Heslop-Harrison JS, Schwarzacher T. 2013. Nucleosomes and centromeric DNA packaging. Proc Nat Acad Sci USA. www.pnas.org/cgi/doi/10.1073/pnas.1319945110 http://dx.doi.org/10.1073/pnas.1319945110. See also http://www.pnas.org/content/early/2013/11/25/1319945110

Commentary – PNAS

Nucleosomes and centromeric DNA packaging

~~Alternative titles:~~
~~Wrapping up histones: packaging DNA in centromeric nucleosomes~~
~~OR~~
~~Phasing and satellite DNA: packaging in centromeric nucleosomes~~
~~OR~~
~~Phasing of CENH3: satellite DNA shapes the centromeric nucleosome landscape~~
~~OR~~
~~It's a wrap for CENH3 in maintaining kinetochore stability~~

JS Heslop-Harrison and Trude Schwarzacher, University of Leicester, UK
Phh4@le.ac.uk and ts32@le.ac.uk

The eukaryotic chromosome is a conserved structure, with the DNA double-helix wrapping around octamers of histone proteins to form the chromatin which is further packaged into chromosomes. The centromere defines the kinetochore, the region of spindle microtubule attachment that pulls the two replicated chromatids of each chromosome apart during cell division (Fig. 1) leading to fidelity in transmission of genetic information. Like the telomere, centromeres are well defined morphologically and functionally, but their DNA sequence shows no conservation between species and their coiling into chromatin is still poorly understood. In PNAS, Zhang et al.(1) give new insight into the CenH3- nucleosomes associated with the rice *CentO* satellite sequence and adjacent regions of DNA~~. The work builds on the detailed studies from the labs of both corresponding authors in assembly of centromeric DNA sequences and protein characterization with~~ having implications for centromere specification, activity and evolution.

Centromeres are readily observed as partial constrictions on the metaphase chromosome by light or electron microscopy (Fig. 1 A,B)~~, and, like the rest of the chromatin, consists of DNA and associated proteins~~. However, few sequence-related characteristics of the DNA at the centromeres are conserved, contrasting with the telomeres or other morphological features of chromosomes such as the rRNA gene sites~~s at sites of the secondary constrictions. There are "exceptions that prove the rule" to most chromosomal features showing that 'solutions' are not unique and can evolve. – For example, *Drosophila* has retroelements inserted near chromosome-ends rather than the telomerase-added~~

~~sequences, while a few plants and animals have holocentric chromosomes where kinetochores are along the whole length of the chromosomes (2). – showing that 'solutions' are not unique and can evolve, but c~~Centromere-defining ~~DNA and related structures have remained enigmatic, despite. In contrast,~~ kinetochores linking centromeres and spindle microtubules with a multiprotein complex involving the histones and other DNA bending proteins, are highly conserved across all kingdoms (3,4).

Typically, there are tandemly repeated satellite DNA sequences at the centromeres of chromosomes of animals and ~~plants – often with similarities sequences on all chromosomes, but showing rapid divergence between even phylogenetically close organisms species~~ (5,6,7). Detailed work on the budding yeast *Schizosaccharomyces pombe* and brewer's yeast *Saccharomyces cerevisiae* identified relatively short DNA sequences and binding protein counterparts that direct chromosome segregation (8,9). ~~From the 1980s to early 2000s, following the demonstration that the DNA sequence itself as is involved in centromere specification in the yeast model systems, t~~The search was then on for critical 'boxes' in plants and animals (5,6). Although some motifs were found to be involved in protein-DNA interactions, no key motif shared across diverse phyla was found within the centromeric tandem repeats.

As sequencing technology advanced, both variants of tandem repeats and almost any other class of DNA – retrotransposons, transposons, genes, transcription factors and microsatellites – were revealed underlying the kinetochore and microtubule attachment sites, and these centromeric and pericentromeric sequences were examined for involvement in centromere specification. Over the years, the evidence for these numerous different sequences as well as their loss or replacement in mutants which still showed centromeric function, the lack of necessity for tandemly repeated satellite sequence, and the presence of facultative neocentromeres forming at unusual points along chromosomes, started to suggest there was no required sequence motif for centromere activity~~. and it became clear that there was no defined sequence for plant and animal centromeres~~.

From the 1980s, the nucleosomal packaging of DNA was being investigated, and in 1991, a protein called CENP-A (Centromere protein A) was found associated exclusively with active centromeres in human cells (10). The protein was characterized as a histone H3 variant, with two monomers replacing those of the non-division-associated or canonical form of H3 present in the histone octamers found in interphase nuclei. It is now clear that presence of the CenH3 histone (terminology here follows 1)) is specifying centromere location independently of DNA sequence – and centromeric function is an epigenetic character of the DNA sequence, where under some conditions a particular DNA sequence associates with the histone and becomes the centromere. Thus, centromere function joins many other aspects of epigenetic control of the nucleus where modification of histone proteins is critical to the DNA behavior (11).

Many centromeric satellite DNA sequences have a repeat motif length of 170 to 180bp, the same as the DNA length forming the 1.67 turns around the histone octamer core (147bp) plus the linker DNA between the nucleosomes, a variable distance but typically 10 to 70bp (Fig. 1F,G). When chromatin is extracted from cells with gentle methods, the DNA remains associated with the histones. The enzyme micrococcal nuclease (MNase) cuts free double-stranded DNA, but not when it is protected by association with the nucleosomal proteins. Digestions and sequence analysis show how tandemly repeated DNA is packaged around nucleosomes (Fig. 1G). For one family of satellites, there is a strictly defined association (phasing) of DNA and nucleosomes with the same sequences in all linkers. By increasing time of digestion with MNase, DNA adjacent to the nucleosome is trimmed, leaving only the core particle (Fig. 1I). The CenH3-nucleosomes occur in multiple blocks

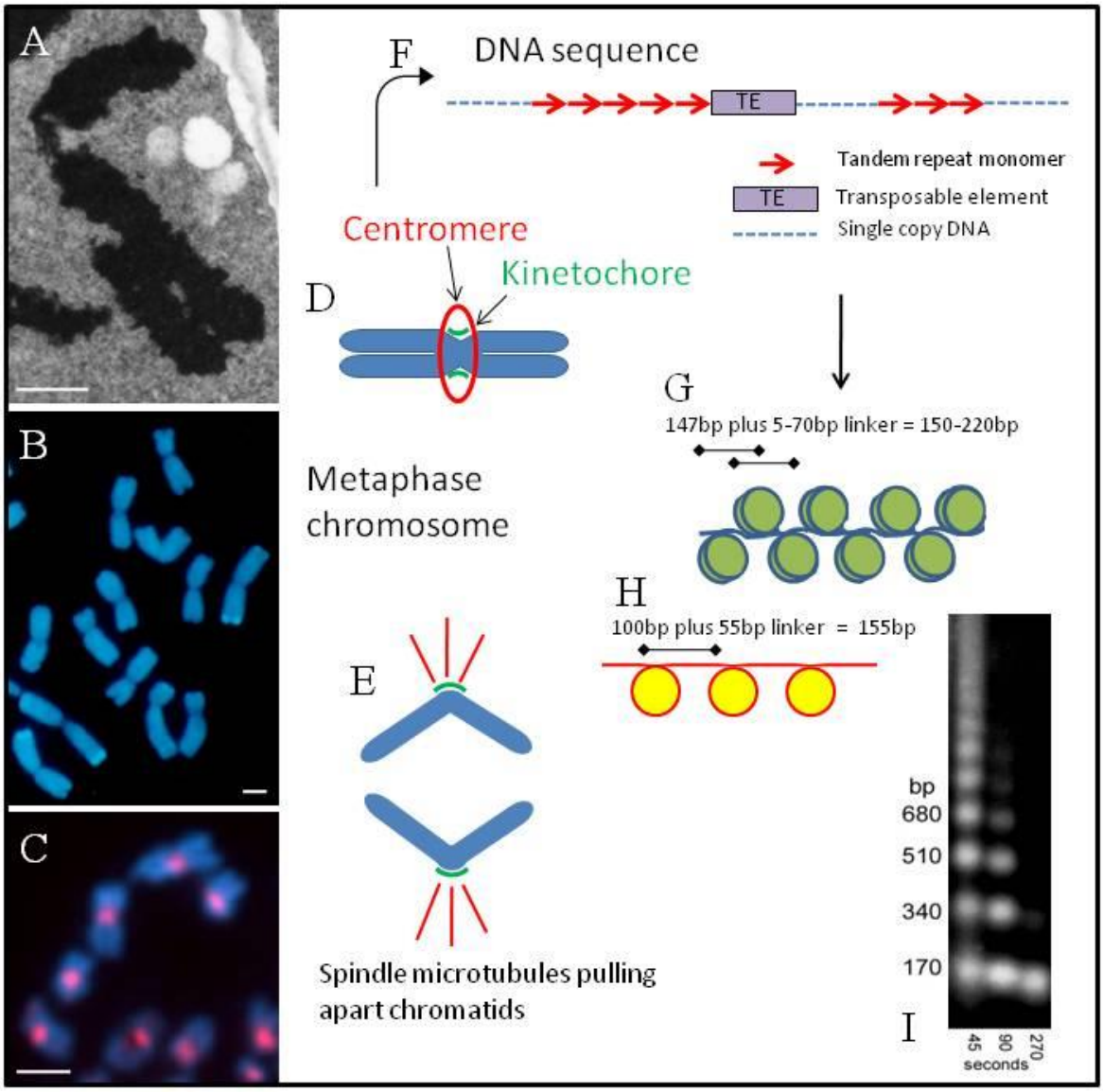
intermingled with nucleosomes with the canonical H3 (12). Some structural models suggest this is due to the positioning of the CenH3 nucleosomes at the periphery of the chromatid, interacting with the kinetochore proteins and the microtubules, while canonical nucleosomes are folded to the inside.

Zhang et al. (1) analyze the CenH3 nucleosome structures associated with the rice centromeric *CentO* satellite sequence and adjacent regions of DNA in unprecedented detail. Rice has proved valuable for understanding centromere structure, where the assembly of the DNA sequence across several centromeres, with exceptional accuracy, has shown the presence of a 155bp long tandemly repeated sequence, *CentO*, with array lengths varying from 65kb to 2,000kb on the different chromosomes, interspersed with transposable element and single-copy sequences (12). Now, the authors digested chromatin to single nucleosomes using the MNase method. They then precipitated the nucleosomes containing CenENH3 by ChIP (immunoprecipitation with a CenH3 antibody), and sequenced the DNA that was associated, thus revealing the length and sequence of the cenH3-packaged nucleosome DNA. Remarkably, only 90 to 100 bp of the *CentO* was protected from digestion on the centromeric nucleosomes, representing only one turn of DNA around the octamer, contrasting with the 147bp typically protected on canonical nucleosomes. The periodicity of the CenH3-nucleosomes was exactly the same as the 155bp length of the satellite, so this set of nucleosomes was in phase with the sequence. A final feature of the analysis was a 10bp periodicity of AT dinucleotides in *CentO*, representing one turn of the DNA double-helix. A 10bp rotational-phasing minimizes the bending energy or strain on the DNA when folded around the nucleosome and hence increases stability (see also 7). The unexpected 100bp result in rice has now also been found in human cells (13) where the CENP-A (CenH3)-nucleosome protects only 110bp of the 171bp long alpha-satellite from MNase digestion. Could there be an alternative packaging structure involving CENH3? Zhang et al. (1) cannot rule out a tetrameric arrangement of histone proteins in CENH3 nucleosomes, reported by Dimitriadis et al. (14), although Hasson et al. (13) conclude that the phased satellite at human centromeres wraps around octameric nucleosomes, with loose and extended spacers.

In rice, Zhang et al. (1) have shown how ChIP-sequencing on a relatively large scale has revealed the phasing and nature of spacing of *CentO* and non-satellite, CenH3 - nucleosomes. This has made a significant advance in our understanding of the DNA:histone variant interaction and the nature of centromeric chromatin packaging, crucial to chromosome structure and the specification of centromeres, assigning a functional role to the abundant but enigmatic class of centromeric satellite DNA. It will now be important to integrate these results with a dynamic picture of cenH3 loading (15). New, super-resolution microscopy methods (16), careful assembly of satellite DNA arrays (9, 17), and comparative studies of different organisms will help with answering the long-standing questions of what defines a centromere. How, and when, is the DNA underlying the centromeric nucleosomes defined in the cell? Why, and how, does this DNA become the target to associate with the centromeric histone variant? These large questions must be partitioned to tractable experiments, and it needs to be considered whether imaging, mutant, sequencing/informatics or biochemical approaches are going to be most informative. Exquisite control of centromere function, and the highest stability of the process, is essential for correct chromosome segregation. This requires both robust and redundant control, includes evolutionary selection for both satellite and other DNA sequences with their interacting counterparts. Application of knowledge of centromere specification and control has potential not only for understanding aneuploidy-related diseases, but also may be exploiteded in generation of new hybrid plant varieties.

1. Zhang et al. The CentO satellite confers translational and rotational phasing on CENH3 nucleosomes in rice centromeres. *Proc Natl Acad Sci USA*.
~~2. Heckmann S, et al. (2011). Holocentric chromosomes of *Luzula elegans* are characterized by a longitudinal centromere groove, chromosome bending, and a terminal nucleolus organizer region. *Cytogenet Genome Res* 134: 220-228.~~
~~3~~2. Talbert PB, et al. (2012) A unified phylogeny-based nomenclature for histone variants *Epigenet Chromatin* 5:7 (19 pages doi:10.1186/1756-8935-5-7).
3. Westermann S, Schleiffer A (2013) Family matters: structural and functional conservation of centromere-associated proteins from yeast to humans. *Trends Cell Biol* 23: 260-269.
4. Henikoff S, Dalal Y (2005) Centromeric chromatin: what makes it unique? *Current Opin Genet Develop* 15:177–184
5. Heslop-Harrison JS, Schwarzacher T (2011) Organisation of the plant genome in chromosomes. *Plant J* 66:18-33.
6. Levitsky VG, Babenko VN, Vershinin AV (2013) The roles of the monomer length and nucleotide context of plant tandem repeats in nucleosome positioning. J Biomol Structure Dynamics, on-line first, http://dx.doi.org/10.1080/07391102.2012.755796
7. Fitzgerald-Hayes MJ, et al. (1982) Nucleotide sequence comparisons and functional analysis of yeast centromere DNAs. *Cell* 29:235-244.
8. Hahnenberger KM, et al. (1991) Identification of DNA regions required for mitotic and meiotic functions within the centromere of *Schizosaccharomyces-pombe* Chromosome-I. *Mol Cell Biol* 11: 2206-2215
9. Nagaki K, et al. (2004) Sequencing of a rice centromere uncovers active genes. *Nature Genet*, *36*: 138-145.
10. Palmer DK, et al. (1991) Purification of the centromere-specific protein CENP-A and demonstration that it is a distinctive histone. *Proc Natl Acad Sci USA* 88:3734-3738
11. Martienssen RA, Chandler VL (2013) Molecular mechanisms of transposon epigenetic regulation. pp 71- 92. In: Plant Transposons and Genome Dynamics in Evolution. NV Fedoroff Ed. John Wiley.
12. Vershinin AV, Heslop-Harrison JS (1998) Comparative analysis of the nucleosomal structure of rye, wheat and their relatives. *Plant Mol Biol* 36: 149-161.
13. Hasson D, et al. (2013) The octamer is the major form of CENP-A nucleosomes at human centromeres. *Nat Struct Mol Biol* 20:687-95.
14. Dimitriadis et al. (2010). Tetrameric organization of vertebrate centromeric nucleosomes. *Proc Natl Acad Sci USA 109:* 1979-1984.
15. Chan FL, et al. (2012) Active transcription and essential role of RNA polymerase II at the centromere during mitosis. *Proc Natl Acad Sci USA* 109: 1979-1984. doi:10.1073/pnas.110870510916. Ribeiro SA, et al. (2010) A super-resolution map of the vertebrate kinetochore. *Proc Natl Acad Sci USA* 107:10484-10489.
17. Kuhn GC, et al. (2011) The 1.688 Repetitive DNA of Drosophila: Concerted evolution at different genomic scales and association with genes. *Mol Biol Evol* 29: 7-11.

Heslop-Harrison JS, Schwarzacher T. 2013. Nucleosomes and centromeric DNA packaging. Proc Nat Acad Sci USA. www.pnas.org/cgi/doi/10.1073/pnas.1319945110 http://dx.doi.org/10.1073/pnas.1319945110.



POWERPOINT VERSION OF FIGURE AVAILABLE FROM WWW.MOLCYT.COM - http://wp.me/p2Ewqp-7h FULLY EDITABLE CC-BY-SA

Fig. 1. Features of centromeric DNA from different viewpoints. (**A)** An electron micrograph of a section of a metaphase chromosome of a wild wheat species, showing two arms with the centromere at the bend. (**B)** Metaphase chromosomes of triticale fluorescing blue in the light microscope. Constrictions at the centromeres are visible on each chromosome, with the two chromatids which will separate as the cell divides. (**C)** Chromosomes from a cell culture line of the model species *Arabidopsis thaliana*, labelled with an antibody labelling a centromeric histone (CENH3) variant. (**D)** A diagram of a metaphase chromosome showing the two arms each of two chromatids, separated at the centromere. (**E)** A diagram of a metaphase chromosome dividing into chromatids which segregate and are pulled by spindle microtubules (red) attached via the kinetochore at the centromere. (**F)** DNA motifs found in many centromeres, with blocks of tandemly repeated satellite DNA monomers interspersed with single copy DNA and transposable elements. (**G)**. A diagram of the packaging of double stranded DNA (blue) into nucleosomes, with 147 bp of DNA wrapping 1.67 times around each octamer of the canonical histone proteins (olive) and fixed phase of the nucleosome within the repeat monomer. (**H)** The novel packaging reported by Zhang et al. (1) with approximately 100bp of the rice *CentO* tandem repeat sequence (red) folding once around the nucleosome core that includes CenH3 (yellow). (**I**). A key method for nucleosome analysis involving micrococcal nuclease (MNase) digestion of chromatin and size separation of the resultant DNA fragments; the enzyme cuts DNA in the linker regions, and over the time course shown isolates more mononucleosomes, and trims overhanging DNA not protected from digestion by the histone proteins (12). Bars in A, B, C 2 µm