Title: Repetitive DNA promotes RNAi-mediated heterochromatin formation via an anti-silencing factor in fission yeast.

Authors: Takahiro Asanuma[1], Soichi Inagaki[2], Tetsuji Kakutani[2], Hiroyuki Aburatani[3], Yota Murakami[1]*

Affiliations:

[1]Department of Chemistry, Faculty of Science, Hokkaido University; Sapporo, 060-0810, Japan.

[2] Department of Biological Sciences, Faculty of Science, University of Tokyo; Tokyo, 113-0033, Japan.

[3] Genome Science Division, Research Center for Advanced Science and Technology, University of Tokyo; Tokyo, 153-8904, Japan.

*Corresponding author. Email: yota@sci.hokudai.ac.jp

Introductory paragraph:

In most eukaryotes, constitutive heterochromatin defined by H3K9me2/me3 is enriched on repetitive DNA, such as the satellite repeats of pericentromeres[1]. Furthermore, repetitive transgenes often cause the formation of silent heterochromatin in diverse model organisms[2]. Although these facts suggest that the repetition itself promotes heterochromatin formation, the mechanism by which this occurs is still unclear[3-7]. Here, using *Schizosaccharomyces pombe*, we show that tandemly repeated reporter genes promote RNA interference (RNAi)-mediated heterochromatin formation in cooperation with an anti-silencing factor. The repeated gene itself does not cause heterochromatin formation; however, once the RNAi recognizes it via artificial small RNAs, the repeated gene starts producing cognate small RNAs to autonomously maintain heterochromatin. This depends on the number of repeats and an anti-silencing factor Epe1, which removes H3K9me and derepresses transcription of genes underlying heterochromatin. Our results suggest that an anti-silencing factor generates sufficient transcripts for the activation of RNAi when repetitive DNA underlies silent heterochromatin.

Main Text:

RNAi pathway is a highly conserved mechanism mediated by small RNAs, which directs not only post-transcriptional gene silencing but also transcriptional silencing by heterochromatin formation. In fission yeast, the RNAi pathway promotes heterochromatin formation at pericentromeric and subtelomeric regions, and at the mating-type locus[8]. These constitutive heterochromatin regions are defined by histone H3 lysine 9 di- or trimethylation (H3K9me) without DNA CpG methylation, and consist of common sequence, namely, *dg*/*dh* elements. From *dg*/*dh* elements, non-coding RNAs (ncRNAs) are transcribed by RNA polymerase II (Pol2) to act as scaffolds for the assembly of the RNAi machinery[9,10]. The RNA-induced transcriptional silencing (RITS) complex, containing argonaute protein, Ago1, recognizes nascent *dg*/*dh* ncRNAs via cognate small interfering RNAs (siRNAs)[11]. This results in recruitment of the CLRC complex, containing the Suv39h homolog, Clr4, which directs all H3K9me in this organism[12]. The RITS complex also recruits the RNA-dependent RNA polymerase complex (RDRC) to synthesize double-stranded RNAs (dsRNAs), which are processed to secondary siRNAs by Dicer (Dcr1)[13]. Thus, this process creates a positive feedback loop of siRNA production[14]. However, how RNAi-mediated heterochromatin formation is restricted to *dg/dh* ncRNAs and does not function on other mRNAs transcribed by the same Pol2 is still not clear[15].

Recent studies using artificial tethering of Clr4 at euchromatin have shown that ectopically deposited H3K9me is actively removed, meaning that the epigenetic inheritance of H3K9me is strictly limited in the cell[16,17]. These studies also demonstrated that the JmjC domain demethylase family protein, Epe1, is responsible for H3K9me demethylation. Notably, despite its function as an eraser, Epe1 is recruited to constitutive heterochromatin through its interaction

with the H3K9me binding protein, Swi6/HP1[18]. Loss of Epe1 (*epe1Δ*) results in fewer siRNAs and defective silencing at constitutive heterochromatin[19], implying its involvement in the RNAi pathway. However, its exact role is still not understood.

To define the link between Epe1 and the RNAi pathway, we examined the effect of Epe1 on the assembly of the RNAi machinery at constitutive heterochromatin. While loss of Epe1 dissociated the RNAi machinery from constitutive heterochromatin (Fig. 1a and Extended Data Fig. 1a-c), overproduction of Epe1 (Epe1 OP) (Extended Data Fig. 1d-g) significantly promoted its assembly, with concomitant up-regulation of siRNAs (Fig. 1a,c and Extended Data Fig. 1h). Consistent with previous reports showing that Epe1 overproduction causes derepression of a reporter gene inserted within constitutive heterochromatin[18-20](Extended Data Fig. 1g), Pol2 occupancy and ncRNA expression at *dg*/*dh* were significantly increased by Epe1 OP (Fig. 1d and Extended Data Fig. 1i). These results suggest that Epe1 promotes assembly of the RNAi machinery at constitutive heterochromatin by expressing scaffold ncRNA. Despite the high expression of *dg*/*dh* ncRNAs, high H3K9me levels were maintained in Epe1 OP cells, probably because the hyper-activated RNAi replenished H3K9me (Fig. 1a-c).

In wild-type cells, *dg*/*dh* ncRNAs are hardly detectable due to silencing by heterochromatin and degradation by Dcr1, making it difficult to analyze them. Therefore, the higher expression of *dg*/*dh* ncRNA in Epe1 OP strain facilitated its analysis. Cap Analysis of Gene Expression (CAGE) -seq, which can profile transcription start sites (TSSs) of capped RNAs transcribed by Pol2, revealed that Epe1 OP activates weak transcription from TSSs that are widespread across constitutive heterochromatin (Fig. 1e and Extended Data Fig. 2a-c). Consensus sequence analysis of those TSSs revealed that many of them share the same features as those of mRNAs. Namely, the Y/R (Y=pyrimidine, R=purine) dinucleotide is enriched at -1/+1 positions (Fig. 1f) and an A/T rich region is observed -25 to 32 nt upstream (Extended Data

Fig. 2d). These motifs correspond to an initiator (Inr) element and a TATA box, respectively[21]. Furthermore, 5'RACE analysis revealed that, while Epe1 activates widespread TSSs in *dg*/*dh* elements, it promotes expression of a reporter gene inserted within constitutive heterochromatin via TSSs that are almost identical to endogenous TSSs (Extended Data Fig. 3 and Supplementary Discussion 1). Thus, Epe1 induces transcription from heterochromatin by following the underlying DNA sequence, and that *dg*/*dh* elements encode a distinctive structure with widespread TSSs that consist of closely spaced canonical core promoter elements (Fig. 1g).

A paradox of RNAi-mediated heterochromatin formation is that "silent" heterochromatin formation requires "transcription" of the corresponding region. Therefore, the multiple TSSs in *dg/dh* elements appear to supply enough RNA template for RNAi without robust transcription that can disrupt heterochromatin[22]. Previous studies using hairpin RNA demonstrated that targeting the RNAi to mRNAs in *trans* with artificial siRNAs (*trans*-acting RNAi) hardly induces heterochromatin formation[15,23,24], and even if heterochromatin is formed, it is not maintained after removal of hairpin RNA[24,25]. We therefore hypothesized that when target genes are repeated, sufficient RNA templates are supplied for the RNAi, as is the case of the multiple TSSs in *dg*/*dh* elements.

To test this hypothesis, we generated *ade6*$^+$x1, x2, x4, x6, and x8 strains that also carry the hairpin RNA construct complementary to a part of the *ade6* ORF (*ade6*-hp) (Fig. 2a and Extended Data Figs. 4 and 5). Silencing of *ade6*$^+$ results in red-pink colony formation on indicator plates. Because post-transcriptional gene silencing is negligible in fission yeast[15,23,24], the red-pink colony reflects heterochromatin formation on *ade6*$^+$. Without *ade6*-hp, both the minimum *ade6*$^+$x1 and the maximum *ade6*$^+$x8 strains formed only white (*ade6*-expressing) colonies (Fig. 2b,c). However, when *ade6*-hp was expressed, red-pink (*ade6*-repressed) colonies appeared and were maintained at a rate that correlated with *ade6*$^+$ copy number (Fig. 2b-d). As

expected from the silencing assay, H3K9me was enriched on *ade6*$^+$ and this correlated with its copy number: H3K9me at the maximum *ade6*$^+$x8 reached the same level as pericentromeric heterochromatin (Fig. 2e,f). Meanwhile, *ade6*$^+$x(8+1) strains, which have the endogenous *ade6*$^+$ gene at a distant site from the *ade6*$^+$x8, did not form red-pink (*ade6*-repressed) colonies by *trans*-acting RNAi (Fig. 2d). Furthermore, when the endogenous *ade6*$^+$ gene was combined with the *ade6*$^+$x8 already heterochromatinized by *trans*-acting RNAi, the cells failed to stably maintain the red-pink (*ade6*-repressed) phenotype (Extended Data Fig. 6a,b), because H3K9me was hardly deposited on an isolated *ade6*$^+$ gene (Extended Data Fig. 6c,d). Thus, repeated genes promote RNAi-mediated heterochromatin formation, but this is not due to their gene dosage.

Importantly, heterochromatin formation at the *ade6*$^+$x8 was accompanied by significant production of novel siRNAs that were not encoded by *ade6*-hp, namely, secondary siRNAs (Fig. 3a,b). This result indicates that *trans*-acting RNAi efficiently activates the RNAi in *cis* on *ade6*$^+$ mRNAs (*cis*-acting RNAi) at *ade6*$^+$x8. By contrast, secondary siRNA was hardly detected at *ade6*$^+$x1 (Fig. 3a), indicating that repeated genes promote the *cis*-acting RNAi more efficiently than a single gene. Indeed, total *ade6* siRNA levels, which include secondary siRNAs that result from *cis*-acting RNAi, correlated with the number of *ade6*$^+$ repeat copies (Fig. 2e).

Furthermore, this effective *cis*-acting RNAi on *ade6*$^+$x8 was autonomously maintained without *trans*-acting RNAi. Even after removal of *ade6*-hp (*ade6*-hp*Δ*), the cells continued to show red-pink (*ade6*-repressed) phenotype (Fig. 4a), and H3K9me on *ade6*$^+$ was maintained at the same level as if *ade6*-hp is present (Fig. 4b). Furthermore, *ade6*$^+$ siRNAs, which exhibited the same properties as *dg*/*dh* siRNAs (Extended Data Fig. 7), were autonomously produced from *ade6*$^+$x8 (Fig. 4c), and resultant heterochromatin was inherited through both mitosis and meiosis in an RNAi-dependent manner (Fig. 4c, d, and f). These results indicate that, although repetition of *ade6*$^+$ itself does not cause heterochromatin formation, once recognized, it starts to function as

a platform for the RNAi pathway, like *dg*/*dh* elements. By contrast, when *ade6*-hp was removed after heterochromatin establishment, *ade6*$^{+}$x1, x2 strains failed to maintain H3K9me, indicating that the establishment of the autonomous *cis*-acting RNAi requires a minimal number of gene repeats (Fig. 5). We have therefore named this autonomous *cis*-acting RNAi "repeat-induced RNAi".

As expected, the RITS complex was localized at *ade6*$^{+}$x8 where repeat-induced RNAi maintains heterochromatin (Fig. 4g). Notably, loss of Epe1 significantly reduced its localization (Fig. 4g), and depleted *ade6* siRNAs (Fig. 4c). These results indicate that the repeat-induced RNAi also depends on Epe1, like the RNAi pathway at *dg*/*dh* elements. Meanwhile, H3K9me on *ade6*$^{+}$x8 was maintained in *epe1Δ* cells with variegated silencing (Fig. 4e,f), consistent with previous reports showing that, in the absence of Epe1, heterochromatin can persist by self-propagation[16,17,19].

While Epe1 removes H3K9me, it also promotes the RNAi pathway to replenish H3K9me. To understand how Epe1 plays such opposing roles, we used our repeat-induced RNAi system. First, we examined the effects of *epe1Δ* using *trans*-acting RNAi. In *epe1Δ* cells, *trans*-acting RNAi established heterochromatin more efficiently than in wild-type cells at both *ade6*$^{+}$x1 and *ade6*$^{+}$x8, suggesting that Epe1 primarily suppresses siRNA-directed heterochromatin formation (Extended Data Fig. 8). This idea was further supported by the robust accumulation of H3K9me at the *ade6*$^{+}$x1 in *epe1Δ* cells (Fig. 6a). At the *ade6*$^{+}$x8, robust ectopic heterochromatin was observed in both wild-type and *epe1Δ* cells (Fig. 6a); however, the ectopic heterochromatin in *epe1Δ* cells was not accompanied by effective production of secondary siRNAs (Fig. 6b). Thus, *trans*-acting RNAi failed to efficiently trigger *cis*-acting RNAi in the absence of Epe1, indicating that Epe1 is required to supply target transcripts for effective activation of *cis*-acting RNAi.

Because Epe1 is localized to heterochromatin via the H3K9me binding protein Swi6/HP1[18], the low levels of H3K9me on the *ade6*$^{+}$x1 in wild-type cells made it difficult to determine whether inefficient production of secondary siRNA on *ade6*$^{+}$x1 is due to defective localization of Epe1 or failure of Epe1 to promote *cis*-acting RNAi (Fig. 6a,b). To address these distinct possibilities, we removed *ade6*-hp from *epe1Δ* cells in which robust ectopic heterochromatin was already established on *ade6*$^{+}$x1, and concurrently restored *epe1*$^{+}$ (Fig. 6c). Without *epe1*$^{+}$ restoration, the ectopic heterochromatin on *ade6*$^{+}$x1 was inherited after removal of *ade6*-hp (Fig. 6c,d). However, when *epe1*$^{+}$ was restored, all progenies formed white (*ade6*-expressing) colonies, and H3K9me on *ade6*$^{+}$x1 was completely removed (Fig. 6c,d). By contrast, when Epe1 was restored in a control experiment with the *ade6*$^{+}$x8, red-pink (*ade6*-repressed) progenies were frequently observed, in which high H3K9me levels were maintained on *ade6*$^{+}$ (Fig. 6c,d). In these clones, restored Epe1 promoted recruitment of the RDRC and the production of *ade6* siRNAs (Fig. 6e). Thus, these results suggest that repeated genes are required for Epe1 to promote *cis*-acting RNAi. Furthermore, when Epe1 failed to activate *cis*-acting RNAi at the *ade6*$^{+}$x1, it removed H3K9me, indicating that success or failure of establishment of the repeat-induced RNAi causes Epe1 to have opposing effects in epigenetic inheritance of H3K9me landscape.

Taken together, our results suggest that RNAi-mediated heterochromatin formation is promoted by repeated transcription units, regardless of whether they encode mRNAs or ncRNAs. The requirement of an anti-silencing factor Epe1 for repeat-induced RNAi provides a model whereby repeated genes enable Epe1 to supply target transcripts enough to be recognized by the RNAi machinery under silent heterochromatin, possibly because the repeated genes can increase local concentration of nascent target RNA without robust transcription (Fig. 6f). We assume that the widespread TSSs present in *dg/dh* elements also function as a variant of repeated genes

(Extended Data Fig. 9 and Supplementary Discussion 2), because closely placed TSSs will produce RNA templates with the same sequences. Notably, in contrast to *dg*/*dh* elements, Epe1 OP did not cause hyper-activation of RNAi at *ade6*$^{+}$x8, but instead removed H3K9me (Extended Data Fig. 10). In addition, higher levels of Epe1 OP also reduced H3K9me at the *dg*/*dh* elements (Extended Data Fig. 1f)[18]. These results suggest that RNAi-mediated heterochromatin formation depends on the balance between replenishment and removal of H3K9me. The number of repeat copies, the amount of Epe1, and possibly, the length of the target transcripts will affect this balance. We speculate that such a balance might also be established when multiple siRNAs target cognate genes that exist in close proximity on the genome[25]. Although Epe1 orthologues have not been identified so far, anti-silencing mechanisms that direct transcription from silent heterochromatin have also been reported in higher eukaryotes[26,27]. Repetitive DNA linked to silent heterochromatin has been observed in diverse biological phenomena, including paramutation alleles and imprinted genes in higher eukaryotes[28-30]. In addition, previous studies showed that repetitive transgenes promote heterochromatin formation in a repeat-number dependent manner, analogous to our present results[2]. The link discovered in this study between repetitive DNA and RNAi via an anti-silencing factor provides a novel insight into this universal phenomenon.

Fig.1

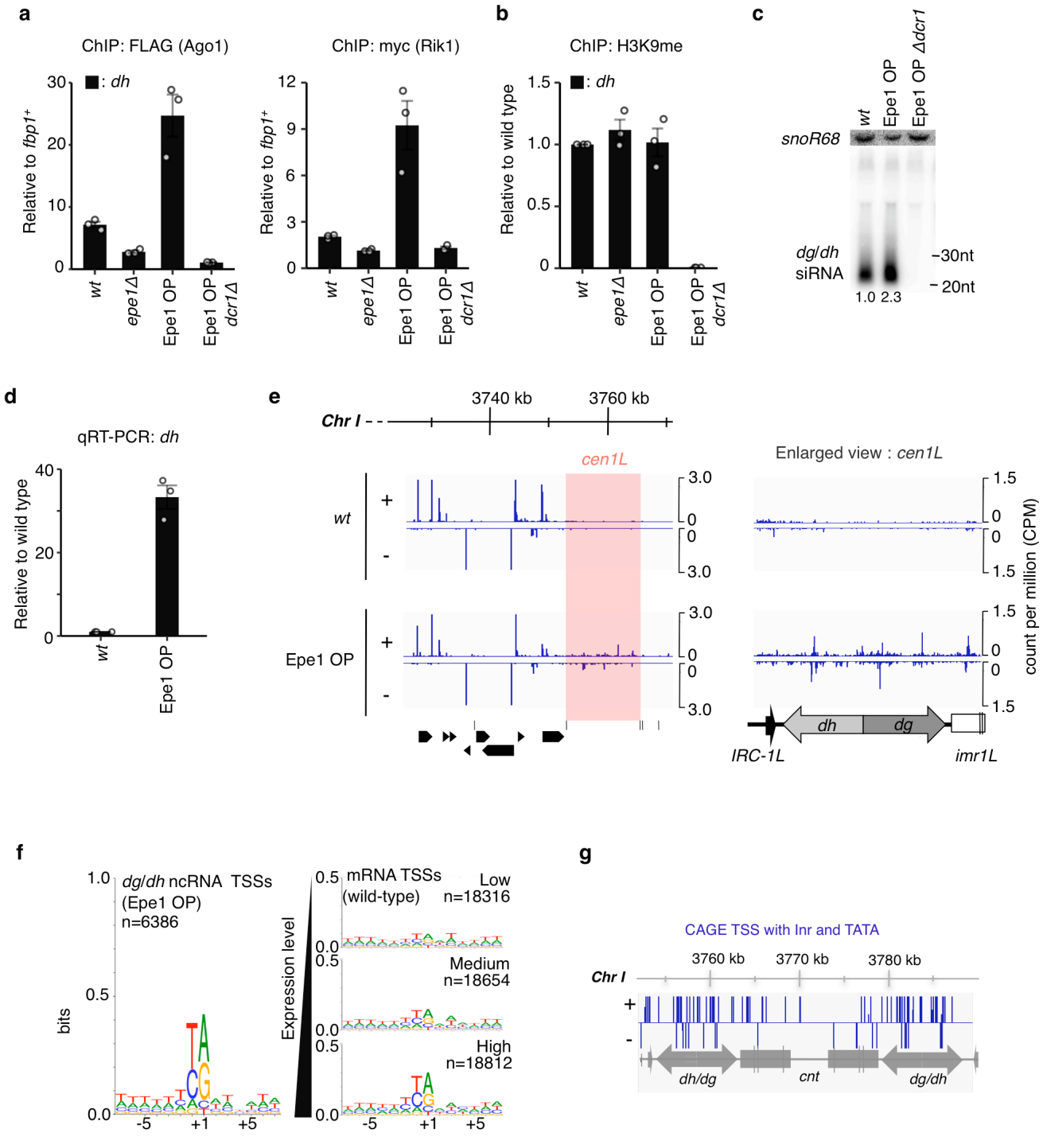


Fig. 1 | Epe1 overproduction promotes assembly of the RNAi machinery by inducing scaffold ncRNA from TSSs that are widespread across constitutive heterochromatin. a, ChIP-qPCR of Ago1, a CLRC component, Rik1, and **b**, H3K9me at pericentromeric *dh*. Error bars represent SEM.; n = 3 biological replicates. **c**, Northern blotting of *dg*/*dh* siRNAs. snoRNA58 (snoR58), loading control. Average signal intensities calculated from three independent experiments are shown below the blot. **d**, qRT-PCR of *dh* transcripts relative to the wild type (*wt*). **e**, CAGE-seq reads of *wt* or Epe1 OP in the vicinity of chromosome I left pericentromere (cen1L) are shown in a strand-specific manner (±). An enlarged view of cen1L is shown in the right hand panel (see also Extended Data Fig. 2a-c). **f**, The consensus sequence of ±7 nt flanking transcription start sites (TSSs) in *dg*/*dh* elements with Epe1 OP (>0.05 CPM) was compared with those of mRNAs in *wt*. TSSs of mRNAs were categorized into three groups according to expression strength. The number of unique TSSs in each category is indicated as "n" (see also Extended Data Fig. 2d). **g**, CAGE reads that have both Inr elements and a TATA-like A/T rich region were extracted, and mapped to the chromosome I pericentromere.

Fig.2

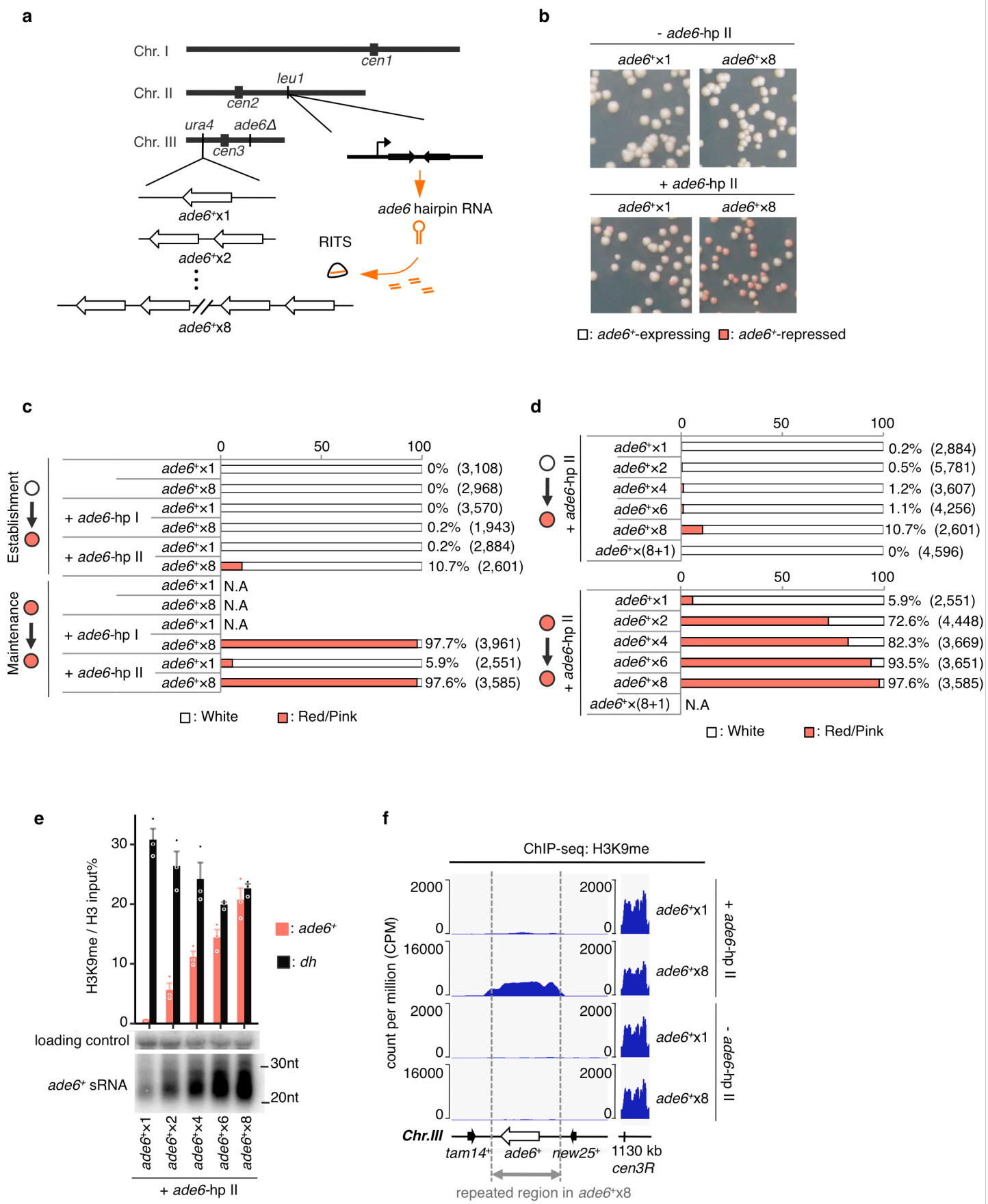


Fig. 2 | Repeated genes promote RNAi-mediated heterochromatin formation. a,Diagram of experimental scheme for artificial targeting of RNAi to repeated reporter gene *ade6*$^+$ on *S.pombe* chromosomes. *ade6* siRNAs derived from hairpin RNAs expression construct at *leu1*$^+$ locus artificially direct the RNAi pathway to tandem *ade6*$^+$ repeat at *ura4*$^+$ locus in *trans* (see also Extended Data Fig. 4 and 5). Note that endogenous *ade6*$^+$ is deleted (*ade6Δ*). **b**, Representative results of *ade6*$^+$ silencing assay. Red-pink colonies represent silencing of *ade6*$^+$. For the results with *ade6*-hpII, progenies derived from red originator of *ade6*$^+$x1 and *ade6*$^+$x8 are shown. **c**,**d**, Efficiency of establishing or maintaining *ade6*-repressed state was evaluated by red-pink colony formation from white or red originator, respectively. Parentheses, numbers of scored colonies. (**c**) *ade6*$^+$x1 and *ade6*$^+$x8 strains without/with two different hairpin RNAs (*ade6*-hp I and II). (**d**) A series of *ade6*$^+$ repeat strains with *ade6*-hp II. **e**, Upper, ChIP-qPCR of H3K9me. Error bars represent SEM; n = 3 biological replicates. Lower, Northern blot of *ade6*$^+$ siRNA using the same cells. A non-specific band was used as a loading control. To equalize a prior condition of each strain, cells after red clone is once selected were used in these experiments. **f**, ChIP-seq analysis of H3K9me with cells used in (**e**) and negative controls. To facilitate visualization, reads were mapped on *ade6*$^+$x1 construct. Therefore, results of *ade6*$^+$x8 strains are scaled by a factor of 8 relative to *ade6*$^+$x1.

Fig.3

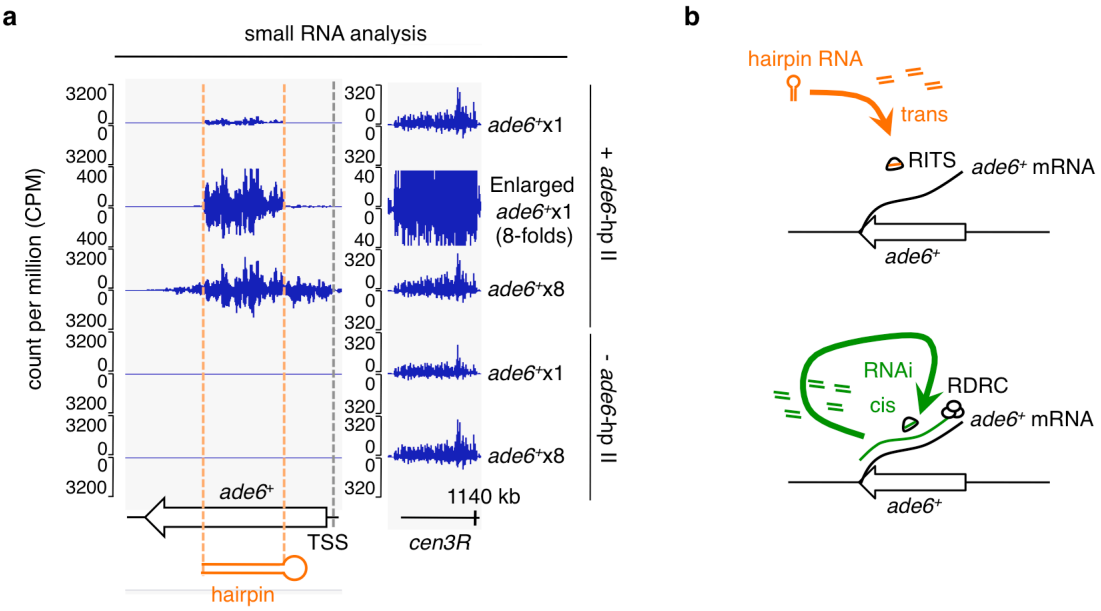


Fig. 3 | Repeated genes promote *cis*-acting RNAi more efficiently than a single gene. a, Small RNA reads mapping to *ade6*$^+$ and part of the chromosome III pericentromere are shown. Dashed orange lines mark a region targeted by *trans*-acting RNAi (*ade6*-hp II). An 8-folds enlarged view of *ade6*$^+$x1 is also shown to take into account the effect of copy number. See also Supplementary Fig. 1 and Supplementary Discussion 3. **b**, Schematic diagram of the RNAi pathway targeted to *ade6*$^+$ mRNAs by two different routes. Upper, when the RITS complex contains siRNAs derived from *ade6* hairpin RNAs, the RNAi pathway is directed to *ade6*$^+$ mRNAs in *trans* (*trans*-acting RNAi). Lower, when the RITS complex contains siRNAs derived from *ade6* dsRNAs, which are synthesized by the RDRC using *ade6*$^+$ mRNAs as templates, the RNAi pathway is directed to *ade6*$^+$ mRNAs in *cis* (*cis*-acting RNAi). When *trans*-acting RNAi activates *cis*-acting RNAi, dsRNA synthesis by the RDRC produces secondary siRNAs that are not encoded by hairpin RNAs.

Fig.4

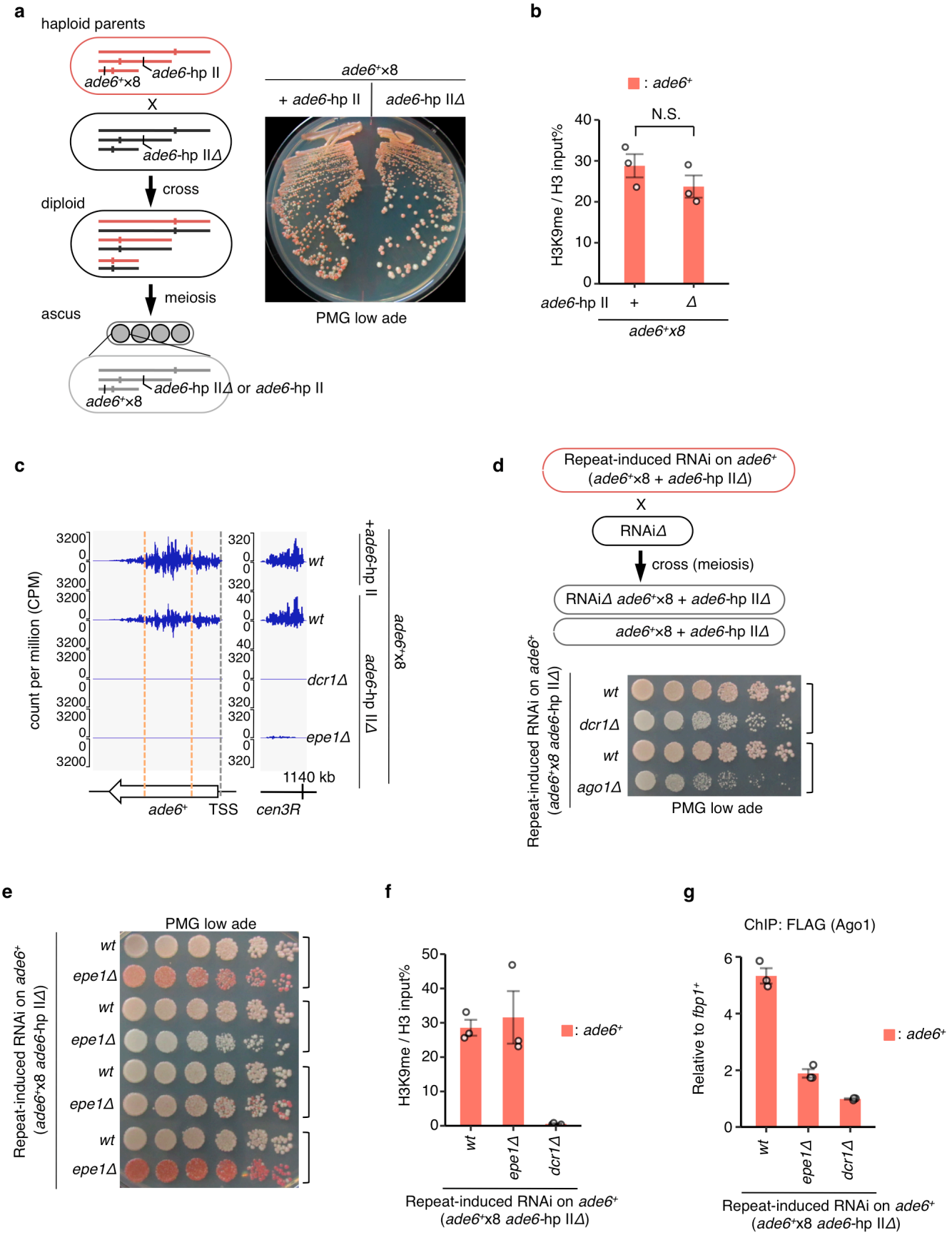


Fig. 4 | Epe1 is required for repeat-induced RNAi. a, Silencing assay of *ade6*$^{+}$x8 strains after removal of the hairpin RNA construct. Red *ade6*$^{+}$x8 (*ade6*-repressed) cells harboring *ade6*-hp II were crossed with *ade6*-hp II*Δ* cells to segregate *ade6*-hp II. Representative results of progenies that inherited cognate *ade6*$^{+}$x8 alleles with/without *ade6*-hp II are shown in the right hand panel. **b**, ChIP-qPCR of H3K9me with progenies that inherited cognate *ade6*$^{+}$x8 alleles with/without *ade6*-hp II. **c**, Small RNA-seq revealed that *cis*-acting RNAi at *ade6*$^{+}$x8 is autonomously maintained even after removal of *ade6*-hp II (Repeat-induced RNAi). Note that the result of *wt* *ade6*$^{+}$x8 with *ade6*-hp II is shared with Fig. 3. **d**,**e**, Silencing assay of *ade6*$^{+}$x8 allele combined with RNAi-defective mutants. Red *ade6*$^{+}$x8 cells, in which repeat-induced RNAi maintains heterochromatin on *ade6*$^{+}$x8, were crossed with (**d**) *dcr1Δ*, *ago1Δ*, and (**e**) *epe1Δ* cells. Wild-type cells derived from the same ascus, which inherit the cognate *ade6*$^{+}$x8 allele are also indicated with close square brackets. **f**,**g**, ChIP-qPCR of (**f**) H3K9me and (**g**) Ago1 at *ade6*$^{+}$x8 allele with indicated strains. Error bars represent SEM; n = 3 biological replicates. N.S., Not significant (P = 0.27, two-sided Student's t-test).

Fig.5

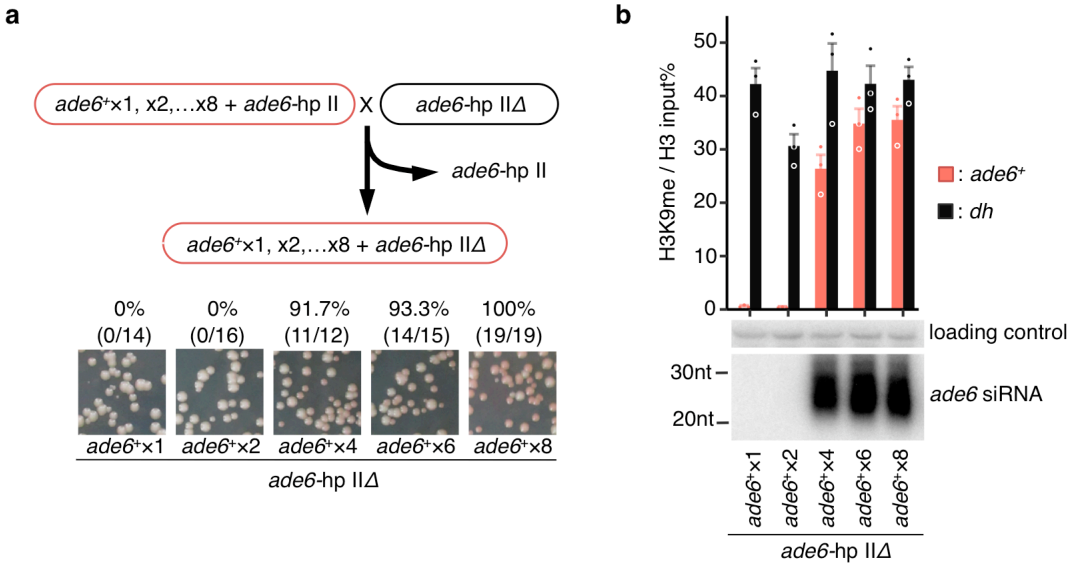


Fig. 5 | Autonomous *cis*-acting RNAi is not established with *ade6*$^+$x1 or x2 alleles. a, Red cells, in which a series of *ade6*$^+$ repeats allele were heterochromatinized by *ade6*-hp II, were crossed with *ade6*-hp II*Δ* cells to segregate the *ade6*$^+$ repeat allele from the *ade6*-hp II. The percentage of progenies that formed red-pink (*ade6*-repressed) colonies without *ade6*-hp II is indicated. These progenies were used for experiments shown in panel (b). **b**, Upper, ChIP-qPCR of H3K9me. Lower, Northern blot analysis for *ade6* siRNAs. Non-specific bands produced by the *ade6* probe were used as a loading control. Following removal of *ade6*-hp II, *ade6*$^+$x1 and x2 strains failed to maintain H3K9me at the *ade6*$^+$ gene, because the autonomous *cis*-acting siRNA was not established.

Fig.6

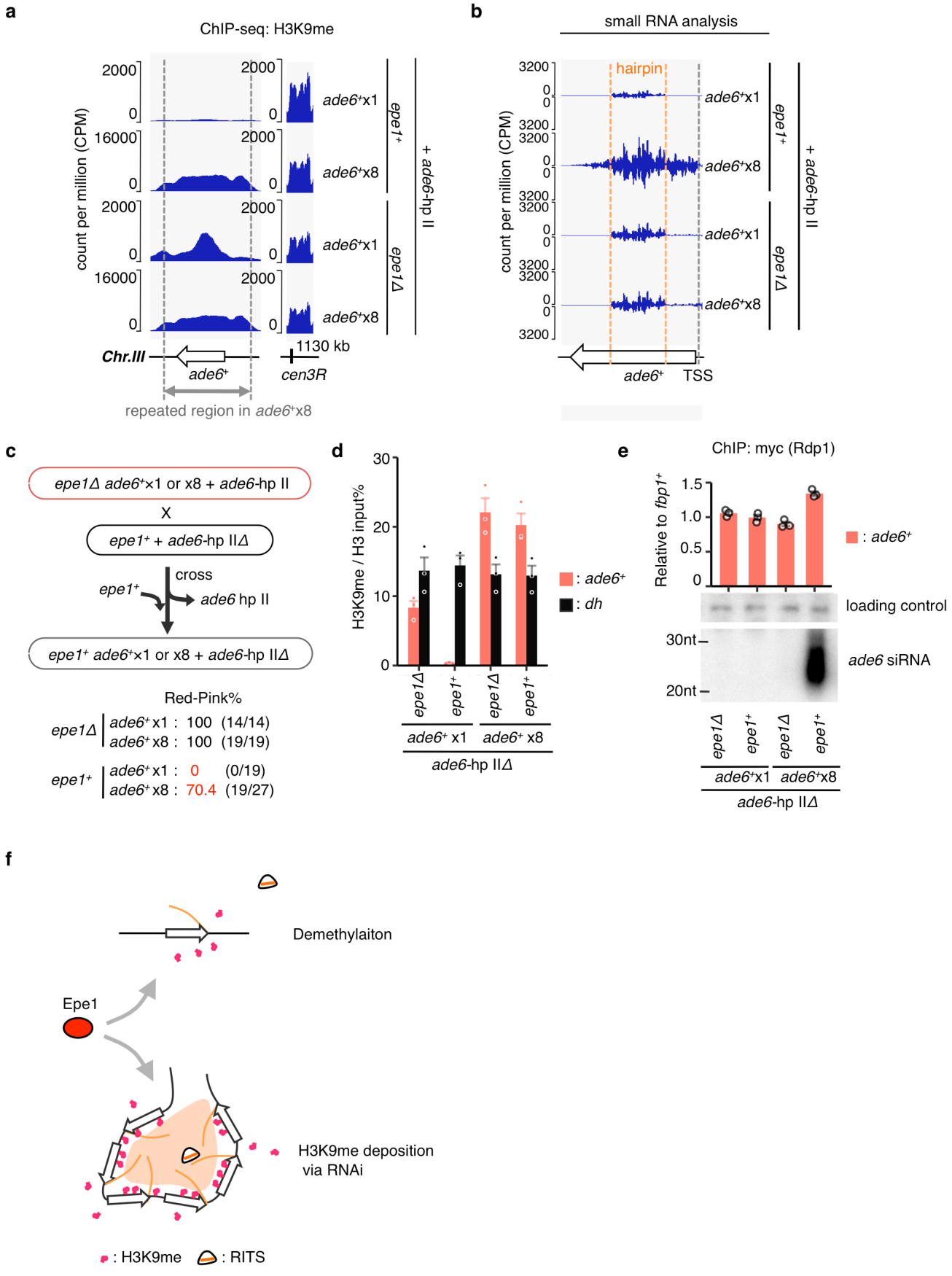
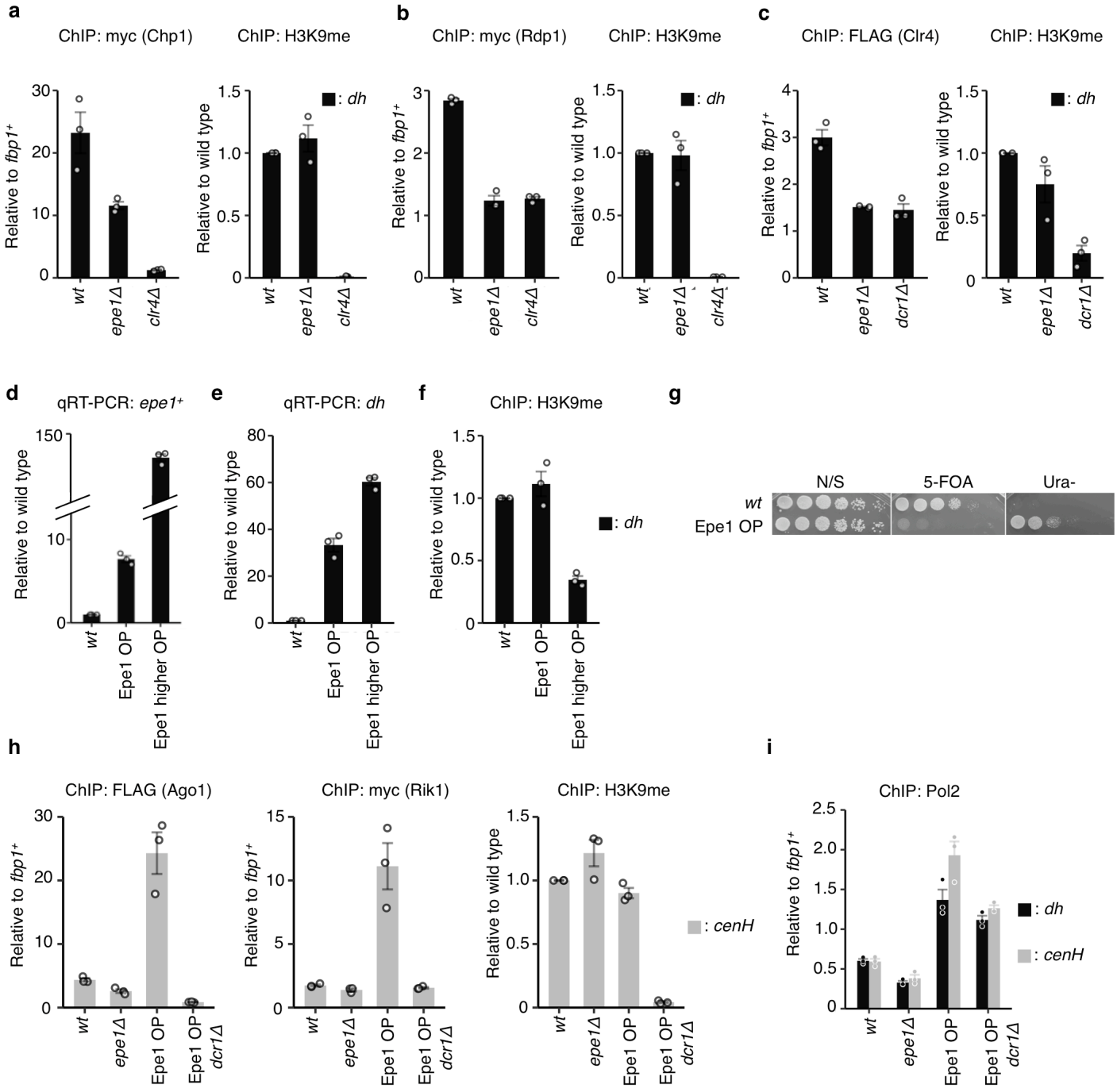


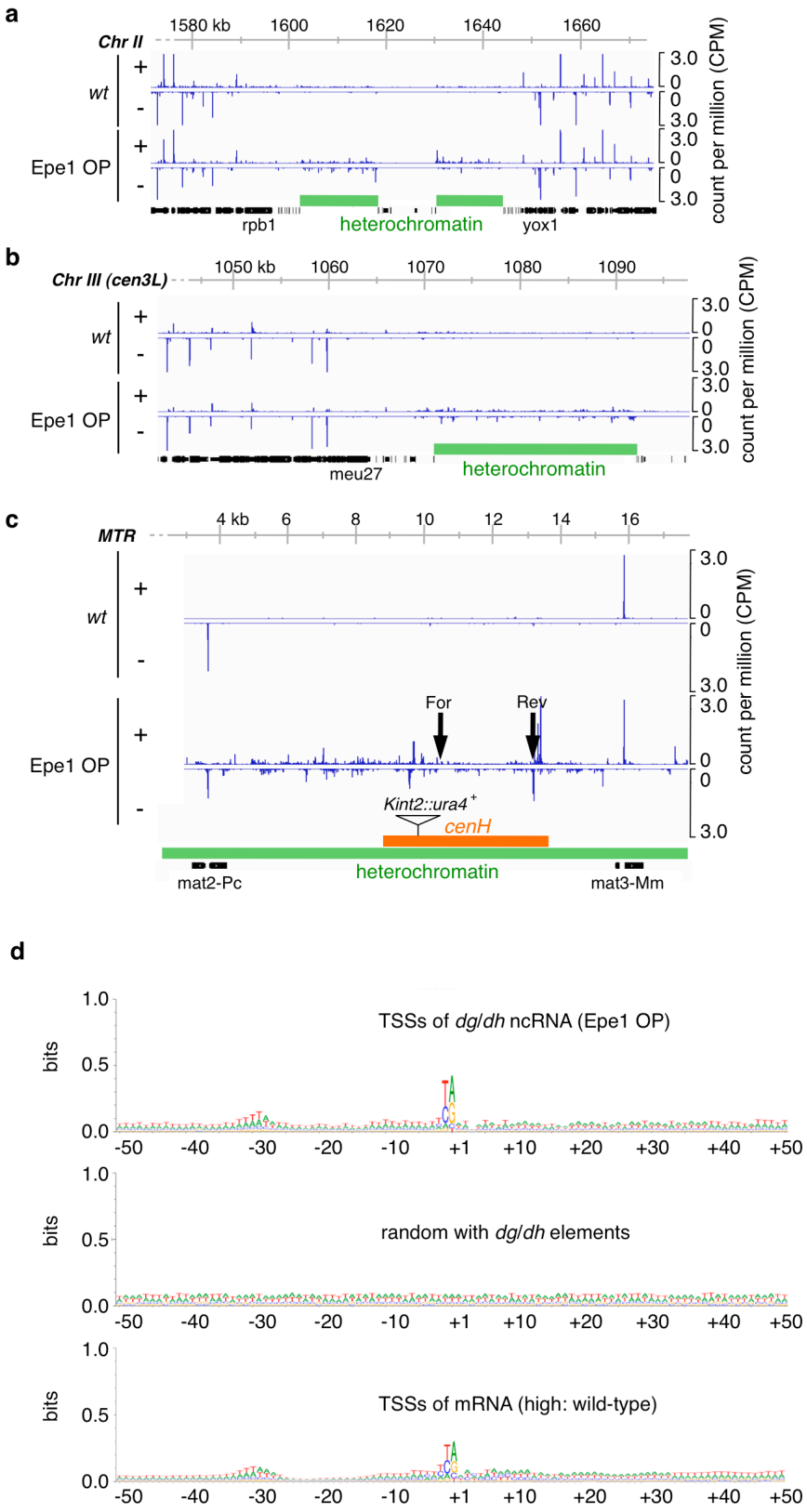
Fig. 6 | Epe1 plays opposing roles in epigenetic inheritance of H3K9me in a repeat-dependent manner. a, ChIP-seq of H3K9me and **b**, Small RNA-seq were performed with *epe1Δ* cells harboring *ade6*-hp II. For comparison, results of *epe1*$^{+}$ cells derived from Fig. 2f and Fig. 3a are also shown, respectively. **c**, Red *epe1Δ* cells harboring *ade6*$^{+}$x1 or *ade6*$^{+}$x8 alleles that were heterochromatinized by *ade6*-hp II, were crossed with *epe1*$^{+}$ *ade6*-hp II*Δ* cells to segregate *ade6*-hp II and concurrently return *epe1*$^{+}$. The percentage of clones that formed red-pink (*ade6*-repressed) colonies without *ade6*-hp II was evaluated for the indicated progenies. These cells were used for ChIP-qPCR of (**d**) H3K9me, and (**e**) an RDRC component Rdp1 (Upper), and Northern blot analysis of *ade6* siRNA (lower). Error bars represent SEM; n = 3 biological replicates. **f**, Model for repeat-induced RNAi via the anti-silencing factor, Epe1. Epe1 removes H3K9me and derepresses transcription of genes underlying heterochromatin, which results in opposing effects in a repeat-dependent manner. When a single target gene underlies silent heterochromatin, the amount of RNA templates induced by Epe1 is not sufficient to assemble the RNAi machinery, therefore resulting in removal of H3K9me. By contrast, repeated target genes enable Epe1 to supply enough RNA templates to be recognized by the RNAi machinery, resulting in replenishment of H3K9me and autonomous maintenance of heterochromatin.



Extended Data Fig. 1 | Epe1 promotes the assembly of the RNAi machinery on constitutive heterochromatin. a, ChIP-qPCR of the RITS complex component, Chp1, **b**, the RDRC component, Rdp1, **c**, the CLRC component, Clr4, at pericentromeric *dh*. Because each *epe1Δ* clone shows different levels of H3K9me[19], results of simultaneously performed ChIP-qPCR of H3K9me are shown side by side. **d**, **e**, and **f**, comparison between Epe1 OP (endogenous *epe1*$^+$ promoter is replaced by Purg1[37]) and higher levels of Epe1 OP (expression from multi-copy plasmid pREP41[32]). (**d**) qRT-PCR of *epe1*$^+$ transcripts relative to wild type. (**e**) qRT-PCR of *dh* transcripts relative to wild type. (**f**) ChIP-qPCR of H3K9me at *dh*. **g**, Silencing assays of Epe1 OP strains, in which a reporter gene, *ura4*$^+$, is inserted within heterochromatin at the mating-type locus (*Kint2::ura4*$^+$) (see also Extended Data Fig. 3a). N/S, non-selective; 5-FOA, medium toxic

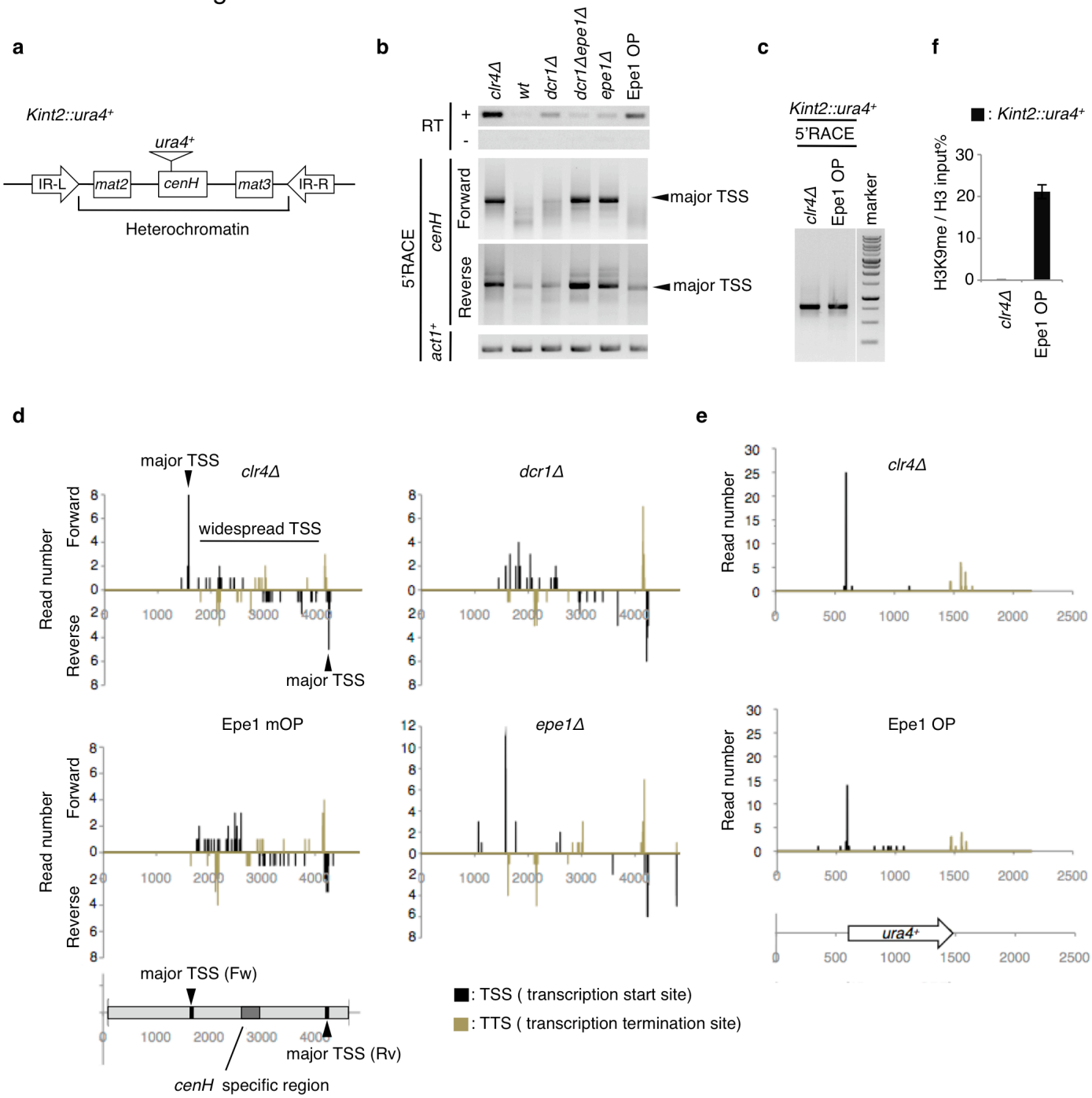
to cells expressing *ura4*$^{+}$; Ura-, medium lacking uracil. **h**, Epe1 OP promotes the assembly of RNAi machinery on *dg*/*dh* elements at the mating type locus (*cenH*). ChIP-qPCR analyses of Ago1, the CLRC component, Rik1, and H3K9me at *cenH* are shown. **i**, ChIP-qPCR of Pol2 at *dh* or *cenH* elements. Error bars represent SEM; n = 3 biological replicates.

Extended Data Fig.2



Extended Data Fig. 2 | Epe1 overproduction promotes ncRNA transcription from TSSs that are widespread across constitutive heterochromatin. a,**b**,**c**, CAGE reads of the indicated strains at (**a**) centromere II, (**b**) centromere III, and (**c**) the mating-type locus (MTR) are shown. Green horizontal bars indicate heterochromatin region. An orange horizontal bar indicates the *cenH* region analyzed in Extended Data Fig. 3. Arrowheads indicate the position of forward (For) and reverse (Rev) major transcription start sites (TSSs) identified by 5'RACE analysis (see also Extended Data Fig. 3). **d**, Consensus sequence analysis of widespread TSSs in *dg*/*dh* elements. The sequences of the ±50 nt flanking TSSs in *dg*/*dh* elements, which showed more than 0.05 CPM under Epe1 OP, were analyzed using WebLogo[34]. The sequences of ±50 nt randomly picked from *dg*/*dh* elements were also analyzed as a control. TSSs of mRNA in wild type cells were categorized into three groups according to expression strength. The sequences of ±50 nt flanking unique TSSs of mRNAs categorized as having high expression were analyzed. The horizontal axis indicates relative position with respect to TSS (+1). TSSs of *dg*/*dh* elements and euchromatic genes commonly exhibit a preference for the consensus initiator (Y/R) sequence at -1/+1 positions and an A/T rich region 25-32 nt upstream.

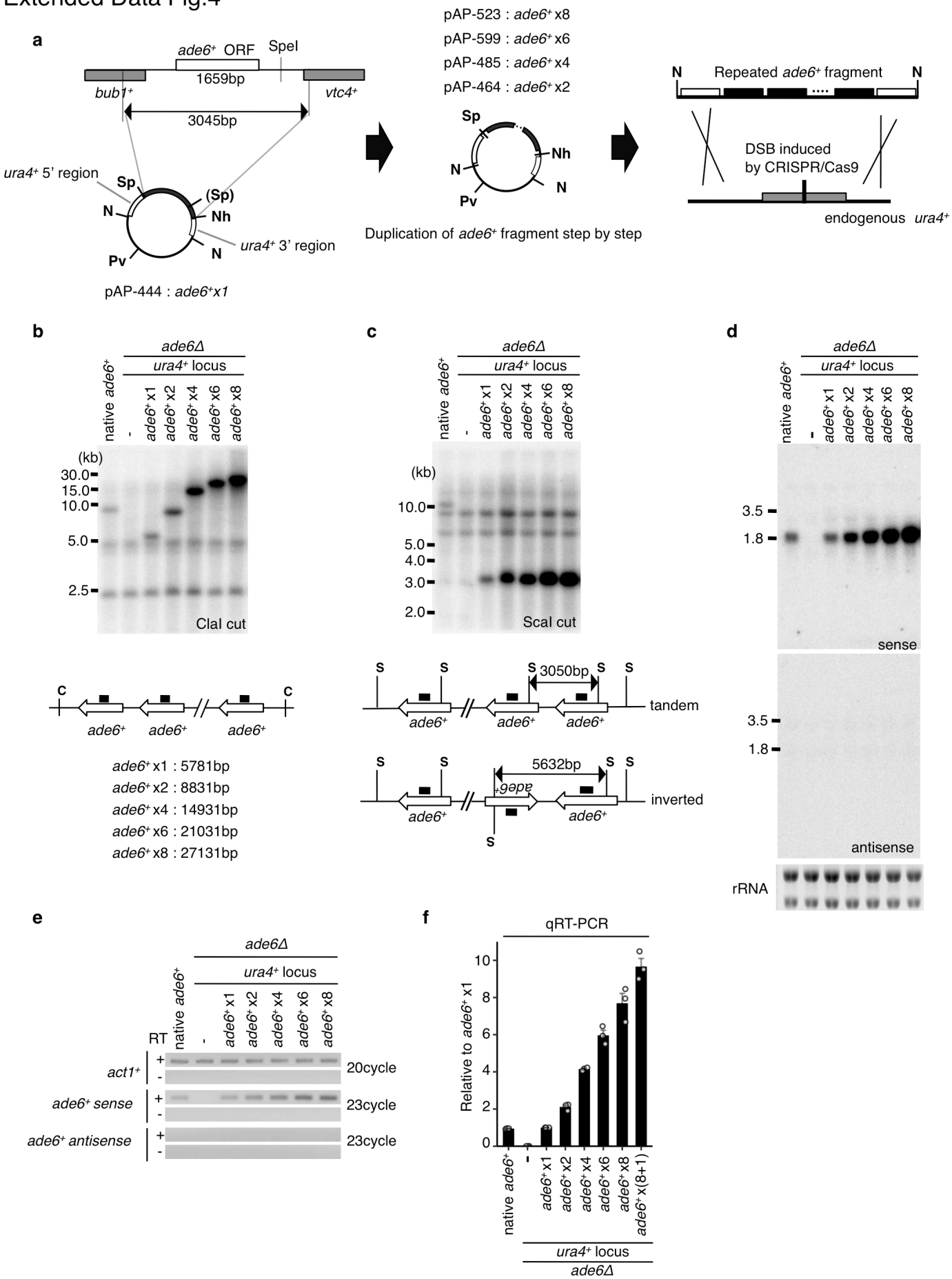
Extended Data Fig.3



Extended Data Fig. 3 | RACE analysis of *dg*/*dh* elements at the mating-type locus (*cenH*) and inserted reporter gene, *ura4*$^+$. **a**, Schematic diagram of the mating-type locus containing *cenH* and inserted *ura4*$^+$ (*Kint2::ura4*$^+$)[38]. The *cenH* has a specific insertion that enables *cenH* ncRNAs to be distinguished from other *dg*/*dh* ncRNAs. **b**, **c**, Representative image of 5'RACE for (**b**) forward and reverse *cenH* ncRNAs, *act1*$^+$ mRNAs, and (**c**) *Kint2::ura4*$^+$. RT-PCR of *cenH* ncRNAs using oligodT primers is also shown at the top of (**b**). **d**,**e**, TSSs Identified by 5'RACE were mapped to genomic sequences of (**d**) *cenH* or (**e**) *ura4*$^+$. Note that TSSs could not be identified with wild-type cells, because their RACE products were poorly cloned. There are two type of TSS for *cenH* ncRNAs: major TSSs and widespread TSSs (d, *clr4Δ*). Notably, while *epe1Δ* inactivated widespread TSSs, Epe1 OP caused their hyper-activation, resulting in smears on the gel (B). By contrast, at *Kint2::ura4*$^+$, Epe1 OP did not cause smearing, but activated TSSs corresponding to those observed in the absence of heterochromatin (*clr4Δ*). Note that Epe1 OP

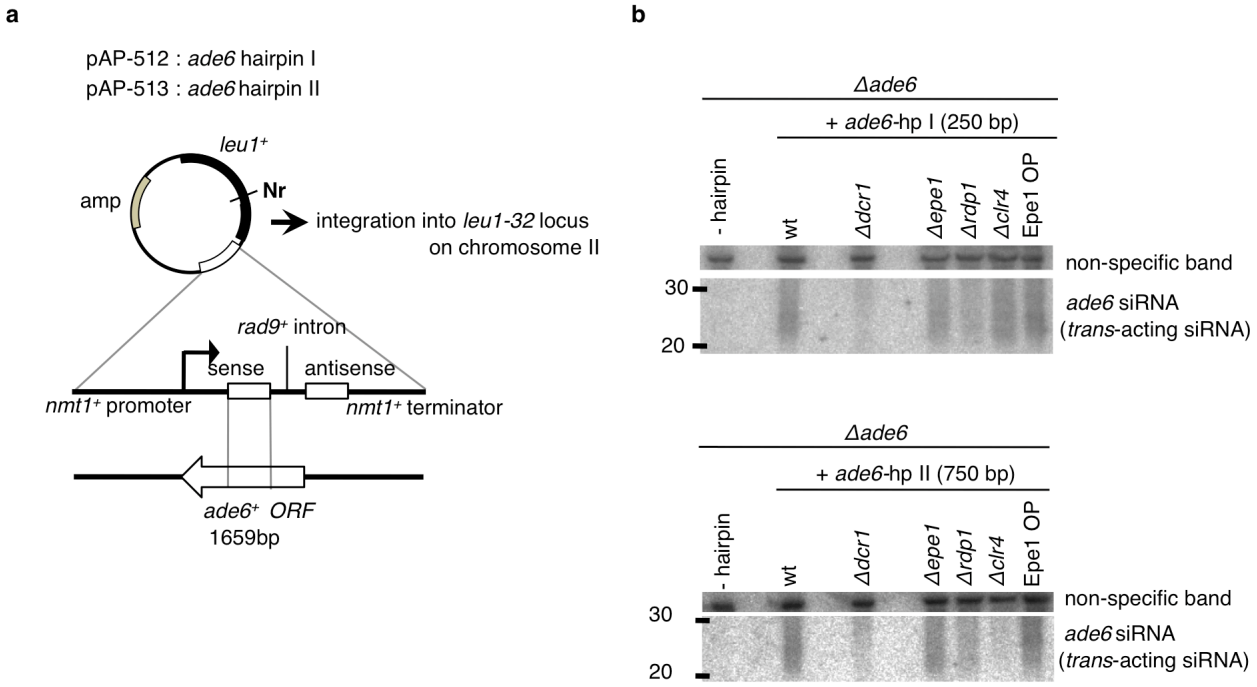
induces the expression of *Kint2::ura4*$^{+}$ in the presence of heterochromatin (**f**). Transcription termination sites (TTSs) identified by 3'RACE were also plotted in (**d**) (**e**). See also Supplementary Discussion 1.

Extended Data Fig.4



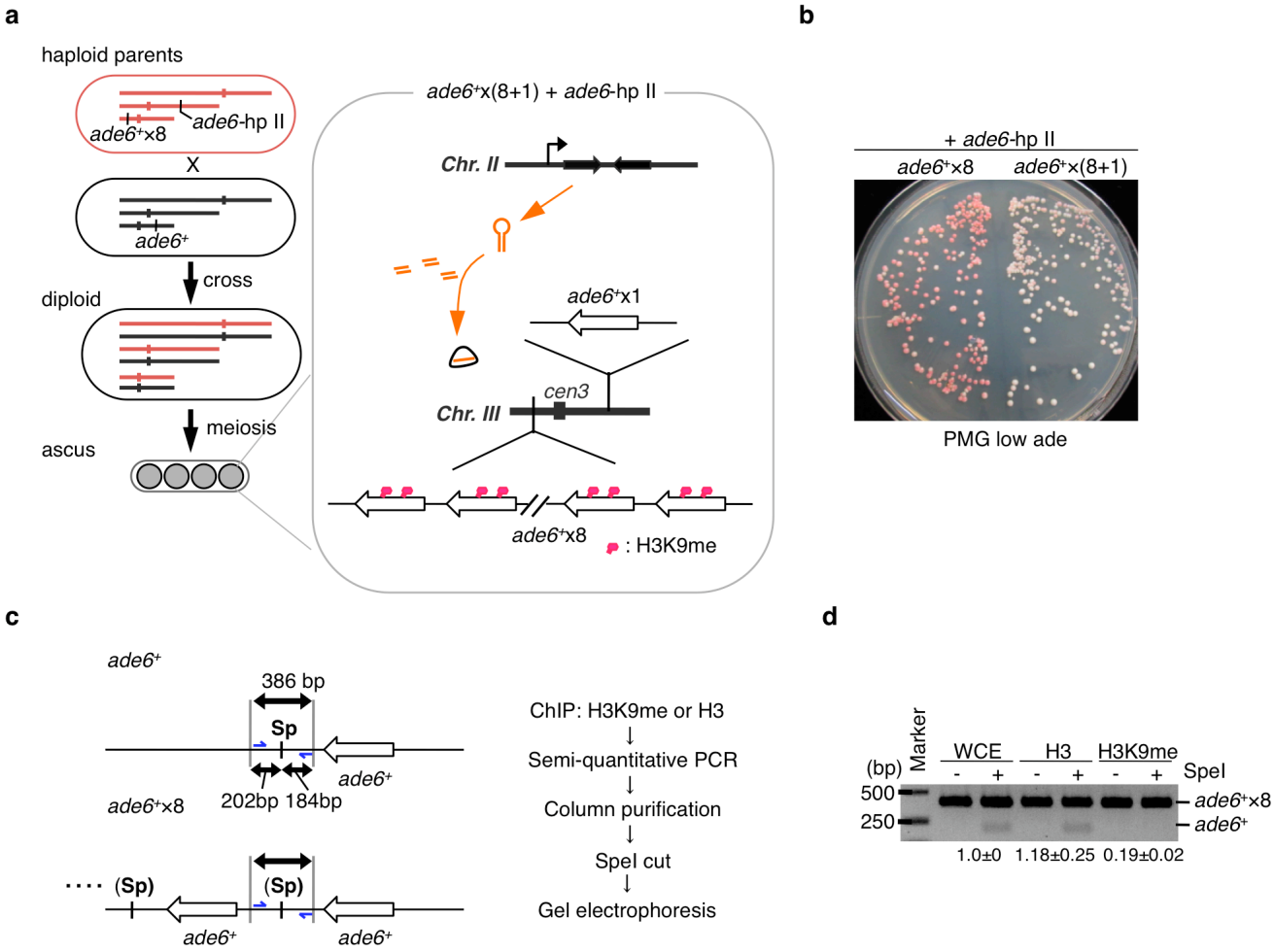
Extended Data Fig. 4 | Generation of *ade6*$^+$ repeat strains. a, Schematic diagram illustrating the construction of *ade6*$^+$ repeat strains. The *ade6*$^+$ fragment was cloned between *ura4*$^+$ homologous sequences on a plasmid, then duplicated stepwise (see also materials and methods). A CRISPR/Cas9-dependent double strand break (DSB) at the endogenous *ura4*$^+$ locus was employed to promote integration of the *ade6*$^+$ repeat fragments by homologous recombination. The restriction sites used for construction and recovery of *ade6*$^+$ repeat fragments are shown: N, NotI; Sp, SpeI; Nh, NheI; Pv, PvuI. Parentheses indicate the restriction site destructed by point mutation. **b**,**c**, Southern blot analyses of strains harboring *ade6*$^+$ repeats at the *ura4*$^+$ locus using an *ade6*$^+$ probe (black rectangle). Note that endogenous *ade6*$^+$ was deleted in these strains. Genomic DNA was digested with (b) ClaI or (c) ScaI, to evaluate the length of the *ade6*$^+$ repeat and its arrangement, respectively. The diagrams below the blots illustrate the scheme for each Southern blot. **d**, Northern blot analysis of strains harboring *ade6*$^+$ repeats at the *ura4*$^+$ locus using sense and antisense probes for *ade6*$^+$. Wild type and *ade6Δ* strain were used as a positive and negative control, respectively. rRNA stained with ethidium bromide is shown as a loading control. **e**, RT-PCR analysis of *ade6*$^+$ transcripts using strand specific primers for reverse transcription in the presence (+) or absence (-) of reverse transcriptase (RT). No antisense transcripts specific for the *ade6*$^+$ repeat are detected even if the number of PCR cycles is increased (data not shown). **f**, qRT-PCR for *ade6*$^+$ mRNA using oligodT primers for reverse transcription. Error bars represent SEM; n = 3 biological replicates.

Extended Data Fig.5

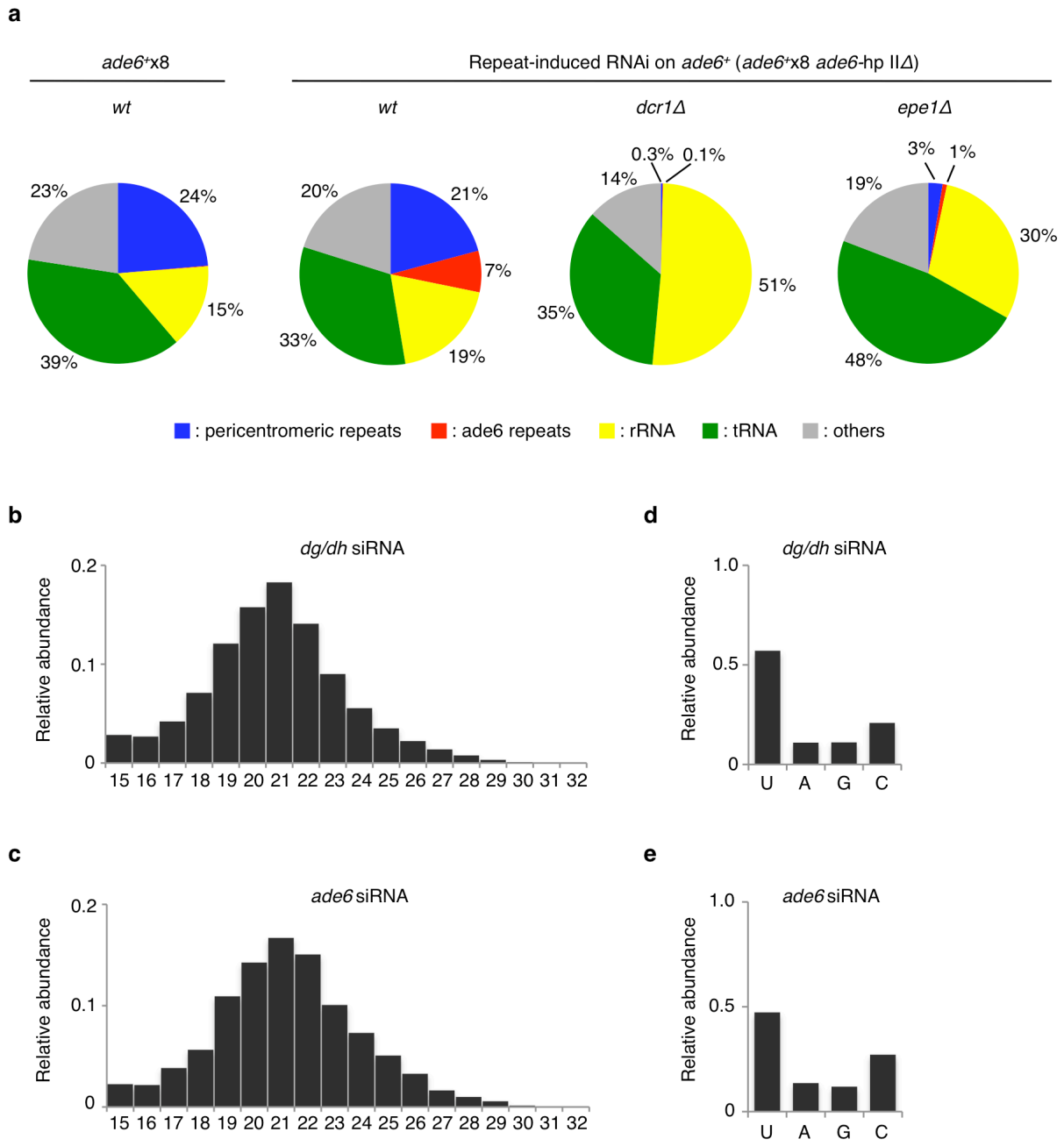


Extended Data Fig. 5 | Design of *ade6* hairpin RNA constructs that drive *trans*-acting RNAi. **a**, Schematic diagram of *ade6* hairpin RNA constructs. The *ade6* hairpin I (*ade6*-hp I) or *ade6* hairpin II (*ade6*-hp II) construct includes a 250bp or 750bp fragment corresponding to 621-871 nt or 386-1132 nt of the *ade6*$^+$ ORF, respectively. A spacer sequence derived from the first intron of *rad9*$^+$ separates the inverted *ade6*$^+$ fragment to form the hairpin construct. The inducible *nmt1*$^+$ promoter drives the expression of hairpin RNAs. The constructed plasmids are integrated into the *leu1-32* locus on chromosome II of host cells. The restriction site used for linearization is shown: Nr, NruI. **b**, Northern blot analysis using *ade6*$^+$ probe to detect *trans*-acting *ade6* siRNAs derived from *ade6*-hp I or II. A non-specific band was used as a loading control. Because these strains do not have both endogenous *ade6*$^+$ and the *ade6*$^+$ repeat allele at the *ura4*$^+$ locus, blots represent siRNAs derived only from hairpin RNA. Note that loss of Epe1 or Epe1 OP does not affect siRNA production from hairpin RNAs. *dcr1Δ* cells were used as a negative control.

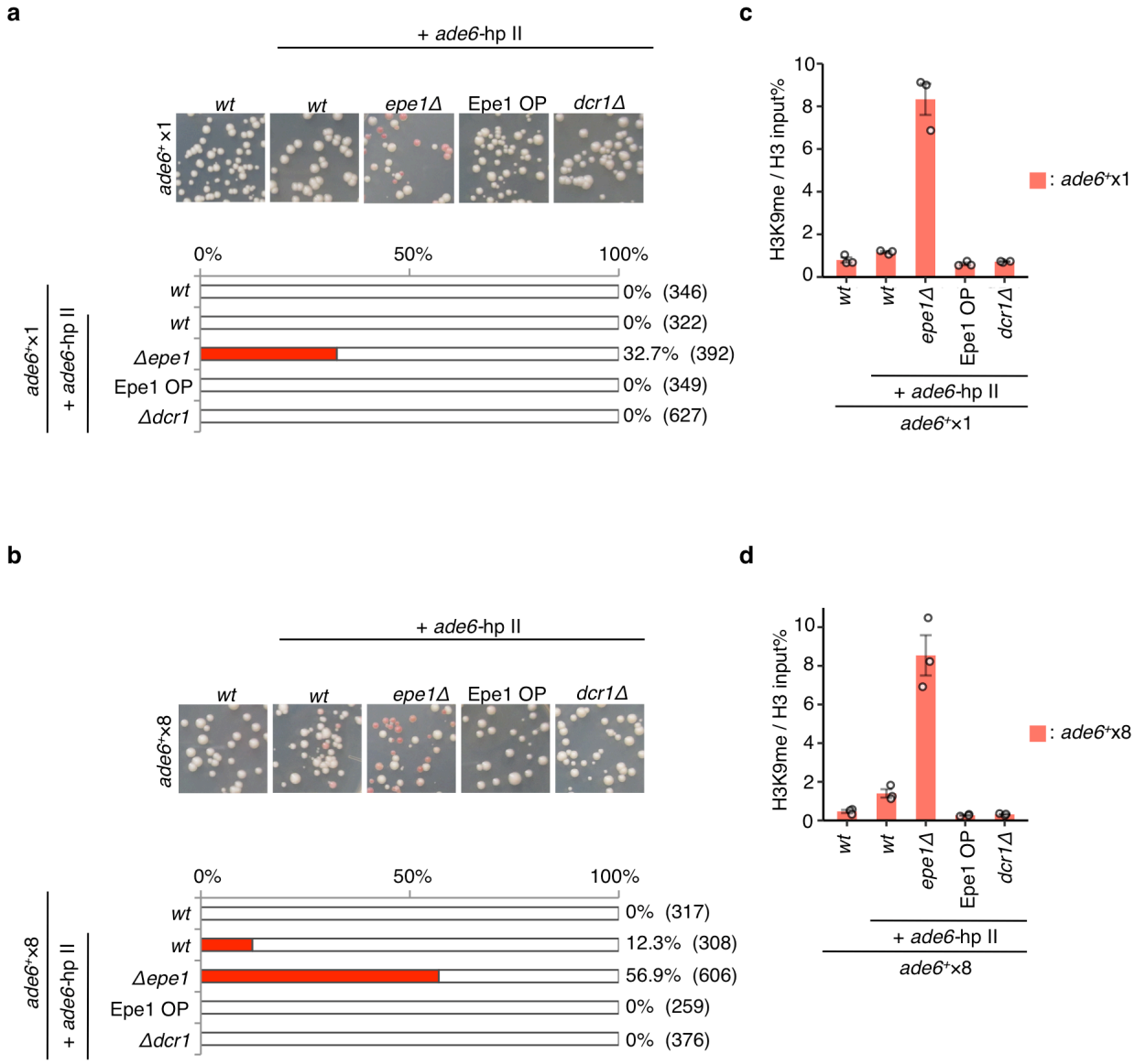
Extended Data Fig.6



Extended Data Fig. 6 | Repeated genes do not promote RNAi-mediated heterochromatin formation by their gene dosage. a, Schematic diagram of *ade6*$^+$x(8+1) strains, in which the endogenous *ade6*$^+$ gene was combined with *ade6*$^+$x8 allele already heterochromatinized by *trans*-acting RNAi. **b**, Silencing assay of *ade6*$^+$x8 and *ade6*$^+$x(8+1) strains. Phenotypes after red (*ade6*-repressed) clones were sequentially selected twice are shown. Compared with the *ade6*$^+$x8 strain, the red-pink phenotype of the *ade6*$^+$x(8+1) strain was hardly maintained. **c**, Diagram of the experimental scheme for semi-quantitative ChIP-PCR of H3K9me to distinguish an isolated *ade6*$^+$ from the *ade6*$^+$x8 allele. The SpeI site situated downstream of the *ade6*$^+$ ORF was removed by point mutation during the *ade6*$^+$x8 construction (see also Extended Data Fig. 4). Therefore, semi-quantitative ChIP-PCR with primers that amplify this restriction site region (blue arrows), followed by SpeI digestion, enabled immunoprecipitated DNAs of an isolated *ade6*$^+$ gene to be distinguished from those of the *ade6*$^+$x8 allele. **d**, Semi-quantitative ChIP-PCR using the indicated antibodies was performed and products were separated by gel electrophoresis with (+)/without (-) SpeI digestion. After electrophoresis, gels were stained with ethidium bromide and the density of each band was measured using imageJ software. The intensity of the band derived from an isolated *ade6*$^+$ relative to that from the *ade6*$^+$x8 allele, and normalized to WCE intensity, is indicated. Representative image was shown of three independent experiments.

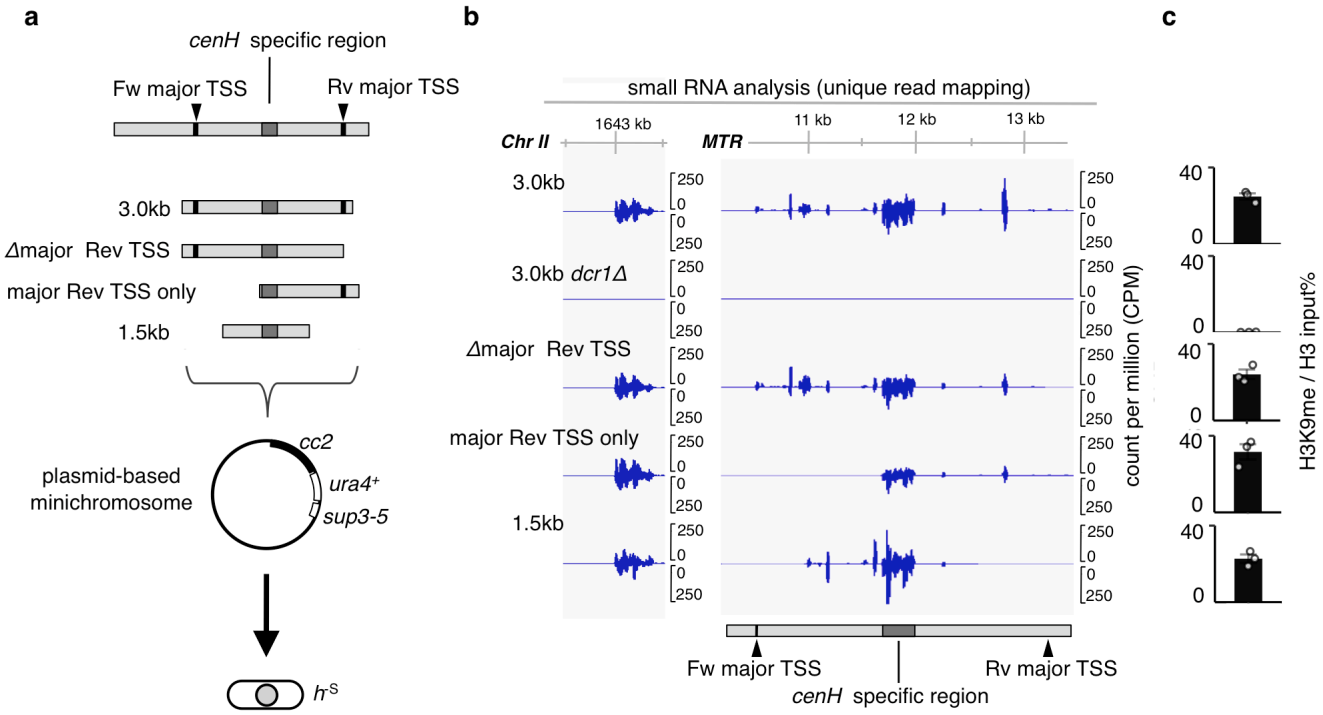


Extended Data Fig. 7 | Repeat-induced RNAi on *ade6*$^+$x8 produces siRNAs that exhibit the same properties as native siRNAs from *dg*/*dh* elements. a, High-throughput sequencing was used to analyze small RNAs in the *ade6*$^+$x8 strain either before establishment of repeat-induced RNAi, or after its establishment in combination with *dcr1Δ* or *epe1Δ*. Note that these strains do not have hairpin RNA constructs expressing *trans*-acting siRNAs. Pie charts illustrate percentages of each small RNA category relative to total small RNA reads. **b**,**c**,**d** and **e**, Small RNA reads of *ade6*$^+$x8 strain after repeat-induced RNAi was established were used for analysis of size distribution, and 5' nucleotide preference. (**b**) (**c**) Size distributions of small RNAs derived from (**b**) *dg*/*dh* elements or (**c**) *ade6*$^+$x8 are shown. Horizontal axes indicate lengths of small RNAs, and vertical axes represent relative abundance. (**d**) (**e**) Relative abundance of the first nucleotide of small RNAs derived from (**d**) *dg*/*dh* elements or (**e**) *ade6*$^+$x8.



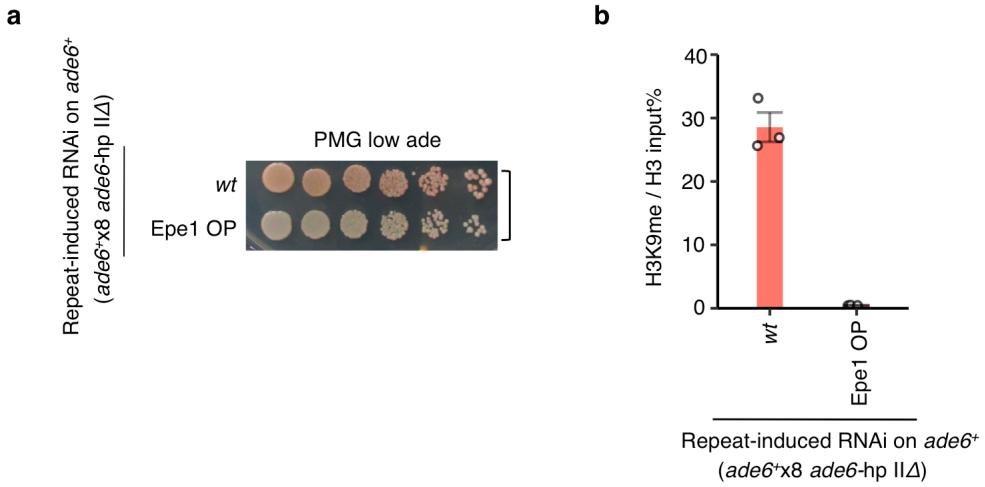
Extended Data Fig. 8 | Epe1 primarily suppresses siRNA-directed heterochromatin formation. a,**b**, Silencing assays of (**a**) *ade6*$^{+}$x1 and (**b**) *ade6*$^{+}$x8 strains harboring *ade6*-hp II. Loss of *epe1*$^{+}$ causes self-propagation of heterochromatin, which is accompanied by the appearance of the red-white variegated phenotype (see also Fig. 4e). Because such effects can perturb the evaluation of heterochromatin maintenance and establishment by the silencing assay, cells were used immediately after induction of hairpin RNA expression. Percentages of red-pink (*ade6*-repressed) colonies observed are shown in the bar graph. The numbers of colonies scored are noted in parentheses. **c**,**d**, ChIP-qPCR of H3K9me with the cells from (**a**) and (**b**). Error bars represent SEM; n = 3 biological replicates. Note that *epe1Δ* or Epe1 OP does not affect the amount of *trans*-acting siRNAs (Extended Data Fig. 5).

Extended Data Fig.9



Extended Data Fig. 9 | A *cenH* fragment that contains only widespread TSSs can establish heterochromatin. a, Diagram of the experimental scheme for truncation analysis of the *cenH* fragment. The positions of forward and reverse major TSSs identified by 5'RACE (see also Extended Data Fig. 3) are indicated. The position of the *cenH*-specific region, which is utilized to evaluate siRNA production and H3K9me accumulation, is also marked. To determine which TSS elements are required for the RNAi, a series of truncated *cenH* fragments were cloned into plasmid-based minichromosomes[36], and transformed into the h^{-S} strain, in which the native *cenH* region at the mating type locus is completely lost. This enabled evaluation of siRNA production and H3K9me accumulation at truncated *cenH* fragment on minichromosomes. **b**, Small RNA-seq mapping to part of the chromosome II pericentromere and *cenH* fragment is shown. To evaluate production of siRNAs on truncated *cenH* fragments easily, only uniquely mapped reads are extracted. Such extraction omits reads that are indistinguishable from those derived from redundant *dg*/*dh* elements at the pericentromere. **c**, ChIP-qPCR of H3K9me using *cenH* specific region primers. Error bars represent SEM; n = 3 biological replicates. These results indicate that a 1.5kb fragment containing only widespread TSSs can establish heterochromatin. See also Supplementary Discussion 2.

Extended Data Fig.10



Extended Data Fig. 10 | Epe1 OP removes H3K9me at *ade6*$^+$x8, where repeat-induced RNAi maintains heterochromatin. a,Silencing assay of the heterochromatinized *ade6*$^+$x8 allele combined with Epe1 OP. Red *ade6*$^+$x8 cells in which repeat-induced RNAi maintains heterochromatin on *ade6*$^+$, were crossed with Epe1 OP cells. Wild-type cells derived from the same ascus, which inherit the cognate *ade6*$^+$x8 allele, are also shown as controls and indicated with close square brackets. When the heterochromatinized *ade6*$^+$x8 allele was combined with Epe1 OP, cells formed only white (*ade6*-expressing) colonies. **b**, ChIP-qPCR of H3K9me shows H3K9me was completely removed from the *ade6*$^+$x8 allele in the Epe1 OP strain. This experiment was performed alongside those in Fig. 4f; therefore, the same data for wild type cells (*wt*) is shown. Error bars represent SEM; n = 3 biological replicates.

Acknowledgments: We thank R. Allshire for providing strains and plasmids; H. Kato for technical advice; T. Urano for providing antibodies; DNAFORM for CAGE-seq; H. Masumoto, M. Siomi, S. Yamanaka, J. Nakayama for critical reading; Our laboratory members for helpful discussions; T. Matsumoto and his lab members for mentoring of TA; A. Kanji, Keiko, and Yoko for their support of TA.

Funding:

JSPS Grant-in-Aid for Scientific Research (A) 10159209 (YM)

JSPS Grant-in-Aid for Scientific Research (A) 12206045 (YM)

MEXT Grant-in-Aid for Scientific Research on Priority Areas 21247001 (YM)

MEXT Grant-in-Aid for Transformative Research Areas (A) JP20H05913 (SI)

JSPS (DC2) (TA)

Author contributions:

Conceptualization: TA, YM

Investigation: TA, SI, TK, HA

Supervision: YM

Writing – original draft: TA, YM

Writing – review & editing: TA, YM

Competing interests: Authors declare that they have no competing interests.

Data availability: CAGE-seq, ChIP-seq, and smRNA-seq data have been submitted to the DDBJ (https://www.ddbj.nig.ac.jp) under accession no. DRA006868, DRA013983, DRA013984, DRA014003, DRA014324.

Methods

Schizosaccharomyces pombe strains and genetic manipulations

All strains used in this study are described in Supplementary Table 1. Standard media and methods were used for genetic manipulations. A PCR-based method was used for deletion or epitope tagging of target genes. All integrations were confirmed by PCR and Southern blot, if required. Silencing assays for *ade6*$^{+}$ gene were performed with PMG-based synthetic media containing limited concentration of adenine (10 mg/L) as the indicator plate. To assess *ura4*$^{+}$ gene silencing, 4.5-fold serial dilutions of log-phase cells were spotted onto each medium based

on Edinburgh minimal medium (EMM). The EMM 5-FOA plate contained 0.2% 5-fluoroorotic acid (FOA).

Generation of *ade6*$^+$ repeat strains

Primers used for construction of *ade6*$^+$ repeat strains are listed in Supplementary Table 2. To generate repeated *ade6*$^+$ DNA fragments targeted to the endogenous *ura4*$^+$ locus, the *ade6*$^+$ gene fragment was duplicated stepwise on the plasmid, in which amplified fragments are flanked by the 5' and 3' end region of *ura4*$^+$. We first created the backbone plasmid pAP-339 as follows. PCR products of both upstream and downstream region of *ura4*$^+$ locus, and of the neomycine resistance gene (neo)-hCVM promoter region derived from pTL2M5, were ligated by In-Fusion cloning (Takara, cat#Z9648N or Clontech, cat#639648) along with BbsI-Bst1107 fragments of pTL2M5 (FYP2047, provided by the National Bio-Resource Project (NBRP), Japan) to provide the origin of replication in *E.coli* and the ampicillin resistance gene. To remove the neomycine resistance gene (AG-714, 715) and unnecessary sequences (AG-720, 721), consecutive rounds of inverse PCR (iPCR) with divergent primers were performed using KOD Fx neo (TOYOBO, cat# KFX-201). Next, the hCMV promoter of pAP-339 was replaced with the *ade6*$^+$ gene fragment as follows: pAP-339 was digested by SpeI and NheI, and ligated with the *ade6*$^+$ PCR fragment (chromosome III, 1315461-1318504) using In-Fusion cloning. The unwanted SpeI site within the *ade6*$^+$ gene fragment was disrupted by site-directed mutagenesis using iPCR (AG-751,752). After each step of iPCR or subcloning, the plasmids were sequenced. The resulting *ade6*$^+$x1 plasmid, pAP-444, has a PvuI site located in the middle of the ampicillin resistance gene, an SpeI site at the 5' end of the *ade6*$^+$ fragment, and an NheI site at the 3' end of the *ade6*$^+$ fragment (Extended Data Fig. 4a, left panel). Thus, when PvuI-*ade6*$^+$-SpeI and PvuI-*ade6*$^+$-NheI fragments derived from pAP-444 were ligated, the resultant plasmid had duplicate *ade6*$^+$ fragments, and the SpeI-NheI junction site became undigestible. Repetition of this duplication process resulted in *ade6*$^+$x2, x4, and x8 plasmids (pAP-464, pAP-485, pAP-523) (Extended Data Fig. 4a, middle panel). Furthermore, ligation of PvuI-*ade6*$^+$x2-SpeI (pAP-464) and PvuI-*ade6*$^+$x4-NheI (pAP-485) fragments resulted in the *ade6*$^+$x6 plasmid (pAP-599). These plasmids have NotI sites at either end of the *ura4*$^+$ fragment that allow retrieval of the *ade6*$^+$ repeat fragment targeting the endogenous *ura4*$^+$ locus. To promote homologous recombination, the CRISPR/Cas9 system was used to induced double strand breaks (DSB) at the endogenous *ura4*$^+$ locus (Extended Data Fig. 4a, right panel). Cloning of a gRNA sequence targeting the *ura4*$^+$ ORF into the gRNA-Cas9 single plasmid, pAH237[31], was performed as previously described, resulting in pAP-562. Co-transformation of the *ade6*$^+$ repeat fragment and gRNA-Cas9 plasmids targeting *ura4*$^+$ was carried out using the classic LiOAc method; $1x10^8$ cells were transformed with 1 µg of AP-562 and 500-1000 ng of *ade6*$^+$ repeat fragments. To obtain clones in which the endogenous *ura4*$^+$ gene is replaced by the *ade6*$^+$ repeat, the transformants selected by EMM-selective plates (-leucine) were subsequently replicated to 5-FOA plates. The cells were again streaked on 5-FOA plates to obtain single colonies and colony PCR was performed to select candidates. To isolate cells that had lost the gRNA-Cas9 plasmid, single colonies of each candidate were repeatedly streaked on YES plates until cells exhibited leucine auxotrophy on EMM (-leu) plates. Finally, genomic DNA was extracted from each candidate, and subjected to qPCR and Southern blot analysis to select clones in which the endogenous *ura4*$^+$ gene had been replaced by the *ade6*$^+$ repeat as planned (Extended Data Fig. 4b, c). Correlation between the copy number of the *ade6*$^+$ gene and the expression level of *ade6*$^+$ mRNA was confirmed by Northern blot and qRT-PCR (Extended Data Fig. 4d-f).

Generation of hairpin RNA constructs

Primers used for generation of *ade6* hairpin constructs are listed in Supplementary Table 2. The structure of the plasmid used to generate cells expressing *ade6* hairpin RNA is shown in Extended Data Fig. 5a. To generate an *ade6* hairpin construct whose expression is driven by the thiamine-repressible *nmt1*$^+$ promoter, the pREP1 plasmid[32] was used as a backbone. The *ade6* hairpin construct was composed of three fragments. The left and right arms were PCR products corresponding to part of the *ade6*$^+$ ORF (*ade6* hairpin I, 250bp: *ade6* hairpin II, 750bp). For the third fragment, PCR products of the first intron of the *S.pombe rad9*$^+$ gene were used as a spacer. Using In-Fusion cloning (Takara, cat#Z9648N or Clontech, cat#639648), these three fragments were assembled and cloned in the pREP1 plasmid, which had first been linearized by SmaI and NdeI. To integrate this hairpin RNA construct at the endogenous *leu1*$^+$ locus, the *leu1*$^+$ targeting plasmid, pAP-509, was created, in which PCR products encoding the *leu1*$^+$ gene were subcloned into the SpeI - NotI site of pBluescript KS II (+). The pREP1-*ade6* hairpin construct was digested with PstI and EcoRI to retrieve fragments encoding the *nmt1*$^+$ promoter -*ade6* hairpin - *nmt1*$^+$ terminator fragment. These fragments were inserted into PstI-EcoRI digested pAP-509, creating the plasmids pAP-512 (*ade6*-hairpin I) and pAP-513 (*ade6*-hairpin II) (Extended Data Fig. 5a). To integrate these plasmids at the endogenous *leu1*$^+$ locus, the NruI site within the *leu1*$^+$ gene fragment was used for linearization, and the resultant fragments were transformed into *leu1-32* strains. Transformants were selected for leucine prototrophy, and integration at *leu1-32* was verified by colony PCR, and subsequent Southern blot analysis. Strains carrying the *ade6* hairpin construct were created in the presence of thiamine to repress hairpin RNA expression during manipulation. To induce expression of hairpin RNA, cells were grown in liquid media without thiamine for 16 h, and spread onto PMG plates without thiamine. Resultant colonies were stocked for use in experiments. Generation of *ade6* small RNAs in the various mutants was verified by Northern blot analysis (Extended Data Fig. 5b).

RT-PCR and qRT-PCR

Total RNA was extracted using the hot phenol method and treated with recombinant DNase I (Takara, cat# 2270A) in the presence of RNasin Plus (Promega, cat# N2611). PrimeScript Reverse Transcriptase (Takara, cat# 2680A) was used for reverse transcription according to the manufacturer's instructions. Quantitative PCR (qPCR) was performed in the same way as ChIP-qPCR (see below). Levels of transcripts of interest relative to *act1*$^+$ mRNA were normalized to those of the *wt* (Fig. 1 and Extended Data Fig. 1) or *ade6*$^+$x1 (Extended Data Fig. 4f). The primers used for analysis are listed in Supplementary Table 2.

Southern blot analysis

Genomic DNA extraction was performed as described previously (https://www.baumann--lab.org/documents/Nurselab_fissionyeasthandbook_000.pdf). Genomic DNA was digested overnight with the indicated restriction enzymes, and separated by agarose gel electrophoresis using 1xTris-Acetate-EDTA buffer, and then transferred to Amersham Hybond-N+ membranes (GE Healthcare, cat# RPN303B) using the TurboBlotter System (Cytiva, cat#10416328). The blots were UV cross-linked using a UVP CL-1000 Ultraviolet Crosslinker (120,000 μJ/cm^2). The DNA probes, which were labeled with [γ-^{32}P] ATP by T4 polynucleotide kinase (Takara, cat# 2021A), were hybridized to the membrane in PerfectHyb Plus Hybridization buffer (Sigma, cat# H7033) at 42℃ overnight. The membrane was subsequently washed with 2× SSC 0.1% SDS buffer at 42℃. The imaging plate was exposed to the membrane for 1-2 days. The primers used for this analysis are listed in Supplementary Table 2.

Northern blot analysis

Total RNA was prepared using the hot phenol procedure and treated with recombinant DNase I (Takara, cat# 2270A) in the presence of RNasin Plus (Promega, cat# N2611). Samples were subjected to electrophoresis in 1% agarose with 6.7% formaldehyde, and transferred to Amersham Hybond-N+ membranes (GE Healthcare, cat# RPN303B) using the TurboBlotter System (Cytiva, cat#10416328). Hybridization was performed as for Southern blot analysis. The primers used for this analysis are listed in Supplementary Table 2.

Northern blot analysis of siRNA

Small RNAs were extracted from 800 × 10^6 cells with the mirVana miRNA Isolation Kit (Ambion, cat# AM1561) according to the manufacturer's instructions. Prior to phenol-chloroform extraction, cells were suspended in 1 mL of Lysis/Binding buffer and homogenized with a bead shocker (Yasui Kikai). Four to twelve micrograms of small RNA was separated on a 15% sequencing gel and transferred to Amersham Hybond-N+ membranes (GE Healthcare, cat# RPN303B) using a Trans-blot SD semi-dry electrophoretic transfer cell (Bio-Rad, cat# 170-3940). After UV crosslinking, *dg*/*dh* elements or *ade6*$^+$ ORF probes labeled with [α-^{32}P] dCTP by random priming (Takara, cat# 6045) were hybridized to the membrane in PerfectHyb Plus Hybridization buffer (Sigma, cat# H7033) at 42℃ overnight. Subsequently, the membrane was washed with 2× SSC 0.1% SDS buffer at 42℃. The imaging plate was exposed to the membrane for 1-2 days. The oligonucleotide probe against snoRNA58 labeled with [γ-^{32}P] ATP by T4 polynucleotide kinase (Takara, cat# 2021A) was also hybridized to the membrane as a loading control. The primers used for this analysis are listed in Supplementary Table 2.

Small RNA-seq

Small RNAs were extracted in the same way as for Northern blot analysis of siRNA. Seventy micrograms of small RNA was separated on a 15% sequencing gel with DynaMarker Prestain Marker for Small RNA Plus (BioDynamics Laboratory Inc., cat# DM253). After electrophoresis, gel fractions corresponding to 20-30 nt were harvested, and crushed in extraction buffer (0.3M NaCl). After overnight rotation at 4℃, the small RNAs were recovered by ethanol precipitation using Ethachinmate (NIPPON GENE, cat# 318-01793) as a carrier. The small RNA (smRNA) libraries were constructed using the SMARTer smRNA-seq kit for Illumina (CLN cat#635029), according to the manufacturer's instructions. After construction, the smRNA libraries were purified by AMPure XP (Beckman Coulter, cat# A63882), and then sequenced on the Illumina HiSeq 4000 system (single-end, 51bp) or HiSeqX system (paired-end, 151bp). Reads (Read 1 in the case of paired-end) were first trimmed for adaptors and A-tailing incorporated during library construction using the cutadapt program[33], with the following parameters; -m 15 -u 3 -a AAAAAAAAAA. Using bowtie (version 1.2.1.1) with -M 1 --best parameters, the trimmed reads were mapped on the modified *S. pombe* genome in which the native *ade6*$^+$ ORF is deleted and the *ura4*$^+$ gene is replaced by the *ade6*$^+$x1 construct. For the truncation analyses of *cenH* (Extended Data Fig. 9), only uniquely mapped reads were used with the -m 1 parameter. The Bedgraph files for visualizing in the IGV were produced with the "genomecov" function of the BEDTools program (version 2.26), which were then normalized by 1 million total mapped reads. Categorization of smRNA reads according to the annotation was performed using the "coverage" function of the BEDTools program. Analyses of smRNA length were conducted after extracting the reads mapped to the *ade6*$^+$x1 construct or pericentromeric *dg*/*dh* regions, using the "intersect" function of the BEDTools program. The length information

of each read is based on bed file. The 5'-end nucleotide analyses were performed using the WebLogo program[34]. In Fig. 3a, Fig. 4c, Fig. 6b, small RNAs mapped on *ade6*$^+$ are shown. For other small RNAs mapped around *ade6*$^+$, see also Supplementary Fig. 1 and Supplementary Discussion 3.

ChIP-qPCR

For ChIP-qPCR, 50 mL of YES or PMG culture at a density of 5 × 10^6 cells/mL was fixed with 1% formaldehyde (Nacalai Tesque, cat# 16223-55) for 20 min at 25℃. After quenching with 150 mM glycine, cells were harvested and washed with buffer A (50 mM HEPES/KOH [pH 7.5], 140 mM NaCl, 1 mM EDTA, 1% Triton-X, and 0.1% sodium-deoxycholate). Cells were then resuspended in buffer A supplemented with protease inhibitors (1 mM PMSF, 0.05 µg/mL bestatin, 0.15 µg/mL benzamidine, 0.25 µg/mL pepstatin and 0.25 µg/mL leupeptin) and homogenized using a bead shocker (Yasui Kikai). After centrifuging lysates at 15,000 rpm for 60 min at 4℃, the pellets were resuspended in 2 mL buffer A and sonicated using a Bioruptor (Cosmo Bio, UCD-250) for 12 min (1 s ON and 2 s OFF at power H). Samples were centrifuged at 15,000 rpm for 15 min, and supernatants were used as whole cell extract (WCE). Each ChIP experiment was performed with 50 µL of 50% slurry of Dynabeads M-280 Sheep anti-Mouse IgG (Thermo Fisher, cat# 11202D) bound to the appropriate antibodies. Antibodies used were: 1 µL of H3K9me2 antibody (mAbProtein, m5.1.1; a gift from Takeshi Urano, Shimane University); 2 µL of anti-Histone H3 antibody (Millipore, cat# 07-690); 1 µL of ANTI-FLAG M2 antibody (Sigma, cat# F1804); 1 µL of anti-Myc tag antibody clone 4A6 (Millipore, cat# 05-724); and 1 µL of the anti-RNA polymerase II antibody clone CTD4H8 (Millipore, cat# 05-623). Note that H3K9me2 antibody (mAbProtein, m5.1.1) recognizes H3K9 mono-, di-, and trimethylation (Takeshi Urano, personal communication). After incubating 500 µL lysates and beads at 4℃ for 2 h with rotation, beads were collected and washed once with each of the following buffers: buffer A, buffer B (50 mM HEPES/KOH [pH 7.5], 500 mM NaCl, 1 mM EDTA, 1% Triton-X, and 0.1% sodium-deoxycholate), buffer C (10 mM Tris-Cl [pH 8.0], 250 mM LiCl, 0.5% NP-40, and 0.5% sodium-deoxycholate), and TE buffer (10 mM Tris-HCl [pH 8.0] and 1 mM EDTA). Beads were resuspended in 100 µL TE buffer containing a final concentration of 100 µg/mL RNase A and incubated at 37℃ for 10 min. Then, an equal volume of Proteinase K solution (0.5% SDS and 0.5 mg/mL Proteinase K [Wako, cat# 166-14003]) was added and incubation continued for 60 min at 37℃. After de-crosslinking by incubation overnight at 65℃, immunoprecipitated DNA was recovered using a FastGene Gel/PCR Extraction Kit (NIPPON Genetics, FG-91302). WCE samples, which were subjected to the same treatments, were purified by phenol-chloroform extraction and ethanol precipitation with glycogen. qPCR was performed using SYBR premix Ex-Taq II (Takara, RR820A) or homemade Taq and qPCR buffer (100 mM Tris-HCl [pH 8.8], 750 mM Trehalose dihydrate, 0.1% [w/v] Tween-20, and 500 mM betaine) containing SYBR Green I, with the Thermal Cycler Dice Real time system TP800 (Takara). Enrichment relative to the euchromatic *fbp1*$^+$ was evaluated for ChIP of RNAi factors, CLRC component, and Pol2. For ChIP of H3K9me, the amount of immunoprecipitated H3K9me relative to the input whole cell extract (input%) was normalized to that of H3, because of its low background signals at *fbp1*$^+$. In Fig. 1 and Extended Data Fig. 1, the amount of immunoprecipitated H3K9me relative to the input whole cell extract (input%) was normalized to that of the wild type. Error bars represent SEM.; n = 3 biological replicates. P-values were determined using a two-sided Student's t-test. The primers used for analysis are listed in Supplementary Table 2.

ChIP-seq.

For ChIP-seq, 600 mL of PMG culture at a density 5×10^6 cells/mL was fixed with 1% formaldehyde for 20 min at 25℃. Subsequent processing was identical to that for ChIP-qPCR with the following modifications. Dynabeads M-280 Sheep anti-Mouse IgG bound to H3K9me antibody (m5.1.1; a gift from Takeshi Urano, Shimane University) were washed once with buffer A before incubation with cell lysate. After incubation at 4℃ for 2 h with rotation, beads were collected and washed twice with each wash buffer except for TE buffer. Beads were resuspended with 100 µL of TE buffer containing a final concentration of 100 µg/mL RNase A. The RNase A solution was also added to 100 µL WCEs at the same final concentration. The samples were incubated at 37℃ for 1 hour. Then, an equal volume of Proteinase K solution was added and samples were incubated at 45℃ for 60 min. After de-crosslinking by incubation overnight at 65℃, immunoprecipitated DNA and WCE were recovered using a QIAquick PCR purification kit (QIAGEN, cat#28104). The ChIP libraries were prepared with a KAPA Hyper Prep Kit (KAPABIOSYSTEMS, cat#KK8504), according to the manufacturer's instructions. The libraries were sequenced on the Illumina HiSeq 2500 system (single-end, 50bp). The sequenced reads were mapped on the modified *S. pombe* genome using bowtie (version 1.2.1.1). The data were processed by SAMtools (version 1.9) and IGVTools to make tdf files that were visualized in the IGV.

5'/3'RACE analysis of *cenH* and *Kint2::ura4*$^+$

Total RNA was extracted using the hot phenol method from 700×10^6 cells, which were derived from exponentially growing cultures. This total RNA was treated with recombinant DNase I (Takara, cat# 2270A) in the presence of RNasin Plus (Promega, cat# N2611), and PolyA+ RNA was purified using the polyA Tract mRNA isolation system IV (Promega, cat# Z5310) according to the manufacturer's instructions. 5'/3'RACE analysis was performed with the SMARTer RACE 5'/3'Kit (Clontech, cat# 634858) following the manufacturer's instructions. The SMART technology takes advantage of a property of terminal transferase activity of reverse transcriptase, called template switching, to select 5'-ends of capped transcripts[35]. Instead of the provided SeqAmpDNA polymerase and linearized pRACE vector, KOD Fx neo (TOYOBO, cat# KFX-201), and a homemade vector derived from pUC18 were used for the amplification and cloning steps, respectively. To enhance specificity, nested PCR was used for analysis of *cenH* ncRNAs. The first PCR of 5'-RACE was performed with 45 cycles at 94℃ for 30 s and 68℃ for 3 min. The subsequent second PCR was performed using 5 µL of 50-fold diluted first-PCR products with 25 cycles at 94℃ for 30 s and 68℃ for 3 min. The first PCR of 3'-RACE was performed with 5 cycles at 94℃ for 30 s and 72℃ for 3 min, followed by 5 cycles at 94℃ for 30 s and 70℃ for 3 min, and finally 25 cycles at 94℃ for 30 s and 68℃ for 3 min. The second PCR was performed using 5 µL of 50-fold diluted first-PCR products with 25 cycles at 94℃ for 30 s and 68℃ for 3 min. For RACE analysis of *Kint2::ura4*$^+$ and *act1*$^+$, PCR was performed in the same way as for the first PCR of *cenH*. The results of sequencing RACE products, which were randomly selected from independent clones, were verified by BLAST, and it was confirmed that they were derived from the *cenH* element, not from the pericentromeric *dg*/*dh* element. For sense transcripts and antisense transcripts of *cenH* ncRNA, thirty clones of 5'RACE or 3'RACE products were identified. The primers used for this analysis are listed in Supplementary Table 2. The Sequence data of RACE products are listed in Supplementary Data 1.

Truncation analysis of *cenH* fragment with a plasmid-based minichromosome

To generate truncated *cenH* fragments, the 4.5kb *cenH* sequence was amplified by PCR from genomic DNA of the h^{90} strain. Diluted PCR products were used as template for the subsequent PCRs, which produced a series of truncated *cenH* fragments. The resultant PCR products had BamHI and NcoI sites derived from primers, and were cloned into the corresponding site of MC-L5[36]. This resulted in a plasmid-based minichromosome[36] in which the L5 fragment is replaced by a truncated *cenH* fragment (pAP-275, pLCC-*cenH* 3.0kb; pAP-315, pLCC-*cenH majorΔ Rev TSS*; pAP-273, pLCC-*cenH major Rev TSS only*; pAP-269, pLCC-*cenH* 1.5kb). Strains containing minichromosomes were grown in PMG medium lacking adenine and uracil. Primers used for generation of the pLCC-*cenH* series are listed in Supplementary Table 2.

CAGE-seq

CAGE library preparation, sequencing, mapping, and gene expression analysis were performed by DNAFORM (Yokohama, Kanagawa, Japan). In brief, RNA quality was assessed by a bioanalyzer (Agilent) to ensure that the RIN (RNA integrity number) was >7.0, and the A260/280 and 260/230 ratios were >1.7. First-strand cDNAs were transcribed to the 5'-end of capped RNAs, attached to CAGE "bar code" tags, and the sequenced CAGE tags were mapped to the *S. pombe* ASM294v2 genome using BWA software (0.7.15-r1140), after discarding ribosomal or non-A/C/G/T base-containing RNAs. To confirm reproducibility, CAGE-seq was repeated twice for total RNA from Epe1 OP cells, and representative data are presented. CAGE-tag data without clustering were used for visualization of TSSs and gene expression profiles to avoid the loss of signals derived from widespread TSSs at *dg*/*dh* elements. For consensus sequence analysis of core promoters in *dg*/*dh*, the sequences of ±7 nt or ±50 nt flanking unique CAGE TSSs (>0.05 CPM) were retrieved from both ±strands of *dg*/*dh* elements, specifically, chromosome I (3752407–3765008, 37772666–3789949), chromosome II (1602261–1618293, 1630123–1643849), chromosome III (1071176–1092437, 1106558–1139536), and MTR (1–20128), and analyzed by Weblogo3[34]. As a control, the same number of pseudo TSSs was randomly chosen from anywhere in the *dg*/*dh* elements, and sequences of ±50 nt flanking those pseudo TSSs were analyzed in the same way. For analysis of mRNA TSSs, CAGE TSSs (>0.05 CPM) of the wild type mapped in the annotated 5'-UTR (including its upstream 100 nt) of euchromatic genes were classified into three groups (high, medium, and low) based on expression levels (CPM). For each group, the sequences of ±7 nt or ±50 nt flanking unique TSSs were retrieved and analyzed in the same way. To extract CAGE TSSs that had both the Inr (Y/R) sequence at -1/+1 and A/T rich region at 25–32 nt upstream of the TSS, a Position-Weight matrices file was created by Biopython (https://biopython.org/) with FASTA files that were used for Weblog analysis. The resultant matrices file was consistent with that of Weblogo. Based on this matrices file, CAGE TSSs with both the Inr sequence and the upstream A/T rich region were extracted using FIMO (http://meme-suite.org/doc/fimo.html?man_type=web), and mapped to ASM294v2 genomes.

References

1. Saksouk, N., Simboeck, E. & Déjardin, J. Constitutive heterochromatin formation and transcription in mammals. *Epigenetics & chromatin* **8,** 3–17 (2015).

2. Henikoff, S. Conspiracy of silence among repeated transgenes. *BioEssays* **20,** 532–535 (1998).

3. Gladyshev, E. & Kleckner, N. DNA sequence homology induces cytosine-to-thymine mutation by a heterochromatin-related pathway in *Neurospora*. *Nat Genet* **49,** 887–894 (2017).

4. Wang, F. *et al.* The assembly and maintenance of heterochromatin initiated by transgene repeats are independent of the RNA interference pathway in mammalian cells. *Mol Cell Biol* **26,** 4028–4040 (2006).

5. Pal-Bhadra, M. *et al.* Heterochromatic silencing and HP1 localization in *Drosophila* are dependent on the RNAi machinery. *Science* **303,** 669–672 (2004).

6. Luo, Z. & Chen, Z. Improperly terminated, unpolyadenylated mRNA of sense transgenes is targeted by RDR6-mediated RNA silencing in *Arabidopsis*. *Plant Cell* **19,** 943–958 (2007).

7. Schubert, D. *et al.* Silencing in Arabidopsis T-DNA transformants: the predominant role of a gene-specific RNA sensing mechanism versus position effects. *Plant Cell* **16,** 2561–2572 (2004).

8. Cam, H. P. *et al.* Comprehensive analysis of heterochromatin- and RNAi-mediated epigenetic control of the fission yeast genome. *Nat Genet* **37,** 809–819 (2005).

9. Kato, H. *et al.* RNA polymerase II is required for RNAi-dependent heterochromatin assembly. *Science* **309,** 467–469 (2005).

10. Djupedal, I. *et al.* RNA Pol II subunit Rpb7 promotes centromeric transcription and RNAi-directed chromatin silencing. *Genes Dev* **19,** 2301–2306 (2005).

11. Verdel, A. *et al.* RNAi-mediated targeting of heterochromatin by the RITS complex. *Science* **303,** 672–676 (2004).

12. Bayne, E. H. *et al.* Stc1: a critical link between RNAi and chromatin modification required for heterochromatin integrity. *Cell* **140,** 666–677 (2010).

13. Motamedi, M. R. *et al.* Two RNAi complexes, RITS and RDRC, physically interact and localize to noncoding centromeric RNAs. *Cell* **119,** 789–802 (2004).

14. Colmenares, S. U., Buker, S. M., Bühler, M., Dlakić, M. & Moazed, D. Coupling of double-stranded RNA synthesis and siRNA generation in fission yeast RNAi. *Moll Cell* **27,** 449–461 (2007).

15. Kowalik, K. M. *et al.* The Paf1 complex represses small-RNA-mediated epigenetic gene silencing. *Nature* **520,** 248–252 (2015).

16. Audergon, P. N. C. B. *et al.* Epigenetics. Restricted epigenetic inheritance of H3K9 methylation. *Science* **348,** 132–135 (2015).

17. Ragunathan, K., Jih, G. & Moazed, D. Epigenetic inheritance uncoupled from sequence-specific recruitment. *Science* **348,** 1258699–1258699 (2015).

18. Zofall, M. & Grewal, S. I. S. Swi6/HP1 recruits a JmjC domain protein to facilitate transcription of heterochromatic repeats. *Moll Cell* **22,** 681–692 (2006).

19. Trewick, S. C., Minc, E., Antonelli, R., Urano, T. & Allshire, R. C. The JmjC domain protein Epe1 prevents unregulated assembly and disassembly of heterochromatin. *EMBO J* **26,** 4670–4682 (2007).

20. Ayoub, N. *et al.* A novel jmjC domain protein modulates heterochromatization in fission yeast. *Mol Cell Biol* **23,** 4356–4370 (2003).

21. Li, H. *et al.* Genome-wide analysis of core promoter structures in *Schizosaccharomyces pombe* with DeepCAGE. *RNA Biol* **12,** 525–537 (2015).

22. Shimada, Y., Mohn, F. & Bühler, M. The RNA-induced transcriptional silencing complex targets chromatin exclusively via interacting with nascent transcripts. *Genes Dev* **30,** 2571–2580 (2016).

23. Simmer, F. *et al.* Hairpin RNA induces secondary small interfering RNA synthesis and silencing in trans in fission yeast. *EMBO Rep* **11,** 112–118 (2010).

24. Iida, T., Nakayama, J.-I. & Moazed, D. siRNA-mediated heterochromatin establishment requires HP1 and is associated with antisense transcription. *Moll Cell* **31,** 178–189 (2008).

25. Yu, R., Wang, X. & Moazed, D. Epigenetic inheritance mediated by coupling of RNAi and histone H3K9 methylation. *Nature* **558,** 615–619 (2018).

26. Andersen, P. R., Tirian, L., Vunjak, M. & Brennecke, J. A heterochromatin-dependent transcription machinery drives piRNA expression. *Nature* **549,** 54–59 (2017).

27. Law, J. A. *et al.* Polymerase IV occupancy at RNA-directed DNA methylation sites requires SHH1. *Nature* **498,** 385–389 (2013).

28. Hövel, I., Pearson, N. A. & Stam, M. Cis-acting determinants of paramutation. *Semin Cell Dev Biol* **44,** 22–32 (2015).

29. Walter, J., Hutter, B., Khare, T. & Paulsen, M. Repetitive elements in imprinted genes. *Cytogenet Genome Res* **113,** 109–115 (2006).

30. Hutter, B., Helms, V. & Paulsen, M. Tandem repeats in the CpG islands of imprinted genes. *Genomics* **88,** 323–332 (2006).

31. Hayashi, A. & Tanaka, K. Short-Homology-Mediated CRISPR/Cas9-Based Method for Genome Editing in Fission Yeast. *G3 (Bethesda)* **9,** 1153–1163 (2019).

32. Maundrell, K. Thiamine-repressible expression vectors pREP and pRIP for fission yeast. *Gene* **123,** 127–130 (1993).

33. M. Martin, Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet Journal* **17,** 10-12 (2011)

34. Crooks, G. E., Hon, G., Chandonia, J.-M. & Brenner, S. E. WebLogo: a sequence logo generator. *Genome Res* **14,** 1188–1190 (2004).

35. Plessy, C. *et al.* Linking promoters to functional transcripts in small samples with nanoCAGE and CAGEscan. *Nat. Methods* **7,** 528–534 (2010).

36. Buscaino, A. *et al.* Distinct roles for Sir2 and RNAi in centromeric heterochromatin nucleation, spreading and maintenance. *EMBO J* **32,** 1250–1264 (2013).

37. Watt, S. *et al.* urg1: a uracil-regulatable promoter system for fission yeast with short induction and repression times. *PLoS ONE* **3,** e1428 (2008).

38. Jia, S., Noma, K.-I. & Grewal, S. I. S. RNAi-independent heterochromatin nucleation by the stress-activated ATF/CREB family proteins. *Science* **304,** 1971–1976 (2004).

39. Yu, R., Jih, G., Iglesias, N. & Moazed, D. Determinants of Heterochromatic siRNA Biogenesis and Function. *Moll Cell* **53,** 262–276 (2014).

40. Volpe, T. A. *et al.* Regulation of heterochromatic silencing and histone H3 lysine-9 methylation by RNAi. *Science* **297,** 1833–1837 (2002).

41. Djupedal, I. *et al.* Analysis of small RNA in fission yeast; centromeric siRNAs are potentially generated through a structured RNA. *EMBO J* **28,** 3832–3844 (2009).

42. Bayne, E. H. *et al.* Splicing factors facilitate RNAi-directed silencing in fission yeast. *Science* **322,** 602–606 (2008).

43. Hayashi, A. *et al.* Heterochromatin protein 1 homologue Swi6 acts in concert with Ers1 to regulate RNAi-directed heterochromatin assembly. *Proc Natl Acad Sci USA* **109,** 6159–6164 (2012).

44. Rougemaille, M. *et al.* Ers1 links HP1 to RNAi. *Proc Natl Acad Sci USA* **109,** 11258–11263 (2012).