




Plant centromeres: genetics, epigenetics and evolution

Ludmila Cristina Oliveira¹ · Giovana Augusta Torres¹ 

Received: 7 October 2016 / Accepted: 26 July 2018 / Published online: 16 August 2018
© Springer Nature B.V. 2018

Abstract

The centromere is a functional locus of the chromosome responsible for chromatid cohesion and segregation in cell division. Usually, the centromeres can be distinguished from the remaining chromosomal regions either in structure, as they are heterochromatic constrictions that divide the chromosome in two arms, or in molecular constitution, as they have an exclusive H3 histone variant and specific DNA sequences. Besides being essential for genetic stability of eukaryotes, centromere is particularly interesting since it plays conserved roles but show high variability on organization and DNA composition. Centromeres are usually composed by satellite repeats and retrotransposons and the sequences can differ even among closely related species. Some unusual configurations containing single copy DNA were also described, including even some transcribed genes. In this review, we discuss molecular constitution, epigenetics and different types of centromere, with emphasis on plant centromeres. We also present recent advances about evolutionary processes involved in origin and differentiation of centromeres.

Keywords Satellite DNA · Retrotransposon · Centromeric DNA · CenH3 · New centromere

Introduction

The functional centromeres of eukaryotic chromosomes, cytologically recognized as primary constrictions, are responsible for cohesion of sister chromatids and regular segregation during cell division, which are essential for genetic stability and development of all organisms [1, 2]. The role of centromeres in segregation lies in the fact that they are the site for kinetochore organization, a proteinaceous structure deposited on centromere surface for spindle fibers binding to promote chromatid or chromosome movements [3].

The centromeres can be simple, also known as point centromere, occupying a region corresponding to one nucleosome, or complex, spanning several megabases of the chromosome [4]. In *Saccharomyces cerevisiae*, the point centromere is a 125 bp region of constriction where a small kinetochore is organized (Fig. 1a) [5]. The complex centromeres can be arranged in two different forms in

eukaryotic chromosomes: located or dispersed [6]. Chromosomes with dispersed centromere, known as holocentrics or holokinetics, do not show constriction and their kinetochore is organized throughout the chromosome (Fig. 1b) [7]. If the chromosome has a located centromere, in most cases this is concentrated in one region where there is a constriction, featuring monocentric chromosomes (Fig. 1c). Eventually, chromosomes with two or more centromeres can be found, presenting, therefore, two or more constrictions [6]. A third type is the meta-polycentric, with three to five well-spaced centromeric sites behaving as a single centromere (Fig. 1d). All of them are located on only one constriction, however elongated and significantly bigger than those observed in monocentric chromosomes [8].

The centromere region activity is determined by an epigenetic modification, the replacement of canonical histone H3 by a specific modified histone H3, called CenH3 (or CENP-A in humans). The centromeric H3 aminoacidic composition is not conserved among species, especially in its tail, but still is a conserved epigenetic mark, unlike the centromeric DNA, which can vary considerably and present a complex organization [reviewed in 1, 9]. The centromeric DNA evolves rapidly and has little similarity among species [10], an unusual feature for structures that perform conserved functions. Efforts have been made to explain why

✉ Giovana Augusta Torres
gatorres@dbi.ufla.br

¹ Laboratório de Citogenética Molecular, Departamento de Biologia, Universidade Federal de Lavras (UFLA), Campus Universitário, CP 3037, Lavras, MG 37200-000, Brazil

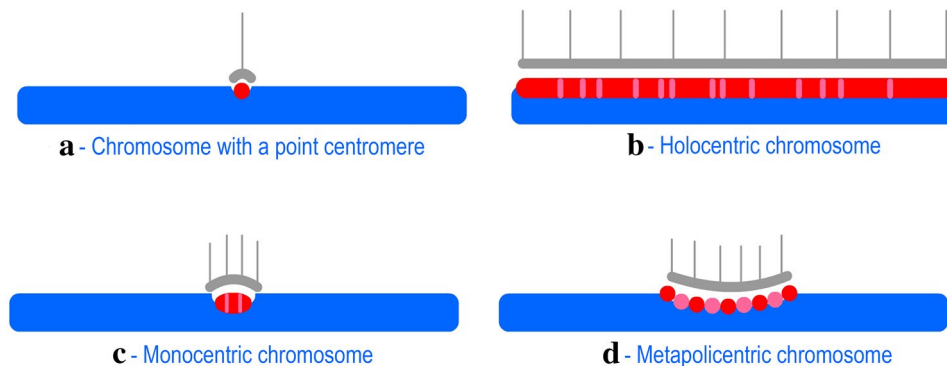


Fig. 1 Graphical representation of centromeric structure and distribution of CenH3 in different types of eukaryotic centromeres: **a** point centromere, **b** holocentromere, **c** monocentromere, **d** metapolycentric

centromere. Elements in blue represent the chromatids, in red the sites with CenH3, in pink the sites with canonical H3, and in gray horizontally and vertically the kinetochore and microtubules, respectively

these variations happen and to better understand the epigenetic aspects involved in centromere function. We present here a review of centromeres, focused on elucidation of plant centromere genetics, epigenetics and evolution.

Epigenetics of centromeres

The centromere identity and inheritance are associated with chromatin organization, that differs from other chromosomal domains by the presence of a histone H3 variant called CenH3 or CENP-A. The difference between the canonical histone H3 and CenH3 is the presence in the latter of an N-terminal tail that is highly divergent with respect to the length and composition, even among closely related organisms, been species-specific in some cases [revised in 1]. Otherwise the C-terminal portion of CenH3 is similar in all eukaryotes and has significant similarity to the canonical H3 [11]. The divergence between the canonical H3 and CenH3 goes beyond the amino acids differences, since the last one also undergoes different types of epigenetic modifications, such as phosphorylation, methylation and acetylation, which contribute to the recruitment of proteins which creates the centromere–kinetochore interface in humans [12]. Recently a new epigenetic modification, the phosphorylation of threonine 120 of histone H2A, was found to be a universal marker for centromeres, working from humans to plants monocentric and polycentric chromosomes [13].

Located exclusively in the centromeres of all studied eukaryotes, the CenH3 is a mark of functionality of this region [revised in 1, 9, 14]. Its associated chromatin is a constituent part of the kinetochore plate which interacts with numerous other proteins to promote segregation [2]. Despite showing a variant specific H3, the centromeric chromatin usually also have H3 nucleosomes interspersed with CenH3 nucleosomes (Fig. 1c) [15].

In addition to the constant presence of CenH3 in active centromeres [16], several other evidences support that this variant histone is the determinant of centromeric function [revised in 14]. Incorporation of centromeric sequences in rice genome by transformation was not enough to induce the formation of active centromeres [17].

Cytogenetic analyses have shown that centromeres eventually adopt new positions in the chromosomes, an event called repositioning. The centromere in this new location is referred as evolutionary new centromeres or neocentromeres [18]. In human chromosomes, chromatin of new centromeres has CenH3 but its DNA is not distinguishable from the rest of the genomic DNA, indicating that there is no requirement in terms of DNA composition for centromere formation and function [19]. Furthermore, CenH3 mutants showed disturbed centromeric chromatin structure and kinetochore assembly, resulting in severe defects in chromosome segregation in *Saccharomyces cerevisiae*, *Caenorhabditis elegans*, *Schizosaccharomyces pombe*, *Drosophila melanogaster* and humans [revised in 2]. Another singularity of the centromeric H3 histone in humans (CENP-A) is the interaction with a chaperone protein to facilitate its assembly into nucleosomes. CENP-A associates to Holliday junction recognition protein (HJURP) that acts as a specific chaperone for CENP-A and is required for the incorporation of newly synthesized CENP-A molecules into nucleosomes at replicated centromeres [12, 20 revised in 12].

Logsdon et al. [21] used H3 chimeras containing different parts of CenH3 to study if the regions of this histone required for maintaining a functional centromere are the same required to the establishment of a new centromere. They found out that the regions of CenH3 required for those situations are actually different and that different proteins are recruited for new centromere formation [21].

Studies with dicentric chromosomes brought important information about how dynamic the CenH3 deposition can

be and about its fundamental role in centromeric activity. In maize, inactive centromeres in dicentric chromosomes, although composed of the characteristic centromeric DNA for the species (CentC and CRM), showed no constriction and no evidence of association with CenH3 [22]. These same inactive centromeres, when separated from their dicentric condition by intrachromosomal recombination, recovered their function in some cells and were stably transmitted to the next generation, due to de novo CenH3 deposition without centromeric DNA changes [23].

A peculiar case for CenH3 organization was described in leguminous species [24]. The authors identified in some species of the group two paralogous genes for CenH3 called CenH3-1 and CenH3-2, which share 75% similarity. Both genes produced functional CenH3 proteins that colocalized in immunofluorescence analysis. Probably, these genes derived from a duplication in the common ancestor of Fabaceae species. One of them was lost or silenced in some genera and both were maintained in *Pisum* L. and in *Lathyrus* L. In species of these two genera meta-polycentric chromosomes were described, with several CenH3 sites located in just one elongated primary constriction. The authors attribute to the gene duplication this unique centromere organization in these genera.

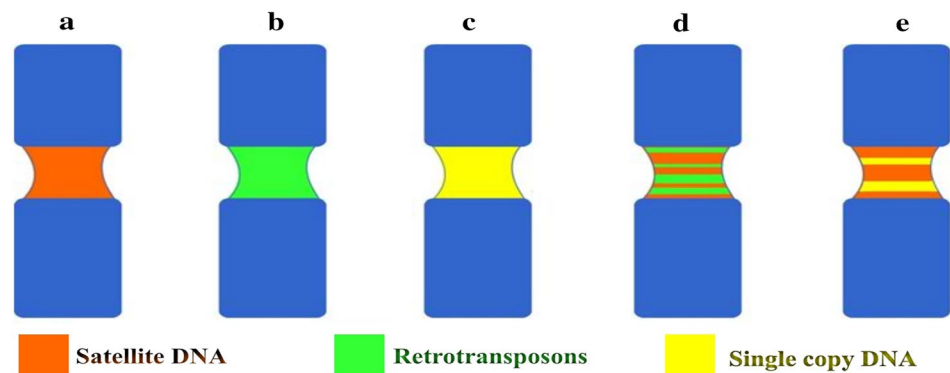
It is unknown why the replacement of canonical H3 by CenH3 turns the centromere into such a distinguishable site of the chromosome, but some studies started showing some insights. A comparison of the active and inactive centromeric chromatin in *Candida albicans* showed that nucleosomal periodicity, which is highly consistent in not centromeric nucleosomes, is suspended in active centromeres [25]. Thermodynamic analyzes performed in *Kluyveromyces lactis* showed that CenH3 nucleosomes are more stable and have more mobility than the canonical ones [26]. Chromatographic studies of CenH3 nucleosomes also showed that they are more compact than the H3 nucleosomes [11]. Therefore, CenH3 nucleosomes have unique physical properties, but it is not clear how this can help in the recruitment of kinetochore components and in determining the centromeric function [2].

CenH3 plays a key role in centromere organization but numerous other factors are involved in the epigenetic regulation of centromeric chromatin [2], like transcription factors [27], remodeling factors that promote histones deacetylation and act to repress and to silence the transcription [28], the chaperone proteins that facilitate transcription in specific situations [29] and cofactors that promote chromatin opening by processing histone acetyltransferases [30]. These factors are certainly also involved in controlling some existing genes in the centromeric region, as described in rice [31] and potato [15]. In rice, active genes were found in several centromeres, however they were located in sub-domains constituted by canonical H3. In potato, although there are cases of genes in the same condition as described in rice, an actively transcribed gene is associated with CenH3 nucleosomes.

The centromeric DNA

The centromeres of higher eukaryotes have conserved kinetochore proteins [32], play a conserved function, but their DNA sequences are highly divergent in model organisms and in most of other studied species (Fig. 2) [10, 33]. Its main constituents are satellite DNAs and retrotransposons (Fig. 2a–c) [2]. In *Arabidopsis thaliana* (L.) Heynh, for example, centromeres of all chromosomes have the same main component, a long DNA satellite array of monomers with 178 bp [34]. In humans, similar size monomers (171 bp), known as α -satellite, are likewise found in all centromeres [35], but their composition in base pairs is totally different from repetitive DNA present in *A. thaliana* centromeres. Satellite DNA sequences that compose the centromeres can be highly divergent even among closely related species, as in rice. Centromeres of *Oryza sativa* L. have a 155 bp satellite DNA repeats that do not share any homology with the 154 bp repeat found in *Oryza brachyantha* A. Chev. & Roehr., a species of wild rice close to *O. sativa* [36]. One aspect that can be considered common to most of studied species is the high similarity of repetitive DNA sequences among the centromeres of the same chromosome

Fig. 2 Schematic representations of different types of DNA sequences and their combinations identified in eukaryotic centromeres. Centromeres composed of **a** satellite DNA, **b** retrotransposons, **c** single copy DNA, **d** satellite DNA and retrotransposons and **e** satellite DNA and single copy DNA



complement, as in *A. thaliana* [34], *O. sativa* [37], *O. brachyantha* [36], humans [35], among others. This is not the case of *Solanum* L. species. In *S. tuberosum* L. (cultivated potato) each of the 12 centromeres have a different genetic composition from the other ones. Six of them are composed by repetitive DNA, five by single copy DNA (Fig. 2d) and the last centromere presents the two types of DNA (Fig. 2e) [15]. For *Solanum verrucosum* Schldl., a wild species already proposed as an ancestral of potato, were also described centromeres composed of either single copy or satellite DNA [38]. Comparison of the two genomes (*S. tuberosum* and *S. verrucosum*) showed common satellites located in different chromosomes and divergent satellites, indicating rapid evolution of centromeric sequences in the genus [38]. Other plant species which also have multiple repeats associated with different centromeres are pea (*Pisum sativum* L.) [8] and bean (*Phaseolus vulgaris* L.) [39]. In pea centromeres, repetitive DNA sequences belonging to 13 different families were described unevenly distributed among chromosomes. They differ extensively in nucleotides composition, in genomic abundance and monomers length, ranging from 50 to 2094 bases [8]. In beans, two satellite repeats were described, CentPv1 and CentPv2, predominantly located in subsets of eight and three centromeres, respectively. The arrangements of both suffered chromosome-specific homogenization processes and are found mixed in the genome [39].

Regarding to the size of centromeric repeats, in most of the studied plant species, they are around 150–180 bp long. This is probably the most common size to be ideal for the monomer to a full turn around a single nucleosome [40]. This hypothesis was supported by an experiment with rice centromeres, composed of Cent-O satellite repeats 155 bp long. When the chromatin was digested with micrococcal nuclease, an enzyme which preferably digests DNA linkers keeping intact the DNA involving nucleosomes, a fragment of 90–100 bp from the satellite repeat was always protected by CenH3 nucleosomes. The authors found, therefore, a matching frequency for the distribution of repeat units and nucleosomes [41]. This relation has been described only in the satellite repeats present in rice and in human centromeres so far [42]. On the other hand, in potato, discrepant data about this size trend were described, with centromeric monomers longer than 390 bp in *S. verrucosum* [38] and ranging from 979 bp to 5.4 kb in *S. tuberosum* [15].

In addition to this structural feature, satellite repeats may contain sequences that trigger interaction between centromeric chromatin and external factors for centromere function. A strong indication of this is the presence of a 17 bp motif called CENP-B box, found in human alpha satellite monomers. This motif binds to CEN-B protein, probably as an important event for kinetochore formation, since this protein plays a key role in microtubules organization [43]. Similar motifs to CENP-B box were observed in several

satellite DNA families, but it is not known yet about their functional significance [44].

Beyond the satellite repeats, long terminal repeats (LTR) retrotransposons accumulate quite often in centromeres and pericentromeres in plants and animals [revised in 1, 33]. There are evidences that the transpositional activity of centromeric LTR elements influence the evolutionary dynamics of centromere, as well as its structure and function, by the generation of new insertions that can later go through illegitimate or unequal homologous recombination processes [33]. Furthermore, it is common to find similarity between retrotransposons and satellite DNA, indicating that they can be a source of new repeats, as reported in potato [15]. Retrotransposons are also found in new centromeres, where they dilute the gene contents over time, contribute to centromere size maintaining and increase the repeat content of new centromeres that lack the tandem repeats [45]. In *Rynchospora* (Cyperaceae), the holocentromeres are composed of multiple centromeric units interspersing euchromatin regions and are highly enriched by a specific satellite family (Tyba) and by centromeric retrotransposons (CRRh). This is the first description of centromeric specific DNA sequences in holokinetic centromeres, showing existence of different types of holocentromeres regarding DNA composition [46].

The success in constructing artificial human chromosomes, obtained only after repetitive DNA was used to compose the centromeres [47], suggested that DNA may play an important but unknown role for centromere function. Experiments showed that plant centromeric DNA do not produce functional centromeres when reintroduced into plant cells [17] and new centromeres are functional even if located in loci with non-centromeric DNA [19], indicating that the DNA, although it may be important, it is not essential for centromeric function.

Origin and evolution

The centromere is a very interesting structure from an evolutionary point of view as it has conserved functions and highly variable DNA. Even CenH3, whose presence is consistent in the studied centromeres, has either conserved or variable domains, especially in their C-terminal tail [10], whereas canonical histone H3 is identical in many eukaryotes [48]. There are evidences that this interesting structure of CenH3, with conserved and not conserved domains, is the result of convergent evolution. In this case, different histones have independently converged to a common centromeric function [49]. On the other hand, there are evidence of divergent evolution originating CenH3 in Brassicaceae [48] and in *Drosophila* [50]. The authors believe that the divergence of CenH3 is an adjustment to the rapidly changing

centromeric DNA. Apparently the CenH3 has suffered both, convergent and divergent evolution [revised in 2].

All existing variation in centromeres has increased the interest of the scientific community to find out why these variations exist and what their evolutionary meaning is. Besides this, the formation of new centromeres has been targeted by many studies. New centromeres have been described extensively in human chromosomes [reviewed in 51] and in some plant species like barley [52], maize [22, 53, 54] and rice [55].

The emergence of new centromeres was also observed in hybrid conditions, such as hybrids between maize and oat. Nine maize chromosomes had their centromeric domains dramatically expanded in the hybrid condition, probably to better compete for spindle fibers with oat chromosomes, which have bigger centromeres. However, in two chromosomes the pericentromeric region had active genes and the expansion would probably harm their expression. In these cases, functional centromeres appeared in another chromosomal locus to allow its expansion without damage to transcription [54].

The repositioning cases show that the establishment of a new centromere does not require specific DNA composition in the target loci [1, 10]. Most new centromeres have no satellite DNA [56]. On the other hand, most of studied mature centromeres are massively composed by repetitive DNA, especially satellite DNA [1, 10]. One hypothesis to explain this apparent contradiction is that new centromeres appear in single copy DNA regions and acquire long arrays of repetitive DNA, especially satellites, during its evolution. According to this theory, a new satellite repeat or one already present in other centromeres may invade and occupy the CenH3 domain of the new centromere. The existence of centromeres with an intermediate configuration containing both single-copy DNA and satellite DNA, observed in rice [31] and potato [15] is an evidence for this evolutionary pathway of satellite DNA acquisition [57]. The satellite DNA invasion mechanism is still unknown and the retrotransposons would be the main source for the origin of new repeats [15].

Conclusion

The observed variations in size, CenH3 and DNA composition of centromeres are intriguing considering their conserved function and reflect the high dynamism of this chromosomal site. CenH3 is required for centromeric function, while the DNA appears not to be, although play an important role in the centromeres maintenance and evolution. The plant centromeres, as in other eukaryotes, are highly complex and dynamic structures. The control of its operation and its heritage involves epigenetic phenomena that seem to be quite distinct from the control of any other chromosomal

region, which increases the interest and the difficulties in centromere studies. The recent technologies of next generation sequencing, chromatin immunoprecipitation, protein immunolocalization and fluorescence in situ hybridization have enabled significant advances in understanding the centromere origin, structure and evolution. Given the diversity found so far, it is very important to increase the number of studied species, especially those with sequenced genomes, so it will be possible to have a more elucidatory picture of this intriguing structure essential to the maintenance of life.

Acknowledgements The authors thank Fundação de Amparo à Pesquisa do Estado de Minas Gerais (FAPEMIG) and Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) for financing the projects and Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES) for granting scholarship to L.C.O.

References

- Jiang J, Birchler JA, Parrott WA, Dawe RK (2003) A molecular view of plant centromeres. *Trends Plant Sci* 8:570–575. <https://doi.org/10.1016/j.tplants.2003.10.011>
- Ekwall K (2007) Epigenetic control of centromere behavior. *Annu Rev Genet* 41:63–81. <https://doi.org/10.1146/annurev.genet.41.110306.130127>
- Cleveland DW, Mao Y, Sullivan KF (2003) Centromeres and kinetochores: from epigenetics to mitotic checkpoint signaling. *Cell* 112:407–421. [https://doi.org/10.1016/S0092-8674\(03\)00115-6](https://doi.org/10.1016/S0092-8674(03)00115-6)
- Birchler JA, Han F (2013) Centromere epigenetics in plants. *J Genet Genomics* 40:201–204. <https://doi.org/10.1016/j.jgg.2013.03.008>
- Furuyama S, Biggins S (2007) Centromere identity is specified by a single centromeric nucleosome in budding yeast. *Proc Natl Acad Sci USA* 104:14706–14711. <https://doi.org/10.1073/pnas.0706985104>
- Pluta AF, Mackay AM, Ainsztein AM, Goldberg IG, Earnshaw WC (1995) The centromere: hub of chromosomal activities. *Science* 270:1591–1594. <https://doi.org/10.1126/science.270.5242.1591>
- Melters DP1, Paliulis LV, Korf IF, Chan SW (2012) Holocentric chromosomes: convergent evolution, meiotic adaptations, and genomic analysis. *Chromosome Res* 20:579–593. <https://doi.org/10.1007/s10577-012-9292-1>
- Neumann P, Navrátilová A, Schroeder-Reiter E, Koblížková A, Steinbauerová V, Chocholová E, Novák P, Wanner G, Macas J (2012) Stretching the rules: monocentric chromosomes with multiple centromere domains. *PLoS Genet* 8:e1002777. <https://doi.org/10.1371/journal.pgen.1002777>
- Wang G, Zhang W, Jin W (2009) An overview of plant centromeres. *J Genet Genomics* 36:529–537. [https://doi.org/10.1016/S1673-8527\(08\)60144-7](https://doi.org/10.1016/S1673-8527(08)60144-7)
- Henikoff S, Ahmad K, Malik HS (2001) The centromere paradox: stable inheritance with rapidly evolving DNA. *Science* 293:1098–1102. <https://doi.org/10.1126/science.1062939>
- Black BE, Foltz DR, Chakravarthy S, Luger K, Woods VL Jr, Cleveland DW (2004) Structural determinants for generating centromeric chromatin. *Nature* 430:578–582. <https://doi.org/10.1038/nature02766>
- Srivastava S, Zasadzińska E, Foltz DR (2018) Posttranslational mechanisms controlling centromere function and assembly.

- Curr Opin Cell Biol 52:126–135. <https://doi.org/10.1016/j.ceb.2018.03.003>
13. Demidov D, Schubert V, Kumke K, Weiss O, Karimi-Ashtiyani R, Buttler J, Heckmann S, Wanner G, Dong Q, Han F, Houben A (2014) Anti-phosphorylated histone H2AThr120: a universal microscopic marker for centromeric chromatin of mono- and holocentric plant species. *Cytogenet Genome Res* 143:150–156. <https://doi.org/10.1159/000360018>
 14. Feng C, Liu Y, Su H, Wang H, Birchler J, Han F (2015) Recent advances in plant centromere biology. *Sci China Life Sci* 58:240–245. <https://doi.org/10.1007/s11427-015-4818-3>
 15. Gong Z, Wu Y, Koblížková A, Torres GA, Wang K, Iovene M, Neumann P, Zhang W, Novák P, Buell CR, Macas J, Jiang J (2012) Repeatless and repeat-based centromeres in potato: implications for centromere evolution. *Plant Cell* 24:3559–3574. <https://doi.org/10.1105/tpc.112.100511>
 16. Vafa O, Sullivan KF (1997) Chromatin containing CENP-A and α -satellite DNA is a major component of the inner kinetochore plate. *Curr Biol* 7:897–900. [https://doi.org/10.1016/S0960-9822\(06\)00381-2](https://doi.org/10.1016/S0960-9822(06)00381-2)
 17. Phan BH, Jin W, Topp CN, Zhong CX, Jiang J, Dawe RK, Parrott WA (2007) Transformation of rice with long DNA-segments consisting of random genomic DNA or centromere-specific DNA. *Transgenic Res* 16:341–351. <https://doi.org/10.1007/s11248-006-9041-3>
 18. McKinley KL, Cheeseman IM (2016) The molecular basis for centromere identity and function. *Nat Rev Mol Cell Biol* 17:16–29. <https://doi.org/10.1038/nrm.2015.5>
 19. Marshall OJ, Chueh AC, Wong LH, Choo KH (2008) Neocentromeres: new insights into centromere structure, disease development, and karyotype evolution. *Am J Hum Genet* 82:261–282. <https://doi.org/10.1016/j.ajhg.2007.11.009>
 20. Foltz DR, Jansen LET, Bailey AO, Yates JR, Bassett EA, Wood S, Black BE, Cleveland DW (2009) Centromere-specific assembly of CENP-A nucleosomes is mediated by HJURP. *Cell* 137:472–484. <https://doi.org/10.1016/j.cell.2009.02.039>
 21. Logsdon GA, Barrey EJ, Bassett EA, DeNizio JE, Guo LY, Panchenko T, Dawicki-McKenna JM, Heun P, Black BE (2015) Both tails and the centromere targeting domain of CENP-A are required for centromere establishment. *J Cell Biol* 208:521–531. <https://doi.org/10.1083/jcb.201412011>
 22. Han F, Lamb JC, Birchler JA (2006) High frequency of centromere inactivation resulting in stable dicentric chromosomes of maize. *Proc Natl Acad Sci USA* 103:3238–3243. <https://doi.org/10.1073/pnas.0509650103>
 23. Han F, Gao Z, Birchler JA. Reactivation of an inactive centromere reveals epigenetic and structural components for centromere specification in maize. *Plant Cell* 21:1929–1939. <https://doi.org/10.1105/tpc.109.066662>
 24. Neumann P, Pavlíková Z, Koblížková A, Fuková I, Jedličková V, Novák P, Macas J (2015) Centromeres off the hook: massive changes in centromere size and structure following duplication of CenH3 gene in Fabaceae species. *Mol Biol Evol* 7:1862–1879. <https://doi.org/10.1093/molbev/msv070>
 25. Baum M, Sanyal K, Mishra PK, Thaler N, Carbon J (2006) Formation of functional centromeric chromatin is specified epigenetically in *Candida albicans*. *Proc Natl Acad Sci USA* 103:14877–14882. <https://doi.org/10.1016/j.jgg.2013.03.008>
 26. Mattei S, Sampaiolese B, De Santis P, Savino M (2002) Nucleosome organization on *Kluyveromyces lactis* centromeric DNAs. *Biophys Chem* 97:173–187. [https://doi.org/10.1016/S0301-4622\(02\)00066-2](https://doi.org/10.1016/S0301-4622(02)00066-2)
 27. Mellor J, Jiang W, Funk M, Rathjen J, Barnes CA, Hinz T, Hegemann JH, Philippsen P (1990) CPF1, a yeast protein which functions in centromeres and promoters. *EMBO J* 9:4017–4026
 28. Walfridsson J, Bjerling P, Thalen M, Yoo EJ, Park SD, Ekwall K (2005) The CHD remodeling factor Hrp1 stimulates CENP-A loading to centromeres. *Nucleic Acids Res* 33:2868–2879. <https://doi.org/10.1093/nar/gki579>
 29. Foltz DR, Jansen LE, Black BE, Bailey AO, Yates JR, Cleveland DW (2006) The human CENP-A centromeric nucleosome-associated complex. *Nat Cell Biol* 8:458–469. <https://doi.org/10.1038/ncb1397>
 30. Minoda A, Saitoh S, Takahashi K, Toda T (2005) BAF53/Arp4 homolog Alp5 in fission yeast is required for histone H4 acetylation, kinetochore-spindle attachment, and gene silencing at centromere. *Mol Biol Cell* 16:316–327. <https://doi.org/10.1091/mbc.E04-06-0519>
 31. Nagaki K, Cheng Z, Ouyang S, Talbert PB, Kim M, Jones KM, Henikoff S, Buell CR, Jiang J (2004) Sequencing of a rice centromere uncovers active genes. *Nat Genet* 36:138–145. <https://doi.org/10.1038/ng1289>
 32. Henikoff S, Furuyama T (2010) Epigenetic inheritance of centromeres. *Cold Spring Harb Symp Quant Biol* 75:51–60. <https://doi.org/10.1101/sqb.2010.75.001>
 33. Ma J, Wing RA, Bennetzen JL, Jackson SA (2007) Plant centromere organization: a dynamic structure with conserved functions. *Trends Genet* 23:134–139. <https://doi.org/10.1016/j.tig.2007.01.004>
 34. Nagaki K, Talbert PB, Zhong CX, Dawe RK, Henikoff S, Jiang J (2003) Chromatin immunoprecipitation reveals that the 180-bp satellite repeat is the key functional DNA element of *Arabidopsis thaliana* centromeres. *Genetics* 163:1221–1225
 35. Willard HF (1998) Centromeres: the missing link in the development of human artificial chromosomes. *Curr Opin Genet Dev* 8:219–225. [https://doi.org/10.1016/S0959-437X\(98\)80144-5](https://doi.org/10.1016/S0959-437X(98)80144-5)
 36. Lee HR, Zhang W, Langdon T, Jin W, Yan H, Cheng Z, Jiang J (2005) Chromatin immunoprecipitation cloning reveals rapid evolutionary patterns of centromeric DNA in *Oryza* species. *Proc Natl Acad Sci USA* 102:11793–11798. <https://doi.org/10.1073/pnas.0503863102>
 37. Cheng Z, Dong F, Langdon T, Ouyang S, Buell CR, Gu M, Blattner FR, Jiang J (2002) Functional rice centromeres are marked by a satellite repeat and a centromere-specific retrotransposon. *Plant Cell* 14:1691–1704. <https://doi.org/10.1105/tpc.003079>
 38. Zhang H, Koblížková A, Wang K, Gong Z, Oliveira L, Torres GA, Wu Y, Zhang W, Novák P, Buell CR, Macas J, Jiang J (2014) Boom-bust turnovers of megabase-sized centromeric DNA in *Solanum* species: rapid evolution of DNA sequences associated with centromeres. *Plant Cell* 26:1436–1447. <https://doi.org/10.1105/tpc.114.123877>
 39. Iwata A, Tek AL, Richard MM, Abernathy B, Fonsêca A, Schmutz J, Chen NW, Thareau V, Magdelenat G, Li Y, Murata M, Pedrosa-Harand A, Geffroy V, Nagaki K, Jackson SA (2013) Identification and characterization of functional centromeres of the common bean. *Plant J* 76:47–60. <https://doi.org/10.1111/tpj.12269>
 40. Heslop-Harrison JSP, Schwarzacher T (2013) Nucleosomes and centromeric DNA packaging. *Proc Natl Acad Sci USA* 110:19974–19975. <https://doi.org/10.1073/pnas.1319945110>
 41. Zhang T, Talbert PB, Zhang W, Wu Y, Yang Z, Henikoff JG, Henikoff S, Jiang J (2013) The CentO satellite confers translational and rotational phasing on cenH3 nucleosomes in rice centromeres. *Proc Natl Acad Sci USA* 110:4875–4883. <https://doi.org/10.1073/pnas.1319548110>
 42. Hasson D, Panchenko T, Salimian KJ, Salman MU, Sekulic N, Alonso A, Warburton PE, Black BE (2013) The octamer is the major form of CENP-A nucleosomes at human centromeres. *Nat Struct Mol Biol* 20:687–695. <https://doi.org/10.1038/nsmb.2562>
 43. Masumoto H, Nakano M, Ohzeki J (2004) The role of CENP-B and alphasatellite DNA: de novo assembly and epigenetic

- maintenance of human centromeres. *Chromosome Res* 12:543–556. <https://doi.org/10.1023/B:CHRO.0000036593.72788.99>
44. Canapa A, Barucca M, Cerioni PN, Olmo E (2000) A satellite DNA containing CENP-B box-like motifs is present in the antarctic scallop *Adamussium colbecki*. *Gene* 247:175–180. [https://doi.org/10.1016/S0378-1119\(00\)00101-3](https://doi.org/10.1016/S0378-1119(00)00101-3)
 45. Presting G (2018) Centromeric retrotransposons and centromere function. *Curr Opin Genet Dev* 49:79–84. <https://doi.org/10.1016/j.gde.2018.03.004>
 46. Marques A, Ribeiro T, Neumann P, Macas J, Novák P, Schubert V, Pellino M, Fuchs J, Ma W, Kuhlmann M, Brandt R, Vanzela ALL, Beseda T, Simková H, Pedrosa-Harand A, Houbenb A (2015) Holocentromeres in *Rhynchospora* are associated with genome-wide centromere-specific repeat arrays interspersed among euchromatin. *Proc Natl Acad Sci USA*. 112:13633–13638. <https://doi.org/10.1073/pnas.1512255112>
 47. Harrington JJ, Van Bokkelen G, Mays RW, Gustashaw K, Willard HF (1997) Formation of de novo centromeres and construction of first-generation human artificial microchromosomes. *Nat Genet* 15:345–355. <https://doi.org/10.1038/ng0497-345>
 48. Cooper JL, Henikoff S (2004) Adaptive evolution of the histone fold domain in centromeric histones. *Mol Biol Evol* 21:1712–1718. <https://doi.org/10.1093/molbev/msh179>
 49. Malik HS, Henikoff S. Phylogenomics of the nucleosome. *Nat Struct Mol Biol* 10:882–891. <https://doi.org/10.1038/nsb996>
 50. Malik HS, Henikoff S (2001) Adaptive evolution of Cid, a centromere-specific histone in *Drosophila*. *Genetics* 157:1293–1298
 51. Warburton PE (2004) Chromosomal dynamics of human neocentromere formation. *Chromosome Res* 12:617–626. <https://doi.org/10.1023/B:CHRO.0000036585.44138.4b>
 52. Nasuda S, Hudakova S, Schubert I, Houben A, Endo TR (2005) Stable barley chromosomes without centromeric repeats. *P Natl Acad Sci USA* 102:9842–9847. <https://doi.org/10.1073/pnas.0504235102>
 53. Topp CN, Okagaki RJ, Melo JR, Kynast RG, Phillips RL, Dawe RK (2009) Identification of a maize neocentromere in an oat-maize addition line. *Cytogenet Genome Res* 124:228–238. <https://doi.org/10.1159/000218128>
 54. Wang K, Wu Y, Zhang W, Dawe RK, Jiang J (2014) Maize centromere expand and adopt a uniform size in the genetic background of oat. *Genome Res* 24:107–116. <https://doi.org/10.1101/gr.160887.113> 2014.
 55. Yan H, Jiang J (2007) Rice as a model for centromere and heterochromatin research. *Chromosome Res* 15:77–84. <https://doi.org/10.1007/s10577-006-1104-z>
 56. Hasson D, Alonso A, Cheung F, Tepperberg JH, Papenhausen PR, Engelen JJ, Warburton PE (2011) Formation of novel CENP-A domains on tandem repetitive DNA and across chromosome breakpoints on human chromosome 8q21 neocentromeres. *Chromosoma* 120:621–632. <https://doi.org/10.1007/s00412-011-0337-6>
 57. Jiang J (2013) Centromere evolution. In: Jiang J, Birchler JA (eds) *Plant centromere biology*, 1st edn. Wiley, New Jersey, pp 159–168