNAPII halts the addition of nucleotide triphosphates (NTPs) to the nascent RNA transcript for a time before resuming elongation on its own. It is self-reversible and a natural mode of transcriptional elongation. Many factors including Lys-rich leukemia (ELL) and primary elongation factors NELF and DSIF facilitate resuming transcription. ELL directly increases the catalytic rate of RNAPII by maintaining 3'-ends of nascent RNAs in proper alignment with the catalytic site of RNAPII (Elmendorf et al. 2001). If pausing persists, such as when NTPs are removed or a physical roadblock is imposed, the pause gradually decays into arrest (Gu and Reines 1995). Transcriptional arrest can be defined as an irreversible halt to RNA synthesis followed by backtracking of RNAPII, a reverse movement of RNAPII on the DNA template. This movement results in a displacement of the 3' end of RNA from the active site and renders the enzyme transcriptionally inactive (Kettenberger et al. 2003; Wang et al. 2009; Lisica et al. 2016). They are most likely caused by combination of identifiable DNA sequences, protein factors, and the nascent transcript. Irreversibly backtracked RNAPII is frequently a target for degradation of RNAPII (Sigurdsson et al. 2010). Restarting transcription requires realigning 3' end of the RNA with the active site. Although RNAPII has a weak intrinsic endonucleolytic cleavage activity to generate a new 3' end aligned with the active site, its cleavage activity is strongly enhanced by TFIIS (Izban and Luse 1992). PAF complex, which is composed of Paf1, Ctr9, Cdc73, Rtf1, and Leo1, is a multifunctional factor that contributes to transcriptional elongation via histone modifications (Sims et al. 2004). PAF complex functions as a mediator or adaptor to facilitate other elongation factors to bind and affect RNAPII. PAF complex has been shown to interact

genetically and physically with many elongation factors, including DSIF and TFIIS, and RNAPII (Kim et al. 2010). Through these interactions, PAF and its partners coordinate the transcriptional elongation at multiple stages.

Transcription and genome instability

Although transcription is important for DNA metabolism via generating proteins or non-coding RNAs, it has potential to trigger genome instability (Kim and Jinks-Robertson 2012). Hyper-transcription of centromere satellite repeats is found in several kinds of tumor cells (Ting et al. 2011). The collision between replication machinery frequently occurs at active transcribed genes (Helmrich et al. 2013) and results in DNA breaks. Such collisions create DNA:RNA hybrid leaving a displaced ssDNA called R-loops. Because the displaced single-stranded DNA (ssDNA) is more accessible to the DNA damage reagents as compared to double-stranded DNA (dsDNA), it is likely that R-loop formation is a source of DNA damages (Santos-Pereira and Aguilera 2015). Double-strand breaks (DSBs) proximal to R-loops are often repaired by mutagenic break-induced replication (BIR) pathway through inhibiting the end resection of one of the two broken strands (Amon and Koshland 2016; Costantino and Koshland 2018). Because R-loops preferentially form at highly expressed RNAPII-transcribed genes and repetitive sequences such as telomeres or transposons (Wahba et al. 2016), suppression of transcription at repetitive sequences may have an important role in genome stability.

The roles of heterochromatin at centromere

Heterochromatin assembles on chromosome landmarks such as centromeres and telomeres. Heterochromatin associates with specific proteins and distinct histone modifications, and have important roles in the functions and the organizations of chromosomes in the nucleus. HP1 recruits cohesin to centromeres to ensure sister chromatid cohesion (Bernard et al. 2001). HP1 assists Aurora B kinase to prevent incorrect attachment of mitotic spindles to kinetochores (Abe et al. 2016). HP1 protects chromosome ends from telomere fusions (Fanti et al. 1998). These evidences support that heterochromatin is essential for centromere and telomere functions.

Whereas, Suv39 methyltransferase knockout mice exhibit developmental abnormality, chromosome aneuploidy and have predisposition to cancer (Peters et al. 2001), suggesting the role of heterochromatin in chromosome stability. However, whether heterochromatin suppresses GCRs mediated by centromere repeats remains elusive.

Here, I found using fission yeast that heterochromatin suppresses GCRs at centromeres. Deletion of Clr4 increased the formation of isochromosomes whose breakpoints were present in centromere repeats. Mutation in the SET domain of Clr4, that is required for its catalytic activity, and amino acid substitutions at H3K9 (i.e. H3K9A and H3K9R) also increased the GCR rate, suggesting that Clr4 suppresses centromeric GCRs through H3K9 methylation. Mutations in HP1 homologs, Swi6 and Chp2, and an RNAi component Chp1 synergistically increased the GCR rate, showing that both HP1 and the RNAi machinery are required to suppress GCRs. Mutations in the CTD of RNAPII impaired chromatin binding of RNAPII and reduced GCRs in *clr4*Δ cells. Strikingly, the deletion of transcription factor Tfs1/TFIIS that facilitates restart of paused and backtracked RNAPII specifically bypassed the requirement of Clr4 for GCR suppression, without changing chromatin binding levels of RNAPII. These data demonstrate that heterochromatin suppresses GCRs by repressing Tfs1/TFIIS-dependent transcription of repetitive sequences.

RESULTS

The Clr4 methyltransferase suppresses GCRs through H3K9 methylation

To know whether heterochromatin affects GCRs, I disrupted the *clr4* gene that is essential for heterochromatin assembly and determined the rate of spontaneous GCRs. Cells harboring ChL (Leu^+ Ura^+ Ade^+) were grown in EMM+UA to keep selecting Leu^+ cells (Fig. 7). After 2-3 days' incubation in EMM+UA liquid media, Leu^+ cells were plated onto YNB+UA and YNB supplemented with 5-fluoroorotic acid, which does not allow the growth of Ura^+ cells, and adenine (5FOA+A) to count the number of Leu^+ and Leu^+ Ura^- colonies, respectively. In wild type, ~90% of cells could grow on YNB+UA plate, while in *clr4*Δ strain, only ~50% of cells could grow on the plate (Fig. 8A), probably because mutation in *clr4* causes high incidence of chromosome loss (Allshire et al. 1995). On 5FOA+A plates, the *clr4*Δ strain formed a large number of colonies as compared to wild type. The Leu^+ Ura^- colonies grown on 5FOA+A plates were further inspected using EMM+A and EMM+U plates and essentially all of the colonies were found to be GCR clones of Leu^+ Ura^- Ade^-. Fluctuation tests showed that *clr4*Δ strongly increases the GCR rate (Lin et al. 1996) (Fig. 8B, gray dots).

In fission yeast, the mating-type is determined by the presence of the plus (P) or minus (M) type allele at the *mat1* locus. The *mat2P* and *mat3M* sequences, which contain P or M information, respectively, are essentially silenced by heterochromatin. The conjugation of h^+ and h^- cells enter meiosis, where Rec12/Spo11 frequently creates DNA double-strand breaks (Keeney et al. 1997) (Fig. 8C). Meiotic recombination between homologous chromosomes predominantly produces crossovers. Importantly, *clr4*Δ de-represses *mat2P-mat3M* and occasionally forms diploid cells (Ekwall and Ruusala 1994). clr4Δ also de-represses the meiotic genes in mitosis (Zhang et al. 2008). Thus, it is possible that *clr4*Δ increases GCRs through diploid formation and/or meiotic factors Rec12/Spo11 (Ellermeier et al. 2010). To exclude any possible effects of diploid formation, I disrupted the *mat2P* and *mat3M* genes. *clr4*Δ increased the GCR rate even in the absence of *mat2P-mat3M* (Fig. 8B, green dots) and *rec12* (Fig. 8B, orange dots), showing that Clr4 suppresses spontaneous GCRs in mitotic cells. Nevertheless, I used *mat2-3*Δ strains hereafter to exclude any possible effects of de-repression of the silent mating-type locus.

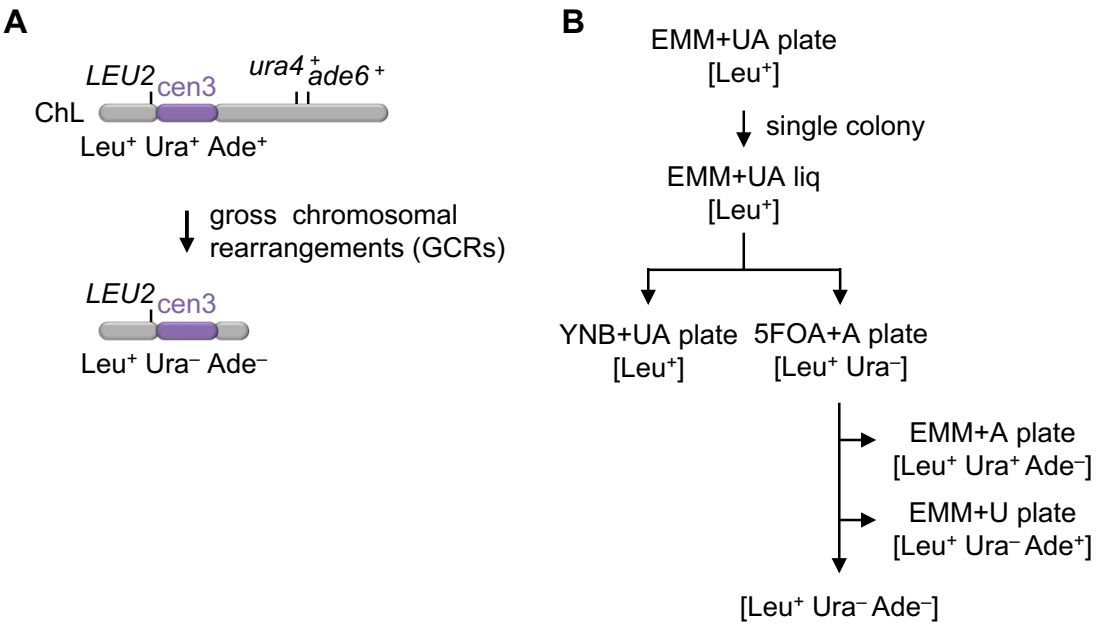


Fig. 7 An assay system to measure the frequency of spontaneous gross chromosomal rearrangement (GCR) using minichromosome ChL.

A. Illustration is an extra-chromosome ChL derived from chromosome 3. Positions of *LEU2*, *ura4*$^+$, *ade6*$^+$, and centromere 3 (cen3) are indicated. When GCRs that are associated with the loss of *ura4*$^+$ and *ade6*$^+$ take place, Leu$^+$ Ura$^+$ Ade$^+$ cells become Leu$^+$ Ura$^-$ Ade$^-$. **B.** A schematic view of a protocol to determine the rate of spontaneous GCRs. A single colony formed on EMM+UA was inoculated into EMM+UA liquid media. After 2 days incubation at 30˚C, cells were plated onto YNB+UA and 5FOA+A plates. After 6-12 days incubation, the colonies were counted to determine the number of Leu$^+$ and Leu$^+$ Ura$^-$ cells. The Leu$^+$ Ura$^-$ colonies formed on 5FOA+A plates were incubated on EMM+UA plates and then replicated onto EMM+A and EMM+U plates to confirm Ura$^-$ and to inspect Ade$^{+/-}$, respectively. The number of Leu$^+$ Ura$^-$ Ade$^-$ cells indicative of GCR was obtained by subtracting the number of Leu$^+$ Ura$^-$ Ade$^+$ cells from that of Leu$^+$ Ura$^-$ cells.

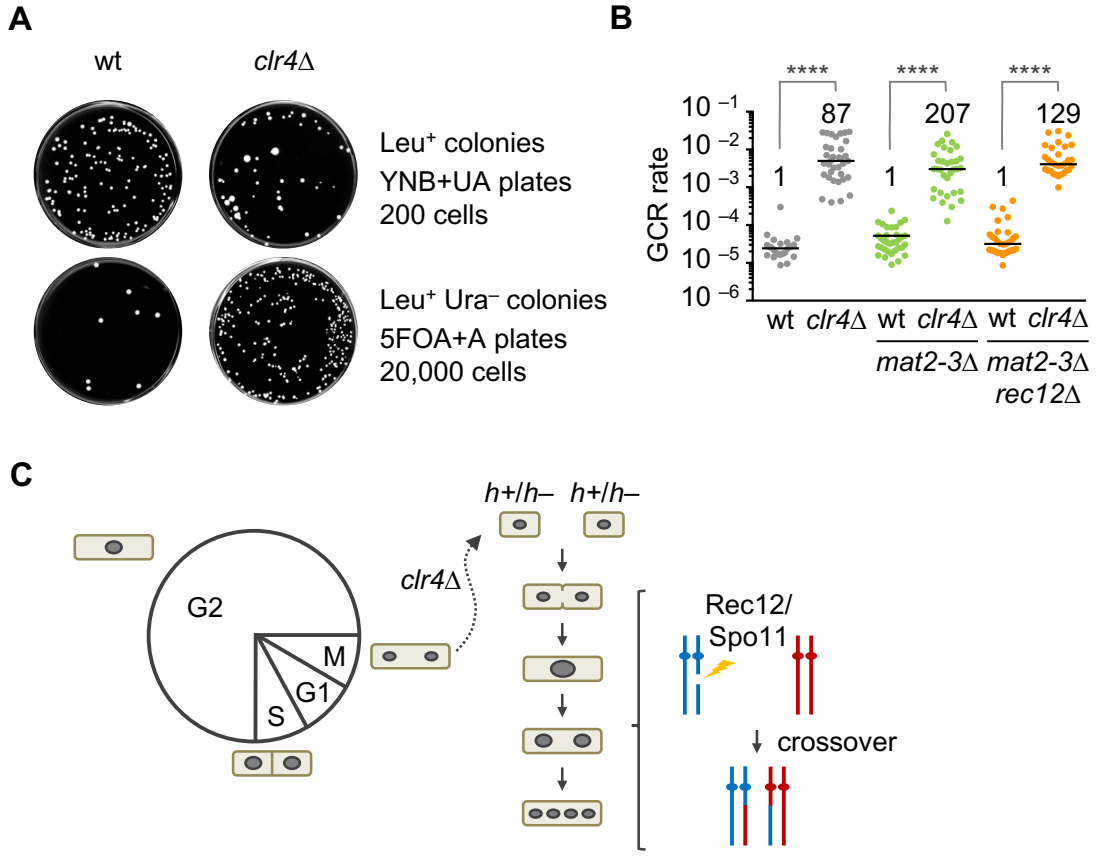


Fig. 8 Clr4 methyltransferase suppresses GCRs in mitosis.

A. Wild-type and *clr4*Δ strains (TNF5676 and 5702, respectively) grown in EMM+UA were plated onto YNB+UA (2×10^2 cells) and 5FOA+A (2×10^4 cells) media to count Leu$^+$ and Leu$^+$ Ura$^-$ colonies, respectively. Plates were incubated at 30 °C for 6–9 days. wt, wild type. **B.** GCR rates of wild-type, *clr4*Δ, *mat2-3*Δ, *mat2-3*Δ *clr4*Δ, *mat2-3*Δ *rec12*Δ, and *mat2-3*Δ *rec12*Δ *clr4*Δ strains (TNF3896, 5440, 5676, 5702, 5701, and 5766, respectively). Each dot represents the GCR rate determined using a single colony formed on EMM+UA plates in scatter plots. Lines represent the median. The GCR rate relative to that of the wild-type *clr4*$^+$ strain is indicated on the top of each column. Statistical significance of differences between pairs of strains was determined using the two-tailed Mann-Whitney test. **** $P < 0.0001$. **C.** A working hypothesis showing that the derepression of transcription at mating-type locus can cause crossover events. Heterochromatin suppresses the transcription at silent mat locus, *mat2P* and *mat3M*. The derepression of the transcription caused by *clr4*Δ may lead to entering into meiosis and the expression of Rec12/Spo11 that induces DNA double strand breaks (DSBs) in the genome. DSBs formed by Rec12/Spo11 are essentially repaired by crossover formation between homologous chromosomes.

The SET domain of Clr4 protein is responsible for its catalytic activity. The R/HφφNH (φ, hydrophobic residues) motif in the SET domain is thought to be an especially important region because it is located on the surface of the binding site of S-adenosyl-L-methionine (SAM), which is essential for the methyl transfer (Rea et al. 2000; Nakayama et al. 2001) (Fig. 9A). Indeed, each amino acid substitution in the R/HφφNH motif of Dim5, a *Neurospora crassa* homolog of Clr4, has been shown to impair its binding ability to SAM *in vitro* (Min et al. 2002). To see whether Clr4 suppresses GCRs through its methyltransferase activity, I introduced alanine substitutions at R406, N409, and H410 in the R/HφφNH motif (Fig. 9A). I performed Chromatin immunoprecipitation (ChIP) experiments to determine H3K9me2, H3K9me3, and H3 levels at centromere repeats (dg and imr3) and at a non-centromere region (*adl1*) (Fig. 9B). In wild type, H3K9me2 and H3K9me3 were specifically detected at dg and imr3, but not at *adl1*, and *clr4*Δ reduced the H3K9me2 and H3K9me3 levels, as expected. Likewise, the *clr4-set* mutation reduced H3K9me levels. Note that similar levels of H3 were observed in each strain, showing that the mutations affect histone modification rather than nucleosome occupancy. These results demonstrate that these amino acids are essential for methyltransferase activity of Clr4. I found that the *clr4-set* mutation increased the GCR rate comparable to *clr4*Δ (Fig. 9C, gray dots), showing that Clr4 suppresses GCRs through its methyltransferase activity. Although H3K9 is the major target of Clr4, it has been shown recently that Clr4 has several targets for methylation such as Mlo3 (Zhang et al. 2011; Kusevic et al. 2017). Methylation of Mlo3-K167 mediated by Clr4 is required for centromeric small RNA production and suppression of antisense RNA (Zhang et al. 2011). Neither alanine (*mlo3KA*) nor arginine (*mlo3KR*) substitution for Mlo3 methylation sites affected the GCR rate (Fig. 9C, gray dots). There are 3 copies of H3 genes in the fission yeast genome. To test the effect of H3K9 mutation on GCRs, I used H3K9A and H3K9R mutant strains, where alanine or arginine were substituted for H3K9, respectively, and 2 out of 3 H3-H4 genes in the genome were disrupted (H3-H4$^{\times 1}$) (Mellone et al. 2003). Reducing the copy number of H3-H4 by itself slightly increased the GCR rate (Fig. 9C, orange dots), probably due to low occupancy of nucleosomes in the genome (Gossett and Lieb 2012). Either alanine (H3K9A) or arginine (H3K9R) substitution further increased the GCR rate, showing the importance of H3K9

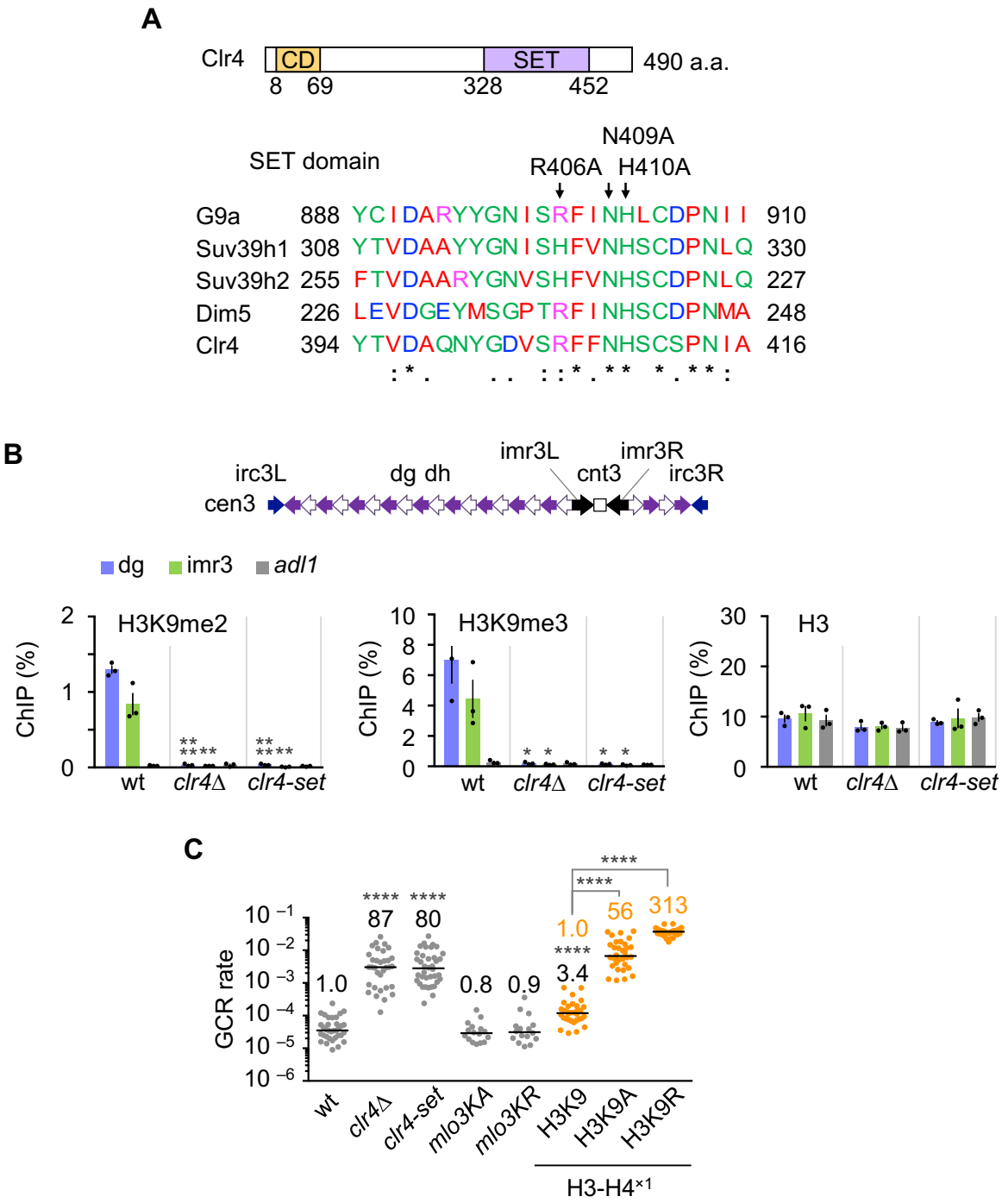


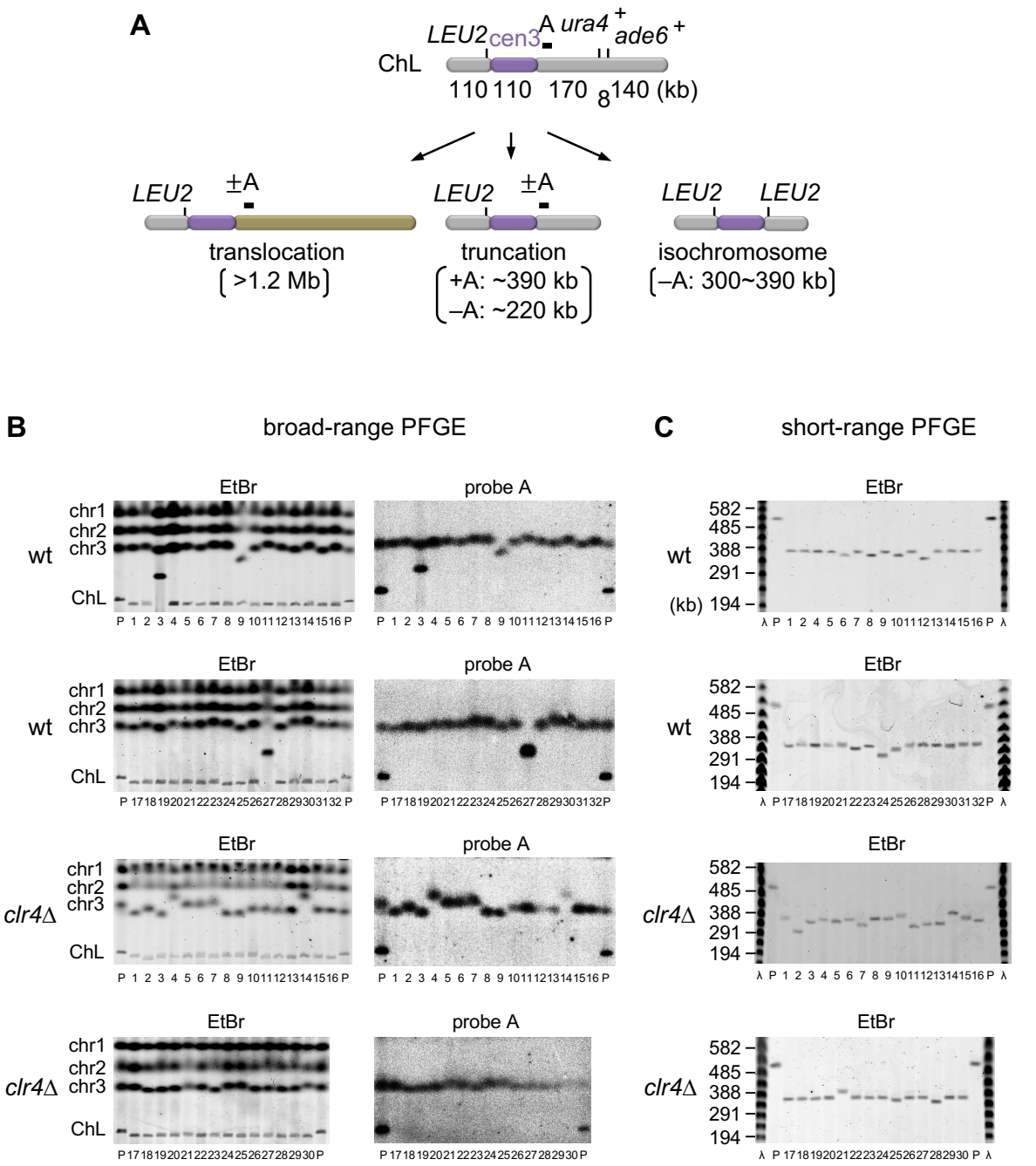
Fig. 9 The Clr4 methyltransferase suppresses GCRs through H3K9 methylation.
A. The Clr4 protein that contains the chromodomain (CD) and the SET domain and the sequence alignment of a portion of the SET domains of *Homo sapiens* G9a, Suv39h1, and Suv39h2, *Neurospora crassa* Dim5, and *Schizosaccharomyces pombe* Clr4, prepared using Clustal Omega (Sievers et al., 2011). The residues altered in the *clr4-set* mutant (R406, N409, and H410) are indicated by arrows. **B.** Chromatin immunoprecipitation (ChIP) analysis of H3K9me2, H3K9me3, and H3 at centromere repeats (dg and imr3) and at a non-centromeric region of chr2 (*adl1)* in wild-type, *clr4Δ*, and *clr4-set* strains (TNF5921, 5948, and 6169, respectively). DNA levels were quantified by real time PCR, and percentages of input DNA were obtained. Data are presented as the mean ± s.e.m. from three biologically independent experiments. Dots represent individual experiments. Statistical significance of differences relative to wild type was determined using the two-tailed Student's *t*-test. * $P < 0.05$, ** $P < 0.01$. **C.** GCR rates of wild-type, *clr4*Δ, *clr4-set*, *mlo3KA*, *mlo3KR*, H3K9, H3K9A, and H3K9R strains in the *mat2-3*Δ background (TNF5676, 5702, 6958, 6155, 6157, 5738, 6223, and 5802, respectively). In the case of H3K9, H3K9A, and H3K9R strains, the GCR rate relative to that of the wild-type H3K9 strain is also shown in orange.

in GCR suppression. These results suggest that Clr4 suppresses GCRs through H3K9 methylation. Interestingly, H3K9R substitution further increased the GCR rate as compared to H3K9A. The difference in electric charge between non-charged alanine and positively-charged arginine residues may affect GCRs (see Discussion).

Clr4 suppresses the formation of isochromosomes mediated by centromere repeats

The Leu^+ Ura^- Ade^- clones result from either translocation, truncation, or isochromosome formation (Fig. 10A) (Nakamura et al. 2008; Onaka et al. 2016; Zafar et al. 2017). Among them, isochromosomes are produced by recombination between inverted repeats at centromeres (Nakamura et al. 2008). To know whether heterochromatin suppresses GCRs that are mediated by centromere repeats, chromosomal DNAs of wild type and *clr4*Δ strains were embedded into agarose plugs, separated by broad-range pulse-field gel electrophoresis (PFGE), and stained with EtBr (Fig. 10B). In wild type, 2 out of 32 GCR products were larger than the parental ChL (Fig. 10B, wt #3 and #27), indicating that they are translocations. The other GCR products were smaller than the parental ChL and they were not detected by Southern hybridization using probe A, suggesting that they have lost the entire right arms (Fig. 10B). These small GCR products could be either truncations or isochromosomes, and they are identified depending on their size; truncations are ~220 kb, while isochromosomes are 300~390kb (Fig. 10A). Short-range PFGE showed that the small GCR products are in the range of 300~390 kb but not ~220 kb (Fig. 10C), indicating that they are isochromosomes but not truncations. Similar to wild type (30 out of 32), all of the GCR products examined in *clr4*Δ strain (30 out of 30) were isochromosomes (94% and 100%, respectively) (Fig. 10D). Given the high rates of GCRs in *clr4*Δ strain (Fig. 8B), these data show that Clr4 suppresses the isochromosome formation. Although I could not detect any translocations or truncations in the *clr4*Δ strain, it is still possible that *clr4*Δ also affect other types of GCRs than isochromosomes.

As mentioned above, Southern analysis suggested that all of the isochromosomes examined in wild type and *clr4*Δ strains have lost the entire right arm of ChL (Fig. 10B). Note that probe A recognize the region ~700bp far from cen3 repeat. To make sure that these



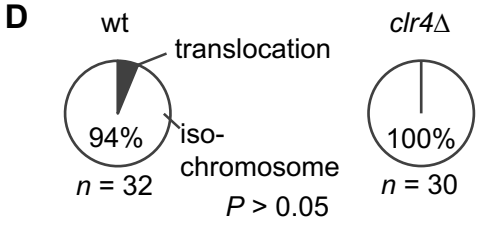


Fig. 10 Clr4 suppresses the formation of isochromosomes whose breakpoints are present in centromere repeats. (to be continued)

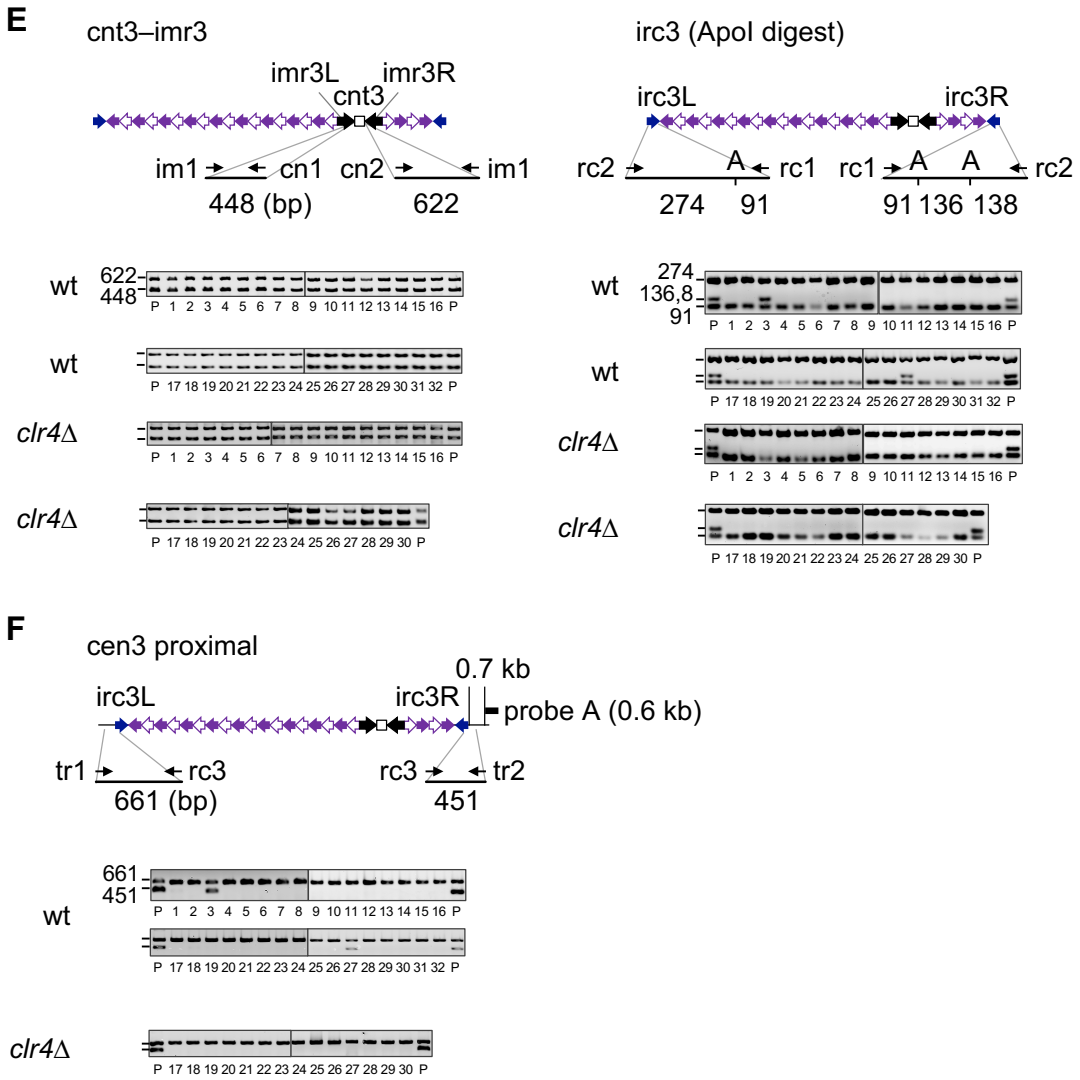


Fig. 10 Clr4 suppresses the formation of isochromosomes whose breakpoints are present in centromere repeats.

A. Illustration of the GCR products that have lost *ura4*$^+$ and *ade6*$^+$ from ChL: translocation, truncation, and isochromosome. The position of probe A used in Southern hybridization is indicated as filled box. Total size of each GCR product is shown at the bottom. **B.** Chromosomal DNAs of wild-type and *clr4*Δ strains (TNF5676 and 5702, respectively) were separated by broad-range pulse field gel electrophoresis (PFGE) and stained with ethidium bromide (EtBr). Positions of chr1, chr2, chr3, and ChL (5.7, 4.6, ~3.5, and 0.5 Mb, respectively) in the parental strain are indicated on the left of the panel. DNAs were transferred onto a nylon membrane and hybridized with probe A. P, Parental. **C.** Chromosomal DNAs were separated by short-range PFGE and stained with EtBr. Sizes of the λ DNA ladder are indicated on the left of the panel. **D.** Pie charts depict proportions of different types of GCRs. The proportion of isochromosomes is indicated. Numbers of isochromosomes and translocations were compared between wild-type and *clr4*Δ strains with the two-tailed Fisher's exact test. **E.** Breakpoints were determined by PCR reactions using GCR products recovered from agarose gel. Both sides of cnt3–imr3 junctions were amplified in the reaction containing im1, cn1, and cn2 primers. irc3L and irc3R were amplified using rc1 and rc2 primers, and the PCR products were digested by ApoI and separated by agarose gel electrophoresis. **F.** Breakpoints were confirmed by PCR using an another set of primers. Both sides of cen3 proximal regions were amplified from the GCR products of wild type and *clr4Δ*.

isochromosomes were generated by recombination between centromere repeats, I determined the breakpoints by PCR analysis of GCR products that were recovered from agarose gel. The both sides of cnt3–imr3 junctions were amplified in all the samples examined (Fig. 10E, cnt3–imr3). However, 136 and 138 bp of ApoI–digested irc3 PCR products, which indicate the right side of irc3 repeats, were not detected in all the isochromosomes (Fig. 10E, irc3 (ApoI digest)). I further confirmed that the boundary between the right side of cen3 and arm regions was specifically missed in all the isochromosomes (Fig. 10F, cen3 proximal). Because the sequence of ChL is essentially homologous with cen3, it is incapable of identifying the breakpoints of translocations between ChL and cen3 using the PCR analysis (Fig. 10E and F, wt #3 and #27). It is needed to insert a unique sequence into cen3 proximal region of ChL to identify the breakpoints of translocations (Nakamura et al. 2008). I conclude that Clr4 suppresses the formation of isochromosomes whose breakpoints are present in centromere repeats.

Both of HP1 and RNAi component are essential for full suppression of GCRs

At heterochromatin region, histone H3K9 is essentially modified by di– or tri–methylation at similar levels (Jih et al. 2017). Clr4, Swi6, Chp2, and Chp1 bind to either H3K9me2 or H3K9me3 through the chromodomain (Bannister et al. 2001; Nakayama et al. 2001; Sadaie et al. 2008; Zhang et al. 2008; Fischer et al. 2009) (Fig. 11A). Chp1, a component of RITS complex, plays a role in RNAi machinery and contributes to establishment of H3K9me2 (Schalch et al. 2009). Swi6 and Chp2, fission yeast homologs of HP1, are associated with chromatin compaction through the chromoshadow domain which is required for polymerization. To identify the chromodomain proteins that are important to suppress GCRs, I determined GCR rates of these chromodomain mutants (Fig. 11B). The *clr4-W31G* and *clr4-W41G* mutations in the Clr4 chromodomain impair Clr4 localization at centromeres and reduce only H3K9me3 levels (Nakayama et al. 2001; Zhang et al. 2008; Jih et al. 2017). As compared to *clr4*Δ, both *clr4-W41G* and *clr4-W31G* mutations only slightly increased the GCR rate, showing that H3K9me3 plays a minor role in GCR suppression. Neither *swi6*Δ nor *chp2*Δ significantly increased the GCR rate, but the *swi6*Δ *chp2*Δ double mutation increased the GCR rate, showing that Swi6 and Chp2

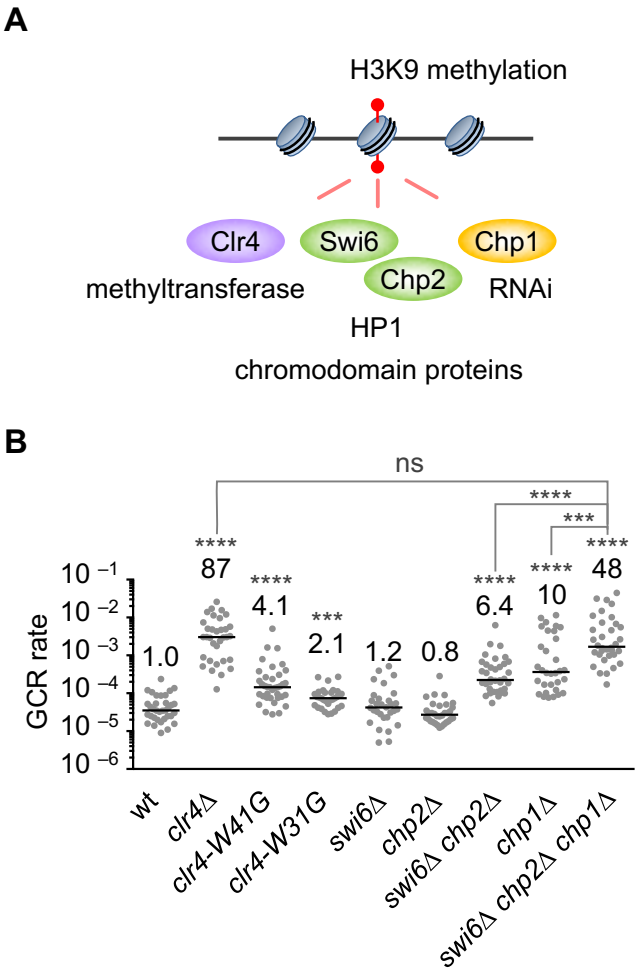


Fig. 11 Both HP1 homologs, Swi6 and Chp2, and the RNAi component Chp1 are the chromodomain proteins that are essential for full suppression of GCRs.

A. The chromodomain proteins Clr4, Swi6, Chp2, and Chp1 that bind to H3K9 methylation marks are illustrated. **B.** GCR rates of wild-type, *clr4*Δ, *clr4-W41G, clr4-W31G, swi6*Δ, *chp2*Δ, *swi6*Δ *chp2*Δ, *chp1*Δ, and *swi6*Δ *chp2*Δ *chp1*Δ strains (TNF5676, 5702, 5992, 6012, 5706, 5685, 5900, 5708, and 6151, respectively) are shown. *** $P < 0.001$; ns, not significant.

redundantly suppress GCRs. Note that the GCR rate of *clr4*Δ is 16-fold higher than that of *swi6*Δ *chp2*Δ ($P < 0.0001$), showing that H3K9 methylation suppresses GCRs only in part through HP1 homologs. Deletion of Chp1 increased the GCR rate. RNAi facilitates methylation of H3K9, providing binding sites for HP1. Interestingly, *swi6*Δ *chp2*Δ and *chp1*Δ synergistically increased the GCR rate to the level similar to that of *clr4*Δ, raising the possibility that RNAi suppresses GCRs through other than facilitating H3K9 methylation. Collectively, these results demonstrate that the chromodomain of Clr4, Swi6, Chp2, and Chp1 are required to suppress GCRs, and at least HP1s and Chp1 redundantly suppress them.

RNAi machinery and RNAi factor Ago1 play essential roles to suppress GCRs at centromeres

At heterochromatin region, transcription is essentially repressed, but centromere repeats are transcribed during a limited phase to generate precursor small RNAs (Chen et al. 2008). Rdp1, the RNA-directed RNA polymerase subunit of the RDRC complex, synthesizes dsRNA, and Dcr1 cleaves dsRNAs to produce small RNAs (Fig. 12A). Ago1 captures small dsRNAs and forms the Argonaute small interfering RNA chaperon (ARC) complex with Arb1 and Arb2 (Holoch and Moazed 2015). Then, Ago1 changes the binding partners into the RITS component Chp1 and Tas3, and localizes to the centromeres through Chp1 and base–pairing between Ago1-captured small RNAs and nascent transcripts at the centromeres (Verdel et al. 2004; Buhler et al. 2006). Note that the transcripts derived from centromere repeats could be processed into small RNA by exosomes independent of Dcr1 and Rdp1 (Halic and Moazed 2010). Thus, Ago1 can localize to heterochromatin through either Dcr1/Rdp1–dependent or independent small RNAs. To see whether the RNAi machinery is required to suppress GCRs at centromeres, I determined the GCR rate of these RNAi mutants (Fig. 12B). As expected, each deletion of RNAi factors examined in this study significantly increased the GCR rate as compared to wild type, showing that RNAi machinery is important to suppress GCRs. Surprisingly, the GCR rate of *ago1*Δ was extremely higher than that of *chp1*Δ, and even higher than *clr4*Δ. Tas3, the other component of RITS complex, partially suppressed the GCR rate as compared to Ago1. Arb1 and Arb2 also partially

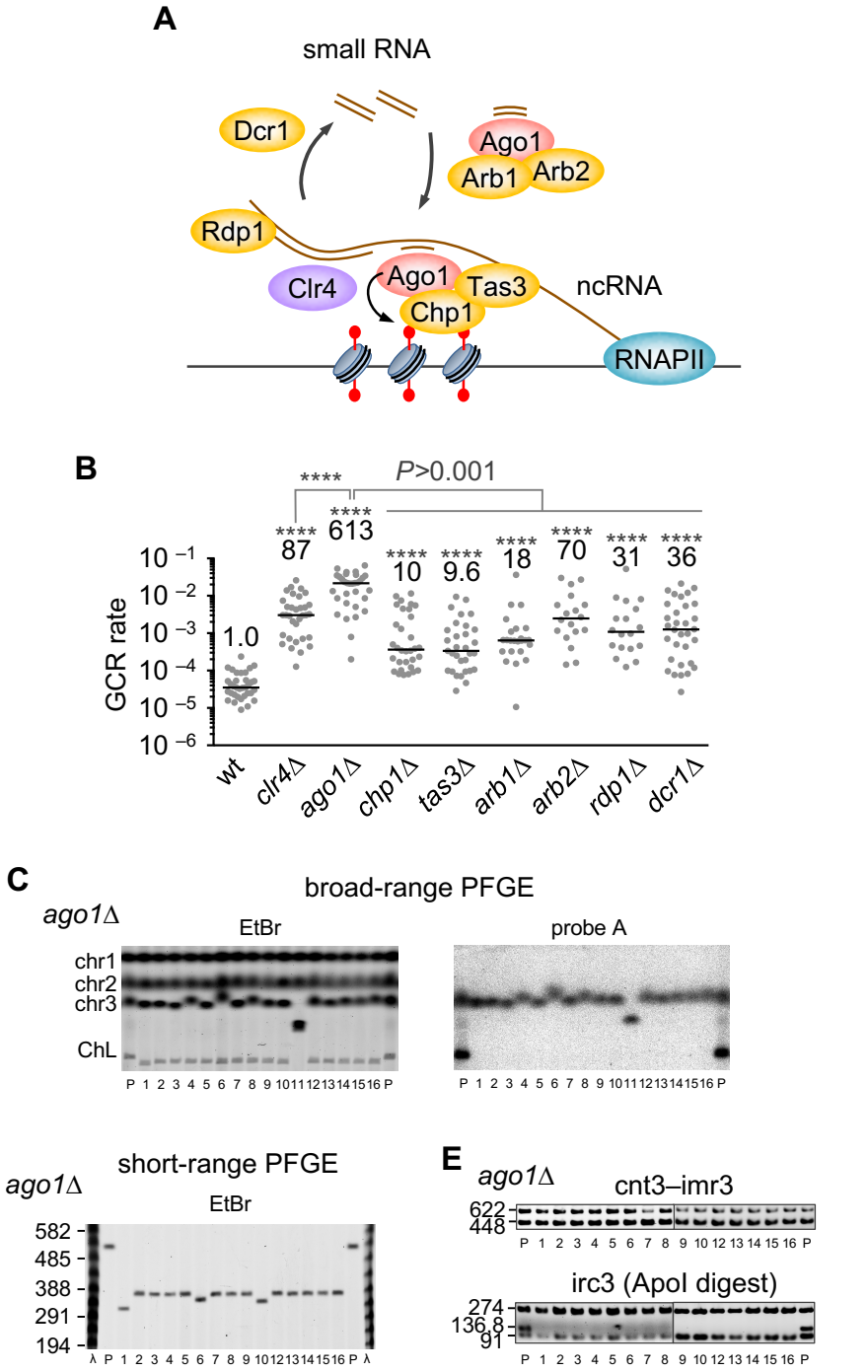


Fig. 12 RNAi machinery plays an essential role to suppress GCRs at the centromeres.
A. Illustrated is the RNAi system that utilizes small RNAs and facilitates H3K9 methylation at the centromeres. ncRNA, noncoding RNA. **B.** GCR rates of wild-type, *clr4*Δ, *ago1*Δ, *chp1*Δ, *tas3*Δ, *arb1*Δ, *arb2*Δ, *rdp1*Δ, and *dcr1*Δ strains (TNF5676, 5702, 5689, 5708, 7335, 7337, 7331, 7333, and 5687, respectively) are shown. **C.** GCR products formed in *ago1*Δ cells (TNF5689). Chromosomal DNAs were separated by broad-range PFGE and stained with EtBr. DNAs transferred onto a nylon membrane were hybridized with probe A. **D.** Chromosomal DNAs were separated by short-range PFGE and stained with EtBr. **E.** Breakpoints were determined by PCR reaction and ApoI digestion.

suppressed the GCR rate. These results suggest that Ago1 suppresses GCRs at least partially independent of RITS and ARC complexes. Rdp1 and Dcr1 also partially suppressed the GCR rate as compared to Ago1, probably due to Dcr1/Rdp1-independent pathway of small RNA production (Halic and Moazed 2010). Most of the GCR products formed in *ago1*Δ cells (15 out of 16) were the isochromosomes whose breakpoints are present in centromere repeats (Fig. 12C, D, and E). These results suggest that Ago1 not only facilitates H3K9 methylation but also plays an important role to suppress GCRs. As the effect of *ago1*Δ on GCRs is higher than that of *clr4*Δ, the additional role of Ago1 may be different from small RNA–mediated H3K9 methylation. *arb2*Δ further increased the GCR rate than *arb1*Δ ($P < 0.05$), suggesting that Arb2 has an additional function that is independent of Arb1 and is important for centromere integrity. Taken together, these results show that the RNAi machinery is required for GCR suppression at centromeres, and Ago1 plays an additional role for it.

Ago1 represses RNAPII chromatin binding to suppress GCRs at centromeres

One of the main functions of Ago1 in RNAi machinery is recruiting Clr4 to heterochromatin via interaction with Stc1, which physically links Ago1 and Clr4 complex (Bayne et al. 2010). To know whether Ago1 suppresses GCRs only by facilitating H3K9 methylation or not, I took advantages of *cid14*Δ, *mlo3*Δ, and *epe1*Δ, which rescue H3K9me2 or silencing defect in *ago1*Δ strain (Zofall and Grewal 2006; Reyes-Turcu et al. 2011). I found that *cid14*Δ does not significantly change the GCR rate of *ago1*Δ cells (Fig. 13A, blue dots). Most of the GCR products formed in *cid14*Δ *ago1*Δ cells were the isochromosomes whose breakpoints are present in centromere repeats (14 out of 16 samples) (Fig. 13B, C, and D). These results suggest that the restoration of the H3K9me2 levels is not sufficient to suppress centromeric GCRs in *ago1*Δ cells. On the other hand, *mlo3*Δ and *epe1*Δ clearly reduced the GCR rate in *ago1*Δ cells. To find the difference among these mutants, I performed ChIP experiments and determined H3K9me2, H3K9me3, and H3 levels (Fig. 13E). As expected, deletion of anti-silencing factor Epe1 increased H3K9me2 levels at centromere repeats (Trewick et al. 2007). *ago1*Δ reduced the H3K9me2 levels and *cid14*Δ, *mlo3*Δ, and *epe1*Δ restored them (Zofall and Grewal 2006; Reyes-Turcu et al. 2011).

*ago1*Δ also reduced the H3K9me3 levels, but *cid14*Δ did not restore it. On the other hand, *mlo3*Δ and *epe1*Δ restored the H3K9me3 in *ago1*Δ cells. Similar levels of H3 were observed in all the strains examined, showing that the mutations affect histone modification rather than nucleosome occupancy. These results suggest that the accumulation of non-degraded centromere transcripts that is mediated by *cid14*Δ or *mlo3*Δ can restore H3K9me2 but not H3K9me3, and *mlo3*Δ has an additional effect to restore H3K9me3. Because Mlo3 is involved in transcription as well as RNA export and degradation (Strasser et al. 2002; MacKellar and Greenleaf 2011; Zhang et al. 2011). H3K9me3 is correlated with the repression of RNAPII binding to heterochromatin (Jih et al. 2017). Thus, it is possible that Mlo3 cause GCRs through transcription. To test the possibility, I determined RNAPII chromatin binding levels. In wild type, the localization of Rpb1, a catalytic subunit of RNAPII, was limited at centromeres as compared to *adl1* (Fig. 13E, RNAPII (Rpb1)). *ago1*Δ increased RNAPII levels at dg and dh to the level comparable to that of *adl1*. Note that *ago1*Δ did not significantly increase the RNAPII levels at imr3, suggesting that an intrinsic transcription activity of imr3 is low (Cam et al. 2005). *cid14*Δ did not significantly change RNAPII levels in *ago1*Δ cells, suggesting that Ago1 acts downstream of H3K9me2 and H3K9me3 to suppress RNAPII localization at centromeres. On the other hand, *mlo3*Δ reduced RNAPII levels at dg and dh in *ago1*Δ, as expected (Reyes-Turcu et al. 2011). I found that *epe1*Δ also decreased RNAPII localization at dg and dh in *ago1*Δ, probably due to restored H3K9me3. Because *clr4-W31G,* which reduces H3K9me3 level at centromere (Jih et al. 2017), did not show marked increase in GCRs (Fig. 11B), repression of RNAPII rather than H3K9me3 is likely to be important to suppress GCRs. Collectively, *cid14*Δ, *mlo3*Δ, and *epe1*Δ restored H3K9me2, but only *mlo3*Δ and *epe1*Δ restored H3K9me3 and reduced RNAPII occupancy and GCRs in *ago1*Δ cells. These results suggest that Ago1 not only facilitates H3K9me2 and H3K9me3 but also represses RNAPII to suppress GCRs at centromeres.

RNAPII induces centromeric GCRs in the absence of H3K9 methylation

Mlo3 as well as RNAPII localizes to centromeres in the absence of Clr4 (Zhang et al. 2011). Yra1 the budding yeast homolog of Mlo3 directly binds to CTD of Rpb1 (MacKellar and Greenleaf

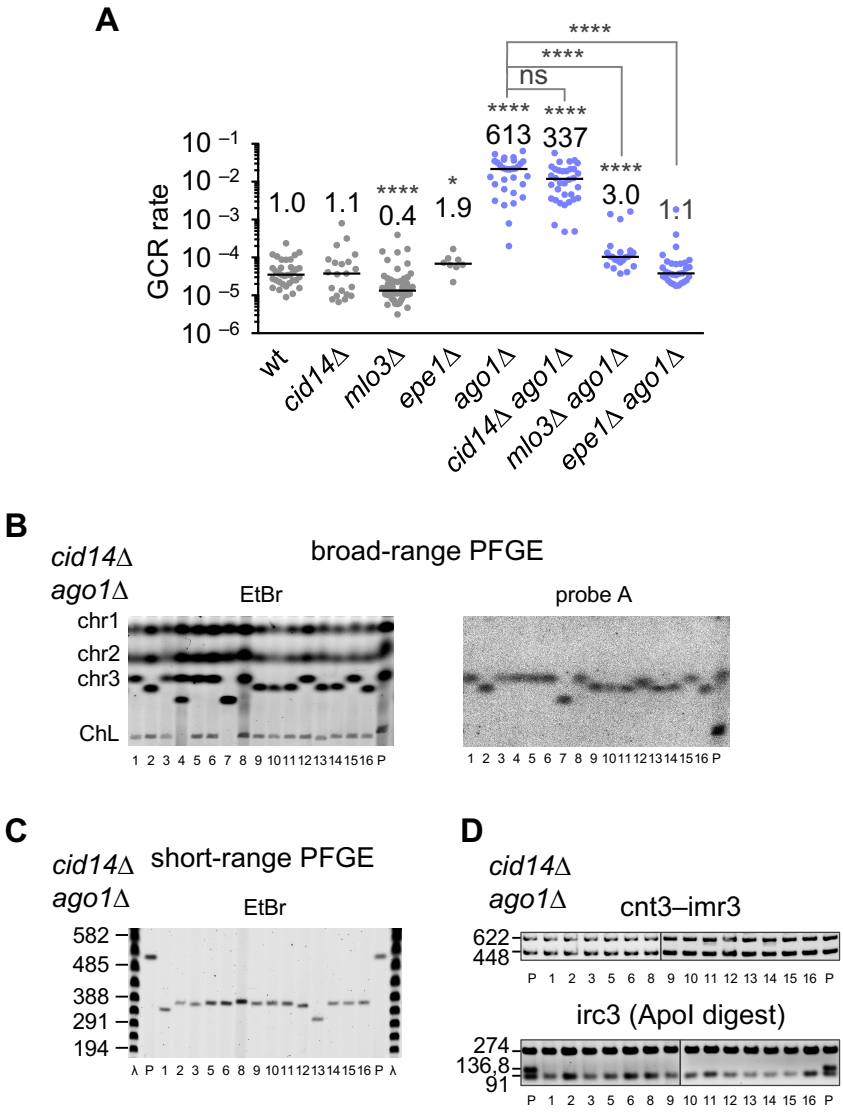


Fig. 13 *mlo3*Δ and *epe1*Δ, but not *cid14*Δ, reduces RNAPII chromatin binding and suppresses GCRs at centromeres in *ago1*Δ cells. (to be continued)

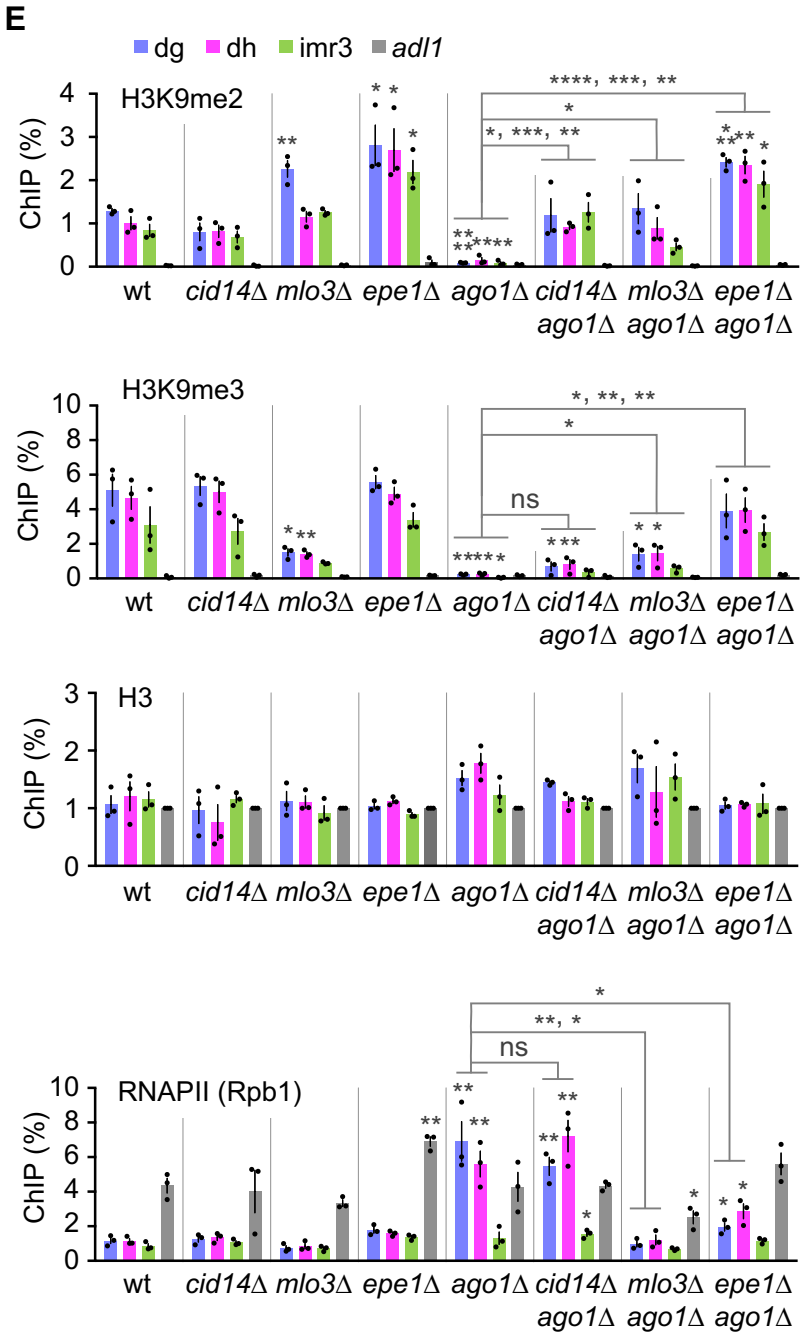


Fig. 13 *mlo3*Δ and *epe1*Δ, but not *cid14*Δ, reduces RNAPII chromatin binding and suppresses GCRs at centromeres in *ago1*Δ cells.

A. GCR rates of wild-type, *cid14*Δ, *mlo3*Δ, *epe1*Δ, *ago1*Δ, *cid14*Δ *ago1*Δ, *mlo3*Δ *ago1*Δ, and *epe1*Δ *ago1*Δ strains (TNF5676, 6153, 5764, 6109, 5689, 6411, 6188, and 7325, respectively). **B.** GCR products formed in *cid14*Δ *ago1*Δ strains (TNF6411). Chromosomal DNAs were separated by broad-range PFGE and stained with EtBr. DNAs transferred onto a nylon membrane were hybridized with probe A. **C.** Chromosomal DNAs were separated by short-range PFGE and stained with EtBr. **D.** Breakpoints were determined by PCR reaction and ApoI digestion. **E.** ChIP analysis was performed to determine H3K9me2, H3K9me3, H3 and RNAPII (Rpb1) levels at centromere repeats (dg, dh, and imr3) and at a non-centromeric region of chr2 (*adl1*) in wild-type, *cid14*Δ, *mlo3*Δ, *epe1*Δ, *ago1*Δ, *cid14*Δ *ago1*Δ, *mlo3*Δ *ago1*Δ, *epe1*Δ *ago1*Δ strains (TNF5921, 6276, 5923, 7349, 5922, 6550, 6210, and 7343, respectively).

2011). Thus, it is possible that Mlo3 facilitates the RNAPII binding to chromatin. To examine whether prevention of RNAPII from centromere is sufficient to suppress GCRs even in the absence of H3K9 methylation, I tested the effect of *mlo3* deletion on GCRs in the *clr4*Δ mutant. I found that *mlo3*Δ greatly reduced the GCR rate in *clr4*Δ cells (Fig. 14A, blue dots). Contrary to *clr4*Δ background, *mlo3*Δ did not reduce the GCR rate in *rad51*Δ cells (Fig.14A, orange dots; Fig. 14B and C), showing that *mlo3*Δ specifically affects GCRs that occur in heterochromatin deficient cells. ChIP experiments showed that *clr4*Δ increased RNAPII binding levels at dg and dh, while *clr4*Δ did not significantly change the H3 levels (Fig. 14D). *mlo3*Δ reduced RNAPII levels at dg, dh, and *adl1* in *clr4*Δ cells. As expected, *mlo3*Δ did not restore H3K9me2 and H3K9me3 in *clr4*Δ cells, suggesting that *mlo3*Δ decreases RNAPII binding at centromeres independent of H3K9me. Repression of histone acetylation is another feature of heterochromatin (Mellone et al. 2003). Sir2 and Clr3 redundantly facilitate H3K9me and silence transcription at the centromeres (Alper et al. 2013; Buscaino et al. 2013). Although neither *sir2*Δ nor *clr3*Δ significantly increased the GCR rate, the *sir2*Δ *clr3*Δ double mutation increased the GCR rate, showing that Sir2 and Clr3 redundantly suppress GCRs (Fig. 14E). Clr6, which deacetylates broad histone residues including H3K9, is essential for cell viability (Bjerling et al. 2002). To test whether Clr6 is required for GCR suppression, I used *clr6-1* temperature-sensitive mutant strain, in which G269 that is located in the consensus sequences among the HDAC proteins was substituted for aspartic acid (Grewal et al. 1998). The *clr6-1* mutation impairs deacetylation of several residues of H3 and H4, including H3K9 (Bjerling et al. 2002). I found that *clr6-1* also increased the GCR rate (Fig. 14F). These results suggest that the deacetylation of H3K9 is required for GCR suppression. Interestingly, *clr6-1* further increased the GCR rate than *sir2*Δ *clr3*Δ double mutant. The deacetylation of histone residues other than H3K9 and K14 may play an important role for GCR suppression (see discussion). On the other hand, *mlo3*Δ did not significantly change H3K9 and H3K14 acetylation levels (H3K9ac and H3K14ac, respectively), suggesting that Mlo3 facilitates RNAPII chromatin binding not through deacetylation of H3K9 or H3K14. These results suggest that Clr4 suppresses centromeric GCRs through repressing RNAPII.

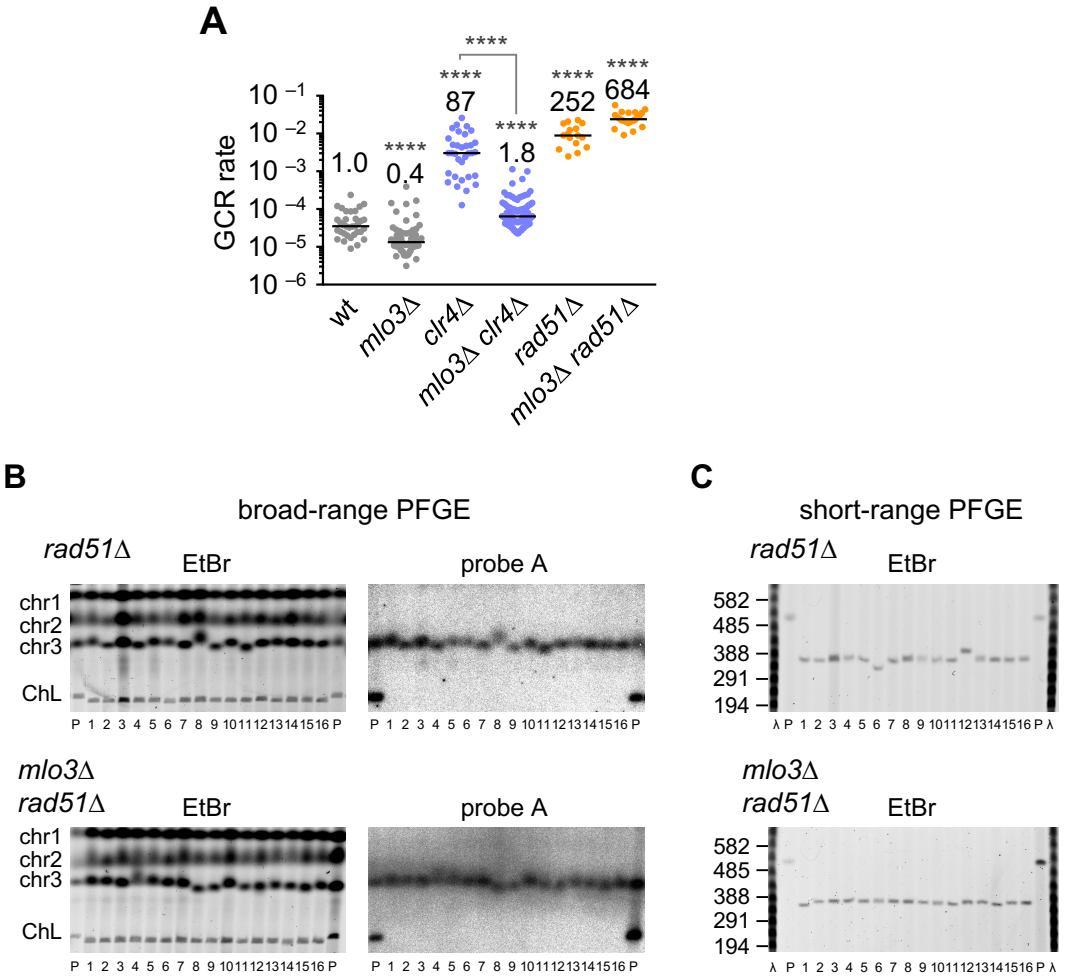


Fig. 14 Repression of RNAPII suppresses GCRs in the absence of H3K9 methylation.

(to be continued)

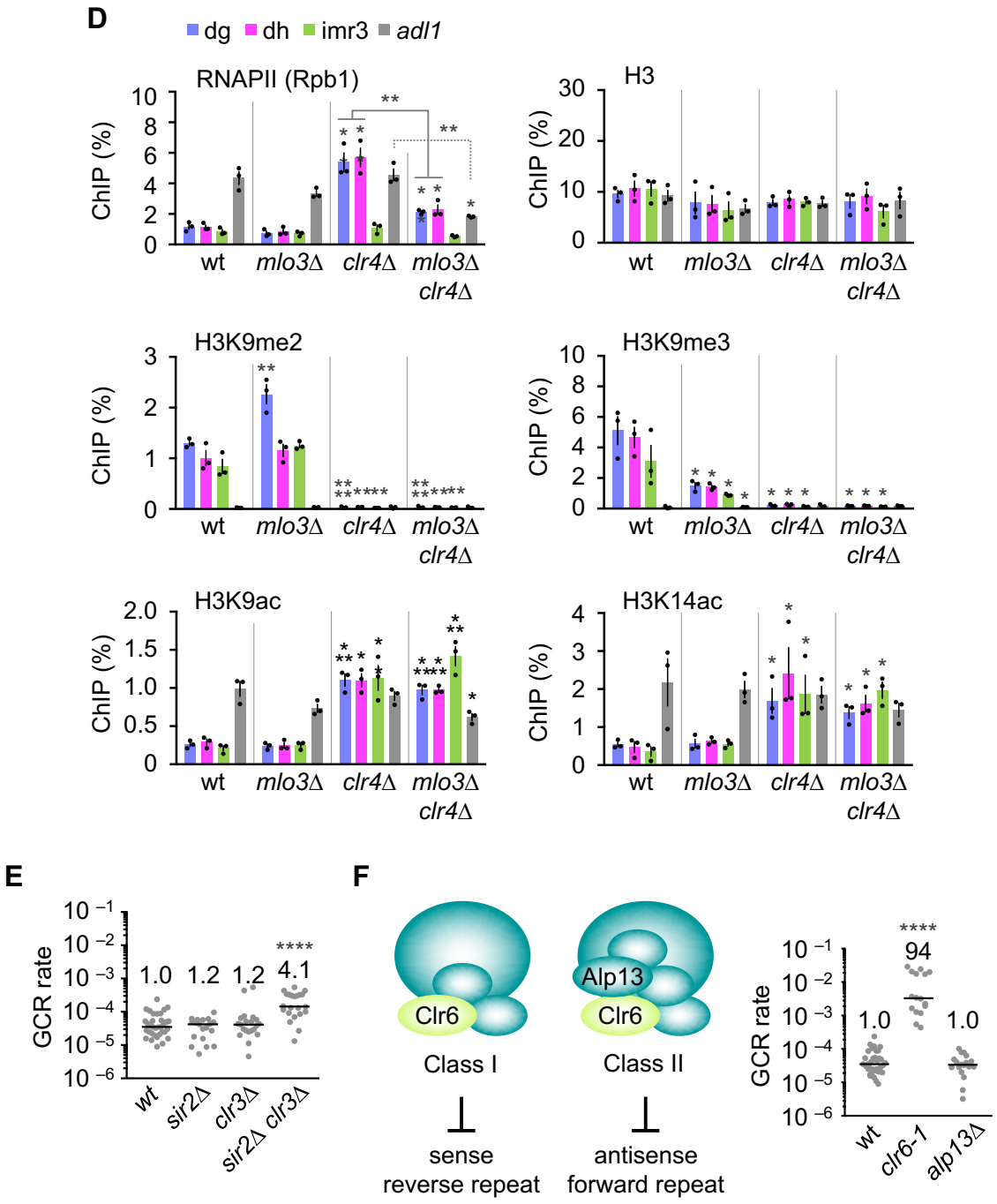


Fig. 14 Repression of RNAPII suppresses GCRs in the absence of H3K9 methylation. **A.** GCR rates of wild-type, *mlo3*Δ, *clr4*Δ, *mlo3*Δ *clr4*Δ, *rad51*Δ, and *mlo3*Δ *rad51*Δ strains (TNF5676, 5764, 5702, 5824, 6244, and 6383, respectively). **B.** GCR products formed in *rad51*Δ and *mlo3*Δ *rad51*Δ strains (TNF6244 and 6383, respectively). Chromosomal DNAs were separated by broad-range PFGE and stained with EtBr. DNAs transferred onto a nylon membrane were hybridized with probe A. **C.** Chromosomal DNAs were separated by short-range PFGE and stained with EtBr. **D.** ChIP analysis of RNAPII (Rpb1), H3, H3K9me2, H3K9me3, H3K9ac, and H3K14ac in wild-type, *mlo3*Δ, *clr4*Δ, and *mlo3*Δ *clr4*Δ strains (TNF5921, 5923, 5948, and 5925, respectively). **E.** GCR rates of wild-type, *sir2*Δ, *clr3*Δ, and *sir2*Δ *clr3*Δ strains (TNF5676, 7341, 7359, and 7357, respectively). **F.** Illustrated shows the two types of histone deacetylase complexes that share Clr6 protein. GCR rates of wild-type, *clr6-1*, and *alp13*Δ strains (TNF5676, 7345, and 5898, respectively).

Clr4 suppresses centromeric GCRs by repressing transcription that is dependent on RNAPII CTD Ser7

The largest subunit of RNAPII, Rpb1 contains repeats of the YSPTSPS heptapeptide in its CTD (Harlen and Churchman 2017) (Fig. 5). Ser7 of CTD is required for transcription of non-coding small RNAs in human cells (Egloff et al. 2007). The *rpb1-S7A* mutation, in which all the serine 7 were substituted with alanine, reduces chromatin-bound RNAs and H3K9me2 levels at centromeres in fission yeast (Cassart et al. 2012; Kajitani et al. 2017). To clarify that RNAPII is involved in centromeric GCRs, I tested the effect of *rpb1-S7A* mutation on GCRs in *clr4*Δ cells. Consistent with low levels of H3K9me2 (Fig. 15A) (Kajitani et al. 2017), *rpb1-S7A* slightly increased the GCR rate as compared to wild-type (Fig. 15B, gray dots). I found that *rpb1-S7A* reduced the GCR rate in *clr4*Δ cells (Fig. 15B, blue dots), although *rpb1-S7A* did not restore H3K9 methylation (Fig. 15A). This result demonstrates that RNAPII is involved in centromeric GCRs that occur in *clr4*Δ cells. *rpb1-S7A* reduced Rpb1 localization but not H3 levels at centromere repeats, *adl1* and highly transcribed region *act1* (Fig. 15C and D). *rpb1-S7A* also reduced chromatin binding of Rpb3 another subunit of RNAPII but not H3 levels (Kimura et al. 2002) (Fig. 15E), suggesting that CTD Ser7 of RNAPII is required for chromatin binding of the RNAPII complex. These results show that Clr4 suppresses RNAPII to prevent from centromeric GCRs.

Transcription elongation associated with Tfs1/TFIIS causes centromeric GCRs in the absence of Clr4

After RNAPII binding to chromatin, transcription proceeds with the aid of many factors, such as Tfs1/TFIIS, Ell1/ELL, Leo1/LEO1, and Spt4/DSIF (Zhou et al. 2012). Among them, Tfs1/TFIIS is the only factor that is required for the restart from backtracked RNAPII (Fig. 5). When RNAPII backtracks on DNA, Tfs1/TFIIS facilitates the restart of transcription elongation by trimming 3'-ends of nascent RNAs (Izban and Luse 1992; Kettenberger et al. 2003). To gain insights into how transcription causes GCRs in the absence of Clr4, I first tested whether Tfs1/TFIIS is involved in centromeric GCRs in *clr4*Δ cells. Interestingly, I found that *tfs1*Δ remarkably reduced the GCR

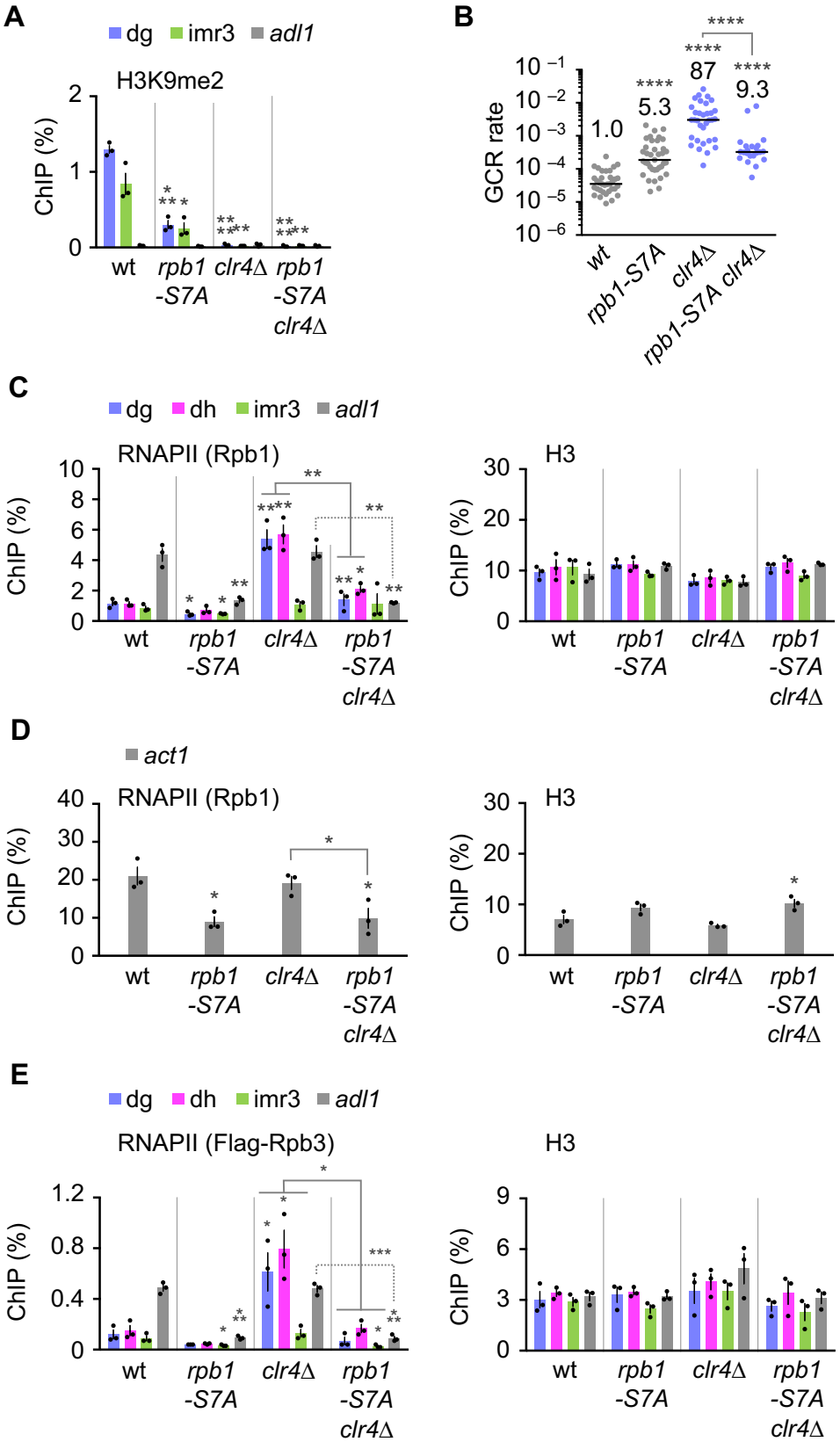


Fig. 15 Clr4 suppresses centromeric GCRs by repressing transcription that is dependent on RNAPII CTD Ser7.

A. ChIP analysis of H3K9me2 in wild-type, *rpb1-S7A*, *clr4*Δ, and *rpb1-S7A clr4*Δ strains (TNF5921, 6862, 5948, and 6864, respectively). **B.** GCR rates of wild-type, *rpb1-S7A*, *clr4*Δ, and *rpb1-S7A clr4*Δ strains (TNF5676, 6848, 5702, and 6850, respectively). **C and D.** ChIP analysis of RNAPII (Rpb1) and H3 at dg, dh, imr3 and *adl1* **(C)** and highly transcribed *act1* gene **(D)** in wild-type, *rpb1-S7A*, *clr4*Δ, and *rpb1-S7A clr4*Δ strains. **E.** ChIP analysis of RNAPII (Flag-Rpb3) and H3 in wild-type, *rpb1-S7A, clr4Δ*, and *rpb1-S7A clr4Δ* strains (TNF6931, 6943, 6933, and 6945, respectively).

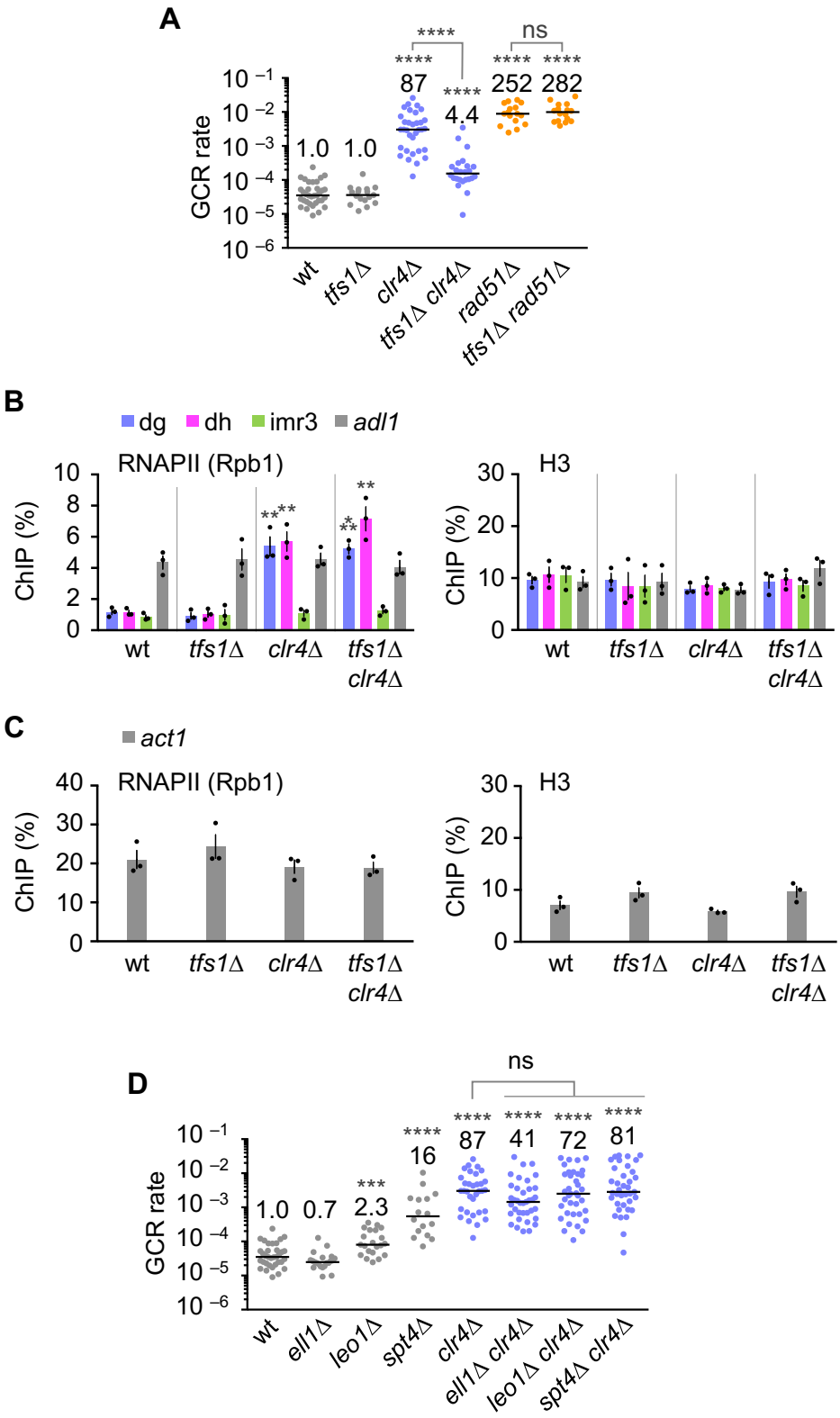


Fig.16 Clr4 suppresses centromeric GCRs by repressing transcription that is dependent on Tfs1/TFIIS.

A. GCR rates of wild-type, *tfs1*Δ, *clr4*Δ, *tfs1*Δ *clr4*Δ, *rad51*Δ, and *tfs1*Δ *rad51*Δ strains (TNF5676, 6688, 5702, 6726, 6244, and 7163, respectively). **B and C.** ChIP analysis of RNAPII (Rpb1) and H3 at dg, dh, imr3 and *adl1* **(B)** and highly transcribed *act1* gene **(C)** in wild-type, *tfs1*Δ, *clr4*Δ, and *tfs1*Δ *clr4*Δ strains (TNF5921, 6722, 5948, and 6799, respectively). **D.** GCR rates of wild-type, *ell1*Δ, *leo1*Δ, *spt4*Δ, *clr4*Δ, *ell1*Δ *clr4*Δ, *leo1*Δ *clr4*Δ, and *spt4*Δ *clr4*Δ strains (TNF5676, 7042, 7130, 7055, 5702, 7063, 7154, and 7057, respectively).

rate in *clr4*Δ cells (Fig. 16A, blue dots), but not in *rad51*Δ cells (Fig. 16A, orange dots), showing that Tfs1 is specifically involved in GCRs that occur in heterochromatin deficient cells. In contrast to *rpb1-S7A* (Fig. 15C, D, and E), *tfs1*Δ did not significantly change the RNAPII levels as well as H3 (Fig. 16B and C), showing that Tfs1-associated transcription elongation but not RNAPII chromatin binding *per se* causes GCR events in the absence of Clr4. The other transcription elongation factors facilitate RNAPII transcription. Ell1/ELL maintains 3'-ends of nascent RNAs in proper alignment with the catalytic site of RNAPII (Elmendorf et al. 2001). Leo1 is a component of the Paf1 complex that is involved in transcription elongation, termination, and histone modification (Tomson and Arndt 2013). Spt4 forms the DSIF complex with Spt5 and positively and negatively affects transcription elongation depending on the phosphorylation state of Spt5 (Wada et al. 1998; Yamada et al. 2006). Spt4 is required for transcriptional silencing at heterochromatin in budding yeast (Crotti and Basrai 2004). Deletion of Spt4 increased GCR rate, as expected (Fig. 16D). I found neither *ell1*Δ, *leo1*Δ, nor *spt4*Δ significantly reduced the GCR rate in *clr4*Δ cells, showing that the specific type of transcription elongation associated with Tfs1/TFIIS causes GCRs.

***rpb1-S7A* and *tfs1*Δ decrease the transcripts from centromere repeats and readthrough RNA**

Steady state levels of RNAPII localization detected by ChIP experiments does not necessarily reflect transcription levels (Buhler et al. 2006), thus it is possible that *tfs1*Δ alter the transcription at centromere repeats. To test the possibility, I performed Northern blotting using total RNA from yeast extracts. *clr4*Δ increased the amount of dg, dh, and less prominently imr3 RNAs (Fig. 17A). As expected, *rpb1-S7A* slightly increased dg and dh RNAs in wild-type background (Kajitani et al. 2017). In the *clr4*Δ background, *rpb1-S7A* partially reduced centromeric transcripts at dg and dh, and most prominently at imr3 where RNAPII binding and transcription levels are low (Cam et al. 2005). Similar to *rpb1-S7A*, *tfs1*Δ only slightly reduced the centromere transcripts. These results show that CTD Ser7 of RNAPII and Tfs1 are specifically required for a subset of transcription in the *clr4*Δ mutant. At the *adl1* region, a ~2.5 kb discrete band and a ~5 kb smeared band could be detected in all the strains examined; the long RNAs is likely to be the readthrough

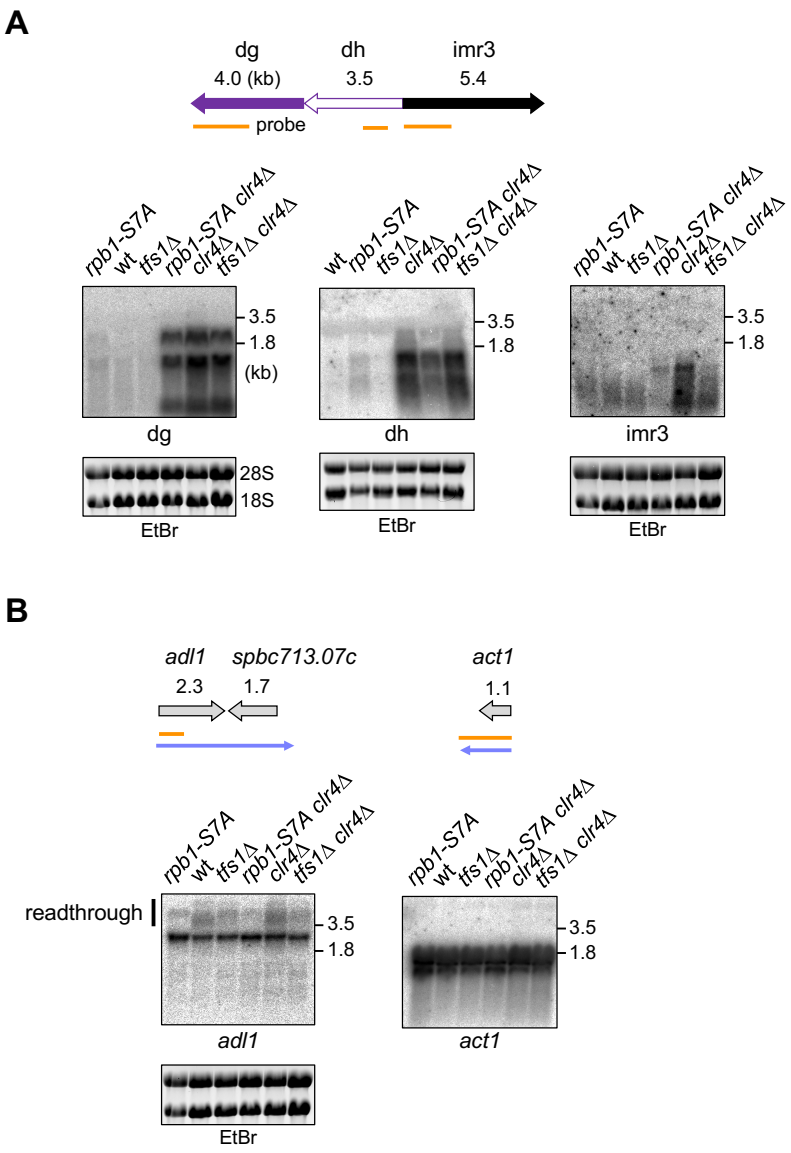


Fig. 17 *rpb1-S7A* and *tfs1*Δ slightly reduce the transcripts from centromere repeats and readthrough RNA.

A. Northern blotting using total RNAs prepared from log phase cultures of *rpb1-S7A*, wild type, *tfs1*Δ, *rpb1-S7A clr4*Δ, *clr4*Δ, and *tfs1*Δ *clr4*Δ strains. Illustrated are the positions of DNA probes used in Northern blotting (orange bars). RNAs were separated by 1.0% agarose gel under denatured condition, stained with EtBr (the bottom panel), transferred onto a nylon membrane, and hybridized with specific probes (the top panel). **B.** Northern blotting using total RNAs prepared from log phase cultures of *rpb1-S7A*, wild type, *tfs1*Δ, *rpb1-S7A clr4*Δ, *clr4*Δ, and *tfs1*Δ *clr4*Δ strains. Illustrated are the positions of DNA probes used in Northern blotting (orange bars) and the readthrough transcript of *adl1* and the transcript of *act1* (blue arrows). The membrane that was hybridized with the *adl1* probe was re-hybridized with the *act1* probe.

transcripts that encompass the downstream converging gene *spbc713.07c* (McDowall et al. 2015). Interestingly, both *rpb1-S7A* and *tfs1*Δ specifically reduced the *adl1* readthrough transcripts (Fig. 17B). Re-hybridization of the membrane using *act1* probe showed that neither *rpb1-S7A* nor *tfs1*Δ affect transcription of the *act1* gene that has no converging genes nearby, suggesting that RNAPII CTD Ser7 and Tfs1 facilitate transcription passing through termination sites. RNAPII might require CTD Ser7 and Tfs1 for passing through the conflict with DNA binding proteins such as other RNAPII or replication factors. Collectively, these results suggest that RNAPII CTD Ser7 and Tfs1 are required for a specific type of transcription at centromeres that causes GCRs in the *clr4*Δ mutant.

_tfs1_Δ reduces chromosome loss and restores hyper-sensitivity to thiabendazole (TBZ) in the _clr4_Δ cells

Recombination events such as crossover and BIR between repetitive elements can give rise to GCRs (Weischenfeldt et al. 2013; Carvalho and Lupski 2016). It has been shown that the recombination using Rad51 protein suppresses isochromosome formation by facilitating non-crossovers rather than crossovers at centromeres (Onaka et al. 2016; Zafar et al. 2017). In the absence of Rad51, Mus81 endonuclease, a central factor that produces crossovers, causes centromeric GCRs but suppresses chromosome loss, probably because of retaining chromosomes (Onaka et al. 2016). Thus, DNA repair factors downstream of crossover or BIR pathway appear to produce isochromosomes to prevent from chromosome loss. To test whether Tfs1, CTD Ser7 of RNAPII, and Mlo3 are involved in such a DNA repair, I determined the loss rate of minichromosome ChL in these mutant strains (Fig. 18A and B). As expected, *clr4*Δ markedly increased the chromosome loss. I found that *tfs1*Δ, *rpb1-S7A* and *mlo3*Δ reduced the chromosome loss in the *clr4*Δ background. These results suggest that Tfs1, CTD Ser7 of RNAPII, and Mlo3 induce chromosome instability, such as isochromosome formation or chromosome loss, rather than that they have roles downstream of crossover or BIR. Because structural abnormalities including isochromosome formation are unstable, they often result in aneuploidy associated with chromosome loss via degradation or missegregation of chromosomes. Clr4 is required for the

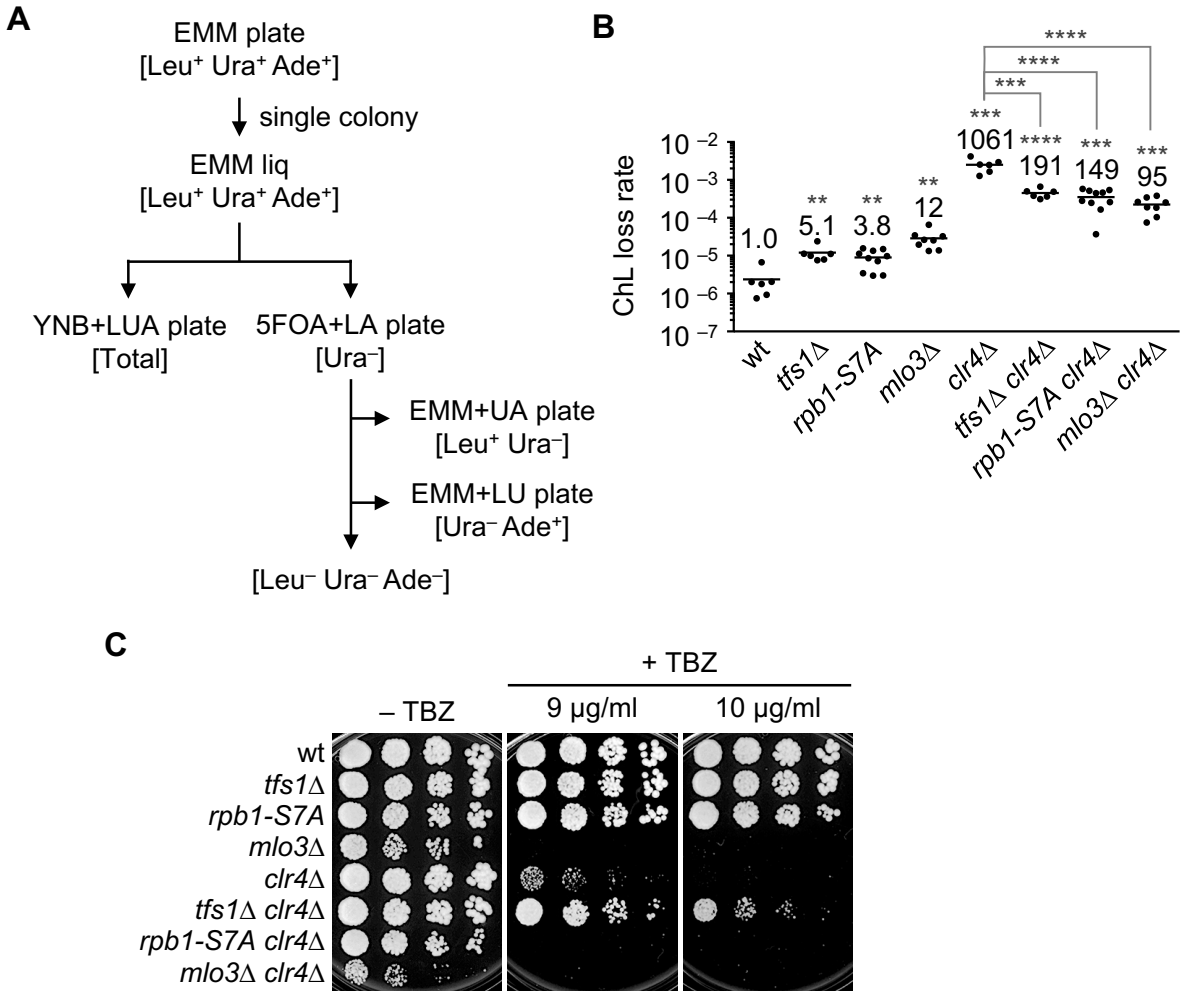


Fig. 18 Effects of *tfs1*Δ, *rpb1-S7A*, and *mlo3*Δ on ChL loss and thiabendazole (TBZ) sensitivity in the *clr4*Δ cells.

A. A schematic view of a protocol to determine the rate of ChL loss. A single colony formed on EMM was inoculated into EMM liquid media. After 2 days incubation at 30˚C, cells were plated onto YNB+LUA and 5FOA+LA plates. After 6-8 days incubation, the colonies were counted to determine the number of total and Ura– cells. The Ura– colonies formed on 5FOA+LA plates were replicated onto EMM+UA and EMM+LU plates to inspect Leu$^{+/-}$, and Ade$^{+/-}$, respectively. The number of Leu– Ura– Ade– cells indicative of ChL loss was obtained by subtracting the number of Leu+ Ura– Ade–, Leu+ Ura– Ade+, and Leu– Ura– Ade+ cells from that of Ura– cells. **B.** Rates of the ChL loss in wild-type, *tfs1Δ*, *rpb1-S7A*, *mlo3Δ*, *clr4Δ*, *tfs1Δ clr4Δ*, *rpb1-S7A clr4Δ*, and *mlo3Δ clr4Δ* strains (TNF5676, 6688, 6848, 5764, 5702, 6726, 6850, and 5824, respectively). Each dot represents the ChL loss rate determined using a single colony formed on EMM plates in scatter plots. Lines represent the mean. The ChL loss rate relative to that of the wild-type strain is indicated on the top of each column. The two-tailed Student's *t*-test. **C.** Log-phase cultures of wild-type, *tfs1*Δ, *rpb1-S7A, mlo3Δ, clr4*Δ, *tfs1Δ clr4Δ, rpb1-S7A clr4Δ,* and *mlo3Δ clr4Δ* strains (TNF5921, 6722, 6862, 5923, 5948, 6799, 6864, and 5925, respectively) were 5-fold serially diluted with distilled water and spotted onto YE+LUA supplemented with TBZ at a final concentration of 9 and 10 μg/mL.

heterochromatin structure in centromeres, and *clr4*Δ cells are hypersensitive to a microtubule-destabilizing drug, thiabendazole (TBZ) (Ekwall et al. 1995; Rea et al. 2000). To know whether Clr4 maintains centromere function by repressing transcription associated with Tfs1, CTD Ser7 of RNAPII, and Mlo3, I tested the sensitivity to TBZ using serial dilution assay. Interestingly, *mlo3*Δ showed hyper sensitivity to TBZ (Fig. 18C). Although Mlo3 is required for the generation of small RNA at centromeres (Zhang et al. 2011), the centromere localization of Mlo3, as well as RNAPII, is not detected in wild-type background (Zhang et al. 2011), and TBZ sensitivity of *mlo3*Δ was even higher than that of *clr4*Δ. Thus, it is unlikely that Mlo3 is important for centromere function further than Clr4. In budding yeast, Yra1/Mlo3 localize to telomeres and play an important role for telomere integrity to prevent transcription-replication collisions (Gavalda et al. 2016; Garcia-Rubio et al. 2018). The hyper sensitivity of *mlo3*Δ to TBZ might be resulted from telomere instability. Interestingly, I found that *tfs1*Δ reduced hypersensitivity to TBZ of *clr4*Δ cells, suggesting the Tfs1/TFIIS-dependent transcription as a critical target of heterochromatin to maintain the function of the centromeres. However, neither *rpb1-S7A* nor *mlo3*Δ restored hypersensitivity to TBZ of *clr4*Δ cells. Thus, the dysfunction of centromere mediated by Tfs1/TFIIS-dependent transcription may not be responsible for GCRs. *tfs1*Δ may restore kinetochore function in *clr4*Δ cells, because *tfs1*Δ is associated with kinetochore assembly by facilitating the deposition of CENP-A nucleosomes (Catania et al. 2015). Further studies are required to clarify the roles of Tfs1/TFIIS-dependent transcription in centromere dysfunction and GCRs mediated by centromere repeats.

DISCUSSION

A few reports have shown that the methylation of H3K9 is important for repeat stability. Loss of H3K9 methyltransferases in *C. elegans* increases expansion and contraction of tandem repeats (Zeller et al. 2016). Loss of Suv39 methyltransferase in Drosophila increases the contraction of satellite and ribosomal DNA repeats that is accompanied with the generation of extrachromosomal circular DNA (Peng and Karpen 2007). However, it has been uncleared how the methylation of H3K9 is involved in chromosome stability. Here, I found that heterochromatin suppresses GCRs that are mediated by centromere repeats. Deletion of Clr4 extremely increased the spontaneous formation of isochromosomes whose breakpoints are present in centromere repeats. Mutations in the catalytic domain of Clr4 or at H3K9 also increased the GCR rate, suggesting that Clr4 suppresses centromeric GCRs through H3K9 methylation. Both HP1 homologs, Swi6 and Chp2, and an RNAi component Chp1 were the chromodomain proteins that are essential for full suppression of GCRs. In clr4Δ strain, mutations in transcription factor Mlo3 and RNAPII impaired chromatin binding of RNAPII and reduced GCRs in the *clr4*Δ mutant, showing that Clr4-dependent H3K9 methylation suppresses GCRs by repressing RNAPII. Strikingly, deletion of Tfs1/TFIIS that facilitates the restart of paused and backtracked RNAPII greatly reduced GCR rate in the *clr4*Δ mutant without changing RNAPII chromatin binding levels. These results suggest that heterochromatin suppresses centromeric GCRs by repressing Tfs1-dependent transcription elongation (Fig. 19A).

Methylation of H3K9 is important to suppress centromeric GCRs

Like the *clr4*Δ and the *clr4-set* mutants, amino acid substitutions of H3K9 for alanine (H3K9A) or arginine (H3K9R) increased the GCR rate (Fig. 9C). Interestingly, H3K9R increased the GCR rate further than H3K9A. Alanine is an uncharged amino acid as well as acetylated or methylated lysine, while arginine is a positively charged amino acid and thereby capable of mimicking the biochemical properties of un-modified lysine (Wang et al. 2003). The H3K9R mutation has been shown to exhibit more severe growth defect and more sensitivity to TBZ, and higher incidence of chromosome missegregation than H3K9A mutation (Mellone et al. 2003). Positively charged histones may affect GCRs and centromere function. On the other hand, neutralization of positively charged histone residues attenuate the interaction with negatively charged DNA,

forming an open chromatin structure that is more accessible to transcription factors. Positively charged histones by H3K9R mutation might obstruct the RNAPII progression and Tfs1 is frequently required to help transcription elongation

Redundant function of HP1 and RNAi for suppression of GCRs

HP1 homolog Swi6 recruits several kinds of factors to heterochromatin. Swi6 is required for stable binding of cohesin complexes at centromeres, and it also facilitates early replication of centromeres by recruiting Dbf4/Dfp1-dependent kinase (DDK) to replication origins (Bernard et al. 2001; Bailis et al. 2003; Hayashi et al. 2009). Swi6 positively regulates transcription that is required for small RNA generation by recruiting Epe1 demethylase (Zofall and Grewal 2006). Because *swi6*Δ single mutant showed similar GCR rate to that of wild type (Fig. 11B), it appears that neither the cohesin enrichment, the replication timing control at centromeres, nor small RNA generation, is essential to suppress GCRs. It is consistent with the fact that *swi6*Δ does not significantly change the GCR rate at centromeres using a similar method to this study (Li et al. 2013), and that *epe1*Δ only slightly increased the GCR rate. Whereas, deletion of both HP1 homologs, Swi6 and Chp2, exhibited increased GCRs as compared to wild type, indicating that the redundant function of Swi6 and Chp2 is important to suppress GCRs. Swi6 and Chp2 redundantly inhibit RNAPII localization at centromeres by recruiting Clr6 and Clr3 (Sadaie et al. 2008; Fischer et al. 2009). Thus, Swi6 and Chp2 may suppress GCRs through inhibiting RNAPII localization at centromeres. Consistent with this, HDACs Sir2, Clr3, and Clr6 were required for GCR suppression (Fig. 14E and 14F). Interestingly, mutation in Clr6 further increased the GCR rate than *sir2*Δ *clr3*Δ double mutant. Clr6 is a shared catalytic subunit of two physically and functionally distinct HDAC complexes (Nicolas et al. 2007). Complex I deacetylates histones on specific genes to regulate promoters and is essential for cell proliferation. By contrast, complex II is not essential, and it acts globally to deacetylate histones across gene–coding regions and represses antisense transcription. Both complex I and complex II suppress transcription at centromeres in a distinct way. The complex I partially suppresses transcription from the reverse strand, which is weakly transcribed for RNAi machinery. The complex II preferentially

suppresses transcription from the forward strand, which is normally silenced. I found that deletion of Alp13, which is the other component of Complex II and affects H3K9ac levels at ~25% of gene-coding regions, did not increase the GCR rate (Fig. 14F), suggesting that Clr6 complex I, but not complex II, is important to suppress the GCRs. As siRNA generation is unlikely to be essential for GCR suppression, *clr6-1* might induce un-controlled transcription of reverse strand of centromere repeats. I found *swi6*Δ *chp2*Δ and *chp1*Δ synergistically increased the GCR rate to the level similar to that of *clr4*Δ, demonstrating that HP1 and the RNAi machinery have non-overlapping roles in suppressing GCRs. Although RNAPII levels at centromeres in *swi6*Δ *chp2*Δ cells are similar to those in *clr4*Δ cells, residual levels of transcriptional silencing have been detected in *swi6*Δ *chp2*Δ as compared to *clr4*Δ cells (Motamedi et al. 2008; Fischer et al. 2009). Both HP1 and RNAi machinery might suppress not only RNAPII loading to heterochromatin but also RNAPII progression (Fig. 19A).

Ago1 has additional role for suppression of GCRs mediated by centromere repeats

Among the mutants related to RNAi machinery, only *ago1*Δ showed extremely high GCR rate, suggesting that Ago1 has an additional role in suppressing GCRs independent of RITS and ARC complexes. Surprisingly, the GCR rate of *ago1*Δ strain was even higher than that of *clr4*Δ strain. *cid14*Δ restored H3K9me2 levels but it did not significantly change the GCR rate in *ago1*Δ cells, suggesting that promoting di-methylation of H3K9 is not the only role of Ago1 in suppressing GCRs. *mlo3*Δ and *epe1*Δ not only restored H3K9me2 and H3K9me3 levels but also reduced RNAPII localization and GCRs at centromeres in *ago1*Δ cells. As *clr4-W31G*, which reduces H3K9me3 levels at centromeres (Jih et al. 2017), did not show marked increase in GCRs (Fig. 11B), repression of RNAPII but not facilitating H3K9me3, is likely to be important to suppress GCRs. Thus, it seems that Ago1 has a direct role to repress transcription for GCR suppression. Consistent with this, human Ago1 directly binds to RNAPII (Kim et al. 2006), and *Drosophila* Ago2 interacts with a negative elongation factor NELF and represses heat-shock genes under the normal condition (Cernilogar et al. 2011). Strikingly, *mlo3*Δ and *rpb1-S7A* reduced RNAPII localization and GCRs even in the absence of Clr4. Thus, transcriptional repression rather than

H3K9 methylation is essential to suppress centromeric GCRs.

How does Tfs1/TFIIS-dependent transcription cause GCRs that are mediated by centromere repeats?

During transcription, RNAPII transiently pauses when it encounters with nucleosomes, replication factors, etc. (Gomez-Herreros et al. 2012). RNAPII resumes transcription with the aid of transcription elongation factors, such as ell1/ELL or Spt4/DSIF complex. If pausing persists, such as when replication factors are imposed, the pause gradually decays into arrest, which is characterized by the backtracking of RNAPII on DNA. Otherwise, RNAPII in an arrested state can be subject to degradation by the ubiquitin/proteasome pathway (Wilson et al. 2013). The backtracked RNAPII resumes transcription by cleaving a nascent 3'RNA. Because the intrinsic endonuclease activity of RNAPII is very weak, RNAPII requires Tfs1/TFIIS to restart transcription (Izban and Luse 1992; Kettenberger et al. 2003). TFIIS comprises and N-terminal domain I, a central domain II, and a C-terminal domain III. A domain II of TFIIS binds near the rim of the RNAPII funnel, and a domain III of TFIIS extends into the central catalytic pore of RNAPII, enabling to help positioning a water molecule for hydrolytic RNA cleavage (Kettenberger et al. 2003).

I found that the deletion of Tfs1 among transcription elongation factors specifically reduced GCRs at centromeres in *clr4*Δ strain (Fig. 16A and D), suggesting that the specific action of transcription, the restart of backtracked RNAPII rather than alleviating RNAPII pausing, causes GCRs. The *rpb1-S7A* mutation, which reduced GCRs in *clr4*Δ strain (Fig. 15B), can cause immature transcription termination with the aid of termination factors in the case when RNAPII encounters with downstream gene (Sanchez et al. 2018), implying that Ser7 of RNAPII CTD is important to overcome transcription factors that bind to DNA. Consistent with this finding, *rpb1-S7A* reduced readthrough transcripts at *adl1* region that encompass the downstream gene (Fig. 17B). Ser7 of Rpb1, as well as Tfs1, might be involved in the release from backtracked RNAPII, resulting in GCRs.

Tfs1, Ser7 of RNAPII CTD, and Mlo3 caused not only GCRs but also loss of chromosomes in the absence of heterochromatin (Fig. 14A, 15B, 16A, and 18B). These results suggest that the restart of backtracked RNAPII cause chromosome instability rather than playing a role downstream of non-conservative repair pathway such as crossover or BIR. It is possible that heterochromatin suppresses DNA damages at centromere repeats, or specifically suppresses

non-conservative repair between non-allelic repeats after DNA damages to prevent from GCRs. Interestingly, *clr4*Δ increases recombination between inverted centromere repeats only by ~2-fold (Zafar et al. 2017), suggesting that the suppression of DNA damage at centromere repeats is at least a part of roles of heterochromatin for GCR suppression. Thus, it appears that heterochromatin suppresses Tfs1-dependent transcription that can cause non-conservative repair between non-allelic repeats after DNA damages.

How does the restart of transcription from backtracked state cause GCRs at centromere repeats? RNAPII binding to chromatin is a major cause of DNA instability that is associated with transcription because of a collision with replication machinery (Kim and Jinks-Robertson 2012). Importantly, *tfs1*Δ did not significantly change the RNAPII chromatin binding levels (Fig. 16B), suggesting that RNAPII binding is unlikely a cause of GCRs at centromeres. It is possible that overproduction of repeat RNA causes GCRs through accumulation of DNA:RNA hybrids at centromeres, because exposed ssDNA is susceptible to DNA damage. However, neither *tfs1*Δ nor *rpb1-S7A* significantly reduced the amount of centromere transcripts in *clr4*Δ cells (Fig. 17A), suggesting that overproduction of centromere repeats is unlikely a cause of GCRs. The other possibility is that a conformational change of DNA structure caused by the restart of backtracked RNAPII results in GCRs at centromeres. In RNAi mutant strain, replication fork prone to stall at centromeres and a recombination factor Rad52 accumulates especially during S phase (Zaratiegui et al. 2011). The de-repressed transcription may frequently encounter with replication factors, resulting in DNA damages. The replication fork protection factors are important to suppress GCRs at centromeres in *swi6*Δ strain (Li et al. 2013). When RNAPII encounters with replication factors, the restarted RNAPII may prone to destabilize the replication complex, resulting in a fork reversal and an exposure of nascent ssDNA (Fig. 19B). The conflicts between transcription and replication machinery result in the accumulation of R-loops behind RNAPII (Santos-Pereira and Aguilera 2015). The accumulated R-loop induces mutagenic BIR repair, which results in continuous replication up to chromosome end, in budding yeast (Amon and Koshland 2016). At repetitive sequences, the nascent repeat ssDNA may invade into homologous sequences within R-loops, following aberrant recombination such as BIR between non-allelic repeats. Supporting with these idea, H3K9 methylation was shown to be required for the suppression of R-loop formation at repetitive elements in *C.elegans* (Zeller et al. 2016). Further studies are required to understand how the restart of transcription from backtracked RNAPII causes GCRs between centromere repeats.

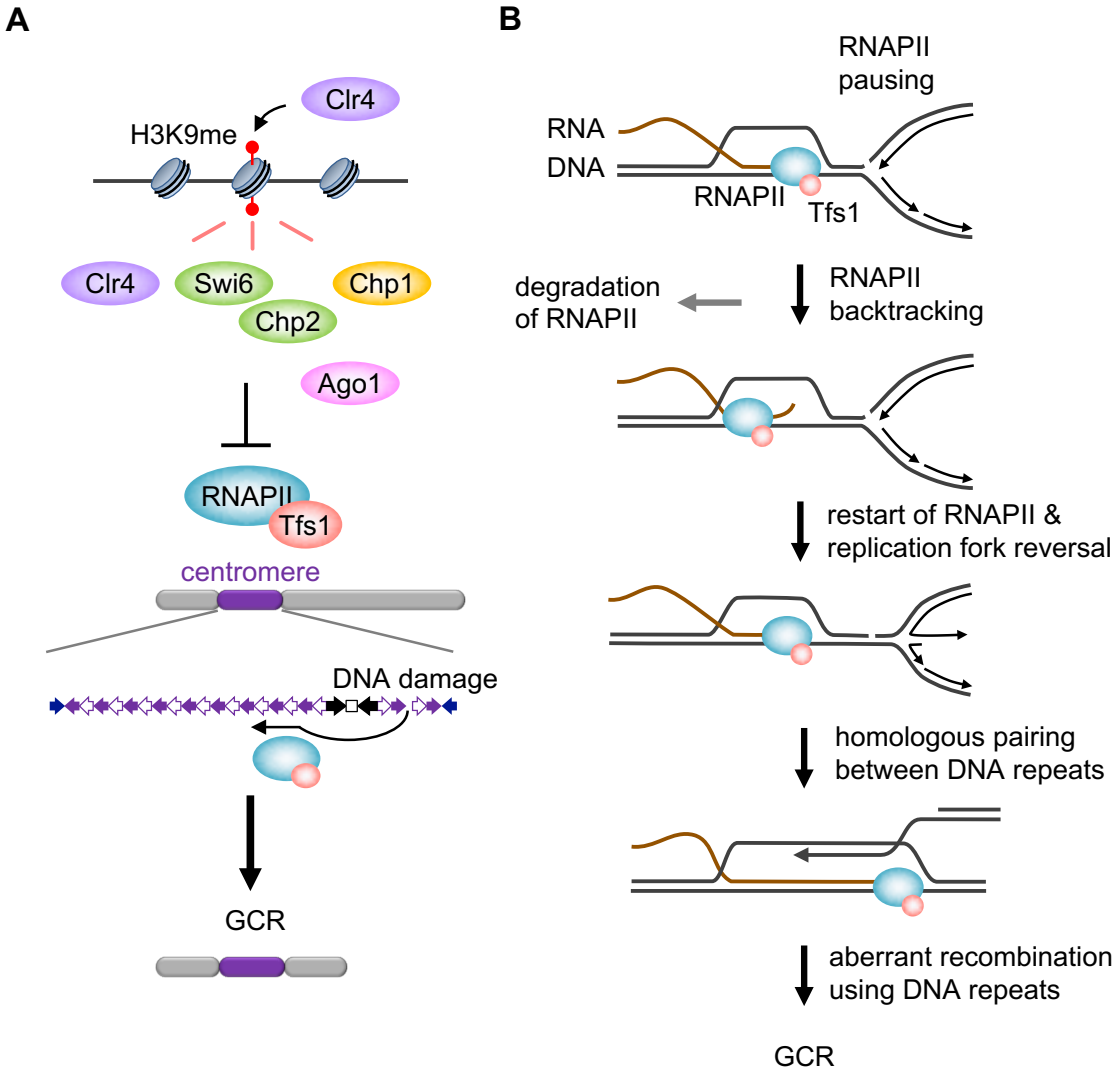


Fig. 19 Heterochromatin suppresses Tfs1/TFIIS-dependent transcription to prevent GCRs at centromere.

A. Illustration is a summary that explains how heterochromatin suppresses GCRs at centromeres. After methylation of H3K9 mediated by Clr4, the chromodomain proteins Clr4, Swi6, Chp2, and Chp1 bind to methylated H3K9 marks and suppresses RNAPII binding and subsequent transcription to prevent GCRs mediated by centromere repeats. The RNAi factor Ago1 also important for GCR suppression through inhibiting RNAPII. In the absence of heterochromatin, Tfs1-mediated transcription reaction facilitates non-conservative recombination between non-allelic repeats that results in GCRs such as isochromosome formation. **B.** When RNAPII pauses during transcription, paused RNAPII is degraded in ubiquitin/proteasome pathway or backtracks on DNA. After the cleavage of nascent RNA that is facilitated by Tfs1, RNAPII resumes transcription. In the case that RNAPII pauses at replication machinery, replication fork reverses on DNA, with exposing nascent ssDNA. At repetitive sequences, the nascent repeat ssDNA may invade into homologous sequences within R-loops, following aberrant recombination such as BIR between non-allelic repeats.

Part II

Regulation of recombination between centromere repeats by kinetochore chromatin

ABSTRACT

Centromere is essential for proper segregation of chromosomes, whereas it consists of repetitive sequences, which have potential to suffer gross chromosomal rearrangements (GCRs). It has been shown in fission yeast that recombination factor Rad51 and Rad54 suppress isochromosome formation generated by recombination between inverted centromere repeats. At centromere inverted repeats, Rad51-dependent homologous recombination (HR) that requires Rad51, Rad54, and Rad52 is predominant and crossovers that can result in isochromosome formation are underrepresented. These previous findings suggest that the regulation of Rad51-dependent HR and suppression of crossovers are essential for centromere integrity. However, what regulates recombination at centromere remains unclear. Using ura4 (cen1) strain, where an entire region of centromere on chromosome 1 (cen1) was integrated into the non-centromeric *ura4* locus, I tested whether pericentromeric repeats, where heterochromatin assembles, affect the regulation of recombination at centromeres. Chromatin immunoprecipitation (ChIP) experiments showed that heterochromatin, but not kinetochore chromatin, can assemble on ectopic centromere repeats. Recombination analysis at ectopic centromere repeats showed that predominance of Rad51-dependent HR and noncrossover are not detected. These results suggest that kinetochore chromatin rather than pericentromeric heterochromatin and repeats is essential for the regulation of recombination at centromeres.

INTRODUCTION

There are Rad51-dependent and -independent homologous recombination (HR). The major HR pathway is Rad51-dependent HR that is associated with Rad51, Rad54, and Rad52. After DNA ends of double strand break (DSB) are resected to yield 3' single-stranded DNA (ssDNA) overhangs, RPA coats on the ssDNA (Fig. 2). Rad51 loads onto RPA-coated ssDNA with the aid of Rad52 in yeast (New et al. 1998; Shinohara and Ogawa 1998). In mammals, BRCA2 stimulates Rad51 nucleofilament formation instead of Rad52 (Jensen et al. 2010). Rad51 catalyzes strand exchange forming D-loop structure. Rad54 stabilizes Rad51 nucleofilament formation and facilitates DNA strand exchange, DNA synthesis, and branch migration (Petukhova et al. 1998; Bugreev et al. 2006; Wright and Heyer 2014). The second end capture by D-loops creates double holliday junctions (dHJs). Resolution of joint molecules including D-loops and HJs by structure-specific endonuclease Mus81 generates either crossover (CO) that exchanges the flanking sequences or non-crossover (NCO) products that maintains the original sequence. Alternatively, dissociation of D-loops before the formation of dHJs leads to synthesis dependent strand annealing (SDSA), which generates only NCO products (Nassif et al. 1994) (Fig. 2). In Rad51-independent HR, Rad52 binds to 3' ssDNA and catalyzes single-strand annealing (SSA) between complementary ssDNA molecules, that is mutagenic DSB repair and cause gross chromosomal rearrangements (GCRs) such as deletion between repeats or translocation (Bhargava et al. 2016).

Previous studies have shown that the regulation of recombination between inverted centromere repeats is essential for centromere integrity. Analysis of recombination between fission yeast centromere inverted repeats showed that Rad51-dependent HR is predominant and crossovers, that results in the inversion of the intervening region, are underrepresented as compared to recombination in arm region (Zafar et al. 2017). In the absence of Rad51, Mus81 facilitates crossovers, demonstrating that Rad51 suppresses Mus81-mediated crossovers at centromeres (Onaka et al. 2016). Analysis of GCRs which are mediated by centromere repeats showed that Rad51 and Rad54 suppress the formation of isochromosomes that is mediated by Mus81 (Nakamura et al. 2008; Onaka et al. 2016). These findings suggest that Rad51-dependent recombination promotes NCOs to prevent Mus81-mediated COs and isochromosome formation. However, it has been uncleared what regulates the recombination at centromere repeats.

In the central domain of centromere, histone H3 variant Cnp1/CENP-A nucleosomes are interspersed with canonical H3-containing nucleosomes. CENP-TWSX nucleosome-like

complex also localize to centromere. CENP-A and CENP-TWSX provide a platform for the assembly of the kinetochore (McKinley and Cheeseman 2016). CENP-C, which binds to CENP-A nucleosomes, and CENP-T recruit microtubule binding proteins to kinetochores. Kinetochore region is flanked by heterochromatin, which is marked by histone H3K9 methylation and enrichment of Swi6/HP1 that binds to H3K9me. Swi6/HP1 recruits many factors including cohesin that is required for sister chromatid cohesion.

It has been shown that temperature-sensitive mutant of Cnp1/CENP-A, that is essential for kinetochore assembly, reduces Cnp1/CENP-A binding to kinetochore chromatin but does not affect crossover suppression at centromeres (Zafar et al. 2017). To know whether pericentromeric repeats, where heterochromatin assembles, affect the regulation of recombination at centromeres, I analyzed recombination properties using ura4 (cen1) strain, where an entire region of centromere on chromosome 1 (cen1) was integrated into the non-centromeric *ura4* locus. I found that pericentromere repeats and heterochromatin are not required for predominance of Rad51-dependent HR and NCO. Although H3K9me2 and Swi6HP1 bound to ectopic pericentromere, Cnp1/CENP-A, Cnp20/CENP-T, and Mhf2/CENP-X did not bind to the central domain of ectopic centromere in the ura4 (cen1) strain. At the ectopic central region of cen1, *rad51*Δ and *rad54*Δ only partially reduced the recombination rate as compared to *rad52*Δ, and the proportion of crossovers was significantly increased as compared to original cen1. These data demonstrate that pericentromere repeats and heterochromatin structure are not responsible for the regulation of recombination in the central region of centromeres, and also suggest that the factors related to kinetochore chromatin except for Cnp1/CENP-A may be responsible for it.

RESULTS and DISCUSSION

Heterochromatin but not kinetochore chromatin assembles at ectopic centromere repeats

At least 1.6 kb of pericentromere sequences is sufficient to form heterochromatin when introduced at an ectopic locus (Partridge et al. 2002). To know the effects of pericentromeric heterochromatin on the regulation of recombination at centromeres, I utilized the ura4 (cen1) strain, where an entire region of centromere on chromosome 1 (cen1) was integrated into the non-centromeric ura4 locus (Fig. 20). *ade6B/X* heteroallelles were inserted into SnaBI sites of the central region of original cen1 (cen1 strain) or ectopic cen1 (ura4 (cen1) strain) to compare HR at centromere and non-centromere regions. I first performed Chromatin immunoprecipitation (ChIP) experiments and determined the levels of H3K9me2, Swi6/HP1, Cnp1/CENP-A, Cnp20/CENP-T, Mhf2/CENP-X, and H3 at the pericentromere region (imr1-out and dg) and the central region (ade6, imr1-in, and cnt2) of centromere, and non-centromere region (*adl1*) (Fig. 21). In the cen1 strain, H3K9me2 and Swi6 chromatin binding were observed specifically at pericentromere, but not at the central region of centromere and non-centromere region. As expected, the levels of H3K9me2 and Swi6 chromatin binding at ectopic cen1 in the ura4 (cen1) strain were comparable to the original cen1 in the cen1 strain, confirming that heterochromatin assembles at ectopic pericentromere repeats. As expected, Cnp1/CENP-A, Cnp20/CENP-T and Mhf2/CENP-X, specifically bound to the central region of centromere (ade6, imr1-in and cnt2) in the cen1 strain (Fig. 22). On the other hand, these kinetochore factors could be detected at background levels at ectopic cen1 (ade6) in the ura4 (cen1) strain. Consistent with this, the localization of these kinetochore factors was reduced by around half of the cen1 level at imr1-in region which is present at both the original and ectopic cen1 in the ura4 (cen1) strain. Whereas, the kinetochore factors bound to the original centromere cnt2 both in the ura4 (cen1) strain and the cen1 strain. These results suggest that the kinetochore factors localize to original but not ectopic centromeres. This may be due to instability of di-centric chromosomes (Sato et al. 2012). Instead of the reduced localization of the kinetochore factors to the ectopic centromere, H3 binding level at

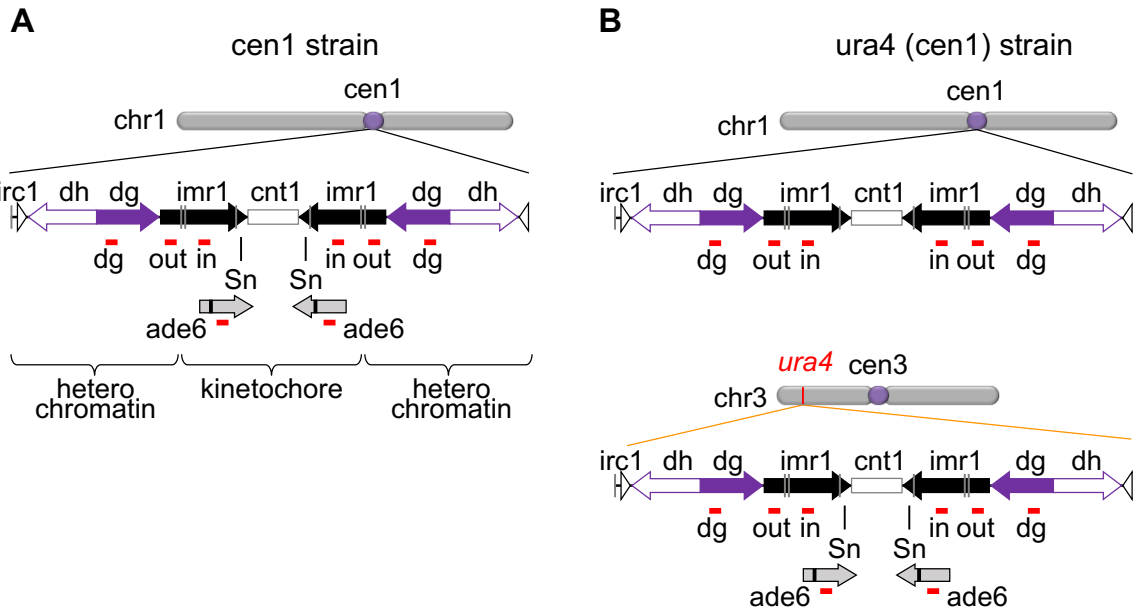


Fig. 20 Recombination between *ade6B* and *ade6X* heteroalleles in centromeres and arm regions.

A. Construction of cen1 strain. Illustrated are the central sequecne *cnt1* and the inverted repeats *imr1*, *dg*, *dh*, and *irc1* in centromere 1 (cen1). *ade6B* and *ade6X* heteroalleles were integrated at the Sn sites in imr1. Kinetochore chromatin and heterochromatin are assembled on the cen1. The positions of PCR amplification in ChIP analysis are shown in red. Sn, SnaBI. **B.** Construction of ura4 (cen1) strain. Illustrated are the cen1 region on chr1 and the ectopic cen1 region introduced at the *ura4* locus of chr3.

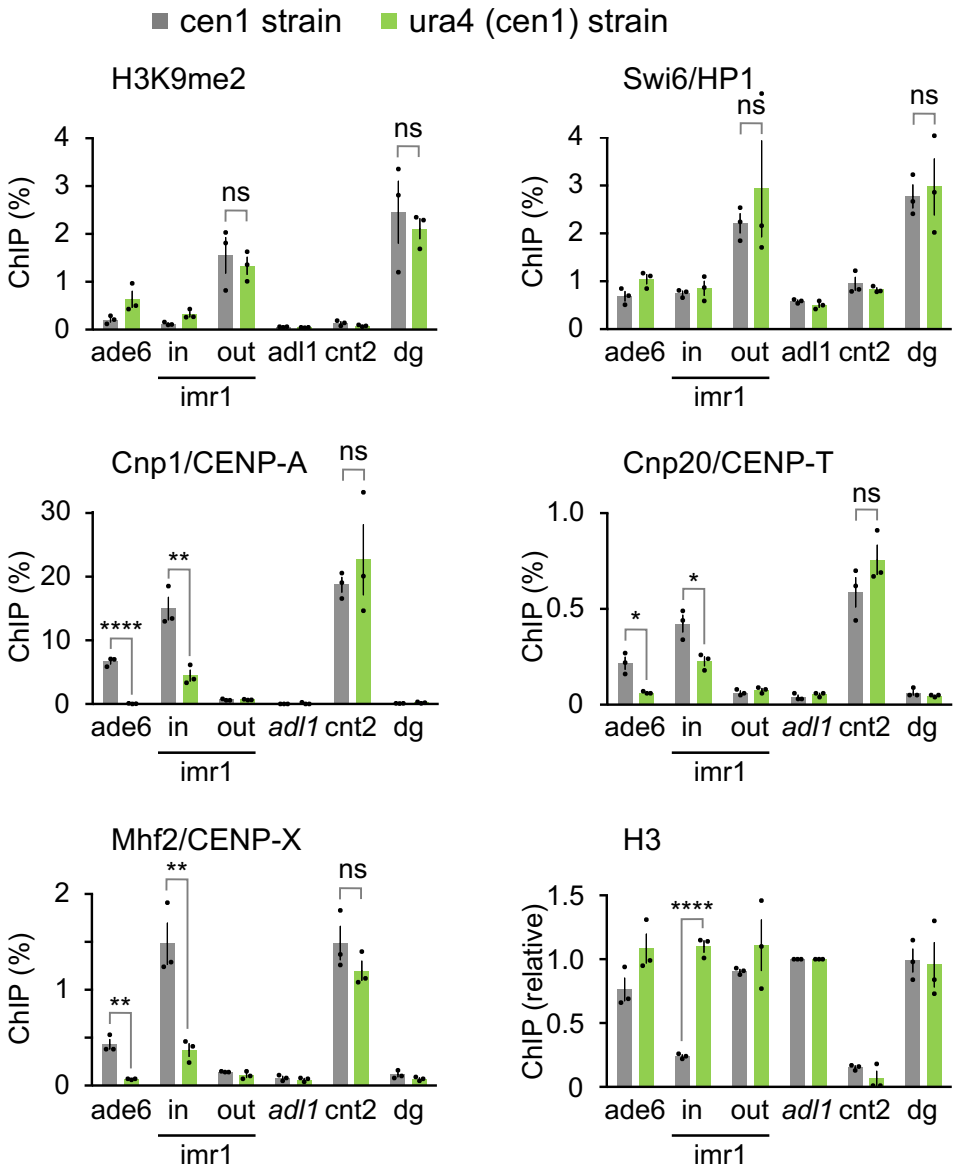


Fig. 21 Heterochromatin but not kinetochore chromatin assembles at ectopic centromere repeats.

Chromatin immunoprecipitation (ChIP) analysis was performed to determine the level of H3K9me2, Swi6, Cnp1/CENP-A, Cnp20/CENP-T. Mhf2/CENP-X, and H3 in the cen1 and the ura4 (cen1) strains (TNF3347 and 4684, respectively). cnt2 and *adl1* are in the centromere and arm regions of chr2, respectively. imr1-in and imr1-out are in the kinetochore and heterochromatin domains, respectively. *ade6* is present in the original cen1 in cen1 strain, while it is present only in the ectopic cen1 in ura4 (cen) strain. The mean ± s.e.m. from 3 biologically independent experiments is shown. The dot represents each experiment. Statistical significance relative to wild type which is shown on the top of bars, and that between pairs of mutant strains were determined using two-tailed Student's t test. * $P < 0.05$. ** $P < 0.01$. **** $P < 0.0001$. ns, not significant.

imr1-in region increased over two-fold. Collectively, these results suggest that heterochromatin, but not kinetochore chromatin, forms on ectopic centromere repeats.

Pericentric heterochromatin does not regulate the recombination in the central region of centromere

Previous studies using cen1 strain showed that Rad51-dependent recombination and non-crossover predominantly occur at central region of centromere (Onaka et al. 2016; Zafar et al. 2017). Analysis of recombination frequencies between *ade6B/X* heteroallelles at ectopic centromere in the ura4 (cen1) strain showed that *rad51*Δ and *rad54*Δ only partially decreased recombination rate as compared to *rad52*Δ, demonstrating that both Rad51-dependent recombination and Rad51-independent recombination occur at the central region of ectopic centromere (Fig. 22A). To determine whether pericentromeric heterochromatin is required for NCO production, chromosomal DNA was prepared from independent Ade^+ recombinants, digested with restriction enzyme AfeI, and separated by pulse-field gel electrophoresis (PFGE). Southern hybridization using probe 1 was carried out to identify the fragments indicative of CO or NCO (Fig. 22B). It has been shown that only 4% of recombinants were COs in the cen1 strain (Zafar et al. 2017). Whereas in the ura4 (cen1) strain, the ratio of COs (24%) was significantly higher than that in the cen1 strain ($P < 0.001$) (Fig. 22C and D). These results show that pericentromeric heterochromatin structure and repetitive sequence do not affect the predominance of Rad51-dependent HR and NCOs. Given that Cnp1/CENP-A has been suggested to be dispensable for NCOs at centromeres, the factors related to kinetochore chromatin except for Cnp1/CENP-A rather than heterochromatin may be responsible for the regulation of recombination in the central region of centromeres (Fig. 23).

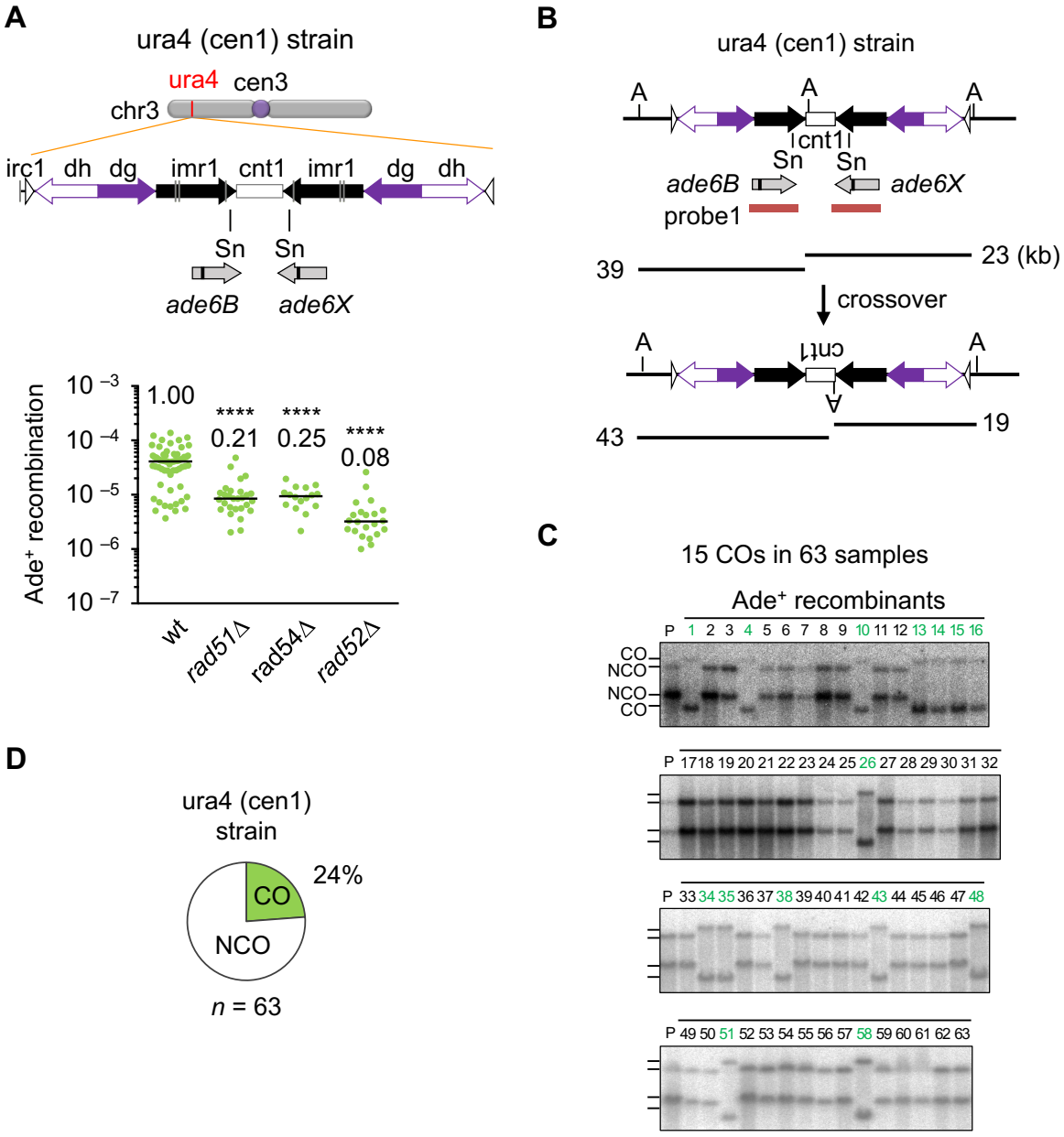


Fig. 22 Pericentromeric heterochromatin is not responsible for Rad51-dependent recombination and non-crosover between *ade6B/ade6X* heteroallels at ectopic centromere.

A. Recombination in the ura4 (cen1) strain construct. Illustrated is the central sequence cnt1, and inverted repeats imr1, dg, dh and irc1 in the centromere 1 (cen1), that was integrated into *ura4* locus on chromosome 3. *ade6B* and *ade6X* mutant genes were integrated at the Sn sites in imr1. Spontaneous rates of Ade+ prototroph formation were determined in wild-type, *rad51*Δ, *rad54*Δ and *rad52*Δ strains (TNF4684, 5814, 5826 and 5829, respectively). Independent experimental values are shown in scatter plots and lines indicate medians. Rates relative to the wild type value are indicated at the top of each column. Statistical significance of differences between pairs of strains was determined using the two-tailed Mann-Whitney test. **B.** Illustrated is the ectopic centromere in the ura4 (cen1) strain and expected sizes of DNA bands that could be detected by Southern analysis. Positions of centromere repeats, AfeI restriction sites, probe1, and the length of AfeI restriction fragments are indicated. ade6B/X were omitted in the bottom part of the illustration for simplicity. **C.** Southern blot data showing the physical detection of crossover (CO) and non-crossover (NCO) in the wild type of ura4 (cen1) strain. **D.** Proportions of crossovers in the cen1 and ura4 (cen1) construct in wild type are indicated in Pie charts.

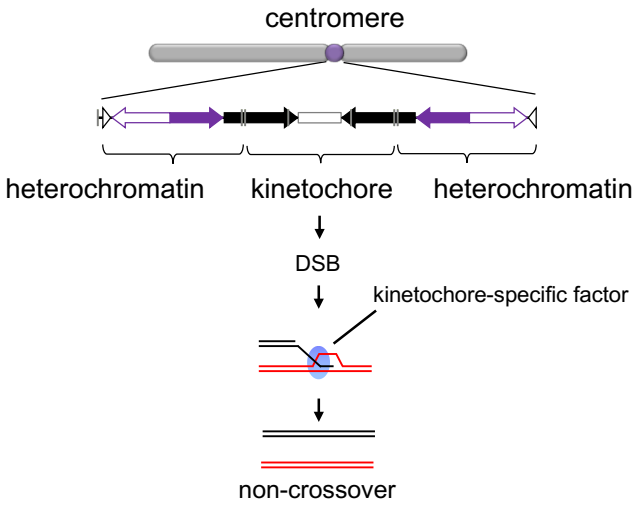


Fig. 23 Pericentromeric heterochromatin is dispensable for centromere-specific recombination at central region of centromere.
At ectopic centromere repeats which is capable of heterochromatin but not kinetochore chromatin formation, pericentromeric heterochromatin does not preferentially occur Rad51-dependent recombination and non-crossovers. The factors associated with kinetochore chromatin may be required for centromere-specific recombination.

MATERIALS and METHODS

Strains and media

Fission yeast strains used in this study are listed in Table 1. Yeast cells were grown in YE, EMM, YNB, and 5FOA media supplemented with appropriate amino acids at a final concentration of 225 mg/l. YNB media contain 1.7 g/l of yeast nitrogen base (BD Biosciences, Difco 233520), 5 g/l of ammonium sulfate (Nacalai Tesque, 02619-15), and 2% glucose. YNB media were supplemented with 1 g/l of 5-fluoroorotic acid (Apollo Scientific, PC4054) and 56 mg/l of uracil to make 5FOA media. Solid media contains 1.5% agarose (Nacalai Tesque, 01028-85). Unless otherwise indicated, cells were grown at 33˚C for recombination assay or 30˚C for GCR assay. Yeast transformation was performed by the lithium acetate method. The gene disruption was carried out using the kanamycin, hygromycin, or nourseothricin resistant gene, and the transformants were selected on the media supplemented with G418 (Nacalai Tesque, 09380-86), hygromycin B (Nacalai Tesque, 09287-84), or clonNAT (Werner BioAgents, 96736-11-7) at a final concentration of 100 μg/ml. *clr4-R406A,N409A,H410A* (*clr4-set*), *mlo3K165A,K167A* (*mlo3KA*), and *mlo3K165R,K167R* (*mlo3KR*) mutant strains were created by a pop-in/pop-out gene replacement (Gao et al. 2014): pTN1220 plasmid containing the wild-type *ura4*$^+$ and the mutant *clr4-set* genes was digested with NgoMIV and introduced into *ura4-D18* mutant cells. pTN1179 containing *ura4*$^+$ and *mlo3KA* and pTN1178 containing *ura4*$^+$ and *mlo3KR* were digested with HpaI and introduced into *ura4-D18* cells. Ura$^+$ transformants were selected on EMM plates, and then Ura$^-$ progenies resulting from *ura4*$^+$ pop-out were selected on 5FOA plates. Correct integration was confirmed by PCR and DNA sequencing.

Plasmids

clr4-set, *mlo3KA*, and *mlo3KR* mutant genes were constructed by a two-step PCR method. From yeast genomic DNA, a 0.7 kb PCR fragment was produced using clr4-1 and clr4-NHR-F primers, and a 1.0 kb fragment using clr4-NHR-R and clr4-2 primers, independently. These partially overlapping PCR fragments were mixed and used for the 2nd PCR in the presence of clr4-1 and

clr4-2 primers. A 1.4 kb SpeI–PvuII restriction fragment prepared from the 2nd PCR product was introduced between SpeI–NaeI sites of pTN782 containing *ura4*$^+$ gene to make pTN1220. A 2.0 kb genomic region that contain the *mlo3*$^+$ gene was amplified using mlo3-1 and mlo3-5, and digested with XbaI at one site. A 1.9 kb restriction fragment with XbaI–blunt ends was introduced between XbaI–NaeI sites of pTN782 to make pTN1169. From yeast genomic DNA, a 1.0 kb PCR fragment was produced using mlo3-1 and mlo3-KA-R primers, and a 0.7 kb fragment using mlo3-KA-F and mlo3-4 primers. These partially overlapping PCR fragments were mixed and used for the 2nd PCR in the presence of mlo3-1 and mlo3-4 primers. A 1.0 kb SacI–XbaI restriction fragment of the 2nd PCR product that contains the *mlo3KA* mutation was introduced between SacI–XbaI sites of pTN1169 to make pTN1179. mlo3-KR-R and mlo3-KR-F primers were used in place of mlo3-KA-R and mlo3-KA-F to make pTN1178 that contains the *mlo3KR* mutation.

The plasmids used for Northern blotting were constructed as below. A 2.3 kb region that contains a portion of imr3 was amplified using otr3-2 and imr3-XhoI-R primers. A 1.7 kb PvuII–MfeI restriction fragment of the PCR product was introduced between HincII–EcoRI sites of pBluescript II KS$^+$ to make pTN1226. A 0.9 kb region that contains a portion of the *adl1* gene was amplified using adl1-F and adl1-R primers. A 0.9 kb XbaI–ApaI restriction fragment of the PCR product was introduced between XbaI–ApaI sites of pBluescript II KS$^+$ to make pTN1227. A 2.1 kb region that contains a portion of the *act1* gene was amplified using act1-F and act1-R primers. A 1.9 kb XhoI–EcoRV restriction fragment of the PCR product was introduced between XhoI–EcoRV sites of pBluescript II KS$^+$ to make pTN1225. DNA sequencing confirmed that no mutations were introduced during PCR amplification. I also used pTN834 and pTN770 for probe preparation from our lab stocks. pTN834: a 9.6 kb XbaI–EcoRI fragment containing cen1 sequence from pRS140 (Chikashige et al. 1989) was introduced between XbaI–EcoRI sites of pUC19. pTN770: a 1.5 kb NheI–ClaI restriction fragment of the PCR product amplified from genomic DNA that contains dh sequence was introduced between SpeI–ClaI sites of pBluescript II KS$^+$.

GCR assay

Yeast cells were incubated for 6–8 days on EMM+UA plates, and 10 ml of EMM+UA was inoculated with a single colony formed on the EMM+UA plates. After 2 days' incubation, 200

cells were plated onto YNB+UA, and either 2,000, 20,000, or 200,000 cells were washed with distilled water and plated onto 5FOA+A media. After 6–9 days' incubation, the number of colonies formed on YNB+UA and 5FOA+A plates were counted to determine the number of Leu^+ and that of Leu^+ Ura^- cells, respectively. Leu^+ Ura^- colonies formed on 5FOA+A plates were incubated on EMM+UA plates and then replicated onto EMM+A and EMM+U plates to confirm Ura^- and to inspect $Ade^{+/-}$, respectively. The number of Leu^+ Ura^- Ade^- cells indicative of GCR was obtained by subtracting the number of Leu^+ Ura^- Ade^+ cells from that of Leu^+ Ura^- cells. Using the number of Leu^+ cells and that of Leu^+ Ura^- Ade^- cells in 10 ml of EMM+UA culture, I determined GCR rate per generation. Note that even if the same number of GCR events occurred in each culture, the number of Leu^+ Ura^- Ade^- cells would differ depending on how early the first GCR event occurred. To avoid inflation of the average number of GCR clones by jackpots arising from an early GCR event, I used fluctuation analysis that examine the number of GCR events in at least 16 cultures to reveal how the number of events fluctuates among each culture (Lea and Coulson 1949; Lin et al. 1996). The analysis is based on the assumption of using a single colony originating from single cell. Using this analysis, the GCR rate (R) is expressed by the ratio of the cell numbers (Np, $Np=2^p$) when the first GCR event occurred at the p-th generation (p): $R=1/Np$ (Lin et al. 1996). Using the number of GCR cells (S) and the cell numbers (N_G, $N_G=2^G$) per generation (G), the GCR rate (R) is given by $R=1/Np=S/[N_G\times(G-p+1)]$ …[A]. Note that only S and N_G could be obtained from the experiments. Because N_G is expressed as 2^G, G can be written as $G=\ln N_G/\ln 2$ …[B]. Likewise, p is given by $p=\ln N_p/\ln 2=-\ln R/\ln 2$ …[C]. To solve [A] using S and N_G, [B] and [C] were substituted into [A], enabling to describe R as follow: $S=RN_G \ln(2RN_G)/\ln 2$. I determined the GCR rate (R) by the least square method. The resultant each GCR rate achieved at least 3-digit accuracy, indicating high-confidence values.

Pulse-field gel electrophoresis (PFGE)

From parental and GCR (Leu^+ Ura^- Ade^-) clones obtained from biologically independent experiments, cells were inoculated in 10 ml of YE3S and incubated at 25°C for 1–2 days. $1.0\text{x}10^8$ cells were collected by centrifugation at 1,580 g for 1 min using a bench-top centrifuge (TOMY,

LC-200) and suspended in 2.5 ml ice cold 50 mM EDTA (pH 8.0). Cells were then harvested by centrifugation at 700 g for 1 min and then suspended in 1ml CSE buffer (20 mM citrate phosphate, 50 mM EDTA, 1 M sorbitol, adjusted pH 5.6 with HCl) and 5 µl Zymolyase 20T (Nacalai Tesque, 07663-91) and 5 µl Lyzing enzyme 25 mg/ml (Sigma-Aldrich, L1412) and incubated at 30°C for 20-50 min. The spheroplast was harvested by centrifugation at 100 g for 10 min and suspended in 120 µl CSE buffer. Pre-melted 140 µl of 1.6% low melting agarose (Nacalai Tesque, 01161-12) was mixed to the cell pellet and the mixture was poured into the mold to make plugs. The plugs were then suspended in 1 ml of SDS-EDTA solution (1% SDS, 0.25 M EDTA (pH8.0)) and incubated at 60°C for 2 hrs. The buffer was then exchanged to 1 ml of ESP buffer (0.5 M EDTA (pH 9.0), 1% N-lauryl sarcosine, 1.5 mM CaAc) supplemented with 1 mg/ml proteinase K (Nacalai Tesque, 29442-85) and incubated at 50°C for 24 hrs. Lastly, the buffer was then exchanged to 1 ml of ice cold TE10:1. Chromosomal DNAs were separated in 0.55% Certified Megabase agarose gel (Bio-Rad, 161-3109) using CHEF-DRII system (Bio-Rad) under the following conditions. Broad-range PFGE: 1,500 sec pulse time at 2 V/cm for 42 hrs and then 180 sec pulse time at 2.4 V/cm for 4 hrs, at 4°C in 1×TAE buffer (40 mM Tris-acetate, 1 mM EDTA). Short-range PFGE: from 40 to 70 sec pulse time at 4.2 V/cm for 24 hrs, at 4°C in 0.5×TBE buffer (89 mM Tris-borate, 2 mM EDTA). After electrophoresis, DNAs were stained with 0.2 µg/ml of EtBr (Nacalai Tesque, 14631-94) and detected using a Typhoon FLA9000 (GE Healthcare).

Southern hybridization

After EtBr staining, agarose gel was irradiated with 300 mJ ultraviolet (UV) light using GS Gene Linker (Bio-Rad) for DNA fragmentation, and then soaked into 800 ml of alkaline buffer (1.2 M NaCl, 0.4 M NaOH) for 40 min with gentle shaking to denature DNA. DNA was transferred to ClearTrans nylon membrane 0.45 µm (Wako, 039-22673) by capillary action in 25 mM sodium phosphate buffer (pH 6.5) and covalently attached to the membrane by 150 mJ UV irradiation. A 0.6 kb EcoRI–EcoRI fragment prepared from pTN755 (Nakamura et al. 2008), α-^{32}P-dCTP (3,000 Ci/mmol, PerkinElmer Life Sciences, NEG013H), and Random primer labeling kit ver. 2 (Takara,

6045) were used to prepare radioactive probes according to manufacturer's instructions. Radioactive signals were detected using BAS2500 phosphorimager (Fuji Film).

PCR analysis of GCR products

After separation of chromosomal DNA by PFGE, GCR products were recovered from agarose gel using the FastGene Gel/PCR Extraction kit (Nippon Genetics, FG-91302). KOD FX Neo polymerase (Toyobo, KFX-201) was utilized to amplify cnt3–imr3 junctions and cen3 proximal regions, while Q5 polymerase (New England Biolabs, M0491) was used to amplify irc3. PCR products were separated by 1.7% Seakem GTG agarose gel (Lonza, 50070) electrophoresis in 1×TBE buffer, stained with 0.2 μg/ml of EtBr, and detected using a Typhoon FLA9000. The PCR primers used in this assay are listed in Table 2.

Chromatin immunoprecipitation (ChIP)

1.7×10^8 cells from log-phase cultures in YE media supplemented with leucine, uracil, adenine, and histidine (YE4S) were collected by centrifugation at 1,580 g for 2 min, washed with distilled water, and suspended in the same volume of EMM. The strains used for recombination assay were grown in EMM+A media, and 1.7×10^8 cells were transferred to a new flask. Cells were fixed in formaldehyde (Sigma-Aldrich, F8775) to a final concentration of 1% by vigorously mixing for 15 min at room temperature, and neutralized the crosslink by further mixing for 5 min following the addition of 2.5 M glycine to a final concentration of 125 mM. Cells were washed twice with cold distilled water and suspended in 500 μl of 0.1% lysis buffer (100 mM HEPES-KOH (pH7.4), 140 mM NaCl, 1 mM EDTA (pH8.0), 0.1% Triton X-100, 0.1% Na-deoxycholate). After centrifugation at 5,100 g for 1 min at 4˚C, the supernatant was discarded and cells were stored at –80˚C. Cells were suspended in 200 μl of 0.1% lysis buffer. An equal volume of glass beads, 2 μl of protease inhibitor cocktail (Sigma-Aldrich, P8215), and 4 μl of 1 mM phenylmethyl-sulfonylfluoride (PMSF) were added to the cell suspension. Cells were disrupted with glass beads for 30 sec for four times at 4˚C using Micro Smash MS-100 (TOMY). After making a tiny hole at the bottom of the tube using a heated needle, the disrupted cell suspensions were isolated by

centrifugation at 700 g for 1 min at 4˚C. The glass beads were washed with 200 µl of 0.1% lysis buffer, following the centrifugation again. After the addition of the 10% Triton X-100 to a final concentration of 1.1% to the cell suspension, the cell extracts were sonicated for 10 sec for four times on ice using Sonifier 250 (Branson). The supernatants containing DNA fragments were recovered after centrifugation at 17,900 g for 10 min at 4˚C. Before immunoprecipitation, Mouse and rabbit antibodies were attached to Dynabeads M-280 sheep anti-Mouse IgG (Invitrogen, 11202D) and Dynabeads M-280 sheep anti-Rabbit IgG (Invitrogen, 11204D), respectively, in 400 µl of PBS supplemented with 2% BSA by incubation at 4˚C overnight. The beads were washed with 400 µl of 1% lysis buffer (100 mM HEPES-KOH (pH7.4), 140 mM NaCl, 1 mM EDTA (pH8.0), 1% Triton X-100, 0.1% Na-deoxycholate) and suspended with 340 µl of 1% lysis buffer. For immunoprecipitation, the beads were mixed with 60 µl of the cell extracts, and incubated for indicated hours at 4˚C with rotation. The period of incubation time, and the antibodies and the magnetic beads used for immunopresipitation were listed in Table 3. After immunoprecipitation, the beads were washed twice with 400 µl of 1% lysis buffer, once with 400 µl of 1% lysis buffer supplemented with 500 mM NaCl, twice with 400 µl of wash buffer (10 mM Tris-HCl (pH8.0), 1 mM EDTA, 250 mM LiCl, 0.5% NP-40, 0.5% Na-deoxycholate), and once with 400 µl of TE10:1 (10 mM Tris-HCl (pH8.0), 1 mM EDTA). The beads were suspended in 100 µl of elution buffer (10 mM Tris-HCl (pH8.0), 1 mM EDTA, 1% SDS), and incubated at 65˚C for 25 min with brief mixing using a vortex every 2 min. The supernatants were transferred to new tubes and incubated at 65˚C overnight to disrupt crosslinks. For preparation of DNA from whole cell extracts, 100 µl of the cell extracts that were diluted 100-fold by elution buffer were incubated at 65˚C overnight and treated same as immunoprecipitated samples. After the treatment with proteinase K at a final concentration of 0.3 mg/ml at 50˚C for 2 hrs, the DNA was purified by phenol/chloroform extraction. The 5 µl of 5 M NaCl, 2 µl of glycogen (Nacalai Tesque, 17110-11), and 300 µl 100% ethanol were sequentially added to the 100 µl of the DNA, following the storage of DNA for 1 hr at –80˚C. The DNA was precipitated by centrifugation at 17,900 g for 15 min at 4˚C, followed by wash with 500 µl of 70% ethanol. The DNA prepared from whole cell extracts and immunoprecipitation fractions were suspended with 60 µl and 30 µl of TE10:1,

respectively. The DNAs were quantified by real time PCR using SYBR FAST (Thermo Fisher, 4385614) in a StepOnePlus real time PCR system (Applied Biosystems). The primers used in ChIP are listed in Table 4.

Northern blotting

1.0×10^9 cells from log-phase cultures in YE4S media were collected by centrifugation at 1,580 g for 2 min, washed with 1 ml of diethylpilocarbonate (DEPC)-treated pure water, and suspended in 400 µl of AE buffer (50 mM NaAc (pH 5.2), 10 mM EDTA (pH8.0)). After adding 40 µl of 10% SDS and 400 µl of phenol equilibrated with AE buffer, the cells were incubated at 65˚C for 4 min, and immediately chilled in a dry ice/ethanol bath for 1 min. RNA was purified by phenol:chloroform:isoamyl alcohol (25:24:1) extraction. The concentration of extracted RNA was quantified by OD260 using GeneQuant 100 (GE Healthcare). Concentration was calculated using the 260 nm reading and a conversion factor (A_{260} of 1.0=40 µg/ml for RNA), and expressed as follow: (RNA)=(OD260)×(40 µg /ml).10 µg of total RNAs were suspended in 8.5 µl of MOPS buffer (20 mM MOPS pH 7.0, 2 mM NaAc, 1 mM EDTA pH8.0) supplemented with 8% formaldehyde, 50% deionized formamide (Nacalai-Tesque, 02020-64), and 10 µg/ml EtBr, and heat denatured by incubation at 55°C for 60 min. RNAs were separated by 1.0% PrimeGel agarose LE (TaKaRa, 5801A) gel / 2.2 M formaldehyde electrophoresis in MOPS buffer. RNAs stained with EtBr were detected using a Typhoon FLA9000. After soaking the gel into 50 mM NaOH (Nacalai Tesque, 31511-05) for 20 min, RNAs were transfer to a ClearTrans nylon membrane 0.45 µm (Wako, 039-22673) by capillary action in alkaline transfer buffer (10 mM NaOH, 3 M NaCl) and covalently attached to the membrane by 150 mJ UV irradiation. A 2.0 kb KpnI–KpnI fragment prepared from pTN834, a 0.9 kb NsiI–XbaI fragment from pTN770, a 1.7 kb PstI–XhoI fragment from pTN1226, a 0.9 kb XbaI–ApaI fragment from pTN1227, and a 1.9 kb XhoI–EcoRV fragment from pTN1225, and α-^{32}P-dCTP (3,000 Ci/mmol, PerkinElmer Life Sciences, NEG013H), and Random primer labeling kit ver. 2 (Takara, 6045) were used to prepare radioactive probes for the detection of dg, dh, imr3, *adl1*, and *act1* RNAs, respectively. Radioactive signals were detected using a BAS2500 phosphorimager.

Chromosome loss assay

Cells were incubated for 6–8 days on EMM plates, and 10 ml of EMM liquid media was inoculated with a single colony from EMM plates. After 2 days' incubation in EMM liquid culture, cells were washed with distilled water and plated onto YNB+LUA and 5FOA+LA media. 8 days after plating, the number of total colonies and that of Ura^- colonies were counted on YNB+LUA and 5FOA+LA plates, respectively. Ura^- colonies formed on 5FOA+LA plates were incubated on EMM+UA and EMM+LU plates to inspect $Leu^{+/-}$ and $Ade^{+/-}$, respectively. The number of Leu^- Ura^- Ade^- cells indicative of chromosome loss was obtained by subtracting the number of Leu^+ Ura^- Ade^+, Leu^+ Ura^- Ade^-, and Leu^- Ura^- Ade^+ cells from that of Ura^- cells. Because the cells retain intact ChL during the incubation in EMM media, the Leu^- Ura^- Ade^- cells grown on 5FOA+LA plates are indicative of ChL loss that have been suffered at first cell cycle soon after plating. The rates of chromosome loss per generation (R) using the number of total cells (T) and that of Leu^- Ura^- Ade^- cells (L) in 10 ml of EMM culture could be expressed: $R=L/T$.

Spot assay

From YE3S plates incubated at 30°C for 3–7 days, single colonies were inoculated into YE4S media. Exponentially growing cells were serially diluted 5-fold with distilled water and aliquots of 8 μl were spotted on YE3S plates supplemented with indicated concentrations of TBZ. TBZ was suspended with N.N-dimethylformamide (Nacalai Tesque, 13016-94). Plates were incubated for 5 days at 30°C.

Recombination assay

Yeast strains containing *ade6B* and *ade6X* hetroalleles were grown on YE+A plates for 3-5 days at 33˚C. Single colonies from YE+A plates were then inoculated into 10 ml of EMM+A and incubated for 1–2 days. The cells were washed with distilled water and plated on EMM+A and EMM+G (EMM supplemented with 50 μg/ml of guanine prevents growth of Ade^- cells) plates and incubated for 3-5 days. The colonies were then counted and total number of viable colonies

and total number of viable recombinants that is Ade^+ formation were determined. The number of colonies formed on EMM+A and EMM+G plates were counted to determine the number of total cells and Ade^+ cells, respectively. Like GCR assay, I determined recombination rate per generation using fluctuation analysis (Lea and Coulson 1949; Lin et al. 1996) The recombination rate (R) is expressed by the ratio of the cell numbers ($N_p=2^p$) when the first recombination occurred at the p-th generation (p): $R=1/Np$ (Lin et al. 1996). Using the number of recombinant cells (S) and the cell numbers (N_G, $N_G=2^G$) per generation (G), the recombination rate (R) is given by $R=1/Np=S/[N_G\times(G-p+1)]$ …[A]. Because N_G is expressed as 2^G, G can be written as $G=\ln N_G/\ln 2$ …[B]. Likewise, p is given by $p=\ln N_p/\ln 2=-\ln R/\ln 2$ …[C]. Then, R is described as follow: $S=RN_G \ln(2RN_G)/\ln 2$. I determined the recombination rate (R) by the least square method. The resultant each recombination rate achieved at least 3-digit accuracy, indicating high-confidence values.

Determination of crossover and non-crossover recombinants

To prepare yeast DNA, single colonies from EMM+G plates were inoculated into YE3S liquid cultures and incubated for 1–2 days at 33˚C. 7×10^8 cells were washed with 5 ml of ice-cold TE10:25 (10 mM Tris-HCl (pH 8.0), 25 mM EDTA), and suspended in 1 ml of SP1 buffer (20 mM sodium citrate, 20 mM di-sodium hydrogenphosphate, 40 mM EDTA, adjusted pH 5.6 with HCl). After adding 20 µl of β-mercaptoethanol, the cells were incubated for 20 min with rotation at 30°C. After centrifugation at 2,900 g for 1 min, the cell pellet was recovered and suspended in 500 µl of SP1 buffer. After the addition of 50 µl of 3.5 mg/ml Lyticase (Sigma-Aldrich, L4025), the cell suspension was incubated for 20-60 min at 37°C. After centrifugation at 700 g for 1 min, the cell spheroplast was recovered and suspended in 300 µl of TE50:20 (50 mM Tris-HCl (pH 8.0), 20 mM EDTA). After the addition of 100 µl of 10% SDS, the suspension was incubated for 60 min at 65˚C. After the addition of 300 µl of 5M KAc, the suspension was mixed gently and left on ice for 10 min. After centrifugation at 17,900 g for 5 min twice, the supernatants were recovered and suspended in 750 µl of isopropanol. After leaving on ice for 10 min, the DNA was precipitated by centrifugation at 17,900 g for 10 min, following wash with 1 ml of 70% ethanol.

DNA was suspended in 100 µl of TE10:1 and treated with RNase A (Sigma-Aldrich, R4642) to a final concentration of 50 µg/ml at 60˚C for 20 min. To discriminate crossover and non-crossover recombinants, 20 µl of DNA was digested with 0.4 µl of AfeI (10,000 units/ml, New England Biolabs, R0652L) at 37˚C for 3 hrs. The DNA was separated in 0.55% Certified Megabase agarose gel in 0.5 × TBE buffer using CHEF-DRII system (Bio-Rad) under the following conditions: from 1 to 6 sec pulse time at 6 V/cm for 15 hrs. After EtBr staining, southern blotting was done in essentially the same way as GCR assay. A 1.9 kb BamHI-PstI fragment of pKT110 that contains the *ade6B* gene was used as DNA template to prepare probe 1. α-^{32}P-dCTP (3,000 Ci/mmol, PerkinElmer Life Sciences, NEG013H), and Random primer labeling kit ver. 2 (Takara, 6045) were used to prepare radioactive probes.

Statistical analyses

The two-tailed Mann-Whitney test and the two-railed Fisher's exact test were performed using GraphPad Prism version 6.0g for Mac (GraphPad Software). The two-tailed Student's *t*-test was performed using Excel (Microsoft).

Table 1. The yeast strains used in this study.

strain	genotype
TNF3896	*h*$^{-}$, *smt0*, *ade6*Δ*-D*, *ura4-D18*, *leu1-32*, ChL
TNF5440	*h*$^{-}$, *smt0*, *ade6*Δ*-D*, *ura4-D18*, *leu1-32*, ChL, *clr4::kanMX6*
TNF5676	*h*$^{-}$, *smt0*, *mat2-3::natMX6*, *ade6*Δ*-D*, *ura4-D18*, *leu1-32*, ChL
TNF5702	*h*$^{-}$, *smt0*, *mat2-3::natMX6*, *ade6*Δ*-D*, *ura4-D18*, *leu1-32*, ChL, *clr4::kanMX6*
TNF5701	*h*$^{-}$, *smt0*, *mat2-3::natMX6*, *ade6*Δ*-D*, *ura4-D18*, *leu1-32*, ChL, *rec12::hphMX6*
TNF5766	*h*$^{-}$, *smt0*, *mat2-3::natMX6*, *ade6*Δ*-D*, *ura4-D18*, *leu1-32*, ChL, *rec12::hphMX6*, *clr4::kanMX6*
TNF6958	*h*$^{-}$, *smt0*, *mat2-3::natMX6*, *ade6*Δ*-D*, *ura4-D18*, *leu1-32*, ChL, *clr4-R406A*,*N409A*,*H410A*
TNF6155	*h*$^{-}$, *smt0*, *mat2-3::natMX6*, *ade6*Δ*-D*, *ura4-D18*, *leu1-32*, ChL, *mlo3-K165A*,*K167A*
TNF6157	*h*$^{-}$, *smt0*, *mat2-3::natMX6*, *ade6*Δ*-D*, *ura4-D18*, *leu1-32*, ChL, *mlo3-K165R*,*K167R*
TNF5738	*h*$^{-}$, *smt0*, *mat2-3::natMX6*, *ade6*Δ*-D*, *ura4-D18*, *leu1-32*, ChL, *h3.1/h4.1::his3*$^{+}$, *h3.3/h4.3::arg3*$^{+}$
TNF6223	*h*$^{-}$, *smt0*, *mat2-3::natMX6*, *ade6*Δ*-D*, *ura4-D18*, *leu1-32*, ChL, *h3.1/h4.1::his3*$^{+}$, *h3.3/h4.3::arg3*$^{+}$, *h3.2-K9A*
TNF5802	*h*$^{-}$, *smt0*, *mat2-3::natMX6*, *ade6*Δ*-D*, *ura4-D18*, *leu1-32*, ChL, *h3.1/h4.1::his3*$^{+}$, *h3.3/h4.3::arg3*$^{+}$, *h3.2-K9R*
TNF5992	*h*$^{-}$, *smt0*, *mat2-3::natMX6*, *ade6*Δ*-D*, *ura4-D18*, *leu1-32*, ChL, *clr4-W41G*
TNF6012	*h*$^{-}$, *smt0*, *mat2-3::natMX6*, *ade6*Δ*-D*, *ura4-D18*, *leu1-32*, ChL, *clr4-W31G*
TNF5706	*h*$^{-}$, *smt0*, *mat2-3::natMX6*, *ade6*Δ*-D*, *ura4-D18*, *leu1-32*, ChL, *swi6::hphMX6*
TNF5685	*h*$^{-}$, *smt0*, *mat2-3::natMX6*, *ade6*Δ*-D*, *ura4-D18*, *leu1-32*, ChL, *chp2::hphMX6*
TNF5900	*h*$^{-}$, *smt0*, *mat2-3::natMX6*, *ade6*Δ*-D*, *ura4-D18*, *leu1-32*, ChL, *swi6::kanMX6*, *chp2::hphMX6*
TNF5708	*h*$^{-}$, *smt0*, *mat2-3::natMX6*, *ade6*Δ*-D*, *ura4-D18*, *leu1-32*, ChL, *chp1::hphMX6*
TNF6151	*h*$^{-}$, *smt0*, *mat2-3::natMX6*, *ade6*Δ*-D*, *ura4-D18*, *leu1-32*, ChL, *swi6::kanMX6*, *chp2::hphMX6*, *chp1::hphMX6*
TNF5689	*h*$^{-}$, *smt0*, *mat2-3::natMX6*, *ade6*Δ*-D*, *ura4-D18*, *leu1-32*, ChL, *ago1::hphMX6*
TNF7335	*h*$^{-}$, *smt0*, *mat2-3::natMX6*, *ade6*Δ*-D*, *ura4-D18*, *leu1-32*, ChL, *tas3::kanMX6*
TNF7337	*h*$^{-}$, *smt0*, *mat2-3::natMX6*, *ade6*Δ*-D*, *ura4-D18*, *leu1-32*, ChL, *arb1::kanMX6*
TNF7331	*h*$^{-}$, *smt0*, *mat2-3::natMX6*, *ade6*Δ*-D*, *ura4-D18*, *leu1-32*, ChL, *arb2::kanMX6*
TNF7333	*h*$^{-}$, *smt0*, *mat2-3::natMX6*, *ade6*Δ*-D*, *ura4-D18*, *leu1-32*, ChL, *rdp1::kanMX6*
TNF5687	*h*$^{-}$, *smt0*, *mat2-3::natMX6*, *ade6*Δ*-D*, *ura4-D18*, *leu1-32*, ChL, *dcr1::hphMX6*
TNF6153	*h*$^{-}$, *smt0*, *mat2-3::natMX6*, *ade6*Δ*-D*, *ura4-D18*, *leu1-32*, ChL, *cid14::kanMX6*
TNF5764	*h*$^{-}$, *smt0*, *mat2-3::natMX6*, *ade6*Δ*-D*, *ura4-D18*, *leu1-32*, ChL, *mlo3::hphMX6*
TNF6109	*h*$^{-}$, *smt0*, *mat2-3::natMX6*, *ade6*Δ*-D*, *ura4-D18*, *leu1-32*, ChL, *epe1::hphMX6*
TNF6411	*h*$^{-}$, *smt0*, *mat2-3::natMX6*, *ade6*Δ*-D*, *ura4-D18*, *leu1-32*, ChL, *cid14::kanMX6*, *ago1::hphMX6*
TNF6188	*h*$^{-}$, *smt0*, *mat2-3::natMX6*, *ade6*Δ*-D*, *ura4-D18*, *leu1-32*, ChL, *mlo3::hphMX6*, *ago1::kanMX6*
TNF7325	*h*$^{-}$, *smt0*, *mat2-3::natMX6*, *ade6*Δ*-D*, *ura4-D18*, *leu1-32*, ChL, *epe1::hphMX6*, *ago1::hphMX6*
TNF5824	*h*$^{-}$, *smt0*, *mat2-3::natMX6*, *ade6*Δ*-D*, *ura4-D18*, *leu1-32*, ChL, *mlo3::kanMX6, clr4::hphMX6*
TNF6244	*h*$^{-}$, *smt0*, *mat2-3::natMX6*, *ade6*Δ*-D*, *ura4-D18*, *leu1-32*, ChL, *rad51::kanMX6*
TNF6383	*h*$^{-}$, *smt0*, *mat2-3::natMX6*, *ade6*Δ*-D*, *ura4-D18*, *leu1-32*, ChL, *mlo3::hphMX6*, *rad51::kanMX6*
TNF7341	*h*$^{-}$, *smt0*, *mat2-3::natMX6*, *ade6*Δ*-D*, *ura4-D18*, *leu1-32*, ChL, *sir2::hphMX6*

TNF7359	*h*⁻, *smt0*, *mat2-3::natMX6*, *ade6Δ-D*, *ura4-D18*, *leu1-32*, ChL, *clr3::hphMX6*
TNF7357	*h*⁻, *smt0*, *mat2-3::natMX6*, *ade6Δ-D*, *ura4-D18*, *leu1-32*, ChL, *sir2::hphMX6, clr3::hphMX6*
TNF7345	*h*⁻, *smt0*, *mat2-3::natMX6*, *ade6Δ-D*, *ura4-D18*, *leu1-32*, ChL, *clr6-1*
TNF5898	*h*⁻, *smt0*, *mat2-3::natMX6*, *ade6Δ-D*, *ura4-D18*, *leu1-32*, ChL, *alp13Δ*
TNF6848	*h*⁻, *smt0*, *mat2-3::natMX6*, *ade6Δ-D*, *ura4-D18*, *leu1-32*, ChL, *rpb1-CTD-S7A:kanMX6*
TNF6850	*h*⁻, *smt0*, *mat2-3::natMX6*, *ade6Δ-D*, *ura4-D18*, *leu1-32*, ChL, *rpb1-CTD-S7A:kanMX6*, *clr4::hphMX6*
TNF6688	*h*⁻, *smt0*, *mat2-3::natMX6*, *ade6Δ-D*, *ura4-D18*, *leu1-32*, ChL, *tfs1::kanMX6*
TNF6726	*h*⁻, *smt0*, *mat2-3::natMX6*, *ade6Δ-D*, *ura4-D18*, *leu1-32*, ChL, *tfs1::kanMX6*, *clr4::hphMX6*
TNF7163	*h*⁻, *smt0*, *mat2-3::natMX6*, *ade6Δ-D*, *ura4-D18*, *leu1-32*, ChL, *tfs1::kanMX6*, *rad51::kanMX6*
TNF7042	*h*⁻, *smt0*, *mat2-3::natMX6*, *ade6Δ-D*, *ura4-D18*, *leu1-32*, ChL, *ell1::kanMX6*
TNF7130	*h*⁻, *smt0*, *mat2-3::natMX6*, *ade6Δ-D*, *ura4-D18*, *leu1-32*, ChL, *leo1::kanMX6*
TNF7055	*h*⁻, *smt0*, *mat2-3::natMX6*, *ade6Δ-D*, *ura4-D18*, *leu1-32*, ChL, *spt4::kanMX6*
TNF7063	*h*⁻, *smt0*, *mat2-3::natMX6*, *ade6Δ-D*, *ura4-D18*, *leu1-32*, ChL, *ell1::kanMX6*, *clr4::hphMX6*
TNF7154	*h*⁻, *smt0*, *mat2-3::natMX6*, *ade6Δ-D*, *ura4-D18*, *leu1-32*, ChL, *leo1::kanMX6*, *clr4::hphMX6*
TNF7057	*h*⁻, *smt0*, *mat2-3::natMX6*, *ade6Δ-D*, *ura4-D18*, *leu1-32*, ChL, *spt4::kanMX6*, *clr4::hphMX6*
TNF5921	*h*⁻, *smt0*, *mat2-3::natMX6*, *ade6Δ-D*, *ura4-D18*, *leu1-32*
TNF5948	*h*⁻, *smt0*, *mat2-3::natMX6*, *ade6Δ-D*, *ura4-D18*, *leu1-32*, *clr4::kanMX6*
TNF6169	*h*⁻, *smt0*, *mat2-3::natMX6*, *ade6Δ-D*, *ura4-D18*, *leu1-32*, *clr4-R406A,N409A,H410A*
TNF6276	*h*⁻, *smt0*, *mat2-3::natMX6*, *ade6Δ-D*, *ura4-D18*, *leu1-32*, *cid14::kanMX6*
TNF5923	*h*⁻, *smt0*, *mat2-3::natMX6*, *ade6Δ-D*, *ura4-D18*, *leu1-32*, *mlo3::hphMX6*
TNF7349	*h*⁻, *smt0*, *mat2-3::natMX6*, *ade6Δ-D*, *ura4-D18*, *leu1-32*, *epe1::hphMX6*
TNF5922	*h*⁻, *smt0*, *mat2-3::natMX6*, *ade6Δ-D*, *ura4-D18*, *leu1-32*, *ago1::hphMX6*
TNF6550	*h*⁻, *smt0*, *mat2-3::natMX6*, *ade6Δ-D*, *ura4-D18*, *leu1-32*, *cid14::kanMX6*, *ago1::hphMX6*
TNF6210	*h*⁻, *smt0*, *mat2-3::natMX6*, *ade6Δ-D*, *ura4-D18*, *leu1-32*, *mlo3::hphMX6*, *ago1::kanMX6*
TNF7343	*h*⁻, *smt0*, *mat2-3::natMX6*, *ade6Δ-D*, *ura4-D18*, *leu1-32*, *epe1::hphMX6*, *ago1::hphMX6*
TNF5925	*h*⁻, *smt0*, *mat2-3::natMX6*, *ade6Δ-D*, *ura4-D18*, *leu1-32*, *clr4::hphMX6*, *mlo3::kanMX6*
TNF6862	*h*⁻, *smt0*, *mat2-3::natMX6*, *ade6Δ-D*, *ura4-D18*, *leu1-32*, *rpb1-CTD-S7A:kanMX6*
TNF6864	*h*⁻, *smt0*, *mat2-3::natMX6*, *ade6Δ-D*, *ura4-D18*, *leu1-32*, *rpb1-CTD-S7A:kanMX6*, *clr4::hphMX6*
TNF6931	*h*⁻, *smt0*, *mat2-3::natMX6*, *ade6Δ-D*, *ura4-D18*, *leu1-32*, *flag-rpb3*
TNF6943	*h*⁻, *smt0*, *mat2-3::natMX6*, *ade6Δ-D*, *ura4-D18*, *leu1-32*, *flag-rpb3*, *rpb1-CTD-S7A:kanMX6*
TNF6933	*h*⁻, *smt0*, *mat2-3::natMX6*, *ade6Δ-D*, *ura4-D18*, *leu1-32*, *flag-rpb3*, *clr4::hphMX6*
TNF6945	*h*⁻, *smt0*, *mat2-3::natMX6*, *ade6Δ-D*, *ura4-D18*, *leu1-32*, *flag-rpb3*, *rpb1-CTD-S7A:kanMX6*, *clr4::hphMX6*
TNF6722	*h*⁻, *smt0*, *mat2-3::natMX6*, *ade6Δ-D*, *ura4-D18*, *leu1-32*, *tfs1::kanMX6*
TNF6799	*h*⁻, *smt0*, *mat2-3::natMX6*, *ade6Δ-D*, *ura4-D18*, *leu1-32*, *tfs1::kanMX6*, *clr4::hphMX6*
TNF3347	*h*⁺, *ade6Δ-D, imr1L (Sn:ade6B), imr1R (Sn:ade6X)*
TNF4684	*h*⁺, *ade6Δ-D, ura4+:cen1 (imr1L (Sn:ade6B), imr1R (Sn:ade6X))*

TNF5814	*h*$^{+}$, *ade6*Δ*-D, ura4+:cen1 (imr1L (Sn:ade6B), imr1R (Sn:ade6X)), rad51::kanMX6*
TNF5826	*h*$^{+}$, *ade6*Δ*-D, ura4+:cen1 (imr1L (Sn:ade6B), imr1R (Sn:ade6X)), rad54::kanMX6*
TNF5829	*h*$^{+}$, *ade6*Δ*-D, ura4+:cen1 (imr1L (Sn:ade6B), imr1R (Sn:ade6X)), rad52::kanMX6*

ChL [*ubc11::LEU2*$^{+}$*, spcc1322.09::ura4*$^{+}$*, ade6*$^{+}$] is a derivative of Ch16.

Table 2. List of primers used to analyze GCR products, carry out site-directed mutagenesis, and prepare plasmids.

primer	stock #	sequences
cn1	TN68	5'- AACCGCAACAAACGATTAGC
cn2	TN69	5'- CGGAATTAGAAAGATTGATGATTTG
im1	TN60	5'- AAGTTTTGATGCTCAACAAATGGC
rc1	TN956	5'- CATTAAAAATCAACAAGTCTTGTCC
rc2	TN1772	5'- GTTACTATGGATAAAGATAATTGTTT
rc3	TN2278	5'- CCGTTAGTGAACGTAAATAATGAAACC
tr1	TN2279	5'- ACAAGCGTACTTGACATGCG
tr2	TN2280	5'- GCTTGCAGCTGAAATGTTTATTCG
clr4-1	TN674	5'- AACTCCAACGCCTCGAACAGCTGC
clr4-NHR-F	TN1950	5'- TATGGAGATGTCTCTGCTTTTTTTGCCGCCTCCTGTTCACC
clr4-NHR-R	TN1951	5'- GGTGAACAGGAGGCGGCAAAAAAAGCAGAGACATCTCCATAG
clr4-2	HM802	5'- GTCAGTGCCTCGTTCTC
mlo3-1	TN1747	5'- TCTGTTGCACTGAATCGTGC
mlo3-5	TN1751	5'- TCATCCAAACAAGCCGTGCC
mlo3-KA-R	TN1914	5'- CGTGTAGTCGCCCGCGCGGAAGATTTGGCGCCATTCTTGC
mlo3-KA-F	TN1913	5'- GCCAAATCTTCCGCGCGGGCGACTACACGCCGCCGTAGAAC
mlo3-KR-R	TN1925	5'- CGTGTAGTCCTCCGCCTGGAAGATTTGGCGCCATTC
mlo3-KR-F	TN1924	5'- GCCAAATCTTCCAGGCGGAGGACTACACGCCG
mlo3-4	TN1750	5'- TAACAGTAGCCGAAGCTACC
dh-1	TN370	5'- TGTCTCCATGTTGTTCGG
dh-2	TN371	5'- ACGCCCATTCATCAAGC
otr3-2	TN1704	5'- CGACAACAAAGCGACAATAGCAGTC
imr3-XhoI-R	TN2193	5'- TTAACAGGTCTCGAGGCCCAATGG
adl1-F	TN2233	5'- GTCTAGAATATGCCTCCCAAAAAGCG
adl1-R	TN2234	5'- TTTACGGTTCTGGGCCCCATTACCG
act1-F	TN2207	5'- GTACATTGCACCACTTCCGC
act1-R	TN2208	5'- AATAGGGACACGCGAGTTGC

Table 3. List of antibodies used for ChIP analysis.

Mouse antibody					
antibody	vol	beads	beads vol	IP	
H3K9me2	5 µl	Mouse	40 µl	6 hrs	(Hayashi-Takanaka et al. 2011)
H3K9me3	5 µl	Mouse	40 µl	6 hrs	
H3K9ac	5 µl	Mouse	40 µl	6 hrs	(Karmodiya et al. 2012)
H3K14ac	5 µl	Mouse	40 µl	6 hrs	
Rpb1	1 µl	Mouse	20 µl	2 hrs	Millipore, CTD4H8, 05-623
FLAG	1.2 µl	Mouse	30 µl	2 hrs	Sigma-Aldrich, F1804
Swi6	2 µl	Mouse	40 µl	6 hrs	(Nakayama et al. 2000)
Mhf2	10 µl	Protein A	40 µl	2 hrs	Sigma-Aldrich
Cnp20	10 µl	Protein A	40 µl	2 hrs	Sigma-Aldrich
Rabbit antibody					
andibody	vol	beads	beads vol	IP	
H3	1 µl	Rabbit	40 µl	2 hrs	Abcam, ab1791
Cnp1	4 µl	Rabbit	40 µl	2 hrs	Sigma-Aldrich

Mouse and Rabbit antibodies were raised against as follows:
Mhf2: NH_2 –CLELEDLENGIAAQLALDFS
Cnp20: NH_2 –CSLMQQYLSREIAPPAIKRT
Cnp1: NH_2 -MAKKSLMAEPGDPIPRPRKKRC

Table 4. List of PCR primers used in real time PCR.

primer	stock #	target site	sequences
RT-dg102-F	HM980	dg	5'- TTGCACTCGGTTTCAGCTAT
RT-dg102-R	HM981		5'- TGCTCTGACTTGGCTTGTCT
dh-F	TN1943	dh	5'- CAACAGTATGGGTATAGAAAGAAGAC
dh-R	TN1944		5'- TGCATGCAAGAAACTCCATAACTT
imr3-out-F1	TN2179	imr3	5'- TGTCCAATTCTAACCACTCTATTACGA
imr3-out-R1	TN2180		5'- CATCATCAGCAACTGTCATTCTCA
spbc713.06_F	TN33	*adl1*	5'- AAATATGGCGATCCAGGAGATG
spbc713.06_R	TN34		5'- GCTTAACGTGCGCACAGACA
act1(ORF)-F	HM3273	*act1*	5'- AGCGTGGTTATACTTTCTCTACT
act1(ORF)-F	HM3273		5'- AGCGTGGTTATACTTTCTCTACT
ade6-D-D-F	TN1154	*ade6*	5'- GCTCGTACCGCAGCTTCAAG
ade6-D-D-R	TN1155		5'- GCAACCATACCAGGCAAATGA
imr1-in-F	TN991	imr1-in	5'- ATTTCCGCTTACAAAATGCCA
imr1-in-R	TN992		5'- TTTCTCAACAGCAAAGCCTGAA
imr1-out-F	TN802	imr1-out	5'- GATGATATCGAGGCTTTCGGTTT
imr1-out-R	TN803		5'- TGTCCCTTCTGTTAAATTCTCGTGTA
cnt2-F	TN946	cnt2	5'- TGCCTCTCCCTTGCCAGTAA
cnt2-R	TN947		5'- TCGTTGCGGTGTTTTGAAAA

ACKNOWLEDGEMENTS

I would like to express my deepest appreciation to Associate Professor Takuro Nakagawa, who gave me the opportunity to study these important questions in the field of centromere integrity. His constructive suggestions have always helped and heartfelt comments encouraged me. His logical way of thinking has been of great value to me. I also appreciate Professor Hisao Masukata, Associate Professors Tatsuro Takahashi and Yumiko Kubota, for their comments and discussions, which was provided an enormous contribution to this work. I also appreciate Professors Chikashi Obuse, Akira Shinohara, Yasushi Hiraoka, and Kojiro Ishii for providing constructive comments on my study and the thesis. I would also like to thank all the Masukata lab members for helpful suggestions and for encouraging me. I am especially grateful to Dr. Faria Zafar, Mr. Jie Su, and Ms. Dayalini Weerasekara for helpful suggestions and supporting my study. Owing to their contribution, I was able to concentrate all of my energy into my study. Finally, I owe my sincere appreciation to my family for their constant support and understanding during the course of my study.

February 2019 Akiko Okita

REFERENCES

Abe Y, Sako K, Takagaki K, Hirayama Y, Uchida KS, Herman JA, DeLuca JG, Hirota T. 2016. HP1-Assisted Aurora B Kinase Activity Prevents Chromosome Segregation Errors. *Dev Cell* **36**: 487-497.

Alexandrov I, Kazakov A, Tumeneva I, Shepelev V, Yurov Y. 2001. Alpha-satellite DNA of primates: old and new families. *Chromosoma* **110**: 253-266.

Allshire RC, Nimmo ER, Ekwall K, Javerzat JP, Cranston G. 1995. Mutations derepressing silent centromeric domains in fission yeast disrupt chromosome segregation. *Genes Dev* **9**: 218-233.

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

Amon JD, Koshland D. 2016. RNase H enables efficient repair of R-loop induced DNA damage. in *Elife*.

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

Bailis JM, Bernard P, Antonelli R, Allshire RC, Forsburg SL. 2003. Hsk1-Dfp1 is required for heterochromatin-mediated cohesion at centromeres. *Nat Cell Biol* **5**: 1111-1116.

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

Barra V, Fachinetti D. 2018. The dark side of centromeres: types, causes and consequences of structural abnormalities implicating centromeric DNA. *Nat Commun* **9**: 4340.

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

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

Bhargava R, Onyango DO, Stark JM. 2016. Regulation of Single-Strand Annealing and its Role in Genome Maintenance. *Trends Genet* **32**: 566-575.

Bjerling P, Silverstein RA, Thon G, Caudy A, Grewal S, Ekwall K. 2002. Functional divergence between histone deacetylases in fission yeast by distinct cellular localization and in vivo

specificity. *Mol Cell Biol* **22**: 2170-2181.

Brown WR, Thomas G, Lee NC, Blythe M, Liti G, Warringer J, Loose MW. 2014. Kinetochore assembly and heterochromatin formation occur autonomously in Schizosaccharomyces pombe. *Proc Natl Acad Sci U S A* **111**: 1903-1908.

Bugreev DV, Mazina OM, Mazin AV. 2006. Rad54 protein promotes branch migration of Holliday junctions. *Nature* **442**: 590-593.

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

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

Buscaino A, Lejeune E, Audergon P, Hamilton G, Pidoux A, Allshire RC. 2013. Distinct roles for Sir2 and RNAi in centromeric heterochromatin nucleation, spreading and maintenance. *EMBO J* **32**: 1250-1264.

Cam HP, Sugiyama T, Chen ES, Chen X, FitzGerald PC, Grewal SI. 2005. Comprehensive analysis of heterochromatin- and RNAi-mediated epigenetic control of the fission yeast genome. *Nat Genet* **37**: 809-819.

Carvalho CM, Lupski JR. 2016. Mechanisms underlying structural variant formation in genomic disorders. *Nat Rev Genet* **17**: 224-238.

Cassart C, Drogat J, Migeot V, Hermand D. 2012. Distinct requirement of RNA polymerase II CTD phosphorylations in budding and fission yeast. *Transcription* **3**: 231-234.

Catania S, Pidoux AL, Allshire RC. 2015. Sequence features and transcriptional stalling within centromere DNA promote establishment of CENP-A chromatin. *PLoS Genet* **11**: e1004986.

Cernilogar FM, Onorati MC, Kothe GO, Burroughs AM, Parsi KM, Breiling A, Lo Sardo F, Saxena A, Miyoshi K, Siomi H et al. 2011. Chromatin-associated RNA interference components contribute to transcriptional regulation in Drosophila. *Nature* **480**: 391-395.

Chen ES, Zhang K, Nicolas E, Cam HP, Zofall M, Grewal SI. 2008. Cell cycle control of centromeric repeat transcription and heterochromatin assembly. *Nature* **451**: 734-737.

Chikashige Y, Kinoshita N, Nakaseko Y, Matsumoto T, Murakami S, Niwa O, Yanagida M. 1989. Composite motifs and repeat symmetry in S. pombe centromeres: direct analysis by integration of NotI restriction sites. *Cell* **57**: 739-751.

Costantino L, Koshland D. 2018. Genome-wide Map of R-Loop-Induced Damage Reveals How a Subset of R-Loops Contributes to Genomic Instability. *Mol Cell* **71**: 487-497 e483.

Crotti LB, Basrai MA. 2004. Functional roles for evolutionarily conserved Spt4p at centromeres and heterochromatin in Saccharomyces cerevisiae. *EMBO J* **23**: 1804-1814.

Egloff S, O'Reilly D, Chapman RD, Taylor A, Tanzhaus K, Pitts L, Eick D, Murphy S. 2007. Serine-7 of the RNA polymerase II CTD is specifically required for snRNA gene expression. *Science* **318**: 1777-1779.

Eick D, Geyer M. 2013. The RNA polymerase II carboxy-terminal domain (CTD) code. *Chem Rev* **113**: 8456-8490.

Ekwall K, Javerzat JP, Lorentz A, Schmidt H, Cranston G, Allshire R. 1995. The chromodomain protein Swi6: a key component at fission yeast centromeres. *Science* **269**: 1429-1431.

Ekwall K, Ruusala T. 1994. Mutations in rik1, clr2, clr3 and clr4 genes asymmetrically derepress the silent mating-type loci in fission yeast. *Genetics* **136**: 53-64.

Ellermeier C, Higuchi EC, Phadnis N, Holm L, Geelhood JL, Thon G, Smith GR. 2010. RNAi and heterochromatin repress centromeric meiotic recombination. *P Natl Acad Sci USA* **107**: 8701-8705.

Elmendorf BJ, Shilatifard A, Yan Q, Conaway JW, Conaway RC. 2001. Transcription factors TFIIF, ELL, and Elongin negatively regulate SII-induced nascent transcript cleavage by non-arrested RNA polymerase II elongation intermediates. *J Biol Chem* **276**: 23109-23114.

Fanti L, Giovinazzo G, Berloco M, Pimpinelli S. 1998. The heterochromatin protein 1 prevents telomere fusions in Drosophila. *Mol Cell* **2**: 527-538.

Fischer T, Cui BW, Dhakshnamoorthy J, Zhou M, Rubin C, Zofall M, Veenstra TD, Grewal SIS. 2009. Diverse roles of HP1 proteins in heterochromatin assembly and functions in fission yeast. *P Natl Acad Sci USA* **106**: 8998-9003.

Gao J, Kan F, Wagnon JL, Storey AJ, Protacio RU, Davidson MK, Wahls WP. 2014. Rapid, efficient and precise allele replacement in the fission yeast Schizosaccharomyces pombe. *Curr Genet* **60**: 109-119.

Garcia-Rubio M, Aguilera P, Lafuente-Barquero J, Ruiz JF, Simon MN, Geli V, Rondon AG, Aguilera A. 2018. Yra1-bound RNA-DNA hybrids cause orientation-independent transcription-replication collisions and telomere instability. *Genes Dev* **32**: 965-977.

Gavalda S, Santos-Pereira JM, Garcia-Rubio ML, Luna R, Aguilera A. 2016. Excess of Yra1 RNA-Binding Factor Causes Transcription-Dependent Genome Instability, Replication Impairment and Telomere Shortening. *PLoS Genet* **12**: e1005966.

Gomez-Herreros F, de Miguel-Jimenez L, Millan-Zambrano G, Penate X, Delgado-Ramos L,

Munoz-Centeno MC, Chavez S. 2012. One step back before moving forward: regulation of transcription elongation by arrest and backtracking. *FEBS Lett* **586**: 2820-2825.

Gordon DJ, Resio B, Pellman D. 2012. Causes and consequences of aneuploidy in cancer. *Nat Rev Genet* **13**: 189-203.

Gossett AJ, Lieb JD. 2012. In vivo effects of histone H3 depletion on nucleosome occupancy and position in Saccharomyces cerevisiae. *PLoS Genet* **8**: e1002771.

Grewal SI, Bonaduce MJ, Klar AJ. 1998. Histone deacetylase homologs regulate epigenetic inheritance of transcriptional silencing and chromosome segregation in fission yeast. *Genetics* **150**: 563-576.

Grewal SI, Jia S. 2007. Heterochromatin revisited. *Nat Rev Genet* **8**: 35-46.

Gu W, Reines D. 1995. Identification of a decay in transcription potential that results in elongation factor dependence of RNA polymerase II. *J Biol Chem* **270**: 11238-11244.

Guenatri M, Bailly D, Maison C, Almouzni G. 2004. Mouse centric and pericentric satellite repeats form distinct functional heterochromatin. *J Cell Biol* **166**: 493-505.

Halic M, Moazed D. 2010. Dicer-independent primal RNAs trigger RNAi and heterochromatin formation. *Cell* **140**: 504-516.

Hamerton JL, Canning N, Ray M, Smith S. 1975. A cytogenetic survey of 14,069 newborn infants. I. Incidence of chromosome abnormalities. *Clin Genet* **8**: 223-243.

Hansen KR, Hazan I, Shanker S, Watt S, Verhein-Hansen J, Bahler J, Martienssen RA, Partridge JF, Cohen A, Thon G. 2011. H3K9me-independent gene silencing in fission yeast heterochromatin by Clr5 and histone deacetylases. *PLoS Genet* **7**: e1001268.

Harlen KM, Churchman LS. 2017. The code and beyond: transcription regulation by the RNA polymerase II carboxy-terminal domain. *Nat Rev Mol Cell Biol* **18**: 263-273.

Hayashi MT, Takahashi TS, Nakagawa T, Nakayama J, Masukata H. 2009. The heterochromatin protein Swi6/HP1 activates replication origins at the pericentromeric region and silent mating-type locus. *Nat Cell Biol* **11**: 357-362.

Hayashi-Takanaka Y, Yamagata K, Wakayama T, Stasevich TJ, Kainuma T, Tsurimoto T, Tachibana M, Shinkai Y, Kurumizaka H, Nozaki N et al. 2011. Tracking epigenetic histone modifications in single cells using Fab-based live endogenous modification labeling. *Nucleic Acids Res* **39**: 6475-6488.

Hayles J, Nurse P. 2018. Introduction to Fission Yeast as a Model System. *Cold Spring Harb Protoc* **2018**: pdb top079749.

Helmrich A, Ballarino M, Nudler E, Tora L. 2013. Transcription-replication encounters,

consequences and genomic instability. *Nat Struct Mol Biol* **20**: 412-418.

Hermsen MA, Joenje H, Arwert F, Welters MJ, Braakhuis BJ, Bagnay M, Westerveld A, Slater R. 1996. Centromeric breakage as a major cause of cytogenetic abnormalities in oral squamous cell carcinoma. *Genes Chromosomes Cancer* **15**: 1-9.

Holoch D, Moazed D. 2015. Small-RNA loading licenses Argonaute for assembly into a transcriptional silencing complex. *Nat Struct Mol Biol* **22**: 328-335.

Izban MG, Luse DS. 1992. The RNA polymerase II ternary complex cleaves the nascent transcript in a 3'----5' direction in the presence of elongation factor SII. *Genes Dev* **6**: 1342-1356.

Jensen RB, Carreira A, Kowalczykowski SC. 2010. Purified human BRCA2 stimulates RAD51-mediated recombination. *Nature* **467**: 678-683.

Jih G, Iglesias N, Currie MA, Bhanu NV, Paulo JA, Gygi SP, Garcia BA, Moazed D. 2017. Unique roles for histone H3K9me states in RNAi and heritable silencing of transcription. *Nature*.

Jones KW. 1970. Chromosomal and nuclear location of mouse satellite DNA in individual cells. *Nature* **225**: 912-915.

Joseph A, Mitchell AR, Miller OJ. 1989. The organization of the mouse satellite DNA at centromeres. *Exp Cell Res* **183**: 494-500.

Kajitani T, Kato H, Chikashige Y, Tsutsumi C, Hiraoka Y, Kimura H, Ohkawa Y, Obuse C, Hermand D, Murakami Y. 2017. Ser7 of RNAPII-CTD facilitates heterochromatin formation by linking ncRNA to RNAi. *Proc Natl Acad Sci U S A* **114**: E11208-E11217.

Kapoor S, Zhu L, Froyd C, Liu T, Rusche LN. 2015. Regional centromeres in the yeast Candida lusitaniae lack pericentromeric heterochromatin. *Proc Natl Acad Sci U S A* **112**: 12139-12144.

Karmodiya K, Krebs AR, Oulad-Abdelghani M, Kimura H, Tora L. 2012. H3K9 and H3K14 acetylation co-occur at many gene regulatory elements, while H3K14ac marks a subset of inactive inducible promoters in mouse embryonic stem cells. *BMC Genomics* **13**: 424.

Keeney S, Giroux CN, Kleckner N. 1997. Meiosis-specific DNA double-strand breaks are catalyzed by Spo11, a member of a widely conserved protein family. *Cell* **88**: 375-384.

Kettenberger H, Armache KJ, Cramer P. 2003. Architecture of the RNA polymerase II-TFIIS complex and implications for mRNA cleavage. *Cell* **114**: 347-357.

Kim DH, Villeneuve LM, Morris KV, Rossi JJ. 2006. Argonaute-1 directs siRNA-mediated transcriptional gene silencing in human cells. *Nat Struct Mol Biol* **13**: 793-797.

Kim J, Guermah M, Roeder RG. 2010. The human PAF1 complex acts in chromatin transcription elongation both independently and cooperatively with SII/TFIIS. *Cell* **140**: 491-503.

Kim N, Jinks-Robertson S. 2012. Transcription as a source of genome instability. *Nat Rev Genet* **13**: 204-214.

Kimura M, Suzuki H, Ishihama A. 2002. Formation of a carboxy-terminal domain phosphatase (Fcp1)/TFIIF/RNA polymerase II (pol II) complex in Schizosaccharomyces pombe involves direct interaction between Fcp1 and the Rpb4 subunit of pol II. *Mol Cell Biol* **22**: 1577-1588.

Kusevic D, Kudithipudi S, Iglesias N, Moazed D, Jeltsch A. 2017. Clr4 specificity and catalytic activity beyond H3K9 methylation. *Biochimie* **135**: 83-88.

Lander ES Linton LM Birren B Nusbaum C Zody MC Baldwin J Devon K Dewar K Doyle M FitzHugh W et al. 2001. Initial sequencing and analysis of the human genome. *Nature* **409**: 860-921.

Larson AG, Elnatan D, Keenen MM, Trnka MJ, Johnston JB, Burlingame AL, Agard DA, Redding S, Narlikar GJ. 2017. Liquid droplet formation by HP1alpha suggests a role for phase separation in heterochromatin. *Nature* **547**: 236-240.

Lea DE, Coulson CA. 1949. The distribution of the numbers of mutants in bacterial populations. *J Genet* **49**: 264-285.

Lejeune J, Gautier M, Turpin R. 1959. [Study of somatic chromosomes from 9 mongoloid children]. *C R Hebd Seances Acad Sci* **248**: 1721-1722.

Li PC, Petreaca RC, Jensen A, Yuan JP, Green MD, Forsburg SL. 2013. Replication fork stability is essential for the maintenance of centromere integrity in the absence of heterochromatin. *Cell Rep* **3**: 638-645.

Lin M, Chang CJ, Green NS. 1996. A new method for estimating high mutation rates in cultured cells. *Mutat Res* **351**: 105-116.

Lisica A, Engel C, Jahnel M, Roldan E, Galburt EA, Cramer P, Grill SW. 2016. Mechanisms of backtrack recovery by RNA polymerases I and II. *Proc Natl Acad Sci U S A* **113**: 2946-2951.

Losada A, Hirano M, Hirano T. 1998. Identification of Xenopus SMC protein complexes required for sister chromatid cohesion. *Genes Dev* **12**: 1986-1997.

MacKellar AL, Greenleaf AL. 2011. Cotranscriptional association of mRNA export factor Yra1 with C-terminal domain of RNA polymerase II. *J Biol Chem* **286**: 36385-36395.

Martinez AC, van Wely KH. 2011. Centromere fission, not telomere erosion, triggers chromosomal instability in human carcinomas. *Carcinogenesis* **32**: 796-803.

Masai H, Matsumoto S, You Z, Yoshizawa-Sugata N, Oda M. 2010. Eukaryotic chromosome

DNA replication: where, when, and how? *Annu Rev Biochem* **79**: 89-130.

McDowall MD, Harris MA, Lock A, Rutherford K, Staines DM, Bahler J, Kersey PJ, Oliver SG, Wood V. 2015. PomBase 2015: updates to the fission yeast database. *Nucleic Acids Res* **43**: D656-661.

McKinley KL, Cheeseman IM. 2016. The molecular basis for centromere identity and function. *Nat Rev Mol Cell Biol* **17**: 16-29.

Mellone BG, Ball L, Suka N, Grunstein MR, Partridge JF, Allshire RC. 2003. Centromere silencing and function in fission yeast is governed by the amino terminus of histone H3. *Curr Biol* **13**: 1748-1757.

Min J, Zhang X, Cheng X, Grewal SI, Xu RM. 2002. Structure of the SET domain histone lysine methyltransferase Clr4. *Nat Struct Biol* **9**: 828-832.

Mirkin EV, Mirkin SM. 2007. Replication fork stalling at natural impediments. *Microbiol Mol Biol Rev* **71**: 13-35.

Motamedi MR, Hong EJ, Li X, Gerber S, Denison C, Gygi S, Moazed D. 2008. HP1 proteins form distinct complexes and mediate heterochromatic gene silencing by nonoverlapping mechanisms. *Mol Cell* **32**: 778-790.

Muller S, Almouzni G. 2017. Chromatin dynamics during the cell cycle at centromeres. *Nat Rev Genet* **18**: 192-208.

Nakamura K, Okamoto A, Katou Y, Yadani C, Shitanda T, Kaweeteerawat C, Takahashi TS, Itoh T, Shirahige K, Masukata H et al. 2008. Rad51 suppresses gross chromosomal rearrangement at centromere in Schizosaccharomyces pombe. *EMBO J* **27**: 3036-3046.

Nakayama J, Klar AJ, Grewal SI. 2000. A chromodomain protein, Swi6, performs imprinting functions in fission yeast during mitosis and meiosis. *Cell* **101**: 307-317.

Nakayama J, Rice JC, Strahl BD, Allis CD, Grewal SI. 2001. Role of histone H3 lysine 9 methylation in epigenetic control of heterochromatin assembly. *Science* **292**: 110-113.

Nassif N, Penney J, Pal S, Engels WR, Gloor GB. 1994. Efficient copying of nonhomologous sequences from ectopic sites via P-element-induced gap repair. *Mol Cell Biol* **14**: 1613-1625.

New JH, Sugiyama T, Zaitseva E, Kowalczykowski SC. 1998. Rad52 protein stimulates DNA strand exchange by Rad51 and replication protein A. *Nature* **391**: 407-410.

Nicolas E, Yamada T, Cam HP, Fitzgerald PC, Kobayashi R, Grewal SI. 2007. Distinct roles of HDAC complexes in promoter silencing, antisense suppression and DNA damage protection. *Nat Struct Mol Biol* **14**: 372-380.

Nielsen J, Wohlert M. 1991. Chromosome abnormalities found among 34,910 newborn children: results from a 13-year incidence study in Arhus, Denmark. *Hum Genet* **87**: 81-83.

Onaka AT, Toyofuku N, Inoue T, Okita AK, Sagawa M, Su J, Shitanda T, Matsuyama R, Zafar F, Takahashi TS et al. 2016. Rad51 and Rad54 promote noncrossover recombination between centromere repeats on the same chromatid to prevent isochromosome formation. *Nucleic Acids Res* **44**: 10744-10757.

Padeken J, Zeller P, Gasser SM. 2015. Repeat DNA in genome organization and stability. *Curr Opin Genet Dev* **31**: 12-19.

Palmer CG, Reichmann A. 1976. Chromosomal and clinical findings in 110 females with Turner syndrome. *Hum Genet* **35**: 35-49.

Pardue ML, Gall JG. 1970. Chromosomal localization of mouse satellite DNA. *Science* **168**: 1356-1358.

Partridge JF, Scott KS, Bannister AJ, Kouzarides T, Allshire RC. 2002. cis-acting DNA from fission yeast centromeres mediates histone H3 methylation and recruitment of silencing factors and cohesin to an ectopic site. *Curr Biol* **12**: 1652-1660.

Peng JC, Karpen GH. 2007. H3K9 methylation and RNA interference regulate nucleolar organization and repeated DNA stability. *Nat Cell Biol* **9**: 25-35.

Peters AH, O'Carroll D, Scherthan H, Mechtler K, Sauer S, Schofer C, Weipoltshammer K, Pagani M, Lachner M, Kohlmaier A et al. 2001. Loss of the Suv39h histone methyltransferases impairs mammalian heterochromatin and genome stability. *Cell* **107**: 323-337.

Petukhova G, Stratton S, Sung P. 1998. Catalysis of homologous DNA pairing by yeast Rad51 and Rad54 proteins. *Nature* **393**: 91-94.

Ragunathan K, Jih G, Moazed D. 2015. Epigenetics. Epigenetic inheritance uncoupled from sequence-specific recruitment. *Science* **348**: 1258699.

Rea S, Eisenhaber F, O'Carroll D, Strahl BD, Sun ZW, Schmid M, Opravil S, Mechtler K, Ponting CP, Allis CD et al. 2000. Regulation of chromatin structure by site-specific histone H3 methyltransferases. *Nature* **406**: 593-599.

Reyes-Turcu FE, Zhang K, Zofall M, Chen E, Grewal SI. 2011. Defects in RNA quality control factors reveal RNAi-independent nucleation of heterochromatin. *Nat Struct Mol Biol* **18**: 1132-1138.

Sadaie M, Kawaguchi R, Ohtani Y, Arisaka F, Tanaka K, Shirahige K, Nakayama J. 2008. Balance between distinct HP1 family proteins controls heterochromatin assembly in fission yeast. *Mol Cell Biol* **28**: 6973-6988.

Sanchez AM, Shuman S, Schwer B. 2018. RNA polymerase II CTD interactome with 3' processing and termination factors in fission yeast and its impact on phosphate homeostasis. *Proc Natl Acad Sci U S A* **115**: E10652-E10661.

Santos-Pereira JM, Aguilera A. 2015. R loops: new modulators of genome dynamics and function. *Nat Rev Genet* **16**: 583-597.

Sato H, Masuda F, Takayama Y, Takahashi K, Saitoh S. 2012. Epigenetic inactivation and subsequent heterochromatinization of a centromere stabilize dicentric chromosomes. *Curr Biol* **22**: 658-667.

Schalch T, Job G, Noffsinger VJ, Shanker S, Kuscu C, Joshua-Tor L, Partridge JF. 2009. High-affinity binding of Chp1 chromodomain to K9 methylated histone H3 is required to establish centromeric heterochromatin. *Mol Cell* **34**: 36-46.

Shang WH, Hori T, Martins NM, Toyoda A, Misu S, Monma N, Hiratani I, Maeshima K, Ikeo K, Fujiyama A et al. 2013. Chromosome engineering allows the efficient isolation of vertebrate neocentromeres. *Dev Cell* **24**: 635-648.

Shankaranarayana GD, Motamedi MR, Moazed D, Grewal SI. 2003. Sir2 regulates histone H3 lysine 9 methylation and heterochromatin assembly in fission yeast. *Curr Biol* **13**: 1240-1246.

Shimada A, Dohke K, Sadaie M, Shinmyozu K, Nakayama J, Urano T, Murakami Y. 2009. Phosphorylation of Swi6/HP1 regulates transcriptional gene silencing at heterochromatin. *Genes Dev* **23**: 18-23.

Shinohara A, Ogawa T. 1998. Stimulation by Rad52 of yeast Rad51-mediated recombination. *Nature* **391**: 404-407.

Sigurdsson S, Dirac-Svejstrup AB, Svejstrup JQ. 2010. Evidence that transcript cleavage is essential for RNA polymerase II transcription and cell viability. *Mol Cell* **38**: 202-210.

Sims RJ, 3rd, Belotserkovskaya R, Reinberg D. 2004. Elongation by RNA polymerase II: the short and long of it. *Genes Dev* **18**: 2437-2468.

Strässer K, Hurt E. 2000. Yra1p, a conserved nuclear RNA-binding protein, interacts directly with Mex67p and is required for mRNA export. *EMBO J* **19**: 410-420.

Strasser K, Masuda S, Mason P, Pfannstiel J, Oppizzi M, Rodriguez-Navarro S, Rondon AG, Aguilera A, Struhl K, Reed R et al. 2002. TREX is a conserved complex coupling transcription with messenger RNA export. *Nature* **417**: 304-308.

Tanaka T, Fuchs J, Loidl J, Nasmyth K. 2000. Cohesin ensures bipolar attachment of microtubules to sister centromeres and resists their precocious separation. *Nat Cell Biol* **2**: 492-499.

Thakurta AG, Gopal G, Yoon JH, Kozak L, Dhar R. 2005. Homolog of BRCA2-interacting Dss1p and Uap56p link Mlo3p and Rae1p for mRNA export in fission yeast. *EMBO J* **24**: 2512-2523.

Ting DT, Lipson D, Paul S, Brannigan BW, Akhavanfard S, Coffman EJ, Contino G, Deshpande V, Iafrate AJ, Letovsky S et al. 2011. Aberrant overexpression of satellite repeats in pancreatic and other epithelial cancers. *Science* **331**: 593-596.

Tomson BN, Arndt KM. 2013. The many roles of the conserved eukaryotic Paf1 complex in regulating transcription, histone modifications, and disease states. *Biochim Biophys Acta* **1829**: 116-126.

Trewick SC, Minc E, Antonelli R, Urano T, Allshire RC. 2007. The JmjC domain protein Epe1 prevents unregulated assembly and disassembly of heterochromatin. *EMBO J* **26**: 4670-4682.

Tsukada Y, Fang J, Erdjument-Bromage H, Warren ME, Borchers CH, Tempst P, Zhang Y. 2006. Histone demethylation by a family of JmjC domain-containing proteins. *Nature* **439**: 811-816.

Verdel A, Jia S, Gerber S, Sugiyama T, Gygi S, Grewal SI, Moazed D. 2004. RNAi-mediated targeting of heterochromatin by the RITS complex. *Science* **303**: 672-676.

Wada T, Takagi T, Yamaguchi Y, Ferdous A, Imai T, Hirose S, Sugimoto S, Yano K, Hartzog GA, Winston F et al. 1998. DSIF, a novel transcription elongation factor that regulates RNA polymerase II processivity, is composed of human Spt4 and Spt5 homologs. *Genes Dev* **12**: 343-356.

Wahba L, Costantino L, Tan FJ, Zimmer A, Koshland D. 2016. S1-DRIP-seq identifies high expression and polyA tracts as major contributors to R-loop formation. *Genes Dev* **30**: 1327-1338.

Wang D, Bushnell DA, Huang X, Westover KD, Levitt M, Kornberg RD. 2009. Structural basis of transcription: backtracked RNA polymerase II at 3.4 angstrom resolution. *Science* **324**: 1203-1206.

Wang YH, Tsay YG, Tan BC, Lo WY, Lee SC. 2003. Identification and characterization of a novel p300-mediated p53 acetylation site, lysine 305. *J Biol Chem* **278**: 25568-25576.

Weckselblatt B, Rudd MK. 2015. Human Structural Variation: Mechanisms of Chromosome Rearrangements. *Trends Genet* **31**: 587-599.

Weischenfeldt J, Symmons O, Spitz F, Korbel JO. 2013. Phenotypic impact of genomic structural variation: insights from and for human disease. *Nat Rev Genet* **14**: 125-138.

Wilson MD, Harreman M, Svejstrup JQ. 2013. Ubiquitylation and degradation of elongating RNA polymerase II: the last resort. *Biochim Biophys Acta* **1829**: 151-157.

Wolff DJ, Miller AP, Van Dyke DL, Schwartz S, Willard HF. 1996. Molecular definition of breakpoints associated with human Xq isochromosomes: implications for mechanisms of formation. *Am J Hum Genet* **58**: 154-160.

Wong AK, Rattner JB. 1988. Sequence organization and cytological localization of the minor satellite of mouse. *Nucleic Acids Res* **16**: 11645-11661.

Wright WD, Heyer WD. 2014. Rad54 functions as a heteroduplex DNA pump modulated by its DNA substrates and Rad51 during D loop formation. *Mol Cell* **53**: 420-432.

Wyllie JP, Wright MJ, Burn J, Hunter S. 1994. Natural history of trisomy 13. *Arch Dis Child* **71**: 343-345.

Yamada T, Fischle W, Sugiyama T, Allis CD, Grewal SI. 2005. The nucleation and maintenance of heterochromatin by a histone deacetylase in fission yeast. *Mol Cell* **20**: 173-185.

Yamada T, Yamaguchi Y, Inukai N, Okamoto S, Mura T, Handa H. 2006. P-TEFb-mediated phosphorylation of hSpt5 C-terminal repeats is critical for processive transcription elongation. *Mol Cell* **21**: 227-237.

Zafar F, Okita AK, Onaka AT, Su J, Katahira Y, Nakayama JI, Takahashi TS, Masukata H, Nakagawa T. 2017. Regulation of mitotic recombination between DNA repeats in centromeres. *Nucleic Acids Res* **45**: 11222-11235.

Zaratiegui M, Castel SE, Irvine DV, Kloc A, Ren J, Li F, de Castro E, Marin L, Chang AY, Goto D et al. 2011. RNAi promotes heterochromatic silencing through replication-coupled release of RNA Pol II. *Nature* **479**: 135-138.

Zeller P, Padeken J, van Schendel R, Kalck V, Tijsterman M, Gasser SM. 2016. Histone H3K9 methylation is dispensable for Caenorhabditis elegans development but suppresses RNA:DNA hybrid-associated repeat instability. *Nat Genet* **48**: 1385-1395.

Zhang K, Fischer T, Porter RL, Dhakshnamoorthy J, Zofall M, Zhou M, Veenstra T, Grewal SI. 2011. Clr4/Suv39 and RNA quality control factors cooperate to trigger RNAi and suppress antisense RNA. *Science* **331**: 1624-1627.

Zhang K, Mosch K, Fischle W, Grewal SI. 2008. Roles of the Clr4 methyltransferase complex in nucleation, spreading and maintenance of heterochromatin. *Nat Struct Mol Biol* **15**: 381-388.

Zhou Q, Li T, Price DH. 2012. RNA polymerase II elongation control. *Annu Rev Biochem* **81**: 119-143.

Zhou Z, Luo MJ, Straesser K, Katahira J, Hurt E, Reed R. 2000. The protein Aly links pre-messenger-RNA splicing to nuclear export in metazoans. *Nature* **407**: 401-405.

Zofall M, Grewal SI. 2006. Swi6/HP1 recruits a JmjC domain protein to facilitate transcription of heterochromatic repeats. *Mol Cell* **22**: 681-692.