Remarkable Evolutionary Plasticity of Centromeric Chromatin

STEVEN HENIKOFF,1,2 JITENDRA THAKUR,1,2 SIVAKANTHAN KASINATHAN,2,3
AND PAUL B. TALBERT1,2
1Howard Hughes Medical Institute, Fred Hutchinson Cancer Research Center, Seattle, Washington 98109
2Basic Sciences Division, Fred Hutchinson Cancer Research Center, Seattle, Washington 98109
3Medical Scientist Training Program, University of Washington School of Medicine, Seattle, Washington 98195
Correspondence: steveh@fhcrc.org

Centromeres were familiar to cell biologists in the late 19th century, but for most eukaryotes the basis for centromere specification has remained enigmatic. Much attention has been focused on the cenH3 (CENP-A) histone variant, which forms the foundation of the centromere. To investigate the DNA sequence requirements for centromere specification, we applied a variety of epigenomic approaches, which have revealed surprising diversity in centromeric chromatin properties. Whereas each point centromere of budding yeast is occupied by a single precisely positioned tetrameric nucleosome with one cenH3 molecule, the “regional” centromeres of fission yeast contain unphased presumably octameric nucleosomes with two cenH3s. In Caenorhabditis elegans, kinetochores assemble all along the chromosome at sites of cenH3 nucleosomes that resemble budding yeast point centromeres, whereas holocentric insects lack cenH3 entirely. The “satellite” centromeres of most animals and plants consist of cenH3-containing particles that are precisely positioned over homogeneous tandem repeats, but in humans, different α-satellite subfamilies are occupied by CENP-A nucleosomes with very different conformations. We suggest that this extraordinary evolutionary diversity of centromeric chromatin architectures can be understood in terms of the simplicity of the task of equal chromosome segregation that is continually subverted by selfish DNA sequences.

Centromeres have been familiar to cell biologists even before the rediscovery of Mendel’s laws. In 1882, Walther Flemming used cytological stains to observe mitotic figures, which he called “chromatin” (Flemming 1882). In his drawings of lily chromosomes at metaphase and anaphase, Flemming darkly shaded the chromatin corresponding to each site of spindle fiber attachment, which later was referred to as the centromere or kinetochore. By current usage, the centromere refers to the genetic locus, whereas the kinetochore is the large complex of proteins specifically bound to centromeric DNA (Talbert et al. 2008). A further distinction can be made between the inner kinetochore proteins that remain associated with the DNA throughout the cell cycle, referred to as the centromeric chromatin-associated network (CCAN), and protein complexes of the outer kinetochore that assemble at mitosis.

The classical work of Morgan and his students genetically defined the centromere as a locus (Sturtevant 1913). But the molecular study of the centromere began in earnest in 1980, when Clarke and Carbon genetically mapped the centromere of Saccharomyces cerevisiae (budding yeast) Chromosome III to a 1.6-kb DNA segment (Clarke and Carbon 1980). When they inserted this DNA segment into an artificial plasmid vector with a replication origin, they observed faithful segregation through mitosis and meiosis.

The molecular description of proteins of the CCAN began with the discovery by Earnshaw and Rothfield of what they termed Centromere Proteins A, B, and C (CENP-A, CENP-B, and CENP-C) using human autoimmune antibodies (Earnshaw and Rothfield 1985). Meanwhile, Palmer and Margolis (1985) showed that bovine CENP-A is a histone H3 variant (Palmer et al. 1987), one that remains with the DNA in mature sperm (Palmer et al. 1990). This was the first suggestion that centromeres might be “epigenetic”—that is, defined by the presence of the centromere-specific histone variant, which we refer to generically as “cenH3” to conform to the well-established histone nomenclature (Bradbury 1977; Talbert et al. 2012). Understanding the basis for this dual definition of centromeres as genetic and/or epigenetic has been a driving force for work in the laboratory for almost two decades.

Centromeres are generally conserved in position along the chromosome over evolutionary time; however, surprisingly this is not the case for centromeric DNA sequence. The realization that centromeric DNA has been rapidly evolving came from studies in the 1960s of satellite DNA, tandem repeats that could be separated from main-band DNA in buoyant density gradients and that differed between even closely related species (Yunis and Yasmineh 1971). The discovery that the sequence of the S. cerevisiae cenH3 (Cse4) is as divergent from H3 as it is from human CENP-A (Meluh et al. 1998) was another clue suggesting that centromeres are not under the same evolutionary constraints as are genes and their DNA-bind-
ing proteins. These considerations led us to identify and characterize cenH3s from diverse model organisms, including *Caenorhabditis elegans* (Buchwitz et al. 1999), *Drosophila melanogaster* (Henikoff et al. 2000), *Arabidopsis thaliana* (Talbert et al. 2002), rice (Nagaki et al. 2004), and maize (Jin et al. 2004). Indeed, we found that evolutionary divergence is a general feature of cenH3s, despite the fact that H3 and its general replacement variant, H3.3, are among the most highly conserved proteins known. We referred to the remarkable divergence of both centromeric satellite DNA and its dedicated histone as the centromere paradox: rapid evolution with a conserved function that is essential through every cell division (Henikoff et al. 2001).

Harmit Malik in the laboratory made the key observation that *Drosophila* cenH3 (Cid) is evolving adaptively using standard population genetic measures, suggesting an arms race (Malik and Henikoff 2001). Reasoning that the arms race would be between centromeric satellites and cenH3, we posited that centromeres compete for inclusion into the egg at female meiosis I (Henikoff et al. 2001). Our “centromere drive” hypothesis combined the concept of female meiotic drive described by Rhoades for the selfish segregation behavior of distal heterochromatic knobs in maize (Rhoades 1952) with that of cenH3 “strength” for the inferred female meiotic orientation of heterochromatin-rich centromeres by Novitski (1955). Centromere drive as an explanation for the centromere paradox has since become the standard explanation for centromere divergence based on evidence from humans (Daniel 2002), monkey flowers (Fishman and Willis 2005), and mice (Chmátal et al. 2014) and an attractive general mechanism for postzygotic reproductive isolation as species diverge (Henikoff and Malik 2002; Burt and Travers 2006).

The realizations that satellite centromeres are selfish and that cenH3s and other CCAN proteins may act to suppress drive (Talbert et al. 2004) raise the question as to how these CCAN proteins interact with their DNA substrates. We started by studying *S. cerevisiae* centromeres, with simple point centromeres that provide a basis for comparison to the more complex regional centromeres of *Schizosaccharomyces pombe*, the holocentromeres of *C. elegans* and holocentric insects, and the satellite centromeres of humans. In each case we have applied high-resolution mapping technologies to address the basic problem of centromere specificity. We find that the evolutionary divergence of centromeric chromatin is reflected in the plasticity of CCAN complex composition between different organisms and sometimes of different centromeres of the same individual.

**BUDDING YEAST CENTROMERES ARE OCCUPIED BY HEMISOMES**

Based on a comparison of centromeres from different budding yeast chromosomes, Clarke and Carbon (1985) identified consensus motifs that they termed centromere-determining element (CDE) I, II, and III. Subsequent work identified the ~120-bp segment spanning CDEI–II–III as the sequence-specific core of each of the 16 *S. cerevisiae* centromeres. The 8-bp CDEI consensus is the binding site for the Cbf1 sequence-specific transcription factor, and the 26-bp CDEIII consensus is the binding site for the centromere-specific multisubunit CBF3 complex. These elements flank the 82 ± 4-bp CDEII, which lacks a specific consensus, but consists of ≥90% AT-rich DNA. Micrococcal nuclease (MNase) mapped a single Cse4 nucleosome to the functional centromere (Furuyama and Biggins 2007). This observation is difficult to reconcile with the availability of only ~80 bp of DNA for wrapping, suggesting the presence of a particle with fewer than the eight subunits in an H3 nucleosome, which wraps 147 bp of DNA.

To address this conundrum, Takehito Furuyama determined the chirality of the DNA superhelical writhe around the Cse4-containing core, using superhelical density mapping of circular chromosomes. This analysis revealed that the budding yeast centromere induces positive DNA supercoils, indicative of a right-handed writhe around the particle, opposite to that of left-handed H3 nucleosomes (Furuyama and Henikoff 2009). Our supercoiling measurements were later confirmed and extended in a study suggesting that the right-handed configuration is enforced by the formation of a DNA loop held together by the CBF3 complex, excluding CDEI (Díaz-Ingelmo et al. 2015). A possible rationale for the right-handed superhelical wrap is that DNA overwinds when stretched (Gore et al. 2006), so that pulling on the centromere at anaphase would be expected to cause the (right-handed) DNA to tighten around the particle.

Using a new high-resolution native chromatin immunoprecipitation (ChIP) method with V-plot analysis (Henikoff et al. 2011), Krassovsky et al. (2012) succeeded in mapping the Cse4-containing particle directly over CDEII. Meanwhile, some cytological studies detected more fluorescent Cse4-GFP particles over centromeres than could plausibly fit within CDEII (Coffman et al. 2011; Lawrimore et al. 2011), challenging the “point” interpretation of yeast centromeres (Furuyama and Biggins 2007). However, our quantitative ChIP-seq data could exclude the possibility that additional molecules are incorporated outside of the genetically defined centromere (Henikoff and Henikoff 2012), rather suggesting that unincorporated Cse4 molecules are closely associated with functional centromeres. Thus, only two plausible models for the budding yeast centromere remained: a (Cse4/H4) tetrasome or a (Cse4/H4/H2A/H2B) “hemisome,” either of which could wrap the ~80-bp CDEII DNA with right-handed chirality.

To determine whether hemisomes can be assembled on short segments of DNA, Takehito Furuyama adopted a classical salt-assembly method (Tatchell and Van Holde 1979), demonstrating that Cse4 but not H3 hemisomes that are stable in 4M urea can be readily produced on CDEII DNA (Furuyama et al. 2013). Our finding suggested that the extreme AT richness of CDEII is an adaptation for exclusion of conventional nucleosomes from the functional centromere, as poly(dA:dT) tracts in the yeast genome are known to exclude nucleosomes (Struhl and Segal 2013). To definitively determine the composition and conformation of the budding yeast centromeric nucleosome in vivo, we applied the Widom chemical cleavage mapping
method, in which histone H4 is derivatized such that it cleaves DNA next to the dyad (Brogaard et al. 2012). By mapping the sites of cleavage, we could determine the precise location of H4 with base-pair resolution and also determine whether there is one H4 or two H4s. Results were unequivocal: Mapping of H4 cleavages at all 16 yeast centromeres showed only a single H4 with cleavage sites within CDEII consistent with a hemisome and only with a hemisome among all proposed structures (Henikoff et al. 2014) (Fig. 1, middle panel). More recently, we have used our novel CUT&RUN chromatin profiling method (Skene and Henikoff 2017) to show the presence of histone H2A in the CDEII particle (Fig. 2), definitively excluding the (Cse4/H4)$_2$ tetrasome model (Mizuguchi et al. 2007; Wisniewski et al. 2014).

Fission yeast centromeres lack positioning and are populated by octameric nucleosomes

Examination of the fungal phylogeny argues that point centromeres, which are exclusive to the Saccharomycetes, have evolved from “regional” centromeres that lack obvious sequence specificity (Malik and Henikoff 2009). For example, centromeres of the parasitic budding yeast Candida albicans show no sequence similarity to those of its close relative Candida dubliniensis (Padmanabhan et al. 2008), and deletion or replacement results in efficient formation of cenH3-enriched neocentromeres nearby (Baum et al. 2006; Ketel et al. 2009; Thakur and Sanyal 2013). Centromeres of the fission yeast S. pombe consist

![Figure 1](https://example.com/figure1.png)

Figure 1. Budding yeast centromeric nucleosomes have one histone H4, but those in fission yeast have two. In vivo chemical cleavage mapping shows one CENP-A H4 molecule for Saccharomyces cerevisiae but two for Schizosaccharomyces pombe. (A–C) (Top) A structure-based model of where H4S47C-anchored cleavages are predicted to occur on the Watson (W, 5′→3′) and Crick (C, 3′→5′) strands around the dyad axis of an H3 nucleosome (Henikoff et al. 2014). Each box represents a single nucleotide position, where the filled boxes are predicted cleavage sites on the W strand (green) and the C strand (pink). The predicted distances between cleavage sites are indicated below. (Middle) Distribution of distances between cleavage fragment ends genome-wide (blue lines) and for budding yeast centromeres (red lines) (Henikoff et al. 2014). (Bottom) Comparable distributions for the central core of fission yeast centromeres (red lines) and euchromatin (blue lines) (Thakur et al. 2015). Distributions were normalized by the total number of combinations within the 40-bp distance range. (A) W–C distances. (B) W–W′ distances. (C) C–C′ distances. (Adapted from Thakur et al. 2015, with permission from the Genetics Society of America.)
of a series of heterochromatic outer repeats flanking a cenH3-rich core that lacks obvious sequence-specific features. Thus it would appear that budding yeast point centromeres have emerged from a fungal ancestor that lacks sequence-specific centromeres, perhaps the result of colonization by the 2-µm parasitic element that segregates autonomously within the Saccharomycetes (Malik and Henikoff 2009).

Consistent with an independent origin of point centromeres from regional centromeres, we have found that the cenH3 nucleosomes at the centromeric core of *S. pombe* regional centromeres are completely different from those of *S. cerevisiae*. Fission yeast cenH3 nucleosomes show no detectable positioning within the central core, and H4 chemical cleavage mapping reveals that there are two H4s per particle (Fig. 1, bottom panel), indicative of octameric nucleosomes (Thakur et al. 2015). Thus, the extreme diversity of centromere architecture is reflected in the diversity of centromeric chromatin within the fungal lineage, where point centromeres are bound by well-phased hemisomes and regional centromeres are bound by unphased octasomes.

**C. elegans Holocentromeres Are Polycentric**

The most fundamental distinction between centromeres is the distinction between familiar monocentromeres and holocentromeres, in which microtubule attachments occur throughout the length of the chromosome. Our cloning and cytological characterization of cenH3 from the nematode worm *C. elegans* revealed that it occupies the leading edge of each sister chromosome as it segregates to the pole at mitosis, and absence of cenH3 within the bulk of chromatin implied that holocentromeres are discontinuous (Buchwitz et al. 1999). Standard ChIP-seq uncovered a large domain structure of low-density cenH3 (Gassmann et al. 2012), and Florian Steiner’s high-resolution native ChIP revealed the presence of approximately 100 well-positioned cenH3 loci dispersed throughout each chromosome (Steiner and Henikoff 2014). Remarkably, these centromeric loci resembled yeast point centromeres with particle sizes of ∼80 bp flanked by well-phased H3 nucleosomes (Fig. 3). Motif analysis showed that holocentromeres correspond to previously described GA-rich transcription factor hotspots (Gerstein et al. 2010), leading us to speculate that low-affinity binding of transcription factors at these sites during interphase prevents encroachment by flanking nucleosomes, thus maintaining holocentromere sites accessible at mitosis throughout development (Steiner and Henikoff 2014).

The mapping of dispersed point centromeres at transcription factor hotspots throughout *C. elegans* chromosomes addressed a long-standing question in chromosome biology by showing that holocentromeres are polycentric as opposed to being diffuse (Schrader 1935). More recently, native ChIP-seq of centromeres of the parasitic nematode *Ascaris suum* suggested that these centromeres are also polycentric, but they do not appear to be point centromeres in that cenH3-rich regions consist of high-density arrays of 1–15 kb, with particle sizes of ∼140 bp, consistent with conventional octasomes (Kang et al. 2016).

**Insect Holocentromeres Lack cenH3**

Holocentricity has evolved in multiple lineages of animals and plants (Melters et al. 2012). For example, holocentromeres have evolved independently in insects at least four times (Drinnenberg et al. 2014). When Ines Drinnenberg was a postdoctoral fellow jointly with the Malik laboratory, she discovered that cenH3 and other
proteins of the CCAN are absent from holocentric Lepidopteran lineages. Remarkably, loss of these CCAN proteins coincided with all four known transitions from monocentricity to holocentricity in insects (Fig. 4). In contrast, proteins of the outer kinetochore are nearly universally conserved, including in the kinetoplastid Trypanosoma, which also lacks canonical CCAN proteins (Akiyoshi and Gull 2014; Senaratne and Drinnenberg 2017). Identifying the chromatin counterparts of the CCAN in these lineages and determining whether insect holocentromeres are polycentric or diffuse are issues that remain to be addressed (Drinnenberg et al. 2016).

**PRECISE POSITIONING OF cenH3 NUCLEOSOMES AT SATELLITE CENTROMERES**

The identification of CENP-A nucleosomes as marking functional mammalian centromeres (Palmer and Margolis 1985) was an important advance in centromere biology. However, it was not until 1997 that the equivalent of the Clarke and Carbon construction of a functional centromere was accomplished for a satellite centromere (Harrington et al. 1997). Most human artificial centromeres require hundreds of kilobases of an ∼170-bp α-satellite repeat array organized as higher-order repeats (HORs), including ∼17-bp consensus binding sites for CENP-B protein (Hayden et al. 2013). Several studies have used ChIP-seq to show that CENP-A nucleosomes occupy α-satellite sequences, although the exact composition and conformation of the particles continue to be debated (Bui et al. 2012; Hasson et al. 2013; Lacoste et al. 2014; Athwal et al. 2015; Henikoff et al. 2015; Thakur and Henikoff 2016; Nechemia-Arbely et al. 2017). We have found that much of the disagreement stems from the difficulty of mapping ChIP-seq reads to tandemly repeated DNA sequences, which have proven to be intractable using standard tools for sequence assembly. To circumvent this problem, we clustered CENP-A ChIP-seq reads de novo, identifying two families of homogeneous dimeric repeats with CENP-B boxes that dominate human centromeres (Henikoff et al. 2015). This extended earlier work of Alexandrov et al. (2001), who originally identified these two “suprachromosomal” α-satellite families, SF1 and SF2, as homogeneous dimeric arrays present on 20 of the 24 different human centromeres. Indeed, we identified a preponderance of precisely positioned 100-bp MNase protected particles for these two dominant families, but we also identified larger particles for other HORs, such as the DXZ1 SF3 subfamily (Fig. 5), that are less homogeneous in CENP-A ChIP-seq data (Henikoff et al. 2015). Our findings suggested that particle conformation detected by native MNase ChIP-seq can differ greatly between different human α-satellite families, with precise positioning a characteristic of the most homogeneous repeats. Precise rotational positioning of cenH3 nucleosomes is a feature of homogeneous satellite centromeres in rice, as determined in a collaborative ChIP-seq study from the Jiming Jiang laboratory (Zhang et al. 2013).

**Figure 3.** Caenorhabditis elegans holocentromeres are polycentric with dispersed point centromeres at transcription factor hotspots. (A) cenH3 native MNase ChIP-seq peaks (n = 707) were called and aligned at their midpoints. The orange curve represents the average cenH3/input profile, and the blue curve represents the input profile for paired-end reads in the 141–500-bp range. (B) For comparison, the equivalent profile is shown for alignment of MNase ChIP-seq and input data from Saccharomyces cerevisiae. (C) Sequence logo representations of the consensus motif at cenH3 peaks (left) and transcription factor hotspots (right) called using MEME (Bailey and Elkan 1994). (Adapted from Steiner and Henikoff 2014.)
A COHERENT INNER KINETOCHORE COMPLEX OCCUPIES YOUNG HUMAN CENTROMERIC REPEATS

Whereas we and others had observed 100-bp α-satellite particles using native MNase-based ChiP-seq (Hasson et al. 2013; Henikoff et al. 2015), we also found that formaldehyde cross-linking resulted in protection of particles that were larger than nucleosome size (Thakur and Henikoff 2016). This observation suggested that MNase cleavage under the conditions we used for native ChiP-seq was disrupting particle integrity, but that cross-linking held together a particle consisting of additional components. In our ChiP-seq study of fission yeast centromeres, we had identified CENP-T as an integral component of CENP-A and CENP-C enriched chromatin (Thakur et al. 2015). However, we initially failed to identify human CENP-T by native MNase-based ChiP-seq, consistent with the prevailing view that in mammals CENP-T makes connections with H3 but not CENP-A nucleosomes (McKinley and Cheeseman 2016). As connections to the outer kinetochore are made independently by both CENP-C and CENP-T, the concept of separate anchors for the kinetochore on the DNA was a central issue in understanding centromere biology. The possibility of independent connections between centromeric DNA and the outer kinetochore seemed plausible, insofar as CENP-T and its three partner proteins in the CENP-TWSX complex should be sufficient to directly anchor the outer kinetochore: All four are histone-fold proteins that can be assembled in vitro into nucleosome-like particles that stably wrap DNA in a right-handed orientation. But we wondered: Could it be that the CENP-TWSX particle accounts for the difference in protection by CENP-A/CENP-C particles using native versus formaldehyde cross-linking MNase-ChIP?

To directly address this possibility, we reasoned that the conditions that are used for MNase ChIP-seq might leave behind more condensed particles. Differential salt solubility is a feature of conventional nucleosomes, which can be fractionated into classical “active” (low-NaCl), histone H1–rich (high-NaCl), and “nuclear matrix” (insoluble) components (Sanders 1978; Henikoff et al. 2009). Indeed, using cross-linking and light sonication after MNase digestion, we found that CENP-T containing α-satellite particles could be recovered in robust amounts, and CENP-T particles precisely co-mapped with CENP-A and CENP-C (Thakur and Henikoff 2016). It is possible that estimated particle size differences reported in studies using native MNase ChiP-seq (Hasson et al. 2013; Lacoste et al. 2014; Henikoff et al. 2015; Nachev-Arbel et al. 2017) can be explained by the lability of CENP-TWSX-containing particles when subjected to MNase treatment.

To confirm that a coherent CCAN particle containing CENP-A, CENP-C, and CENP-T occupies functional human centromeres, we performed tandem ChiP-seq on FLAG-CENP-A particles subjected to cross-linking and heavy digestion with MNase to produce single coherent particles. After first immunoprecipitating with an anti-FLAG antibody and eluting with FLAG peptide, we immunoprecipitated with CENP-A, CENP-B, CENP-C, and CENP-T antibodies, obtaining strong enrichment for each component (Fig. 6). Enrichment was seen for consensus sequences representing SF1, SF2, and SF3 subfamilies and for the Y centromere, which indicates that at all human centromeres a single CCAN complex contains four of these centromere-specific DNA-binding components.
Figure 5. Precise positioning of cenH3 nucleosomes at human α-satellite centromeres. Young α-satellite dimers precisely position ~100-bp CENP-A nucleosome particles. (A–C) Size distributions of fragments mapping to an SF1 consensus dimer (A) and an SF2 consensus dimer (B) and to the most proximal 6-kb region of DXZ1 (C), which belongs to SF3. Graphs on the right are expansions of graphs on the left (indicated by brackets). The y-axis scale is for input normalized counts, and the areas under the other curves were equalized to that for input. (Reprinted from Henikoff et al. 2015.)
These studies of centromeric chromatin in a wide variety of organisms have uncovered an astonishing diversity that is difficult to reconcile with the conserved function of centromeric chromatin, which is to anchor the outer kinetochore complex (Steiner and Henikoff 2015). What accounts for such remarkable centromeric chromatin plasticity?

Evolutionary plasticity of centromeric chromatin might arise from selfish processes. Centromere drive is a cyclical process, in which satellite sequences expand and compete for the egg pole at meiosis I but are suppressed by mutations in host CCAN and other proteins (Malik and Henikoff 2001). Successful drive of a satellite centromere may eventually result in fixation of that satellite sequence in the species evolving into a genetically defined centromere, whereas successful suppression of drive over evolutionary time tends to reduce centromere sequence specificity evolving into an epigenetically defined centromere (Dawe and Henikoff 2006). This telokinetic form of chromosome disjunction may be less effective at avoiding centromere drive than the holokinetic disjunction of equatorially oriented \textit{Luzula} chromosomes (Heckmann et al. 2014). Although the role of cenH3 in the cup-like kinetochore of \textit{Caenorhabditis} meiosis is disputed (Chan et al. 2004; Monen et al. 2005), it differs between oocyte and spermatocyte meiosis (Shakes et al. 2009), and cenH3 shows evidence of positive selection (Zedek and Bereš 2012), suggesting that centromere drive may affect telokinetic holocentromeres. This may explain the massive accumulation of heterochromatic sequences at the ends of the single chromosome in \textit{Parascaris univalens} (Talbert et al. 2008).

Whereas native human centromeres are genetically defined in that they are dominated by specific \(\alpha\)-satellite arrays on each centromere, rare cases of epigenetically defined human neocentromeres suggests that centromeric chromatin plasticity is an inevitable consequence of an ever-changing DNA sequence substrate for CCAN assembly. Other selfish processes may be responsible for transitions between different fungal centromere types, such as the proposed colonization of Saccharomycetes centromeres by its 2-\(\mu\)m plasmid (Malik and Henikoff 2009).
One common feature of genetically defined centromeres, whether the point centromeres of Saccharomyces or the satellite centromeres of animals and plants, is that they are well-positioned. In contrast, the central core of fission yeast regional centromeres lacks obvious sequence specificity and shows random positioning of nucleosomes (Thakur et al. 2015). Precise positioning of nucleosomes, which is especially evident in the satellite centromeres of rice (Zhang et al. 2013), is a sequence-specific adaptation of satellite DNA that has been exploited for the first successful crystallization of H3 nucleosomes, which were produced using a human α-satellite DNA derivative (Harp et al. 1996; Luger et al. 1997). Precise phasing, together with the action of satellite-specific DNA-binding proteins, such as CENP-B, may represent the means whereby selfish centromeres perpetuate themselves. CENP-B is derived from the transposase encoded by the pogo DNA transposon, and it is attractive to imagine that its original domestication as a centromere protein was via recruitment by a driving centromere.

The domestication of selfish elements appears to be responsible for the most remarkable example of centromere evolution, the partial or complete replacement of the satellite centromeres of the wild ancestor of maize by a selfish element in maize inbred lines (Schneider et al. 2016). During domestication of maize from teosinte over the past ~9000 years, selection for agronomic traits was accompanied by dense insertion of CR2 centromere-specific retrotransposons and frequent loss by small inversions and deletions of the ancestral 155-bp CentC satellite in all 10 maize centromeres, 57 times independently in 26 inbred lines. Domestication and inbreeding would thus appear to be a powerful driver of centromere plasticity and could account for the near-complete loss of satellite arrays at several rice (Nagaki et al. 2004) and potato (Gong et al. 2012) centromeres and the complete loss of repetitive sequences from horse Centromere 11 (Purgato et al. 2015). Thus, whether the direct result of centromere competition, the colonization by selfish elements, or the indirect result of domestication, strong selective forces appear to be drivers of the remarkable plasticity of centromeric chromatin.

**CONCLUSION**

Our investigations into the molecular basis for centromere function have uncovered an unexpected plasticity in chromatin architecture, with evidence for histone-containing particles that are unlike any nucleosomes observed on chromosome arms. These studies have led to evidence for centromeres with right-handed writhes, hemisomes, and coherent CCAN particles containing multiple centromere-specific histone-fold and nonhistone proteins. Such structural diversity might seem surprising considering the simplicity of the conserved function of centromeric chromatin to connect to the outer kinetochore. However, the centromere’s control of its chromosome’s fate at every cell division has unleashed powerful selfish and selective forces over evolutionary time and even during human domes-
tication that have resulted in such interesting molecular complexity.

ACKNOWLEDGMENTS

We thank Mitchell Smith, who approximately 20 years ago pointed out that yeast Cse4 and human CENP-A are no more similar to one another than to H3, a curious fact that led to our subsequent focus on histone variants and centromere evolution. We also thank former members of our laboratory whose work and insights over the years have contributed to this narrative.

REFERENCES


Thakur J, Talbert PB, Henikoff S. 2015. Interactions of inner
kinetochore proteins with fission yeast regional centromeres. 
*Genetics* **201**: 543–561.

Wisniewski J, Hajj B, Chen J, Mizuguchi G, Xiao H, Wei D, 
Dahan M, Wu C. 2014. Imaging the fate of histone Cse4 
reveals de novo replacement in S phase and subsequent stable 

Yunis JJ, Yasmineh WG. 1971. Heterochromatin, satellite DNA, 
and cell function. Structural DNA of eucaryotes may support 
and protect genes and aid in speciation. *Science* **174**: 1200–
1209.

Zedek F, Bureš P. 2012. Evidence for centromere drive in the 
e30496.

in Luzula suggests that holokinetic chromosomes may sup-

Zhang T, Talbert PB, Zhang W, Wu Y, Yang Z, Henikoff JG, 
Henikoff S, Jiang J. 2013. The CentO satellite confers 
translational and rotational phasing on cenH3 nucleosomes 
E4883.