Swi6/HP1 binding to RNA-DNA hybrids initiates heterochromatin assembly at the centromeric dg-dh repeats in Fission Yeast

Jyotsna Kumar[1,2, 3], Swati Haldar[1,2,4], Neelima Gupta[1,2,5], Viney Kumar[1,6], Manisha Thakur[1,7], Keerthivasan Raanin Chandradoss[8], Debarghya Ghose[1], Dipak Dutta[1], Kuljeet Singh Sandhu[8] and Jagmohan Singh[1,9]

[2]: Equal contribution

[3]: Present address: Laboratory of Biochemistry and Genetics, National Institute of Diabetes, Digestive, and Kidney Diseases, National Institutes of Health, Bethesda, Maryland 20892-0851

[4]: Present address: Microbiology Division, Patanjali Research Institute, Haridwar, Uttarakhand, INDIA

[5]: Present address: Department of Anatomy, Yong Loo Lin School of Medicine, National University of Singapore, Singapore.

[6]: Present address: Molecular Endocrinology Laboratory, Indian Institute of Technology at Roorkee, Roorkee, UP, India

[7]: Present address: 1/47 Pine Lane, Heatherton 3202 Melbourne, Victoria.

[8]: Department of Biological Sciences, Indian Institute of Science Education and Research (IISER)-Mohali, Knowledge City, Sector 81, SAS Nagar 140306, Punjab

[1]: CSIR-Institute of Microbial Technology, Chandigarh-160036. India.

[9]: Corresponding author.

ABSTRACT

Canonically, heterochromatin formation in fission yeast and metazoans involves di/trimethylation of histone H3 at lysine 9 position (me2/me3-K9-H3) by the histone methyltransferase (HMT) Suv39/Clr4, followed by binding of Swi6/HP1 to me2/me3-K9-H3 via its chromodomain. Subsequent self-association of Swi6/HP1 on adjacent nucleosomes leads to folded heterochromatin structure. An alternate model suggests a cooperative interaction between Clr4 and Swi6/HP1 in heterochromatin assembly. HP1 binding to RNA has also been invoked for heterochromatin silencing in metazoans. Recruitment of Swi6/HP1 to centromere has been shown to be dependent on the RNAi pathway in fission yeast. Here we show that Swi6/HP1 exhibits a hierarchy of binding affinity to RNAs, ranging from promiscuous, low-affinity binding to mRNAs, to moderate-affinity binding to the RNAi-generated siRNAs corresponding to the dg-dh repeats present in pericentromeric heterochromatin regions, to high affinity binding to the RNA-DNA hybrids to the cognate dg-dh repeats. Together with sensitivity of Swi6 localization and silencing to RNaseH, our results suggest a dynamic control of localization of Swi6/HP1 towards the dg-dh repeats versus euchromatic regions. This is mediated by its binding to RNA-DNA hybrid at the dg-dh repeats, as an RNAi-dependent and Me2/me3-K9-H3-independent mechanism of recruitment, leading to heterochromatin formation and silencing.

INTRODUCTION

According to the histone-code hypothesis, eukaryotic chromatin is organized into expressed/euchromatic and repressed/heterochromatic regions through a gamut of distinct post-translational histone modifications (1). In fission yeast (*Schizosaccharomyces pombe*) and metazoans, freshly assembled chromatin contains histone H3 acetylated at K9 and K14 residues (2,3). At heterochromatin regions, the H3-acetyl groups are removed by the deacetylases, Clr3, Clr6 and Sir2, followed by di- and tri-methylation of K9 by Suv39/Clr4 (3). Heterochromatin protein Swi6/HP1 binds to Me2/Me3-H3-K9 through its chromodomain (4). Subsequent self-association of HP1/Swi6 through its chromoshadow domain results in a folded, transcriptionally inactive heterochromatin (5). An alternative model has invoked a mutually cooperative recruitment of Swi6/HP1 and Clr4/Suv3-9 in heterochromatin formation and spreading (6,7).

The RNAi pathway is also involved in heterochromatin assembly as the *rnai* mutants show impaired H3K9 dimethylation and Swi6 recruitment at the mating type and pericentromeric regions (8). According to the current model, the double-stranded (ds) RNAs formed by bidirectional transcription from the *dg-dh* repeat regions are cleaved by Dcr1 to produce the ds siRNAs. These siRNAs are bound to the ARC (Argonaute small interfering RNA chaperone) complex, which contains the endoribonuclease 1 (Ago1), Arb1 and Arb2 subunits (9). After the passenger strand release, the guide strand is transferred to the chromatin-bound complex, called RNA-induced transcriptional silencing (RITS) complex (10), which consists of Ago1, Chp1- a chromodomain protein and Tas3. The binding of Chp1 to Me2/Me3-K9-H3 tethers the RITS complex to heterochromatin. Subsequently, Rdp1 (RNA-dependent RNA polymerase) copies the ncRNA strand synthesized by RNA PolII, using the guide strand as a primer, to generate the double-stranded (ds) RNA (11). In turn, the ds RNA, is cleaved by Dcr1 to produce the siRNA, which start the new amplification cycle (12). The RNAi link to the formation of Me2/Me3-K9-H3 possibly occurs through Stc1-dependent recruitment of the HMT Clr4-containing CLRC (CLR like Clr4 complex; 13,14).

As mentioned above, RNAi pathway in *S. pombe* is linked to recruitment of Swi6/HP1 at the pericentromeric regions. These regions contain the *dg-dh* repeats, which are also present at the cenH region between the *mat2* and *mat3* loci and the subtelomeric genes *tlh1* and *tlh2* (15,16).

However, the exact mechanism by which the RNAi pathway is involved in heterochromatin recruitment of Swi6/HP1 is not known.

In this regard, earlier studies showed that apart from binding to Me2/Me3-Lys9-H3, the chromodomain present in Swi6/HP1, also acts as RNA-binding domain (17). Furthermore, RNA binding by the murine HP1 through its chromo- and hinge domains is required for heterochromatin assembly (18). However, another study claimed that heterochromatin-transcribed RNA competes with Lys9 methylated H3 for binding to Swi6 and is carried to the exosome (19).

Here, we demonstrate that Swi6/HP1 binds with moderate affinity to the dg-dh specific siRNAs engendered by RNAi machinery-mediated processing and high affinity to the cognate RNA-DNA hybrids. We also show that his mode of binding appears to be an autonomous mode of Swi6/HP1 recruitment to heterohromatin, which is independent of the canonical mode of recruitment by binding to Me2/Me3-K9-H3.

MATERIALS AND METHODS

Strains and plasmids

The strains used in this study are listed in Table S1

The *rnh1* gene (RNase H1SPBC.336.06c) cDNA was amplified by PCR from genomic DNA and ligated between the *Xho*I and *Bam*HI sites in pREP3X vector to obtain a thiamine repressible *rnh1* expression construct. Similarly, the construct *nmt1-rnh20*1 (a gift of J-I. Nakayama; 20), encoding the subunit A of the three subunit enzyme RNase H2 (SPAC4G9.02), was PCR amplified from genomic DNA and cloned between the restriction sites *Bam*HI and *Sal*I upstream of the *nmt1* promoter in the vector pREP1. All fission yeast media were prepared according to Moreno *et.al*. (21).

Genetic Techniques

Sporulation was checked either microscopically or by staining the colonies with iodine vapours. Efficiently switching strains produce equal number of cells of opposite mating type, which mate to form zygotic asci that sporulate on minimal media. The spore cell wall contains a starchy compound that stains dark purple with iodine vapours. Thus, efficiently switching cells of an h^{90} strain grow into colonies that stain dark purple with iodine while colonies of cells that switch inefficiently give light yellowish staining (21). Determining the percentage of dark colonies provided a semi-quantitative measure of the switching efficiency. After iodine staining the colonies were photographed under Olympus StereoZoom Microscope.

Silencing status of any reporter gene was assessed by monitoring its expression using different plate assays. Expression of the *ura4* reporter was detected by dilution spotting, wherein 5µl of ten-fold serial dilutions of the overnight cultures of the required strains were spotted on non-selective plates, plates lacking uracil and those containing 5'FOA. FOA provided counter-selection to ura^- cells. Thus, ura^+ strains that grow well on ura minus but grow poorly on FOA plates, while ura- cells grow poorly on ura minus but grow well on FOA plates. Cell densities of all the cultures were normalized before serial dilution.

Expression of heterochromatic *ade6* was assessed by streaking the cells for single colonies on plates containing limited adenine. Cells expressing *ade6* produced light pink or white colonies on these plates and cells failing to express *ade6* gene produced dark red colonies. The fraction of colonies giving pink/white colour provide a qualitative measure of the silencing defect. Similarly, the extent of silencing of the *his3* locus at subtelomeric locus, denoted as *his3-telo* (22) was measured qualitatively by growth on plates lacking histidine.

Cloning, site-directed mutagenesis, protein over-expression and purification

Lysine residues (KKK) at positions 242-244 in Swi6 were mutated to alanine (AAA) in the *nmt1-GFP-swi6*$^{+}$ expression construct (kindly gifted by Dr. Alison Pidoux) using the QuickChange site-directed mutagenesis kit (Stratagene), as per the manufacturer's instructions, and confirmed by sequencing. The mutant is referred to as *swi6*$^{3K\rightarrow 3A}$. The wild type full length Swi6 and Swi6$^{3K\rightarrow 3A}$ were tagged with GST at amino-terminus by cloning into pGEX-KG expression vector as BamHI-HindIII fragments.

Expression and purification of GST-tagged Swi6 protein

The recombinant constructs containing GST, GST-Swi6, GST-Swi6$^{3K\rightarrow 3A}$, GST-Swi6 -CD, GST-Swi6-CD-Hinge and GST-Swi6 (CSD) were transformed in to BL21(DE3) host cells. The transformed cells were cultured in the LB-Amp medium at 37°C. At mid-log phase, cells were induced with 1mM IPTG and incubated further at 25°C for 8hrs. For purification of GST-tagged proteins, 2-4 mg protein extract was added to 50 µl of Glutathione Sepharose beads (Cat. no. 17527901, GE healthcare Ltd.) equilibrated with 1XPBS and incubated for 10-14 hours at 4°C. Purified proteins were collected after subjecting the mixture to elution for 2hrs at 4°C with 250 µl elution buffer containing 10mM reduced glutathione. Purification was confirmed by SDS-PAGE.

Expression of the *nmt1* promoter driven GFP tagged *swi6*$^{+}$ as well as *swi6*$^{3K\rightarrow -3A}$ gene was induced by removing thiamine from the medium of an overnight grown culture by washing the cells thoroughly and growing them further in medium lacking thiamine. It took 17-18 hrs of

induction to achieve maximum expression level. Swi6 protein was detected by Western blotting with either anti-Swi6 (in-house, 1:10000), anti-GFP [Santa Cruz (sc-9996) 1:1000] or anti-GST [Santa Cruz (sc-138), 1:1000] antibodies. Wherever required, anti-α-tubulin antibody (Sigma, cat# T-9026; 1:4000) was used to probe the loading control. Alkaline phosphatase conjugated anti-rabbit (Sigma, cat# A-3562; 1:20,000) antibody was used, as per instructions. Western blots were developed using nitroblue tetrazolium (NBT) and bromo-chloro-indolyl phosphate (BCIP) as chromogenic substrates.

Fluorescence Microscopy

10ml cultures of cells expressing the empty vector, *nmt1-GFP-swi6*$^{+}$ or *nmt1-GFP-swi6*$^{3K\rightarrow\text{-}3A}$ in *swi6Δ* background were harvested by centrifugation at 5000 rpm at RT. Cells were fixed in chilled 70% ethanol, rehydrated sequentially in 50%, 30%, 20% ethanol and finally in water. The pellets were resuspended in 50µl of 1X PBS. 5µl of this suspension was then visualized under fluorescence microscope. Similar method was followed for visualizing subcellular localization of chromosomally GFP-tagged Swi6 under the influence of over-expressed *rnh1* and *rnh201*.

Generation of [γ-P^{32}] labelled single stranded *in vitro* transcribed RNAs (ssRNA) and RNA-DNA hybrids

The protocol for synthesis and labeling of the single stranded RNA complementary to the heterochromatin-specific DNA sequences (23,24) and the cognate RNA-DNA hybrids, is provided in supplementary data.

Electrophoretic mobility shift assay (EMSA)

For EMSA, purified protein extract was mixed with ss siRNA or siRNA-DNA hybrid that was radiolabeled with [γ-^{32}P]ATP with polynucleotide kinase in the binding buffer containing 20mM HEPES (pH 7.5), 50mM KCl, 2mM EDTA, 0.01% NP40, 1mM DTT, 5 units RNase inhibitor, 1 µg BSA and 5% glycerol alongwith ~5 µg budding yeast total RNA or poly dI-dC as non-specific competitor, respectively, followed by incubation on ice. The samples were loaded on native TGE/ polyacrylamide (the tracking dyes Bromophenol blue and Xylene cyanol were not

used in loading buffer as they interfere with the binding of proteins with nucleic acids). A negative control (without protein) was also loaded on the gel. 5X loading buffer containing both tracking dyes was added to negative control to track migration. For competition experiment, 0.1-10 pmol of 'cold' unlabeled RNAs and DNAs were used. The gel was dried in gel dryer and exposed to X-ray film or scanned in a phosphoimager.

The band intensities of the DNA-protein complex and free DNA for each lane were quantified using Scion Image software and the data were plotted according to a non-linear regression equation corresponding to one site specific-binding model as follows:
Y = Bmax.X /(Kd+X), where Y represents the amount of protein-bound radioactivity, X, the protein concentration, Bmax –the total change in radioactivity counts and Kd, the equilibtium binding constant (25).

Chromatin Immunoprecipitation (ChIP) Assay

Cells of appropriate strains were grown to an OD_{600} of 0.6 and crosslinked for 30 min with 3% formaldehyde at 18°C. Immunoprecipitation was performed with 1 µl anti-H3K9me$_2$ (Upstate, 07-353) or 2 µl of polyclonal anti-Swi6 antibody per 400 µl reaction. Multiplex radioactive PCR amplifications of the immunoprecipitated chromatin samples, in the presence of [α-P^{32}]ATP as the source of radiolabelled ATP in addition to non-radioactive dNTPs, were done using the following primers
ADE6F: 5'-TGCGATGCACCTGACCAGGAAAGT-3';
ADE6R: 5'-AGAGTTGGGTGTTGATTTCGCTGA-3'
URA4F: 5'-GAGGGGATGAAAAATCCCAT-3'
URA4R: 5'-TTCGACAACAGGATTACGACC-3'
HIS3F: 5'-AGGTGTCCTTCTTGATGCCA-3'
HIS3R: 5'-CGAATTCCTGCTAGACCGAA-3'
dh For: 5'-GGAGTTGCGCAAACGAAGTT-3'
dh Rev: 5'-CTCACTCAAGTCCAATCGCA-3'
act1F: 5'-TCCTACGTTGGTGATGAAGC-3'
act1R: 5'-TCCGATAGTGATAACTTGAC-3'

The products are resolved on 4% 1X TBE polyacrylamide gel electrophoresis. The gel was exposed to the magnetic screen and scanned in Fuji ImageProcessor. The bands are quantified densitometrically from three independent sets of PCR using inbuilt MultiGauge Software. Data presented as mean SE. Data was analyzed through one-way ANOVA with Turkey's post-hoc test where *** denotes $p<0.001$ when compared with vector and ns stands for non-significant.

DNA-RNA hybrid Immunoprecipitation (DRIP) Analysis

Cells harvested from 500 ml log phase culture were washed with 1X PBS, resuspended in 2 ml PEMS (100mM PIPES, 1mM $MgCl_2$, 1mM EDTA, 1.2 M sorbitol) supplemented with 20µl lyticase/zymolyase and incubated at 37°C for 30 min for spheroplasting. Spheroplasts were resuspended in 1.6ml TE containing 42µl of 20% SDS and 10µl of 10mg/ml Proteinase K and overnight at 37ºC. Supernatant was extracted once with phenol:CIA and genomic DNA in the aqueous phase ethanol precipitated. 10µg of this DNA was digested overnight with Sau3AI at 37°C in 150 µl reaction. Following a phenol clean-up ethanol precipitated Sau3AI digested DNA was dissolved in TE without RNase A. One aliquot of the digested DNA was subjected to overnight RNase H (NEB cat#M0297) treatment at 37°C. Subsequently, the RNAse H treated samples were purified by phenol extraction and ethanol precipitation and dissolved in 50 µl TE mixed with binding buffer [10mM NaPO4 (pH 7.0), 0.14M NaCl, 0.05% TritonX-100]. 50 µl of both RNase H treated and untreated samples were saved as non-immunoprecipitated samples (NIPs). The DNA was immunoprecipitated overnight with 2 µl of anti-DNA-RNA hybrid mouse mono clonal antibody S9.6 (ATCC HB8730; ref 26) with protein G Sepharose beads and eluted in a buffer containing 50mM Tris (pH 8.0), 10mM EDTA, 0.5% SDS and 1.5 µg/ml Proteinase K at 55°C for 45 min. DNA from this complex was retrieved through phenol:CIA clean-up followed by ethanol precipitation. 2-3ng of this DNA was used for each RT-qPCR reaction set up with primers specific to the centromeric *dh* repeat region using SYBR Master Mix (ThernoFisher) in QuantStudio3 (ThermoFisher). Fold enrichment was calculated using $\Delta\Delta C_t$ method (27).

RESULTS

We investigated whether Swi6/HP1 may be recruited to the dg-dh repeats by binding to the respective si RNAs in *S. pombe*. This scenario is consistent with the loss of heterochromatin recruitment of Swi6 as well as siRNA generation in *rnai* mutants (8). Indeed, the siRNAs found *in vivo*, labeled A-K (23) or generated *in vitro* (24) were mapped to the *dg* and *dh* repeats that are shared among the outer repeats (*otr*) on all three centromeres (24), the *cenH* (15) region (*mat2-mat3* interval) on *chrII* and the subtelomeric *tlh1* and *tlh2* genes on *chrI* and *chrII* (16) (Figure 1A,B; Supplementary Data Figure S1 and S2; Supplementary Data Tables S2 and S3).

Specific binding of Swi6/HP1 to the dg-dh- specific siRNAs

By pairwise alignment of the siRNA sequences (23,24) (Supplementary Data Figure S2) we constructed DNA templates comprising the T7 promoter linked with the *dg-dh* DNA sequences cognate to the siRNA sequences, denoted as B,D,E H and K (23) (Figure 1B), and I-IX (24) generated by Dicer mediated cleavage of "RevCen" RNA (Supplementary Data Table S2). The respective single stranded (ss)-siRNAs were generated by using T7 RNA polymerase (Supplementary Data Figure 3A, Table S4). Of these, the siRNAs, denoted as 'D-For' and 'E-For' were radiolabeled and used to study the interaction with Swi6 (Supplementary Data Figure S3B, C) by Electrophoretic Mobility Shift Assay (EMSA; Figure 1C). Binding of Swi6 with the 'D-For' RNA was indicated by slow migrating bands I, II and III, representing different complexes (Figure 1C). The binding was specific, as it was competed out by excess of unlabeled D-For RNA (Figure 1C). No binding was detected with the complementary D-Rev RNA (not shown). Other ssRNAs-B, E and H (Figure 1C) and K (not shown), as well as I-IX; (Supplementary Data Figure S4A,B) also competed efficiently, suggesting the presence of conserved sequence elements in the 'A-K' siRNAs (Figure 1D). The strength of binding/competition efficiency was in the order: B>D=E>H. A consensus Swi6/HP1-binding sequence obtained by sequence alignment (28) was: G/C.A.G.T/C/A.A.G/T/A.G/C.G/T.G/C/A.T (Supplementary Data Fig. S5).

The equilibrium binding constant, Kd, for Swi6/HP1 binding to 'D-For' RNA was estimated to be 2.15±0.1μM (Figure 1E). This binding affinity is ~17.5-fold stronger than that reported for the euchromatic, 'Cen100' RNA corresponding to the thymidylate kinase gene (Kd=38μM; 19). We noted that in the latter study EMSA was performed in agarose gel, which can detect weaker binding due to greater pore size and less hindrance to the protein-RNA complex from migrating (19). Indeed, an EMSA using polyacrylamide gel failed to detect any binding of Swi6 to the 'Cen100' RNA (Supplementary Data Figure S6A) or the 'RevCen" RNA (Supplementary Data Figure S6B), the putative heterochromatic si-RNA precursor (24). Thus, Swi6-HP1 shows stronger affinity towards the small RNAs like 'D-For' and 'E-For' than the longer euchromatic and heterochromatic precursor RNAs

Chromodomain-mediated binding of Swi6 to siRNAs is influenced by the hinge domain

RNA-binding by the chromodomain in murine HP1α was influenced by the hinge domain, in particular, by the lysine triplet sequence in the hinge domain (18). Sequence comparison showed that this lysine triplet is conserved between the murine HP1 and Swi6/HP1 (residues 242-244; Figure 1F). We found that, like HP1α, both the CD and CD-hinge but not the CSD region of Swi6 exhibit strong binding to the 'E-For' (used interchangeably with D-For because of similar affinity; Figure 1G; 1H)). Furthermore, the siRNA binding is drastically reduced in case of the Swi6$^{3K\rightarrow 3A}$ protein (lysine triplet mutated to alanine; Figure 1D, right panel).

Swi6-HP1 is involved in protection but not generation of siRNA

We validated these results *in vivo* using RNA immunoprecipitation (RIP-seq) analysis. The radiolabelled RNA isolated from immunoprecipitates of extracts of cells expressing *GFP-swi6*$^{+}$, but not *GFP-swi6*$^{3K\rightarrow 3A}$ or empty vector with anti-GFP antibody was hybridized with blots containing DNA fragments of *dg*, *dh*, *cenH* (*dhk*) but not with a*ct1* (Supplementary Data Figure S7B). Visualization of the radiolabelled size-fractionated siRNA by denaturing urea/acrylamide gel electrophoresis indicated a qualitative and quantitative change in siRNAs.in *swi6*$^{3K\rightarrow 3A}$ mutant, as compared with *swi6*$^{+}$ (Supplementary Data Figure S7C). We infer that as compared to the Swi6-bound siRNAs, the siRNAs bound to Swi6$^{3K\rightarrow 3A}$ are depleted with respect to the *dg-dh* sequences and may represent non-specifically associated smaller RNAs similar to *swi6Δ* strain

(Vector control, Supplementary Data, Figure S7C). These results indicate that Swi6^{+} protein binds and protects the *dg-dh* siRNA from degradation but Swi6$^{3K\rightarrow 3A}$ does not. RT-PCR analysis revealed the lack of bidirectional transcription in *swi6Δ* strain as well as the transformed strains (Supplementary Data Figure S7D), while *dcr1Δ* strain showed accumulation of bidirectional transcripts specific for the *dg-dh* repeats, as shown earlier[7](Supplementary Data Figure S7D; ref. 8). These results argue against a role of Swi6 and Swi6$^{3K\rightarrow 3A}$ in generation of siRNAs.

Swi6-HP1 interacts with *dg-dh* specific siRNAs *in vivo*

Further, the size-fractionated small RNAs were subjected to RIP-Seq analysis. We observed no change in the genome wide occupancy of siRNAs bound to Swi6p and Swi6$^{3K\rightarrow 3A}$p (Figure 2A, left panel). However, we did observe a considerable depletion of siRNA corresponding to the *dg* and *dh* repeats in all three centromeres (Figure 2A,B; p= 0.001; Supplementary Data Figure S8B), the *dh* repeats in the *tlh1* and *tlh2* genes[21] (Figure 2A, B; Supplementary Data Figure S8B; p=0.02), and *cenH* region of the *mat* locus (Figure 2A,B) in the *swi6$^{3K\rightarrow 3A}$* mutant as compared to *swi6^{+}* strain. This is consistent with the report that Swi6 remains bound to sub-telomeric domains in the absence of telomeric repeats (29), implying prominent subtelomeric heterochromatinization by Swi6 and associated siRNAs.

To validate the EMSA results we analyzed the RIP-seq signals in the genomics bins containing the sequence motifs A-K (23). The analysis clearly showed significantly reduced enrichment of the reported siRNAs (23) bound to Swi6$^{3K\rightarrow 3A}$ versus the Swi6^{+} protein (Figure 2C), confirming that these siRNAs do interact with Swi6 but not Swi6$^{3K\rightarrow 3A}$ *in vivo*.

Loss of siRNA binding abrogates heterochromatin localization of Swi6 and silencing

We further analyzed whether the *swi6$^{3K\rightarrow 3A}$* mutation affects the heterochromatin localization of GFP tagged Swi6. As reported earlier (30), ~65% of *swi6^{+}* cells contain 3 spots, ~25% cells two spots and 10% cells one spot GFP-Swi6 (Supplementary Data, Figure S9A,B). However, in *swi6$^{3K\rightarrow 3A}$* mutant, only ~20% cells each contained one, two or three spots or diffuse appearance (Supplementary Data Figure S9A,B), consistent with delocalization of Swi6 from

heterochromatin and loss of silencing (30). We performed plate and ChIP assay to assess the effect of the *swi6*$^{3K\rightarrow 3A}$ mutation on silencing. Indeed, we observed derepression of the *ade6* reporter at the outer repeat *otr1R* and *ura4* at the inner repeat *imr1L* of *cenI* in *swi6*$^{3K\rightarrow 3A}$ mutant, as indicated by growth of pink colonies on adenine limiting media (Figure 3A,B) or enhanced growth of cells on plates lacking uracil, respectively (Figure 3E; 31). Complementing these results, ChIP assay showed reduced localization of Swi6 and Me2-K9-H3 at the *ade6* (Figure 3C,D) and Swi6 at the *ura4* reporter, respectively (Figure 3F,G). Similar loss of silencing at the *his3-telo* locus (22) was indicated by enhanced growth of the *swi6*$^{3K\rightarrow 3A}$ strain on plates lacking histidine (Figure 3H,I); this as accompanied by reduced localization of mutant Swi6 at the *his3-telo* locus (Figure 3J,K). The abrogation of silencing of the *mat2*-linked *ura4* reporter in the *swi6*$^{3K\rightarrow 3A}$ mutant, was indicated by lack of restoration of FOA sensitivity in *swi6Δ* strain by the *swi6*$^{3K\rightarrow 3A}$ gene (Supplementary Figure S9C,D). Furthermore, in the homothallic, efficiently switching background (*h*^{90}), the *swi6*$^{+}$ gene restored efficient switching to the *h*^{90} *swi6Δ* mutant, as indicated by enhanced iodine staining of the transformants (ref. 21; Materials and Methods; Figure 3E), but the mutant *swi6*$^{3K\rightarrow 3A}$ did not (Supplementary Data, Figure S9E).

Swi6-HP1 exhibits strong binding to siRNA-DNA hybrids *in vitro* and *in vivo*

We envisaged that the si-RNA binding by Swi6/HP1 could facilitate its recruitment to the cognate complementary sequences in the genome. We speculated that Swi6^{+} may also bind to the siRNAs as RNA-DNA hybrid. Indeed, EMSA experiment confirmed the binding of Swi6 to the 'D-For' RNA-DNA hybrid with very high affinity (Kd=0.17$\pm$0.05μM; Figure 4B), while Swi6$^{3K\rightarrow 3A}$ showed no binding (Figure 4A).

To confirm the RNA-DNA binding by Swi6/HP1 *in vivo*, we checked the susceptibility of heterochromatin formation to RNaseH, which cleaves the RNA-DNA hybrids. Indeed, overexpression of *rnh1* (RNaseH1) but not *rnh201* (RNaseH2 subunit A) in cells expressing GFP-tagged Swi6 caused delocalization of GFP-Swi6 as visualized by confocal microscopy (Figure 4C,D; Supplementary Figure S10). This was accompanied by loss of silencing of the

ade6 reporter at the *otrR* repeat of *cenI* and *mat3* locus, as indicated by growth of pink colonies on adenine limiting medium (Figure 4E-4G; Supplementary Figure S11A-11C), while ChIP assay showed reduced localization of Swi6 at both loci (Figure 4H; Supplementary Figure S11D) in cells transformed with *rnh1* but not by empty vector or *rnh201*. Overexpression of *rnh1* also elicited loss of silencing at the *mat2P* locus in a strain containing a stable *mat1M* locus, as indicated by iodine staining assay (21; Materials and Methods; Supplementary Figure S11E-G) and ChIP assay showed loss of Swi6 localization at the *mat2*-linked *ura4* reporter (Supplementary Figure S11H). These results confirm that localization of Swi6 to the *dg-dh* repeats (and spreading, as in case of *mat2::ura4* and *mat3:: ade6*) is dependent on its binding to RNA-DNA hybrid.

To check whether Swi6/HP1 binds and protects the *dg-dh* repeats existing as RNA-DNA hybrid *in vivo*, we performed the DNA RNA Immunoprecipitation (DRIP) experiment (32) using the monoclonal antibody against RNA-DNA hybrid. Chromatin samples that were immunoprecipitated with the antibody were treated with RNaseH and the RNaseH-resistant *dh* regions were quantitated by real time PCR (light grey bars; Figure 5A). The treatment with RNaseH *in vitro* caused a 2-fold reduction of the *dh* signal in $swi6^{3K\rightarrow 3A}$ as compared to the $swi6^{+}$ cells and also in *swi6Δ* cells transformed with the empty vector and $swi6^{3K\rightarrow 3A}$, as compared to $swi6^{+}$ gene (Figure 5A), supporting a role of Swi6 in binding and protecting the RNA-DNA hybrid.

Heterochromatin recruitment of Swi6-HP1 depends on siRNA generation

It was shown earlier that the level of siRNAs originating from the outer repeats of centromere is reduced in *cid14Δ* mutant, which is defective in polyadenylation of RNAs and their subsequent degradation by the exosome pathway, resulting in accumulation of pre-mRNAs and ribosomal RNAs (33). Results of real time PCR confirm that the level of Swi6 recruitment to the *dh* repeats is reduced in the *cid14Δ* mutant (Figure 5B). As the siRNAs' level is reduced in *cid14Δ* mutant, we speculate that siRNA binding may be the first step in recruitment of Swi6 to the *dg-dh*

repeats. This is in agreement with the reported dependence of Swi6-HP1 localization to centromeres on Dicer (34).

siRNA-DNA-Swi6 complex binds specifically to Me2-K9-H3

Earlier, the "Cen100" RNA (corresponding to thymidylate kinase gene) was shown to compete with Me2/Me3/-K9-H3 for Swi6 binding (19). We performed EMSA assay to determine whether the binding of Swi6 to the D-For RNA/DNA hybrid affected its binding to the Me2/Me3-K9-H3 peptide. Surprisingly, we observed a dose-dependent super-shift in the binding pattern of Swi6 in presence of Me2-K9-H3 (Figure 5C, M2) and to a lesser extent with Me3-K9-H3 peptide (Figure 5C, M3) but not with the unmethylated H3 peptide (Figure 5C, M0). Thus, the high affinity binding of Swi6 to the *dg-dh* specific RNA/DNA hybrid facilitates the binding to Me2-K9-H3 rather than competing against it. We got similar results with D For RNA indicating that Swi6 binding with Cen100 is different from D For RN (not shown).

Mutant Swi6$^{3K\rightarrow3A}$ shows shift in binding to euchromatic transcripts

Next we asked whether the loss of binding by Swi6$^{3K\rightarrow3A}$ mutant protein to the siRNAs generated from the *dg-dh* repeats may shift its binding to non-heterochromatic RNAs. We compared the regions with >2-fold increase in siRNA enrichment in mutant with the regions exhibiting >2 fold increase in WT, for the enrichment of Swi6, RNA PolII, Me3-K36-H3, from the ChIP-seq signals and the gene expression values. As expected, siRNAs bound to the Swi6$^{3K\rightarrow3A}$ protein were found to map to regions that show less association with Swi6^{+} (Figure 6). Interestingly, mutant-bound siRNAs mapped to regions showing greater enrichment of RNA PolII and Me3-K36-H3, which is associated with transcription elongation (35), as well as higher gene expression values, implying the loss of specificity of Swi6$^{3K\rightarrow3A}$ for *dg-dh* siRNAs and a general non-specific shift towards abundantly transcribing regions (Figure 6). Interestingly, a similar shift occurs in case of the 'Cen100' RNA (SPAC15E1.04; Supplementary Figure S12).

DISCUSSION

Here, we report that Swi6/HP1 displays a hierarchy of binding affinities to RNA and DNA sequences: i) weak affinity to the euchromatic transcripts like "Cen100", ii) moderate affinity with sequence specificity towards the ss-siRNAs corresponding to the *dg-dh* repeats, which is >15-fold stronger than that towards 'Cen100' and iii) strongest binding (over 220-fold) to the *dg-dh* sequences as RNA-DNA hybrid. This property of Swi6/HP1 governs its localization, in a *cis*-acting manner, to the cognate sites in DNA and helps to initiate heterochromatin formation and silencing. This mechanism of Swi6/HP1 recruitment is independent of histone methylation but dependent on RNAi pathway. We speculate that the siRNAs bound to Swi6/HP1 may be chaperoned to the *dg-dh* sequences by sequence complementarity. The cognate sites may be unwound due to transcription by RNA PolII and/or Rdp1 and/or association with the RITS complex, which exists as a RNA-DNA hybrid and is susceptible to RNaseH1 (36). This window of opportunity may allow Swi6 trafficking from the RNA-bound to the RNA-DNA-bound form. The stronger binding may help to lock Swi6 stably to RNA-DNA hybrid rendering it resistant to RNaseH. Subsequently, the Swi6-bound Clr4 may initiate lysine methylation, as proposed earlier (6,7), initiating heterochromatin spreading by further binding of Swi6/HP1 to Me2-K9-H3 (Figure 7). This scenario can explain the apparent discrepancy between the delocalization of Swi6 upon overexpression of RNaseh1 and protection of RNA-DNA hybrids by Swi6. Independent support comes from the enrichment of the *dg-dh* and *cenH* regions as RNA-DNA hybrids (37). The Swi6$^{3K\rightarrow 3A}$ protein's binding sites may shift from the *dg-dh* repeats to the euchromatin regions because of low affinity promiscuous RNA binding property.

Among chromodomain proteins, Chp1 binds to centromeric nc RNA, while Clr4 binding depends on Lys9-Methylated H3 (38). We suggest that Chp1 may also bind more strongly to ncRNA-DNA hybrid than to ncRNA. Because of evolutionary conservation, the metazoan orthologs of Swi6/HP1 may effect heterchromatinization of centromeric repeats by similar mechanism.

Interestingly, the chromo- and hinge domains of Swi6/HP1 are conserved between *S. pombe* and higher eukaryotes. Furthermore, the centromeres in higher eukaryotes also contain repeat sequences. Thus, the variations in the sequences in these domains in the metazoan orthologs of Swi6-HP1 may allow for their species-specific binding to the repeat-specific siRNAs and thereby regulate their binding to the cognate sites in the genome as RNA-DNA hybrids. Given the conservation of components of RNAi and heterochromatin between the fission yeast and higher eukaryotes, the findings of this study should serve as a paradigm for understanding the connection between RNAi and heterochromatin assembly in eukaryotes.

Figure Legends

Figure 1. Sequence-specific binding of Swi6 to the *dg.dh* specific siRNAs *in vitro*.
(A) Schematic representation of the *dg* and *dh* regions in the centromere, mating type (*cenH*) and the subtelomeric genes. *tlh1* and *tlh2*. Right panel, a schematic showing siRNAs being generated from the *dg-dh* sequences. (**B**) List of si RNA sequences, reported by Reinhardt and Bartel (23), that were synthesized *in vitro*. (**C**) EMSA assay showing sequence specific binding of Swi6 to the ss D-For RNA and its competition by excess col RNA sequences B,D, E and K. Arrows indicate the complexes I, II and II between Swi6 and 'D-For' RNA. (**D**) The mutant Swi6$^{3K\rightarrow 3A}$ shows weaker binding to the D-For RNA. EMSA was performed as in (**C**). (**E**) Estimation of equilibrium binding constant (Kd) for the binding of Swi6 to D-For RNA. The data of protein concentration of Swi6 along X axis and the fraction of siRNA bound was plotted according to Ref. 25. (**F**) Schematic depiction of the domain structure of Swi6 showing the chromo-, chromo-hinge and chromoshadow domains. Alignment of the sequences of the hinge domain between HP1α, HP1β, HP1γ and Swi6, showing conservation of the lysine triplet located at residues 242-244. (**G**) SDS-PAGE showing the purified GST, GST-Swi6 and GST-Swi6$^{3K\rightarrow 3A}$. (**H**) EMSA showing that the RNA binding could be ascribed primarily to the CD and CD-hinge but not the CSD.

Figure 2.Localization of Swi6-bound siRNAs to *dg-dh* regions is abrogated in *swi6$^{3K\rightarrow 3A}$* mutant.
(A) Overall distribution of normalized small-RNA accumulation at centromeres, telomeres and mating-type locus in WT and *swi6$^{3K\rightarrow 3A}$* mutant. (**B**) Chromosomal track of RIP-seq fold change (log2 3K-3A/WT ratio) at and around centromere of chr-I at 5kb (top) and 200bp (bottom) resolutions, left telomere of chr-I (200bp resolution) and the mating locus (200bp resolution). Locus annotations are given at the bottom of the plot. Plots for other centromeres and telomeres are given in the Supplementary Figure S8. (**C**) Distributions of normalized RIP-seq signals in *swi6^{+}* and *swi6$^{3K\rightarrow 3A}$* strains for regions containing the sequence motif(s) known to interact with Swi6. The p-values were calculated using two-tailed Mann-Whitney U tests.

Figure 3. Abrogation of silencing in *swi6$^{3K\rightarrow 3A}$* mutant accompanies Swi6 delocalization at the outer repeat *otr1R* and *imr1L* regions of *cenI* and *his3-telo*. (**A**) Diagrammatic representation of

the genotype of the strain. Centromere of each fission yeast chromosome comprises of a central core (*cnt*), immediately flanked on both sides by inverted repeats (*imr*); *imr* on either side are flanked by outer repeat regions (*otr*). Reporter genes *ura4, ade6* and *his3* have been inserted into the *imr*, *otr* of centromere and telomeric repeats, respectively of chrI. The strain has a *swi6Δ* mutation and carries internal deletions in native *ura4* locus, *denoted as ura4DS/E,* and in *ade6*, denoted as *ade6DN/N*, which act as euchromatic controls in ChIP assay. (**B-D)** Derepression of the *otr1R::ade6* locus in *swi6*$^{3K\rightarrow 3A}$ mutant, showing pink/white colonies on adenine limiting plates (**B**). (**C**) ChIP assay showing reduction of Me2-K9-H3 and Swi6 at the *ade6* locus in *swi6*$^{3K\rightarrow 3A}$ mutant. Ratio of the heterochromatic *ade6* and euchromatic *ade6DN/N* locus is shown (**D**). Quantitation of data shown in (**C**). (**E**) Spotting assay showing inability of *swi6*$^{3K\rightarrow 3A}$ to restore silencing at the *imr1::ura4* locus in *swi6Δ* mutant, as indicated by growth on plates lacking uracil and lack of growth on FOA plates. (**F)**, ChIP assay showing delocalization of Swi6 from the *imr1::ura4* locus in the *swi6*$^{3K\rightarrow 3A}$ mutant, where *ura4* represents the heterochromatin and *ura4DS/E* the euchromatin. (**G**) Quantitation of the data showing enrichment of Swi6 at *ura4* versus *ura4DS/E* in (**F)**. (**H-K**) Derepression of the subtelomeric *his3* locus in *swi6*$^{3K\rightarrow 3A}$ mutant. (**H**) gives a schematic representation of the *his3* gene inserted distal to the telomeric repeats, denoted as *his3-telo* (22). (**I**) Spotting assay showing the enhanced growth of the strain having *his3-telo* reporter in presence of the *swi6*$^{3K\rightarrow 3A}$ mutation on plates lacking histidine. (**J**) ChIP assay reduced localization of mutant Swi6$^{3K\rightarrow 3A}$ at the *his3-telo* locus as compared to wt. *act1* was used as a control. (**K**) Quantitation of data shown in (**J)**.

Figure 4. Role of Swi6 binding to the *dg.dh* sequences as RNA-DNA hybrids in silencing. (A) Swi6 shows strong affinity towards the RNA-DNA hybrid. EMSA assay was performed using the hybrid between the D-For RNA and its complementary DNA sequence and purified GST-Swi6 or GST-Swi6$^{3K\rightarrow 3A}$. GST was used as control. (**B**) Data in (**A)** was plotted according to Ref. 25 to determine the equilibrium binding constant, Kd. (**C**) Delocalization of GFP-Swi6 upon overexpression of *rnh1* but not *rnh201*. Fluorescence micrographs showing the effect of overexpressed *nmt1-rnh1* and *nmt1-rnh201* on the subcellular localization of Swi6 with respect to the three heterochromatic loci, *mat*, *cen* and *tel* as distinct green foci. The strain had the native copy of *swi6* gene chromosomally tagged with GFP. Nuclei were stained with DAPI. Strains

expressing GFP-tagged Swi6 was transformed with vector alone, *rnh1* or *rnh201* under the control of *nmt1* promoter. Transformed cells were grown in the presence (+) or absence (-) of the repressor, thiamine, visualized under confocal microscope and photographed. (**D**) Microscopic data shown in (**C**) was quantified as distinct green foci pertaining to GFP-Swi6 subcellular localizing to the heterochromatic loci. Quantitation of the localization of GFP-Swi6 as three, two or one spots. (**E-H**) Overexpression of *rnh1* causes loss of silencing at the *otr1R::ade6* locus. (**E**) Schematic representation of the centromere I having an *ade6* reporter gene inserted into the *otr1R*. **(F**) The effect of expression of empty vector alone (*nmt1*), *nmt1*-GFP-*swi6* or *nmt1*-GFP-*swi6*$^{3K\rightarrow 3A}$ on the expression of the *ade6* reporter. Colors of the colonies indicated the state of the *ade6* reporter. Red colonies indicate suppression of *ade6* repression while pint/white colonies indicated repression. The data are represented in terms of percent white colonies. (**G**) Plot showing the per cent colonies giving pink/white appearance in different transformants, as shown in (**F**). **(H)** Delocalization of Swi6 from the *otr1R::ade6* locus in the transformants shown in (**F)**. Derepression of the *ade6* reporter at *otr1R* repeat as a result of overexpression of *rnh1* and *rnh201*, as observed in (**G)** was biochemically verified through Swi6 enrichment at the locus through ChIP assay.

.

Figure 5. Swi6 binds to and protects the RNA-DNA hybrid at the *dh* region *in vitro*. (**A**) ChIP assay was performed to quantitate the *dh* regions enriched with anti-RNA-DNA hybrid antibody in the indicated strains. (**B**) si RNA generation is required for Swi6 enrichment at *dh* region. ChIP assay was performed to quantitate the *dh* region associated with Swi6 in wt versus *cid14Δ* mutant. (**C**) The RNA-DNA hybrid shows preferential binding to the Me2-K9-H3. EMSA was performed to detect binding of Swi6 to the RNA-DNA hybrid (D-For RNA/DNA duplex) in presence of unmethylated (M0), Lys9 dimethylated (M2) and Lys9 trimethylated H3 (M3) peptides. No binding of the hybrid was observed with any of the peptides. To be able to observe the supershift (indicated by arrowheads) in presence of the peptides, electrophoresis had to be performed longer to run out the labeled hybrid, which also led to a diffuse signal.

Figure 6. Enrichment densities of Swi6 (A), RNA PolII (B), H3K36me3 (C), and gene expression levels (D) (FPKM) in the 5Kb genomic bins exhibiting 2-fold increase (log2

swi6$^{3K\rightarrow 3A}$/*swi6*$^{+}$ >=1, orange) or 2-fold decrease (log2 *swi6*$^{3K\rightarrow 3A}$/$^{swi6+}$ <=-1, blue) in siRNA Enrichment in *swi6*$^{3K\rightarrow 3A}$ mutant when compared with that of WT strain. P-values were calculated using Mann-Whitney U tests.

Figure 7. Speculative model about the role of nucleic acid binding function of Swi6/HP1 in initiating RNAi dependent heterochromatin assembly in Fission Yeast. Based on data described here, Swi6/HP1 binds with moderate affinity to the single stranded siRNA mapping to the dg-dh repeats. This helps it to be recruited to chromatin at the stage in the RNAi pathway where the RITS complex is already tethered through the nc transcript. This stage is sensitive to RNaseH. The recruitment of Swi6-ss-siRNA complex is facilitated by the occurrence of the complementary DNA strand in single stranded form. The high affinity binding of Swi6 to the RNA-DNA hybrid helps to retain Swi6 and makes its localization RNaseH-resistant. Subsequent recruitment of the CLRC complex helps to methylate the histones and promote formation of a more stable ternary complex with Swi6-HP1.

REFERENCES

1. Jenuwein, T. and Allis, C.D. Translating the histone code. (2001) *Science*. **293**, 1074-1080.
2. Grewal,S.I and Jia,S. (2007) Heterochromatin Revisited. *Nat. Genet.* **8**, 35-46.
3. Nakayama,J., Rice,J.C., Strahl,B.D., Allis,C.D. and Grewal,S.I. (2001) Role of histone H3 lysine 9 methylation in epigenetic control of heterochromatin assembly. *Science* **292**, 110-113.
4. Bannister,A.J., Zegerman,P., Partridge,J.F., Miska,E.A., Thomas,J.O., Allshire,R.C. and Kouzarides,T. (2001) Selective recognition of methylated lysine 9 on histone H3 by the HP1 chromo domain. *Nature* **410**, 120-124.
5. Canzio,D., Chang,E.Y., Shankar,S., Kuchenbecker,K.M., Simon,M.D., Madhani, H.D., Narlikar, G.J. and Al-Sady B. (2011) Chromodomain-mediated oligomerization of HP1 suggests a nucleosome-bridging mechanism for heterochromatin assembly. *Mol. Cell*. **41**, 67-81.
6. Haldar, S., Saini,A., Nanda,J.S., Saini,S. and Singh,J. (2011) Role of Swi6/HP1 self-association-mediated recruitment of Clr4/Suv39 in establishment and maintenance of heterochromatin in fission yeast. .*J. Biol. Chem.* **286**, 9308-9320.
7. Al-Sady,B., Madhani,H.D. and Narlikar,G.J. (2013) Division of labor between the chromodomains of HP1 and Suv39 methylase enables coordination of heterochromatin spread. *Mol. Cell.* **51**, 80-91.
8. Volpe,T.A., Kidner,C., Hall,I.M., Teng,G., Grewal,S.I. and Martienssen,R.A. (2002) Regulation of heterochromatic silencing and histone H3 lysine-9 methylation by RNAi. *Science*. **297**, 1833-1837.
9. Holoch,D. and Moazed,D. (2015) Small-RNA loading licenses Argonaute for assembly into a transcriptional silencing complex. *Nat. Struct. Mol. Biol.* **22**, 328-35.
10. Verdel,A., Jia,S., Gerber,S., Sugiyama,T., Gygi,S., Grewal,S.I. and Moazed. D. (2004) RNAi-mediated targeting of heterochromatin by the RITS complex. *Science.* **303**, 672-676.
11. Sugiyama,T., Cam,H., Verdel,A., Moazed,D. and Grewal, S.I. (2005) RNA-dependent RNA polymerase is an essential component of a self-enforcing loop coupling heterochromatin assembly to siRNA production. *Proc. Natl. Acad. Sci. U.S.A.* **102**, 152-157.
12. Martienssen,R. and Moazed, D. (2015) RNAi and Heterochromatin Assembly. *Cold Spring Harb. Perspect. Biol*. **7**, a019323.
13. Bayne,E.H., White,S.A., Kagansky,A, Bijos,D.A., Sanchez-Pulido,L., Hoe,K.L., Kim,D. U, Park,H.O, Ponting,C.P, Rappsilber,J, and Allshire, R.C.(2010) Stc1: a critical link between RNAi and chromatin modification required for heterochromatin integrity. *Cell*. **140**, 666-677.
14. Hong, E-J., Villén,J., Gerace,E.L., Gygi,S.P. and Moazed,D. (2005) A Cullin E3 Ubiquitin Ligase Complex Associates With Rik1 and the Clr4 Histone H3-K9 Methyltransferase and Is Required for RNAi-mediated Heterochromatin Formation. *RNA Biol.* **2**, 106-111.
15. Grewal,S.I. and Klar,A.J. (1997) A recombinationally repressed region between mat2 and mat3 loci shares homology to centromeric repeats and regulates directionality of mating-type switching in fission yeast. *Genetics* **146**,1221-1238.

16. Hansen,K.R., Ibarra,P.T. and Thon,G. (2006) Evolutionary-conserved telomere-linked helicase genes of fission yeast are repressed by silencing factors, RNAi components and the telomere-binding protein Taz1. *Nucleic Acids Res*. **34**, 78-88.
17. Akhtar,A., Zink,D. and Becker,P.B. (2000) Chromodomains are protein-RNA interaction modules. *Nature*. **407**,405-409.
18. Muchardt,C., Guilleme,M., Seeler,J.S., Trouche,D., Dejean,A. and Yaniv, M. (2002) Coordinated methyl and RNA binding is required for heterochromatin localization of mammalian HP1alpha. *EMBO Rep.* **3**, 975-981.
19. Keller,C., Adaixo,R., Stunnenberg,R., Woolcock,K.J., Hiller,S. and Bühler, M. (2012) HP1(Swi6) mediates the recognition and destruction of heterochromatic RNA transcripts. *Mol. Cell.* **47**, 215-227.
20. Nakama,M., Kawakami,K., Kajitani,T., Urano,T. and Murakami,Y. (2012) DNA-RNA hybrid formation mediates RNAi-directed heterochromatin formation. *Genes Cells*. **17**, 218-233.
21. Moreno,S., Klar,A. and Nurse,P. (1991) Molecular Genetic Analysis of Fission Yeast Schizosaccharomyces pombe. *Methods Enzymol.***194**,795-823.
22. Cooper,J.P., Nimmo,E.R., Allshire,R.C.and Cech,T.R. (1997) Regulation of telomere length and function by a Myb-domain protein in fission yeast. *Nature*. **385**, 744-747.
23. Reinhart,B.J. and Bartel,D.P. (2002) Small RNAs correspond to centromere heterochromatic repeats. *Science*. **297**, 1831.
24. Djupedal,I., Kos-Braun,I.C., Mosher,R.A., Söderholm,N., Simmer,F., Hardcastle,T.J., Fender,A., Heidrich,N., Kagansky,A., Bayne,E., Wagner,E.G., Baulcombe,D.C., Allshire,R.C. and Ekwall,K. (2009) Analysis of small RNA in fission yeast; centromeric siRNAs are potentially generated through a structured RNA. *EMBO J.* **28**, 3832-3844.
25. Heffler,M.A., Walters,R.D. and Kugel,J.F. (2012) Using electrophoretic mobility shift assays to measure equilibrium dissociation constants: GAL4-p53 binding DNA as a model system. *Biochem. Mol, Biol. Educ.* **40**, 383-387.
26. Boguslawski,S.J., Smith,D.E., Michalak,M.A., Mickelson,K.E., Yehle,C.O., Patterson, W.L. and Carrico,R.J. (1986) Characterization of monoclonal antibody to DNA-RNA and its application to immunodetection of hybrids. *J. Imm. Methods.* **89**, 123-130.
27. Livak,K.J. and Schmittgen,T.D. (2001) Analysis of Relative Gene Expression Data Using Real-Time Quantitative PCR and the 22DDC T Method. *Methods* **25**, 402–408.
28. Bailey,T.L., Boden,M., Buske,F.A., Frith,M., Grant,C.E., Clementi,L., Jingyuan Ren,J., Li,W.W. and Noble, W.S. (2009) MEME SUITE: tools for motif discovery and searching. *Nucleic Acids Res.* 37(Web Server issue): W202–W208 (2009).
29. Kanoh,J., Sadaie,M., Urano,T. and Ishikawa,F. (2005) Telomere binding protein Taz1 establishes Swi6 heterochromatin independently of RNAi at telomeres. *Curr. Biol.* **15**, 1808-1819.
30. Pidoux,A.L., Uzawa,S., Perry,P.E., Cande,W.Z. and Allshire,R.C. (2000) Live analysis of lagging chromosomes during anaphase and their effect on spindle elongation rate in fission yeast. *J. Cell. Sci.* **113**, 4177-4191.
31. Allshire,R.C., Javerzat,J-P., Redhead,N.J. and Cranston, G. (1994) Position effect variegation at fission yeast centromeres. *Cell* **76**, 157-169.

32. Ginno,P.A.,Lott,PL., Christensen.H.C., Korf,I. and Chedin,F. (2012) R-loop formation is a distinctive characteristic of unmethylated human CpG island promoters. *Mol. Cell.* **45**, 977-986.
33. Wang,S.W., Stevenson,A.L., Kearsey,S.E., Watt,S. and Bähler,J. (2008) Global Role for Polyadenylation-Assisted Nuclear RNA Degradation in Posttranscriptional Gene Silencing. *Mol. Cell. Biol.* **28**, 656-665.
34. Woolcock, K.J., Gaidatzis,D., Punga,T. and Buhler,M. (2011) Dicer associates with chromatin to repress genome activity in Schizosaccharomyces pombe. *Nat. Str. Mol. Biol.* **18**, 94-99.
35. Morris,S.A., Shibata,Y., Noma,K-I., Tsukamoto,Y., Warren,E., Temple,B., Grewal,S.I.S. and Strahl, B. D. (2005) Histone H3 K36 methylation is associated with transcription elongation in *Schizosaccharomyces pombe*. *Eukaryotic Cell*. **4**, 1446-1454.
36. Shimada,Y., Mohn,F. and Bühler,M. (2016) The RNA-induced transcriptional silencing complex targets chromatin exclusively via interacting with nascent transcripts. *Genes Dev*. **30**, 2571-2580.
37. Dutrow,N., Nix,D.A., Holt,D., Milash,B., Dalley,B., Westbroek,E., Parnell,T.J. and Cairns,B.R. (2008) Dynamic transcriptome of *Schizosaccharomyces pombe* shown by RNA-DNA hybrid mapping. *Nat. Genet.* **40**, 977–986.
38. Ishida,M., Shimojo,H., Hayashi,A., Kawaguchi,R., Ohtani,Y., Uegaki,K., Nishimura,Y. and Nakayama,J. (2012) Intrinsic nucleic acid-binding activity of Chp1 chromodomain is required for heterochromatic gene silencing. *Mol. Cell.* **47**, 228-241.

ACKNOWLEDGEMENTS

We thank J.I. Nakayama for the gift of the plasmids expressing rnh1 and rnh201 and R. Allshire for the strains with ade6 and ura4 reporters at the cenI locus.

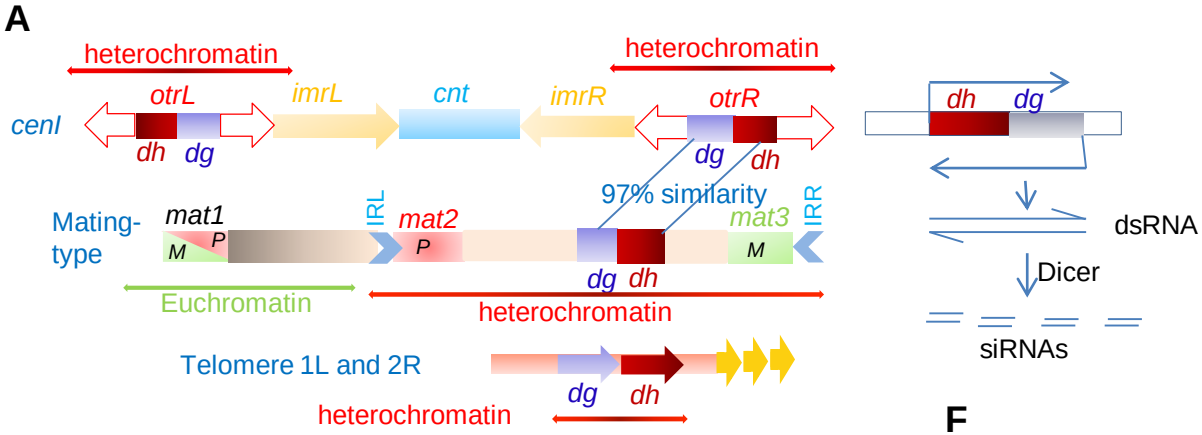
.

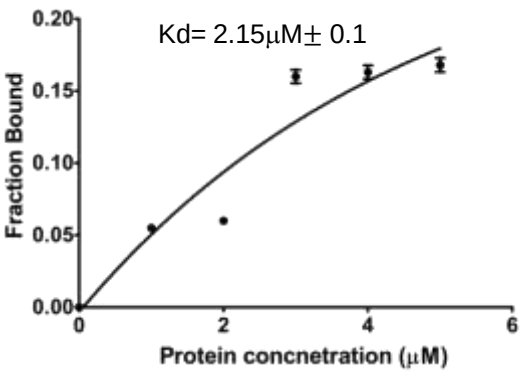
FUNDING

The project was supported by extramural funding by the Department of Biotechnology, New Delhi, India and Department of Science and Technology, New Delhi, India and intramural funding by Council of Scientific and Industrial Research, New Delhi, India. K.R.C was supported by UGC fellowship.

Conflict of Interest Statement:

The authors declare no conflict of interest.

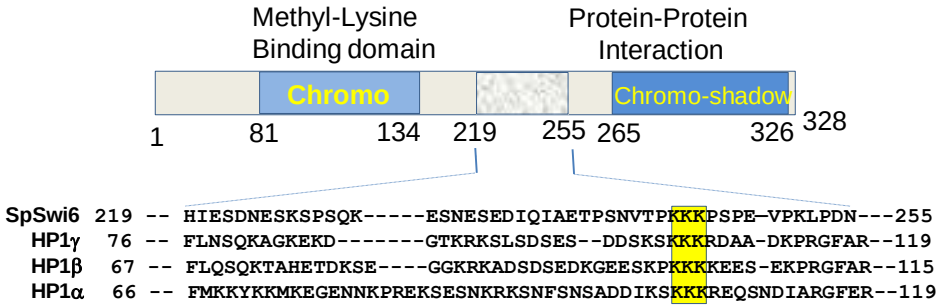




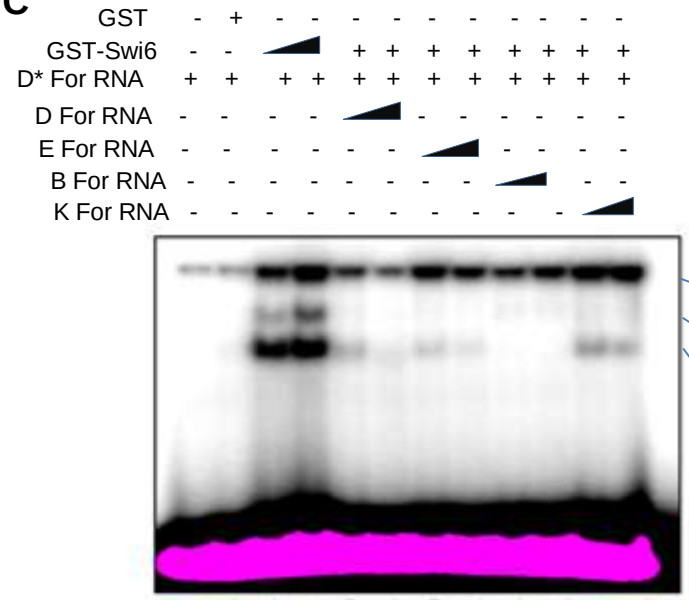
B

Name	Sequence	Location
B	AATGCGGAGTAAGGCTAATCACGGTA	cenI, cenII, cenIII
D	TGGATTAAGGAGAAGCGGTA	cenI, cenII, cenIII
E	ACAAGTGATAAGAGTAGGTGT	cenI, cenII, cenIII, mat
K	TGTGCCATATCCATGCTGTGTCCA	cenIII

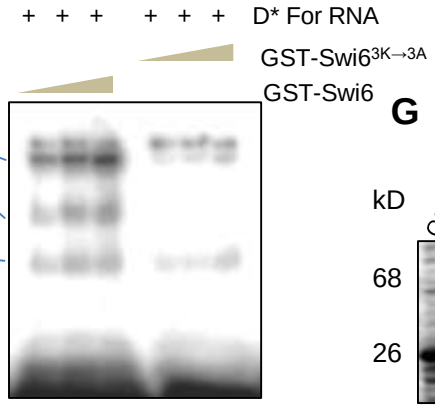
F



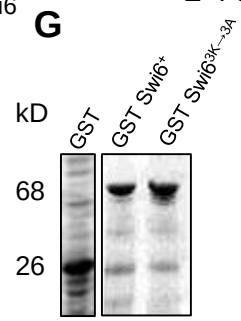
C



D



G



H

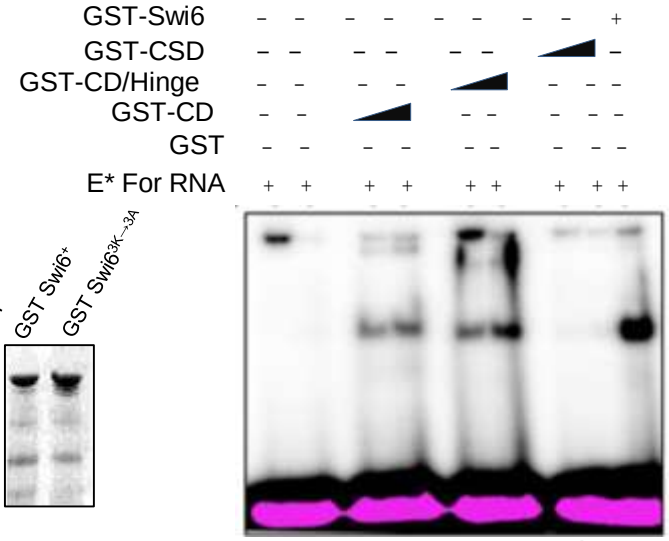


Figure 1

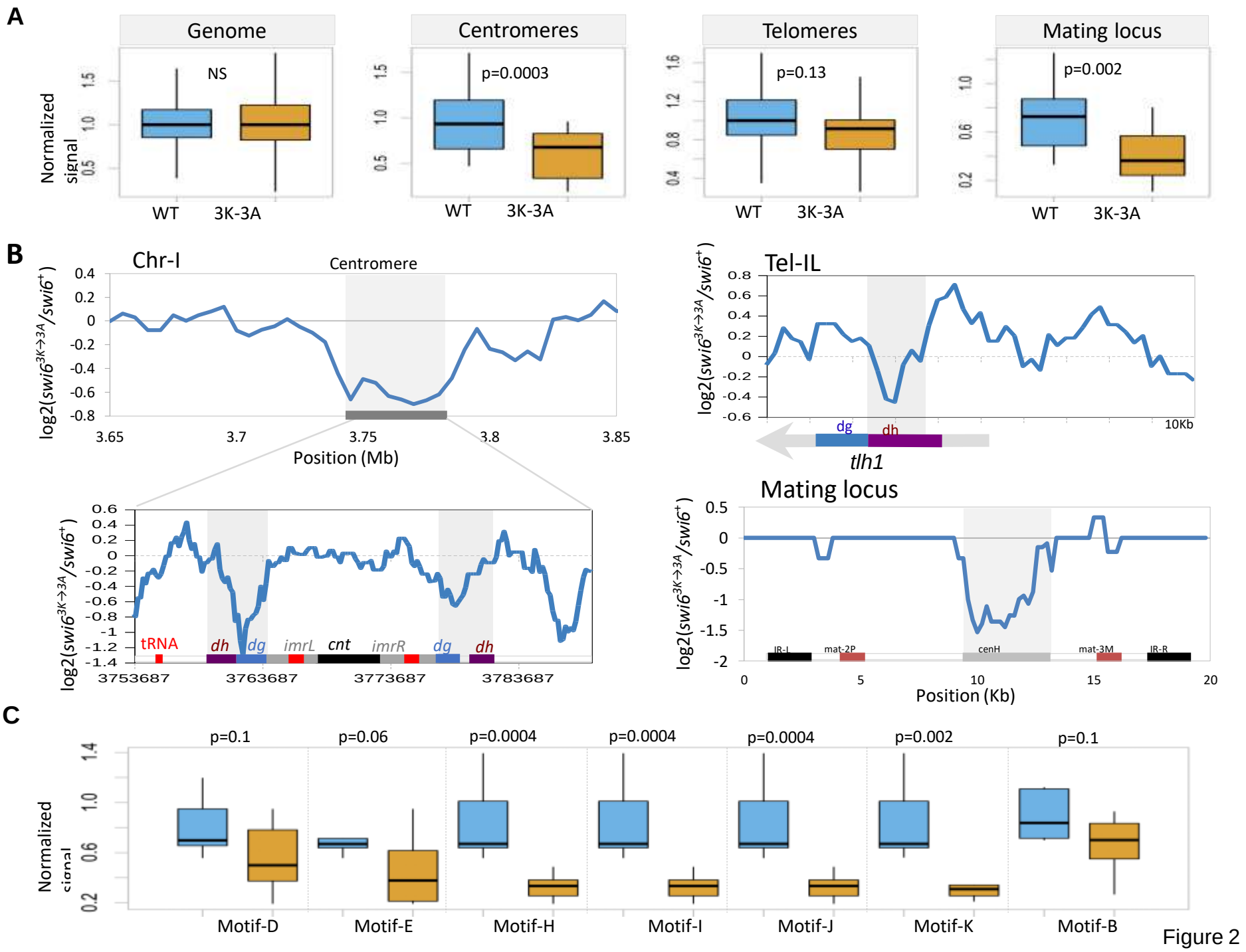

A
Genome
Centromeres
Telomeres
Mating locus
Normalized signal
NS
p=0.0003
p=0.13
p=0.002
WT
3K-3A
B
Chr-I
Centromere
log2(swi6^3K→3A/swi6^+)
Position (Mb)
Tel-IL
dg
dh
tlh1
10Kb
tRNA
dh
dg
imrL
cnt
imrR
dg
dh
Mating locus
IR-L
mat-2P
cenH
mat-3M
IR-R
Position (Kb)
C
p=0.1
p=0.06
p=0.0004
p=0.0004
p=0.0004
p=0.002
p=0.1
Motif-D
Motif-E
Motif-H
Motif-I
Motif-J
Motif-K
Motif-B

Figure 2

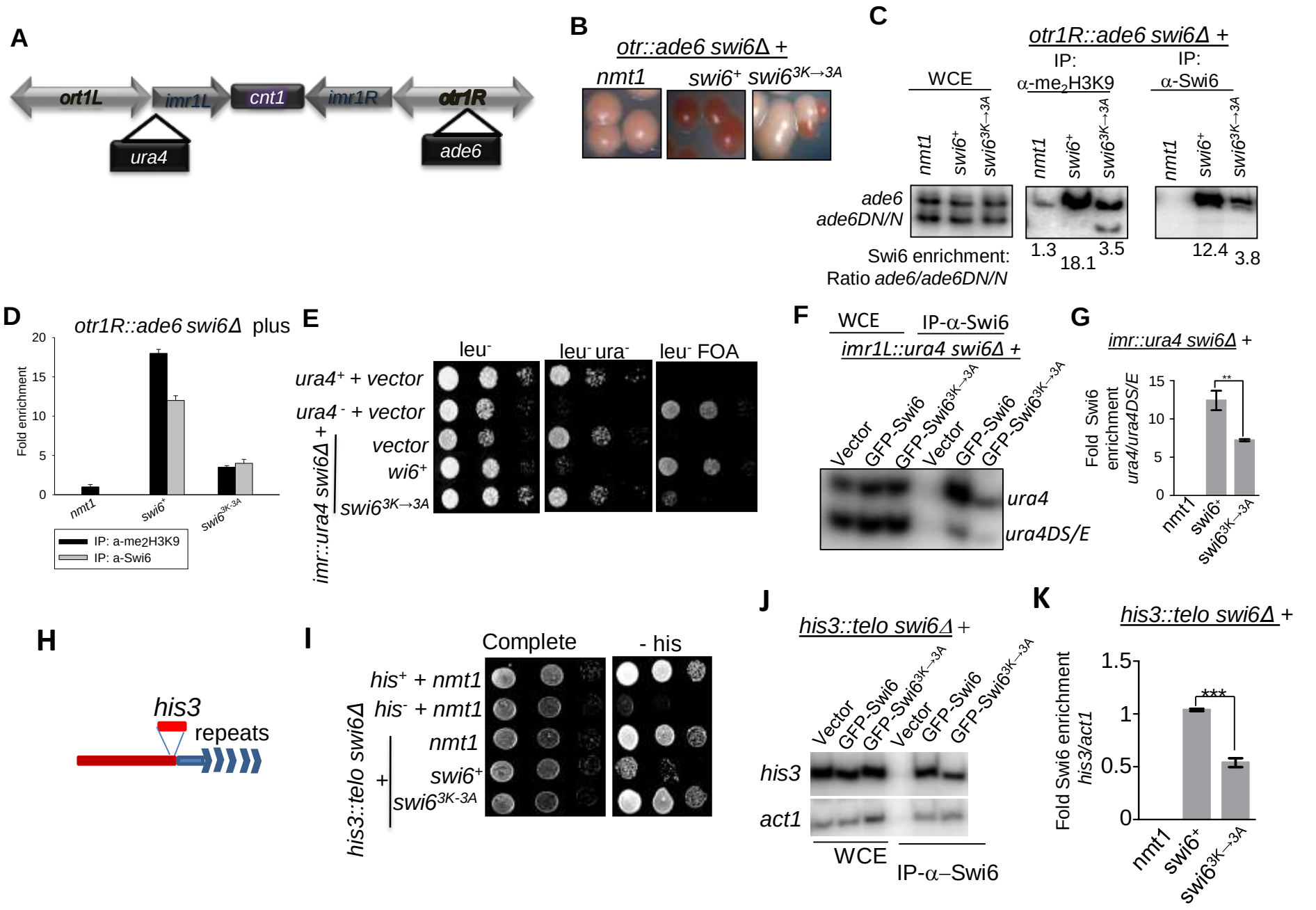

A
ort1L
imr1L
cnt1
imr1R
otr1R
ura4
ade6
B
otr::ade6 swi6Δ +
nmt1
swi6+
swi6 3K→3A
C
otr1R::ade6 swi6Δ +
WCE
IP: α-me2H3K9
IP: α-Swi6
nmt1
swi6+
swi6 3K→3A
ade6
ade6DN/N
Swi6 enrichment:
Ratio ade6/ade6DN/N
1.3
18.1
3.5
12.4
3.8
D
otr1R::ade6 swi6Δ plus
Fold enrichment
20
15
10
5
0
nmt1
swi6+
swi6 3K-3A
IP: a-me2H3K9
IP: a-Swi6
E
leu-
leu- ura-
leu- FOA
ura4+ + vector
ura4- + vector
vector
wi6+
swi6 3K→3A
imr::ura4 swi6Δ +
F
WCE
IP-α-Swi6
imr1L::ura4 swi6Δ +
Vector
GFP-Swi6
GFP-Swi6 3K→3A
Vector
GFP-Swi6
GFP-Swi6 3K→3A
ura4
ura4DS/E
G
imr::ura4 swi6Δ +
Fold Swi6 enrichment ura4/ura4DS/E
15
10
5
0
**
nmt1
swi6+
swi6 3K→3A
H
his3
repeats
I
Complete
- his
his+ + nmt1
his- + nmt1
nmt1
swi6+
swi6 3K-3A
his3::telo swi6Δ
J
his3::telo swi6Δ +
Vector
GFP-Swi6
GFP-Swi6 3K→3A
Vector
GFP-Swi6
GFP-Swi6 3K→3A
his3
act1
WCE
IP-α–Swi6
K
his3::telo swi6Δ +
Fold Swi6 enrichment his3/act1
1.5
1
0.5
0
***
nmt1
swi6+
swi6 3K→3A

Figure 3

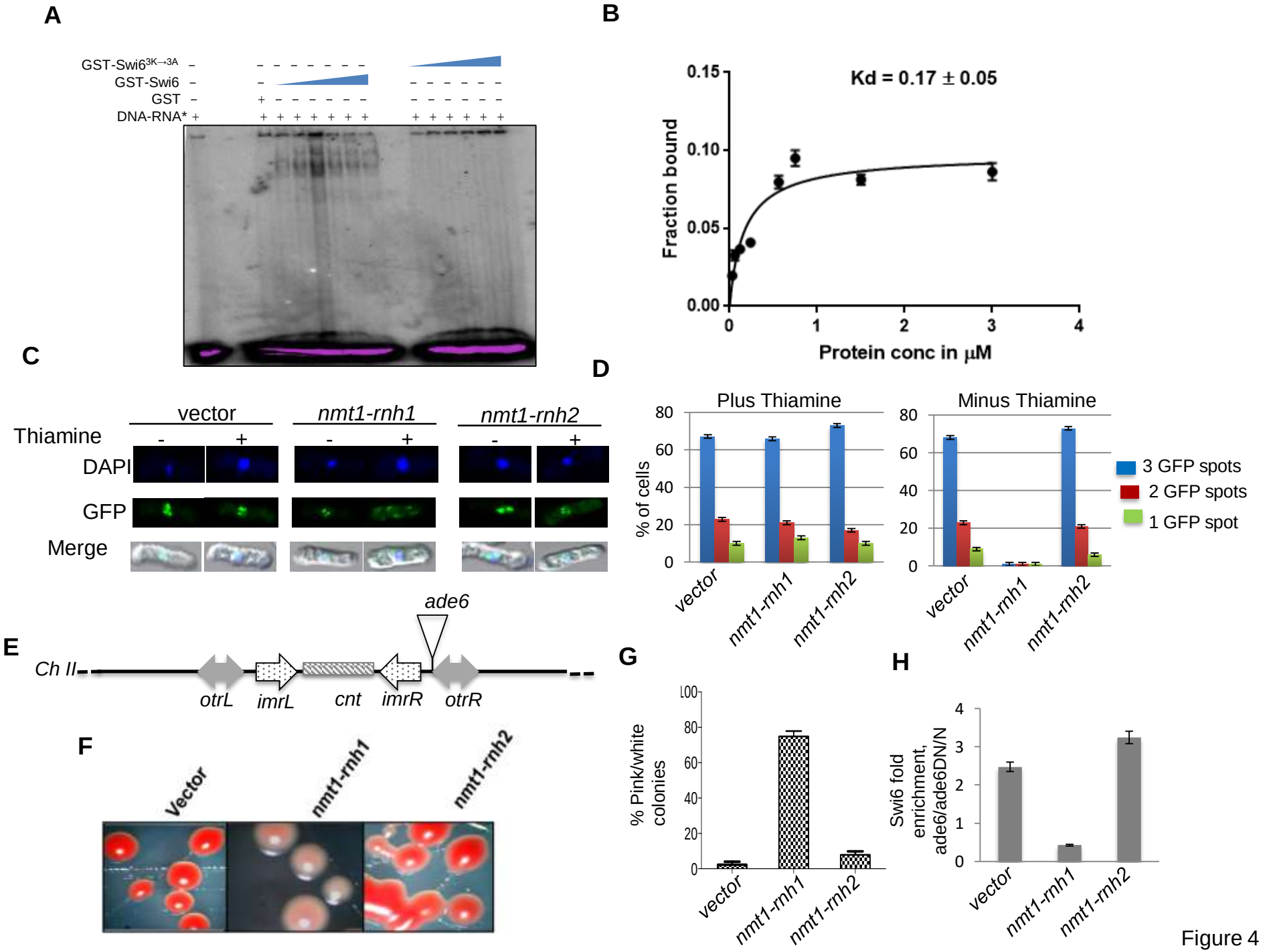

A
GST-Swi6$^{3K\rightarrow 3A}$
GST-Swi6
GST
DNA-RNA*
B
Kd = 0.17 ± 0.05
Fraction bound
Protein conc in μM
C
vector
nmt1-rnh1
nmt1-rnh2
Thiamine
DAPI
GFP
Merge
D
Plus Thiamine
Minus Thiamine
% of cells
3 GFP spots
2 GFP spots
1 GFP spot
E
ade6
Ch II
otrL
imrL
cnt
imrR
otrR
F
Vector
nmt1-rnh1
nmt1-rnh2
G
% Pink/white colonies
H
Swi6 fold enrichment, ade6/ade6DN/N

Figure 4

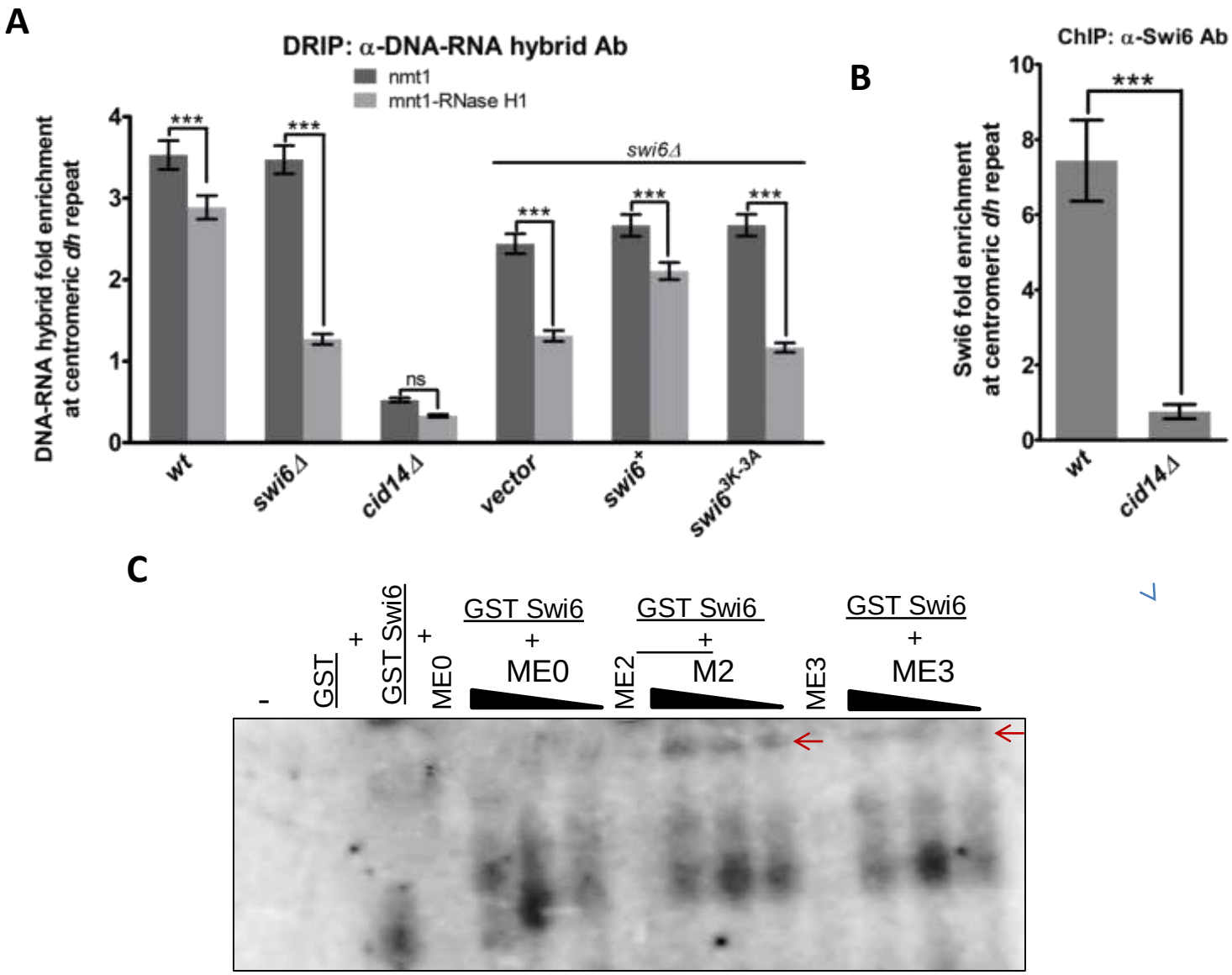

A
DRIP: α-DNA-RNA hybrid Ab
nmt1
mnt1-RNase H1
DNA-RNA hybrid fold enrichment at centromeric dh repeat
swi6Δ
***
ns
wt
swi6Δ
cid14Δ
vector
swi6*
swi6 3K-3A
B
ChIP: α-Swi6 Ab
Swi6 fold enrichment at centromeric dh repeat
wt
cid14Δ
C
-
GST
+
GST Swi6
+
ME0
GST Swi6
+
ME0
ME2
GST Swi6
+
M2
ME3
GST Swi6
+
ME3

Figure 5

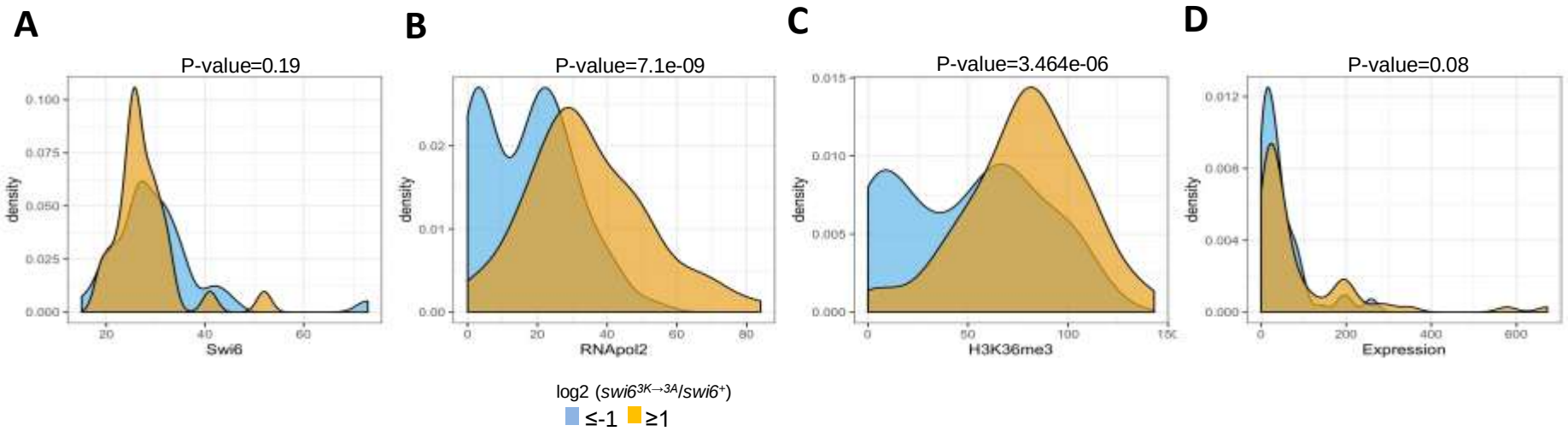

A
P-value=0.19
density
Swi6
B
P-value=7.1e-09
density
RNApol2
C
P-value=3.464e-06
density
H3K36me3
D
P-value=0.08
density
Expression
log2 (swi6^3K→3A/swi6+)
≤-1
≥1

Figure 6

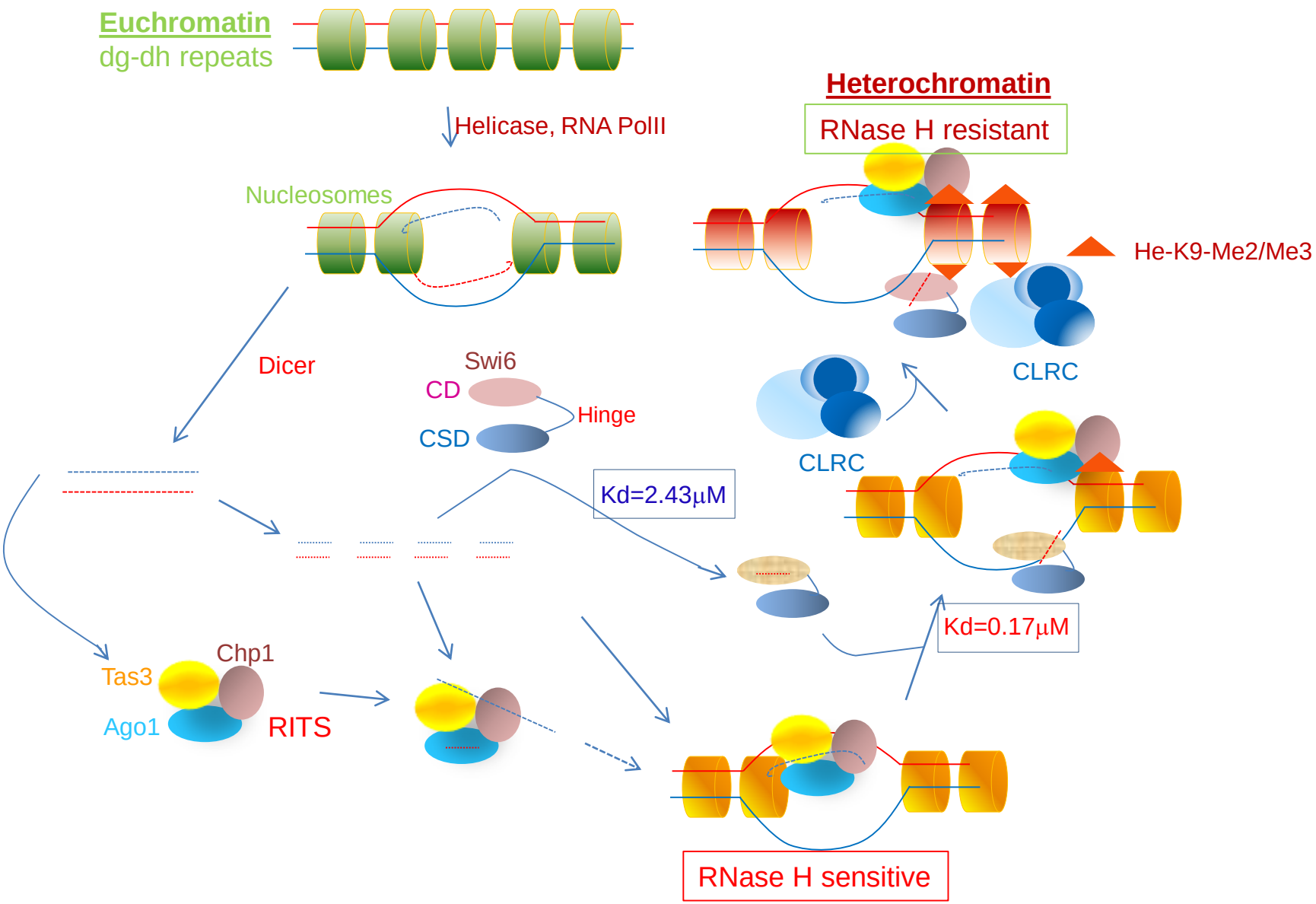

Euchromatin
dg-dh repeats
Helicase, RNA PolII
Heterochromatin
RNase H resistant
Nucleosomes
He-K9-Me2/Me3
Dicer
Swi6
CD
Hinge
CSD
CLRC
CLRC
Kd=2.43μM
Kd=0.17μM
Chp1
Tas3
Ago1
RITS
RNase H sensitive

Figure 7

SUPPLEMENTARY DATA

Supplementary Methods

An *in vitro* transcription protocol was adapted from Wilusz lab protocol (http://csu-cvmbs.colostate.edu/academics/mip/wilusz-lab/Pages/lab-protocols.aspx) with suitable modifications. The reaction was carried out by assembling nuclease free water, 1µl of DNA template (1µg/µl conc.), 4 µl of ribonucleotides mix (2.5mM each of ATP, GTP, CTP, and UTP) to a final concentration of 0.5mM, 2 µl of 10X RNA polymerase buffer and 1 µl of T7 RNA polymerase enzyme (Cat. no. AM2178, Ambion) were mixed and diluted to 20 µl with nuclease-free water and incubated at 37˚C for exactly 3 hours. The reaction was subjected to Phenol (pH 5.2)/Chloroform extraction, followed by ethanol precipitation.

a) Polyacrylamide gel elution of *in vitro* transcribed RNAs

To obtain ~500 ng of *in vitro* transcribed RNA, products of 10 reactions were pooled. The pellet obtained after precipitation was resuspended in 20 µl 1X formamide dye and loaded on 10- 20% polyacrylamide/8M urea gels depending on the size of the required transcript. Subsequently, the gel was stained with Syber gold nucleic acid stain (Cat. no. S11494, Invitrogen Life Technologies) for visualization under UV transilluminator. The RNA bands were located along the RNA marker (Cat. no. AM7778, Ambion), excised and frozen at -20ºC for 20 min. Thereafter, the gel pieces were crushed and resuspended in 1 ml of Crush and Soak buffer. The tubes were incubated at 37˚C and 700 rpm for 12-14 hrs on a thermo mixer. The supernatant was then subjected to ethanol precipitation.

b) Dephosphorylation of 5'ends of RNA

The ssRNA generated by *in vitro* transcription was dephosphorylated before labeling. The dephosphorylation reaction was set up with 500ng of ssRNA, Alkaline phosphatase calf intestinal enzyme, CIP (Cat. no. M0290, New England BioLabs Inc., USA), 1X CIP buffer, RNase inhibitor (Cat. no. AM2696, Ambion) and nuclease-free water up to 25 µl. The reaction was incubated at 37˚C for 1 hour, phenol/chloroform extracted and ethanol precipitated.

c) 5'-end labeling of RNA

To perform 5'-end labeling of dephosphorylated RNA, a reaction mix containing 500ng of dephosphorylated RNA, T4 Polynucleotide kinase (PNK) enzyme (Cat. no. M0201S, New England BioLabs Inc., USA), 1X T4 PNK buffer, RNase inhibitor, [γ-^{32}P]-dATP (>3300 Ci/mmol) from Board of Radiation and Isotope Technology (BARC), Mumbai, and nuclease-free water up to 50µl was prepared and incubated at 37ºC for 1hr. Unincorporated isotope was removed by passing through microspin G-50 columns (Cat. no. GE27-5330-01, GE Healthcare Ltd.). The eluent was stored at -70ºC and used for performing RNA-EMSA.

e) RNA-DNA hybrid formation

For producing RNA-DNA hybrids, the ssRNA was annealed with equimolar amount of complementary DNA oligo in presence of 1X transcription buffer. The reactions were incubated at 85ºC for 5 minutes to remove the secondary structures and then subjected to gradual cooling to 55ºC. The reaction was incubated at 55ºC overnight for annealing and checked on 10% native TBE gel before using for the EMSA experiments.

f) Synthesis of long RNAs

For *in vitro* synthesis of 'Revcen' RNA, a long centromeric sequence shown in Djupedal *et al* (24) was PCR amplified and cloned into the vector Litmus-28i. The clone containing the DNA sequence complementary to '*RevCen'* RNA was digested with *Hin*dIII to obtain a linear plasmid. The linearized plasmid was used as a template for *in vitro* transcription, as described earlier. The *in vitro* transcribed RNA was gel purified and subjected to alkaline phosphatase treatment followed by 5'end-labeling with [γ-^{32}P] dATP (10 µCi/µl, specific activity 3300 Ci/mmol) and used for EMSA experiments. The 'Cen100' transcript, as reported by Keller *et al* (SPAC15E1.04; ref. 19), was also synthesized and labeled by the same protocol. The radio-labelled oligonucleotides were purified by passing through a BioGel P-6 column (Biorad Inc.).

RIP-Seq Analysis

Small RNA was isolated from immunoprecipitated samples from cells expressing GFP-Swi6^{+}, GFP-Swi6$^{3K\rightarrow 3A}$ and TAP-Tas3. Cells from 500 ml cultures were harvested at log-phase (OD595= 0.5-0.7) were washed once with 10 ml ice-cold water and once with ice-cold STOP

buffer (0.5% SDS, 5mM EDTA pH 8, 100 µg proteinase K) at 4°C. Cells were resuspended in (one-fifth the culture volume) HB buffer (25mM MOPS pH 7.2, 15mM MgCl2, 15mM EGTA, 1mM DTT, 1% Triton-X100, 10% glycerol) supplemented with PMSF (to a final concentration of 1mM) and protease inhibitor cocktail (PIC) (at 1:100 dilution) and dispensed as 1 ml aliquots into bead beater tubes. Ice-cold zirconium or acid-washed RNase-free glass beads were added upto 500µl mark, given 10-12 pulses of 1 min each alternated with 5 min incubations on ice. The lysate was centrifuged at 13500 rpm at 4°C for 1 hr. The clarified supernatants were removed to fresh tubes and mixed with 100% glycerol to a final concentration of 10%. The expression of the protein of interest can be checked at this point by Western analysis (as described previously). The extracts were kept at -80°C until further use. For immunoprecipitating TAP tagged proteins, commercially available TAP antibody conjugated agarose beads (Santa Cruz, sc-32319 AC) were used and no coupling was needed. These beads were directly used after equilibration with IP buffer. 2µl of anti-Swi6 antibody was conjugated to 50 µl protein-A sepharose beads (Sigma, P9424-1ml) in a 500 µl 1X PBS containing coupling reaction for 16hrs at 4°C with continuous mixing. The antibody bound beads are then washed twice with 500µl of 0.2M borate buffer (pH 9) and subsequently incubated in 500µl of 0.2M respective borate buffer containing dimethylpimelimidate (DMP) at a concentration of 5mg/ml for 30 minutes at RT with mixing. DMP-mediated coupling was stopped by washing the beads once with 500µl of 0.2M ethanoleamine (pH 8.0). The residual DMP was quenched by incubating the beads in 500µl 0.2M of ethanoleamine (pH 8.0) for 2 hrs at RT. The beads were then washed once with 1X IP buffer [50mM Tris-Cl (pH 7.5), 150mM NaCl, 0.5% NP40 and 0.5mg/ml BSA). Approximately 3 mg of crude extract was bound in the presence of 1X IP buffer for 16hrs at 4°C with mixing. Beads were then washed once with IP wash buffer [50mM Tris-Cl (pH 7.5), 150mM NaCl, 0.1% NP40 (or 0.1% Tween-20) and 1mM EDTA (pH 8.0)]. To each 50 µl beads loaded with immunoprecipitate, 200µl each of LETS buffer [100mM LiCl, 10mM EDTA (pH8.0), 10mM Tris (pH 7.4), 0.2% SDS] and phenol (pH 5.2) were added. Samples were vortexed gently with intermittent ice incubation (30 sec on, 30sec off) followed by 1 hr at 37°C. The beads are then centrifuged at RT for 5min at 12000rpm. The collected supernatants were extracted once with equal volumes (~400µl) of phenol (pH 5.2): chloroform (1:1). RNA in the upper aqueous phase was precipitated with 20µg glycogen, one-tenth volume of 5M LiCl (40µl) and 2.5 volume of

chilled absolute ethanol (1100µl) at -80ºC for 16hrs. Pellet was washed with 300µl chilled 80% ethanol and dissolved in 250µl of RNase-free water. To the dissolved RNA, 250µl PEG precipitation solution (20% PEG (MW8000) in 2M NaCl) was added. Contents were incubated on ice for 30 min and centrifuged at 12000rpm for 20 min at 4ºC. The supernatants were transferred to fresh tubes while the pellets were first dissolved in TE (pH 8.0), ethanol precipitated, washed with 80% ethanol, dried and resuspended in RNase-free water. This contains the population of long RNA while the supernatant isolated after PEG precipitation contains the small RNAs. Supernatant obtained after PEG precipitation was extracted once with phenol (pH 5.2): chloroform and small RNA present in the aqueous layer was precipitated using LiCl-ethanol system as earlier. The isolated small RNA was checked by running on a 15% 7M denaturing urea gel. This RNA was subjected to Northern blotting with probes complementary to *dh* sequences of the *S. pombe* centromeric repeat regions. Alternatively, they were used as riboprobes for Southern blots of PCR products generated from the centromeres and the mating type loci of fission yeast, like, *dh, dhk* and *dg*; *act1* was taken as a negative control. Finally, one set of the isolate RNA was subjected to next generation sequencing.

The NEBNext® Multiplex Small RNA Library Prep Set for Illumina based system was used for 3' adaptor ligation, cDNA synthesis, 5' adaptor ligation and PCR, followed by size selection and sequencing using the Illumina platform was performed by Bionivid Technology Pvt. Ltd., Bangalore.

The sequencing data was analyzed using bowtie2 and R-packages as detailed in the following section.

RIP-seq data processing

Gene annotations and reference genome of fission yeast were downloaded from https://www.pombase.org/. Reference genome was indexed using 'bowtie2-build' * http://bowtie-bio.sourceforge.net/bowtie2/index.shtml). We trimmed the 5' and 3' Illumina small RNA-seq adapters sequences (GUUCAGAGUUCUACAGUCCGACGAUC and

TGGAATTCTCGGGTGCCAAGG, respectively) using 'cutadapt' (https://cutadapt.readthedocs.io/). Duplicates were removed using 'rmdup' utility of SeqKit package (https://bioinf.shenwei.me/seqkit). The clean reads were mapped onto the indexed reference genome using bowtie2 with '–sensitive' setting. The quality of the reads was assessed using FastQC (https://www.bioinformatics.babraham.ac.uk/projects/fastqc/) at each step. We binned the raw reads at 5kb (for boxplot) and 200 bp (for line plot) resolutions and quantile normalized the read counts across samples using the function 'normalize.quantiles' of 'preprocessCore' R-package (https://www.bioconductor.org/packages//2.11/bioc/html/preprocessCore.html). The centromere annotations were obtained from https://www.pombase.org/status/centromeres. The telomeric and sub-telomeric regions were defined as 0-50Kb and 50-100Kb regions from the chromosome ends. The read-counts for *swi6*$^{+}$ and *swi6*$^{3K\rightarrow 3A}$ samples were divided by those of control ('vector') sample. We calculated the log2 ratio of read counts of *swi6*$^{3K\rightarrow 3A}$ to *swi6*$^{+}$ for plotting purpose. Sequence and annotations for mating-type region were obtained from https://www.ncbi.nlm.nih.gov/nuccore/FP565355.1. The reads were independently mapped onto mating-type sequence following the same procedure as earlier and processed at 200 bp resolution. We further identified the 5kb bins that had at-least one of the sequence motifs known to interact with Swi6 (Table X) using in-house script. P-values were calculated using two-tailed Mann-Whitney U tests on R.

The quality scores of RIP-Seq data are shown in Supplementary Table S6.

ChIP-seq data processing

We downloaded the 'sra' files of the ChIP-seq experiments, such as H3K4me3, H3K36me3, H3K9me2, H3K9me3, RNA PolII and Swi6 from GSE83495. We converted the files to 'fastq' format using 'fastq-dump' of NCBI SRAtoolkit. The reads were then mapped on the indexed genome reference using bowtie2 with default parameters. We converted the 'sam' files into 'bam' files using 'samtools' and 'bam' to 'bed' using 'bedtools' with default parameters. We binned the raw reads at 5kb resolution and labeled bins appropriately when their coordinates mapped into telomeres, centromeres and mating region. We plotted the density of read counts for

the bins with RIP-seq log2 ratio (mutant to wild-type) > +1 or < -1. The p-values were calculated using Mann Whitney U tests.

Gene expression analysis

FPKM values for WT pombe strain were downloaded from GSE104546 and the average gene expression was calculated for each 5kb bin. We plotted the distribution of average gene expression values for the bins with RIP-seq log2 ratio (mutant to wild-type) > +1 or < -1. The p-value was calculated using Mann Whitney U test.

REFERENCES

1. Thon,G., Cohen,A. and Klar,A. (1994) Three Additional Linkage Groups That Repress Transcription and Meiotic Recombination in the Mating-Type Region of Schizosaccharomyces pombe. *Genetics* **138**, 29-38.
2. Thon,G., Bjerling,K.P. and Nielsen,I.S. (1999) Localization and Properties of a Silencing Element Near the mat3-M Mating-Type Cassette of Schizosaccharomyces pombe. *Genetics* **151**, 945-963.

Supplementary Figure S1. Mapping of the siRNA sequences to the *dg.dh* regions in a, *cenI*, b, *cenH* in the mating type locus and c, *tlh1-tlh2* genes.

Supplementary Figure S2..Pairwise alignment of different siRNAs used in this study.

Supplementary Figure S3. (**A**) The double stranded template, wherein the top DNA strand comprised of the T7 promoter sequence fused to DNA sequence cognate to the siRNA sequences, is transcribed in presence of ribonucleotides and T7 RNA polymerase to synthesize the respective siRNA. Red bar denotes the T7 promoter and green bar the DNA sequence cognate to the siRNA sequence. (**B**) The full length and domains of Swi6: CD, CD-hinge and CSD, expressed as GST fusion, as visualized by SDS-PAGE in (**C**).

Supplementary Figure S4. Binding of Swi6/HP1 to the 'E-For' siRNA and competition for binding by sequences I-IX (according to Ref. 24). (**A**) EMSA assay was performed for binding of Swi6 to the radiolabelled 'E-For' RNA in presence of excess of unlabeled 'cold' RNAs with sequences of E and I-IX. (**B**) Binding to 'E-For' RNA could also be competed out by excess of 'V' RNA that maps to the subtelomeric genes, *tlh1* and *tlh2* (Figure. 1).

Supplementary Figure S5..Consensus binding sequence for Swi6. The sequences, labeled B, D, E and K were aligned according to Ref. 23..

Supplementary Figure S6. The siRNA precursor RNA 'RevCen' and 'Cen100', fail to show binding to Swi6in EMSA assay. The EMSA assay was performed by incubating the radiolabelled RNAs, 'RevCen' (**A**) and 'Cen100' (**B**) with GST or GST-Swi6, and subjecting to electrophoresis in 8% acrylamide gel, as in Figure 1.

Supplementary Figure S7. Specificity of binding of Swi6 to the *dg.dh* repeat sequences *in vivo*. (**A**) Expression of GFP-Swi6 and GFP-Swi6$^{3K\rightarrow3A}$ in *swi6Δ* cells. Extracts prepared from cells of *swi6Δ* strain expressing the control vector, GFP-Swi6 and GFP-Swi6$^{3K\rightarrow3A}$ were subjected to western blotting with anti-GFP (upper panel) and α-tubulin (lower panel) antibody. (**B**) Slot blot

hybridization.RNAs extracted from cells expressing TAP-tagged Tas3 immunprecipitated with anti-CBP antibody and those expressing GFP-tagged Swi6 or GFP-tagged Swi6$^{3K\rightarrow3A}$, immunoprecipitated with anti-GFP antibody, were radiolabeled at the 5'-end with [γ–^{32}P]ATP and T4 polynucleotide kinase and used as a probe for slot blots on which equal amounts of the DNAs (100ng) corresponding to *act1*, *dg*, *dh* and *dhk* (*cenH*) were blotted. The bound signal was visualized by autoradiography. (**C**) The size fractionated radiolabelled siRNA samples from indicated strains were subjected to electrophoresis in 8% acrylamide/7M urea gels and visualized by autoradiography. In parallel, unlabeled samples were similarly processed blotted and probed with radiolabelled snoRNAs (lower panel), followed by autoradiography. (**D**) No bidirectional transcription is observed in cells expressing *swi6$^{3K\rightarrow3A}$* mutant. RNAs extracted from wt, *dcr1Δ*, and *swi6Δ* strain harboring the empty vector, or vector expressing GFP-Swi6 or GFP-Swi6$^{3K\rightarrow3A}$ were subjected to RTPCR for the Forward strand and Reverse strand for the *dh* and *act1*. The PCR products were visualized by agarose gel electrophoresis and staining with ethidium bromide.

Supplementary Figure S8. (**A**) Chromosomal tracks of RIP-seq fold change (log2 *swi6$^{3K\rightarrow3A}$*/*swi6^{+}* ratio) at and around centromeres of chr-II and chr-III at 5kb (top) and 200bp (bottom) resolutions. (**B**) RIP-seq fold change fold change at right telomere of chr-II (200b resolution). Locus annotations are given at the bottom of the plot.

Supplementary Figure S9. The mutant *swi6$^{3K\rightarrow3A}$* is defective in heterochromatin localization. **(A**) Transformants of *swi6Δ* strain harboring empty vector, vector expressing GFP-Swi6 and GFP-Swi6$^{3K\rightarrow3A}$, were visualized by confocal microscopy. Arrowheads indicate the spots (normally three in over 65% of cells) where GFP Swi6 is localized (middle panel, right) but not in case of GFP-Swi6$^{3K\rightarrow3A}$. (**B**) Quantitation of per cent cells containing one, two or three fluorescent spots of GFPSwi6, while in case of GFPSwi6$^{3K\rightarrow3A}$, fewer cells (~20%) contain 3 spots, and more cells have two or one spot or diffuse fluorescence. (**C**) Schematic representation of the mating-type locus, indicating the insertion of *ura4* reporter at centromere distal location with respect to the *mat2P*. Also show are the homology boxes H2 and H1 at all three loci: *mat1*, *mat2* and *mat3* and H3 box at *mat2* and *mat3* loci. *REII* represents the repression element II and

Δ sign indicates deletion of the *REII* site (1). (**D**) Spotting assay as a qualitative measure of the level of expression of the *ura4* reporter. Cells of *swi6Δ* strain with the genotype shown in (**C**) transformed with empty vector, GFP-Swi6 and GFP-Swi6$^{3K\rightarrow 3A}$, were serially diluted and spotted on plates lacking leucine (for plasmid selection) and those lacking leucine and containing FOA. The cells were allowed to grow for 4 days at 30ºC and photographed. (**E**) Swi6$^{3K\rightarrow 3A}$ fails to complement the switching defect in the h^{90} *swi6Δ* mutant. Cells of indicated transformants of *swi6Δ* strain were spotted on minimal medium lacking iodine. After 4 days' growth at 30ºC, the colonies were stained with iodine and photographed.

Supplementary Figure S10. Expression of *rnh1* abrogates the localization of GFP-Swi6. Cells expressing chromosomally tagged GFP-Swi6 were transformed with empty vector, and vector expressing *rnh1* and *rnh201* under the control of *nmt1* promoter, which is repressed in presence of thiamine. Cultures of the transformants grown in presence (+; repressed) and absence (-, derepressed) of thiamine were visualized by confocal microscopy. Representative data shown in Figure 4**C**.

Supplementary Figure S11. RNA-DNA hybrid formation is essential for heterochromatin silencing and Swi6 localization. (**A)**Schematic representation of mating type locus, showing the deletion of the *mat3*-linked repression element *REIII* (denoted by the Δ sign) and insertion of the *ade6* reporter distal to *mat3* (2). (**B**) The silencing of the *ade6* reporter is abrogated by overexpression of *rnh1* but not *rnh201*. Transformants of the indicated strain were streaked on selective plates containing limiting amount of adenine. The plates were grown for 4 days at 30ºC and photographed. (**C**) The fraction of transformant colonies giving pink/white colour were counted and plotted. (**D**) ChIP assay was performed to quantitate the enrichment of Swi6 at the *mat3::ade6* locus in the transformants shown in (**B)**. (**E**) Schematic representation of the mating type locus *mat2*, indicating the deletion of repression element *REII* (denoted by the sign Δ) and insertion of *ura4* reporter at centromere distal location (1). (**F**) Overexpression of *rnh1* but not *rnh201* causes loss of silencing at the *mat2* locus as well, as indicated by growth of colonies that give dark staining with iodine (1; Materials and Methods).(**G**) Fraction of transformants with the different vectors giving dark staining with iodine was plotted. (**H**) Delocalization of Swi6 at

mat2 locus upon transformation with *rnh1*. Level of Swi6 enrichment at the *mat2*-linked *ura4* locus in different transformants was plotted.

Supplementary Figure S12. Log2 ratio ($swi6^{3K\rightarrow 3A}/swi6^{+}$) of RIP-seq profile at centromere of *chrI* and the flanking regions at 200 bp resolution. The thymidylate kinase locus (SPAC15E1.04) is highlighted in yellow colour. The zoomed image shows normalized RIP-seq signals at thymidylate kinase in $swi6^{+}$ and $swi6^{3K\rightarrow 3A}$ strains.

Supplementary Table S1

List of strains used in this study

Strain Name	Genotype
SPA236	*mat1Msmt0 leu1-32 uraD18 ade6216 REIIΔmat2::ura4*
SPW146	*mat1Msmt0 leu1-32 uraD18 ade6216 REIIΔmat2::ura4 swi6Δ::kan*^{r}
FY2002	h^{+} *leu1-32 uraDS/E ade6DN/N imr1L::ura4 otr1R::ade6*
SPW142	h^{+} *leu1-32 uraDS/E ade6DN/N imr1L::ura4 otr1R::ade6 swi6Δ::kan*^{r}
FY1862	h^{90} *leu1-32 his3D ade6-210 ura4D18 his3-tel(1L)*
SPW170	h^{90} *leu1-32 his3D ade6-210 ura4D18 his3-tel(1L) swi6Δ::kan*^{r}
SPJ541	h^{90} *leu1-32 uraD18 ade6-216*
FY13724	h^{90} *leu1-32 ade6-216 swi6Δ::kan*^{r}
FY13723	h^{90} *leu1-32 ade6-216 swi6::GFP-kan*^{r}

Supplementary Table S2

List of siRNAs according to Reinhardt and Bartel (23)

Assigned Name	Sequence	Length	Location
A	GAGGCUUUCGGUUUAGUCGC	20	dg-otr1 (Rev), otr2 (For), otr3 (For)
B	AAUGCGGAGUAAGGCUAAUCACGGUA	26	dg-otr1 (Rev), otr2 (For), otr3 (For)
C	UCUAGCUUCGCCAUCAAUAAGUA	23	dg-otr1 (For), dg-otr2(Rev), otr3 (Rev), dh-otr2(Rev)
D	UGGAUUAAGGAGAAGCGGUA	20	dh-otr1 (Rev), otr2 (Rev)
E	ACAAGUGAUAAGAGUAGGUGU	21	dh-otr1 (Rev), otr2 (Rev)
F	UGCGCAACUCCUGCUUAUCGUC	22	dh-otr1 (For), otr2 (For), otr3 (For)
G	UACAAGAUAUAGCGCCACACU	21	dh-otr1 (For), otr2 (For), otr3 (For)
H	UGAGCAUAUCCUAAUGACAGUA	22	dh-otr1 (For), otr2 (For), otr3 (For), tlh1/tlh2 (partial)
I	UGCCUAUUUAUACAUUUCCC	20	dh-otr1 (For), otr2 (For), otr3 (For)
J	UCUACCUCAGCAGUCCUUGGGAAA	24	dh-otr1 (Rev), otr2 (Rev), otr3 (Rev)
K	UGUGUCCAUAUCCAUGCUGUGUCCA	25	dh- otr3 (Rev)
L	UAAACAACUUGCAAUAUCUGCCA	23	dh-otr1 (Rev)

Supplementary Table S3

List of siRNA sequences generated by Dicer cleavage of RevCen RNA by Djupedal et al (24)

Assigned Name	Sequence	Length	Location	Sense
I	GUCCUACUCCUACACCUACUCUUA	24	otr1L-dh, otr1R-dh, otr2L-dh, otr2R-dh, otr3L-dh, otr3R-dh, cenH	Rev
II	GAUUGUAUUUGAUUCCAUCGUA	22	otr1L-dh, otr1R-dh, otr2L-dh, otr2R-dh, otr3L-dh, otr3R-dh, cenH	Rev
III	UAUCCCGUGUUCUUUCCUUCUUCCA	25	otr1L-dh, otr1R-dh, otr2L-dh, otr2R-dh, otr3L-dh, otr3R-dh, cenH	Rev
IV	UAGCCUUUUAUACAAUGAUUCCUCU	25	otr1L-dh, otr1R-dh, otr2L-dh, otr2R-dh, otr3L-dh, otr3R-dh, cenH	Rev
V	UUUCGUGAAUGCCCAGCCUGUAAAU	25	otr1L-dh, otr1R-dh, otr2L-dh, otr2R-dh, otr3L-dh, otr3R-dh, cenH. tlh1, tlh2	For
VI	CAUAACAGAGAGGUUCAGAAAGGAA	25	otr1R-dh, otr2L-dh, otr2R-dh, otr3L-dh, otr3R-dh, cenH	Fo
VII	GUUUUGAAGUAGACAUUCCGCACAA	25	otr1R-dh, otr2L-dh, otr2R-dh, otr3L-dh, otr3R-dh	Rev
VIII	GUCUAGUACACCAUGUCUCUUGUGCUCAG	29	otr1R-dh, otr2L-dh, otr2R-dh, otr3L-dh, otr3R-dh, tlh1/tlh2 (partial)	Rev
IX	GUACAACAGGC/AAAAUGGAUGUUUC	24	otr1R-dh, otr2L-dh, otr2R-dh, otr3L-dh, otr3R-dh	For
X	AAGTCCATCGCACCATGTTAATT	23	otr1R, otr2L,otr2R, otr3L, otr3R, tlh1/tlh2 (partial)	Rev
XI	CCTTCTGTTCGA/GATGTATCCCTA	23	otr1R, otr2L,otr2R, otr3L, otr3R, cenH, tlh1/tlh2 (partial)	For

Supplementary Table S4

List of DNA templates used for synthesizing si RNAs

Name	Sequence
T7-dh20D For	5’CCAAGCTTCTAATACGACTCACTATAGGTGGATTAAGGAGAAGCGGTA3’
T7 dh20D For compl	5’TACCGCTTCTCCTTAATCCACCTATAGTGAGTCGTATTAGAAGCTTGG3’
T7-dh26B For compl	5’TACCGTGATTAGCCTTACTCCGCACCTATAGTGAGTCGTATTAGAAGCTTGG3’
T7-dh20E For	5’CCAAGCTTCTAATACGACTCACTATAGGACAAGTGATAAGAGTAGGTGT-3’
T7-dh20E For compl.	5’ACACCTACTCTTATCACTTGTCCTATAGTGAGTCGTATTAGAAGCTTGG3’
T7-dh22H For	5’CCAAGCTTCTAATACGACTCACTATAGGTGAGCATATCCTAATGACAGTA-3’
T7-dh22H For compl.	5’TACTGTCATTAGGATATGCTCACCTATAGTGAGTCGTATTAGAAGCTTGG3’
T7-dh25K For:	5’CCAAGCTTCTAATACGACTCACTATAGGTGTGTCCATATCCATGCTGTGTCCA-3’
T7-dh25K For compl.	5’TGGACACAGCATGGATATGGACACACCTATAGTGAGTCGTATTAGAAGCTTGG3’
T7-I24For	5’CCAAGCTTCTAATACGACTCACTATAGGGTCCTACTCCTACACCTACTCTTA3’
T7-I24For compl	5’TAAGAGTAGGTGTAGGAGTAGGACCCTATAGTGAGTCGTATTAGAAGCTTGG 3’
T7-II24For	5’CCAAGCTTCTAATACGACTCACTATAGG’GATTGTATTTGGATTCCATCGGTA3’
T7-II24For compl	5’TACCGATGGAATCCAAATACAATCCCTATAGTGAGTCGTATTAGAAGCTTGG 3’
T7-III25For	5’CCAAGCTTCTAATACGACTCACTATAGGTATCCCGTGTTCTTTCCTTCTTCCA3’
T7-III25For compl	5’TGGAAGAAGGAAAGAACACGGGATACCTATAGTGAGTCGTATTAGAAGCTTGG 3’
T7-IV25For	5’CCAAGCTTCTAATACGACTCACTATAGG’TAGCCTTTTATACAATGATTCCTCT3’
T7-IV25For compl	5’ AGAGGAATCATTGTATAAAAGGCTACCTATAGTGAGTCGTATTAGAAGCTTGG3’
T7-V25For	5’CCAAGCTTCTAATACGACTCACTATAGG’TTTCGTGAATGCCCAGCCTGTAAAT3’
T7-V25For compl	5’ATTTACAGGCTGGGCATTCACGAAACCTATAGTGAGTCGTATTAGAAGCTTGG3’
T7-VI25For	5’CCAAGCTTCTAATACGACTCACTATAGG’CATAACAGAGAGGTTCAGAAAGGAA3’
T7-VI25For compl	5’TTCCTTTCTGAACCTCTCTGTTATGCCTATAGTGAGTCGTATTAGAAGCTTGG3
T7-VII25For	5’CCAAGCTTCTAATACGACTCACTATAGG’GTTTTGAAGTAGACATTCCGCACAA3’
T7-VII25For compl	5’ TTGTGCGGAATGTCTACTTCAAAACCCTATAGTGAGTCGTATTAGAAGCTTGG3’
T7-IX24For	5’CCAAGCTTCTAATACGACTCACTATAGG’GTACAACAGGCAAATGGATGTTTC3’
T7-IX24For compl	5’GAAACATCCATTTGCCTGTTGTACCCTATAGTGAGTCGTATTAGAAGCTTGG3’
presiRNA-Revcen for (EcoRI)	5’ATCTAGAATTCATGGTACCAAAGCTCG3’
presiRNA-Revcen rev hindIII	5’CTGCAAGCTTGTTTTCATTGTGTATCATCT3’

Supplementary Table S5

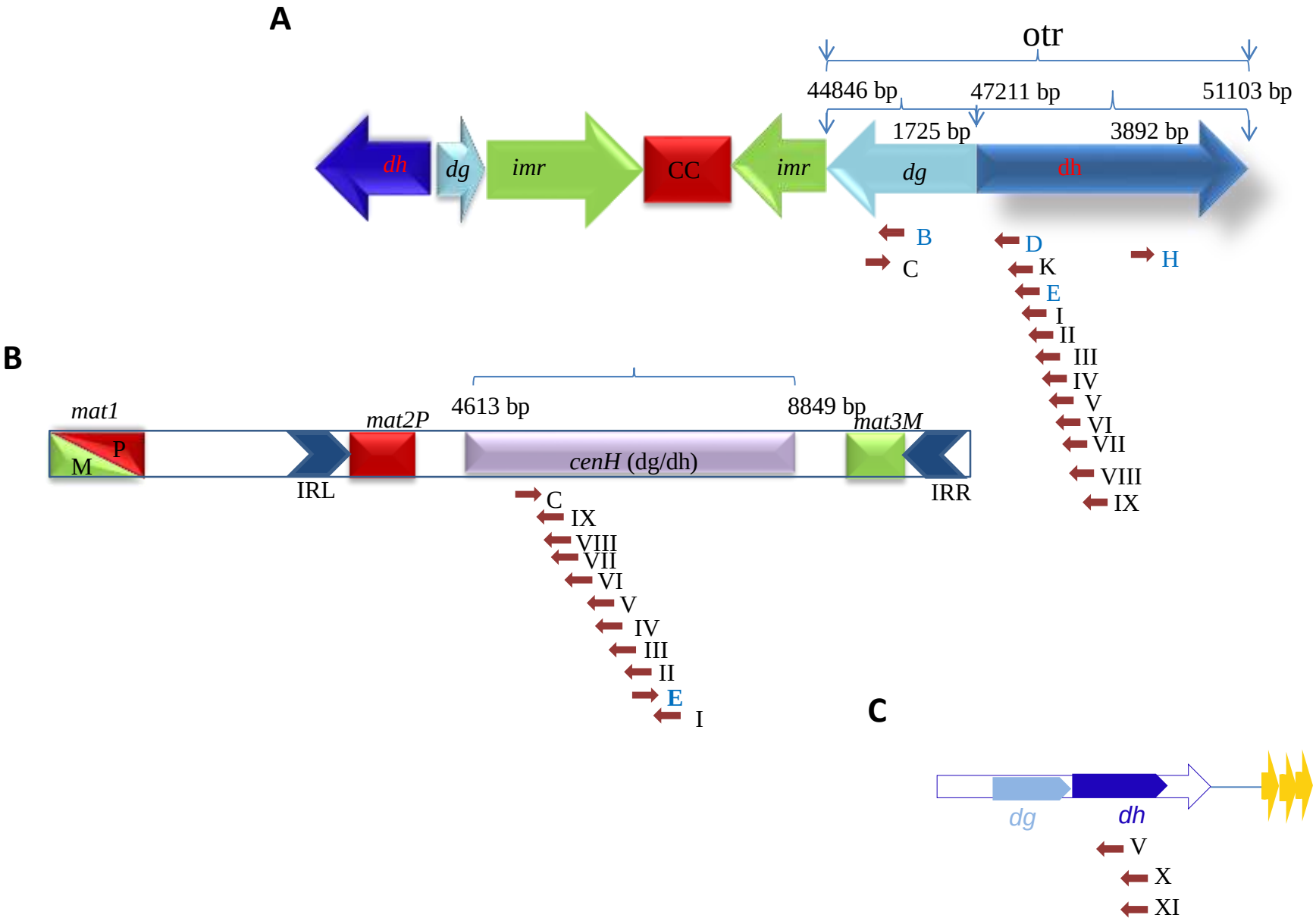
Expression vector constructs used in this study.

Plasmid construct	Brief description
pGEX-2T	GST-fusion gene expression system procured from GE Healthcare
pGEX-2T-swi6	Full length swi6 gene cloned in pGEX-2T vector at BamHI/EcoRI sites
pGEX-2T-swi6^{3K-3A}	Full length swi6 (KKK) mutant gene cloned in pGEX-2T vector at BamHI/EcoRI sites,
pGEX-2T-swi6-CD	Chromodomain (CD) of swi6 gene cloned in pGEX-2T vector at BamHI/EcoRI sites
pGEX-2T-swi6-Hinge	Hinge region of swi6 gene cloned in pGEX-2T vector at BamHI/EcoRI sites
pGEX-2T-swi6-CSD	Chromoshadow domain (CSD) of swi6 gene cloned in pGEX-2T vector at BamHI/EcoRI sites
Litmus 28i-Revcen	RevCen fragment cloned between EcoRI and HindIII restriction sites
Litmus 28i-Cen100	Cen100 fragment clones between EcoRI and HindIII restriction sites

Supplementary Table 6. RIP-seq statistics

Distribution per-base quality scores of final clean reads. The plots are obtained from FastQC.

Sample	# total raw reads	# trimmed and de-duplicated reads (clean set)	Alignment rate (% reads aligned to S. pombe genome)
Vector	11716052	383068	77.02%
swi6^{+}	13072769	385039	74.32%
swi6$^{3K\rightarrow 3A}$	10992694	344485	83.04%

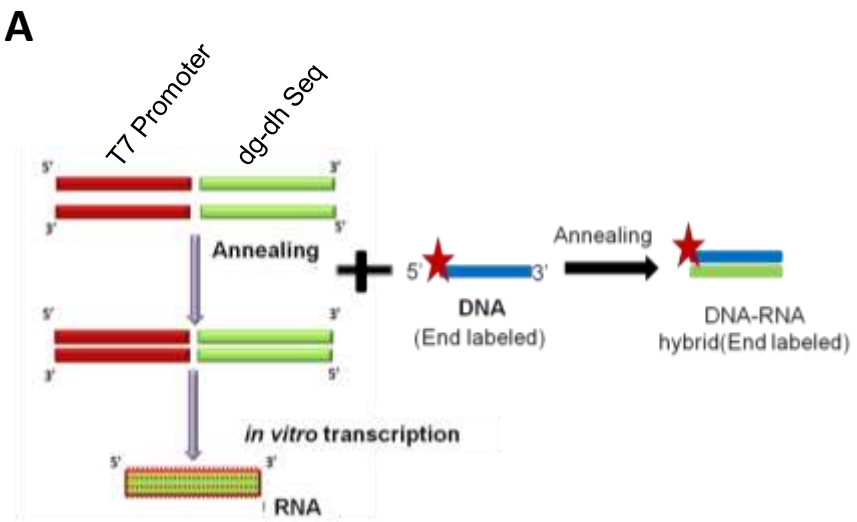


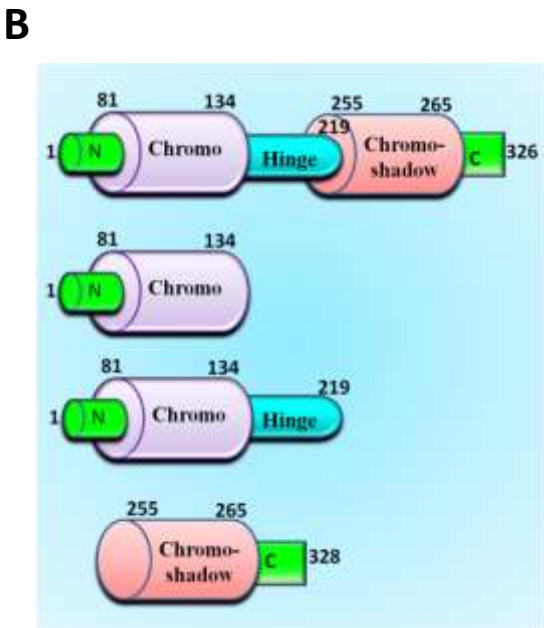
Supplementary Fig. S1

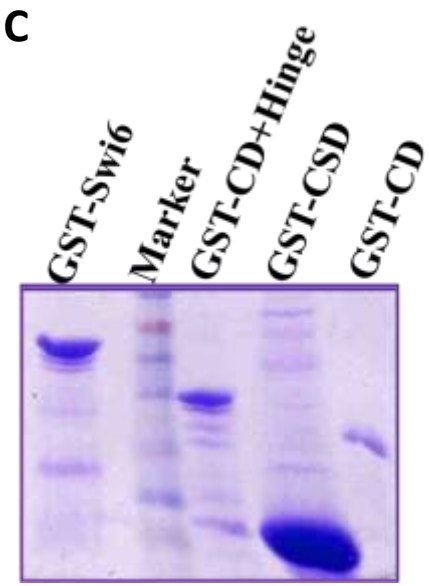
Group 1 ‘D’ & ‘J’	D	--UGGAUUAAGGAGAAGCGGUA--
	J	UCUACCUCAGCAGUCCUUGGGAAA
		* * * ** *
Group 2 ‘F’ & ‘H’	F	UGCGCAACUCCUGCUUAUCGUC
	H	UGAGCAUAUCCUAAUGACAGUA
		** *** **** * * **
Group 3 ‘I’ & ‘K’	I	----UGCCUAUUUAUACAUUUCCC-
	K	UGUGUCCAUAUCCAUGCUGUGUCCA
		* * *** ** * * **
Group 4 ‘E’, ‘G’ & ‘L’	E	---ACAAGUGAUAAGAGUAGGUG
	G	--UACAAGAUAUAGCGCCACACU
	L	UAAACAACUUGCAAUAUCUGCCA
		**** *

RNA sequences from ‘dh’ repeat region :**D, E, H** and **K**
RNA sequence from ‘dg’ repeat region: **B**

Supplementary Fig. S2

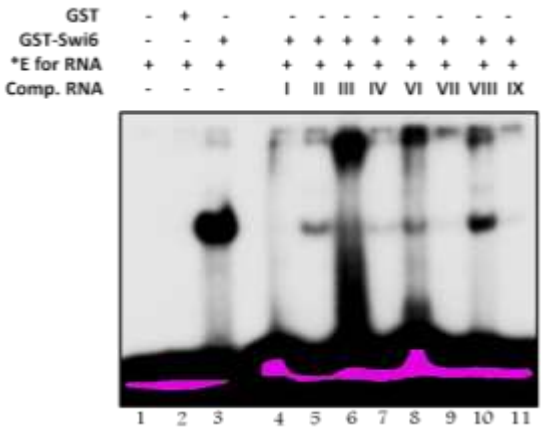




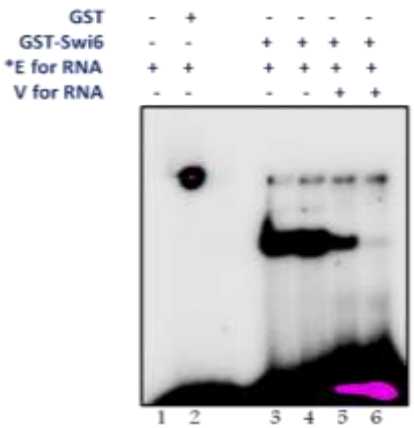


Supplementary Fig. S3

A



B



Supplementary Fig. S4

Discovered Motifs

1.

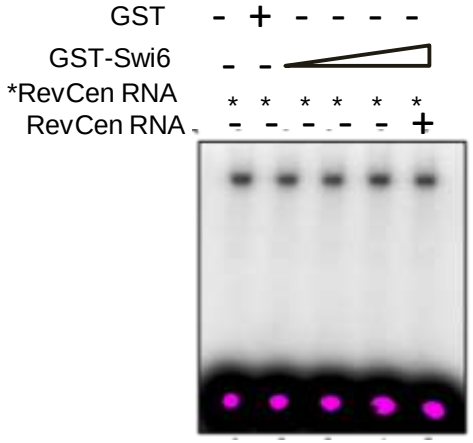
***E*-value:** 1.9e+002 **Site Count:** 4 **Width:** 10

Log Likelihood Ratio: 37 **Information Content:** 13.1 **Relative Entropy:** 13.2 **Bayes Threshold:** 3.70044

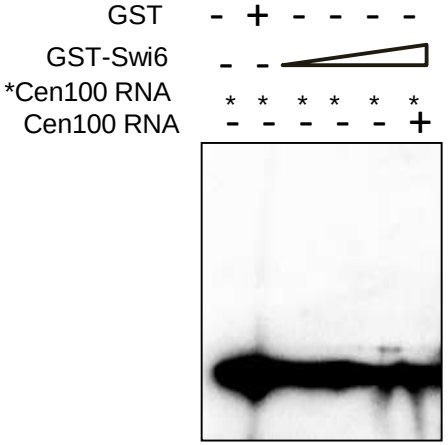
Name	Strand	Start	*p*-value	Sites
3. E	+	12	8.38e-6	CAAGTGATAA GAGTAGGTGT
1. B	+	7	1.69e-5	AATGCG GAGTAAGGCT AATCACGGTA
2. D	+	10	2.06e-5	TGGATTAAG GAGAAGCGGT A
4. K	-	10	1.05e-4	TGGACA CAGCATGGAT ATGGACACA

Supplementary Fig. S5

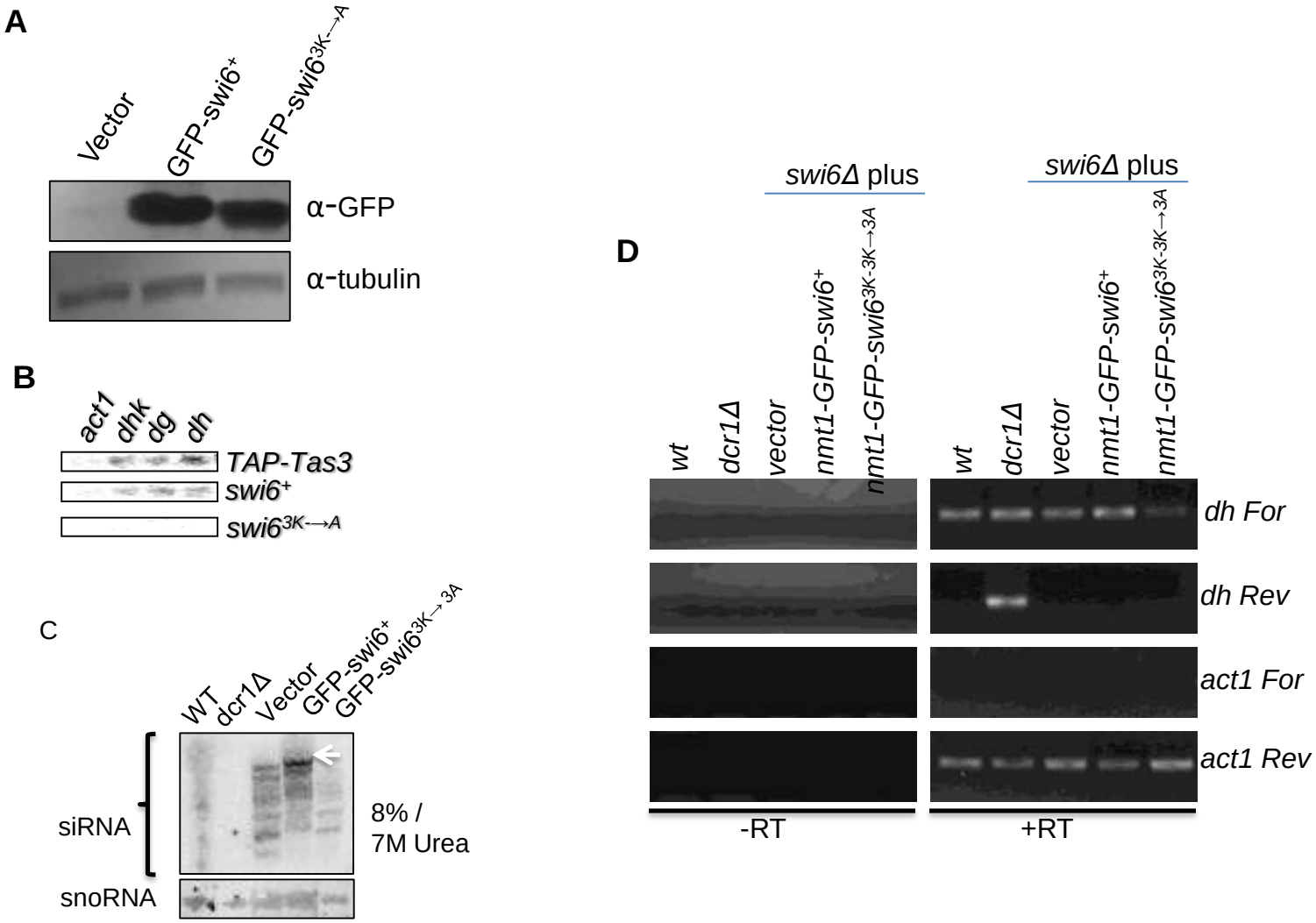
A



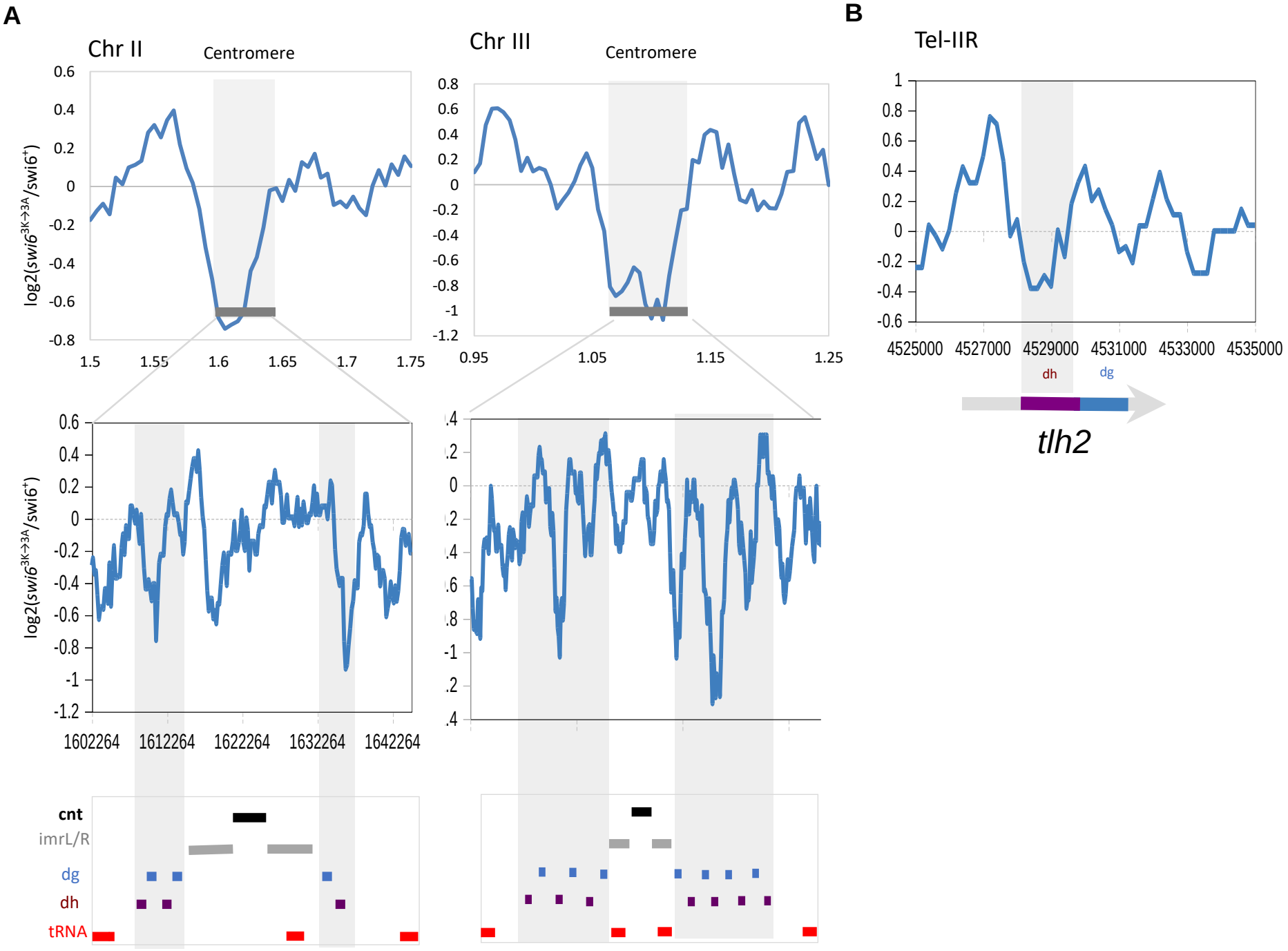
B



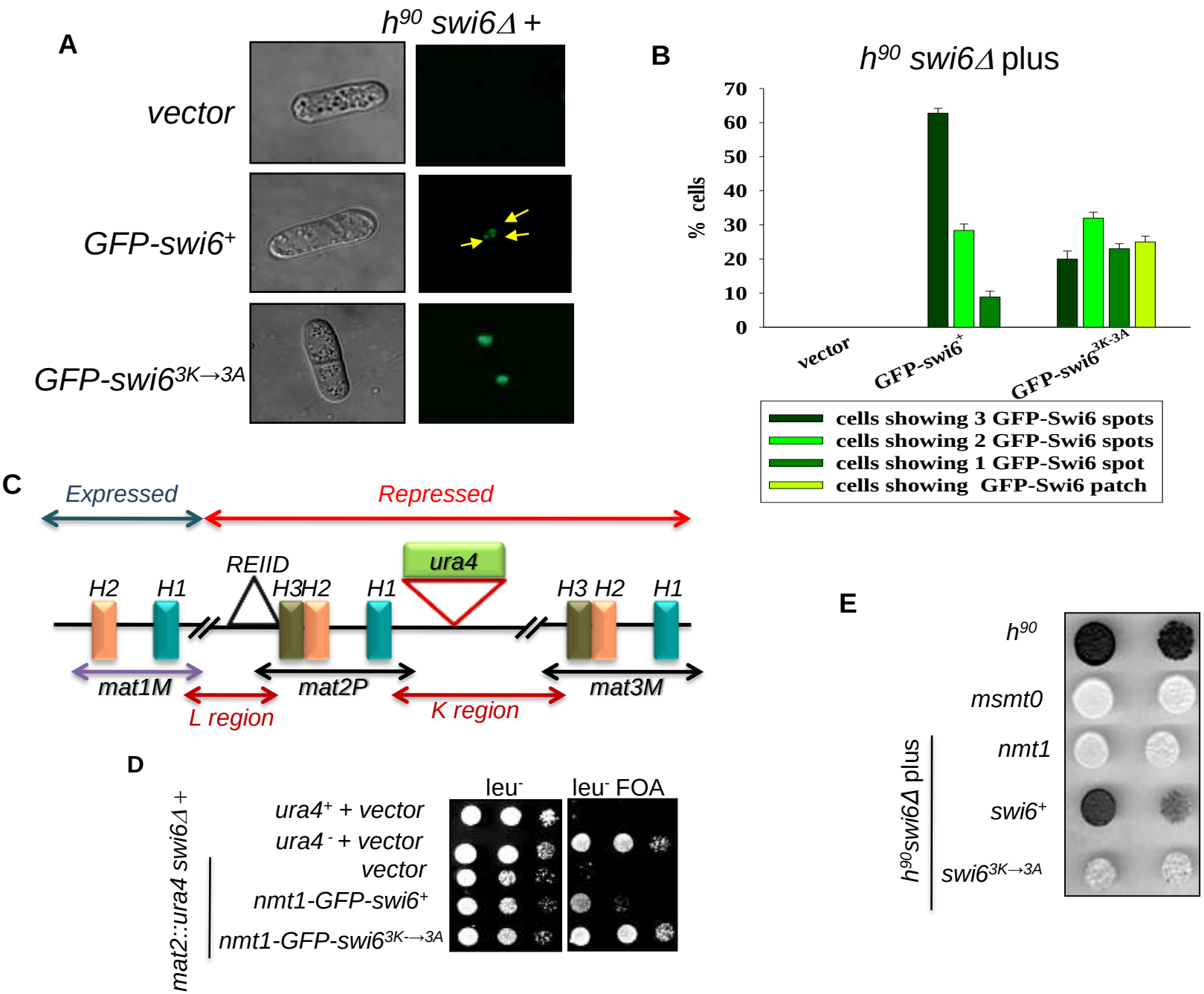
Supplementary Fig. S6



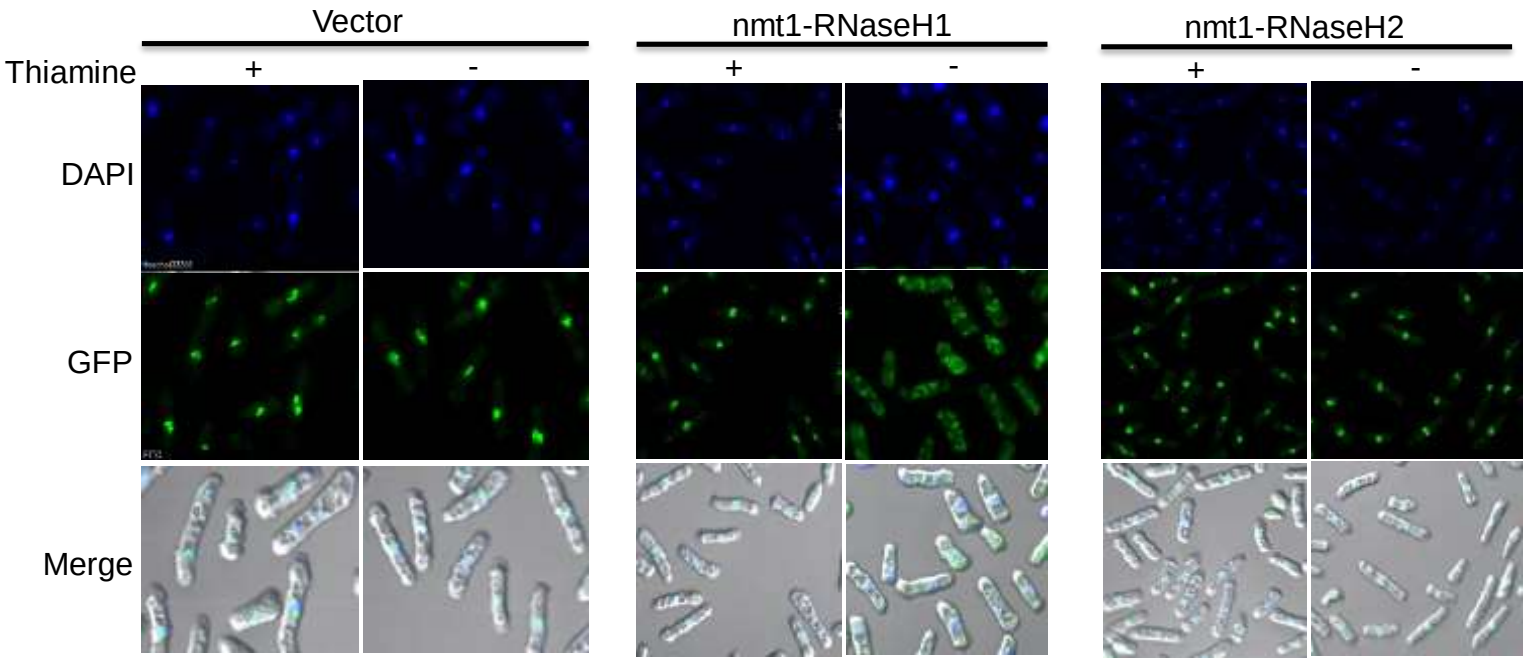
Supplementary Fig. S7



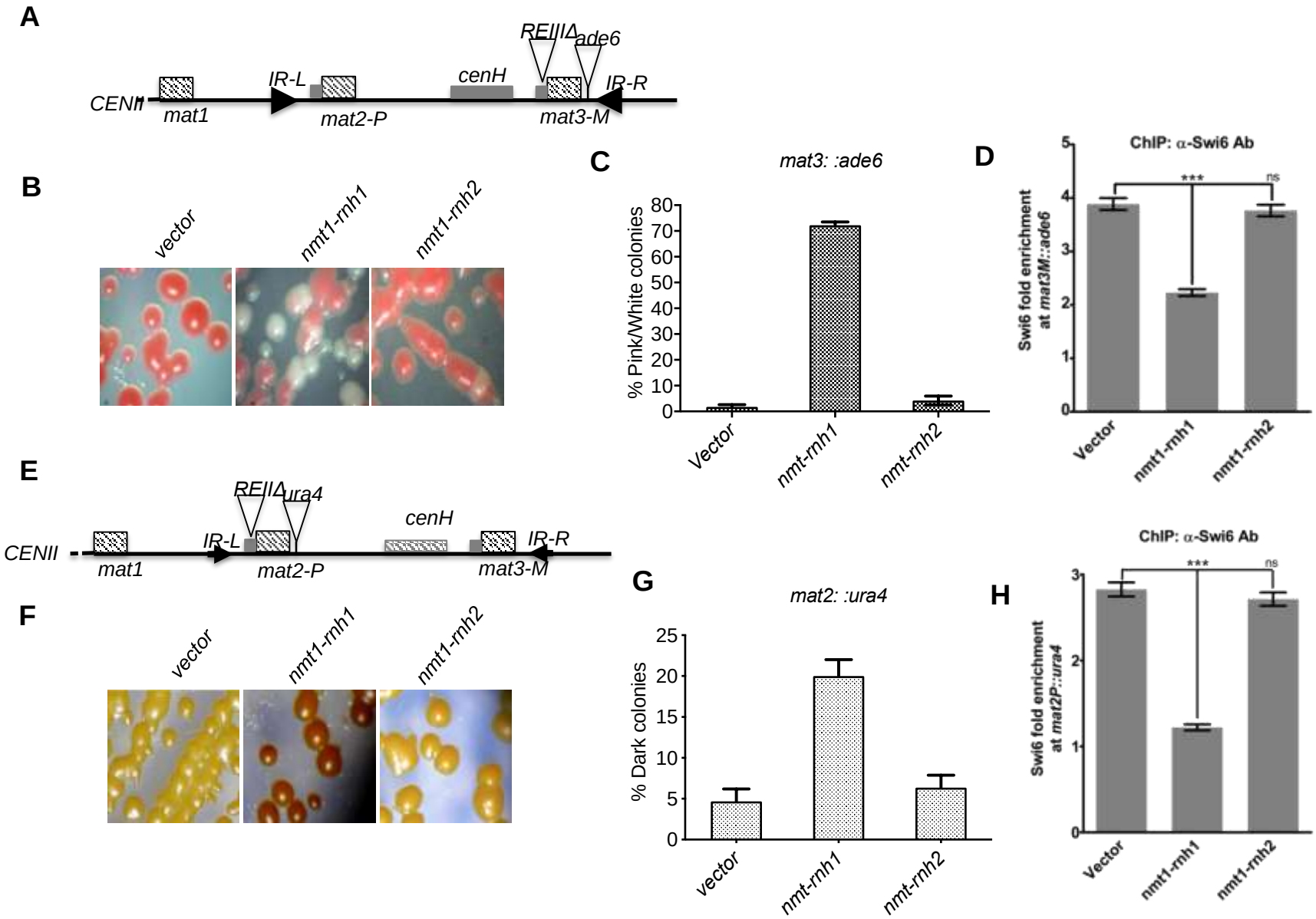
Supplementary Fig. S8



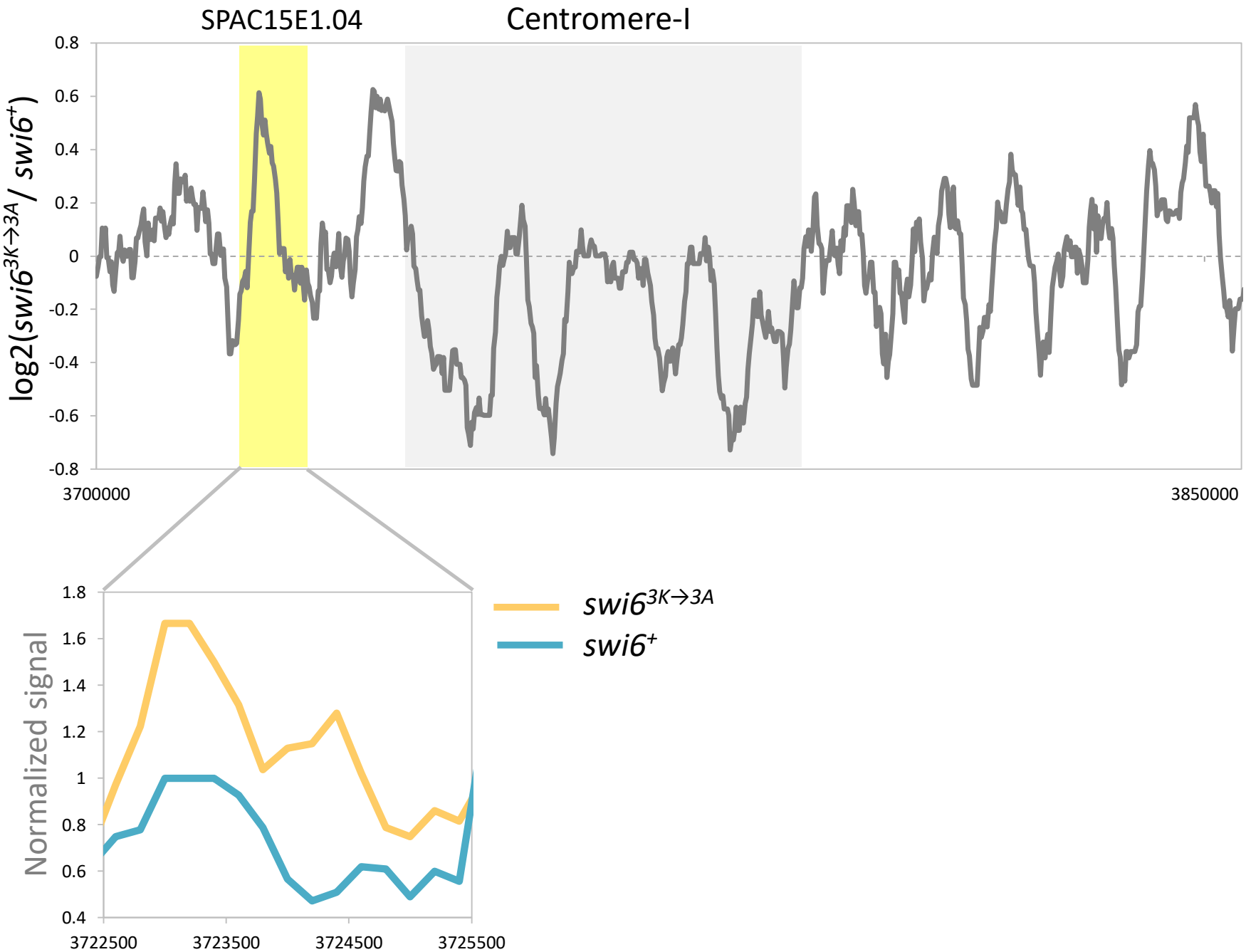
Supplementary Fig. S9



Supplementary Fig. S10



Supplementary Fig. S11



Supplementary Fig. S12