Vol. 10, No. 11

Diversity in Requirement of Genetic and Epigenetic Factors for Centromere Function in Fungi[▽]

Babhrubahan Roy and Kaustuv Sanyal*

Molecular Mycology Laboratory, Molecular Biology and Genetics Unit, Jawaharlal Nehru Centre for Advanced Scientific Research, Jakkur Post, Bangalore 560 064, India

A centromere is a chromosomal region on which several proteins assemble to form the kinetochore. The centromere-kinetochore complex helps in the attachment of chromosomes to spindle microtubules to mediate segregation of chromosomes to daughter cells during mitosis and meiosis. In several budding yeast species, the centromere forms in a DNA sequence-dependent manner, whereas in most other fungi, factors other than the DNA sequence also determine the centromere location, as centromeres were able to form on nonnative sequences (neocentromeres) when native centromeres were deleted in engineered strains. Thus, in the absence of a common DNA sequence, the cues that have facilitated centromere formation on a specific DNA sequence for millions of years remain a mystery. Kinetochore formation is facilitated by binding of a centromere-specific histone protein member of the centromeric protein A (CENP-A) family that replaces a canonical histone H3 to form a specialized centromeric chromatin structure. However, the process of kinetochore formation on the rapidly evolving and seemingly diverse centromere DNAs in different fungal species is largely unknown. More interestingly, studies in various yeasts suggest that the factors required for *de novo* centromere formation (establishment) may be different from those required for maintenance (propagation) of an already established centromere. Apart from the DNA sequence and CENP-A, many other factors, such as posttranslational modification (PTM) of histones at centric and pericentric chromatin, RNA interference, and DNA methylation, are also involved in centromere formation, albeit in a species-specific manner. In this review, we discuss how several genetic and epigenetic factors influence the evolution of structure and function of centromeres in fungal species.

A complete understanding of the complexities of the process of cellular differentiation requires a thorough analysis of the molecular events occurring during eukaryotic cell division. As an important part of this process, a cell has to ensure accurate segregation of duplicated chromosomes into its progeny cells. In eukaryotes, specific DNA sequences, and the factors that bind to them immediately after replication, partly dictate the state of chromatin. Apart from the genetic factors, many epigenetic phenomena also contribute to formation of specialized chromatin required for specific functions. One such specialized chromatin domain, the centromere (*CEN*)-kinetochore complex, plays a crucial role in high-fidelity chromosome segregation and has great implications for human health. A consequence of improper chromosome segregation is the abnormal chromosome numbers associated with most human cancers. For example, in human colorectal cancers, almost 85% of cells are aneuploid, with 60 to 90 chromosomes (128).

The high-fidelity chromosome segregation that occurs during mitosis and meiosis requires a functional centromere, defined as the primary constriction on a chromosome. The centromere is a region where spindle fibers attach to bring about congression and subsequent movement of the chromatids to opposite poles during the anaphase stage of the cell cycle. In organisms for which meiotic analysis is possible, a centromere can be defined genetically as a chromosomal region where the mutant and wild-type alleles of a diploid heterozygous genetic marker always separate from each other in the first meiotic division. A centromere region is generally transcriptionally inactive, gene-poor, recombination deficient, and heterochromatic in nature. From the cell biologist's point of view, the centromere is an essential specialized chromosomal locus upon which the kinetochore, a macromolecular proteinaceous structure that links a chromosome to spindle microtubules, assembles and powers segregation of chromosomes in high fidelity during cell division.

Even though we cannot visualize a centromere on the small and relatively uncondensed chromosomes of certain lower eukaryotes such as yeasts, we can recognize its presence, localize it accurately, isolate it physically on a cloned segment of DNA, and study its structure and function. More than 30 years ago, the discovery of a functional centromere in a short sequence in the unicellular budding yeast *Saccharomyces cerevisiae* first revealed the molecular nature of a centromere at the DNA sequence level (26, 27). Molecular determination of the presence of centromeres in a few other model organisms, followed by the advent of the availability of genome sequence information, fuelled identification and characterization of centromeres for a large number of eukaryotes. Functional identification of centromeres from various organisms relied on many distinct properties, such as the presence of a region defined by tetrad analysis or binding of an evolutionarily conserved kinetochore protein or even a minimal region that can provide stable inheritance through mitosis and meiosis to an otherwise unstable piece of naked DNA introduced into a cell by chemical transformation. However, the factors that are required for *de novo*

* Corresponding author. Mailing address: Molecular Mycology Laboratory, Molecular Biology and Genetics Unit, Jawaharlal Nehru Centre for Advanced Scientific Research, Jakkur Post, Bangalore 560 064, India. Phone: 91 80 2208 2878. Fax: 91 80 2208 2766. E-mail: sanyal@jncasr.ac.in.
[▽] Published ahead of print on 9 September 2011.

assembly of a functional centromere on an artificially introduced piece of naked DNA (centromere establishment) and those that support stable inheritance of a natural or artificial chromosome with an already-established centromere (centromere propagation) may not always be the same.

Centromeres usually occur once per chromosome. Cloning and characterization of centromeres from many organisms have shown that, while functional centromeres are contained in short sequences with conserved DNA motifs in a few budding yeasts, most other organisms carry longer centromere sequences that are usually rich in repeat DNA elements. Interestingly, the length of an individual repeat unit in centromeres in many organisms, including humans (171 bp), mice (120 bp), and *Arabidopsis* spp. (178 bp), is approximately equal to that of a nucleosome (1, 32, 52, 68, 136). However, centromere functions are strictly species specific, and there is no common sequence determinant that specifies the centromeres of all organisms. Even after decades of research since its discovery, these diverse properties mark the centromere as one of the most mysterious regions of a chromosome.

What could be the most relevant general definition of a centromere today? A serendipitous finding of anti-centromere antibodies isolated from human patients with an autoimmune disease called CREST led to identification of several centromeric proteins (CENPs), including CENP-A (36, 96). Subsequently, members of the CENP-A family of proteins were identified in many organisms as centromere-specific histone H3 (CenH3) variants. CENP-A molecules form specialized centromeric chromatin and are present at the functional centromeres. Thus, a centromere can be redefined as the binding region of CENP-A in a eukaryotic chromosome. It has been observed that, in *S. cerevisiae*, CENP-A/Cse4 also gets recruited at a low level to other noncentromeric loci, indicating that there must be tight epigenetic regulation of formation of the complete kinetochore architecture exclusively at the native centromere locus that acts as the sole microtubule attachment site (16, 71). In this review, we focus on the wide diversity and evolution of structure and function of the centromere (*CEN*) DNA and of the associated CENP-A-containing chromatin in various fungal species.

ORGANIZATION OF DNA ELEMENTS IN DIFFERENT FUNGAL CENTROMERES

Point centromeres. The centromere DNAs of a few budding yeasts of the *Hemiascomycetes*, including *S. cerevisiae*, have been characterized and found to be contained in a relatively short stretch (<400 bp) of DNA. Members of this class of centromeres are often referred to as "point" centromeres. Since conserved DNA motifs serve as the binding sites of specific kinetochore proteins, formation of point centromeres is directed by the DNA sequence and is thus genetically determined. Plasmids carrying an autonomously replicating sequence (*ARS*) that serves as the DNA replication origin and the centromere (*CEN*) behave like aneuploid chromosomes (minichromosomes) that segregate normally through mitosis and meiosis. Yeast Artificial Chromosomes (YACs) can be circular or linear with telomeres at the ends and are useful tools for cloning large pieces of heterologous DNA that can transmit faithfully in mitosis.

The centromere DNA in *S. cerevisiae* is the best-studied example in this class (Fig. 1A). The 125-bp *CEN* DNA contains three consensus Centromeric DNA Elements (CDEs). The central element (CDEII) is a nonconserved 78- to 86-bp-long AT-rich (>86%) sequence (41, 57). CDEII acts as a "spacer" element and is flanked by two conserved motifs, the 8-bp CDEI (PuTCACPuTG) and the 25-bp CDEIII [TGTTT(T/A)TGNT TTCCGAAANNNAAAAA]. The CDEIII sequence is an imperfect palindrome. The role of these conserved elements has been studied extensively (28, 33). Deletion of CDEI causes a marginal (20-fold) increase in chromosome missegregation in mitosis, suggesting that CDEI is not absolutely essential for centromere function. Interestingly, deletion of CDEI or certain alterations of CDEI cause premature sister chromatid separation in meiosis I (33). Changes in the length or base composition (AT richness) of CDEII cause only partial inactivation, but a portion of CDEII is essential for *CEN* function. CDEIII is absolutely essential. Deletion of CDEIII, or even single-base substitutions in the central CCG sequence, completely abolishes *CEN* function (79, 89).

Genome sequences of three other species of the *Saccharomyces* genus, *S. bayanus*, *S. mikatae, and S. paradoxus*, are available but they are not completely annotated. On the basis of the sequence similarity and synteny of genes of these species in comparison to *S. cerevisiae*, putative *CEN* regions with CDEI-CDEII-CDEIII motifs in all 16 chromosomes in each of these three organisms were identified (65, 75, 81). In contrast, *S. castellii* does not seem to have a point centromere-like sequence (29). Centromeres in a pathogenic budding yeast, *Candida glabrata*, also contain elements highly homologous to those of *S. cerevisiae* (69). In this case, a 153-bp fragment accommodates all three DNA elements: CDEI (8 bp) and CDEIII (18 bp) are well conserved among all the chromosomes, while CDEII (77 to 79 bp) is AT rich (83 to 93%) (69). Like *S. cerevisiae* CDEIIIs, *C. glabrata* CDEIIIs have a central CCG sequence (35, 69). Both CDEI and CDEIII are required for centromere function. Some mutations in CDEIII cause a complete loss of centromere activity. The *C. glabrata* centromeres do not function in *S. cerevisiae*, suggesting that, in spite of the highly similar sequences, the centromere function is species specific. *Kluyveromyces lactis* and *Ashbya gossypii* (*Eremothecium gossypii*) also contain point centromeres. Mutual comparisons to nucleotide sequences of centromeres of *K. lactis* led to identification of consensus elements similar to those of *S. cerevisiae* (54, 55, 81). In *K. lactis*, a longer CDEII element (KlCDEII [160 bp long; ~90% AT rich]) that is almost double the size of the *S. cerevisiae* CDEII is flanked by two short, highly conserved boxes, CDEI and CDEIII. CDEIII also contains a conserved CCG element. An additional, 100-bp common AT-rich element (CDE0) is present at 150 bp upstream of CDEI. Centromeres of *K. lactis* are nonfunctional in *S. cerevisiae* and vice versa. Both the lengths and sequences of CDEI and CDEIII are essential for proper centromere function in *K. lactis*. Centromeres in *A. gossypii* are very similar to those in *K. lactis*, as the lengths of CDEII in both these organisms are approximately the same and double the length of the *S. cerevisiae* CDEII (34, 81). As with that of *S. cerevisiae*, the *CEN* sequences of *C. glabrata* and *K. lactis* are able to provide mitotic stability to an *ARS* plasmid (54, 69).

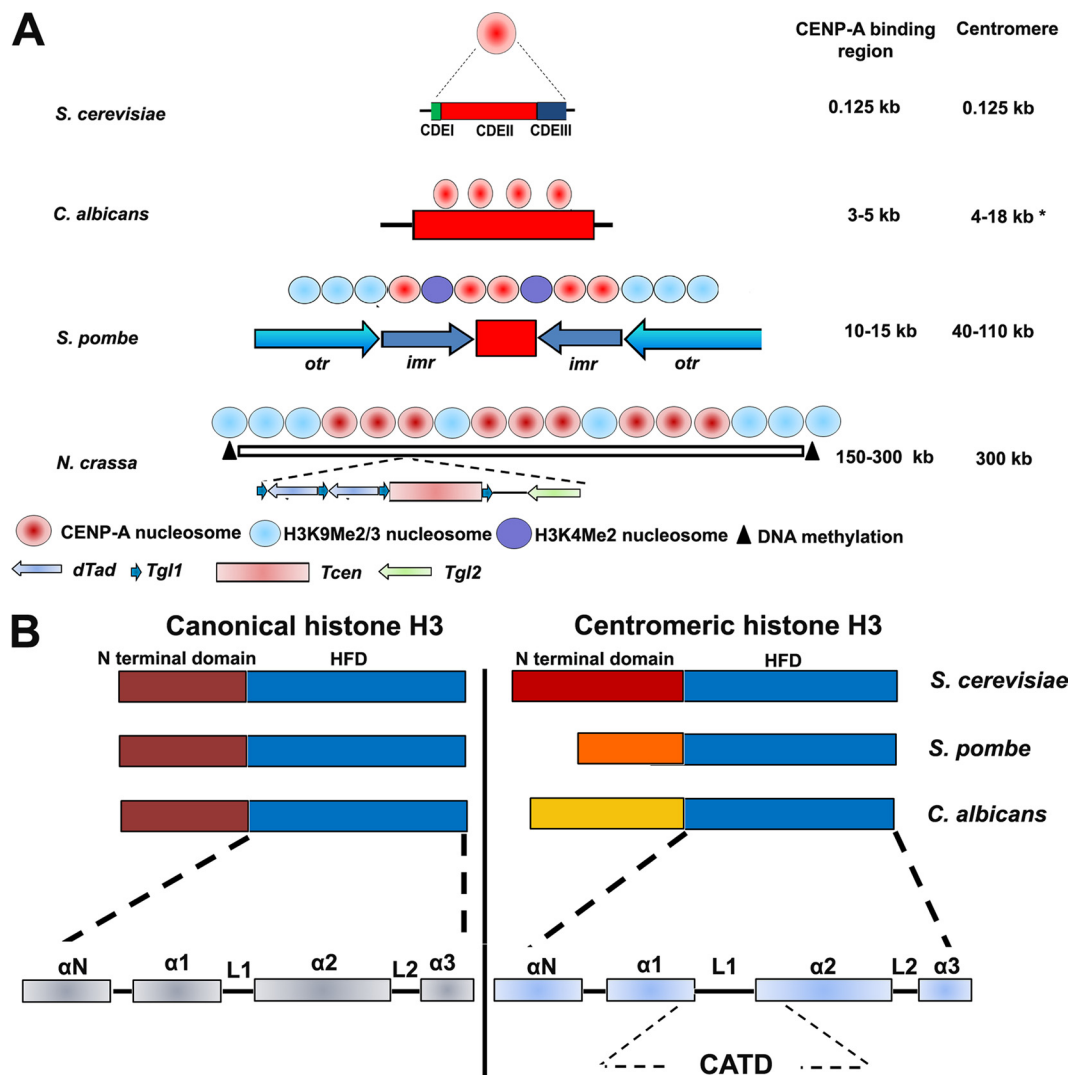Centromeres in two other budding yeasts, *Yarrowia lipolytica*

FIG. 1. Centromeres are highly diverse in fungi. (A) Organization of DNA elements and associated chromatin in point (*S. cerevisiae*), small regional (*C. albicans*), and large regional (*S. pombe* and *N. crassa*) centromeres. *, ORF-free regions containing CENP-A-rich centromeres are 4 to 18 kb long. (B) Domains of canonical histone H3 and centromere-specific histone H3 of the CENP-A family. While both the N-terminal domain and C-terminal histone fold domain (HFD) in histone H3 are highly invariant, the former is highly variable among centromeric histone H3 (CENP-A) proteins. Even the HFD is more variable in CenH3 than in histone H3. The α-helices (α-N, α-1, α-2, and α-3), loops (L1 and L2), and CENP-A targeting domain (CATD) present in the L1-α2 region of CenH3 are shown.

and *Candida maltosa*, have features that are slightly different from those of a point centromere. Centromeres in *Y. lipolytica* span up to ~200 bp but, surprisingly, lack any consensus sequences such as those seen in CDEI or CDEIII. Nonetheless, the AT richness of the *Y. lipolytica* centromere DNA is similar to that of *S. cerevisiae* CDEII (35, 43). Interestingly, the functions of *ARS* and *CEN* are interdependent in *Y. lipolytica* (131). The centromere DNA in *C. maltosa* contains consensus CDEI and AT-rich CDEII elements, but a conserved element equivalent to CDEIII is absent. However, a 325-bp region is sufficient to provide an *ARS* plasmid with high mitotic stability (94).

Large regional centromeres. Most organisms, except certain budding yeasts with short sequence-dependent point centromeres as described above, possess longer (>40 kb to a few megabases) centromere DNAs that are usually highly repeti-

tive and heterochromatic in nature. These centromeres are often referred to as "large regional" centromeres. Due to the absence of the motifs of the DNA binding proteins that are exclusively present in all centromere regions, the formation of centromeres in this class is probably mediated by a sequence-independent epigenetic mechanism that may or may not be conserved among all the species.

Centromeres of the fission yeast *S. pombe* are the best-studied centromeres in this class. Each centromere in this organism, originally identified in a DNA region of 40 to 110 kb, is organized symmetrically in a 10-to-15-kb CENP-A-rich chromatin that is flanked by an ~10-to-60-kb pericentric heterochromatin (119, 124, 138) (Fig. 1A). A nonhomologous unique sequence of 4 to 7 kb forms the central core (*cnt* or *cc*). The central cores of chromosome 1 (*cc1*) and 3 (*cc3*) have almost identical 3.3-kb-long "*tm*" elements. *cc2* contains a 1.5-kb-long

DNA element that is 48% identical to *tm* (138). Each *cc* is flanked on either side by innermost repeats (*imr*; also called *B* repeats) that have a unique sequence in each centromere. The DNA that contains *cc* and *imr* is flanked on either side by outer repeats (*otr*) that include "*dg*" and "*dh*" elements (also called *K* and *L* repeats) (10, 25, 40, 87, 99). The requirement of various centromere DNA elements for its activity was determined by a plasmid-based assay using *cen2* of *S. pombe* (10). That study revealed that *cc* and *K* repeats are essential and sufficient for maintaining an active centromere in *S. pombe*. All other elements are dispensable and not sufficient to complement the function of *cc* and *K* repeats. Even the *cc* region reveals functional redundancy, because deletion of short sequences from any part of this region does not cause a significant reduction in centromere function. Intriguingly, deletion of the entire *cc* results in abolishment of centromere function. The lack of a conserved sequence of any considerable length among all the three *cc* regions implies that the part of *cc* that is critical for *CEN* function in *S. pombe* is very small or degenerate.

The centromere in *Neurospora crassa*, a muticellular filamentous ascomycete, is another example of this class (21, 24, 115) (Fig. 1A). The centromere DNA is contained within a 300-kb region which is AT rich and recombination deficient and consists of centromere-specific repeat sequences. Unlike *S. pombe*, *N. crassa* *CEN*s have no inverted repeats. The centromere-specific repeats are highly divergent, indicating that repeat induced mutation (RIP) (111, 112), which causes C-to-T transitions at repetitive sequences with high frequency, is active in these loci (24). Further sequence analysis of a 16.1-kb region in *N. crassa* *CEN7* revealed that it contains a cluster of three retrotransposon-like elements (*Tcen*, *TglI*, and *Tgl2*) along with sequences of degenerate fragments of *Tad*, a previously characterized LINE retrotransposon (21).

Putative centromeres in another filamentous ascomycetous fungus, *Aspergillus nidulans*, are predicted to be associated with large repeat sequences which are less prone to meiotic recombination and consist of two degenerated LTR retrotransposons, *Dane1* and *Dane2* (Degenerated Aspergillus nidulans element). The prototype of degeneration suggests that, as seen with *N. crassa*, a process of RIP may be active in these sequences (2, 91).

The centromeric regions in the basidiomycetous fungus *Cryptococcus neoformans*, a human pathogen, were predicted by sequence analysis (74). Transposon (Tcn5, Tcn6)-rich gene-free regions of 40 to 110 kb, present once in each chromosome, are considered presumptive centromeres. This analysis was further supported by linkage studies of two marker genes, *URA5* and *ADE2*, which have been found to be linked to centromeres (60). Direct experimental evidence to define these putative regions as centromeres is still lacking.

The putative centromeres in another basidiomycete, *Coprinus cinereus*, a mushroom, lie within the transposon clusters that are localized at the less meiotic recombination-prone regions of the genome (117). These DNA elements can correspond to the sequence-independent regional centromeres, which are common in the fungi such as *C. neoformans* and *N. crassa* (24, 74, 115).

Small regional centromeres. Apart from the two major classes of fungal centromeres described above, centromeres of a third type, "small regional" centromeres, were identified in *Candida* species. Two closely related pathogenic *Candida* species, *Candida albicans* and *Candida dubliniensis*, have centromeres with several properties that are unique and thus absent from point and large regional centromeres. Each of the eight chromosomes of these two *Candida* species has a 3-to-5-kb-long CENP-A-rich centromere DNA, present in 4-to-18-kb open reading frame (ORF)-free regions, that has no common sequence motifs or repeats (95, 110) (Fig. 1A). Thus, each chromosome has a unique and different *CEN* DNA sequence. Unlike those in other budding yeasts, a circular *ARS* plasmid with a CENP-A-rich *CEN* region does not produce a stable minichromosome in *C. albicans*, suggesting that centromere formation is not strictly sequence dependent in this organism (11). Although chromosome-specific short inverted repeats were found at the centromeres of some of the chromosomes, a long inverted repeat is present only at *CEN5*. *CEN5* of *C. dubliniensis* is also flanked with similar long inverted-repeat sequences. It is also important that many point centromere-specific proteins, which are absent from organisms having regional centromeres, are also absent from these two *Candida* species (81).

The locations of the putative centromeres in three other yeasts of the *Candida* clade, *Candida lusitaniae*, *Pichia stipitis*, and *Debaryomyces hansenii* (75), have been predicted by sequence analysis. Analysis of the GC content of the total genome of these organisms identified a single GC-poor region in every chromosome. In *C. lusitaniae*, GC-poor troughs are located in intergenic regions approximately 4 kb in length. *C. albicans* and *C. dubliniensis* centromeres, which were identified experimentally, are also located in intergenic regions of similar length, although those DNA regions are not GC-poor (96, 110). Putative centromeres in *P. stipitis* and *D. hansenii* have clusters of retrotransposons.

CENTROMERE DNA IS RAPIDLY EVOLVING

Analysis of chromosome III of three closely related lineages of the wild yeast *Saccharomyces paradoxus* suggests that centromeres are the most rapidly evolving regions (12). Centromeres of the other chromosomes also show a higher rate of nucleotide substitutions than noncentromeric regions. More recently, rapid evolution of centromeres has been demonstrated by comparative genomic analyses in different species of fission yeast in the *Schizosaccharomyces* clade (104). *S. pombe* and *Schizosaccharomyces octosporus* contain no transposons at the centromere, whereas *Schizosaccharomyces japonicus* has clusters of transposons at the pericentric region. Although repeats are present at the centromeres of all the three *Schizosaccharomyces* species, the repeat sequences are completely dissimilar. Repeats present at the centromeres in different chromosomes of *S. pombe* or *S. octosporus* share high homology, suggesting homogenization of repeats by nonreciprocal recombination in the large array of inverted repeat structures. However, the lack of symmetry observed in the repeats of different chromosomes of *S. japonicus* indicates that transposition occurred more rapidly than homogenization. It has been argued that evolution of symmetric centromeric repeats occurred by suppression of transposition. Although *C. albicans* and *C. dubliniensis* exhibit a high degree of sequence homology through-

out the genome and the relative centromere locations in these two species are conserved, comparison of the *CEN* DNA sequences of orthologous chromosomes reveals very little or no conservation. Genome-wide analysis of orthologous noncoding regions in these two species shows a significantly lower rate of sequence diversity. Thus, comparative genome analysis suggests that centromeres are the most rapidly evolving loci in those two *Candida* species (95). Although several lines of evidence suggest that centromeres are rapidly evolving, the reason and the underlying mechanism of such rapid evolution are not clearly understood. It is possible that rapid changes in the centromere sequence may help in speciation, as centromeres function in a species-specific manner. It is worth noting that rapid evolution of centromeres has been shown to occur in plants and animals as well (77).

CENP-A: THE UNIVERSAL COMPONENT OF CENTROMERIC CHROMATIN

The lengths, sequences, and organizations of DNA elements at various centromeres in different fungi are seemingly diverse. In spite of this diversity in the characteristics of the *cis*-acting DNA elements, one common feature of all the functional centromeres is the exclusive presence of specialized chromatin marked by the centromere-specific histone H3 (CenH3) variant of the CENP-A family: Cse4 in *S. cerevisiae*, Cnp1 in *Schizosaccharomyces pombe*, CID in *Drosophila melanogaster*, and CENP-A in humans. Localization of most kinetochore proteins in many well-studied organisms has been shown to be regulated by CENP-A, confirming that CENP-A initiates kinetochore formation; thus, it is considered the epigenetic hallmark of a functional centromere (50). However, the presence of CENP-A alone is insufficient to initiate kinetochore formation in humans (46, 129). The mechanism by which only a few molecules of CENP-A compete with a high molar excess of canonical histone H3 molecules to make unique nucleosome structures at centromeric chromatin remains an enigma.

Strikingly, unlike the almost invariant histone H3 proteins, CENP-A proteins show a high degree of sequence diversity across species (Fig. 1B). The N-terminal domain (the essential N-terminal domain [END] in *S. cerevisiae*) of CENP-A (Cse4) is hypervariable among most species (51) except for members of the *Saccharomyces-Kluyveromyces* clade (7). Divergence of this region probably reflects differences in the underlying centromere DNA sequence that is believed to be undergoing rapid evolution (50). In addition, CENP-A proteins are known to be loaded at the centromere by the activity of a rapidly evolving chaperone Scm3 in fungi (5, 19, 100, 120, 137). CENP-A and H3 share a C-terminal globular <u>H</u>istone <u>F</u>old <u>D</u>omain (HFD), where amino acid sequence identity is ~50%. The HFD is shared by all histone proteins and is composed of three alpha helices ($\alpha 1$, $\alpha 2$, and $\alpha 3$) connected by loops (L1 and L2) (6). H3 and CENP-A contain an additional helix, the N-helix, present upstream of $\alpha 1$. L1 in CENP-A is longer than that of histone H3 (76). In *S. cerevisiae*, the L1-$\alpha 2$ region of CENP-A/Cse4, which contains the <u>CENP-A</u> <u>T</u>argeting <u>D</u>omain (CATD), interacts directly with the Cse4 binding domain of Scm3 (140).

Partial micrococcal nuclease (MNase) digestion along with DNase I digestion of chromatin in *S. cerevisiae* revealed more distinct ladder patterns at the centromeric chromatin than were seen in bulk chromatin (15). A distinctly protected region of 220 to 250 bp of centromeric chromatin was observed. A similar centromeric chromatin structure in *K. lactis* was reported as well (53). Partial MNase digestion analysis of both *S. pombe* and *C. albicans* revealed that the centromeric chromatin structure in each case is unusual. The bulk chromatin showed ~150-bp ladder patterns, while the centromeric chromatin exhibited smeary patterns of digestion (11, 101). The reason for this unusual pattern of centromeric chromatin seen *in vivo* is still unknown, but it may be attributable to replacement of histone H3 by CENP-A. CENP-A molecules have been shown to replace histone H3 molecules either partially or fully in centromeric chromatin in different fungi (17, 80, 110, 123). Studies in *S. pombe* have suggested that the unusual smeary pattern of partial MNase digestion could be due to protection provided by an intact kinetochore. Random access of MNase could potentially yield different sizes of *CEN* DNA fragments that appear to have a continuous smeary pattern due to blocking by the complex kinetochore architecture that is formed on CENP-A nucleosomes (116).

The unusual features of centromeric chromatin in different organisms suggest that the composition of CENP-A-containing nucleosomes is different from that of canonical nucleosomes (discussed in more detail in a review in reference 3). Strikingly, Scm3 replaces H2A-H2B dimers at centromeres in *S. cerevisiae* and forms hexameric nucleosomes containing two molecules each of Scm3, Cse4, and H4 (83). However, the existence of octameric nucleosomes is evident from another recent study on CENP-A/Cse4 nucleosomes in *S. cerevisiae* (20). A previous proteomic study performed with purified centromeric nucleosomes of *S. cerevisiae* revealed that H2A and H2B copurified with CENP-A/Cse4 (135). Not even the formation of intermediate or hybrid nucleosomes of CENP-A–H3 can be completely ruled out. The occurrence of tetrameric or half-nucleosomes is also a possibility, since CENP-A octamers have a higher propensity for disassembly (31). Although the exact composition of CENP-A nucleosomes in *S. cerevisiae* is still uncertain, it is plausible that the composition of the CENP-A nucleosome changes in different stages of the cell cycle (3).

Propagation of centromeric chromatin for newly replicated *CEN* DNA is another active area of research. Unlike the results seen in studies of humans and flies, replication of the centromere DNA in *C. albicans* and *S. pombe* occurs at the early S phase (67, 70). If CENP-A nucleosomes behave as do canonical H3-containing nucleosomes at the time of replication, one would assume that half of the preexisting CENP-A molecules would be segregated to each of the two sister chromatids. Using a fluorescence recovery after photobleaching (FRAP)-based method, it has been shown that old Cse4 molecules are replaced by new ones at the kinetochore in *S. cerevisiae* during the early S phase of the cell cycle (98). In *S. pombe*, incorporation of CENP-A/Cnp1 molecules occurs at both the S and late G_2 phases (125). Hence, it appears that CENP-A deposition may be coincident with replication, at least in unicellular yeasts (23).

Like incorporation of other histones, recruitment of CENP-A can also be affected by various events (including CENP-A expression, turnover of CENP-A protein or RNA, expression of other histones, and posttranslational modifica-
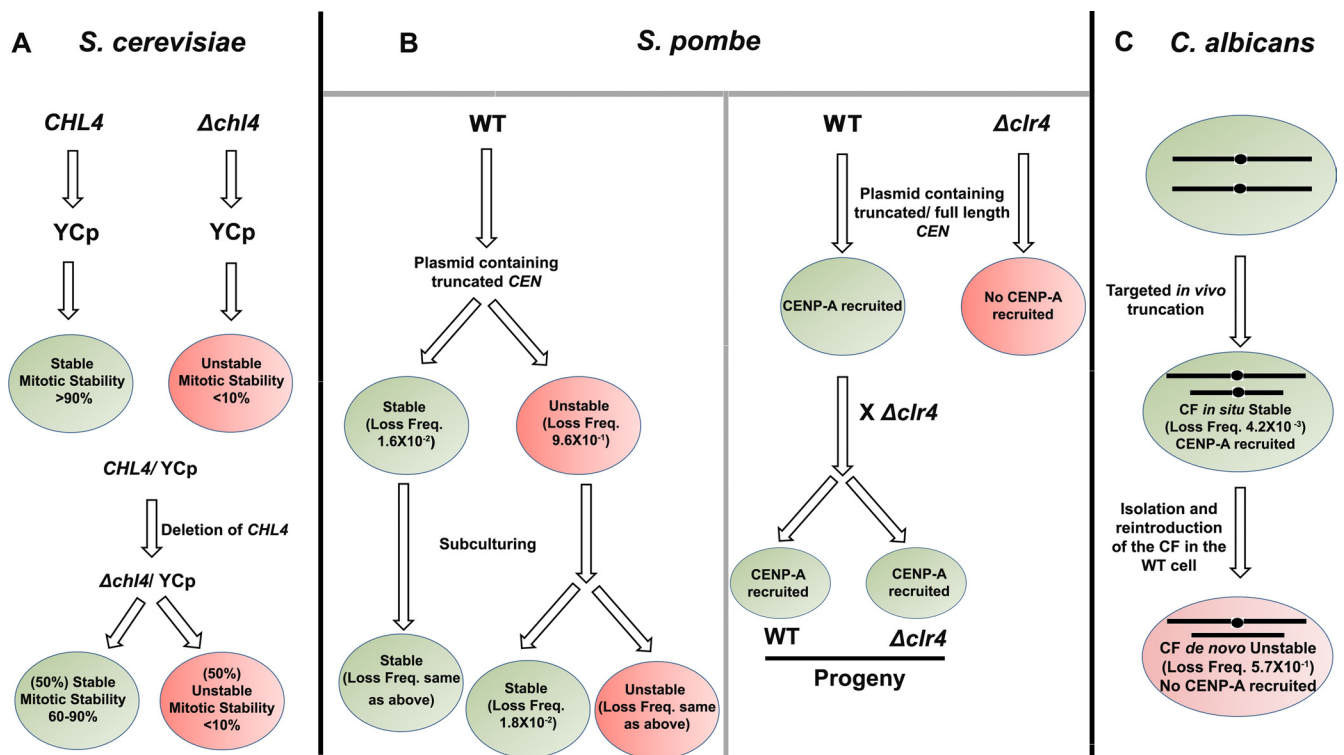
FIG. 2. Centromere "establishment" versus "propagation": epigenetic control of centromere function (see text for details). WT, wild type; Loss freq., loss frequency.

tions [PTMs]) and factors (including chaperons, assembly factors, essential components of kinetochore architecture, and histone modifiers) (3). In *S. cerevisiae*, CENP-A/Cse4 proteolysis, which is mediated by Psh1, an E3 ubiquitin ligase, restricts its localization primarily to the centromere (30, 56, 102). Overexpression of CENP-A/CaCse4 results in recruitment of more CENP-A/CaCse4 molecules along with other kinetochore proteins such as Mtw1 at the centromere in *C. albicans* (17, 106). Posttranslational modifications associated with the N-terminal histone tail are related to different functional states of chromatin-like repression (silencing) or activation of transcription. Although most of the longer centromeres are flanked by pericentric heterochromatin, surprisingly, the interspersed canonical H3 molecules show modifications that are characteristic of both euchromatin and heterochromatin (122) (Fig. 1A). These include molecules of H3K4Me2 (dimethylated lysine at the fourth position in histone H3), which are commonly observed in transcriptionally active regions, although these histone H3 molecules do not harbor any acetylation marks related to open chromatin. CENP-A-rich core centromere regions usually do not contain H3K9Me2 (dimethylated lysine at the ninth position in histone H3), a signature mark of closed or silent chromatin. The pericentric chromatin of long repetitive centromeres in various eukaryotes, however, harbors H3K9Me2/3 (13, 88, 92, 93, 103). Such H3K9Me2 marks are present but in *S. pombe* are restricted solely to pericentric heterochromatin (18, 23). Strikingly, H3K4Me2/3 marks, often found in centromeres of other eukaryotes, are absent from the *N. crassa* centromeres. Instead, H3K9Me3 marks are present and have been shown to colocalize with other kinetochore proteins at the centromeres in that organism. The flanking pericentric region has H3K9Me3 along with cytosine DNA methylation, which partially overlaps with kinetochore protein distribution. Interestingly, DNA methylation is directed by Dim5, a histone H3 methyltransferase that is also responsible for the presence of H3K9Me3 in *N. crassa*. Moreover, heterochromatin formation is essential for CENP-A localization at the centromere in that organism (115). Thus, the factors required for centromere formation in *N. crassa*, with a centromere organization more similar to that seen with higher eukaryotes, are different from those in *S. pombe*, carrying relatively simple long centromeres.

ESTABLISHMENT VERSUS PROPAGATION: EPIGENETIC REGULATION IN FUNGAL CENTROMERES

The epigenetic phenomenon has been defined as a change in phenotype that is heritable but does not involve DNA mutations. Fungal centromeres are excellent systems for the study of such epigenetic regulation (Fig. 2).

Epigenetic regulation in the point centromere has been documented to illustrate the differential requirements of centromere establishment and propagation of an already established centromere. The classic example of such regulation became obvious from studies of the behavior of a centromeric plasmid in the *chl4/mcm17/ctf17* kinetochore mutant of *S. cerevisiae* (86) (Fig. 2A). A centromeric plasmid (YCp) stably propagates in the wild-type strain but exhibited very low mitotic stability when introduced into this mutant, suggesting that *de novo* assembly of a kinetochore on *CEN* DNA requires the Chl4

protein (107). Interestingly, when the same centromeric plasmid was introduced into a wild-type cell and the *CHL4* gene was then deleted, two classes of transformants were obtained. About 50% of the transformants showed high mitotic stability of the centromeric plasmid, while the rest exhibited lower mitotic stability. That study revealed that two different chromatin states can exist and that Chl4 is not absolutely required for propagation of an already established centromere (Fig. 2A). Differential recruitment of cohesion (the molecular glue that prevents sister chromatid separation until anaphase onset) due to mutations in conserved DNA elements in a naive versus an established centromere also illustrates epigenetic propagation of point centromeres (126).

The epigenetic regulation of the *in vivo* centromere function of various *in vitro*-constructed minichromosomes was demonstrated in *S. pombe* (118). Cells carrying circular plasmids with incomplete centromere regions exhibited two states of mitotic stability: either unstable, as seen with an *ARS* plasmid without any centromere, or stable, as seen with a plasmid having a full-length centromere. The most striking observation from that study is that, when an unstable plasmid with a nonfunctional centromere switches to a stable state in which the centromere is fully activated, the active state is faithfully inherited for many generations (Fig. 2B). Once active, the plasmids with reduced centromeres exhibit mitotic stability that is indistinguishable from that seen with plasmids having a larger centromere DNA. The precise mechanism for this delayed or slow activation could not be determined, but it is postulated that formation of the higher-order centromeric chromatin structure required for its function, which is readily achieved when a plasmid having the full-length centromere is present, is a time-intensive process when the full-length centromere DNA is absent. That study thus set the stage for investigation of nongenetic factors that control centromere function. A striking observation was reported from a later study performed by Folco and coworkers (42). When plasmids containing the essential elements of the *S. pombe cen2* gene were transformed in *clr4* mutant cells (a histone H3 methyl transferase; see below), CENP-A/Cnp1 did not get recruited on the *cc2* region of *cen2* in plasmids. However, when wild-type cells were transformed with the same set of plasmids, proper CENP-A chromatin was established on the *cc2* region of the plasmids, confirming that Clr4 is essential for establishment of CENP-A chromatin. However, when the wild-type strain transformed with these *CEN* plasmids was crossed with *clr4* mutant strain, surprisingly, both the wild-type and *clr4* mutant progeny strains exhibited proper recruitment of CENP-A on the plasmid centromere, implying that inheritance of an established CENP-A chromatin does not require Clr4.

The most striking example of epigenetic regulation of centromere function was observed in *C. albicans* (11). In that organism, attempts to construct a minichromosome with an *ARS* and a *CEN* sequence failed. Even ectopic insertion of the centromere DNA into a nearby locus could not deposit CENP-A to form a functional kinetochore, suggesting that *de novo* centromere formation on an introduced *CEN* DNA does not occur in *C. albicans*. However, a short 85-kb chromosome fragment (CF) carrying a functional centromere obtained by *in vivo* telomere-mediated truncation of native chromosome 7 was found to be mitotically stable (Fig. 2C). Strikingly, when total DNA isolated from this strain was used to reintroduce the CF in wild-type *C. albicans*, it was found to be unstable in the transformed cells. MNase digestion assays confirmed that, after reintroduction, CENP-A molecules could not be deposited on the introduced *CEN* DNA sequence of the CF, implying that an epigenetic memory is required for *de novo* assembly of a functional centromere.

Thus, regulation of the centromere function of each of the three types of fungal centromeres (point, large regional, and small regional) clearly demonstrates that the requirement for *de novo* centromere assembly can be different from that for centromere propagation, which is largely epigenetically determined.

THE NEOCENTROMERE: FORMATION OF A CENTROMERE ON A NONNATIVE LOCUS

A centromere can occasionally form on a nonnative location on the same chromosome when a native centromere is deleted or inactivated. Formation of a new centromere, more popularly known as the neocentromere, was first evidenced in humans several years ago (9, 109). Since neocentromeres can form on DNA sequences that share no common sequence features (i.e., neither DNA sequence nor features of DNA elements such as repetitive, gene-free regions) with a native centromere, the discovery of neocentromeres suggested redundancy in the sequence requirement for formation of a functional regional centromere. Subsequently, neocentromere formation has been demonstrated in fungi, including both *S. pombe* and *C. albicans*, each of which has been shown to carry epigenetically determined centromeres (61, 66).

The requirement for centromere formation in the fission yeast *S. pombe*, which carries a metazoan-like centromere organization with pericentric heterochromatic repeats on both sides of a nonrepetitive CENP-A-rich core region, is more complex but better understood. In a study using a conditional strain, *cen1*-deleted so-called "acentric" chromosome-containing survivors were recovered (61). Whereas the acentric chromosome has been found in most cases to be fused to one of the other two chromosomes, in rare instances neocentromeres have formed. The neocentromeric regions did not have centromeric repeats but rather formed on regions with several ORFs that are normally upregulated under conditions of nitrogen starvation. Interestingly, once neocentromeres are formed, these ORFs are found to be largely silenced, suggesting a possible correlation of transcriptional repression and neocentromere seeding. Binding of several key kinetochore proteins such as CENP-A, Mis12, and CENP-C was found to be associated with a 20-kb region of the neocentromere, as seen with the native centromere. One of the neocentromere regions is in the proximity of a telomere where the chromatin had already been modified by H3K9Me2 (see below). The other neocentromere also formed close to a telomere. In this case, interestingly, the formation of the neocentromere stimulated the accumulation of H3K9Me2 at a proximal region that is devoid of such marks in wild-type strains. These phenomena again imply that there is a possible interdependency between neocentromere formation and heterochromatinization. The neocentromere-containing chromosomes appeared to have been replicated and segregated normally through the cell cycle.

In the absence of a protein involved in heterochromatin formation such as Swi6/HP1, Dicer (RNA interference [RNAi] machinery), or Clr4 (H3K9 methyltransferase), however, neocentromeres can still form, but the ratio of neocentromere formation to telomere-telomere fusion is drastically reduced compared to wild-type observations. This result indicates that the presence of neighboring heterochromatin can provide assistance but is not essential for *de novo* neocentromere formation. One of the striking observations from that study is the role of telomere-proximal regions in serving the platform for centromere formation. This may support the hypothesis that centromeres originated from telomeres during karyotype evolution (132).

Native centromere sequences of *C. albicans* do not show any common sequence motifs or repeats (110). Deletion of the CENP-A-rich centromere region from chromosome 7 in *C. albicans* leads to a high rate of chromosome loss (11). However, an inverted repeat is associated only with the centromere of chromosome 5. When the *CEN5* region was replaced with a marker gene, the so-called "acentric" chromosome remained stable in mitosis (66). Further investigation suggested formation at an astonishingly high rate of a neocentromere on the chromosome devoid of the native centromere. Two classes of neocentromeres, centromere-proximal and centromere-distal neocentromeres, were observed. In both cases, the extent of CENP-A binding (~3 kb) was found to be similar to that seen with the native centromere, albeit the number of CENP-A molecules on neocentromeres was found to be less than on the native centromeres. Moreover, a chromosome with a neocentromere shows mitotic stability as high as that seen with the native centromere. Interestingly, as with humans, neocentromeres show no sequence similarities with the native centromere, suggesting that factors other than the DNA sequence alone determine centromere and neocentromere location. Since neocentromeres can form very efficiently at multiple locations, at least on chromosome 5 in *C. albicans*, it is intriguing that centromere location was found to be conserved in a variety of the clinical isolates of *C. albicans* examined (82).

RNA INTERFERENCE: ROLE OF RNAi IN THE CENTROMERE FUNCTION IN FUNGI

RNA interference (RNAi) is a conserved eukaryotic process that is triggered by double-stranded RNA (dsRNA) and results in transcriptional repression or silencing of genes with complementary sequences. The term "RNAi" was first used to represent silencing mediated by exogenous dsRNA in *Caenorhabditis elegans* (39). However, it is now used broadly to describe both endogenous and exogenous gene silencing mediated by small RNAs of various types: short interfering RNAs (siRNAs), micro-RNA (miRNA), and PIWI-interacting RNAs (piRNAs) (reviewed in references 72 and 84). Among these, siRNAs are found in fungi. More recently, a new type of small RNA degradation product, referred to as primal RNA (priRNA), has been shown to be associated with pericentric heterochromatin formation in *S. pombe* centromeres (48).

S. pombe has two distinct chromatin subdomains at the centromere: a core CENP-A-rich central domain where most histone H3 molecules are replaced by CenH3 molecules and the pericentric heterochromatic domain on outer repeats that

form chromatin with H3K9Me2 molecules (13, 88, 93). However, a transgene placed in the central region that supports kinetochore formation shows reversible gene silencing, indicating the facultative nature of heterochromatin at the core centromere (4, 37, 134). As mentioned above, the histone H3 molecules that remain at the central regions have been found to be H3K4Me2 molecules, a mark associated with chromatin with active transcription (18). A landmark discovery of about a decade ago demonstrated a relation between RNAi machinery and pericentric heterochromatin formation in *S. pombe* (133, 134). The RNAi pathway in *S. pombe* is mediated by three key members, each encoded by a single gene, Dicer ($dcr1^+$), Argonaute ($ago1^+$), or RNA-dependent RNA polymerase ($rdp1^+$) (141)(Table 1). Heterochromatin formation at the centromere requires a number of other *trans*-acting factors: histone deacetylase (HDACs), Clr4 (histone H3K9 methyltransferase [HKMT]), and the histone H3K9-methyl binding proteins such as Swi6 (an HP1 homolog) and Chp1 (8, 44, 62, 88, 90, 103, 108, 113, 121, 139). Both strands of the outer repeats are transcribed by RNA polymerase II (RNA PolII) during the S phase; the dsRNA is recognized by Dicer (Dcr1), the RNase III protein, which cleaves dsRNAs into 21-to-23-nucleotide-long complementary siRNAs (14). Dcr1, which presumably shuttles between the cytoplasm and nucleus, needs to be retained in the nucleus for production of siRNAs (38, 49, 78). Argonaute-related proteins bind to the Dcr1-processed siRNAs through their PIWI domains and thus serve as the core components of RNA-induced silencing complexes (RISCs) (22). Ago1, Chp1, and Tas3 are components of RNA-induced transcriptional silencing (RITS) complexes (130). Ago1 "slices" one strand of the dsRNA, releasing the other strand to form an activated RISC with a short single-stranded RNA (97, 105). This nascent siRNA transcript then binds to the homologous chromatin of the *otr* repeats to load an RITS complex. The RNAi pathway is eventually amplified by the RNA-dependent RNA polymerase complex (RDRC) composed of Rdp1, Hrr1, and Cid12 (85). This cyclic process makes further dsRNA and triggers synthesis of siRNA (85). Finally, the *Clr4-Rik1-Cul* (CLRC) complex is recruited. Clr4, a member of the CLRC complex, mediates H3K9 methylation on the chromatin bound to it, resulting in heterochromatin formation and spreading and transcription silencing (58, 59, 63, 73, 127). The spreading is mediated by other chromodomain proteins such as Swi6, Chp2 and Chp1 (8, 108, 139).

Is the RNAi machinery essential for pericentric heterochromatin formation that is required for centromere function? Deletion of any of the key genes ($dcr1^+$, $ago1^+$, or $rdp1^+$) results in accumulation of forward and reverse transcripts from *otr*, loss of centromeric heterochromatin, loss of H3K9 methylation, and increased chromosome loss (133). Thus, it was believed that the RNAi pathway is essential for recruitment of Clr4 and Swi6 for establishment, spreading, and propagation of centromeric heterochromatin. However, other studies suggest alternative mechanisms. It has been previously shown that the RNAi machinery is absolutely required for establishment of H3K9 methylation on the pericentric repeats but that propagation of an already-established heterochromatin is epigenetic and thus does not depend on RNAi (108). It has been formally demonstrated that, in the absence of Ago1, overexpression of Clr4 alone can result in H3K9 methylation at the centromere

TABLE 1. Genetic and epigenetic determinants of various representative fungal centromeres

Centromere type	Organism	Genetic determinant (DNA feature)			Epigenetic determinant				
		Length	CDE(s)	Repeat(s)	H3 PTM/DNA methylation	Presence of RNAi component: Ago	Dcr	RdRp	RNAi needed for heterochromatin
Point	S. cerevisiae	~125 bp	I-II-III	No	ND[a]	No	No	No	NA[b]
	K. lactis	~186 bp	0-I-II-III	No	ND	No	No	No	NA
	Y. lipolytica	~200 bp	CDEI and CDEIII absent, CDEII-like	No	ND	ND	ND	ND	NA
Large regional	S. pombe	40–110 kb	No	Unique cc flanked by inverted repeats (imr and otr)	H3K4Mc2 (cc), H3K9Me2 (otr)	Yes	Yes	Yes	Yes
	N. crassa	~300 kb	No	Nonconserved AT-rich, transposon-rich repeats	H3K9Me3 CEN and pericentric region, cytosine DNA methylation	Yes	Yes	Yes	No
	C. neoformans	40–110 kb	No	Transposon-rich repeats	ND	Yes	Yes	Yes	ND
Small regional	C. albicans	4–18 kb[c]	No	No (except CEN5)	ND	Yes	Dcr-like protein	No	ND

[a] ND, not determined.
[b] NA, not applicable.
[c] In C. albicans, a 3- to 5-kb CENP-A/CaCse4-rich CEN sequence is located within a 4- to 18-kb ORF-free region.

(114). Even tethering Clr4 to a noncentromeric region could induce expression of artificial heterochromatin (64). Thus, the dependency on RNAi and heterochromatin formation at the centromere in *S. pombe* has not been fully understood. Nevertheless, the involvement of siRNA produced from centromeric repeats has been shown to be essential for centromere function and transcriptional silencing via heterochromatin formation in *S. pombe*, which completely lacks transposons at centromeres (47). However, a recent study revealed that the process of siRNA production is active in transposon-rich regions of centromeres and subtelomeric regions in another *Schizosaccharomyces* species, *S. japonicus*. This observation suggests that, in *S. pombe*, either the RNAi machinery targets repetitive elements instead of mobile elements or the pathway has changed from its previous role of transposon silencing to heterochromatinization (104). In addition, the relationship between RNAi and DNA methylation, a common feature of heterochromatin observed in *N. crassa* and mammals but not in fission yeast or flies, has been examined. Interestingly, in *N. crassa*, HP1 recruitment and DNA methylation, components of heterochromatin, were found to be independent of the RNAi machinery (45). That study indicates that, in contrast to the results seen with *S. pombe*, *Neurospora* centromeric chromatin formation is independent of components of RNAi machinery.

CONCLUSIONS

In spite of performing a well-conserved function, centromere DNAs are substantially divergent in length, sequence, and composition as well as in the requirement for various genetic and epigenetic factors (Table 1). Some of the essential factors required for *de novo* assembly of a kinetochore for propagation of an established centromere in a given organism appear to be redundant. Since new centromere seeding must occur at the time of neocentromere formation, further mechanistic insights in this process may provide us with useful information on factors required for the establishment of epigenetically determined centromeres. Rapid coevolution of centromere DNA sequences, the N-terminal domain of CENP-A, and its chaperone Scm3 is particularly striking. Thus, it is of great interest to find the biological implications associated with rapid evolution of centromeres, previously known to occur in plants and animals and now evident in fungi as well. We are now at an exciting stage of investigation of this mysterious darkly stained region of a chromosome.

REFERENCES

1. **Aker, M., and H. Huang.** 1996. Extreme heterogeneity of minor satellite repeat arrays in inbred strains of mice. Mamm. Genome **7:**62–64.
2. **Aleksenko, A., M. L. Nielsen, and A. J. Clutterbuck.** 2001. Genetic and physical mapping of two centromere-proximal regions of chromosome IV in *Aspergillus nidulans*. Fungal Genet. Biol. **32:**45–54.
3. **Allshire, R. C., and G. H. Karpen.** 2008. Epigenetic regulation of centromeric chromatin: old dogs, new tricks? Nat. Rev. Genet. **9:**923–937.
4. **Allshire, R. C., E. R. Nimmo, K. Ekwall, J. P. Javerzat, and G. Cranston.** 1995. Mutations derepressing silent centromeric domains in fission yeast disrupt chromosome segregation. Genes Dev. **9:**218–233.

5. **Aravind, L., L. M. Iyer, and C. Wu.** 2007. Domain architectures of the Scm3p protein provide insights into centromere function and evolution. Cell Cycle **6:**2511–2515.

6. **Arents, G., R. W. Burlingame, B. C. Wang, W. E. Love, and E. N. Moud-rianakis.** 1991. The nucleosomal core histone octamer at 3.1 A resolution: a tripartite protein assembly and a left-handed superhelix. Proc. Natl. Acad. Sci. U. S. A. **88:**10148–10152.

7. **Baker, R. E., and K. Rogers.** 2006. Phylogenetic analysis of fungal centromere H3 proteins. Genetics **174:**1481–1492.

8. **Bannister, A. J., et al.** 2001. Selective recognition of methylated lysine 9 on histone H3 by the HP1 chromo domain. Nature **410:**120–124.

9. **Barry, A. E., E. V. Howman, M. R. Cancilla, R. Saffery, and K. H. Choo.** 1999. Sequence analysis of an 80 kb human neocentromere. Hum. Mol. Genet. **8:**217–227.

10. **Baum, M., V. K. Ngan, and L. Clarke.** 1994. The centromere K-type repeat and the central core are together sufficient to establish a functional *Schizosaccharomyces pombe* centromere. Mol. Biol. Cell **5:**747–761.

11. **Baum, M., K. Sanyal, P. K. Mishra, N. Thaler, and J. Carbon.** 2006. Formation of functional centromeric chromatin is specified epigenetically in *Candida albicans*. Proc. Natl. Acad. Sci. U. S. A. **103:**14877–14882.

12. **Bensasson, D., M. Zarowiecki, A. Burt, and V. Koufopanou.** 2008. Rapid evolution of yeast centromeres in the absence of drive. Genetics **178:**2161–2167.

13. **Bernard, P., et al.** 2001. Requirement of heterochromatin for cohesion at centromeres. Science **294:**2539–2542.

14. **Bernstein, E., A. A. Caudy, S. M. Hammond, and G. J. Hannon.** 2001. Role for a bidentate ribonuclease in the initiation step of RNA interference. Nature **409:**363–366.

15. **Bloom, K. S., and J. Carbon.** 1982. Yeast centromere DNA is in a unique and highly ordered structure in chromosomes and small circular minichromosomes. Cell **29:**305–317.

16. **Brown, W. R. A., and Z.-Y. Xu.** 2009. The 'kinetochore maintenance loop'—the mark of regulation? BioEssays **31:**228–236.

17. **Burrack, L., S. Shelly, E. Applen, and J. Berman.** 2011. The requirement for the Dam1 complex is dependent upon the number of kinetochore proteins and microtubules. Curr. Biol. **21:**889–896.

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

19. **Camahort, R., et al.** 2007. Scm3 is essential to recruit the histone H3 variant Cse4 to centromeres and to maintain a functional kinetochore. Mol. Cell **26:**853–865.

20. **Camahort, R., et al.** 2009. Cse4 is part of an octameric nucleosome in budding yeast. Mol. Cell **35:**794–805.

21. **Cambareri, E. B., R. Aisner, and J. Carbon.** 1998. Structure of the chromosome VII centromere region in *Neurospora crassa*: degenerate transposons and simple repeats. Mol. Cell. Biol. **18:**5465–5477.

22. **Carmell, M. A., Z. Xuan, M. Q. Zhang, and G. J. Hannon.** 2002. The Argonaute family: tentacles that reach into RNAi, developmental control, stem cell maintenance, and tumorigenesis. Genes Dev. **16:**2733–2742.

23. **Castillo, A. G., et al.** 2007. Plasticity of fission yeast CENP-A chromatin driven by relative levels of histone H3 and H4. PLoS Genet. **3:**e121.

24. **Centola, M., and J. Carbon.** 1994. Cloning and characterization of centromeric DNA from *Neurospora crassa*. Mol. Cell. Biol. **14:**1510–1519.

25. **Clarke, L., H. Amstutz, B. Fishel, and J. Carbon.** 1986. Analysis of centromeric DNA in the fission yeast *Schizosaccharomyces pombe*. Proc. Natl. Acad. Sci. U. S. A. **83:**8253–8257.

26. **Clarke, L., and J. Carbon.** 1980. Isolation of a yeast centromere and construction of functional small circular chromosomes. Nature **287:**504–509.

27. **Clarke, L., and J. Carbon.** 1980. Isolation of the centromere-linked CDC10 gene by complementation in yeast. Proc. Natl. Acad. Sci. U. S. A. **77:**2173–2177.

28. **Clarke, L., and J. Carbon.** 1983. Genomic substitutions of centromeres in *Saccharomyces cerevisiae*. Nature **305:**23–28.

29. **Cliften, P. F., R. S. Fulton, R. K. Wilson, and M. Johnston.** 2006. After the duplication: gene loss and adaptation in *Saccharomyces* genomes. Genetics **172:**863–872.

30. **Collins, K. A., S. Furuyama, and S. Biggins.** 2004. Proteolysis contributes to the exclusive centromere localization of the yeast Cse4/CENP-A histone H3 variant. Curr. Biol. **14:**1968–1972.

31. **Conde e Silva, N., et al.** 2007. CENP-A-containing nucleosomes: easier disassembly versus exclusive centromeric localization. J. Mol. Biol. **370:**555–573.

32. **Copenhaver, G. P., et al.** 1999. Genetic definition and sequence analysis of *Arabidopsis* centromeres. Science **286:**2468–2474.

33. **Cumberledge, S., and J. Carbon.** 1987. Mutational analysis of meiotic and mitotic centromere function in *Saccharomyces cerevisiae*. Genetics **117:**203–212.

34. **Dietrich, F. S., et al.** 2004. The *Ashbya gossypii* genome as a tool for mapping the ancient *Saccharomyces cerevisiae* genome. Science **304:**304–307.

35. **Dujon, B., et al.** 2004. Genome evolution in yeasts. Nature **430:**35–44.

36. **Earnshaw, W. C., and N. Rothfield.** 1985. Identification of a family of human centromere proteins using autoimmune sera from patients with scleroderma. Chromosoma **91:**313–321.

37. **Ekwall, K., et al.** 1996. Mutations in the fission yeast silencing factors clr4+ and rik1+ disrupt the localisation of the chromo domain protein Swi6p and impair centromere function. J. Cell Sci. **109:**2637–2648.

38. **Emmerth, S., et al.** 2010. Nuclear retention of fission yeast Dicer is a prerequisite for RNAi-mediated heterochromatin assembly. Dev. Cell **18:**102–113.

39. **Fire, A., et al.** 1998. Potent and specific genetic interference by double-stranded RNA in *Caenorhabditis elegans*. Nature **391:**806–811.

40. **Fishel, B., H. Amstutz, M. Baum, J. Carbon, and L. Clarke.** 1988. Structural organization and functional analysis of centromeric DNA in the fission yeast *Schizosaccharomyces pombe*. Mol. Cell. Biol. **8:**754–763.

41. **Fitzgerald-Hayes, M., L. Clarke, and J. Carbon.** 1982. Nucleotide sequence comparisons and functional analysis of yeast centromere DNAs. Cell **29:**235–244.

42. **Folco, H. D., A. L. Pidoux, T. Urano, and R. C. Allshire.** 2008. Heterochromatin and RNAi are required to establish CENP-A chromatin at centromeres. Science **319:**94–97.

43. **Fournier, P., et al.** 1993. Colocalization of centromeric and replicative functions on autonomously replicating sequences isolated from the yeast *Yarrowia lipolytica*. Proc. Natl. Acad. Sci. U. S. A. **90:**4912–4916.

44. **Freeman-Cook, L. L., et al.** 2005. Conserved locus-specific silencing functions of *Schizosaccharomyces pombe* sir2+. Genetics **169:**1243–1260.

45. **Freitag, M., et al.** 2004. DNA methylation is independent of RNA interference in *Neurospora*. Science **304:**1939.

46. **Gascoigne, K., et al.** 2011. Induced ectopic kinetochore assembly bypasses the requirement for CENP-A nucleosomes. Cell **145:**410–422.

47. **Grewal, S. I. S.** 2010. RNAi-dependent formation of heterochromatin and its diverse functions. Curr. Opin. Genet. Dev. **20:**134–141.

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

49. **Hayashi, A., et al.** 2009. Localization of gene products using a chromosomally tagged GFP-fusion library in the fission yeast *Schizosaccharomyces pombe*. Genes Cells **14:**217–225.

50. **Henikoff, S., K. Ahmad, and H. S. Malik.** 2001. The centromere paradox: stable inheritance with rapidly evolving DNA. Science **293:**1098–1102.

51. **Henikoff, S., and Y. Dalal.** 2005. Centromeric chromatin: what makes it unique? Curr. Opin. Genet. Dev. **15:**177–184.

52. **Heslop-Harrison, J. S., A. Brandes, and T. Schwarzacher.** 2003. Tandemly repeated DNA sequences and centromeric chromosomal regions of *Arabidopsis* species. Chromosome Res. **11:**241–253.

53. **Heus, J. J., K. S. Bloom, B. J. M. Zonneveld, H. Y. Steensma, and J. A. Berg.** 1993. Chromatin structures of *Kluyveromyces lactis* centromeres in *K. lactis* and *Saccharomyces cerevisiae*. Chromosoma **102:**660–667.

54. **Heus, J. J., Z. B. H. Y. de Steensma, and J. A. van den Berg.** 1993. The consensus sequence of *Kluyveromyces lactis* centromeres shows homology to functional centromeric DNA from *Saccharomyces cerevisiae*. Mol. Gen. Genet. **236:**355–362.

55. **Heus, J. J., Z. B. H. Y. Steensma, and J. A. Van den Berg.** 1994. Mutational analysis of centromeric DNA elements of *Kluyveromyces lactis* and their role in determining the species specificity of the highly homologous centromeres from *K. lactis* and *Saccharomyces cerevisiae*. Mol. Gen. Genet. **243:**325–333.

56. **Hewawasam, G., et al.** 2010. Psh1 is an E3 ubiquitin ligase that targets the centromeric histone variant Cse4. Mol. Cell **40:**444–454.

57. **Hieter, P., et al.** 1985. Functional selection and analysis of yeast centromeric DNA. Cell **42:**913–921.

58. **Hong, E.-J., J. Villén, E. L. Gerace, S. P. Gygi, and D. Moazed.** 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.

59. **Horn, P. J., J.-N. Bastie, and C. L. Peterson.** 2005. A Rik1-associated, cullin-dependent E3 ubiquitin ligase is essential for heterochromatin formation. Genes Dev. **19:**1705–1714.

60. **Idnurm, A.** 2010. A tetrad analysis of the basidiomycete fungus *Cryptococcus neoformans*. Genetics **185:**153–163.

61. **Ishii, K., et al.** 2008. Heterochromatin integrity affects chromosome reorganization after centromere dysfunction. Science **321:**1088–1091.

62. **Ivanova, A. V., M. J. Bonaduce, S. V. Ivanov, and A. J. S. Klar.** 1998. The chromo and SET domains of the Clr4 protein are essential for silencing in fission yeast. Nat. Genet. **19:**192–195.

63. **Jia, S., R. Kobayashi, and S. I. S. Grewal.** 2005. Ubiquitin ligase component Cul4 associates with Clr4 histone methyltransferase to assemble heterochromatin. Nat. Cell Biol. **7:**1007–1013.

64. **Kagansky, A., et al.** 2009. Synthetic heterochromatin bypasses RNAi and centromeric repeats to establish functional centromeres. Science **324:**1716–1719.

65. **Kellis, M., N. Patterson, M. Endrizzi, B. Birren, and E. S. Lander.** 2003. Sequencing and comparison of yeast species to identify genes and regulatory elements. Nature **423:**241–254.

66. **Ketel, C., et al.** 2009. Neocentromeres form efficiently at multiple possible loci in *Candida albicans*. PLoS Genet. **5:**e1000400.

67. **Kim, S.-M., D. D. Dubey, and J. A. Huberman.** 2003. Early-replicating heterochromatin. Genes Dev. **17:**330–335.

68. **Kipling, D., H. E. Ackford, B. A. Taylor, and H. J. Cooke.** 1991. Mouse minor satellite DNA genetically maps to the centromere and is physically linked to the proximal telomere. Genomics **11:**235–241.

69. **Kitada, K., E. Yamaguchi, K. Hamada, and M. Arisawa.** 1997. Structural analysis of a *Candida glabrata* centromere and its functional homology to the *Saccharomyces cerevisiae* centromere. Curr. Genet. **31:**122–127.

70. **Koren, A., et al.** 2010. Epigenetically-inherited centromere and neocentromere DNA replicates earliest in S-phase. PLoS Genet. **6:**e1001068.

71. **Lefrançois, P., et al.,** 2009. Efficient yeast ChIP-Seq using multiplex short-read DNA sequencing. BMC Genomics **10:**37.

72. **Lejeune, E., and R. C. Allshire.** 2011. Common ground: small RNA programming and chromatin modifications. Curr. Opin. Cell Biol. **23:**258–265.

73. **Li, F., et al.** 2005. Two novel proteins, Dos1 and Dos2, interact with Rik1 to regulate heterochromatic RNA interference and histone modification. Curr. Biol. **15:**1448–1457.

74. **Loftus, B. J., et al.,** 2005. The genome of the basidiomycetous yeast and human pathogen *Cryptococcus neoformans*. Science **307:**1321–1324.

75. **Lynch, D. B., M. E. Logue, G. Butler, and K. H. Wolfe.** 2010. Chromosomal G + C content evolution in yeasts: systematic interspecies differences, and GC-poor troughs at centromeres. Genome Biol. Evol. **2:**572–583.

76. **Malik, H. S., and S. Henikoff.** 2003. Phylogenomics of the nucleosome. Nat. Struct. Biol. **10:**882–891.

77. **Malik, H. S., and S. Henikoff.** 2009. Major evolutionary transitions in centromere complexity. Cell **138:**1067–1082.

78. **Matsuyama, A., et al.** 2006. ORFeome cloning and global analysis of protein localization in the fission yeast *Schizosaccharomyces pombe*. Nat. Biotechnol. **24:**841–847.

79. **McGrew, J., B. Diehl, and M. Fitzgerald-Hayes.** 1986. Single base-pair mutations in centromere element III cause aberrant chromosome segregation in *Saccharomyces cerevisiae*. Mol. Cell. Biol. **6:**530–538.

80. **Meluh, P. B., P. Yang, L. Glowczewski, D. Koshland, and M. M. Smith.** 1998. Cse4p is a component of the core centromere of *Saccharomyces cerevisiae*. Cell **94:**607–613.

81. **Meraldi, P., A. McAinsh, E. Rheinbay, and P. Sorger.** 2006. Phylogenetic and structural analysis of centromeric DNA and kinetochore proteins. Genome Biol. **7:**R23.

82. **Mishra, P., M. Baum, and J. Carbon.** 2007. Centromere size and position in *Candida albicans* are evolutionarily conserved independent of DNA sequence heterogeneity. Mol. Genet. Genomics **278:**455–465.

83. **Mizuguchi, G., H. Xiao, J. Wisniewski, M. M. Smith, and C. Wu.** 2007. Nonhistone Scm3 and histones CenH3-H4 assemble the core of centromere-specific nucleosomes. Cell **129:**1153–1164.

84. **Moazed, D.** 2009. Small RNAs in transcriptional gene silencing and genome defence. Nature **457:**413–420.

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

86. **Mythreye, K., and K. S. Bloom.** 2003. Differential kinetochore protein requirements for establishment versus propagation of centromere activity in *Saccharomyces cerevisiae*. J. Cell Biol. **160:**833–843.

87. **Nakaseko, Y., N. Kinoshita, and M. Yanagida.** 1987. A novel sequence common to the centromere regions of *Schizosaccharomyces pombe* chromosomes. Nucleic Acids Res. **15:**4705–4715.

88. **Nakayama, J., J. C. Rice, B. D. Strahl, C. D. Allis, and S. I. Grewal.** 2001. Role of histone H3 lysine 9 methylation in epigenetic control of heterochromatin assembly. Science **292:**110–113.

89. **Ng, R., and J. Carbon.** 1987. Mutational and in vitro protein-binding studies on centromere DNA from *Saccharomyces cerevisiae*. Mol. Cell. Biol. **7:**4522–4534.

90. **Nicolas, E., et al.** 2007. Distinct roles of HDAC complexes in promoter silencing, antisense suppression and DNA damage protection. Nat. Struct. Mol. Biol. **14:**372–380.

91. **Nielsen, M. L., T. D. Hermansen, and A. Aleksenko.** 2001. A family of DNA repeats in *Aspergillus nidulans* has assimilated degenerated retrotransposons. Mol. Genet. Genomics **265:**883–887.

92. **Noma, K., C. D. Allis, and S. I. S. Grewal.** 2001. Transitions in distinct histone H3 methylation patterns at the heterochromatin domain boundaries. Science **293:**1150–1155.

93. **Nonaka, N., et al.** 2002. Recruitment of cohesin to heterochromatic regions by Swi6/HP1 in fission yeast. Nat. Cell Biol. **4:**89–93.

94. **Ohkuma, M., K. K. S. Kawai, C. W. Hwang, A. Ohta, and M. Takagi.** 1995. Identification of a centromeric activity in the autonomously replicating TRA region allows improvement of the host-vector system for *Candida maltosa*. Mol. Gen. Genet. **249:**447–455.

95. **Padmanabhan, S., J. Thakur, R. Siddharthan, and K. Sanyal.** 2008. Rapid evolution of Cse4p-rich centromeric DNA sequences in closely related pathogenic yeasts, *Candida albicans* and *Candida dubliniensis*. Proc. Natl. Acad. Sci. U. S. A. **105:**19797–19802.

96. **Palmer, D. K., K. O'Day, M. H. Wener, B. S. Andrews, and R. L. Margolis.** 1987. A 17-kD centromere protein (CENP-A) copurifies with nucleosome core particles and with histones. J. Cell Biol. **104:**805–815.

97. **Parker, J.** 2010. How to slice: snapshots of Argonaute in action. Silence **1:**3.

98. **Pearson, C. G., et al.** 2004. Stable kinetochore-microtubule attachment constrains centromere positioning in metaphase. Curr. Biol. **14:**1962–1967.

99. **Pidoux, A., and R. Allshire.** 2004. Kinetochore and heterochromatin domains of the fission yeast centromere. Chromosome Res. **12:**521–534.

100. **Pidoux, A. L., et al.** 2009. Fission yeast Scm3: a CENP-A receptor required for integrity of subkinetochore chromatin. Mol. Cell **33:**299–311.

101. **Polizzi, C., and L. Clarke.** 1991. The chromatin structure of centromeres from fission yeast: differentiation of the central core that correlates with function. J. Cell Biol. **112:**191–201.

102. **Ranjitkar, P., et al.** 2010. An E3 ubiquitin ligase prevents ectopic localization of the centromeric histone H3 variant via the centromere targeting domain. Mol. Cell **40:**455–464.

103. **Rea, S., et al.** 2000. Regulation of chromatin structure by site-specific histone H3 methyltransferases. Nature **406:**593–599.

104. **Rhind, N., et al.** 2011. Comparative functional genomics of the fission yeasts. Science **332:**930–936.

105. **Rivas, F. V., et al.** 2005. Purified Argonaute2 and an siRNA form recombinant human RISC. Nat. Struct. Mol. Biol. **12:**340–349.

106. **Roy, B., L. S. Burrack, M. A. Lone, J. Berman, and K. Sanyal.** 2011. CaMtw1, a member of the evolutionarily conserved Mis12 kinetochore protein family, is required for efficient inner kinetochore assembly in the pathogenic yeast *Candida albicans*. Mol. Microbiol. **80:**14–32.

107. **Roy, N., A. Poddar, A. Lohia, and P. Sinha.** 1997. The mcm17 mutation of yeast shows a size-dependent segregational defect of a mini-chromosome. Curr. Genet. **32:**182–189.

108. **Sadaie, M., T. Iida, T. Urano, and J.-i, Nakayama.** 2004. A chromodomain protein, Chp1, is required for the establishment of heterochromatin in fission yeast. EMBO J. **23:**3825–3835.

109. **Saffery, R., et al.** 2000. Human centromeres and neocentromeres show identical distribution patterns of >20 functionally important kinetochore-associated proteins. Hum. Mol. Genet. **9:**175–185.

110. **Sanyal, K., M. Baum, and J. Carbon.** 2004. Centromeric DNA sequences in the pathogenic yeast *Candida albicans* are all different and unique. Proc. Natl. Acad. Sci. U. S. A. **101:**11374–11379.

111. **Selker, E. U.** 1990. Premeiotic instability of repeated sequences in *Neurospora Crassa*. Annu. Rev. Genet. **24:**579–613.

112. **Selker, E. U., E. B. Cambareri, B. C. Jensen, and K. R. Haack.** 1987. Rearrangement of duplicated DNA in specialized cells of *Neurospora*. Cell **51:**741–752.

113. **Shankaranarayana, G. D., M. R. Motamedi, D. Moazed, and S. I. S. Grewal.** 2003. Sir2 regulates histone H3 lysine 9 methylation and heterochromatin assembly in fission yeast. Curr. Biol. **13:**1240–1246.

114. **Shanker, S., et al.** 2010. Continuous requirement for the Clr4 complex but not RNAi for centromeric heterochromatin assembly in fission yeast harboring a disrupted RITS complex. PLoS Genet. **6:**e1001174.

115. **Smith, K. M., P. A. Phatale, C. M. Sullivan, K. R. Pomraning, and M. Freitag.** 2011. Heterochromatin is required for normal distribution of *Neurospora crassa* CenH3. Mol. Cell. Biol. **31:**2528–2542.

116. **Song, J. S., X. Liu, X. S. Liu, and X. He.** 2008. A high-resolution map of nucleosome positioning on a fission yeast centromere. Genome Res. **18:**1064–1072.

117. **Stajich, J. E., et al.** 2010. Insights into evolution of multicellular fungi from the assembled chromosomes of the mushroom *Coprinopsis cinerea* (*Coprinus cinereus*). Proc. Natl. Acad. Sci. U. S. A. **107:**11889–11894.

118. **Steiner, N. C., and L. Clarke.** 1994. A novel epigenetic effect can alter centromere function in fission yeast. Cell **79:**865–874.

119. **Steiner, N. C., K. M. Hahnenberger, and L. Clarke.** 1993. Centromeres of the fission yeast *Schizosaccharomyces pombe* are highly variable genetic loci. Mol. Biol. Cell **13:**4578–4587.

120. **Stoler, S., et al.** 2007. Scm3, an essential *Saccharomyces cerevisiae* centromere protein required for G2/M progression and Cse4 localization. Proc. Natl. Acad. Sci. U. S. A. **104:**10571–10576.

121. **Sugiyama, T., et al.** 2007. SHREC, an effector complex for heterochromatic transcriptional silencing. Cell **128:**491–504.

122. **Sullivan, B., and G. Karpen.** 2001. Centromere identity in *Drosophila* is not determined in vivo by replication timing. J. Cell Biol. **154:**683–690.

123. **Takahashi, K., E. S. Chen, and M. Yanagida.** 2000. Requirement of Mis6 centromere connector for localizing a CENP-A-like protein in fission yeast. Science **288:**2215–2219.

124. **Takahashi, K., et al.** 1992. A low copy number central sequence with strict symmetry and unusual chromatin structure in fission yeast centromere. Mol. Biol. Cell **3:**819–835.

125. **Takayama, Y., et al.** 2008. Biphasic incorporation of centromeric histone CENP-A in fission yeast. Mol. Biol. Cell **19:**682–690.

126. **Tanaka, T., M. P. Cosma, K. Wirth, and K. Nasmyth.** 1999. Identification of cohesin association sites at centromeres and along chromosome arms. Cell **98:**847–858.

127. **Thon, G., et al.** 2005. The Clr7 and Clr8 directionality factors and the Pcu4

cullin mediate heterochromatin formation in the fission yeast *Schizosaccharomyces pombe*. Genetics **171:**1583–1595.

128. **Tomonaga, T., et al.** 2003. Overexpression and mistargeting of centromere protein-A in human primary colorectal cancer. Cancer Res. **63:**3511–3516.

129. **Van Hooser, A. A., et al.** 2001. Specification of kinetochore-forming chromatin by the histone H3 variant CENP-A. J. Cell Sci. **114:**3529–3542.

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

131. **Vernis, L., et al.** 2001. Only centromeres can supply the partition system required for ARS function in the yeast *Yarrowia lipolytica*. J. Mol. Biol. **305:**203–217.

132. **Villasante, A., J. P. Abad, and M. Méndez-Lago.** 2007. Centromeres were derived from telomeres during the evolution of the eukaryotic chromosome. Proc. Natl. Acad. Sci. U. S. A. **104:**10542–10547.

133. **Volpe, T., et al.** 2003. RNA interference is required for normal centromere function in fission yeast. Chromosome Res. **11:**137–146.

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

135. **Westermann, S., et al.** 2003. Architecture of the budding yeast kinetochore reveals a conserved molecular core. J. Cell Biol. **163:**215–222.

136. **Willard, H. F.** 1990. Centromeres of mammalian chromosomes. Trends Genet. **6:**410–415.

137. **Williams, J. S., T. Hayashi, M. Yanagida, and P. Russell.** 2009. Fission yeast Scm3 mediates stable assembly of Cnp1/CENP-A into centromeric chromatin. Mol. Cell **33:**287–298.

138. **Wood, V., et al.** 2002. The genome sequence of *Schizosaccharomyces pombe*. Nature **415:**871–880.

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

140. **Zhou, Z., et al.** 2011. Structural basis for recognition of centromere histone variant CenH3 by the chaperone Scm3. Nature **472:**234–237.

141. **Zofall, M., and S. I. S. Grewal.** 2006. RNAi-mediated heterochromatin assembly in fission yeast. Cold Spring Harb. Symp. Quant. Biol. **71:**487–496.