Journal of Biology

BioMed Central

Minireview

Who's driving the centromere?

Gregory P Copenhaver

Address: Department of Biology and The Carolina Center for Genome Sciences, The University of North Carolina at Chapel Hill, Chapel Hill, NC 27599, USA. E-mail: gcopenhaver@bio.unc.edu

Published: 11 October 2004

Journal of Biology 2004, **3:**17

The electronic version of this article is the complete one and can be found online at http://jbiol.com/content/3/4/17

Abstract

Centromere function is remarkably conserved between species, yet the satellite sequences that make up centromeric DNA are highly divergent. Proteins that bind these sequences appear to be evolving under positive selection, supporting a model wherein the interplay between centromeric repeats and the proteins that bind them creates an opportunity for an intriguing phenomenon known as centromere-based meiotic drive.

Centromeres are crucial regions on each chromosome that direct several functions that are essential for life. They aggregate the proteins that make up the kinetochore, the chromosomal attachment point for the microtubules that form the meiotic or mitotic spindle, and they are responsible for the proper segregation of chromosomes into daughter cells at each division, as well as ensuring proper pairing between sister chromatids [1]. The DNA sequences that direct these functions are strikingly divergent between even closely related species [2,3]. With the exception of holocentric species, such as *Caenorhabditis elegans*, which assemble a kinetochore along the entire length of each chromosome [4], higher eukaryotic centromeres can be visualized as the site of the primary constriction on metaphase chromosomes. The DNA in these regions is often densely methylated and typically harbors a variety of repetitive sequences including, in most species, large arrays of repeats known as satellite sequences. Until recently, it has been difficult to determine the relative importance of specific sequence classes in centromere function because of the size, complexity, and repetitive content of centromere sequences.

Recent advances, including whole-genome sequencing efforts, chromatin immunoprecipitation using kinetochore proteins, and tetrad analysis in the model plant *Arabidopsis thaliana*, have bolstered the belief that centromere satellite repeats may be functionally important sequences [5-8]. This focus on satellite repeats presents an intriguing paradox. As more examples of satellite repeats are gathered it has become apparent that these sequences, which are involved in a highly conserved function necessary for each cell division, are remarkably different between different species. *A priori* one would expect each member of a repetitive sequence array to acquire mutations independently and diverge from one another just as single-copy sequences diverge from one another between species. But despite the fact that centromere satellites are so divergent between species, they show a high level of sequence conservation between the repeats within a species. The same phenomenon was first observed in clustered multi-gene families, and is now referred to as concerted evolution [9,10]. Presumably, variants that arise in any repetitive class of DNA through mutational processes can be either eliminated or spread and homogenized by mechanisms that include unequal sister-chromatid exchange, gene conversion, replication slippage due to misalignment of repeats, and replicative transposition of transposable elements [11]. While this explanation satisfies the observed patterns of variation of

centromeric repeats, it merely emphasizes the central paradox of centromere biology. How do DNA repeats that are inherently prone to rapid sequence divergence between species successfully interact with a reasonably well-conserved set of kinetochore proteins to execute a critical cellular function?

An important piece of this puzzle may come from understanding how kinetochore proteins evolve in concert with the centromere sequences. Along with a better understanding of centromere sequences has come a better understanding of the proteins that interact with them [12]. Two important centromere-associated proteins are centromere proteins A (CENP-A) and C (CENP-C). Both have highly conserved domains with clear homology between the yeast and human counterparts, both have been shown by chromatin immunoprecipitation to bind directly to centromeric satellite repeat sequences, and both proteins associate with active but not inactive centromeres. CENP-A, also called CenH3, is a centromere-specific histone H3 variant and CENP-C is an integral component of the inner kinetochore [13,14].

As reported in this journal [15], Talbert and colleagues set out to examine how CENP-A and CENP-C have evolved in light of their association with rapidly diverging centromere DNA. Most proteins that are both ancient and functionally conserved are expected to show little variation between species, given that there should be strong selection against changing a well-tuned molecule involved in a critical function. By comparison, selectively neutral or unconstrained genes are expected to show significant sequence variation between distantly related species. Kimura's neutral theory [16] predicts that the rate of effectively neutral mutation for DNA sequence changes leading to nonsynonymous amino-acid changes in the encoded protein is expected to be proportional to the rate of synonymous changes. The ratio of these rates - denoted K_a and K_s, respectively - is commonly used as a metric for testing the assumption of neutral selection. If the K_a/K_s ratio is 1 then the protein is behaving according to neutral theory and selection is unlikely to be an important evolutionary force acting on it. More frequently, K_a is significantly less than K_s because a mutation that changes the amino-acid sequence of a protein is much less likely to be fixed between two species than one that is silent: nonsynonymous changes are typically deleterious and are therefore expected to be eliminated by selection; evolutionary biologists often refer to this as negative or purifying selection. More rarely, the K_a/K_s ratio is larger than 1, usually indicating that selection has acted to change the protein; this is referred to as positive or adaptive selection [17].

Previous work by the Henikoff lab [18] had revealed that the CENP-A homologs in *Drosophila melanogaster* and *Arabidopsis* are evolving adaptively ($K_a/K_s > 1$). This was a satisfying finding: the adaptive nature of CENP-A could be proposed to provide a bridge between the rapidly diverging centromere repeats and the conserved cellular machinery with which these sequences interact (Figure 1a). Surprisingly, in the current study [15], the CENP-A homologs from mammals and from grasses showed evidence of negative selection ($K_a/K_s < 1$). How do these organisms achieve the same bridge between diverging centromere repeats and conserved cellular machinery when the evolution of their CENP-A homologs appears to be constrained? To answer this question, Talbert *et al.* [15] looked at CENP-C, the other ancient kinetochore protein that has a conserved core domain. Using the same type of analysis, they show that significant portions of both the mammalian and grass CENP-C proteins show positive selection. Notably, the regions of CENP-C that showed evidence for positive selection included the DNA-binding domain. Thus it appears that different lineages have solved the need to have an evolutionary flexible kinetochore component in two distinct ways, both utilizing proteins that directly interface with the highly divergent centromeric satellites (Figure 1a) [15].

Talbert and colleagues note that the tendency of the centromere DNA to rapidly diverge, combined with the ability of CENP-A and CENP-C to evolve adaptively, creates an opportunity for a phenomenon known as meiotic drive. Meiotic drive describes any mechanism that favors the inheritance of one nuclear component (gene, centromere, or chromosome, for example) over another [19,20]. Female meiosis provides a particularly rich occasion for drive, as only one of four meiotic products typically develops into a gamete in higher eukaryotes; the other three degenerate. The phenomena that allow concerted evolution (mutation, unequal exchange, and gene conversion) could also create arrays of satellite sequences with different numbers of repeats, and/or repeats with different affinities for the proteins that bind them. Thus, during female meiosis there may be 'stronger' and 'weaker' centromeres, and meiotic drive could occur if stronger centromeres are preferentially segregated to the egg (Figure 1b). This process would tend to generate stronger and stronger centromeres. Given that all the centromeres of all four meiotic products survive male meiosis, however, weak centromeres would also be maintained in the population. Thus female-specific meiotic drive of centromeres could potentially create significant imbalances in the strength of centromeres present during male meioses. Such imbalances might lead to inefficient divisions. To mitigate this imbalance, CENP-A and CENP-C could adaptively evolve in a manner that abrogates the imbalance during male meiosis (Figure 1c). In an elegant *coup de grace* the authors demonstrate that both CENP-A and CENP-C homologs in the budding yeast *Saccharomyces*

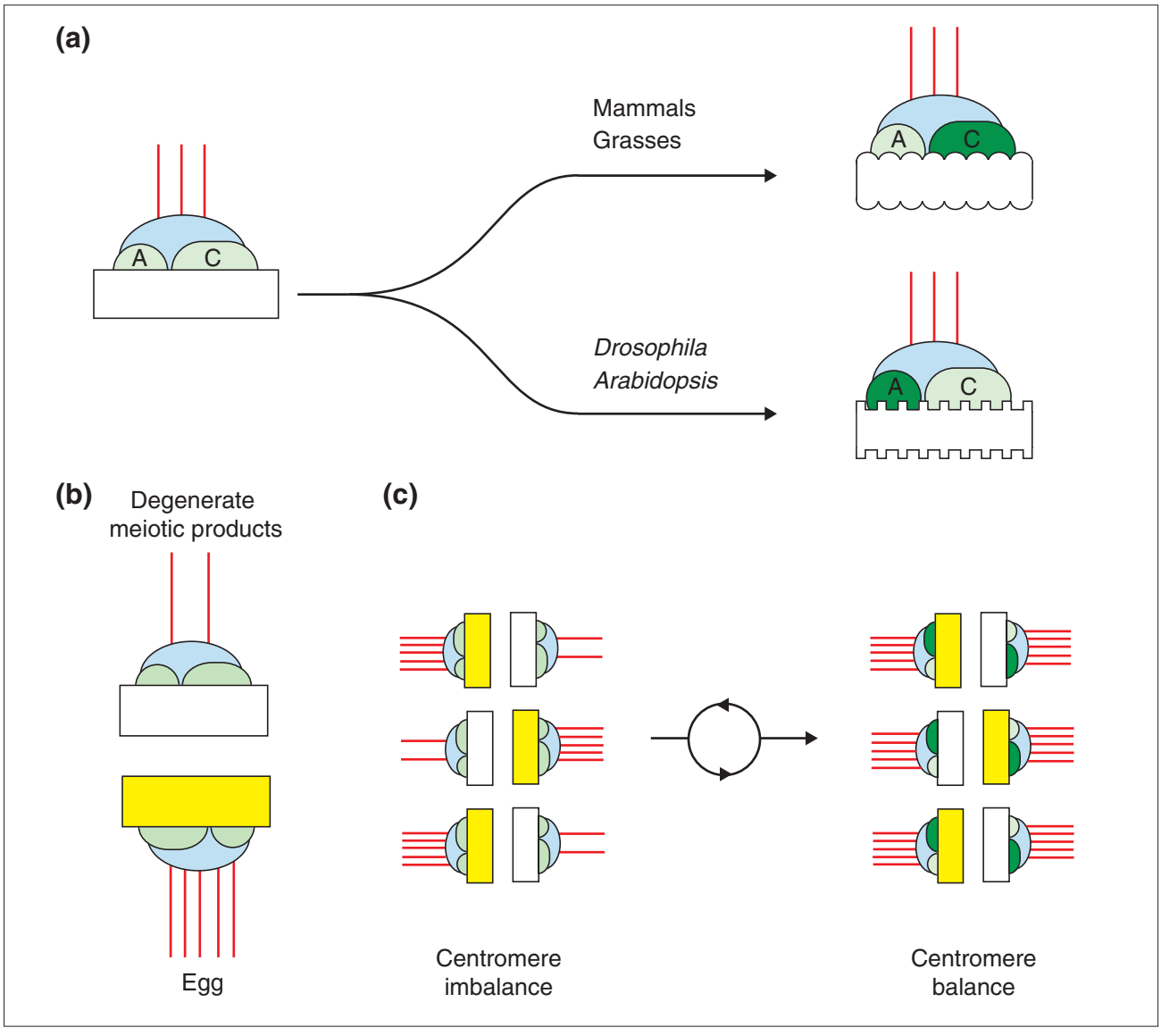

Figure 1
Centromere DNA sequences diverge rapidly and kinetochore components evolve adaptively. **(a)** Centromeres (white boxes) nucleate a specialized set of proteins called the kinetochore (blue), including CENP-A and CENP-C (green), which in turn interact with the spindle microtubules (red lines). As centromere sequences diverge, either CENP-A or CENP-C has evolved adaptively (darker green) in different lineages. **(b)** Changes in the sequence or organization of centromere DNA (yellow boxes) may create stronger centromeres (indicated, for convenience, by a greater number of microtubule interactions). Selective segregation of stronger centromeres into egg cells could, theoretically, lead to meiotic drive (see text for details). **(c)** Meiotic drive in female meioses could generate imbalances in subsequent male meioses. Such imbalances might in turn be neutralized by adaptive evolution (circular arrows) of kinetochore proteins such as CENP-A or CENP-C.

cerevisiae show signs of negative selection. This fits nicely with the meiotic centromere drive model as all four products normally survive each meiosis in *S. cerevisiae*.

While this model is attractive, many aspects remain to be elucidated. Is there a mechanism that selectively promotes the segregation of stronger centromeres into the egg during female meiosis? If so, how does the selfish nature of centromere meiotic drive balance with the complement of alleles linked to a centromere - in other words, what happens if a strong centromere drives the egg to propagate an undesirable haplotype? How might CENP-A and CENP-C execute their

mitigating role in male meiosis? The fact that the exciting developments in the article [15] raise so many intriguing questions simply emphasizes the complexity and beauty of centromere biology. This study in particular highlights the need to integrate molecular, bioinformatic and evolutionary approaches in multiple organisms when studying how centromeres and kinetochores interact.

References

1. Choo KHA: *The Centromere*. Oxford: Oxford University Press; 1997.
2. Waye JS, Willard HF: **Concerted evolution of alpha satellite DNA: evidence for species specificity and a general lack of sequence conservation among alphoid sequences of higher primates.** *Chromosoma* 1989, **98:**273-279.
3. Hall SE, Kettler G, Preuss D: **Centromere satellites from *Arabidopsis* populations: maintenance of conserved and variable domains.** *Genome Res* 2003, **13:**195-205.
4. Dernburg AF: **Here, there, and everywhere: kinetochore function on holocentric chromosomes.** *J Cell Biol* 2001, **153:**F33-F38.
5. Nagaki K, Talbert PB, Zhong CX, Dawe RK, Henikoff S, Jiang J: **Chromatin immunoprecipitation reveals that the 180-bp satellite repeat is the key functional DNA element of *Arabidopsis thaliana* centromeres.** *Genetics* 2003, **163:**1221-1225.
6. Politi V, Perini G, Trazzi S, Pliss A, Raska I, Earnshaw WC, Della Valle G: **CENP-C binds the alpha-satellite DNA *in vivo* at specific centromere domains.** *J Cell Sci* 2002, **115:**2317-2327.
7. Copenhaver GP, Nickel K, Kuromori T, Benito MI, Kaul S, Lin X, Bevan M, Murphy G, Harris B, Parnell LD, *et al.*: **Genetic definition and sequence analysis of *Arabidopsis* centromeres.** *Science* 1999, **286:**2468-2474.
8. Meluh PB, Yang P, Glowczewski L, Koshland D, Smith MM: **Cse4p is a component of the core centromere of *Saccharomyces cerevisiae*.** *Cell* 1998, **94:**607-613.
9. Brown DD, Wensink PC, Jordan E: **A comparison of the ribosomal DNAs of *Xenopus laevis* and *Xenopus mulleri*: the evolution of tandem genes.** *J Mol Biol* 1972, **63:**57-73.
10. Zimmer EA, Martin SL, Beverley SM, Kan YW, Wilson AC: **Rapid duplication and loss of genes coding for the alpha chains of hemoglobin.** *Proc Natl Acad Sci USA* 1980, **77:**2158-2162.
11. Li W-H: *Molecular Evolution*. Sunderland: Sinauer Associates; 1997.
12. Cleveland DW, Mao Y, Sullivan KF: **Centromeres and kinetochores: from epigenetics to mitotic checkpoint signaling.** *Cell* 2003, **112:**407-421.
13. Earnshaw WC, Cooke CA: **Proteins of the inner and outer centromere of mitotic chromosomes.** *Genome* 1989, **31:**541-552.
14. Smith MM: **Centromeres and variant histones: what, where, when and why?** *Curr Opin Cell Biol* 2002, **14:**279-285.
15. Talbert PB, Bryson TD, Henikoff S: **Adaptive evolution of centromere proteins in plants and animals.** *J Biol* 2004, **3:**18.
16. Kimura M: **Evolutionary rate at the molecular level.** *Nature* 1968, **217:**624-626.
17. Li WH, Wu CI, Luo CC: **A new method for estimating synonymous and nonsynonymous rates of nucleotide substitution considering the relative likelihood of nucleotide and codon changes.** *Mol Biol Evol* 1985, **2:**150-174.
18. Malik HS, Henikoff S: **Adaptive evolution of Cid, a centromere-specific histone in *Drosophila*.** *Genetics* 2001, **157:**1293-1298.
19. Zimmering S, Sandler L, Nicoletti B: **Mechanisms of meiotic drive.** *Annu Rev Genet* 1970, **4:**409-436.
20. Sandler L, Novitski E: **Meiotic drive as an evolutionary force.** *Am Nat* 1957, **91:**105-110.