

Recurrent loss of CenH3 is associated with independent transitions to holocentricity in insects

Ines A Drinnenberg¹, Dakota deYoung², Steven Henikoff^{1*}, Harmit Singh Malik^{1,3*}

¹Division of Basic Sciences, Fred Hutchinson Cancer Research Center, Seattle, United States; ²Department of Biology, University of Washington, Seattle, United States; ³Howard Hughes Medical Institute, Fred Hutchinson Cancer Research Center, Seattle, United States

Abstract Faithful chromosome segregation in all eukaryotes relies on centromeres, the chromosomal sites that recruit kinetochore proteins and mediate spindle attachment during cell division. The centromeric histone H3 variant, CenH3, is the defining chromatin component of centromeres in most eukaryotes, including animals, fungi, plants, and protists. In this study, using detailed genomic and transcriptome analyses, we show that CenH3 was lost independently in at least four lineages of insects. Each of these lineages represents an independent transition from monocentricity (centromeric determinants localized to a single chromosomal region) to holocentricity (centromeric determinants extended over the entire chromosomal length) as ancient as 300 million years ago. Holocentric insects therefore contain a CenH3-independent centromere, different from almost all the other eukaryotes. We propose that ancient transitions to holocentricity in insects obviated the need to maintain CenH3, which is otherwise essential in most eukaryotes, including other holocentrics.

DOI: 10.7554/eLife.03676.001

*For correspondence: steveh@fhcrc.org (SH); hsmalik@fhcrc.org (HSM)

Competing interests: The authors declare that no competing interests exist.


Funding: See page 14

Received: 13 June 2014

Accepted: 25 August 2014

Published: 23 September 2014

Reviewing editor: Anthony A Hyman, Max Planck Institute of Molecular Cell Biology and Genetics, Germany

 Copyright Drinnenberg et al. This article is distributed under the terms of the [Creative Commons Attribution License](#), which permits unrestricted use and redistribution provided that the original author and source are credited.

Introduction

In eukaryotes, accurate chromosome segregation relies on specific chromosomal regions called centromeres that recruit components of the proteinaceous kinetochore complex to mediate spindle attachments and ensure high-fidelity segregation (*DeWulf and Earnshaw, 2008*). In animals, fungi, and plants, kinetochore assembly onto centromeres occurs in a hierarchical process that strictly depends on the presence of the specialized histone H3 variant CenH3 (first identified as Cenp-A in mammals [*Earnshaw and Rothfield, 1985; Palmer et al., 1991*]), which replaces the canonical H3 in centromeric nucleosomes (*Sullivan et al., 1994; Yoda et al., 2000*). In all organisms that have been studied, CenH3 deletions are lethal and lead to catastrophic defects in chromosome segregation (*Stoler et al., 1995; Buchwitz et al., 1999; Howman et al., 2000; Blower and Karpen, 2001; Talbert et al., 2002*). Moreover, the presence of CenH3 defines both canonical centromeres and neocentromeres in diverse organisms (*Dawson et al., 2007; Malik and Henikoff, 2009*). The presence of CenH3 homologs in all animals, fungi, and plants studied so far, together with their identification in distantly branching protist lineages has established the paradigm that CenH3-containing chromatin is an absolute requirement for centromere function (*Malik and Henikoff, 2003; Panchenko and Black, 2009*). In contrast to this prevailing paradigm, we show that multiple insect lineages have independently lost CenH3 despite preserving some other canonical kinetochore components. These CenH3 losses coincide with dramatic changes in centromere architecture, suggesting that alternate centromere configurations may render the CenH3 protein dispensable.

eLife digest Cell division is a fundamentally important process for living organisms. In eukaryotes, such as plants and animals, genomic DNA is tightly packaged into chromosomes, which needs to be copied and faithfully divided into daughter cells. Segregating the chromosomes is accomplished by the kinetochore, a protein complex that assembles on the chromosome and forms attachments to the machinery that provides the force for chromosome segregation. Kinetochores assemble on specialized chromosomal regions called centromeres. In most eukaryotes, kinetochore assembly relies on a centromeric protein called CenH3 that is essential for the process of chromosome segregation.

Most animal and plant species are monocentric—one part of the chromosome is dedicated to CenH3 loading and centromere function and paired chromosomes appear to be joined at a single point, or primary constriction. In contrast, holocentric species instead have centromeric activity distributed along the entire length of the paired chromosomes. How holocentricity arose from monocentricity over the course of evolution remains unclear.

Drinnenberg et al. took advantage of the fact that insects represent at least four independent transitions from monocentric to holocentric chromosomes. Several species of insects are holocentric—including butterflies and moths, bugs and lice, earwigs, and dragonflies—while others are monocentric—such as flies, bees, and beetles.

Drinnenberg et al. compared the repertoire of kinetochore proteins from each of these insect lineages and found that CenH3 was absent in all the holocentric insects examined but present in all the monocentric insect species. Despite the loss of CenH3 in the holocentric insects, they still had many of the kinetochore proteins—particularly, those proteins that attach to the machinery that forces chromosomes apart. Based on an evolutionary reconstruction, Drinnenberg et al. infer that each independent transition to holocentricity in insects likely introduced changes to the centromere that eliminated the need for the otherwise essential CenH3 protein.

This study challenges the notion that CenH3 is essential in all eukaryotes. Indeed, holocentric insects, which make-up 16% of the biodiversity of the currently known eukaryote species, appear to have evolved a completely novel way to define their centromeres, distinct from all the other eukaryotes.

DOI: [10.7554/eLife.03676.002](https://doi.org/10.7554/eLife.03676.002)

Despite the universality of CenH3 in eukaryotes, centromeres are remarkably diverse (**Malik and Henikoff, 2009**). Most eukaryotic chromosomes are monocentric, that is, centromeres and kinetochore assembly are restricted to a defined chromosomal region. Monocentromeres can range dramatically in size, from 125 bp in budding yeasts to megabases in humans, and can be either genetically (sequence-dependent) or epigenetically defined. In contrast, holocentromeres have kinetochores attached along the extensive segments or even the entire length of chromosomes. First described by **Schrader (1935)**, holocentromeres have been best studied in the nematode *Caenorhabditis elegans* (**Dernburg, 2001; Maddox et al., 2004; Gassmann et al., 2012; Steiner and Henikoff, 2014**). However, holocentricity appears to have evolved independently in multiple eukaryotic lineages by convergent evolution (**Melters et al., 2012**).

Results and discussion

To gain insight into kinetochore changes associated with transitions to holocentricity, we focused on insects in which holocentricity is believed to have evolved at least four times from monocentric ancestors (**Figure 1A, Figure 1—figure supplement 1**) (**Whiting, 2002; Grimaldi and Engel, 2005; Savard et al., 2006**) (modified from **Melters et al., 2012**). The strength of evidence about holocentricity varies among different insect orders; we briefly summarize this evidence in the ‘Materials and methods’ section. For instance, a strong consensus has emerged from many different species that both Lepidoptera and Hemiptera represent holocentric insect orders. In contrast, there is relatively modest evidence for holocentricity in Dermaptera, Odonata, and Phthiraptera. Nevertheless, the currently held consensus view is that each of the insect orders indicated in blue in **Figure 1A** is holocentric.

Using homology searches, we analyzed all the sequenced (monocentric and holocentric) insect genomes to identify protein components of the inner kinetochore complex. Initially, we focused on CenH3

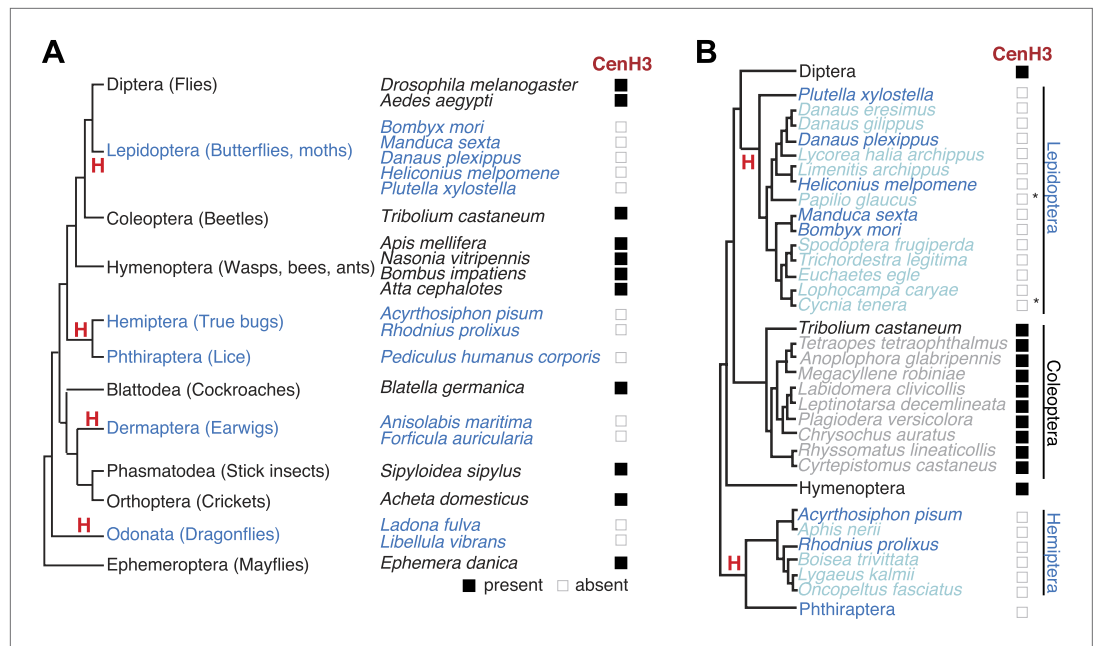


Figure 1. Holocentric insects lack CenH3. **(A)** Phylogeny of insect orders (species) examined in this study. Holocentric insect orders are indicated in blue, and inferred multiple transitions to holocentricity in insects are labeled with 'H'. Using protein homology searches of genomes or assembled transcriptomes, we inferred either the presence (black box) or absence (empty box) of CenH3. **(B)** CenH3 loss is widespread in Lepidoptera and Hemiptera, but not Coleoptera. Phylogenetic relationship of holocentric insects used for transcriptome assemblies (light blue), holocentric insects with sequenced genomes (blue), monocentric insects used for transcriptome assemblies (gray), and monocentric insects with sequenced genomes (black). The presence of contaminating microsporidian CenH3 transcripts is indicated with an asterisk.

DOI: [10.7554/eLife.03676.003](https://doi.org/10.7554/eLife.03676.003)

The following source data and figure supplements are available for figure 1:

Source data 1. List of species with sequenced genomes and information to their corresponding analyzed proteomes.

DOI: [10.7554/eLife.03676.004](https://doi.org/10.7554/eLife.03676.004)

Source data 2. Statistics of mRNA-Seq assemblies.

DOI: [10.7554/eLife.03676.005](https://doi.org/10.7554/eLife.03676.005)

Figure supplement 1. Evolution of holocentricity in insects.

DOI: [10.7554/eLife.03676.006](https://doi.org/10.7554/eLife.03676.006)

Figure supplement 2. Insect CenH3 comparison.

DOI: [10.7554/eLife.03676.007](https://doi.org/10.7554/eLife.03676.007)

Figure supplement 3. Additional H3-like variants in holocentric insects.

DOI: [10.7554/eLife.03676.008](https://doi.org/10.7554/eLife.03676.008)

Figure supplement 4. Fungal CenH3 transcript contaminants in the cockroach assembly.

DOI: [10.7554/eLife.03676.009](https://doi.org/10.7554/eLife.03676.009)

Figure supplement 5. CenH3 transcript abundance in mRNA-Seq assemblies.

DOI: [10.7554/eLife.03676.010](https://doi.org/10.7554/eLife.03676.010)

Figure supplement 6. CenH3 phylogeny in insects.

DOI: [10.7554/eLife.03676.011](https://doi.org/10.7554/eLife.03676.011)

Figure supplement 7. CenH3 homologs in Coleoptera (beetles).

DOI: [10.7554/eLife.03676.012](https://doi.org/10.7554/eLife.03676.012)

Figure supplement 8. Contaminating microsporidian CenH3 transcripts in the two lepidopteran assemblies.

DOI: [10.7554/eLife.03676.013](https://doi.org/10.7554/eLife.03676.013)

Figure supplement 9. Chromosome spreads from various lepidopteran species generously provided by Frantisek Marec and Atsuo Yoshido.

DOI: [10.7554/eLife.03676.014](https://doi.org/10.7554/eLife.03676.014)

that can be identified by its homology to histone H3, but distinguished from canonical H3 and other H3 histone variants via their distinct phylogenetic grouping and features such as a longer loop1 region and a highly divergent N-terminal tail (Blower and Karpen, 2001; Malik and Henikoff, 2003). These criteria have allowed the correct identification of CenH3 in all the eukaryotic genomes examined so far (Malik and Henikoff, 2003; Talbert and Henikoff, 2010). In addition to the well-studied CenH3 homologs in *Drosophila* species (Henikoff et al., 2000), we were able to identify CenH3 homologs in the monocentric insect orders: Diptera (flies and mosquitoes), Hymenoptera (wasps, bees, and ants), and Coleoptera (beetles) (Figure 1A, Figure 1—figure supplement 2). In contrast, we were unable to find CenH3 homologs in any of the five sequenced holocentric lepidopteran species (butterflies and moths) (Figure 1A). Extending our survey, we were similarly unable to find CenH3 in two additional insect orders previously reported to be holocentric: Hemiptera (true bugs) and Phthiraptera (lice). Together, these two orders represent an independent transition to holocentricity in insects (Figure 1A, 'Materials and methods') (Whiting, 2002). Although we found an additional H3-like gene in the hemipteran pea aphid *Acyrtosiphon pisum*, we found that it unambiguously phylogenetically clustered with canonical H3 proteins, rather than CenH3s (Figure 1—figure supplement 3A,C). Thus, CenH3 is missing from all sequenced genomes from eight holocentric insect species, but present in all sequenced monocentric insects.

We wanted to test this correlation between holocentricity and the absence of CenH3 in other mono- and holocentric insect orders beyond those with sequenced genomes. To search for CenH3, we carried out mRNA-sequencing (mRNA-Seq) analyses and transcriptome assemblies of five insects, and combined our analyses with transcriptome assemblies of three additional available insect mRNA-Seq data sets. Our analyses yielded assemblies from insects of four monocentric orders, *Blattodea germanica* (Blattodea, cockroaches), *Acheta domesticus* (Orthoptera, crickets), *Sipylodea sipylus* (Phasmatodea, stick insects), and *Ephemera danica* (Ephemeroptera, mayflies) (<https://www.hgsc.bcm.edu/>). We were able to identify putative CenH3 homologs in the assemblies of all monocentric insects (Figure 1A, Figure 1—figure supplement 2). Identification of two bona fide CenH3 transcripts in the cockroach assembly led us to clearly attribute one of them to a fungal contamination via phylogenetic analyses (Figure 1—figure supplement 4) whereas the other was the cockroach CenH3 ortholog. We also obtained transcriptomes of three additional insects from two additional insect orders previously reported to be holocentric: *Anisolabis maritima* (Dermaptera, earwigs), *Libellula vibrans* and *Ladona fulva* (Odonata, dragonflies and damselflies) (<https://www.hgsc.bcm.edu/>). Additionally, we included a transcriptome assembly from *Forficula auricularia* (Dermaptera, earwigs) that had recently become available (Roulin et al., 2014). Together with Lepidoptera and Hemiptera/Phthiraptera, these new holocentric insect orders (Dermaptera, Odonata) represent all independent transitions to holocentricity in insects (Figure 1A, 'Materials and methods'). In contrast to the monocentric insects, we could not detect CenH3 in the assemblies derived from any of the holocentric insects. Although we did obtain an H3-like gene in the *L. vibrans* assembly, evolutionary analyses support its phylogenetic grouping with H3 rather than CenH3 (Figure 1—figure supplement 3B,C).

We considered the possibilities that either insufficient coverage of the transcriptome assembly and/or low expression of CenH3 led to our inability to detect CenH3 transcripts. To address the first possibility, we compared the probabilities of finding a CenH3 homolog between mono- and holocentric assemblies. Using the well-characterized proteome from *Tribolium castaneum* as a benchmark, we tested if similar numbers of proteins could be predicted in our assemblies (Figure 1—source data 2). Indeed, our analyses revealed comparable or even higher number of predicted proteins in the holocentric compared to the monocentric assemblies, implying at least equivalent transcriptome coverages of the holocentric and monocentric assemblies. Second, we found that CenH3 transcripts are not rare in monocentric insects; CenH3 transcript abundance was at least at the 50th percentile in all instances except the cockroach assembly (Figure 1—figure supplement 5). We therefore conclude that CenH3 is likely absent in the four holocentric odonatan and dermapteran insects examined as it is absent in the genome sequences of the holocentric lepidopteran and hemipteran/phthirapteran insects. Thus, we conclude that each of the four independent transitions to holocentricity in insects was associated with CenH3 loss.

Phylogenetic analyses of CenH3 and other H3 proteins based on their homologous histone-fold domains reveal a topology of insect CenH3s that is largely consistent with the expected branching order of the insect species (Whiting, 2002; Grimaldi and Engel, 2005; Savard et al., 2006) (Figure 1—figure supplement 6), confirming that the absence of CenH3 in holocentric insects is due to recurrent loss rather than reinvention or horizontal transfer of CenH3 in monocentric insect lineages.

To precisely date when CenH3 loss occurred in holocentric insects, we took advantage of additional transcriptome assemblies of mRNA-Seq data sets from nine coleopteran, nine lepidopteran, and four hemipteran species (Negre *et al.*, 2006; Zhen *et al.*, 2012) (<https://www.hgsc.bcm.edu/>). These additional assemblies confirmed not only CenH3 presence in all (monocentric) coleopteran species (Figure 1—figure supplement 7), but also confirmed its absence in all hemipteran and lepidopteran species. Although we obtained CenH3-like sequences in two lepidopteran assemblies, we could unambiguously attribute them to microsporidian contaminants via phylogenetic analyses (Figure 1—figure supplement 8A,B). As a result, we estimate that CenH3 loss probably preceded or occurred close to the emergence of these insect orders, at least 120 million years ago in Lepidoptera (Hedges *et al.*, 2006) and likely 300 million years ago in the common ancestor of Hemiptera and Phthiraptera (Grimaldi and Engel, 2005).

We next investigated the genome and transcriptome data sets to ask whether the onset of holocentricity also correlated with the loss of other inner kinetochore components, or CCAN (constitutive centromere associated network) components as previously defined in vertebrates and fungi (Hori *et al.*, 2008; Westermann and Schleiffer, 2013). We identified CenpI, CenpL/M/N, and the DNA-proximal CenpS/X proteins. We note that since the CenpT/W proteins, which are obligatory for the CenpS/X kinetochore function, are absent in all insects, it is likely that the CenpS/X complex serves roles in DNA damage (Singh *et al.*, 2010) rather than at the kinetochore. Nonetheless, several inner kinetochore components continue to be present even in the CenH3-deficient species (Figure 2A). This is somewhat unexpected since the recruitment of these components has been shown to depend on CenH3 in animals and fungi (Westhorpe and Straight, 2013). It will therefore be important in future to test if these kinetochore components localize to centromeres in the absence of CenH3; this would imply that holocentric insects have adopted a CenH3-independent inner kinetochore assembly pathway.

We did find the inner kinetochore protein CenpC to be absent from all insects that lack CenH3 (Figure 2A). CenpC is the direct DNA-binding partner of CenH3 and together with CenH3 constitutes the centromere core components (Carroll *et al.*, 2010). Although CenpC homologs are not as easily identifiable as CenH3, two motifs are conserved enough to allow cross-species homology searches. Of these, the CenpC motif is universally found in CenpC proteins and has been shown to interact with the CenH3 C-terminus (Kato *et al.*, 2013). In addition, most animal and fungal CenpC proteins contain a cupin fold domain (Pfam PF00190), about 100 amino acids downstream of the CenpC motif (Talbert *et al.*, 2004). Although cupin domains are not unique to CenpC proteins, CenpC-borne cupin domains can be easily distinguished from non-CenpC cupin domains based on phylogeny (Dunwell *et al.*, 2001). Homology searches using the CenpC motif revealed high-scoring CenpC-like motifs in most monocentric insects (HHpred p value range 2×10^{-9} to 4.6×10^{-22}) except in the *Tribolium castaneum* (genome) and cricket (transcriptome) (Figure 2A). We found that CenpC proteins from mayflies and stick insects (both monocentrics) possess both the CenpC motif and the associated cupin domain (Figure 2C). In contrast, putative CenpC proteins from Hymenoptera and cockroach (also monocentric) contained the CenpC motif but not the cupin domain. We were unable to find the CenpC motif in any of the holocentric species (genomes or transcriptomes). Intriguingly, although holocentric dragonflies encode a protein with a clearly identifiable CenpC cupin domain (Figure 2D, Figure 2—figure supplement 1), its CenpC motif has decayed (Figure 2C). We speculate that since the CenpC motif interacts with CenH3, CenH3 loss in dragonflies may have allowed the degeneration of its CenpC motif. The absence of CenH3-interacting CenpC motifs in dragonflies and other holocentric insects also serves as an independent confirmation for the losses of CenH3 in these lineages.

In contrast to the variation in inner kinetochore protein repertoires, homologs of outer kinetochore components including Ndc80 and Mis12 are almost universally conserved in most insect orders (Figure 2A). This suggests that although insects vary widely in the DNA-proximal part of their kinetochore, they have largely conserved the means to attach the kinetochore to the spindle apparatus. This further implies that holocentric insects likely employ CenH3-independent ways of connecting centromeric DNA to the rest of a relatively conserved kinetochore complex. Our findings in holocentric insects contrast with the recent findings of CenH3 absence in kinetoplastids, which not only lack CenH3 but also possess an entirely different repertoire of kinetochore proteins (Akiyoshi and Gull, 2013, 2014). Our findings are more reminiscent of male meiosis in holocentric *C. elegans*, in which the outer kinetochore components appear to localize independent of CenH3 (Moen *et al.*, 2005); CenH3

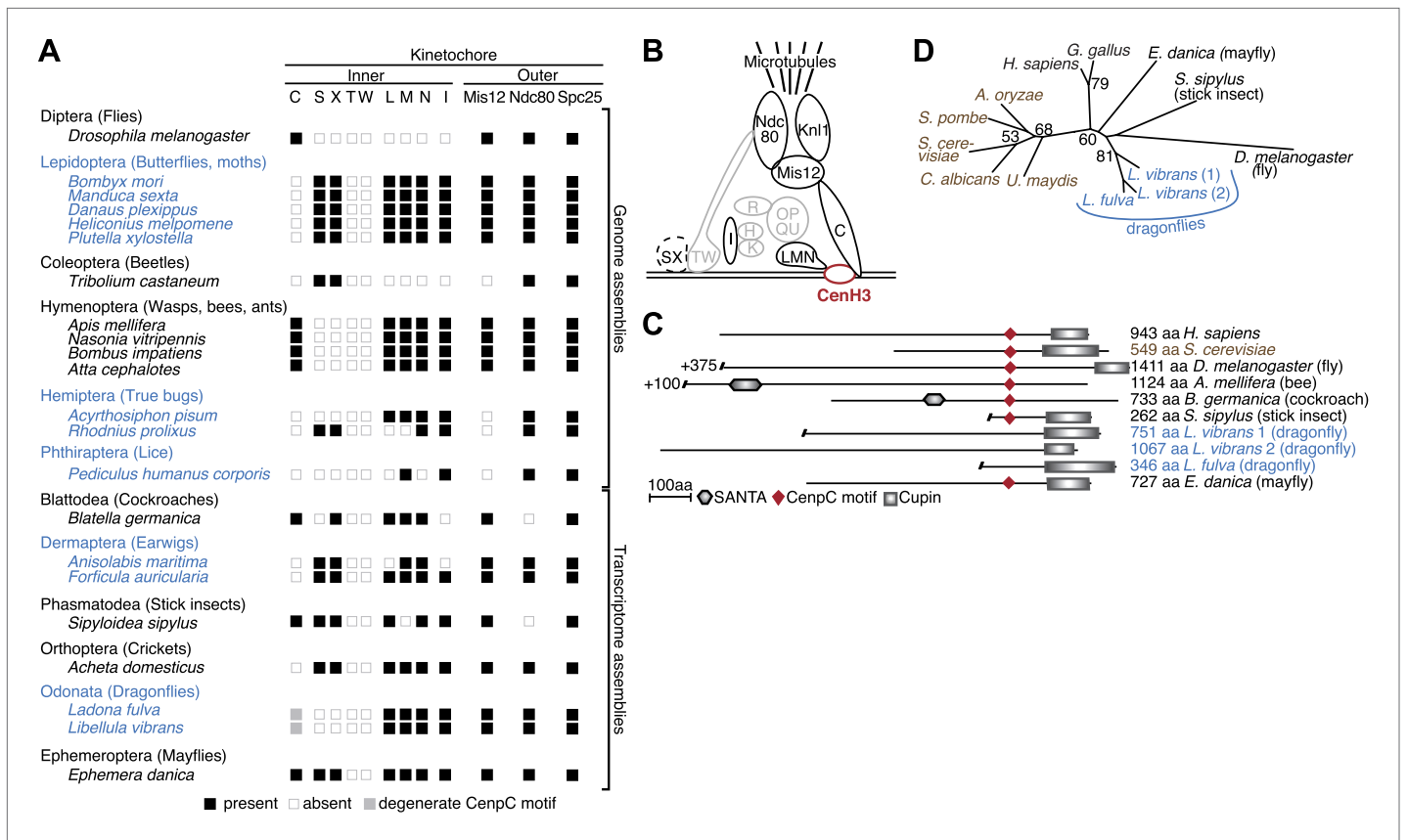


Figure 2. Evolution of the kinetochore composition in insects. **(A)** Kinetochores in insects. Using protein homology searches of the genome or assembled transcriptomes of monocentric (indicated in black) and holocentric insects (indicated in blue), we inferred either the presence (black box) or absence (empty box) of inner and outer kinetochore protein components. The presence of a putative CenpC homolog without a recognizable CenpC motif in Odonata is indicated by a filled gray box. We also found a weak match to a putative Mis12 homolog in *T. castaneum* (TC001997); however its weak homology relative to other insect Mis12 proteins leads us to assign this to be only a tentative match. **(B)** Schematic structure of a kinetochore largely based on its characterization in vertebrates and fungi (Hori et al., 2008; Westermann and Schleiffer, 2013). Components found in insects are highlighted in black, whereas components that are absent in all insects examined are in gray. The dashed line around the CenpS/X complex indicates that its kinetochore localization is unlikely in insects. **(C)** Putative CenpC proteins in insects. Schematics of structural domains and the CenpC motif of human, yeast, and insect CenpC proteins are shown. **(D)** Maximum likelihood tree of cupin domains of dragonfly (blue), fungal (brown), and other animal (black) CenpC proteins. Bootstrap percentages above 50 are indicated.

DOI: 10.7554/eLife.03676.015

The following source data and figure supplements are available for figure 2:

Source data 1. Accession numbers or sequences for all insect kinetochore proteins analyzed or described in this study.

DOI: 10.7554/eLife.03676.016

Figure supplement 1. CenpC-Cupin domain alignment.

DOI: 10.7554/eLife.03676.017

is nevertheless essential for *C. elegans* mitosis (Buchwitz et al., 1999). Thus, holocentric insects, which account for 16% of all named eukaryotic species (Grimaldi and Engel, 2005), appear to be unique among complex eukaryotes in entirely losing their dependence on CenH3s.

Concluding comments

Given the essential requirement of CenH3 for chromosome segregation in monocentric contexts as shown in *Drosophila melanogaster* (Blower and Karpen, 2001), we hypothesize that independent transitions to holocentricity preceded the losses of CenH3s in insects. However, since other holocentric animals (including other arthropods) and plants still encode for CenH3 (Buchwitz et al., 1999; Heckmann et al., 2011) (Figure 2—source data 1), we speculate that another, distinct event allowing CenH3 loss in holocentric insects must have occurred early in insect evolution, in the common ancestor of dragonflies and flies (Figure 3). Such a determinant may have been the lineage-specific evolutionary

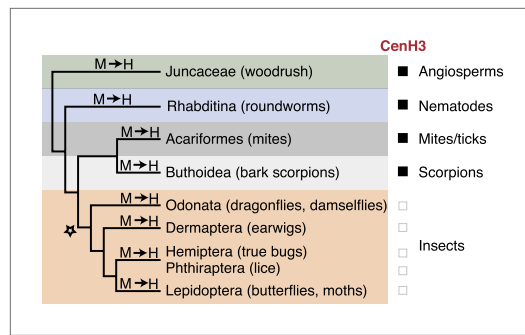


Figure 3. CenH3 losses associated with holocentric transitions are unique to insects. Phylogenetic relationship of holocentric lineages are schematized. Transitions to holocentricity (M → H) and the absence (black box) or presence (empty box) of CenH3 are indicated. Inferred evolutionary origin of a first ‘potentiating event’ (indicated with a star) together with subsequent recurrent transitions to holocentricity allowing the loss of CenH3 in four insect lineages. DOI: [10.7554/eLife.03676.018](https://doi.org/10.7554/eLife.03676.018)

diation, and (4) the presence of kinetochore plates covering large surfaces of each chromatid during mitosis. In addition to these four primary mitotic criteria for holocentricity, we consider certain properties at meiosis suggestive of holocentricity. Even though mitosis of holocentric chromosomes is straightforward, holocentric chromosomes would face kinetochore geometry problems during meiosis. Therefore, meiosis in holocentric species requires unique adaptations and can therefore be used as indications for holocentricity. Such adaptations include ‘inverted meiosis’ (i.e., inverting the order of the two meiotic divisions) or the change of kinetic activity on chromosomes from one region to another between meiotic divisions (Melters et al., 2012).

Lepidoptera (butterflies, moths)

Most Lepidoptera have numerous, small chromosomes, challenging conclusions about the nature of centromeres purely on the basis of light microscopy. Nonetheless, convincing evidence for holocentric chromosomes exists in *Bombyx mori*, one of the most well-studied holocentric insects (Murakami and Imai, 1974). *B. mori* chromosomes are rod- and dot-shaped and lack primary constrictions. Also, during anaphase, sister chromatids migrate in parallel. Furthermore, chromosomal fragments induced by X-ray irradiation behave normally through a number of cell generations (Murakami and Imai, 1974). More recently, immuno-fluorescence studies showed the localization of *B. mori* chromosomal passenger complex protein INCENP along the whole length of mitotic chromosome (Mon et al., 2014).

Studies from additional lepidopteran insects support the general conclusion of holocentricity in all members of Lepidoptera. Mitotic chromosomes of *Orthosia gracilis* appear to lack a primary constriction (Traut and Mosbacher, 1968). Electron microscopy (EM) in *Origya thyellina* and *Origya antiqua* chromosomes report spindle microtubule attachments to large kinetochore plates, covering more than 70% of the entire surface of the chromosomes (Wolf, 1994; Wolf et al., 1997). This has led to the proposal of a polykinetic organization of *Origya* chromosomes. Original images of mitotic chromosomes of *Samia cynthia* (unpublished), *Ectomyelois ceratoniae* (Mediouni et al., 2004), and *Cydia pomonella* (Fukova et al., 2007), as well as an image of metaphase I chromosomes of *Ephestia kuehniella* (unpublished) (Figure 1—figure supplement 9), generously provided by Frantisek Marec and Atsuo Yoshido, show that the chromosomes of these species lack primary constrictions. *E. kuehniella* chromosomes have also been subject to EM studies that have reported large kinetochore plates covering the chromosomal surface. However, compared to *Origya* species, the extent of kinetochore occupancy in *E. kuehniella* was estimated to be not as wide. Therefore, the kinetic organization of *E. kuehniella* chromosomes was interpreted to represent an intermediate between monocentricity and complete holocentricity (Wolf, 1994). Finally, the presence of holocentric chromosomes is also consistent with

invention of a centromeric protein (Ross et al., 2013), which may have conferred these independently derived holocentric lineages with a unique ability to carry out mitosis in a CenH3-independent manner, thereby relaxing the selective pressure to maintain CenH3 in the genome. Thus, changes in centromere architecture may have rendered dispensable one of the most defining proteins associated with centromere function in almost all eukaryotes.

Materials and methods

Evidence for holocentric insect orders

Four main criteria have been used to diagnose holocentricity by studying the mitotic behavior of chromosomes. These include (1) the absence of a primary constriction in metaphase chromosomes, (2) parallel migration of sister chromatids in mitotic anaphase, (3) persistence of chromosomal fragments and/or low rate of lethality upon X-ray irradiation, and (4) the presence of kinetochore plates covering large surfaces of each chromatid during mitosis.

experimental data reporting high resistance to radiation in multiple lepidopteran species (**LaChance and Graham, 1984; Koval, 1996; Marec et al., 2001**).

In contrast to the general conclusion about holocentricity in Lepidoptera, two lepidopteran species, *Pieris brassica* and *Polyommatus icarus*, have been reported to be monocentric in mitosis inferred by the presence of primary chromosomal constrictions observed by light microscopy (**Bigger, 1975; Rishi and Rishi, 1978**). However, meiotic chromosomes from one of these species were reported to lack a localized centromere (**Bigger, 1975**). Moreover, due to low zygotic lethality and frequent presence of chromosomal translocations after X-ray irradiation in *Pieris brassicae*, it has been proposed that *P. brassicae* chromosomes are in fact holocentric (**Bauer, 1967**).

Taken together the cytogenetic results on Lepidoptera, Bauer proposed that 'sufficient proof exists to conclude that all Lepidoptera have holokinetic chromosomes' (**Bauer, 1967**). We therefore favor this general conclusion that Lepidoptera are holocentric.

Trichoptera (caddis flies)

Trichoptera are a sister lineage to the Lepidoptera. An early study suggested the possibility of holocentric chromosomes in the trichopteran species *Limnophilus borealis* and *Limnophilus decipiens* (**Suomalainen, 1966**). More recent EM data of *L. decipiens* instead proposed an intermediate form between the holo- and monocentric type (**Wolf et al., 1997**). Additional trichopteran species have not been examined. Thus, although the consensus view remains that members of Trichoptera are holo- rather than monocentric, this conclusion is not as well supported as for some of the other insect orders.

As a result, it remains unknown whether the transition to holocentricity occurred in the common ancestor of all Lepidoptera or in the common ancestor of Lepidoptera and Trichoptera. We have not assessed the status of *CenH3* genes in Trichoptera.

Hemiptera (true bugs)

Similar to Lepidoptera, there are numerous reports of holocentric chromosomes in several species of Hemiptera (**Melters et al., 2012**). Chromosomes of multiple hemipteran species belonging to the hemipteran suborder Heteroptera (*Oncopeltus fasciatus* [**Wolfe and John, 1965**], *Euschistus servus*, *Euschistus tristigmus*, and *Sotubea pugnax* [**Hughes-Schrader and Schrader, 1961**]), as well as from the suborder Sternorrhyncha (aphids [**Sun and Robinson, 1965; Mandrioli et al., 2011**]) lack primary constrictions. EM studies in *O. fasciatus* (**Comings and Okada, 1972**) and *Rhodnius prolixus* (**Buck, 1967**) reporting kinetochore occupancy covering 75% and 100% of mitotic chromosomes respectively, further support the presence of holocentric chromosomes in Heteroptera. Experimental support for holocentricity was also obtained using chromosomal fragmentation studies, which demonstrated that fragments induced by X-rays are able to propagate themselves mitotically through many cell generations (**Hughes-Schrader and Schrader, 1961; Khuda-Bukhsh and Datta, 1981; Kuznetsova and Sapunov, 1985**). Furthermore, C-banding of several aphid species revealed that heterochromatic areas are interspersed with euchromatic areas whereas heterochromatin of monocentric chromosomes is concentrated in one part of the chromosome (reviewed in **Manicardi et al., 2002**).

There also exists ample evidence of holocentricity in Hemiptera based on meiotic observations that reveals inverted meiosis of sex chromosomes (**Schrader, 1935; Viera et al., 2009**) and/or change of kinetic activity on autosomes from one region to another between meiosis I and II (**Perez et al., 1997**).

As a result of all these studies, we agree with the previous conclusion that 'the diffuse nature of the hemipteran kinetochore is attested by a wealth of observational evidence' (**Hughes-Schrader and Schrader, 1961**).

Phthiraptera (lice)

Phthiraptera is a sister order with Hemiptera. However, compared to Hemiptera, there is relatively little data characterizing the nature of the centromere in this order. Chromosome structure has been investigated by cytological analysis in *Haematopinus suis* and *Menacanthus stramineus* (**Tombesi and Papeschi, 1993**) and in *Bovicola limbata* and *Bovicola caprae* (**Tombesi et al., 1999**). The presence of holocentric chromosomes was reported based on the cytological observations, including the lack of primary constrictions and parallel sister chromatid migration during anaphase. Experimental support has been obtained in *H. suis* by Bayreuther (**Bayreuther, 1955**), who observed a regular behavior of chromosome fragments during cell division after irradiating first and second instar individuals.

Although Hindle and Pontecorvo reported the existence of primary constrictions in *Pediculus corporis* chromosomes (without providing cytological data) indicating the presence of monocentric

chromosomes (**Hindle and Pontecorvo, 1942**), this view was challenged by Bayreuther, who instead proposed holocentric chromosomes in *Pediculus*, as this is more consistent with observations following DNA fragmentation experiments (**Pontecorvo, 1944**), including the lack of harmful effects on *Pediculus* spermatozoa and the absence of reports about akinetic fragments or mitotic chromosome bridges (**Bayreuther, 1955**). To our knowledge no photographic evidence exists showing the presence of monocentric chromosomes in Phthiraptera. Though the listed photographic support leading to the author's interpretation of holocentricity is difficult to assess today, we favor the consensus interpretation that Phthiraptera are holocentric, particularly due to the strength of evidence based on studies assessing chromosomal inheritance following irradiation.

Thus, the common ancestor of Hemiptera and Phthiraptera may have undergone a single transition to holocentricity.

Dermaptera (earwigs)

Multiple species of Dermaptera have been reported to be holocentric. Ortiz studied chromosome cytology of seven species of Dermaptera by microscopy and reported that '*the chromosomes show no localized centromere, as inferred from their structure, mode of division and anaphase movement*' (**Ortiz, 1969**). In this study, the most conclusive data are derived from *Labidura ripura*, a close relative to *Forficula auricularia* belonging to the same suborder. *L. ripura* has the smallest number of chromosomes with the largest relative length facilitating the assessment of lack of primary constrictions (**Ortiz, 1969**). Furthermore, the chromosomes of two additional earwigs have been examined. In one study in *Labidura truncata*, Webb reported that '*the primary constrictions of fixed centromeres do not show, and uninterrupted chromosomes, characteristic of holocentric chromosomes...*' (**Webb, 2009**). In another study, MJD White examined *Hemimerus bouvieri* (**White, 1971**) and observed rod-shaped chromosomes with no constriction and parallel migration of sister chromatids during anaphase.

Two studies from the earwig *F. auricularia* challenged the consensus view that Dermaptera are holocentric (**Callan, 1941; Henderson, 1970**). Both Hendersen and Callan reported the presence of primary constrictions in *F. auricularia* chromosomes. The view of these authors, however, has been challenged by Ortiz who found '*chromosomal spindle fibers exist along the mitotic chromosomes*' thus proposing that '*the chromosomes of Dermaptera have a diffuse centromere*' (**Ortiz, 1969**). We also favor the more general conclusion that Dermaptera are indeed holocentric.

Odonata (dragonflies)

The nature of centromeres in Odonata has been a subject of great debate. Early studies concluded that chromosomes in this order are metacentric (**Oksala, 1943**), acrocentric (**Chauduri and Das Gupta, 1949; Seschachar and Bagga, 1962**), or holocentric (**Cumming, 1964; Kiauta, 1969a, 1969b**). The most recent study on odonatan chromosomes, Nokkala et al. applied silver staining methods to visualize metaphase chromosomes and the behavior of male meiotic chromosomes in two dragonfly species (**Nokkala et al., 2002**). This study clearly demonstrated the absence of a primary constriction and parallel alignment of metaphase chromosomes. Furthermore, this study also showed that during meiosis different chromosomal regions (telomeres in the first division and the middle parts of chromosomes in the second) showed kinetic activity. At present, although the two species studied most recently appear to be holocentric, we cannot conclude that this is representative of the entire order. Indeed, it is formally possible that holocentricity might have arisen within Odonata rather than in the common ancestor. Although the centromere status of the species that we sequenced has not been determined, they belong to the same suborder as the holocentric dragonflies studied by Nokkala.

Zoraptera (zorapterans)

Zoraptera represent an insect order that is relatively poorly studied especially with respect to their chromosome structure. Although there is also a significant uncertainty of their phylogenetic position, they may represent a sister lineage to the Dermaptera. In investigations of male meiotic chromosomes in *Zorotypus hubbardi*, no primary constrictions were observed at any meiotic stage (**Kuznetsova et al., 2002**) leading to the conclusion that they are holocentric. We have been unable to obtain samples of Zoraptera to assess the status of their *CenH3* genes.

Ephemeroptera (mayflies)

Ephemeropteran insects have been previously referred to as being holocentric (**Melters et al., 2012**). However, cytogenetic studies of many species of this order (including *E. danica* that we included in our

transcriptome analyses) have clearly reported that the centromeres are metacentric or acrocentric (Kiauta and Mol, 1977; Soldan and Putz, 2000). Kiauta and Mol compared the cytological evidence for monocentric chromosomes in Ephemeroptera directly to their cytological evidence of holocentric chromosomes in species belonging to the sister order Odonata (Kiauta, 1969a, 1969b). As a result, the authors conclude that 'any cytogenetic similarities between the two orders are completely lacking' (Kiauta and Mol, 1977). We therefore conclude that the assignment of holocentricity to the Ephemeroptera is incorrect and they should be assigned a monocentric status instead.

RNA isolation and sample preparation

We obtained German cockroaches (*Blattella germanica*), crickets (*Acheta domesticus*), and dragonfly nymphs (*Libellula vibrans*) from Carolina Biological Supply Company. Earwigs (*Anisobasis maritima*) and stick insects (*Sipyloidea sipyilus*) were collected from the field and obtained from captive populations, respectively. For each insect, we dissected the head and thorax (to minimize the contamination by gut microbiota), which were then homogenized and total RNA was isolated using the TRIZOL reagent according to the manufacturer's instructions (Invitrogen, Carlsbad, CA). RNA was treated with DNaseI (Ambion) to remove DNA contamination and purified using Qiagen RNeasy Mini Kit. mRNA-sequencing (mRNA-Seq) libraries were prepared using Illumina's mRNA-Seq Sample Prep Kit (Illumina, San Diego, CA) following the manufacturer's protocol. Barcoded libraries were multiplexed and sequenced as 50-bp paired-end reads on the HiSeq 2000 platform. Reads were parsed by barcodes, and the number of reads for each species is listed in **Figure 1—source data 2**.

Assemblies of mRNA-Seq reads

We carried out de novo assembly of mRNA-Seq reads using a combination of Velvet (Zerbino and Birney, 2008) and Oases (Schulz et al., 2012) programs. We generated assemblies of the publically available 100-bp paired-end mRNA-Seq libraries from mayfly (*Ephemera danica*) and scarce chaser (*Ladona fulva*) (<https://www.hgsc.bcm.edu/>), as well as our own 50 bp paired-end mRNA-Seq libraries, using a k-mer length of 21. K-mer lengths were set to 31 for additional assemblies of publically available 100 bp single-end or paired-end mRNA-Seq data sets from several lepidopteran, coleopteran, and hemipteran species (Figure 1B [Zhen et al., 2012]), as well as from *Drosophila melanogaster* strain Oregon-R ovaries (SRR384962, <http://www.ncbi.nlm.nih.gov/sra/SRX109311>). We generated additional assemblies for the cockroach (*B. germanica*), dragonfly (*L. vibrans*), earwig (*A. maritima*), and cricket (*A. domesticus*) libraries with k-mer lengths set to 17, 23, and 33 using the SPAdes assembler (Nurk et al., 2013). A transcriptome assembly of a second dermapteran mRNA-Seq data set from the earwig *Forficula auricularia* (Roulin et al., 2014) (SRR1043671, SRR1048074, SRR1051467) was also included in our analyses.

Comparing transcriptome coverages in mRNA-Seq assemblies

Transcripts at least 250 nucleotides in length were selected. In cases where multiple isoforms were present, we chose the longest isoform. We used the *Tribolium castaneum* proteome (16,644 proteins, version as of December 2013) and the *D. melanogaster* proteome (30,305 proteins, version as of May 2014) to query the assembled transcripts using BLAST (tblastn). Significant alignments (E value $< 10^{-10}$) were counted considering only one alignment per transcript of *T. castaneum*/*D. melanogaster* protein. The results are listed in **Figure 1—source data 2**.

We predicted similar fractions of *T. castaneum* (and *D. melanogaster*) proteins in each of our assemblies from mono- or holocentric species implying comparable transcriptome coverages. Overall, we could predict 24–48% of the *T. castaneum* proteome in assemblies from monocentric insects. Each of these assemblies also contained CenH3. As expected due to their closer evolutionary relationship compared to other insect orders, the biggest *T. castaneum* proteome fractions were predicted in assemblies generated from beetle mRNA-Seq data sets. Protein predictions in assemblies from holocentric insects yield fractions ranging from 25% to 44% (Figure 1—source data 2). Proteins that were not found either did not exist in the corresponding species or were too divergent to pass our alignment threshold (E value $< 10^{-10}$), or were not expressed in the sample (or tissues analyzed). The assembly of *D. melanogaster*, an insect species with an annotated proteome, was used to assess the expected fraction of predicted *T. castaneum* proteins. 28% of all *T. castaneum* proteins were predicted (total 16,644 proteins) in the *D. melanogaster* assembly, which falls within the range of predicted fractions in other assemblies. This assembly covers 21% of all *D. melanogaster* proteins (total 30,305 proteins) and includes the *D. melanogaster* CenH3 homolog.

Estimating transcript abundance in mRNA-Seq assemblies

The assembled transcripts were parsed into a non-redundant set selecting the longest isoform per transcript. We mapped the mRNA-Seq reads to the assembled set of transcripts using Bowtie 2 using default parameters (Langmead and Salzberg, 2012). Mapped reads were normalized to transcript lengths to estimate transcript abundances. Transcripts (longer than 200 nt) were rank-ordered based on their abundances. We determined the ranked percentiles corresponding to transcripts encoding for CenH3 homologs and kinetochore proteins, where low percentiles correspond to low abundances, while high percentiles correspond to high abundances (Figure 1—figure supplement 5). Although CenH3 was only at the bottom 10th percentile in terms of abundance in the cockroach transcriptome assembly, CenH3 transcript abundance was in the top 50th percentile or better in the mayfly, the stick insect, the cricket, and the *D. melanogaster* assemblies. Thus, the absence of CenH3 in the assemblies of all holocentric insect species is likely not due to low CenH3 expression levels.

Searches for CenH3 and kinetochore proteins

CenH3 homology search in insects with sequenced genomes

NCBI non-redundant protein databases or specialized proteomes (Figure 1—source data 1) and whole genomes were searched using BLAST (psi-blast, blastp, and tblastn). We used *D. melanogaster* histone H3 (canonical histone H3 sequence is almost invariant among insects) and *D. melanogaster* CenH3 homolog Cid (FBgn0040477) as queries for our initial searches. Newly obtained CenH3 homologs were subsequently added to the list of query proteins, in iterative searches. We tested any putative homolog thus obtained for histone folds by HHpred (Soding et al., 2005) and for diagnostic features of known CenH3 homologs (Malik and Henikoff, 2003) including an extended loop one region, absence of specific amino acids including glutamine, phenylalanine, and threonine at positions 69, 85, and 118, respectively (as compared to canonical H3). These criteria have allowed the successful identification of CenH3 homologs in animals, plants, and even in distantly related protozoa (Malik and Henikoff, 2003), while revealing the absence of CenH3 in kinetoplastids (Talbert and Henikoff, 2010) that has recently been confirmed (Akiyoshi and Gull, 2014). In addition to those features, phylogenetic analyses support homology of the newly identified CenH3s to other known CenH3 proteins (Figure 1—figure supplement 6). All identified homologous proteins or protein fragments are listed in Figure 2—source data 1.

In addition to canonical H3, we found the universally conserved H3 variant H3.3 that is almost identical to H3 in all genomes of mono- and holocentric species. All genomes examined from monocentric insects contained an additional H3-like protein with blastp *E* values in the order of 10⁻¹⁰, which corresponds to CenH3. In contrast, apart from H3 and H3.3, no additional functional H3-like protein could be detected in any of the genomes of holocentric insects except for a putatively nonfunctional H3-derived variant in the *Acyrtosiphon pisum* genome (Figure 1—figure supplement 3A,C).

Although a putative *B. mori* CenH3 homolog with 23% identity to the histone-fold domain of the *Caenorhabditis elegans* CenH3 homolog (hcp-3) was described in the *B. mori* genome sequence publication (Xia et al., 2004), we were unable to replicate these findings; our analyses revealed no CenH3 in the *B. mori* genome, consistent with a more recent study (Mon et al., 2011) that also was unable to identify a CenH3 homolog in *B. mori* (however no details were provided about the prediction criteria and results leading to this conclusion).

CenH3 searches in transcriptome assemblies and expressed sequence tag (EST) databases

We surveyed transcriptome assemblies and EST databases via tblastn searches using canonical H3 and any identified insect CenH3 homolog as query proteins. As expected, those searches revealed the presence of the universally conserved H3 variant H3.3 in all assemblies. In addition to H3.3, assemblies from all monocentric species contained another H3-like protein corresponding to their CenH3 homologs. CenH3 encoding transcript fragments in the monocentric coleopteran *Tetraopes tetraophthalmus* and *Megacyllene robiniae* assemblies were detected in the raw 100 bp reads using the *Anoplophora glabripennis* CenH3 as query. Assemblies or EST data (Negre et al., 2006) from holocentric insects did not contain any additional insect H3-like hits apart from H3.3 and a recent H3-derived variant in the *L. vibrans* assembly (Figure 1—figure supplement 3B,C). Two lepidopteran assemblies revealed CenH3 encoding transcripts that could easily be attributed to microsporidan contamination

via phylogenetic analyses (**Figure 1—figure supplement 8A,B**). To rule out the possibility that CenH3 encoding read fragments are present in the mRNA-Seq libraries but not assembled to transcripts, raw 100 bp reads from any of the Lepidoptera as well as from *L. fulva* library were searched for H3-like protein fragments using BLAST (tblastn). Apart from H3 or H3.3 encoding read fragments, no additional CenH3 encoding read fragments could be identified in those libraries, whereas the same search yields significant hits (with *E* values down to 5×10^{-9}) to the CenH3 homolog in the monocentric mayfly mRNA-Seq data set. The 50 bp reads of the other insect libraries were not suitable for these analyses because the translated protein fragment is too short to yield significant alignments to the CenH3 homologous fragments using H3-like query proteins. Consistent with our analyses based on the Velvet/Oases assemblies, searches for CenH3 homologs using BLAST (tblastn) in the SPAdes assemblies from two holocentric species (earwig and dragonfly) did also not reveal any hits apart from H3.3.

The transcriptome assembly of the second earwig, *F. auricularia*, was recently generated using a combination of Roche 454 reads, Illumina 100 bp and 150 bp paired-end reads (**Roulin et al., 2014**). We searched H3-like proteins in the *F. auricularia* database using the associated BLAST server (http://gdcws1.ethz.ch/blastdb_fau/blast.html) and in the transcripts using tblastn. No CenH3 homolog was found.

CenpC homology search

Transcripts, genomes, or annotated proteomes were searched (psi-blast, blastp, and tblastn) as described above using CenpC motifs, CenpC-C-terminal regions including the cupin domain or full-length CenpC proteins of human, yeast, and insect homologs. Newly identified CenpC homologs were added to the list of query proteins for iterative searches. A recent study searched selected insect genomes for CenpC homologs and other kinetochore proteins, without providing sequence information or accession numbers (**Schleiffer et al., 2012**). A CenpC homolog was reported in the beetle *T. castaneum*, while CenpC was not found in several species of Hymenoptera. Our searches did not reveal a recognizable CenpC homolog in *T. castaneum*; however, we identified several putative hymenopteran CenpC homologs.

The assembled *L. fulva* (scarce chaser) CenpC fragment only covered the 3' end of the mRNA transcript including the stop codon of the open reading frame. The 5' end of the fragment was manually extended with unassembled read fragments of the *L. fulva* library, identified by sequence similarity (tblastn) to the homologous *L. vibrans* CenpC protein (SPAdes Node 13,580).

Cupin domains are found in many protein subfamilies, which form distinct clades in phylogenetic analyses (**Dunwell et al., 2001**). Reciprocal best hit analyses using the dragonfly cupin domains identified CenpC-like cupin domains of various eukaryotic CenpC homologs, therefore confirming the homology of the dragonfly cupins to CenpC-like cupin subfamily.

5' RACE analyses

The dragonfly CenpC fragment (SPAdes node 13,580) identified in the assemblies was extended towards the 5' end of the transcript by 5' RACE analyses using Invitrogen's 5'RACE System for Rapid Amplification of cDNA Ends. 10 µg total RNA was treated with DNaseI (Ambion) to remove DNA contamination. Gene-specific primers complementary to the assembled CenpC fragment were used for first strand cDNA syntheses using Superscript III (Invitrogen, Carlsbad, CA) following manufacturer's instructions. The cDNA was treated with RNase H and purified using S.N.A.P. columns (Invitrogen, Carlsbad, CA). Half of the purified cDNA was used for terminal deoxynucleotidyl transferase tailing reaction with dCTP. Tailed cDNA was precipitated and used for final amplification with gene-specific primers and 5' RACE abridged anchor primer provided in the kit.

Kinetochore protein homology searches

While most inner kinetochore components found in vertebrates and fungi are absent in *D. melanogaster* (**Przewloka et al., 2007**), our own analyses are consistent with previous data (**Schleiffer et al., 2012**) that have revealed the presence of several kinetochore components in other insects, many of which appear to be highly divergent among insects. We therefore applied sensitive protein predictions using a combination of psi-blast, blastp, and tblastn in genomes, NCBI non-redundant protein databases, or within specialized proteomes (**Figure 1—source data.1**). Searches in the mRNA-Seq assemblies were performed using tblastn using human and any newly identified insect homologs. Hits were verified by identification of corresponding reciprocal best hits and HHpred prediction analyses.

CenpS and CenpX

Using human and chicken CenpS and CenpX proteins as queries, we identified several insect homologs. We further verified those candidates using the structural based HHpred analyses aligning their conserved histone fold to known homologs. The annotation of the CenpX encoding gene in *B. mori* was not correct, but we could infer the correct transcript using *B. mori* EST data.

CenpT and CenpW homologs form a complex with CenpS and CenpX in kinetochores of vertebrates and fungi (*Perpelescu and Fukagawa, 2011*). CenpT and CenpW could not be found in any of the insect genomes or transcriptomes. We therefore conclude that CenpS and CenpX are conserved in insects due to their role in DNA repair (*Singh et al., 2010*), rather than as kinetochore components due to the absence of their binding partners CenpT and CenpW.

Cenpl

Vertebrate Cenpl proteins reveal significant alignments to insect homologs that could be verified by reciprocal best-hit analyses. The Cenpl encoding gene in *B. mori* genome was not correctly annotated, but we could correctly infer it using EST data. Three non-overlapping transcript fragments in the cricket assembly revealed significant Cenpl alignments and both are included in **Figure 2—source data 1**. Homologs of the Cenpl protein-binding partners, CenpH and CenpK, could not be identified in any insect genome.

CenpL/M/N

CenpL and CenpM proteins appear to evolve rapidly, making cross-species prediction between insect and vertebrate homologs difficult. Still, insect homologs could be detected using annotated *Pediculus humanus corporis* CenpM and hymenopteran CenpL homologs and verified by HHpred and reciprocal best-hit analyses. CenpN homologs are unstructured proteins that rapidly evolve in animals and fungi. Nevertheless, insect CenpN homologs could be identified using psi-blast and verified by reciprocal best-hit analyses.

The N-terminus of CenpN interacts with CenH3 in humans (*Carroll et al., 2009*), and two arginine residues conserved across vertebrates have been shown to be important for CenpN–CenH3 interaction (*Carroll et al., 2009*). However, we could not identify any corresponding arginines in any of the insect CenpN homologs due to weak conservation between vertebrate and insect CenpN homologs. We therefore do not know if insect CenpN proteins retain the property of CenH3-binding, even in monocentric insects.

CenpO/P/Q/R

The CenpO/P/Q/R complex is not essential in vertebrates (*Perpelescu and Fukagawa, 2011*); however recent findings in mice indicate that CenpU is essential in embryonic stem cells and in embryonic development (*Kagawa et al., 2014*). Homologs of the CenpO/P/Q/R complex could not be identified in any insect genome or transcriptome.

Outer kinetochore proteins

We used *D. melanogaster*, vertebrate, and lepidopteran Mis12 and Ndc80 proteins to find their insect homologs. Both proteins were widely present in most but not all insects. Mis12 and Ndc80 are the central components of the outer kinetochore protein complexes. Using several rounds of psi-blast, we detected putative homologs of additional Mis12 complex components, Nnf1 and Dsn1, in insect genomes that also encode for Mis12 but not in the Mis12-deficient orders including the Coleoptera, the Hemiptera, and the Phthiraptera, supporting the loss of Mis12 in those orders. Still, experimental validation as performed in the *D. melanogaster* S2 cells (*Przewlaka et al., 2007*) will be necessary to verify the presence of additional Mis12 complex components or confirm their absence. Ndc80 homologs could not be found in the stick insect and the cockroach transcriptome assemblies. We searched for Spc25 that is part of the Ndc80 complex and found homologs to be present in all insect genomes and transcriptomes, including the Ndc80-deficient stick insect and the cockroach assembly. Spc25 transcripts are highly abundant in all assemblies facilitating their detection. In contrast, Ndc80 abundance is more variable. Thus, it is possible that our inability to detect Ndc80 (**Figure 2**) was a result of lower expression, particularly since data only from one representative species was analyzed from either of these orders. It will be necessary to confirm this finding with analyses of additional representatives in the future.

Phylogenetic analyses

Amino acid sequences of histone fold domains for CenH3 homologs or cupin domains of CenpC homologs were aligned using MUSCLE (*Edgar, 2004*). For **Figure 2D**, **Figure 1—figure supplement 3C**,

Figure 1—figure supplement 4, and **Figure 1—figure supplement 8B** the optimal amino acid substitution model was determined using protpstest (Abascal et al., 2005). Maximum likelihood trees were generated using Phyml (Guindon et al., 2010). Bootstrap values supporting the topology with the highest maximum likelihood were obtained using Phylip (Felsenstein, 1989). For **Figure 1—figure supplement 6**, the maximum likelihood tree was built using the PhyML server (Dereeper et al., 2008) using default parameters and bootstraps set to 100. (Suomalainen, 1966; Jarvis et al., 2005; Terry and Whiting, 2005).

Acknowledgements

We thank S Biggins, I Cheeseman, M Daugherty, L Kursel, A Marty, S Ramachandran, B Ross, F Steiner, P Talbert, J Young, and S Zanders for advice, helpful discussions, and comments on the manuscript. We are grateful to P Andolfatto for suggesting transcriptome assemblies as a search strategy, and to J Young and M Eickbush for help with transcriptome assembly analyses. We thank the Baylor I5K Project for providing access to four RNA-Seq data sets, and JC Walser for providing the *F. auricularis* assembly and early access to the transcriptome blast server. We would also like to acknowledge our colleagues S Nokkala, W Traut, F Marec, A Yoshido, and J Rufas for generously providing images of holocentric chromosome spreads for inclusion in our supplementary discussion. This study was supported by a postdoctoral Fellowship of the Jane Coffin Childs Memorial Fund for Medical Research (IAD), NIH grants R01-GM74108 (HSM). HSM and SH are Investigators of the Howard Hughes Medical Institute. Sequence data are available in the NCBI Short Read Archive (SRA) (PRNJNA258192).

Additional information

Funding

| Funder | Grant reference number | Author |
|---|-------------------------|-------------------------------------|
| Jane Coffin Childs Memorial Fund for Medical Research | Postdoctoral fellowship | Ines A Drinnenberg |
| National Institute of General Medical Sciences | R01 GM74108 | Harmit Singh Malik |
| Howard Hughes Medical Institute | Investigator | Steven Henikoff, Harmit Singh Malik |
| G Harold and Leila Y. Mathers Foundation | | Harmit Singh Malik |

The funders had no role in study design, data collection and interpretation, or the decision to submit the work for publication.

Author contributions

IAD, Conception and design, Acquisition of data, Analysis and interpretation of data, Drafting or revising the article; DY, Acquisition of data, Contributed unpublished essential data or reagents; SH, HSM, Conception and design, Analysis and interpretation of data, Drafting or revising the article

Author ORCIDs

Harmit Singh Malik,  <http://orcid.org/0000-0001-6005-0016>

Additional files

Major datasets

The following dataset was generated:

| Author(s) | Year | Dataset title | Dataset ID and/or URL | Database, license, and accessibility information |
|---|------|--|--|--|
| Drinnenberg IA, deYoung D, Henikoff S, Malik HS | 2014 | Insects Transcriptome or Gene expression | PRJNA258192; http://www.ncbi.nlm.nih.gov/sra/PRJNA258192 | Publicly available at NCBI Short Reads Archive. |

The following previously published datasets were used:

| Author(s) | Year | Dataset title | Dataset ID and/or URL | Database, license, and accessibility information |
|--|------|--|---|---|
| The modENCODE consortium | 2011 | Functional genomics project for the <i>Drosophila</i> modENCODE Project | http://www.ncbi.nlm.nih.gov/sra/SRX109311 | Publicly available at NCBI Short Reads Archive. |
| Roulin AC, Wu M, Pichon S, Arbore R, Kühn-Bühlmann S, Kölliker M, Walsler JC | 2013 | De novo transcriptome hybrid assembly and validation in the European earwig (<i>Dermaptera</i> , <i>Forficula auricularia</i>) | http://www.ncbi.nlm.nih.gov/bioproject?LinkName=sra_bioproject&from_uid=557129 | Publicly available at NCBI Short Reads Archive, https://www.hgsc.bcm.edu/arthropods . |
| | N/A | Arthropod Sequencing at the BCM-HGSC | https://www.hgsc.bcm.edu/arthropods | Publicly available. |
| Zhen Y, Aardema ML, Medina EM, Schumer M, Andolfatto P | 2012 | Insect Genomics | http://genomics-pubs.princeton.edu/insect_genomics/data.shtml | Publicly available. |

Reporting standards: Standard used to collect data: Datasets reported as per guidelines of the NCBI Short Read Archive (SRA):

References

- Abascal F, Zardoya R, Posada D. 2005. ProtTest: selection of best-fit models of protein evolution. *Bioinformatics* **21**:2104–2105. doi: [10.1093/bioinformatics/bti263](https://doi.org/10.1093/bioinformatics/bti263).
- Akiyoshi B, Gull K. 2013. Evolutionary cell biology of chromosome segregation: insights from trypanosomes. *Open Biology* **3**:130023. doi: [10.1098/rsob.130023](https://doi.org/10.1098/rsob.130023).
- Akiyoshi B, Gull K. 2014. Discovery of unconventional kinetochores in kinetoplastids. *Cell* **156**:1247–1258. doi: [10.1016/j.cell.2014.01.049](https://doi.org/10.1016/j.cell.2014.01.049).
- Bauer H. 1967. Die kinetische Organisation der Lepidopteren-Chromosomen. *Chromosoma* **22**:101–125. doi: [10.1007/BF00326724](https://doi.org/10.1007/BF00326724).
- Bayreuther K. 1955. Holokinetische Chromosomen bei *Haematopinus suis* (Anoplura, Haematopinidae). *Chromosoma* **7**:260–270. doi: [10.1007/BF00329726](https://doi.org/10.1007/BF00329726).
- Bigger TRL. 1975. Karyotypes of some Lepidoptera chromosomes and changes in their holokinetic Organisation as revealed by NewCytological Techniques. *Cytologia* **40**:713–726.
- Blower MD, Karpen GH. 2001. The role of *Drosophila* CID in kinetochore formation, cell-cycle progression and heterochromatin interactions. *Nature Cell Biology* **3**:730–739. doi: [10.1038/35087045](https://doi.org/10.1038/35087045).
- Buchwitz BJ, Ahmad K, Moore LL, Roth MB, Henikoff S. 1999. A histone-H3-like protein in *C. elegans*. *Nature* **401**:547–548. doi: [10.1038/44062](https://doi.org/10.1038/44062).
- Buck RC. 1967. Mitosis and eiosis in *Rhodnius Prolixus*: the Fine structure of the spindle and diffuse kinetochore. *Ultrastructure Research* **18**:489–501. doi: [10.1016/S0022-5320\(67\)80199-0](https://doi.org/10.1016/S0022-5320(67)80199-0).
- Callan HG. 1941. The sex-determining mechanism of the earwig, *Forficula auricularia*. *Journal of Genetics* **41**:350–374. doi: [10.1007/BF02983027](https://doi.org/10.1007/BF02983027).
- Carroll CW, Milks KJ, Straight AF. 2010. Dual recognition of CENP-A nucleosomes is required for centromere assembly. *The Journal of Cell Biology* **189**:1143–1155. doi: [10.1083/jcb.201001013](https://doi.org/10.1083/jcb.201001013).
- Carroll CW, Silva MC, Godek KM, Jansen LE, Straight AF. 2009. Centromere assembly requires the direct recognition of CENP-A nucleosomes by CENP-N. *Nature Cell Biology* **11**:896–902. doi: [10.1038/ncb1899](https://doi.org/10.1038/ncb1899).
- Chaudhuri SP, Das Gupta J. 1949. Cytological studies on the Indian dragonflies. I. Structure and behaviour of the chromosomes in six species of dragonflies (Odonata). *Proceedings of the Zoological Society of Bengal* **2**:81–93.
- Comings DE, Okada TA. 1972. Holocentric chromosomes in oncopeltus: kinetochore plates are present in mitosis but absent in meiosis. *Chromosoma* **37**:177–192. doi: [10.1007/BF00284937](https://doi.org/10.1007/BF00284937).
- Cumming RB. 1964. *Cytogenetic studies in the order Odonata*. Austin, Texas. University of Texas.
- Dawson SC, Sagolla MS, Cande WZ. 2007. The cenH3 histone variant defines centromeres in *Giardia intestinalis*. *Chromosoma* **116**:175–184. doi: [10.1007/s00412-006-0091-3](https://doi.org/10.1007/s00412-006-0091-3).
- Dereeper A, Guignon V, Blanc G, Audic S, Buffet S, Chevenet F, Dufayard JF, Guindon S, Lefort V, Lescot M, Claverie JM, Gascuel O. 2008. Phylogeny.fr: robust phylogenetic analysis for the non-specialist. *Nucleic Acids Research* **36**:W465–W469. doi: [10.1093/nar/gkn180](https://doi.org/10.1093/nar/gkn180).
- Dernburg AF. 2001. Here, there, and everywhere: kinetochore function on holocentric chromosomes. *The Journal of Cell Biology* **153**:F33–F38. doi: [10.1083/jcb.153.6.F33](https://doi.org/10.1083/jcb.153.6.F33).
- DeWulf P, Earnshaw W. 2008. *The Kinetochore::From Molecular Discoveries to Cancer Therapy*. Springer.
- Dunwell JM, Culham A, Carter CE, Sosa-Aguirre CR, Goodenough PW. 2001. Evolution of functional diversity in the cupin superfamily. *Trends in Biochemical Sciences* **26**:740–746. doi: [10.1016/S0968-0004\(01\)01981-8](https://doi.org/10.1016/S0968-0004(01)01981-8).
- Earnshaw WC, Rothfield N. 1985. Identification of a family of human centromere proteins using autoimmune sera from patients with scleroderma. *Chromosoma* **91**:313–321. doi: [10.1007/BF00328227](https://doi.org/10.1007/BF00328227).

- Edgar RC.** 2004. MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Research* **32**:1792–1797. doi: [10.1093/nar/gkh340](https://doi.org/10.1093/nar/gkh340).
- Felsenstein J.** 1989. PHYLIP - phylogeny Inference Package (Version 3.2). *Cladistics* **5**:164–166.
- Fukova I, Traut W, Vitkova M, Ngyuen P, Kubickova S, Marec F.** 2007. Probing the W chromosome of the codling moth, *Cydia pomonella*, with sequences from microdissected sex chromatin. *Chromosoma* **116**:135–145. doi: [10.1007/s00412-006-0086-0](https://doi.org/10.1007/s00412-006-0086-0).
- Gassmann R, Rechtsteiner A, Yuen KW, Muroyama A, Egelhofer T, Gaydos L, Barron F, Maddox P, Essex A, Monen J, Ercan S, Lieb JD, Oegema K, Strome S, Desai A.** 2012. An inverse relationship to germline transcription defines centromeric chromatin in *C. elegans*. *Nature* **484**:534–537. doi: [10.1038/nature10973](https://doi.org/10.1038/nature10973).
- Grimaldi D, Engel MS.** 2005. *The Evolution of the Insects*.
- Guindon S, Dufayard JF, Lefort V, Anisimova M, Hordijk W, Gascuel O.** 2010. New algorithms and methods to estimate maximum-likelihood phylogenies: assessing the performance of PhyML 3.0. *Systematic Biology* **59**:307–321. doi: [10.1093/sysbio/syq010](https://doi.org/10.1093/sysbio/syq010).
- Heckmann S, Schroeder-Reiter E, Kumke K, Ma L, Nagaki K, Murata M, Wanner G, Houben A.** 2011. Holocentric chromosomes of *Luzula elegans* are characterized by a longitudinal centromere groove, chromosome bending, and a terminal nucleolus organizer region. *Cytogenetic and Genome Research* **134**:220–228. doi: [10.1159/000327713](https://doi.org/10.1159/000327713).
- Hedges SB, Dudley J, Kumar S.** 2006. TimeTree: a public knowledge-base of divergence times among organisms. *Bioinformatics* **22**:2971–2972. doi: [10.1093/bioinformatics/btl505](https://doi.org/10.1093/bioinformatics/btl505).
- Henderson SA.** 1970. Sex chromosomal polymorphism in the earwig Forficula. *Chromosoma* **31**:139–164. doi: [10.1007/BF00285145](https://doi.org/10.1007/BF00285145).
- Henikoff S, Ahmad K, Platero JS, van Steensel B.** 2000. Heterochromatic deposition of centromeric histone H3-like proteins. *Proceedings of the National Academy of Sciences of USA* **97**:716–721. doi: [10.1073/pnas.97.2.716](https://doi.org/10.1073/pnas.97.2.716).
- Hindle E, Pontecorvo G.** 1942. Mitotic divisions following meiosis in *Pediculus corporis* males. *Nature* **149**:668. doi: [10.1038/149668a0](https://doi.org/10.1038/149668a0).
- Hori T, Amano M, Suzuki A, Backer CB, Welburn JP, Dong Y, McEwen BF, Shang WH, Suzuki E, Okawa K, Cheeseman IM, Fukagawa T.** 2008. CCAN makes multiple contacts with centromeric DNA to provide distinct pathways to the outer kinetochore. *Cell* **135**:1039–1052. doi: [10.1016/j.cell.2008.10.019](https://doi.org/10.1016/j.cell.2008.10.019).
- Howman EV, Fowler KJ, Newson AJ, Redward S, MacDonald AC, Kalitsis P, Choo KH.** 2000. Early disruption of centromeric chromatin organization in centromere protein A (Cenpa) null mice. *Proceedings of the National Academy of Sciences of USA* **97**:1148–1153. doi: [10.1073/pnas.97.3.1148](https://doi.org/10.1073/pnas.97.3.1148).
- Hughes-Schrader S, Schrader F.** 1961. The Kinetochore of the Hemiptera. *Chromosoma* **12**:327–350. doi: [10.1007/BF00328928](https://doi.org/10.1007/BF00328928).
- Jarvis KJ, Haas F, Whiting MF.** 2005. Phylogeny of earwigs (Insecta: dermaptera) based on molecular and morphological evidence: reconsidering the classification of Dermaptera. *Systematic Entomology* **30**:442–453. doi: [10.1111/j.1365-3113.2004.00276.x](https://doi.org/10.1111/j.1365-3113.2004.00276.x).
- Kagawa N, Hori T, Hoki Y, Hosoya O, Tsutsui K, Saga Y, Sado T, Fukagawa T.** 2014. The CENP-O complex requirement varies among different cell types. *Chromosome Research* **22**:293–303. doi: [10.1007/s10577-014-9404-1](https://doi.org/10.1007/s10577-014-9404-1).
- Kato H, Jiang J, Zhou BR, 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**:1110–1113. doi: [10.1126/science.1235532](https://doi.org/10.1126/science.1235532).
- Khuda-Bukhsh AR, Datta S.** 1981. A preliminary note on the mitotic chromosome aberrations in embryos of whole-body X-irradiated parthenogenetic viviparous *Aphis gossypii* (Homoptera, Aphididae). *Chromosome Information Service* **30**:4–5.
- Kiauta B.** 1969a. Autosomal fragmentations and fusions in Odonata and their evolutionary implications. *Genetica* **40**:158–180. doi: [10.1007/BF01787347](https://doi.org/10.1007/BF01787347).
- Kiauta B.** 1969b. Sex chromosomes and sex determining mechanisms in Odonata with a review of the cytological conditions in the family Gomphidae and references to the karyotypic evolution in the order. *Genetica* **40**:127–157. doi: [10.1007/BF01787346](https://doi.org/10.1007/BF01787346).
- Kiauta B, Mol AWM.** 1977. Behaviour of the spermatocyte chromosomes of the mayfly, *Cloeon dipterum* (Linnaeus, 1761) s.1. (Ephemeroptera: Baetidae) with a note on the cytology of the order. *Genen en phaenen* **19**:31–39.
- Koval TM.** 1996. Moths: myths and mysteries of stress resistance. *Bioessays* **18**:149–156. doi: [10.1002/bies.950180211](https://doi.org/10.1002/bies.950180211).
- Kuznetsova VG, Nokkala S, Shcherbakov DE.** 2002. Karyotypes, reproductive organs, and pattern of gametogenesis in *Zorotypus hubbardi* Caudell (Insecta: Zoraptera, Zorotypidae), with discussion on relationships of the order. *Canadian Journal of Zoology* **80**:1047–1054. doi: [10.1139/z02-074](https://doi.org/10.1139/z02-074).
- Kuznetsova VG, Sapunov VB.** 1987. Effect of X-rays on the morphological and karyological inconstancy of aphids. Population Structure, Genetics and Taxonomy of Aphids and Thysanoptera. *Proceedings of the International Symposia held at Smolenice Czechoslovakia* **1985**:134–138.
- LaChance LE, Graham CK.** 1984. Insect radiosensitivity: Dose curves and dose fractionation studies of dominant lethal mutations in the mature sperm of 4 insect species. *Mutational Research* **127**:49–59. doi: [10.1016/0027-5107\(84\)90139-8](https://doi.org/10.1016/0027-5107(84)90139-8).
- Langmead B, Salzberg SL.** 2012. Fast gapped-read alignment with Bowtie 2. *Nature Methods* **9**:357–359. doi: [10.1038/nmeth.1923](https://doi.org/10.1038/nmeth.1923).

- Maddox PS**, Oegema K, Desai A, Cheeseman IM. 2004. 'Holo'er than thou: chromosome segregation and kinetochore function in *C. elegans*. *Chromosome Research* **12**:641–653. doi: [10.1023/B:CHRO.0000036588.42225.2f](https://doi.org/10.1023/B:CHRO.0000036588.42225.2f).
- Malik HS**, Henikoff S. 2003. Phylogenomics of the nucleosome. *Nature Structural Biology* **10**:882–891. doi: [10.1038/nsb996](https://doi.org/10.1038/nsb996).
- Malik HS**, Henikoff S. 2009. Major evolutionary transitions in centromere complexity. *Cell* **138**:1067–1082. doi: [10.1016/j.cell.2009.08.036](https://doi.org/10.1016/j.cell.2009.08.036).
- Mandrioli M**, Azzoni P, Lombardo G, Manicardi GC. 2011. Composition and epigenetic markers of heterochromatin in the aphid *Aphis nerii* (Hemiptera: Aphididae). *Cytogenetic and Genome Research* **133**:67–77. doi: [10.1159/000323510](https://doi.org/10.1159/000323510).
- Manicardi GC**, Mandrioli M, Bizzaro D, Bianchi U. 2002. Cytogenetic and molecular analysis of heterochromatic areas in the holocentric chromosomes of different aphid species. *Some Aspects of Chromosome Structure and Functions* 47–56.
- Marec F**, Tothova A, Sahara K, Traut W. 2001. Meiotic pairing of sex chromosome fragments and its relation to atypical transmission of a sex-linked marker in *Ephestia kuehniella* (Insecta: Lepidoptera). *Heredity* **87**:659–671. doi: [10.1046/j.1365-2540.2001.00958.x](https://doi.org/10.1046/j.1365-2540.2001.00958.x).
- Mediouni J**, Fukova I, Frydrychova R, Dhoubi MH, Marec F. 2004. Karyotype, sex chromatin and sex chromosome differentiation in the carob moth, *Ectomyelois ceratoniae* (Lepidoptera: Pyralidae). *Caryologia* **57**:184–194. doi: [10.1080/00087114.2004.10589391](https://doi.org/10.1080/00087114.2004.10589391).
- Melters DP**, Paliulis LV, Korf IF, Chan SW. 2012. Holocentric chromosomes: convergent evolution, meiotic adaptations, and genomic analysis. *Chromosome Research* **20**:579–593. doi: [10.1007/s10577-012-9292-1](https://doi.org/10.1007/s10577-012-9292-1).
- Mon H**, Izumi M, Mitsunobu H, Tatsuke T, Iiyama K, Jikuya H, Lee JM, Kusakabe T. 2011. Post-translational modifications of the N-terminal tail of histone H3 in holocentric chromosomes of *Bombyx mori*. *Insect Biochemistry and Molecular Biology* **41**:902–908. doi: [10.1016/j.ibmb.2011.08.004](https://doi.org/10.1016/j.ibmb.2011.08.004).
- Mon H**, Lee JM, Mita K, Goldsmith MR, Kusakabe T. 2014. Chromatin-induced spindle assembly plays an important role in metaphase congression of silkworm holocentric chromosomes. *Insect Biochemistry and Molecular Biology* **45**:40–50. doi: [10.1016/j.ibmb.2013.11.007](https://doi.org/10.1016/j.ibmb.2013.11.007).
- Monen J**, Maddox PS, Hyndman F, Oegema K, Desai A. 2005. Differential role of CENP-A in the segregation of holocentric *C. elegans* chromosomes during meiosis and mitosis. *Nature Cell Biology* **7**:1248–1255. doi: [10.1038/ncb1331](https://doi.org/10.1038/ncb1331).
- Murakami A**, Imai HT. 1974. Cytological evidence for holocentric chromosomes of the silkworms, *Bombyx mori* and *B. mandarina*, (Bombycidae, Lepidoptera). *Chromosoma* **47**:167–178. doi: [10.1007/BF00331804](https://doi.org/10.1007/BF00331804).
- Negre N**, Hotelier T, Volkoff AN, Gimenez S, Cousserans F, Mita K, Sabau X, Rocher J, Lopez-Ferber M, d'Alencon E, Audant P, Sabourault C, Bidegainberry V, Hilliou F, Fournier P. 2006. Spodobase: an EST database for the lepidopteran crop pest Spodoptera. *BMC Bioinformatics* **7**:1471–210. doi: [10.1186/1471-2105-7-322](https://doi.org/10.1186/1471-2105-7-322).
- Nokkala S**, Laukkanen A, Nokkala C. 2002. Mitotic and meiotic chromosomes in *Somatochlora metallica* (Cordulidae, Odonata). The absence of localized centromeres and inverted meiosis. *Heredity* **136**:7–12. doi: [10.1034/j.1601-5223.2002.1360102.x](https://doi.org/10.1034/j.1601-5223.2002.1360102.x).
- Nurk S**, Bankevich A, Antipov D, Gurevich AA, Korobeynikov A, Lapidus A, Pribelski AD, Pyshkin A, Sirotkin A, Sirotkin Y, Stepanauskas R, Clingenpeel SR, Woyke T, McLean JS, Lasken R, Tesler G, Alekseyev MA, Pevzner PA. 2013. Assembling single-cell genomes and mini-metagenomes from chimeric MDA products. *Journal of Computational Biology* **20**:714–737. doi: [10.1089/cmb.2013.0084](https://doi.org/10.1089/cmb.2013.0084).
- Oksala T**. 1943. Zytologische Studien an Odonaten. *Annales Academiae scientiarum Fennicae: Biologica* **4**:1–64.
- Ortiz E**. 1969. Chromosomes and meiosis in dermaptera. *Chromosomes today* **2**:33–40.
- 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. *Proceedings of the National Academy of Sciences of USA* **88**:3734–3738. doi: [10.1073/pnas.88.9.3734](https://doi.org/10.1073/pnas.88.9.3734).
- Panchenko T**, Black BE. 2009. The Epigenetic basis for centromere identity. *Progress in Molecular and Subcellular Biology* **48**:1–32. doi: [10.1007/978-3-642-00182-6_1](https://doi.org/10.1007/978-3-642-00182-6_1).
- Perez R**, Panzera F, Page J, Suja JA, Rufas JS. 1997. Meiotic behaviour of holocentric chromosomes: orientation and segregation of autosomes in *Triatoma infestans* (Heteroptera). *Chromosome Research* **5**:47–56. doi: [10.1023/A:1018493419208](https://doi.org/10.1023/A:1018493419208).
- Perpelescu M**, Fukagawa T. 2011. The ABCs of CENPs. *Chromosoma* **120**:425–446. doi: [10.1007/s00412-011-0330-0](https://doi.org/10.1007/s00412-011-0330-0).
- Pontecorvo G**. 1944. Synchronous mitosis and differentiation, sheltering the germ track. *Drosophila Information Service* **18**:54–55.
- Przewlaka 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**:e478. doi: [10.1371/journal.pone.0000478](https://doi.org/10.1371/journal.pone.0000478).
- Rishi S**, Rishi KK. 1978. Occurrence of monocentric chromosomes in *Pieris brassicae* L. (Lepidoptera, Pieridae). *Specialia* 451–452.
- Ross BD**, Rosin L, Thomae AW, Hiatt MA, Vermaak D, de la Cruz AF, Imhof A, Mellone BG, Malik HS. 2013. Stepwise evolution of essential centromere function in a *Drosophila* neogene. *Science* **340**:1211–1214. doi: [10.1126/science.1234393](https://doi.org/10.1126/science.1234393).
- Roulin AC**, Wu M, Pichon S, Arbore R, Kuhn-Buhlmann S, Kölliker M, Walser JC. 2014. De novo transcriptome hybrid assembly and validation in the European earwig (Dermaptera, *Forficula auricularia*). *PLOS ONE* **9**:e94098. doi: [10.1371/journal.pone.0094098](https://doi.org/10.1371/journal.pone.0094098).

- Savard J**, Tautz D, Richards S, Weinstock GM, Gibbs RA, Werren JH, Tettelin H, Lercher MJ. 2006. Phylogenomic analysis reveals bees and wasps (Hymenoptera) at the base of the radiation of Holometabolous insects. *Genome Research* **16**:1334–1338. doi: [10.1101/gr.5204306](https://doi.org/10.1101/gr.5204306).
- Schleiffer A**, Maier M, Litos G, Lampert F, Hornung P, Mechtler K, Westermann S. 2012. CENP-T proteins are conserved centromere receptors of the Ndc80 complex. *Nature Cell Biology* **14**:604–613. doi: [10.1038/ncb2493](https://doi.org/10.1038/ncb2493).
- Schrader F**. 1935. Notes on the mitotic behavior of long chromosomes. *Cytologia* **6**:422–430. doi: [10.1508/cytologia.6.422](https://doi.org/10.1508/cytologia.6.422).
- Schulz MH**, Zerbino DR, Vingron M, Birney E. 2012. Oases: robust de novo RNA-seq assembly across the dynamic range of expression levels. *Bioinformatics* **28**:1086–1092. doi: [10.1093/bioinformatics/bts094](https://doi.org/10.1093/bioinformatics/bts094).
- Seschachar BR**, Bagga S. 1962. Chromosome number and sex determining mechanism in the dragonfly *Hemianax ephippiger* (Burmeister). *Cytologia* **27**:443–449. doi: [10.1508/cytologia.27.443](https://doi.org/10.1508/cytologia.27.443).
- Singh TR**, Saro D, Ali AM, Zheng XF, Du CH, Killen MW, Sachpatzidis A, Wahengbam K, Pierce AJ, Xiong Y, Sung P, Meetei AR. 2010. MHF1-MHF2, a histone-fold-containing protein complex, participates in the Fanconi anemia pathway via FANCM. *Molecular Cell* **37**:879–886. doi: [10.1016/j.molcel.2010.01.036](https://doi.org/10.1016/j.molcel.2010.01.036).
- Soding J**, Biegert A, Lupas AN. 2005. The HHpred interactive server for protein homology detection and structure prediction. *Nucleic Acids Research* **33**:W244–W248. doi: [10.1093/nar/gki408](https://doi.org/10.1093/nar/gki408).
- Soldan T**, Putz M. 2000. Karyotypes of some Central European mayflies (Ephemeroptera) and their contribution to phylogeny of the order. *Acta Societatis Zoologicae Bohemicae* **64**:437–445.
- Steiner FA**, Henikoff S. 2014. Holocentromeres are dispersed point centromeres localized at transcription factor hotspots. *eLife* **3**:e02025. doi: [10.7554/eLife.02025](https://doi.org/10.7554/eLife.02025).
- 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 & Development* **9**:573–586. doi: [10.1101/gad.9.5.573](https://doi.org/10.1101/gad.9.5.573).
- Sullivan KF**, Hechenberger M, Masri K. 1994. Human CENP-A contains a histone H3 related histone fold domain that is required for targeting to the centromere. *The Journal of Cell Biology* **127**:581–592. doi: [10.1083/jcb.127.3.581](https://doi.org/10.1083/jcb.127.3.581).
- Sun RY**, Robinson AG. 1965. Chromosome studies of 50 species of aphids. *Canadian Journal of Zoology* **44**:649–653. doi: [10.1139/z66-063](https://doi.org/10.1139/z66-063).
- Suomalainen E**. 1966. Achiastatische Oogenese bei Trichopteren. *Chromosoma* **18**:201–207. doi: [10.1007/BF00326868](https://doi.org/10.1007/BF00326868).
- Talbert PB**, Bryson TD, Henikoff S. 2004. Adaptive evolution of centromere proteins in plants and animals. *Journal of Biology* **3**:18. doi: [10.1186/jbiol11](https://doi.org/10.1186/jbiol11).
- Talbert PB**, Henikoff S. 2010. Histone variants—ancient wrap artists of the epigenome. *Nature Reviews Molecular Cell Biology* **11**:264–275. doi: [10.1038/nrm2861](https://doi.org/10.1038/nrm2861).
- Talbert PB**, Masuelli R, Tyagi AP, Comai L, Henikoff S. 2002. Centromeric localization and adaptive evolution of an Arabidopsis histone H3 variant. *The Plant Cell* **14**:1053–1066. doi: [10.1105/tpc.010425](https://doi.org/10.1105/tpc.010425).
- Terry MD**, Whiting MF. 2005. Mantophasmatodea and phylogeny of the lower neopterous insects. *Cladistics* **21**:240–257. doi: [10.1111/j.1096-0031.2005.00062.x](https://doi.org/10.1111/j.1096-0031.2005.00062.x).
- Tombesi ML**, Papeschi AG. 1993. Meiosis in *Haematopinus suis* and *Menacanthus stramineus* (Phthiraptera, Insecta). *Hereditas* **119**:31–38. doi: [10.1111/j.1601-5223.1993.00031.x](https://doi.org/10.1111/j.1601-5223.1993.00031.x).
- Tombesi ML**, Papeschi AG, Mola LM. 1999. Spermatogenesis in *Bovicola limbata gervais*, 1844 and *B. caprae* Gurtl, 1843 (Phthiraptera, Ischnocera). *Cytologia* **64**:25–27. doi: [10.1508/cytologia.64.25](https://doi.org/10.1508/cytologia.64.25).
- Traut W**, Mosbacher GC. 1968. [Sex chromatin in Lepidoptera]. *Chromosoma* **25**:343–356. doi: [10.1007/BF01183125](https://doi.org/10.1007/BF01183125).
- Viera A**, Page J, Rufas JS. 2009. Inverted meiosis: the true bugs as a model to study. *Genome Dynamics* **5**:137–156. doi: [10.1159/000166639](https://doi.org/10.1159/000166639).
- Webb GC**. 2009. Types of chromosomes in insects. In: Resh VH, Carde RT. editors. *Encyclopedia of Insects*. p. 158–165.
- Westermann S**, Schleiffer A. 2013. Family matters: structural and functional conservation of centromere-associated proteins from yeast to humans. *Trends in Cell Biology* **23**:260–269. doi: [10.1016/j.tcb.2013.01.010](https://doi.org/10.1016/j.tcb.2013.01.010).
- Westhorpe FG**, Straight AF. 2013. Functions of the centromere and kinetochore in chromosome segregation. *Current Opinion in Cell Biology* **25**:334–340. doi: [10.1016/j.ceb.2013.02.001](https://doi.org/10.1016/j.ceb.2013.02.001).
- White MJD**. 1971. The chromosomes of *Hemimerus bouvieri* Chopard (Dermaptera). *Chromosoma* **34**:183–189. doi: [10.1007/BF00285185](https://doi.org/10.1007/BF00285185).
- Whiting MF**. 2002. Phylogeny of the holometabolous insect orders based on 18S ribosomal DNA: when bad things happen to good data. *EXS* **69**:69–83.
- Wolf KW**. 1994. The unique structure of Lepidopteran spindles. *International Review of Cytology* **152**:1–48.
- Wolf KW**, Novac K, Marec F. 1997. Kinetic organization of metaphase I bivalents in spermatogenesis of Lepidoptera and Trichoptera species with small chromosome numbers. *Heredity* **79**:135–143. doi: [10.1038/hdy.1997.136](https://doi.org/10.1038/hdy.1997.136).
- Wolfe SL**, John B. 1965. The organization and ultrastructure of male meiotic chromosomes in *Oncopeltus fasciatus*. *Chromosoma* **17**:85–103. doi: [10.1007/BF00330076](https://doi.org/10.1007/BF00330076).
- Xia Q**, Zhou Z, Lu C, Cheng D, Dai F, Li B, Zhao P, Zha X, Cheng T, Chai C, Pan G, Xu J, Liu C, Lin Y, Qian J, Hou Y, Wu Z, Li G, Pan M, Li C, Shen Y, Lan X, Yuan L, Li T, Xu H, Yang G, Wan Y, Zhu Y, Yu M, Shen W, Wu D, Xiang Z, Yu J, Wang J, Li R, Shi J, Li H, Li G, Su J, Wang X, Li G, Zhang Z, Wu Q, Li J, Zhang Q, Wei N, Xu J, Sun H, Dong L, Liu D, Zhao S, Zhao X, Meng Q, Lan F, Huang X, Li Y, Fang L, Li C, Li D, Sun Y, Zhang Z, Yang Z, Huang Y, Xi Y, Qi Q, He D, Huang H, Zhang X, Wang Z, Li W, Cao Y, Yu Y, Yu H, Li J, Ye J, Chen H, Zhou Y, Liu B, Wang J,

- Ye J, Ji H, Li S, Ni P, Zhang J, Zhang Y, Zheng H, Mao B, Wang W, Ye C, Li S, Wang J, Wong GK, Yang H, Biology Analysis Group. 2004. A draft sequence for the genome of the domesticated silkworm (*Bombyx mori*). *Science* **306**:1937–1940. doi: [10.1126/science.1102210](https://doi.org/10.1126/science.1102210).
- Yoda K**, Ando S, Morishita S, Houmura K, Hashimoto K, Takeyasu K, Okazaki T. 2000. Human centromere protein A (CENP-A) can replace histone H3 in nucleosome reconstitution in vitro. *Proceedings of the National Academy of Sciences of USA* **97**:7266–7271. doi: [10.1073/pnas.130189697](https://doi.org/10.1073/pnas.130189697).
- Zerbino DR**, Birney E. 2008. Velvet: algorithms for de novo short read assembly using de Bruijn graphs. *Genome Research* **18**:821–829. doi: [10.1101/gr.074492.107](https://doi.org/10.1101/gr.074492.107).
- Zhen Y**, Aardema ML, Medina EM, Schumer M, Andolfatto P. 2012. Parallel molecular evolution in an herbivore community. *Science* **337**:1634–1637. doi: [10.1126/science.1226630](https://doi.org/10.1126/science.1226630).