HALLMARK FINDING IN EPIGENETIC GENE REGULATION

THE RNAi MACHINERY AND REGULATION OF HETEROCHROMATIN

siRNAs as endogenous

epigenetic regulators

Small ncRNA and gene/chromatin regulation

micro-RNAs = miRNAs

short interfering RNAs = siRNAs

miRNAs and siRNAs are generated by the same machinery

1. Double stranded precursor RNA

2. Processing into small RNAs by Type 3 RNAses
siRNAs (DICER) and miRNAs (DROSHA+DICER) (21-23 nt): still double-stranded

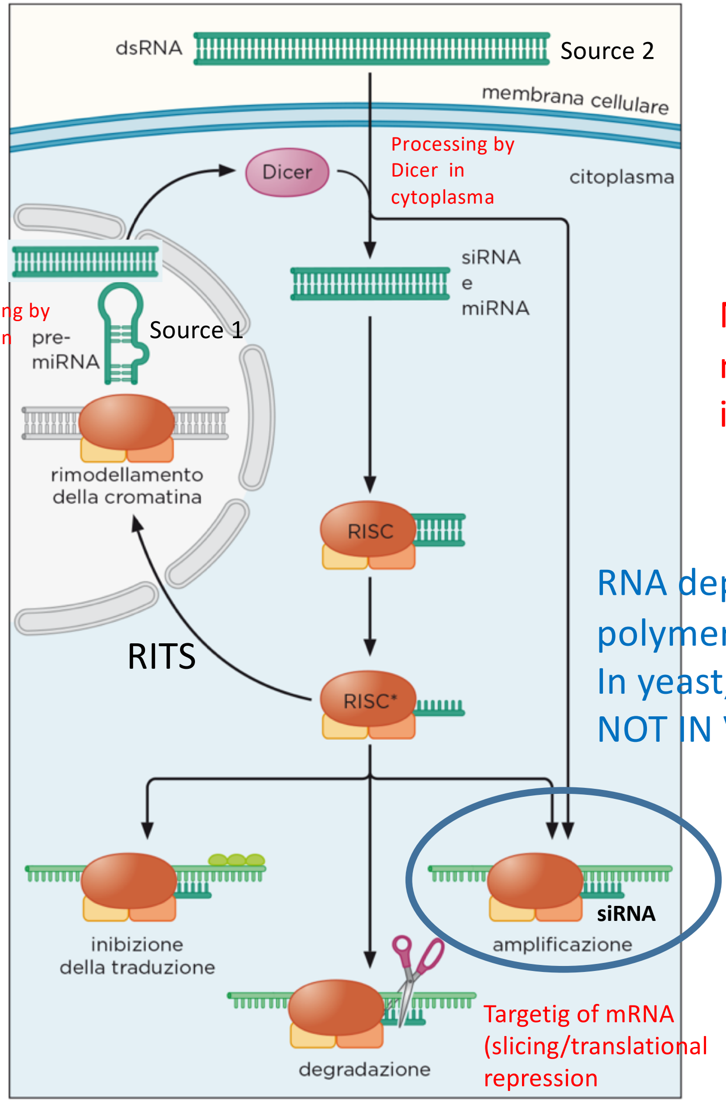
3. Processing by RISC complex
(RNA induced silencing complex)

4. guide RNA → regualtory RNA
passenger RNA → will be eliminated

5. RISC complex+guide RNA → regulatory function

A. RNA degradation = siRNA effect (cutting = “slicing”
B. inhibition of mRNA translation =mRNA effect
C. transfer to nucleus and chromatin regulation = siRNA mediated silencing

miRNAs: always “trans”-acting on mRNAs
siRNAs: mostly “cis” acting on chromatin (S. pombe)



Processing by Dicer in cytoplasma

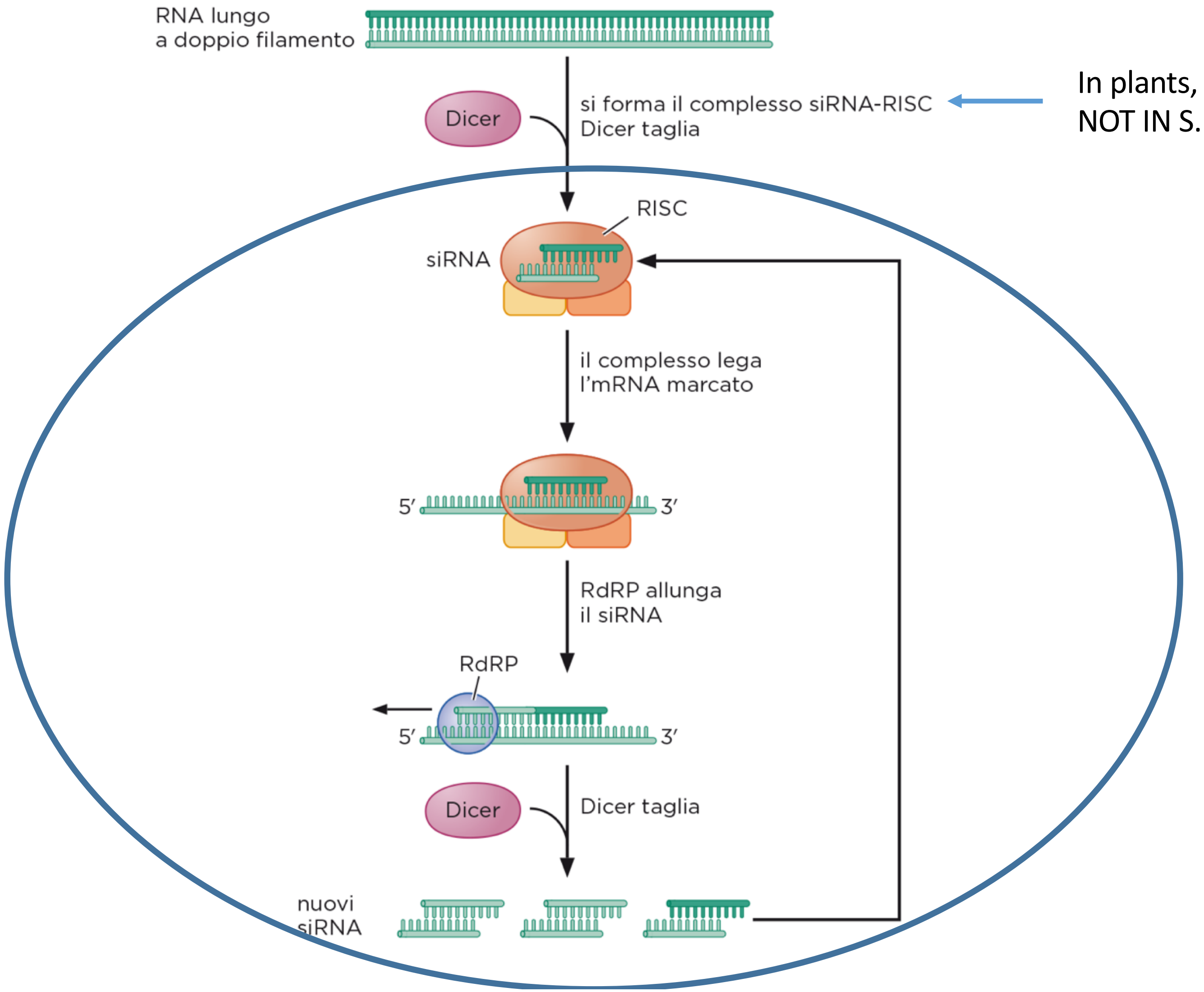
Processing by Drosha in nucleus

Targetig of mRNA (slicing/translational repression

NO miRNAs in yeast

RNA dependent RNA polymerase
In yeast, C.elegans
NOT IN VERTEBRATES

RNA dependent RNA polymerase amplifies siRNAs



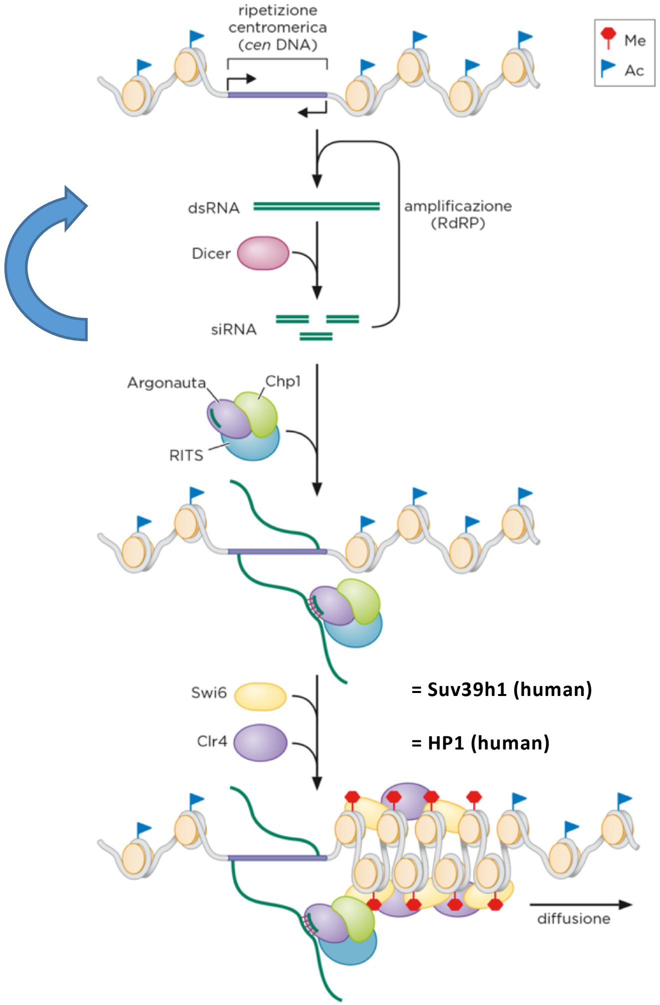
In plants, S. pombe, nematodes, human,
NOT IN S. CEREVISIAE

In plants, S. pombe, nematodes

- Works primer less (de novo)
- With primers (virus)

- NO RNA dependent RNA polymerase in Vertebrates

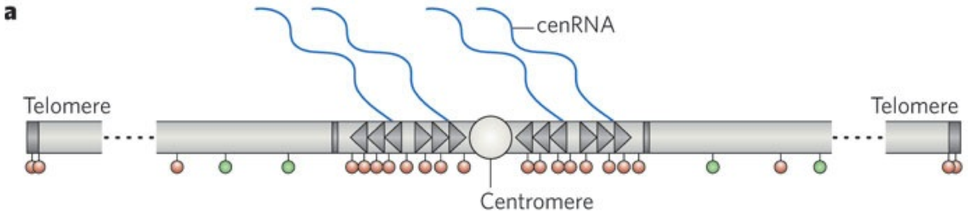
The first model for siRNAs and chromatin regulation (induction):
S. pombe: siRNA mediated chromatin regulation (silencing)



RNA dependent RNA polymerase

= Suv39h1 (human)

= HP1 (human)



Centromeres in S. pombe:

-Heterochromatin (H3K9me3, Clr4 (Suv39h1); Swi 6 (HP1))

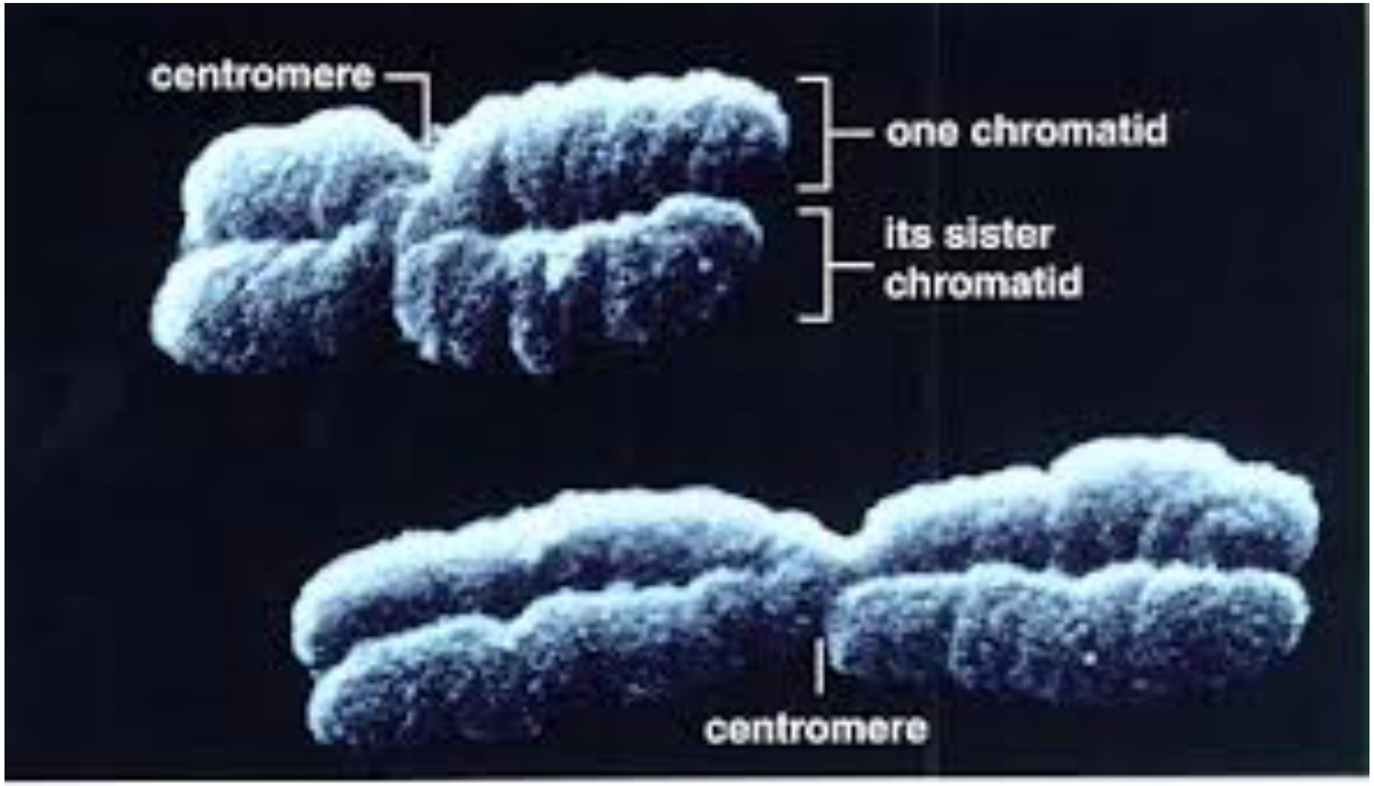
-Reporter genes inserted: repression

-**Discovery:**

RNAi mutant result in loss of H3K9m3/Clr4 and reactivation of reporter gene that was inserted into centromeric region

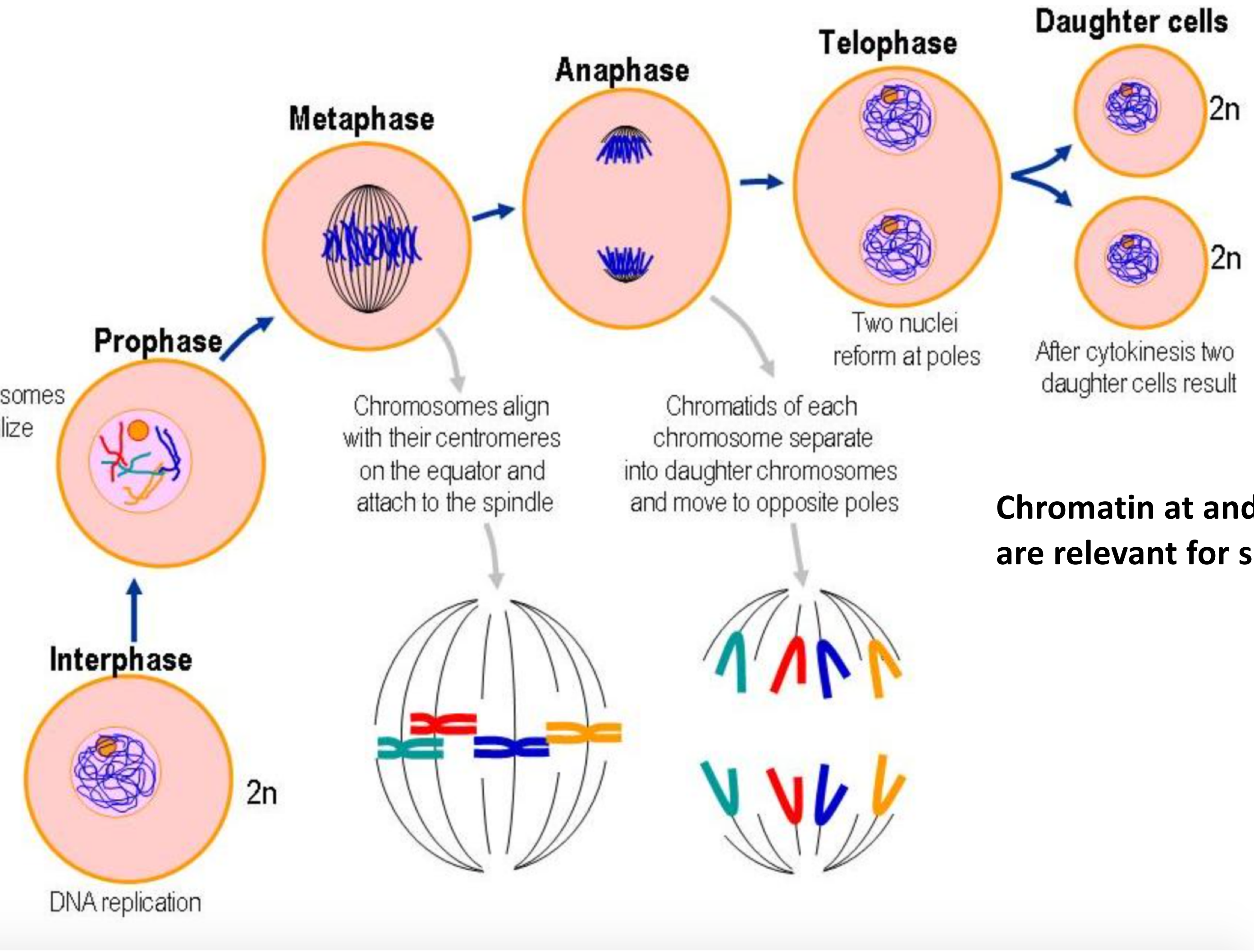
=RNAi mediated gene silencing

Eukaryotic centromeres



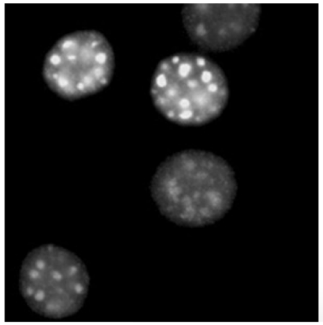
Chromatin at and around centromeres centromeres are relevant for successful mitosis

Eukaryotic centromeres

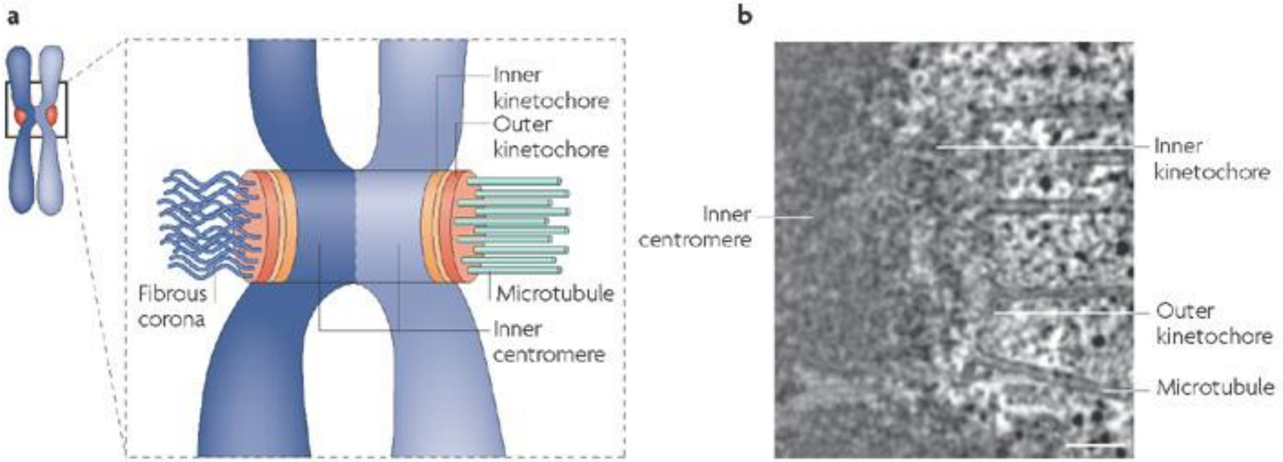


Chromatin at and around centromeres are relevant for successful mitosis

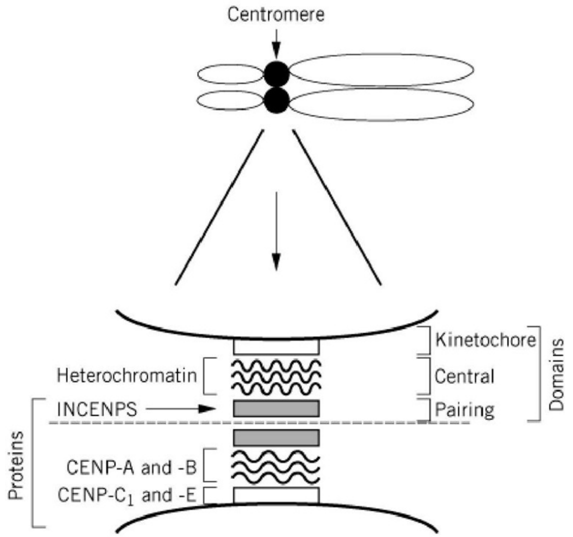
Chromocenters, aggregation of centromeres



Functional regions at centromeres

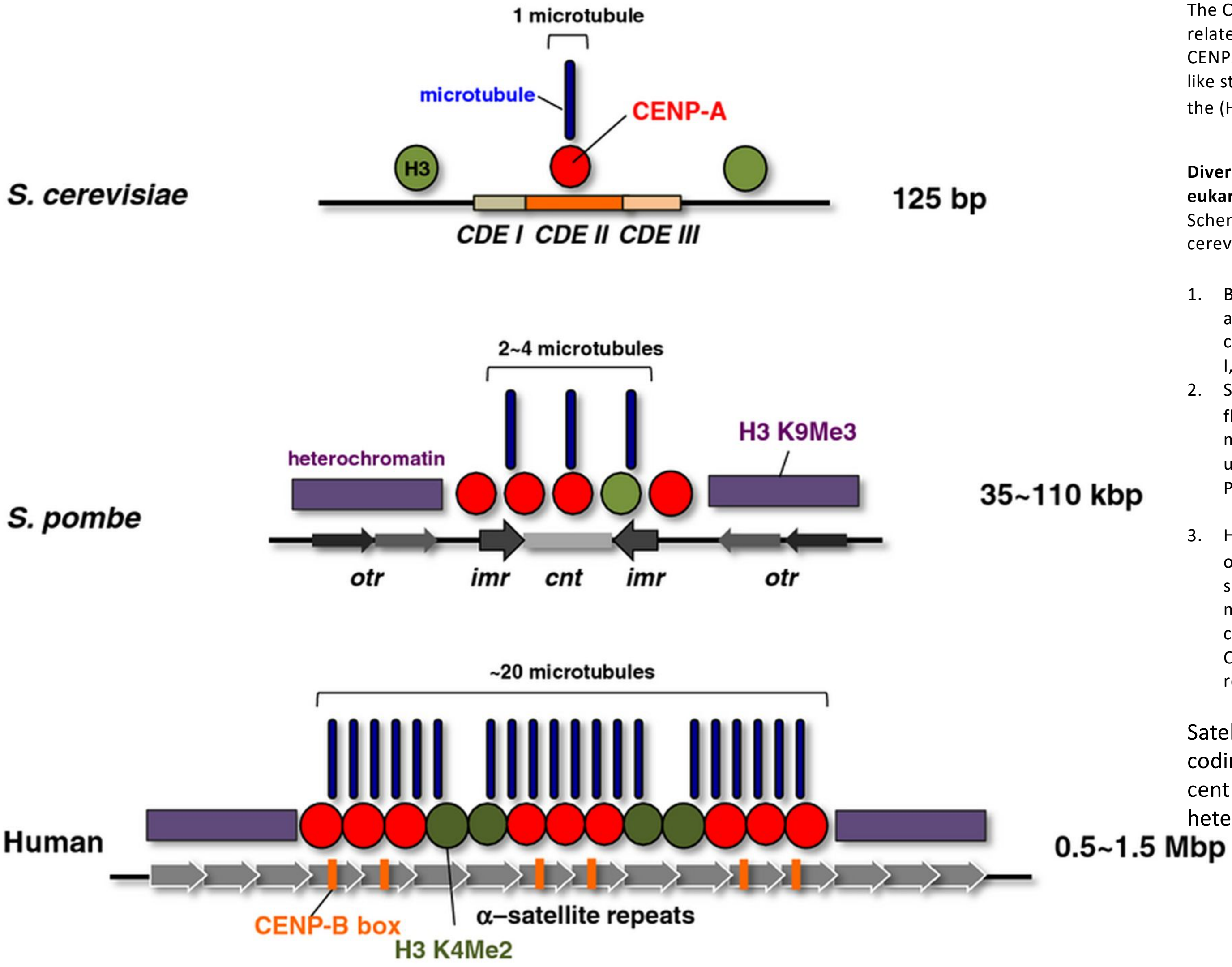


Nature Reviews | Molecular Cell Biology



**(Peri)-centromeric DNA consists of constitutive
Heterochromatin (H3K9me3, HP1, DNA methylation**

Centromere structure in diverse organisms



The CENPA gene encodes a centromere protein which contains a histone H3 related histone fold domain that is required for targeting to the centromere. CENPA is proposed to be a component of a modified nucleosome or nucleosome-like structure in which it replaces 1 or both copies of conventional histone H3 in the (H3-H4)2 tetrameric core of the nucleosome particle.

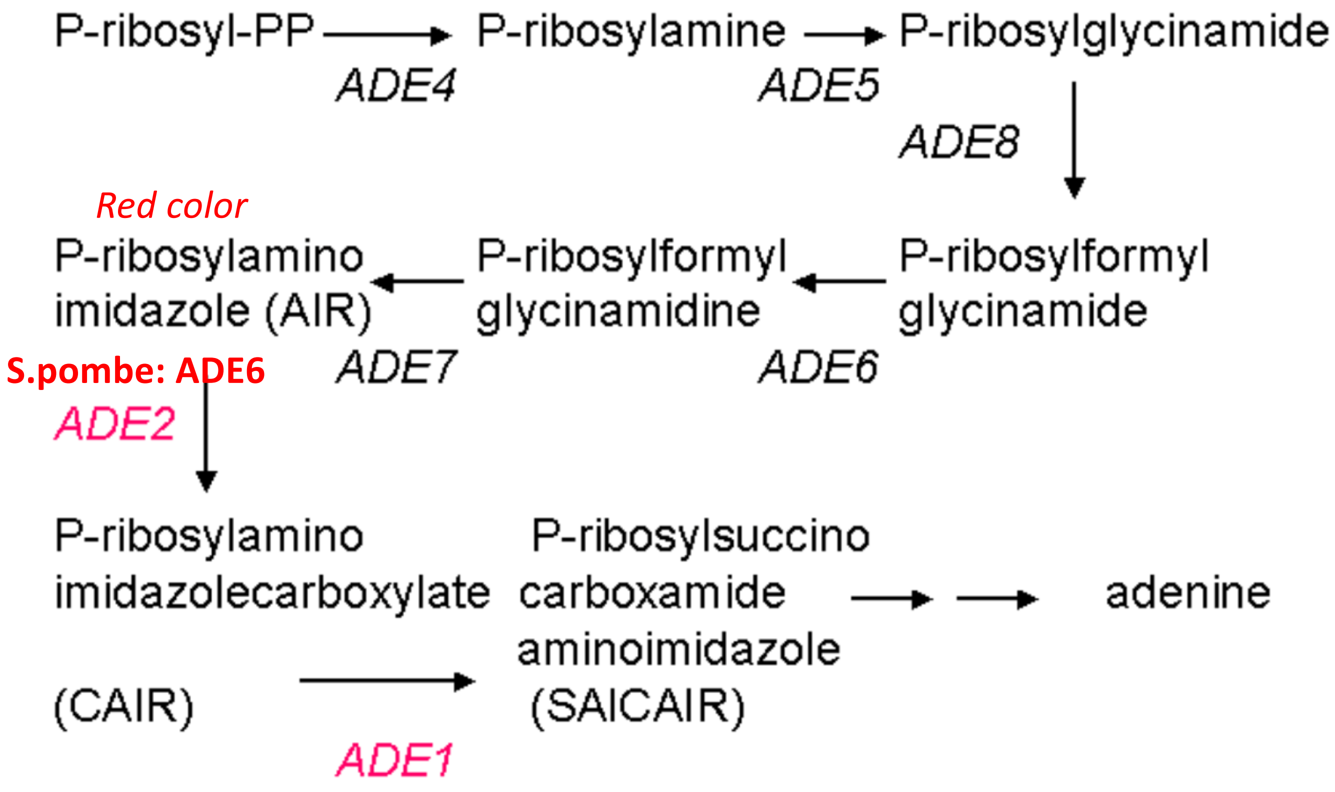
Diversity of centromeric structure and CENP-A nucleosome distribution among eukaryotes.
Schematic representation of centromeres in budding yeast (Saccharomyces cerevisiae), fission yeast (Schizosaccharomyces pombe) and human.

1. Budding yeast has a point centromere, which is defined by a 125-bp sequence and is occupied by a single CENP-A containing nucleosome. There are three conserved elements in budding yeast CEN DNA: CDE (centromere DNA element) I, CDEII (AT riche element) and CDEIII
2. Schizosaccharomyces pombe and human have regional centromeres that are flanked by pericentromeric heterochromatin. In Schizosaccharomyces pombe, a multiple number of CENP-A (Cnp1)-containing nucleosomes are assembled onto unique sequences (cnt – centromeric sequence and imr – innermost repeats). Pericentric,heterochromatic repeats are called “outer repeats” (otr)

3. Human centromeres consist of α-satellite DNA arranged in tandem into higher order repeats (each arrow), and some α-satellite DNA contains CENP-B binding sites (CENP-B box). CENP-A localizes to a portion of these arrays. The number of microtubule attachment sites also varies among organisms. CENP-B is a highly conserved centromere protein in mammals and binds to a 17-bp motif in a CENP-B box. It has been shown that α-satellite DNA with a CENP-B box is responsible for de novo centromere assembly in human somatic cells.

Satellite DNA consists of very large arrays of tandemly repeating, non-coding DNA. Satellite DNA is the main component of functional centromeres, and form the main structural constituent of heterochromatin.

Ade6 mutants in *S.pombe* (ADE2 in S. cerevisiae) generate red-colored intermediates of the adenine biosynthesis pathway


P-ribosyl-PP → P-ribosylamine → P-ribosylglycinamide

ADE4

ADE5

ADE8

Red color

P-ribosylamino imidazole (AIR) ← P-ribosylformyl glycinamidine ← P-ribosylformyl glycinamide

ADE7

ADE6

S.pombe: ADE6

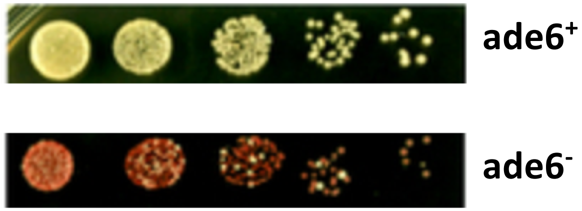
ADE2

P-ribosylamino imidazolecarboxylate (CAIR) → P-ribosylsuccino carboxamide aminoimidazole (SAICAIR) → → adenine

ADE1

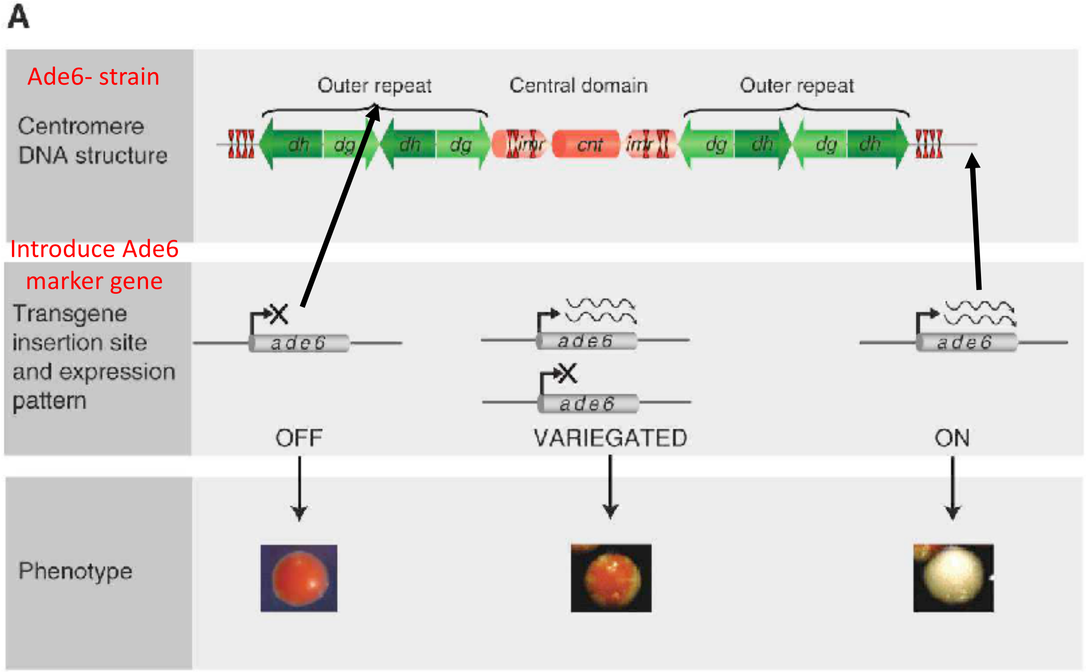
The adenine biosynthesis pathway in <u>S.cerevisiae</u>

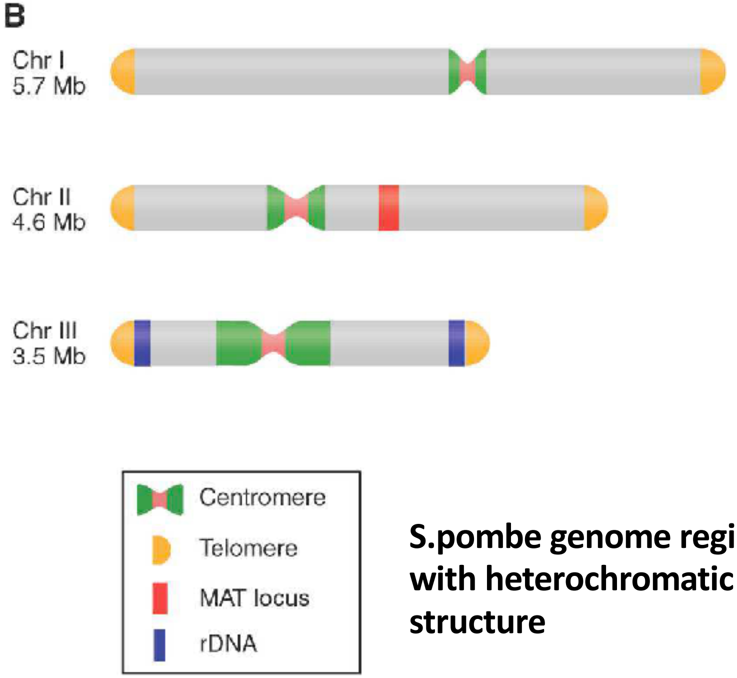
S. pombe



Strategy:
Generate S.pombe strain with Ade6 mutation.
Re-introduce Ade6 maker into the genome.
If located in euchroamtin: White colony
If located in pericentric heterocrhomatin: Red colony

Centromere structure in S.pombe



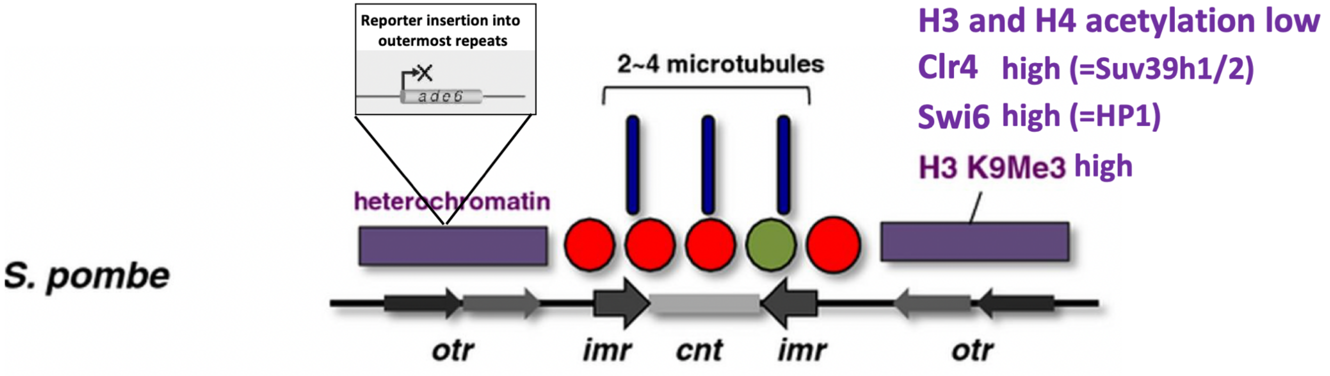


S.pombe genome regions with heterochromatic structure

Figure 2. Distinct outer repeat heterochromatin and central kinetochore domains at fission yeast centromeres. (*A*, *top*) Representation of a fission yeast centromere. The central domain (pink, kinetochore) is composed of *imr* and *cnt* elements, the outer repeats contain transcribed *dg* and *dh* repeats (green, heterochromatin). All three centromeres have a similar overall arrangement; however, the number of outer repeats differs: *cen1* (40 kb) has two, *cen2* (65 kb) has three, and *cen3* (110 kb) has approximately 13. Clusters of transfer RNA (tRNA) genes (double arrowheads) occur in the *imr* region and at the extremities of all three centromeres. (*Middle*) Schematically shows transcription patterns of marker genes placed within the outer repeats, central domain, or beyond the centromere. (*Bottom*) Images showing the phenotype of *S. pombe* colonies of *ade6*$^{+}$ transgenics inserted at various sites within the centromere. Cells expressing *ade6*$^{+}$ from a transgene inserted in sequences outside the centromere form white colonies. When *ade6*$^{+}$ is inserted at sites within the outer repeats, expression is silenced and red colonies are formed. Expression of *ade6*$^{+}$ from the central domain is typically variegated, resulting in red, white, and sectored colonies. (*B*) A schematic representation of *S. pombe* chromosomes. The three chromosomes are depicted showing the four main regions of heterochromatin: centromere, telomere, *mat2/3*, and rDNA regions.

Ade6 reporter gene system to measure chromatin structure:
When expressed - euchromatin: white colony
When silenced - heterochromatin: red colony
Ura4 reporter system: Ura4 converts agent to toxin →
cell dies when Ura3 is expressed

Centromere structure in S.pombe

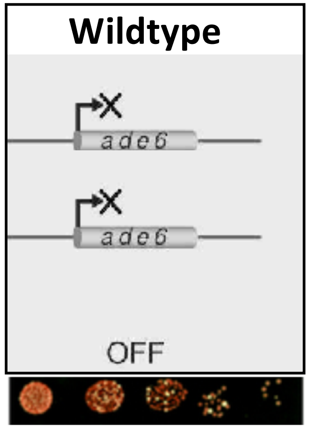


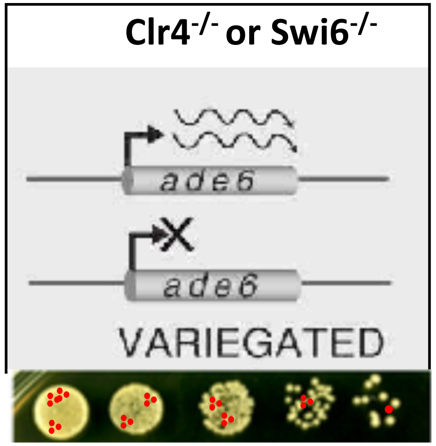
H3 and H4 acetylation low

Clr4 high (=Suv39h1/2)

Swi6 high (=HP1)

H3 K9Me3 high

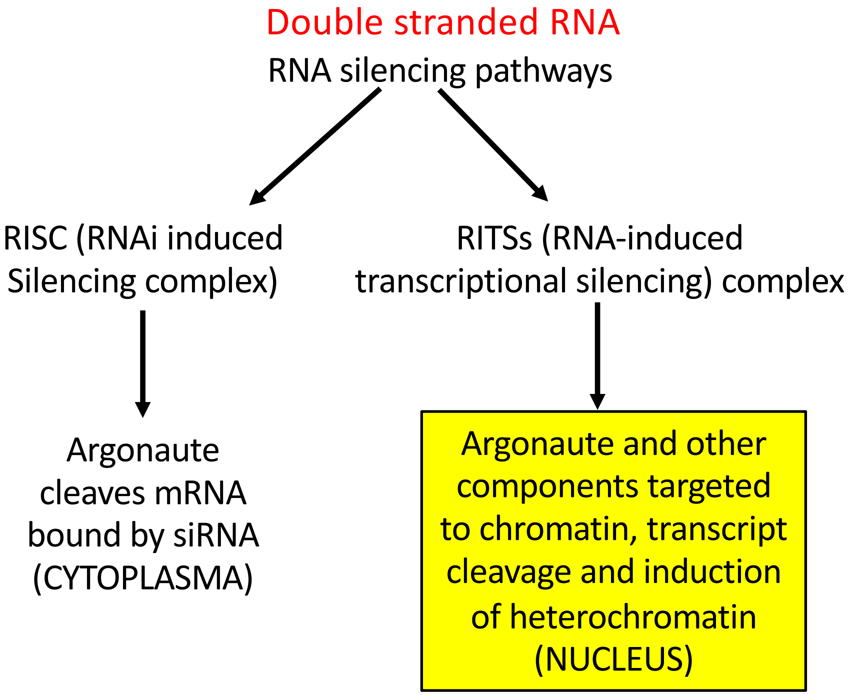




How is heterochromatin recruited
To centromeres??

BUT ALSO IN RNAi MUTANTs (for example Dicer $^{-/-}$)

RNAi silencing pathways



Double stranded RNA

RNA silencing pathways

RISC (RNAi induced Silencing complex)

Argonaute cleaves mRNA bound by siRNA (CYTOPLASMA)

RITSs (RNA-induced transcriptional silencing) complex

Argonaute and other components targeted to chromatin, transcript cleavage and induction of heterochromatin (NUCLEUS)

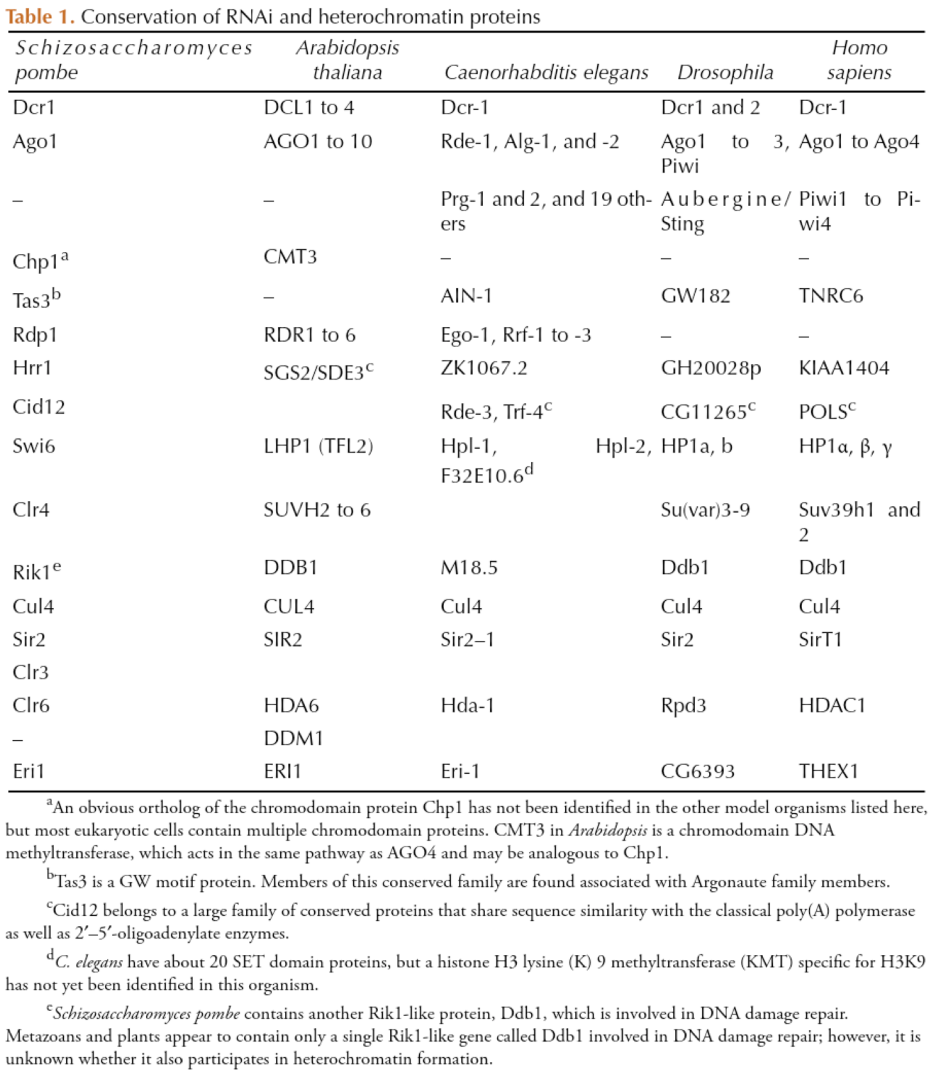


Table 1. Conservation of RNAi and heterochromatin proteins

Schizosaccharomyces pombe	*Arabidopsis thaliana*	*Caenorhabditis elegans*	*Drosophila*	*Homo sapiens*
Dcr1	DCL1 to 4	Dcr-1	Dcr1 and 2	Dcr-1
Ago1	AGO1 to 10	Rde-1, Alg-1, and -2	Ago1 to 3, Piwi	Ago1 to Ago4
–	–	Prg-1 and 2, and 19 others	Aubergine/Sting	Piwi1 to Piwi4
Chp1[a]	CMT3	–	–	–
Tas3[b]	–	AIN-1	GW182	TNRC6
Rdp1	RDR1 to 6	Ego-1, Rrf-1 to -3	–	–
Hrr1	SGS2/SDE3[c]	ZK1067.2	GH20028p	KIAA1404
Cid12		Rde-3, Trf-4[c]	CG11265[c]	POLS[c]
Swi6	LHP1 (TFL2)	Hpl-1, Hpl-2, F32E10.6[d]	HP1a, b	HP1α, β, γ
Clr4	SUVH2 to 6		Su(var)3-9	Suv39h1 and 2
Rik1[e]	DDB1	M18.5	Ddb1	Ddb1
Cul4	CUL4	Cul4	Cul4	Cul4
Sir2	SIR2	Sir2–1	Sir2	SirT1
Clr3				
Clr6	HDA6	Hda-1	Rpd3	HDAC1
–	DDM1			
Eri1	ERI1	Eri-1	CG6393	THEX1

[a] An obvious ortholog of the chromodomain protein Chp1 has not been identified in the other model organisms listed here, but most eukaryotic cells contain multiple chromodomain proteins. CMT3 in *Arabidopsis* is a chromodomain DNA methyltransferase, which acts in the same pathway as AGO4 and may be analogous to Chp1.

[b] Tas3 is a GW motif protein. Members of this conserved family are found associated with Argonaute family members.

[c] Cid12 belongs to a large family of conserved proteins that share sequence similarity with the classical poly(A) polymerase as well as 2′–5′-oligoadenylate enzymes.

[d] *C. elegans* have about 20 SET domain proteins, but a histone H3 lysine (K) 9 methyltransferase (KMT) specific for H3K9 has not yet been identified in this organism.

[e] *Schizosaccharomyces pombe* contains another Rik1-like protein, Ddb1, which is involved in DNA damage repair. Metazoans and plants appear to contain only a single Rik1-like gene called Ddb1 involved in DNA damage repair; however, it is unknown whether it also participates in heterochromatin formation.

Sources of RNAs for siRNA generation

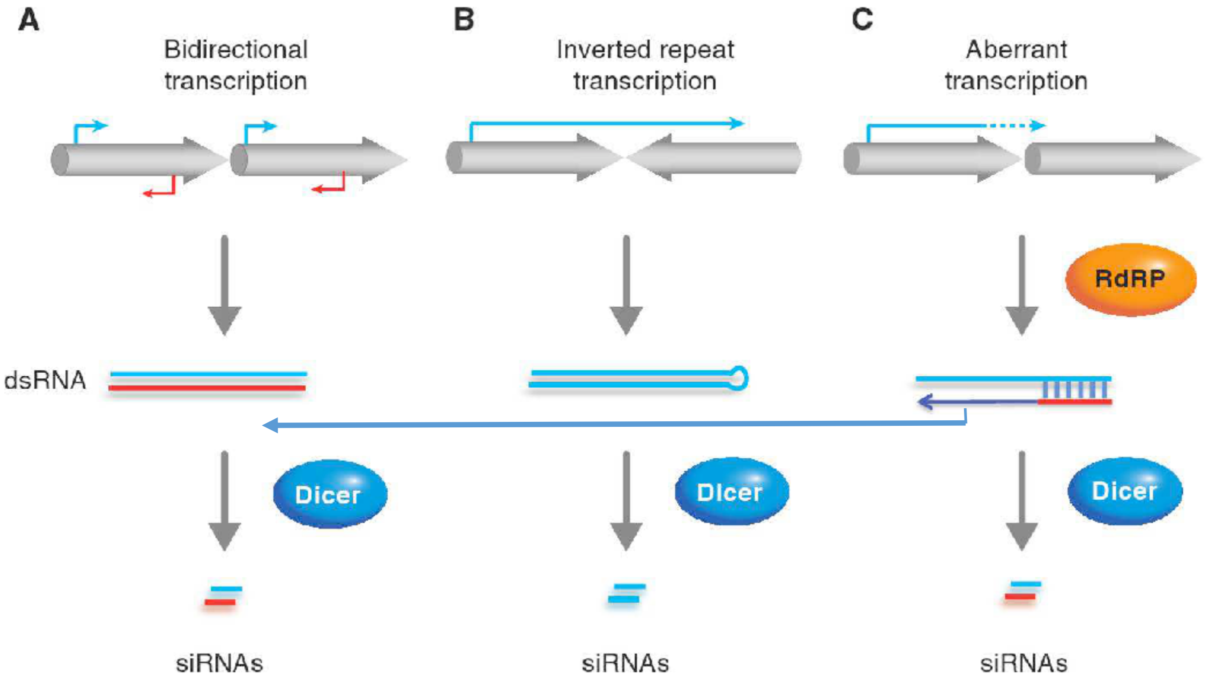


Figure 2. Pathways for the biogenesis of primary small RNAs that mediate silencing. (*A*) Bidirectional transcription has been observed at the *Schizosaccharomyces pombe* centromeric repeats and the *cenH* region of the silent mating-type locus and may provide a dsRNA substrate for the Dicer ribonuclease. (*B*) Transcription through inverted repeats found in many plant and animal cells can potentially produce dsRNA. (*C*) Transcription of aberrant RNAs that may lack proper processing signals may trigger dsRNA synthesis by RNA-dependent RNA polymerases (RdRPs). (*D*) Transcription from several driver loci gives rise to Piwi-associated small RNAs (piRNAs) that silence dispersed transposons. Piwi proteins together with other ribonucleases, which are not fully defined (represented by the gray dotted line), mediate primary piRNA generation.

The RNAi machinery degrades RNAs derived from centromeres

Volpe et al. 2002
Science

Step1: $ura4^-$ strain

***$ura4^-$* renders cells resistant to 5-fluoroorotic acid (5-FOA)**

Step2: insert a URA4 mini gene into chromosomal arm
(produces a short, non-functional URA4 RNA) – always expressed

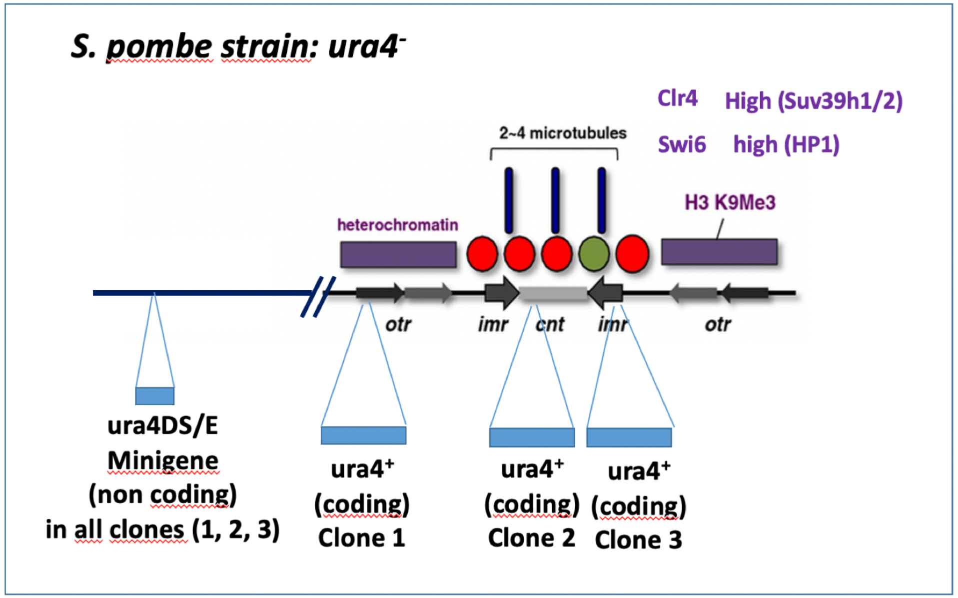
Step3: insert a URA4 full length gene into centromeric Positions (cnt, imr, otr)
(produces a full length, functional URA4 mRNA (=marker) when expressed)

→ In centromeric positions URA4 marker gene is silenced

→ =5FOA resistant

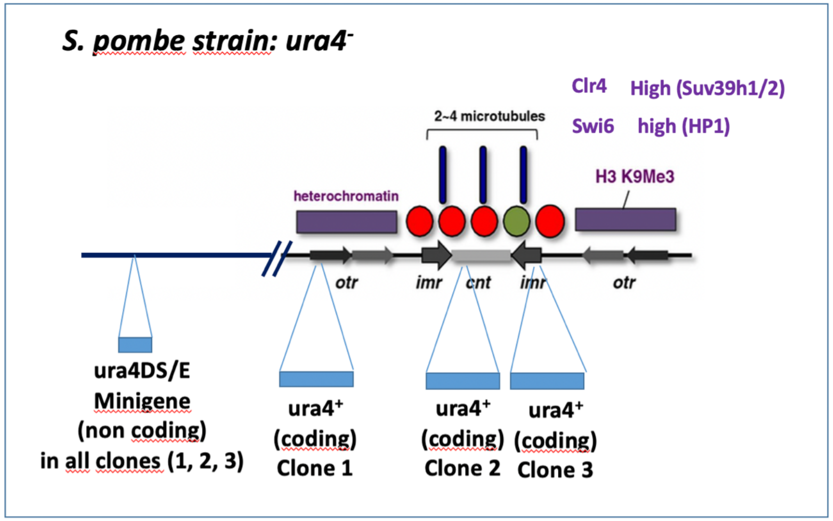
Step4: Elimination of RNAi components and HC components or: random mutagenesis screen

Step5: Interesting yeast strains: 5FOA sensitive; Identify mutated genes



The RNAi machinery degrades RNAs derived from centromeres

Volpe et al. 2002
Science



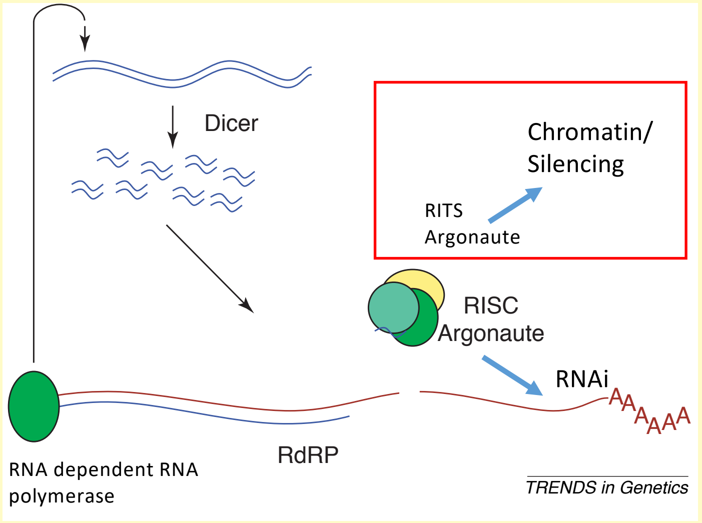
MUTATIONAL SCREEN

STEP 1: Readout: FAO re-sensitzation

STEP 2: Molecular and functional dissection

Components of the RNAi machinery in S.pombe:

1. **Dicer: cleavage of dsRNA into siRNAs**
2. **RISC/RITS with Argonaute (RNase domain): target RNA cleavage guided by siRNA**
3. **RNA dependent RNA polymerase: RdRP: reverse transcription of RNA, guided by siRNA → creates dsRNA → RNAi amplification**



TRENDS in Genetics

Figure I. RNA interference. A simplified scheme illustrating how Dicer, Argonaute and RNA-dependent RNA polymerase (RdRP) participate in RNA silencing through RNA interference. Dicer processes dsRNA into 20–24 nt siRNA, which guides 'slicing' by Argonaute, via the RNA-induced silencing complex (RISC). Slicing and/or priming by siRNA could guide RdRP, which regenerates dsRNA, amplifying the cycle.

or RITS

The RNAi machinery degrades RNAs derived from centromeres

Volpe et al. 2002 Science

Step1: ura4$^-$ strain

ura4$^-$ renders cells resistant to 5-fluoroorotic acid (5-FOA)

Step2: insert a URA4 mini gene (URA DS/E into chromosomal arm
(produces a short, non-functional URA4 RNA)

Step3: insert a URA4 full length gene into centromeric Positions (cnt, imr, otr)
(produces a full length, functional URA4 mRNA (=marker) when expressed)

- → In centromeric positions URA4 marker gene is silenced
- → =5FOA resistant

Step4: Elimination of RNAi components and HC components
Or random mutagenesis screen

Step5: Interesting yeast strains: 5FOA sensitive; identify mutated gene(s)

Model system: 3 URA reporter genes:

1. URA4 reporter in centromere regions (otd, imr, cnt) → ura4+
2. URA4 minigene (shorter then centromeric URA4 located in a chromosomal arm (ura4 DS/E)

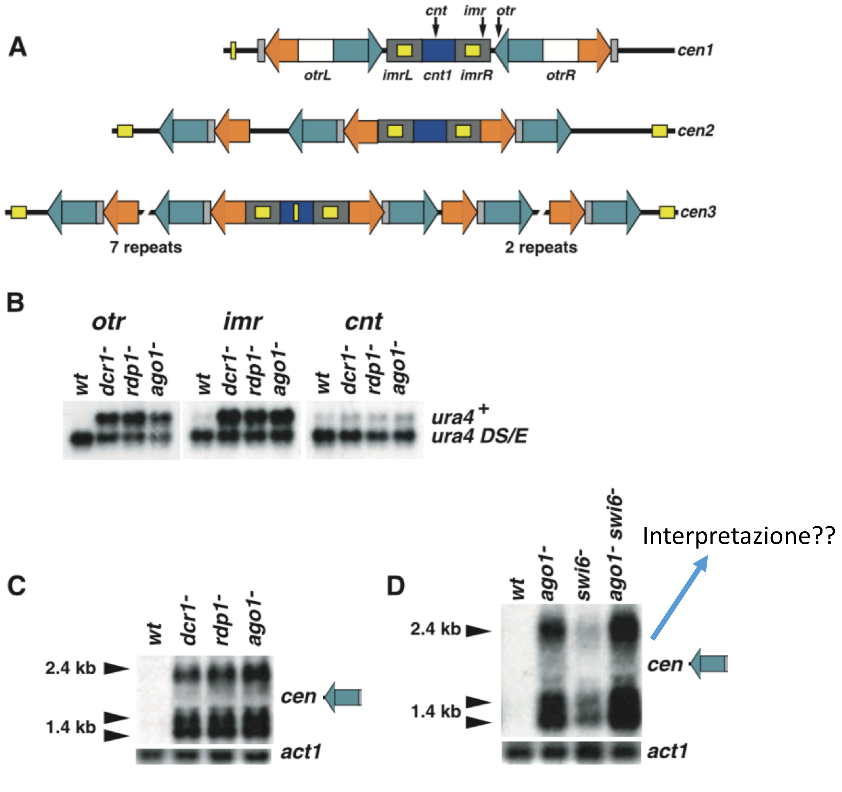


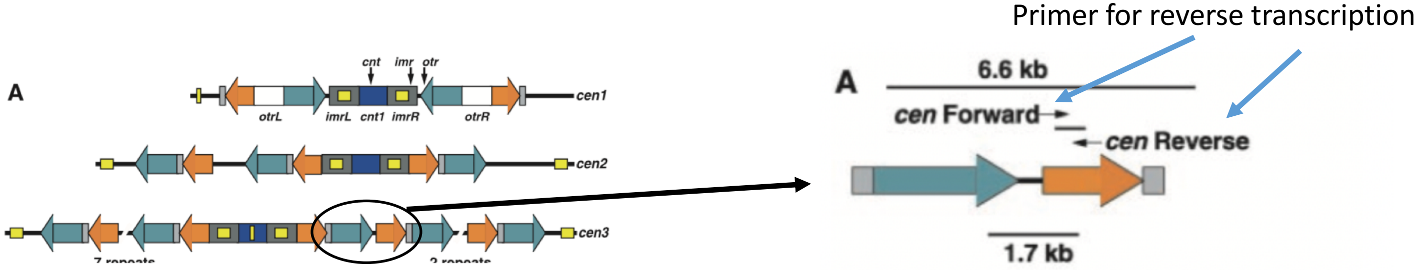
Fig. 1. Centromeric silencing is relieved in $ago1^-$, $dcr1^-$, and $rdp1^-$ mutant strains compared to wild type. Diagram of the three *S. pombe* centromeres (**A**) including locations of $ura4^+$ transgenes as well as outermost (*otr*), innermost (*imr*), and central (*cnt*) centromeric regions (*24*, *25*). Conserved *dg* (green) and *dh* (red) repeats are indicated as arrows. Regions containing one or more tRNA genes are indicated by yellow boxes. Northern analysis (**B**) of RNA transcripts transcribed from centromeric $ura4^+$ transgenes and a $ura4^+$ (*DS/E*) mini-gene located on the chromosome arm. Transcripts derived from centromeric repeats were detected by Northern blotting (**C** and **D**) using probes specific for *dg* centromeric repeats (Fig. 5).

Interpretazione??

B: loss of RNAi machinery results in increases URA4 reporter expression
C: loss of RNAi machinery results in increased centromeric transcript levels
D: loss of RNAi and loss of Swi6 results in further increased centromeric transcript levels

The RNAi machinery degrades RNAs derived from centromeres

What strand is transribed to produce RNA as guide for heterochromatin production??

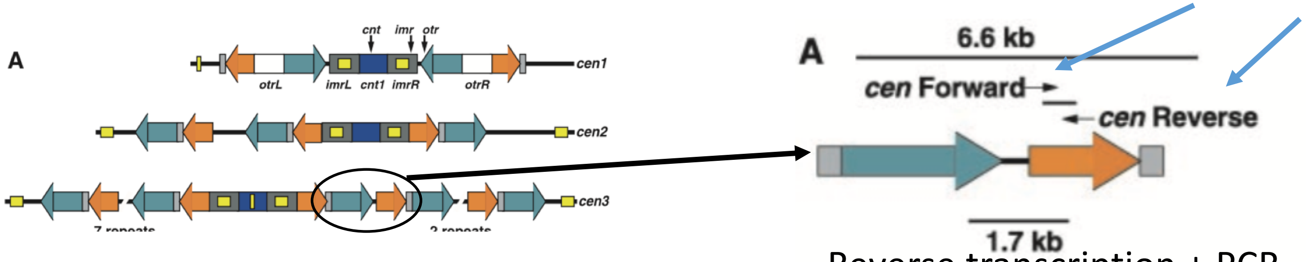


Primer for reverse transcription

Reverse transcription + PCR

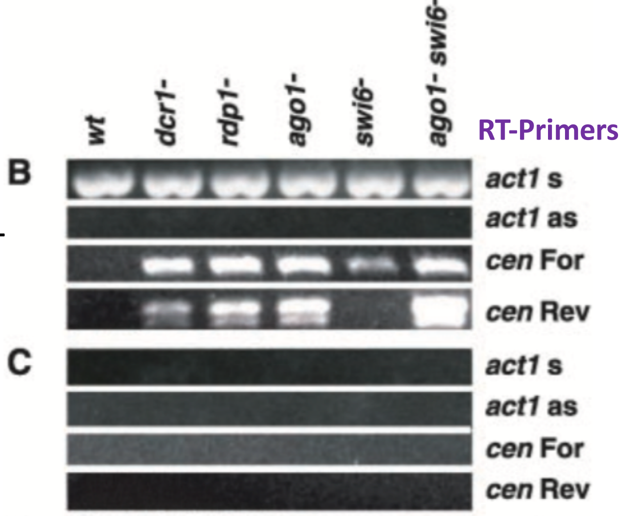
Fig. 2. Centromeric *otr* transcripts are both transcriptionally and posttranscriptionally regulated. (**A**) A schematic diagram showing the direction of transcription of forward and reverse transcripts corresponding to the *dh* repeat. Strand-specific RT-PCR analysis was performed in the presence (**B**) or absence (**C**) of reverse transcriptase. Samples were incubated with primers from the *dh* repeat complementary to either the forward (*cen* For) or reverse (*cen* Rev) centromeric transcripts in first strand cDNA synthesis reaction (primer locations are summarized in Fig. 5). Both primers were present in subsequent cycles of PCR amplification after heat inactivation of the reverse transcriptase. Treatment of control reactions lacking reverse transcriptase (C) was identical except that these samples were not subjected to first strand synthesis. Strand-specific control reactions were also conducted using primers specific for *act1* sense (*act1* s) or *act1* antisense (*act1* as) transcripts. Strand-specific analysis of nascent RNA transcripts was performed by nuclear run-on assay (**D**). Radiolabeled nascent RNA purified from mutant and wild-type strains was hybridized to nylon membranes containing strand-specific probes made using the same primer pair as in (A). These probes recognized either forward (*cen* For) or reverse (*cen* Rev) centromeric transcripts. Control probes recognized either sense (*act1* s) or antisense (*act1* as) actin transcripts.

The RNAi machinery degrades RNAs derived from centromeres



Primer for reverse transcription

Reverse transcription + PCR



RT-Primers

Total RNA prep.

+RT

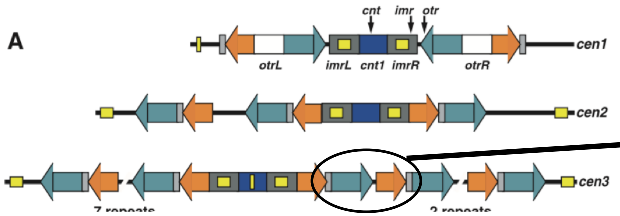
-RT

In wt cells production of centromeric reverse transcripts

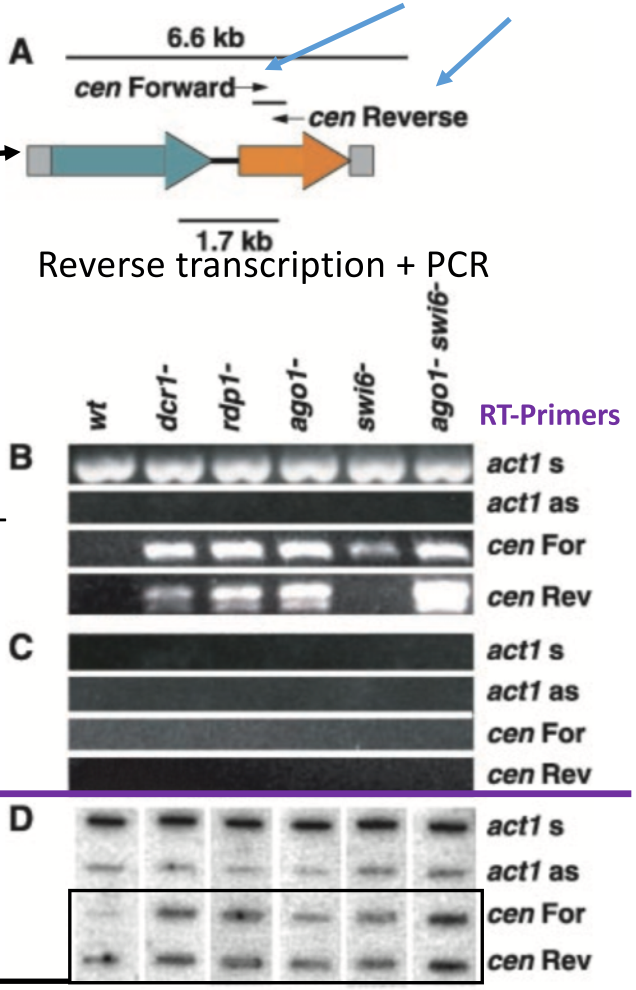
Loss of RNAi machinery results in an increase of *forward* – *reverse* RNAs derived from pericentric repeats

Fig. 2. Centromeric *otr* transcripts are both transcriptionally and posttranscriptionally regulated. **(A)** A schematic diagram showing the direction of transcription of forward and reverse transcripts corresponding to the *dh* repeat. Strand-specific RT-PCR analysis was performed in the presence **(B)** or absence **(C)** of reverse transcriptase. Samples were incubated with primers from the *dh* repeat complementary to either the forward (*cen* For) or reverse (*cen* Rev) centromeric transcripts in first strand cDNA synthesis reaction (primer locations are summarized in Fig. 5). Both primers were present in subsequent cycles of PCR amplification after heat inactivation of the reverse transcriptase. Treatment of control reactions lacking reverse transcriptase (C) was identical except that these samples were not subjected to first strand synthesis. Strand-specific control reactions were also conducted using primers specific for *act1* sense (*act1* s) or *act1* antisense (*act1* as) transcripts. Strand-specific analysis of nascent RNA transcripts was performed by nuclear run-on assay **(D)**. Radiolabeled nascent RNA purified from mutant and wild-type strains was hybridized to nylon membranes containing strand-specific probes made using the same primer pair as in (A). These probes recognized either forward (*cen* For) or reverse (*cen* Rev) centromeric transcripts. Control probes recognized either sense (*act1* s) or antisense (*act1* as) actin transcripts.

The RNAi machinery degrades RNAs derived from centromeres



Primer for reverse transcription



Reverse transcription + PCR

RT-Primers

Total RNA prep.

+RT

-RT

Nascent RNA analysis:

Preparation of nascent RNA
RNA that is in progressing transcription.
RNA still associated at chromatin.
WT: Rev strand major RNA species
Mutants: more Rev and and also For strands RNA become visible → processing in wild-type conditions

Blotting using radioactive centromere probe (detects both strands)

wt: long rev transcripts
mutants: rev and for transcripts

Fig. 2. Centromeric *otr* transcripts are both transcriptionally and posttranscriptionally regulated. **(A)** A schematic diagram showing the direction of transcription of forward and reverse transcripts corresponding to the *dh* repeat. Strand-specific RT-PCR analysis was performed in the presence **(B)** or absence **(C)** of reverse transcriptase. Samples were incubated with primers from the *dh* repeat complementary to either the forward (*cen* For) or reverse (*cen* Rev) centromeric transcripts in first strand cDNA synthesis reaction (primer locations are summarized in Fig. 5). Both primers were present in subsequent cycles of PCR amplification after heat inactivation of the reverse transcriptase. Treatment of control reactions lacking reverse transcriptase (C) was identical except that these samples were not subjected to first strand synthesis. Strand-specific control reactions were also conducted using primers specific for *act1* sense (*act1* s) or *act1* antisense (*act1* as) transcripts. Strand-specific analysis of nascent RNA transcripts was performed by nuclear run-on assay **(D)**. Radiolabeled nascent RNA purified from mutant and wild-type strains was hybridized to nylon membranes containing strand-specific probes made using the same primer pair as in (A). These probes recognized either forward (*cen* For) or reverse (*cen* Rev) centromeric transcripts. Control probes recognized either sense (*act1* s) or antisense (*act1* as) actin transcripts.

Loss of the RNAi machinery results in the loss of H3K9me3 at outermost repeats

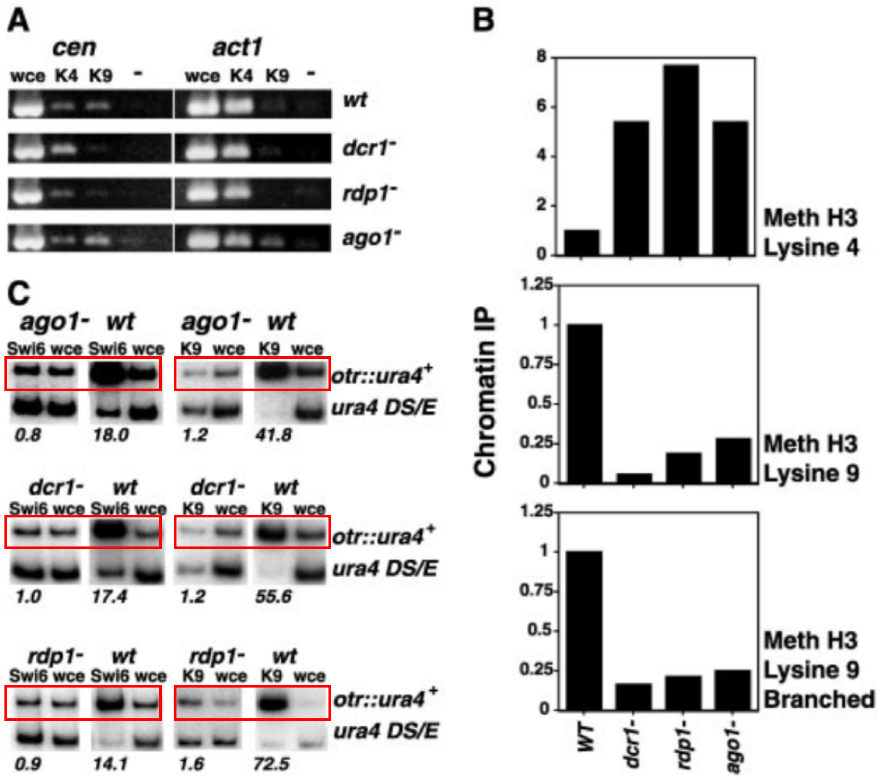
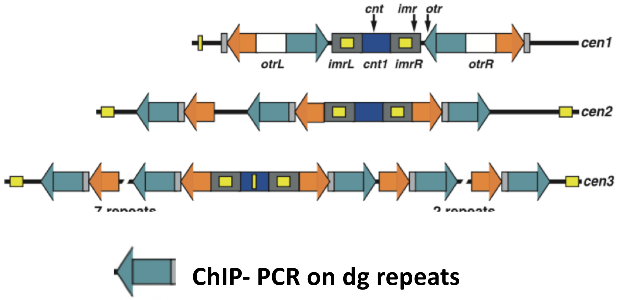


Fig. 3. Chromatin structure at centromeric repeats is altered in *$ago1^-$*, *$dcr1^-$*, and *$rdp1^-$* mutants. Chromatin immunoprecipitation (ChIP) was performed on extracts from formaldehyde fixed mutant or wild-type cells. DNA fragments purified from whole-cell extracts (wce) or co-precipitated with antibodies to K4 or K9 histone H3 were amplified by PCR using primers specific for centromeric *dg* repeats (Fig. 5) or *act1* (**A**). Quantitation was performed using quantitative PCR (**B**) and normalized to actin (K4) and mating type region (K9) controls, which were amplified from the same ChIPs in each genotype (*21*). DNA fragments from wce or co-precipitated with antibodies to Swi6 or K9 histone H3 were amplified by multiplex PCR using *ura4*-specific primers (**C**). These primers amplify both a *$ura4^+$* transgene located in the *otr* of *cen1* (*otr::$ura4^+$*) and the *ura4 DS/E* minigene (*ura4 DS/E*) located on the chromosome arm. Relative levels were estimated using a FUJI phosphoimager and are indicated below each lane.



ChIP- PCR on dg repeats

- **RNAi mutants:**

Loss of H3K9me3 methylation
Loss of Swi4 recruitment

Loss of the RNAi machinery results in the loss of H3K9me3 at outermost repeats

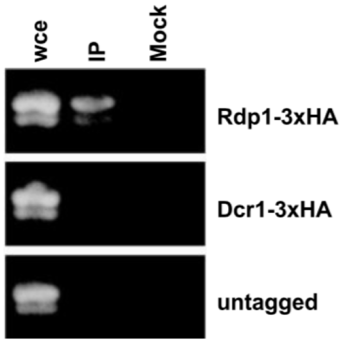
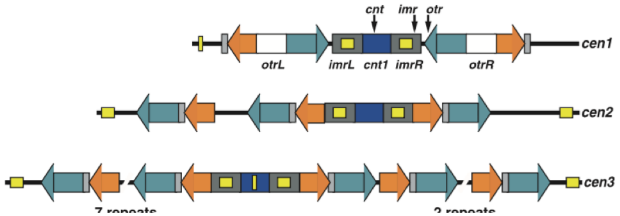


Fig. 4. Rdp1, but not Dcr1, is bound to centromeric chromatin. Chromatin immunoprecipitation (ChIP) was performed on extracts derived from triple-HA tagged Dcr1 or Rdp1 strains. DNA fragments from wce or co-precipitated with antibodies raised against the triple-HA tag were amplified by PCR using centromere-specific primers and compared to untagged strains. Mock reactions were identical except without the addition of primary antibody.



ChIP- PCR on dg repeats

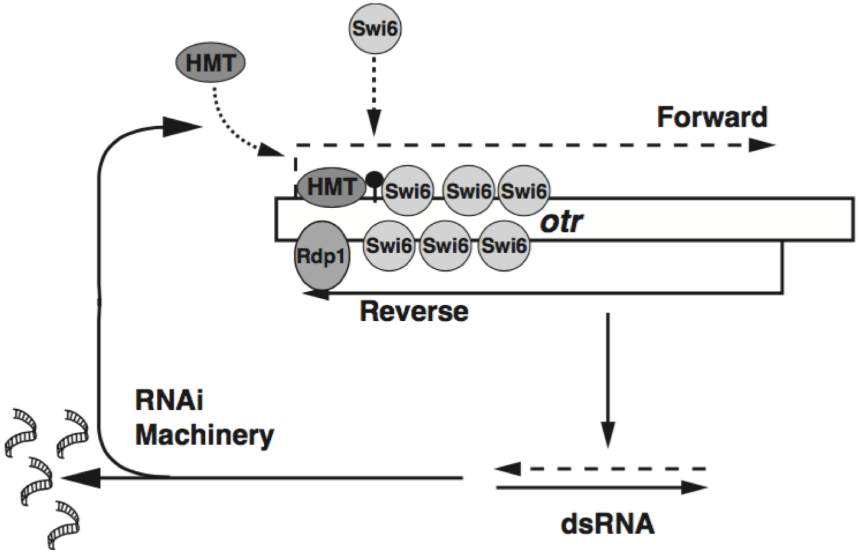


Fig. 6. The RNAi machinery is required for the initiation and maintenance of the heterochromatic state of centromeric repeats. Reverse strand centromeric transcription occurs in wild-type cells and is degraded posttranscriptionally by the RNAi machinery. Low-level transcription from the forward strand and/or amplification by Rdp1 results in generation of dsRNA, which is converted to siRNA by RNAi. Rdp1, bound to the chromatin, promotes targeting of histone modifications to specific sequences via siRNA, resulting in maintenance of the heterochromatic state (HMT, histone methyl transferase).

Rdp1: RNA synthesis performed by means of a primer-independent (de novo) transcription using an RNA molecule as template

4 critical steps in heterochromatin formation at centromeres in S. pombe

A combination of genetic and biochemical experiments identify RNA dependent RNA silencing pathways at S.pombe centromere

1. Production of pericentric sense - anti-sense RNA in outermost regions

2. RNAi mediated degradation of pericentromeric dsRNA

3. Amplification of siRNA loop by RNA dependent RNA polymerase

4. Heterochromatinization

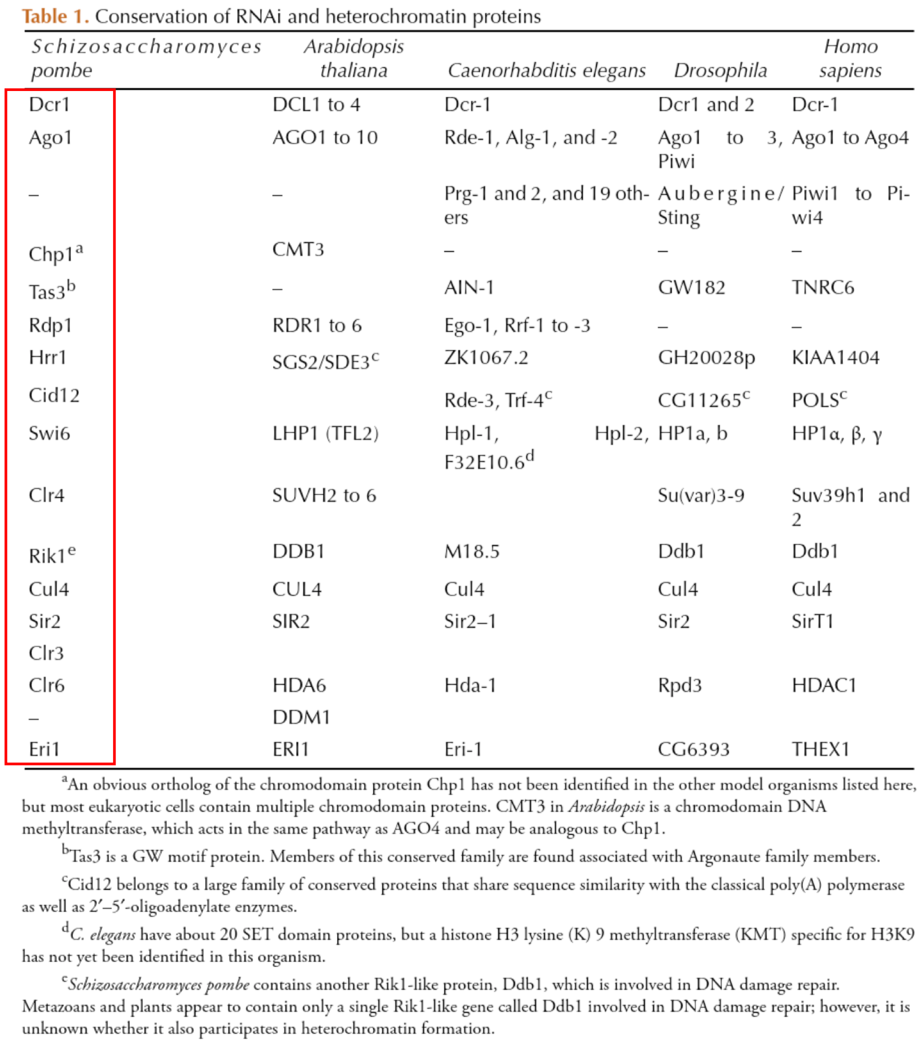


Table 1. Conservation of RNAi and heterochromatin proteins

Schizosaccharomyces pombe	*Arabidopsis thaliana*	*Caenorhabditis elegans*	*Drosophila*	*Homo sapiens*
Dcr1	DCL1 to 4	Dcr-1	Dcr1 and 2	Dcr-1
Ago1	AGO1 to 10	Rde-1, Alg-1, and -2	Ago1 to 3, Piwi	Ago1 to Ago4
–	–	Prg-1 and 2, and 19 others	Aubergine/Sting	Piwi1 to Piwi4
Chp1[a]	CMT3	–	–	–
Tas3[b]	–	AIN-1	GW182	TNRC6
Rdp1	RDR1 to 6	Ego-1, Rrf-1 to -3	–	–
Hrr1	SGS2/SDE3[c]	ZK1067.2	GH20028p	KIAA1404
Cid12		Rde-3, Trf-4[c]	CG11265[c]	POLS[c]
Swi6	LHP1 (TFL2)	Hpl-1, Hpl-2, F32E10.6[d]	HP1a, b	HP1α, β, γ
Clr4	SUVH2 to 6		Su(var)3-9	Suv39h1 and 2
Rik1[e]	DDB1	M18.5	Ddb1	Ddb1
Cul4	CUL4	Cul4	Cul4	Cul4
Sir2	SIR2	Sir2–1	Sir2	SirT1
Clr3				
Clr6	HDA6	Hda-1	Rpd3	HDAC1
–	DDM1			
Eri1	ERI1	Eri-1	CG6393	THEX1

[a] An obvious ortholog of the chromodomain protein Chp1 has not been identified in the other model organisms listed here, but most eukaryotic cells contain multiple chromodomain proteins. CMT3 in *Arabidopsis* is a chromodomain DNA methyltransferase, which acts in the same pathway as AGO4 and may be analogous to Chp1.

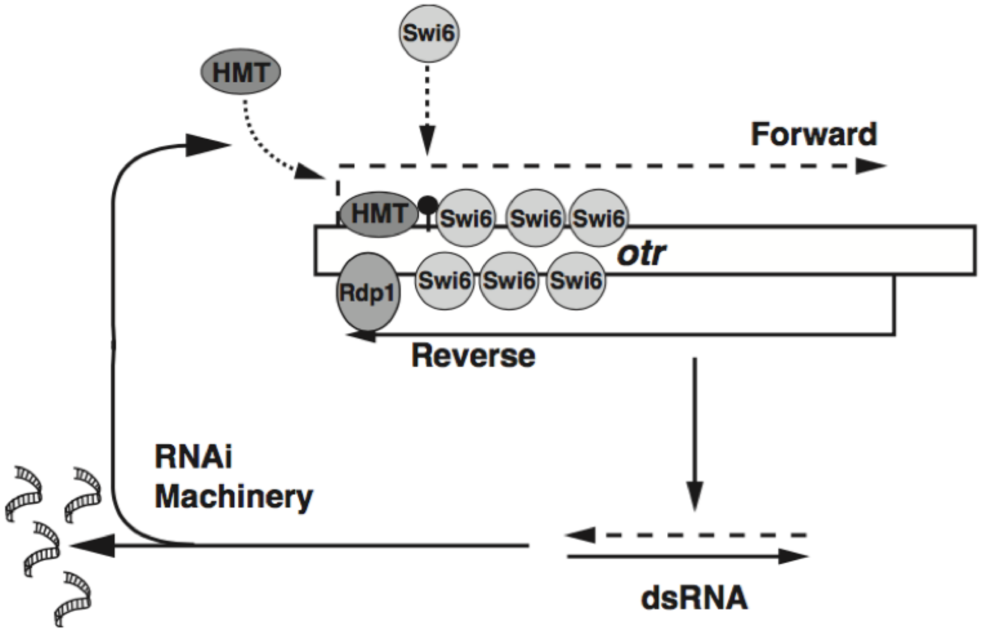
[b] Tas3 is a GW motif protein. Members of this conserved family are found associated with Argonaute family members.

[c] Cid12 belongs to a large family of conserved proteins that share sequence similarity with the classical poly(A) polymerase as well as 2′–5′-oligoadenylate enzymes.

[d] *C. elegans* have about 20 SET domain proteins, but a histone H3 lysine (K) 9 methyltransferase (KMT) specific for H3K9 has not yet been identified in this organism.

[e] *Schizosaccharomyces pombe* contains another Rik1-like protein, Ddb1, which is involved in DNA damage repair. Metazoans and plants appear to contain only a single Rik1-like gene called Ddb1 involved in DNA damage repair; however, it is unknown whether it also participates in heterochromatin formation.

Step 1. Production of pericentric sense anti-sense RNA in outermost regions



1. Predominant transcription of reverse strand
2. Low level of sense strand transcription
3. Amplification by Rdp
4. Processing by Dicer

Dissection of RNAi mediated formation of heterochromatin

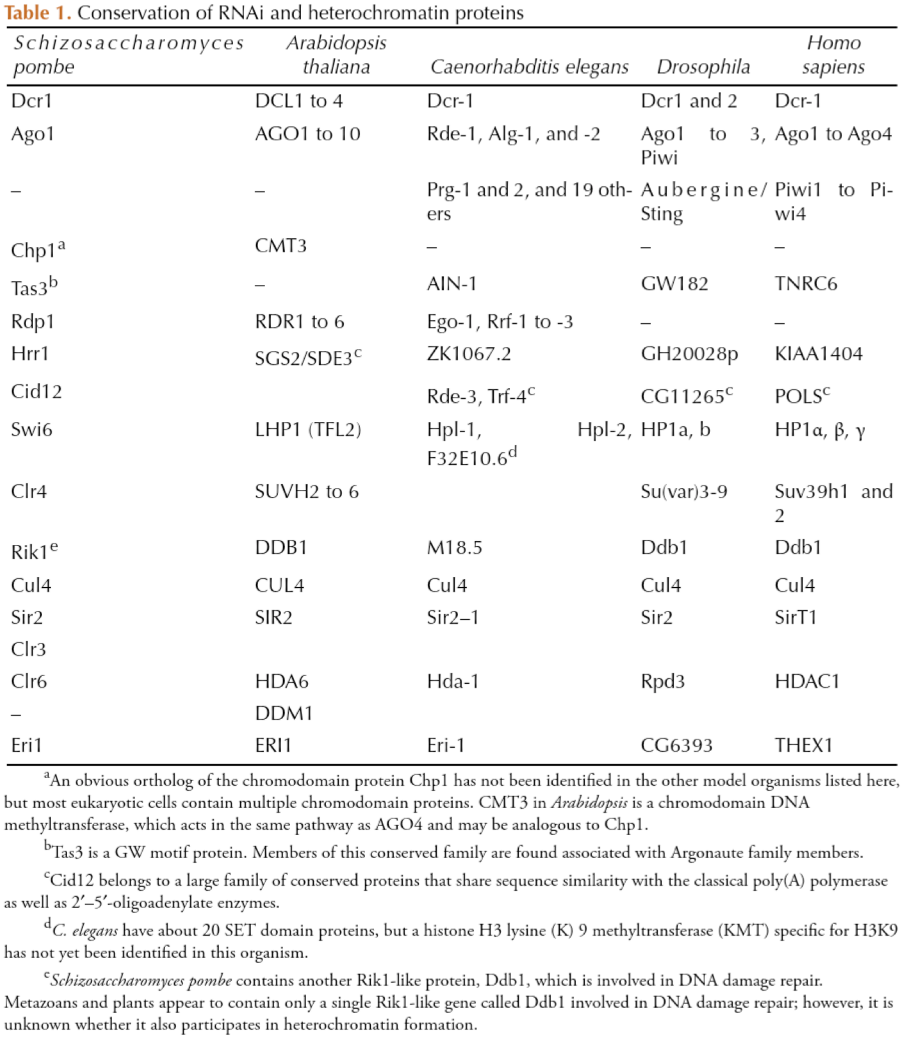


Table 1. Conservation of RNAi and heterochromatin proteins

Schizosaccharomyces pombe	*Arabidopsis thaliana*	*Caenorhabditis elegans*	*Drosophila*	*Homo sapiens*
Dcr1	DCL1 to 4	Dcr-1	Dcr1 and 2	Dcr-1
Ago1	AGO1 to 10	Rde-1, Alg-1, and -2	Ago1 to 3, Piwi	Ago1 to Ago4
–	–	Prg-1 and 2, and 19 others	Aubergine/Sting	Piwi1 to Piwi4
Chp1[a]	CMT3	–	–	–
Tas3[b]	–	AIN-1	GW182	TNRC6
Rdp1	RDR1 to 6	Ego-1, Rrf-1 to -3	–	–
Hrr1	SGS2/SDE3[c]	ZK1067.2	GH20028p	KIAA1404
Cid12		Rde-3, Trf-4[c]	CG11265[c]	POLS[c]
Swi6	LHP1 (TFL2)	Hpl-1, Hpl-2, F32E10.6[d]	HP1a, b	HP1α, β, γ
Clr4	SUVH2 to 6		Su(var)3-9	Suv39h1 and 2
Rik1[e]	DDB1	M18.5	Ddb1	Ddb1
Cul4	CUL4	Cul4	Cul4	Cul4
Sir2	SIR2	Sir2–1	Sir2	SirT1
Clr3				
Clr6	HDA6	Hda-1	Rpd3	HDAC1
–	DDM1			
Eri1	ERI1	Eri-1	CG6393	THEX1

[a] An obvious ortholog of the chromodomain protein Chp1 has not been identified in the other model organisms listed here, but most eukaryotic cells contain multiple chromodomain proteins. CMT3 in *Arabidopsis* is a chromodomain DNA methyltransferase, which acts in the same pathway as AGO4 and may be analogous to Chp1.

[b] Tas3 is a GW motif protein. Members of this conserved family are found associated with Argonaute family members.

[c] Cid12 belongs to a large family of conserved proteins that share sequence similarity with the classical poly(A) polymerase as well as 2′–5′-oligoadenylate enzymes.

[d] *C. elegans* have about 20 SET domain proteins, but a histone H3 lysine (K) 9 methyltransferase (KMT) specific for H3K9 has not yet been identified in this organism.

[e] *Schizosaccharomyces pombe* contains another Rik1-like protein, Ddb1, which is involved in DNA damage repair. Metazoans and plants appear to contain only a single Rik1-like gene called Ddb1 involved in DNA damage repair; however, it is unknown whether it also participates in heterochromatin formation.

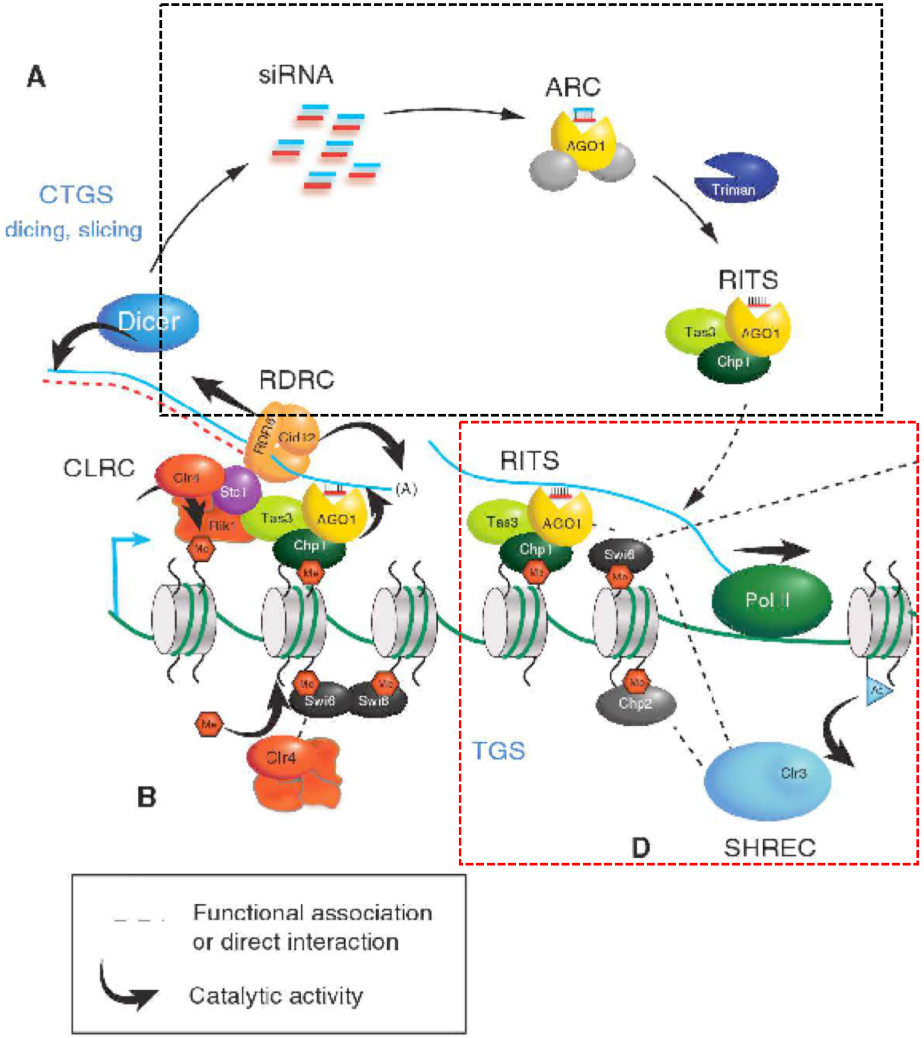
Gentetic approach:

- Make mutation and evaluate transcripts and heterochromatin
- Individuate key players

Biochemical approaches:

- Purify complexes that contain key regulators identified in genetic experiments and new factors
- Functionally validate new proteins
- Check interdependence of complexes

Step 2: RNAi – dsRNA formation – RITS assembly



1. Dicer processed pericentric dsRNAs generating siRNAs (nucleus); generation of the Argonaute siRNA chaperone (ARC) complex

2. siRNA incorporated into the RITS complex consisting of Ago1, Chp1 and Tas3

3. RITS uses siRNA to target nascent transcripts from centromeres (base pairing) and induces heterochromatin formation → localization of complex!

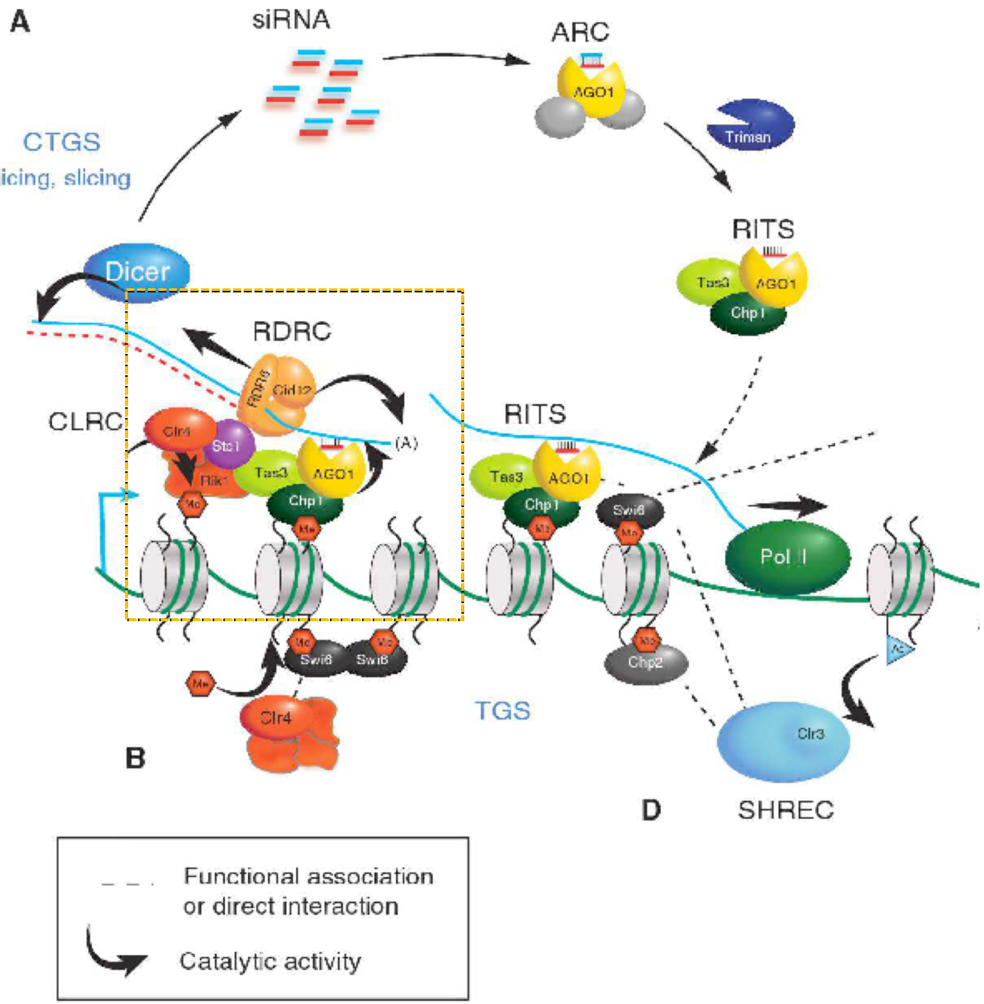
4. Chp1 contains a chromodomain that binds to H3K9me3 that was previously imposed by Clr4

5. Clr4 has a chromodomain that binds H3K9me3 and Clr4 interacts with RITS

6. H3K9me3 and recruitment of Swi6, that stabilizes the RITS complex

7. Chp2 chromodomain protein binds H3K9me3 and recruits the SHREC HDAC complex (**H4K14 specific Clr3 HDAC**) that shuts down transcription in S-Phase

Step 3: RNA dependent RNA Pol activity



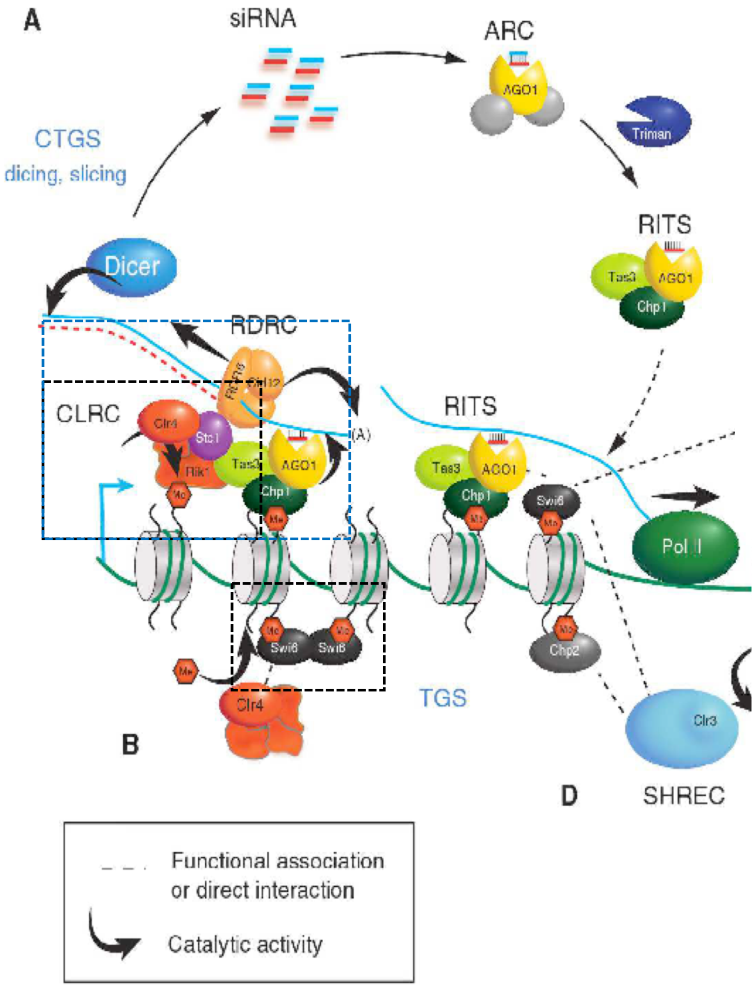
1. Swi6 interacts with Ers1 that binds to the RDRC (RNA directed RNA polymerase complex); Swi6 has affinity for RNA

2. RDRC contains Rdp1 and other components.

3. Rdp1 does not require a primer for dsRNA synthesis. siRNA/RITS gives information for specific interaction of Rdp1 complex with template (RITS: RNA-induced transcriptional silencing complex).

4. RDRC interacts with chromatin at centromere in a siRNA and Clr4 dependent manner –> siRNAs and Clr4 target RDRC to peri-centromeres

Step 4: Heterochromatinization (Clr4, Swi6)



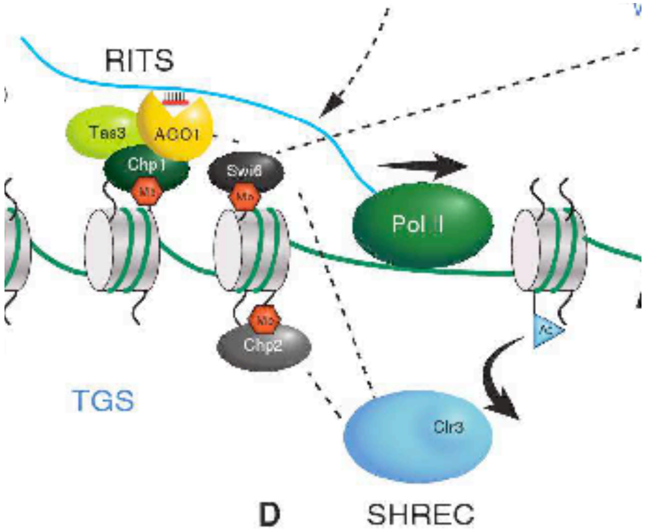
1. Clr4 is required for stabilizing RITS at centromere. → some Clr4-H3K9me3 is required to start heterochromatin formation. *Note: loss of RNAi leads to loss of centromeric silencing → RNAi is required to boost H3K9me3/Swi6*

2. Clr4 is a component of the CLRC complex= “Clr4 methyltransferase complex” (Clr4-Rik1-Cul4) –> essential for heterochromatin formation

3. CLRC and RITS interact with each other via Stc1. *Note: in RNAi-/- cells Stc1 levels are reduced → siRNAs contribute to complex structure that favors the action of Stc1 in ensuring the interaction between RITS and CLRC (ev. via other factors)*

4. Efficient H3K9me3 by Clr4; Swi6 is recruited

A remark: The role of RNA polymerase in centromeric silencing



….complex interplay of many factors

S.pombe: Transcription by RNA-polymerase I is essential for heterochromatin formation

Evidence: Mutations in RNA-Pol in S.pombe:

Rpb7-1: Reduced levels of transcription at pericentric repeats → less siRNAs → loss of silent chromatin

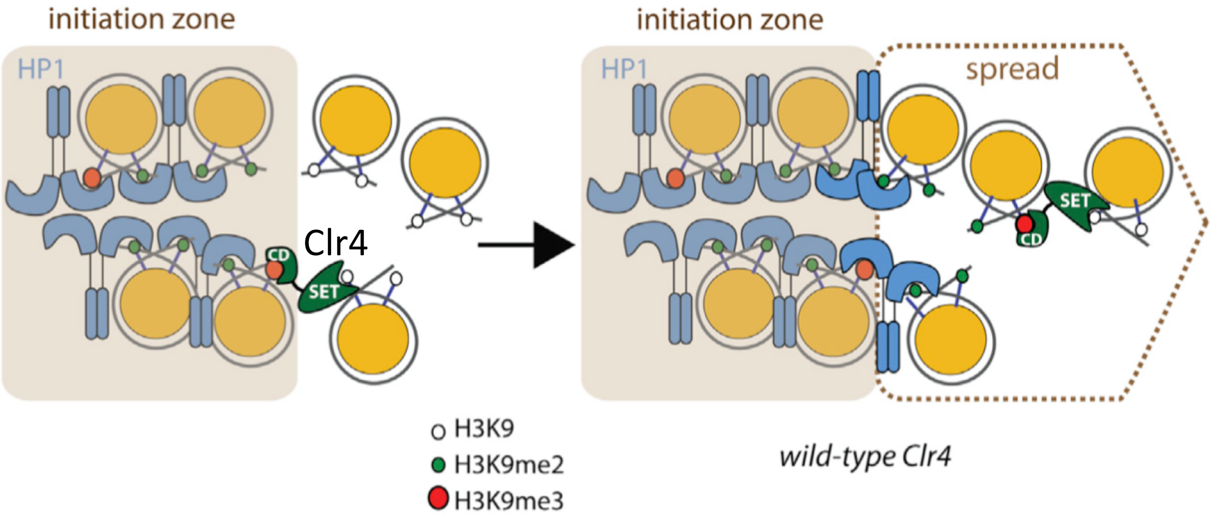
Rpb2-m203: RNA is produced, but not processed into siRNAs → reduced heterochromatin

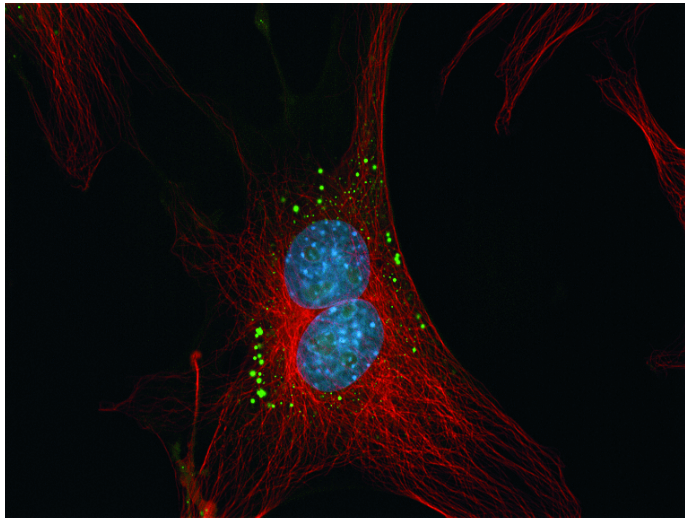
→ RNA polymerase is not exclusively responsible for transcription, but also has a role in integrating transcription with the downstream processing of the produced RNA

→ Remember the RNA-Pol holoenzyme is a huge complex that integrates many proteins: i.e. FACT: chromatin modelling factor that colocalizes to centromeres i.e. Spt6 (histone chaperon): RNA pol interactor that is required to establish H3K9me3

→ precise mechanisms: unknown – the complexity of protein interaction with the RNA-Pol holoenzyme is extremly high

Step 4: Heterochromatinization (Clr4, Swi6)





Clr4 mediates H3K9m3; Clr4 has a chromo-domain that interacts with H3K9me2/3

Swi6 interacts with H3K9me2/3 and forms homodimer with another Swi6;

Clr4 bind H3K9me2 + interacts with Swi6.

Spreading of Clr4/Swi6 along chromatin

Swi6 mediates nucleosome bridging and shortening of nucleosome to nucleosome distance –> efficient chromatin compaction

MECHANISM CONSERVED FROM S.pombe → humans

Clr4	**Suv39h1/h2**
Swi6	**HP1**
H3K9me3	**H3K9me3**

The maintenance of heterochromatin through numerous cell cycles

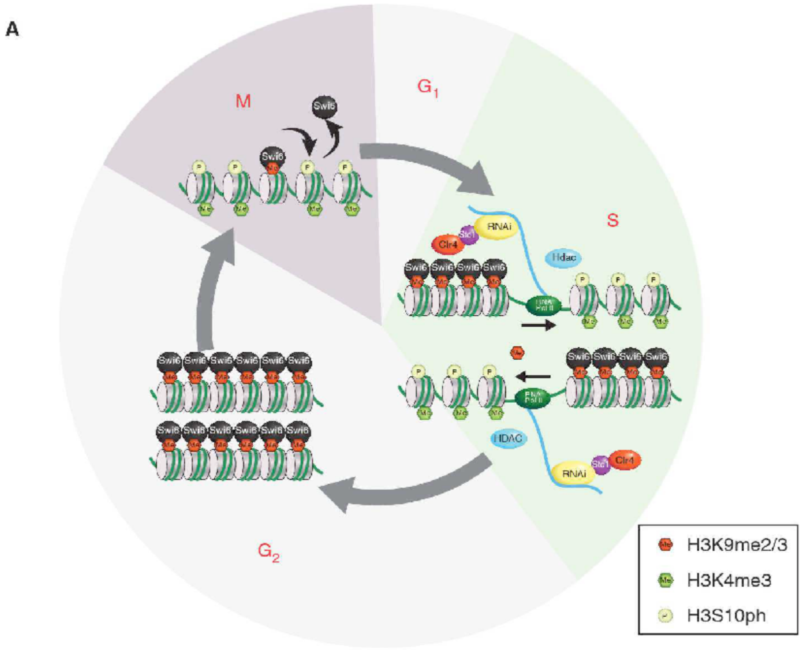


Figure 4. Cell-cycle regulation of centromere heterochromatin assembly. (*A*) Heterochromatin located at chromosomal centromeres becomes differentially methylated and phosphorylated on histones throughout the cell cycle as indicated. These modifications control the binding of the heterochromatin protein Swi6. During mitosis Swi6 is displaced by H3S10 phosphorylation. Swi6 binding is reestablished during subsequent DNA replication (S phase) when a more accessible chromatin structure permits RNA Pol II to transcribe centromeric DNA. This, in turn, recruits the RNAi machinery to direct H3K9me methylation. (*B*) Replication-coupled RNAi model (Li et al. 2011). This figure

Swi6 (HP1) binding to chromosomes is regulated during the cell cycle

At the start of M-Phase H3S10 is phosphorylated at the entire chromosomes
Resulting in the detachment of Swi6 (HP1) from H3K9me3

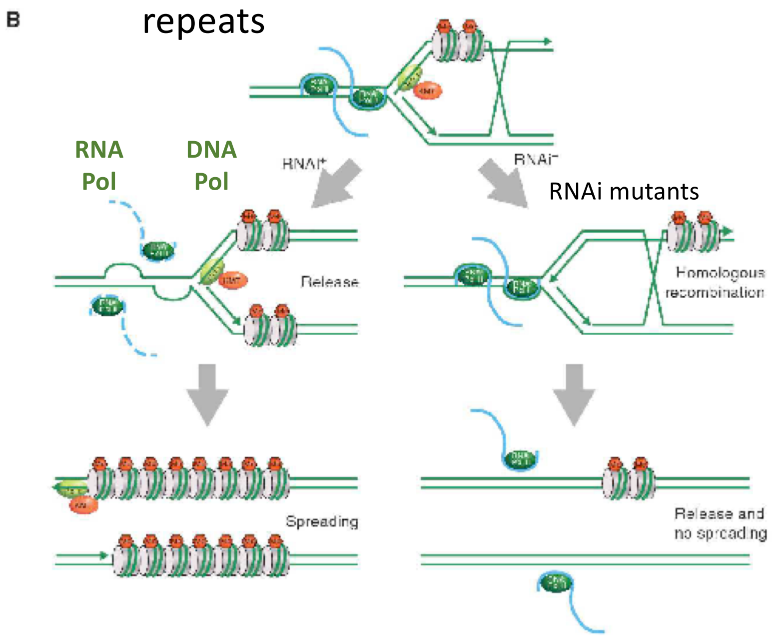
Swi6 re-enters to chromosome at high efficiency in S-Phase

Importantly: RNA Pol is associated with repeat sequences mainly in S-Phase (not at genes)

→ **burst of pericentric transcription**
→ Rapid recruitment of the RNAi-machinery by nascent transcripts
→ Burst in heterochromatin formation!!!
→ G2: reduced transcription → heterochromatin already formed

Note: Stc1 interacts with RITS (Ago1) and the CLRC complex (Clr4)
Note: SHREC complex contains the HDAC Clr3 that is important to subsequently shut down transcription

The maintenance of heterochromatin



H3K9me methylation. (*B*) Replication-coupled RNAi model (Li et al. 2011). This figure illustrates an alternative model for how RNAi works at centromeres. Here, RNAi serves to release RNA Pol II from chromatin to avoid collision with DNA replication machinery during S phase. See text for further details. (*A*, Adapted from Djupedal and Ekwall 2008.)

Swi6 (HP1) binding to chromosomes is regulated during the cell cycle

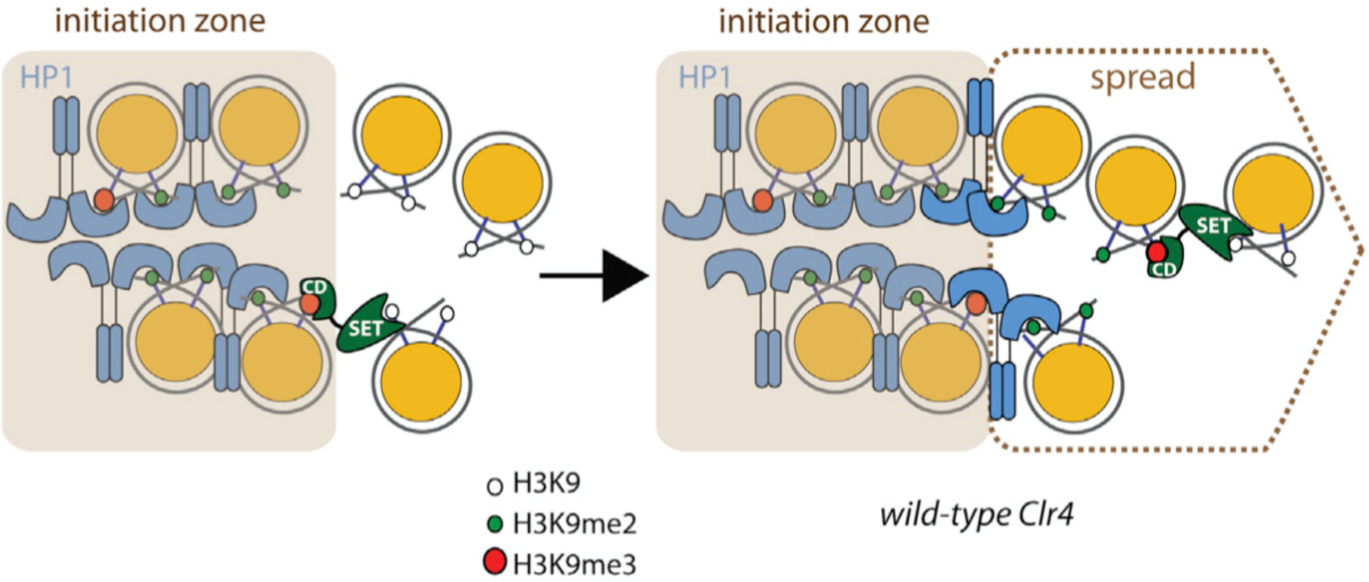
DNA replication represents high risk for genome stability:
If RNA Polymerase can encounter DNA Polymerase → collision of major protein complexes that are linked to open DNA double strands (fragile DNA) → sterical stess on DNA → risk of breakage → genomic instability

RNAi mutants result in stalling DNA replication forks and activate recombination

Pol epsilon (leading strand synthesis) has a role in recruiting the CLRC complex (Clr4) – H3K9me3 –> recruitment of RNAi machinery
RNA dependent RNA polymerase → amplification of RNAi pathway
→ Heterochromatin formation (SHREC and CLRC complex)
→ Silencing of transcription

RNAi dependent formation of heterochromatin preserves genomic stability

HOW IS CHROMATIN SPREADING LIMITED



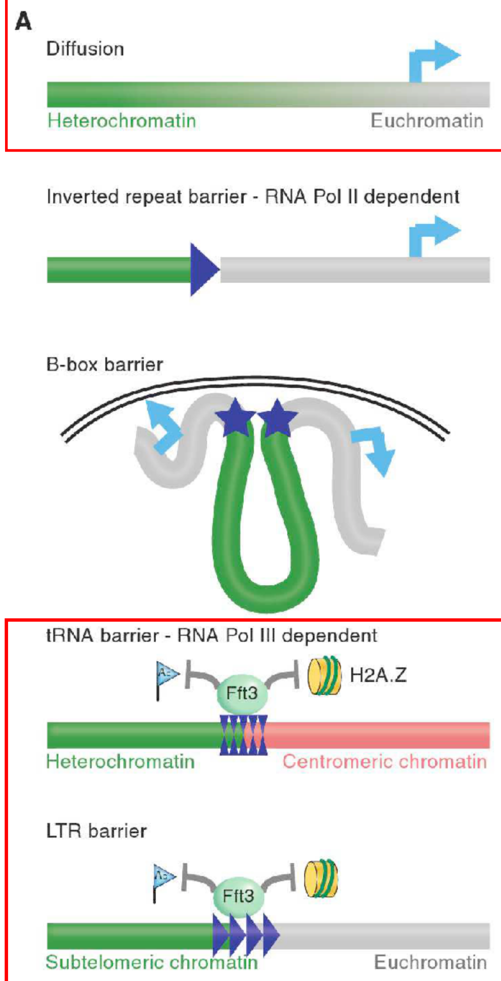
1. Clr4 mediates H3K9m3; Clr4 has a chromodomain that interacts with H3K9me3
2. Swi6 interacts with H3K9me3 and forms homodimer with another Swi6
3. Swi6 directly interacts with Clr4
4. These interactions mediate the spreading of Clr4/Swi6/H3K9me3 along chromatin.

Swi6 mediates nucleosome bridging and shortening of nucleosome to nucleosome distance
–> Chromatin compaction and SPREADING

HOW CAN THIS PROCESS BEING STOPPED????

HOW IS CHROMATIN SPREADING LIMITED

Boundary regions separate genome segments with different chromatin structure (euchromatin/heterochromatin)



S.pombe

DIFFUSE BOUNDARY

RNA Pol II DEPENDENT INVERTED REPEAT BARRIER

B-BOX BARRIERS

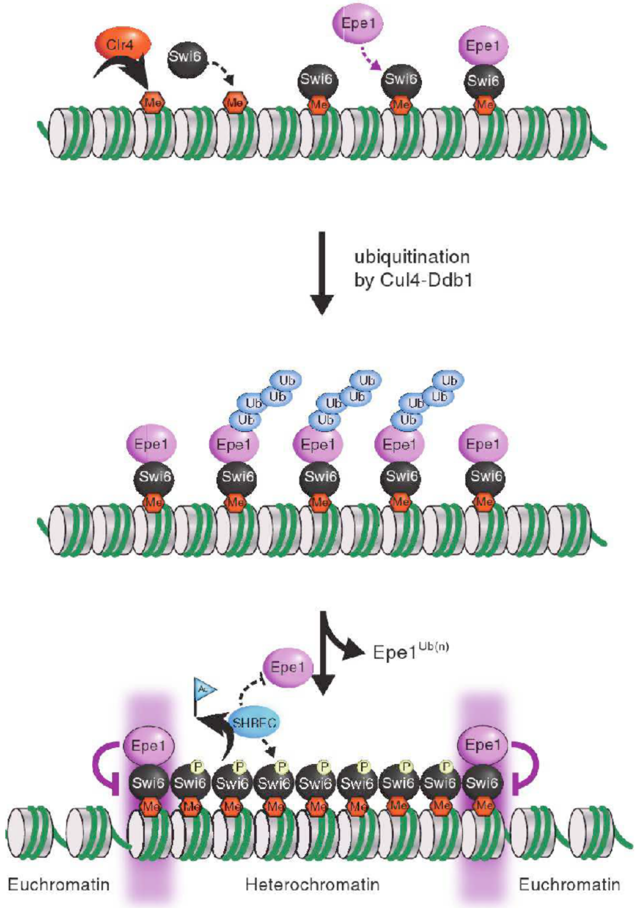
RNA Pol III DEPENDENT tRNA BOUNDARIES

LTR BOUNDARIES (from retrotransposons)

Figure 5. Chromatin boundaries and the boundary mechanism involving Epe1 (for enhancement of position effect) in *S. pombe*. (*A*) A schematic representation of different types of boundary elements in *S. pombe*. (*B*) The mechanism of boundary function by

HOW IS CHROMATIN SPREADING LIMITED: Epe1

Diffuse boundary: Epe1



Epe 1 is an “anti-silencing” factor = counteracts heterochromatin building.

S. pombe Epe1-/- : more heterochromatin spreading

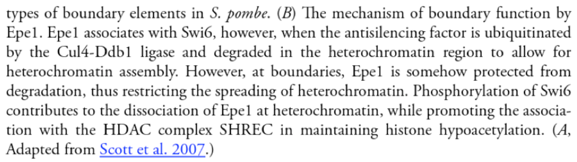
1. Epe1 forms a complex with Swi6 in heterochromatic regions

2. However: **Epe1 is ubiquitinated and degraded by a cullin dependent ubiquitinase in regions with heterochromatin and remains only at boundaries between heterochromatin and euchromatin**
(not known how ubiquitination is prevented at boundaries)

3. After removal of Epe1, Swi6 becomes phosphorylated (casein kinase 2)

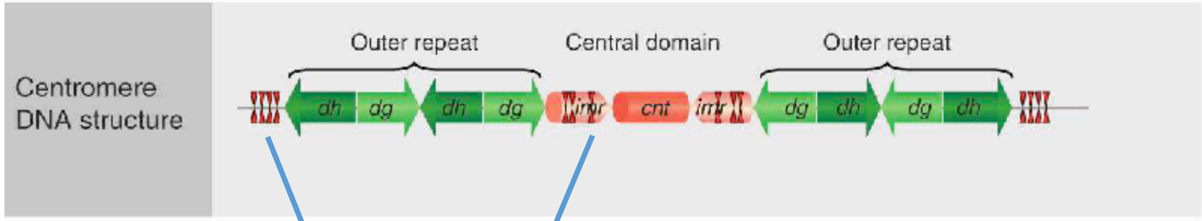
4. Swi6 phosphorylation causes the recruitment of the SHREC complex

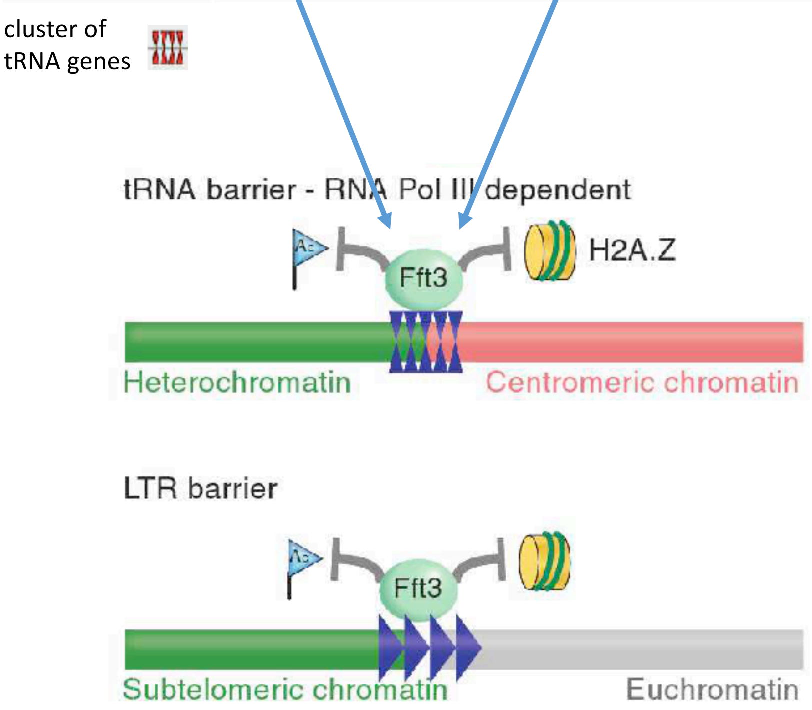
5. Epe1 remains associated at periphery and blocks phosphorylation of HP1



types of boundary elements in *S. pombe*. (*B*) The mechanism of boundary function by Epe1. Epe1 associates with Swi6, however, when the antisilencing factor is ubiquitinated by the Cul4-Ddb1 ligase and degraded in the heterochromatin region to allow for heterochromatin assembly. However, at boundaries, Epe1 is somehow protected from degradation, thus restricting the spreading of heterochromatin. Phosphorylation of Swi6 contributes to the dissociation of Epe1 at heterochromatin, while promoting the association with the HDAC complex SHREC in maintaining histone hypoacetylation. (*A*, Adapted from Scott et al. 2007.)

HOW IS CHROMATIN SPREADING LIMITED: tRNA clusters and Fft3





tRNA boundaries

A factor that prevents the spreading of euchromatin into heterochromatin

1. Despite their proximity to the repressive pericentric heterochromatin, the tRNA genes were actively transcribed by RNA Polymerase III (RNAPIII) = Eurchromatin. Deletion of the tRNA clusters resulted in expansion of heterochromatin <u>into the CENP-A chromatin at cnt region</u> —indicating that the tRNA genes functioned as chromatin barriers at the centromere.
2. Centromeric repeats represent HETEROCHROMATIN
3. The chromatin remodeler Fft3 is restricted to tRNA genes (mechanism not known); blocks nucleosome turnover and maintains repressed chromatin state
4. **Fft3** is localized to the centromeric tRNA and subtelomeric LTR boundary regions, where it **prevents euchromatin formation in the centromeric domain and subtelomeric regions** (Fft3=nucleosome remodelling factor; ATPase domain)
5. Loss of Fft3 → extension of euchromatin into centromeric regions (H3/H4 acetylation and H2AZ = histone variant in euchromatin)
6. Mechanism ??? , also works at other sites → LTR barriers and subtelomeres

HOW IS CHROMATIN SPREADING LIMITED: tRNA clusters and Fft3

Cnp1 =
S.Pombe CENPA

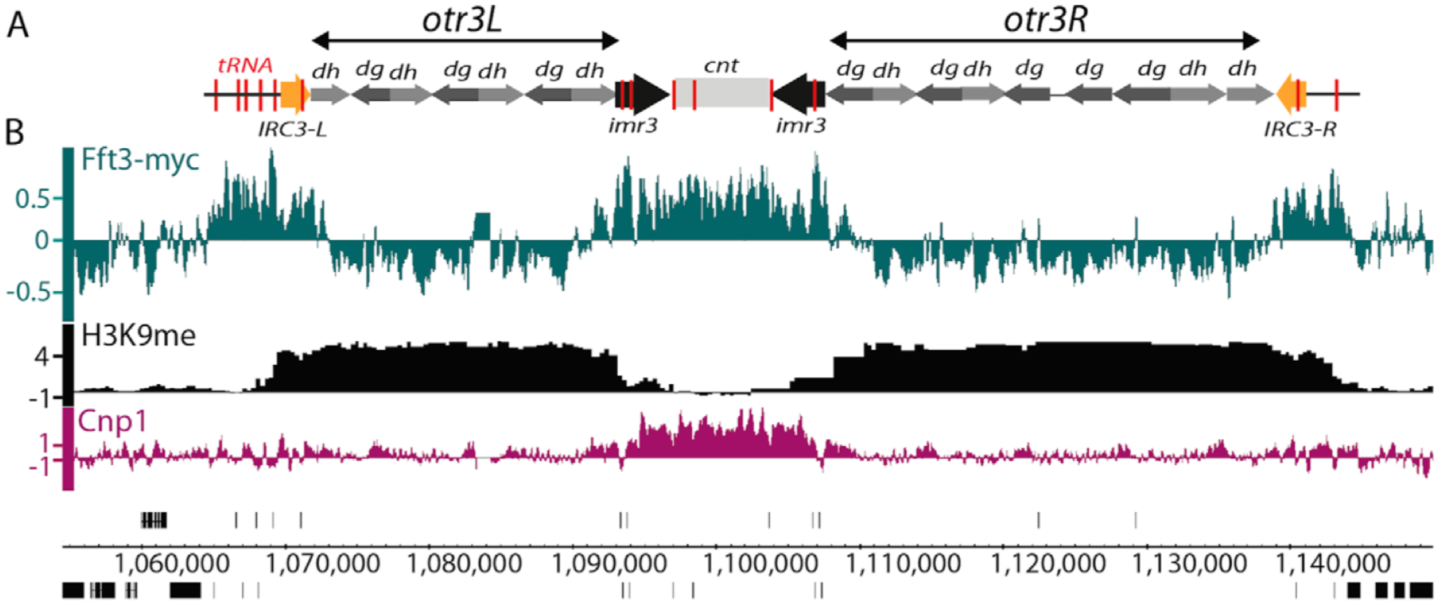
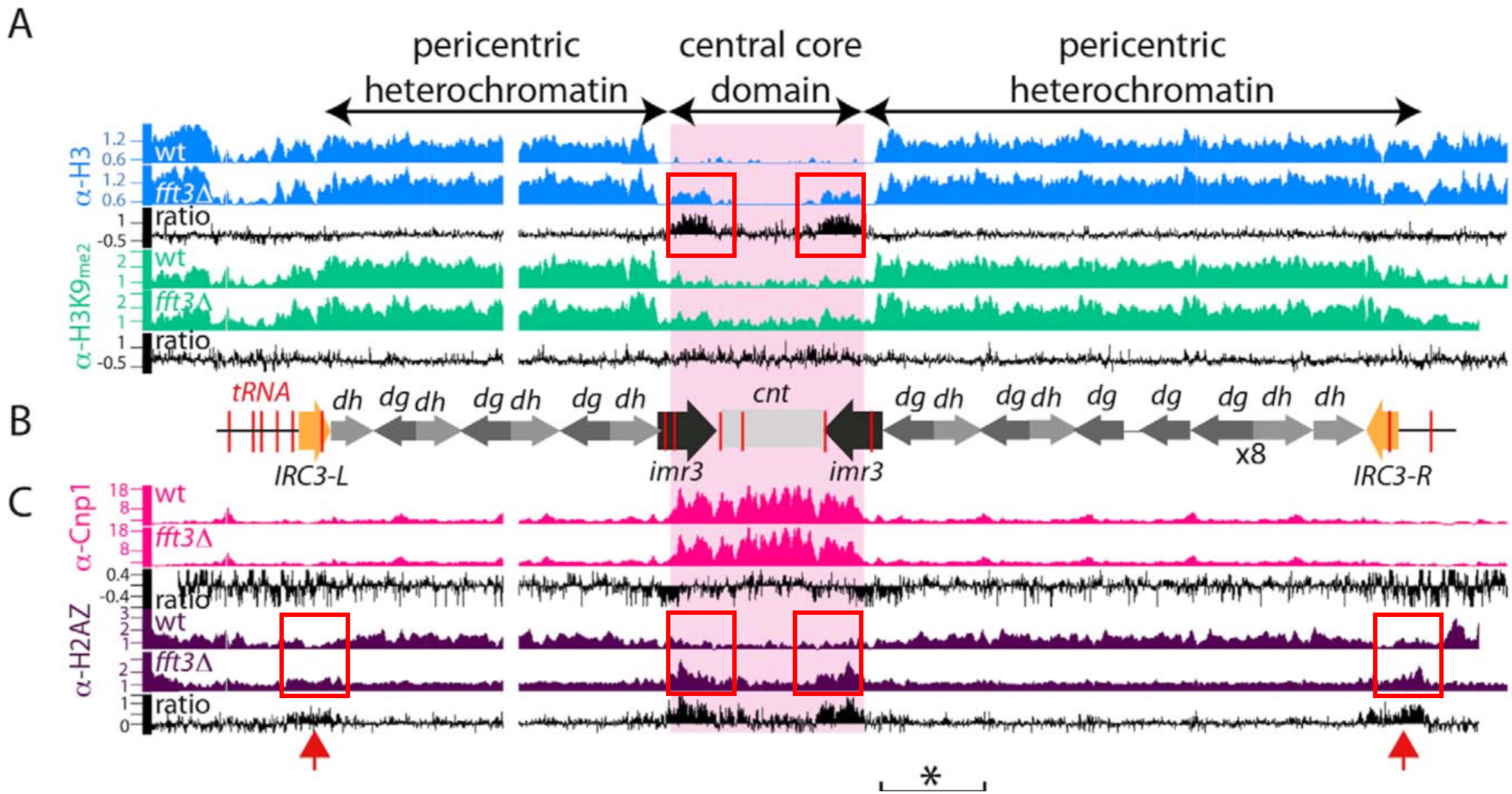


Figure 2. Fft3 associates with known centromeric insulators and the central core domain. A) A schematic presentation of centromere 3. tRNA genes are marked in red and the *IRC* elements in yellow. B) A genome browser view of *cen3* showing the ChIP-chip occupancy profile for Fft3-myc (green), H3K9me2 (black) and Cnp1 (purple). Data on the Y-axis are presented in log2 scale and the X-axis shows genome positions in base pairs. H3K9me2 data are from [1].
doi:10.1371/journal.pgen.1001334.g002

Cnp1 = S.pombe CENP-A

HOW IS CHROMATIN SPREADING LIMITED: tRNA clusters and Fft3



Normally <u>H2A.Z is absent from all centromeric regions</u>, including both the Cnp1 containing inner domain and the IRC and IMR. H2A.Z marks active genes. Surprisingly, absence of Fft3, the H2A.Z levels increase by 4-fold at imr3 and increase in IRC and IMR

HOW IS CHROMATIN SPREADING LIMITED: tRNA clusters and Fft3

ChIP on
Imr – region

Loss of Ffts3 results in
An extension of euchromatin
to imr repeats
(H2A-Z, H4K12Ac, H3K9Ac)

Ffts has a chromatin
boundary function

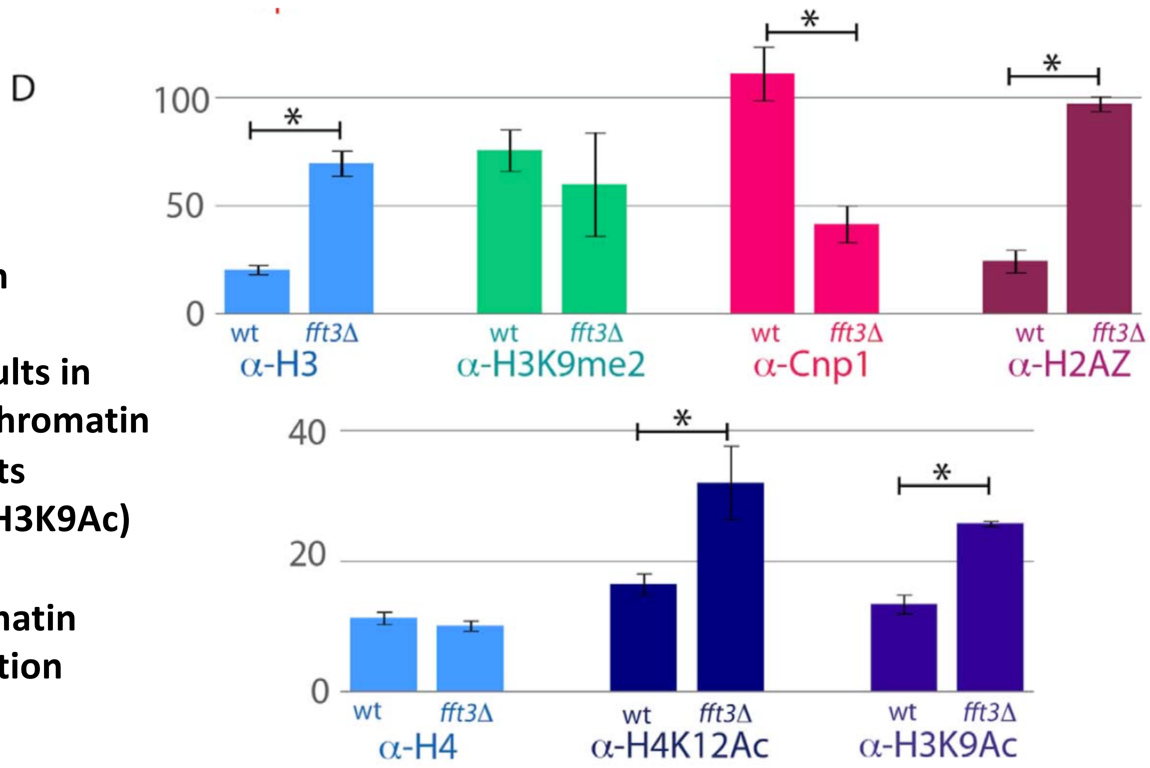
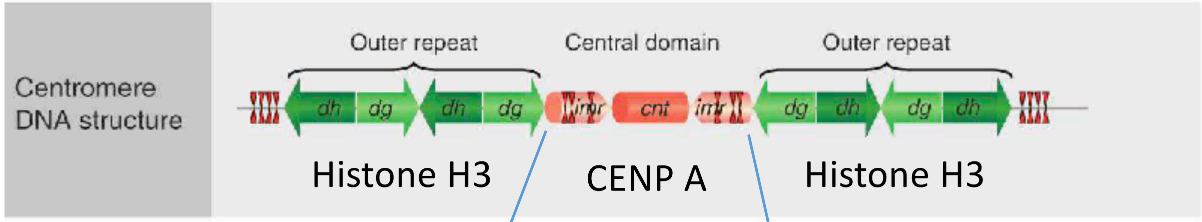


Figure 3. Centromeric boundary function is impaired in *fft3Δ* cells. A) The ChIP-chip distributions of histone H3 (blue) and H3K9me2 (green) at centromere 3 in wt and *fft3Δ* cells are shown in Genome Browser images. Y-axis: Linear scale. Mutant/wt ratios are indicated in black in log2 scale. B) A schematic diagram of centromere 3. tRNA genes are marked in red and the *IRC* elements in yellow. C) Genome browser images showing ChIP-chip data for Cnp1 (red) and H2A.Z^{Pht1} (purple). Y-axis: Linear scale. Mutant/wt ratios are shown in black in log2 scale. D) Bar diagrams showing the results from real-time quantitative PCR analysis of ChIP signals. The enrichments at *imr3* are relative to the *act1+* euchromatic control locus except for the Cnp1 enrichment which is relative to *cen1*. The ChIP signals were normalized to input samples from the same chromatin extract. The error bars represent S.D. values from triplicate samples. *indicates a significant difference between wt and mutant ($P < 0.01$; T-test, two-tailed, unpaired).
doi:10.1371/journal.pgen.1001334.g003

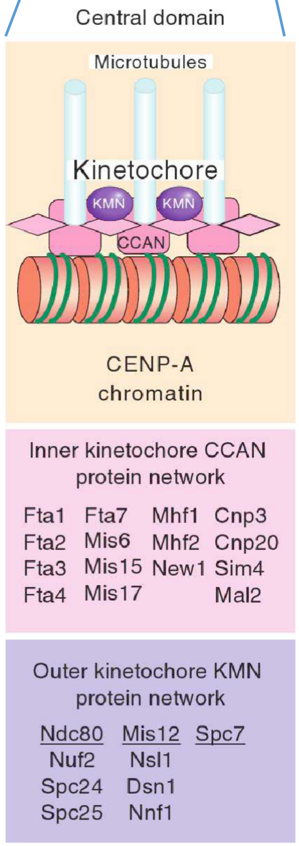
Loss of Fft3:
More H2AZ, H4K12Ac and H3K9Ac
- → Active chromatin state at Imr repeates
- → Euchromatin spreading
- → Barrier lost

WHY IS CENTROMERIC HETEROCHROMATIN IMPORTANT FOR CENTROMERE FUNCTION?



RNAi dependent formation of heterochromatin

RNAi dependent formation of heterochromatin



Structural kinetochore complex that prevents the entry of RNA polymerase

At the central domain the majority of histone H3 is exchanged for Cnp1 = CENP A
→ ca 19-24 CENP A nucleosomes per centromere

The CENPA gene encodes a centromere protein which contains a histone H3 related histone fold domain that is required for targeting to the centromere. CENPA is proposed to be a component of a modified nucleosome or nucleosome-like structure in which it replaces 1 or both copies of conventional histone H3 in the (H3-H4)2 tetrameric core of the nucleosome particle. Silencing in central domain is thought to be mediated by the altered nucleosome composition and kinetochore construction (sterical hindrance)

Recruitment of CENP A requires RNAi and the presence of heterochromatin

WHY IS CENTROMERIC HETEROCHROMATIN IMPORTANT?

Experiment:

Make a yeast plasmid without origin of replication, but instead, insert yeast centromere containing cen, imr, omr sequneces

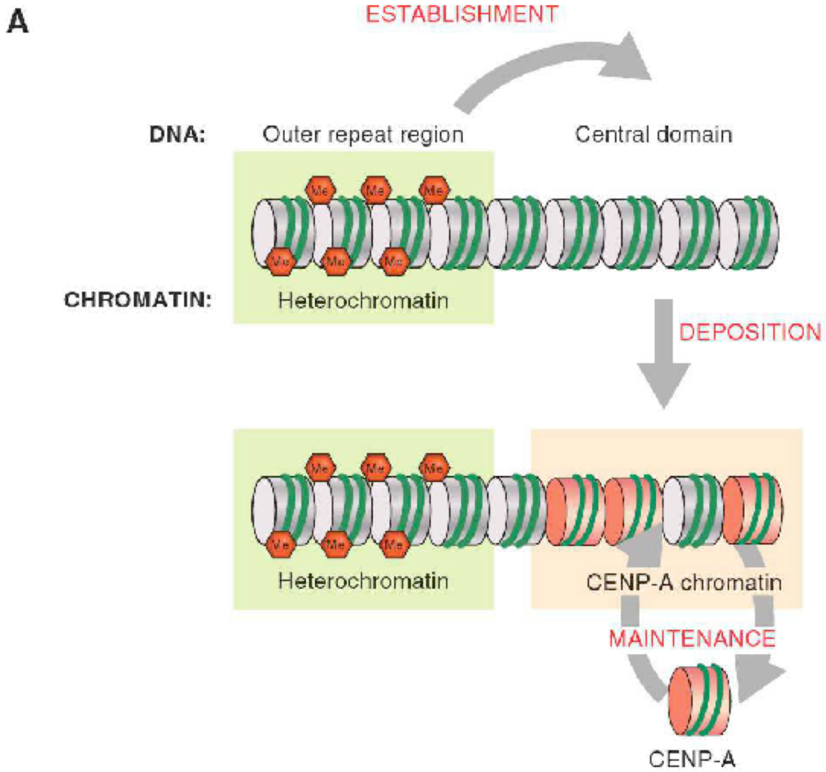
→ Formation centromere+kinetochor (CenpA incorporation) and maintainance during mitosis

  → Plasmid with only central repeats (cen) cannot be maintained in S. pombe

  → No CENP-A incorporation

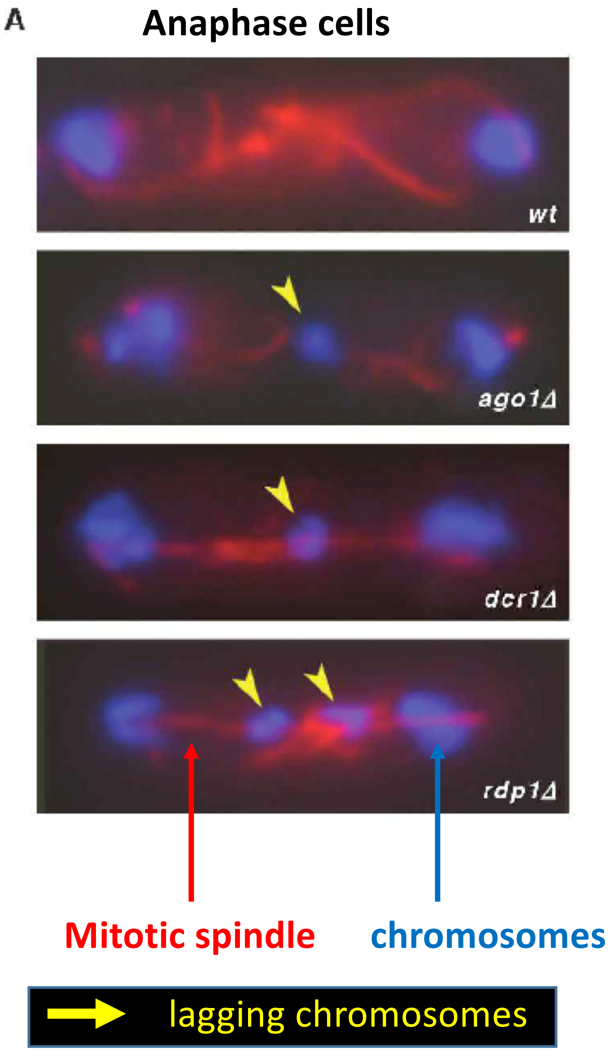
  → Addition of outer repeat + central repeats: CENP A incorporation and plasmid maintenance

  → Heterochromatin at outer repeats stabilize centromere function.



HETEROCHROMATIN IS IMPORTANT TO ACTIVATE CENTROMERE FUNCTION

WHY IS CENTROMERIC HETEROCHROMATIN IMPORTANT?



Heterochromatin is essential for controlled chromosome segregation

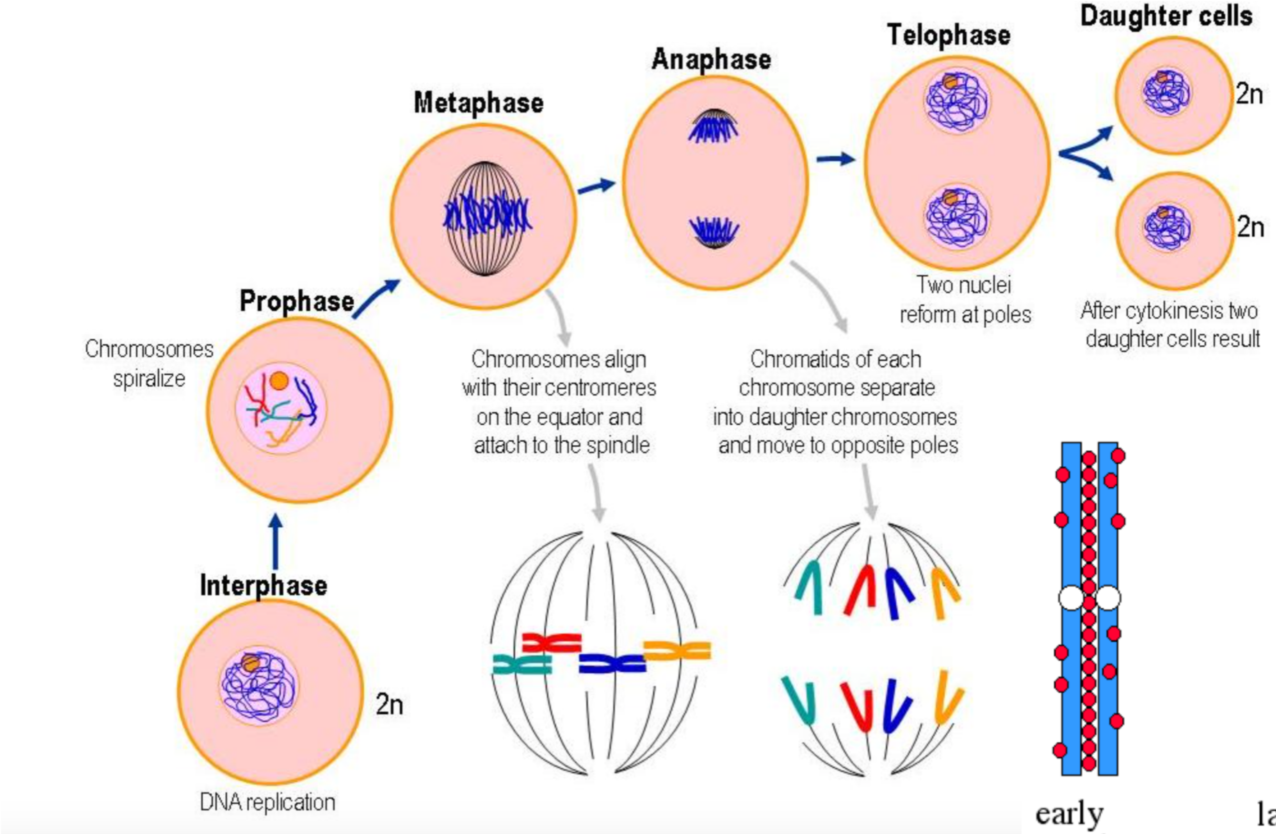
Loss of_
RNAi
machinery
Ago1
Dcr1
Swi6
Clr4

→

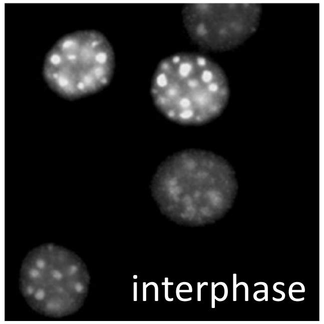
Chromosome segregation defects:
= sister chromatids are not faithfully separated in to daughter cells

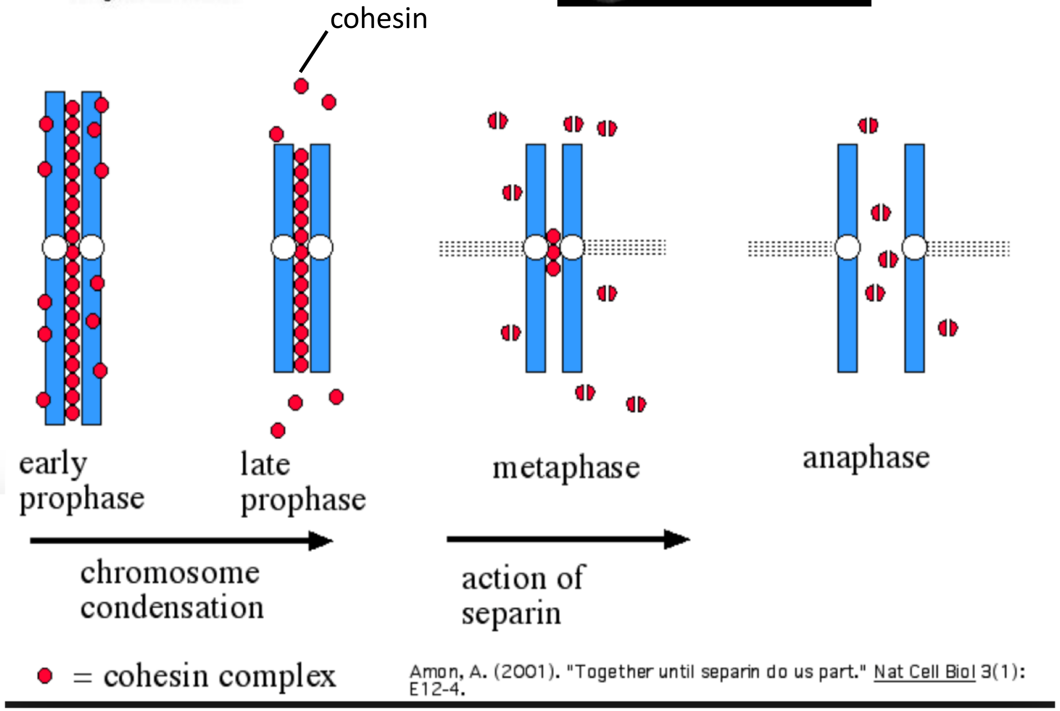
- Loss of chromosomes
- Lagging anaphase chromosomes
- Sister chromatid cohesion defects

Eukaryotic centromeres hold together sister chromatids until chromosome segregation



Chromocenters,
aggregation of centromeres



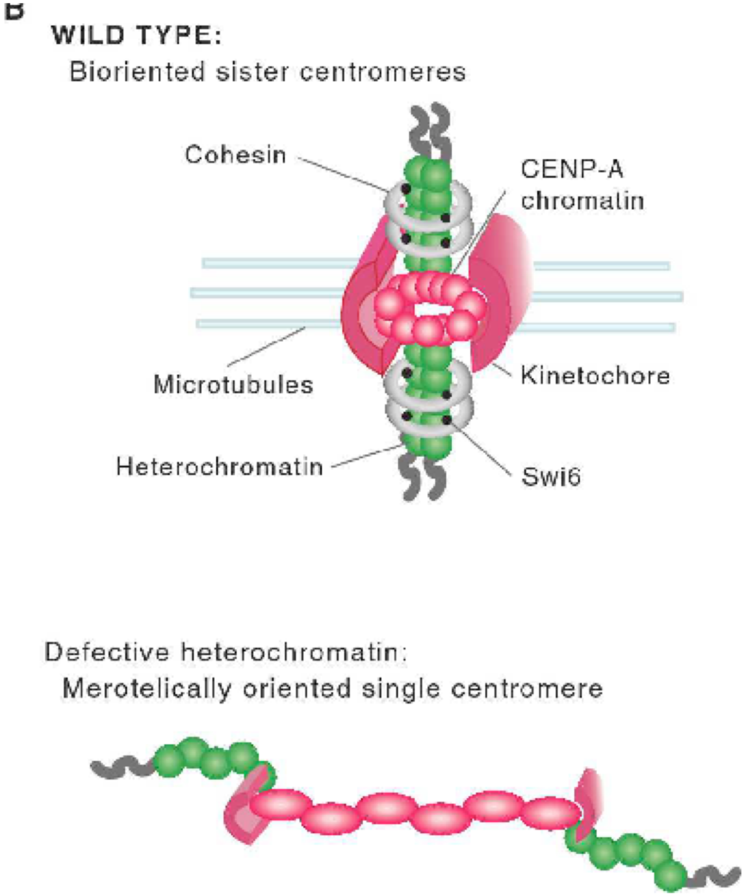


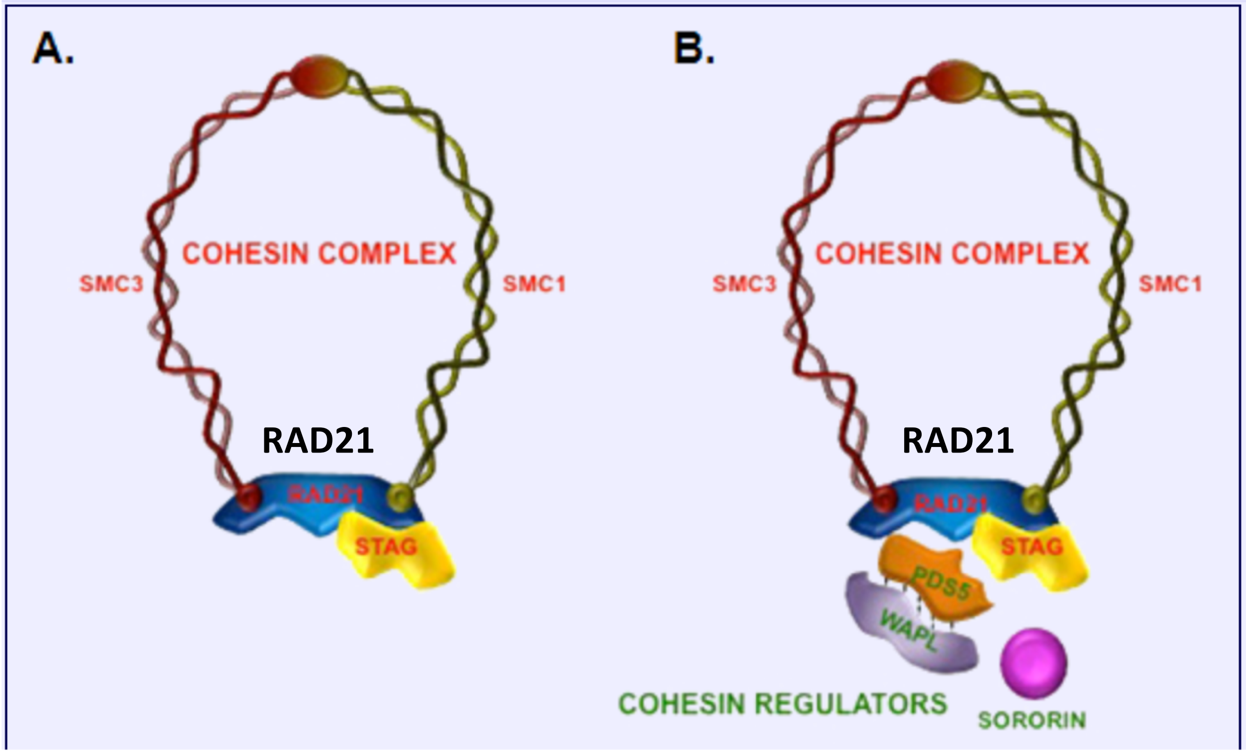
Amon, A. (2001). "Together until separin do us part." Nat Cell Biol 3(1): E12-4.

WHY IS CENTROMERIC HETEROCHROMATIN IMPORTANT?

Cohesin forms a ring around sister-chromatids to organize chromosome segregation

cohesin at centromere is very resistant and is thought to hold chromosomes well together until the spindel is perfectly attached to kinetochores



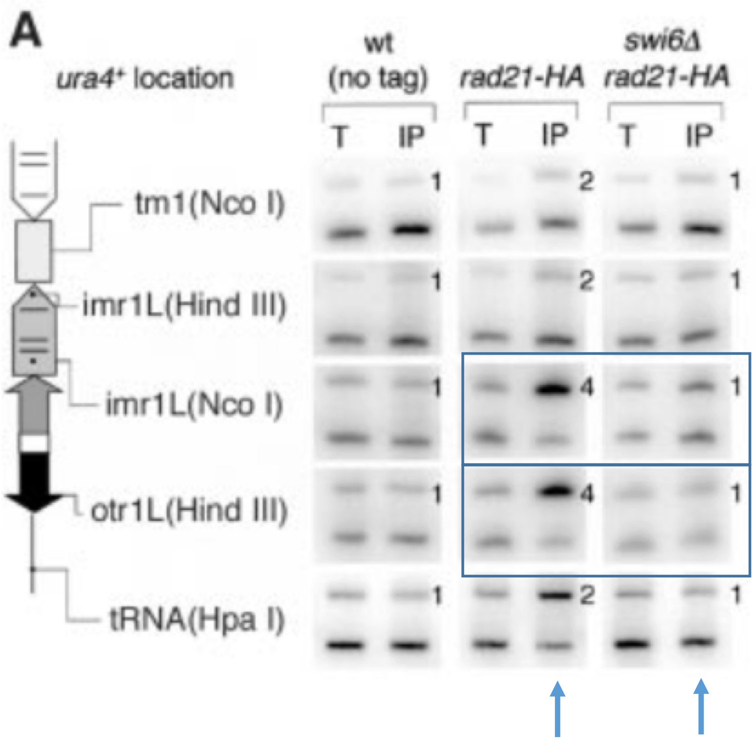


Centromeric heterochromatin is required to load cohesin to centromeres

WHY IS CENTROMERIC HETEROCHROMATIN IMPORTANT?

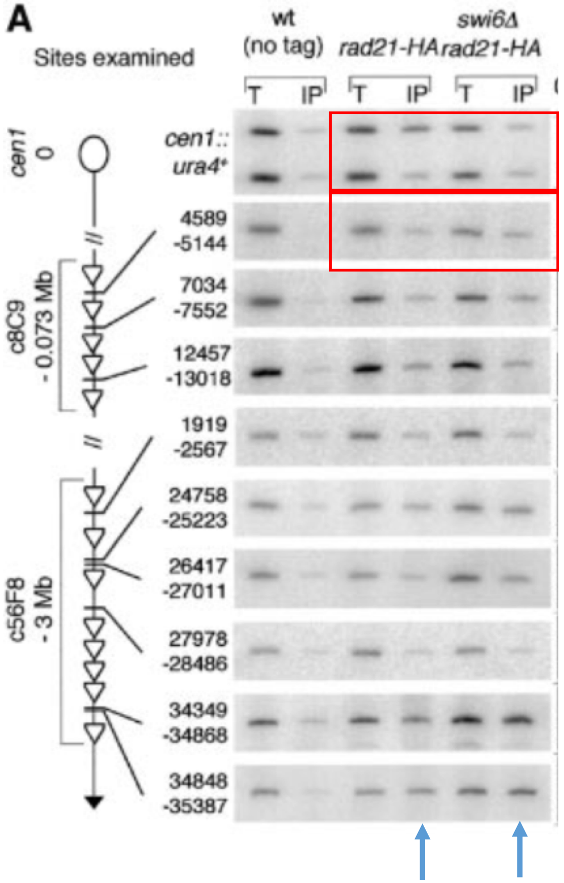
Swi6 is required to recruit cohesin subunit Rad21 to centromeres…..

→ Centromeric heterochromatin is required to recruit cohesin to centromeres



Repeat DNA
internal PCR control
Repeat DNA
internal PCR control
Repeat DNA
internal PCR control
Repeat DNA
internal PCR control
Repeat DNA
internal PCR control

S. pombe: Wt
S.pombe: wt+HA-rad21
S.pombe: Swi6-/- HA-rad21



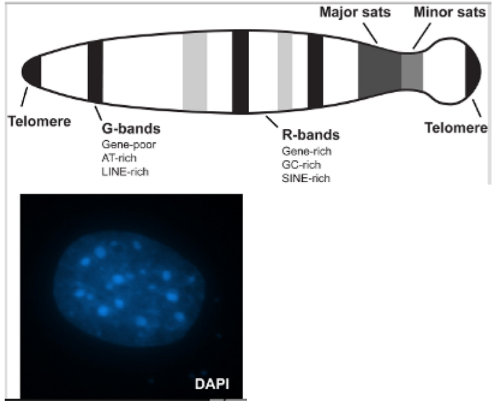
S. pombe: Wt
S.pombe: wt+HA-Rad21
S.pombe: Swi6-/- HA-Rad21

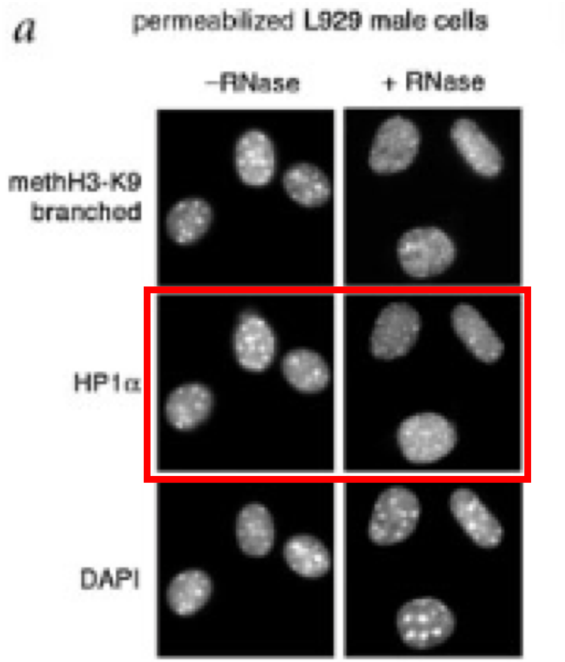
… but Swi6 is not important to recruit cohesin subunit Rad21 to other chromosomal positions

Pericentric heterochromatin in vertebrates: the mouse as a model system

Pericentric heterochromatin in vertebrates: the mouse as a model system

A role of pericentric RNAs in heterochromatin formation





Permeabilize cells with soft detergent – without cell killing

Digest RNA with RNase

Fix cells with formaldehyde

Immunofluorescence with anti-HP1 antibodies

HP1 dissociate from pericentric DNA

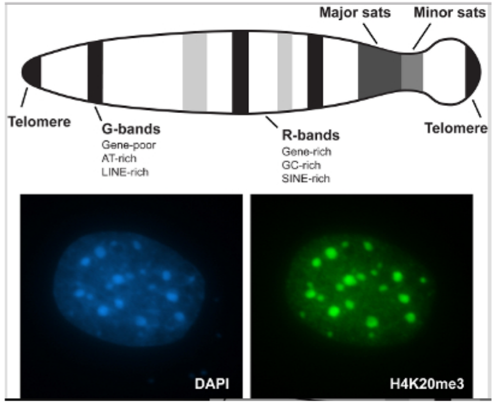
RNA component is important for HP1 recruitment at pericentric heterochromatin in mouse cells

Presence of sense-antisense RNA

What about specific transcripts and DICER?

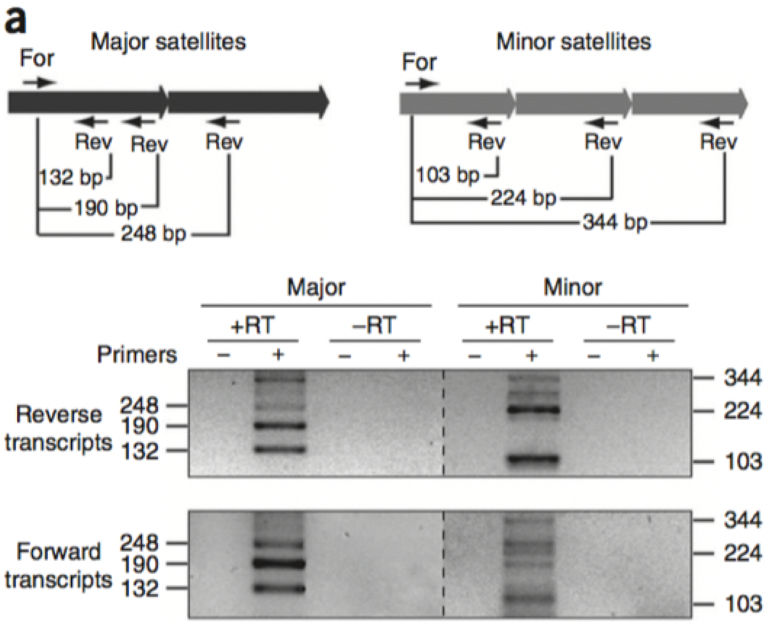
Pericentric heterochromatin in vertebrates: the mouse as a model system

Conserved pericentric heterochromatin



Suv39h1/h2 → **H3K9m3 + HP1**
H3K9me3 calls in Suv420h
Suv420h1/h2 → **H4K20me3**
DNA methylation
Bi-directional transcription of satellite repeats

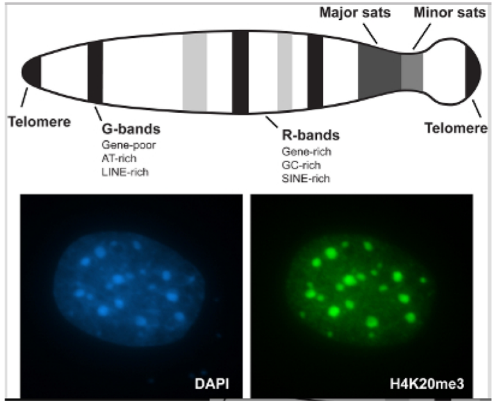
Forward and reverse transcripts from mouse pericentric heterochromatin
Strand specific RT-PCR (end point PCR)



Mouse pericentric repeats are transcribed in sense anti-sense orientation

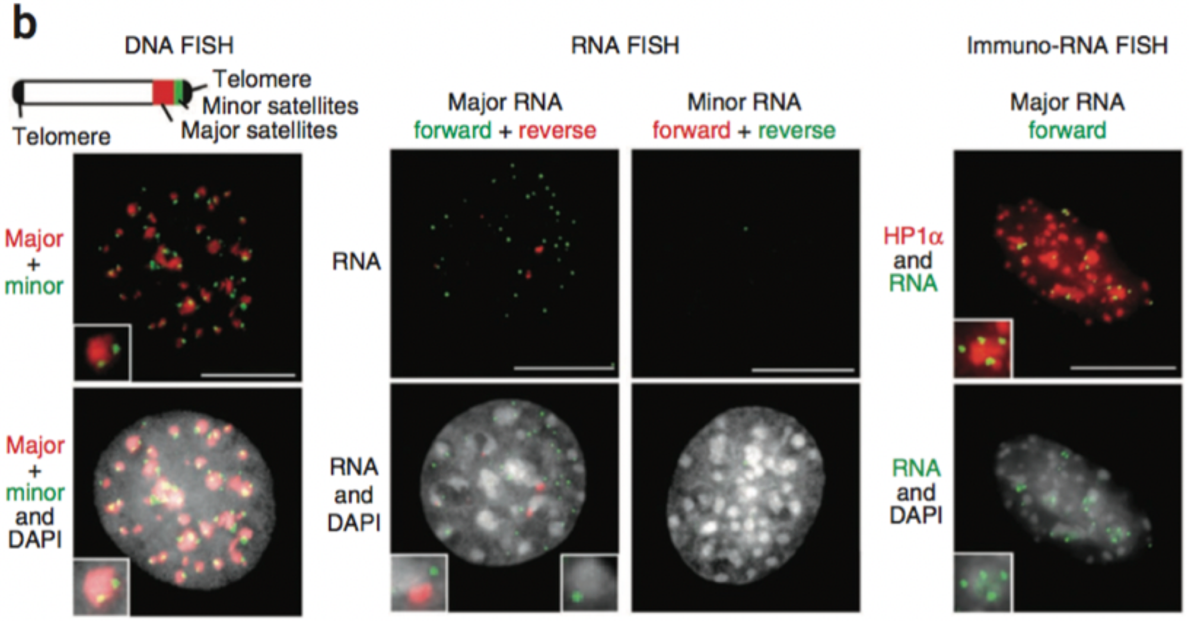
Pericentric heterochromatin in vertebrates: the mouse as a model system

Conserved pericentric heterochromatin



Suv39h1/h2 → **H3K9m3 + HP1**
H3K9me3 calls in Suv420h
Suv420h1/h2 → **H4K20me3**
DNA methylation
Bi-directional transcription of satellite repeats

Forward and reverse transcripts from mouse pericentric heterochromatin
RNA-FISH: nascent transcripts

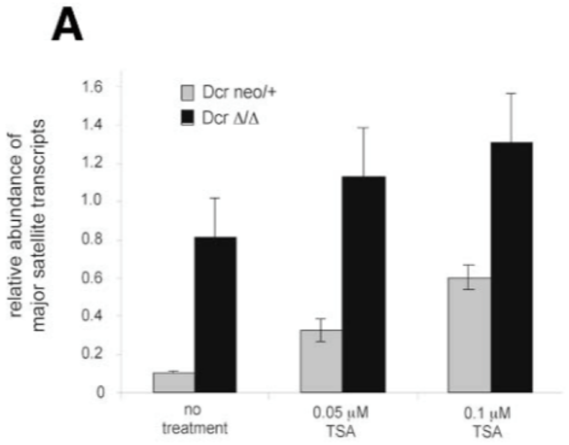


FW+REV FW+REV FW+REV FW

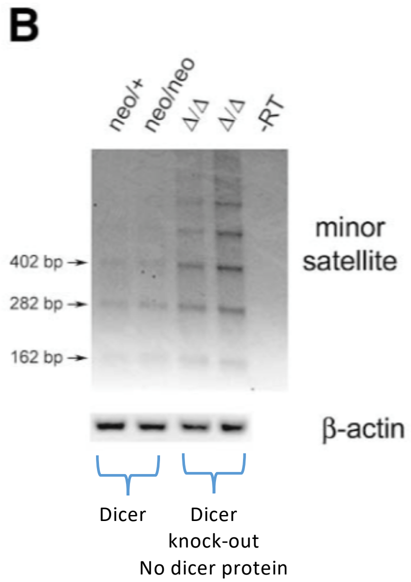
Mouse pericentric repeats are transcribed in sense anti-sense orientation (preferentially FW)

A role for Dicer at mouse centromeres??

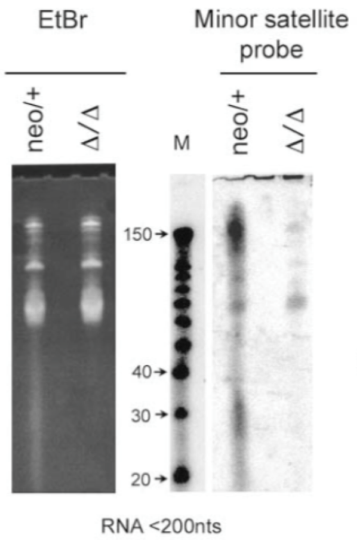
Loss of Dicer: more major and minor satellite transcripts (long)….and less <u>small</u> minor-satellite RNAs



TSA: HDAC inhibitor



RT-PCR

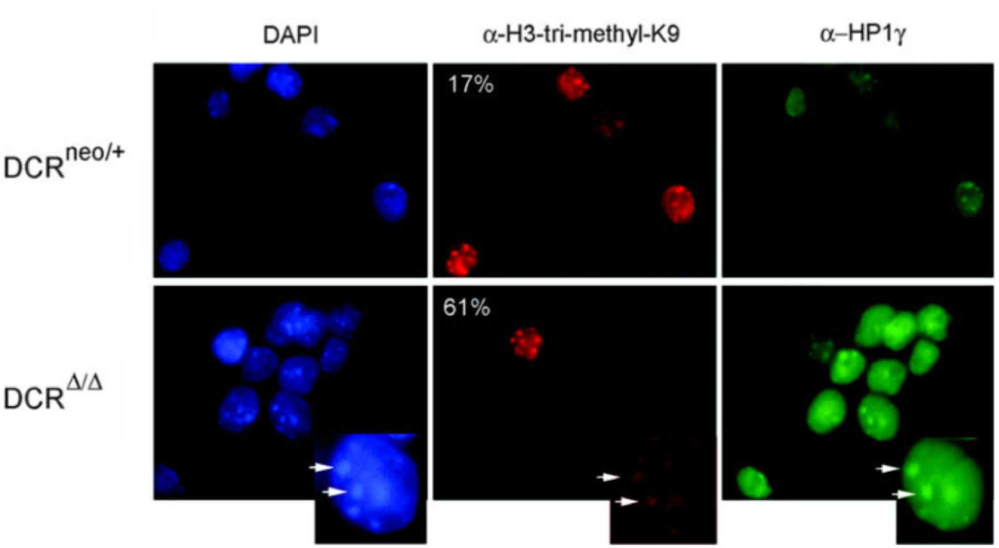


Northern blot on small RNA fraction

In the presence of ds pericentric RNAs → Dicer dependent siRNA production

A role for Dicer at mouse centromeres??

Loss of Dicer reduced that abundance of pericentric heterochromatin in mouse embryonic stem cells



17% of tri-methyl H3K9 negative cells

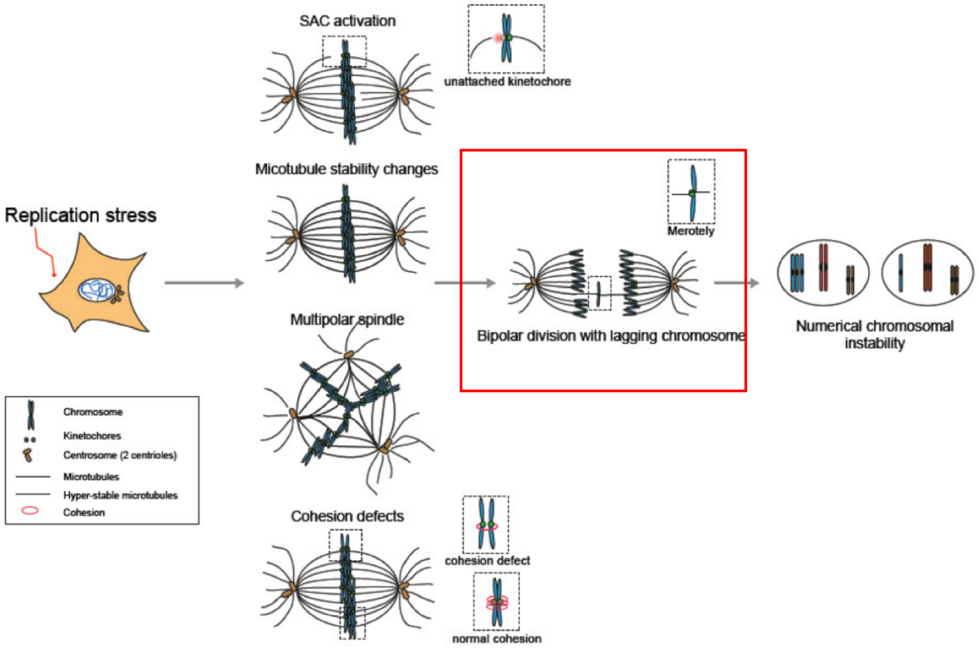
61% of tri-methyl H3K9 negative cells

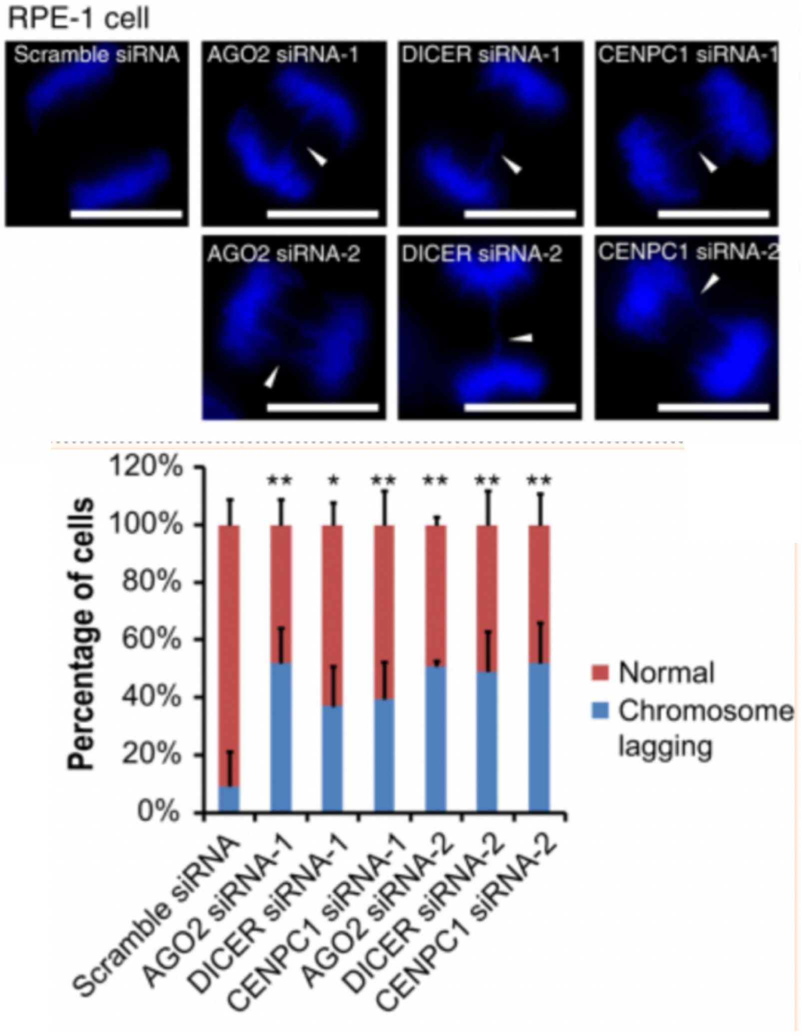
Loss of Dicer only reduces H3K9me3 and HP1 heterochromatin but

DOES NOT HAVE AN IMPACT ON CHROMOSOME SEGREGATION

DICER IS DISPENSABLE FOR HETEROCHROMATIN FORMATION

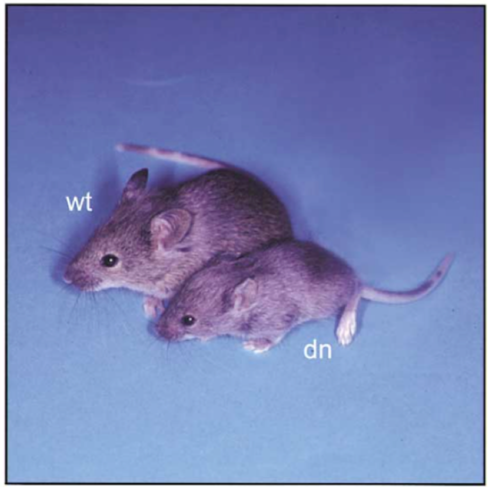
Loss of Dicer leads to mitotic instability in vertebrate cells





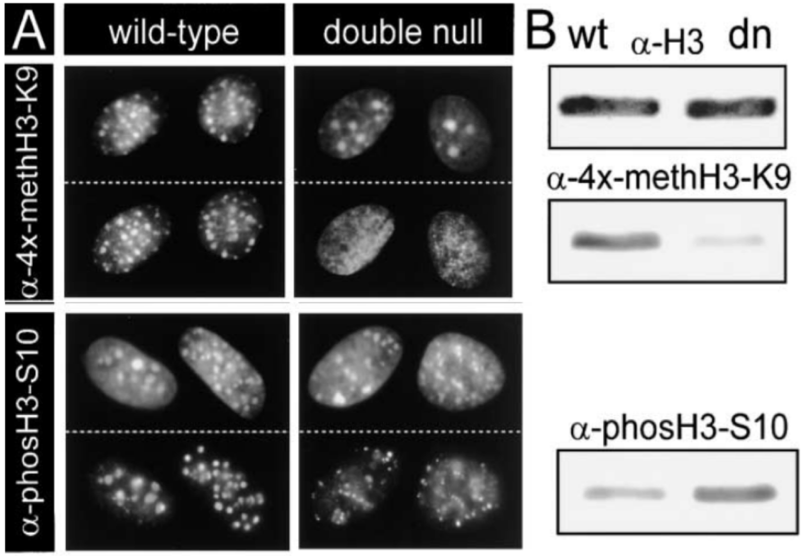
A role for Dicer at mouse centromeres??

Suv39h1/h2 KO

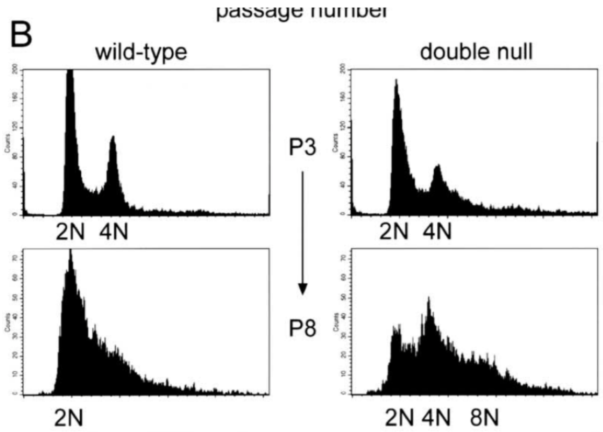


Suv39h1/h2 dn mice are smaller

Loss of pericentric H3K9 tri-methylation And abnormal H3S10 phoshporylation



FACS – scan: loss of Suv39h1/2 results in aneuploidy – increased chromosome numbers



In mouse cells:

- forward and reverse transcripts locate to pericentric regions
- Recruit Suv39h1/h2
- Drive H3K9me3 and HP1
- Preserve mitotic stability

A role for Dicer??