


Centromere structure and function: lessons from *Drosophila*

Eftychia Kyriacou,¹ Patrick Heun ^{2,3,*}

¹Swiss Institute for Experimental Cancer Research (ISREC), School of Life Sciences, Ecole Polytechnique Fédérale de Lausanne (EPFL), 1015 Lausanne, Switzerland

²Wellcome Centre of Cell Biology, School of Biological Sciences, University of Edinburgh, EH9 3BF Edinburgh, UK

³Department of Biology, Technische Universität Darmstadt, 64287 Darmstadt, Germany

*Corresponding author: Department of Biology, Technical University Darmstadt, Schnittspahnstr. 10 (B2|03 Biologie), 64287 Darmstadt, Germany.

Email: patrick.heun@tu-darmstadt.de

The fruit fly *Drosophila melanogaster* serves as a powerful model organism for advancing our understanding of biological processes, not just by studying its similarities with other organisms including ourselves but also by investigating its differences to unravel the underlying strategies that evolved to achieve a common goal. This is particularly true for centromeres, specialized genomic regions present on all eukaryotic chromosomes that function as the platform for the assembly of kinetochores. These multiprotein structures play an essential role during cell division by connecting chromosomes to spindle microtubules in mitosis and meiosis to mediate accurate chromosome segregation. Here, we will take a historical perspective on the study of fly centromeres, aiming to highlight not only the important similarities but also the differences identified that contributed to advancing centromere biology. We will discuss the current knowledge on the sequence and chromatin organization of fly centromeres together with advances for identification of centromeric proteins. Then, we will describe both the factors and processes involved in centromere organization and how they work together to provide an epigenetic identity to the centromeric locus. Lastly, we will take an evolutionary point of view of centromeres and briefly discuss current views on centromere drive.

Keywords: centromeres; fruit fly; *Drosophila*; chromosomes; epigenetics; FlyBook

Introduction

Model organisms are used to study a particular biological question by taking advantage of a simpler and more tractable system. Nonetheless, in addition to learning about the similarities between species used as model organisms, important insights can be gained by investigating their differences. In this chapter, we are focusing on fly centromeres as a good example for a simple centromere that encompasses both of the above: employing a set of fly-specific components yet serving to ensure accurate chromosome segregation during cell division, a function universally shared among all eukaryotes.

One of the most fundamental processes during the cell cycle is the distribution of equal numbers of the replicated chromosomes into the two daughter cells, to ensure maintenance of the genetic material. Centromeres, loci with specialized chromatin, play a key role in this process. They serve as the platform for the assembly of the kinetochore, a complex proteinaceous structure that is required for centromeric chromatin to attach to the spindle apparatus. Errors in the process of chromosome segregation can lead to aneuploidy, a major cause of miscarriage and a hallmark of cancer. Therefore, to fully elucidate centromere function, it is imperative to understand in great detail their structure and organization.

Drosophila has proven to be a powerful model for centromere biology. Fly centromeres are less complex compared to their

human counterparts. They contain a much smaller number of constitutively associated centromere proteins, yet they perform the same critical function in chromosome segregation. Importantly, several notions and processes discovered for *Drosophila* centromeres were proven to be simplifications of what holds true in human centromeres and were useful to direct studies carried out in human cells, as we will discuss later.

Here, we are reviewing older and more recent work of various groups that used *Drosophila* as a model organism and focused on one common theme: understanding the centromere. We will start by describing *Drosophila* centromeres from the DNA sequence to the chromatin composition, how the centromere is assembled, and which proteins associate with it. We will discuss studies that investigated the epigenetic nature of centromeres as well as the evolutionary history of *Drosophila* centromeres and their components. All through, we will be comparing fly centromeres to other species, especially human, highlighting the significance of employing *Drosophila* as a model organism.

Centromeric DNA in *Drosophila*

As a functional element of the chromosome, elucidating the DNA composition of centromeres was an essential step to begin the quest to unveiling their function. In the early 70s, it was evident that *Drosophila* centromeric sequences, like in most eukaryotes,

are highly repetitive based on hybridization experiments on salivary gland squashes with in vitro transcribed repetitive sequences (Jones and Robertson 1970; Rae 1970; Botchan et al. 1971). Parallel studies in mouse and human cells revealed that sequences comprising centromeres in metazoans also contain repetitive sequences (Pardue and Gall 1970; Manuelidis 1978), a centromeric feature that is now accepted to be conserved among many species (Fig. 1). Importantly, the difficulties associated with assembling repetitive DNA into larger sequence reads represented up to recent years a major challenge for the molecular study of centromeres.

Centromeric sequences obtained by sequencing X-derived minichromosomes

While sequences enriched in human centromeres, namely alpha-satellites, were discovered years ago (Fig. 1b) (Manuelidis 1978) and found to be organized in arrays of higher-order repeats (HORs), studies in *Drosophila* to identify centromeric sequences lagged. However, the use of small, nonessential X-chromosome-derived mini chromosomes that could be fragmented without affecting the fly viability (Karpen and Spradling 1990) allowed for determining the minimal sequences that were necessary and sufficient for inheritance and proper centromere function. Detailed molecular analysis of the X-derived Dp1187 minichromosome centromere suggested that it is primarily composed of the simple repeats AATAT and TTCTC (or AAGAG), interrupted by short islands of complex DNA, named after the Polynesian islands *Maupiti*, *Bora Bora*, *Tahiti*, and *Moorea* and transposable elements (retrotransposons H.M.S. Beagle, 412, BEL, F Doc, and G-like) (Sun et al. 1997, 2003; Murphy and Karpen 1995; Le et al. 1995). However, these complex X-centromere sequences were not found on all the other autosomal *Drosophila* centromeres (Le et al. 1995; Sun et al. 1997, 2003), suggesting that unlike in humans, centromeres of different chromosomes are composed of different genetic motifs (Fig. 1c).

Sequences at native centromeres

Even though minichromosomes provided valuable insight into centromeric DNA sequence structure, the specific composition of centromeric sequences on each chromosome remained elusive. For clarity, in this review, we define the centromere as the underlying DNA sequence spanning from the first to the last genomic region that is associated with the centromere-specific histone H3-variant CENP-A of a single chromosome. The centromere is flanked at either side by large arrays of simple satellite repeats that are organized into pericentric heterochromatin (Fig. 1a). Centromeres of the four different chromosomes in *Drosophila* were studied and compared to understand their evolution and conservation. Mapping of the Y centromere revealed that it has evolved from a telomere, suggesting that an ancestral telocentric chromosome has acquired a pericentric insertion (Méndez-Lago et al. 2009). In further support of this, the 18HT satellite present on this chromosome is a tandem array of sequences related to the telomeric retrotransposon-based HeT-A and TART sequences, together with another ancient telomere-related set of complex repeats (Agudo et al. 1999; DeBaryshe and Pardue 2011).

Tandem GC-rich arrays called dodeca satellite sequences were shown to be present on the centromere of the third chromosome (Carmena et al. 1993; Ferrer et al. 1995; Ortiz-Lombardía et al. 1998). A 10-mer AATAACATAG satellite called Prodsat is found upstream of the dodeca satellite (Garavís et al. 2015) and is bound by the proliferation disrupter protein (Prod) during mitosis, which is thought to play a role in clustering centromeres into chromocenters (Platero et al. 1998; Török et al. 2000; Jagannathan et al. 2019). However, the

above studies did not examine specifically which sequences are bound by centromere-specific proteins, like dCENP-A. Thus, using microscopy and native dCENP-A chromatin immunoprecipitation (ChIP), Talbert et al. identified sequences specifically bound by dCENP-A, with Prodsat as the most enriched sequence, followed by AATAG, as well as the AATAT sequences. The previously centromere-enriched dodeca satellite and the TTCTC repeats were depleted from dCENP-A precipitates (Talbert et al. 2018). Despite this elaborate work, still centromeres were represented as major gaps in genomic assemblies of *Drosophila* chromosomes, not only because of their repetitive nature, like in other organisms, but also because of the lack of information regarding the sequences present on centromeres of different chromosomes.

Major advances in sequencing technologies allowing the production of long sequencing reads enabled the sequencing of all *Drosophila* centromeres in a breakthrough study in 2019 (Chang et al. 2019) (Fig. 1c). By assembling long reads of sequences precipitated with dCENP-A, Chang et al. found that dCENP-A is assembled over islands of complex DNA sequences, enriched for retrotransposable elements, and flanked by large blocks of simple satellite repeats. The centromeres range between 101 and 171 kb, and while the overall sequence organization of all centromeres is similar, each contains unique islands of complex DNA. Following the nomenclature of Polynesian islands previously exploited for the islands of retrotransposable elements in *Drosophila*, like *Maupiti* for the island present on the X chromosome (Le et al. 1995; Sun et al. 1997), the newly described islands were named after the Italian islands *Lipari*, *Capri*, *Giglio*, and *Lampedusa* for chromosomes Y, 2, 3, and 4, respectively (Fig. 1c). Interestingly, the non-LTR retroelement *G2/Jockey-3* is a sequence enriched in all centromeres, regardless of its presence in other genomic locations, and associates with dCENP-A. It is noteworthy that even though DNA sequence and organization per se differ between different organisms, it seems that similar to humans, only a portion of the underlying centromeric sequence is occupied by dCENP-A (Sullivan et al. 2011; Chang et al. 2019).

In addition to its specialized repetitive composition, centromeric DNA was proposed to be fit for engaging in particular conformations. The TTCTC (or AAGAG) satellite can form a parallel pyrimidine triple helix under nearly physiological pH conditions, and this conformation is promoted and stabilized by copper ions (Horn et al. 2004; Paris et al. 2007). Whether this structure is present in vivo remains to be determined. However, an increase in the number of satellites (as it is the case in vivo) was shown to highly increase the stability of these helices (Horn et al. 2004), implying that these structures could be formed and found in vivo (Paris et al. 2007). The dodeca satellite can form 4-stranded intercalated structures, called i-motifs, through the C-rich strand. These are composed by 2 parallel duplexes organized in an anti-parallel fashion (Garavís et al. 2015). Based on the above data and given that centromeric sequences on different chromosomes are not conserved, it is tempting to speculate that the structure of centromeric DNA might contribute to mark a locus for attracting centromeric proteins and be established as the centromere. Advancement of new tools and technologies taking advantage of long-read DNA sequencing will allow for this to be determined in future studies and clarify whether such structures are indeed present in vivo, in endogenous centromeres.

Centromeric proteins in *Drosophila*

Through the fulfillment of its function as the platform for the assembly of the kinetochore, the centromeric proteome is highly

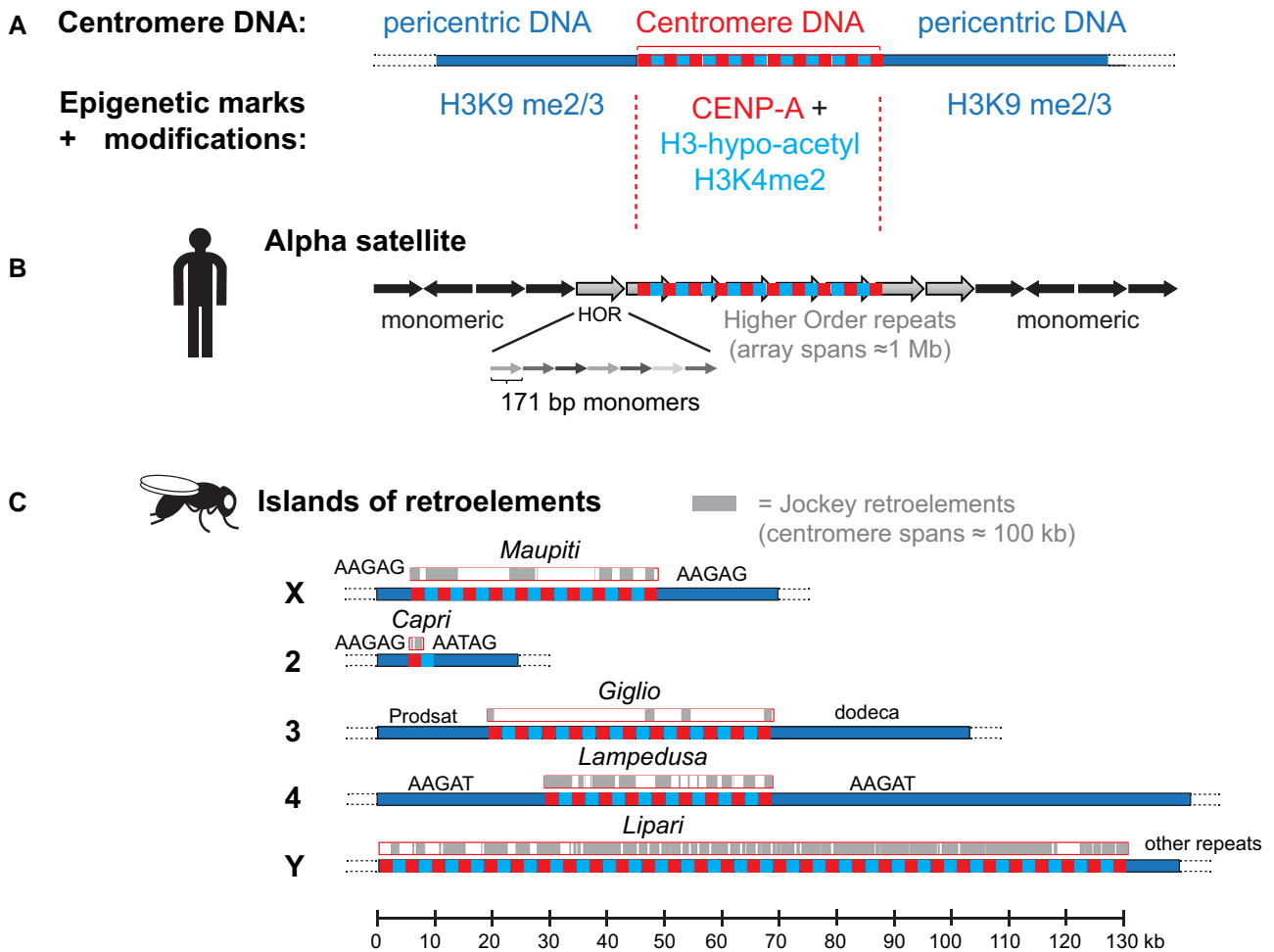


Fig. 1. Centromeric DNA in humans and *Drosophila*. a) Domain organization of the centromere DNA and association with epigenetic marks and histone modifications. Note that the functional centromere is defined from the first to the last genomic region occupied by CENP-A (red) and that CENP-A and Histone H3 (light blue) regions are interspersed at the centromere. b) Cartoon of a generic human centromere containing alpha satellite sequences. c) Cartoon of the different fly centromeres containing islands of retroelements interspersed and flanked by simple repeats. Many different types of retroelements are found, but for simplicity, only the Jockey retroelement is shown, which includes the G2/Jockey-3 elements, shared among all centromeres and is the highest enriched sequence for dCENP-A association (Chang et al. 2019).

specialized. Starting from the proteins that bind centromeric DNA all the way to the outer kinetochore components that mediate interactions with microtubules, in this section, we will discuss the intriguing complex protein network surrounding *Drosophila* centromeres.

While a centromere-specific H3 histone was identified as marking the centromeres already at the end of the 80s in humans (CENP-A) and budding yeast (Cse4p) (Earnshaw and Rothfield 1985; Palmer et al. 1987, 1991; Stoler et al. 1995; Meluh et al. 1998), its homolog in *Drosophila* (hereafter referred to as dCENP-A) remained unknown for another decade. In 2000, Henikoff et al. (2000) described an H3-like protein that localizes specifically to fly centromeres, called Cid, for centromere identifier. Because human and yeast CENP-A have a histone core divergent from canonical H3 as well as a very dissimilar N-terminal tail, they correctly predicted that a *Drosophila* ORF with these characteristics may encode a centromere-specific histone in flies. Indeed, by raising an antibody against the predicted peptide of this ORF, a staining on the primary constriction site of mitotic chromosomes was observed, colocalizing with Polo kinase, a known kinetochore-related protein (Henikoff et al. 2000). Subsequent functional studies on dCENP-A confirmed its critical

role in centromere and kinetochore function. dCENP-A was shown to be associated with centromeric DNA in or near the inner kinetochore plate and is spatially separated both from outer kinetochore components that are closer to spindle microtubules and pericentric heterochromatin (Blower and Karpen 2001). Interestingly, dCENP-A localization at neocentromeres suggested that its presence correlates with centromere function and activity, independently of the underlying DNA sequence (Blower and Karpen 2001). Indeed, dCENP-A was shown to be required for the recruitment of all outer kinetochore proteins tested and cell-cycle progression.

dCENP-A is profoundly different from canonical H3, and significant identity is only found in the amino acid sequence of the histone fold domain (HFD) (Ahmad and Henikoff 2002). By taking advantage of the adaptive evolution of CENP-A in different *Drosophila* species and the fact that dCENP-A from *Drosophila bipunctinata* does not localize to *melanogaster* centromeres, domain swap experiments have identified the loop 1 (L1), within the HFD of dCENP-A, to be a necessary and sufficient element for centromere targeting (Vermaak et al. 2002). Several amino acids located at both ends of the L1 loop as well as its minimal length are required for providing centromere targeting (Vermaak et al. 2002). As we

will discuss later, these evolutionary features of the CENP-A HFD in flies were shown to critically affect the potential deposition of CENP-A into centromeric chromatin, by altering its interaction with its specific loading machinery (Rosin and Mellone 2016).

The dCENP-A-containing nucleosome

Since centromeric DNA composition did not seem to be the determinant for centromere identity, but rather CENP-A was a common feature between centromeres, the field has seen extensive efforts for the dCENP-A-containing nucleosome structural characterization. Was the presence of a H3 variant in centromeric nucleosomes enough to provide it with specialized properties or was it something else? Studies of the structure of the dCENP-A-containing nucleosomes have yielded contradictory results. Biochemical studies in combination with electron microscopy (EM) and atomic force microscopy (AFM) suggested that dCENP-A nucleosomes have only half the expected height. Based on this data, it was proposed that dCENP-A-containing nucleosomes are heterotypic, containing only one copy of each of the histones dCENP-A, H4, H2A, and H2B, forming a tetramer or “hemisome,” instead of a canonical octameric nucleosome (Dalal et al. 2007; Wang et al. 2008). Additional experiments led to the proposal that centromeric nucleosomes not only are heterotypic but also wrap DNA right-handedly, inducing positive supercoiling, in contrast to canonical H3-containing nucleosomes (Furuyama and Henikoff 2009). Contrary to these findings, *in vivo* dCENP-A-containing nucleosomes were shown to contain dCENP-A dimers, suggesting that centromeric nucleosomes are octameric, while this dimerization of dCENP-A was proposed to be essential for correct centromere assembly, further opposing to the suggestion that centromeric nucleosomes were of a “hemisome” nature (Zhang et al. 2012). In particular, salt extraction and sucrose gradient ultracentrifugation of centromeric chromatin suggested that centromeric nucleosomes contain two dCENP-A molecules *in vivo*, consistent with an octameric structure. Importantly, a conserved residue within the 4-helix bundle of dCENP-A was found to play an important role for this dimerization and promoting integrity of dCENP-A-containing nucleosomes (Zhang et al. 2012). Copurification of endogenous and tagged dCENP-A from cells retrieves equal amounts of each, suggesting that more than one copy of dCENP-A is present per nucleosome (Erhardt et al. 2008), while extended micrococcal nuclease digestions of dCENP-A nucleosomes assembled *in vitro* by the dCENP-A-specific chaperone suggested that ~120 bp of DNA is protected, instead of ~65 bp expected for a half octamer (Chen et al. 2014). Thus, nucleosomes containing CENP-A are likely octameric independent of the species they originate from but with specialized physical properties for height and rigidity that distinguish them from nucleosomes containing canonical H3 (Sekulic et al. 2010; Miell et al. 2013; Ali-Ahmad and Sekulic 2020).

Centromeric chromatin organization

The CENP-A-containing nucleosome is in fact the major building block of centromeric chromatin. However, using extended chromatin fibers, it was discovered that centromeric chromatin is not solely composed of dCENP-A nucleosomes—as expected from cytological studies of lower resolution (Ahmad and Henikoff 2001)—but rather of interspersed blocks of dCENP-A and H3 nucleosomes, spanning 200–500 kb, with each block comprising ~14–50 kb (Blower et al. 2002). This organization is likely retained throughout the cell cycle, and even in mitosis, as mitosis-specific phosphorylated H3 is also interspersed with

dCENP-A and found to be conserved between humans and flies (Blower et al. 2002).

Initial reports have demonstrated that like in mammals, fly centromeric chromatin is also embedded in large arrays of heterochromatin (Bonaccorsi and Lohe 1991; Lohe et al. 1993; Murphy and Karpen 1995). But even though it is embedded in heterochromatin, which is characterized by the presence of H3K9 dimethylation, this mark is absent from centromeric H3 nucleosomes (Blower et al. 2002; Sullivan and Karpen 2004). The boundaries between centromeric and pericentric heterochromatin seem to be sharp with no visible presence of H3K9 dimethylation in the chromatin regions containing dCENP-A (Sullivan and Karpen 2004; Olszak et al. 2011; Blower et al. 2002; Kharchenko et al. 2011; ModENCODE et al. 2011). Chromatin at centromeres was also shown to be generally hypoacetylated (Turner et al. 1992; Sullivan and Karpen 2004), a feature reminiscent of heterochromatin. However, centromeric H3 nucleosomes are enriched for H3K4 dimethylation—but not H3K4 trimethylation—a mark known to be associated with “open” but not necessarily active chromatin states (Sullivan and Karpen 2004) (Fig. 1a). Thus, histones at the centromere are decorated by both heterochromatic and euchromatic marks, maintaining a unique pattern referred to as “cen-chromatin,” a property also found to be conserved in human centromeres.

How the centromeric chromatin posttranslational modification profile is established, how the boundaries between the 2 types of chromatin are maintained, and their physiological relevance are not clear and remain important questions to be addressed in the future (Fig. 1).

Factors associating with the dCENP-A nucleosome

As the basis for the complex kinetochore structure, centromeric chromatin interacts with multiple proteins. Several studies attempted to identify fly-specific centromere-associated factors, to extend studies carried out in human cells (Foltz et al. 2006). Genetic screens identified interactors of dCENP-A such as the condensin subunit Cap-G (Jäger et al. 2005) or the long sought-after homolog of CENP-C, a constitutive centromere protein, colocalizing with dCENP-A during interphase and mitosis (Heeger et al. 2005). dCENP-C’s localization to the centromere is mediated by its C-terminal domain, specifically through a region encompassing amino acids 1009–1205 that is sufficient to target the centromere (Heeger et al. 2005). Contrary to the fly homolog, mammalian CENP-C has been shown to bind CENP-A nucleosomes through its central domain (amino acids 426–537) (Carroll et al. 2010; Kato et al. 2013) and its CENP-C motif (727–767) (Kato et al. 2013; Ali-Ahmad et al. 2019). dCENP-C shows moderate to low sequence conservation with the central domain (19% identity), but its CENP-C motif is highly similar to other species (amino acids 1086–1117, 29% identity with human) suggesting a similar ability to bind CENP-A nucleosomes directly. Despite some differences in its C-terminal domain, dCENP-C is highly conserved between Drosophilids, particularly at its N-terminus (Heeger et al. 2005). Interestingly, similar to its budding yeast homolog Mif2 but unlike human, dCENP-C also contains an AT hook. In Mif2, the AT hook is part of the DNA- and histone-binding domain (DHBD) (Xiao et al. 2017) and contributes to centromere binding, but whether it performs a similar role in dCENP-C remains to be shown (Heeger et al. 2005).

Attempts to identify proteins possibly homologous to the Constitutive Centromere Associated Network (CCAN) (Foltz et al. 2006) (Fig. 2b; Table 1) by analyzing the interactome of the immunoprecipitated dCENP-A nucleosome have failed to pinpoint

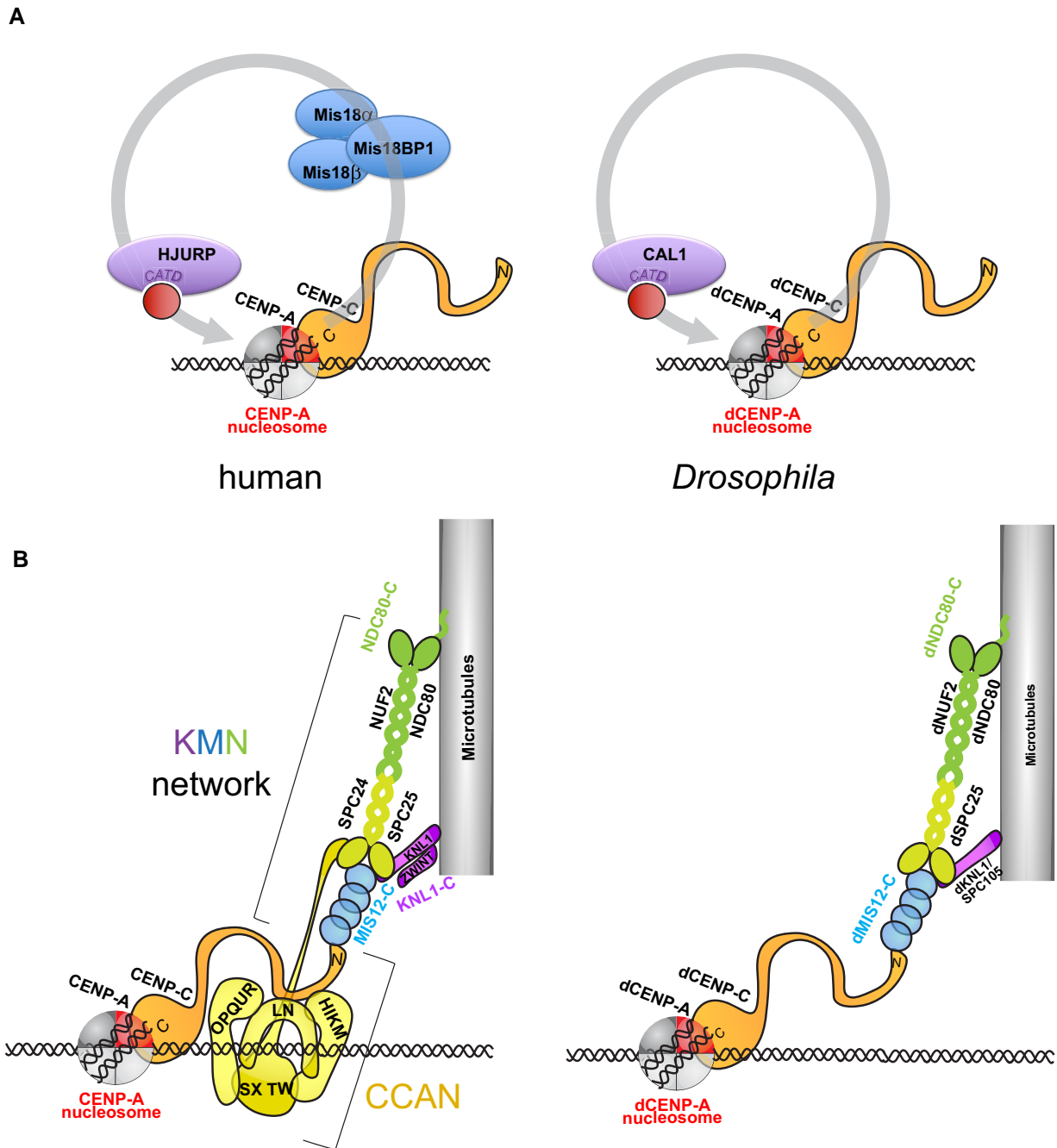


Fig. 2. CENP-A maintenance and kinetochore organization in human and *Drosophila*. a) Cartoons show the factors involved in the epigenetic inheritance of CENP-A. Human CENP-C is associated with a CENP-A nucleosome and recruits the MIS18 complex, which in turn targets the CENP-A chaperone holliday junction recognition protein (HJURP) to centromeres to load new CENP-A. dCENP-C directly recruits the dCENP-A chaperone CAL1 to load new dCENP-A to centromeric chromatin. b) Cartoons show the kinetochore organization at the centromere in mitosis with connection to spindle microtubules. CCAN is shown in orange (CENP-C) and yellow (HIKMLNOPQRSXTW); the KNL1-MIS12-NDC80 (KMN) network is shown in purple (KNL1-complex), blue (MIS12-complex), and light and dark green (NDC80-complex). Note that dCENP-C is the only CCAN member in *Drosophila*.

proteins exclusively and constitutively present at centromeres (Barth *et al.* 2015, 2014). Nonetheless, other factors enriched at centromeres were identified, including the histone acetyltransferase 1 (HAT1) homolog, the proteasome regulator REG, and the transcription elongation factor Spt6 (Barth *et al.* 2014; Bobkov *et al.* 2020). Curiously, although the possibility remains that functional homologs of the CCAN proteins are yet to be discovered, it appears that *Drosophila* has evolved a very reduced inner kinetochore, with dCENP-C as the only component fulfilling all functions

of the complex and multiprotein CCAN human counterpart (Fig. 2c; Table 1).

A breakthrough in our understanding of CCAN association with human centromeric nucleosomes has recently been achieved through CryoEM studies (Fig. 2b; Pesenti *et al.* 2022; Yatskevich *et al.* 2022). Interestingly, the structure of the CCAN bound to the CENP-A nucleosome by the Barford Lab reveals a tight association of CCAN proteins with linker DNA protruding from the CENP-A nucleosome. It is suggested that this type of anchoring might help to

Table 1. List of *Drosophila* centromere proteins including mention of components missing compared to human centromeres.

	Protein name	Family	Functions at centromeres	References
Centromeric chromatin	dCENP-A	Histone H3 variant	1) Replaces canonical H3 and wraps a portion of centromeric DNA to nucleosomes 2) Specifies centromeric identity	Earnshaw and Rothfield (1985) Henikoff et al. (2000)
CCAN (inner kinetochore)	dCENP-C		1) Binds the dCENP-A nucleosome and bridges centromeric chromatin to outer kinetochore proteins 2) Interacts and recruits Cal1 to centromeres for dCENP-A propagation (only CCAN member conserved in <i>D. melanogaster</i>)	Heeger et al. (2005)
		Members		
KMN network (outer kinetochore proteins)	MIND/dMIS12 complex dNDC80 complex SPC105/KNL-1	dMIS12, NSL1, NNF1a, NNF1b dSPC25 dNDC80 dNUF2	Binds dCENP-C, KNL-1 and dSPC25 (DSN1 not found) Interacts with dMIS12 (SPC24 not found) Interacts with dSPC25 and microtubules Interacts with dNDC80 and dSPC25 Anchors the dMIS12 complex to microtubules (ZWINT not found)	Przewlaka et al. (2007) Schittenhelm et al. (2007)
CENP-A assembly	CAL1	dCENP-A chaperone	Incorporates dCENP-A into centromeric nucleosomes (functional homolog of HJURP)	Goshima et al. (2007) Erhardt et al. (2008)

stabilize the centromere structure on chromatin, when exposed to the forces exerted by the mitotic spindle. Which proteins might play a similar function in *Drosophila* remains unclear at this point, but dCENP-C would be a potential candidate to fulfill this role.

Unlike the CCAN, outer kinetochores or KNL1-MIS18-NDC80 (KMN) network is more conserved in *Drosophila* (Fig. 2b) (Przewlaka et al. 2007; Schittenhelm et al. 2007; Liu et al. 2016). Numerous approaches have been taken to characterize them in molecular detail. Phylogenetic analyses of the eukaryotic kinetochores have identified potential orthologs of the outer kinetochore members Mis12, Ndc80, and Nuf2 (Meraldi et al. 2006), which were afterward confirmed to be proteins colocalizing with dCENP-A at centromeres (Przewlaka et al. 2007; Schittenhelm et al. 2007; Fig. 2b). A proteomic-based approach using tagged proteins of known members of the outer kinetochore components Mis12, Ndc80, and Nuf2 proposed that the three proteins often copurify, in addition to identifying novel factors, probably members of the potential *Drosophila* MIND/MIS12, NDC80, and SPC105/KNL-1 complexes (Przewlaka et al. 2007; Schittenhelm et al. 2007). Those are assembled sequentially, with the MIND/Mis12 complex being recruited to the centromeres first by the core kinetochore proteins dCENP-A/CENP-C, followed by the recruitment of the NDC80 complexes. In yeast and human, the microtubule-binding Ndc80 complex contains four proteins (NDC80, NUF2, SPC24, and SPC25) (DeLuca et al. 2006; Wei et al. 2007; Ciferri et al. 2008). A canonical Spc24 homolog could not be identified in *Drosophila* (Cheeseman et al. 2006; Przewlaka et al. 2007; Schittenhelm et al. 2007), and only SPC25 appears to anchor the NDC80 complex to the underlying KNL1 SPC105/KLN-1 and MIND/MIS-12 complex (Przewlaka et al. 2007). dCENP-C is closer to centromeric chromatin and dCENP-A via its C-terminus, which also contains the region responsible for dCENP-A interaction, and its N-terminus extends to the outer kinetochore (Schittenhelm et al. 2007; Przewlaka et al. 2011). Closer to dCENP-C N-terminus is the dSPC25 subunit, followed closely by dMIS12, while in the outermost surface, and possibly closer to the microtubules, are dNUF2 (Schittenhelm et al. 2007) and dNDC80 (Fig. 2b; Table 1).

The quantitative architecture of *Drosophila* kinetochores has also been explored. Counting the number of molecules per

kinetochore of different kinetochore components (dSPC105, dMIS12, ddSPC25, and NUF2) (Fig. 2b) has suggested that these proteins are present in similar amounts, lower than the amount of dCENP-A at the centromere, and about half the amount of dCENP-C, which can dimerize (Schittenhelm et al. 2010). Interestingly, these numbers seem to correlate with the calculated number of kinetochore microtubules in flies, and this correlation is consistent and suggests conservation with observations in yeast kinetochores (Maiato et al. 2006; Schittenhelm et al. 2010).

The above studies provided insights in the layered structure of the kinetochore and its quantitative organization. *Drosophila* seems to employ a simplified version of an inner kinetochore, but it appears that the outer kinetochore layer and the one directly responsible for mediating attachments to spindle microtubules is largely conserved and resembles the one from other species.

3D structure of the kinetochore

Centromeric chromatin constitutes the platform for the assembly of the kinetochore, and despite its 2D organization being elucidated, its 3D organization has not been extensively studied yet. Only a few models have been proposed for the 3D organization of the centromere and the kinetochore. Early reports based on a repeat subunit model (Zinkowski et al. 1991) proposed that centromeric chromatin might be arranged in a spiral solenoid or a looping structure, in which dCENP-A-containing blocks are exposed to the poleward faces of the chromosomes, allowing interactions with the CCAN and eventually outer kinetochore proteins, while the H3-containing nucleosome blocks are facing the inner centromere region in the space between the 2 sister chromatids (Blower et al. 2002). EM experiments on human centromeres and neocentromeres suggested that the 3D organization of the primary constrictions of chromosomes is perhaps even more complex, possibly implicating further coiling or folding of the coiled centromeric chromatin, to achieve a higher-order compact structure (Marshall et al. 2008). The most recent model that was put forward for the 3D organization of centromeres is taking into account all the above, and based on results from chicken DT-40 cells and stretched kinetochore fibers, kinetochores are proposed to be arranged in a “boustrophedon” model (Ribeiro et al. 2010). In such

a model, centromeric chromatin is anticipated to be arranged in a sinusoidal layered arrangement, linked to heterochromatin on each side, and allowing for interactions with the CCAN and outer kinetochore members in the outer surface (Ribeiro et al. 2010). A key feature of all models for the higher-order organization of the centromere is that outer kinetochore components and pericentric heterochromatin bound by HP1 are distinct domains (Blower and Karpen 2001). Interestingly, pericentric heterochromatin appears to be located between two sister kinetochores and in the chromatin immediately flanking the centromere (Sullivan and Karpen 2004).

The localization pattern of specific kinetochore members on mitotic chromosomes was visualized in high resolution to decipher their specific spatial arrangement. Interestingly, dCENP-C does not have the same pattern as dCENP-A as seen on mitotic kinetochores imaged in high resolution. It colocalizes with dCENP-A at the core of the kinetochore and is seen as two discrete foci on two sister chromatids; however, it seems to extend further than the dCENP-A discrete dot, possibly reaching to the outer kinetochore (Blower et al. 2002; Maiato et al. 2006; Schittenhelm et al. 2007). Unlike dCENP-C, outer kinetochore proteins like CENP-E were found to be forming a layer above the dCENP-A spot, and the whole kinetochore forms a bilaminar hemispheric structure (Maiato et al. 2006). Although the molecular details regarding the interactions of centromeric chromatin and kinetochore proteins have been extensively demonstrated, it remains an open question how exactly centromeric chromatin is folded in vivo in interphase (since most of the above studies focused on mitosis) and how this organization is maintained before and after cell division. In addition, whether heterochromatin flanking the complex kinetochore structure plays any role to the 3D arrangement or functionality during mitosis remains to be determined.

Epigenetic identity of centromeres

dCENP-A is sufficient for centromere identity

While centromeric DNA sequences show little conservation between different organisms, CENP-A is present across many species in the opisthokont (including fungi and animals) and plant group of eukaryotes. It was therefore speculated that CENP-A could play an important role in determining centromere identity. Indeed, it has now been extensively demonstrated that centromeres in most organisms—with the exception of budding yeast—are epigenetically defined (Karpen and Allshire 1997), with CENP-A being the central epigenetic mark.

In support of this, the plasticity and spreading of a centromere mark were particularly highlighted in studies using the propagation of chromosome fragments after X-ray irradiation in *Drosophila*. Neocentromeres only formed on fragments that were located proximally to endogenous centromeres, suggesting that centromere identity can spread to nearby DNA sequences (Maggert and Karpen 2001). Indeed, it was later shown that mislocalization of dCENP-A promotes the formation of functional kinetochores both in cells and animals (Heun et al. 2006), contradicting studies performed in human cells where overexpression of CENP-A was not sufficient to induce formation of functional ectopic kinetochores (Hooser et al. 2001). This induction of functional ectopic kinetochores led to chromosome segregation defects, resulting in aneuploidy and growth defects (Heun et al. 2006). Overexpressed dCENP-A is incorporated in euchromatic regions but not in pericentric heterochromatin, supporting a previous hypothesis that heterochromatin antagonizes the spreading of centromeric chromatin, controlling its size and distribution

(Maggert and Karpen 2001; Heun et al. 2006). Heterochromatin does, however, promote de novo functional centromere formation, as dCENP-A-overexpression-mediated functional kinetochores preferentially assemble close to pericentric heterochromatin or telomeres (Olszak et al. 2011), in agreement with findings in *Schizosaccharomyces pombe* and attempts for building human artificial chromosomes (Barrey and Heun 2017).

Further supporting a central role for CENP-A in conferring centromere identity was the finding that artificial tethering of dCENP-A to an ectopic site is sufficient to promote functional kinetochore formation as well as self-propagation and inheritance of the epigenetic mark (Mendiburo et al. 2011), alongside similar findings in human cells and in vitro (Barnhart et al. 2011; Guse et al. 2011). Using a heterologous system to dissect the molecular determinants for *Drosophila* centromere identity, it was shown that the three factors dCENP-A, dCENP-C, and the chaperone Cal1, are sufficient to promote heritable centromere identity, implying a 3-component epigenetic loop for centromere propagation (Roure et al. 2019).

dCENP-A inheritance

dCENP-A is present at centromeres in postmeiotic spermatids (Dunleavy et al. 2012; Raychaudhuri et al. 2012), suggesting that it survives the nucleosome replacement by protamines during spermatogenesis. It also serves as a mark for the transgenerational inheritance of centromere identity, as well as being important for the quantity of dCENP-A loaded to the centromeres, as seen in studies where dCENP-A levels were manipulated in males and the levels of dCENP-A on paternal chromosomes were quantified in the offspring (Raychaudhuri et al. 2012). Those altered dCENP-A levels were inherited on the paternal chromosomes of the offspring throughout development. When dCENP-A was eliminated in the males, dCENP-A was not present at the centromeres of paternal chromosomes of the embryos. Altogether, these data suggest that already present dCENP-A is a prerequisite for the recruitment of more dCENP-A, but it also regulates the amount of more dCENP-A that is recruited (Raychaudhuri et al. 2012).

Timing of dCENP-A assembly

Unlike canonical H3, dCENP-A incorporation into chromatin is replication independent (Ahmad and Henikoff 2001; Sullivan and Karpen 2001; Blower et al. 2002; Schuh et al. 2007). Following mitosis, dCENP-A is equally distributed to the daughter cells, as evinced by pulse-chase experiments after one cell division where a close to 50% decrease of labeled dCENP-A levels is measured (Mellone et al. 2011; Bobkov et al. 2020). The timing of dCENP-A assembly into centromeres was first studied by live imaging of syncytial embryos, where it was suggested that dCENP-A is incorporated into centromeric chromatin during anaphase (Schuh et al. 2007). The assembly of dCENP-A during that time is accompanied by an increase in dCENP-C levels at the centromeres, and the whole process requires progression through mitosis (Schuh et al. 2007). In *Kc* cells, dCENP-A assembly was shown to be tightly coupled to the cell cycle, since cyclin A and regulator of cyclin A1 (RCA1) depletion was shown to influence dCENP-A localization to the centromere (Erhardt et al. 2008). In human cells, the timing of CENP-A deposition has been determined to start at the end of mitosis in anaphase and continue through G1 (Jansen et al. 2007). For *Drosophila*, the exact timing of dCENP-A deposition remains controversial. Studies in *Drosophila* cell culture cells (S2 and *Kc*) proposed that dCENP-A loading occurs only in mitosis, specifically starting already in metaphase, following the increase of Cal1 at centromeres during prophase

(Mellone *et al.* 2011; Pauleau *et al.* 2019). However, in vivo imaging of EGFP-tagged dCENP-A revealed that its loading in S2R+ cells occurs mostly during early G1 with only a transient increase in dCENP-A levels at centromeres during metaphase (Lidsky *et al.* 2013). This is similar to observations in somatic tissues of the fly (Dunleavy *et al.* 2012) and the release of cytoplasm-anchored dCENP-A in S2 cells (Bobkov *et al.* 2018), matching more closely the window for CENP-A loading in human cells (Jansen *et al.* 2007; Dunleavy *et al.* 2012; Lidsky *et al.* 2013; Bobkov *et al.* 2018). Interestingly, in meiosis, dCENP-A assembly timing is different than mitosis of cultured cells and somatic tissues, since, in females, it occurs during prophase I while, in males, it occurs in 2 stages, during prophase I and after exit from meiosis II, in spermatids (Dunleavy *et al.* 2012).

Proteins involved in dCENP-A assembly and regulation

In humans, CENP-A assembly is mediated by a complex set of proteins priming chromatin before incorporation of CENP-A. This involves the CENP-A-specific chaperone holliday junction recognition protein (HJURP) and CENP-C and connecting these two proteins, the Mis18 complex, which is composed of Mis18 α , Mis18 β and Mis18BP1 (McKinley and Cheesman 2014) (Fig. 2a). *Drosophila* instead seems to have simplified this process for achieving the same goal. To identify the factors mediating dCENP-A assembly into centromeric chromatin, different approaches have been taken in multiple studies. First, the complex interacting with soluble dCENP-A in S2 cells was described to contain H4 and the chaperone RbAp48, and the latter was shown to be able to assemble dCENP-A-containing nucleosomes in vitro (Furuyama and Henikoff 2006). It was earlier shown that *argonaute-2* mutant embryos have defects in dCENP-A assembly, since they presented lower levels of dCENP-A at centromeres, suggesting that the RNAi machinery might play a role in this process (Deshpande *et al.* 2005). In a genome-wide RNAi screen looking for factors affecting dCENP-A centromeric localization, dCENP-C and Cal1 were the two proteins showing the strongest effect, along with the cell-cycle regulators cyclin A and the RCA1 (Erhardt *et al.* 2008). A previous study had already identified Cal1 as a factor important for centromeric localization of dCENP-A and dCENP-C (Goshima *et al.* 2007), and depletion of dCENP-C also impaired dCENP-A localization (Orr and Sunkel 2011).

As studied by in vitro nucleosome assembly assays, Cal1 can specifically assemble left-handed octameric nucleosomes containing dCENP-A through its N-terminal domain and was therefore called the dCENP-A-specific chaperone (Chen *et al.* 2014). Interestingly, tethering of Cal1 to an ectopic site is also sufficient to promote dCENP-A and kinetochore assembly, similar to what was observed in tethering experiments of HJURP (Hori *et al.* 2013; Chen *et al.* 2014; Palladino *et al.* 2020). Curiously, unlike HJURP, Cal1 is present at centromeres throughout the cell cycle, almost constitutively, in addition to the nucleolus, where it was suggested to be sequestered through its middle region (Erhardt *et al.* 2008), when not in complex with dCENP-A and dCENP-C (Schittenhelm *et al.* 2010; Lidsky *et al.* 2013) (Fig. 2a). The interactions between Cal1, dCENP-A, and dCENP-C were further dissected by yeast 2- and 3-hybrid assays, which showed that Cal1 interacts via its N-terminus with dCENP-A, while its C-terminus interacts with the C-terminus of dCENP-C. Since Cal1 can directly interact with dCENP-A and dCENP-C, it was suggested that it might be a bridging factor between the 3 proteins at centromeres (Schittenhelm *et al.* 2010). To answer the question of whether Cal1 acts as a stoichiometric bridge between dCENP-A and dCENP-C

molecules, the number of molecules per centromere was counted in wing imaginal discs. Interestingly, it was found that while dCENP-A is present in 84 copies per centromere, dCENP-C was present as 135 molecules and Cal1 as only 2.5 molecules per centromere, excluding the possibility that all dCENP-A and dCENP-C molecules are bridged by Cal1 (Schittenhelm *et al.* 2010).

Together, the N- and C-terminal regions of Cal1, responsible for interaction with dCENP-A and dCENP-C respectively, are required for its centromeric function (Schittenhelm *et al.* 2010), and its N-terminus was shown to contain an "Scm3 domain"-like region (Phansalkar *et al.* 2012), suggesting possible conservation in the domain organization of Cal1, as in CENP-A chaperones in other species. Despite its evolutionary uniqueness, Cal1 participates in a process that is evolutionarily conserved, that is, however, as it seems, highly plastic. Even though Cal1 is structurally distinct from other CENP-A chaperones, and especially the human ortholog HJURP, a recent study reporting on the crystal structure of Cal1 together with dCENP-A-H4 and dCENP-C surprisingly suggests that Cal1 engages with its substrates using modes employed by both HJURP and Smc3 (Cho and Harrison 2011; Hu *et al.* 2011; Chik *et al.* 2019; Medina-Pritchard *et al.* 2020).

In addition to Cal1, soluble dCENP-A was found to interact with Modulo (Chen *et al.* 2012), the chaperones Caf1 and FACT, and Hat1, which binds it in a complex distinct from the Cal1/FACT-CENP-A-H4 complex, but together with acetylated H4 and Caf1 (Boltegangen *et al.* 2015). This interaction between dCENP-A and Hat1 does not affect its acetylation status but rather was shown to play a role in dCENP-A assembly.

In summary, it appears that Cal1 or dCENP-C in *Drosophila* fulfills some functions of the Mis18 complex in human cells, yet the molecular details of their action remain to be established.

Regulation of dCENP-A assembly

A crucial question in the field has been how the deposition of the histone H3 variant CENP-A is restricted specifically to the centromeres and how it is replenished after DNA replication. While dCENP-A is expressed in the early S phase (Henikoff *et al.* 2000), its levels and availability are effectively regulated by proteasomal degradation, and even overexpressed and mislocalized dCENP-A will be constrained at the centromeres within a few cell cycles (Moreno-Moreno *et al.* 2006; Olszak *et al.* 2011). Indeed, the F-box protein, partner of paired (Ppa), was found to interact with the L1/ α 2 of the HFD of dCENP-A, in both soluble and nucleosomal states, and regulate its stability (Moreno-Moreno *et al.* 2011). Ppa is a component of the E3 ubiquitin ligase SCF, which interestingly is inactive during mitosis (Nakayama and Nakayama 2006), the preferred cell-cycle stage for dCENP-A incorporation into centromeric chromatin. This suggests that a tight regulation for specific incorporation of dCENP-A only at centromeres is achieved because dCENP-A is only made available in mitosis when its proteolysis is inhibited. A fine balance between dCENP-A degradation and availability during a specific cell-cycle stage ensures that dCENP-A is not misincorporated elsewhere in the genome and that it is restricted only at centromeres. Proteasome-mediated degradation of cyclin A and other targets is required for dCENP-A assembly in *Drosophila* (Mellone *et al.* 2011).

To add another level of regulation in dCENP-A assembly, Cal1 was found to interact directly with RDX, a protein that acts as an adaptor for CUL3-mediated ubiquitylation. Cal1 is not itself a target of the complex; it rather helps in specifying dCENP-A as a substrate, and interestingly, ubiquitylation of the latter does not promote its degradation but rather stabilizes it together with Cal1 (Bade *et al.* 2014). CHRAC14, another interacting factor of

dCENP-A, prevents its incorporation at sites of DNA damage (Mathew *et al.* 2014).

Like dCENP-A, dCENP-C is also equally distributed to daughter cells while Cal1 is turned over at a higher rate per cell cycle (66% of preexisting Cal1 is replenished in each cell cycle) (Mellone *et al.* 2011). Cal1 is highly dynamic at centromeres, with its levels being decreased during progression to the G1 phase but increase again in the transition between G1 and S phases and do not change during S phase and G2 (Lidsky *et al.* 2013). During the late M phase, early G1, when dCENP-A loading occurs, centromeric Cal1 levels remain high (Lidsky *et al.* 2013). In contrast, dCENP-C is mostly stably incorporated at centromeres with some fluctuation of its levels seen throughout the cell cycle. dCENP-C was found to be increased mainly during the late S phase, G2, and mitosis, differently from dCENP-A and also differently from the timing of its centromere loading in early embryos (Schuh *et al.* 2007; Lidsky *et al.* 2013). While Cal1 dynamics correlate with the time dCENP-A loading occurs, it remains to be determined what is the physiological relevance of the changes in dCENP-C levels throughout the cell cycle. Particularly, does the increase seen during the late S phase, G2, and M play any role in ensuring dCENP-A stability/maintenance of centromeric position after dilution of dCENP-A nucleosomes following DNA replication?

Centromeres are transcriptionally active, and this affects dCENP-A assembly

Several hypotheses attempted to explain the preferential incorporation of dCENP-A over H3 into centromeric nucleosomes. One such proposal was that centromeres are replicated with different timing compared to other loci, allowing for the distinction of the sites where dCENP-A should be assembled (Ahmad and Henikoff 2001). Although an attractive model, experiments using different pulse labeling and chase timings showed that centromeres are asynchronously replicated in the mid and late S phase in *Drosophila* tissue culture cells and larval neuroblasts, suggesting that replication timing is not the determinant of centromeric nucleosome assembly (Sullivan and Karpen 2001).

dCENP-A is distributed to the two sister centromeres during every S phase, and gaps emerge at sites previously occupied by dCENP-A. These are possibly filled by placeholder H3/H3.3 nucleosomes as it has been suggested by studies in human cells and fission yeast (Dunleavy *et al.* 2011; Shukla *et al.* 2018). Histone deposition during the S phase is easily realized since gaps are being formed that can be subsequently filled by new histones. Since CENP-A deposition in *Drosophila*, and most organisms, is uncoupled from DNA replication, it is hard to envision how dCENP-A can be deposited at centromeres. Recent studies in accordance with the notion that fission yeast and human centromeres are transcribed (Choi *et al.* 2011; 2012; Chan *et al.* 2012; McNulty *et al.* 2017) have shown that similarly fly centromeres are transcribed and that FACT-mediated transcription is important for dCENP-A assembly in somatic cells (Rošić *et al.* 2014; Chen *et al.* 2015; Bobkov *et al.* 2018). This provides an attractive model for the specific incorporation of dCENP-A at centromeres. dCENP-A becomes available during mitosis when chromatin is largely compacted. Conveniently, centromeres are specifically transcribed during mitosis, potentially evicting H3 placeholder nucleosomes, allowing for the generation of chromatin gaps for the specific incorporation of dCENP-A nucleosomes. It is important to note that during early *Drosophila* embryogenesis, the loading of dCENP-A does not require RNAPII-mediated transcription (Ghosh and Lehner 2022). It is tempting to speculate that this might reflect more open, nucleosome gap-containing chromatin because of the very fast nuclear

divisions in early embryogenesis, while somatic chromatin might have more time for full nucleosome occupancy and hence be more “sealed” for dCENP-A incorporation. Moreover, it is still unclear whether the act of transcription per se or the RNA products are important for dCENP-A incorporation, and this remains to be experimentally determined. However, that transcription-mediated chromatin remodeling can evict nucleosomes at the centromere, including those containing old dCENP-A, has been recently demonstrated in cells depleted for the transcriptional chaperone Spt6 that serves as CENP-A maintenance factor (Bobkov *et al.* 2020).

Drosophila as a model organism to understand neocentromeres and evolution

Lessons from neocentromeres and dicentric chromosomes

Concurrently, with neocentromere cases being reported in human cells (du Voullaire *et al.* 1993; Depinet *et al.* 1997; Sart *et al.* 1997), a study using minichromosomes lacking centromeric sequences in *Drosophila* aimed at understanding the relationship between DNA sequence and centromere function (Williams *et al.* 1998). Interestingly, these seemingly structurally acentric minichromosomes were recovered after irradiation mutagenesis (Murphy and Karpen 1995), even though they are lacking centromeric sequences, raising the possibility that they acquire centromere functions and thus acquire neocentromeres (Williams *et al.* 1998). Indeed, these minichromosomes were not lost upon cell division, mostly in male meiosis and less in mitosis and female meiosis. They were shown to bind outer kinetochore proteins like ZW10 and the anaphase spindle equally well as endogenous centromeres, and when checked, they were negative on centromeric DNA sequences (Williams *et al.* 1998). Several lines of evidence suggest that there are no specific sequence characteristics that would determine centromere function; however, the specific chromatin state of a region was suggested to play a role in the specification of centromeres. Particularly, Henikoff and colleagues (Platero *et al.* 1999) using flies containing a distal heterochromatic block on the right arm of their second chromosome demonstrated that this heterochromatic block occasionally acquires centromere activity, suggesting that even though sequence per se is not capable of determining centromere identity, heterochromatic blocks composed of satellites display centromere competence (Platero *et al.* 1999). *Drosophila*, in contrast to mammalian systems, not only served as a powerful tool to study both neocentromere formation but also allowed for manipulation of the system.

While natural selection favored either monocentric or holocentric chromosomes, the presence of intermediate variants usually leads to breakage–fusion events, causing major chromosomal rearrangements. Interestingly, a dicentric transmissible chromosome was recovered in *Drosophila* (Y S · Y L 2Rh4), which consists essentially of the entire Y and fourth chromosomes joined by 2R heterochromatin, even though earlier studies have suggested that this is not possible (Ault and Lyttle 1988). It was discussed that this dicentric was recovered due to differences in strength of the two centromeres and that since the Y chromosome might have a stronger kinetochore has contributed to the recovery and not breakage of this dicentric chromosome (Ault and Lyttle 1988). Another dicentric chromosome C(1)A was recovered in flies, with one of the two centromeres carrying all the sequences of the centromere of the Y chromosome and the other carrying only a part of the Y centromeric region, rich in telomere-related sequences (Agudo *et al.* 2000). Similar to what has been reported in

other organisms (Fisher *et al.* 1997; Sullivan and Wilard 1998), one of its two centromeres is inactivated in most cases. Interestingly, in the less frequent cases of dicentrics, one of the two kinetochores is more faintly stained than the other, suggesting that it is possible that kinetochore strength might play a role in the successful segregation of dicentric chromosomes (Agudo *et al.* 2000). Importantly, the above studies exhibit strong lines of evidence that indeed centromeres are epigenetically, rather than genetically specified. Intriguingly, the same principles seem to be conserved across species, except for *Saccharomyces cerevisiae*, where there is a specific centromeric sequence determining centromere position.

***Drosophila* centromeres as a model to understand meiotic drive**

The hypothesis for a meiotic drive in *Drosophila* was put forward by Sandler and Novitski back in 1957 (Sandler and Novitski, 1957), as an attempt to explain a phenomenon in which species create certain gametes in unequal frequencies. Such an occurrence is common in female gametes of most animals and plants, where, of the four meiotic products, only one survives and makes it to the oocyte (Henikoff and Malik 2002; Rosin and Mellone 2017). Centromeres were characterized as “selfish” genomic elements, which, by taking advantage of the meiotic drive, try to make it into the oocyte. Rapidly evolving centromeres that usually acquire expanded satellite arrays were proposed to have an advantage over smaller centromeres, since they can attract more kinetochore proteins, resulting in their kinetochores being “stronger” and thus more likely to attach to the spindle microtubules and make it to the oocyte. In the female meiosis, this is advantageous to the “newly expanded” centromere; however, this could have deleterious consequences in male meiosis, since such a discrepancy in the size and sequence of the evolved centromere would not allow for efficient homolog pairing and lead to increased non-disjunctions (Rosin and Mellone 2017). It was proposed that the proteins that bind centromeric sequences, like CENP-A, also rapidly evolve, maybe to balance the effects of these phenomena, taking part in an evolutionary tug-of-war game. This was described as “centromere drive” (Zwick *et al.* 1999; Henikoff and Malik 2002).

The centromere paradox and evolution of centromeres

From an evolutionary point of view, centromeres represent a fascinating paradigm. *Drosophila* constitutes an ideal model organism to study evolution, since many genomes have been sequenced and the phylogenetic relationships between species have been extensively determined (Kursel and Malik 2017). Although centromere function is highly conserved, centromere specification appears to be plastic as demonstrated by the lack of evolutionary conserved DNA sequences and highly divergent paralogs of CENP-A. Only the network of outer kinetochore proteins that mediate interactions with spindle microtubules reveals a higher level of evolutionary conservation. It is paradoxical that a locus with such an essential function has evolved to be present in so many different flavors. *Drosophila* has been used as a model to study not only the evolution of both dCENP-A and centromeric DNA but also the specific dCENP-A chaperone in flies, Cal1, providing important insights into the evolution of centromeres (Malik and Henikoff 2001; Phansalkar *et al.* 2012; Rosin and Mellone 2016; Kursel and Malik 2017). While, in *melanogaster*, dCENP-A is encoded by a single copy gene, two recent studies have reported that in other species, there have been five independent duplications of the genes during evolution, with the products retaining

their centromeric localization but seems that their functions might be nonredundant, while some are expressed primarily in the male germline (Kursel and Malik 2017; Teixeira *et al.* 2018). Interestingly, the cenp-c gene was also found to have duplicated during evolution, something that was probably expected, given its association with Cal1 and dCENP-A (Rosin and Mellone 2016; Teixeira *et al.* 2018). Like some of the cid genes, one cenp-c duplication product also presents a male bias, raising the exciting possibility that kinetochore structure is diverse in specific contexts (Teixeira *et al.* 2018).

The gene-encoding dCENP-A (cid) was shown to have undergone adaptive evolution mostly not only in its N-terminal tail but also in part of the core region, including the L1 of the HFD (Malik and Henikoff 2001; Malik *et al.* 2002). The adaptive changes in regions that mediate DNA contacts, like L1 or the α_N helix on the N-terminus of the HFD, are particularly interesting and raised the exciting possibility that these changes might reflect changes in DNA binding specificity (Malik *et al.* 2002). An interesting explanation for this paradox considers the fact that centromeric DNA itself contains the sequences that are most rapidly evolving across eukaryotes (Haaf and Wilard 1997; Csink and Henikoff 1998; Murphy and Karpen 1998; Henikoff *et al.* 2001). Henikoff *et al.* (2001) proposed that these coevolution events happen in parallel to relieve centromeric imbalance, inhibit the accumulation of deleterious expansion of satellite repeats, and also provide reproductive isolation between different populations of a species leading to speciation. The latter hypothesis could, however, be rejected, since cid sequences in *simulans* and *melanogaster* strains do not correlate with the ability of these strains to rescue hybrid fitness and suggesting that evolution of the gene-encoding cid is not a cause for speciation (Sainz *et al.* 2003). Another challenge of the model suggesting that cid and centromeric DNA coevolve in response to the centromere drive was a recent study reporting that Cal1 Scm3-like region coevolves with the L1 of cid, and this concerted interplay is important for the deposition of dCENP-A into centromeric nucleosomes (Rosin and Mellone 2016). These data suggest that while centromeric sequences evolve, the L1 of cid also rapidly evolves. However, cal1 evolves at a slower rate (Phansalkar *et al.* 2012). As a result, during the intermediate stage of evolution, Cal1 cannot deposit dCENP-A to the whole length of the expanded centromeric sequences, balancing like this the possible effect of a “stronger” kinetochore due to the larger size of repeats. In the next step, Cal1 Scm3-like region also evolves to be able to deposit dCENP-A into the selected centromeric sequences (Rosin and Mellone 2016).

In addition to dCENP-A and centromeric DNA, other centromere-associated proteins in flies, like members of the condensin complex and HP1 proteins, were also shown to rapidly evolve, suggesting that they may have to compete for spindle attachment that will be optimal (Beck and Liopart 2015). Kinetochore protein evolution is greater in flies and worms, with most of the CCAN components missing from these species, and some of the outer kinetochore components are divergent compared to yeasts and mammals (Meraldi *et al.* 2006).

Concluding Remarks

In this chapter, we tried to give an overview of our current understanding of the *Drosophila* centromere in comparison with humans and other eukaryotic organisms. Using the fruit fly as a model organism to study essential and conserved processes like chromosome segregation can reveal intriguing differences between distantly related eukaryotes. More importantly, these studies

help to unmask the underlying common basic principles that connect all living beings to ensure the passing on of genetic material from one generation to the next.

Funding

E.K. and P.H. were supported by the Wellcome Trust (103897/Z/14/Z) and P.H. by the BBSRC, UK [grant number BB/W00674X/1]. The Wellcome Trust Centre for Cell Biology is supported by core funding from the Wellcome Trust (092076 and 203149).

Conflicts of interest

The author(s) declare no conflict of interest.

Literature cited

- Agudo M, Abad JP, Molina I, Losada A, Ripoll P, Villasante A. 2000. A dicentric chromosome of *Drosophila melanogaster* showing alternate centromere inactivation. *Chromosoma*. 109(3):190–196. doi:10.1007/s004120050427.
- Agudo M, Losada A, Abad JP, Pimpinelli S, Ripoll P, Villasante A. 1999. Centromeres from telomeres? The centromeric region of the Y chromosome of *Drosophila melanogaster* contains a tandem array of telomeric HeT-A- and TART-related sequences. *Nucleic Acids Res*. 27(16):3318–3324. doi:10.1093/nar/27.16.3318.
- Ahmad K, Henikoff S. 2001. Centromeres are specialized replication domains in heterochromatin. *J Cell Biol*. 153(1):101–109. doi:10.1083/jcb.153.1.101.
- Ahmad K, Henikoff S. 2002. Histone H3 variants specify modes of chromatin assembly. *Proc Natl Acad Sci U S A*. 99(suppl_4):16477–16484. doi:10.1073/pnas.172403699.
- Ali-Ahmad A, Bilokapic S, Schäfer IB, Halić M, Sekulić N. 2019. CENP-C unwraps the human CENP-A nucleosome through the H2A C-terminal tail. *EMBO Rep*. 20(10):e48913. doi:10.15252/embr.201948913.
- Ali-ahmad A, Sekulić N. 2020. CENP-A nucleosome—a chromatin-embedded pedestal for the centromere: lessons learned from structural biology. *Essays Biochem*. 64:205–221. doi:10.1042/EBC20190074.
- Ault JG, Lyttle TW. 1988. A transmissible dicentric chromosome in *Drosophila melanogaster*. *Chromosoma*. 97:71–79. doi:10.1007/BF00331796.
- Bade D, Pauleau AL, Wendler A, Erhardt S. 2014. The E3 ligase CUL3/RDX controls centromere maintenance by ubiquitylating and stabilizing CENP-A in a CAL1-dependent manner. *Dev Cell*. 28:508–519. doi:10.1016/j.devcel.2014.01.031.
- Barnhart MC, Kuich PHJL, Stellfox ME, Ward JA, Bassett EA, Black BE, Foltz DR. 2011. HJURP is a CENP-A chromatin assembly factor sufficient to form a functional de novo kinetochore. *J Cell Biol*. 194(2):229–243. doi:10.1083/jcb.201012017.
- Barrey EJ, Heun P. 2017. Artificial chromosomes and strategies to initiate epigenetic centromere establishment, In: Black BE, editor. *Centromeres and Kinetochores: Discovering the Molecular Mechanisms Underlying Chromosome Inheritance*. Cham: Springer International Publishing. p. 193–212. https://doi.org/10.1007/978-3-319-58592-5_8.
- Barth TK, Schade GOM, Schmidt A, Vetter I, Wirth M, Heun P, Imhof A, Thomae AW. 2015. Identification of *Drosophila* centromere associated proteins by quantitative affinity purification-mass spectrometry. *Data Brief*. 4:544–550. doi:10.1016/j.dib.2015.07.016.
- Barth TK, Schade GOM, Schmidt A, Vetter I, Wirth M, Heun P, Thomae AW, Imhof A. 2014. Identification of novel *Drosophila* centromere-associated proteins. *Proteomics*. 14(19):2167–2178. doi:10.1002/pmic.201400052.
- Beck EA, Llopart A. 2015. Widespread positive selection drives differentiation of centromeric proteins in the *Drosophila melanogaster* subgroup. *Sci Rep*. 5:17197. doi:10.1038/srep17197.
- Blower MD, Karpen GH. 2001. The role of *Drosophila* CID in kinetochore formation, cell-cycle progression and heterochromatin interactions. *Nat Cell Biol*. 3(8):730–739. doi:10.1038/35087045.
- Blower MD, Sullivan BA, Karpen GH. 2002. Conserved organization of centromeric chromatin in flies and humans. *Dev Cell*. 2:319–330. doi:10.1016/S1534-5807(02)00135-1.
- Bobkov GOM, Gilbert N, Heun P. 2018. Centromere transcription allows CENP-A to transit from chromatin association to stable incorporation. *J Cell Biol*. 217(6):1957–1972. doi:10.1083/jcb.201611087.
- Bobkov GOM, Huang A, Berg SJW, van den Berg SJW, Mitra S, Anselm E, Lazou V, Schunter S, et al. 2020. Spt6 is a maintenance factor for centromeric CENP-A. *Nat Commun*. 11(1):2919. doi:10.1038/s41467-020-16695-7.
- Boltengagen M, Huang A, Boltengagen A, Trixl L, Lindner H, Kremser L, Offterdinger M, Lusser A. 2015. A novel role for the histone acetyltransferase Hat1 in the CENP-A/CID assembly pathway in *Drosophila melanogaster*. *Nucleic Acids Res*. 44(5):2145–2159. doi:10.1093/nar/gkv1235.
- Bonaccorsi S, Lohe A. 1991. Fine mapping of satellite DNA sequences along the Y chromosome of *Drosophila melanogaster*: relationships between satellite sequences and fertility factors. *Genetics*. 129(1):177–189. doi:10.1093/genetics/129.1.177.
- Botchan M, Kram R, Schmid CW, Hearst JE. 1971. Isolation and chromosomal localization of highly repeated DNA sequences in *Drosophila melanogaster*. *Proc Natl Acad Sci U S A*. 68(6):1125–1129. doi:10.1073/pnas.68.6.1125.
- Carmena M, Abad JP, Villasante A, Gonzalez C. 1993. The *Drosophila melanogaster* dodecasatellite sequence is closely linked to the centromere and can form connections between sister chromatids during mitosis. *J Cell Sci*. 105(1):41–50. doi:10.1242/jcs.105.1.41.
- Carroll CW, Milks KJ, Straight AF. 2010. Dual recognition of CENP-A nucleosomes is required for centromere assembly. *J Cell Biol*. 189:1143–1155. doi:10.1083/jcb.201001013.
- Chan FL, Marshall OJ, Saffery R, Won Kim B, Earle E, Choo KHA, Wong LH. 2012. Active transcription and essential role of RNA polymerase II at the centromere during mitosis. *Proc Natl Acad Sci U S A*. 109(6):1979–1984. doi:10.1073/pnas.1108705109.
- Chang C-H, Chavan A, Palladino J, Wei X, Martins NMC, Santinello B, Chen C-C, Erceg J, Beliveau BJ, Wu CT, et al. 2019. Islands of retroelements are the major components of *Drosophila* centromeres. *PLoS Biol*. 17(5). doi:10.1371/journal.pbio.3000241.
- Cheeseman IM, Chappie JS, Wilson-Kubalek EM, Desai A. 2006. The conserved KMN network constitutes the core microtubule-binding site of the kinetochore. *Cell*. 127:983–997. doi:10.1016/j.cell.2006.09.039.
- Chen CC, Bowers S, Lipinszki Z, Palladino J, Trusiak S, Bettini E, Rosin L, Przewloka MR, Glover DM, O'Neill RJ, et al. 2015. Establishment of centromeric chromatin by the CENP-A assembly factor CAL1 requires FACT-mediated transcription. *Dev Cell*. 34:73–84. doi:10.1016/j.devcel.2015.05.012.
- Chen CC, Dechassa ML, Bettini E, Ledoux MB, Belisario C, Heun P, Luger K, Mellone BG. 2014. CAL1 is the *Drosophila* CENP-A assembly factor. *J Cell Biol*. 204(3):313–329. doi:10.1083/jcb.201305036.
- Chen CC, Greene E, Bowers SR, Mellone BG. 2012. A role for the CAL1-partner Modulo in centromere integrity and accurate chromosome segregation in *Drosophila*. *PLoS One*. 7:e45094. doi:10.1371/journal.pone.0045094.

- Chik JK, Moiseeva V, Goel PK, Meinen BA, Koldewey P, An S, Mellone BG, Subramanian L, Cho US. 2019. Structures of CENP-C cupin domains at regional centromeres reveal unique patterns of dimerization and recruitment functions for the inner pocket. *J Biol Chem.* 294(38):14119–14134. doi:[10.1074/jbc.RA119.008464](https://doi.org/10.1074/jbc.RA119.008464).
- Cho US, Harrison SC. 2011. Recognition of the centromere-specific histone Cse4 by the chaperone Scm3. *Proc Natl Acad Sci U S A.* 108(23):9367–9371. doi:[10.1073/pnas.1106389108](https://doi.org/10.1073/pnas.1106389108).
- Choi ES, Strålfors A, Castillo AG, Durand-Dubief M, Ekwall K, Allshire RC. 2011. Identification of noncoding transcripts from within CENP-A chromatin at fission yeast centromeres. *J Biol Chem.* 286(26):23600–23607. doi:[10.1074/jbc.M111.228510](https://doi.org/10.1074/jbc.M111.228510).
- Choi ES, Strålfors A, Catania S, Castillo AG, Svensson JP, Pidoux AL, Ekwall K, Allshire RC. 2012. Factors that promote H3 chromatin integrity during transcription prevent promiscuous deposition of CENP-ACnp1 in fission yeast. *PLoS Genet.* 8(9):e1002985. doi:[10.1371/journal.pgen.1002985](https://doi.org/10.1371/journal.pgen.1002985).
- Ciferri C, Pasqualato S, Screpanti E, Varetto G, Santaguida S, Dos Reis G, Maiolica A, Polka J, De Luca JG, De Wulf P, et al. 2008. Implications for kinetochore-microtubule attachment from the structure of an engineered Ndc80 complex. *Cell.* 133:427–439. doi:[10.1016/j.cell.2008.03.020](https://doi.org/10.1016/j.cell.2008.03.020).
- Csink AK, Henikoff S. 1998. Something from nothing: the evolution and utility of satellite repeats. *Trends Genet.* 14(5):200–204. doi:[10.1016/S0168-9525\(98\)01444-9](https://doi.org/10.1016/S0168-9525(98)01444-9).
- Dalal Y, Wang H, Lindsay S, Henikoff S. 2007. Tetrameric structure of centromeric nucleosomes in interphase *Drosophila* cells. *PLoS Biol.* 5(8):e218. doi:[10.1371/journal.pbio.0050218](https://doi.org/10.1371/journal.pbio.0050218).
- DeBaryshe PG, Pardue M. 2011. Differential maintenance of DNA sequences in telomeric and centromeric heterochromatin. *Genetics.* 187(1):51–60. doi:[10.1534/genetics.110.122994](https://doi.org/10.1534/genetics.110.122994).
- DeLuca JG, Gall WE, Ciferri C, Cimini D, Musacchio A, Salmon ED. 2006. Kinetochore microtubule dynamics and attachment stability are regulated by *hec1*. *Cell.* 127:969–982. doi:[10.1016/j.cell.2006.09.047](https://doi.org/10.1016/j.cell.2006.09.047).
- Depinet TW, Zackowski JL, Earnshaw WC, Kaffe S, Sekhon GS, Stallard R, Sullivan BA, et al. 1997. Characterization of neo-centromeres in marker chromosomes lacking detectable alpha-satellite DNA. *Hum Mol Genet.* 6(8):1195–1204. doi:[10.1093/hmg/6.8.1195](https://doi.org/10.1093/hmg/6.8.1195).
- Deshpande G, Calhoun G, Schedl P. 2005. *Drosophila* Argonaute-2 is required early in embryogenesis for the assembly of centric/centromeric heterochromatin, nuclear division, nuclear migration, and germ-cell formation. *Genes Dev.* 19(14):1680–1685. doi:[10.1101/gad.1316805](https://doi.org/10.1101/gad.1316805).
- Dunleavy EM, Almouzni G, Karpen GH. 2011. H3.3 is deposited at centromeres in S phase as a placeholder for newly assembled CENP-A in G₁ phase. *Nucleus.* 2(2):146–157. doi:[10.4161/nucl.2.2.15211](https://doi.org/10.4161/nucl.2.2.15211).
- Dunleavy EM, Beier NL, Gorgescu W, Tang J, Costes SV, Karpen GH. 2012. The cell cycle timing of centromeric chromatin assembly in *Drosophila* meiosis is distinct from mitosis yet requires CAL1 and CENP-C. *PLoS Biol.* 10:e1001460. doi:[10.1371/journal.pbio.1001460](https://doi.org/10.1371/journal.pbio.1001460).
- Earnshaw WC, Rothfield N. 1985. Identification of a family of human centromere proteins using autoimmune sera from patients with scleroderma. *Chromosoma.* 91(3–4):313–321. doi:[10.1007/BF00328227](https://doi.org/10.1007/BF00328227).
- Erhardt S, Mellone BG, Betts CM, Zhang W, Karpen GH, Straight AF. 2008. Genome-wide analysis reveals a cell cycle-dependent mechanism controlling centromere propagation. *J Cell Biol.* 183:805–818. doi:[10.1083/jcb.200806038](https://doi.org/10.1083/jcb.200806038).
- Ferrer N, Azorin F, Villasante A, Gutierrez C, Abad JP. 1995. Centromeric dodeca-satellite DNA sequences form fold-back structures. *J Mol Biol.* 245(1):8–21. doi:[10.1016/S0022-2836\(95\)80034-4](https://doi.org/10.1016/S0022-2836(95)80034-4).
- Fisher AM, Al-Gazali L, Pramathan T, Quaife R, Cockwell AE, Barber JCK, Earnshaw WC, Axelman J, Migeon BR, Tyler-Smith C. 1997. Centromeric inactivation in a dicentric human Y; 21 translocation chromosome. *Chromosoma.* 106(4):199–206. doi:[10.1007/s004120050240](https://doi.org/10.1007/s004120050240).
- Foltz DR, Jansen LET, Black BE, Bailey AO, Yates Iii JR, Cleveland DW. 2006. The human CENP-A centromeric nucleosome-associated complex. *Nat Cell Biol.* 8(5):458–469. doi:[10.1038/ncb1397](https://doi.org/10.1038/ncb1397).
- Furuyama T, Henikoff S. 2006. Biotin-tag affinity purification of a centromeric nucleosome assembly complex. *Cell Cycle.* 5(12):1269–1274. doi:[10.4161/cc.5.12.2889](https://doi.org/10.4161/cc.5.12.2889).
- Furuyama T, Henikoff S. 2009. Centromeric nucleosomes induce positive DNA supercoils. *Cell.* 138(1):104–113. doi:[10.1016/j.cell.2009.04.049](https://doi.org/10.1016/j.cell.2009.04.049).
- Garavis M, Méndez-Lago M, Gabelica V, Whitehead SL, González C, Villasante A. 2015. The structure of an endogenous *Drosophila* centromere reveals the prevalence of tandemly repeated sequences able to form I-motifs. *Sci Rep.* 5(1):13307. doi:[10.1038/srep13307](https://doi.org/10.1038/srep13307).
- Ghosh S, Lehner CF. 2022. Incorporation of CENP-A/CID into centromeres during early *Drosophila* embryogenesis does not require RNA polymerase II-mediated transcription. *Chromosoma.* 131(1–2):1–17. doi:[10.1007/s00412-022-00767-2](https://doi.org/10.1007/s00412-022-00767-2).
- Goshima G, Wollman R, Goodwin SS, Zhang N, Scholey JM, Vale RD, Stuurman N. 2007. Genes required for mitotic spindle assembly in *Drosophila* S2 cells. *Science.* 316(5823):417–421. doi:[10.1126/science.1141314](https://doi.org/10.1126/science.1141314).
- Guse A, Carroll CW, Moree B, Fuller CJ, Straight AF. 2011. In vitro centromere and kinetochore assembly on defined chromatin templates. *Nature.* 477(7364):354–358. doi:[10.1038/nature10379](https://doi.org/10.1038/nature10379).
- Haaf T, Willard HF. 1997. Chromosome-specific alpha-satellite DNA from the centromere of chimpanzee chromosome 4. *Chromosoma.* 106(4):226–232. doi:[10.1007/s004120050243](https://doi.org/10.1007/s004120050243).
- Heeger S, Leismann O, Schittenhelm R, Schraidt O, Heidmann S, Lehner CF. 2005. Genetic interactions of separase regulatory subunits reveal the diverged *Drosophila* Cenp-C homolog. *Genes Dev.* 19(17):2041–2053. doi:[10.1101/gad.347805](https://doi.org/10.1101/gad.347805).
- Henikoff S, Ahmad K, Malik HS. 2001. The centromere paradox: stable inheritance with rapidly evolving DNA. *Science.* 293(5532):1098–1102. doi:[10.1126/science.1062939](https://doi.org/10.1126/science.1062939).
- Henikoff S, Ahmad K, Platero JS, Van Steensel B, Rubin GM. 2000. Heterochromatic deposition of centromeric histone H3-like proteins. *Proc Natl Acad Sci.* 97(2):716–721. doi:[10.1073/pnas.97.2.716](https://doi.org/10.1073/pnas.97.2.716).
- Henikoff S, Malik HS. 2002. Centromeres: selfish drivers. *Nature.* 417(6886):227. doi:[10.1038/417227a](https://doi.org/10.1038/417227a).
- Heun P, Erhardt S, Blower MD, Weiss S, Skora AD, Karpen GH. 2006. Mislocalization of the *Drosophila* centromere-specific histone CID promotes formation of functional ectopic kinetochores. *Dev Cell.* 10:303–315. doi:[10.1016/j.devcel.2006.01.014](https://doi.org/10.1016/j.devcel.2006.01.014).
- Hooser AA, Van Ouspenski II, Gregson HC, Starr DA, Yen TJ, Goldberg ML, Yokomori K, Earnshaw WC, Sullivan KF, and Brinkley BR. 2001. Specification of kinetochore-forming chromatin by the histone H3 variant CENP-A. *J Cell Sci.* 114(Pt 19):3529–3542. doi:[10.1007/bf00332792](https://doi.org/10.1007/bf00332792).
- Hori T, Shang WH, Takeuchi K, Fukagawa T. 2013. The CCAN recruits CENP-A to the centromere and forms the structural core for kinetochore assembly. *J Cell Biol.* 200:45–60. doi:[10.1083/jcb.201210106](https://doi.org/10.1083/jcb.201210106).
- Horn V, Lacroix L, Gautier T, Takasugi M, Mergny JL, Lacoste J. 2004. Triple helix formation with *Drosophila* satellite repeats. Unexpected stabilization by copper ions. *Biochemistry.* 43(35):11196–11205. doi:[10.1021/bi049287t](https://doi.org/10.1021/bi049287t).

- Hu H, Liu Y, Wang M, Fang J, Huang H, Yang N, Li Y, Wang J, Yao X, Shi Y, et al. 2011. Structure of a CENP-A—histone H4 heterodimer in complex with chaperone HJURP. *Genes Dev.* 25(9):901–906. doi:10.1101/gad.2045111.
- Jagannathan M, Cummings R, Yamashita YM. 2019. The modular mechanism of chromocenter formation in *Drosophila*. *eLife.* 8: e43938. doi:10.7554/eLife.43938.
- Jäger H, Rauch M, Heidmann S. 2005. The *Drosophila melanogaster* condensin subunit cap-G interacts with the centromere-specific histone H3 variant CID. *Chromosoma.* 113(7):350–361. doi:10.1007/s00412-004-0322-4.
- Jansen LET, Black BE, Foltz DR, Cleveland DW. 2007. Propagation of centromeric chromatin requires exit from mitosis. *J Cell Biol.* 176:795–805. doi:10.1083/jcb.200701066.
- Jones KW, Robertson FW. 1970. Localisation of reiterated nucleotide sequences in *Drosophila* and mouse by in situ hybridisation of complementary RNA. *Chromosoma.* 31(3):331–345. doi:10.1007/BF00321229.
- Karpen GH, Allshire RC. 1997. The case for epigenetic effects on centromere identity and function. *Trends Genet.* 13:489–496. doi:10.1016/S0168-9525(97)01298-5.
- Karpen GH, Spradling AC. 1990. Reduced DNA polytenization of a minichromosome region undergoing position-effect variegation in *Drosophila*. *Cell.* 63(1):97–107. doi:10.1016/0092-8674(90)90291-L.
- Kato H, Jiang J, Zhou B-R, Rozendaal M, Feng H, Ghirlando R, Xiao TS, Straight AF, Bai Y. 2013. A conserved mechanism for centromeric nucleosome recognition by centromere protein CENP-C. *Science.* 340(6136):1110–1113. doi:10.1126/science.1235532.
- Kharchenko PV, Alekseyenko AA, Schwartz YB, Minoda A, Riddle NC, Ernst J, Sabo PJ, Larschan E, Gorchakov AA, Gu T, et al. 2011. Comprehensive analysis of the chromatin landscape in *Drosophila melanogaster*. *Nature.* 471(7339):480–486. doi:10.1038/nature09725.
- Kursel LE, Malik HS. 2017. Recurrent gene duplication leads to diverse repertoires of centromeric histones in *Drosophila* species. *Mol Biol Evol.* 34(6):1445–1462. doi:10.1093/molbev/msx091.
- Le M, Duricka D, Karpen GH. 1995. Islands of complex DNA are widespread in *Drosophila* centric heterochromatin. *Genetics.* 141(1):283–303. doi:10.1093/genetics/141.1.283.
- Lidsky PV, Sprenger F, Lehner CF. 2013. Distinct modes of centromere protein dynamics during cell cycle progression in *Drosophila* S2R+ cells. *J Cell Sci.* 126(Pt 20):4782–4793. doi:10.1242/jcs.134122.
- Liu Y, Petrovic A, Rombaut P, Mosalaganti S, Keller J, Raunser S, Herzog F, Musacchio A. 2016. Insights from the reconstitution of the divergent outer kinetochore of *Drosophila melanogaster*. *Open Biol.* 6(2):150236. doi:10.1098/rsob.150236.
- Lohe AR, Hillikert AJ, Roberts PA. 1993. Mapping simple repeated DNA sequences in heterochromatin of *Drosophila melanogaster*. *Genetics.* 134(4):1149–1174. doi:10.1093/genetics/134.4.1149.
- Maggert KA, Karpen GH. 2001. The activation of a neocentromere in *Drosophila* requires proximity to an endogenous centromere. *Genetics.* 158(4):1615–1628. doi:10.1093/genetics/158.4.1615.
- Maiato H, Hergert PJ, Moutinho-Pereira S, Dong Y, Vandenbeldt KJ, Rieder CL, McEwen BF. 2006. The ultrastructure of the kinetochore and kinetochore fiber in *Drosophila* somatic cells. *Chromosoma.* 115(6):469–480. doi:10.1007/s00412-006-0076-2. [The](#).
- Malik HS, Henikoff S. 2001. Adaptive evolution of Cid, a centromere-specific histone in *Drosophila*. *Genetics.* 157(3):1293–1298. doi:10.1093/genetics/157.3.1293.
- Malik HS, Vermaak D, Henikoff S. 2002. Recurrent evolution of DNA-binding motifs in the *Drosophila* centromeric histone. *Proc Natl Acad Sci U S A.* 99(3):1449–1454. doi:10.1073/pnas.032664299.
- Manuelidis L. 1978. Chromosomal localization of complex and simple repeated human DNAs. *Chromosoma.* 66(1):23–32. doi:10.1007/BF00285813.
- Marshall OJ, Marshall AT, Choo KHA. 2008. Three-dimensional localization of CENP-A suggests a complex higher order structure of centromeric chromatin. *J Cell Biol.* 183(7):1193–1202. doi:10.1083/jcb.200804078.
- Mathew V, Pauleau AL, Steffen N, Bergner A, Becker PB, Erhardt S. 2014. The histone-fold protein CHRAC14 influences chromatin composition in response to DNA damage. *Cell Rep.* 7:321–330. doi:10.1016/j.celrep.2014.03.008.
- McKinley KL, Cheeseman IM. 2014. Polo-like kinase 1 licenses CENP-A deposition at centromeres. *Cell.* 158:397–411. doi:10.1016/j.cell.2014.06.016.
- McNulty SM, Sullivan LL, Sullivan BA. 2017. Human centromeres produce chromosome-specific and array-specific alpha satellite transcripts that are complexed with CENP-A and CENP-C. *Dev Cell.* 42(3):226–240.e6. doi:10.1016/j.devcel.2017.07.001.
- Medina-Pritchard B, Lazou V, Zou J, Byron O, Abad MA, Rappsilber J, Heun P, Jeyaprakash AA. 2020. Structural basis for centromere maintenance by *Drosophila* CENP-A chaperone CAL 1. *EMBO J.* 39(7):e103234. doi:10.15252/embj.2019103234.
- Mellone BG, Grive KJ, Shteyn V, Bowers SR, Oderberg I, Karpen GH. 2011. Assembly of *Drosophila* centromeric chromatin proteins during mitosis. *PLoS Genet.* 7:e1002068. doi:10.1371/journal.pgen.1002068.
- Meluh PB, Yang P, Glowczewski L, Koshland D, Smith MM. 1998. Cse4p is a component of the core centromere of *Saccharomyces cerevisiae*. *Cell.* 94(5):607–613. doi:10.1016/S0092-8674(00)81602-5.
- Méndez-Lago M, Wild J, Whitehead SL, Tracey A, de Pablos B, de Rogers J, Szybalski W, Villasante A. 2009. Novel sequencing strategy for repetitive DNA in a *Drosophila* BAC clone reveals that the centromeric region of the Y chromosome evolved from a telomere. *Nucleic Acids Res.* 37(7):2264–2273. doi:10.1093/nar/gkp085.
- Mendiburo MJ, Padeken J, Fülöp S, Schepers A, Heun P. 2011. *Drosophila* CENH3 is sufficient for centromere formation. *Science.* 334(6056):686–690. doi:10.1126/science.1206034.
- Meraldi P, McAinsh AD, Rheinbay E, Sorger PK. 2006. Phylogenetic and structural analysis of centromeric DNA and kinetochore proteins. *Genome Biol.* 7(3):R23. doi:10.1186/gb-2006-7-3-r23.
- Miell MDD, Fuller CJ, Guse A, Barysz HM, Downes A, Owen-Hughes T, Rappsilber J, Straight AF, Allshire RC. 2013. CENP-A confers a reduction in height on octameric nucleosomes. *Nat Struct Mol Biol.* 20(6):763–765. doi:10.1038/nsmb.2574.
- ModENCODE Consortium; Roy S, Ernst J, Kharchenko PV, Kheradpour P, Negre N, Eaton ML, Landolin JM, Bristow CA, Ma L, et al. 2011. Identification of functional elements and regulatory circuits by *Drosophila* ModENCODE. *Science.* 330(6012):1787–1797. doi:10.1126/science.1198374. [Identification](#).
- Moreno-Moreno O, Medina-Giró S, Torras-Llort M, Azorín F. 2011. The F box protein partner of paired regulates stability of *Drosophila* centromeric histone H3, CenH3 CID. *Curr Biol.* 21(17):1488–1493. doi:10.1016/j.cub.2011.07.041.
- Moreno-Moreno O, Torras-Llort M, Azorín F. 2006. Proteolysis restricts localization of CID, the centromere-specific histone H3 variant of *Drosophila*, to centromeres. *Nucleic Acids Res.* 34(21):6247–6255. doi:10.1093/nar/gkl902.
- Murphy TD, Karpen GH. 1995. Localization of centromere function in a *Drosophila* minichromosome. *Cell.* 82:599–609. doi:10.1016/0092-8674(95)90032-2.

- Murphy TD, Karpen GH. 1998. Centromeres take flight: alpha satellite and the quest for the human centromere. *Cell*. 93:317–320. doi:[10.1016/S0092-8674\(00\)81158-7](https://doi.org/10.1016/S0092-8674(00)81158-7).
- Nakayama KI, Nakayama K. 2006. Ubiquitin ligases: cell-cycle control and cancer. *Nat Rev Cancer*. 6(5):369–381. doi:[10.1038/nrc1881](https://doi.org/10.1038/nrc1881).
- Olszak AM, van Essen D, Pereira AJ, Diehl S, Manke T, Maiato H, Saccani S, Heun P. 2011. Heterochromatin boundaries are hotspots for de novo kinetochore formation. *Nat Cell Biol*. 13(7):799–808. doi:[10.1038/ncb2272](https://doi.org/10.1038/ncb2272).
- Orr B, Sunkel CE. 2011. *Drosophila* CENP-C is essential for centromere identity. *Chromosoma*. 120(1):83–96. doi:[10.1007/s00412-010-0293-6](https://doi.org/10.1007/s00412-010-0293-6).
- Ortiz-Lombardía M, Cortés A, Huertas D, Eritja R, Azorín F. 1998. Tandem 5'-GA:GA-3' mismatches account for the high stability of the fold-back structures formed by the centromeric *Drosophila* dodeca-satellite. *J Mol Biol*. 277(4):757–762. doi:[10.1006/jmbi.1998.1646](https://doi.org/10.1006/jmbi.1998.1646).
- Palladino J, Chavan A, Sposato A, Mason TD, Mellone BG. 2020. Targeted de novo centromere formation in *Drosophila* reveals plasticity and maintenance potential of CENP-A chromatin. *Dev Cell*. 52(3):379–394.e7. doi:[10.1016/j.devcel.2020.01.005](https://doi.org/10.1016/j.devcel.2020.01.005).
- Palmer DK, O'Day K, Trong HL, Charbonneau H, Margolis RL. 1991. Purification of the centromere-specific protein CENP-A and demonstration that it is a distinctive histone. *Proc Natl Acad Sci U S A*. 88(9):3734–3738. doi:[10.1073/pnas.88.9.3734](https://doi.org/10.1073/pnas.88.9.3734).
- Palmer DK, O'Day K, Wener MH, Andrews BS, Margolis RL. 1987. A 17-kD centromere protein (CENP-A) copurifies with nucleosome core particles and with histones. *J Cell Biol*. 104(4):805–815. doi:[10.1083/jcb.104.4.805](https://doi.org/10.1083/jcb.104.4.805).
- Pardue M, Gall JG. 1970. Chromosomal localization of mouse satellite DNA. *Science*. 168(3937):1356–1358. doi:[10.1126/science.168.3937.1356](https://doi.org/10.1126/science.168.3937.1356).
- Paris C, Geinguenaud F, Gouyette C, Liquier J, Lacoste J. 2007. Mechanism of copper mediated triple helix formation at neutral PH in *Drosophila* satellite repeats. *Biophysical J*. 92(7):2498–2506. doi:[10.1529/biophysj.106.099788](https://doi.org/10.1529/biophysj.106.099788).
- Pauleau AL, Bergner A, Kajtez J, Erhardt S. 2019. The checkpoint protein Zw10 connects CAL1-dependent CENP-A centromeric loading and mitosis duration in *Drosophila* cells. *PLoS Genet*. 15(9). doi:[10.1371/journal.pgen.1008380](https://doi.org/10.1371/journal.pgen.1008380).
- Pesenti ME, Raisch T, Conti D, Walstein K, Hoffmann I, Vogt D, Prumbaum D, Vetter IR, Raunser S, Musacchio A. 2022. Structure of the human inner kinetochore CCAN complex and its significance for human centromere organization. *Mol Cell*. 82:2113–2131.e8. doi:[10.1016/j.molcel.2022.04.027](https://doi.org/10.1016/j.molcel.2022.04.027).
- Phansalkar R, Lapierre P, Mellone BG. 2012. Evolutionary insights into the role of the essential centromere protein CAL1 in *Drosophila*. *Chromosome Res*. 20:493–504. doi:[10.1007/s10577-012-9299-7](https://doi.org/10.1007/s10577-012-9299-7).
- Platero JS, Ahmad K, Henikoff S. 1999. A distal heterochromatic block displays centromeric activity when detached from a natural centromere. *Mol Cell*. 4(6):995–1004. doi:[10.1016/S1097-2765\(00\)80228-2](https://doi.org/10.1016/S1097-2765(00)80228-2).
- Platero JS, Csink AK, Quintanilla A, Henikoff S. 1998. Changes in chromosomal localization of heterochromatin-binding proteins during the cell cycle in *Drosophila*. *J Cell Biol*. 140(6):1297–1306. doi:[10.1083/jcb.140.6.1297](https://doi.org/10.1083/jcb.140.6.1297).
- Przewloka MR, Venkei Z, Bolanos-Garcia VM, Debski J, Dadlez M, Glover DM. 2011. CENP-C is a structural platform for kinetochore assembly. *Curr Biol*. 21:399–405. doi:[10.1016/j.cub.2011.02.005](https://doi.org/10.1016/j.cub.2011.02.005).
- Przewloka MR, Zhang W, Costa P, Archambault V, D'Avino PP, Lilley KS, Laue ED, McAinsh AD, Glover DM. 2007. Molecular analysis of core kinetochore composition and assembly in *Drosophila melanogaster*. *PLoS One*. 2(5):e478. doi:[10.1371/journal.pone.0000478](https://doi.org/10.1371/journal.pone.0000478).
- Rae PMM. 1970. Chromosomal distribution of rapidly reannealing DNA in *Drosophila melanogaster*. *Proc Natl Acad Sci U S A*. 67(2):1018–1025. doi:[10.1073/pnas.67.2.1018](https://doi.org/10.1073/pnas.67.2.1018).
- Raychaudhuri N, Dubruielle R, Orsi GA, Bagheri HC, Loppin B, Lehner CF. 2012. Transgenerational propagation and quantitative maintenance of paternal centromeres depends on Cid/Cenp-A presence in *Drosophila* sperm. *PLoS Biol*. 10(12):e1001434. doi:[10.1371/journal.pbio.1001434](https://doi.org/10.1371/journal.pbio.1001434).
- Ribeiro SA, Vagnarelli P, Dong Y, Hori T, McEwen BF, Fukagawa T, Flors C, Earnshaw WC. 2010. A super-resolution map of the vertebrate kinetochore. *Proc Natl Acad Sci U S A*. 107(23):10484–10489. doi:[10.1073/pnas.1002325107](https://doi.org/10.1073/pnas.1002325107).
- Rošić S, Köhler F, Erhardt S. 2014. Repetitive centromeric satellite RNA is essential for kinetochore formation and cell division. *J Cell Biol*. 207(3):335–349. doi:[10.1083/jcb.201404097](https://doi.org/10.1083/jcb.201404097).
- Rosin L, Mellone BG. 2016. Co-evolving CENP-A and CAL1 domains mediate centromeric CENP-A deposition across *Drosophila* species. *Dev Cell*. 37:136–147. doi:[10.1016/j.devcel.2016.03.021](https://doi.org/10.1016/j.devcel.2016.03.021).
- Rosin LF, Mellone BG. 2017. Centromeres drive a hard bargain. *Trends Genet*. 33:101–117. doi:[10.1016/j.tig.2016.12.001](https://doi.org/10.1016/j.tig.2016.12.001).
- Roure V, Medina-Pritchard B, Lazou V, Rago L, Anselm E, Venegas D, Jeyaprakash AA, Heun P. 2019. Reconstituting *Drosophila* centromere identity in human cells. *Cell Rep*. 29(2):464–479.e5. doi:[10.1016/j.celrep.2019.08.067](https://doi.org/10.1016/j.celrep.2019.08.067).
- Sainz A, Wilder JA, Wolf M, Hollocher H. 2003. *Drosophila melanogaster* and *D simulans* rescue strains produce fit offspring, despite divergent centromere-specific histone alleles. *Heredity (Edinb)*. 91(1):28–35. doi:[10.1038/sj.hdy.6800275](https://doi.org/10.1038/sj.hdy.6800275).
- Sandler L, Novitski E. 1957. Meiotic drive as an evolutionary force source. *Am Naturalist*. 91(857):105–110. doi:[10.1086/281969](https://doi.org/10.1086/281969).
- Sart D, Cancelli MR, Earle E, Mao J, Saffery R, Tainton KM, Kalitsis P, Martyn J, Barry AE, Choo KHA. 1997. A functional neo-centromere formed through activation of a latent human centromere and consisting of non-alpha-satellite DNA. *Nat Genet*. 16(2):144–153. doi:[10.1038/ng0697-144](https://doi.org/10.1038/ng0697-144).
- Schittenhelm RB, Althoff F, Heidmann S, Lehner CF. 2010. Detrimental incorporation of excess Cenp-A/Cid and Cenp-C into *Drosophila* centromeres is prevented by limiting amounts of the bridging factor cal1. *J Cell Sci*. 123(21):3768–3779. doi:[10.1242/jcs.067934](https://doi.org/10.1242/jcs.067934).
- Schittenhelm RB, Heeger S, Althoff F, Walter A, Heidmann S, Mechtler K, Lehner CF. 2007. Spatial organization of a ubiquitous eukaryotic kinetochore protein network in *Drosophila* chromosomes. *Chromosoma*. 116:385–402. doi:[10.1007/s00412-007-0103-y](https://doi.org/10.1007/s00412-007-0103-y).
- Schuh M, Lehner CF, Heidmann S. 2007. Incorporation of *Drosophila* CID/CENP-A and CENP-C into centromeres during early embryonic anaphase. *Curr Biol*. 17:237–243. doi:[10.1016/j.cub.2006.11.051](https://doi.org/10.1016/j.cub.2006.11.051).
- Sekulic N, Bassett EA, Rogers DJ, Black BE. 2010. The structure of (CENP-A-H4)₂ reveals physical features that mark centromeres. *Nature*. 467(7313):347–351. doi:[10.1038/nature09323](https://doi.org/10.1038/nature09323).
- Shukla M, Tong P, White SA, Singh PP, Reid AM, Catania S, Pidoux AL, Allshire RC. 2018. Centromere DNA destabilizes H3 nucleosomes to promote CENP-A deposition during the cell cycle. *Curr Biol*. 28(24):3924–3936.e4. doi:[10.1016/j.cub.2018.10.049](https://doi.org/10.1016/j.cub.2018.10.049).
- Stoler S, Keith KC, Curnick KE, Fitzgerald-Hayes M. 1995. A mutation in CSE4, an essential gene encoding a novel chromatin-associated protein in yeast, causes chromosome nondisjunction and cell cycle arrest at mitosis. *Genes Dev*. 9(5):573–586. doi:[10.1101/gad.9.5.573](https://doi.org/10.1101/gad.9.5.573).

- Sullivan LL, Boivin CD, Mravinac B, Song IY, Sullivan BA. 2011. Genomic size of CENP-A domain is proportional to total alpha satellite array size at human centromeres and expands in cancer cells. *Chromosome Res.* 19:457–470. doi:[10.1007/s10577-011-9208-5](https://doi.org/10.1007/s10577-011-9208-5).
- Sullivan B, Karpen G. 2001. Centromere identity in *Drosophila* is not determined in vivo by replication timing. *J Cell Biol.* 154(4): 683–690. doi:[10.1083/jcb.200103001](https://doi.org/10.1083/jcb.200103001).
- Sullivan BA, Karpen GH. 2004. Centromeric chromatin exhibits a histone modification pattern that is distinct from both euchromatin and heterochromatin. *Nat Struct Mol Biol.* 11:1076–1083. doi:[10.1038/nsmb845](https://doi.org/10.1038/nsmb845).
- Sullivan BA, Willard HF. 1998. Stable dicentric X chromosomes with two functional centromeres. *Nat Genet.* 20(3):227–228. doi:[10.1038/3024](https://doi.org/10.1038/3024).
- Sun X, Le HD, Wahlstrom JM, Karpen GH. 2003. Sequence analysis of a functional *Drosophila* centromere. *Genome Res.* 13(2):82–94. doi:[10.1101/gr.681703](https://doi.org/10.1101/gr.681703).
- Sun X, Wahlstrom J, Karpen G. 1997. Molecular structure of a functional *Drosophila* centromere. *Cell.* 91:1007–1019. doi:[10.1016/S0092-8674\(00\)80491-2](https://doi.org/10.1016/S0092-8674(00)80491-2).
- Talbert PB, Kasinathan S, Henikoff S. 2018. Simple and complex centromeric satellites in *Drosophila* sibling species. *Genetics.* 208(3):977–990. doi:[10.1534/genetics.117.300620](https://doi.org/10.1534/genetics.117.300620).
- Teixeira JR, Dias GB, Svartman M, Ruiz A, Kuhn GCS. 2018. Concurrent duplication of *Drosophila* Cid and Cenp-C genes resulted in accelerated evolution and male germline-biased expression of the new copies. *J Mol Evol.* 86(6):353–364. doi:[10.1007/s00239-018-9851-y](https://doi.org/10.1007/s00239-018-9851-y).
- Török T, Gorjánác M, Bryant PJ, Kiss I. 2000. Prod is a novel DNA-binding protein that binds to the 1.686 g/cm³ 10 bp satellite repeat of *Drosophila melanogaster*. *Nucleic Acids Res.* 28(18): 3551–3557. doi:[10.1093/nar/28.18.3551](https://doi.org/10.1093/nar/28.18.3551).
- Turner BM, Birley AJ, Lavender J. 1992. Histone H4 isoforms acetylated at specific lysine residues define individual chromosomes and chromatin domains in *Drosophila* polytene nuclei. *Cell.* 69: 375–384. doi:[10.1016/0092-8674\(92\)90417-b](https://doi.org/10.1016/0092-8674(92)90417-b).
- Vermaak D, Hayden HS, Henikoff S. 2002. Centromere targeting element within the histone fold domain of Cid. *Mol Cell Biol.* 22(21):7553–7561. doi:[10.1128/MCB.22.21.7553-7561.2002](https://doi.org/10.1128/MCB.22.21.7553-7561.2002).
- Voullaire LE, Slater HR, Petrovic V, Choo KHA. 1993. A functional marker centromere with no detectable alpha-satellite, satellite ill, or CENP-B protein: activation of a latent centromere? *Am J Hum Genet.* 52:1153–1163.
- Wang H, Dalal Y, Henikoff S, Lindsay S. 2008. Single-epitope recognition imaging of native chromatin. *Epigenetics Chromatin.* 1(1):10. doi:[10.1186/1756-8935-1-10](https://doi.org/10.1186/1756-8935-1-10).
- Wei RR, Al-Bassam J, Harrison SC. 2007. The Ndc80/HEC1 complex is a contact point for kinetochore-microtubule attachment. *Nat Struct Mol Biol.* 14(1):54–59. doi:[10.1038/nsmb1186](https://doi.org/10.1038/nsmb1186).
- Williams BC, Murphy TD, Goldberg ML, Karpen GH. 1998. Neocentromere activity of structurally acentric mini-chromosomes in *Drosophila*. *Nat Genet.* 18(1):231–236. doi:[10.1038/ng0598-51](https://doi.org/10.1038/ng0598-51).
- Xiao H, Wang F, Wisniewski J, Shaytan AK, Ghirlando R, Fitzgerald PC, Huang Y, Wei D, Li S, Landsman D, et al. 2017. Molecular basis of CENP-C association with the CENP-A nucleosome at yeast centromeres. *Genes Dev.* 31(19):1958–1972. doi:[10.1101/gad.304782.117](https://doi.org/10.1101/gad.304782.117).
- Yatskevich S, Muir KW, Bellini D, Zhang Z, Yang J, Tischer T, Predin M, Dendooven T, McLaughlin SH, Barford D. 2022. Structure of the human inner kinetochore bound to a centromeric CENP-A nucleosome. *Science.* 376(6595):844–852. doi:[10.1126/science.abn3810](https://doi.org/10.1126/science.abn3810).
- Zhang W, Colmenares SU, Karpen GH. 2012. Assembly of *Drosophila* centromeric nucleosomes requires CID dimerization. *Mol Cell.* 45:263–269. doi:[10.1016/j.molcel.2011.12.010](https://doi.org/10.1016/j.molcel.2011.12.010).
- Zinkowski RP, Meyne J, Brinkley BR. 1991. The centromere-kinetochore complex: a repeat subunit model. *J Cell Biol.* 113: 1091–1110. doi:[10.1083/jcb.113.5.1091](https://doi.org/10.1083/jcb.113.5.1091).
- Zwick ME, Salstrom JL, Langley CH. 1999. Genetic variation in rates of nondisjunction: association of two naturally occurring polymorphisms in the chromokinesin nod with increased rates of nondisjunction in *Drosophila melanogaster*. *Genetics.* 152: 1605–1614. doi:[10.1093/genetics/152.4.1605](https://doi.org/10.1093/genetics/152.4.1605).